POD-Galerkin Modeling for Incompressible Flows with Stochastic Boundary Conditions

Dem Fachbereich Mathematik
der Technischen Universität Darmstadt
zur Erlangung des Grades eines
Doktors der Naturwissenschaften (Dr. rer. nat.)
genehmigte

Dissertation

von
M. Sc. Sebastian Ullmann
aus Schlema

Referent:	Prof. Dr. Jens Lang
1. Korreferent:	Prof. Dr. Stefan Ulbrich
2. Korreferent:	Prof. Max D. Gunzburger
Tag der Einreichung:	12. Dezember 2013
Tag der mündlichen Prüfung:	5. Februar 2014

2014 Darmstadt D17

to Mia

Abstract

In the context of the numerical solution of parametrized partial differential equations, a proper orthogonal decomposition (POD) provides a basis of a subspace of the solution space. The method relies on a singular value decomposition of a snapshot matrix, which contains the numerical solutions at predefined parameter values. Often a sufficiently accurate representation of the solution can be given by a linear combination of a small number of POD basis functions. In this case, using POD basis functions as test and trial functions in a Galerkin projection leads to POD-Galerkin reduced-order models. Such models are derived and tested in this thesis for flow problems governed by the incompressible Navier-Stokes equations with stochastic Dirichlet boundary conditions.

In the first part of the thesis, POD-Galerkin reduced-order models are developed for unsteady deterministic problems of increasing complexity: heat conduction, isothermal flow, and thermoconvective flow. Here, time acts as a parameter, so that the snapshot matrix consists of discrete solutions at different times. Special attention is paid to the reduced-order computation of the pressure field, which is realized by projecting a discrete pressure Poisson equation onto a pressure POD basis. It is demonstrated that the reduced-order solutions of the considered problems converge toward the underlying snapshots when the dimension of the POD basis is increased.

The second part of the thesis is devoted to a steady thermally driven flow problem with a temperature Dirichlet boundary condition given by a spatially correlated random field. In order to compute statistical quantities of interest, the stochastic problem is split into separate deterministic sub-problems by means of a Karhunen-Loève parametrization of the boundary data and subsequent stochastic collocation on a sparse grid. The sub-problems are solved with suitable POD-Galerkin models. Different methods to handle the parametrized Dirichlet conditions are introduced and compared. The use of POD-Galerkin reduced-order models leads to a significant speed-up of the overall computational process compared to a standard finite element model.

Zusammenfassung

Die Proper Orthogonal Decomposition (POD) ist eine Methode, die für die numerische Berechnung parametrisierter partieller Differentialgleichungen eine Basis eines Unterraum des Lösungsraums bereitstellt. Sie beruht auf der Singulärwertzerlegung einer Snapshot-Matrix, welche aus numerischen Lösungen für ausgewählte Parameterwerte besteht. Oft kann die Lösung ausreichend genau als Linearkombination weniger POD-Basisfunktionen dargestellt werden. Werden in einem solchen Fall die POD-Basisfunktionen als Test- und Ansatzfunktionen in einer Galerkin-Projektion verwendet, führt dies zu einem ordnungsreduzierten POD-Galerkin-Modell. In dieser Arbeit werden solche Modelle hergeleitet und erprobt für Strömungsprobleme, die beschrieben werden durch die inkompressiblen Navier-Stokes-Gleichungen mit stochastischen Dirichlet-Randbedingungen.

Im ersten Teil der Arbeit werden ordnungsreduzierte POD-Galerkin-Modelle entwickelt für instationäre deterministische Szenarien zunehmender Komplexität: Wärmeleitung, isotherme Strömung und thermo-konvektive Strömung. Hier übernimmt die Zeit die Rolle eines Parameters, sodass die Snapshot-Matrix aus diskreten Lösungen zu unterschiedlichen Zeiten besteht. Besondere Beachtung findet die ordnungsreduzierte Berechnung des Druckfelds, welche mithilfe der Projektion einer diskreten Druck-Poisson-Gleichung auf eine POD-Basis des Drucks realisiert wird. Es wird gezeigt, dass die ordnungsreduzierten Lösungen der betrachteten Probleme gegen die zugrunde liegenden Snapshots konvergieren, wenn die Dimension der POD-Basis vergrößert wird.

Der zweite Teil der Arbeit widmet sich einer stationären thermo-konvektiven Strömung mit einer Temperatur-Dirichlet-Bedingung, die durch ein räumlich korreliertes Zufallsfeld gegeben ist. Um statistische Zielgrößen zu berechnen, wird das stochastische Problem zunächst in einzelne deterministische Unterprobleme aufgespaltet. Dies wird mittels einer Karhunen-Loève-Parametrisierung der Randdaten und anschließender stochastischer Kollokation auf einem dünn besetzten Gitter erreicht. Die Unterprobleme werden mit geeigneten ordnungsreduzierten POD-Galerkin-Modellen gelöst. Für die

hierbei notwendige Behandlung der parametrisierten Dirichlet-Bedingungen werden verschiedene Ansätze eingeführt und verglichen. Die Verwendung ordnungsreduzierter POD-Galerkin-Modelle führt zu einer signifikanten Beschleunigung des gesamten Berechnungsprozesses im Vergleich zu herkömmlichen Finite-Elemente-Modellen.

Acknowledgements

First and foremost, I wish to thank my supervisor Jens Lang for his enthusiasm, guidance and encouragement along the way that has lead to this thesis. Thank you for your confidence in my abilities and your helpful advice.

Thank you, Bettina Schieche, for introducing me to stochastic collocation and for your comments on the second part of this thesis. Your expertise has helped me a lot in stepping from the certain to the uncertain world. Furthermore, I would like to thank Alf Gerisch, Karen Kuhn, Jan-Frederik Pietschmann, Matthias Schlottbom and Jens Wehner for proof-reading various parts of this thesis. My thanks go to the whole research group Numerical Analysis and Scientific Computing for a stimulating research atmosphere and inspiring lunch discussions.

I gratefully acknowledge the financial support by the Deutsche Forschungsgemeinschaft. This research has been conducted under the grants SFB 568, EXC 259 and SPP 1276.

I want to express my gratitude to my parents, Simona and Tobias Ullmann, who have encouraged me in all stages of my studies. Finally, I want to thank my wife Maria for her unshakable belief in a successful finish of this project. I am grateful that you share your life with me.

Contents

Chapter 1

Introduction

1.1 Galerkin reduced-order modeling

Galerkin's method is often used for the numerical solution of boundary value problems involving partial differential equations (PDEs). It requires the considered PDE problem to be presented in a variational form, which can be obtained by multiplying the governing equations with suitable test functions and integrating over the domain. When the unknown fields are represented as linear combinations of test functions, this leads to a system of equations – the Galerkin model – with the coefficients of the linear combinations as unknowns.

Initial-boundary value PDE problems can be solved, for instance, by the vertical method of lines. In this approach, Galerkin's method is applied only to the spatial dimensions of the problem. The resulting system of ordinary differential equations (ODEs) can be solved with suitable time integration schemes.

While the choice of the variational form and the handling of the time dependency already lead to a variety of different Galerkin approaches for a single PDE problem, the test functions must be chosen as well. The standard finite element method is characterized by test functions which are piecewise polynomial functions within the cells of some computational mesh, continuous over the cell edges and non-zero only within a small number of cells. The spatial resolution of a finite element model can be adapted to the problem of interest. For instance, the underlying mesh can be locally refined (h-adaptivity), the polynomial degree can be locally adjusted (p-adaptivity), or the mesh points can be relocated (r-adaptivity). Because the finite element

method relies on Galerkin's method, the discrete finite element solution space is equal to the span of the test functions. This solution space can become quite high-dimensional for some applications, even if adaptivity is employed.

Compared to a standard finite element model, a Galerkin reduced-order model relies on test functions which are more closely linked to the solutions of the considered PDE problem. The particular test functions considered in this thesis are computed from a set of predetermined numerical solutions. In this context, the test functions are called *reduced basis functions* and the underlying numerical solutions are called *snapshots*. The following list presents some types of basis functions, which have been used to create reduced-order Galerkin models, previously:

- Lagrange: A set of snapshots is taken as basis functions (Peterson, 1989).
- Taylor: A single snapshot and its derivatives with respect to a parameter are taken as basis functions (Noor and Peters, 1980).
- Hermite: A set of snapshots and their first derivatives with respect to a parameter are taken as basis functions (Ito and Ravindran, 1998).
- Proper orthogonal decomposition (POD): For a set of snapshots, orthonormal basis functions are computed which minimize the average energy of the error between the snapshots and their orthogonal projection on the basis (Holmes et al., 1996).
- Centroidal Voronoi tessellation (CVT): A set of snapshots is arranged in clusters, so that the snapshots are close to the respective cluster centers in an average sense. The cluster centers are taken as basis functions (Burkardt et al., 2006b). Comparisons between POD and CVT reduced-order models can be found in Burkardt et al. (2006a) for the flow in a T-cell and in Ullmann and Lang (2012) and Ullmann et al. (2013) for the flow around a cylinder.
- Greedy: The basis is built by iteratively adding basis functions. In each iteration a parameter domain is probed with an error estimator. A snapshot is computed at the position with the maximum estimated error. The resulting snapshots are used as basis functions in a Lagrangian fashion (Prud'homme et al., 2002a).
- POD-Greedy: For time-dependent parametric problems, in each step of a greedy method an error estimator is used to choose a snapshot position in the parameter domain. A time series of snapshots is generated at

this position. A POD is applied to these snapshot, and the leading POD basis function is added to the POD-Greedy basis (Haasdonk and Ohlberger, 2008).

Because of the cost associated with the computation of the snapshots, it is often computationally more expensive to *create* a Galerkin reduced-order model than to *create and solve* a respective finite element model. However, there are PDE problems whose solutions can be approximated well with a reduced-order model that can be *solved* many times faster than a respective finite element model. Therefore, Galerkin reduced-order modeling can decrease the total computation time for problems which have to be solved repeatedly, e.g. for different parameter values. The application of POD-Galerkin modeling to optimization and control with PDE constraints, for instance, has lead to the development of the trust-region POD (TR-POD) method and the optimality system POD (OS-POD) method, see Sachs and Volkwein (2010) for an overview. Another area of application, examined in the second part of this thesis, is uncertainty quantification for PDE problems with parametrized random inputs. The stochastic collocation method relies on deterministic simulations for given realizations of the random variables to compute the statistics of a quantity of interest. In this framework, expensive finite element computations can possibly be replaced by solutions of a POD-Galerkin surrogate model.

1.2 Outline of the thesis

Chapter 2 presents the proper orthogonal decomposition as the solution of a minimization problem. It is shown that the POD of a set of finite element snapshots can be computed with the help of a singular value decomposition. By modifying the POD basis functions or the underlying snapshots, it is possible to obtain functions which can be used in a Galerkin projection for problems involving Dirichlet boundary conditions.

In chapter 3, POD-Galerkin modeling is introduced for an unsteady nonlinear heat conduction problem. First, a numerical discretization by means of quadratic finite elements in space and the Crank-Nicolson method in time is presented and validated. The purpose of the finite element model is to generate a snapshot set and a reference solution for subsequent reduced-order modeling. Several approaches to derive a respective POD-Galerkin model are presented. Their equivalence is shown under the condition that the finite element discretization schemes are used in the set-up of the reduced-order model equations. The heat conduction in a square, proposed in Gunzburger

et al. (2007), is used as a test case.

The subject of chapter 4 is POD-Galerkin modeling for an unsteady isothermal flow problem governed by the incompressible Navier-Stokes equations. A finite element model is introduced, employing Taylor-Hood elements in connection with the Crank-Nicolson method. Relying on the discrete divergence-freeness of the velocity POD basis functions, a POD-Galerkin model for the velocity is derived. An auxiliary pressure model is developed for the purpose of computing pressure-related quantities of interest. A validation is performed for the flow around a cylinder, as described in Posdziech and Grundmann (2007).

Based on the material presented in the previous chapters, in chapter 5 an unsteady thermoconvective incompressible flow problem is investigated. The reduced-order model of chapter chapter 4 is complemented with a convection-diffusion equation for the temperature and with coupling terms in the equations for the velocity and the pressure. The model is validated for the test case of Evans and Paolucci (1990), which features the flow in a duct heated from below.

Chapter 6 introduces the sparse grid stochastic collocation method for steady PDE problems with boundary conditions given by spatially correlated random fields. The Karhunen-Loève expansion is presented as a means to parametrize the random field. The resulting multi-dimensional stochastic problem is transformed a set of independent deterministic problems by applying a collocation method. Smolyak's algorithm is introduced for the creation of collocation points which are suited for problems with a moderate number of stochastic parameters. It is explained how reduced-order modeling can be used to accelerate the sparse grid stochastic collocation method.

In chapter 7, the POD-aided stochastic collocation method of the previous chapter is applied to the steady thermoconvective incompressible flow driven by stochastic Dirichlet conditions. The presentation of the finite element and reduced-order models relies largely on chapter 5. However, dependency on time is replaced by dependency on the random parameters. The implementation of the stochastic Dirichlet conditions relies on the techniques presented in chapter 2. The accuracy and efficiency of the POD-aided stochastic collocation method is assessed for the problem proposed by Ganapathysubramanian and Zabaras (2007).

1.3 Main contributions

A modification of the method of Akhtar et al. (2009) is presented for the computation of pressure-related quantities in the reduced-order models involving

incompressible flow. In the original method of Akhtar et al., a continuous pressure Poisson equation is projected onto a pressure POD basis. Such a reduced-order pressure model may lead to accurate results when the snapshot computation involves a discretized continuous pressure Poisson equation, like in Akhtar et al. (2009). However, it is not consistent with a snapshot computation employing a discretized continuity equation, like in this thesis. As a remedy, in the modified method a discrete pressure Poisson equation according to Gresho and Sani (2000) is projected onto a set of pressure POD basis vectors. The reduced-order pressure fields resulting from the new pressure model converge toward the underlying pressure snapshots as the number of basis functions is increased. This is a consequence of the numerics of the reduced-order model being consistent with the numerics of the underlying finite element simulation.

The boundary treatment presented in Gunzburger et al. (2007) is generalized in order to be applicable to problems with parametrized stochastic Dirichlet boundary conditions. The method of Gunzburger et al. recombines the POD basis functions so that homogeneous Dirichlet conditions are enforced at selected boundary points. This leads to modified POD basis functions which fulfill homogeneous Dirichlet conditions under the following conditions: Firstly, the Dirichlet boundary must consist of non-overlapping segments. Secondly, the Dirichlet data on each segment must be given by a spatial functions times a possible time-dependent parameter. Additionally, the original method requires the selection of one boundary point per segment, which corresponds to inhomogeneous Dirichlet data unless homogeneous conditions are prescribed at the complete segment. The generalized method builds on the idea that homogeneous Dirichlet conditions can be enforced directly on the union of all Dirichlet boundaries. This leads to a relaxation of the necessary conditions. In particular, the Dirichlet data on the union of all boundary segments must be expressible as a linear combination of spatial functions with possibly time-dependent coefficients. This exactly corresponds to the way in which random fields are presented after a Karhunen-Loève expansion has been applied. Moreover, in the new method it is not necessary anymore to select certain points at which homogeneous conditions are enforced.

A new POD-aided stochastic collocation method is introduced for the discretization of the stochastic dimensions resulting from parametrized random fields. The method relies on a POD-Galerkin reduced-order model created from snapshots at the collocation points of a coarse sparse grid. The reduced-order model is evaluated at the collocation points of a finer sparse grid. Finally, the interpolant on the fine grid is used for uncertainty quantification. In essence, a collocation on a fine stochastic grid is performed at the cost of

finite element solutions at the collocation points of a rough stochastic grid. The method differs from the recently published approach of Chen et al. (2012) in the following aspects: Chen et al. locate the snapshots in the stochastic parameter domain with a greedy algorithm, while in the present thesis the snapshot locations are determined by a coarse sparse grid. Furthermore, Chen et al. use the orthogonalized snapshots as a basis while the present method employs a truncated POD of the snapshots.

1.4 Mathematical notation

For some bounded domain $\Omega \subset \mathbb{R}^2$ and two functions $a, b : \Omega \to \mathbb{R}$, the $L^2(\Omega)$ inner product is defined as

$$(a, b)_{L^2(\Omega)} = \int_\Omega a(\vec{x}) b(\vec{x}) \, \mathrm{d}\vec{x},$$

where the vector $\vec{x} = (x, y)^T$ denotes a position in $\mathbb{R}^2$ and the integral is interpreted in the Lebesgue sense. The $L^2(\Omega)$ norm is specified as

$$\|a\|_{L^2(\Omega)} = (a, a)^{\frac{1}{2}}_{L^2(\Omega)}$$

and the $L^2(\Omega)$ function space is given by

$$L^2(\Omega) = \{a : \|a\|_{L^2(\Omega)} < \infty\}.$$

Wherever inner products and norms are written without subscripts, $L^2(\Omega)$ is implied. Also, where the domain of definition is not of central importance, L^2 is written instead of $L^2(\Omega)$. The definitions are extended to vector-valued variables without change of notation: For $\vec{a} = (a_1, \ldots, a_D)^T$ and $\vec{b} = (b_1, \ldots, b_D)^T$ with $D \geq 1$ it is defined that

$$(\vec{a}, \vec{b}) = \sum_{d=1}^{D} (a_d, b_d), \qquad \|\vec{a}\| = (\vec{a}, \vec{a})^{\frac{1}{2}}.$$

In the formulation of the Galerkin methods, weak first derivatives are applied frequently. Let $C_0^1(\Omega)$ be the linear space of all functions ϕ which have a compact support in Ω, which are continuously differentiable in Ω and whose partial derivatives $\partial_x \phi$ and $\partial_y \phi$ are continuous on the closure $\bar{\Omega}$. For some not necessarily continuous function $a \in L^2(\Omega)$, the functions $\partial_x a$ and

$\partial_y a$ are called weak first derivatives of a if, respectively,

$$\int_\Omega \partial_x a \phi \, \mathrm{d}\Omega = -\int_\Omega a \partial_x \phi \, \mathrm{d}\Omega, \qquad \forall \phi \in C_0^1(\Omega),$$
$$\int_\Omega \partial_y a \phi \, \mathrm{d}\Omega = -\int_\Omega a \partial_y \phi \, \mathrm{d}\Omega, \qquad \forall \phi \in C_0^1(\Omega).$$

The Sobolev space $H^1(\Omega)$ is defined as the linear space of functions $a \in L^2(\Omega)$ whose weak derivatives fulfill $\partial_x a \in L^2(\Omega)$ and $\partial_y a \in L^2(\Omega)$. More details about Sobolev spaces and their use in the context of finite element methods can be found in Larsson and Thomée (2003) and Knabner and Angermann (2003).

Part I

Deterministic problems

Chapter 2

Proper orthogonal decomposition

The proper orthogonal decomposition (POD) is employed in many disciplines of applied mathematics, albeit in different mathematical settings and under different names, e.g. 'principal component analysis' (statistical analysis), 'empirical orthogonal functions' (meteorology) and 'Karhunen-Loève expansion' (stochastics). The principle of POD is to approximate a large number of vectors or fields by a linear combination of a small number of representatives, which are chosen with respect to optimality and orthonormality constraints. Such low-dimensional representations can often be analyzed much easier than the raw data. For example, the POD has been used to extract coherent structures from turbulent flow fields. The study of the interactions between these structures has originally led to the development of POD-Galerkin reduced-order flow models (Sirovich, 1987).

In this chapter the proper orthogonal decomposition is first introduced as an approximation problem in the L^2 function space. In case the considered functions are members of a finite element subspace, an equivalent discrete approximation problem for the respective finite element coefficient vectors is derived in section 2.1. It is proved in section 2.2 that solutions of the discrete problem can be computed by means of a singular value decomposition. Section 2.3 presents modifications of the POD which lead to low-dimensional representations which exactly fulfill inhomogeneous Dirichlet conditions.

2.1 Formulation in the finite element context

The POD approximation problem is first presented in the classical $L^2(\Omega)$ setting for a closed domain $\Omega \in \mathbb{R}^d$, where d is a small positive integer (Holmes et al., 1996):

Problem 1. *For given $u_1, \dots, u_N \in L^2(\Omega)$, find functions $\phi_1, \dots, \phi_N \in L^2(\Omega)$ and coefficients $B_1^1, \dots, B_1^N, \dots, B_N^1, \dots, B_N^N \in \mathbb{R}$ which minimize*

$$J_R = \sum_{n=1}^{N} \left\| u_n - \sum_{r=1}^{R} \tilde{\phi}_r \tilde{B}_n^r \right\|_{L^2(\Omega)}^2, \tag{2.1}$$

for all $R = 1, \dots, N$ under the constraints

$$(\phi_i, \phi_j)_{L^2(\Omega)} = \delta_{ij}, \qquad i, j = 1, \dots, N. \tag{2.2}$$

In the literature, $u_1, \dots, u_N$ are typically called snapshots, $\phi_1, \dots, \phi_N$ are called POD basis functions and $B_1^1, \dots, B_1^N, \dots, B_N^1, \dots, B_N^N$ are called POD coefficients. Also, J_R is often called POD energy.

Throughout this thesis, the snapshots and POD basis functions are assumed to be members of a finite element space. This assumption transforms problem 1 into a discrete problem which can be solved with a computer. To this end, let $\vec{x}_1, \dots, \vec{x}_{M'} \in \bar{\Omega}$ be a set of distinct mesh points and let $\psi_1, \dots, \psi_{M'} : \bar{\Omega} \to \mathbb{R}$ be respective continuous finite element basis functions with $\psi_i(\vec{x}_j) = \delta_{ij}$ for $i, j = 1, \dots, M'$. A finite element subspace $L_h^2(\Omega) \subset L^2(\Omega)$ is defined as the span of the finite element basis functions, $L_h^2(\Omega) = \text{span}(\psi_1, \dots, \psi_{M'})$. Therefore, any snapshots $u_1, \dots, u_N \in L_h^2$ can be written as a linear combination of finite element basis functions,

$$u_n(\vec{x}) = \sum_{m=1}^{M'} \psi_m(\vec{x}) U_n^m, \qquad n = 1, \dots, N, \qquad \vec{x} \in \bar{\Omega}, \tag{2.3}$$

where $U_1^1, \dots, U_1^{M'}, \dots, U_N^1 \dots, U_N^{M'} \in \mathbb{R}$ are called finite element coefficients. The coefficients can be arranged in vector form,

$$U_n = \begin{pmatrix} U_n^1 \\ \vdots \\ U_n^{M'} \end{pmatrix}, \qquad n = 1, \dots, N, \tag{2.4}$$

$$U^m = (U_1^m, \dots, U_N^m), \qquad m = 1, \dots, M', \tag{2.5}$$

or in matrix form,

$$\mathcal{U} = (U_1, \ldots, U_N) = \begin{pmatrix} U^1 \\ \vdots \\ U^{M'} \end{pmatrix} = \begin{pmatrix} U_1^1 & \ldots & U_N^1 \\ \vdots & & \vdots \\ U_1^{M'} & \ldots & U_N^{M'} \end{pmatrix}. \tag{2.6}$$

A mass matrix $\mathcal{M} \in \mathbb{R}^{M' \times M'}$ is defined as the matrix that contains the L^2 inner products (ψ_i, ψ_j) in the ith row and jth column for each $i, j \in \{1, \ldots, M'\}$. The matrix $\mathcal{M}$ is symmetric and positive definite. The symmetry follows from the definition of $\mathcal{M}$ and the symmetry of the inner product. To show the positive definiteness (see Larsson and Thomée, 2003), one can write

$$U_k^T \mathcal{M} U_k = \sum_{i,j=1}^{M'} U_k^i (\psi_i, \psi_j) U_k^j = \left(\sum_{i=1}^{M'} \psi_i U_k^i, \sum_{j=1}^{M'} \psi_j U_k^j \right) = \left\| \sum_{i=1}^{M'} \psi_i U_k^i \right\|^2 \geq 0$$

for an arbitrary column vector $U_k \in \mathbb{R}^{M'}$. Because the functions $\psi_1, \ldots, \psi_{M'}$ are assumed to form a basis of a linear space, they are linearly independent. Therefore the left side of the inequality is zero if and only if U_k is a zero vector. Consequently, $\mathcal{M}$ is positive definite.

The L^2 inner product of any two finite element snapshots $u_k, u_l \in L_h^2$ for $k, l \in \{1, \ldots, N\}$ can be expressed via their finite element coefficients and the finite element mass matrix:

$$(u_k, u_l) = \left(\sum_{i=1}^{M'} \psi_i U_k^i, \sum_{j=1}^{M'} \psi_j U_l^j \right) = U_k^T \mathcal{M} U_l = (U_k, U_l)_{\mathcal{M}}.$$

Hence the L^2 norm of any snapshot $u_k \in L_h^2$ for $k \in \{1, \ldots, N\}$ can be defined via

$$\|u_k\|^2 = (u_k, u_k) = U_k^T \mathcal{M} U_k = (U_k, U_k)_{\mathcal{M}} = \|U_k\|_{\mathcal{M}}^2.$$

The symmetry and positive definiteness of the matrix $\mathcal{M}$ justify that $(\cdot, \cdot)_{\mathcal{M}}$ is called the $\mathcal{M}$ inner product and $\|\cdot\|_{\mathcal{M}}$ is called the $\mathcal{M}$ norm. The notation presented here for the finite element snapshots can be applied to arbitrary elements of L_h^2.

The continuous POD problem has been introduced as problem 1 above. In case the snapshots and basis functions are members of L_h^2, an equivalent discrete POD problem can be formulated using the finite element coefficient vectors and the mass matrix $\mathcal{M}$. The finite element notation is first introduced for

an extended set of basis functions $\phi_1, \ldots, \phi_{M'} \in L_h^2$. Indeed, for any set of basis functions $\phi_1, \ldots, \phi_N \in L_h^2$ with $(\phi_i, \phi_j) = \delta_{ij}$ for $i, j = 1, \ldots, N$ it is possible to find additional basis functions $\phi_{N+1}, \ldots, \phi_{M'} \in L_h^2$ via a Gram-Schmidt procedure (Leon et al., 2013), so that $(\phi_i, \phi_j) = \delta_{ij}$ for $i, j = 1, \ldots, M'$. A finite element representation of the extended set of basis functions is given as

$$\phi_n(\vec{x}) = \sum_{m=1}^{M'} \psi_m(\vec{x}) \Phi_n^m, \qquad n = 1, \ldots, M', \qquad \vec{x} \in \bar{\Omega}, \tag{2.7}$$

for suitable finite element coefficients $\Phi_1^1, \ldots, \Phi_1^{M'}, \ldots, \Phi_{M'}^1 \ldots, \Phi_{M'}^{M'}$. The coefficients can be arranged in vector form,

$$\Phi_n = \begin{pmatrix} \Phi_n^1 \\ \vdots \\ \Phi_n^{M'} \end{pmatrix}, \qquad n = 1, \ldots, M', \tag{2.8}$$

$$\Phi^m = (\Phi_1^m, \ldots, \Phi_{M'}^m), \qquad m = 1, \ldots, M', \tag{2.9}$$

or in matrix form,

$$\Phi = (\Phi_1, \ldots, \Phi_{M'}) = \begin{pmatrix} \Phi^1 \\ \vdots \\ \Phi^{M'} \end{pmatrix} = \begin{pmatrix} \Phi_1^1 & \ldots & \Phi_{M'}^1 \\ \vdots & & \vdots \\ \Phi_1^{M'} & \ldots & \Phi_{M'}^{M'} \end{pmatrix}. \tag{2.10}$$

The same type of notation is applied to general functions $\tilde{\phi}_1, \ldots, \tilde{\phi}_{M'} \in L_{\tilde{h}}^2$.

Substituting the finite element representations of $u_1, \ldots, u_n$ and $\tilde{\phi}_1, \ldots, \tilde{\phi}_N$ into the POD functionals J_R, see (2.1), for $R = 1, \ldots, N$ leads to

$$\begin{aligned}
J_R &= \sum_{n=1}^{N} \left\| u_n - \sum_{r=1}^{R} \tilde{\phi}_r \tilde{B}_n^r \right\|^2 = \sum_{n=1}^{N} \Big(u_n - \sum_{r=1}^{R} \tilde{\phi}_r \tilde{B}_n^r, u_n - \sum_{r=1}^{R} \tilde{\phi}_r \tilde{B}_n^r \Big) \\
&= \sum_{n=1}^{N} \left(\sum_{k=1}^{M'} \psi_k \Big(U_n^k - \sum_{r=1}^{R} \tilde{\Phi}_r^k \tilde{B}_n^r \Big), \sum_{l=1}^{M'} \psi_l \Big(U_n^l - \sum_{r=1}^{R} \tilde{\Phi}_r^l \tilde{B}_n^r \Big) \right) \\
&= \sum_{n=1}^{N} \sum_{k,l=1}^{M'} \Big(U_n^k - \sum_{r=1}^{R} \tilde{\Phi}_r^k \tilde{B}_n^r \Big) (\psi_k, \psi_l) \Big(U_n^l - \sum_{r=1}^{R} \tilde{\Phi}_r^l \tilde{B}_n^r \Big) \\
&= \sum_{n=1}^{N} \Big(U_n - \sum_{r=1}^{R} \tilde{\Phi}_r \tilde{B}_n^r \Big)^T \mathcal{M} \Big(U_n - \sum_{r=1}^{R} \tilde{\Phi}_r \tilde{B}_n^r \Big) = \sum_{n=1}^{N} \left\| U_n - \sum_{r=1}^{R} \tilde{\Phi}_r \tilde{B}_n^r \right\|_{\mathcal{M}}^2
\end{aligned}$$

and a substitution of the finite element representations into the orthogonality constraints, see (2.2), for any $i, j = 1, \ldots, N$ leads to

$$\delta_{ij} = (\tilde{\phi}_i, \tilde{\phi}_j) = \left(\sum_{k=1}^{M'} \psi_k \tilde{\Phi}_i^k, \sum_{l=1}^{M'} \psi_l \tilde{\Phi}_j^l \right) = \sum_{k,l=1}^{M'} \tilde{\Phi}_i^k (\psi_k, \psi_l) \tilde{\Phi}_j^l = (\tilde{\Phi}_i, \tilde{\Phi}_j)_{\mathcal{M}}.$$

These finite element versions of the POD functionals and orthogonality constraints are substituted in problem 1. It is taken into account that at most M' mutually $\mathcal{M}$ orthogonal members of $\mathbb{R}^{M'}$ can be found. This leads to the finite element formulation of the POD approximation problem:

Problem 2. *For given $U_1, \ldots, U_N \in \mathbb{R}^{M'}$, find vectors $\Phi_1, \ldots, \Phi_L \in \mathbb{R}^{M'}$ and coefficients $B_1^1, \ldots, B_1^L, \ldots, B_N^1, \ldots, B_N^L \in \mathbb{R}$ which minimize*

$$J_R = \sum_{n=1}^{N} \left\| U_n - \sum_{r=1}^{R} \tilde{\Phi}_r \tilde{B}_n^r \right\|_{\mathcal{M}}^2 \tag{2.11}$$

for all $R = 1, \ldots, L$ and $L = \min(M', N)$ under the constraints

$$(\Phi_i, \Phi_j)_{\mathcal{M}} = \delta_{ij}, \qquad i, j = 1, \ldots, L. \tag{2.12}$$

Here, $U_1, \ldots, U_N$ are called snapshot vectors and $\Phi_1, \ldots, \Phi_L$ are called POD basis vectors.

For any set of basis vectors, optimal coefficients are provided by the $\mathcal{M}$ orthogonal projection of the snapshots onto the space spanned by the basis vectors:

Proposition 1. *Given $U_1, \ldots, U_N \in \mathbb{R}^{M'^q}$ and $L = \min(M', N)$, the coefficients*

$$B_n^r = (\tilde{\Phi}_r, U_n)_{\mathcal{M}}, \qquad r = 1, \ldots, L, \qquad n = 1, \ldots, N. \tag{2.13}$$

minimize the functionals $J_1, \ldots, J_L$ of (2.11) for any vectors $\tilde{\Phi}_1, \ldots, \tilde{\Phi}_L \in \mathbb{R}^{M'}$ satisfying $\tilde{\Phi}_i \mathcal{M} \tilde{\Phi}_j = \delta_{ij}$ for all $i, j = 1, \ldots, L$.

Proof. The proof is carried over from Meyer (2000, p. 435–436) and Trefethen and Bau (1997, p. 80–81) to the current setting. For any $R \in \{1, \ldots, L\}$ and arbitrary $\tilde{B}_1^1, \ldots, \tilde{B}_1^L, \ldots, \tilde{B}_N^1, \ldots, \tilde{B}_N^L \in \mathbb{R}$, the functional can be written as

$$\begin{aligned} J_R &= \sum_{n=1}^{N} \left\| U_n - \sum_{r=1}^{R} \tilde{\Phi}_r \tilde{B}_n^r \right\|_{\mathcal{M}}^2 \\ &= \sum_{n=1}^{N} \left\| U_n - \sum_{r=1}^{R} \tilde{\Phi}_r (\tilde{\Phi}_r, U_n)_{\mathcal{M}} + \sum_{r=1}^{R} \tilde{\Phi}_r (\tilde{\Phi}_r, U_n)_{\mathcal{M}} - \sum_{r=1}^{R} \tilde{\Phi}_r \tilde{B}_n^r \right\|_{\mathcal{M}}^2 . \end{aligned}$$

From the orthonormality constraints it follows that $U_n - \sum_{r=1}^{R} \tilde{\Phi}_r(\tilde{\Phi}_r, U_n)_{\mathcal{M}}$ is the projection of U_n onto the orthogonal complement of $\text{span}(\tilde{\Phi}_1, \ldots, \tilde{\Phi}_R)$. The orthogonality allows the application of the Pythagorean theorem:

$$\begin{aligned} J_R &= \sum_{n=1}^{N} \left\| U_n - \sum_{r=1}^{R} \tilde{\Phi}_r(\tilde{\Phi}_r, U_n)_{\mathcal{M}} \right\|_{\mathcal{M}}^2 + \left\| \sum_{r=1}^{R} \tilde{\Phi}_r(\tilde{\Phi}_r, U_n)_{\mathcal{M}} - \sum_{r=1}^{R} \tilde{\Phi}_r \tilde{B}_n^r \right\|_{\mathcal{M}}^2 \\ &\geq \sum_{n=1}^{N} \left\| U_n - \sum_{r=1}^{R} \tilde{\Phi}_r(\tilde{\Phi}_r, U_n)_{\mathcal{M}} \right\|_{\mathcal{M}}^2 . \end{aligned}$$

Thus, $B_n^r = (\tilde{\Phi}_r, U_n)_{\mathcal{M}}$ for $r = 1, \ldots, L$ and $n = 1, \ldots, N$ minimizes the functionals. □

Applying proposition 1 to problem 2 leaves only the basis vectors as unknowns:

Problem 3. *For given $U_1, \ldots, U_N \in \mathbb{R}^{M'}$, find $\Phi_1, \ldots, \Phi_L \in \mathbb{R}^{M'}$ which minimize*

$$J_R = \sum_{n=1}^{N} \left\| U_n - \sum_{r=1}^{R} \tilde{\Phi}_r(\tilde{\Phi}_r, U_n)_{\mathcal{M}} \right\|_{\mathcal{M}}^2$$

for all $R = 1, \ldots, L$ and $L = \min(M', N)$ under the constraints

$$(\Phi_i, \Phi_j)_{\mathcal{M}} = \delta_{ij}, \qquad \forall i, j = 1, \ldots, L.$$

If solutions $\Phi_1, \ldots, \Phi_L$ to the minimization problem have been found, then POD basis functions $\phi_1, \ldots, \phi_L$ can be formulated via (2.7). Also, coefficients $B_1^1, \ldots, B_1^L, \ldots, B_N^1, \ldots, B_N^L$ can be computed via proposition 1. The following section presents how a singular value decomposition can be employed to solve the minimization problem.

2.2 Computation via singular value decomposition

The singular value decomposition for real rectangular matrices is given in the following theorem. Proofs can be found e.g. in Trefethen and Bau (1997) and other textbooks.

Theorem 1. *For a given matrix* $\hat{\mathcal{U}} \in \mathbb{R}^{M' \times N}$ *there exist orthogonal matrices* $\hat{\Phi} \in \mathbb{R}^{M' \times M'}$ *and* $\mathcal{V} \in \mathbb{R}^{N \times N}$ *such that*

$$\hat{\mathcal{U}} = \hat{\Phi} \Sigma \mathcal{V}^T,$$

where $\Sigma \in \mathbb{R}^{M' \times N}$ *is a diagonal matrix with diagonal entries* $\sigma_1, \ldots, \sigma_L$ *which satisfy* $\sigma_1 \geq \cdots \geq \sigma_R > 0 = \sigma_{R+1} = \cdots = \sigma_L$ *for* $R = \operatorname{rank}(\hat{\mathcal{U}})$ *and* $L = \min(M', N)$.

To include orthogonality with respect to the mass matrix $\mathcal{M}$, a slight modification is necessary, which leads to the following theorem. The symbol $\mathcal{I}$ is used to denote identity matrices of different dimensions, which can be immediately deduced from the context.

Theorem 2. *Let* $\mathcal{M} \in \mathbb{R}^{M' \times M'}$ *be a symmetric and positive definite matrix with a Cholesky decomposition* $\mathcal{M} = \mathcal{C}^T \mathcal{C}$ *and let a singular value decomposition of* $\hat{\mathcal{U}} = \mathcal{C}\mathcal{U}$ *for some matrix* $\mathcal{U} \in \mathbb{R}^{M' \times N}$ *be given as*

$$\hat{\mathcal{U}} = \hat{\Phi} \Sigma \mathcal{V}^T, \qquad \hat{\Phi}^T \hat{\Phi} = \mathcal{I}, \qquad \hat{\Phi} \hat{\Phi}^T = \mathcal{I}, \qquad \mathcal{V}^T \mathcal{V} = \mathcal{I}, \qquad \mathcal{V} \mathcal{V}^T = \mathcal{I}.$$

Then $\hat{\Phi} = \mathcal{C}\Phi$ *gives rise to*

$$\mathcal{U} = \Phi \Sigma \mathcal{V}^T, \quad \Phi^T \mathcal{M} \Phi = \mathcal{I}, \quad \Phi \Phi^T = \mathcal{M}^{-1}, \quad \mathcal{V}^T \mathcal{V} = \mathcal{I}, \quad \mathcal{V} \mathcal{V}^T = \mathcal{I}.$$

Proof. The Cholesky factors are invertible and, therefore, of full rank (see, e.g. Trefethen and Bau, 1997). Therefore, the matrix $\hat{\mathcal{U}}$ has the same rank as $\mathcal{U}$. The decomposition of $\hat{\mathcal{U}}$ and the statements about $\hat{\Phi}$ and $\mathcal{V}$ follow from theorem 1. The other statements follow from substitution and inverting the Cholesky factor $\mathcal{C}$. □

The vector and matrix notations of (2.5)–(2.6) and (2.8)–(2.10) are applied also for $\mathcal{V}$:

$$V_n = \begin{pmatrix} V_n^1 \\ \vdots \\ V_n^N \end{pmatrix}, \qquad n = 1, \ldots, N,$$

$$V^m = (V_1^m, \ldots, V_N^m), \qquad m = 1, \ldots, N,$$

$$\mathcal{V} = (V_1, \ldots, V_N) = \begin{pmatrix} V^1 \\ \vdots \\ V^N \end{pmatrix} = \begin{pmatrix} V_1^1 & \ldots & V_N^1 \\ \vdots & & \vdots \\ V_1^N & \ldots & V_N^N \end{pmatrix}.$$

Taking into account the special structure of the matrix Σ, it is possible to write

$$\mathcal{U} = \sum_{r=1}^{L} \sigma_r \Phi_r V_r^T, \quad U_n = \sum_{r=1}^{L} \sigma_r \Phi_r V_r^n, \quad n = 1, \ldots, N, \quad L = \min(M', N).$$

In the following it is shown that a solution to problem 3 is given by the columns of the matrix Φ of theorem 2. The derivation is based in parts on Volkwein (2008) and Kahlbacher (2006). The work of Kahlbacher refers to Volkwein (2001) for a proof. The derivation in Volkwein (2001), however, builds on Stewart (1993), where the proof of the approximation theorem of Schmidt (1907) is reproduced for a singular value decomposition of a square matrix.

