ESAIM: M2AN 54 (2020) 2011–2043
https://doi.org/10.1051/m2an/2020046

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

PROJECTION-BASED MODEL REDUCTION WITH DYNAMICALLY TRANSFORMED MODES

Felix Black, Philipp Schulze and Benjamin Unger*

Abstract. We propose a new model reduction framework for problems that exhibit transport phenomena. As in the moving finite element method (MFEM), our method employs time-dependent transformation operators and, especially, generalizes MFEM to arbitrary basis functions. The new framework is suitable to obtain a low-dimensional approximation with small errors even in situations where classical model order reduction techniques require much higher dimensions for a similar approximation quality. Analogously to the MFEM framework, the reduced model is designed to minimize the residual, which is also the basis for an *a posteriori* error bound. Moreover, since the dependence of the transformation operators on the reduced state is nonlinear, the resulting reduced order model is obtained by projecting the original evolution equation onto a nonlinear manifold. Furthermore, for a special case, we show a connection between our approach and the method of freezing, which is also known as symmetry reduction. Besides the construction of the reduced order model, we also analyze the problem of finding optimal basis functions based on given data of the full order solution. Especially, we show that the corresponding minimization problem has a solution and reduces to the proper orthogonal decomposition of transformed data in a special case. Finally, we demonstrate the effectiveness of our method with several analytical and numerical examples.

Mathematics Subject Classification. 35Q35, 65M15, 37L65.

Received December 23, 2019. Accepted July 1, 2020.

1. Introduction

Over the past three decades, *model order reduction* (MOR) has become an established tool to reduce the computational cost for obtaining high fidelity solutions of (partial) differential-algebraic equations that are required in parameter studies, controller design, and optimization. The key observation that is used in MOR is that in many applications, the solution evolves in a low-dimensional manifold, which itself can be embedded approximately in a low-dimensional subspace. The projection of the dynamical system onto this low-dimensional subspace diminishes the computational cost while maintaining a high fidelity solution. A successful MOR scheme can identify a suitable low-dimensional subspace and quantify the error of the solution of the resulting surrogate model with respect to the solution of the original dynamics. For an overview of such methods we refer to [1, 6, 33, 58, 59] and the recent surveys [4, 5]. Most of these MOR methods are formulated in a projection

Keywords and phrases. Transport dominated phenomena, model order reduction, error bound, nonlinear Galerkin, residual minimization.

Institut für Mathematik, Technische Universität Berlin, Str. des 17. Juni 136, 10623 Berlin, Germany.
*Corresponding author: unger@math.tu-berlin.de

Article published by EDP Sciences

framework. In more detail, consider a separable Hilbert space $(\mathscr{X}, \langle\cdot,\cdot\rangle_{\mathscr{X}})$ with induced norm $\|\cdot\|_{\mathscr{X}}$ and an evolution equation of the form

$$\dot{z}(t) = \mathcal{F}(z(t)), \quad z(0) = z_0, \quad t \in (0,T), \tag{1.1}$$

where the operator $\mathcal{F}$ is defined on a dense subspace $\mathscr{Y} \subseteq \mathscr{X}$,

$$\mathcal{F}\colon \mathscr{Y} \to \mathscr{X}, \quad y \mapsto \mathcal{F}(y).$$

We assume $z_0 \in \mathscr{Y}$ and call z a solution of (1.1) if $z(t) \in \mathscr{Y}$ for all $t \geq 0$, z is continuous in $[0,T)$, differentiable in $(0,T)$, and (1.1) is satisfied.

Standard projection-based MOR is based on the identification of a suitable r-dimensional subspace $\mathcal{Y} \subseteq \mathscr{Y}$, conveniently described by a basis $(\varphi_1,\ldots,\varphi_r)$, and the assumption that the solution z of (1.1) is well-approximated within this space, *i.e.*, there exist scalar functions $\tilde{z}_i$ such that

$$z(t) \approx \hat{z}(t) := \sum_{i=1}^{r} \tilde{z}_i(t)\varphi_i \qquad \text{for } t \in [0,T). \tag{1.2}$$

Substituting $\hat{z}$ into (1.1) and requiring that the residual is orthogonal to the approximation space $\mathcal{Y}$ results in the *reduced order model* (ROM)

$$\sum_{j=1}^{r} \dot{\tilde{z}}_j \langle\varphi_i,\varphi_j\rangle_{\mathscr{X}} = \left\langle \varphi_i, \mathcal{F}\left(\sum_{j=1}^{r}\tilde{z}_j\varphi_j\right)\right\rangle_{\mathscr{X}}, \qquad i = 1,\ldots,r. \tag{1.3}$$

Note that the matrix $M := [\langle\varphi_i,\varphi_j\rangle_{\mathscr{X}}]_{i,j=1}^{r}$ is symmetric and nonsingular such that under reasonable assumptions on $\mathcal{F}$ the ROM (1.3) is uniquely solvable for any initial condition. For numerical reasons, the basis $(\varphi_1,\ldots,\varphi_r)$ is usually chosen to be orthonormal such that M equals the r-dimensional identity matrix. The question that remains to be answered is how to choose the subspace $\mathcal{Y}$. The particular choice of this subspace is one of the distinguishing features of the different MOR methods.

The best subspace of a given dimension – in the sense that the worst-case approximation error is minimized – is described by the Kolmogorov n-widths [40,56] (or equivalently as shown in [78] by the Hankel singular values). For specific equations, it can be shown (*cf.* [45,46,59]) that the n-widths decay exponentially, enabling MOR to succeed. Unfortunately, there are also several problem classes, where the n-widths do not decay exponentially. This is typically the case if the dynamical system features strong convection or transport of a quantity within the spatial domain. For examples we refer to [11,26]. For such problems, standard MOR methods cannot produce a low-dimensional model that, at the same time, is very accurate. Several authors have observed this, yielding the emerging field of so-called MOR for transport dominated phenomena. We give a detailed overview of the current state of the art in Section 2.

The slow decay of the Kolmogorov n-widths is often related to sharp gradients of the solution that travel through the physical domain. The large gradients might develop over time (for instance, in Burgers' equation with large Reynolds number) or are enforced within the system by the initial condition. The *moving finite element method* (MFEM) [21,48] accounts for these typically local effects by moving the discretization nodes in an automated fashion to the critical areas. Therefore the basis functions are transformed based on the current position of the respective nodes, and the node positions are considered as additional unknowns. The equation for the node position is derived from the necessary condition for minimizing the squared norm of the residual. In this paper we transfer this idea to the context of model order reduction by replacing the approximation (1.2) with

$$z(t) \approx \hat{z}(t) := \sum_{i=1}^{r} \tilde{z}_i(t)\mathcal{T}_i(p_i(t))\varphi_i \qquad \text{for } t \in [0,T) \tag{1.4}$$

with suitable transformation operators $\mathcal{T}_i\colon \mathcal{P}_i \to \mathcal{B}(\mathscr{X})$ for $i = 1, \ldots, r$ that are $\mathscr{Y}$-invariant, *i.e.*, $\mathcal{T}_i(p)\mathscr{Y} \subseteq \mathscr{Y}$ for all $p \in \mathcal{P}_i$. Here, $\mathcal{B}(\mathscr{X})$ denotes the space of all linear and bounded operators on $\mathscr{X}$. For given paths p_i the adaptive low-dimensional approximation manifold is thus given as the linear span of $\{\mathcal{T}_i\left(p_i(t)\right)\varphi_i\}_{i=1}^{r}$. Since our main goal is to use the approximation (1.4) to derive a ROM, we have to ensure that we can differentiate with respect to time. In particular, we have to choose suitable sets $\mathcal{P}_i$. Natural choices are to assume that the $\mathcal{P}_i$'s are Lie groups or real Banach spaces. For the main applications we have in mind, it suffices to choose $\mathcal{P}_i = \mathbb{R}^{q_i}$ and for the ease of presentation we do not consider the more general cases described before. We immediately arise at three questions:

(1) What are suitable families of transformation operators $\mathcal{T}_i$?
(2) For given transformations $\mathcal{T}_i$, how to determine $\tilde{z}_i$, φ_i, and p_i such that the approximation error in (1.4) is minimized?
(3) How to use (1.4) in a projection framework such that the resulting ROM can be evaluated efficiently?

In practice, we often have knowledge about the solution behavior of the equations at hand and can thus design appropriate transformation operators. For instance, for wave-like phenomena, we can use translation operators that shift the basis functions or modes φ_i in the spatial domain (see the forthcoming Example 4.3 for the advection equation and Rem. 6.2 for a more abstract discussion). Therefore, we assume for the remainder of the manuscript that the transformations $\mathcal{T}_i$ are given such that we only have to deal with the second and the third question, which we address in Sections 4 and 5, respectively. More precisely, we extend in Section 4 the residual minimization version of the *shifted proper orthogonal decomposition* (shifted POD) as presented in [71] to the infinite-dimensional setting and, thereby, answer parts of the second question. The resulting adaptive basis is then used to obtain a ROM *via* Galerkin projection. In most applications, the paths p_i are unknown and have to be computed during the online phase along with the coefficient functions $\tilde{z}_i$, which renders the approximation (1.4) nonlinear. Consequently, the ROM that is obtained *via* Galerkin projection of (1.1) is an underdetermined nonlinear *differential-algebraic equation* (DAE) that needs to be completed with additional equations. This is discussed in detail in Section 5. Our main contributions are the following:

(1) In Section 2, we give a detailed overview of the literature on model reduction methods with special emphasis on transport dominated problems or problems with slowly decaying Kolmogorov n-widths.
(2) We show (*cf.* Thm. 4.6) that for given paths p_i, the optimization problem that minimizes the residual in (1.4) has a solution. In the special case that all modes are transformed with the same operator and the same path we establish in Theorem 4.8 that the minimization problem is equivalent to standard *proper orthogonal decomposition* (POD) (see Sect. 3.1) with transformed data.
(3) Based on the approximation ansatz (1.4), we construct a ROM that minimizes the residual, see Theorem 5.5, and thus extend the idea of the MFEM to the model reduction context. The ROM is certified in the sense that we provide an *a posteriori* residual-based error bound in Theorem 5.10.
(4) For the special case that all modes are transformed with the same operator and the same path, we discuss in Section 6 the close connection to the symmetry reduction framework [7,64]. In particular, Theorem 6.5 details the connection of the phase condition (*cf.* (5.4)) that is obtained from minimizing the residual and the phase condition that minimizes the temporal change of the reduced state.

We close our discussion with four numerical examples. In the case of the advection-diffusion equation, which is discussed in Section 7.1, our ROM is capable of yielding a good approximation of the dynamics with only two modes. In addition, this example numerically justifies the approach to use the same transformation for several modes, as proposed in the forthcoming equation (1.5). This idea is not only advantageous for the actual implementation, but also for the theory; see Proposition 5.7 and Remark 5.8. In Section 7.2, we demonstrate that our method can also obtain accurate low-dimensional models for the advection-diffusion equation with non-periodic boundary conditions. In this context, we also illustrate that the full-order model's discretization error has a significant influence on the approximation quality of our reduced-order model. In Section 7.3, we consider the linear wave equation and obtain an excellent agreement with the *full order model* (FOM) by using an approximation with only two modes. Furthermore, we observe that the ROM allows choosing significantly larger time steps than the full order model, which is an advantage of performing the model reduction on the time-continuous level. Finally, we consider the viscous Burgers' equation in Section 7.4 as an example of a nonlinear full order model. In this context, we discuss how the evaluation of the reduced order model can be rendered independent from the full order dimension. Furthermore, our approach outperforms the classical POD-Galerkin approach and is able to yield a decent approximation of the dynamics with just seven modes.

Remark 1.1. In [43], the authors use a general nonlinear approximation of the form $z(t) \approx g(\tilde{z}_1(t), \dots, \tilde{z}_r(t))$, which certainly generalizes our approximation approach presented in (1.4). We believe however that our parametrization of the general nonlinear approximation offers some advantages. First, in the general setting with a possibly infinite-dimensional Banach space $\mathscr{X}$ it seems rather difficult to prescribe a suitable nonlinear operator $g\colon \mathbb{R}^r \to \mathscr{X}$, even within a deep-learning framework as proposed in [43]. Working in the infinite-dimensional setting often allows using equivariance (see Sect. 6 for further details), which may not be possible anymore after semi-discretization in space. Moreover, the separation of amplitudes and transformed modes as in (1.4), gives a direct physical interpretation of the ROM, where the transformed mode $\mathcal{T}_i(p_i(t))\varphi_i$ may be related to a reference frame. For additional numerical and theoretical benefits, this may be exploited even further by prescribing the same transformation operator with the same path for different modes φ_i, *i.e.*, to replace (1.4) with

$$z(t) \approx \sum_{i=1}^{\hat{r}} \sum_{j=1}^{r_i} \tilde{z}_{i,j}(t) \mathcal{T}_i(p_i(t)) \varphi_{i,j}. \tag{1.5}$$

On the one hand, this approach reduces the computational cost required to solve the ROM, since less path variables need to be computed and, on the other hand, provides more flexibility for theoretical investigations, see the forthcoming Remark 5.8.

2. State of the art

As mentioned in the introduction, classical model order reduction techniques are usually not able to provide accurate low-dimensional models when applied to systems exhibiting the transport of structures with large gradients, such as shocks. In the past years, there has been an increasing effort in the model reduction community to develop new methods that can solve this problem. The most relevant approaches can be roughly subdivided into three classes. The first class of methods aims at describing the transport by appropriately chosen time-dependent coordinate transformations, which we hence refer to as *reference frame methods*. For instance, in the case of a simple advection problem, an apparent coordinate transformation is given by a time-dependent translation that describes the advective behavior. The second class incorporates a more general time-dependent update of the basis functions, which is not necessarily described by a coordinate transformation. This update of the basis functions may be, *e.g.*, realized by an h-refinement-like enrichment of the basis functions or by an equation describing the evolution of the basis functions. The third class considers general nonlinear approximation ansatzes accompanied by a nonlinear Petrov–Galerkin projection to construct the ROM. The corresponding nonlinear mappings are usually constructed *via* techniques from machine learning.

Reference frame methods

A common approach among the methods using coordinate transformations consists of formulating the FOM in a new coordinate system, which we refer to as the *reference frame*. The goal is to choose the coordinate system in such a way that the Kolmogorov n-widths for the transformed problem decay fast and thus enable standard MOR methods to succeed. The first developments in this direction are presented within the symmetry reduction framework, *cf.* [7, 64, 65]. The main idea is to approximate the solution by a composition of the action of a time-dependent group element and a so-called *frozen solution*, *cf.* Section 6. Ideally, the group action and the time-dependent element are chosen such that the frozen solution is almost constant over time, which supports a low-dimensional approximation. The group action can, for example, be chosen based on physical considerations or from snapshot data of the full order solution. For instance, in [74], the authors present a method to pre-process snapshot data to align them in such a way that the singular values decay fast and, thus, a low-dimensional description of the dynamics can be obtained. To this end, they assume the snapshots to be almost identical up to the action of some underlying symmetry group. Here, the coordinate transformation is described by the group action, and the task is to assign group elements to the snapshots optimally. The proposed solution for this problem involves the recently introduced eigenvector alignment method [73], which, in comparison to other methods, has the advantage that no template snapshot needs to be chosen. Similar ideas to pre-process the snapshot data by some kind of alignment or calibration have been applied, for instance, in [52,72]. A different approach for determining a low-dimensional description of given snapshot data was presented in [36] and is based on solving optimal mass transfer problems.

The contributions mentioned so far mainly consider the task to find a low-dimensional approximation of snapshot data, but they do not construct a dynamic ROM based on the low-dimensional description. An approach that considers both tasks is, for instance, presented in [50]. At the first step, the authors apply a symmetry reduction tool, called the method of slices [64], to formulate the FOM in the reference frame. Afterward, they apply a standard model reduction scheme to the evolution equation in the reference frame. Similar approaches based on the idea to formulate the FOM in a different coordinate system and apply standard model reduction schemes afterward are discussed, *e.g.*, in [49, 53, 64, 75].

Instead of first transforming the full order model and then reducing the transformed system, the authors in [11] directly reduce the untransformed FOM using an approximation ansatz that includes a coordinate transformation. Their approach for the compression of snapshot data is similar to the ones mentioned above. However, they also present a method for constructing a ROM based on the identified low-dimensional description of the snapshot data. To this end, the authors first discretize in time and then substitute the approximation ansatz into the semi-discrete full order model. They present an algorithm for updating the time-discrete states of the ROM by minimizing the time-discrete residual. The evaluation of the ROM still scales with the full dimension, but they present an additional approximation of the ROM, which allows achieving an efficient offline/online decomposition. Further methods which are based on the idea to enhance the approximation ansatz *via* a coordinate transformation are discussed in [10, 23, 38, 51, 63].

All methods mentioned so far have in common that they apply a single coordinate transform in order to get a faster decay of the singular values or the Kolmogorov n-widths. However, one coordinate transformation may not be sufficient to obtain a low-dimensional description in the case of multiple transport velocities in the system. For instance, the analytical solution of the linear wave equation can be represented by two traveling waves, which cannot be described by two modes and a single transformation. So far, there are only a few methods that consider the case of multiple transport velocities and treat them with different coordinate transformations. For example, the shifted POD [60, 61, 71] considers an approximation ansatz, which allows to have different coordinate transformations for different sets of basis functions. To compute a low-dimensional approximation of given snapshot data, the authors in [61] propose a heuristic method that iteratively transforms the snapshots into the different reference frames and compresses the data in the respective frame using a singular value decomposition. In [60], the dominant modes are identified by solving an optimization problem that maximizes the leading singular values in each reference frame. Also, in [71], a shifted POD approximation is computed by solving an optimization problem, which aims at minimizing the deviation of the approximation from the original snapshot data. An alternative approach is considered in [62], where the authors also present a method for identifying modes in several reference frames similar to the shifted POD. However, instead of applying an iterative or optimization procedure, they extract the dominant modes one after another in a greedy fashion. Thus, this method has a lower computational complexity in comparison to the shifted POD. However, it fails to give the minimum number of modes even for examples like the linear wave equation where the analytic solution is known. Both the shifted POD and the transport reversal presented in [62] are methods for the identification of dominant modes only, and there is no dynamic ROM constructed based on the determined modes.

In this paper, we close this gap by introducing a framework that allows constructing ROMs based on transformed modes, which can, for instance, be computed by one of the contributions mentioned above. Notably, our framework is not restricted to the case that all modes are transformed by the same coordinate transformation, but we consider the general ansatz (1.4), which allows incorporating several transformation operators.

Adaptive basis methods

The second class of MOR methods for transport dominated systems is based on an online adaptation of the modes during the simulation of the ROM. For instance, in [16], the authors combine the reduced basis method with a segmentation of the time interval, *i.e.*, they subdivide the total time interval of interest into subintervals and then compute a reduced basis on each of these subintervals. The authors of [16] propose an adaptive segmentation such that a prescribed error tolerance and a maximum number of basis functions per time interval is not exceeded. Depending on the time interval segmentation, this may lead to a considerable reduction of the number of required basis functions in each time interval in comparison to a non-segmented approach. In the online phase, they simulate the ROM subsequently on each subinterval. When the interface between two adjacent subintervals is reached, the initial value for the new subinterval is computed by an orthogonal projection of the current approximation of the full order state onto the span of the basis functions of the new subinterval. Thus, the ROM is a switched system, where the switching condition depends solely on time. Furthermore, they provide an *a posteriori* error estimator, which they use to drive a greedy parameter sampling and the adaptive time partitioning.