Theorem 3. *Let $\Phi = (\Phi_1, \ldots, \Phi_{M'}) \in \mathbb{R}^{M' \times M'}$ and $\mathcal{U} = (U_1, \ldots, U_N) \in \mathbb{R}^{M' \times N}$ be given as in theorem 2 and let $L = \min(M', N)$. For any set of mutually $\mathcal{M}$-orthonormal vectors $\tilde{\Phi}_1, \ldots, \tilde{\Phi}_L \in \mathbb{R}^{M'}$ and any $R = 1, \ldots, L$ it holds that*

$$\sum_{n=1}^{N} \left\| U_n - \sum_{r=1}^{R} \tilde{\Phi}_r \tilde{\Phi}_r^T \mathcal{M} U_n \right\|_{\mathcal{M}}^2 \geq \sum_{n=1}^{N} \left\| U_n - \sum_{r=1}^{R} \Phi_r \Phi_r^T \mathcal{M} U_n \right\|_{\mathcal{M}}^2. \tag{2.14}$$

Proof. It holds that

$$\begin{aligned}
\sum_{n=1}^{N} \|U_n\|_{\mathcal{M}}^2 &= \sum_{n=1}^{N} \left\| \sum_{r=1}^{L} \sigma_r \Phi_r V_r^n \right\|_{\mathcal{M}}^2 \\
&= \sum_{n=1}^{N} \left(\sum_{i=1}^{L} \sigma_i \Phi_i V_i^n \right)^T \mathcal{M} \left(\sum_{j=1}^{L} \sigma_j \Phi_j V_j^n \right) \\
&= \sum_{n=1}^{N} \sum_{i,j=1}^{L} \sigma_i V_i^n \Phi_i^T \mathcal{M} \Phi_j V_j^n \sigma_j \\
&= \sum_{n=1}^{N} \sum_{i=1}^{L} \sigma_i V_i^n V_i^n \sigma_i \\
&= \sum_{i=1}^{L} \sigma_i^2.
\end{aligned}$$

The right-hand side of the statement of the theorem can be reformulated as

$$\begin{aligned}
&\sum_{n=1}^{N}\left\|U_n-\sum_{r=1}^{R}\Phi_r\Phi_r^T\mathcal{M}U_n\right\|_{\mathcal{M}}^2=\sum_{n=1}^{N}\left\|U_n-\sum_{r=1}^{R}\Phi_r\Phi_r^T\mathcal{M}\Phi\Sigma(V^n)^T\right\|_{\mathcal{M}}^2\\
&=\sum_{n=1}^{N}\left\|\sum_{r=1}^{L}\sigma_r\Phi_rV_r^n-\sum_{r=1}^{R}\sigma_r\Phi_rV_r^n\right\|_{\mathcal{M}}^2=\sum_{n=1}^{N}\left\|\sum_{r=R+1}^{L}\sigma_r\Phi_rV_r^T\right\|_{\mathcal{M}}^2\\
&=\sum_{n=1}^{N}\left(\sum_{i=R+1}^{L}\sigma_i\Phi_iV_i^n\right)^T\mathcal{M}\left(\sum_{j=R+1}^{L}\sigma_j\Phi_jV_j^n\right)\\
&=\sum_{n=1}^{N}\sum_{i,j=R+1}^{L}\sigma_iV_i^n\Phi_i^T\mathcal{M}\Phi_jV_j^n\sigma_j=\sum_{n=1}^{N}\sum_{i=R+1}^{L}\sigma_iV_i^nV_i^n\sigma_i=\sum_{i=R+1}^{L}\sigma_i^2\\
&=\sum_{n=1}^{N}\|U_n\|_{\mathcal{M}}^2-\sum_{i=1}^{R}\sigma_i^2.
\end{aligned}$$

The left-hand side of the statement of the theorem can be reformulated as

$$\begin{aligned}
&\sum_{n=1}^{N}\|U_n-\sum_{r=1}^{R}\tilde{\Phi}_r\tilde{\Phi}_r^T\mathcal{M}U_n\|^2\\
&=\sum_{n=1}^{N}(U_n-\sum_{r=1}^{R}\tilde{\Phi}_r\tilde{\Phi}_r^T\mathcal{M}U_n)^T\mathcal{M}(U_n-\sum_{r=1}^{R}\tilde{\Phi}_r\tilde{\Phi}_r^T\mathcal{M}U_n)\\
&=\sum_{n=1}^{N}\Big(U_n^T\mathcal{M}U_n-2\sum_{r=1}^{R}U_n^T\mathcal{M}\tilde{\Phi}_r\tilde{\Phi}_r^T\mathcal{M}U_n+\sum_{i,j=1}^{R}U_n^T\mathcal{M}\tilde{\Phi}_i^T\tilde{\Phi}_i\mathcal{M}\tilde{\Phi}_j\tilde{\Phi}_j^T\mathcal{M}U_n\Big)\\
&=\sum_{n=1}^{N}\|U_n\|_{\mathcal{M}}^2-\sum_{r=1}^{R}\sum_{n=1}^{N}U_n^T\mathcal{M}\tilde{\Phi}_r\tilde{\Phi}_r^T\mathcal{M}U_n.
\end{aligned}$$

After substituting the results in the statement, it remains to be shown that for any $R=1,\ldots,L$

$$\sum_{r=1}^{R}\sum_{n=1}^{N}U_n^T\mathcal{M}\tilde{\Phi}_r\tilde{\Phi}_r^T\mathcal{M}U_n\leq\sum_{i=1}^{R}\sigma_i^2. \tag{2.15}$$

Focusing on the individual terms in the outer sum of the left-hand side of

(2.15), it holds for any $r = 1, \dots, R$ that

$$
\begin{aligned}
&\sum_{n=1}^{N} U_n^T \mathcal{M} \tilde{\Phi}_r \tilde{\Phi}_r^T \mathcal{M} U_n \\
&\quad = \sum_{n=1}^{N} \tilde{\Phi}_r^T \mathcal{M} U_n U_n^T \mathcal{M} \tilde{\Phi}_r \\
&\quad = \tilde{\Phi}_r^T \mathcal{M} U U^T \mathcal{M} \tilde{\Phi}_r \\
&\quad = \tilde{\Phi}_r^T \mathcal{M} \Phi \Sigma V^T V \Sigma^T \Phi^T \mathcal{M} \tilde{\Phi}_r \\
&\quad = \tilde{\Phi}_r^T \mathcal{M} \Phi \Sigma \Sigma^T \Phi^T \mathcal{M} \tilde{\Phi}_r \\
&\quad = \sum_{i=1}^{L} \tilde{\Phi}_r^T \mathcal{M} \Phi_i \sigma_i^2 \Phi_i^T \mathcal{M} \tilde{\Phi}_r \\
&\quad = \sum_{i=1}^{L} \sigma_i^2 (\tilde{\Phi}_r^T \mathcal{M} \Phi_i)^2 \\
&\quad = \sigma_R^2 + \Big(\sum_{i=1}^{R} \sigma_i^2 (\tilde{\Phi}_r^T \mathcal{M} \Phi_i)^2 - \sigma_R^2 \sum_{i=1}^{R} (\tilde{\Phi}_r^T \mathcal{M} \Phi_i)^2 \Big) \\
&\qquad - \Big(\sigma_R^2 \sum_{i=R+1}^{L} (\tilde{\Phi}_r^T \mathcal{M} \Phi_i)^2 - \sum_{i=R+1}^{L} \sigma_i^2 (\tilde{\Phi}_r^T \mathcal{M} \Phi_i)^2 \Big) \\
&\qquad - \sigma_R^2 \Big(1 - \sum_{i=1}^{L} (\tilde{\Phi}_r^T \mathcal{M} \Phi_i)^2 \Big).
\end{aligned} \tag{2.16}
$$

In the following, the last two lines of (2.16) are examined. Due to the ordering of the singular values, see theorem 1, for any $R = 1, \dots, L$ and $r = 1, \dots, L$ it is true that

$$
\sigma_R^2 (\tilde{\Phi}_r^T \mathcal{M} \Phi_i)^2 \geq \sigma_i^2 (\tilde{\Phi}_r^T \mathcal{M} \Phi_i)^2, \qquad i = R+1, \dots, L
$$

From theorem 2 it follows for any $r = 1, \dots, L$ that

$$
\sum_{i=1}^{L} (\tilde{\Phi}_r^T \mathcal{M} \Phi_i)^2 \leq \sum_{i=1}^{M'} (\tilde{\Phi}_r^T \mathcal{M} \Phi_i)^2 = \tilde{\Phi}_r^T \mathcal{M} \Phi \Phi^T \mathcal{M} \tilde{\Phi}_r = \tilde{\Phi}_r^T \mathcal{M} \tilde{\Phi}_r = 1.
$$

Applying these results to (2.16) gives rise to the inequality

$$\sum_{n=1}^{N} U_n^T \mathcal{M}\tilde{\Phi}_r \tilde{\Phi}_r^T \mathcal{M} U_n \leq \sigma_R^2 + \Big(\sum_{i=1}^{R} \sigma_i^2 (\tilde{\Phi}_r^T \mathcal{M}\Phi_i)^2 - \sigma_R^2 \sum_{i=1}^{R} (\tilde{\Phi}_r^T \mathcal{M}\Phi_i)^2 \Big)$$
$$= \sigma_R^2 + \sum_{i=1}^{R} (\sigma_i^2 - \sigma_R^2)(\tilde{\Phi}_r^T \mathcal{M}\Phi_i)^2$$

for each $r = 1, \ldots, R$, and

$$\sum_{r=1}^{R}\sum_{n=1}^{N} U_n^T \mathcal{M}\tilde{\Phi}_r \tilde{\Phi}_r^T \mathcal{M} U_n \leq R\sigma_R^2 + \sum_{i,r=1}^{R} (\sigma_i^2 - \sigma_R^2)(\tilde{\Phi}_r^T \mathcal{M}\Phi_i)^2$$
$$= \sum_{i=1}^{R} \Big(\sigma_R^2 + (\sigma_i^2 - \sigma_R^2) \sum_{r=1}^{R} (\tilde{\Phi}_r^T \mathcal{M}\Phi_i)^2 \Big).$$

If $L < M'$, a Gram-Schmidt procedure is able to provide additional vectors $\tilde{\Phi}_{L+1}, \ldots, \tilde{\Phi}_{M'}$, so that $\tilde{\Phi} = (\tilde{\Phi}_1, \ldots, \tilde{\Phi}_{M'})$ fulfills

$$\mathcal{I} = \tilde{\Phi}^T \mathcal{M} \tilde{\Phi} = (\mathcal{C}\tilde{\Phi})^T (\mathcal{C}\tilde{\Phi}) = (\mathcal{C}\tilde{\Phi})(\mathcal{C}\tilde{\Phi})^T = \mathcal{C}\tilde{\Phi}\tilde{\Phi}^T \mathcal{C}^T.$$

Here it was used that a square matrix with mutually orthonormal columns has mutually orthonormal rows. Multiplying with the inverses of the Cholesky factors leads to

$$\tilde{\Phi}\tilde{\Phi}^T = \mathcal{M}^{-1}. \tag{2.17}$$

Therefore,

$$\sum_{r=1}^{R} (\tilde{\Phi}_r^T \mathcal{M}\Phi_i)^2 \leq \sum_{r=1}^{M'} (\tilde{\Phi}_r^T \mathcal{M}\Phi_i)^2 = \Phi_i^T \mathcal{M}\tilde{\Phi}\tilde{\Phi}^T \mathcal{M}\Phi_i = \Phi_i^T \mathcal{M}\Phi_i = 1.$$

Using $0 \leq \sigma_i^2 - \sigma_R^2$ for all $i \leq R$, one obtains

$$\sum_{r=1}^{R}\sum_{n=1}^{N} U_n^T \mathcal{M}\tilde{\Phi}_r \tilde{\Phi}_r^T \mathcal{M} U_n \leq \sum_{i=1}^{R} \Big(\sigma_R^2 + (\sigma_i^2 - \sigma_R^2) \sum_{r=1}^{R} (\tilde{\Phi}_r^T \mathcal{M}\Phi_i)^2 \Big)$$
$$\leq \sum_{i=1}^{R} \Big(\sigma_R^2 + (\sigma_i^2 - \sigma_R^2) \Big)$$
$$= \sum_{i=1}^{R} \sigma_i^2,$$

so (2.15) is valid. □

Theorem 2 provides solutions to the POD minimization problem 3, and so the following definitions are justified:

Definition 1. *For $L = \min(M', N)$, the decomposition*

$$\mathcal{U} = \sum_{r=1}^{L} \sigma_r \Phi_r V_r^T = \sum_{r=1}^{L} \Phi_r \Phi_r^T \mathcal{M}\mathcal{U}$$

according to theorem 2 with $\mathcal{M}$ equal to the finite element mass matrix is called a proper orthogonal decomposition.

Definition 2. *For any $R \leq \min(M', N)$, the decomposition*

$$\mathcal{U} \approx \sum_{r=1}^{R} \sigma_r \Phi_r V_r^T = \sum_{r=1}^{R} \Phi_r \Phi_r^T \mathcal{M}\mathcal{U}$$

according to theorem 2 with $\mathcal{M}$ equal to the finite element mass matrix is called a truncated proper orthogonal decomposition.

The first POD basis vectors can be represented as a linear combination of snapshots:

Corollary 1. *For $\mathcal{U}$, Φ, Σ and $\mathcal{V}$ as given in theorem 2, it holds that*

$$\Phi_r = \mathcal{U} V_r / \sigma_r, \qquad r = 1, \ldots, \operatorname{rank}(\mathcal{U}).$$

Proof. Theorem 2 states $\mathcal{U} = \Phi \Sigma \mathcal{V}^T$ and $\mathcal{V}^T \mathcal{V} = \mathcal{I}$. Consequently, $\mathcal{U}\mathcal{V} = \Phi\Sigma$. From the definition of Σ follows $\mathcal{U} V_r = \Phi_r \sigma_r$ for $r = 1, \ldots, \operatorname{rank}(\mathcal{U})$. As the considered singular values are non-zero, division is allowed, which proves the claim. □

Corollary 2. *For $\mathcal{U}$, Φ, Σ and $\mathcal{V}$ as given in theorem 2, it holds that*

$$\Phi_r = \mathcal{U}\mathcal{U}^T \mathcal{M} \Phi_r / \sigma_r^2, \qquad r = 1, \ldots, \operatorname{rank}(\mathcal{U}).$$

Proof. The substituting of $\mathcal{U}^T = \mathcal{V}\Sigma\Phi^T$ and subsequent simplification leads to the statement of corollary 1. □

The corollaries 1 and 2 are useful for the formulation of the POD-Galerkin models, as they relate properties of the snapshots with properties of the basis functions. In particular, if some row of $\mathcal{U}$ is zero, then the same rows of the first $\operatorname{rank}(\mathcal{U})$ POD basis vector are zero, which is useful for the implementation of Dirichlet conditions. Moreover, if $\mathcal{A}\mathcal{U} = 0$ for a matrix $\mathcal{A}$

of corresponding size, then $\mathcal{A}\Phi_r = 0$ for $r = 1, \ldots, \text{rank}(\mathcal{U})$. This is helpful for the implementation of the incompressible flow models, where the general matrix $\mathcal{A}$ is replaced by a discretized divergence operator.

Because theorem 2 implies an algorithm for the computation of a POD, a few remarks on the actual computations are given: For the computation of the POD basis functions from finite element snapshots via a singular value decomposition, not necessarily a Cholesky decomposition of the mass matrix needs to be performed. For instance Kunisch and Volkwein (1999) use the positive square root of the mass matrix. A Cholesky factorization with reordering is quite efficient in the computations shown in later chapters, so that the computation time and storage requirements for the matrix factorization are no issue. In the literature it is often argued that, depending on the dimensions of the snapshot matrix, it may be cheaper to obtain a singular value decomposition via the related eigenvalue decomposition of the correlation matrix $\mathcal{U}^T\mathcal{M}\mathcal{U}$ (Kunisch and Volkwein, 1999). However, in the presented computations the singular value decompositions took only a small fraction of the total computational time, which justifies the approach of a direct computation of the singular value decomposition.

2.3 Implementation of Dirichlet conditions

The next chapters feature reduced-order models for PDE problems where the values of the solution are prescribed at some parts of the boundary. As a preparation, this section introduces three different implementations of inhomogeneous Dirichlet boundary conditions: the control function method, the modified basis function method and the penalty method.

The first and most common way to incorporate inhomogeneous Dirichlet conditions in POD reduced-order models is to subtract a known lifting function from the snapshots, so that the modified snapshots fulfill homogeneous Dirichlet conditions. The resulting POD basis functions are linear combinations of the modified snapshots and, therefore, individually fulfill homogeneous Dirichlet conditions, and so does any linear combination of POD basis functions. Adding the lifting function to a linear combination of POD basis functions results in a function which fulfills the original inhomogeneous boundary conditions. In this way the inhomogeneous Dirichlet conditions are automatically included in the POD approximation. The method is computationally equivalent to subtracting the lifting function from the unknown in the original PDE problem, solving for the modified unknown and adding the lifting function to the solution. Implementing inhomogeneous Dirichlet conditions with a known lifting function is termed 'control function method'

in Graham et al. (1999) for PDE problems whose Dirichlet conditions can be parametrized with a single time-dependent coefficient. Gunzburger et al. (2007) generalize this method for problems with multiple parameters at distinct boundary sections. In subsection 2.3.1 a generalization of the procedure of Gunzburger et al. is introduced and typical choices of lifting functions are listed.

Gunzburger et al. (2007) introduce another implementation of inhomogeneous Dirichlet conditions, which is called 'modified basis function method' in the present thesis. The method relies on a POD of the original snapshots. The resulting basis functions are modified so that some of them fulfill inhomogeneous Dirichlet conditions, while the others fulfill homogeneous Dirichlet conditions. A linear combination of the former is used as a lifting function, while the latter are used exactly like the POD basis functions in the control function method. The modified basis functions are orthonormal and span the same space as the unmodified basis functions. A derivation of the method is presented in subsection 2.3.2.

The penalty method leaves the snapshots and basis functions unmodified, but enforces inhomogeneous Dirichlet conditions with a penalty term in the model equations. While being relatively common in the area of finite elements, the method has not been applied to reduced-order modeling very often. An overview of a few recent publications is given in subsection 2.3.3.

The control function method and the modified basis function method can be viewed in a common framework. They both provide solutions to the following problem:

Problem 4. *Assuming*

- $\psi_1, \ldots, \psi_{M'}$ *form a continuous Lagrangian finite element basis,*
- $u_1, \ldots, u_N \in \operatorname{span}(\psi_1, \ldots, \psi_{M'})$ *are given,*
- $u_1 - u^h_{D,1}, \ldots, u_N - u^h_{D,N} \in \operatorname{span}(\psi_1, \ldots, \psi_M)$*,*
- $u^h_{D,1}, \ldots, u^h_{D,N} \in \operatorname{span}(g_1, \ldots, g_K)$ *for* $K \leq \min(M', N)$*,*
- $g_1, \ldots, g_K \in \operatorname{span}(\psi_1, \ldots, \psi_{M'})$*,*

find functions

$$\phi_1, \ldots \phi_R \in \operatorname{span}(\psi_1, \ldots, \psi_M), \tag{2.18}$$

$$\phi_{R+1}, \ldots, \phi_{R'} \in \operatorname{span}(\psi_1, \ldots, \psi_M, \psi_{M+1}, \ldots, \psi_{M'}), \tag{2.19}$$

for $R' = R + K \leq \min(M', N)$*, so that there exist*

$$u_1^R, \dots, u_N^R \in \operatorname{span}(\phi_1, \dots, \phi_{R'}) \tag{2.20}$$

which satisfy

$$u_1 - u_1^R, \dots, u_N - u_N^R \in \operatorname{span}(\psi_1, \dots, \psi_M). \tag{2.21}$$

The two methods provide functions which, apart from solving problem 4, are solutions to a minimization problem. In particular, the control function method (see subsection 2.3.1) yields mutually $L^2(\Omega)$-orthonormal functions $\phi_1, \dots \phi_R$ which minimize

$$\sum_{n=1}^{N} \left\| u_n - \sum_{r=R+1}^{R'} \phi_r b_n^r - \sum_{r=1}^{R} \phi_r \Big(\phi_r, u_n - \sum_{r=R+1}^{R'} \phi_r b_n^r \Big) \right\|_{L^2(\Omega)}^2, \tag{2.22}$$

where $\phi_{R+1}, \dots, \phi_{R'}$ and $b_n^{R+1}, \dots, b_n^{R'}$ for $n = 1, \dots, N$ are fixed. On the contrary, the modified basis function method (subsection 2.3.2) finds mutually L^2-orthonormal functions $\phi_1, \dots \phi_{R'}$ which minimize

$$\sum_{n=1}^{N} \left\| u_n - \sum_{r=1}^{R'} \phi_r \left(\phi_r, u_n \right) \right\|_{L^2(\Omega)}^2. \tag{2.23}$$

2.3.1 Control function method

The standard method for inhomogeneous Dirichlet conditions in reduced-order models is often introduced using continuous snapshots and lifting functions. Sometimes the inhomogeneous Dirichlet conditions are already eliminated by modifying the original PDE problem. Here, on the contrary, the control function method is described on the linear algebraic level. While this may seem overly complicated, it paves the way for the subsequent modified basis function method, which can be derived more naturally from the point of view of linear algebra.

A snapshot matrix $\mathcal{U} = (U_1, \dots, U_N) \in \mathbb{R}^{M' \times N}$ is introduced like in (2.6). The columns of $\mathcal{U}$ contain the finite element coefficients vector of the snapshots $u_1, \dots, u_N \in L_h^2$. In a similar way, a matrix $\bar{\Phi} = (\bar{\Phi}_1, \dots, \bar{\Phi}_K) \in \mathbb{R}^{M' \times K}$ is defined, whose columns represent the finite element coefficient vectors of the functions $\bar{\phi}_1, \dots, \bar{\phi}_K$. The practical choice of these functions is discussed at

the end of this section. The matrices are partitioned as

$$\mathcal{U} = \begin{pmatrix} \mathcal{U}^{\mathrm{F}} \\ \mathcal{U}^{\mathrm{D}} \end{pmatrix}, \qquad \mathcal{U}^{\mathrm{F}} \in \mathbb{R}^{M \times N}, \qquad \mathcal{U}^{\mathrm{D}} \in \mathbb{R}^{\bar{M} \times N}$$
$$\bar{\Phi} = \begin{pmatrix} \bar{\Phi}^{\mathrm{F}} \\ \bar{\Phi}^{\mathrm{D}} \end{pmatrix}, \qquad \bar{\Phi}^{\mathrm{F}} \in \mathbb{R}^{M \times K}, \qquad \bar{\Phi}^{\mathrm{D}} \in \mathbb{R}^{\bar{M} \times K}$$

for $M' = M + \bar{M}$. The rows of the upper parts of the matrices, $\mathcal{U}^{\mathrm{F}}$ and $\bar{\Phi}^{\mathrm{F}}$, correspond to the mesh nodes where the unknowns of the finite element simulation are situated. The rows of the lower parts of the matrices, $\mathcal{U}^{\mathrm{D}}$ and $\bar{\Phi}^{\mathrm{D}}$, correspond to the Dirichlet mesh nodes. The superscripts F and D stand for 'free' and 'Dirichlet', respectively. It is assumed that the columns of $\mathcal{U}^{\mathrm{D}}$ can be formed by a linear combination of the rows of $\bar{\Phi}^{\mathrm{D}}$. In particular, $\mathcal{U}^{\mathrm{D}} = \bar{\Phi}^{\mathrm{D}} \bar{\mathcal{B}}^T$ for some known coefficient matrix $\bar{\mathcal{B}} \in \mathbb{R}^{N \times K}$. A modified snapshot matrix is defined as $\tilde{\mathcal{U}} = \mathcal{U} - \bar{\Phi} \bar{\mathcal{B}}^T$, so that

$$\tilde{\mathcal{U}} = \begin{pmatrix} \tilde{\mathcal{U}}^{\mathrm{F}} \\ \tilde{\mathcal{U}}^{\mathrm{D}} \end{pmatrix} = \begin{pmatrix} \mathcal{U}^{\mathrm{F}} - \bar{\Phi}^{\mathrm{F}} \bar{\mathcal{B}}^T \\ \mathcal{U}^{\mathrm{D}} - \bar{\Phi}^{\mathrm{D}} \bar{\mathcal{B}}^T \end{pmatrix} = \begin{pmatrix} \mathcal{U}^{\mathrm{F}} - \bar{\Phi}^{\mathrm{F}} \bar{\mathcal{B}}^T \\ 0 \end{pmatrix}. \tag{2.24}$$

A truncated POD of the modified snapshot matrix $\tilde{\mathcal{U}}$ is computed via theorem 2, with $\mathcal{M}$ equal to the finite element mass matrix. The number of retained basis vectors is $R \leq \operatorname{rank}(\tilde{\mathcal{U}})$, so that

$$\tilde{\mathcal{U}} \approx \hat{\mathcal{U}} = \Phi \Sigma \mathcal{V}^T, \tag{2.25}$$

where $\Phi \in \mathbb{R}^{M' \times R}$, $\Sigma \in \mathbb{R}^{R \times R}$ and $\mathcal{V} \in \mathbb{R}^{N \times R}$. Equality holds in (2.25) if $R = \operatorname{rank}(\tilde{\mathcal{U}})$. By definition 1 and theorem 2 it holds that $\Phi^T \mathcal{M} \Phi = \mathcal{I}$ and $\mathcal{V}^T \mathcal{V} = \mathcal{I}$, where $\mathcal{I}$ denotes an $R \times R$ identity matrix. However, no statements can be made about $\Phi \Phi^T$ and $\mathcal{V} \mathcal{V}^T$.

Corollary 1 implies that the columns of Φ are linear combinations of the columns of $\tilde{\mathcal{U}}$, i.e. $\Phi = \tilde{\mathcal{U}} \mathcal{V} \Sigma^{-1}$. Let

$$\Phi = \begin{pmatrix} \Phi^{\mathrm{F}} \\ \Phi^{\mathrm{D}} \end{pmatrix}, \qquad \Phi^{\mathrm{F}} \in \mathbb{R}^{M \times R}, \qquad \Phi^{\mathrm{D}} \in \mathbb{R}^{\bar{M} \times R},$$

so that $\Phi^{\mathrm{D}} = \tilde{\mathcal{U}}^{\mathrm{D}} \mathcal{V} \Sigma^{-1}$. From $\tilde{\mathcal{U}}^{\mathrm{D}} = 0$, see (2.24), follows that $\Phi^{\mathrm{D}} = 0$. In view of the finite element representation (2.7) it is clear that $\phi_1, \ldots, \phi_R \in \operatorname{span}(\psi_1, \ldots, \psi_M)$, so (2.18) is fulfilled. Moreover, the columns of $\bar{\Phi}$ can be used as coefficients in the finite element representation of $\phi_{R+1}, \ldots, \phi_{R'}$, which satisfy (2.19).

Now define $\mathcal{U}_R = \Phi\mathcal{B}^T + \bar{\Phi}\bar{\mathcal{B}}^T$, where Φ, $\bar{\Phi}$ and $\bar{\mathcal{B}}$ have been introduced above and $\mathcal{B} \in \mathbb{R}^{N\times R}$ is arbitrary. From the results above it can be deduced that

$$\mathcal{U}_R = \begin{pmatrix} \Phi^{\mathrm{F}}\mathcal{B}^T \\ 0 \end{pmatrix} + \begin{pmatrix} \bar{\Phi}^{\mathrm{F}}\bar{\mathcal{B}}^T \\ \mathcal{U}^{\mathrm{D}} \end{pmatrix}$$

and

$$\mathcal{U} - \mathcal{U}_R = \begin{pmatrix} \mathcal{U}^{\mathrm{F}} \\ \mathcal{U}^{\mathrm{D}} \end{pmatrix} - \begin{pmatrix} \Phi^{\mathrm{F}}\mathcal{B}^T + \bar{\Phi}^{\mathrm{F}}\bar{\mathcal{B}}^T \\ \mathcal{U}^{\mathrm{D}} \end{pmatrix} = \begin{pmatrix} \mathcal{U}^{\mathrm{F}} - \Phi^{\mathrm{F}}\mathcal{B}^T - \bar{\Phi}^{\mathrm{F}}\bar{\mathcal{B}}^T \\ 0 \end{pmatrix}.$$

The columns of $\mathcal{U}_R \in \mathbb{R}^{M'\times N}$ are used as finite element coefficients in combination with the finite element basis functions $\psi_1, \ldots, \psi_{M'}$ to construct reduced-order approximations $u_1^R, \ldots, u_N^R$ of the snapshots $u_1, \ldots, u_N$, see (2.20). Because the finite element coefficients of $\mathcal{U} - \mathcal{U}_R$ corresponding to $\psi_{M+1}, \ldots, \psi_{M'}$ are zero, it follows that $u_1 - u_1^R, \ldots, u_N - u_N^R \in \operatorname{span}(\psi_1, \ldots, \psi_M)$, so (2.21) is satisfied. Finally, substituting the definition of modified snapshot matrix $\tilde{\mathcal{U}}$ in the POD minimization problem 3 reveals that Φ minimizes

$$\sum_{n=1}^{N} \left\| U_n - \bar{\Phi}(\bar{B}^n)^T - \Phi\Phi^T\mathcal{M}(U_n - \bar{\Phi}(\bar{B}^n)^T) \right\|_{\mathcal{M}}^2$$

for prescribed $\bar{\Phi} \in \mathbb{R}^{M'\times K}$ and $\bar{B}^1, \ldots, \bar{B}^N \in \mathbb{R}^K$ under the condition that the columns of Φ are $\mathcal{M}$-orthonormal. This corresponds to (2.22) written in terms of the finite element coefficients vectors.

While it has been assumed that $\mathcal{U}^{\mathrm{D}} = \bar{\Phi}^{\mathrm{D}}\bar{\mathcal{B}}^T$, no restriction was imposed on $\bar{\Phi}^{\mathrm{F}}$. In the following, some popular choices are listed:

- The trivial choice is $\bar{\Phi}^{\mathrm{F}} = 0$. This reflects the standard method in the finite element context (Gresho and Sani, 2000).

- For some problems it is possible to manually construct suitable columns of $\bar{\Phi}^{\mathrm{F}}$, e.g. by using polynomial extensions of the boundary data into the interior domain. This leads to a smoother continuous extension of the boundary data, which might be favorable when gradients of the reduced-order solution must be computed at the boundary.

- In the special case of time-independent Dirichlet boundary conditions it is sufficient to choose $\bar{\Phi}$ as one column vector and $\bar{\mathcal{B}}$ as a vector of ones in order to satisfy $\mathcal{U}^{\mathrm{D}} = \bar{\Phi}^{\mathrm{D}}\bar{\mathcal{B}}^T$. A standard approach is to set $\bar{\Phi}$ equal to a snapshot average. For instance, the arithmetic mean of a set of snapshots with time-constant Dirichlet conditions fulfills the Dirichlet conditions, too.

- Suitable columns of $\bar{\Phi}$ can also be computed by solving auxiliary problems. Examples are the stationary version of the original problem (Burkardt et al., 2006a) or the adjoint equations (Eftang and Rønquist, 2010).

In some applications the reduced-order solution is required to fulfill additional conditions. Most prominently in the context of incompressible flows, the reduced-order solution is often required to be discretely divergence-free by construction. To achieve this, $\bar{\Phi}$ must be chosen so that $\phi_{R+1}, \ldots, \phi_{R'}$ are discretely divergence-free. In this case using a snapshot average or the solution of a stationary problem is advantageous, because $\bar{\Phi}^{\mathrm{F}} = 0$ does not lead to discretely divergence-free functions in general and constructing divergence-free functions manually can be challenging.

2.3.2 Modified basis function method

Gunzburger et al. (2007) describe how inhomogeneous boundary conditions in POD-Galerkin reduced-order models can be implemented by modifying the POD basis functions computed from a set of snapshots, without the need to subtract any predefined lifting function. Their method is limited, however, to problems where Dirichlet conditions at multiple non-overlapping boundary sections are characterized by functions which are products of time-dependent and space dependent functions. The method relies on a manual choice of one point on each boundary section, at which an equation has to be solved to enforce the fulfillment of the boundary condition. In the following, the method is generalized to problems with possibly *overlapping* boundary sections, at which inhomogeneous Dirichlet boundary conditions are given by *linear combinations* of products of time-dependent and space-dependent functions. The presented approach does *not* require a manual choice of certain points to enforce the boundary conditions.

The theory is provided in terms of the finite element coefficient vectors. At any time in the derivation it is possible to switch to the finite element fields by employing the finite element basis functions. A set of snapshot vectors is assumed to be given in terms of the columns of a snapshot matrix $\mathcal{U} \in \mathbb{R}^{M' \times N}$. A POD of $\mathcal{U}$ is performed via theorem 2, with $\mathcal{M}$ equal to the finite element mass matrix. A truncated POD basis is stored as columns in a matrix $\Phi_{\mathcal{U}} \in \mathbb{R}^{M' \times R'}$ for some fixed $R' \leq \min(M', N)$. Now modified POD basis vectors are sought, which span the column space of $\Phi_{\mathcal{U}}$. The modified basis functions are required to solve problem 4 and at the same time minimize (2.23).

The modified POD basis vectors are computed with an additional singular

value decomposition applied to a submatrix of $\Phi_{\mathcal{U}}$. The following theorem provides the foundation.

Theorem 4. *Let a matrix $\Phi_{\mathcal{U}} \in \mathbb{R}^{M' \times R'}$ be given and let $\Phi_{\mathcal{U}}^F$ and $\Phi_{\mathcal{U}}^D$ be submatrices of $\Phi_{\mathcal{U}}$, so that*

$$\Phi_{\mathcal{U}} = \begin{pmatrix} \Phi_{\mathcal{U}}^F \\ \Phi_{\mathcal{U}}^D \end{pmatrix}, \qquad \Phi_{\mathcal{U}}^F \in \mathbb{R}^{M \times R'}, \qquad \Phi_{\mathcal{U}}^D \in \mathbb{R}^{\bar{M} \times R'}.$$

Then a singular value decomposition

$$\Phi_{\mathcal{U}}^D = \hat{\Phi}^D \hat{\Sigma} \hat{\mathcal{V}}^T, \qquad \hat{\Phi}^D \in \mathbb{R}^{\bar{M} \times \bar{M}}, \qquad \hat{\Sigma} \in \mathbb{R}^{\bar{M} \times R'}, \qquad \hat{\mathcal{V}} \in \mathbb{R}^{R' \times R'} \tag{2.26}$$

gives rise to

$$\Phi_{\mathcal{U}} \hat{\mathcal{V}} = \begin{pmatrix} \bar{\Phi}^F & \Phi^F \\ \bar{\Phi}^D & \Phi^D \end{pmatrix}, \qquad \bar{\Phi} = \begin{pmatrix} \bar{\Phi}^F \\ \bar{\Phi}^D \end{pmatrix}, \qquad \Phi = \begin{pmatrix} \Phi^F \\ \Phi^D \end{pmatrix},$$

where Φ^D is a zero matrix with $\bar{M}$ rows and $R' - \operatorname{rank}(\Phi_{\mathcal{U}}^D)$ columns.

Proof. The singular value decomposition can be performed for any matrix $\Phi_{\mathcal{U}}^{\mathrm{D}} \in \mathbb{R}^{\bar{M} \times R'}$. It must be shown that the product $\Phi_{\mathcal{U}} \hat{\mathcal{V}}$ has the claimed block structure. It follows from the statements of the theorem and the orthogonality of $\hat{\mathcal{V}}$ that

$$\begin{pmatrix} \bar{\Phi}^{\mathrm{D}} & \Phi^{\mathrm{D}} \end{pmatrix} = \Phi_{\mathcal{U}}^{\mathrm{D}} \hat{\mathcal{V}} = \hat{\Phi}^{\mathrm{D}} \hat{\Sigma} \hat{\mathcal{V}}^T \hat{\mathcal{V}} = \hat{\Phi}^{\mathrm{D}} \hat{\Sigma}.$$

The right $R' - \operatorname{rank}(\Phi_{\mathcal{U}}^{\mathrm{D}})$ columns of the matrix $\hat{\Sigma}$ are zero vectors due to the properties of the singular value decomposition. As a consequence, the right $R' - \operatorname{rank}(\Phi_{\mathcal{U}}^{\mathrm{D}})$ columns of $\begin{pmatrix} \bar{\Phi}^{\mathrm{D}} & \Phi^{\mathrm{D}} \end{pmatrix}$ are zero vectors, which means Φ^{D} is a zero matrix. □

Theorem 4 almost ensures that the columns of Φ and $\bar{\Phi}$ qualify as finite element coefficient vectors of functions fulfilling (2.18) and (2.19), respectively. It is still left to show, however, that at least R functions belong to $\operatorname{span}(\psi_1, \ldots, \psi_M)$. But first a few helpful auxiliary results are presented.

Corollary 3. *If $\Phi_{\mathcal{U}}^T \mathcal{M} \Phi_{\mathcal{U}} = \mathcal{I}$, where $\mathcal{M}$ is a symmetric and positive definite matrix and $\mathcal{I}$ is an identity matrix, then $\begin{pmatrix} \bar{\Phi} & \Phi \end{pmatrix}^T \mathcal{M} \begin{pmatrix} \bar{\Phi} & \Phi \end{pmatrix} = \mathcal{I}$.*

Proof. The statements of theorem 4 and the orthogonality of the matrix $\hat{\mathcal{V}}$ result in

$$\begin{pmatrix} \bar{\Phi} & \Phi \end{pmatrix}^T \mathcal{M} \begin{pmatrix} \bar{\Phi} & \Phi \end{pmatrix} = \hat{\mathcal{V}}^T \Phi_{\mathcal{U}}^T \mathcal{M} \Phi_{\mathcal{U}} \hat{\mathcal{V}} = \hat{\mathcal{V}}^T \hat{\mathcal{V}} = \mathcal{I}.$$

□

Corollary 4. *It holds that* $\Phi_{\mathcal{U}} = \begin{pmatrix}\bar{\Phi} & \Phi\end{pmatrix} \hat{\mathcal{V}}^T$.

Proof. Theorem 4 and the orthogonality of $\hat{\mathcal{V}}$ lead to

$$\Phi_{\mathcal{U}} = \Phi_{\mathcal{U}} \hat{\mathcal{V}} \hat{\mathcal{V}}^T = \begin{pmatrix}\bar{\Phi} & \Phi\end{pmatrix} \hat{\mathcal{V}}^T.$$

□

Corollary 5. *The matrix* $\bar{\Phi}^D$ *has full rank.*

Proof. It has been shown in the proof of theorem 4 that the matrix $\bar{\Phi}^{\mathrm{D}}$ consists of the left $\mathrm{rank}(\Phi_{\mathcal{U}}^{\mathrm{D}})$ columns of $\hat{\Phi}^{\mathrm{D}}\hat{\Sigma}$. The singular value decomposition ensures that $\hat{\Phi}^{\mathrm{D}}$ is orthogonal and $\hat{\Sigma}$ is diagonal, with the first $\mathrm{rank}(\Phi_{\mathcal{U}}^{\mathrm{D}})$ diagonal entries larger than zero. The left $\mathrm{rank}(\Phi_{\mathcal{U}}^{\mathrm{D}})$ columns of $\hat{\Phi}^{\mathrm{D}}\hat{\Sigma}$ are therefore orthogonal to each other and not equal to zero vectors. Thus, $\bar{\Phi}^{\mathrm{D}}$ has full rank. □

Now theorem 4 is applied to the POD of a set of finite element snapshot vectors given by the columns of a matrix $\mathcal{U} \in \mathbb{R}^{M' \times N}$. Let $\Phi_{\mathcal{U}} \in \mathbb{R}^{M' \times R'}$ denote a matrix containing the basis vectors of a truncated POD of the snapshot matrix $\mathcal{U}$ and let $\begin{pmatrix}\bar{\Phi} & \Phi\end{pmatrix} \in \mathbb{R}^{M' \times R'}$ denote a matrix of corresponding modified POD basis vectors. The modified basis vectors can be represented by linear combinations of snapshots if the number of basis vectors is not too large:

Corollary 6. *Let* $\Phi_{\mathcal{U}}\Sigma_{\mathcal{U}}\mathcal{V}_{\mathcal{U}}^T$ *be a truncated POD of a matrix* $\mathcal{U} \in \mathbb{R}^{M' \times N}$ *with* $\Phi_{\mathcal{U}} \in \mathbb{R}^{M' \times R'}$, $\Sigma_{\mathcal{U}} \in \mathbb{R}^{R' \times R'}$ *and* $\mathcal{V}_{\mathcal{U}} \in \mathbb{R}^{N \times R'}$, *and let* $\begin{pmatrix}\bar{\Phi} & \Phi\end{pmatrix}$ *be the matrix of modified basis vectors according to theorem 4. Then*

$$1 \leq R' \leq \mathrm{rank}(\mathcal{U}) \quad \Rightarrow \quad \begin{pmatrix}\bar{\Phi} & \Phi\end{pmatrix} = \mathcal{U}\mathcal{V}_{\mathcal{U}}\Sigma_{\mathcal{U}}^{-1}\hat{\mathcal{V}}.$$

Proof. Corollary 1 implies $1 \leq R' \leq \mathrm{rank}(\mathcal{U}) \Rightarrow \Phi_{\mathcal{U}} = \mathcal{U}\mathcal{V}_{\mathcal{U}}\Sigma_{\mathcal{U}}^{-1}$, which can be substituted into the statement $\begin{pmatrix}\bar{\Phi} & \Phi\end{pmatrix} = \Phi_{\mathcal{U}}\hat{\mathcal{V}}$ of theorem 4. □

Assume that the upper M rows of $\bar{\Phi}$, Φ and $\mathcal{U}$ correspond to the non-Dirichlet mesh nodes, while the lower $\bar{M}$ rows correspond to the Dirichlet mesh nodes. Theorem 4 states that $R' - \mathrm{rank}(\Phi_{\mathcal{U}}^{\mathrm{D}})$ modified POD basis functions are zero at the Dirichlet boundary nodes. Corollary 5 states that $\mathrm{rank}(\Phi_{\mathcal{U}}^{\mathrm{D}})$ modified POD basis functions are non-zero at the Dirichlet boundary nodes. The rank of $\Phi_{\mathcal{U}}^{\mathrm{D}}$ can be deduced from the entries of the diagonal matrix $\hat{\Sigma}$. It is related to the rank of the Dirichlet part of the snapshot matrix:

Theorem 5. *Let a truncated POD of a matrix* $\mathcal{U} \in \mathbb{R}^{M' \times N}$ *be given by* $\Phi_\mathcal{U} \Sigma_\mathcal{U} \mathcal{V}_\mathcal{U}^T$ *with* $\Phi_\mathcal{U} \in \mathbb{R}^{M' \times R'}$, $\Sigma_\mathcal{U} \in \mathbb{R}^{R' \times R'}$ *and* $\mathcal{V}_\mathcal{U} \in \mathbb{R}^{N \times R'}$ *and let*

$$\mathcal{U} = \begin{pmatrix} \mathcal{U}^F \\ \mathcal{U}^D \end{pmatrix}, \qquad \mathcal{U}^F \in \mathbb{R}^{M \times N}, \qquad \mathcal{U}^D \in \mathbb{R}^{\bar{M} \times N}$$

as well as

$$\Phi_\mathcal{U} = \begin{pmatrix} \Phi_\mathcal{U}^F \\ \Phi_\mathcal{U}^D \end{pmatrix}, \qquad \Phi_\mathcal{U}^F \in \mathbb{R}^{M \times R'}, \qquad \Phi_\mathcal{U}^D \in \mathbb{R}^{\bar{M} \times R'}.$$

Then the following holds:

$$\operatorname{rank}(\mathcal{U}^D) \leq R' \leq \operatorname{rank}(\mathcal{U}) \quad \Rightarrow \quad \operatorname{rank}(\Phi_\mathcal{U}^D) = \operatorname{rank}(\mathcal{U}^D).$$

Proof. From corollary 1 and the structure of the matrices it follows that

$$\begin{aligned} R' \leq \operatorname{rank}(\mathcal{U}) \quad &\Rightarrow \quad \Phi_\mathcal{U} = \mathcal{U} \mathcal{V}_\mathcal{U} \Sigma_\mathcal{U}^{-1} \\ &\Rightarrow \quad \begin{pmatrix} \Phi_\mathcal{U}^\mathrm{F} \\ \Phi_\mathcal{U}^\mathrm{D} \end{pmatrix} = \begin{pmatrix} \mathcal{U}^\mathrm{F} \\ \mathcal{U}^\mathrm{D} \end{pmatrix} \mathcal{V}_\mathcal{U} \Sigma_\mathcal{U}^{-1} \\ &\Rightarrow \quad \Phi_\mathcal{U}^\mathrm{D} = \mathcal{U}^\mathrm{D} \mathcal{V}_\mathcal{U} \Sigma_\mathcal{U}^{-1}. \end{aligned}$$

In general, the rank of a product of matrices is the minimum of the ranks of the factors:

$$\operatorname{rank}(\Phi_\mathcal{U}^\mathrm{D}) = \min(\operatorname{rank}(\mathcal{U}^\mathrm{D}), \operatorname{rank}(\mathcal{V}_\mathcal{U} \Sigma_\mathcal{U}^{-1})).$$

By the properties of the singular value decomposition,

$$R' \leq \operatorname{rank}(\mathcal{U}) \quad \Rightarrow \quad \operatorname{rank}(\mathcal{V}_\mathcal{U} \Sigma_\mathcal{U}^{-1}) = R',$$

which leads to

$$R' \leq \operatorname{rank}(\mathcal{U}) \quad \Rightarrow \quad \operatorname{rank}(\Phi_\mathcal{U}^\mathrm{D}) = \min(\operatorname{rank}(\mathcal{U}^\mathrm{D}), R').$$

Finally, after introducing the additional assumption of $\operatorname{rank}(\mathcal{U}^\mathrm{D}) \leq R'$,

$$\operatorname{rank}(\mathcal{U}^\mathrm{D}) \leq R' \leq \operatorname{rank}(\mathcal{U}) \quad \Rightarrow \quad \operatorname{rank}(\Phi_\mathcal{U}^\mathrm{D}) = \operatorname{rank}(\mathcal{U}^\mathrm{D}).$$

□

The snapshot matrix $\mathcal{U}$ is assumed to result from problem 4, where Dirichlet conditions were implemented with a linear combination of given functions $g_1^h, \ldots, g_K^h$. This means that for $\mathcal{U}^\mathrm{D} \in \mathbb{R}^{\bar{M} \times N}$ containing the finite element

coefficients of the Dirichlet nodes it holds that $\operatorname{rank}(\mathcal{U}^{\mathrm{D}}) \leq K$. In the following, theorem 5 and 4 are applied to this matrix and the resulting POD. Recalling $R = R' - K$, theorem 5 leads to

$$R \leq \operatorname{rank}(\mathcal{U}) - K \quad \Rightarrow \quad \operatorname{rank}(\Phi^{\mathrm{D}}_{\mathcal{U}}) \leq K. \tag{2.27}$$

In theorem 4, Φ^{D} is a zero matrix with $\bar{M}$ rows and $R' - \operatorname{rank}(\Phi^{\mathrm{D}}_{\mathcal{U}})$ columns. Therefore, (2.27) implies that Φ^{D} has at least R columns if $R \leq \operatorname{rank}(\mathcal{U}) - K$. Assume at first that Φ^{D} is a matrix which consists of *exactly* R columns, so $\Phi \in \mathbb{R}^{M' \times R}$ and $\bar{\Phi} \in \mathbb{R}^{M' \times \bar{R}}$. Using the elements of the matrices as finite element coefficient vectors in combination with the respective finite element basis functions, it can be observed that the columns of Φ and $\bar{\Phi}$ result in finite element function which fulfill (2.18) and (2.19), respectively. Assume now that Φ^{D} consists of *more than* R columns. The first R columns of Φ can be used to form finite element functions which fulfill (2.18). The remaining columns of Φ together with all columns of $\bar{\Phi}$ can be used to form finite element functions which fulfill (2.19).

The Dirichlet part of the snapshot matrix can be shown to be equal to a linear combination of modified POD basis vectors:

Theorem 6. *If* $\operatorname{rank}(\mathcal{U}^{D}) \leq R' \leq \operatorname{rank}(\mathcal{U})$*, then each column of* $\mathcal{U}^{D}$ *is a linear combination of the columns of* $\bar{\Phi}^{D}$*.*

Proof. A truncated singular value decomposition $\mathcal{U}^{\mathrm{D}} = \tilde{\Phi}^{\mathrm{D}} \tilde{\Sigma} \tilde{\mathcal{V}}^T$ is performed while keeping all non-zero singular values. From corollary 6 follows

$$\begin{pmatrix} \bar{\Phi}^{\mathrm{D}} & \Phi^{\mathrm{D}} \end{pmatrix} = \mathcal{U}^{\mathrm{D}} \mathcal{V}_{\mathcal{U}} \Sigma_{\mathcal{U}}^{-1} \hat{\mathcal{V}} = \tilde{\Phi}^{\mathrm{D}} \mathcal{D}, \qquad \mathcal{D} = \tilde{\Sigma} \tilde{\mathcal{V}}^T \mathcal{V}_{\mathcal{U}} \Sigma_{\mathcal{U}}^{-1} \hat{\mathcal{V}}.$$

By inspecting each of the factors of $\mathcal{D}$, it can be shown that for $\operatorname{rank}(\mathcal{U}^{\mathrm{D}}) \leq R' \leq \operatorname{rank}(\mathcal{U})$ the rank of the matrix $\mathcal{D} \in \mathbb{R}^{\operatorname{rank}(\mathcal{U}^{\mathrm{D}}) \times R'}$ equals the rank of $\mathcal{U}^{\mathrm{D}}$, so that $\mathcal{D}\mathcal{D}^T$ is of full rank and, thus, invertible. By verifying

$$\begin{pmatrix} \bar{\Phi}^{\mathrm{D}} & \Phi^{\mathrm{D}} \end{pmatrix} \mathcal{D}^T (\mathcal{D}\mathcal{D}^T)^{-1} \tilde{\Sigma} \tilde{\mathcal{V}}^T = \tilde{\Phi}^{\mathrm{D}} \mathcal{D}\mathcal{D}^T (\mathcal{D}\mathcal{D}^T)^{-1} \tilde{\Sigma} \tilde{\mathcal{V}}^T = \tilde{\Phi}^{\mathrm{D}} \tilde{\Sigma} \tilde{\mathcal{V}}^T = \mathcal{U}^{\mathrm{D}},$$

it can be shown that each column of $\mathcal{U}^{\mathrm{D}}$ is a linear combination of the columns of $\begin{pmatrix} \bar{\Phi}^{\mathrm{D}} & \Phi^{\mathrm{D}} \end{pmatrix}$. Because Φ^{D} is a zero matrix, each column of $\mathcal{U}^{\mathrm{D}}$ is a linear combination of the columns of $\bar{\Phi}^{\mathrm{D}}$. □

Let $\bar{\mathcal{B}} \in \mathbb{R}^{N \times K}$ be any matrix for which $\mathcal{U}^{\mathrm{D}} = \bar{\Phi}^{\mathrm{D}} \bar{\mathcal{B}}^T$. The proof of theorem 6 provides an example for $\bar{\mathcal{B}}$. It can be shown like in the previous subsection that (2.20) and (2.21) are satisfied for $u_1^R, \ldots, u_N^R$ resulting from $\mathcal{U}_R = \Phi \mathcal{B}^T + \bar{\Phi} \bar{\mathcal{B}}^T$ with arbitrary $\mathcal{B} \in \mathbb{R}^{N \times R}$.