Differently from [16], the authors in [18] propose a scheme that adapts the reduced basis only in the online phase and only if the error estimator returns values which are too high or too low. In the offline phase, they define a tree structure that represents hierarchical orthogonal decompositions of the underlying vector space of the FOM. Based on this tree structure, they are able to adaptively refine the basis if the error of the ROM is too high or to compress the basis if the error is smaller than a prescribed error threshold. This refinement or compression of the basis corresponds to moving downwards or upwards in the tree structure, respectively. As in [16], the obtained ROM is a switched system, but in contrast to [16], the switching condition is state-dependent since the error estimators are based on the current state of the ROM. Thus, the switching times are *a priori* not known. The work in [18] is based on the ideas presented in [12] and extends them towards more general refinement trees and a more general and efficient basis compression scheme. Another online-adaptive scheme is proposed in [55], where the basis functions are regularly modified *via* a low-rank update, where the number of basis functions remains constant in contrast to [18].

Remark 2.1. Model reduction for switched systems is an active research area for itself, with several contributions within the last years. For an overview we mention [2, 25, 66, 69] and the references therein.

In [22], the authors present an approach that is also based on a reduced basis, which is adapted as time evolves. However, in contrast to the works mentioned above, the rules for updating the basis are more problem-specific. In concreto, they propose to use the eigenfunctions of a linear Schrödinger operator associated with the initial value of the FOM as basis functions. Then, the time evolution of the basis functions is performed in such a way that the basis functions remain eigenfunctions of a linear Schrödinger operator associated with the time-dependent FOM. Consequently, they obtain an additional evolution equation for the basis functions. In contrast to the works mentioned above, the ROM is not a switched system with time-discrete changes in the basis functions, but instead they obtain a time-continuous equation for the evolution of the basis functions.

In the last paragraphs, we have discussed model reduction methods, which are based on time-dependent basis functions, while there are different ways of establishing this time dependency. Formally, also the reference frame methods fit into this rather general class of approaches since applying time-dependent coordinate transformations is a particular way of inducing a time dependency into the modes. Nevertheless, we distinguish between these two classes since the main ideas are somewhat different. The reference frame methods are motivated by physics and incorporate coordinate transformations to model the advection behavior present in the dynamics. On the other hand, the adaptive basis methods propose rather general ways of updating the basis functions without attempting to model the advection explicitly.

Nonlinear projection methods

The idea of the third class of methods is to approximate the solution *via* a general nonlinear approximation ansatz. In [43], the authors use an autoencoder, which is a type of artificial neural network, to obtain a low-dimensional description of the FOM solution. Based on the snapshot data, a decoder and an encoder mapping are learned, where the decoder is a mapping from the reduced state space to the full state space and the encoder *vice versa*. Especially, the lifting of the reduced state to an approximation of the full order state is performed by the decoder mapping, which thus describes the approximation ansatz. The projection of the FOM is carried out by substituting the approximation ansatz into the FOM and then constructing the ROM *via* minimization of the residual. They propose two different approaches: One is based on minimizing the residual for the time-continuous FOM while the other one considers the FOM in the time-discrete setting. The idea to use autoencoders for the purpose of model order reduction has been previously presented in [32,39].

Remark 2.2. For several hyperbolic problems, it is possible to derive an equivalent system of delay equations, see, for instance [8, 15, 44]. Although these delay equations are still infinite-dimensional, they are – from a computational perspective – much easier to solve. From a model reduction perspective, this provides a different approach to the approximation of transport dominated phenomena by searching for a surrogate model that captures the transport by including a time-delay. Some first results in this direction are obtained in [20,57,67, 68,70].

3. Notation and preliminaries

Throughout the paper we denote the Bochner space of square integrable functions and locally square integrable functions in the time interval $[0, T]$ with values in a Banach space $\mathscr{X}$ by $L^2(0, T; \mathscr{X})$ and $L^2_{\mathrm{loc}}(-\infty, \infty; \mathscr{X})$, respectively. If $\mathscr{X} = \mathbb{R}$, we simply write $L^2(0, T)$. The Sobolev space of square integrable functions in a domain $\Omega \subseteq \mathbb{R}^d$ with square integrable derivative is denoted by $H^1(\Omega)$. If additionally periodic boundary conditions are assumed, we write $H^1_{\mathrm{per}}(\Omega)$. For an operator $\mathcal{A}$ we denote its adjoint operator by $\mathcal{A}^\star$. The induced operator norm of $\mathcal{A}$ is denoted by $|||\mathcal{A}|||$ and the standard Euclidean norm on $\mathbb{R}^n$ is written as $\|\cdot\|_2$. The transpose of a matrix $M = [m_{ij}] \in \mathbb{R}^{n\times m}$ is denoted by M^T and δ_{ij} is the Kronecker delta.

3.1. Proper orthogonal decomposition

To motivate what follows, we consider one classical way of constructing the MOR projection basis $(\varphi_1, \ldots, \varphi_r)$, namely POD, see for instance [29, 35] and the references therein. Our starting point for the discussion is the approximation (1.2). Suppose that we have access to a solution trajectory z of (1.1), for instance *via* a numerical simulation. For a given dimension $r < \dim(\mathscr{X})$, the approximation error in (1.2) is minimized *via* the optimization problem

$$\begin{cases} \min \dfrac{1}{2}\displaystyle\int_0^T \left\| z(t) - \sum_{j=1}^r \tilde{z}_j(t)\varphi_j \right\|_{\mathscr{X}}^2 \mathrm{d}t \\ \text{s.t. } \{\varphi_j\}_{j=1}^r \subseteq \mathscr{Y} \text{ and } \langle \varphi_i, \varphi_j\rangle_{\mathscr{X}} = \delta_{ij},\ i,j = 1,\ldots,r. \end{cases} \tag{3.1}$$

For later referencing, we call the solution $\{\varphi_j\}_{j=1}^r$ of (3.1) the *dominant modes* for the solution trajectory z. Notice that the optimization problem (3.1) depends on the coefficient functions $\tilde{z}_j$ for $j = 1, \ldots, r$ and in principle one should also minimize over the $\tilde{z}_j$'s. It is however well-known that the best approximation in a subspace is given by the orthogonal projection onto this subspace and hence we have $\tilde{z}_j(t) = \langle z(t), \varphi_j\rangle_{\mathscr{X}}$ for $j = 1, \ldots, r$. Thus, we are interested in the optimization problem

$$\begin{cases} \min \dfrac{1}{2}\displaystyle\int_0^T \left\| z(t) - \sum_{j=1}^r \langle z(t), \varphi_j\rangle_{\mathscr{X}}\, \varphi_j \right\|_{\mathscr{X}}^2 \mathrm{d}t \\ \text{s.t. } \{\varphi_j\}_{j=1}^r \subseteq \mathscr{Y} \text{ and } \langle \varphi_i, \varphi_j\rangle_{\mathscr{X}} = \delta_{ij},\ i,j = 1,\ldots,r. \end{cases} \tag{3.2}$$

Following [29], the optimization problem can be solved by computing the eigenvalues of the nonnegative, self-adjoint compact operator

$$\mathcal{R}\colon \mathscr{X} \to \mathscr{X}, \qquad \mathcal{R}\varphi = \int_0^T \langle z(t), \varphi\rangle_{\mathscr{X}}\, z(t)\,\mathrm{d}t. \tag{3.3}$$

Observe that $z \in L^2(0,T;\mathscr{Y})$ implies (see also [29], Lem. 1.24) $\mathcal{R}\varphi \in \mathscr{Y}$ for all $\varphi \in \mathscr{X}$. In this case, a solution of (3.2) can be obtained as follows.

Theorem 3.1 ([29], Thm. 1.15)**.** *Let $\mathscr{X}$ be a separable real Hilbert space and suppose that $z \in L^2(0,T;\mathscr{X})$ is given. Then there exist nonnegative eigenvalues λ_i and associated orthonormal eigenvectors $\varphi_i \in \mathscr{X}$ for $i \in \mathcal{I}$ with*

$$\mathcal{I} := \begin{cases} \{1,\ldots,\dim(\mathscr{X})\}, & \text{if } \dim(\mathscr{X}) < \infty,\\ \mathbb{N}, & \text{otherwise,}\end{cases}$$

satisfying

$$\mathcal{R}\varphi_i = \lambda_i\varphi_i \qquad \text{for } i \in \mathcal{I}, \tag{3.4}$$

$$\textit{and} \qquad \lambda_i \geq \lambda_{i+1} \geq 0 \qquad \text{for } i < \dim(\mathscr{X}), \tag{3.5}$$

with $\mathcal{R}$ as defined in (3.3). If in addition $z \in L^2(0,T;\mathscr{Y})$ with a dense subspace $\mathscr{Y} \subseteq \mathscr{X}$, then for any $r \in \mathcal{I}$ with $\lambda_r > 0$, the set of the r leading eigenvectors $\{\varphi_i\}_{i=1}^r$ is a solution of (3.2).

Remark 3.2. One can show (see for instance [80]) that (3.4) equals the first-order necessary optimality condition for the minimization problem (3.2). In particular, the minimizer of (3.2) is unique if, and only if, the first $r+1$ eigenvalues of $\mathcal{R}$ are simple, that is $\lambda_1 > \ldots > \lambda_r > \lambda_{r+1}$.

To illustrate Theorem 3.1 we consider the following simple example with an advection equation, see also [35, 77].

Example 3.3. The one-dimensional linear advection equation with constant coefficients and periodic boundary condition is given by

$$\begin{cases} \partial_t z(t,\xi) + \partial_\xi z(t,\xi) = 0, & (t,\xi) \in (0,1)\times(0,1),\\ z(t,0) = z(t,1), & t \in (0,1),\\ z(0,\xi) = z_0(\xi), & \xi \in (0,1),\end{cases} \tag{3.6}$$

with given initial value

$$z_0 \in \mathscr{Y} := H^1_{\mathrm{per}}(0,1).$$

For notational convenience, we consider z_0 as an element of $L^2_{\mathrm{loc}}(\mathbb{R})$ *via* periodic continuation. It is well-known that the solution of (3.6) is given by $z(t,\xi) = z_0(\xi - t)$ for $(t,\xi) \in (0,1)\times(0,1)$. To compute the POD basis as in Theorem 3.1 we set $\mathscr{X} = L^2(0,1)$ and observe that the operator in (3.3) is given by

$$(\mathcal{R}\varphi)(\zeta) = \int_0^1 R(\zeta,\xi)\varphi(\xi)\,\mathrm{d}\xi$$

with auto-correlation function $R(\zeta,\xi) = \int_0^1 z(t,\zeta)z(t,\xi)\,\mathrm{d}t = \int_0^1 z_0(t)z_0(t+(\zeta-\xi))\,\mathrm{d}t$. In particular, $R(\zeta,\xi)$ depends only on the distance between ζ and ξ. Since R is periodic in the first argument, and thus also in the distance $\xi - \zeta$, we can consider its representation in the Fourier basis, *i.e.*,

$$R(\zeta,\xi) = \sum_{k=-\infty}^{\infty} c_k \exp\left(2\pi\imath k(\xi-\zeta)\right) \qquad \text{with } c_k \in \mathbb{R}.$$

Using this expression, we observe that the eigenvectors of $\mathcal{R}$, which correspond to the solution of the minimization problem (3.2) for the advection equation (3.6), are given by the functions $\varphi_i(\xi) = \exp(2\pi\imath i\xi)$. Note that the eigenvectors of $\mathcal{R}$ are independent of the initial value, which completely describes the solution of (3.6) and the initial value only influences the ordering of the dominant modes.

In practice, the solution of the FOM (1.1) usually depends on an additional variable μ, *i.e.*, $z(t) = z(t;\mu)$, which may represent physical or geometry parameters or a control function. In any case, the dominant modes should reflect the dynamics for a large range of the additional variable μ, which we hereafter refer to as *parameter*. The current state of the art is to sample the parameter space, *i.e.*, to pick parameters μ_j for $j = 1,\ldots,M$ and solve (1.1) for each parameter value μ_j. The dominant modes can then be computed by solving (3.2) simultaneously for all parameters [29], which is equivalent to concatenating the solution trajectories for different parameters and solve (3.2) based on the concatenated solution. An alternative approach is to determine dominant modes for each parameter and then combine the different dominant modes. In general, the first approach provides a smaller set of dominant modes, while the second approach allows to pick the parameters iteratively, for instance by a greedy selection procedure. Let us mention that convergence rates for the greedy approach can be related to the decay of the Kolmogorov n-widths, see for instance [30, 78].

3.2. Galerkin projection and offline/online decomposition

Having identified a set of dominant modes $\{\varphi_j\}_{j=1}^r$ we substitute the Galerkin ansatz (1.2) into (1.1) and obtain at time $t > 0$ the residual

$$\sum_{i=1}^r \dot{\tilde{z}}_i(t)\varphi_i - \mathcal{F}\left(\sum_{i=1}^r \tilde{z}_i(t)\varphi_i\right).$$

Note that the choice of the modes $\{\varphi_j\}_{j=1}^r$ fixes the initial condition and thus the coefficient functions $\tilde{z}_i$ at time $t = 0$. Thus, if we want to minimize the norm of the residual we can do so only by optimizing over the slope of the coefficient functions. The concept of minimizing the norm of the residual with respect to the slope of the coefficient functions is called *continuous optimality* in [13]. We deduce that the partial derivative with respect to $\dot{\tilde{z}}_\ell$ is given by

$$\frac{\partial}{\partial \dot{\tilde{z}}_\ell}\left\|\sum_{i=1}^r \dot{\tilde{z}}_i\varphi_i - \mathcal{F}\left(\sum_{i=1}^r \tilde{z}_i\varphi_i\right)\right\|_{\mathscr{X}}^2 = 2\sum_{i=1}^r \dot{\tilde{z}}_i\left\langle \varphi_\ell,\varphi_i\right\rangle_{\mathscr{X}} - 2\left\langle \varphi_\ell, \mathcal{F}\left(\sum_{i=1}^r \tilde{z}_i\varphi_i\right)\right\rangle_{\mathscr{X}}.$$

As a consequence the ROM (1.3), or equivalently

$$M\dot{\tilde{z}}(t) = \widetilde{F}(\tilde{z}(t)), \quad \tilde{z}(0) = \widetilde{z}_0, \quad t \in (0,T) \tag{3.7}$$

with $M := [\langle\varphi_i,\varphi_j\rangle_{\mathscr{X}}]_{i,j=1}^r \in \mathbb{R}^{r\times r}$, reduced state $\tilde{z}(t) \in \mathbb{R}^r$, $\widetilde{F}\colon \mathbb{R}^r \to \mathbb{R}^r$, and initial value $\widetilde{z}_0 \in \mathbb{R}^r$ defined as

$$\tilde{z}(t) := \begin{bmatrix}\tilde{z}_1(t)\\ \vdots\\ \tilde{z}_r(t)\end{bmatrix}, \qquad \widetilde{F}(\tilde{z}) := \begin{bmatrix}\langle\varphi_1,\mathcal{F}\left(\sum_{i=1}^r \tilde{z}_i\varphi_i\right)\rangle_{\mathscr{X}}\\ \vdots\\ \langle\varphi_r,\mathcal{F}\left(\sum_{i=1}^r \tilde{z}_i\varphi_i\right)\rangle_{\mathscr{X}}\end{bmatrix}, \qquad \text{and} \qquad \widetilde{z}_0 := \begin{bmatrix}\langle\varphi_1,z_0\rangle_{\mathscr{X}}\\ \vdots\\ \langle\varphi_r,z_0\rangle_{\mathscr{X}}\end{bmatrix},$$

satisfies the following result, see also [13].

Lemma 3.4 (Continuous optimality). *The ROM (3.7) is continuously optimal in the sense that if $\tilde{z}$ is a solution of (3.7), then for each $t > 0$, the velocity $\dot{\tilde{z}}(t)$ is the unique minimizer of*

$$\min_{\alpha=[\alpha_i]_{i=1}^r}\left\|\sum_{i=1}^r \alpha_i\varphi_i - \mathcal{F}\left(\sum_{i=1}^r \tilde{z}_i(t)\varphi_i\right)\right\|_{\mathscr{X}}.$$

Although the ROM is formally stated in $\mathbb{R}^r$, it still depends on the evaluation of $\mathcal{F}$ in the original space $\mathscr{Y}$ and thus it might still be computationally intractable to solve (3.7) efficiently. However, in many cases we can precompute all quantities that depend on $\mathscr{Y}$, which allows a fast evaluation of (3.7). This process is called *efficient offline/online decomposition* in the MOR literature. For instance for a linear operator $\mathcal{A}\colon \mathscr{Y} \to \mathscr{X}$ we have

$$\begin{bmatrix} \langle \varphi_1, \mathcal{A}(\sum_{i=1}^r \tilde{z}_i\varphi_i)\rangle_{\mathscr{X}} \\ \vdots \\ \langle \varphi_r, \mathcal{A}(\sum_{i=1}^r \tilde{z}_i\varphi_i)\rangle_{\mathscr{X}} \end{bmatrix} = \begin{bmatrix} \langle \varphi_1, \mathcal{A}\varphi_1\rangle_{\mathscr{X}} & \dots & \langle \varphi_1, \mathcal{A}\varphi_r\rangle_{\mathscr{X}} \\ \vdots & & \vdots \\ \langle \varphi_r, \mathcal{A}\varphi_1\rangle_{\mathscr{X}} & \dots & \langle \varphi_r, \mathcal{A}\varphi_r\rangle_{\mathscr{X}} \end{bmatrix} \begin{bmatrix} \tilde{z}_1 \\ \vdots \\ \tilde{z}_r \end{bmatrix} =: \widetilde{\mathcal{A}}\tilde{z}$$

with $\widetilde{\mathcal{A}} \in \mathbb{R}^{r\times r}$. Notice that the precomputation is possible for all polynomial structures, see for instance [41] and the references therein. Indeed, let us present the details for a quadratic polynomial, exemplified by a linear operator $\widetilde{\mathcal{N}}\colon (\mathscr{Y} \otimes \mathscr{Y}) \to \mathscr{X}$, where $\mathscr{Y} \otimes \mathscr{Y}$ denotes the tensor product of $\mathscr{Y}$ with itself. In this case, we have

$$\mathcal{N}\left(\left(\sum_{i=1}^r \tilde{z}_i\varphi_i\right) \otimes \left(\sum_{\ell=1}^r \tilde{z}_\ell\varphi_\ell\right)\right) = \sum_{i=1}^r\sum_{\ell=1}^r \tilde{z}_i\tilde{z}_\ell\mathcal{N}(\varphi_i \otimes \varphi_\ell).$$

Defining

$$\widetilde{\mathcal{N}} := \begin{bmatrix} \langle \varphi_1, \mathcal{N}(\varphi_1 \otimes \varphi_1)\rangle_{\mathscr{X}} & \langle \varphi_1, \mathcal{N}(\varphi_1 \otimes \varphi_2)\rangle_{\mathscr{X}} & \dots & \langle \varphi_1, \mathcal{N}(\varphi_r \otimes \varphi_r)\rangle_{\mathscr{X}} \\ \vdots & \vdots & & \vdots \\ \langle \varphi_r, \mathcal{N}(\varphi_1 \otimes \varphi_1)\rangle_{\mathscr{X}} & \langle \varphi_r, \mathcal{N}(\varphi_1 \otimes \varphi_2)\rangle_{\mathscr{X}} & \dots & \langle \varphi_r, \mathcal{N}(\varphi_r \otimes \varphi_r)\rangle_{\mathscr{X}} \end{bmatrix} \in \mathbb{R}^{r\times r^2}$$

implies

$$\begin{bmatrix} \langle \varphi_1, \mathcal{N}\left((\sum_{i=1}^r \tilde{z}_i\varphi_i) \otimes (\sum_{\ell=1}^r \tilde{z}_\ell\varphi_\ell)\right)\rangle_{\mathscr{X}} \\ \vdots \\ \langle \varphi_r, \mathcal{N}\left((\sum_{i=1}^r \tilde{z}_i\varphi_i) \otimes (\sum_{\ell=1}^r \tilde{z}_\ell\varphi_\ell)\right)\rangle_{\mathscr{X}} \end{bmatrix} = \widetilde{\mathcal{N}}(\tilde{z} \otimes \tilde{z}),$$

where $\widetilde{\mathcal{N}}$ can be precomputed in the offline phase and, thus, the right-hand side can be computed independently of the original space $\mathscr{Y}$. Recall that, for each time instance, $\tilde{z}$ is an r-dimensional real vector and thus the tensor product $\tilde{z}\otimes\tilde{z}$ reduces (up to an isomorphism) to the standard Kronecker product. The procedure extends directly for a general polynomial but for the sake of notation we omit the details.

Remark 3.5. If the Hilbert space $\mathscr{X}$ is finite-dimensional it is isomorphic to $\mathbb{R}^N$ (with standard inner product) such that without loss of generality we may assume $\mathscr{X} = \mathbb{R}^N$. In this case, the basis vectors φ_j $(j = 1, \dots, r)$ form a matrix

$$\Phi := \begin{bmatrix}\varphi_1 & \dots & \varphi_r\end{bmatrix} \in \mathbb{R}^{N\times r}.$$

In this case, the polynomial operators are associated with matrices, *i.e.*, $\mathcal{A} \in \mathbb{R}^{N\times N}$ and $\mathcal{N} \in \mathbb{R}^{N\times N^2}$ and the reduced analogues are given by

$$\widetilde{\mathcal{A}} = \Phi^T\mathcal{A}\Phi \qquad \text{and} \qquad \widetilde{\mathcal{N}} = \Phi^T\mathcal{N}(\Phi \otimes \Phi),$$

where again $\otimes$ denotes the Kronecker product.

Albeit many nonlinear systems can be rewritten as polynomial systems by introducing additional states [28], it may not be possible to reduce the computational complexity to a satisfactory level with the approach presented above. To remedy this problem a standard approach is to further approximate the nonlinear function, for instance *via* the *empirical interpolation method* (EIM) [3] or the *discrete empirical interpolation method* (DEIM) [14]. Although the extension of these methods to our methodology presented in the forthcoming Section 5 is certainly an interesting aspect we consider this a second step and postpone the extension to a future work.

Remark 3.6 (Parameter separability)**.** If the right-hand side in (1.1) depends on a parameter μ and is separable with respect to this parameter, that is $\mathcal{F}(z, \mu) = \sum_{k=1}^K \theta_k(\mu)\mathcal{F}_k(z)$ with suitable scalar-valued functions θ_k, then this structure is retained in the ROM by setting $\widetilde{F}(\tilde{z}, \mu) = \sum_{k=1}^K \theta_k(\mu)\widetilde{F}_k(\tilde{z})$. This facilitates the efficient usage of the ROM in a many-query context, where the ROM has to be evaluated for many different parameter values.

4. Identification of dominant modes

Similarly as in Section 3.1, we aim for identifying the dominant modes of the system, which capture most of the dynamics. However, instead of considering a linear Galerkin approach as in (1.2), we use the more general ansatz (1.4) here. Thus, similarly to Section 3.1, we assume that we are given $z \in L^2(0,T;\mathscr{X})$, which we want to approximate *via* (1.4), *i.e.*, we want to solve the minimization problem

$$\begin{cases} \min \dfrac{1}{2} \displaystyle\int_0^T \left\| z(t) - \sum_{i=1}^r \tilde{z}_i(t)\mathcal{T}_i\left(p_i(t)\right)\varphi_i \right\|_{\mathscr{X}}^2 \,\mathrm{d}t, \\ \text{s.t. } \varphi \in \mathscr{Y}^r,\ \|\varphi_i\|_{\mathscr{X}} = 1,\ \tilde{z}_i \in L^2(0,T),\ p_i \in L^2(0,T;\mathcal{P}_i) \ \text{ for } i = 1,\ldots,r. \end{cases} \tag{4.1}$$

Here, we assume that the mappings $\mathcal{T}_i\colon \mathcal{P}_i \to \mathcal{B}(\mathscr{X})$ for $i = 1,\ldots,r$ are given and satisfy the following assumption.

Assumption 4.1. *For each $\varphi_i \in \mathscr{Y}$ and each $i \in \{1,\ldots,r\}$, the mappings $\mathcal{T}_i(\cdot)\varphi_i\colon \mathcal{P}_i \to \mathscr{Y}$ are continuous. Moreover, there exists a constant $\overline{c} > 0$ such that*

$$|\!|\!|\mathcal{T}_i(\eta)|\!|\!| \leq \overline{c}, \quad \forall\, \eta \in \mathcal{P}_i \tag{4.2}$$

for $i = 1,\ldots,r$, where $|\!|\!|\cdot|\!|\!|$ denotes the induced operator norm.

Lemma 4.2. *Let the mappings $\mathcal{T}_i$ (i=1,...,r) satisfy Assumption 4.1 and assume that $z \in L^2(0,T;\mathscr{X})$, $\varphi \in \mathscr{Y}^r$, $\tilde{z} \in L^2(0,T;\mathbb{R}^r)$, and $p \in L^2(0,T);\mathcal{P}_1 \times \cdots \times \mathcal{P}_r$ are given. Then the integal in (4.1) is defined.*

Proof. We define for $i = 1,\ldots,r$ the mapping

$$\alpha_i\colon (0,T) \times \mathcal{P}_i \to \mathscr{X}, \qquad (t,\eta_i) \mapsto \tilde{z}_i(t)\mathcal{T}_i(\eta_i)\varphi_i$$

and the associated Nemytskij operator $\mathcal{A}_i(p_i)(t) = \alpha_i(t,p_i(t))$ for almost all $t \in (0,T)$. Assumption 4.1 implies that $\alpha_i(t,\cdot)\colon \mathcal{P}_i \to \mathscr{X}$ is continuous for almost all $t \in (0,T)$. By assumption, $\tilde{z}$ is measurable and thus $\alpha_i(\cdot,\eta_i)$ is measurable for all $\eta_i \in \mathcal{P}_i$. In particular, α_i satisfies the Carathéodory condition and thus $\mathcal{A}_i(p_i)$ is measurable [24]. We conclude the proof by observing

$$\|\mathcal{A}_i(p_i)\|_{L^2(0,T;\mathscr{X})}^2 = \int_0^T |\tilde{z}_i(t)|^2 \|\mathcal{T}_i(p_i(t))\varphi_i\|_{\mathscr{X}}^2 \,\mathrm{d}t \leq \overline{c}^2 \|\tilde{z}_i\|_{L^2(0,T)}^2$$

and thus $\mathcal{A}_i(p_i) \in L^2(0,T;\mathscr{X})$. □

Before we discuss existence of a minimizer of (4.1) let us illustrate the usefulness of the ansatz (1.4) by revisiting the linear advection equation discussed in Example 3.3.

Example 4.3. Recall that the solution z of the linear advection equation in Example 3.3 is given by a shift of the initial condition, *i.e.*, $z(t,\xi) = z_0(\xi - t)$ for all $(t,\xi) \in (0,1) \times (0,1)$. Defining for $\eta \in \mathbb{R}$ the shift operator

$$\mathcal{S}(\eta)\colon L^2(0,1) \to L^2(0,1), \qquad \mathcal{S}(\eta)f := f(\cdot - \eta)$$

via periodic continuation, we observe that the solution of the advection equation can be written as

$$z(t,\xi) = \mathcal{S}(t)z_0(\xi) \qquad \text{for all } (t,\xi) \in (0,1) \times (0,1).$$

Thus, a minimizer of (4.1) is given by the choice $r = 1$, $\mathcal{T}_1 = \mathcal{S}$, $\varphi_1 = z_0/\|z_0\|_{\mathscr{X}}$, $\tilde{z}_1(t) \equiv \|z_0\|_{\mathscr{X}}$, and $p_1(t) = t$ for $t \in (0,1)$. Thus, the solution can be described without approximation error with just one mode when using the ansatz (1.4). Furthermore, while the dominant modes determined *via* POD are independent from the initial condition (*cf.* Example 3.3), here, the dominant mode is given by the (normalized) initial condition itself, which in turn fully describes the solution. Especially, it is possible to construct initial conditions which result in a need of arbitrarily many POD modes to capture the solution, *cf.* [11], while using (1.4) only one mode is needed regardless of which initial condition is chosen.

Let us emphasize that in contrast to the POD minimization problem (3.1) discussed in Section 3.1, we only require the modes to be normalized but not necessarily to form an orthonormal set. The proof of Theorem 3.1 relies heavily on the fact that the modes are orthogonal. Mimicking this proof would require that $\mathcal{T}_i(p_i(t))\varphi_i$ is orthogonal to $\mathcal{T}_j(p_j(t))\varphi_j$ for all $i \neq j$ and all $t \in [0,T]$. The next example highlights that in general, this is not a reasonable assumption. Instead, we only require the modes to be normalized in $\mathscr{X}$.

Example 4.4 (Wave equation)**.** We consider the linear acoustic wave equation in $\Omega := (0,1)$ with periodic boundary conditions for the density ρ and the velocity v given by

$$\begin{cases} \partial_t \rho(t,\xi) + \partial_\xi v(t,\xi) = 0, & (t,\xi) \in (0,1)\times\Omega, \\ \partial_t v(t,\xi) + \partial_\xi \rho(t,\xi) = 0, & (t,\xi) \in (0,1)\times\Omega, \\ \rho(t,0) = \rho(t,1), & t \in (0,1), \\ v(t,0) = v(t,1), & t \in (0,1), \\ \rho(0,\xi) = \rho_0(\xi), & \xi \in \Omega, \\ v(0,\xi) = 0, & \xi \in \Omega, \end{cases} \tag{4.3}$$

with given initial value

$$z_0 = \begin{bmatrix} \rho_0 \\ 0 \end{bmatrix} \in \mathscr{Y} := \left(H^1_{\mathrm{per}}(\Omega)\right)^2.$$

Similar to Example 3.3, we consider z_0 as an element of $L^2_{\mathrm{loc}}\left(\mathbb{R};\mathbb{R}^2\right)$ *via* periodic continuation. The analytic solution can be expressed as

$$\begin{bmatrix} \rho(t,\xi) \\ v(t,\xi) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} q_+(\xi - t) + \begin{bmatrix} 1 \\ -1 \end{bmatrix} q_-(\xi + t), \tag{4.4}$$

where $q_+, q_- \in L^2_{\mathrm{loc}}(\mathbb{R})$ are functions determined *via* the initial value and the boundary conditions, *cf.* [34]. In the case of a homogeneous initial condition for v, the values of q_+ and q_- in Ω are determined *via*

$$q_+(\xi) = q_-(\xi) = \frac{1}{2}\rho_0(\xi), \quad \xi \in \Omega,$$

and are periodically extended to $L^2_{\mathrm{loc}}(\mathbb{R})$. Now let us assume that we are only interested in a low-dimensional approximation of the density which is given by

$$\rho(t,\xi) = \frac{1}{2}\left(\rho_0(\xi - t) + \rho_0(\xi + t)\right) = \frac{1}{2}\left(\mathcal{S}(t)\rho_0(\xi) + \mathcal{S}(-t)\rho_0(\xi)\right) =: z(t,\xi)$$

for $(t,\xi) \in (0,1)\times\Omega$ with the shift operator as defined in Example 4.3. Thus, a minimizer of (4.1) is given by the choice $r = 2$, $\mathcal{T}_1 = \mathcal{T}_2 = \mathcal{S}$, $\varphi_1 = \varphi_2 = \rho_0/\|\rho_0\|_{\mathscr{X}}$, and

$$\tilde{z}_1(t) = \tilde{z}_2(t) \equiv \frac{\|\rho_0\|_{\mathscr{X}}}{2}, \qquad p_1(t) = t, \qquad p_2(t) = -t \quad \text{for } t \in (0,1),$$

and we conclude that the solution can be described without approximation error with just two modes with (1.4). Especially, we observe that the transformed modes $\mathcal{T}_1(p_1(t))\varphi_1$ and $\mathcal{T}_2(p_2(t))\varphi_2$ become linearly dependent for $t = 0$ and $t = 1$. Thus, even a minimizer of the cost functional may lead to transformed modes which become linearly dependent. This observation indicates that it is not reasonable to enforce orthogonality of the transformed modes in contrast to the POD minimization problem addressed in Section 3.1. In this example, we may still obtain linearly independent modes as long as $\rho_0 \neq 0$, by adding the velocity data, *i.e.*, by considering $z = [\rho \ \ v]^T$ instead of $z = \rho$.

By assumption, $\mathscr{Y}$ is a dense subspace of $\mathscr{X}$. Since we only require the modes to be normalized in $\mathscr{X}$, it is not clear that (4.1) possesses a minimizer in $\mathscr{Y}^r$. Instead, we assume that $\mathscr{Y}$ itself is a reflexive Banach space with norm $\|\cdot\|_{\mathscr{Y}}$. Moreover, we assume that $\mathscr{Y}$ is compactly embedded in $\mathscr{X}$ (*cf.* [81], Def. 21.13) and propose to restrict the admissible set of (4.1) by imposing a bound on the modes with respect to $\|\cdot\|_{\mathscr{Y}}$. Another drawback of not enforcing that the modes are linearly independent is that the coefficient functions $\tilde{z}_i$ may become unbounded. To prevent this, we further restrict the admissible set by imposing a bound on $\tilde{z}$ in the L^2 norm. Finally, we simplify the minimization problem (4.1) by assuming that the paths are known *a priori* or have been determined in a pre-processing step, see for instance [47,61,62,71]. In summary, we assume that the cost functional is defined as

$$J\colon L^2(0,T;\mathbb{R}^r)\times\mathscr{X}^r\to\mathbb{R},\qquad (\tilde{z},\varphi)\mapsto\frac{1}{2}\int_0^T\left\|z(t)-\sum_{i=1}^r\tilde{z}_i(t)\mathcal{T}_i(p_i(t))\varphi_i\right\|_{\mathscr{X}}^2\mathrm{d}t,\tag{4.5}$$

where we use the notation $\tilde{z}=(\tilde{z}_1,\ldots,\tilde{z}_r)$ and $\varphi=(\varphi_1,\ldots,\varphi_r)$. The admissible set is given by

$$\mathcal{A}:=\left\{(\tilde{z},\varphi)\in L^2(0,T;\mathbb{R}^r)\times\mathscr{X}^r\;\middle|\;\begin{array}{c}\varphi\in\mathscr{Y}^r,\max\{\|\varphi_i\|_{\mathscr{Y}},\|\tilde{z}_i\|_{L^2(0,T;\mathbb{R})}\}\leq C,\\ \text{and }\|\varphi_i\|_{\mathscr{X}}=1\text{ for }i=1,\ldots,r\end{array}\right\}\tag{4.6}$$

with given constant $C>0$ that is large enough. Instead of the minimization problem (4.1) we thus consider

$$\min J(\tilde{z},\varphi)\qquad\text{s.t. }(\tilde{z},\varphi)\in\mathcal{A}.\tag{4.7}$$

Example 4.5. For a bounded Lipschitz domain $\Omega\subseteq\mathbb{R}^d$ set $\mathscr{X}=L^2(\Omega)$ and $\mathscr{Y}=H^1(\Omega)$. Then $\mathscr{Y}$ is compactly embedded in $\mathscr{X}$, see for instance Theorem 9.16 of [9] or Theorem 21.A of [81]. Note that in this case, we can replace $\|\varphi_i\|_{\mathscr{Y}}\leq C$ by $\|\partial_\xi\varphi_i\|_{\mathscr{X}}\leq C$ in (4.6).

Theorem 4.6. *Let $(\mathscr{Y},\|\cdot\|_{\mathscr{Y}})$ be a reflexive Banach space which is compactly embedded in $\mathscr{X}$. Moreover, assume that paths $p_i\in L^2(0,T;\mathcal{P}_i)$ are given and the transformation operators satisfy Assumption 4.1. Then (4.7) has a solution.*

Proof. Let $J^\star\geq 0$ denote the infimum of J over $\mathcal{A}$ and let $(\tilde{z}^k,\varphi^k)$ denote a sequence in $\mathcal{A}$ with $\lim_{k\to\infty}J(\tilde{z}^k,\varphi^k)=J^\star$. Since $(\tilde{z}^k,\varphi^k)$ is bounded in $L^2(0,T,\mathbb{R}^r)\times\mathscr{Y}^r$, the Eberlein–Šmuljian theorem ([81], Thm. 21.D) ensures that $(\tilde{z}^k,\varphi^k)$ possesses a weakly convergent subsequence $(\tilde{z}^{k_n},\varphi^{k_n})$ with limit $(\tilde{z}^\star,\varphi^\star)$, *i.e.*, $(\tilde{z}^{k_n},\varphi^{k_n})\rightharpoonup(\tilde{z}^\star,\varphi^\star)$ in $L^2(0,T;\mathbb{R}^r)\times\mathscr{Y}^r$ for $n\to\infty$. Since $\mathscr{Y}$ is compactly embedded into $\mathscr{X}$ we conclude strong convergence in $\mathscr{X}$ ([81], Prop. 21.35), *i.e.*, $\varphi^{k_n}\to\varphi^\star$ in $\mathscr{X}^r$. This immediately implies $\|\varphi_i^\star\|_{\mathscr{X}}=1$ for $i=1,\ldots,r$. Define the bilinear mapping

$$\beta\colon L^2(0,T;\mathbb{R}^r)\times\mathscr{X}^r\to L^2(0,T;\mathscr{X}),\qquad(\tilde{z},\varphi)\mapsto\sum_{i=1}^r\tilde{z}_i(\cdot)\mathcal{T}_i(p_i(\cdot))\varphi_i.$$

For $f\in L^2(0,T;\mathscr{X})$, $\tilde{z}\in L^2(0,T;\mathbb{R}^r)$, and $\varphi\in\mathscr{X}^r$ we compute

$$\langle f,\beta(\tilde{z},\varphi)\rangle_{L^2(0,T;\mathscr{X})}=\sum_{i=1}^r\int_0^T\tilde{z}_i(t)\,\langle f(t),\mathcal{T}_i(p_i(t))\varphi_i\rangle_{\mathscr{X}}\,\mathrm{d}t.$$

Since $\varphi_i^{k_n}\to\varphi_i^\star$ for $n\to\infty$ implies $\left\langle f(t),\mathcal{T}_i(p_i(t))\varphi_i^{k_n}\right\rangle_{\mathscr{X}}\to\langle f(t),\mathcal{T}_i(p_i(t))\varphi_i^\star\rangle_{\mathscr{X}}$ for $n\to\infty$ and almost all $t\in(0,T)$, we use Proposition 21.23(j) of [81] to infer $\beta(\tilde{z}^{k_n},\varphi^{k_n})\rightharpoonup\beta(\tilde{z}^\star,\varphi^\star)$. The claim now follows from the fact that the norm is weakly sequentially lower semi-continuous ([81], Prop. 21.23(c)) and thus

$$J^\star\leq J(\tilde{z}^\star,\varphi^\star)\leq\liminf_{n\to\infty}J(\tilde{z}^{k_n},\varphi^{k_n})=\lim_{k\to\infty}J(\tilde{z}^k,\varphi^k)=J^\star.$$

□

Remark 4.7. Instead of restricting the admissible set in (4.6) to bounded coefficients and modes, we can alternatively regularize the cost functional. In more detail, we consider

$$\begin{cases} \min J(\tilde{z},\varphi) + \dfrac{\gamma_1}{2}\|\tilde{z}\|^2_{L^2(0,T;\mathbb{R}^r)} + \dfrac{\gamma_2}{2}\|\varphi\|^2_{\mathscr{Y}^r} \\ \text{s.t. } \varphi \in \mathscr{Y}^r \text{ and } \|\varphi_i\|_{\mathscr{X}} = 1 \text{ for } i = 1,\dots,r \end{cases} \tag{4.8}$$

with given regularization parameters $\gamma_1, \gamma_2 > 0$ instead of (4.7). Note that $\gamma_1, \gamma_2 > 0$ implies that a minimizing sequence $(\tilde{z}^k, \varphi^k)$ is bounded and thus one can show existence of a minimizer of (4.8) as in the proof of Theorem 4.6. In general, we expect that (4.8) is favorable compared to (4.7) from a numerical point of view. This is subject to further investigation.