Regarding optimality with respect to (2.23), note that $\Phi_{\mathcal{U}}$ solves the discrete POD minimization problem 3 and the respective functional can be reformulated using theorem 4:

$$\begin{aligned}
&\sum_{n=1}^{N} \left\| U_n - \Phi_{\mathcal{U}} \Phi_{\mathcal{U}}^T \mathcal{M} U_n \right\|_{\mathcal{M}}^2 \\
&\quad = \sum_{n=1}^{N} \left\| U_n - \Phi_{\mathcal{U}} \hat{\mathcal{V}} \hat{\mathcal{V}}^T \Phi_{\mathcal{U}}^T \mathcal{M} U_n \right\|_{\mathcal{M}}^2 \\
&\quad = \sum_{n=1}^{N} \left\| U_n - \begin{pmatrix} \bar{\Phi} & \Phi \end{pmatrix} \begin{pmatrix} \bar{\Phi} & \Phi \end{pmatrix}^T \mathcal{M} U_n \right\|_{\mathcal{M}}^2,
\end{aligned}$$

which is a discrete version of (2.23).

2.3.3 Penalty method

A third approach to incorporate inhomogeneous Dirichlet boundary conditions in POD-Galerkin reduced-order models is a penalty method, which enforces the boundary conditions weakly using a penalty term. The method has been studied in the context of the finite element method already e.g. by Babuška (1973) and Barrett and Elliott (1986) for elliptic boundary value problems and by Gunzburger and Hou (1992) for steady Stokes and Navier-Stokes problems. Although the majority of papers on reduced-order modeling employ a strong enforcement of inhomogeneous Dirichlet conditions, the penalty method has been used as well. In the following, some references to studies are given, which focus on reduced-order models with penalty enforcement of the boundary conditions. For mathematical details, the reader is referred to these references.

Graham et al. (1999) introduce a penalty method to incorporate inhomogeneous time-dependent Dirichlet boundary conditions for the incompressible flow around a cylinder at a Reynolds number of 100. They compare the method with the more conventional control function method, concluding that 'no clear winner emerges'. The authors mention the introduction of an additional parameter as a disadvantage of the penalty method and suggest that good parameter values should be determined by numerical experiments.

Also Sirisup and Karniadakis (2005) employ a penalty method to incorporate time-dependent Dirichlet boundary conditions in a POD-Galerkin flow model. They consider two examples, the simulation of the flow around a cylinder based on a POD of a set of numerically generated snapshots as well as a wave-structure interaction problem with snapshots taken from particle velocimetry measurements. In the paper the accuracy and stability of the

method depending on the value of the penalty parameters are studied numerically.

Kalashnikova and Barone (2010) consider POD-Galerkin models of the linearized compressible Euler equations. The boundary conditions are implemented with a penalty-like formulation. The penalty parameters are derived from the governing equations. The authors present a convergence analysis and a computation of a priori error bounds for the solution of the reduced-order model.

In a subsequent work, Kalashnikova and Barone (2012) present a penalty POD-Galerkin model for non-linear problems, using the Allan-Cahn equation and a convection-diffusion-reaction system as examples, both in an unsteady one-dimensional setting. The core of the method is the determination of the penalty parameters by a linear stability analysis around a stable state.

Janon et al. (2013) present a reduced-order model with penalty enforcement of the boundary conditions for the viscous Burgers' equation, where the viscosity field as well as the initial and boundary data are parametrized. To construct the underlying reduced spaces, the authors consider the POD method, the Greedy method and the POD-Greedy method. The value of the penalty parameter is chosen manually. Different values are compared to illustrate the effect of the parameter on the error bounds and the actual error.

Chapter 3

Heat conduction in a square

A nonlinear parabolic PDE problem is studied as a first example of reduced-order modeling. The setting was proposed by Gunzburger et al. (2007) to compare different implementations of inhomogeneous Dirichlet conditions. The problem is described by the transient heat equation augmented with a quadratic term. The unit square is taken as a computational domain, with unsteady inhomogeneous Dirichlet boundary conditions at two opposing sides and homogeneous Dirichlet conditions at the other sides. Section 3.1 presents the governing equations in their strong form and in a weak form which is suitable for finite element and reduced-order modeling. A finite element model of the problem is derived and validated in section 3.2. Based on snapshots of the finite element model, in section 3.3 a reduced-order model of the heat conduction problem is derived and tested.

The reduced-order models presented in this thesis all rely on finite element snapshots. Still, direct access to the discretized operators is not strictly necessary to create a reduced-order model whose solution converges toward the underlying snapshots. In fact it is possible to derive a reduced-order model directly from a weak form of the problem. This approach is sometimes called *non-intrusive*. On the other hand, one can derive a reduced-order model from the discretized equations. This can be viewed as an *intrusive* approach, as it requires access to the discretized operators. To compare these approaches, different derivations of the reduced-order heat conduction model are presented: one starting from the weak form, one starting from the spatially semi-discretized equations and another one starting from the fully discretized equations. An outline is given in Figure 3.1. All presented derivations lead to the same reduced-order model under the conditions that that all spatial

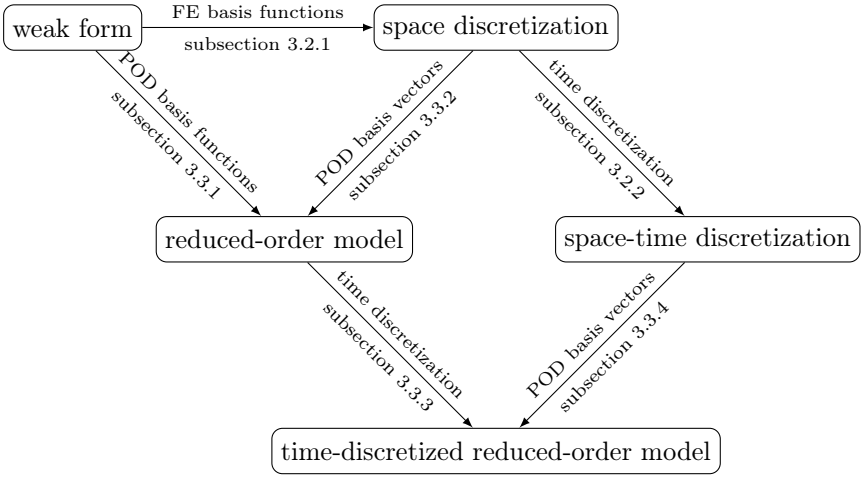

Figure 3.1: Outline of the derivation of the discretized reduced-order model.

integrals and derivatives are implemented with finite elements on the snapshot mesh and all time derivatives are realized with the time integrator of the snapshot simulation. The reduced-order model is validated with respect to the underlying snapshots and with respect to finite element solutions obtained with a higher resolution in space and time.

This goal of this chapter is to prepare the ground for reduced-order models of more complex problems. The finite element and reduced-order terminology is introduced by means of a scalar PDE problem. Peculiarities like inhomogeneous Dirichlet boundary conditions and quadratic non-linearities are already present in this otherwise simple problem. The notation is kept general enough to carry over to the problems in the following chapters in a straight-forward manner.

For an error analysis of POD-Galerkin reduced-order models for parabolic problems, the reader is referred to Kunisch and Volkwein (2001), where forward and backward Euler as well as Crank-Nicolson discretizations are studied. Reduced-order modeling based on a greedy method is considered in Grepl and Patera (2005) for parametrized linear parabolic problems. Haasdonk and Ohlberger (2008) introduce a combination of a greedy method with a POD for a general parametrized linear evolution equation. Generalizations of such 'reduced basis methods' to non-linear problems exist (Grepl et al., 2007). All

these references provide examples of parabolic problems for which efficient reduced-order models can be developed.

3.1 Governing equations

Let $\Omega \subset \mathbb{R}^2$ be a bounded spatial domain given by a non-empty open connected set with boundary Γ and closure $\bar{\Omega}$, and let the spatial coordinates be denoted by $\vec{x} = (x, y)^T$. The boundary Γ is split in a Dirichlet part Γ_{D}^q and a Neumann part Γ_{N}^q so that $\Gamma = \Gamma_{\mathrm{D}}^q \cup \Gamma_{\mathrm{N}}^q$ with $\Gamma_{\mathrm{D}}^q \cap \Gamma_{\mathrm{N}}^q = \emptyset$. The surface unit normal vector at the boundary is given by $\vec{n} = (n_x, n_y)^T$. Furthermore, let $[0, T]$ be a time interval with a final time $T > 0$ and let t denote the time variable. A two-dimensional unsteady heat conduction problem has to be solved for the temperature $q : [0, T] \times \bar{\Omega} \to \mathbb{R}$. As initial data a function $q_0 : \bar{\Omega} \to \mathbb{R}$ is given, whose values depend continuously on $\vec{x}$. An extension of the temperature at the Dirichlet boundary into the domain is given by a function $q_{\mathrm{D}} : [0, T] \times \bar{\Omega} \to \mathbb{R}$, whose values depend continuously on space and time. The initial and boundary data is assumed to be consistent in that $q_0 = q_{\mathrm{D}}$ on $\{0\} \times \Gamma_{\mathrm{D}}^q$.

The problem is governed by the semilinear parabolic partial differential equation

$$\partial_t q(t, \vec{x}) - \Delta q(t, \vec{x}) + q(t, \vec{x})^2 = 0, \qquad (t, \vec{x}) \in (0, T] \times \Omega, \tag{3.1}$$

with initial and boundary conditions

$$q(t, \vec{x}) - q_0(\vec{x}) = 0, \qquad (t, \vec{x}) \in \{0\} \times \Omega, \tag{3.2}$$

$$q(t, \vec{x}) - q_{\mathrm{D}}(t, \vec{x}) = 0, \qquad (t, \vec{x}) \in (0, T] \times \Gamma_{\mathrm{D}}^q, \tag{3.3}$$

$$\nabla q(t, \vec{x}) \cdot \vec{n} = 0, \qquad (t, \vec{x}) \in (0, T] \times \Gamma_{\mathrm{N}}^q. \tag{3.4}$$

As an additional restriction, the continuous extension of the Dirichlet data is assumed to be a linear combination with K^q terms,

$$q_{\mathrm{D}}(t, \vec{x}) = \sum_{k=1}^{K^q} g_k^q(\vec{x}) \eta_q^k(t), \qquad (t, \vec{x}) \in [0, T] \times \bar{\Omega}, \tag{3.5}$$

for $g_1^q, \ldots, g_{K^q}^q : \bar{\Omega} \to \mathbb{R}$ being continuous in space and $\eta_q^1, \ldots, \eta_q^{K^q} : [0, T] \to \mathbb{R}$ being continuous in time.

The finite element method is based on a weak form of the governing equations, which is derived in the following. As a prerequisite, the subspace of $H^1(\Omega)$ functions which vanish on the Dirichlet boundary is given by

$$H_{q,0}^1(\Omega) = \{a \in H^1(\Omega) : a|_{\Gamma_{\mathrm{D}}^q} = 0\}, \tag{3.6}$$

where $q|_{\Gamma_{\mathrm{D}}^q}$ is interpreted as the trace of q (see Knabner and Angermann, 2003, chapter 3). The partial differential equation (3.1) and initial condition (3.2) are multiplied by some test function $\psi^q \in H^1_{q,0}(\Omega)$ and integrated over Ω, so that

$$\begin{aligned}(\psi^q, \partial_t q) - (\psi^q, \Delta q) + (\psi^q, q^2) &= 0, & t &\in (0, T],\\ (\psi^q, q) - (\psi^q, q_0) &= 0, & t &\in \{0\}.\end{aligned}$$

Applying the divergence theorem to the second term of the first equation gives

$$(\psi^q, \partial_t q) + (\nabla\psi^q, \nabla q) + (\psi^q, q^2) = \int_\Gamma \psi^q \nabla q \cdot \vec{n}\, \mathrm{d}\Gamma,$$

where the boundary integral is interpreted in the sense of traces, too. The boundary integral can be eliminated by using (3.4) at the Neumann boundary Γ_{N}^q and using (3.6) at the Dirichlet boundary Γ_{D}^q. The unknown q is assumed to belong to the Sobolev space $H^1(\Omega)$. The inhomogeneous Dirichlet condition is implemented by imposing that the difference between the unknown and the continuous extension of the Dirichlet boundary belongs to $H^1_{q,0}(\Omega)$. This standard approach leads to the following weak form: Find $q - q_{\mathrm{D}} \in H^1_{q,0}(\Omega)$ such that

$$(\psi^q, \partial_t q) + (\nabla\psi^q, \nabla q) + (\psi^q, q^2) = 0, \quad \forall\psi^q \in H^1_{q,0}(\Omega), \quad t \in (0, T], \tag{3.7}$$

$$(\psi^q, q) - (\psi^q, q_0) = 0, \quad \forall\psi^q \in H^1_{q,0}(\Omega), \quad t \in \{0\}. \tag{3.8}$$

3.2 Finite element modeling

The derivation of the finite element model is carried out using the vertical method of lines (see Knabner and Angermann, 2003), which consists of first discretizing the equations in space (subsection 3.2.1), and then applying a time-stepping scheme (subsection 3.2.2). The simulation is validated with a space and time step refinement study (subsection 3.2.3).

3.2.1 Finite element discretization in space

The problem is discretized in space with quadratic Lagrangian finite elements on a triangular grid. To this end, the domain Ω is subdivided into triangles so that no vertex is positioned at the edge of another triangle. The union of vertices and edge midpoints is called mesh nodes and denoted by $\vec{x}_1^q, \ldots, \vec{x}_{M'^q}^q$.

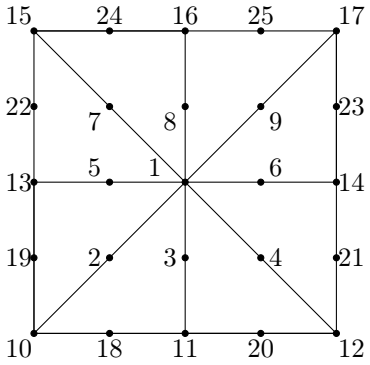

Figure 3.2: Sketch of a possible numbering of the mesh nodes for a quadratic domain divided into 8 triangles. The first indices (1–9) refer to non-Dirichlet nodes, while the other (10–25) refer to Dirichlet nodes.

The finite element solution is assumed to be a quadratic polynomial within each triangle and continuous over the edges, so that it is uniquely defined by the values at the mesh nodes. The solution can be represented as a linear combination of quadratic Lagrangian finite element basis functions $\psi_1^q, \ldots, \psi_{M'^q}^q$. These functions are piecewise quadratic within each triangle, continuous over the edges and fulfill $\psi_i^q(\vec{x}_j^q) = \delta_{ij}$ for all $i, j = 1, \ldots, M'^q$. The mesh nodes are assumed to be sorted, so that the first nodes $\vec{x}_1^q, \ldots, \vec{x}_{M^q}^q$ are situated in the interior domain or on the Neumann boundary and the other nodes $\vec{x}_{M^q+1}^q, \ldots, \vec{x}_{M'^q}^q$ are situated on the Dirichlet boundary. A possible numbering is sketched in Figure 3.2, where $M^q = 9$, $M'^q = 25$ and where the whole boundary is a Dirichlet boundary. More details about quadratic finite elements in two dimensions can be found in Babuška et al. (2011) and other textbooks.

The finite element method relies on a subspace of the function space used in the weak form. This subspace can be defined as the span of finite element basis functions. In particular, the finite element subspaces $H_{q,h}^1$ and $H_{q,0,h}^1$ are introduced as

$$H_{q,h}^1 = \operatorname{span}(\psi_1^q, \ldots, \psi_{M^q}^q, \psi_{M^q+1}^q, \ldots, \psi_{M'^q}^q), \tag{3.9}$$

$$H_{q,0,h}^1 = \operatorname{span}(\psi_1^q, \ldots, \psi_{M^q}^q). \tag{3.10}$$

A finite element approximation q^h of a temperature field q is then defined as

the linear combination of finite element basis functions

$$q(t,\vec{x}) \approx q^h(t,\vec{x}) = \sum_{i=1}^{M'^q} \psi_i^q(\vec{x}) Q^i(t), \qquad (t,\vec{x}) \in [0,T] \times \bar{\Omega}, \tag{3.11}$$

with the finite element coefficient vector $Q : [0,T] \to \mathbb{R}^{M'^q}$ written in components as

$$Q(t) = (Q^1(t), \ldots, Q^{M'^q}(t))^T, \qquad t \in [0,T].$$

For simplicity, the functions $g_1^q, \ldots, g_{K^q}^q$ and q_0, see (3.2) and (3.5), are also approximated by their finite element interpolations,

$$g_k^q(\vec{x}) \approx g_k^h(\vec{x}) = \sum_{i=1}^{M'^q} \psi_i^q(\vec{x}) g_k^q(\vec{x}_i^q), \qquad k = 1, \ldots, K^q, \qquad \vec{x} \in \bar{\Omega}, \tag{3.12}$$

$$q_0(\vec{x}) \approx q_0^h(\vec{x}) = \sum_{i=1}^{M'^q} \psi_i^q(\vec{x}) q_0(\vec{x}_i^q), \qquad \vec{x} \in \bar{\Omega}. \tag{3.13}$$

The discrete equivalent q_{D}^h of q_{D}, see (3.5), is given as

$$q_{\mathrm{D}}(t,\vec{x}) \approx q_{\mathrm{D}}^h(t,\vec{x}) = \sum_{k=1}^{K^q} g_k^h(\vec{x}) \eta_q^k(t), \qquad (t,\vec{x}) \in [0,T] \times \bar{\Omega}. \tag{3.14}$$

Although it is customary to choose $q_{\mathrm{D}}^h(t,\vec{x}_i^q) = 0$ for $i = 1, \ldots, M^q$ (see Gresho and Sani, 2000, section 2.2), the general case is considered here.

By employing the finite element spaces, the weak form (3.7)–(3.8) can be replaced by their discrete counterpart: Find $q^h - q_{\mathrm{D}}^h \in H_{q,0,h}^1$ such that

$$\begin{aligned} &(\psi^q, \partial_t q^h) + (\nabla\psi^q, \nabla q^h) \\ &\quad + (\psi^q, q^h q^h) = 0, \qquad \forall \psi^q \in H_{q,0,h}^1, \qquad t \in (0,T], \end{aligned} \tag{3.15}$$

$$(\psi^q, q^h) - (\psi^q, q_0^h) = 0, \qquad \forall \psi^q \in H_{q,0,h}^1, \qquad t \in \{0\}. \tag{3.16}$$

To ensure $q^h - q_{\mathrm{D}}^h \in H_{q,0,h}^1$ it is sufficient to provide for

$$Q^i(t) = \sum_{k=1}^{K^q} g_k^q(\vec{x}_i^q) \eta_q^k(t), \qquad i = M^q + 1, \ldots, M'^q, \qquad t \in [0,T]. \tag{3.17}$$

This can be shown by first substituting (3.11), (3.14) and (3.12) in the expression $q^h - q^h_{\mathrm{D}}$, reformulating, and finally inserting (3.17):

$$\begin{aligned} q^h - q^h_{\mathrm{D}} &= \sum_{i=1}^{M'^q} \psi_i^q Q^i - \sum_{k=1}^{K^q} \sum_{i=1}^{M'^q} \psi_i^q g_k^q(\vec{x}_i^q) \eta_q^k \\ &= \sum_{i=1}^{M'^q} \psi_i^q \left(Q^i - \sum_{k=1}^{K^q} g_k^q(\vec{x}_i^q) \eta_q^k \right) \\ &= \sum_{i=1}^{M^q} \psi_i^q \left(Q^i - \sum_{k=1}^{K^q} g_k^q(\vec{x}_i^q) \eta_q^k \right). \end{aligned}$$

The result employs only the first M^q finite element basis functions. From (3.10) it follows that $q^h - q^h_{\mathrm{D}}$ is an element of $H^1_{q,0,h}$ if the condition (3.17) is fulfilled.

As $H^1_{q,0,h}$ is spanned by $\psi_1^q, \ldots, \psi_{M^q}^q$, it is allowed to replace ψ^q in (3.15) and (3.16) with the finite element basis functions in turn. After expanding the finite element approximations (3.11) and (3.13), the following discrete problem is obtained: Find $Q^1, \ldots, Q^{M^q}$ satisfying

$$\begin{aligned} &\sum_{j=1}^{M'^q} \underbrace{(\psi_i^q, \psi_j^q)}_{\mathcal{M}_q^{\mathrm{F}}} \dot{Q}^j + \sum_{j=1}^{M'^q} \underbrace{(\nabla\psi_i^q, \nabla\psi_j^q)}_{\mathcal{K}_q^{\mathrm{F}}} Q^j \\ &\quad + \sum_{j=1}^{M'^q} \underbrace{\left(\psi_i^q, \sum_{k=1}^{M'^q} Q^k \psi_k^q \psi_j^q\right)}_{\mathcal{F}_q^{\mathrm{F}}} Q^j = 0, \qquad i = 1, \ldots, M^q, \quad t \in (0, T], \end{aligned} \tag{3.18}$$

$$\sum_{j=1}^{M'^q} \underbrace{(\psi_i^q, \psi_j^q)}_{\mathcal{M}_q^{\mathrm{F}}} Q^j - \sum_{j=1}^{M'^q} \underbrace{(\psi_i^q, \psi_j^q)}_{\mathcal{M}_q^{\mathrm{F}}} q_0(\vec{x}_i^q) = 0, \quad i = 1, \ldots, M^q, \quad t \in \{0\}, \tag{3.19}$$

with Dirichlet boundary values given by (3.17). The braces indicate which terms enter which matrices in the subsequent finite element assembly. The subscript Q implies that the matrices are associated with the finite element coefficient vector of the temperature field. The superscript F implies that only the columns corresponding to the 'free' nodes $\vec{x}_1^q, \ldots, \vec{x}_{M^q}^q$ are employed.

For a given set of finite element basis functions, the scalar products of

(3.18) and (3.19) can be computed. An additional vector

$$Q_0 = (q_0(\vec{x}_1^q), \ldots, q_0(\vec{x}_{M'^q}^q))^T$$

is introduced to denote the finite element coefficients of the initial data. The resulting vector form of the equations reads

$$\mathcal{M}_q^{\mathrm{F}} \dot{Q} + \mathcal{K}_q^{\mathrm{F}}\, Q + \mathcal{F}_q^{\mathrm{F}}(Q) Q = 0, \qquad t \in (0, T], \tag{3.20}$$

$$\mathcal{M}_q^{\mathrm{F}} Q - \mathcal{M}_q^{\mathrm{F}} Q_0 = 0, \qquad t \in \{0\}, \tag{3.21}$$

with Dirichlet values prescribed by (3.17) and with the definitions of $\mathcal{M}_q^{\mathrm{F}}$, $\mathcal{K}_q^{\mathrm{F}}$, $\mathcal{F}_q^{\mathrm{F}}(Q) \in \mathbb{R}^{M^q \times M'^q}$ implied by (3.18) and (3.19). It can be shown that $\mathcal{M}_q^{\mathrm{F}}$ is positive definite and, thus, invertible. Nevertheless, (3.21) is not further simplified in order to be consistent with the equations of the reduced-order models later on.

3.2.2 Finite difference discretization in time

The system of equations (3.20), (3.21) and (3.17) is discretized in time using the trapezoidal rule, also known as Crank-Nicolson scheme (Crank and Nicolson, 1947). At first, the time instances $t_1, \ldots, t_N$ of the numerical solution are defined, which satisfy $0 = t_1 < \cdots < t_N = T$. Then, the fully discrete approximate solution vectors are introduced as $Q_n = (Q_n^1, \ldots, Q_n^{M'^q})^T$ for $n = 1, \ldots, N$. They are supposed to approximate the semi-discrete solution at the time instances, i.e. $Q_n \approx Q(t_n)$ for $n = 1, \ldots, N$. Finite element representations of the space-time discretized solutions are given by

$$q(t_n, \vec{x}) \approx q_n^h(\vec{x}) = \sum_{i=1}^{M'^q} \psi_i^q(\vec{x}) Q_n^i, \qquad n = 1, \ldots, N, \qquad \vec{x} \in \bar{\Omega}. \tag{3.22}$$

The Crank-Nicolson method applied to the (3.20), (3.21) and (3.17) leads to the time stepping scheme

$$\mathcal{M}_q^{\mathrm{F}} \frac{Q_n - Q_{n-1}}{t_n - t_{n-1}} + \mathcal{K}_q^{\mathrm{F}} \frac{Q_n + Q_{n-1}}{2} + \frac{\mathcal{F}_q^{\mathrm{F}}(Q_n) Q_n + \mathcal{F}_q^{\mathrm{F}}(Q_{n-1}) Q_{n-1}}{2} = 0, \qquad n = 2, \ldots, N, \tag{3.23}$$

$$\mathcal{M}_q^{\mathrm{F}} Q_n - \mathcal{M}_q^{\mathrm{F}} Q_0 = 0, \qquad n = 1 \tag{3.24}$$

with Dirichlet boundary values

$$Q_n^i = \sum_{k=1}^{K^q} g_k^q(\vec{x}_i^q) \eta_q^k(t_n), \qquad i = M^q + 1, \ldots, M'^q, \qquad n = 1, \ldots, N. \tag{3.25}$$

Table 3.1: Parameters of the non-linear heat conduction problem.

description	symbol	expression
domain	Ω	$(0,1) \times (0,1)$
Dirichlet boundary	Γ_{D}^{q}	Γ
Neumann boundary	Γ_{N}^{q}	$\emptyset$
simulation time	T	1
initial condition	q_0	0
Dirichlet condition	q_{D}	$g_1^q(\vec{x})\eta_q^1(t)$
	$g_1^q(\vec{x})$	$4x(1-x)$
	$\eta_q^1(t)$	$\begin{cases} 2t & \text{if } 0 \leq t \leq 0.5 \\ 2(1-t) & \text{if } 0.5 \leq t \leq 1 \end{cases}$

In each step of the time-marching procedure, a non-linear system of equations must be solved for the non-Dirichlet components of the solution vector at the new time instance. Newton's method is used to split the non-linear problems into sequences of linear algebraic sub-problems, which can be solved via LU-decomposition.

3.2.3 Validation

A implementation of the finite element model of the previous sections was created with Matlab. In the following, the resulting model is validated with a refinement study for one of the scenarios given in Gunzburger et al. (2007). The respective parameters are listed in Table 3.1.

For the time discretization a uniform grid with a constant time step size of $\Delta t = 0.05 \cdot 2^{-i}$ is chosen, where i is the time refinement level and $i = 0$ corresponds to the initial time grid. The spatial domain is discretized with a uniform mesh that resembles a union jack pattern. A series of refined spatial grids is depicted in Figure 3.3. The mesh width h is defined as the shortest edge length and satisfies $h = 0.5 \cdot 2^{-j}$, where j is the initial refinement level.

A series of Matlab simulations is performed for all combinations of $i = 0, \ldots, 6$ and $j = 0, \ldots, 6$, employing the numerical scheme described above. In Figure 3.4 the solutions at $t = 0, 0.25, 0.5, 0.75, 1$ obtained with $\Delta t = 1/320$ ($i = 4$) and $h = 1/32$ ($j = 4$) are shown. The respective time grid consists

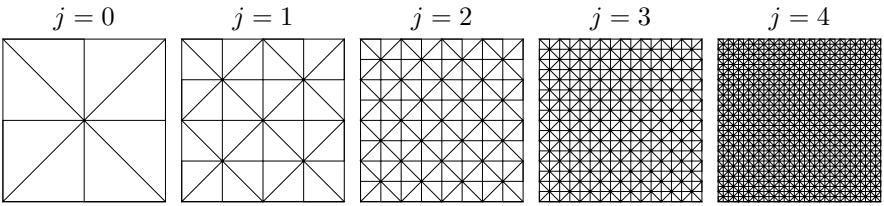

Figure 3.3: Uniformly refined spatial grids.

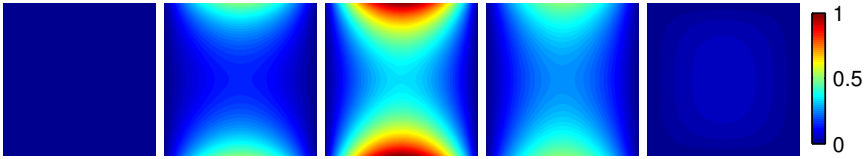

Figure 3.4: Numerical solutions at times $t = 0, 0.25, 0.5, 0.75, 1$ from left to right.

of 321 time instances including the initial time. The space grid consists of 2048 triangles and 4225 nodes. The simulation took a wall-clock time of approximately 16 s using one thread of an Intel Xeon E5-4650 CPU with Matlab R2012b.

The validation is performed by computing an estimate of the error between the unknown solution q and its numerical approximation q^h for the various simulations. In particular, $q - q^h$ is approximated by $q^\star - q^h$, where $q^\star$ is a numerical solution with a higher numerical resolution, which is used as a best guess for the unknown true solution. The space-time L^2 norm

$$\|\cdot\| = \left(\int_0^T \int_\Omega (\cdot)^2 \, \mathrm{d}\Omega \, \mathrm{d}t \right)^{\frac{1}{2}}$$

is used to quantify the error. For sufficiently smooth data the error norm is expected to decay with a certain rate depending on the time step size and the space step size. A decay with a lower rate indicates a programming bug or inadequate data.

The computation of the error norm from the discrete solution values requires a few numerical tools itself. For simplicity, standard Matlab routines are used with their default options. The time integral is computed with an

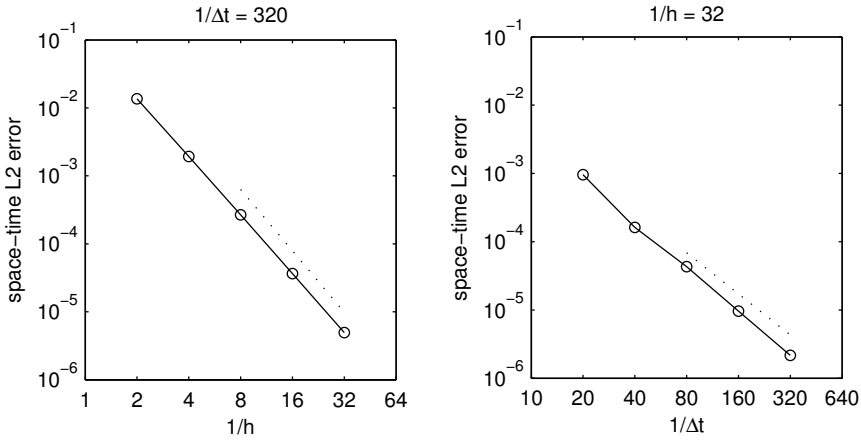

Figure 3.5: Convergence of the error norms with respect to the spatial refinement (left) and the temporal refinement (right). The dotted lines indicate the slopes of cubic (left) and quadratic (right) convergence.

adaptive Simpson quadrature (`quad`) using cubic spline interpolations of q^h and $q^\star$ (`spline`). The cubic splines satisfy not-a-knot conditions, which prescribe continuity of the third derivative at the points next to an end point. The spatial integrals are computed exactly via the finite element representations of q^h and $q^\star$.

The convergence of $\|q^\star - q^h\|$ with respect to the spatial refinement level ($j = 0, \ldots, 4$) is presented in Figure 3.5 (left) for a fixed time step refinement level of $i = 4$. A solution obtained with $i = 4$ and $j = 6$ is used as a reference finite element solution $q^\star$, in order to isolate the effect of the spatial error. An order of convergence of approximately 2.9 is achieved for the spatial refinement. This is close to the expected cubic convergence rate for quadratic finite elements.

The convergence of $\|q^\star - q^h\|$ with respect to the temporal refinement ($i = 0, \ldots, 4$) is presented in Figure 3.5 (right) for a fixed spatial mesh refinement level of $j = 4$. A solution obtained with $i = 6$ and $j = 4$ is used as a reference finite element solution $q^\star$, in order to isolate the effect of the temporal error. The order of convergence is approximately 2.2 for the time step refinement, which is close to the expected quadratic convergence rate of the Crank-Nicolson scheme.

In summary, the finite element code provides a numerical solution which

converges with the expected rates when the resolution is increased in time or in space. Figure 3.5 shows that the minimum spatial and temporal errors are quite balanced for $h = 1/32$ and $\Delta t = 1/320$. In lack of more knowledge about what magnitude of error might be acceptable, the respective simulation is used to produce snapshots for the subsequent reduced-order modeling.

3.3 Reduced-order modeling

Reduced-order models of the heat conduction problem are derived in three different ways, starting from the weak form of the problem, from the spatially semi-discretized equations and from the space-time discretized equations. In case of pure Galerkin modeling all approaches lead to the same set of equations constituting the reduced-order model.

In the first approach (subsection 3.3.1) the weak form of the equations is projected onto the space spanned by the POD basis functions. This leads to a system of ordinary differential equations. In order to assemble these reduced-order equations on a computer, integrals over the spatial domain have to be computed. If these integrations are carried out exactly, the reduced-order equations can be written in terms of the matrices of the finite element semi-discretization.

The second approach (subsection 3.3.2) uses a spatial finite element semi-discretization as a starting point. The reduced-order model is derived by projecting the semi-discretized equations on the space spanned by a set of reduced basis vectors. This leads to the same reduced-order system of ODEs as in the first approach.

A time-discretization is applied to solve the reduced-order system of ODEs that appears in the first two approaches (subsection 3.3.3). The same time-stepping scheme as in the snapshot simulation is used.

The third approach (subsection 3.3.4) obtains the discretized reduced-order equations from a projection of a given space-time discretized finite element model on the space spanned by a set of reduced basis vectors. It is shown that if the reduced-order models of the first two approaches are discretized using the same time integration method and time step sizes as in the full-order simulation, all approaches lead to the same reduced-order time-stepping scheme.

In the validation of the reduced-order model (subsection 3.3.5), the ability of the reduced-order model to approximate the underlying snapshots is studied independently from its ability to approximate a finite element reference solution obtained with a finer step size in time and space.

3.3.1 Temperature model via weak form

The output of the finite element simulation of the heat conduction problem consists of a set of solution time instances $0 = t_1 < \cdots < t_N = T$ and respective discrete solutions $q_1^h, \ldots, q_N^h \in H^1_{q,h}(\Omega)$ of the numerical solution. In the following, a POD-Galerkin reduced-order model with R^q degrees of freedom is created from these data.

As a first step, the POD basis functions and a resulting POD representation are introduced for the heat conduction setting. The goal is to obtain a reduced-order representation which fulfills the discrete Dirichlet conditions of the snapshot simulation. The POD representation is presented in a general context which is valid for the control function approach and the modified basis function approach at the same time. Each of the procedures provides two sets of functions, $\phi_1^q, \ldots, \phi_{R^q}^q \in H^1_{q,0,h}(\Omega)$ and $\phi_{R^q+1}^q, \ldots, \phi_{R'^q}^q \in H^1_{q,h}(\Omega)$, where $R'^q = R^q + K^q$. They define reduced-order spaces $H^1_{q,R} \subset H^1_{q,h}$ and $H^1_{q,0,R} \subset H^1_{q,0,h}$ via

$$H^1_{q,R} = \operatorname{span}\{\phi_1^q, \ldots, \phi_{R^q}^q, \phi_{R^q+1}^q, \ldots, \phi_{R'^q}^q\}, \tag{3.26}$$

$$H^1_{q,0,R} = \operatorname{span}\{\phi_1^q, \ldots, \phi_{R^q}^q\}. \tag{3.27}$$

A linear combination of basis functions

$$q^R(t, \vec{x}) = \sum_{i=1}^{R'^q} \phi_i^q(\vec{x}) b^i(t), \qquad t \in [0,T], \qquad \vec{x} \in \bar{\Omega} \tag{3.28}$$

with arbitrary coefficient functions $b^1, \ldots, b^{R'^q} : [0,T] \to \mathbb{R}$ is called a reduced-order representation. Let particular coefficients $\bar{b}^{R+1}, \ldots, \bar{b}^{R'^q} : [0,T] \to \mathbb{R}$ be given so that the linear combination

$$q^R_{\mathrm{D}}(t, \vec{x}) = \sum_{i=R^q+1}^{R'^q} \phi_i^q(\vec{x}) \bar{b}^i(t), \qquad t \in [0,T], \qquad \vec{x} \in \bar{\Omega} \tag{3.29}$$

results in $q^h_{\mathrm{D}} - q^R_{\mathrm{D}} \in H^1_{q,0,h}$ for any $t \in [0,T]$. If $b^r(t) = \bar{b}^r(t)$ for $r = R^q, \ldots R'^q$ and $t \in [0,T]$ in (3.28), then q^R fulfills the discrete Dirichlet conditions of the snapshot simulation. In other words, if $q^R - q^R_{\mathrm{D}} \in H^1_{q,0,R}$ for all $t \in [0,T]$, then $q^R - q^h_{\mathrm{D}} \in H^1_{q,0,h}$ for all $t \in [0,T]$. This can be deduced from

$$q^R - q^R_{\mathrm{D}} \in H^1_{q,0,R} \Rightarrow q^R - q^R_{\mathrm{D}} \in H^1_{q,0,h}$$

and $q^h_{\mathrm{D}} - q^R_{\mathrm{D}} \in H^1_{q,0,h}$, so that

$$q^R - q^h_{\mathrm{D}} = (q^R - q^R_{\mathrm{D}}) - (q^h_{\mathrm{D}} - q^R_{\mathrm{D}}) \in H^1_{q,0,h}.$$

A general finite element representation of $\phi_1^q, \ldots, \phi_{R'^q}^q \in H_{q,h}^1$ is given by

$$\phi_i^q(\vec{x}) = \sum_{m=1}^{M'^q} \psi_m^q(\vec{x}) \Phi_i^{q,m}, \qquad i = 1, \ldots, R'^q, \qquad \vec{x} \in \bar{\Omega} \tag{3.30}$$

with finite element coefficient vectors $\Phi_1^q, \ldots, \Phi_{R'^q}^q \in \mathbb{R}^{M'^q}$ defined as

$$\Phi_i^q = \begin{pmatrix} \Phi_i^{q,1} \\ \vdots \\ \Phi_i^{q,M'^q} \end{pmatrix}, \qquad i = 1, \ldots, R'^q. \tag{3.31}$$

For $\phi_1^q, \ldots, \phi_{R^q}^q \in H_{q,0,h}^1$ the finite element basis coefficients at the Dirichlet nodes are zero. Consequently,

$$\phi_i^q(\vec{x}) = \sum_{m=1}^{M^q} \psi_m^q(\vec{x}) \Phi_i^{q,m}, \qquad i = 1, \ldots, R^q \tag{3.32}$$

with finite element coefficient vectors $\Phi_1^{q,\mathrm{F}}, \ldots, \Phi_{R^q}^{q,\mathrm{F}} \in \mathbb{R}^{M^q}$ given by

$$\Phi_i^{q,\mathrm{F}} = \begin{pmatrix} \Phi_i^{q,1} \\ \vdots \\ \Phi_i^{q,M^q} \end{pmatrix}, \qquad i = 1, \ldots, R^q. \tag{3.33}$$

The derivation of the reduced-order model is based on the weak form (3.7)–(3.8). The space $H_{q,0,R}^1$, see (3.27), is chosen as a reduced-order subspace of $H_{q,0}^1$ and the finite element approximation q_0^h of the initial condition is employed, see (3.13). The Dirichlet conditions are enforced using q_{D}^R, see (3.29) and the remarks thereafter, in particular using the known coefficients $\bar{b}^{R^q+1}, \ldots, \bar{b}^{R'^q}$. The following reduced-order weak form is obtained for q^R according to (3.28): Find $q^R - q_{\mathrm{D}}^R \in H_{q,0,R}^1$ such that

$$\begin{aligned} &(\phi^q, \partial_t q^R) + (\nabla \phi^q, \nabla q^R) + (\phi^q, q^R q^R) = 0 \quad && \forall \phi^q \in H_{q,0,R}^1, \quad t \in (0, T], \\ &(\phi^q, q^R) = (\phi^q, q_0^h) && \forall \phi^q \in H_{q,0,R}^1, \quad t \in \{0\}. \end{aligned}$$

Expanding q^R and rearranging the equations leads to the following problem:

For given $b^{R^q+1}, \ldots, b^{R'^q}$, find $b^1, \ldots, b^{R^q}$ which satisfy

$$\sum_{i=1}^{R'^q} (\phi_r^q, \phi_i^q)\dot{b}^i + \sum_{i=1}^{R'^q} (\nabla\phi_r^q, \nabla\phi_i^q) b^i + \sum_{i,j=1}^{R'^q} (\phi_r^q, \phi_i^q \phi_j^q) b^i b^j = 0, \qquad r = 1, \ldots, R^q, \quad t \in (0, T], \tag{3.34}$$

$$\sum_{i=1}^{R'^q} (\phi_r^q, \phi_i^q) b^i - (\phi_r^q, q_0^h) = 0, \qquad r = 1, \ldots, R^q, \quad t \in \{0\}. \tag{3.35}$$

To solve these equations efficiently it is necessary to evaluate the inner products beforehand, as is presented term by term in the following.

For the expression (ϕ_r^q, ϕ_i^q), the equations are first presented and then explained in detail:

$$\begin{aligned}(\phi_r^q, \phi_i^q) &= \left(\sum_{m=1}^{M^q} \psi_m^q \Phi_r^{q,m}, \sum_{n=1}^{M'^q} \psi_n^q \Phi_i^{q,n} \right) \\ &= \sum_{m=1}^{M^q} \sum_{n=1}^{M'^q} \Phi_r^{q,m} (\psi_m^q, \psi_n^q) \Phi_i^{q,n} \\ &= (\Phi_r^{q,\mathrm{F}})^T \mathcal{M}_q^{\mathrm{F}} \Phi_i^q, \quad r = 1, \ldots, R^q, \quad i = 1, \ldots, R'^q.\end{aligned}$$

In the first step the finite element representations (3.30) and (3.32) are substituted in the inner product. In the second step the equations are rearranged. Finally, the sums are written as matrix-vector products, using the finite element coefficient vectors (3.31) and (3.33) of the reduced basis functions. The matrix $\mathcal{M}_q^{\mathrm{F}}$ is equal to the rectangular finite element mass matrix (see subsection 3.2.1).

The expression $(\nabla\phi_r^q, \nabla\phi_i^q)$ is reformulated as

$$\begin{aligned}(\nabla\phi_r^q, \nabla\phi_i^q) &= \left(\nabla \sum_{m=1}^{M^q} \psi_m^q \Phi_r^{q,m}, \nabla \sum_{n=1}^{M'^q} \psi_n^q \Phi_i^{q,n} \right) \\ &= \sum_{m=1}^{M^q} \sum_{n=1}^{M'^q} \Phi_r^{q,m} (\nabla\psi_m^q, \nabla\psi_n^q) \Phi_i^{q,n} \\ &= (\Phi_r^{q,\mathrm{F}})^T \mathcal{K}_q^{\mathrm{F}}\, \Phi_i^q, \quad r = 1, \ldots, R^q, \quad i = 1, \ldots, R'^q,\end{aligned}$$

where $\mathcal{K}_q^{\mathrm{F}}$ is identical to the rectangular finite element stiffness matrix.

Now $(\phi_r^q, \phi_i^q \phi_j^q)$ is written as

$$\begin{aligned}
\left(\phi_r^q, \phi_i^q \phi_j^q\right) &= \left(\sum_{k=1}^{M^q} \psi_k^q \Phi_r^{q,k}, \left(\sum_{l=1}^{M'^q} \psi_l^q \Phi_i^{q,l}\right)\left(\sum_{m=1}^{M'^q} \psi_m^q \Phi_j^{q,m}\right)\right) \\
&= \sum_{k=1}^{M^q} \sum_{m=1}^{M'^q} \Phi_r^{q,k} \left(\psi_k^q, \sum_{l=1}^{M'^q} \Phi_i^{q,l} \psi_l^q \psi_m^q\right) \Phi_j^{q,m} \\
&= (\Phi_r^{q,\mathrm{F}})^T \mathcal{F}_q^{\mathrm{F}}(\Phi_i^q) \Phi_j^q, \quad r = 1, \ldots, R^q, \quad i = 1, \ldots, R'^q,
\end{aligned}$$

where $\mathcal{F}_q^{\mathrm{F}}(\Phi_i^q)$ is identical to the finite element matrix of the nonlinear term, evaluated for Φ_i^q.

Finally for (ϕ_r^q, q_0^h), using the finite element representation (3.13) of q_0^h,

$$\begin{aligned}
(\phi_r^q, q_0^h) &= \left(\sum_{m=1}^{M^q} \psi_m^q \Phi_r^{q,m}, \sum_{n=1}^{M'^q} \psi_n^q Q_0^n\right) \\
&= \sum_{m=1}^{M^q} \sum_{n=1}^{M'^q} \Phi_r^{q,m} (\psi_m^q, \psi_n^q) Q_0^n \\
&= (\Phi_r^{q,\mathrm{F}})^T \mathcal{M}_q^{\mathrm{F}} Q_0, \quad r = 1, \ldots, R^q
\end{aligned}$$

with the finite element mass matrix $\mathcal{M}_q^{\mathrm{F}}$ and the finite element coefficient vector Q_0 of the initial condition.

After collecting the reduced-order expressions of all terms, one obtains the following problem: For given $b^{R^q+1}, \ldots, b^{R'^q}$, find $b^1, \ldots, b^{R^q}$ such that

$$\begin{aligned}
&\sum_{i=1}^{R'^q} (\Phi_r^{q,\mathrm{F}})^T \mathcal{M}_q^{\mathrm{F}} \Phi_i^q \dot{b}^i \\
&\quad + \sum_{i=1}^{R'^q} (\Phi_r^{q,\mathrm{F}})^T \mathcal{K}_q^{\mathrm{F}} \, \Phi_i^q b^i \\
&\quad + \sum_{i,j=1}^{R'^q} (\Phi_r^{q,\mathrm{F}})^T \mathcal{F}_q^{\mathrm{F}}(\Phi_i^q) \Phi_j^q b^i b^j = 0, \quad r = 1, \ldots, R^q, \quad t \in (0, T], \quad (3.36)
\end{aligned}$$

$$\sum_{i=1}^{R'^q} (\Phi_r^{q,\mathrm{F}})^T \mathcal{M}_q^{\mathrm{F}} \Phi_i^q b^i - (\Phi_r^{q,\mathrm{F}})^T \mathcal{M}_q^{\mathrm{F}} Q_0 = 0, \quad r = 1, \ldots, R^q, \quad t \in \{0\}. \quad (3.37)$$

3.3.2 Temperature model via space discretization

The reduced-order model is now created using the semi-discretization (3.20) and (3.21) with Dirichlet boundary values given by (3.17). First the discrete PDE and the initial condition is treated and then the Dirichlet conditions are considered separately.