Let us emphasize that even in the case that all transformation operators are given by the identity, it is not clear that the minimizer of (4.7) is unique, see Remark 3.2. In particular the uniqueness depends on the data z and thus without further restrictions on z we cannot expect to establish that the minimizer of (4.7) is unique.

In order to numerically solve the minimization problem (4.7), it needs to be discretized in space and time. In [71], the authors present a method which first discretizes the problem and afterwards solves the fully discrete minimization problem numerically. Related approaches are presented in [60–62], where the transformed modes are also identified based on fully discrete problems. These approaches have in common that the problem to be solved is already formulated based on a given discretization in space and time.

In the case that all modes are transformed by the same operator, *i.e.*, J is given by

$$J(\tilde{z},\varphi) := \frac{1}{2}\int_0^T \left\| z(t) - \mathcal{T}(p(t))\sum_{i=1}^r \tilde{z}_i(t)\varphi_i \right\|^2_{\mathscr{X}} \mathrm{d}t, \tag{4.9}$$

we can use the following observation to compute a minimizer of (4.7). For the special case that the transformation $\mathcal{T}$ is given by the shift operator, this was recognized for instance in [11, 61, 71].

Theorem 4.8. *For given data $z \in L^2(0,T;\mathscr{Y})$ and a given path $p \in L^2(0,T;\mathcal{P})$, consider the minimization problem (4.7) with J as defined in (4.9). Suppose that the operator $\mathcal{T}$ is isometric and satisfies Assumption 4.1. Let $\varphi^\star := (\varphi_1^\star,\dots,\varphi_r^\star)$ denote a solution of the POD minimization problem (3.2) with transformed data $\mathcal{T}^*(p)z$ with corresponding eigenvalues $\lambda_1 \geq \ldots \geq \lambda_r > 0$ as defined in Theorem 3.1. Define $\tilde{z}^\star = (\tilde{z}_1^\star,\dots,\tilde{z}_r^\star)$ via $\tilde{z}_i^\star := \langle z, \varphi_i^\star\rangle_{\mathscr{X}}$ for $i = 1,\dots,r$. If C in (4.6) satisfies*

$$\max\left\{ \frac{1}{2\lambda_r}\left(\|z\|^2_{L^2(0,T,\mathscr{X})} + \|z\|^2_{L^2(0,T,\mathscr{Y})}\right), \|z\|_{L^2(0,T,\mathscr{X})} \right\} \leq C,$$

then $(\tilde{z}^\star,\varphi^\star)$ is a minimizer of (4.7).

Proof. Since $\mathcal{T}$ is isometric, we have

$$J(\tilde{z},\varphi) = \frac{1}{2}\int_0^T \left\| \mathcal{T}^*(p(t))z(t) - \sum_{i=1}^r \tilde{z}_i(t)\varphi_i \right\|^2_{\mathscr{X}} \mathrm{d}t.$$

It is easy to see that we can substitute the condition $\|\varphi_i\|_{\mathscr{X}} = 1$ for $i = 1,\dots,r$ in the admissible set (4.6) by the condition $\langle\varphi_i,\varphi_j\rangle_{\mathscr{X}} = \delta_{ij}$ for $i,j = 1,\dots,r$ without changing the minimum. It thus remains to show that $(\tilde{z}^\star,\varphi^\star)$ is an element of the admissible set defined in (4.6). We immediately obtain

$$\|\tilde{z}_i^\star\|^2_{L^2(0,T)} \leq \int_0^T \|z(t)\|^2_{\mathscr{X}}\|\varphi_i^\star\|^2_{\mathscr{X}}\,\mathrm{d}t = \|z\|^2_{L^2(0,T;\mathscr{X})} \leq C.$$

For the estimate of the modes we use the operator $\mathcal{R}$ defined in (3.3) and Young's inequality to obtain

$$\begin{aligned}\|\varphi_i^\star\|_{\mathscr{Y}} &= \frac{1}{\lambda_i}\|\mathcal{R}\varphi_i^\star\|_{\mathscr{Y}} \leq \frac{1}{\lambda_i}\int_0^T |\tilde{z}_i^\star(t)|\,\|z(t)\|_{\mathscr{Y}}\,\mathrm{d}t \leq \frac{1}{2\lambda_i}\left(\|\tilde{z}_i^\star\|_{L^2(0,T)}^2 + \|z\|_{L^2(0,T;\mathscr{Y})}^2\right)\\ &\leq \frac{1}{2\lambda_i}\left(\|z\|_{L^2(0,T;\mathscr{X})}^2 + \|z\|_{L^2(0,T;\mathscr{Y})}^2\right) \leq C.\end{aligned}$$

□

5. Reduced order model with transformed modes

Suppose now that we are given suitable transformation operators $\mathcal{T}_i$ and have identified a set of dominant modes φ_i, for instance *via* the procedure described in Section 4. Then we are able to construct a ROM for (1.1) *via* Galerkin projection, *i.e.*, by substituting the approximation (1.4) in (1.1) and formally project the resulting equations onto the time-dependent approximation space

$$\operatorname{span}\{\mathcal{T}_j(p_j(t))\varphi_j \mid j = 1,\ldots,r\}. \tag{5.1}$$

Since the abstract differential equation (1.1) involves a differentiation with respect to time, we have to assume that the transformation operators are continuously differentiable. Indeed, we only require that the transformation applied to the respective mode is continuously differentiable and thus make the following assumption.

Assumption 5.1. *The mappings $\mathcal{T}_i(\cdot)\varphi_i\colon \mathcal{P}_i \to \mathscr{Y}$ are continuously differentiable.*

Example 5.2. The shift operator $\mathcal{T}(p)z = z(\cdot - p)$ from Example 4.3 with periodic embedding into $L^2(0,1)$ is a strongly continuous semigroup ([17], Example I.5.4). In particular, semigroup theory implies that $\mathcal{T}(\cdot)\varphi$ is continuously differentiable for all

$$\varphi \in D(\mathcal{A}) = H^1_{\mathrm{per}}(0,1)$$

see for instance Chapter 1, Theorem 2.4 of [54]. We conclude that the shift operator satisfies Assumption 5.1, whenever $\mathscr{Y} \subseteq D(\mathcal{A})$.

By abuse of notation we denote the derivative of $\mathcal{T}_i(\cdot)\varphi_i$ at $p_i \in \mathcal{P}_i$ by $[\mathcal{T}_i'(p_i)\varphi_i] \in \mathcal{L}(\mathcal{P}_i, \mathscr{X})$. Recall that for the sake of notation we assume $\mathcal{P}_i = \mathbb{R}^{q_i}$ and thus $\mathcal{P} := \mathcal{P}_1 \times \cdots \times \mathcal{P}_r = \mathbb{R}^q$ with $q := \sum_{i=1}^r q_i$. The Galerkin projection of (1.1) onto (5.1) is then given by

$$M_{\tilde{z}}(p(t))\dot{\tilde{z}}(t) + N(p(t))D(\tilde{z}(t))\dot{p}(t) = \widetilde{F}_{\tilde{z}}(p(t),\tilde{z}(t)) \tag{5.2}$$

with state and path vectors

$$\tilde{z}(t) := \left[\tilde{z}_1(t) \cdots \tilde{z}_r(t)\right]^T \in \mathbb{R}^r, \qquad p(t) := \left[p_1^T(t) \cdots p_r^T(t)\right]^T \in \mathbb{R}^q, \tag{5.3}$$

mass matrix $M_{\tilde{z}}(p) := [\langle \mathcal{T}_i(p_i)\varphi_i, \mathcal{T}_j(p_j)\varphi_j\rangle_{\mathscr{X}}]_{i,j=1}^r \in \mathbb{R}^{r\times r}$, correlation block matrix

$$N(p) := \left[\left\langle \mathcal{T}_i(p_i)\varphi_i, [\mathcal{T}_j'(p_j)\varphi_j]e_1\right\rangle_{\mathscr{X}} \cdots \left\langle \mathcal{T}_i(p_i)\varphi_i, [\mathcal{T}_j'(p_j)\varphi_j]e_{q_j}\right\rangle_{\mathscr{X}}\right]_{i,j=1}^r \in \mathbb{R}^{r\times q},$$

diagonal matrix $D(\tilde{z}) := \operatorname{diag}(\tilde{z}_1 I_{q_1}, \ldots, \tilde{z}_r I_{q_r}) \in \mathbb{R}^{q\times q}$, and right-hand side

$$\widetilde{F}_{\tilde{z}}(p,\tilde{z}) := \begin{bmatrix} \langle \mathcal{T}_1(p_1)\varphi_1, \mathcal{F}(\sum_{i=1}^r \tilde{z}_i\mathcal{T}_i(p_i)\varphi_i\rangle_{\mathscr{X}} \\ \vdots \\ \langle \mathcal{T}_r(p_r)\varphi_r, \mathcal{F}(\sum_{i=1}^r \tilde{z}_i\mathcal{T}_i(p_i)\varphi_i\rangle_{\mathscr{X}} \end{bmatrix}.$$

Here, e_i denotes the ith unit vector of suitable dimension, such that $(e_1,\ldots,e_{q_i})$ forms a basis of $\mathcal{P}_i = \mathbb{R}^{q_i}$.

As in Section 3.2, the right-hand side $\widetilde{F}_{\tilde{\mathrm{z}}}$ still depends on the original space $\mathscr{Y}$ and requires further simplifications. For instance, if $\mathcal{F}$ is given by a quadratic polynomial of the form $\mathcal{F}(z) = \mathcal{N}(z \otimes z)$ with linear operator $\mathcal{N}\colon \mathscr{Y} \otimes \mathscr{Y} \to \mathscr{X}$ we can (for given p) precompute the quantities $\langle \mathcal{T}_j(p_j)\varphi_j, \mathcal{N}(\mathcal{T}_i(p_i)\varphi_i \otimes \mathcal{T}_\ell(p_\ell)\varphi_\ell)\rangle_{\mathscr{X}}$ for $i, j, \ell = 1, \dots, r$. A further simplification is possible if $\mathcal{F}$ is *equivariant* with respect to the transformation operators $\mathcal{T}_i$ (see the upcoming Assumption 6.1 and the discussion thereafter for further details). Let us mention that parameter separability (see Rem. 3.6) is easily retained in the ROM (5.2).

Remark 5.3. In contrast to POD we cannot ensure that $M_{\tilde{\mathrm{z}}}(p)$ is nonsingular for every $p \in \mathcal{P}$, since some of the modes $\mathcal{T}_i(p_i)\varphi_i$ may become linearly dependent (see Example 4.4). This may happen either at single time points or at a complete time-interval. In the latter case this implies that some of the modes are redundant and can be removed during this interval. In any case, whenever $M_{\tilde{\mathrm{z}}}(p)$ becomes singular we have to restart the computation of the reduced model.

It is clear that (5.2) is not sufficient to compute $\tilde{z}$ and p and hence can be understood as underdetermined DAE, *cf.* [42]. To complete the underdetermined DAE (5.2) we have to add additional equations

$$\Psi(p, \dot{p}, \tilde{z}, \dot{\tilde{z}}) = 0 \tag{5.4}$$

and consider the coupled system (5.2) and (5.4). In the literature, these equations are called *phase conditions* [7,53] or *reconstruction equations* [64,65] and are used to determine the path $p(t)$ along the solution $\tilde{z}$. Although several choices for Ψ are proposed in [7], it is not clear *a priori*, which phase condition benefits the model the most. Since our ROM is obtained *via* Galerkin projection, which satisfies the continuous optimality principle (see Lem. 3.4 and [13, 43]), we propose to construct the phase condition also *via* continuous optimality. More precisely we define

$$\Psi_{\mathrm{Res}}(p, \dot{p}, \tilde{z}, \dot{\tilde{z}}) := D(\tilde{z})^T \left(N(p)^T \dot{\tilde{z}} + M_p(p) D(\tilde{z}) \dot{p} - \widetilde{F}_p(p, \tilde{z}) \right) \tag{5.5}$$

with block mass matrix

$$M_p(p) := \begin{bmatrix} \left\langle [\mathcal{T}_i'(p_i)\varphi_i]e_1, [\mathcal{T}_j'(p_j)\varphi_j]e_1 \right\rangle_{\mathscr{X}} & \cdots & \left\langle [\mathcal{T}_i'(p_i)\varphi_i]e_1, [\mathcal{T}_j'(p_j)\varphi_j]e_{q_j} \right\rangle_{\mathscr{X}} \\ \vdots & & \vdots \\ \left\langle [\mathcal{T}_i'(p_i)\varphi_i]e_{q_i}, [\mathcal{T}_j'(p_j)\varphi_j]e_1 \right\rangle_{\mathscr{X}} & \cdots & \left\langle [\mathcal{T}_i'(p_i)\varphi_i]e_{q_i}, [\mathcal{T}_j'(p_j)\varphi_j]e_{q_j} \right\rangle_{\mathscr{X}} \end{bmatrix}_{i,j=1}^{r} \in \mathbb{R}^{q \times q} \tag{5.6}$$

and reduced right-hand side

$$\widetilde{F}_p(p, \tilde{z}) := \begin{bmatrix} \left\langle [\mathcal{T}_i'(p_i)\varphi_i]e_1, \mathcal{F}\left(\sum_{j=1}^r \tilde{z}_j \mathcal{T}_j(p_j)\varphi_j\right) \right\rangle_{\mathscr{X}} \\ \vdots \\ \left\langle [\mathcal{T}_i'(p_i)\varphi_i]e_{q_i}, \mathcal{F}\left(\sum_{j=1}^r \tilde{z}_j \mathcal{T}_j(p_j)\varphi_j\right) \right\rangle_{\mathscr{X}} \end{bmatrix}_{i=1}^{r} \in \mathbb{R}^{q}.$$

The coupled ROM for the reduced state $\tilde{z}$ and the transformation path p is thus given by

$$M_{\tilde{\mathrm{z}}}(p(t))\dot{\tilde{z}}(t) + N(p(t))D(\tilde{z}(t))\dot{p}(t) = \widetilde{F}_{\tilde{\mathrm{z}}}(p(t), \tilde{z}(t)), \tag{5.7a}$$

$$D(\tilde{z}(t))^T N(p(t))^T \dot{\tilde{z}}(t) + D(\tilde{z}(t))^T M_p(p(t)) D(\tilde{z}(t))\dot{p}(t) = D(\tilde{z}(t))^T \widetilde{F}_p(p(t), \tilde{z}(t)), \tag{5.7b}$$

or equivalently in matrix notation

$$\begin{bmatrix} I_r & 0 \\ 0 & D(\tilde{z})^T \end{bmatrix} \begin{bmatrix} M_{\tilde{\mathrm{z}}}(p) & N(p) \\ N(p)^T & M_p(p) \end{bmatrix} \begin{bmatrix} I_r & 0 \\ 0 & D(\tilde{z}) \end{bmatrix} \begin{bmatrix} \dot{\tilde{z}} \\ \dot{p} \end{bmatrix} = \begin{bmatrix} \widetilde{F}_{\tilde{\mathrm{z}}}(p, \tilde{z}) \\ D(\tilde{z})^T \widetilde{F}_p(p, \tilde{z}) \end{bmatrix}. \tag{5.8}$$

Remark 5.4. The phase condition (5.7b) can be obtained from (1.1) by substituting the ansatz (1.4) and enforcing that the residual is orthogonal to

$$\operatorname{span}\left\{\tilde{z}_i[\mathcal{T}_i'(p_i)\varphi_i]e_j \mid i=1,\ldots,r, j=1,\ldots,q_i\right\}. \tag{5.9}$$

Thus the phase condition (5.7b) can be obtained *via* projection onto the space in (5.9).

For $t>0$ we define $\mathscr{R}\colon (0,T]\times\mathbb{R}^r\times\mathcal{P}\to\mathscr{X}$ *via*

$$\mathscr{R}(t,x,\eta)=\sum_{i=1}^r x_i\mathcal{T}_i(p_i(t))\varphi_i+\sum_{i=1}^r\tilde{z}_i(t)\left[\mathcal{T}_i'(p_i(t))\varphi_i\right]\eta_i-\mathcal{F}\left(\sum_{i=1}^r\tilde{z}_i(t)\mathcal{T}_i(p_i(t))\varphi_i\right)$$

such that the residual that is obtained at time $t>0$ by substituting the ansatz (1.4) into the evolution equation (1.1) is given by $\mathscr{R}(t,\dot{\tilde{z}}(t),\dot{p}(t))$.

Theorem 5.5 (Continuous optimality)**.** *The ROM* (5.7) *is continuously optimal in the sense that if* $(\tilde{z},p)$ *is a solution of* (5.7)*, then for each* $t>0$*, the pair* $(\dot{\tilde{z}}(t),\dot{p}(t))$ *is a minimizer of the norm of* $\mathscr{R}$*, i.e.*

$$\|\mathscr{R}(t,\dot{\tilde{z}}(t),\dot{p}(t))\|_{\mathscr{X}}\leq\|\mathscr{R}(t,x,\eta)\|_{\mathscr{X}}\qquad\text{for all }(x,\eta)\in\mathbb{R}^r\times\mathcal{P}.$$

Proof. Let $t>0$. We first notice that $\|\mathscr{R}(t,x,\eta)\|_{\mathscr{X}}^2$ is convex in (x,η) and hence the first-order necessary optimality condition is also sufficient. The partial derivatives with respect to the first variable are given by

$$\begin{aligned}\frac{\partial}{\partial x_\ell}\|\mathscr{R}(t,x,\eta)\|_{\mathscr{X}}^2&=2\sum_{i=1}^r x_i\left\langle\mathcal{T}_\ell(p_\ell(t))\varphi_\ell,\mathcal{T}_i(p_i(t))\varphi_i\right\rangle_{\mathscr{X}}+2\sum_{i=1}^r\tilde{z}_i(t)\left\langle\mathcal{T}_\ell(p_\ell(t))\varphi_\ell,\left[\mathcal{T}_i'(p_i(t))\varphi_i\right]\eta_i\right\rangle_{\mathscr{X}}\\&\quad-2\left\langle\mathcal{T}_\ell(p_\ell(t))\varphi_\ell,\mathcal{F}\left(\sum_{i=1}^r\tilde{z}_i(t)\mathcal{T}_i(p_i(t))\varphi_i\right)\right\rangle_{\mathscr{X}}\end{aligned}$$

for $\ell=1,\ldots,r$. The partial derivatives with respect to the second variable constitute linear mappings $\frac{\partial}{\partial\eta_\ell}\|\mathscr{R}(t,x,\eta)\|_{\mathscr{X}}^2\colon\mathcal{P}_\ell\to\mathbb{R}$ given by

$$\begin{aligned}\frac{\partial}{\partial\eta_\ell}\|\mathscr{R}(t,x,\eta)\|_{\mathscr{X}}^2(\lambda_\ell)&=2\sum_{i=1}^r\tilde{z}_\ell(t)\tilde{z}_i(t)\left\langle\left[\mathcal{T}_\ell'(p_\ell(t))\varphi_\ell\right]\lambda_\ell,\left[\mathcal{T}_i'(p_i(t))\varphi_i\right]\eta_i\right\rangle_{\mathscr{X}}\\&\quad+2\sum_{i=1}^r x_i\tilde{z}_\ell(t)\left\langle\left[\mathcal{T}_\ell'(p_\ell(t))\varphi_\ell\right]\lambda_\ell,\mathcal{T}_i(p_i(t))\varphi_i\right\rangle_{\mathscr{X}}\\&\quad-2\tilde{z}_\ell(t)\left\langle\left[\mathcal{T}_\ell'(p_\ell(t))\varphi_\ell\right]\lambda_\ell,\mathcal{F}\left(\sum_{i=1}^r\tilde{z}_i(t)\mathcal{T}_i(p_i(t))\varphi_i\right)\right\rangle_{\mathscr{X}}.\end{aligned}$$

Choosing the standard basis $(e_1,\ldots,e_{q_\ell})$ for $\mathcal{P}_\ell$ and using the notation above implies that the first-order necessary condition is given by

$$\begin{aligned}M_{\tilde{z}}(p(t))x+N(p(t))D(\tilde{z}(t))\eta&=\widetilde{F}_{\tilde{z}}(p(t),\tilde{z}(t)),\\D(\tilde{z}(t))^TN(p(t))^Tx+D(\tilde{z}(t))^TM_p(p(t))D(\tilde{z}(t))\eta&=D(\tilde{z}(t))^T\widetilde{F}_p(p(t),\tilde{z}(t)).\end{aligned}\tag{5.10}$$

Since $(\tilde{z},p)$ is a solution of (5.7) we conclude that $(\dot{\tilde{z}},\dot{p})$ is a solution of (5.10) and thus a minimizer of $\|\mathscr{R}(t,x,\eta)\|_{\mathscr{X}}$. □

Remark 5.6. The proof of Theorem 5.5 shows that instead of using the standard basis of $\mathcal{P}_i=\mathbb{R}^{q_i}$ it is possible to use any basis of $\mathcal{P}_i$ for the construction of the ROM (5.7).

If all transformations are chosen constant, then it is easy to see that the phase condition (5.4) is satisfied for any $(\tilde{z}, p)$ and hence the ROM (5.7) may not have a unique solution. We immediately conclude that the minimizer in Theorem 5.5 may not be unique. On the other hand, by virtue of Assumption 5.1, the matrix

$$M(p) := \begin{bmatrix} M_{\tilde{\mathrm{z}}}(p) & N(p) \\ N(p)^T & M_{\mathrm{p}}(p) \end{bmatrix} \in \mathbb{R}^{(r+q)\times(r+q)} \tag{5.11}$$

is continuous with respect to p and thus, if we assume that $M(p(0))$ is nonsingular, then there exists a neighborhood $\mathcal{U} \subseteq \mathbb{R}^q$ around $p(0)$ such that $M(p(0))$ is nonsingular for all $p \in \mathcal{U}$. We conclude that $M(p)^{-1}$ is continuous for all $p \in \mathcal{U}$. As a direct consequence we have proven the following result.

Proposition 5.7. *Let $(\tilde{z}_{(0)}, p_{(0)})$ denote the initial value for (5.7), i.e.,*

$$\tilde{z}(0) = \tilde{z}_{(0)} \qquad \text{and} \qquad p(0) = p_{(0)}. \tag{5.12}$$

Assume that $M(p_{(0)})$ in (5.11) is nonsingular, $e_i^T \tilde{z}_{(0)} \neq 0$ for all $i = 1, \dots, r$, and the transformation operators satisfy Assumption 5.1. If $\mathcal{F}$ is continuous, then there exists $\widetilde{T} > 0$ such that the ROM (5.7) has a (classical) solution on $[0, \widetilde{T})$. If the transformation operators and $\mathcal{F}$ are sufficiently smooth, then the solution is unique.

Remark 5.8. The approximation ansatz (1.4) suffers from the fact that whenever $\tilde{z}_i(t) = 0$ for some $t \in [0, T)$ we cannot expect to determine any information on $p_i(t)$. This drawback results in the rather restrictive assumption $e_i^T \tilde{z}_{(0)} \neq 0$ for all $i = 1, \dots, r$ in Proposition 5.7. We can mitigate this restriction by enforcing the same transformation and the same path for a couple of modes as proposed in (1.5). In this case, we only have to ensure that for each of the transformations one single coefficient of the initial value is nonzero. This means that the initial condition has to contribute to every reference frame that we are interested in.

Remark 5.9. In principle, the MFEM may also suffer from a possible degenerate mass matrix. To circumvent this issue, a regularization is proposed in [48] to prevent nodes to move arbitrarily. In our setting, this corresponds to adding a regularization term for the path variable, respectively its derivative.

In order to apply Proposition 5.7, respectively Remark 5.8, we have to discuss how to choose the initial values $\tilde{z}_{(0)}$ and $p_{(0)}$. Following our general approximation (1.4), we thus have to find a minimizer for the optimization problem

$$\min_{p_{(0)}, \tilde{z}_{(0)}} J_{\mathrm{IV}} := \left\| z_0 - \sum_{i=1}^{r} \tilde{z}_{(0),i} \mathcal{T}_i(p_{(0),i}) \varphi_i \right\|_{\mathscr{X}}^2 . \tag{5.13}$$

The first-order optimality condition is given by

$$M_{\tilde{\mathrm{z}}}(p_{(0)}) \tilde{z}_{(0)} = \left[\langle z_0, \mathcal{T}_i(p_{(0),i}) \varphi_i \rangle_{\mathscr{X}} \right]_{i=1}^{r} =: b_{\mathrm{z}}(p_{(0)}), \tag{5.14a}$$

$$D(\tilde{z}_{(0)})^T N(p_{(0)})^T \tilde{z}_{(0)} = D(\tilde{z}_{(0)})^T \begin{bmatrix} \langle z_0, [\mathcal{T}_i'(p_{(0),i})\varphi_i] e_1 \rangle_{\mathscr{X}} \\ \vdots \\ \langle z_0, [\mathcal{T}_i'(p_{(0),i})\varphi_i] e_{q_i} \rangle_{\mathscr{X}} \end{bmatrix}_{i=1}^{r} =: D(\tilde{z}_{(0)})^T b_{\mathrm{p}}(p_{(0)}). \tag{5.14b}$$

We immediately notice that if $M_{\tilde{\mathrm{z}}}(p_{(0)})$ is singular, then in general we cannot expect a solution of (5.14a). On the other hand, if $M_{\tilde{\mathrm{z}}}(p_{(0)})$ is nonsingular, then we can solve the first equation for $\tilde{z}(0)$ and it is easy to see that in this case $p_{(0)}$ has to satisfy

$$D(\tilde{z}_{(0)})^T N(p_{(0)})^T M_{\tilde{\mathrm{z}}}(p_{(0)})^{-1} b_{\mathrm{z}}(p_{(0)}) - D(\tilde{z}_{(0)})^T b_{\mathrm{p}}(p_{(0)}) = 0. \tag{5.15}$$

In order to apply the inverse function theorem, we need additional smoothness of the mappings in Assumption 5.1. In the context of semigroups, this imposes stronger restrictions on the modes φ_i. Instead, we simply assume that we have initial values $\tilde{z}_{(0)}$ and $p_{(0)}$ available, such that the approximation error $J_{\mathrm{IV}}(\tilde{z}_{(0)}, p_{(0)})$ is sufficiently small. If we pick $p_{(0)}$ such that $\mathcal{T}_i(p_{(0),i}) = \mathrm{Id}_{\mathscr{X}}$ is the identity on $\mathscr{X}$, then (5.14a) simply describes the projection of the initial condition on the dominant modes.

For the remainder of this section we analyze the special case, where the right-hand side of (1.1) is given by a linear operator $\mathcal{A}\colon D(\mathcal{A}) \to \mathscr{X}$, *i.e.*, $\mathcal{F}(z) = \mathcal{A}z$. Assuming additionally that $\mathcal{A}$ is the generator of a strongly continuous semigroup ([54], Chap. 1, Def. 2.1) allows us immediately to establish a simple *a posteriori* error bound. Note that in this case $D(\mathcal{A})$ is a dense subspace of $\mathscr{X}$ such that we can choose $\mathscr{Y} := D(\mathcal{A})$.

Theorem 5.10 (*A posteriori* error bound)**.** *Let z denote the solution of the FOM (1.1) with initial value $z_0 \in D(\mathcal{A}) = \mathscr{Y}$, linear right-hand side $\mathcal{F}(z) = \mathcal{A}z$ and suppose that $\mathcal{A}$ is the generator of a strongly continuous semigroup $\{S(t)\}_{t\geq 0}$. Assume that the assumptions from Proposition 5.7 are satisfied and $t \mapsto \mathcal{T}_i(p_i(t))\varphi_i$ is twice continuously differentiable for $i = 1, \ldots, r$. For the unique solution $(\tilde{z}, p)$ of the ROM (5.7) with initial condition (5.12) define the error $\varepsilon := z - \sum_{i=1}^r \tilde{z}_i \mathcal{T}_i(p_i)\varphi_i$. Then there exist constants $\widetilde{T}, \widetilde{C}, \omega \geq 0$ independent of the modes φ_i and the transformation operators $\mathcal{T}_i$ such that*

$$\|\varepsilon(t)\|_{\mathscr{X}} \leq \widetilde{C}\mathrm{e}^{\omega t}\left(J_{\mathrm{IV}}(\tilde{z}_{(0)}, p_{(0)}) + t\|\mathscr{R}(\cdot, \dot{\tilde{z}}, \dot{p})\|_{L^\infty(0,t;\mathscr{X})}\right) \tag{5.16}$$

for $t \in [0, \widetilde{T})$.

Proof. For the proof we first note that Proposition 5.7 implies the existence of $\widetilde{T}$ such that (5.7) possesses a unique solution for $t \in [0, \widetilde{T})$ and observe a standard error residual relation given by the (abstract) differential equation

$$\begin{aligned}\dot{\varepsilon}(t) &= \dot{z}(t) - \sum_{i=1}^r \dot{\tilde{z}}_i(t)\mathcal{T}_i(p_i(t))\varphi_i - \sum_{i=1}^r \tilde{z}_i(t)[\mathcal{T}_i'(p_i(t))\varphi_i]\dot{p}_i(t) \\ &= \mathcal{A}\varepsilon(t) - \mathscr{R}(t, \dot{\tilde{z}}(t), \dot{p}(t))\end{aligned} \tag{5.17}$$

together with the initial condition

$$\varepsilon(0) = \varepsilon_0 := z_0 - \sum_{i=1}^r \tilde{z}_{(0),i}\mathcal{T}_i(p_{(0),i})\varphi_i \in D(\mathcal{A}).$$

Since $\{S(t)\}_{t\geq 0}$ is a strongly continuous semigroup, there exist (*cf.* [54], Chap. 1, Thm. 2.2) constants $\omega \geq 0$ and $\widetilde{C} \geq 1$ with $|||S(t)||| \leq \widetilde{C}\mathrm{e}^{\omega t}$ for all $t \geq 0$. Since $(\tilde{z}, p)$ is a solution of (5.7), we infer that the residual $\mathscr{R}$ is continuously differentiable and thus ([54], Chap. 4, Cor. 2.5) ensures that the (classical) solution of (5.17) is given by

$$\varepsilon(t) = S(t)\varepsilon_0 - \int_0^t S(t-\tau)\mathscr{R}(\tau, \dot{\tilde{z}}(\tau), \dot{p}(\tau))\,\mathrm{d}\tau.$$

We conclude the proof by estimating the integral with the supremum norm and the fact that for all $|||S(\tau)||| \leq \widetilde{C}\mathrm{e}^{\omega\tau} \leq \widetilde{C}\mathrm{e}^{\omega t}$ for all $0 \leq \tau \leq t$. □

Remark 5.11. In many applications, we expect that $\mathcal{A}$ is a semigroup of contractions or even an analytic semigroup with negative spectral abscissa. In either case, we can remove the exponential factor in (5.16) and hence obtain a linear growth factor in the time variable. Moreover, we can modify the result to include the L^2-norm of the residual instead of the L^∞-norm by using Young's inequality in the last step of the proof. Further possible modifications are the incorporation of the time-discretization error [31,37], using a weighted norm [27], or a space-time formulation [79]. Moreover, the error bound might be extended to nonlinear systems by using similar techniques as in [76]. Modifications of the error bound are considered future work.

Assume again that $\mathcal{A}$ is the generator of a strongly continuous semigroup $\{S(t)\}_{t\geq 0}$. For the moment, let us further assume that for some $p_i(t)$ we have $\mathcal{T}_i(p_i(t)) = S(t)$. Using $\frac{\mathrm{d}}{\mathrm{d}t}S(t)z = \mathcal{A}S(t)z = S(t)\mathcal{A}z$ for all $z \in D(\mathcal{A})$ (*cf.* [54], Chap. 1, Thm. 2.4), we observe $\widetilde{F}_{\tilde{\mathrm{z}}}(p,\tilde{z}) = N(p)D(\tilde{z})\dot{p}$ and thus (5.2) is given by

$$M(p(t))\dot{\tilde{z}}(t) = 0. \tag{5.18}$$

Clearly, $\tilde{z}(t) = \tilde{z}_0$ for all $t \geq 0$ is a solution of (5.18) that is particularly easy to compute.

Example 5.12. Let us reconsider the advection equation (*cf.* Example 3.3). For the ROM (5.7) we use the shift operator $\mathcal{T}(p)z = z(\cdot - p)$ with the usual embedding into $\mathscr{X} = L^2(0,1)$, *i.e.*, $\mathcal{T}_i(p_i) := \mathcal{T}(p_i)$. It is well-known that $\mathcal{T}$ is a semi-group and $-\partial_\xi \colon \mathscr{Y} \to \mathscr{X}$ is its generator ([17], Sect. II.2.10). For the right-hand side in (5.7b) we obtain

$$\begin{aligned}
\widetilde{F}_p(p,\tilde{z}) &= \begin{bmatrix} \left\langle \partial_\xi \mathcal{T}(p_1)\varphi_1, \partial_\xi \left(\sum_{j=1}^r \tilde{z}_j \mathcal{T}(p_j)\varphi_j\right)\right\rangle_{\mathscr{X}} \\ \vdots \\ \left\langle \partial_\xi \mathcal{T}(p_r)\varphi_r, \partial_\xi \left(\sum_{j=1}^r \tilde{z}_j \mathcal{T}(p_j)\varphi_j\right)\right\rangle_{\mathscr{X}} \end{bmatrix} \\
&= \begin{bmatrix} \langle \partial_\xi \mathcal{T}(p_1)\varphi_1, \partial_\xi \mathcal{T}(p_1)\varphi_1\rangle_{\mathscr{X}} & \cdots & \langle \partial_\xi \mathcal{T}(p_1)\varphi_1, \partial_\xi \mathcal{T}(p_r)\varphi_r\rangle_{\mathscr{X}} \\ \vdots & & \vdots \\ \langle \partial_\xi \mathcal{T}(p_r)\varphi_r, \partial_\xi \mathcal{T}(p_1)\varphi_1\rangle_{\mathscr{X}} & \cdots & \langle \partial_\xi \mathcal{T}(p_r)\varphi_r, \partial_\xi \mathcal{T}(p_r)\varphi_r\rangle_{\mathscr{X}} \end{bmatrix} \tilde{z} \\
&= M_p(p)D(\tilde{z})e,
\end{aligned}$$

with $e := \begin{bmatrix}1 & \cdots & 1\end{bmatrix}^T \in \mathbb{R}^r$. Similarly, $\widetilde{F}_{\tilde{\mathrm{z}}}(p,\tilde{z}) = N(p)D(\tilde{z})e$. Substituting (5.7a) into (5.7b) thus implies

$$D(\tilde{z})^T \left(N(p)^T M_{\tilde{\mathrm{z}}}^{-1}(p)N(p) - M_p(p)\right) D(\tilde{z})(\dot{p} - e) = 0.$$

Note that $N(p)^T M_{\tilde{\mathrm{z}}}^{-1}(p)N(p) - M_p(p)$ is the Schur complement of $M_{\tilde{\mathrm{z}}}(p)$ in $M(p)$ defined in (5.11). In particular, the assumptions of Proposition 5.7 imply $\dot{p} = e$ and hence the initial condition $p(0) = 0$ implies that $\mathcal{T}(p_i)$ equals the semigroup associated with the advection equation (3.6).

6. Connection with symmetry reduction

In this section, we investigate how the methodology outlined in Section 5 compares to the symmetry reduction framework as described in [7,53,64,65]. To this end we focus on the special case that all modes are transformed with the same transformation $\mathcal{T}$ and the same path $p(t) \in \mathcal{P} = \mathbb{R}^q$. More precisely, we assume $\hat{r} = 1$ in (1.5), *i.e.*, we consider the approximation.

$$z(t) \approx \sum_{i=1}^r \tilde{z}_i(t)\mathcal{T}(p(t))\varphi_i. \tag{6.1}$$

In this case, the matrices in the ROM (5.7), which we recall here

$$M_{\tilde{\mathrm{z}}}(p(t))\dot{\tilde{z}}(t) + N(p(t))D(\tilde{z}(t))\dot{p}(t) = \widetilde{F}_{\tilde{\mathrm{z}}}(p(t),\tilde{z}(t)), \tag{6.2a}$$

$$D(\tilde{z}(t))^T N(p(t))^T \dot{\tilde{z}}(t) + D(\tilde{z}(t))^T M_p(p(t))D(\tilde{z}(t))\dot{p}(t) = D(\tilde{z}(t))^T \widetilde{F}_p(p(t),\tilde{z}(t)), \tag{6.2b}$$

simplify as follows: The matrices $M_{\tilde{z}}$ and $\widetilde{F}_{\tilde{z}}$ are defined as in Section 5 but with $\mathcal{T}(p)$ instead of $\mathcal{T}_i(p_i)$, $D(\tilde{z}) := \tilde{z} \otimes I_q \in \mathbb{R}^{rq\times q}$, where $\otimes$ denotes the Kronecker product,

$$N(p) := \left[\langle \mathcal{T}(p)\varphi_i, [\mathcal{T}'(p)\varphi_j]e_1\rangle_{\mathscr{X}} \ \cdots \ \langle \mathcal{T}(p)\varphi_i, [\mathcal{T}'(p)\varphi_j]e_q\rangle_{\mathscr{X}}\right]_{i,j=1}^{r} \in \mathbb{R}^{r\times rq},$$

$$M_p(p) := \begin{bmatrix} \langle [\mathcal{T}'(p)\varphi_i]e_1, [\mathcal{T}'(p)\varphi_j]e_1\rangle_{\mathscr{X}} & \cdots & \langle [\mathcal{T}'(p)\varphi_i]e_1, [\mathcal{T}'(p)\varphi_j]e_q\rangle_{\mathscr{X}} \\ \vdots & & \vdots \\ \langle [\mathcal{T}'(p)\varphi_i]e_q, [\mathcal{T}'(p)\varphi_j]e_1\rangle_{\mathscr{X}} & \cdots & \langle [\mathcal{T}'(p)\varphi_i]e_q, [\mathcal{T}'(p)\varphi_j]e_q\rangle_{\mathscr{X}} \end{bmatrix}_{i,j=1}^{r} \in \mathbb{R}^{rq\times rq},$$

$$\widetilde{F}_p(p,\tilde{z}) := \begin{bmatrix} \left\langle [\mathcal{T}'(p)\varphi_i]e_1, \mathcal{F}\left(\sum_{j=1}^r \tilde{z}_j \mathcal{T}(p)\varphi_j\right)\right\rangle_{\mathscr{X}} \\ \vdots \\ \left\langle [\mathcal{T}'(p)\varphi_i]e_q, \mathcal{F}\left(\sum_{j=1}^r \tilde{z}_j \mathcal{T}(p)\varphi_j\right)\right\rangle_{\mathscr{X}} \end{bmatrix}_{i=1}^{r} \in \mathbb{R}^{rq}.$$

Recall that (6.2b) is the phase condition that corresponds to minimizing the residual with respect to $\dot{p}$ and is therefore referred to as Ψ_{Res}.

Another common assumption in the symmetry reduction framework is that $\mathcal{T}$ is a group action and the right-hand side $\mathcal{F}$ is equivariant with respect to $\mathcal{T}$, *cf.* Assumption 6.1. Furthermore, we assume in the following that $\mathcal{T}(\eta)$ is isometric for all $\eta \in \mathcal{P}$.