The finite element representation (3.28) of the reduced-order approximation can be written as

$$\begin{aligned} q^h(t,\vec{x}) \approx q^R(t,\vec{x}) &= \sum_{i=1}^{R'^q} \sum_{m=1}^{M'^q} \psi_m^q(\vec{x}) \Phi_i^{q,m} b^i(t), \\ &= \sum_{m=1}^{M'^q} \psi_m^q(\vec{x}) \sum_{i=1}^{R'^q} \Phi_i^{q,m} b^i(t), \\ &= \sum_{m=1}^{M'^q} \psi_m^q(\vec{x}) Q^{R,m}(t), \qquad\qquad t \in [0,T], \qquad \vec{x} \in \bar{\Omega} \end{aligned}$$

with the finite element coefficient vector $Q^R : [0,T] \to \mathbb{R}^{M'^q}$ given by

$$Q(t) \approx Q^R(t) = \sum_{i=1}^{R'^q} \Phi_i^q b^i(t), \qquad\qquad t \in [0,T]. \tag{3.38}$$

Substituting the reduced-order approximation (3.38) in (3.20) and (3.21) gives rise to the residuals

$$\begin{aligned} \mathcal{R}^q &= \mathcal{M}_q^{\mathrm{F}} \dot{Q}^R + \mathcal{K}_q^{\mathrm{F}}\, Q^R + \mathcal{F}_q^{\mathrm{F}}(Q^R) Q^R, && t \in (0,T], \\ \mathcal{R}_0^q &= \mathcal{M}_q^{\mathrm{F}} Q^R - \mathcal{M}_q^{\mathrm{F}} Q_0, && t \in \{0\}. \end{aligned}$$

The residuals are required to be orthogonal to the finite element coefficient vectors $\Phi_1^{q,\mathrm{F}}, \ldots, \Phi_R^{q,\mathrm{F}}$ of the reduced basis functions, see (3.33). In particular, this means that

$$\begin{aligned} &(\Phi_r^{q,\mathrm{F}})^T \mathcal{M}_q^{\mathrm{F}} \dot{Q}^R \\ &\quad + (\Phi_r^{q,\mathrm{F}})^T \mathcal{K}_q^{\mathrm{F}}\, Q^R \\ &\quad + (\Phi_r^{q,\mathrm{F}})^T \mathcal{F}_q^{\mathrm{F}}(Q^R) Q^R = 0, \qquad r = 1,\ldots,R^q, \qquad t \in (0,T], \\ &(\Phi_r^{q,\mathrm{F}})^T \mathcal{M}_q^{\mathrm{F}} Q^R \\ &\quad - (\Phi_r^{q,\mathrm{F}})^T \mathcal{M}_q^{\mathrm{F}} Q_0 = 0, \qquad r = 1,\ldots,R^q, \qquad t \in \{0\}. \end{aligned}$$

After expanding Q^R via (3.38), the reduced-order equations (3.36) and (3.37) are recovered.

In the previous subsection, suitable functions $\phi^q_{R^q+1}, \ldots, \phi^q_{R'^q}$ in space and $\bar{b}^{R^q+1}, \ldots, \bar{b}^{R'^q}$ in time were assumed to be available in order to implement the condition $q^R - q^h_{\mathrm{D}} \in H^1_{q,0,h}$ for $t \in [0,T]$. The Lagrangian property implies

$$q^R(t, \vec{x}^q_i) = \sum_{m=1}^{M'^q} \psi^q_m(\vec{x}^q_i) Q^{R,m}(t) = Q^{R,i}(t), \quad i = 1, \ldots, M'^q, \quad t \in [0,T].$$

One can deduce, via (3.28), (3.14) and (3.10), that $q^R(t, \vec{x}^q_i) = q^h_{\mathrm{D}}(t, \vec{x}^q_i)$ for $i = M^q + 1, \ldots, M'^q$ and $t \in [0,T]$. Therefore, by (3.14), (3.12) and the Lagrangian property,

$$\begin{aligned} Q^{R,i}(t) = q^h_{\mathrm{D}}(t, \vec{x}^q_i) &= \sum_{k=1}^{K^q} g^h_k(\vec{x}^q_i) \eta^k_q(t) \\ &= \sum_{k=1}^{K^q} \sum_{j=1}^{M'^q} \psi^q_j(\vec{x}^q_i) g^q_k(\vec{x}^q_j) \eta^k_q(t) \\ &= \sum_{k=1}^{K^q} g^q_k(\vec{x}^q_i) \eta^k_q(t), \quad i = 1, \ldots, M'^q, \quad t \in [0,T]. \end{aligned}$$

Therefore, the discrete Dirichlet condition (3.17) is fulfilled by the reduced-order representation.

3.3.3 Finite difference discretization in time

The goal of finite element based POD-Galerkin modeling is a set of equations that can be solved with a computation time that is independent of the number of mesh nodes. To achieve this, it is necessary to evaluate the coefficients of the equations (3.36) and (3.37) already in the setup-phase, so that only low-dimensional matrices and tensors remain. The following reduced-order problem is obtained: For given $b^{R^q+1}, \ldots, b^{R'^q}$, find $b^1, \ldots, b^{R^q}$ such that

$$\begin{aligned} &\sum_{i=1}^{R'^q} \mathcal{M}^b_{ri} \dot{b}^i + \sum_{i=1}^{R'^q} \mathcal{K}^b_{ri} b^i + \sum_{i,j=1}^{R'^q} F^b_{rij} b^i b^j = 0, \quad r = 1, \ldots, R^q, \quad t \in (0,T], \\ &\sum_{i=1}^{R'^q} \mathcal{M}^b_{ri} b^i - b^r_0 = 0, \qquad\qquad\qquad\qquad\qquad\qquad r = 1, \ldots, R^q, \quad t \in \{0\}. \end{aligned}$$

with

$$\begin{aligned}
\mathcal{M}^b_{ri} &= (\Phi^{q,\mathrm{F}}_r)^T \mathcal{M}^{\mathrm{F}}_q \Phi^q_i, && r = 1,\dots,R^q, && i = 1,\dots,R'^q,\\
\mathcal{K}^b_{ri} &= (\Phi^{q,\mathrm{F}}_r)^T \mathcal{K}^{\mathrm{F}}_q \, \Phi^q_i, && r = 1,\dots,R^q, && i = 1,\dots,R'^q,\\
F^b_{rij} &= (\Phi^{q,\mathrm{F}}_r)^T \mathcal{F}^{\mathrm{F}}_q(\Phi^q_i)\Phi^q_j, && r = 1,\dots,R^q, && i,j = 1,\dots,R'^q,\\
b^r_0 &= (\Phi^{q,\mathrm{F}}_r)^T \mathcal{M}^{\mathrm{F}}_q Q_0, && r = 1,\dots,R^q.
\end{aligned}$$

A time discretization method must be applied to solve this system of ordinary differential equations numerically. The numerical scheme and the time step size of the snapshot simulation are used for this purpose. The notation $b^r_n \approx b^r(t_n)$ for all $r = 1,\dots,R^q$ and $n = 1,\dots,N$ is introduced for the unknown time-discrete reduced-order coefficients. Similarly, $b^r_n = b^r(t_n)$ for all $r = R^q+1,\dots,R'^q$ and $n = 1,\dots,N$ is used for the given coefficients, so that via (3.29) the Dirichlet conditions are exactly fulfilled at any discrete time instance $t_1,\dots,t_N$. The Crank-Nicolson method applied to the reduced-order model results in the following time stepping scheme: For given $\{b^{R^q+1}_n,\dots,b^{R'^q}_n\}_{n=1}^N$, find $\{b^1_n,\dots,b^{R^q}_n\}_{n=1}^N$, such that

$$\begin{aligned}
&\sum_{i=1}^{R'^q} \mathcal{M}^b_{ri} \frac{b^i_n - b^i_{n-1}}{t_n - t_{n-1}}\\
&\qquad + \sum_{i=1}^{R'^q} \mathcal{K}^b_{ri} \frac{b^i_n + b^i_{n-1}}{2}\\
&\qquad + \sum_{i,j=1}^{R'^q} F^b_{rij} \frac{b^i_n b^j_n + b^i_{n-1} b^j_{n-1}}{2} = 0, \quad r = 1,\dots,R^q, \quad n = 2,\dots,N,\\
&\sum_{i=1}^{R'^q} \mathcal{M}^b_{ri} b^i_n - b^r_0 = 0 \qquad\qquad r = 1,\dots,R^q, \quad n = 1.
\end{aligned}$$

Hence, in each time step a non-linear system of algebraic equations needs to be solved. For this purpose Newton's method is employed with an LU decomposition for the linear sub-steps. If R^q is chosen small enough, the computation of the reduced-order model is much faster than the computation of the underlying finite element model.

3.3.4 Temperature model via space-time discretization

A space-time discretized reduced-order model can be obtained directly from the space-time discretized finite element model given by (3.23), (3.24) and

(3.25). In these equations the reduced-order approximations

$$Q_n \approx Q_n^R = \sum_{i=1}^{R'^q} \Phi_i^q b_n^i, \qquad\qquad n = 1, \ldots, N$$

are substituted, where again $b_n^r = b^r(t_n)$ for $r = R^q + 1, \ldots, R'^q$ and $n = 1, \ldots, N$ is used for the known coefficients, so that the Dirichlet conditions are exactly fulfilled at $t = t_1, \ldots, t_N$. The substitution gives rise to the residuals

$$\begin{aligned}
\mathcal{R}_n^q &= \mathcal{M}_q^{\mathrm{F}} \frac{Q_n^R - Q_{n-1}^R}{t_n - t_{n-1}} + \mathcal{K}_q^{\mathrm{F}} \frac{Q_n^R + Q_{n-1}^R}{2} \\
&\quad + \frac{\mathcal{F}_q^{\mathrm{F}}(Q_n^R) Q_n^R + \mathcal{F}_q^{\mathrm{F}}(Q_{n-1}^R) Q_{n-1}^R}{2}, \qquad n = 2, \ldots, N, \\
\mathcal{R}_n^q &= \mathcal{M}_q^{\mathrm{F}} Q_n^R - \mathcal{M}_q^{\mathrm{F}} Q_0, \qquad\qquad n = 1.
\end{aligned}$$

For each time step the residuals are multiplied by the reduced basis functions in turn, as to require orthogonality with respect to the reduced space. This leads to the following reduced-order problem: For given coefficients $\{b_n^{R^q+1}, \ldots, b_n^{R'^q}\}_{n=1}^N$, find $\{b_n^1, \ldots, b_n^{R^q}\}_{n=1}^N$, such that, for $r = 1, \ldots, R^q$,

$$\begin{aligned}
&\sum_{i=1}^{R'^q} (\Phi_r^{q,\mathrm{F}})^T \mathcal{M}_q^{\mathrm{F}} \Phi_i^q \frac{b_n^i - b_{n-1}^i}{t_n - t_{n-1}} + \sum_{i=1}^{R'^q} (\Phi_r^{q,\mathrm{F}})^T \mathcal{K}_q^{\mathrm{F}} \Phi_i^q \frac{b_n^i + b_{n-1}^i}{2} \\
&\qquad + \sum_{i,j=1}^{R'^q} (\Phi_r^{q,\mathrm{F}})^T \mathcal{F}_q^{\mathrm{F}}(\Phi_i^q) \Phi_j^q \frac{b_n^i b_n^j + b_{n-1}^i b_{n-1}^j}{2} = 0, \qquad n = 2, \ldots, N, \\
&\sum_{i=1}^{R'^q} (\Phi_r^{q,\mathrm{F}})^T \mathcal{M}_q^{\mathrm{F}} \Phi_i^q b_n^i - (\Phi_r^{q,\mathrm{F}})^T \mathcal{M}_q^{\mathrm{F}} Q_0 = 0, \qquad n = 1.
\end{aligned}$$

After substituting the definitions of $\mathcal{M}_{ri}^b$, $\mathcal{K}_{ri}^b$, F_{ri}^b and b_0^r, see subsection 3.3.3, it turns out that the Galerkin projection of the fully discretized equations on the reduced basis vectors gives the same time stepping scheme as the time discretization applied to the reduced-order system of ODEs.

3.3.5 Validation

In subsection 3.2.3 finite element simulations were performed for different refinement levels in space and time. The POD-Galerkin reduced-order models presented in the previous sections were also implemented in Matlab. The validation of the reduced-order models builds upon the results of the finite

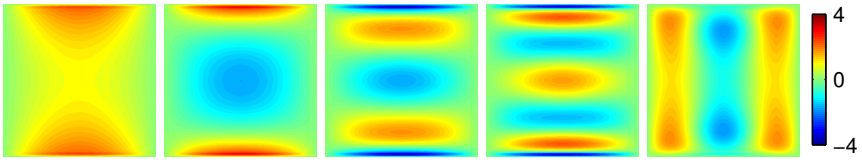

Figure 3.6: First 5 temperature POD basis functions.

element simulation with a time refinement level of $i = 4$ and a space refinement level of $j = 4$, corresponding to 321 time instances and 4225 mesh nodes. All 321 snapshots are used to generate reduced bases with varying dimensions $R^q = 0, \ldots, 50$. The Dirichlet conditions are implemented with the control function approach using a lifting function which is zero at all non-Dirichlet nodes. A few reduced basis functions are shown in Figure 3.6. Although this is hardly visible in the plot, all basis functions are zero at the boundary.

The accuracy of the reduced-order approximations is investigated using two types of errors,

- the error $q^h - q^R$ between the reduced-order solution and the underlying snapshot solution and
- the error $q^\star - q^R$ between the reduced-order solution and a reference finite element solution obtained with a higher numerical resolution than the snapshots,

which are measured in the space-time L^2 norm used already for the validation of the finite element simulation before. For comparison, the norm of the finite element error between the snapshot solution and the reference solution is provided, too. Two types of reduced-order solutions q^R are distinguished, the solution of reduced-order model and the projection of the snapshots on the reduced basis. While the error of the former measures the accuracy of the model, the error of the latter is the minimum error achievable with any linear combination of any R basis functions. The results are presented in Figure 3.7.

The graph on the left-hand side of Figure 3.7 shows the dependence of the error norm $\|q^h - q^R\|$ on the dimension R^q of the reduced basis. The red circles correspond to the reduced-order solution obtained by projection while the blue crosses correspond to the reduced-order solution obtained with the reduced-order model. Both errors are very close to each other, which means that the model provides a solution which is close to optimal. The error decreases by about 5 orders of magnitude from $R^q = 0$ to $R^q = 7$, which

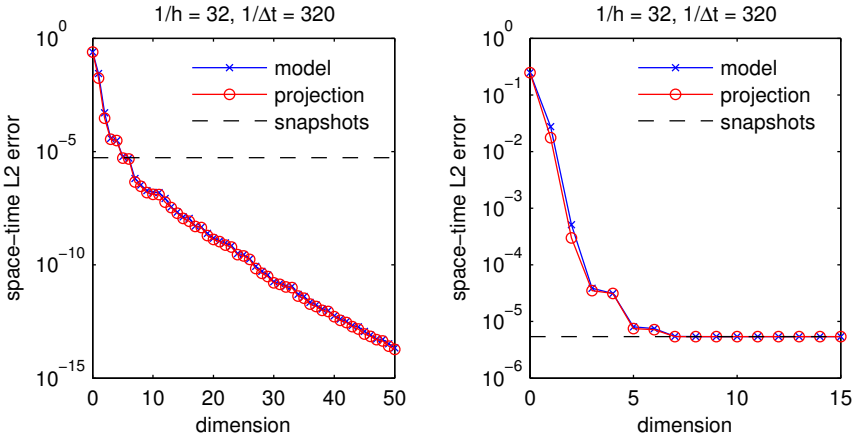

Figure 3.7: Semi-logarithmic plots of the space-time L^2 norms of the projection errors and the model errors depending on the dimension of the reduced basis. The dashed line indicates the error between the snapshots and the reference solution. Left: Error between the reduced-order solutions and the underlying snapshots. Right: Error between the reduced-order solutions and the reference solution.

suggests that the problem is very amenable to reduced-order modeling. In fact, the error for $R^q = 7$ is already well below the error of the finite element snapshots. For $R^q > 7$ an exponential decay of the error can be observed, as the graph in the semi-logarithmic plot is close to a straight line.

The graph on the right-hand side of Figure 3.7 shows the dependence of the error norm $\|q^\star - q^R\|$ on the dimension of the reduced basis, where again in one instance q^R is obtained by projection (red circles) and in another instance q^R is obtained by solving the reduced-order model (blue crosses). Again, both errors are very close to each other and a fast initial decay of the error can be seen. An increase of the dimension beyond $R^q = 7$, however, does not reduce the error below the error of the finite element simulation with which the snapshots were generated.

The observations regarding the accuracy of the reduced-order solution can be summarized with the following statements: By increasing the dimension of the reduced-order model it is possible to approximate the underlying snapshots better and better. Still, the total error with respect to a reference solution of higher accuracy cannot be decreased below the accuracy of the snapshots.

The wall-clock time to solve the reduced-order model with $R^q = 7$ was approximately 0.10 s using one thread of an Intel Xeon E5-4650 CPU with Matlab R2012b. The computationally most expensive part of the solution was the evaluation of the non-linear term. The creation of the POD from the snapshots took about 0.31 s and the creation of the reduced-order model from the POD basis functions took about 0.046 s. Therefore, the reduced-order model is able to compute a solution with the same order of accuracy as the finite element simulation, but about two orders of magnitude faster, when considering only the solution times. The computation of the snapshots, however, involved a comparably expensive finite element simulation. Hence, the model should be applied in settings where the fast evaluation time pays off, for example when repetitive calls to the model are made or when a real-time response is desired.

Chapter 4

Isothermal flow around a cylinder

This chapter focuses on POD-Galerkin reduced-order modeling for the incompressible isothermal flow of a Newtonian fluid around a circular cylinder, which is a common benchmark problem in computational fluid dynamics. An overview of experimental and numerical studies is given in Norberg (2003). The flow around a circular cylinder has also been widely used to test reduced-order models for incompressible flow (see, e.g. Deane et al. (1991), Ma and Karniadakis (2002), Noack et al. (2003), Bergmann et al. (2005), Sirisup et al. (2005), Sirisup and Karniadakis (2005), Cordier et al. (2010)). An error analysis of a general equations in fluid dynamics, which includes the incompressible Navier-Stokes equations, is presented in Kunisch and Volkwein (2002a) for the implicit Euler method and in Kunisch and Volkwein (2002b) for the Crank-Nicolson method.

The ideal case of an infinitely long cylinder in an open domain is characterized by a single parameter, namely the Reynolds number Re, which is based on the cylinder diameter, the free stream velocity and the kinematic viscosity of the fluid. A two-dimensional vortex shedding flow appears between $Re \approx 47$ and $Re \approx 190$. In this parameter interval, the problem can be simplified using a two-dimensional setting. In typical experimental or numerical setups the domain is bounded. In this case, additional geometric parameters enter the problem description. Perhaps the most important parameter is the blockage ratio, which is defined as the ratio between the height of the domain and the diameter of the cylinder. Often it is argued that blockage ratio is large enough so that confinement effects are negligible. Posdziech and Grundmann (2007)

study such blockage effects in a systematic way. Their investigations reveal that the discrepancies between simulation results found in the literature can often by explained by the different sizes of the computational domains. Posdziech and Grundmann compute characteristic flow quantities for geometries of varying dimensions and varying Reynolds numbers. These data are used to validate the numerical codes of the finite element model and the reduced-order model.

In section 4.1 the governing equations of the problem are given in their strong form and subsequently and transferred to a weak form. A finite element model based on Taylor-Hood elements in connection with the Crank-Nicolson scheme is presented in section 4.2. The derivation of the reduced-order model, which is the subject of section 4.3, consists of two parts. The first part is dedicated to a model for the velocity field. The equivalence of the derivations from the continuous, semi-discrete and fully discrete equations is demonstrated. In all approaches, the discrete divergence-freeness of the POD basis functions eliminates the continuity equation and the pressure term. The second part consists of an additional model for the pressure, which is derived from the semi-discrete and fully discrete equations.

4.1 Governing equations

Let $\Omega \subset \mathbb{R}^2$ be a bounded spatial domain given by a non-empty open connected set with boundary Γ and closure $\bar{\Omega}$, and let $[0,T]$ be a time interval with a final time $T > 0$. The spatial coordinates are denoted by $\vec{x} = (x,y)^T$ and time is denoted by t. The unknowns of the considered unsteady incompressible Navier-Stokes flow problem are the dimensionless velocity vector $\vec{u} = (u,v)^T$, defined as a function $\vec{u} : [0,T] \times \bar{\Omega} \to \mathbb{R}^2$, and the scaled dimensionless pressure $p : [0,T] \times \bar{\Omega} \to \mathbb{R}$. The problem is governed by the set of partial differential algebraic equations

$$\partial_t u + \vec{u} \cdot \nabla u + \partial_x p - \frac{1}{Re} \nabla \cdot (\nabla u + \partial_x \vec{u}) = 0, \quad (t,\vec{x}) \in (0,T] \times \Omega, \tag{4.1}$$

$$\partial_t v + \vec{u} \cdot \nabla v + \partial_y p - \frac{1}{Re} \nabla \cdot (\nabla v + \partial_y \vec{u}) = 0, \quad (t,\vec{x}) \in (0,T] \times \Omega, \tag{4.2}$$

$$\nabla \cdot \vec{u} = 0, \quad (t,\vec{x}) \in (0,T] \times \Omega, \tag{4.3}$$

where Re is the positive dimensionless Reynolds number. The first two equations are called momentum equations. The third equation is called continuity equation.

The boundary is partitioned as $\Gamma = \Gamma_{\mathrm{D}}^u \cup \Gamma_{\mathrm{N}}^u = \Gamma_{\mathrm{D}}^v \cup \Gamma_{\mathrm{N}}^v$ with $\Gamma_{\mathrm{D}}^u \cap \Gamma_{\mathrm{N}}^u = \Gamma_{\mathrm{D}}^v \cap \Gamma_{\mathrm{N}}^v = \emptyset$. The surface unit normal vector pointing outward of the domain is

denoted by $\vec{n} = (n_x, n_y)^T$. A space-time continuous extension of the velocity Dirichlet boundary data into the domain is given by a vector $\vec{u}_\mathrm{D} = (u_\mathrm{D}, v_\mathrm{D})^T$ with $\vec{u}_\mathrm{D} : [0,T] \times \bar{\Omega} \to \mathbb{R}^2$, complying with the particular form

$$\vec{u}_\mathrm{D}(t,\vec{x}) = \sum_{k=1}^{K^{\vec{u}}} \vec{g}_k^{\vec{u}}(\vec{x}) \eta_{\vec{u}}^k(t), \qquad (t,\vec{x}) \in [0,T] \times \bar{\Omega},$$

where $K^{\vec{u}}$ is a small positive integer. The functions $\vec{g}_1^{\vec{u}}, \ldots, \vec{g}_{K^{\vec{u}}}^{\vec{u}} : \bar{\Omega} \to \mathbb{R}^2$ are continuous in space and $\eta_{\vec{u}}^1, \ldots, \eta_{\vec{u}}^{K^{\vec{u}}} : [0,T] \to \mathbb{R}$ are continuous in time. The following boundary conditions are prescribed:

$$u = u_\mathrm{D}(t,\vec{x}), \qquad (t,\vec{x}) \in (0,T] \times \Gamma_\mathrm{D}^u, \tag{4.4}$$

$$v = v_\mathrm{D}(t,\vec{x}), \qquad (t,\vec{x}) \in (0,T] \times \Gamma_\mathrm{D}^v, \tag{4.5}$$

$$pn_x - \frac{1}{Re}(\nabla u + \partial_x \vec{u}) \cdot \vec{n} = 0, \qquad (t,\vec{x}) \in (0,T] \times \Gamma_\mathrm{N}^u, \tag{4.6}$$

$$pn_y - \frac{1}{Re}(\nabla v + \partial_y \vec{u}) \cdot \vec{n} = 0, \qquad (t,\vec{x}) \in (0,T] \times \Gamma_\mathrm{N}^v. \tag{4.7}$$

If at least one of the velocity Neumann boundaries Γ_N^u or Γ_N^v is non-empty, the pressure field can be uniquely determined, because the value of the pressure enters the computation via (4.6) and (4.7). Otherwise the pressure field can be determined up to an arbitrary additive constant, because only spatial derivatives of the pressure field are entering the system of equations. In the latter case, the pressure can be fixed e.g. by introducing an artificial pressure Dirichlet point. This can be implemented on a discrete level in a straightforward manner and is not detailed, here.

Based on a continuous initial velocity $\vec{u}_0 = (u_0, v_0)^T$ with $\vec{u}_0 : \bar{\Omega} \to \mathbb{R}^2$, the initial conditions

$$u = u_0, \qquad (t,\vec{x}) \in \{0\} \times \Omega, \tag{4.8}$$

$$v = v_0, \qquad (t,\vec{x}) \in \{0\} \times \Omega, \tag{4.9}$$

are prescribed. It is assumed that the initial velocity is divergence free, $\nabla \cdot \vec{u}_0 = 0$ in $\bar{\Omega}$, and that it is compatible with the boundary conditions, $u_0 = u_\mathrm{D}$ on $\{0\} \times \Gamma_\mathrm{D}^u$ and $v_0 = v_\mathrm{D}$ on $\{0\} \times \Gamma_\mathrm{D}^v$.

A weak form of the problem provides the foundation of the finite element model and the reduced-order model of the flow simulation. As a preliminary, the subspaces $H_{u,0}^1(\Omega)$ and $H_{v,0}^1(\Omega)$ of the Sobolev space $H^1(\Omega)$ are defined:

$$H_{u,0}^1(\Omega) := \{a \in H^1(\Omega) : a|_{\Gamma_\mathrm{D}^u} = 0\},$$
$$H_{v,0}^1(\Omega) := \{a \in H^1(\Omega) : a|_{\Gamma_\mathrm{D}^v} = 0\}.$$

The equations (4.1), (4.2) and (4.3) as well as the initial conditions (4.8) and (4.9) are multiplied by respective test functions $\psi^u \in H^1_{u,0}(\Omega)$, $\psi^v \in H^1_{v,0}(\Omega)$ and $\psi^p \in L^2(\Omega)$. A subsequent integration over the domain yields

$$
\begin{aligned}
(\psi^u, \partial_t u + \vec{u} \cdot \nabla u) + (\psi^u, \partial_x p) - \frac{1}{Re}(\psi^u, \nabla \cdot (\nabla u + \partial_x \vec{u})) = 0, \quad t \in (0,T], \\
(\psi^v, \partial_t v + \vec{u} \cdot \nabla v) + (\psi^v, \partial_y p) - \frac{1}{Re}(\psi^v, \nabla \cdot (\nabla v + \partial_y \vec{u})) = 0, \quad t \in (0,T], \\
(\psi^p, \nabla \cdot \vec{u}) = 0, \quad t \in (0,T], \\
(\psi^u, u - u_0) = 0, \quad t \in \{0\}, \\
(\psi^v, v - v_0) = 0, \quad t \in \{0\}.
\end{aligned}
$$

The divergence theorem, applied to the pressure and viscous terms of the first two equations, leads to

$$
\begin{aligned}
&(\psi^u, \partial_t u + \vec{u} \cdot \nabla u) - (\partial_x \psi^u, p) + \frac{1}{Re}(\nabla \psi^u, \nabla u + \partial_x \vec{u}) \\
&\quad = \int_\Gamma -\psi^u p n_x \, \mathrm{d}\vec{x} + \int_\Gamma \frac{1}{Re} \psi^u (\nabla u + \partial_x \vec{u}) \cdot \vec{n} \, \mathrm{d}\vec{x}, \qquad (4.10) \\
&(\psi^v, \partial_t v + \vec{u} \cdot \nabla v) - (\partial_y \psi^v, p) + \frac{1}{Re}(\nabla \psi^v, \nabla v + \partial_y \vec{u}) \\
&\quad = \int_\Gamma -\psi^v p n_y \, \mathrm{d}\vec{x} + \int_\Gamma \frac{1}{Re} \psi^v (\nabla v + \partial_y \vec{u}) \cdot \vec{n} \, \mathrm{d}\vec{x}. \qquad (4.11)
\end{aligned}
$$

The boundary terms can be eliminated: At the Dirichlet boundaries they can be dropped because the test functions fulfill zero Dirichlet conditions. At the Neumann boundaries the conditions (4.6) and (4.7) can be substituted.

The pressure is assumed to be a member of $L^2(\Omega)$ and the velocity components are assumed to be members of $H^1(\Omega)$. The fulfillment of the Dirichlet conditions is assured explicitly. This leads to the final weak form of the flow problem (4.1) to (4.9): Find $p \in L^2(\Omega)$, $u - u_\mathrm{D} \in H^1_{u,0}(\Omega)$ and $v - v_\mathrm{D} \in H^1_{v,0}(\Omega)$, such that

$$
\begin{aligned}
&(\psi^u, \partial_t u + \vec{u} \cdot \nabla u) - (\partial_x \psi^u, p) \\
&\qquad + \frac{1}{Re}(\nabla \psi^u, \nabla u + \partial_x \vec{u}) = 0 \quad \forall \psi^u \in H^1_{u,0}(\Omega), \quad t \in (0,T], \qquad (4.12) \\
&(\psi^v, \partial_t v + \vec{u} \cdot \nabla v) - (\partial_y \psi^v, p) \\
&\qquad + \frac{1}{Re}(\nabla \psi^v, \nabla v + \partial_y \vec{u}) = 0 \quad \forall \psi^v \in H^1_{v,0}(\Omega), \quad t \in (0,T], \qquad (4.13) \\
&\qquad (\psi^p, \nabla \cdot \vec{u}) = 0 \quad \forall \psi^p \in L^2(\Omega), \quad t \in (0,T], \qquad (4.14)
\end{aligned}
$$

$$(\psi^u, u - u_0) = 0 \quad \forall \psi^u \in H^1_{u,0}(\Omega), \quad t \in \{0\}, \tag{4.15}$$
$$(\psi^v, v - v_0) = 0 \quad \forall \psi^v \in H^1_{v,0}(\Omega), \quad t \in \{0\}. \tag{4.16}$$

4.2 Finite element modeling

For the discretization of the flow problem in space, see subsection 4.2.1, Taylor-Hood mixed finite elements are used (Taylor and Hood, 1973). They comprise a continuous, piecewise linear approximation for the pressure and a continuous, piecewise quadratic approximation for the velocity components. Taylor-Hood elements are known to satisfy the inf-sup stability condition, see Gunzburger (1989) and references therein. Following the vertical method of lines, a subsequent time discretization is performed in subsection 4.2.2 using the Crank-Nicolson method (Crank and Nicolson, 1947). The finite element model is validated in subsection 4.2.3 using the drag, lift and base-pressure coefficients as well as the Strouhal number.

4.2.1 Finite element discretization in space

The domain Ω is triangulated so that no vertex is positioned at the edge of another triangle. Velocity mesh nodes are placed at the triangle vertices and the midpoints of each side. For notational consistency, the mesh nodes of the velocity components u and v are given separate symbols $\vec{x}^u_1, \ldots, \vec{x}^u_{M'^u}$ and $\vec{x}^v_1, \ldots, \vec{x}^v_{M'^v}$, respectively. Pressure mesh nodes $\vec{x}^p_1, \ldots, \vec{x}^p_{M'^p}$ are placed at the triangle vertices only. A distinction is made between the non-Dirichlet nodes represented by $\vec{x}^u_1, \ldots, \vec{x}^u_{M^u}$, $\vec{x}^v_1, \ldots, \vec{x}^v_{M^v}$, $\vec{x}^p_1, \ldots, \vec{x}^p_{M^p}$ and the Dirichlet nodes represented by $\vec{x}^u_{M^u+1}, \ldots, \vec{x}^u_{M'^u}$, $\vec{x}^v_{M^v+1}, \ldots, \vec{x}^v_{M'^v}$, $\vec{x}^p_{M^p+1}, \ldots, \vec{x}^p_{M'^p}$. Continuous and piecewise quadratic velocity finite element basis functions are denoted by $\psi^u_1, \ldots, \psi^u_{M'^u}$ and $\psi^v_1, \ldots, \psi^v_{M'^v}$, and continuous and piecewise linear pressure finite element basis functions are denoted by $\psi^p_1, \ldots, \psi^p_{M'^p}$. The basis functions fulfill the Lagrangian properties

$$\begin{aligned}
\psi^u_i(\vec{x}^u_j) &= \delta_{ij}, && \forall i, j = 1, \ldots, M'^u, \\
\psi^v_i(\vec{x}^v_j) &= \delta_{ij}, && \forall i, j = 1, \ldots, M'^v, \\
\psi^p_i(\vec{x}^p_j) &= \delta_{ij}, && \forall i, j = 1, \ldots, M'^p.
\end{aligned}$$

The finite element spaces are defined by

$$H^1_{u,h} := \operatorname{span}(\psi^u_1, \ldots, \psi^u_{M^u}, \psi^u_{M^u+1}, \ldots, \psi^u_{M'^u}), \tag{4.17}$$

$$H^1_{u,0,h} := \operatorname{span}(\psi^u_1, \ldots, \psi^u_{M^u}), \tag{4.18}$$

$$H^1_{v,h} := \operatorname{span}(\psi^v_1, \ldots, \psi^v_{M^v}, \psi^v_{M^v+1}, \ldots, \psi^v_{M'^v}), \tag{4.19}$$

$$H^1_{v,0,h} := \operatorname{span}(\psi^v_1, \ldots, \psi^v_{M^v}), \tag{4.20}$$

$$L^2_{p,h} := \operatorname{span}(\psi^p_1, \ldots, \psi^p_{M^p}, \psi^p_{M^p+1}, \ldots, \psi^p_{M'^p}), \tag{4.21}$$

$$H^1_{\vec{u},h} := \{\vec{u} = (u, v)^T : u \in H^1_{u,h}, v \in H^1_{v,h}\}, \tag{4.22}$$

$$H^1_{\vec{u},0,h} := \{\vec{u} = (u, v)^T : u \in H^1_{u,0,h}, v \in H^1_{v,0,h}\}, \tag{4.23}$$

where h symbolizes a spatial discretization parameter. The finite element coefficient vectors of the spatial semi-discretization are given by the time-dependent vector-valued functions $U : [0,T] \to \mathbb{R}^{M'^u}$, $V : [0,T] \to \mathbb{R}^{M'^v}$, $P : [0,T] \to \mathbb{R}^{M'^p}$ and $\vec{U} : [0,T] \to \mathbb{R}^{M'^u+M'^v}$ with the components

$$\begin{aligned}
U(t) &= (U^1(t), \ldots, U^{M'^u}(t))^T, && t = [0,T],\\
V(t) &= (V^1(t), \ldots, V^{M'^v}(t))^T, && t = [0,T],\\
P(t) &= (P^1(t), \ldots, P^{M'^p}(t))^T, && t = [0,T],\\
\vec{U}(t) &= (U^1(t), \ldots, U^{M'^u}(t), V^1(t), \ldots, V^{M'^v}(t))^T, && t = [0,T].
\end{aligned}$$

The finite element approximations u^h, v^h, p^h and $\vec{u}^h$ of the unknown fields are defined by

$$u(t,\vec{x}) \approx u^h(t,\vec{x}) = \sum_{m=1}^{M'^u} \psi^u_m(\vec{x}) U^m(t), \qquad (t,\vec{x}) \in [0,T] \times \bar{\Omega}, \tag{4.24}$$

$$v(t,\vec{x}) \approx v^h(t,\vec{x}) = \sum_{m=1}^{M'^v} \psi^v_m(\vec{x}) V^m(t), \qquad (t,\vec{x}) \in [0,T] \times \bar{\Omega}, \tag{4.25}$$

$$p(t,\vec{x}) \approx p^h(t,\vec{x}) = \sum_{m=1}^{M'^p} \psi^p_m(\vec{x}) P^m(t), \qquad (t,\vec{x}) \in [0,T] \times \bar{\Omega}, \tag{4.26}$$

$$\vec{u}(t,\vec{x}) \approx \vec{u}^h(t,\vec{x}) = (u^h(t,\vec{x}), v^h(t,\vec{x}))^T, \qquad (t,\vec{x}) \in [0,T] \times \bar{\Omega}. \tag{4.27}$$

The Dirichlet conditions are implemented using finite element interpolations at the Dirichlet mesh nodes:

$$u^h_{\mathrm{D}}(t,\vec{x}) = \sum_{m=1}^{M'^u} \psi^u_m(\vec{x}) u_{\mathrm{D}}(t, \vec{x}^u_m), \qquad (t,\vec{x}) \in [0,T] \times \bar{\Omega}, \tag{4.28}$$

$$v^h_{\mathrm{D}}(t,\vec{x}) = \sum_{m=1}^{M'^v} \psi^v_m(\vec{x}) v_{\mathrm{D}}(t, \vec{x}^v_m), \qquad (t,\vec{x}) \in [0,T] \times \bar{\Omega}. \tag{4.29}$$

The finite element representations of the initial data are also realized via finite element interpolations:

$$u_0^h(\vec{x}) = \sum_{m=1}^{M'^u} \psi_m^u u_0(\vec{x}_m^u), \qquad \vec{x} \in \bar{\Omega}, \tag{4.30}$$

$$v_0^h(\vec{x}) = \sum_{m=1}^{M'^v} \psi_m^v v_0(\vec{x}_m^v), \qquad \vec{x} \in \bar{\Omega}. \tag{4.31}$$

The respective finite element coefficient vectors are defined by

$$U_0 = \begin{pmatrix} u_0(\vec{x}_1^u) \\ \vdots \\ u_0(\vec{x}_{M'^u}^u) \end{pmatrix}, \qquad V_0 = \begin{pmatrix} v_0(\vec{x}_1^v) \\ \vdots \\ v_0(\vec{x}_{M'^v}^v) \end{pmatrix}.$$

A substitution of the finite element spaces (4.17) to (4.21) in the weak form (4.12) to (4.16) leads to the following problem: Find $p^h \in L_{p,h}^2$, $u^h - u_{\mathrm{D}}^h \in H_{u,0,h}^1$ and $v^h - v_{\mathrm{D}}^h \in H_{v,0,h}^1$, such that

$$\begin{aligned}
&(\psi^u, \partial_t u^h + \vec{u}^h \cdot \nabla u^h) - (\partial_x \psi^u, p^h) \\
&\quad + \frac{1}{Re}(\nabla \psi^u, \nabla u^h + \partial_x \vec{u}^h) = 0 \qquad \forall \psi^u \in H_{u,0,h}^1, \quad t \in (0,T], && (4.32)\\
&(\psi^v, \partial_t v^h + \vec{u}^h \cdot \nabla v^h) - (\partial_y \psi^v, p^h) \\
&\quad + \frac{1}{Re}(\nabla \psi^v, \nabla v^h + \partial_y \vec{u}^h) = 0 \qquad \forall \psi^v \in H_{v,0,h}^1, \quad t \in (0,T], && (4.33)\\
&(\psi^p, \nabla \cdot \vec{u}^h) = 0 \qquad \forall \psi^p \in L_{p,h}^2, \quad t \in (0,T], && (4.34)\\
&(\psi^u, u^h - u_0^h) = 0 \qquad \forall \psi^u \in H_{u,0,h}^1, \quad t \in \{0\}, && (4.35)\\
&(\psi^v, v^h - v_0^h) = 0 \qquad \forall \psi^v \in H_{v,0,h}^1, \quad t \in \{0\}. && (4.36)
\end{aligned}$$

The functions ψ^u, ψ^v and ψ^p in the discrete weak formulation (4.32) to (4.36) are substituted by the respective finite element basis functions in turn and the unknowns u^h, v^h, p^h and $\vec{u}^h$ are substituted by their finite element approximations (4.24) to (4.27). The resulting equations are:

- Momentum equation in x-direction for $t \in (0,T]$ and $i = 1, \dots, M^u$:

$$\sum_{j=1}^{M'^u} \underbrace{(\psi_i^u, \psi_j^u)}_{\mathcal{M}_u^{\mathrm{F}}} \dot{U}_j + \sum_{j=1}^{M'^u} \underbrace{\Big(\psi_i^u, \sum_{k=1}^{M'^u} U_k \psi_k^u \partial_x \psi_j^u + \sum_{k=1}^{M'^v} V_k \psi_k^v \partial_y \psi_j^u\Big)}_{\mathcal{N}_u^{\mathrm{F}}(\vec{U})} U_j$$
$$+ \sum_{j=1}^{M'^p} \underbrace{-(\partial_x \psi_i^u, \psi_j^p)}_{\mathcal{C}_u^{\mathrm{F}}} P_j + \sum_{j=1}^{M'^v} \underbrace{\frac{1}{Re}(\partial_y \psi_i^u, \partial_x \psi_j^v)}_{\mathcal{K}_{uv}^{\mathrm{F}}} V_j$$
$$+ \sum_{j=1}^{M'^u} \underbrace{\frac{1}{Re}\Big(2(\partial_x \psi_i^u, \partial_x \psi_j^u) + (\partial_y \psi_i^u, \partial_y \psi_j^u)\Big)}_{\mathcal{K}_{uu}^{\mathrm{F}}} U_j = 0.$$

- Momentum equation in y-direction for $t \in (0,T]$ and $i = 1, \dots, M^v$:

$$\sum_{j=1}^{M'^v} \underbrace{(\psi_i^v, \psi_j^v)}_{\mathcal{M}_v^{\mathrm{F}}} \dot{V}_j + \sum_{j=1}^{M'^u} \underbrace{\Big(\psi_i^v, \sum_{k=1}^{M'^u} U_k \psi_k^u \partial_x \psi_j^v + \sum_{k=1}^{M'^v} V_k \psi_k^v \partial_y \psi_j^v\Big)}_{\mathcal{N}_v^{\mathrm{F}}(\vec{U})} V_j$$
$$+ \sum_{j=1}^{M'^p} \underbrace{-(\partial_y \psi_i^v, \psi_j^p)}_{\mathcal{C}_v^{\mathrm{F}}} P_j + \sum_{j=1}^{M'^u} \underbrace{\frac{1}{Re}(\partial_x \psi_i^v, \partial_y \psi_j^u)}_{\mathcal{K}_{vu}^{\mathrm{F}}} U_j$$
$$+ \sum_{j=1}^{M'^v} \underbrace{\frac{1}{Re}\Big((\partial_x \psi_i^v, \partial_x \psi_j^v) + 2(\partial_y \psi_i^v, \partial_y \psi_j^v)\Big)}_{\mathcal{K}_{vv}^{\mathrm{F}}} V_j = 0.$$

- Continuity equation for $t \in (0,T]$ and $i = 1, \dots, M^p$:

$$\sum_{j=1}^{M'^u} \underbrace{(\psi_i^p, \partial_x \psi_j^u)}_{\mathcal{D}_u^{\mathrm{F}}} U_j + \sum_{j=1}^{M'^v} \underbrace{(\psi_i^p, \partial_y \psi_j^v)}_{\mathcal{D}_v^{\mathrm{F}}} V_j = 0.$$

- Initial condition in x-direction for $t \in \{0\}$ and $i = 1, \dots, M^u$:

$$\sum_{j=1}^{M'^u} \underbrace{(\psi_i^u, \psi_j^u)}_{\mathcal{M}_u^{\mathrm{F}}} (U_j - u_0(\vec{x}_j^u)) = 0.$$

- Initial condition in y-direction for $t \in \{0\}$ and $i = 1, \ldots, M^v$:

$$\sum_{j=1}^{M'^v} \underbrace{(\psi_i^v, \psi_j^v)}_{\mathcal{M}_v^{\mathrm{F}}} (V_j - v_0(\vec{x}_j^v)) = 0.$$

Knowing the finite element basis functions for a given mesh, the inner products can be evaluated. The braces in the formulas above indicate which terms enter which matrices. The number of matrix columns is equal to the total number of nodes. The number of matrix rows is equal to the number of non-Dirichlet nodes. The superscript F indicates that only these 'free' rows are contained in the matrices. In the actual implementation on a computer, the square matrices containing all rows are computed. The non-Dirichlet matrix rows are extracted when needed. For the computation of the drag and the lift coefficients later on, it is handy to have the matrices containing the Dirichlet rows available as well. These are marked with a superscript D.