Assumption 6.1. *The right-hand side $\mathcal{F}$ is* equivariant *with respect to $\mathcal{T}$, i.e.,*

$$\mathcal{F}(\mathcal{T}(p)\varphi) = \mathcal{T}(p)\mathcal{F}(\varphi) \qquad \text{for all } \varphi \in \mathscr{Y} \text{ and } p \in \mathcal{P}. \tag{6.3}$$

Moreover the mapping $\mathcal{T}\colon \mathcal{P}\times\mathscr{X} \to \mathscr{X}$ is a group action*, i.e.*

$$\mathcal{T}(0) = \mathrm{Id}_{\mathscr{X}} \qquad \text{and} \qquad \mathcal{T}(\tilde{p})\mathcal{T}(p)\varphi = \mathcal{T}(\tilde{p}+p)\varphi \qquad \text{for all } \varphi \in \mathscr{X} \text{ and } p,\tilde{p} \in \mathcal{P}. \tag{6.4}$$

Remark 6.2. The properties of $\mathcal{F}$ and $\mathcal{T}$ stated in Assumption 6.1 can also be motivated by considering them from the perspective of the semigroup theory. In Section 5, we demonstrated that if the right-hand side $\mathcal{F}$ is linear and the corresponding linear operator is the generator of a strongly continuous semigroup, choosing the transformation $\mathcal{T}$ to be the action of this semigroup leads to a particularly simple ROM. However, in practice we do not expect to have direct access to the semigroup $\{S(t)\}_{t\geq 0}$ or, more generally speaking, to the flow of the differential equation. Still, this discussion provides a good starting point for the construction of the transformation operator $\mathcal{T}$, which should reflect the characteristic features of the expected solution behavior as encoded in $\{S(t)\}_{t\geq 0}$. In fact, by Assumption 6.1 we inherit two important properties of this semigroup.

Remark 6.3. It is important to note that due to the isometry property of $\mathcal{T}(p(t))$ and due to the equivariance assumption (6.3), the matrix $M_{\tilde{z}}$ and the right-hand side $\widetilde{F}_{\tilde{z}}$ do not depend on the path p anymore. If additionally, $\mathcal{T}(-p)[\mathcal{T}'(p)\phi]$ and $\langle [\mathcal{T}'(p)\phi]v, [\mathcal{T}'(p)\psi]w\rangle_{\mathscr{X}}$ do not depend on p for all $\phi,\psi \in \mathscr{Y}$ and $v,w \in \mathcal{P} = \mathbb{R}^q$, then also N, $\widetilde{F}_p$, and M_p are independent from the path. For instance, the shift operator as discussed in Example 5.12 satisfies these properties. In this case, the coefficient matrices $M_{\tilde{z}}$, N, and M_p can be precomputed in the offline phase. If additionally $\mathcal{F}$ is linear, or more generally, a polynomial mapping, the corresponding reduced operator can also be precomputed in the offline phase and, thus, the ROM can be evaluated efficiently without requiring computations that scale with the dimension of the FOM.

Using a single transformation $\mathcal{T}$ establishes a *reference frame*, *cf.* Section 2. In more detail assume that we are given a smooth path p and that we can split the dynamics as

$$z(t) = \mathcal{T}(p(t))v(t), \tag{6.5}$$

where we refer to v as the *frozen solution*. Especially, if (6.4) from Assumption 6.1 is satisfied, the frozen solution is given by $v(t) = \mathcal{T}(-p(t))z(t)$. Substituting (6.5) into the evolution equation (1.1) yields

$$\mathcal{T}(p(t))\dot{v}(t) + [\mathcal{T}'(p)v(t)]\dot{p}(t) = \mathcal{F}(\mathcal{T}(p(t))v(t)).$$

If Assumption 6.1 is satisfied, we can employ the identities $\mathcal{F}(\mathcal{T}(p(t))v(t)) = \mathcal{T}(p(t))\mathcal{F}(v(t))$ and $\mathcal{T}(-p(t))\mathcal{T}(p(t)) = \mathrm{Id}_{\mathscr{X}}$ to arrive at the reference frame equation

$$\dot{v}(t) = \mathcal{F}(v(t)) - \mathcal{T}(-p(t))[\mathcal{T}'(p)v(t)]\dot{p}(t). \tag{6.6}$$

Especially, given a continuously differentiable path p the assumptions imply that v is a solution of (6.6) if and only if z is a solution of (1.1), see Theorem 2.6 of [7] for further details. Based on the reference frame equation (6.6), we can construct a ROM *via* Galerkin projection onto the span of a suitable orthonormal basis $(\varphi_1, \ldots, \varphi_r)$ and obtain

$$\dot{\tilde{v}}_i(t) = \left\langle \varphi_i, \mathcal{F}\left(\sum_{j=1}^{r} \tilde{v}_j(t)\varphi_j\right)\right\rangle_{\mathscr{X}} - \sum_{j=1}^{r} \langle \varphi_i, \mathcal{T}(-p(t))[\mathcal{T}'(p)\varphi_j]\dot{p}(t)\rangle_{\mathscr{X}}\, \tilde{v}_j(t) \tag{6.7}$$

for $i = 1, \ldots, r$. Note that the approximation is given by $v(t) \approx \sum_{i=1}^{r} \tilde{v}_i(t)\varphi_i$ with associated residual

$$\mathscr{R}(t) = \sum_{i=1}^{r} \dot{\tilde{v}}_i(t)\varphi_i - \mathcal{F}\left(\sum_{i=1}^{r} \tilde{v}_i(t)\varphi_i\right) + \sum_{i=1}^{r} \tilde{v}_i(t)\mathcal{T}(-p(t))[\mathcal{T}'(p(t))\varphi_i]\dot{p}(t). \tag{6.8}$$

For convenience we introduce the reduced frozen state $\tilde{v} := [\tilde{v}_1 \ \cdots \ \tilde{v}_r]^T$. Using the notation from (6.2), we can write the ROM (6.7) of the reference frame equation as

$$\dot{\tilde{v}}(t) = \widetilde{F}_{\tilde{\mathrm{z}}}(p(t), \tilde{v}(t)) - N(p(t))D(\tilde{v}(t))\dot{p}(t). \tag{6.9}$$

Thus, we immediately arrive at the following relation between the symmetry reduction ROM (6.7) and the continuously optimal ROM (6.2).

Lemma 6.4. *Let $(\varphi_1, \ldots, \varphi_r)$ be an orthonormal basis of an r-dimensional subspace of $\mathscr{Y}$ and consider the single-frame approximation* (6.1)*. If the transformation operator $\mathcal{T}$ is an isometry and satisfies Assumptions 5.1 and 6.1, then* (6.2a) *and* (6.7) *are equivalent in the sense that for every continuously differentiable path p, any solution of* (6.2a) *is a solution of* (6.7) *and vice versa.*

Lemma 6.4 establishes the equivalence between the ROM obtained by symmetry reduction and the one obtained by our framework in the case that the same path is chosen for both ROMs. In our framework, the path is fixed by adding the phase condition (6.2b), which ensures to minimize the residual. In the symmetry reduction framework, different phase conditions have been proposed in the literature. For instance, in [7] the authors propose to derive a phase condition for the reference frame equation (6.6) by minimizing the temporal change of the frozen solution, *i.e.*, by minimizing $\frac{1}{2}\|\dot{v}(t)\|_{\mathscr{X}}^2$ over $\dot{p}(t)$. This idea of choosing the path is usually referred to as *freezing*. The associated phase condition $\Psi_{\mathrm{freezeFOM}}(p, \dot{p}, v)$[2] is given by the first-order necessary optimality condition

$$[\mathcal{T}'(p(t))v(t)]^*[\mathcal{T}'(p(t))v(t)]\dot{p}(t) = [\mathcal{T}'(p(t))v(t)]^*\mathcal{F}(\mathcal{T}(p(t))v(t)), \tag{6.10}$$

[2] The authors of [7] denote the phase condition $\Psi_{\mathrm{freezeFOM}}$ as Ψ_{min}. Since all phase conditions discussed in our exposition are based on a minimization problem, we use a different name here.

where we use that $\mathcal{T}(p(t))$ is isometric as well as Assumption 6.1. In [7], the authors propose to discretize the phase condition in space by replacing the occurring operators and the frozen solution v by their finite difference approximations. In the context of model reduction, this strategy corresponds to enforcing

$$\Psi_{\text{freeze}}(p, \dot{p}, \tilde{v}) := \Psi_{\text{freezeFOM}}\left(p, \dot{p}, \sum_{i=1}^{r} \tilde{v}_i \varphi_i\right) = 0$$

or equivalently,

$$D(\tilde{v}(t))^T M_p(p(t)) D(\tilde{v}(t)) \dot{p}(t) = D(\tilde{v}(t))^T \widetilde{F}_p(p(t), \tilde{v}(t)). \tag{6.11}$$

It is important to note that in general, (6.11) is not equivalent to the first-order necessary optimality condition for minimizing the temporal change of the reduced state, *i.e.*, by minimizing $\frac{1}{2}\|\dot{\tilde{v}}(t)\|_2^2$ over $\dot{p}(t)$. This optimality condition, which we call $\widetilde{\Psi}_{\text{freeze}}$, is given by

$$D(\tilde{v}(t))^T N(p(t))^T N(p(t)) D(\tilde{v}(t)) \dot{p}(t) = D(\tilde{v}(t))^T N(p(t))^T \widetilde{F}_{\tilde{z}}(p(t), \tilde{v}(t)). \tag{6.12}$$

The relation between the three different phase conditions Ψ_{Res}, Ψ_{freeze}, and $\widetilde{\Psi}_{\text{freeze}}$, defined in (6.2b), (6.11), and (6.12), respectively, is provided in the next result.

Theorem 6.5. *Let the assumptions of Lemma 6.4 be satisfied. Then, for each $t > 0$, the phase condition Ψ_{freeze} given in* (6.11) *is the necessary first-order optimality condition for the optimization problem*

$$\min_{\dot{p}(t)} \frac{1}{2}\|\mathscr{R}(t)\|_{\mathscr{X}}^2 + \frac{1}{2}\|\dot{\tilde{v}}(t)\|_2^2,$$

where $\mathscr{R}$ denotes the residual defined in (6.8) *and $\dot{\tilde{v}}$ is given by* (6.9).

Proof. Using (6.9), we calculate for $t > 0$

$$\begin{aligned}
J_{\text{freeze}}(\dot{p}(t)) &:= \frac{1}{2}\|\mathscr{R}(t)\|_{\mathscr{X}}^2 + \frac{1}{2}\|\dot{\tilde{v}}(t)\|_2^2 = \frac{1}{2}\|\mathscr{R}(t)\|_{\mathscr{X}}^2 + \frac{1}{2}\left\|\sum_{i=1}^{r} \dot{\tilde{v}}_i(t)\varphi_i\right\|_{\mathscr{X}}^2 \\
&= \frac{1}{2}\|\mathscr{R}(t)\|_{\mathscr{X}}^2 + \frac{1}{2}\left\|\sum_{i=1}^{r} \dot{\tilde{v}}_i(t)\varphi_i\right\|_{\mathscr{X}}^2 - \left\langle \mathscr{R}(t), \sum_{i=1}^{r} \dot{\tilde{v}}_i(t)\varphi_i \right\rangle_{\mathscr{X}} \\
&= \frac{1}{2}\left\|\mathscr{R}(t) - \sum_{i=1}^{r} \dot{\tilde{v}}_i(t)\varphi_i\right\|_{\mathscr{X}}^2 \\
&= \frac{1}{2}\left\|\sum_{i=1}^{r} \tilde{v}_i(t)[\mathcal{T}'(p(t))\varphi_i]\dot{p}(t) - \mathcal{F}\left(\sum_{j=1}^{r} \tilde{v}_j(t)\mathcal{T}(p(t))\varphi_j\right)\right\|_{\mathscr{X}}^2,
\end{aligned}$$

where the second equality follows from the fact that the Galerkin ROM (6.9) enforces the residual to be orthogonal to $\operatorname{span}\{\varphi_1, \ldots, \varphi_r\}$. The necessary first-order optimality condition for this minimization problem is obtained by differentiating J_{freeze} with respect to $\dot{p}(t)$ and setting the derivative to zero. The resulting equation is (6.11) and, thus, the claim follows. □

The different phase conditions together with their associated optimization problems are summarized in Table 1. Note that (6.2a) implies that (6.2b) is equivalent to

$$\begin{aligned}
&D(\tilde{z}(t))^T \left(M_p(p(t)) - N(p(t))^T N(p(t))\right) D(\tilde{z}(t))\dot{p}(t) \\
&\quad = D(\tilde{z}(t))^T \left(\widetilde{F}_p(p(t), \tilde{z}(t)) - N(p(t))^T \widetilde{F}_{\tilde{z}}(p(t), \tilde{z}(t))\right),
\end{aligned} \tag{6.13}$$

where we used that $\mathcal{T}(p(t))$ is an isometry. In particular, we have $\Psi_{\text{Res}} = \Psi_{\text{freeze}} - \widetilde{\Psi}_{\text{freeze}}$.

TABLE 1. Phase conditions.

Phase condition	Equation	Optimization problem
$\Psi_{\mathrm{Res}} = 0$	(6.2b)	$\min_{\dot{p}} \frac{1}{2}\|\mathscr{R}\|^2$
$\Psi_{\mathrm{freeze}} = 0$	(6.11)	$\min_{\dot{p}} \frac{1}{2}\|\mathscr{R}\|^2 + \frac{1}{2}\|\dot{\tilde{v}}\|^2$
$\widetilde{\Psi}_{\mathrm{freeze}} = 0$	(6.12)	$\min_{\dot{p}} \frac{1}{2}\|\dot{\tilde{v}}\|^2$

7. Numerical examples

In this section, we illustrate the ROM (5.7) derived in Section 5 for several test cases. In Section 7.1, we discuss the one-dimensional advection-diffusion equation with periodic boundary condition and illustrate the advantages of a restriction of the transformations as outlined in Remark 5.8. A numerical study with non-periodic boundary conditions is performed in Section 7.2. For an example that exhibits more than one transport we discuss the linear wave equation in Section 7.3. We conclude our numerical results with the nonlinear (viscous) Burgers' equation (*cf.* Sect. 7.4) and demonstrate the advantages in comparison to standard POD. For all our examples with periodic boundary conditions we use the shift operator discussed in Example 5.2 for $\mathcal{T}_i$. In particular, Assumption 5.1 is satisfied provided that the data is sufficiently smooth. Moreover, this immediately implies $\mathcal{P}_i = \mathbb{R}$, *i.e.*, $q_i = 1$. For the case of non-periodic boundary conditions the transformation operator has to be modified, which is discussed in Section 7.2.

As is common in the model-order reduction literature, we refer to the solution of the spatially discretized equation (1.1) as the *truth solution* and, by abuse of notation, denote it by z. For all examples except for the one in Section 7.2, we use a 6-th order central finite difference scheme (see for instance [19]) on a grid with 200 equidistant points for the spatial discretization. The stencils for the finite-difference matrices $\partial_\xi \approx D_1$ and $\partial_\xi^2 \approx D_2$ approximating the spatial derivatives are given by

$$\left[-\frac{1}{60}, \frac{3}{20}, -\frac{3}{4}, 0, \frac{3}{4}, -\frac{3}{20}, \frac{1}{60}\right] \qquad \text{and} \qquad \left[\frac{1}{90}, -\frac{3}{20}, \frac{3}{2}, -\frac{49}{18}, \frac{3}{2}, -\frac{3}{20}, \frac{1}{90}\right],$$

respectively. For the time integration, we use the implicit trapezoidal rule with constant stepsize of $\tau = 5 \times 10^{-3}$ and we use the MATLAB function `fsolve` with standard tolerances to solve the resulting implicit system of equations. The $L^2(0, T; \mathscr{X})$ norm is approximated with the trapezoidal rule, and for a truth solution z with approximation $\hat{z}$ we denote by

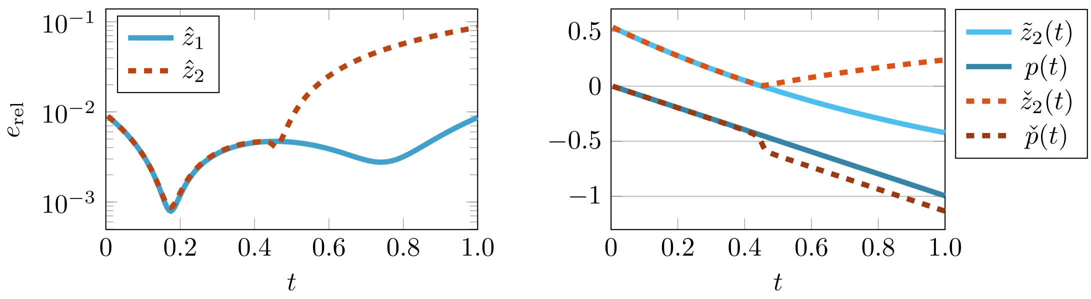
$$e_{\mathrm{rel}} := \frac{\|z - \hat{z}\|_{L^2(0,T;\mathscr{X})}}{\|z\|_{L^2(0,T;\mathscr{X})}}$$

the relative error. The computation of the dominant modes for our approach, *i.e.*, the solution of the minimization problem (4.1) is performed with the algorithm described in [61]. Consequently, we denote the ROM (5.7) as *shifted POD.*

The shift operator's application may require the evaluation of the modes at points not captured by the spatial grid. In the numerical experiments, we compute these evaluations utilizing polynomial interpolation with cubic Lagrange polynomials. Furthermore, we also need to evaluate the spatial derivatives of the shifted modes in the online phase, *cf.* Example 5.12. To this end, we use the same discretization of the space derivative as in the simulation of the FOM.

Code availability

The MATLAB source code for the numerical examples can be obtained from the doi 10.5281/zenodo.3902924.



(a) **(b)**

FIGURE 1. Advection-diffusion equation – Comparison of the ROM with a single transformation for all modes ($\hat{z}_1$, solid blue) and the ROM with one transformation per mode ($\hat{z}_2$, dashed red). (a) Time evolution of the relative error for the ROMs $\hat{z}_1$ and $\hat{z}_2$. (b) Evolution of the coefficient for the second mode and the corresponding path for $\hat{z}_1$ and $\hat{z}_2$.

7.1. Advection-diffusion equation

The one-dimensional advection-diffusion equation with periodic boundary conditions is given by

$$\begin{cases} \partial_t z(t,\xi) + c\partial_\xi z(t,\xi) - \mu\partial_\xi^2 z\,(t,\xi) = 0, & (t,\xi) \in (0,1)\times(0,1), \\ z(t,0) = z(t,1), & t \in (0,1), \\ z(0,\xi) = \exp\left(-\left(\frac{\xi-0.5}{0.1}\right)^2\right), & \xi \in (0,1). \end{cases} \tag{7.1}$$

The parameters c and μ denote the transport velocity and the diffusion coefficient, respectively, and are chosen as $c = 1$ and $\mu = 0.002$. Choosing $r = 2$ modes, the algorithm described in [61], yields a relative offline error of 8.6×10^{-3}.

For the online phase (*cf.* Sect. 5) we compare two different ROMs: the first one uses only one path p for both modes as described in (6.1), *i.e.*, the approximate solution $\hat{z}_1$ of (7.1) is given by $\hat{z}_1(t) := \sum_{i=1}^{r} \tilde{z}_i(t)\mathcal{T}(p)\varphi_i$. The second ROM is constructed as proposed in (5.7) and the corresponding approximation is denoted by $\hat{z}_2$. The relative errors are

$$e_{\mathrm{rel},\hat{z}_1} \approx 4.3 \times 10^{-3} \qquad \text{and} \qquad e_{\mathrm{rel},\hat{z}_2} \approx 3.6 \times 10^{-2}$$

and the evolution of the relative error with respect to time is presented in Figure 1a. To understand the larger error in the second approximation, we consider the coefficient functions for the second mode for both approaches, which are presented in Figure 1b. We observe that at time $t \approx 0.45$ the coefficient functions are almost equal to zero. Following the discussion in Remark 5.8, the solution of the ROM (5.7) in a neighborhood of such a point is ill-conditioned, which results in the weaker approximation quality of $\hat{z}_2$.