The resulting matrix form of the discretized equations is given by

$$\begin{aligned}
\mathcal{M}_u^{\mathrm{F}} \dot{U} + \mathcal{N}_u^{\mathrm{F}}(\vec{U})U + \mathcal{K}_{uu}^{\mathrm{F}} U + \mathcal{K}_{uv}^{\mathrm{F}} V + \mathcal{C}_u^{\mathrm{F}}\, P = 0, \qquad & t \in (0,T], \\
\mathcal{M}_v^{\mathrm{F}} \dot{V} + \mathcal{N}_u^{\mathrm{F}}(\vec{U})V + \mathcal{K}_{vu}^{\mathrm{F}} U + \mathcal{K}_{vv}^{\mathrm{F}} V + \mathcal{C}_v^{\mathrm{F}}\, P = 0, \qquad & t \in (0,T], \\
\mathcal{D}_u^{\mathrm{F}}\, U + \mathcal{D}_v^{\mathrm{F}}\, V = 0, \qquad & t \in (0,T], \\
\mathcal{M}_u^{\mathrm{F}} U - \mathcal{M}_u^{\mathrm{F}} U_0 = 0, \qquad & t \in \{0\}, \\
\mathcal{M}_v^{\mathrm{F}} V - \mathcal{M}_v^{\mathrm{F}} V_0 = 0, \qquad & t \in \{0\},
\end{aligned}$$

and the Dirichlet entries of the solution vectors are prescribed using

$$U^m(t) = u_{\mathrm{D}}(t, \vec{x}_m^u), \qquad m = M^u + 1, \ldots, M'^u, \qquad t \in [0,T], \tag{4.37}$$

$$V^m(t) = v_{\mathrm{D}}(t, \vec{x}_m^v), \qquad m = M^v + 1, \ldots, M'^v, \qquad t \in [0,T]. \tag{4.38}$$

A more compact notation can be obtained after defining

$$\begin{aligned}
\mathcal{M}_{\vec{u}}^{\mathrm{F}} &= \begin{pmatrix} \mathcal{M}_u^{\mathrm{F}} & 0 \\ 0 & \mathcal{M}_v^{\mathrm{F}} \end{pmatrix}, & \mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{U}) &= \begin{pmatrix} \mathcal{N}_u^{\mathrm{F}}(\vec{U}) & 0 \\ 0 & \mathcal{N}_v^{\mathrm{F}}(\vec{U}) \end{pmatrix}, \\
\mathcal{C}_{\vec{u}}^{\mathrm{F}} &= \begin{pmatrix} \mathcal{C}_u^{\mathrm{F}} \\ \mathcal{C}_v^{\mathrm{F}} \end{pmatrix}, & \mathcal{K}_{\vec{u}}^{\mathrm{F}} &= \begin{pmatrix} \mathcal{K}_{uu}^{\mathrm{F}} & \mathcal{K}_{uv}^{\mathrm{F}} \\ \mathcal{K}_{vu}^{\mathrm{F}} & \mathcal{K}_{vv}^{\mathrm{F}} \end{pmatrix}, \\
\mathcal{D}_{\vec{u}}^{\mathrm{F}} &= \begin{pmatrix} \mathcal{D}_u^{\mathrm{F}} & \mathcal{D}_v^{\mathrm{F}} \end{pmatrix}, & \vec{U}_0 &= (U_0, V_0)^T.
\end{aligned}$$

From the definitions of the matrix elements it can be deduced that

$$M^p = M'^p \Rightarrow \mathcal{D}_{\vec{u}}^{\mathrm{F}} = (\mathcal{C}_{\vec{u}}^{\mathrm{F}})^T, \tag{4.39}$$

which can be applied if no artificial pressure Dirichlet points are involved. The compact notation leads to the following form of the finite element semi-discretization of the flow problem:

$$\mathcal{M}_{\vec{u}}^{\mathrm{F}}\dot{\vec{U}} + \mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{U})\vec{U} + \mathcal{C}_{\vec{u}}^{\mathrm{F}}\, P + \mathcal{K}_{\vec{u}}^{\mathrm{F}}\, \vec{U} = \vec{0}, \qquad t \in (0,T], \tag{4.40}$$

$$\mathcal{D}_{\vec{u}}^{\mathrm{F}}\, \vec{U} = 0, \qquad t \in (0,T], \tag{4.41}$$

$$\mathcal{M}_{\vec{u}}^{\mathrm{F}}\vec{U} - \mathcal{M}_{\vec{u}}^{\mathrm{F}}\vec{U}_0 = \vec{0}, \qquad t \in \{0\}. \tag{4.42}$$

The first two equations form a system of differential algebraic equations, with (4.40) being the differential part, (4.41) being the algebraic part. The spatially discretized problem is completed by the initial conditions (4.42) and the boundary conditions (4.37), (4.38).

Note that a time derivative of the pressure appears neither in the strong or weak form of the governing equations nor in the semi-discretized equations. Still, the time discretization scheme presented in the next section requires the pressure field at the initial time. As a remedy, the initial pressure field is determined by a pressure Poisson equation, which depends on the initial velocity field and the initial acceleration at the Dirichlet boundary. A pressure Poisson equation is not only valid at the initial time but also throughout the time domain, and is used in other numerical schemes to substitute the continuity equation. An overview and discussion of such schemes can be found in Gresho and Sani (2000). In the finite element model presented in the current work, however, the pressure equation is only used to compute valid starting values for the pressure. In the reduced-order model presented later, a pressure Poisson equation is employed throughout the time domain in order to compute pressure-related quantities.

In the following, the derivation of a pressure Poisson equation is presented, starting from the spatially semi-discretized momentum equation (4.40). The accelerations at the free nodes are separated on the left-hand side of the system of equations, so that

$$\mathcal{M}_{\vec{u}}^{\mathrm{FF}}\dot{\vec{U}}^{\mathrm{F}} = -\mathcal{M}_{\vec{u}}^{\mathrm{FD}}\dot{\vec{U}}^{\mathrm{D}} - \mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{U})\vec{U} - \mathcal{K}_{\vec{u}}^{\mathrm{F}}\, \vec{U} - \mathcal{C}_{\vec{u}}^{\mathrm{F}}\, P, \qquad t \in (0,T],$$

with the coefficient vectors given by

$$\vec{U}^{\mathrm{F}}(t) = (U^1(t), \ldots, U^{M^u}(t), V^1(t), \ldots, V^{M^v}(t))^T, \qquad t \in [0,T]$$

$$\vec{U}^{\mathrm{D}}(t) = (U^{M^u+1}(t), \ldots, U^{M'^u}(t), V^{M^u+1}(t), \ldots, V^{M'^v}(t))^T, \quad t \in [0,T]$$

and the matrices $\mathcal{M}_{\vec{u}}^{\mathrm{FF}}$ and $\mathcal{M}_{\vec{u}}^{\mathrm{FD}}$ defined in correspondence. Assuming $\mathcal{M}_{\vec{u}}^{\mathrm{FF}}$

is invertible, the equation can be recast as

$$\dot{\vec{U}}^{\mathrm{F}} = -(\mathcal{M}_{\vec{u}}^{\mathrm{FF}})^{-1}(\mathcal{M}_{\vec{u}}^{\mathrm{FD}}\dot{\vec{U}}^{\mathrm{D}} + \mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{U})\vec{U} + \mathcal{K}_{\vec{u}}^{\mathrm{F}}\,\vec{U} + \mathcal{C}_{\vec{u}}^{\mathrm{F}}\,P), \qquad t \in (0,T]. \tag{4.43}$$

The time derivative of the spatially semi-discretized continuity equation (4.41) is given by

$$\mathcal{D}_{\vec{u}}^{\mathrm{F}}\,\dot{\vec{U}} = \mathcal{D}_{\vec{u}}^{\mathrm{FD}}\dot{\vec{U}}^{\mathrm{D}} + \mathcal{D}_{\vec{u}}^{\mathrm{FF}}\dot{\vec{U}}^{\mathrm{F}} = 0, \qquad t \in (0,T]. \tag{4.44}$$

It is assumed that (4.43) and (4.44) are valid at the initial time $t = 0$ as well. Then, substituting the former in the latter gives

$$0 = \mathcal{D}_{\vec{u}}^{\mathrm{FD}}\dot{\vec{U}}^{\mathrm{D}} - \mathcal{D}_{\vec{u}}^{\mathrm{FF}}(\mathcal{M}_{\vec{u}}^{\mathrm{FF}})^{-1}\Big(\mathcal{M}_{\vec{u}}^{\mathrm{FD}}\dot{\vec{U}}^{\mathrm{D}} + \mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{U})\vec{U} + \mathcal{K}_{\vec{u}}^{\mathrm{F}}\,\vec{U} + \mathcal{C}_{\vec{u}}^{\mathrm{F}}\,P\Big), \qquad t \in [0,T]. \tag{4.45}$$

This equation is often called a discrete pressure Poisson equation. The similarity to a classical Poisson equation can be discovered when the pressure term is shifted to the left-hand side. The equations can be used as a discrete initial condition for the pressure field,

$$\mathcal{D}_{\vec{u}}^{\mathrm{FF}}(\mathcal{M}_{\vec{u}}^{\mathrm{FF}})^{-1}\mathcal{C}_{\vec{u}}^{\mathrm{F}}\,P = \mathcal{D}_{\vec{u}}^{\mathrm{FD}}\dot{\vec{U}}_0^{\mathrm{D}} - \mathcal{D}_{\vec{u}}^{\mathrm{FF}}(\mathcal{M}_{\vec{u}}^{\mathrm{FF}})^{-1}\Big(\mathcal{M}_{\vec{u}}^{\mathrm{FD}}\dot{\vec{U}}_0^{\mathrm{D}} + \mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{U}_0)\vec{U}_0 + \mathcal{K}_{\vec{u}}^{\mathrm{F}}\,\vec{U}_0\Big), \qquad t \in \{0\}. \tag{4.46}$$

The vector $\dot{\vec{U}}_0^{\mathrm{D}}$ contains the initial accelerations at the Dirichlet boundary points, which can be computed analytically or numerically from the boundary data u_{D}. An important special case is provided by problems which start from a fluid at rest with zero velocity and zero acceleration at the Dirichlet boundary. In this case all terms on the right-hand side of (4.46) are zero and, hence, the initial pressure is zero.

The final coupled semi-discrete problem consists of the discrete velocity boundary conditions (4.37) and (4.38), the set of ordinary differential equations (4.40), the set of linear algebraic equations (4.41), the initial conditions for the velocity (4.42) and the initial conditions for the pressure (4.46).

4.2.2 Finite difference discretization in time

For the time discretization of the spatially semi-discretized problem, at first a time grid $0 = t_1 < \cdots < t_N = T$ is introduced. Then, for $n = 1, \ldots, N$ the

fully discrete solution vectors U_n, V_n, P_n and $\vec{U}_n$ are introduced as approximations to the solution vectors of the spatial semi-discretization:

$$U(t_n) \approx U_n = \begin{pmatrix} U_n^1 \\ \vdots \\ U_n^{M'^u} \end{pmatrix}, \qquad V(t_n) \approx V_n = \begin{pmatrix} V_n^1 \\ \vdots \\ V_n^{M'^v} \end{pmatrix}, \tag{4.47}$$

$$P(t_n) \approx P_n = \begin{pmatrix} P_n^1 \\ \vdots \\ P_n^{M'^p} \end{pmatrix}, \qquad \vec{U}(t_n) \approx \vec{U}_n = \begin{pmatrix} U_n \\ V_n \end{pmatrix}. \tag{4.48}$$

The fully discretized equations are derived from (4.37), (4.38), (4.40), (4.41), (4.42) and (4.46). The trapezoidal rule, also called Crank-Nicolson method, is applied to the equations defined over the time interval. It consists of approximating the time derivatives by central finite differences and replacing the remaining terms by an average of evaluations at the old and at the new time level. This leads to the computational scheme

$$\begin{aligned} &\mathcal{M}_{\vec{u}}^{\mathrm{F}} \frac{\vec{U}_n - \vec{U}_{n-1}}{t_n - t_{n-1}} \\ &\quad + \frac{\mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{U}_n)\vec{U}_n + \mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{U}_{n-1})\vec{U}_{n-1}}{2} \\ &\quad + \mathcal{K}_{\vec{u}}^{\mathrm{F}} \frac{\vec{U}_n + \vec{U}_{n-1}}{2} \\ &\quad + \mathcal{C}_{\vec{u}}^{\mathrm{F}} \frac{P_n + P_{n-1}}{2} = \vec{0}, \qquad n = 2, \ldots, N, \end{aligned} \tag{4.49}$$

$$\mathcal{D}_{\vec{u}}^{\mathrm{F}} \frac{\vec{U}_n + \vec{U}_{n-1}}{2} = 0, \qquad n = 2, \ldots, N, \tag{4.50}$$

$$\mathcal{M}_{\vec{u}}^{\mathrm{F}} \vec{U}_n - \mathcal{M}_{\vec{u}}^{\mathrm{F}} \vec{U}_0 = \vec{0}, \qquad n = 1. \tag{4.51}$$

$$\begin{aligned} &\mathcal{D}_{\vec{u}}^{\mathrm{FF}} (\mathcal{M}_{\vec{u}}^{\mathrm{FF}})^{-1} \mathcal{C}_{\vec{u}}^{\mathrm{F}} P_n = \mathcal{D}_{\vec{u}}^{\mathrm{FD}} \dot{\vec{U}}_0^{\mathrm{D}} - \mathcal{D}_{\vec{u}}^{\mathrm{FF}} (\mathcal{M}_{\vec{u}}^{\mathrm{FF}})^{-1} \Big(\\ &\qquad \mathcal{M}_{\vec{u}}^{\mathrm{FD}} \dot{\vec{U}}_0^{\mathrm{D}} + \mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{U}_0)\vec{U}_0 + \mathcal{K}_{\vec{u}}^{\mathrm{F}} \vec{U}_0 \Big), \qquad n = 1, \end{aligned} \tag{4.52}$$

$$U_n^m = u_{\mathrm{D}}(t_n, \vec{x}_m^u), \qquad m = M^u + 1, \ldots, M'^u, \qquad n = 1, \ldots, N, \tag{4.53}$$

$$V_n^m = v_{\mathrm{D}}(t_n, \vec{x}_m^v), \qquad m = M^v + 1, \ldots, M'^v, \qquad n = 1, \ldots, N. \tag{4.54}$$

At first, $\vec{U}_1$ and P_1 are computed from $\vec{U}_0$ and $\dot{\vec{U}}_0^{\mathrm{D}}$ via (4.51) and (4.52). In each subsequent step of the time-marching procedure, the coupled non-linear system of (4.49) and (4.50) is solved for the non-Dirichlet components of $\vec{U}_n$

and P_n. To this end, a simplified Newton method is applied, which transforms the non-linear problem to a sequence of linear algebraic sub-problems. An LU-decomposition is used to solve these linear systems of equations.

In order to obtain discretely divergence-free velocity approximations for $n = 2, \ldots, N$, the initial velocity must be discretely divergence-free, i.e. $\mathcal{D}_{\vec{u}}^{\mathrm{F}} \vec{U}_1 = 0$ (see Gresho and Sani, 2000). Under this condition, the discrete continuity equations (4.50) can also be reformulated into $\mathcal{D}_{\vec{u}}^{\mathrm{F}} \vec{U}_n = 0$ for $n = 2, \ldots, N$. The numerical results presented below are not influenced by the whether the original or the reformulated equation is used. However, the convergence of the reduced-order models toward the finite element snapshots, shown in section 4.3, is better if the reformulated version has been used to compute the snapshots. The reason is that the pressure finite element solution resulting from the original version contains a slight wiggle signal in time, which is not captured by the reduced-order model.

4.2.3 Validation

In this section, the finite element discretization of the isothermal flow problem is validated using a mesh and time step refinement study for a model problem defined in Posdziech and Grundmann (2007). Like in the previous chapter, all computations are performed with Matlab. The geometry and the assignment of boundary conditions is sketched in Figure 4.1. The diameter of the cylinder is chosen as $D = 1$, the length of the domain as $L = 70$ and the height of the domain as $H = 40$. This is equivalent to the case $L_{b|i}/D = 20$ of Posdziech and Grundmann. A simulation time of $T = 400$ is used. As initial velocity components, $u_0 = 0$ and $v_0 = 0$ were chosen. The Dirichlet boundary condition for the vertical velocity was set to $v_{\mathrm{D}} = 0$ at all Dirichlet boundaries. For the horizontal velocity component, $u_{\mathrm{D}} = 0$ was used at the cylinder boundary, while at the outer boundaries an initial run-up from $u_{\mathrm{D}} = 0$ to $u_{\mathrm{D}} = U_\infty$ was implemented using

$$u_{\mathrm{D}} = \begin{cases} \frac{1}{2}\left(1 - \cos\left(\frac{\pi t}{10}\right)\right) U_\infty & \text{if } t < 10, \\ U_\infty & \text{if } t \geq 10. \end{cases}$$

Note that the Dirichlet condition at $t = 0$ complies with the initial condition and that the time-derivative of the Dirichlet boundary data is zero. The free-stream velocity is chosen as $U_\infty = 1$ and the Reynolds number as $Re = 100$.

The space domain is discretized with non-uniform grids that are manually refined close to the cylinder and in the wake region. Four refinement levels are considered, with grids that consist of $M'^p = 324, 1\,308, 5\,754, 23\,582$ pressure nodes. For the construction of the grids, the Matlab PDE toolbox is used,

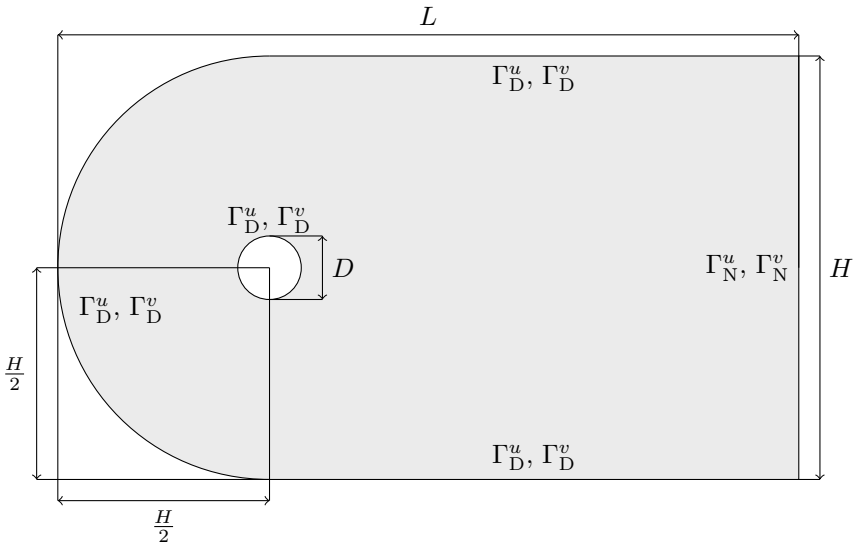

Figure 4.1: Sketch of the geometry and assignment of boundary conditions for the isothermal flow problem.

which automatically refines the cylinder boundary, so that the cylinder is approximated by a polygon with an increasing number of sides. The mesh of the spatial discretization with $M'^p = 5\,754$ is presented in Figure 4.2. The time domain is discretized with uniform step sizes of $\Delta t = 0.4, 0.2, 0.1, 0.05$.

The components of the solution at $t = 300$ are presented in Figure 4.3 for the simulation with $M'^p = 23\,582$ and $\Delta t = 0.05$. The velocity field exhibits a thin boundary layer at the cylinder and a prominent von Kármán vortex street.

Plots of the velocity components and the pressure at a sample point are shown in Figure 4.4. The sample point was located two cylinder diameters downstream of the cylinder midpoint. It can be seen that after the run-up phase the solution undergoes a transition to a time-periodic behavior.

The accuracy of the solution is assessed using the Strouhal number, the drag and lift coefficients as well as the base-pressure coefficients. The Strouhal

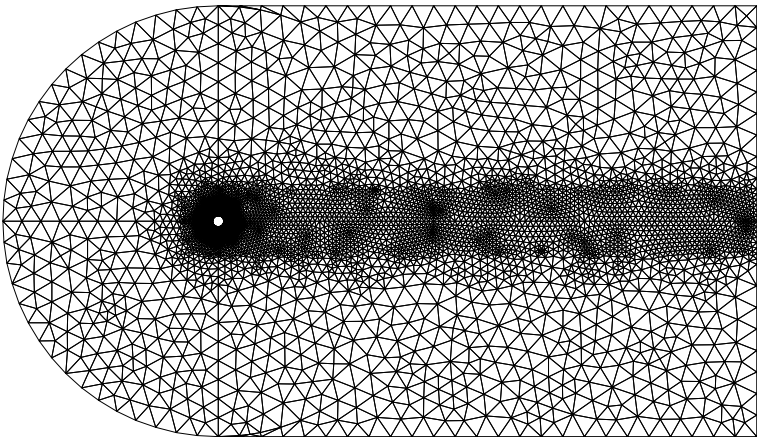

Figure 4.2: Mesh of the isothermal flow problem with $M'^p = 5\,754$.

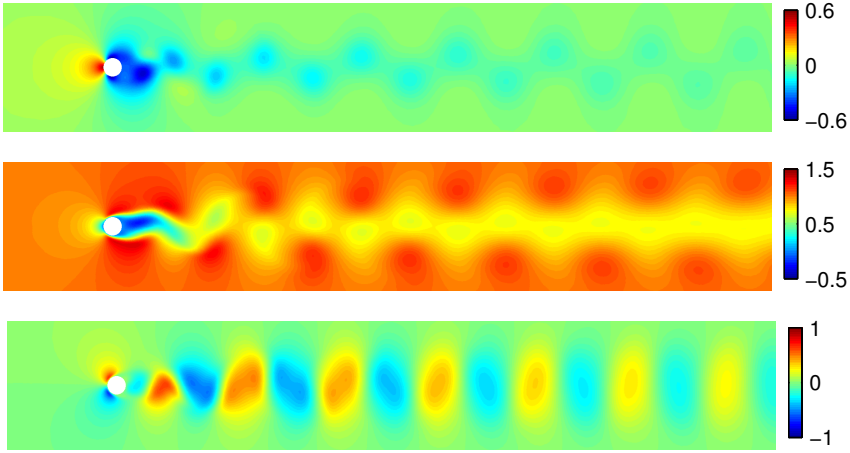

Figure 4.3: Cutouts of the solutions close to the cylinder at $t = 300$. From top to bottom: p, u, v.

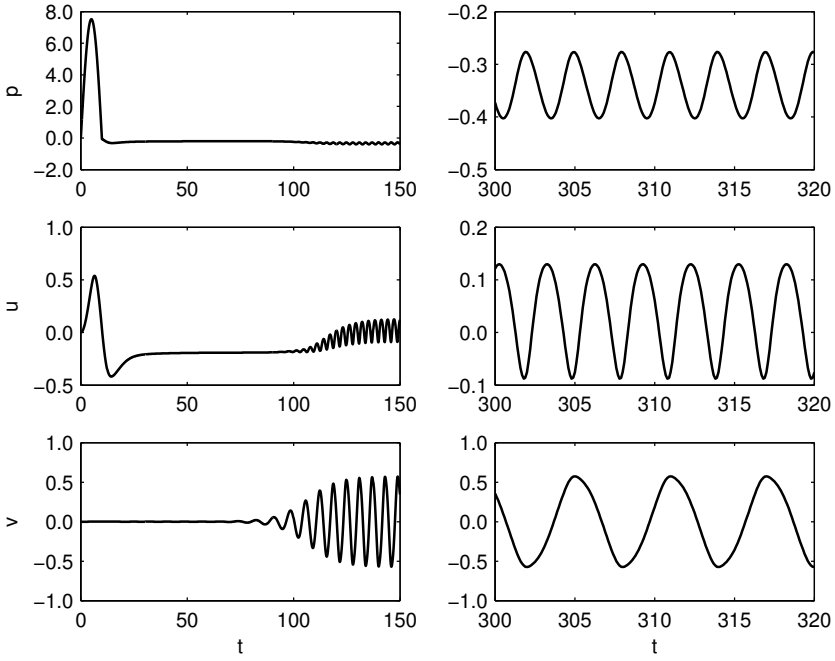

Figure 4.4: Time evolution of the components of the solution at a sample point in the cylinder wake. Left: Transient solutions at the startup of the simulation. Right: Periodic solutions at later simulation times.

number St of a time-periodic solution is defined by

$$St = \frac{fD}{U_\infty},$$

where f is the frequency of the solution. The drag coefficient C_D and the lift coefficient C_L are defined by

$$C_\mathrm{D} = \frac{2}{DU_\infty^2} \int_S \left(pn_x - \frac{1}{Re} \left(\nabla u + \partial_x \vec{u} \right) \cdot \vec{n} \right) \mathrm{d}S,$$
$$C_\mathrm{L} = \frac{2}{DU_\infty^2} \int_S \left(pn_y - \frac{1}{Re} \left(\nabla v + \partial_y \vec{u} \right) \cdot \vec{n} \right) \mathrm{d}S,$$

respectively, where S denotes the surface of the cylinder. The computation of the drag and lift coefficients is realized using a weak implementation (see John, 2004), which leads to the discrete approximations

$$C_\mathrm{D}^h(t) = \frac{2}{DU_\infty^2} \mathbb{1} \left(\mathcal{N}_u^\mathrm{D}(\vec{U})U + \mathcal{K}_{uu}^\mathrm{D} U + \mathcal{K}_{uv}^\mathrm{D} V + \mathcal{C}_u^\mathrm{D}\ P \right), \tag{4.55}$$
$$C_\mathrm{L}^h(t) = \frac{2}{DU_\infty^2} \mathbb{1} \left(\mathcal{N}_u^\mathrm{D}(\vec{U})V + \mathcal{K}_{vu}^\mathrm{D} U + \mathcal{K}_{vv}^\mathrm{D} V + \mathcal{C}_v^\mathrm{D}\ P \right), \tag{4.56}$$

where the matrices correspond to the Dirichlet rows of the finite element matrices and $\mathbb{1}$ denotes a vector of ones of corresponding size. The mean drag and the maximum absolute lift coefficients are used as quantities of interest, in accordance with Posdziech and Grundmann (2007). Finally, the base-pressure coefficient is defined by

$$C_\mathrm{P}^h(t) = 2\frac{p^h(t, \vec{x}_\mathrm{base}) - p^h(t, \vec{x}_\mathrm{inlet})}{U_\infty^2}, \tag{4.57}$$

where $\vec{x}_\mathrm{base}$ denotes the rear stagnation point at the cylinder and $\vec{x}_\mathrm{inlet}$ is the left-most point of the domain. The time average of the base-pressure coefficient is used as a quantity of interest.

With increasing resolution in space and time, the quantities of interest are supposed to converge toward the reference values provided by Posdziech and Grundmann (2007) with four decimal digits. In Table 4.1, their values depending on the spatial grid resolution are given for a fixed time step size of $\Delta t = 0.05$. The results are already relatively well converged for $M'^p = 5\,754$, because the results change only in the fourth decimal place compared

Table 4.1: Convergence of the quantities of interest with respect to the spatial refinement for a fixed temporal mesh with a high resolution.

M'^p	324	1308	5754	23582	ref.
N	8001	8001	8001	8001	ref.
Strouhal	0.1344	0.1658	0.1666	0.1667	0.1667
lift	0.1668	0.3261	0.3308	0.3309	0.3309
drag	1.2781	1.3425	1.3501	1.3504	1.3504
base-pressure	0.5937	0.7393	0.7447	0.7448	0.7448

Table 4.2: Convergence of the quantities of interest with respect to the time refinement for a fixed spatial mesh with a high resolution.

M'^p	23582	23582	23582	23582	ref.
N	1001	2001	4001	8001	ref.
Strouhal	0.1646	0.1660	0.1665	0.1667	0.1667
lift	0.3303	0.3309	0.3309	0.3309	0.3309
drag	1.3534	1.3504	1.3504	1.3504	1.3504
base-pressure	0.7484	0.7448	0.7448	0.7448	0.7448

to $M'^p = 23\,582$. The finest resolution, however, is needed to recover all quantities of interest with the accuracy provided by the reference.

The values of quantities of interest depending on the time resolution are given in Table 4.2 for the spatial grid with $M'^p = 23\,582$. It appears that the drag, lift and base-pressure coefficients are relatively insensitive to the time step size, because the simulation with $N = 2001$ or, equivalently, $\Delta t = 0.2$ was already sufficient to recover the reference results up to four decimal digits. For an accurate prediction of the Strouhal number, however, the smallest time step size of $\Delta t = 0.05$ was necessary ($N = 8001$).

4.3 Reduced-order modeling

In this section a reduced-order model for the velocity is derived in three different ways, similarly to the reduced-order heat conduction model of the previous chapter. In the first derivation (subsection 4.3.1), the weak formulation of the flow problem is projected onto the space spanned by the reduced basis functions. In the second derivation (subsection 4.3.2), the spatial finite element semi-discretization is projected onto the space spanned by the reduced basis vectors. Both approaches lead to the same reduced-order ODE system that is subsequently discretized in time (subsection 4.3.4). In the third derivation (subsection 4.3.5), the fully discretized finite element model is projected onto the space spanned by the reduced basis vectors, which leads to a time-discrete reduced-order model that is equal to the discretized reduced-order ODE system.

The derivations of the velocity model feature an elimination of the pressure term and the continuity equation due to the discrete divergence-free property of the basis functions. This standard procedure is used in many publications on reduced-order models for incompressible flow in fixed domains. Still, some applications require the computation of the pressure field or related quantities. For this purpose, a reduced-order model for the pressure is derived, building on the ideas of Akhtar et al. (2009), by projecting a discrete pressure Poisson equation onto the space spanned by the pressure reduced basis vectors. The derivation is given for the spatially semi-discretized setting (subsection 4.3.3) and the fully discretized setting (subsection 4.3.6).

Some alternative approaches to reduced-order modeling for the pressure field are mentioned in the following. In Rempfer (1996) and Noack et al. (2005) POD approximations of the pressure are substituted in the right-hand side of a continuous pressure Poisson equation. The resulting decoupled Poisson equations are solved for a set of partial pressures. A reduced-order pressure approximation is obtained with a linear combination of these partial pressures,

which involves the solution of a velocity reduced-order model. In Ullmann et al. (2013) this approach is transferred to a discrete pressure Poisson equation (Gresho and Sani, 2000). In Akhtar et al. (2009) the Galerkin projection of a continuous pressure Poisson equation on a set of pressure reduced basis functions is presented and tested for the flow around a circular cylinder at $Re = 100$. This approach is applied in Ullmann and Lang (2010) to the flows around circular cylinders at $Re = 100$ and $Re = 3\,900$. For the latter case, a Smagorinsky model (Smagorinsky, 1963) was included to model the contributions of small scale fluctuations. In Ullmann and Lang (2012) the approach of Akhtar et al. (2009) is applied to a discrete pressure Poisson equation. All these methods have in common that they use a velocity reduced basis with divergence-free basis functions. If the divergence-free property can not be guaranteed, a coupled velocity-pressure system with a saddle point structure arises. A treatment of such problems in the context of the reduced basis method for the Stokes equations is presented in Rozza and Veroy (2007).

4.3.1 Velocity model via weak form

A set of solution time instances $0 = t_1 < \cdots < t_N = T$ and respective velocity snapshot fields $\vec{u}_1^h, \ldots, \vec{u}_N^h \in H^1_{\vec{u},h}$ are provided by the finite element simulation. The goal is to derive a reduced-order model with $R^{\vec{u}}$ degrees of freedom. For given snapshots, the control function approach or the modified basis function approach provide sets of functions $\vec{\phi}_1^{\vec{u}}, \ldots, \vec{\phi}_{R^{\vec{u}}}^{\vec{u}} \in H^1_{\vec{u},0,h}$ and $\vec{\phi}_{R^{\vec{u}}+1}^{\vec{u}}, \ldots, \vec{\phi}_{R'^{\vec{u}}}^{\vec{u}} \in H^1_{\vec{u},h}$ for $R'^{\vec{u}} = R^{\vec{u}} + K^{\vec{u}}$. The respective reduced-order spaces $H^1_{\vec{u},R} \subset H^1_{\vec{u},h}$ and $H^1_{\vec{u},0,R} \subset H^1_{\vec{u},0,h}$ are given by

$$
\begin{aligned}
H^1_{\vec{u},R} &= \operatorname{span}\{\vec{\phi}_1^{\vec{u}}, \ldots, \vec{\phi}_{R^{\vec{u}}}^{\vec{u}}, \vec{\phi}_{R^{\vec{u}}+1}^{\vec{u}}, \ldots, \vec{\phi}_{R'^{\vec{u}}}^{\vec{u}}\}, \\
H^1_{\vec{u},0,R} &= \operatorname{span}\{\vec{\phi}_1^{\vec{u}}, \ldots, \vec{\phi}_{R^{\vec{u}}}^{\vec{u}}\}.
\end{aligned}
$$

A reduced-order representation is defined by the linear combination of basis functions

$$
\vec{u}^h(t) \approx \vec{u}^R(t) = \sum_{i=1}^{R'^{\vec{u}}} \vec{\phi}_i^{\vec{u}} a^i(t), \qquad t \in [0,T] \tag{4.58}
$$

with arbitrary coefficient functions $a^1, \ldots, a^{R'^{\vec{u}}} : [0,T] \to \mathbb{R}$.

Particular coefficient functions $\bar{a}^{R^{\vec{u}}+1}, \ldots, \bar{a}^{R'^{\vec{u}}} : [0,T] \to \mathbb{R}$ are assumed

to exist, so that the linear combination

$$\vec{u}_{\mathrm{D}}^{R}(t) = \sum_{i=R^{\vec{u}}+1}^{R'^{\vec{u}}} \vec{\phi}_i^{\vec{u}} a^i(t), \qquad t \in [0,T] \tag{4.59}$$

fulfills $\vec{u}_{\mathrm{D}}^{h} - \vec{u}_{\mathrm{D}}^{R} \in H^1_{\vec{u},0,h}$ for any $t \in [0,T]$. It can be shown that if $\vec{u}^R - \vec{u}_{\mathrm{D}}^{R} \in H^1_{\vec{u},0,R}$ for all $t \in [0,T]$, then $\vec{u}^R - \vec{u}_{\mathrm{D}}^{h} \in H^1_{\vec{u},0,h}$ for all $t \in [0,T]$.

The following finite element representations are required for the derivation of the reduced-order model:

$$\phi_i^u = \sum_{m=1}^{M'^u} \Phi_i^{u,m} \psi_m^u, \qquad \Phi_i^u = \begin{pmatrix} \Phi_i^{u,1} \\ \vdots \\ \Phi_i^{u,M'^u} \end{pmatrix}, \qquad i = 1, \ldots, R'^{\vec{u}}, \tag{4.60}$$

$$\phi_i^v = \sum_{m=1}^{M'^v} \Phi_i^{v,m} \psi_m^v, \qquad \Phi_i^v = \begin{pmatrix} \Phi_i^{v,1} \\ \vdots \\ \Phi_i^{v,M'^v} \end{pmatrix}, \qquad i = 1, \ldots, R'^{\vec{u}}, \tag{4.61}$$

$$\vec{\phi}_i^{\vec{u}} = \begin{pmatrix} \phi_i^u \\ \phi_i^v \end{pmatrix}, \qquad \vec{\Phi}_i^{\vec{u}} = \begin{pmatrix} \Phi_i^u \\ \Phi_i^v \end{pmatrix}, \qquad i = 1, \ldots, R'^{\vec{u}}. \tag{4.62}$$

Some of the reduced basis functions are zero at the Dirichlet points. Thus,

$$\phi_i^u = \sum_{m=1}^{M^u} \Phi_i^{u,m} \psi_m^u, \qquad \Phi_i^{u,\mathrm{F}} = \begin{pmatrix} \Phi_i^{u,1} \\ \vdots \\ \Phi_i^{u,M^u} \end{pmatrix}, \qquad i = 1, \ldots, R^{\vec{u}}, \tag{4.63}$$

$$\phi_i^v = \sum_{m=1}^{M^v} \Phi_i^{v,m} \psi_m^v, \qquad \Phi_i^{v,\mathrm{F}} = \begin{pmatrix} \Phi_i^{v,1} \\ \vdots \\ \Phi_i^{v,M^v} \end{pmatrix}, \qquad i = 1, \ldots, R^{\vec{u}}, \tag{4.64}$$

$$\vec{\phi}_i^{\vec{u}} = \begin{pmatrix} \phi_i^u \\ \phi_i^v \end{pmatrix}, \qquad \vec{\Phi}_i^{\vec{u},\mathrm{F}} = \begin{pmatrix} \Phi_i^{u,\mathrm{F}} \\ \Phi_i^{v,\mathrm{F}} \end{pmatrix}, \qquad i = 1, \ldots, R'^{\vec{u}}. \tag{4.65}$$

The weak form (4.12)–(4.16) is the starting point for the derivation. It is

written in vector notation: Find $p \in L^2$ and $\vec{u} - \vec{u}_\mathrm{D} \in H^1_{\vec{u},0}$, such that

$$
\begin{aligned}
&(\vec{\psi^u}, \partial_t \vec{u} + \vec{u} \cdot \nabla \vec{u}) - (\nabla \cdot \vec{\psi^u}, p) \\
&\quad + \frac{1}{Re}(\nabla \vec{\psi^u}, \nabla \vec{u} + (\nabla \vec{u})^T) = 0 \qquad \forall \vec{\psi^u} \in H^1_{\vec{u},0}, \qquad t \in (0,T], \\
&\qquad\qquad (\psi^p, \nabla \cdot \vec{u}) = 0 \qquad \forall \psi^p \in L^2, \qquad t \in (0,T], \\
&\qquad\qquad (\vec{\psi^u}, \vec{u} - \vec{u}_0) = 0 \qquad \forall \vec{\psi^u} \in H^1_{\vec{u},0}, \qquad t \in \{0\}.
\end{aligned}
$$

In the reduced-order model, the initial conditions are approximated using the finite element interpolations u_0^h and v_0^h, see (4.30) and (4.31). The Dirichlet conditions are implemented using $\vec{u}^R_\mathrm{D}$, see (4.59), using the coefficients $\bar{a}^{R^{\vec{u}}+1}, \ldots, \bar{a}^{R'^{\vec{u}}}$ in the reduced-order representation. The space $H^1_{\vec{u},0,R}$ is used as a test space for the momentum equation and the initial condition. The pressure is assumed to be a member of the pressure finite element space $L^2_{p,h}$. A further restriction of the pressure field to a reduced-order subspace is not necessary. Consequently, the weak form becomes: For any $p^h \in L^2_{p,h}$, find $\vec{u}^R - \vec{u}^R_\mathrm{D} \in H^1_{\vec{u},0,R}$, such that

$$
\begin{aligned}
&(\vec{\phi^u}, \partial_t \vec{u}^R + \vec{u}^R \cdot \nabla \vec{u}^R) - (\nabla \cdot \vec{\phi^u}, p^h) \\
&\quad + \frac{1}{Re}(\nabla \vec{\phi^u}, \nabla \vec{u}^R + (\nabla \vec{u}^R)^T) = 0 \qquad \forall \vec{\phi^u} \in H^1_{\vec{u},0,R}, \qquad t \in (0,T], \\
&(\psi^p, \nabla \cdot \vec{u}^R) = 0 \qquad \forall \psi^p \in L^2_{p,h}, \qquad t \in (0,T], \\
&(\vec{\phi^u}, \vec{u}^R - \vec{u}^h_0) = 0 \qquad \forall \vec{\phi^u} \in H^1_{\vec{u},0,R}, \qquad t \in \{0\}.
\end{aligned}
$$

The following assumption allows for significant simplifications of the flow models. Its validity is studied separately for the control function approach and the modified basis function approach.

Assumption 1. *For all $p^h \in L^2_{p,h}$ and $\vec{\phi^u} \in H^1_{\vec{u},R}$ it holds that $(\nabla \cdot \vec{\phi^u}, p^h) = 0$.*

Firstly, the modified basis function approach is considered. In the case of discretely divergence-free snapshots it can be shown that the assumption is fulfilled.

Proposition 2. *If*

- *$\vec{u}^h_1, \ldots, \vec{u}^h_N$ are obtained with the finite element model of subsection 4.2.2,*
- *$H^1_{\vec{u},R} = \mathrm{span}(\vec{\phi}^{\vec{u}}_1, \ldots, \vec{\phi}^{\vec{u}}_{R'^{\vec{u}}})$ with $\vec{\phi}^{\vec{u}}_1, \ldots, \vec{\phi}^{\vec{u}}_{R'^{\vec{u}}}$ obtained by applying the modified basis function method of subsection 2.3.2 to $\vec{u}^h_1, \ldots, \vec{u}^h_N$,*

- $R'^{\vec{u}}$ *is small enough, so that all singular values equal to zero have been excluded,*

then assumption 1 is fulfilled.

Proof. Any function $\vec{\phi}^{\vec{u}} \in H^1_{\vec{u},R}$ can be written as a linear combination of $\vec{\phi}^{\vec{u}}_1, \ldots, \vec{\phi}^{\vec{u}}_{R'^{\vec{u}}}$. Due to linearity it suffices to show that $(\nabla \cdot \vec{\phi}^{\vec{u}}_r, p^h) = 0$ for all $p^h \in L^2_{p,h}$ and $r = 1, \ldots, R'^{\vec{u}}$. It is a property of the modified basis function method that $\vec{\phi}^{\vec{u}}_1, \ldots, \vec{\phi}^{\vec{u}}_{R'^{\vec{u}}}$ are a linear combination of the snapshots $\vec{u}^h_1, \ldots, \vec{u}^h_N$ if no zero singular values have been kept. Due to linearity it is sufficient to show that $(\nabla \cdot \vec{u}^h_n, p^h) = 0$ for all $n = 1, \ldots, N$ and $p^h \in L^2_{p,h}$, which is implied by (4.50) and the conditions on the initial data of the finite element model. □

Secondly, the control function approach is considered, where the additional assumption of discretely divergence-free control functions is necessary.

Proposition 3. *If*

- $\vec{u}^h_1, \ldots, \vec{u}^h_N$ *are obtained with the finite element model of subsection 4.2.2,*
- $H^1_{\vec{u},R} = \operatorname{span}(\vec{\phi}^{\vec{u}}_1, \ldots, \vec{\phi}^{\vec{u}}_{R'^{\vec{u}}})$ *with* $\vec{\phi}^{\vec{u}}_1, \ldots, \vec{\phi}^{\vec{u}}_{R'^{\vec{u}}}$ *obtained by applying the control function method of subsection 2.3.1 to* $\vec{u}^h_1, \ldots, \vec{u}^h_N$,
- $R'^{\vec{u}}$ *is small enough, so that all singular values equal to zero have been excluded,*
- $(\nabla \cdot \vec{\phi}^{\vec{u}}_r, \psi^p) = 0$ *for all* $\psi^p \in L^2_{p,h}$ *and* $r = R^{\vec{u}} + 1, \ldots, R'^{\vec{u}}$,

then assumption 1 is fulfilled.

Proof. Any function $\vec{\phi}^{\vec{u}} \in H^1_{\vec{u},R}$ can be written as a linear combination of $\vec{\phi}^{\vec{u}}_1, \ldots, \vec{\phi}^{\vec{u}}_{R'^{\vec{u}}}$. Due to linearity it suffices to show that $(\nabla \cdot \vec{\phi}^{\vec{u}}_r, p^h) = 0$ for all $p^h \in L^2_{p,h}$ and $r = 1, \ldots, R'^{\vec{u}}$. Because $(\nabla \cdot \vec{\phi}^{\vec{u}}_r, \psi^p) = 0$ was assumed for all $\psi^p \in L^2_{p,h}$ and $r = R^{\vec{u}}+1, \ldots, R'^{\vec{u}}$, it is left to show that $(\nabla \cdot \vec{\phi}^{\vec{u}}_r, p^h) = 0$ for all $p^h \in L^2_{p,h}$ and $r = 1, \ldots, R^{\vec{u}}$. It is a property of the control function method that $\vec{\phi}^{\vec{u}}_1, \ldots, \vec{\phi}^{\vec{u}}_{R^{\vec{u}}}$ are a linear combination of the modified snapshots if no zero singular values have been kept. The modified snapshots themselves are obtained by subtracting linear combinations of $\vec{\phi}^{\vec{u}}_{R^{\vec{u}}+1}, \ldots, \vec{\phi}^{\vec{u}}_{R'^{\vec{u}}}$ from $\vec{u}^h_1, \ldots, \vec{u}^h_N$. Due to linearity it is sufficient to show that $(\nabla \cdot \vec{u}^h_n, , p^h) = 0$ for all $n = 1, \ldots, N$ and $p^h \in L^2_{p,h}$, which is implied by (4.50) and the conditions on the initial data of the finite element model. □

As a consequence of assumption 1, the pressure term and the continuity equation in the weak form drop out. Because of the linear independence of $\vec{\phi}^{\vec{u}}_1, \ldots, \vec{\phi}^{\vec{u}}_{R'^{\vec{u}}}$, it suffices to test against each of the reduced basis functions, individually. Thus, the weak form reduces to the following: Find $\vec{u} - \vec{u}_{\mathrm{D}} \in H^1_{\vec{u},0,R}$, such that

$$
\begin{aligned}
&(\vec{\phi}^{\vec{u}}_r, \partial_t \vec{u}^R + \vec{u}^R \cdot \nabla \vec{u}^R) \\
&\qquad + \frac{1}{Re}(\nabla \vec{\phi}^{\vec{u}}_r, \nabla \vec{u}^R + (\nabla \vec{u}^R)^T) = 0, \qquad r = 1, \ldots, R^{\vec{u}}, \qquad t \in (0, T], \\
&(\vec{\phi}^{\vec{u}}_r, \vec{u}^R - \vec{u}^h_0) = 0, \qquad\qquad\qquad\qquad\quad r = 1, \ldots, R^{\vec{u}}, \qquad t \in \{0\}.
\end{aligned}
$$

After substitution of the reduced-order representation (4.58) and elementary reformulations, the resulting equations are given by

$$
\begin{aligned}
&\sum_{i=1}^{R'^{\vec{u}}} (\vec{\phi}^{\vec{u}}_r, \vec{\phi}^{\vec{u}}_i) \dot{a}^i + \sum_{i,j=1}^{R'^{\vec{u}}} (\vec{\phi}^{\vec{u}}_r, \vec{\phi}^{\vec{u}}_i \cdot \nabla \vec{\phi}^{\vec{u}}_j) a^i a^j \\
&\qquad + \frac{1}{Re} \sum_{i=1}^{R'^{\vec{u}}} (\nabla \vec{\phi}^{\vec{u}}_r, \nabla \vec{\phi}^{\vec{u}}_i + (\nabla \vec{\phi}^{\vec{u}}_i)^T) a^i = 0, \quad r = 1, \ldots, R^{\vec{u}}, \quad t \in (0, T], \\
&\sum_{i=1}^{R'^{\vec{u}}} (\vec{\phi}^{\vec{u}}_r, \vec{\phi}^{\vec{u}}_i) a^i - (\vec{\phi}^{\vec{u}}_r, \vec{u}^h_0) = 0, \qquad\qquad r = 1, \ldots, R^{\vec{u}}, \quad t \in \{0\}.
\end{aligned}
$$

The evaluation of the inner products is carried out term by term in the appendix. They can be written using the matrices of the spatial finite element semi-discretization. Therefore, the velocity reduced-order model amounts to solving the following problem: For given $a^{R^{\vec{u}}+1}, \ldots, a^{R'^{\vec{u}}}$, find $a^1, \ldots, a^{R^{\vec{u}}}$ which satisfy

$$
\begin{aligned}
&\sum_{i=1}^{R'^{\vec{u}}} (\vec{\Phi}^{\vec{u},\mathrm{F}}_r)^T \mathcal{M}^{\mathrm{F}}_{\vec{u}} \vec{\Phi}^{\vec{u}}_i \dot{a}^i \\
&\qquad + \sum_{i,j=1}^{R'^{\vec{u}}} (\vec{\Phi}^{\vec{u},\mathrm{F}}_r)^T \mathcal{N}^{\mathrm{F}}_{\vec{u}} (\vec{\Phi}^{\vec{u}}_i) \vec{\Phi}^{\vec{u}}_j a^i a^j \\
&\qquad + \sum_{i=1}^{R'^{\vec{u}}} (\vec{\Phi}^{\vec{u},\mathrm{F}}_r)^T \mathcal{K}^{\mathrm{F}}_{\vec{u}} \, \vec{\Phi}^{\vec{u}}_i a^i = 0, \qquad r = 1, \ldots, R^{\vec{u}}, \quad t \in (0, T], \qquad (4.66)
\end{aligned}
$$

$$\sum_{i=1}^{R'^{\vec{u}}} (\vec{\Phi}_r^{\vec{u},\mathrm{F}})^T \mathcal{M}_{\vec{u}}^{\mathrm{F}} \vec{\Phi}_i^{\vec{u}} a^i - (\vec{\Phi}_r^{\vec{u},\mathrm{F}})^T \mathcal{M}_{\vec{u}}^{\mathrm{F}} \vec{U}_0 = 0 \qquad r = 1, \ldots, R^{\vec{u}}. \quad t \in \{0\} \tag{4.67}$$

Before continuing with the time discretization of this set of equations in subsection 4.3.4, an alternative derivation of the velocity reduced-order model is presented in subsection 4.3.2 and a pressure reduced-order model is introduced in subsection 4.3.3.

4.3.2 Velocity model via space discretization

In the following, the reduced-order model for the velocity is derived using the spatially semi-discretized equations of the flow problem. To this end, the reduced-order approximation of the velocity finite element coefficient vector is defined as

$$\vec{U}(t) \approx \vec{U}^R(t) = \sum_{i=1}^{R'^{\vec{u}}} \vec{\Phi}_i^{\vec{u}} a^i(t), \qquad t \in [0, T], \tag{4.68}$$

where $\vec{\Phi}_1^{\vec{u}}, \ldots, \vec{\Phi}_{R'^q}^{\vec{u}} \in \mathbb{R}^{M'^q}$ are determined like in the previous section and $a^1, \ldots, a^{R'^q} : [0, T] \to \mathbb{R}$ are arbitrary. The reduced-order approximation $\vec{U}^R$ is substituted in the finite element semi-discretization of the weak form of the incompressible Navier-Stokes problem, (4.40)–(4.42). This gives rise to the residuals

$$\vec{\mathcal{R}}^{\vec{u}} = \mathcal{M}_{\vec{u}}^{\mathrm{F}} \dot{\vec{U}}^R + \mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{U}^R)\vec{U}^R + \mathcal{C}_{\vec{u}}^{\mathrm{F}} P + \mathcal{K}_{\vec{u}}^{\mathrm{F}} \vec{U}^R, \qquad t \in (0, T], \tag{4.69}$$

$$\mathcal{R}^p = \mathcal{D}_{\vec{u}}^{\mathrm{F}} \vec{U}^R, \qquad t \in (0, T], \tag{4.70}$$

$$\vec{\mathcal{R}}_0^{\vec{u}} = \mathcal{M}_{\vec{u}}^{\mathrm{F}} \vec{U}^R - \mathcal{M}_{\vec{u}}^{\mathrm{F}} \vec{U}_0, \qquad t \in \{0\}. \tag{4.71}$$

The case of a non-empty Neumann boundary is assumed, which leads to a definite pressure. Recalling (4.39), it holds that $\mathcal{D}_{\vec{u}}^{\mathrm{F}} = (\mathcal{C}_{\vec{u}}^{\mathrm{F}})^T$.

The discrete equivalents to assumption 1 and propositions 2 and 3 are given as follows:

Assumption 2. *For $\vec{U}^R$ and $\vec{\Phi}_1^{\vec{u}}, \ldots, \vec{\Phi}_{R'^q}^{\vec{u}}$ according to (4.68) it holds that $\mathcal{D}_{\vec{u}}^F \vec{U}^R = 0$ and $(\vec{\Phi}_r^{\vec{u}})^T \mathcal{C}_{\vec{u}}^F P = 0$ for any $P \in \mathbb{R}^{M'^q}$ and $r = 1, \ldots, R^q$.*

Proposition 4. *If*

- *$\vec{U}_1, \ldots, \vec{U}_N$ are obtained with the finite element model of subsection 4.2.2,*

- $\vec{\Phi}_1^{\vec{u}}, \ldots, \vec{\Phi}_{R'^q}^{\vec{u}}$ *are obtained by applying the modified basis function method of subsection 2.3.2,*

- $R'^{\vec{u}}$ *is small enough so that all singular values equal to zero have been excluded,*

then assumption 2 is fulfilled.

Proof. The vector $\vec{U}^R$ is a linear combination of $\vec{\Phi}_1^{\vec{u}}, \ldots, \vec{\Phi}_{R'^{\vec{u}}}^{\vec{u}}$. Due to linearity, to show $\mathcal{D}_{\vec{u}}^{\mathrm{F}} \vec{U}^R = 0$ it suffices to show that $\mathcal{D}_{\vec{u}}^{\mathrm{F}} \vec{\Phi}_r^{\vec{u}} = 0$ for all $r = 1, \ldots, R'^{\vec{u}}$. Moreover, to show $(\vec{\Phi}_r^{\vec{u}})^T \mathcal{C}_{\vec{u}}^{\mathrm{F}} P = (\vec{\Phi}_r^{\vec{u}})^T (\mathcal{D}_{\vec{u}}^{\mathrm{F}})^T P = (\mathcal{D}_{\vec{u}}^{\mathrm{F}} \vec{\Phi}_r^{\vec{u}})^T P = 0$ for any $P \in \mathbb{R}^{M'^q}$ and $r = 1, \ldots, R^q$, it also suffices to show $\mathcal{D}_{\vec{u}}^{\mathrm{F}} \vec{\Phi}_r^{\vec{u}} = 0$ for all $r = 1, \ldots, R'^{\vec{u}}$. It is a property of the modified basis function method that $\vec{\Phi}_1^{\vec{u}}, \ldots, \vec{\Phi}_{R'^{\vec{u}}}^{\vec{u}}$ are a linear combination of the snapshots $\vec{U}_1, \ldots, \vec{U}_N$ if no zero singular values have been kept. Due to linearity it is sufficient to show that $\mathcal{D}_{\vec{u}}^{\mathrm{F}} \vec{U}_n = 0$ for all $n = 1, \ldots, N$, which is implied by (4.50) and the conditions on the initial data in the finite element model. □

Proposition 5. *If*

- $\vec{U}_1, \ldots, \vec{U}_N$ *are obtained with the finite element model of subsection 4.2.2,*

- $\vec{\Phi}_1^{\vec{u}}, \ldots, \vec{\Phi}_{R'^q}^{\vec{u}}$ *are obtained with the control function method of subsection 2.3.1,*

- $R'^{\vec{u}}$ *is small enough so that all singular values equal to zero have been excluded,*

- $\mathcal{D}_{\vec{u}}^{F} \vec{\Phi}_r^{\vec{u}} = 0$ *for all* $r = R^{\vec{u}} + 1, \ldots, R'^{\vec{u}}$,

then assumption 2 is fulfilled.

Proof. The vector $\vec{U}^R$ is a linear combination of $\vec{\Phi}_1^{\vec{u}}, \ldots, \vec{\Phi}_{R'^{\vec{u}}}^{\vec{u}}$. Due to linearity, to show $\mathcal{D}_{\vec{u}}^{\mathrm{F}} \vec{U}^R = 0$ it suffices to show that $\mathcal{D}_{\vec{u}}^{\mathrm{F}} \vec{\Phi}_r^{\vec{u}} = 0$ for all $r = 1, \ldots, R'^{\vec{u}}$. Moreover, to show $(\vec{\Phi}_r^{\vec{u}})^T \mathcal{C}_{\vec{u}}^{\mathrm{F}} P = (\vec{\Phi}_r^{\vec{u}})^T (\mathcal{D}_{\vec{u}}^{\mathrm{F}})^T P = (\mathcal{D}_{\vec{u}}^{\mathrm{F}} \vec{\Phi}_r^{\vec{u}})^T P = 0$ for any $P \in \mathbb{R}^{M'^q}$ and $r = 1, \ldots, R^q$, it also suffices to show $\mathcal{D}_{\vec{u}}^{\mathrm{F}} \vec{\Phi}_r^{\vec{u}} = 0$ for all $r = 1, \ldots, R'^{\vec{u}}$. It is a property of the control function method that $\vec{\Phi}_1^{\vec{u}}, \ldots, \vec{\Phi}_{R^q}^{\vec{u}}$ are a linear combination of the modified snapshots if no zero singular values have been kept. The modified snapshots themselves are obtained by subtracting linear combinations of $\vec{\Phi}_{R^{\vec{u}}+1}^{\vec{u}}, \ldots, \vec{\Phi}_{R'^{\vec{u}}}^{\vec{u}}$ from $\vec{U}_1, \ldots, \vec{U}_N$. Due to linearity it is sufficient to show that $\mathcal{D}_{\vec{u}}^{\mathrm{F}} \vec{U}_n = 0$ for all $n = 1, \ldots, N$, which is implied by (4.50) and the conditions on the initial data in the finite element model. □

Equation (4.70) can be omitted from the further considerations because $\mathcal{R}^p = 0$ due to assumption 2. By expanding $\vec{U}^R$ in (4.69) and (4.71) via (4.68) and by using the linear dependence of $\mathcal{N}^{\mathrm{F}}_{\vec{u}}(\vec{U}^R)$ on its argument (see subsection 4.2.1), one obtains

$$\vec{\mathcal{R}}^{\vec{u}} = \sum_{i=1}^{R'^{\vec{u}}} \mathcal{M}^{\mathrm{F}}_{\vec{u}} \vec{\Phi}^{\vec{u}}_i \dot{a}^i + \sum_{i,j=1}^{R'^{\vec{u}}} \mathcal{N}^{\mathrm{F}}_{\vec{u}}(\vec{\Phi}^{\vec{u}}_i) \vec{\Phi}^{\vec{u}}_j a^i a^j + \mathcal{C}^{\mathrm{F}}_{\vec{u}}\, P + \sum_{i=1}^{R'^{\vec{u}}} \mathcal{K}^{\mathrm{F}}_{\vec{u}}\, \vec{\Phi}^{\vec{u}}_i a^i, \quad t \in (0,T],$$

$$\vec{\mathcal{R}}^{\vec{u}}_0 = \sum_{i=1}^{R'^{\vec{u}}} \mathcal{M}^{\mathrm{F}}_{\vec{u}} \vec{\Phi}^{\vec{u}}_i a^i - \mathcal{M}^{\mathrm{F}}_{\vec{u}} \vec{U}_0 \qquad t \in \{0\}.$$

As the second step of the Galerkin procedure, it is demanded that the residuals are orthogonal to the non-Dirichlet parts of the velocity reduced basis functions,

$$(\vec{\Phi}^{\vec{u},\mathrm{F}}_r)^T \vec{\mathcal{R}}^{\vec{u}} = 0, \qquad r = 1,\dots,R^{\vec{u}}, \qquad t \in (0,T],$$
$$(\vec{\Phi}^{\vec{u},\mathrm{F}}_r)^T \vec{\mathcal{R}}^{\vec{u}}_0 = 0, \qquad r = 1,\dots,R^{\vec{u}}, \qquad t \in \{0\},$$

which results in the reduced-order momentum equation

$$\sum_{i=1}^{R'^{\vec{u}}} (\vec{\Phi}^{\vec{u},\mathrm{F}}_r)^T \mathcal{M}^{\mathrm{F}}_{\vec{u}} \vec{\Phi}^{\vec{u}}_i \dot{a}^i + \sum_{i,j=1}^{R'^{\vec{u}}} (\vec{\Phi}^{\vec{u},\mathrm{F}}_r)^T \mathcal{N}^{\mathrm{F}}_{\vec{u}}(\vec{\Phi}^{\vec{u}}_i) \vec{\Phi}^{\vec{u}}_j a^i a^j + (\vec{\Phi}^{\vec{u},\mathrm{F}}_r)^T \mathcal{C}^{\mathrm{F}}_{\vec{u}}\, P + \sum_{i=1}^{R'^{\vec{u}}} (\vec{\Phi}^{\vec{u},\mathrm{F}}_r)^T \mathcal{K}^{\mathrm{F}}_{\vec{u}}\, \vec{\Phi}^{\vec{u}}_i a^i = 0$$

for $r = 1,\dots,R^{\vec{u}}$ and $t \in (0,T]$, and the initial condition

$$\sum_{i=1}^{R'^{\vec{u}}} (\vec{\Phi}^{\vec{u},\mathrm{F}}_r)^T \mathcal{M}^{\mathrm{F}}_{\vec{u}} \vec{\Phi}^{\vec{u}}_i a^i - \sum_{i=1}^{R'^{\vec{u}}} (\vec{\Phi}^{\vec{u},\mathrm{F}}_r)^T \mathcal{M}^{\mathrm{F}}_{\vec{u}} \vec{U}_0 = 0$$

for $r = 1,\dots,R^{\vec{u}}$ and $t \in \{0\}$. By assumption 2, the pressure term can be eliminated. As a consequence, (4.66) and (4.67) are recovered.

4.3.3 Pressure model via space discretization

The pressure has been eliminated from the velocity model. For the computation of the drag and lift coefficients, however, the pressure is necessary. In the following, a reduced-order model is described which can be used to compute a pressure field from the solution of a reduced-order velocity model. The

method is based on the method of Akhtar et al. (2009), who substituted the reduced-order approximation of the velocity field in the right-hand side of a continuous pressure Poisson equation and projected the resulting equations on a set of pressure POD basis functions using Galerkin's method. Here, the method of Akhtar et al. is extended to the case of a discrete pressure Poisson equation. Note that here 'continuous' and 'discrete' relate to the spatial dimensions of the problem and not to time. For a discussion about continuous and discrete pressure Poisson equations, see Gresho and Sani (2000).

A POD of the pressure snapshots vectors $P_1, \ldots, P_N \in \mathbb{R}^{M'^p}$ provides pressure reduced basis vectors $\Phi_1^p, \ldots, \Phi_{R'^p}^p \in \mathbb{R}^{M'^p}$. In the absence of inhomogeneous Dirichlet conditions it is not necessary to employ the control function method or the modified basis function method, so $R'^p = R^p$. Still, it is possible to use, for example, the snapshot mean as a control function, which leads to a POD of the pressure fluctuations and $R'^p = R^p + 1$. In any case, a reduced-order representation P^R of the pressure finite element coefficient vector is given by

$$P(t) \approx P^R(t) = \sum_{i=1}^{R'^p} \Phi_i^p c^i(t), \qquad t \in [0,T]. \tag{4.72}$$

As the first step of the Galerkin procedure, the substitution of the reduced-order approximations of the velocity and the pressure into the discrete pressure Poisson equation (4.45) gives rise to the residuum

$$\mathcal{R}^{\mathrm{PP}} = \mathcal{D}_{\vec{u}}^{\mathrm{FD}} \dot{\vec{U}}^{R,\mathrm{D}} - \mathcal{D}_{\vec{u}}^{\mathrm{FF}} (\mathcal{M}_{\vec{u}}^{\mathrm{FF}})^{-1} \Big(\mathcal{M}_{\vec{u}}^{\mathrm{FD}} \dot{\vec{U}}^{R,\mathrm{D}} + \mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{U}^R)\vec{U}^R + \mathcal{K}_{\vec{u}}^{\mathrm{F}} \vec{U}^R + \mathcal{C}_{\vec{u}}^{\mathrm{F}} P^R \Big), \quad t \in [0,T].$$

As the second step, the residuum is required to be orthogonal to the free parts of the pressure reduced basis functions,

$$(\Phi_r^{p,\mathrm{F}})^T \mathcal{R}^{\mathrm{PP}} = 0, \qquad r = 1, \ldots, R^p, \qquad t \in [0,T],$$

which results in

$$(\Phi_r^{p,\mathrm{F}})^T \mathcal{D}_{\vec{u}}^{\mathrm{FD}} \dot{\vec{U}}^{R,\mathrm{D}} - (\Phi_r^{p,\mathrm{F}})^T \mathcal{D}_{\vec{u}}^{\mathrm{FF}} (\mathcal{M}_{\vec{u}}^{\mathrm{FF}})^{-1} \Big(\mathcal{M}_{\vec{u}}^{\mathrm{FD}} \dot{\vec{U}}^{R,\mathrm{D}} + \mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{U}^R)\vec{U}^R + \mathcal{K}_{\vec{u}}^{\mathrm{F}} \vec{U}^R + \mathcal{C}_{\vec{u}}^{\mathrm{F}} P^R \Big) = 0, \quad t \in [0,T]$$

for $r = 1, \ldots, R^p$. After expanding the expressions for the reduced-order pressure and velocity, and using $\vec{\Phi}_i^{\vec{u},\mathrm{D}} = 0$ for $i = 1, \ldots, R^{\vec{u}}$, one obtains the

following reduced-order model for the pressure: For given $a^1, \ldots, a^{R'^{\vec{u}}}$ and $c^{R^p+1}, \ldots, c^{R'^p}$, find $c^1, \ldots, c^{R^p}$ such that

$$\sum_{i=R^{\vec{u}}+1}^{R'^{\vec{u}}} (\Phi_r^{p,\mathrm{F}})^T \mathcal{D}_{\vec{u}}^{\mathrm{FD}} \vec{\Phi}_i^{\vec{u},\mathrm{D}} \dot{a}^i - (\Phi_r^{p,\mathrm{F}})^T \mathcal{D}_{\vec{u}}^{\mathrm{FF}} (\mathcal{M}_{\vec{u}}^{\mathrm{FF}})^{-1} \Bigg(\sum_{i=R^{\vec{u}}+1}^{R'^{\vec{u}}} \mathcal{M}_{\vec{u}}^{\mathrm{FD}} \vec{\Phi}_i^{\vec{u},\mathrm{D}} \dot{a}^i + \sum_{i,j=1}^{R'^{\vec{u}}} \mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{\Phi}_i^{\vec{u}}) \vec{\Phi}_j^{\vec{u}} a^i a^j + \sum_{i=1}^{R'^{\vec{u}}} \mathcal{K}_{\vec{u}}^{\mathrm{F}} \vec{\Phi}_i^{\vec{u}} a^i + \sum_{i=1}^{R'^p} \mathcal{C}_{\vec{u}}^{\mathrm{F}} \Phi_i^p c^i \Bigg) = 0 \tag{4.73}$$

for $r = 1, \ldots, R^p$ and $t \in [0, T]$.

4.3.4 Finite difference discretization in time

The model given by (4.66), (4.73) and (4.67) can be written in the following way:

- Velocity model for $r = 1, \ldots, R^{\vec{u}}$ and $t \in (0, T]$:

$$\sum_{i=1}^{R'^{\vec{u}}} \mathcal{M}_{ri}^{a,\mathrm{F}} \dot{a}^i + \sum_{i,j=1}^{R'^{\vec{u}}} \mathcal{N}_{rij}^{a,\mathrm{F}} a^i a^j + \sum_{i=1}^{R'^{\vec{u}}} \mathcal{K}_{ri}^{a,\mathrm{F}} a^i = 0, \tag{4.74}$$

- Pressure model for $r = 1, \ldots, R^p$ and $t \in [0, T]$:

$$\sum_{i=1}^{R'^p} \mathcal{C}_{ri}^{c,\mathrm{F}} c^i + \sum_{i=R^{\vec{u}}+1}^{R'^{\vec{u}}} \mathcal{M}_{ri}^{c,\mathrm{F}} \dot{a}^i + \sum_{i,j=1}^{R'^{\vec{u}}} \mathcal{N}_{rij}^{c,\mathrm{F}} a^i a^j + \sum_{i=1}^{R'^{\vec{u}}} \mathcal{K}_{ri}^{c,\mathrm{F}} a^i = 0, \tag{4.75}$$

- Initial condition for $r = 1, \ldots, R^{\vec{u}}$ and $t \in \{0\}$:

$$\sum_{i=1}^{R'^{\vec{u}}} \mathcal{M}_{ri}^{a,\mathrm{F}} a^i - (\vec{\Phi}_r^{\vec{u},\mathrm{F}})^T \mathcal{M}_{\vec{u}}^{\mathrm{F}} \vec{U}_0 = 0. \tag{4.76}$$

The coefficients of the system of equations are

$$
\begin{aligned}
\mathcal{M}^{a,\mathrm{F}}_{ri} &= (\vec{\Phi}^{\vec{u},\mathrm{F}}_r)^T \mathcal{M}^{\mathrm{F}}_{\vec{u}} \vec{\Phi}^{\vec{u}}_i, \\
\mathcal{N}^{a,\mathrm{F}}_{rij} &= (\vec{\Phi}^{\vec{u}}_r)^T \mathcal{N}^{\mathrm{F}}_{\vec{u}} (\vec{\Phi}^{\vec{u}}_i) \vec{\Phi}^{\vec{u}}_j, \\
\mathcal{K}^{a,\mathrm{F}}_{ri} &= (\vec{\Phi}^{\vec{u}}_r)^T \mathcal{K}^{\mathrm{F}}_{\vec{u}}\ \vec{\Phi}^{\vec{u}}_i, \\
\mathcal{M}^{c,\mathrm{F}}_{ri} &= -(\Phi^{p,\mathrm{F}}_r)^T \mathcal{D}^{\mathrm{FF}}_{\vec{u}} (\mathcal{M}^{\mathrm{FF}}_{\vec{u}})^{-1} \mathcal{M}^{\mathrm{FD}}_{\vec{u}} \vec{\Phi}^{\vec{u},\mathrm{D}}_i + (\Phi^{p,\mathrm{F}}_r)^T \mathcal{D}^{\mathrm{FD}}_{\vec{u}} \vec{\Phi}^{\vec{u},\mathrm{D}}_i, \\
\mathcal{C}^{c,\mathrm{F}}_{ri} &= -(\Phi^{p,\mathrm{F}}_r)^T \mathcal{D}^{\mathrm{FF}}_{\vec{u}} (\mathcal{M}^{\mathrm{FF}}_{\vec{u}})^{-1} \mathcal{C}^{\mathrm{F}}_{\vec{u}}\ \Phi^p_i, \\
\mathcal{N}^{c,\mathrm{F}}_{rij} &= -(\Phi^{p,\mathrm{F}}_r)^T \mathcal{D}^{\mathrm{FF}}_{\vec{u}} (\mathcal{M}^{\mathrm{FF}}_{\vec{u}})^{-1} \mathcal{N}^{\mathrm{F}}_{\vec{u}} (\vec{\Phi}^{\vec{u}}_i) \vec{\Phi}^{\vec{u}}_j, \\
\mathcal{K}^{c,\mathrm{F}}_{ri} &= -(\Phi^{p,\mathrm{F}}_r)^T \mathcal{D}^{\mathrm{FF}}_{\vec{u}} (\mathcal{M}^{\mathrm{FF}}_{\vec{u}})^{-1} \mathcal{K}^{\mathrm{F}}_{\vec{u}}\ \vec{\Phi}^{\vec{u}}_i
\end{aligned}
$$

and $a^{R^{\vec{u}}+1}, \ldots, a^{R'^{\vec{u}}}$ as well as $c^{R^p+1}, \ldots, c^{R'^p}$ are given as input data. Note that $(\Phi^{p,\mathrm{F}}_r)^T (\mathcal{C}^{\mathrm{FF}}_{\vec{u}})^T (\mathcal{M}^{\mathrm{FF}}_{\vec{u}})^{-1}$ for $r = 1, \ldots, R^p$ can be computed efficiently by solving a linear algebraic system of equations with R^p different right-hand sides.

The problem (4.74)–(4.76) is now discretized in time with the Crank-Nicolson method using the time grid of the snapshot simulation. The unknown time-discrete reduced-order coefficients for $n = 1, \ldots, N$ are denoted by $a^r_n \approx a^r(t_n)$ for all $r = 1, \ldots, R^{\vec{u}}$ and $c^r_n \approx c^r(t_n)$ for all $r = 1, \ldots, R^p$. Similarly, $a^r_n = a^r(t_n)$ for all $r = R^{\vec{u}} + 1, \ldots, R'^{\vec{u}}$ as well as $c^r_n = c^r(t_n)$ for all $r = R^p + 1, \ldots, R'^p$ are used to denote the given coefficients for $n = 1, \ldots, N$. The Dirichlet conditions are exactly fulfilled at any discrete time instance $t_1, \ldots, t_N$ via (4.59).

The time derivatives in (4.74) and (4.75) are approximated by central finite differences, and the other terms are interpreted as an average of evaluations at the old time instance t_{n-1} and the new time instance t_n. This amounts to the following coupled scheme:

- Discrete momentum equation for $r = 1, \ldots, R^{\vec{u}}$ and $n = 2, \ldots, N$:

$$
\begin{aligned}
\sum_{i=1}^{R'^{\vec{u}}} \mathcal{M}^{a,\mathrm{F}}_{ri} \frac{a^i_n - a^i_{n-1}}{t_n - t_{n-1}} + \sum_{i,j=1}^{R'^{\vec{u}}} \mathcal{N}^{a,\mathrm{F}}_{rij} \frac{a^i_n a^j_n + a^i_{n-1} a^j_{n-1}}{2} \\
+ \sum_{i=1}^{R'^{\vec{u}}} \mathcal{K}^{a,\mathrm{F}}_{ri} \frac{a^i_n + a^i_{n-1}}{2} = 0,
\end{aligned}
\tag{4.77}
$$

- Discrete pressure Poisson equation for $r = 1, \ldots, R^p$ and $n = 2, \ldots, N$:

$$\sum_{i=1}^{R'^p} \mathcal{C}_{ri}^{c,\mathrm{F}} \frac{c_n^i + c_{n-1}^i}{2} + \sum_{i=1}^{R'^{\vec{u}}} \mathcal{M}_{ri}^{c,\mathrm{F}} \frac{a_n^i - a_{n-1}^i}{t_n - t_{n-1}} + \sum_{i,j=1}^{R'^{\vec{u}}} \mathcal{N}_{rij}^{c,\mathrm{F}} \frac{a_n^i a_n^j + a_{n-1}^i a_{n-1}^j}{2} + \sum_{i=1}^{R'^{\vec{u}}} \mathcal{K}_{ri}^{c,\mathrm{F}} \frac{a_n^i + a_{n-1}^i}{2} = 0, \tag{4.78}$$

- Discrete velocity initial condition for $r = 1, \ldots, R^{\vec{u}}$ and $n = 1$:

$$\sum_{i=1}^{R'^{\vec{u}}} \mathcal{M}_{ri}^{a,\mathrm{F}} a_n^i - (\vec{\Phi}_r^{\vec{u},\mathrm{F}})^T \mathcal{M}_{\vec{u}}^{\mathrm{F}} \vec{U}_0 = 0, \tag{4.79}$$

- Discrete pressure initial condition for $r = 1, \ldots, R^p$ and $n = 1$:

$$\sum_{i=1}^{R'^p} \mathcal{C}_{ri}^{c,\mathrm{F}} c_n^i + \sum_{i=R^{\vec{u}}+1}^{R'^{\vec{u}}} \mathcal{M}_{ri}^{c,\mathrm{F}} \dot{a}^i(t_n) + \sum_{i,j=1}^{R'^{\vec{u}}} \mathcal{N}_{rij}^{c,\mathrm{F}} a_n^i a_n^j + \sum_{i=1}^{R'^{\vec{u}}} \mathcal{K}_{ri}^{c,\mathrm{F}} a_n^i = 0, \tag{4.80}$$

The coefficients $a_n^{R^{\vec{u}}+1}, \ldots, a_n^{R'^{\vec{u}}}$ and $c_n^{R^p+1}, \ldots, c_n^{R'^p}$ for $n = 1, \ldots, N$ as well as the time derivatives $\dot{a}^{R^{\vec{u}}+1}(t_1), \ldots, \dot{a}^{R'^{\vec{u}}}(t_1)$ are given as input data.

Now the system of equations (4.77) – (4.80) can be solved step by step. It is possible to compute only the velocity via (4.77) and (4.79), if a pressure is not needed. The time-stepping equations are non-linear. Like in case of the finite element element computation, a standard or simplified Newton method in connection with an LU decomposition in each linear sub-step is sufficient to solve the systems of equations.

4.3.5 Velocity model via space-time discretization

To derive a time-discretized reduced-order model from the fully discretized finite element model, a time-discrete reduced-order representation of the velocity vector is introduced:

$$\vec{U}(t_n) \approx \vec{U}_n^R = \sum_{i=1}^{R'^{\vec{u}}} \vec{\Phi}_i^{\vec{u}} a_n^i, \qquad n = 1, \ldots, N. \tag{4.81}$$

Substitution in the fully discretized equations (4.49), (4.50) and (4.51) leads to the residuals

$$\begin{aligned}
\vec{\mathcal{R}}_n^{\vec{u}} &= \mathcal{M}_{\vec{u}}^{\mathrm{F}}(\vec{U}_n^R - \vec{U}_{n-1}^R) \\
&\quad + \frac{\Delta t}{2}(\mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{U}_n^R)\vec{U}_n^R + \mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{U}_{n-1}^R)\vec{U}_{n-1}^R) \\
&\quad + \frac{\Delta t}{2}\mathcal{K}_{\vec{u}}^{\mathrm{F}}\,(\vec{U}_n^R + \vec{U}_{n-1}^R) \\
&\quad + \frac{\Delta t}{2}\mathcal{C}_{\vec{u}}^{\mathrm{F}}\,(P_n + P_{n-1}), \qquad\qquad n = 2, \ldots, N, \\
\mathcal{R}_n^p &= \frac{\Delta t}{2}\mathcal{D}_{\vec{u}}^{\mathrm{F}}\,(\vec{U}_n^R + \vec{U}_{n-1}^R), \qquad\qquad n = 2, \ldots, N. \\
\vec{\mathcal{R}}_n^{\vec{u}} &= \mathcal{M}_{\vec{u}}^{\mathrm{F}}\vec{U}_n^R - \mathcal{M}_{\vec{u}}^{\mathrm{F}}\vec{U}_0, \qquad\qquad n = 1.
\end{aligned}$$

The residuals of the discrete continuity equation, $\mathcal{R}_2^p, \ldots, \mathcal{R}_N^p$, are equal to the zero vector because $\mathcal{D}_{\vec{u}}^{\mathrm{F}}\,\vec{U}_n^R = 0$ for $n = 1, \ldots, N$, which can be shown in the same way as presented for assumption 2 in subsection 4.3.2. By requiring the residuals of the momentum equation and initial condition, $\vec{\mathcal{R}}_1^{\vec{u}}, \ldots, \vec{\mathcal{R}}_N^{\vec{u}}$, to be orthogonal to the free parts of the velocity reduced basis functions, $\vec{\Phi}_1^{\vec{u},\mathrm{F}}, \ldots, \vec{\Phi}_{R^{\vec{u}}}^{\vec{u},\mathrm{F}}$, one obtains

$$\begin{aligned}
&(\vec{\Phi}_r^{\vec{u},\mathrm{F}})^T\mathcal{M}_{\vec{u}}^{\mathrm{F}}(\vec{U}_n^R - \vec{U}_{n-1}^R) \\
&\quad + \frac{\Delta t}{2}(\vec{\Phi}_r^{\vec{u},\mathrm{F}})^T(\mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{U}_n^R)\vec{U}_n^R + \mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{U}_{n-1}^R)\vec{U}_{n-1}^R) \\
&\quad + \frac{\Delta t}{2}(\vec{\Phi}_r^{\vec{u},\mathrm{F}})^T\mathcal{K}_{\vec{u}}^{\mathrm{F}}\,(\vec{U}_n^R + \vec{U}_{n-1}^R) \\
&\quad + \frac{\Delta t}{2}(\vec{\Phi}_r^{\vec{u},\mathrm{F}})^T\mathcal{C}_{\vec{u}}^{\mathrm{F}}\,(P_n + P_{n-1}) = 0, \qquad n = 2, \ldots, N, \qquad r = 1, \ldots, R^{\vec{u}}, \\
&(\vec{\Phi}_r^{\vec{u},\mathrm{F}})^T\mathcal{M}_{\vec{u}}^{\mathrm{F}}(\vec{U}_n - \vec{U}_0) = 0, \qquad n = 1, \qquad r = 1, \ldots, R^{\vec{u}}.
\end{aligned}$$

For the same reasons as in subsection 4.3.2 the pressure term drops out, which also justifies that the pressure initial condition (4.52) has been disregarded. After substitution of (4.81) and basic reformulations, (4.77) and (4.79) are obtained.

4.3.6 Pressure model via space-time discretization

To derive a reduced-order model for the pressure from the fully discretized finite element model, at first a time-discretized pressure Poisson equation is derived. Then the time-discretized pressure Poisson equation is projected on the pressure reduced basis vectors.

In the discrete momentum equation (4.49), the term stemming from the time derivative is split in a Dirichlet and a non-Dirichlet part,

$$\mathcal{M}_{\vec{u}}^{\mathrm{FF}} \frac{\vec{U}_n^{\mathrm{F}} - \vec{U}_{n-1}^{\mathrm{F}}}{t_n - t_{n-1}} = -\mathcal{M}_{\vec{u}}^{\mathrm{FD}} \frac{\vec{U}_n^{\mathrm{D}} - \vec{U}_{n-1}^{\mathrm{D}}}{t_n - t_{n-1}} - \frac{\mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{U}_n)\vec{U}_n + \mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{U}_{n-1})\vec{U}_{n-1}}{2} - \mathcal{K}_{\vec{u}}^{\mathrm{F}} \frac{\vec{U}_n + \vec{U}_{n-1}}{2} - \mathcal{C}_{\vec{u}}^{\mathrm{F}} \frac{P_n + P_{n-1}}{2}$$

for $n = 2, \ldots, N$, where

$$\vec{U}_n^{\mathrm{F}} = (U_n^1, \ldots, U_n^{M^u}, V_n^1, \ldots, V_n^{M^v})^T, \qquad n = 1, \ldots, N,$$
$$\vec{U}_n^{\mathrm{D}} = (U_n^{M^u+1}, \ldots, U_n^{M'^u}, V_n^{M^u+1}, \ldots, V_n^{M'^v})^T, \qquad n = 1, \ldots, N.$$

After multiplying with ${\mathcal{M}_{\vec{u}}^{\mathrm{FF}}}^{-1}$,

$$\frac{\vec{U}_n^{\mathrm{F}} - \vec{U}_{n-1}^{\mathrm{F}}}{t_n - t_{n-1}} = -(\mathcal{M}_{\vec{u}}^{\mathrm{FF}})^{-1} \left(\mathcal{M}_{\vec{u}}^{\mathrm{FD}} \frac{\vec{U}_n^{\mathrm{D}} - \vec{U}_{n-1}^{\mathrm{D}}}{t_n - t_{n-1}} + \frac{\mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{U}_n)\vec{U}_n + \mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{U}_{n-1})\vec{U}_{n-1}}{2} + \mathcal{K}_{\vec{u}}^{\mathrm{F}} \frac{\vec{U}_n + \vec{U}_{n-1}}{2} + \mathcal{C}_{\vec{u}}^{\mathrm{F}} \frac{P_n + P_{n-1}}{2} \right) \tag{4.82}$$

for $n = 2, \ldots, N$. It holds that $\mathcal{D}_{\vec{u}}^{\mathrm{F}} \vec{U}_n = 0$ for all $n = 1, \ldots, N$ and, therefore, also

$$\mathcal{D}_{\vec{u}}^{\mathrm{FD}} \frac{\vec{U}_n^{\mathrm{D}} - \vec{U}_{n-1}^{\mathrm{D}}}{t_n - t_{n-1}} + \mathcal{D}_{\vec{u}}^{\mathrm{FF}} \frac{\vec{U}_n^{\mathrm{F}} - \vec{U}_{n-1}^{\mathrm{F}}}{t_n - t_{n-1}} = 0, \qquad n = 2, \ldots, N. \tag{4.83}$$

Substitution of (4.82) in (4.83) gives

$$\mathcal{D}_{\vec{u}}^{\mathrm{FD}} \frac{\vec{U}_n^{\mathrm{D}} - \vec{U}_{n-1}^{\mathrm{D}}}{t_n - t_{n-1}} - \mathcal{D}_{\vec{u}}^{\mathrm{FF}} (\mathcal{M}_{\vec{u}}^{\mathrm{FF}})^{-1} \left(\mathcal{M}_{\vec{u}}^{\mathrm{FD}} \frac{\vec{U}_n^{\mathrm{D}} - \vec{U}_{n-1}^{\mathrm{D}}}{t_n - t_{n-1}} + \frac{\mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{U}_n)\vec{U}_n + \mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{U}_{n-1})\vec{U}_{n-1}}{2} + \mathcal{K}_{\vec{u}}^{\mathrm{F}} \frac{\vec{U}_n + \vec{U}_{n-1}}{2} + \mathcal{C}_{\vec{u}}^{\mathrm{F}} \frac{P_n + P_{n-1}}{2} \right) = 0, \quad n = 2, \ldots, N, \tag{4.84}$$

which can be viewed as a discrete pressure Poisson equation.

A time-discrete reduced-order representation of the pressure field is denoted by

$$P(t_n) \approx P_n^R = \sum_{i=1}^{R'^p} \Phi_i^p c_n^i, \qquad n = 1, \ldots, N. \tag{4.85}$$

Substituting the time-discrete reduced-order representations (4.81) and (4.85) in (4.84) gives rise to the residuals

$$\mathcal{R}_n^{\mathrm{pp}} = \mathcal{D}_{\vec{u}}^{\mathrm{FD}} \frac{\vec{U}_n^{R,\mathrm{D}} - \vec{U}_{n-1}^{R,\mathrm{D}}}{t_n - t_{n-1}} - \mathcal{D}_{\vec{u}}^{\mathrm{FF}} (\mathcal{M}_{\vec{u}}^{\mathrm{FF}})^{-1} \Bigg(\mathcal{M}_{\vec{u}}^{\mathrm{FD}} \frac{\vec{U}_n^{R,\mathrm{D}} - \vec{U}_{n-1}^{R,\mathrm{D}}}{t_n - t_{n-1}} + \frac{\mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{U}_n)\vec{U}_n^R + \mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{U}_{n-1})\vec{U}_{n-1}^R}{2} + \mathcal{K}_{\vec{u}}^{\mathrm{F}} \frac{\vec{U}_n^R + \vec{U}_{n-1}^R}{2} + \mathcal{C}_{\vec{u}}^{\mathrm{F}} \frac{P_n^R + P_{n-1}^R}{2} \Bigg), \quad n = 2, \ldots, N.$$

The residuals are required to be orthogonal to the free parts of the pressure reduced basis functions, i.e. $(\Phi_r^{p,\mathrm{F}})^T \mathcal{R}_n^{\mathrm{pp}} = 0$ for $n = 2, \ldots, N$ and $r = 1, \ldots, R^p$. By inserting the expressions for the reduced-order approximations, (4.78) is obtained for $n = 2, \ldots, N$.

Two alternative methods can be devised to compute P_1^R, both employing the pressure Poisson equation (4.52) at the initial time. Firstly, substituting the reduced-order approximations of the initial pressure, velocity and acceleration at the same time, directly leads to (4.80). Secondly, substituting only the reduced-order approximation of the initial pressure results in the residual

$$\mathcal{R}_n^{\mathrm{pp}} = \mathcal{D}_{\vec{u}}^{\mathrm{FD}} \dot{\vec{U}}_0^{\mathrm{D}} - \mathcal{D}_{\vec{u}}^{\mathrm{FF}} (\mathcal{M}_{\vec{u}}^{\mathrm{FF}})^{-1} \Big(\mathcal{M}_{\vec{u}}^{\mathrm{FD}} \dot{\vec{U}}_0^{\mathrm{D}} + \mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{U}_0)\vec{U}_0 + \mathcal{K}_{\vec{u}}^{\mathrm{F}} \vec{U}_0 + \mathcal{C}_{\vec{u}}^{\mathrm{F}} P_n^R \Big), \qquad n = 1.$$

Orthogonality with respect to the free parts of the pressure reduced basis functions results in

$$\sum_{i=1}^{R'^p} \mathcal{C}_{ri}^{c,\mathrm{F}} c_n^i = (\Phi_r^{p,\mathrm{F}})^T \mathcal{D}_{\vec{u}}^{\mathrm{FD}} \dot{\vec{U}}_0^{\mathrm{D}} - (\Phi_r^{p,\mathrm{F}})^T \mathcal{D}_{\vec{u}}^{\mathrm{FF}} (\mathcal{M}_{\vec{u}}^{\mathrm{FF}})^{-1} \Big(\mathcal{M}_{\vec{u}}^{\mathrm{FD}} \dot{\vec{U}}_0^{\mathrm{D}} + \mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{U}_0)\vec{U}_0 + \mathcal{K}_{\vec{u}}^{\mathrm{F}} \vec{U}_0 \Big), \quad n = 1, \quad r = 1, \ldots, R^p.$$

However, in the following only the first method represented by (4.80) is considered.

4.3.7 Computation of flow quantities

The goal of reduced-order flow modeling is often the computation of a functional of the solution, such as the drag, lift or the Strouhal number. A straightforward approach would be to compute the functionals via the finite element

representation of the solution fields. The respective solution time is dependent on the number of unknowns of the finite element model, however, which deteriorates the computation time of the reduced-order model. Therefore, the reduced-order representation of the solution fields is first substituted in the expressions of the functional. This leads to a reduced-order expression for the functionals, without the need to create the finite element representations of the velocity or pressure fields. Moreover, the Strouhal number can be determined directly by the periodicity of the solutions of the reduced-order model.

Regarding the drag and lift coefficients, the reduced-order approximations are substituted in the discrete equations (4.55) and (4.56), which gives

$$C_{\mathrm{D}}^{R}(t) = \frac{2}{DU_{\infty}^{2}} \mathbb{1} \left(\mathcal{N}_{u}^{\mathrm{D}}(\vec{U}^{R}) U^{R} + \mathcal{K}_{uu}^{\mathrm{D}} U^{R} + \mathcal{K}_{uv}^{\mathrm{D}} V^{R} + \mathcal{C}_{u}^{\mathrm{D}} \; P^{R} \right),$$
$$C_{\mathrm{L}}^{R}(t) = \frac{2}{DU_{\infty}^{2}} \mathbb{1} \left(\mathcal{N}_{u}^{\mathrm{D}}(\vec{U}^{R}) V^{R} + \mathcal{K}_{vu}^{\mathrm{D}} U^{R} + \mathcal{K}_{vv}^{\mathrm{D}} V^{R} + \mathcal{C}_{v}^{\mathrm{D}} \; P^{R} \right).$$

The reduced-order representations are expanded, so that

$$C_{\mathrm{D}}^{R}(t) = \frac{2}{DU_{\infty}^{2}} \mathbb{1} \Big(\sum_{i,j=1}^{R'^{\vec{u}}} \mathcal{N}_{u}^{\mathrm{D}}(\vec{\Phi}_{i}^{\vec{u}}) \Phi_{j}^{u} a^{i} a^{j} + \sum_{i=1}^{R'^{\vec{u}}} \mathcal{K}_{uu}^{\mathrm{D}} \Phi_{i}^{u} a^{i} + \sum_{i=1}^{R'^{\vec{u}}} \mathcal{K}_{uv}^{\mathrm{D}} \Phi_{i}^{v} a^{i} + \sum_{i=1}^{R'^{p}} \mathcal{C}_{u}^{\mathrm{D}} \; \Phi_{i}^{p} c^{i} \Big),$$
$$C_{\mathrm{L}}^{R}(t) = \frac{2}{DU_{\infty}^{2}} \mathbb{1} \Big(\sum_{i,j=1}^{R'^{\vec{u}}} \mathcal{N}_{v}^{\mathrm{D}}(\vec{\Phi}_{i}^{\vec{u}}) \Phi_{j}^{v} a^{i} a^{j} + \sum_{i=1}^{R'^{\vec{u}}} \mathcal{K}_{vu}^{\mathrm{D}} \Phi_{i}^{u} a^{i} + \sum_{i=1}^{R'^{\vec{u}}} \mathcal{K}_{vv}^{\mathrm{D}} \Phi_{i}^{v} a^{i} + \sum_{i=1}^{R'^{p}} \mathcal{C}_{v}^{\mathrm{D}} \; \Phi_{i}^{p} c^{i} \Big).$$

When all computable multiplications are carried out and the resulting products are stored before the model is actually solved, then the drag and lift forces can be computed without having to build the velocity and pressure fields from the solutions of the reduced-order model. To compute the base-pressure coefficient, one can substitute the reduced-order approximation of the pressure in (4.57) to obtain

$$C_{\mathrm{P}}^{R}(t) = 2 \frac{p^{R}(t, \vec{x}_{\mathrm{base}}) - p^{R}(t, \vec{x}_{\mathrm{inlet}})}{U_{\infty}^{2}}.$$

Expanding p^R leads to

$$C_{\mathrm{P}}^R(t) = 2\sum_{i=1}^{R'^p} \frac{\phi_i^p(\vec{x}_{\mathrm{base}}) - \phi_i^p(\vec{x}_{\mathrm{inlet}})}{U_\infty^2} c^i(t).$$

The evaluations of the reduced-basis functions at the sampling points can be carried out before solving the reduced-order model in order to enable a reduced-order computation of the base-pressure coefficient.

4.3.8 Validation

A reduced-order model was created with Matlab from snapshots of the finite element simulation with $M'^p = 23\,582$ and $\Delta t = 0.05$. The time span of interest was $t \in [300, 400]$, which corresponded to approximately 16.7 cycles of the periodic solution. To build the reduced basis functions, two different sets of snapshots were considered. In the first case (A) only the 120 discrete solutions at $t = 300, 300.05, \ldots, 305.95$ were taken as snapshots, so that approximately one period was sampled equidistantly. In the second case (B), the discrete solutions at all 2001 time instances of the time span of interest were chosen. The respective velocity snapshot averages were taken as reference solutions in the context of the control function method to implement the inhomogeneous velocity Dirichlet conditions. For consistency, the control function method was also applied to the pressure. In Figure 4.5 the reference solutions of case (A) for the pressure and the velocity components are presented.

From the snapshot fluctuations, i.e. the snapshots minus their respective snapshot average, L^2-orthogonal reduced basis functions were computed separately for the velocity vector and the pressure field using the control function approach. The first 8 reduced basis functions of the pressure are presented in Figure 4.6 for case (A).