The ROM is not only able to reproduce the FOM with the same set of parameters, but can also accurately predict the behavior of the FOM for different parameters. The relative error of the shifted POD approximation with $r = 2$ for different values of the transport velocity is presented in the left image in Figure 2. Note that POD requires $r = 11$ modes to obtain a similar accuracy as the shifted POD albeit with a larger computational time (*cf.* right image in Fig. 2). Here, the computational time is the median time of 5 simulation runs. A similar computational time as the shifted POD is achieved with POD with $r = 3$ modes although with a larger relative error. For the comparison of the computational time we use the MATLAB solver `ode45` (see also the forthcoming discussion in Sect. 7.3) and the shifted POD approach is implemented utilizing equivariance, see Remark 6.3. Nevertheless, the numerical simulation code for the FOM, POD, and shifted POD are not optimized for performance, which is the main reason for omitting the computational times in the following examples. An efficient implementation of our framework is subject to further research.

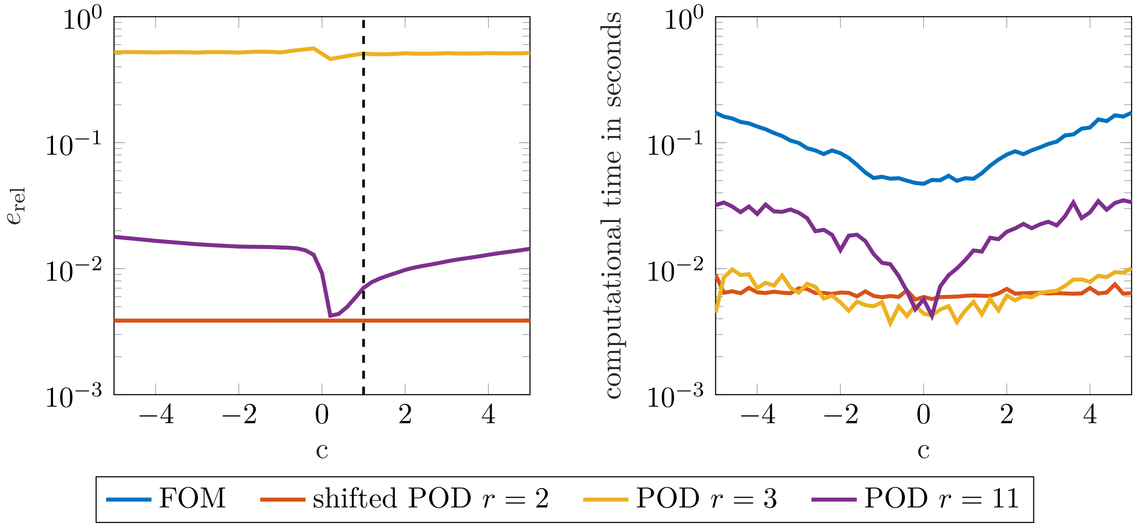


FIGURE 2. Advection-diffusion equation – Approximation quality of the shifted POD reduced model with $r = 2$ and POD reduced model with $r \in \{3, 11\}$ for different values of the transport velocity c. The dashed line at $c = 1$ denotes the parameter that is used to construct the ROMs. *Left image*: relative online error. *Right image*: median computational time in seconds.

7.2. Non-periodic boundary conditions

To illustrate that the presented framework is not restricted to problems with periodic boundary conditions, we consider the advection-diffusion equation with Dirichlet–Neumann boundary conditions of the form

$$\begin{cases} \partial_t z(t,\xi) + c\partial_\xi z(t,\xi) - \mu\partial_\xi^2 z\,(t,\xi) = 0, & (t,\xi) \in (0,1.5) \times (0,1), \\ \qquad z(t,0) = \dfrac{1}{2}\exp\left(-\left(\frac{t-0.2}{0.03}\right)^2\right), & t \in (0,1.5), \\ \qquad \partial_\xi z(t,1) = 0, & t \in (0,1.5), \\ \qquad z(0,\xi) = \dfrac{1}{2}\exp\left(-\left(\frac{\xi-0.5}{0.02}\right)^2\right), & \xi \in (0,1.5). \end{cases}$$

In the following, we use the values $c = 1$ and $\mu = 0.001$. For the spatial discretization we employ a standard central second-order finite difference scheme with mesh width $\Delta\xi = 1.25 \times 10^{-3}$. The solution of the FOM and the error of the corresponding ROM with our framework with $r = 4$ and a single transformation operator for all modes are depicted in Figure 3.

Note that a proper treatment of the non-periodic boundary conditions requires a modification of the (periodic) shift operator introduced in Example 5.2. To mimic the behavior of the upstream boundary, we propose the following strategy: Instead of using modes on the original domain $\Omega := [0,1]$, we define the modes on a *virtual* domain Ω_{virt} satisfying

$$\{\xi - p(t) \mid \xi \in \Omega, t \in [0,T)\} \subseteq \Omega_{\text{virt}}.$$

The modified shift operator can then be defined as

$$(\mathcal{T}(\eta)\varphi_i)\,(\cdot) := \varphi_i(\cdot - \eta)|_\Omega\,, \tag{7.2}$$

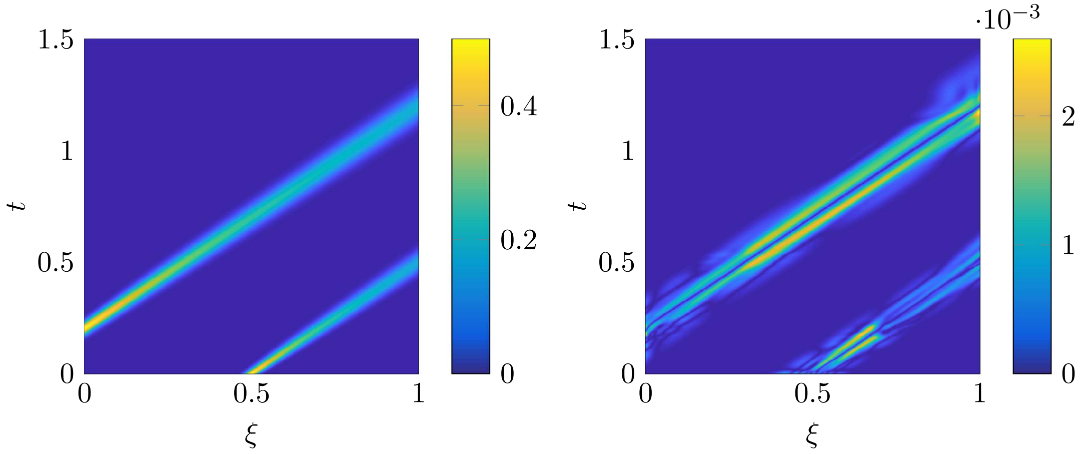


FIGURE 3. Advection-diffusion equation with non-periodic boundary conditions – FOM (*left*) and absolute error between the FOM and the shifted POD ROM (*right*) with $r = 4$.

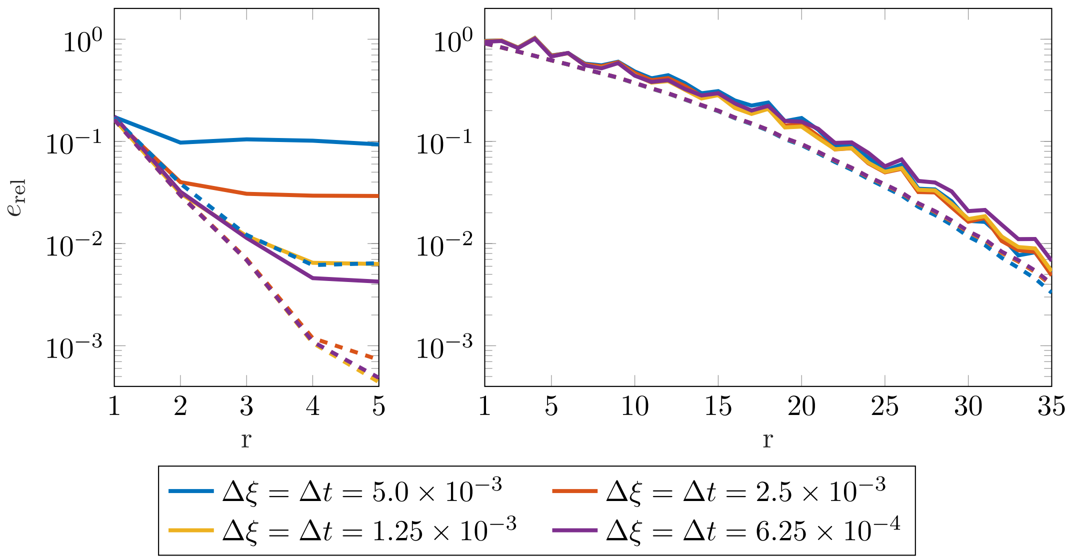


FIGURE 4. Advection-diffusion equation with non-periodic boundary conditions – Decay of the relative error for the shifted POD ROM (*left*) and the POD ROM (*right*). The solid lines represent the online error, while the dashed lines represent the offline error.

where $\varphi_i(\cdot - \eta)|_\Omega$ denotes the restriction of $\varphi_i(\cdot - \eta)$ to Ω. If the path p is known, this virtual domain can, for instance, be chosen as

$$\Omega_{\text{virt}} = \left[-\sup_{t\in[0,T]} p(t),\ 1 - \inf_{t\in[0,T]} p(t) \right].$$

With the parameters defined above we thus obtain $\Omega_{\text{virt}} = [-1.5, 1]$. At this point, we emphasize that the shift operator (7.2) does formally not fit into the framework presented in this paper. In particular, Assumptions 4.1 and 5.1 are not satisfied. Nevertheless, the right image in Figure 3 details that our approach with $r = 4$ modes

TABLE 2. Number of adaptive timesteps for the FOM, the POD reduced model, and the shifted POD reduced model.

	FOM	POD	Shifted POD
`ode45`	202	51 (25%)	15 (7%)
`ode23`	189	194 (103%)	16 (8%)

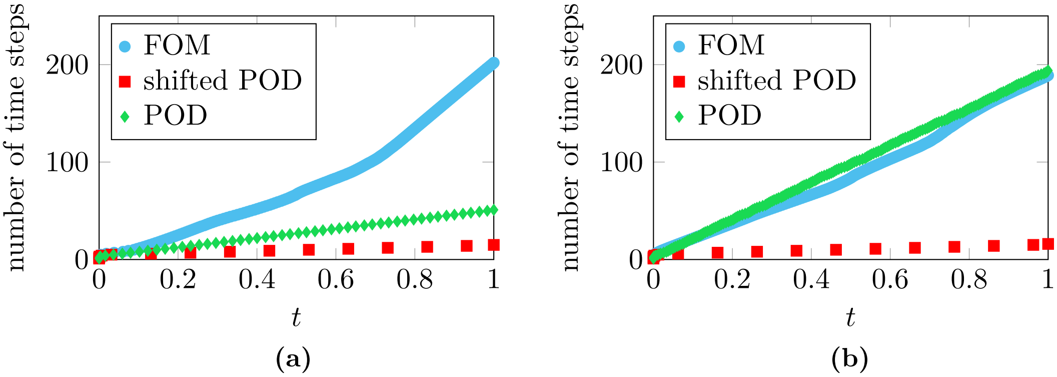


FIGURE 5. Linear wave equation – Number of time steps chosen by an adaptive time integrator in the computation of the FOM (blue circles), the shifted POD reduced model (red squares), and POD reduced model (green diamonds). (a) MATLAB solver `ode45`. (b) MATLAB solver `ode23`.

can represent the solution of the FOM accurately. It is important to note that the approximation quality of the shifted POD strongly depends on the accuracy of the FOM. If the discretization of the FOM is not sufficiently fine, then the online and the offline error of shifted POD may stagnate. The specific case where the FOM is computed with the same spatial and temporal step size is presented in the left plot of Figure 4. In contrast, the approximation with POD is less sensitive, albeit with a larger error. A detailed analysis of the impact of the discretization error is subject to further research.

7.3. Linear wave equation

We consider the one-dimensional acoustic wave equation with periodic boundary conditions as discussed in Example 4.4. For the offline phase, we evaluate the analytical solution (4.4) on an equally-distributed space-time grid with 200 points in space and time, respectively. We use the algorithm from [61] to compute $r = 2$ modes with a relative offline approximation error of 7×10^{-16}. Based on the two dominant modes determined in the offline stage, the ROM (5.7) yields an online error of $e_{\mathrm{rel}} \approx 3.2 \times 10^{-10}$ and thus accurately represents the original equation.

Let us emphasize that our ROM not only achieves a considerable reduction in the dimension of the spatial variable, but also benefits the time integration. To detail this, we compare the numerical solution of the FOM, a POD reduced model and our approach with the MATLAB solvers `ode45` and `ode23`. Both solvers choose an adaptive step size and the corresponding numbers for the three different models are presented in Table 2 and Figure 5.

Clearly, our approach requires considerably less time steps for the numerical integration compared to the FOM and the POD reduced model. The main reason for this behavior is the fact that the reduced state $\tilde{z}$ in our framework is changing only slightly, *i.e.*, $\|\dot{\tilde{z}}(t)\|_2$ is small. In view of the different phase conditions (*cf.* Tab. 1) this is interesting, since our ROM is based on the minimization of the residual.

TABLE 3. Approximation quality of ROMs for Burgers' equation.

	Shifted POD $r = 7$	POD $r = 7$	POD $r = 32$
Relative offline error	4.4×10^{-3}	7.2×10^{-2}	1.7×10^{-3}
Relative online error	3.8×10^{-3}	2.1×10^{-1}	3.5×10^{-3}

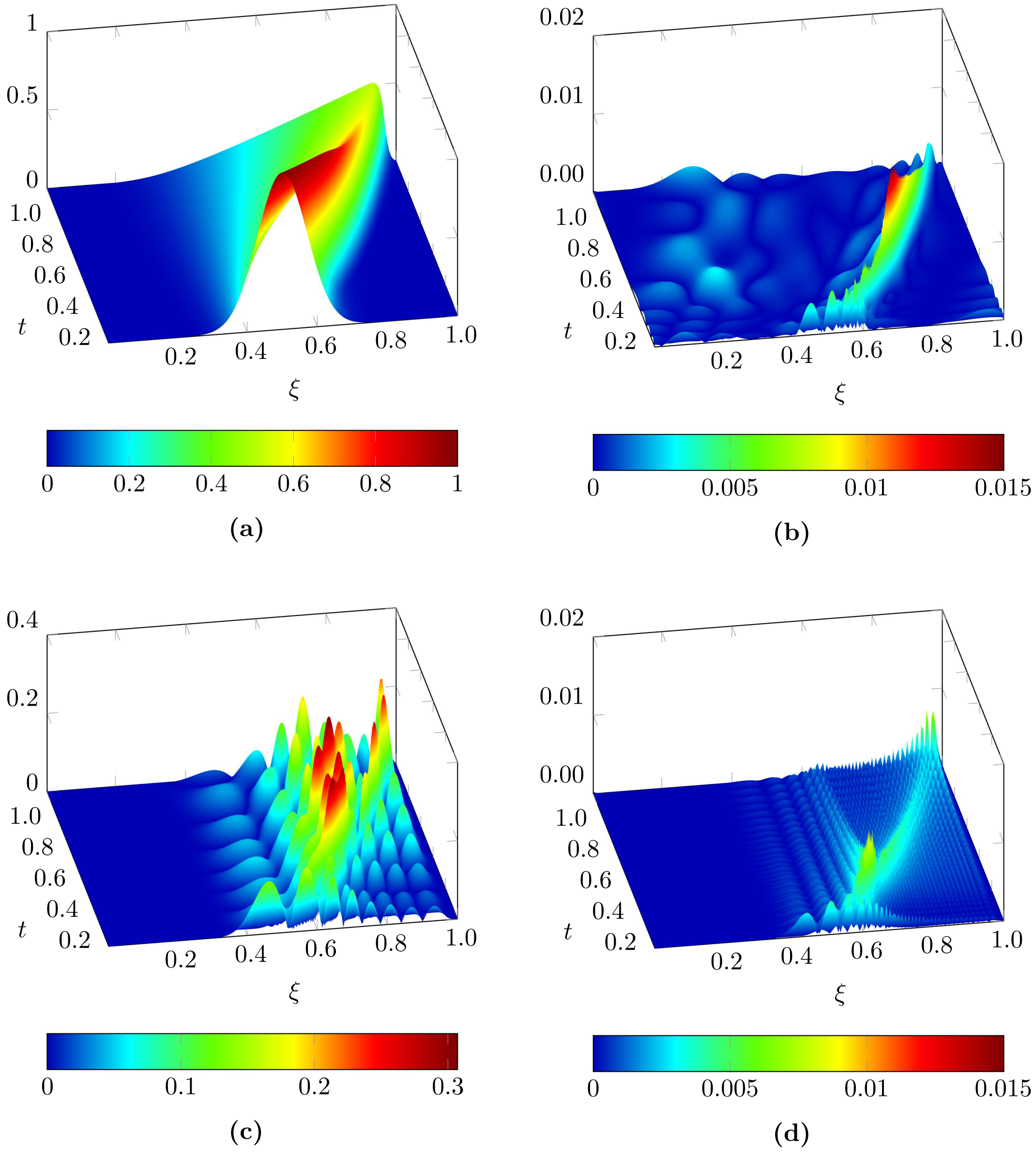


FIGURE 6. Burgers' equation – Solution of the FOM and absolute errors for different ROMs. (a) FOM with parameter $\mu = 2 \times 10^{-3}$ and 200 degrees of freedom. (b) Absolute error for the shifted POD approximation with $r = 7$ modes. (c) Absolute error for the POD approximation with $r = 7$ modes. (d) Absolute error for the POD approximation with $r = 32$ modes.

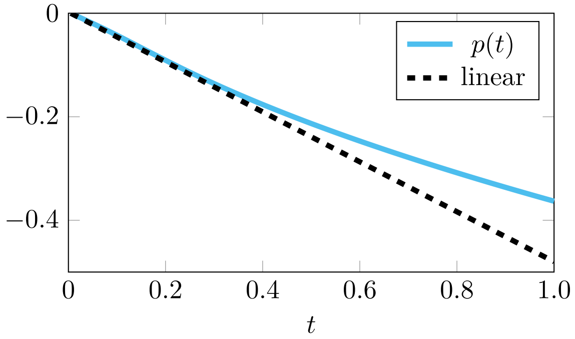


FIGURE 7. Burgers' equation – The shift variable is nonlinear.