Reduced-order models were created with $R^{\vec{u}} = R^p = 0, \ldots, 20$. The models were solved for the time interval $t \in [300, 400]$. The first velocity snapshot at $t = 300$ was used to initialize the velocity reduced-order model, i.e. the respective finite element coefficients were used as $\vec{U}_0$ in (4.79). The pressure initial condition was computed with (4.80), where the term containing the time derivative could be neglected as it was equal to zero.

The velocity and pressure fields were computed from the reduced-order solution and compared to the finite element solutions in the time interval $t \in [300, 400]$ using the relative space-time L^2 norms. The convergence graphs for the pressure and velocity are presented in Figure 4.7 for case (A), using

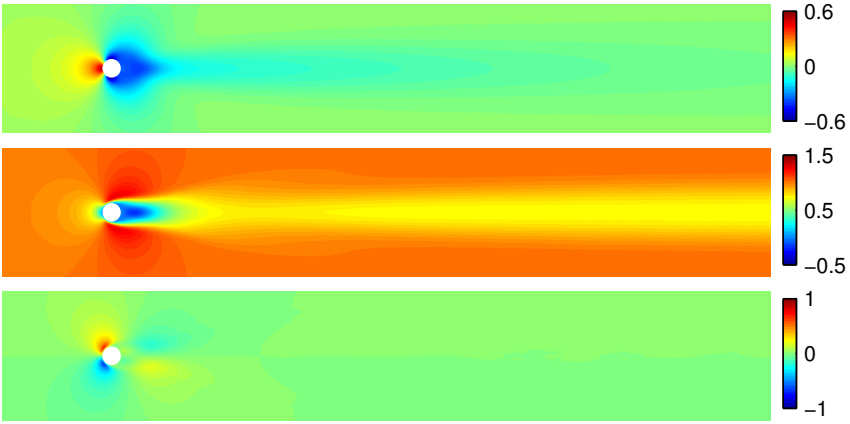

Figure 4.5: Cutouts of the snapshot averages close to the cylinder. From top to bottom: p, u, v.

only snapshots of one cycle, and in Figure 4.8 for case (B), using snapshots of the whole time interval from $t = 300$ to $t = 400$. As a reference, the errors between the snapshot fluctuations and their L^2 projections on the reduced basis are given in all graphs. It can be observed that all errors decay close to exponentially in the beginning, and the model and projection errors differ by not more than about one order of magnitude. While the convergence of the projection error is strictly monotonic, the convergence of the model error is not monotonic. In case (A) the decay of the model and projection errors stagnate when about 25 basis functions or more are used. This means that the snapshots can not be approximated better by further increasing the number of basis functions, rather the number of snapshots must be increased first. This statement is underlined by the results for case (B), where more snapshots have been taken into account. Here the projection error fully converges – in fact it is possible to decrease the error almost down to machine precision when the number of reduced basis functions is increased further. The relative modeling errors reach a minimum in the order magnitude of about 10^{-9} at around 70 basis functions, which can be most probably related to the error in the solution of the non-linear systems of equations. While the numerics regarding the time and space discretization were fully compatible between the finite element model and the reduced-order model, the non-linear systems of equations were solved with different methods using finite precision.

A computational peculiarity in the computation of the pressure should

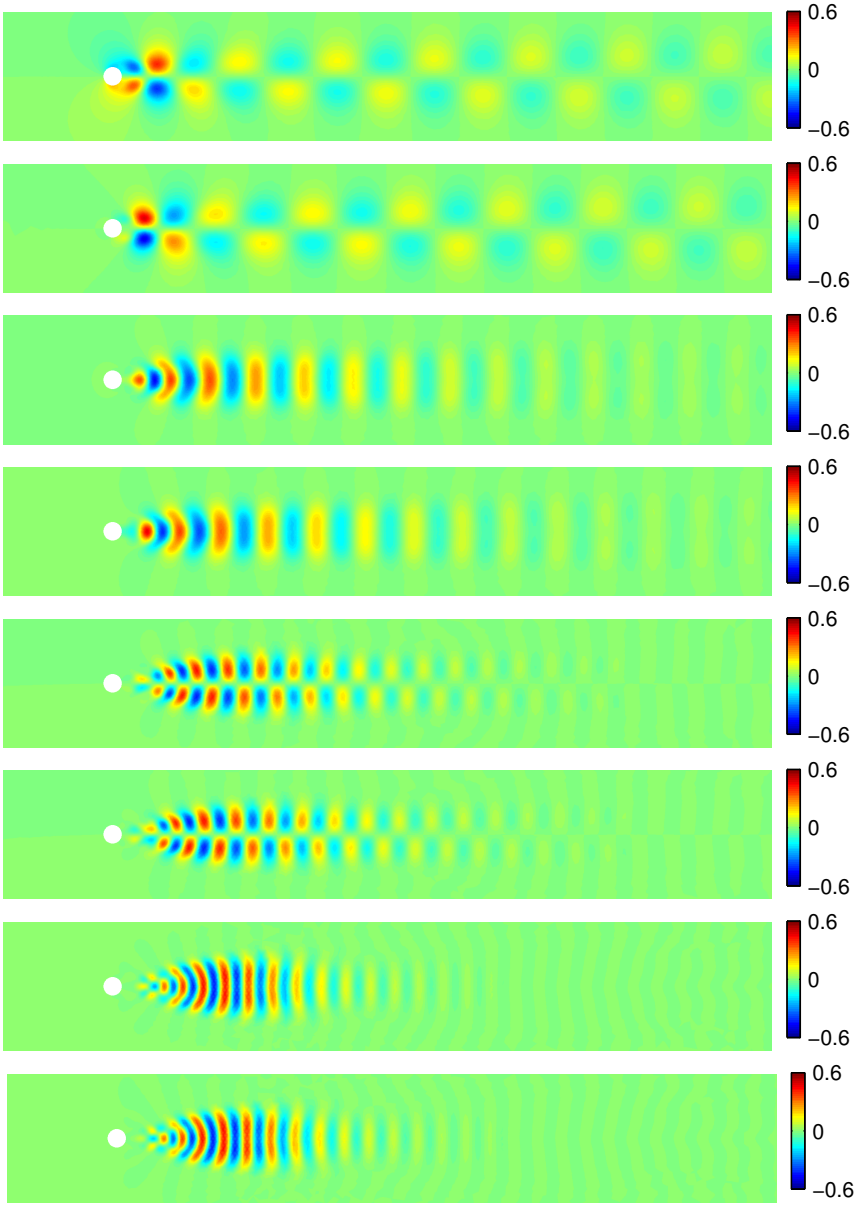

Figure 4.6: Cutouts of the first 8 pressure POD basis functions close to the cylinder.

be noted. The snapshots had been first computed with the trapezoidal rule, which demands that the average of two consecutive snapshots is discretely divergence free. In this case the pressure error of a resulting reduced-order model did not converge as well as in Figure 4.8, even if all snapshots were used to compute the basis functions. The reason was a growing pressure 'wiggle', as anticipated by Gresho and Sani (2000, sect. 3.16.1.a.4). The authors of the book blame the initial divergence error as the reason for the wiggles, but in the present case the initial discrete divergence error is exactly zero. By changing the tolerance of the nonlinear solver, however, it could be verified that the accumulating error of this solver is the source of the wiggle signal. The consequence of the pressure wiggle is the appearance of one or more POD basis functions corresponding to this wiggle signal. The reduced-order model does not mimic the behavior of the nonlinear solver of the finite element model and is therefore not able to reproduce the behavior of the finite element solution. To reduce the appearance of wiggles in time, Gresho and Sani (2000) suggest to use a simplified trapezoidal rule, which demands that at any time step the solution at the new time level is discretely divergence-free. The simplified trapezoidal rule is employed in the finite element model used to create the numerical results of this chapter. Nevertheless, using the standard trapezoidal rule in the time integration of the reduced-order model did not cause any issues.

In Table 4.3, the values of the drag, lift and base-pressure coefficients as well as the Strouhal number are given for different dimensions of the reduced space. The average drag and base-pressure coefficients were already well captured with the reference solution. These quantities depend linearly on the average solution, which is very closely resembled by the reference solution. Notably, adding basis functions can have a negative effect on the accuracy, as can be seen in the base-pressure coefficient computed from the solution with 4 reduced basis functions. The maximum lift relies on the correct time-dependent behavior of the solution. Here, at least about 8 basis functions were required to achieve an error that was less than 1 percent. The Strouhal number was captured already quite well using 2 basis functions. Less than 2 basis functions, however, did not result in a periodic solution. The model with 20 basis functions was able to reproduce the reference values of citet PosdziechGrundmann2007 for all quantities of interest up to the fourth digit.

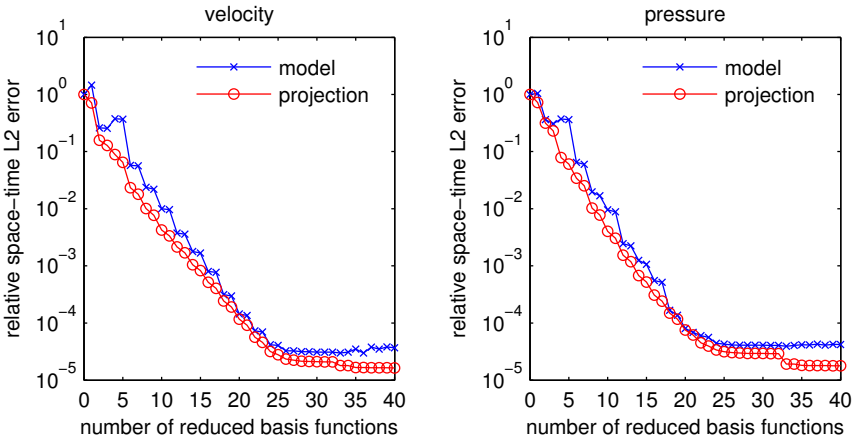

Figure 4.7: Relative space-time L^2 norms of the model error and the projection error depending on the numbers of reduced basis functions for the velocity vector and the pressure. The reduced basis functions were computed using 120 snapshots – case (A).

Table 4.3: Convergence of the quantities of interest with respect to the number of basis functions. The right-most column contains the results of Posdziech and Grundmann (2007)

$R^p = R^{\vec{u}}$	0	2	4	8	20	ref.
Strouhal	–	0.1670	0.1677	0.1666	0.1667	0.1667
lift	0.0000	0.3937	0.3263	0.3325	0.3309	0.3309
drag	1.3504	1.3504	1.3504	1.3504	1.3504	1.3504
base-pressure	0.7448	0.7448	0.7446	0.7448	0.7448	0.7448

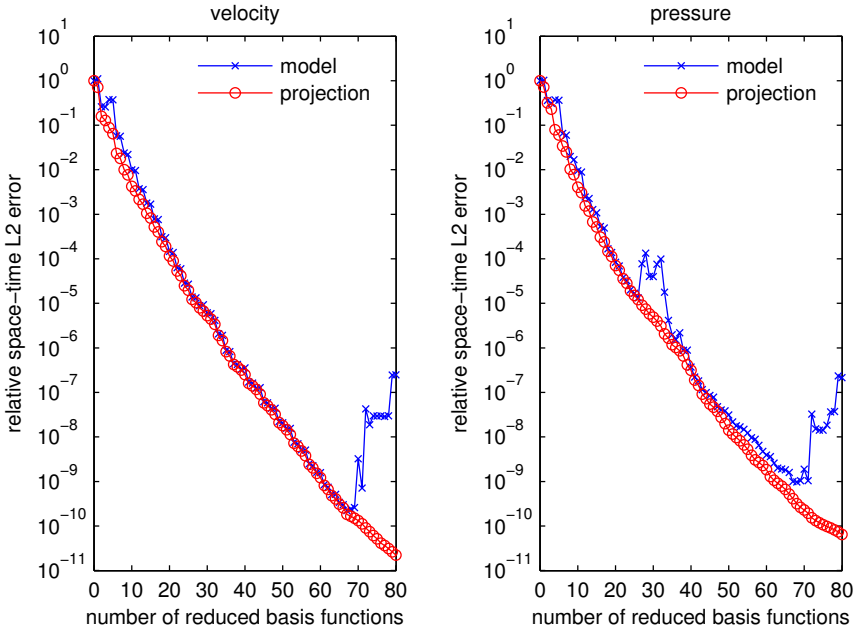

Figure 4.8: Relative space-time L^2 norms of the model error and the projection error depending on the numbers of reduced basis functions for the velocity vector and the pressure. The reduced basis functions were computed using 2001 snapshots – case (B).

Chapter 5

Thermoconvective flow in a channel

The third scenario for POD-Galerkin reduced-order modeling involves the coupling of an incompressible flow with heat transfer by convection and diffusion. The setting is a laminar, transient flow through a horizontal channel heated from below. The problem was proposed by Evans and Paolucci (1990) as a benchmark for outflow conditions in the presence of strong buoyancy effects. An interplay between natural and forced convection leads to a regular pattern of convection cells which are transported through the channel. After the transient effects have decayed, the solution eventually becomes periodic in time. Evans and Paolucci suggest a set of quantities of interest, for which they provide reference values and confidence intervals. The test case was used in Lang (1998) and Gottermeier and Lang (2010) to demonstrate the accuracy of the time integration methods implemented in the KARDOS finite element software (Erdmann et al., 2002). In Schieche and Lang (2013) the flow problem is modified to account for stochastic variations.

Much of the material covered in the previous two chapters is reused for the present setting. Additional features of the problem are the coupling of the heat and momentum equations via the Boussinesq forcing term in the momentum equations and via the convective term in the heat equation. In section 5.1, the governing equations are given in their strong and weak forms. The numerical discretization is presented in section 5.2 and validated with a space and time step refinement study and a comparison to the reference values of Evans and Paolucci. In section 5.3, a reduced-order flow model for the velocity and the temperature is derived from the spatially semi-discretized

problem. A reduced-order pressure model is derived from a discrete pressure Poisson equation. The reduced-order models are validated using the space-time L^2 norms of the errors in the velocity, temperature and pressure with respect to the snapshots and with respect to reference solutions obtained with a higher resolution in space and time.

An error analysis of POD-Galerkin models for the unsteady Boussinesq equations is presented in Ravindran (2011). Respective reduced-order models based on a greedy approach are developed and analyzed in Nguyen et al. (2005) for the steady and in Knezevic et al. (2011) for the unsteady case.

5.1 Governing equations

Let $\Omega \subset \mathbb{R}^2$ be a bounded spatial domain with boundary Γ and closure $\bar{\Omega}$ and let the spatial coordinate vector be $\vec{x} = (x, y)^T$. The boundary is split as $\Gamma = \Gamma_{\mathrm{D}}^u \cup \Gamma_{\mathrm{N}}^u = \Gamma_{\mathrm{D}}^v \cup \Gamma_{\mathrm{N}}^v = \Gamma_{\mathrm{D}}^q \cup \Gamma_{\mathrm{N}}^q$ with $\Gamma_{\mathrm{D}}^u \cap \Gamma_{\mathrm{N}}^u = \Gamma_{\mathrm{D}}^v \cap \Gamma_{\mathrm{N}}^v = \Gamma_{\mathrm{D}}^q \cap \Gamma_{\mathrm{N}}^q = \emptyset$. Furthermore, let $[0, T]$ be a time interval with a final time $T > 0$ and with the time variable denoted by t. An unsteady two-dimensional incompressible flow problem is considered. The unknowns are the velocity vector $\vec{u} = (u, v)^T : [0, T] \times \bar{\Omega} \to \mathbb{R}^2$, the pressure $p : [0, T] \times \bar{\Omega} \to \mathbb{R}$ and the temperature $q : [0, T] \times \bar{\Omega} \to \mathbb{R}$. Buoyancy effects are included in the equations using a Boussinesq approximation. The problem is defined by the set of partial differential algebraic equations

$$\partial_t u + \vec{u} \cdot \nabla u + \partial_x p - \frac{1}{Re} \nabla \cdot (\nabla u + \partial_x \vec{u}) + \frac{g_x}{Fr} q = 0, \tag{5.1}$$

$$\partial_t v + \vec{u} \cdot \nabla v + \partial_y p - \frac{1}{Re} \nabla \cdot (\nabla v + \partial_y \vec{u}) + \frac{g_y}{Fr} q = 0, \tag{5.2}$$

$$\partial_t q + \vec{u} \cdot \nabla q - \frac{1}{Pe} \nabla \cdot \nabla q = 0, \tag{5.3}$$

$$\nabla \cdot \vec{u} = 0, \tag{5.4}$$

for $(t, \vec{x}) \in (0, T] \times \Omega$. The direction of gravity is denoted by the unit vector $\vec{g} = (g_x, g_y)^T$ and the dimensionless parameters are the Reynolds number Re, the Péclet number Pe and the Froude number Fr.

Space-time continuous extensions of the velocity and temperature Dirichlet boundary data into the domain are given by $\vec{u}_{\mathrm{D}} = (u_{\mathrm{D}}, v_{\mathrm{D}})^T : [0, T] \times \bar{\Omega} \to \mathbb{R}^2$ and $q_{\mathrm{D}} : [0, T] \times \bar{\Omega} \to \mathbb{R}$, respectively, which can be written as

$$\vec{u}_{\mathrm{D}}(t, \vec{x}) = \sum_{k=1}^{K^{\vec{u}}} \vec{g}_k^{\vec{u}}(\vec{x}) \eta_{\vec{u}}^k(t), \qquad (t, \vec{x}) \in [0, T] \times \bar{\Omega}, \tag{5.5}$$

$$q_{\mathrm{D}}(t,\vec{x}) = \sum_{k=1}^{K^q} g_k^q(\vec{x})\eta_q^k(t), \qquad (t,\vec{x}) \in [0,T]\times\bar{\Omega}, \tag{5.6}$$

where $K^{\vec{u}}$ and K^q are small positive integers, $\vec{g}_1^{\vec{u}},\ldots,\vec{g}_{K^{\vec{u}}}^{\vec{u}} : \bar{\Omega} \to \mathbb{R}^2$ and $g_1^q,\ldots,g_{K^q}^q : \bar{\Omega}\to\mathbb{R}$ are continuous in space, and $\eta_{\vec{u}}^1,\ldots,\eta_{\vec{u}}^{K^{\vec{u}}} : [0,T]\to\mathbb{R}$ and $\eta_q^1,\ldots,\eta_q^{K^q} : [0,T]\to\mathbb{R}$ are continuous in time. Under these assumptions, Dirichlet boundary conditions are imposed by

$$u = u_{\mathrm{D}}, \qquad (t,\vec{x}) \in (0,T]\times\Gamma_{\mathrm{D}}^u, \tag{5.7}$$
$$v = v_{\mathrm{D}}, \qquad (t,\vec{x}) \in (0,T]\times\Gamma_{\mathrm{D}}^v, \tag{5.8}$$
$$q = q_{\mathrm{D}}, \qquad (t,\vec{x}) \in (0,T]\times\Gamma_{\mathrm{D}}^q, \tag{5.9}$$

and Neumann boundary conditions are imposed by

$$pn_x - \frac{1}{Re}(\nabla u + \partial_x\vec{u})\cdot\vec{n} = 0, \qquad (t,\vec{x}) \in (0,T]\times\Gamma_{\mathrm{N}}^u, \tag{5.10}$$
$$pn_y - \frac{1}{Re}(\nabla v + \partial_y\vec{u})\cdot\vec{n} = 0, \qquad (t,\vec{x}) \in (0,T]\times\Gamma_{\mathrm{N}}^v, \tag{5.11}$$
$$-\frac{1}{Pe}\nabla q\cdot\vec{n} = 0, \qquad (t,\vec{x}) \in (0,T]\times\Gamma_{\mathrm{N}}^q, \tag{5.12}$$

where $\vec{n} = (n_x,n_y)^T$ denotes the surface unit normal vector at the boundary pointing outward of the domain. At least one of the velocity Neumann boundaries is assumed non-empty, in order to have a unique pressure field.

With a given continuous initial velocity $\vec{u}_0 = (u_0,v_0)^T : \bar{\Omega}\to\mathbb{R}^2$ and a given continuous initial temperature $q_0 : \bar{\Omega}\to\mathbb{R}$, the following initial conditions are prescribed:

$$u = u_0, \qquad (t,\vec{x}) \in \{0\}\times\Omega, \tag{5.13}$$
$$u = u_0, \qquad (t,\vec{x}) \in \{0\}\times\Omega, \tag{5.14}$$
$$q = q_0, \qquad (t,\vec{x}) \in \{0\}\times\Omega. \tag{5.15}$$

The initial velocity is assumed to be divergence free, $\nabla\cdot\vec{u}_0 = 0$ in $\bar{\Omega}$. Also, the initial data is assumed to be compatible with the boundary data, $u_0 = u_{\mathrm{D}}$ on $\{0\}\times\Gamma_{\mathrm{D}}^u$, $v_0 = v_{\mathrm{D}}$ on $\{0\}\times\Gamma_{\mathrm{D}}^v$ and $q_0 = q_{\mathrm{D}}$ on $\{0\}\times\Gamma_{\mathrm{D}}^q$.

To derive a weak form of the Boussinesq flow problem, the subspaces $H_{u,0}^1$, $H_{v,0}^1$ and $H_{q,0}^1$ of the Sobolev space H^1 are defined as

$$H_{u,0}^1 := \{a\in H^1 : a|_{\Gamma_{\mathrm{D}}^u} = 0\},$$
$$H_{v,0}^1 := \{a\in H^1 : a|_{\Gamma_{\mathrm{D}}^v} = 0\},$$
$$H_{q,0}^1 := \{a\in H^1 : a|_{\Gamma_{\mathrm{D}}^q} = 0\}.$$

The equations (5.1)–(5.4) are multiplied by functions $\psi^u \in H^1_{u,0}$, $\psi^v \in H^1_{v,0}$, $\psi^q \in H^1_{q,0}$ and $\psi^p \in L^2$, respectively, and integrated over the domain, which yields

$$\begin{aligned}
(\psi^u, \partial_t u + \vec{u} \cdot \nabla u) + (\psi^u, \partial_x p) - \frac{1}{Re}(\psi^u, \nabla \cdot (\nabla u + \partial_x \vec{u})) + \frac{g_x}{Fr}(\psi^u, q) &= 0, \\
(\psi^v, \partial_t v + \vec{u} \cdot \nabla v) + (\psi^v, \partial_y p) - \frac{1}{Re}(\psi^v, \nabla \cdot (\nabla v + \partial_y \vec{u})) + \frac{g_y}{Fr}(\psi^v, q) &= 0, \\
(\psi^q, \partial_t q + \vec{u} \cdot \nabla q) - \frac{1}{Pe}(\psi^q, \nabla \cdot \nabla q) &= 0, \\
(\psi^p, \nabla \cdot \vec{u}) &= 0.
\end{aligned}$$

The divergence theorem, applied to the pressure, viscous and diffusive terms, leads to

$$\begin{aligned}
&(\psi^u, \partial_t u + \vec{u} \cdot \nabla u) - (\partial_x \psi^u, p) + \frac{1}{Re}(\nabla \psi^u, \nabla u + \partial_x \vec{u}) + \frac{g_x}{Fr}(\psi^u, q) \\
&\quad = \int_\Gamma -\psi^u p n_x \, \mathrm{d}\vec{x} + \int_\Gamma \frac{1}{Re} \psi^u (\nabla u + \partial_x \vec{u}) \cdot \vec{n} \, \mathrm{d}\vec{x}, \qquad (5.16) \\
&(\psi^v, \partial_t v + \vec{u} \cdot \nabla v) - (\partial_y \psi^v, p) + \frac{1}{Re}(\nabla \psi^v, \nabla v + \partial_y \vec{u}) + \frac{g_y}{Fr}(\psi^v, q) \\
&\quad = \int_\Gamma -\psi^v p n_y \, \mathrm{d}\vec{x} + \int_\Gamma \frac{1}{Re} \psi^u (\nabla v + \partial_y \vec{u}) \cdot \vec{n} \, \mathrm{d}\vec{x}, \qquad (5.17) \\
&(\psi^q, \partial_t q + \vec{u} \cdot \nabla q) + \frac{1}{Pe}(\nabla \psi^q, \nabla q) \\
&\quad = \int_\Gamma \frac{1}{Pe} \psi^q \nabla q \cdot \vec{n} \, \mathrm{d}\vec{x}, \qquad (5.18) \\
&(\psi^p, \nabla \cdot \vec{u}) = 0. \qquad (5.19)
\end{aligned}$$

Now the boundary terms of (5.16)–(5.18) can be eliminated. At the Dirichlet boundaries they can be dropped, because the basis functions fulfill zero Dirichlet conditions. At the Neumann boundaries the conditions (5.10)–(5.12) can be substituted. The weak form of the flow problem can now be constituted: Find $p \in L^2$, $u - u_\mathrm{D} \in H^1_{u,0}$, $v - v_\mathrm{D} \in H^1_{v,0}$ and $q - q_\mathrm{D} \in H^1_{q,0}$, such that

$$\begin{aligned}
&(\psi^u, \partial_t u + \vec{u} \cdot \nabla u) - (\partial_x \psi^u, p) \\
&\quad + \frac{1}{Re}(\nabla \psi^u, \nabla u + \partial_x \vec{u}) + \frac{g_x}{Fr}(\psi^u, q) = 0 \qquad \forall \psi^u \in H^1_{u,0}, \qquad (5.20) \\
&(\psi^v, \partial_t v + \vec{u} \cdot \nabla v) - (\partial_y \psi^v, p) \\
&\quad + \frac{1}{Re}(\nabla \psi^v, \nabla v + \partial_y \vec{u}) + \frac{g_y}{Fr}(\psi^v, q) = 0 \qquad \forall \psi^v \in H^1_{v,0}, \qquad (5.21)
\end{aligned}$$

$$(\psi^q, \partial_t q + \vec{u} \cdot \nabla q) + \frac{1}{Pe}(\nabla \psi^q, \nabla q) = 0 \qquad \forall \psi^q \in H^1_{q,0}, \tag{5.22}$$

$$(\psi^p, \nabla \cdot \vec{u}) = 0 \qquad \forall \psi^p \in L^2, \tag{5.23}$$

as well as $\vec{u}(0, \vec{x}) = \vec{u}_0(\vec{x})$ and $q(0, \vec{x}) = q_0(\vec{x})$.

5.2 Finite element modeling

The finite element discretization features continuous Lagrangian finite elements with piecewise linear basis functions for the pressure and piecewise quadratic finite elements for the velocity and the temperature. Most of the spatially discretized operators are identical to the ones of the heat conduction and isothermal flow problems of the preceding chapters. However, new terms in the equations are given by the thermal forcing in the momentum equations and by the convective transport in the heat equation. For the time discretization, the Crank-Nicolson scheme is employed again.

5.2.1 Finite element discretization in space

Using the definitions of the finite element discretization of the heat conduction problem and the isothermal flow problem presented in the previous chapters, the discrete weak form of the thermoconvective flow is given as follows: Find $p^h \in L^2_{p,h}$, $u^h - u^h_{\mathrm{D}} \in H^1_{u,0,h}$, $v^h - v^h_{\mathrm{D}} \in H^1_{v,0,h}$ and $q^h - q^h_{\mathrm{D}} \in H^1_{q,0,h}$, such that

$$\begin{aligned}
&(\psi^u, \partial_t u^h + \vec{u}^h \cdot \nabla u^h) - (\partial_x \psi^u, p^h) \\
&\quad + \frac{1}{Re}(\nabla \psi^u, \nabla u^h + \partial_x \vec{u}^h) + \frac{g_x}{Fr}(\psi^u, q^h) = 0 \quad \forall \psi^u \in H^1_{u,0,h}, \quad t \in (0, T], \\
&(\psi^v, \partial_t v^h + \vec{u}^h \cdot \nabla v^h) - (\partial_y \psi^v, p^h) \\
&\quad + \frac{1}{Re}(\nabla \psi^v, \nabla v^h + \partial_y \vec{u}^h) + \frac{g_y}{Fr}(\psi^v, q^h) = 0 \quad \forall \psi^v \in H^1_{v,0,h}, \quad t \in (0, T], \\
&(\psi^q, \partial_t q^h + \vec{u}^h \cdot \nabla q^h) + \frac{1}{Pe}(\nabla \psi^q, \nabla q^h) = 0 \qquad \forall \psi^q \in H^1_{q,0,h}, \quad t \in (0, T], \\
&(\psi^p, \nabla \cdot \vec{u}^h) = 0 \qquad \forall \psi^p \in L^2_{p,h}, \quad t \in (0, T], \\
&(\psi^u, u^h - u^h_0) = 0 \qquad \forall \psi^u \in H^1_{u,0,h}, \quad t \in \{0\}, \\
&(\psi^v, v^h - v^h_0) = 0 \qquad \forall \psi^v \in H^1_{v,0,h}, \quad t \in \{0\}, \\
&(\psi^q, q^h - q^h_0) = 0 \qquad \forall \psi^q \in H^1_{q,0,h}, \quad t \in \{0\}.
\end{aligned}$$

The unknowns and initial conditions are substituted by their finite element representations (3.11), (3.13), (4.24)–(4.27), (4.30) and (4.31), and the

functions ψ^u, ψ^v, ψ^p and ψ^q are replaced by the respective finite element basis functions in turn. The resulting equations are:

- Momentum equation in x-direction for $t \in (0, T]$ and $i = 1, \ldots, M^u$:

$$\sum_{j=1}^{M'^u} \underbrace{(\psi_i^u, \psi_j^u)}_{\mathcal{M}_u^{\mathrm{F}}} \dot{U}_j + \sum_{j=1}^{M'^u} \underbrace{\left(\psi_i^u, \sum_{k=1}^{M'^u} U_k \psi_k^u \partial_x \psi_j^u + \sum_{k=1}^{M'^v} V_k \psi_k^v \partial_y \psi_j^u\right)}_{\mathcal{N}_u^{\mathrm{F}}(\vec{U})} U_j$$
$$+ \sum_{j=1}^{M'^p} \underbrace{-(\partial_x \psi_i^u, \psi_j^p)}_{\mathcal{C}_u^{\mathrm{F}}} P_j + \sum_{j=1}^{M'^v} \underbrace{\frac{1}{Re}(\partial_y \psi_i^u, \partial_x \psi_j^v)}_{\mathcal{K}_{uv}^{\mathrm{F}}} V_j$$
$$+ \sum_{j=1}^{M'^u} \underbrace{\frac{1}{Re}\Big(2(\partial_x \psi_i^u, \partial_x \psi_j^u) + (\partial_y \psi_i^u, \partial_y \psi_j^u)\Big)}_{\mathcal{K}_{uu}^{\mathrm{F}}} U_j$$
$$+ \sum_{j=1}^{M'^q} \underbrace{\frac{g_x}{Pe}(\psi_i^u, \psi_j^q)}_{\mathcal{B}_u^{\mathrm{F}}} Q_j = 0.$$

- Momentum equation in y-direction for $t \in (0, T]$ and $i = 1, \ldots, M^v$:

$$\sum_{j=1}^{M'^v} \underbrace{(\psi_i^v, \psi_j^v)}_{\mathcal{M}_v^{\mathrm{F}}} \dot{V}_j + \sum_{j=1}^{M'^v} \underbrace{\left(\psi_i^v, \sum_{k=1}^{M'^u} U_k \psi_k^u \partial_x \psi_j^v + \sum_{k=1}^{M'^v} V_k \psi_k^v \partial_y \psi_j^v\right)}_{\mathcal{N}_v^{\mathrm{F}}(\vec{U})} V_j$$
$$+ \sum_{j=1}^{M'^p} \underbrace{-(\partial_y \psi_i^v, \psi_j^p)}_{\mathcal{C}_v^{\mathrm{F}}} P_j + \sum_{j=1}^{M'^u} \underbrace{\frac{1}{Re}(\partial_x \psi_i^v, \partial_y \psi_j^u)}_{\mathcal{K}_{vu}^{\mathrm{F}}} U_j$$
$$+ \sum_{j=1}^{M'^v} \underbrace{\frac{1}{Re}\Big((\partial_x \psi_i^v, \partial_x \psi_j^v) + 2(\partial_y \psi_i^v, \partial_y \psi_j^v)\Big)}_{\mathcal{K}_{vv}^{\mathrm{F}}} V_j$$
$$+ \sum_{j=1}^{M'^q} \underbrace{\frac{g_y}{Pe}(\psi_i^v, \psi_j^q)}_{\mathcal{B}_v^{\mathrm{F}}} Q_j = 0.$$

- Heat equation for $t \in (0, T]$ and $i = 1, \ldots, M^q$:

$$\sum_{j=1}^{M'^q} \underbrace{(\psi_i^q, \psi_j^q)}_{\mathcal{M}_q^{\mathrm{F}}} \dot{Q}_j + \sum_{j=1}^{M'^q} \underbrace{\Big(\psi_i^q, \sum_{k=1}^{M'^u} U_k \psi_k^u \partial_x \psi_j^q + \sum_{k=1}^{M'^v} V_k \psi_k^v \partial_y \psi_j^q\Big)}_{\mathcal{N}_q^{\mathrm{F}}(\vec{U})} Q_j$$
$$+ \sum_{j=1}^{M'^q} \underbrace{\frac{1}{Pe}\Big((\partial_x \psi_i^q, \partial_x \psi_j^q) + (\partial_y \psi_i^q, \partial_y \psi_j^q)\Big)}_{\mathcal{K}_q^{\mathrm{F}}} Q_j = 0.$$

- Continuity equation for $t \in (0, T]$ and $i = 1, \ldots, M^p$:

$$\sum_{j=1}^{M'^u} \underbrace{(\psi_i^p, \partial_x \psi_j^u)}_{\mathcal{D}_u^{\mathrm{F}}} U_j + \sum_{j=1}^{M'^v} \underbrace{(\psi_i^p, \partial_y \psi_j^v)}_{\mathcal{D}_v^{\mathrm{F}}} V_j = 0.$$

- Initial condition for u at $t \in \{0\}$ and $i = 1, \ldots, M^u$:

$$\sum_{j=1}^{M'^u} \underbrace{(\psi_i^u, \psi_j^u)}_{\mathcal{M}_u^{\mathrm{F}}} (U_j - u_0(\vec{x}_j^u)) = 0.$$

- Initial condition for v at $t \in \{0\}$ and $i = 1, \ldots, M^v$:

$$\sum_{j=1}^{M'^v} \underbrace{(\psi_i^v, \psi_j^v)}_{\mathcal{M}_v^{\mathrm{F}}} (V_j - v_0(\vec{x}_j^v)) = 0.$$

- Initial condition for q at $t \in \{0\}$ and $i = 1, \ldots, M^q$:

$$\sum_{j=1}^{M'^q} \underbrace{(\psi_i^q, \psi_j^q)}_{\mathcal{M}_q^{\mathrm{F}}} (Q_j - q_0(\vec{x}_j^q)) = 0.$$

Knowing the finite element basis functions for a given mesh, the inner products can be evaluated and assembled in sparse matrices. The braces in

the formulas above indicate which terms enter which matrices. The resulting matrix form of the discretized equations is

$$\begin{aligned}
\mathcal{M}_u^{\mathrm{F}}\dot{U} + \mathcal{N}_u^{\mathrm{F}}(\vec{U})U + \mathcal{C}_u^{\mathrm{F}}\ P + \mathcal{K}_{uu}^{\mathrm{F}}U + \mathcal{K}_{uv}^{\mathrm{F}}V + \mathcal{B}_u^{\mathrm{F}}Q = 0, \qquad & t \in (0,T],\\
\mathcal{M}_v^{\mathrm{F}}\dot{V} + \mathcal{N}_v^{\mathrm{F}}(\vec{U})V + \mathcal{C}_v^{\mathrm{F}}\ P + \mathcal{K}_{vu}^{\mathrm{F}}U + \mathcal{K}_{vv}^{\mathrm{F}}V + \mathcal{B}_v^{\mathrm{F}}Q = 0, \qquad & t \in (0,T],\\
\mathcal{M}_q^{\mathrm{F}}\dot{Q} + \mathcal{N}_q^{\mathrm{F}}(\vec{U})Q + \mathcal{K}_q^{\mathrm{F}}\ Q = 0, \qquad & t \in (0,T],\\
\mathcal{D}_u^{\mathrm{F}}\ U + \mathcal{D}_v^{\mathrm{F}}\ V = 0, \qquad & t \in (0,T],\\
\mathcal{M}_u^{\mathrm{F}}U - \mathcal{M}_u^{\mathrm{F}}U_0 = 0, \qquad & t \in \{0\},\\
\mathcal{M}_v^{\mathrm{F}}V - \mathcal{M}_v^{\mathrm{F}}V_0 = 0, \qquad & t \in \{0\},\\
\mathcal{M}_q^{\mathrm{F}}Q - \mathcal{M}_q^{\mathrm{F}}Q_0 = 0, \qquad & t \in \{0\}.
\end{aligned}$$

Using the combined matrices

$$\begin{aligned}
\mathcal{M}_{\vec{u}}^{\mathrm{F}} &= \begin{pmatrix} \mathcal{M}_u^{\mathrm{F}} & 0 \\ 0 & \mathcal{M}_v^{\mathrm{F}} \end{pmatrix}, & \mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{U}) &= \begin{pmatrix} \mathcal{N}_u^{\mathrm{F}}(\vec{U}) & 0 \\ 0 & \mathcal{N}_v^{\mathrm{F}}(\vec{U}) \end{pmatrix},\\
\mathcal{C}_{\vec{u}}^{\mathrm{F}} &= \begin{pmatrix} \mathcal{C}_u^{\mathrm{F}} \\ \mathcal{C}_v^{\mathrm{F}} \end{pmatrix}, & \mathcal{K}_{\vec{u}}^{\mathrm{F}} &= \begin{pmatrix} \mathcal{K}_{uu}^{\mathrm{F}} & \mathcal{K}_{uv}^{\mathrm{F}} \\ \mathcal{K}_{vu}^{\mathrm{F}} & \mathcal{K}_{vv}^{\mathrm{F}} \end{pmatrix},\\
\mathcal{D}_{\vec{u}}^{\mathrm{F}} &= \begin{pmatrix} \mathcal{D}_u^{\mathrm{F}} & \mathcal{D}_v^{\mathrm{F}} \end{pmatrix}, & \mathcal{B}_{\vec{u}}^{\mathrm{F}} &= \begin{pmatrix} \mathcal{B}_u^{\mathrm{F}} \\ \mathcal{B}_v^{\mathrm{F}} \end{pmatrix},
\end{aligned}$$

one obtains the short form of the finite element semi-discretization of the flow problem,

$$\mathcal{M}_{\vec{u}}^{\mathrm{F}}\dot{\vec{U}} + \mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{U})\vec{U} + \mathcal{C}_{\vec{u}}^{\mathrm{F}}\ P + \mathcal{K}_{\vec{u}}^{\mathrm{F}}\ \vec{U} + \mathcal{B}_{\vec{u}}^{\mathrm{F}}Q = \vec{0}, \qquad t \in (0,T], \tag{5.24}$$

$$\mathcal{M}_q^{\mathrm{F}}\dot{Q} + \mathcal{N}_q^{\mathrm{F}}(\vec{U})Q + \mathcal{K}_q^{\mathrm{F}}\ Q = 0, \qquad t \in (0,T], \tag{5.25}$$

$$\mathcal{D}_{\vec{u}}^{\mathrm{F}}\ \vec{U} = 0, \qquad t \in (0,T], \tag{5.26}$$

$$\mathcal{M}_{\vec{u}}^{\mathrm{F}}\vec{U} - \mathcal{M}_{\vec{u}}^{\mathrm{F}}\vec{U}_0 = \vec{0}, \qquad t \in \{0\}, \tag{5.27}$$

$$\mathcal{M}_q^{\mathrm{F}}Q - \mathcal{M}_q^{\mathrm{F}}Q_0 = 0, \qquad t \in \{0\}, \tag{5.28}$$

with Dirichlet conditions are implemented via

$$Q^m(t) = q_{\mathrm{D}}(t, \vec{x}_m^q), \qquad m = M^q + 1, \ldots, M'^q, \qquad t \in [0,T], \tag{5.29}$$

$$U^m(t) = u_{\mathrm{D}}(t, \vec{x}_m^u), \qquad m = M^u + 1, \ldots, M'^u, \qquad t \in [0,T], \tag{5.30}$$

$$V^m(t) = v_{\mathrm{D}}(t, \vec{x}_m^v), \qquad m = M^v + 1, \ldots, M'^v, \qquad t \in [0,T]. \tag{5.31}$$

In just the same way as for the isothermal flow problem of the last chapter,

a discrete pressure Poisson equation can be derived,

$$0 = \mathcal{D}_{\vec{u}}^{\mathrm{FD}} \dot{\vec{U}}^{\mathrm{D}} - \mathcal{D}_{\vec{u}}^{\mathrm{FF}} (\mathcal{M}_{\vec{u}}^{\mathrm{FF}})^{-1} \Big(\mathcal{M}_{\vec{u}}^{\mathrm{FD}} \dot{\vec{U}}^{\mathrm{D}} + \mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{U}) \vec{U} + \mathcal{K}_{\vec{u}}^{\mathrm{F}} \vec{U} + \mathcal{B}_{\vec{u}}^{\mathrm{F}} Q + \mathcal{C}_{\vec{u}}^{\mathrm{F}} P \Big), \qquad t \in [0, T],$$

which in the present case contains a Boussinesq forcing term. When applied as an initial condition for the pressure, the equations becomes

$$\mathcal{D}_{\vec{u}}^{\mathrm{FF}} (\mathcal{M}_{\vec{u}}^{\mathrm{FF}})^{-1} \mathcal{C}_{\vec{u}}^{\mathrm{F}} P = \mathcal{D}_{\vec{u}}^{\mathrm{FD}} \dot{\vec{U}}_0^{\mathrm{D}} - \mathcal{D}_{\vec{u}}^{\mathrm{FF}} (\mathcal{M}_{\vec{u}}^{\mathrm{FF}})^{-1} \Big(\mathcal{M}_{\vec{u}}^{\mathrm{FD}} \dot{\vec{U}}_0^{\mathrm{D}} + \mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{U}_0) \vec{U}_0 + \mathcal{K}_{\vec{u}}^{\mathrm{F}} \vec{U}_0 + \mathcal{B}_{\vec{u}}^{\mathrm{F}} Q_0 \Big), \qquad t \in \{0\}. \tag{5.32}$$

The initial pressure is zero if the problem starts from a fluid at rest with zero initial temperature as well as zero velocity and zero acceleration at the Dirichlet boundary. In this case all terms on the right-hand side are zero.

The final coupled semi-discrete problem consists of the discrete Dirichlet boundary conditions (5.29), (5.30) and (5.31), the ordinary differential equations (5.24) and (5.25), the linear algebraic equation (5.26) as well as the initial conditions (5.27), (5.28) and (5.32).

5.2.2 Finite difference discretization in time

In the preceding chapters, the fully discrete solution vectors U_n, V_n, Q_n, P_n and $\vec{U}_n$ with respect to the time instances $t_1, \ldots, t_N$ with $0 = t_1 < \cdots < t_N = T$ have been introduced. The fully discretized equations for the thermoconvective flow problem are derived from (5.24) – (5.32). The Crank-Nicolson methods results in the following computational scheme:

- Momentum equation for $n = 2, \ldots, N$:

$$\mathcal{M}_{\vec{u}}^{\mathrm{F}} \frac{\vec{U}_n - \vec{U}_{n-1}}{t_n - t_{n-1}} + \frac{\mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{U}_n)\vec{U}_n + \mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{U}_{n-1})\vec{U}_{n-1}}{2} + \mathcal{K}_{\vec{u}}^{\mathrm{F}} \frac{\vec{U}_n + \vec{U}_{n-1}}{2} + \mathcal{C}_{\vec{u}}^{\mathrm{F}} \frac{P_n + P_{n-1}}{2} + \mathcal{B}_{\vec{u}}^{\mathrm{F}} \frac{Q_n + Q_{n-1}}{2} = \vec{0}. \tag{5.33}$$

- Heat equation for $n = 2, \ldots, N$:

$$\mathcal{M}_q^{\mathrm{F}} \frac{Q_n - Q_{n-1}}{t_n - t_{n-1}} + \frac{\mathcal{N}_q^{\mathrm{F}}(\vec{U}_n)Q_n + \mathcal{N}_q^{\mathrm{F}}(\vec{U}_{n-1})Q_{n-1}}{2} + \mathcal{K}_q^{\mathrm{F}} \frac{Q_n + Q_{n-1}}{2} = 0. \tag{5.34}$$

- Continuity equation for $n = 2, \ldots, N$:

$$\mathcal{D}^{\mathrm{F}}_{\vec{u}} \, \frac{\vec{U}_n + \vec{U}_{n-1}}{2} = 0. \tag{5.35}$$

- Initial equation of the velocity for $n = 1$:

$$\mathcal{M}^{\mathrm{F}}_{\vec{u}} \vec{U}_n - \mathcal{M}^{\mathrm{F}}_{\vec{u}} \vec{U}_0 = \vec{0}. \tag{5.36}$$

- Initial equation of the temperature for $n = 1$:

$$\mathcal{M}^{\mathrm{F}}_{q} Q_n - \mathcal{M}^{\mathrm{F}}_{q} Q_0 = 0. \tag{5.37}$$

- Initial equation of the pressure for $n = 1$:

$$\begin{aligned} \mathcal{D}^{\mathrm{FF}}_{\vec{u}} (\mathcal{M}^{\mathrm{FF}}_{\vec{u}})^{-1} \mathcal{C}^{\mathrm{F}}_{\vec{u}} \, P_n = \mathcal{D}^{\mathrm{FD}}_{\vec{u}} \dot{\vec{U}}^{\mathrm{D}}_0 - \mathcal{D}^{\mathrm{FF}}_{\vec{u}} (\mathcal{M}^{\mathrm{FF}}_{\vec{u}})^{-1} \Big(\\ \mathcal{M}^{\mathrm{FD}}_{\vec{u}} \dot{\vec{U}}^{\mathrm{D}}_0 + \mathcal{N}^{\mathrm{F}}_{\vec{u}} (\vec{U}_0) \vec{U}_0 + \mathcal{K}^{\mathrm{F}}_{\vec{u}} \, \vec{U}_0 \Big). \end{aligned} \tag{5.38}$$

- Dirichlet conditions of the velocity and temperature for $n = 1, \ldots, N$:

$$\begin{aligned} U^m_n &= u_{\mathrm{D}}(t_n, \vec{x}^u_m), && m = M^u + 1, \ldots, M'^u \\ V^m_n &= v_{\mathrm{D}}(t_n, \vec{x}^v_m), && m = M^v + 1, \ldots, M'^v \\ Q^m_n &= q_{\mathrm{D}}(t_n, \vec{x}^q_m), && m = M^q + 1, \ldots, M'^q. \end{aligned}$$

In each time step a non-linear system of equations is solved with a simplified Newton method for the non-Dirichlet components of the solution.