7.4. Burgers' equation

We consider the one-dimensional (viscous) Burgers' equation with periodic boundary conditions, given by

$$\begin{cases} \partial_t z(t,\xi) = \mu\partial_\xi^2 z(t,\xi) - z(t,\xi)\,\partial_\xi z(t,\xi)\,, & (t,\xi)\in(0,1)\times(0,1)\,,\\ z(t,0) = z(t,1)\,, & t\in(0,1)\,,\\ z(0,\xi) = \exp\left(\left(\frac{\xi-0.5}{0.1}\right)^2\right), & \xi\in(0,1)\,, \end{cases} \tag{7.3}$$

as a nonlinear test case for the methodology presented within this paper. For our experiments we use the viscosity parameter $\mu = 2\times 10^{-3}$. The solution is presented in Figure 6a. As for the other examples, we use the shift operator $\mathcal{T}(p)\,z = z(\cdot - p)$ from Example 4.3 as transformation operator, and transform each mode with the same transformation, *i.e.*, we use the approximation (6.1). Let us emphasize that the right-hand side of the Burgers' equation is equivariant with respect to the shift operator (*cf.* Assumption 6.1) and hence the ROM simplifies as described in Remark 6.3. Note that we can exploit the polynomial form of the nonlinear right-hand side to eliminate the dependency of the original space by computing the reduced right-hand side *via*

$$\begin{aligned}\widetilde{F}_{\tilde{z}}(\tilde{z}) = &-\begin{bmatrix} \langle\varphi_1,\varphi_1\partial_\xi\varphi_1\rangle_{\mathscr{X}} & \dots & \langle\varphi_1,\varphi_1\partial_\xi\varphi_r\rangle_{\mathscr{X}} & \langle\varphi_1,\varphi_2\partial_\xi\varphi_1\rangle_{\mathscr{X}} & \dots & \langle\varphi_1,\varphi_r\partial_\xi\varphi_r\rangle_{\mathscr{X}}\\ \vdots & & \vdots & \vdots & & \vdots\\ \langle\varphi_r,\varphi_1\partial_\xi\varphi_1\rangle_{\mathscr{X}} & \dots & \langle\varphi_r,\varphi_1\partial_\xi\varphi_r\rangle_{\mathscr{X}} & \langle\varphi_r,\varphi_2\partial_\xi\varphi_1\rangle_{\mathscr{X}} & \dots & \langle\varphi_r,\varphi_r\partial_\xi\varphi_r\rangle_{\mathscr{X}} \end{bmatrix}(\tilde{z}\otimes\tilde{z})\\ &-\mu\begin{bmatrix} \langle\partial_\xi\varphi_1,\partial_\xi\varphi_1\rangle_{\mathscr{X}} & \dots & \langle\partial_\xi\varphi_1,\partial_\xi\varphi_r\rangle_{\mathscr{X}}\\ \vdots & & \vdots\\ \langle\partial_\xi\varphi_r,\partial_\xi\varphi_1\rangle_{\mathscr{X}} & \dots & \langle\partial_\xi\varphi_r,\partial_\xi\varphi_r\rangle_{\mathscr{X}} \end{bmatrix}\tilde{z}.\end{aligned}$$

For our ROM, we identify $r = 7$ modes with a relative offline error of 4.4×10^{-3}. We compare our ROM with a POD reduced model with $r = 7$ and $r = 32$. The results are presented in Table 3 and Figure 6. Clearly, our approach outperforms POD with the same number of modes. To achieve a comparable error, POD requires more than four times as many modes as our framework.

Comparing the error plots in Figure 6 we observe, in agreement with Example 3.3, strong oscillations in the case of the POD approximation with $r = 7$ modes, and they remain present even for the approximation with $r = 32$ modes. In contrast, the shifted POD absolute error does not exhibit this behaviour, the error is dominated by the region where the verge of the transport front lies. We remark that the shift $p(t)$ computed from the ROM is indeed nonlinear (see Fig. 7), as is expected, since the transport velocity for the nonlinear test case of the Burgers' equation depends on the solution itself.

8. Conclusions

In this paper, we introduce a new framework for constructing reduced order models based on the approximation ansatz (1.4), which features multiple time-dependent transformation operators. This ansatz allows obtaining accurate low-dimensional surrogate models even for systems whose dynamics are dominated by multiple transport modes with potentially large gradients. The construction of the ROM is based on residual minimization and extends the ideas of the moving finite element method to model reduction. Furthermore, we provide a residual-based *a posteriori* error bound. For the particular case that only one isometric transformation operator is employed, we show a connection between our method and the symmetry reduction framework, *cf.* [7, 64]. Further contributions include a thorough literature review of related approaches and analysis for the identification of optimal basis functions on the infinite-dimensional level. We illustrate our theoretical findings with several analytical and numerical examples.

The problem of identifying optimal basis functions is currently solved by a first-discretize-then-optimize approach. For future work, it is interesting to analyze the first-optimize-then-discretize approach and compare the results with the current strategy. Furthermore, we plan to investigate the efficient implementation of the ROM. Notably, the combination with hyper-reduction techniques is a promising research direction to obtain an efficient offline/online decomposition.

Acknowledgements. We are grateful for helpful discussions with Mathieu Rosière, Fredi Tröltzsch, and Christoph Zimmer from TU Berlin. The work of Felix Black and Philipp Schulze is supported by the DFG Collaborative Research Center 1029 *Substantial efficiency increase in gas turbines through direct use of coupled unsteady combustion and flow dynamics*, project number 200291049. The work of Benjamin Unger is supported by the DFG Collaborative Research Center 910 *Control of self-organizing nonlinear systems: Theoretical methods and concepts of application*, project number 163436311.

References

[1] A.C. Antoulas, Approximation of Large-Scale Dynamical Systems. *Advances in Design and Control.* SIAM, Philadelphia, PA, USA (2005).

[2] M. Baştuğ, M. Petreczky, R. Wisniewski and J. Leth, Model reduction by nice selections for linear switched systems. *IEEE Trans. Autom. Control* **61** (2016) 3422–3437.

[3] M. Barrault, Y. Maday, N.C. Nguyen and A.T. Patera, An "empirical interpolation" method: application to efficient reduced-basis discretization of partial differential equations. *C. R. Math. Acad. Sci. Paris* **339** (2004) 667–672.

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

[5] P. Benner, S. Gugercin and K. Willcox, A survey of projection-based model reduction methods for parametric dynamical systems. *SIAM Rev.* **57** (2015) 483–531.

[6] P. Benner, A. Cohen, M. Ohlberger and K. Willcox, Model Reduction and Approximation. *Advances in Design and Control.* SIAM, Philadelphia, PA, USA (2017).

[7] W.-J. Beyn and V. Thümmler, Freezing solutions of equivariant evolution equations. *SIAM J. Appl. Dyn. Syst.* **3** (2004) 85–116.

[8] R.K. Brayton, Nonlinear oscillations in a distributed network. *Q. Appl. Math.* **24** (1967) 289–301.

[9] H. Brezis, Functional Analysis, Sobolev Spaces and Partial Differential Equations. Universitext. Springer, New York, NY, USA (2011).

[10] N. Cagniart, R. Crisovan, Y. Maday and R. Abgrall, Model order reduction for hyperbolic problems: a new framework. Preprint https://hal.archives-ouvertes.fr/hal-01583224 (2017).

[11] N. Cagniart, Y. Maday and B. Stamm, Model order reduction for problems with large convection effects. In: *Computational Methods in Applied Sciences.* Springer, Cham, Switzerland (2019) 131–150.

[12] K. Carlberg, Adaptive h-refinement for reduced-order models. *Int. J. Numer. Methods Eng.* **102** (2015) 1192–1210.

[13] K. Carlberg, M. Barone and H. Antil, Galerkin v. least-squares Petrov–Galerkin projection in nonlinear model reduction. *J. Comput. Phys.* **330** (2017) 693–734.

[14] S. Chaturantabut and D. Sorensen, Nonlinear model reduction *via* discrete empirical interpolation. *SIAM J. Sci. Comput.* **32** (2010) 2737–2764.

[15] K.L. Cooke and D.W. Krumme, Differential-difference equations and nonlinear initial-boundary value problems for linear hyperbolic partial differential equations. *J. Math. Anal. App.* **24** (1968) 372–387.

[16] M. Dihlmann, M. Drohmann and B. Haasdonk, Model reduction of parametrized evolution problems using the reduced basis method with adaptive time partitioning. In: *International Conference on Adaptive Modeling and Simulation* (2011) 156–167.

[17] K.J. Engel and R. Nagel, One-Parameter Semigroups for Linear Evolution Equations. *Graduate Texts in Mathematics*. Springer, New York, NY, USA (2000).

[18] P.A. Etter and K.T. Carlberg, Online adaptive basis refinement and compression for reduced-order models via vector-space sieving. Preprint arXiv:1902.10659v2 (2019).

[19] B. Fornberg, Generation of finite difference formulas on arbitrarily spaced grids. *Math. Comput.* **51** (1988) 699–706.

[20] E. Fosong, P. Schulze and B. Unger, From time-domain data to low-dimensional structured models. Preprint arXiv:1902.05112 (2019).

[21] R.J. Gelinas, S.K. Doss and K. Miller, The moving finite element method: applications to general partial differential equations with multiple large gradients. *J. Comput. Phys.* **40** (1981) 202–249.

[22] J.-F. Gerbeau and D. Lombardi, Approximated Lax pairs for the reduced order integration of nonlinear evolution equations. *J. Comput. Phys.* **265** (2014) 246–269.

[23] S. Glavaski, J.E. Marsden and R.M. Murray, Model reduction, centering, and the Karhunen–Loeve expansion. In: Vol. 2 of *Proceedings of the 37th IEEE Conference on Decision and Control.* Tampa, FL, USA (1998) 2071–2076.

[24] H. Goldberg, W. Kampowsky and F. Tröltzsch, On Nemytskij operators in L_p-spaces of abstract functions. *Math. Nachr.* **155** (1992) 127–140.

[25] I.V. Gosea, M. Petreczky, A.C. Antoulas and C. Fiter, Balanced truncation for linear switched systems. *Adv. Comput. Math.* **44** (2018) 1845–1886.

[26] C. Greif and K. Urban, Decay of the Kolmogorov n-width for wave problems. *Appl. Math. Lett.* **96** (2019) 216–222.

[27] M.A. Grepl, *Reduced-basis approximations and a posteriori error estimation for parabolic partial differential equations*. Ph.D. thesis, Massachusetts Institute of Technology (2005).

[28] C. Gu, QLMOR: a projection-based nonlinear model order reduction approach using quadratic-linear representation of nonlinear systems. *IEEE Trans. Comput. Aided Des. Integr. Circuits Syst.* **30** (2011) 1307–1320.

[29] M. Gubisch and S. Volkwein, Proper orthogonal decomposition for linear-quadratic optimal control, chapter 1, edited by P. Benner, A. Cohen, M. Ohlberger and K. Willcox. In: Model Reduction and Approximation. SIAM, Philadelphia, PA, USA (2017) 3–63.

[30] B. Haasdonk, Convergence rates of the POD – greedy method. *ESAIM: M2AN* **47** (2013) 859–873.

[31] B. Haasdonk and M. Ohlberger, Reduced basis method for finite volume approximations of parametrized linear evolution equations. *ESAIM: M2AN* **42** (2008) 277–302.

[32] D. Hartman and L.K. Mestha, A deep learning framework for model reduction of dynamical systems. In: *IEEE Conference on Control Technology and Applications (CCTA). Kohala Coast, HI, USA* (2017) 1917–1922.

[33] J.S. Hesthaven, G. Rozza and B. Stamm, Certified Reduced Basis Methods for Parametrized Partial Differential Equations. *Springer Briefs in Mathematics.* Springer, Cham, Switzerland (2016).

[34] A. Hirschberg and S. Rienstra, Theoretical background: aeroacoustics, In: Large-Eddy Simulation for Acoustics, edited by C.A. Wagner, T. Hüttl and P. Sagaut. Cambridge University Press, Cambridge, UK (2007) 24–88.

[35] P. Holmes, J.L. Lumley, G. Berkooz and C.W. Rowley, Turbulence, Coherent Structures, Dynamical Systems and Symmetry, 2nd edition. *Cambridge Monographs on Mechanics.* Cambridge University Press, New York, NY, USA (2012).

[36] A. Iollo and D. Lombardi, Advection modes by optimal mass transfer. *Phys. Rev. E* **89** (2014) 022923.

[37] A. Janon, M. Nodet and C. Prieur, Certified reduced-basis solutions of viscous Burgers equation parametrized by initial and boundary values. *ESAIM: M2AN* **47** (2013) 317–348.

[38] E.N. Karatzas, F. Ballarin and G. Rozza, Projection-based reduced order models for a cut finite element method in parametrized domains. Preprint arXiv:1901.03846v1 (2019).

[39] K. Kashima, Nonlinear model reduction by deep autoencoder of noise response data. In: *55th IEEE Conference on Decision and Control (CDC). Las Vegas, USA* (2016) 5750–5755.

[40] A. Kolmogoroff, Über die beste Annäherung von Funktionen einer gegebenen Funktionenklasse. *Ann. Math.* **37** (1936) 107–110.

[41] B. Kramer and K.E. Willcox, Nonlinear model order reduction via lifting transformations and proper orthogonal decomposition. *AIAA J.* **57** (2019) 2297–2307.

[42] P. Kunkel and V. Mehrmann, Differential-Algebraic Equations. Analysis and Numerical Solution. European Mathematical Society, Zürich, Switzerland (2006).

[43] K. Lee and K.T. Carlberg, Model reduction of dynamical systems on nonlinear manifolds using deep convolutional autoencoders. *J. Comput. Phys.* **404** (2019) 108973.

[44] O. Lopes, Stability and forced oscillations. *J. Math. Anal. Appl.* **55** (1976) 686–698.

[45] Y. Maday, A.T. Patera and G. Turinici, Global a priori convergence theory for reduced-basis approximations of single-parameter symmetric coercive elliptic partial differential equations. *C. R. Math. Acad. Sci. Paris* **335** (2002) 289–294.

[46] Y. Maday, A.T. Patera and G. Turinici, A priori convergence theory for reduced-basis approximations of single-parameter elliptic partial differential equations. *J. Sci. Comput.* **17** (2002) 437–446.

[47] A. Mendible, S.L. Brunton, A.Y. Aravkin, W. Lowrie and J.N. Kutz, Dimensionality reduction and reduced order modeling for traveling wave physics. Preprint arXiv:1911.00565v1 (2019).

[48] K. Miller and R.N. Miller, Moving finite elements. I. *SIAM J. Numer. Anal.* **18** (1981) 1019–1032.

[49] R. Mojgani and M. Balajewicz, Lagrangian basis method for dimensionality reduction of convection dominated nonlinear flows. Preprint arXiv:1701.04343v1 (2017).

[50] S. Mowlavi and T.P. Sapsis, Model order reduction for stochastic dynamical systems with continuous symmetries. *SIAM J. Sci. Comput.* **40** (2018) A1669–A1695.

[51] N.J. Nair and M. Balajewicz, Transported snapshot model order reduction approach for parametric, steady-state fluid flows containing parameter-dependent shocks. *Int. J. Numer. Methods Eng.* **117** (2019) 1234–1262.

[52] M. Nonino, F. Ballarin, G. Rozza and Y. Maday, Overcoming slowly decaying Kolmogorov n-width by transport maps: application to model order reduction of fluid dynamics and fluid–structure interaction problems. Preprint arXiv:1911.06598 (2019).

[53] M. Ohlberger and S. Rave, Nonlinear reduced basis approximation of parameterized evolution equations via the method of freezing. *C. R. Math. Acad. Sci. Paris* **351** (2013) 901–906.

[54] A. Pazy, Semigroups of Linear Operators and Applications to Partial Differential Equations. *Applied Mathematical Sciences.* Springer, New York, NY, USA (1983).

[55] B. Peherstorfer, Model reduction for transport-dominated problems via online adaptive bases and adaptive sampling. Preprint arXiv:1812.02094 (2018).

[56] A. Pinkus, N-Widths in Approximation Theory. Ergebnisse der Mathematik und ihrer Grenzgebiete. Springer, Heidelberg, Germany (1985).

[57] I. Pontes Duff, C. Poussot-Vassal and C. Seren, $\mathcal{H}_2$-optimal model approximation by input/output-delay structured reduced order models. *Syst. Control Lett.* **117** (2018) 60–67.

[58] A. Quarteroni and G. Rozza, Reduced Order Methods for Modeling and Computational Reduction. In Vol. 9 of *MS&A – Modeling, Simulation and Applications.* Springer, Cham, Switzerland (2014).

[59] A. Quarteroni, A. Manzoni and F. Negri, Reduced Basis Methods for Partial Differential Equations: An Introduction. UNITEXT. Springer, Cham, Switzerland (2016).

[60] J. Reiss, Model reduction for convective problems: formulation and application. *IFAC-PapersOnLine* **51** (2018) 186–189.

[61] J. Reiss, P. Schulze, J. Sesterhenn and V. Mehrmann, The shifted proper orthogonal decomposition: a mode decomposition for multiple transport phenomena. *SIAM J. Sci. Comput.* **40** (2018) A1322–A1344.

[62] D. Rim, S. Moe and R.J. LeVeque, Transport reversal for model reduction of hyperbolic partial differential equations. *SIAM/ASA J. Uncertain. Quantif.* **6** (2018) 118–150.

[63] D. Rim, B. Peherstorfer and K.T. Mandli, Manifold approximations *via* transported subspaces: model reduction for transport-dominated problems. Preprint 1912.13024v2 (2020).

[64] C.W. Rowley and J.E. Marsden, Reconstruction equations and the Karhunen–Loève expansion for systems with symmetry. *Phys. D* **142** (2000) 1–19.

[65] C.W. Rowley, I.G. Kevrekidis, J.E. Marsden and K. Lust, Reduction and reconstruction for self-similar dynamical systems. *Nonlinearity* **16** (2003) 1257–1275.

[66] G. Scarciotti and A. Astolfi, Model reduction for hybrid systems with state-dependent jumps. *IFAC-PapersOnLine* **49** (2016) 850–855.

[67] G. Scarciotti and A. Astolfi, Model reduction of neutral linear and nonlinear time-invariant time-delay systems with discrete and distributed delays. *IEEE Trans. Automat. Contr.* **61** (2016) 1438–1451.

[68] P. Schulze and B. Unger, Data-driven interpolation of dynamical systems with delay. *Syst. Control Lett.* **97** (2016) 125–131.

[69] P. Schulze and B. Unger, Model reduction for linear systems with low-rank switching. *SIAM J. Control Optim.* **56** (2018) 4365–4384.

[70] P. Schulze, B. Unger, C. Beattie and S. Gugercin, Data-driven structured realization. *Linear Algebra Appl.* **537** (2018) 250–286.

[71] P. Schulze, J. Reiss and V. Mehrmann, Model reduction for a pulsed detonation combuster via shifted proper orthogonal decomposition, In: Active Flow and Combustion Control 2018, edited by R. King. Springer, Cham, Switzerland (2019) 271–286.

[72] J. Sesterhenn and A. Shahirpour, A characteristic dynamic mode decomposition. *Theor. Comput. Fluid Dyn.* **33** (2019) 281–305.

[73] A. Singer, Angular synchronization by eigenvectors and semidefinite programming. *Appl. Comput. Harmon. Anal.* **30** (2011) 20–36.

[74] B. Sonday, A. Singer and I.G. Kevrekidis, Noisy dynamic simulations in the presence of symmetry: data alignment and model reduction. *Comput. Math. Appl.* **65** (2013) 1535–1557.

[75] T. Taddei, A registration method for model order reduction: data compression and geometry reduction. Preprint arXiv:1906.11008v1 (2019).

[76] T. Taddei, S. Perotto and A. Quarteroni, Reduced basis techniques for nonlinear conservation laws. *ESAIM:M2AN* **49** (2015) 787–814.

[77] B. Unger, *Impact of discretization techniques on nonlinear model reduction and analysis of the structure of the POD basis.* Master's thesis, Virginia Polytechnic and State University, Blacksburg, Virginia, USA (2013).

[78] B. Unger and S. Gugercin, Kolmogorov n-widths for linear dynamical systems. *Adv. Comput. Math.* **45** (2019) 2273–2286.

[79] K. Urban and A. Patera, An improved error bound for reduced basis approximation of linear parabolic problems. *Math. Comput.* **83** (2014) 1599–1615.

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

[81] E. Zeidler, Nonlinear Functional Analysis and its Applications IIa: Linear Monotone Operators. Springer, New York, NY, USA (1990).