5.2.3 Validation

The finite element discretization presented in the previous section is implemented with Matlab and applied to the test case proposed by Evans and Paolucci (1990). The parameters of the problem are listed in Table 5.1. Differently from the original problem description, a smooth start-up from zero initial conditions is used, which is obtained by multiplying the original Dirichlet data with a shifted cosine function in time. Moreover, while the reference simulation of Evans and Paolucci consists of a domain of length 20, the current simulation consists of a domain of length 10, like in Lang (1998), Gottermeier and Lang (2010) and Schieche and Lang (2013). Otherwise, the problem is identical to the one given by Evans and Paolucci. Figure 5.1 provides a sketch

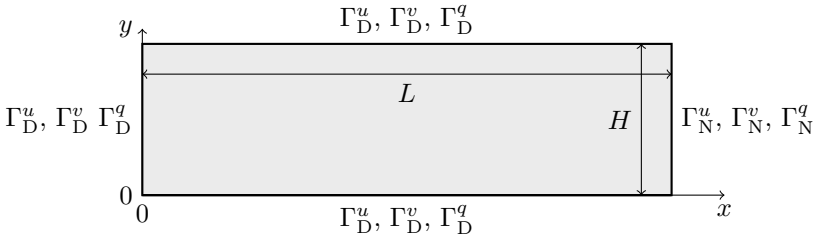

Figure 5.1: Sketch of the geometry of the thermally coupled flow problem.

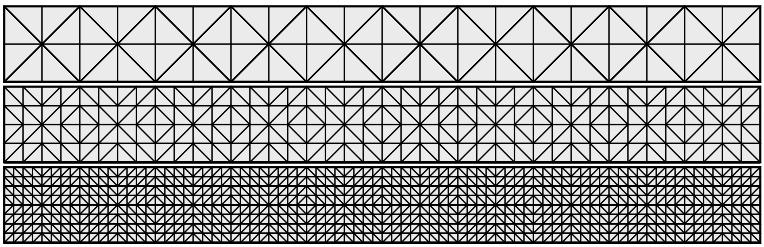

Figure 5.2: Mesh refinement of the thermally coupled flow problem. Starting from an initial mesh (top), the mesh is refined successively by splitting each triangle into four sub-triangles, which halves the minimum edge length in each refinement step (middle and bottom).

of the geometry and the assignment of the boundary conditions. For the validation of the finite element discretization, the space domain is discretized with a uniform mesh, as sketched in Figure 5.2. The minimum edge length defines the mesh width parameter Δx. The time domain is discretized with a uniform time step size Δt.

The components of the solution at the final simulation time are presented in Figure 5.3, where a regular pattern of rising hot plumes with co-existing falling cold plumes can be observed. Near the inflow and the outflow, the solutions deviate from this regular pattern due to the boundary conditions imposed on the problem. Particularly at the outflow, the open boundary condition induces velocities and temperatures that are higher than in the

Table 5.1: Parameters of the thermally coupled flow problem.

description	symbol	expression
domain height	H	1
domain length	L	10
simulation time	T	15
Reynolds number	Re	10
Peclet number	Pe	$20/3$
Froude number	Fr	$1/150$
gravity vector	$\vec{g}$	$(0, -1)$
initial velocity	$\vec{u}_0$	$(0, 0)^T$
initial temperature	q_0	0
x-velocity at Γ_{D}^u	u_{D}	$6y(1-y)\eta_{\vec{u}}$
y-velocity at Γ_{D}^v	v_{D}	0
temperature at Γ_{D}^q	q_{D}	$(1-y)\eta_q$
start-up functions	$\eta_{\vec{u}}$, η_q	$\begin{cases} 1 & \text{if } t \geq 1, \\ \frac{1}{2}\left(1 - \cos(\pi t)\right) & \text{if } t < 1. \end{cases}$

interior domain. The maximum rotating velocity of the eddies is larger than the inflow velocity. Therefore, fluid is entering the domain at the boundary at $x = 10$, which means that this boundary is not a true outflow boundary. Still, the effect of the outflow boundary on the largest part of the interior domain is small, as verified below in the comparison with the reference simulation.

Figure 5.4 shows the time-dependent behavior of the solution at $\vec{x} = (5, 0.5)$. During the run-up time of one time unit, the Dirichlet conditions change from zero to the final values, which is reflected by the plot of the velocity component u. The convective instability, however, causes an unsteady behavior even if the Dirichlet conditions do not change anymore. After a transient solution time of approximately 10 time units, the effect of the initial run-up has decayed and the solution has become time-periodic.

The simulation is validated using a series of discretizations with increasing resolution in space and time. The validation consists of two steps: a comparison with reference quantities of interest from the literature and a convergence study with respect to an L^2 space-time error norm, using a high-resolution finite element solution as a reference.

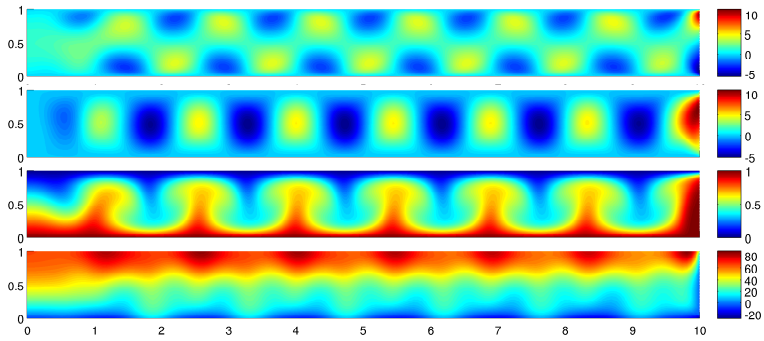

Figure 5.3: Solutions at $t = 15$, from top to bottom: u, v, q, p.

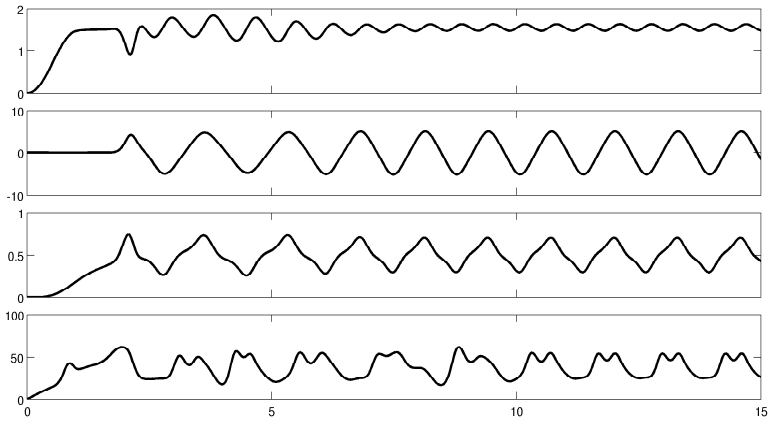

Figure 5.4: Solutions at $\vec{x} = (5, 0.5)$, from top to bottom: u, v, q, p.

In Evans and Paolucci (1990), reference values are given for the time period τ, the space period λ, the average Nusselt number $\langle\overline{Nu}\rangle$, as well as the maxima and minima of each of the velocity components together with the vertical positions at which these extrema have been obtained. The time period was computed via the minimum of the L^2-distance of the solution at $t = 15$ to the solution at the rest of the time interval, using a cubic spline interpolation of the solution between the discrete solution times. The space period was computed via the difference between the temperature at $x = 5$ and $y = 0.5$ and the temperatures at $x > 5$ and $y = 0.5$, using exact finite element interpolations of the solution between the mesh points. The Nusselt number was computed for both the bottom and top boundary. It was averaged in time over the last period and in space over the horizontal interval $2 \leq x \leq 8$. The time averaging employed a cubic spline interpolation between the discrete time instances. The space averaging employed exact finite element interpolations. The computation of the extrema was restricted to $t = 15$ and $2 \leq x \leq 8$ and computed via the finite element representation of the solution.

It is not expected that the results comply exactly with the source for the following reasons:

- There are differences between the problem settings, in particular the run-up phase and the domain length.
- Evans and Paolucci believe that the reference values contain an error of up to about 1% to 2%

A comparison of the values of the functionals given by Evans and Paolucci with the values of the current simulations are presented in Table 5.2 and Table 5.3. The first table demonstrates the convergence of the quantities of interest with respect to the spatial resolution, holding the temporal resolution fixed. The second table demonstrates the convergence with respect to the temporal resolution, holding the spatial resolution fixed. The number of digits is chosen in accordance with the reference. By using a fine enough resolution in space and time, most of the reference values could be reproduced sufficiently well. Only the Nusselt numbers and the spatial period deviate by approximately 3 percent from the reference values, which is slightly more than the assumed error bounds given for the reference values. The cause for this deviation is unclear, because the quantities of interest are numerically well converged. However, as the statement of Evans and Paolucci about the accuracy of their results is formulated quite vaguely, the current results seem acceptable.

Table 5.2: Comparison of the results of the present computation with the reference values of Evans and Paolucci (1990). The results of the finest time resolution with $\Delta t = 1/320$ are shown for mesh widths varying between $\Delta x = 1/4$ and $\Delta x = 1/64$, as to demonstrate the convergence of the results with respect to the spatial resolution. The rightmost column contains the reference values.

Δt	1/320	1/320	1/320	1/320	1/320	reference
Δx	1/4	1/8	1/16	1/32	1/64	reference
τ	1.3233	1.2954	1.2904	1.2900	1.2900	1.3319
λ	1.5227	1.4508	1.4444	1.4441	1.4441	1.4465
Nu_{b}	2.8402	2.8068	2.6688	2.6435	2.6400	2.5583
Nu_{t}	2.8402	2.8068	2.6688	2.6435	2.6400	2.5583
$u_{\max}$	4.5485	4.4281	4.3957	4.3961	4.3964	4.3958
$y\|_{u_{\max}}$	0.7500	0.1875	0.8059	0.8049	0.8048	0.8040
$u_{\min}$	−2.4046	−2.7363	−2.7257	−2.7281	−2.7280	−2.7329
$y\|_{u_{\min}}$	0.8750	0.8571	0.1418	0.1417	0.1412	0.1444
$v_{\max}$	4.6871	5.0122	5.0257	5.0256	5.0256	5.0319
$y\|_{v_{\max}}$	0.5000	0.5000	0.5070	0.5085	0.5088	0.5094
$v_{\min}$	−4.7375	−5.0149	−5.0254	−5.0257	−5.0257	−5.0587
$y\|_{v_{\min}}$	0.5000	0.5000	0.4930	0.4915	0.4912	0.4907

Table 5.3: Comparison of the results of the present computation with the reference values of Evans and Paolucci (1990). The results of the finest space resolution with $\Delta x = 1/64$ are shown for time step sizes varying between $\Delta t = 1/20$ and $\Delta t = 1/320$, as to demonstrate the convergence of the results with respect to the temporal resolution. The rightmost column contains the reference values.

Δt	1/20	1/40	1/80	1/160	1/320	reference
Δx	1/64	1/64	1/64	1/64	1/64	reference
τ	1.2960	1.2915	1.2904	1.2901	1.2900	1.3319
λ	1.4438	1.4440	1.4441	1.4441	1.4441	1.4465
Nu_{b}	2.6398	2.6399	2.6399	2.6400	2.6400	2.5583
Nu_{t}	2.6398	2.6399	2.6400	2.6400	2.6400	2.5583
$u_{\max}$	4.3930	4.3958	4.3962	4.3963	4.3964	4.3958
$y\|_{u_{\max}}$	0.8047	0.8048	0.8048	0.8048	0.8048	0.8040
$u_{\min}$	−2.7354	−2.7299	−2.7285	−2.7281	−2.7280	−2.7329
$y\|_{u_{\min}}$	0.1412	0.1412	0.1412	0.1412	0.1412	0.1444
$v_{\max}$	5.0257	5.0256	5.0256	5.0256	5.0256	5.0319
$y\|_{v_{\max}}$	0.5090	0.5090	0.5088	0.5088	0.5088	0.5094
$v_{\min}$	−5.0250	−5.0256	−5.0257	−5.0257	−5.0257	−5.0587
$y\|_{v_{\min}}$	0.4911	0.4912	0.4912	0.4912	0.4912	0.4907

As a second step of the validation, the simulation with the finest resolution in space and time is taken as a reference solution, which corresponds to 41665 mesh nodes for the pressure and 165249 mesh nodes for the other unknowns, as well as 4800 time steps. The difference between the reference solution and the other solutions is taken as an error estimate for the other solutions. The error is measured via an approximation of the relative space-time L^2 error norm.

The convergence of the errors of the solution components with respect to the number of mesh nodes is plotted in figure Figure 5.5 for fixed numbers of time steps. The graph on the left-hand side was obtained using $N = 601$, or, equivalently, $\Delta t = 1/40$. It shows that the convergence with respect to the spatial resolution is hindered by the time discretization error. The graph on the right-hand side was obtained using $N = 2401$, or, equivalently, $\Delta t = 1/160$. It displays quadratic convergence of the pressure for all considered values of Δx. The graph displays cubic convergence of the velocity components and the temperature between $\Delta x = 1/8$ and $\Delta x = 1/16$. For finer mesh resolutions the error from the time discretization becomes apparent.

The convergence with respect to the number of time steps is plotted in figure Figure 5.6 for fixed numbers of unknowns. The graph on the left-hand side was obtained using 729 pressure nodes and 2737 nodes of the other unknowns, which corresponds to $\Delta x = 1/8$. It shows that the convergence with respect to the time resolution is hindered by the space error. The graph on the right-hand side was obtained using 10593 pressure nodes and 41665 nodes of the other unknowns, which corresponds to $\Delta x = 1/32$. The graph displays quadratic convergence of all solution components. Only for the pressure the spatial discretization error appears between $\Delta t = 1/80$ and $\Delta t = 1/160$.

For the considered discretization schemes, choosing $\Delta x = 1/32$ and $\Delta t = 1/160$ leads to a balancing of the space and time discretization errors, when all components of the solution are of interest. This implies that the finite element model is not significantly over-resolving the time or space dependency.

5.3 Reduced-order modeling

The derivation of the reduced-order model for the temperature-driven flow is carried out only with the semi-discretized approach. Other derivations are possible as well, following the steps presented in the previous chapter. Again, the discrete divergence-freeness of the basis functions for the velocity leads to

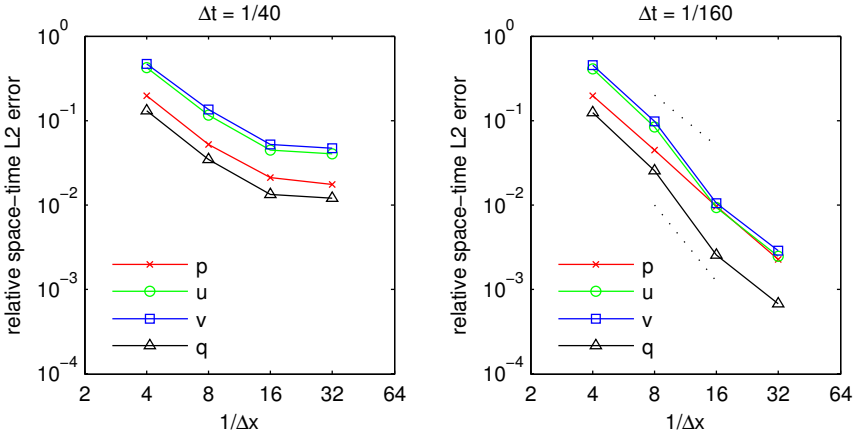

Figure 5.5: Relative space-time L^2 error norms depending on the spatial resolution. The errors are computed with respect to a reference finite element solution with $\Delta t = 1/320$ and $\Delta x = 1/64$. The dotted lines denote the slopes for quadratic (above) and cubic (below) convergence.

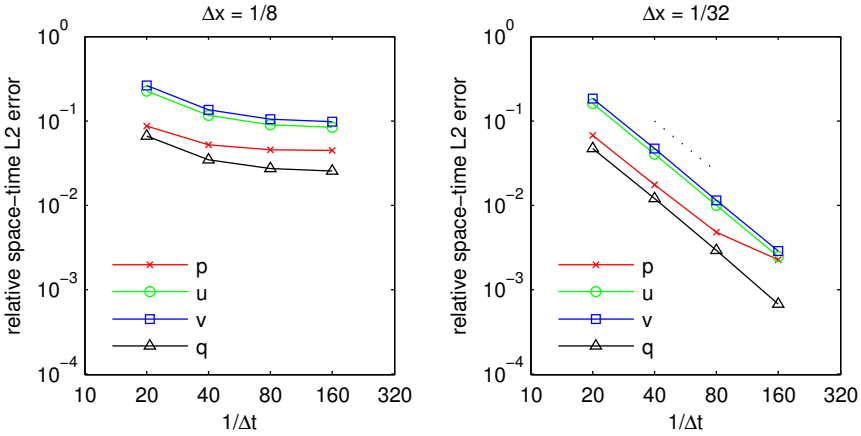

Figure 5.6: Relative space-time L^2 error norms depending on the time step size. The errors are computed with respect to a reference finite element solution with $\Delta t = 1/320$ and $\Delta x = 1/64$. The dotted line denotes the slope for quadratic convergence.

an elimination of the pressure term and the continuity equations. A pressure model derived from a discrete pressure Poisson equations is now depending on the reduced-order coefficients of the velocity and the temperature. A validation of the models is presented with respect to the underlying snapshots and with respect to a simulation with a higher resolution in space and time.

5.3.1 Velocity-temperature model

The derivation of the reduced-order model for the velocity and the temperature is based on the spatially semi-discretized equations (5.24)–(5.28). Independent sets of reduced basis vectors are used for the velocity vector and the temperature. The reduced-order approximations of the temperature and velocity finite element coefficient vectors, (3.38) and (4.68), are substituted in the semi-discretized equations, which gives rise to the residuals of the discretized differential equations,

$$\vec{\mathcal{R}}^{\vec{u}} = \mathcal{M}_{\vec{u}}^{\mathrm{F}} \dot{\vec{U}}^R + \mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{U}^R)\vec{U}^R + \mathcal{C}_{\vec{u}}^{\mathrm{F}}\, P + \mathcal{K}_{\vec{u}}^{\mathrm{F}}\, \vec{U}^R + \mathcal{B}_{\vec{u}}^{\mathrm{F}} Q^R, \quad t \in (0,T], \tag{5.39}$$

$$\mathcal{R}^q = \mathcal{M}_q^{\mathrm{F}} \dot{Q}^R + \mathcal{N}_q^{\mathrm{F}}(\vec{U}^R)Q^R + \mathcal{K}_q^{\mathrm{F}}\, Q^R, \quad t \in (0,T], \tag{5.40}$$

$$\mathcal{R}^p = \mathcal{D}_{\vec{u}}^{\mathrm{F}}\, \vec{U}^R, \quad t \in (0,T], \tag{5.41}$$

and the residuals of the initial conditions

$$\vec{\mathcal{R}}_0^{\vec{u}} = \mathcal{M}_{\vec{u}}^{\mathrm{F}} \vec{U}^R - \mathcal{M}_{\vec{u}}^{\mathrm{F}} \vec{U}_0, \quad t \in \{0\}, \tag{5.42}$$

$$\mathcal{R}_0^q = \mathcal{M}_q^{\mathrm{F}} Q^R - \mathcal{M}_q^{\mathrm{F}} Q_0^R, \quad t \in \{0\}. \tag{5.43}$$

Like in the isothermal case, it can be shown that the residuum (5.41) of the continuity equation is zero. Therefore the continuity equation can be neglected. The reduced-order approximations are expanded and the linear dependence of the discretized convection operators on their arguments is used. The residuals are required to be orthogonal to the non-Dirichlet parts of the respective reduced basis vectors. Finally, the pressure term can be skipped for the same reasons as in the isothermal case. As a consequence, the following set of equations is obtained:

- Velocity equation for $r = 1, \ldots, R^{\vec{u}}$ and $t \in (0,T]$:

$$\sum_{i=1}^{R'^{\vec{u}}} (\vec{\Phi}_r^{\vec{u},\mathrm{F}})^T \mathcal{M}_{\vec{u}}^{\mathrm{F}} \vec{\Phi}_i^{\vec{u}} \dot{a}^i + \sum_{i=1}^{R'^{\vec{u}}} \sum_{j=1}^{R'^{\vec{u}}} (\vec{\Phi}_r^{\vec{u},\mathrm{F}})^T \mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{\Phi}_i^{\vec{u}}) \vec{\Phi}_j^{\vec{u}} a^i a^j + \sum_{i=1}^{R'^{\vec{u}}} (\vec{\Phi}_r^{\vec{u},\mathrm{F}})^T \mathcal{K}_{\vec{u}}^{\mathrm{F}}\, \vec{\Phi}_i^{\vec{u}} a^i + \sum_{i=1}^{R'^{q}} (\vec{\Phi}_r^{\vec{u},\mathrm{F}})^T \mathcal{B}_{\vec{u}}^{\mathrm{F}} \Phi_i^q b^i = 0, \tag{5.44}$$

- Temperature equation for $r = 1, \ldots, R^q$ and $t \in (0, T]$:

$$\sum_{i=1}^{R'^q} (\Phi_r^{q,\mathrm{F}})^T \mathcal{M}_q^{\mathrm{F}} \Phi_i^q \dot{b}^i + \sum_{i=1}^{R'^{\vec{u}}} \sum_{j=1}^{R'^q} (\Phi_r^{q,\mathrm{F}})^T \mathcal{N}_q^{\mathrm{F}}(\vec{\Phi}_i^{\vec{u}}) \Phi_j^q a^i b^j + \sum_{i=1}^{R'^q} (\Phi_r^{q,\mathrm{F}})^T \mathcal{K}_q^{\mathrm{F}} \Phi_j^q b^j = 0, \tag{5.45}$$

- Velocity initial condition for $r = 1, \ldots, R^{\vec{u}}$ and $t \in \{0\}$:

$$\sum_{i=1}^{R'^{\vec{u}}} (\vec{\Phi}_r^{\vec{u},\mathrm{F}})^T \mathcal{M}_{\vec{u}}^{\mathrm{F}} \vec{\Phi}_i^{\vec{u}} a^i - \sum_{i=1}^{R'^{\vec{u}}} (\vec{\Phi}_r^{\vec{u},\mathrm{F}})^T \mathcal{M}_{\vec{u}}^{\mathrm{F}} \vec{U}_0 = 0, \tag{5.46}$$

- Temperature initial condition for $r = 1, \ldots, R^q$ and $t \in \{0\}$:

$$\sum_{i=1}^{R'^q} (\Phi_r^{q,\mathrm{F}})^T \mathcal{M}_q^{\mathrm{F}} \Phi_i^q b^i - \sum_{i=1}^{R'^q} (\Phi_r^{q,\mathrm{F}})^T \mathcal{M}_q^{\mathrm{F}} Q_0 = 0, \tag{5.47}$$

with input data $a^{R^{\vec{u}}+1}, \ldots, a^{R'^{\vec{u}}}$ and $b^{R^q+1}, \ldots, b^{R'^q}$.

5.3.2 Pressure model

A discrete pressure Poisson equation for the Boussinesq problem is derived like in the Navier-Stokes case, with the only difference being the additional temperature term. The resulting discrete pressure Poisson equation is

$$\mathcal{D}_{\vec{u}}^{\mathrm{FD}} \dot{\vec{U}}^{\mathrm{D}} - \mathcal{D}_{\vec{u}}^{\mathrm{FF}} (\mathcal{M}_{\vec{u}}^{\mathrm{FF}})^{-1} \Big(\mathcal{M}_{\vec{u}}^{\mathrm{FD}} \dot{\vec{U}}^{\mathrm{D}} + \mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{U}) \vec{U} + \mathcal{K}_{\vec{u}}^{\mathrm{F}} \vec{U} + \mathcal{B}_{\vec{u}}^{\mathrm{F}} Q + \mathcal{C}_{\vec{u}}^{\mathrm{F}} P \Big) = 0, \qquad t \in [0, T].$$

The reduced-order approximations of the velocity, temperature and the pressure are substituted. The resulting residual is required to be orthogonal to

the pressure reduced basis vectors. This leads to

$$\sum_{i=R^{\vec{u}}+1}^{R'^{\vec{u}}} (\Phi_r^{p,\mathrm{F}})^T \mathcal{D}_{\vec{u}}^{\mathrm{FD}} \vec{\Phi}_i^{\vec{u},\mathrm{D}} \dot{a}^i - (\Phi_r^{p,\mathrm{F}})^T \mathcal{D}_{\vec{u}}^{\mathrm{FF}} (\mathcal{M}_{\vec{u}}^{\mathrm{FF}})^{-1} \Bigg(\sum_{i=R^{\vec{u}}+1}^{R'^{\vec{u}}} \mathcal{M}_{\vec{u}}^{\mathrm{FD}} \vec{\Phi}_i^{\vec{u},\mathrm{D}} \dot{a}^i + \sum_{i,j=1}^{R'^{\vec{u}}} \mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{\Phi}_i^{\vec{u}}) \vec{\Phi}_j^{\vec{u}} a^i a^j + \sum_{i=1}^{R'^{\vec{u}}} \mathcal{K}_{\vec{u}}^{\mathrm{F}} \vec{\Phi}_i^{\vec{u}} a^i + \sum_{i=1}^{R'^{\vec{u}}} \mathcal{B}_{\vec{u}}^{\mathrm{F}} \Phi_i^q b^i + \sum_{i=1}^{R'^{p}} \mathcal{C}_{\vec{u}}^{\mathrm{F}} \Phi_i^p c^i \Bigg) = 0 \tag{5.48}$$

for $r = 1, \ldots, R^p$ and $t \in [0, T]$ with input data $a^1, \ldots, a^{R'^{\vec{u}}}$, $b^1, \ldots, b^{R'^q}$ and $c^{R^p+1}, \ldots, c^{R'^p}$.

5.3.3 Finite difference discretization in time

The set of equations (5.44)–(5.48) can be recast in the following form:

- Velocity model for $r = 1, \ldots, R^{\vec{u}}$ and $t \in (0, T]$:

$$\sum_{i=1}^{R'^{\vec{u}}} \mathcal{M}_{ri}^{a,\mathrm{F}} \dot{a}^i + \sum_{i=1}^{R'^{\vec{u}}} \sum_{j=1}^{R'^{\vec{u}}} \mathcal{N}_{rij}^{a,\mathrm{F}} a^i a^j + \sum_{i=1}^{R'^{\vec{u}}} \mathcal{K}_{ri}^{a,\mathrm{F}} a^i + \sum_{i=1}^{R'^q} \mathcal{B}_{ri}^{a,\mathrm{F}} b^i = 0, \tag{5.49}$$

- Temperature model for $r = 1, \ldots, R^q$ and $t \in (0, T]$:

$$\sum_{i=1}^{R'^q} \mathcal{M}_{ri}^{b,\mathrm{F}} \dot{b}^i + \sum_{i=1}^{R'^{\vec{u}}} \sum_{j=1}^{R'^q} \mathcal{N}_{rij}^{b,\mathrm{F}} a^i b^j + \sum_{i=1}^{R'^q} \mathcal{K}_{ri}^{b,\mathrm{F}} b^i = 0, \tag{5.50}$$

- Pressure model for $r = 1, \ldots, R^p$ and $t \in [0, T]$:

$$\sum_{i=1}^{R'^p} \mathcal{C}_{ri}^{c,\mathrm{F}} c^i + \sum_{i=R^{\vec{u}}+1}^{R'^{\vec{u}}} \mathcal{M}_{ri}^{c,\mathrm{F}} \dot{a}^i + \sum_{i=1}^{R'^{\vec{u}}} \sum_{j=1}^{R'^{\vec{u}}} \mathcal{N}_{rij}^{c,\mathrm{F}} a^i a^j + \sum_{i=1}^{R'^{\vec{u}}} \mathcal{K}_{ri}^{c,\mathrm{F}} a^i + \sum_{i=1}^{R'^q} \mathcal{B}_{ri}^{c,\mathrm{F}} b^i = 0, \tag{5.51}$$

- Velocity initial condition for $r = 1, \ldots, R^{\vec{u}}$ and $t \in \{0\}$:

$$\sum_{i=1}^{R'^{\vec{u}}} \mathcal{M}_{ri}^{a,\mathrm{F}} a^i - (\vec{\Phi}_r^{\vec{u},\mathrm{F}})^T \mathcal{M}_{\vec{u}}^{\mathrm{F}} \vec{U}_0 = 0, \tag{5.52}$$

- Temperature initial condition for $r = 1, \ldots, R^q$ and $t \in \{0\}$:

$$\sum_{i=1}^{R'^q} \mathcal{M}_{ri}^{b,\mathrm{F}} b^i - (\Phi_r^{q,\mathrm{F}})^T \mathcal{M}_q^{\mathrm{F}} Q_0 = 0. \tag{5.53}$$

The model coefficients are given by

$$\begin{aligned}
\mathcal{M}_{ri}^{a,\mathrm{F}} &= (\vec{\Phi}_r^{\vec{u},\mathrm{F}})^T \mathcal{M}_{\vec{u}}^{\mathrm{F}} \vec{\Phi}_i^{\vec{u}}, & \mathcal{M}_{ri}^{b,\mathrm{F}} &= (\Phi_r^{q,\mathrm{F}})^T \mathcal{M}_q^{\mathrm{F}} \Phi_i^q, \\
\mathcal{N}_{rij}^{a,\mathrm{F}} &= (\vec{\Phi}_r^{\vec{u},\mathrm{F}})^T \mathcal{N}_{\vec{u}}^{\mathrm{F}} (\vec{\Phi}_i^{\vec{u}}) \vec{\Phi}_j^{\vec{u}}, & \mathcal{N}_{rij}^{b,\mathrm{F}} &= (\Phi_r^{q,\mathrm{F}})^T \mathcal{N}_q^{\mathrm{F}} (\vec{\Phi}_i^{\vec{u}}) \Phi_j^q, \\
\mathcal{K}_{ri}^{a,\mathrm{F}} &= (\vec{\Phi}_r^{\vec{u},\mathrm{F}})^T \mathcal{K}_{\vec{u}}^{\mathrm{F}} \vec{\Phi}_i^{\vec{u}}, & \mathcal{K}_{ri}^{b,\mathrm{F}} &= (\Phi_r^{q,\mathrm{F}})^T \mathcal{K}_q^{\mathrm{F}} \Phi_j^q, \\
\mathcal{B}_{ri}^{a,\mathrm{F}} &= (\vec{\Phi}_r^{\vec{u},\mathrm{F}})^T \mathcal{B}_{\vec{u}}^{\mathrm{F}} \Phi_i^q, \\
\mathcal{M}_{ri}^{c,\mathrm{F}} &= -(\Phi_r^{p,\mathrm{F}})^T \mathcal{D}_{\vec{u}}^{\mathrm{FF}} (\mathcal{M}_{\vec{u}}^{\mathrm{FF}})^{-1} \mathcal{M}_{\vec{u}}^{\mathrm{FD}} \vec{\Phi}_i^{\vec{u},\mathrm{D}} + (\Phi_r^{p,\mathrm{F}})^T \mathcal{D}_{\vec{u}}^{\mathrm{FD}} \vec{\Phi}_i^{\vec{u},\mathrm{D}}, \\
\mathcal{C}_{ri}^{c,\mathrm{F}} &= -(\Phi_r^{p,\mathrm{F}})^T \mathcal{D}_{\vec{u}}^{\mathrm{FF}} (\mathcal{M}_{\vec{u}}^{\mathrm{FF}})^{-1} \mathcal{C}_{\vec{u}}^{\mathrm{F}} \Phi_i^p, \\
\mathcal{N}_{rij}^{c,\mathrm{F}} &= -(\Phi_r^{p,\mathrm{F}})^T \mathcal{D}_{\vec{u}}^{\mathrm{FF}} (\mathcal{M}_{\vec{u}}^{\mathrm{FF}})^{-1} \mathcal{N}_{\vec{u}}^{\mathrm{F}} (\vec{\Phi}_i^{\vec{u}}) \vec{\Phi}_j^{\vec{u}}, \\
\mathcal{K}_{ri}^{c,\mathrm{F}} &= -(\Phi_r^{p,\mathrm{F}})^T \mathcal{D}_{\vec{u}}^{\mathrm{FF}} (\mathcal{M}_{\vec{u}}^{\mathrm{FF}})^{-1} \mathcal{K}_{\vec{u}}^{\mathrm{F}} \vec{\Phi}_i^{\vec{u}}, \\
\mathcal{B}_{ri}^{c,\mathrm{F}} &= -(\Phi_r^{p,\mathrm{F}})^T \mathcal{D}_{\vec{u}}^{\mathrm{FF}} (\mathcal{M}_{\vec{u}}^{\mathrm{FF}})^{-1} \mathcal{B}_{\vec{u}}^{\mathrm{F}} \Phi_i^q.
\end{aligned}$$

The input values are $a^{R^{\vec{u}}+1}, \ldots, a^{R'^{\vec{u}}}$, $b^{R^q+1}, \ldots, b^{R'^q}$,$c^{R^p+1}, \ldots, c^{R'^p}$. To solve this system of equations, the Crank-Nicolson method is applied. This amounts to the following time stepping scheme:

- Momentum equation for $r = 1, \ldots, R^{\vec{u}}$ and $n = 2, \ldots, N$:

$$\begin{aligned}
&\sum_{i=1}^{R'^{\vec{u}}} \mathcal{M}_{ri}^{a,\mathrm{F}} \frac{a_n^i - a_{n-1}^i}{t_n - t_{n-1}} + \sum_{i=1}^{R'^{\vec{u}}} \sum_{j=1}^{R'^{\vec{u}}} \mathcal{N}_{rij}^{a,\mathrm{F}} \frac{a_n^i a_n^j + a_{n-1}^i a_{n-1}^j}{2} \\
&\qquad + \sum_{i=1}^{R'^{\vec{u}}} \mathcal{K}_{ri}^{a,\mathrm{F}} \frac{a_n^i + a_{n-1}^i}{2} + \sum_{i=1}^{R'^q} \mathcal{B}_{ri}^{a,\mathrm{F}} \frac{b_n^i + b_{n-1}^i}{2} = 0.
\end{aligned} \tag{5.54}$$

- Heat equation for $r = 1, \ldots, R^{\vec{u}}$ and $n = 2, \ldots, N$:

$$\begin{aligned}
&\sum_{i=1}^{R'^q} \mathcal{M}_{ri}^{b,\mathrm{F}} \frac{b_n^i - b_{n-1}^i}{t_n - t_{n-1}} + \sum_{i=1}^{R'^{\vec{u}}} \sum_{j=1}^{R'^q} \mathcal{N}_{rij}^{b,\mathrm{F}} \frac{a_n^i b_n^j + a_{n-1}^i b_{n-1}^j}{2} \\
&\qquad + \sum_{i=1}^{R'^q} \mathcal{K}_{ri}^{b,\mathrm{F}} \frac{b_n^i + b_{n-1}^i}{2} = 0.
\end{aligned} \tag{5.55}$$

- Pressure Poisson equation for $r = 1, \ldots, R^p$ and $n = 2, \ldots, N$:

$$\sum_{i=1}^{R'^p} \mathcal{C}_{ri}^{c,\mathrm{F}} \frac{c_n^i + c_{n-1}^i}{2} + \sum_{i=R^{\vec{u}}+1}^{R'^{\vec{u}}} \mathcal{M}_{ri}^{c,\mathrm{F}} \frac{a_n^i - a_{n-1}^i}{t_n - t_{n-1}} + \sum_{i=1}^{R'^{\vec{u}}} \sum_{j=1}^{R'^{\vec{u}}} \mathcal{N}_{rij}^{c,\mathrm{F}} \frac{a_n^i a_n^j + a_{n-1}^i a_{n-1}^j}{2} + \sum_{i=1}^{R'^{\vec{u}}} \mathcal{K}_{ri}^{c,\mathrm{F}} \frac{a_n^i + a_{n-1}^i}{2} + \sum_{i=1}^{R'^q} \mathcal{B}_{ri}^{c,\mathrm{F}} \frac{b_n^i + b_{n-1}^i}{2} = 0. \tag{5.56}$$

- Velocity initial condition for $r = 1, \ldots, R^{\vec{u}}$ and $n = 1$:

$$\sum_{i=1}^{R'^{\vec{u}}} \mathcal{M}_{ri}^{a,\mathrm{F}} a_n^i - (\vec{\Phi}_r^{\vec{u},\mathrm{F}})^T \mathcal{M}_{\vec{u}}^{\mathrm{F}} \vec{U}_0 = 0. \tag{5.57}$$

- Temperature initial condition for $r = 1, \ldots, R^q$ and $n = 1$:

$$\sum_{i=1}^{R'^q} \mathcal{M}_{ri}^{b,\mathrm{F}} b_n^i - (\Phi_r^{q,\mathrm{F}})^T \mathcal{M}_q^{\mathrm{F}} Q_0 = 0. \tag{5.58}$$

- Pressure initial condition for $r = 1, \ldots, R^p$ and $n = 1$:

$$\sum_{i=1}^{R'^p} \mathcal{C}_{ri}^{c,\mathrm{F}} c_n^i + \sum_{i=R^{\vec{u}}+1}^{R'^{\vec{u}}} \mathcal{M}_{ri}^{c,\mathrm{F}} \dot{a}^i(t_n) + \sum_{i,j=1}^{R'^{\vec{u}}} \mathcal{N}_{rij}^{c,\mathrm{F}} a_n^i a_n^j + \sum_{i=1}^{R'^{\vec{u}}} \mathcal{K}_{ri}^{c,\mathrm{F}} a_n^i + \sum_{i=1}^{R'^q} \mathcal{B}_{ri}^{c,\mathrm{F}} b_n^i = 0. \tag{5.59}$$

The input values are $a_n^{R^{\vec{u}}+1}, \ldots, a_n^{R'^{\vec{u}}}$ and $b_n^{R^q+1}, \ldots, b_n^{R'^q}$ for $n = 1, \ldots, N$ as well as $\dot{a}^{R^{\vec{u}}+1}(t_1), \ldots, \dot{a}^{R'^{\vec{u}}}(t_1)$.

The system of equations (5.54) – (5.59) can be solved step by step. It is possible to exclude (5.56) and (5.59), if a pressure output is not needed. The non-linear equations in each time step are solved with Newton's method. In each linear Newton step, an LU decomposition is used to solve the systems of equations.

5.3.4 Validation

The POD-Galerkin reduced-order models introduced above were implemented with Matlab. The Dirichlet conditions were implemented using the control functions approach, employing the snapshots at $t = 1$ as control functions for all components of the solution. To validate the reduced-order models, the relative space-time L^2 error norms are approximated for the reduced-order approximations of p, u, v and q. Firstly, the error between the snapshots and their L^2-projection on the reduced basis is considered. This error results from approximating the snapshots by a linear combination of a small number of reduced basis functions, using coefficients which result from a minimization problem. Secondly, the error between the snapshots and the solution of the reduced-order model is considered. While the reduced-basis functions and Dirichlet functions are the same as in the first case, the coefficients are computed with the reduced-order model. In Figure 5.7, both errors are presented for varying dimensions of the reduced space and for all components of the solution. It can be observed that all errors converge almost exponentially. While the error of the projected snapshots decreases monotonously, this is not the case for the error of the solution of the reduced-order model. Moreover, the error in the solution of the reduced-order model is about one order of magnitude larger than the error of the projected snapshots.

While the errors with respect to the snapshots are good indicators of the correctness of the computer code, the errors with respect to the unknown true solution are ultimately of interest. To estimate the magnitude of these errors, a high-fidelity simulation with a finer grid in space and time was employed to compute a reference solution. The errors of the reduced-order solutions with respect to the reference solution was used to approximate the true error. Figure 5.7 shows that the reduced-order solutions converge toward the underlying finite element snapshots as the number of reduced-basis functions is increased. Therefore, the error of the reduced-order solutions with respect to the true solutions is expected to converge to the error of the snapshots with respect to the true solutions. Exactly this behavior can be observed in Figure 5.8. Here the error of the finite element solution with respect to the reference solution is indicated with a dashed line. At the point where the error between the reduced-order approximations and the underlying snapshots becomes smaller than the error between the snapshots and the reference solution, a further increase in the number of basis functions does not lead to a significant decrease of the error between the reduced-order approximations and the reference solution.

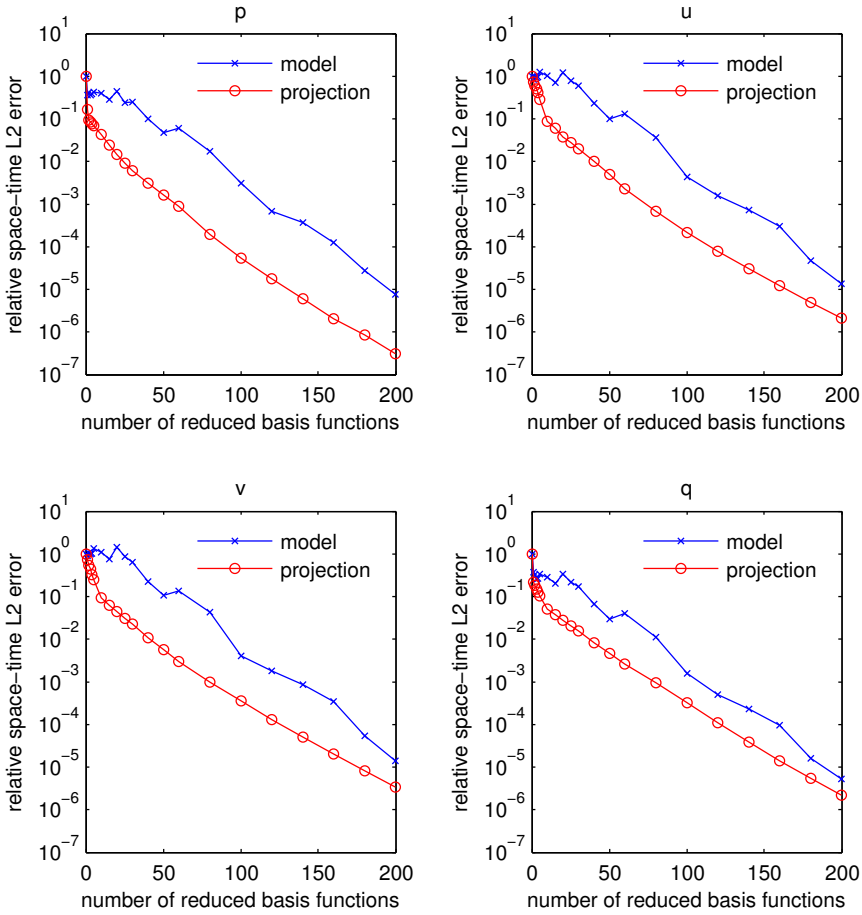

Figure 5.7: Relative space-time L^2 norms of the projection error and the model error depending on the numbers of reduced basis functions for all four components of the solution. The reduced basis functions were computed on a mesh with 41665 mesh nodes (10593 pressure nodes) and 2400 time steps. The errors were computed with respect to these snapshots.

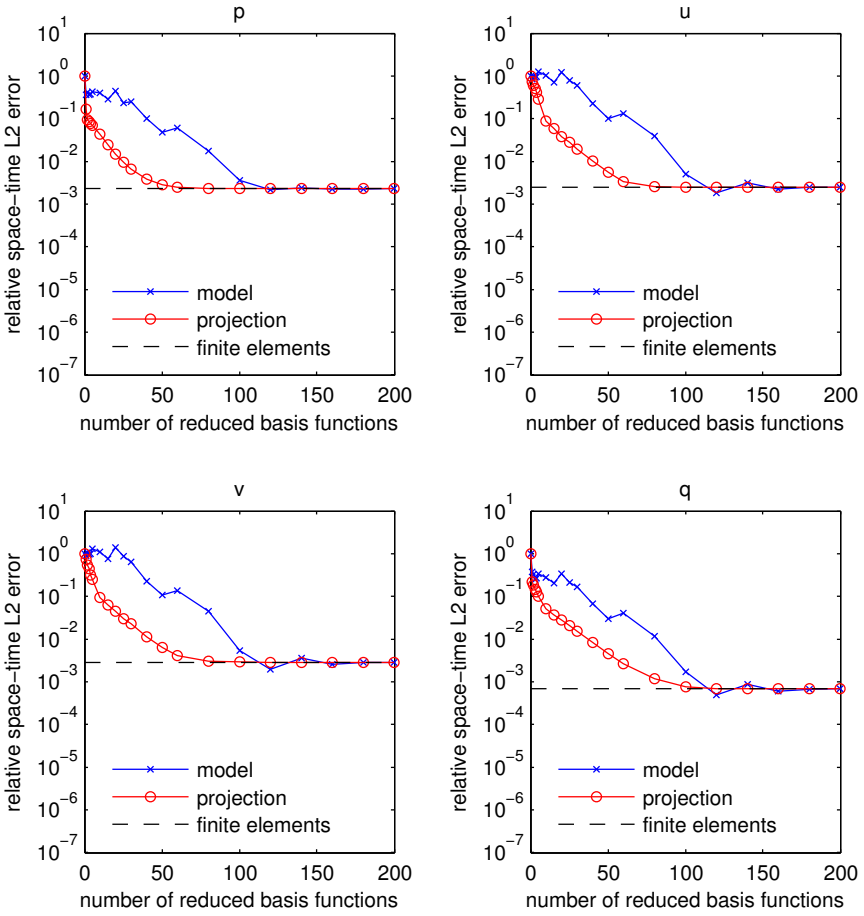

Figure 5.8: Relative space-time L^2 norms of the projection error and the model error depending on the numbers of reduced basis functions for all four components of the solution. The reduced basis functions were computed on a mesh with 41665 mesh nodes (10593 pressure nodes) and 2400 time steps. The errors were computed with respect to a reference solution using 165249 mesh nodes (41665 pressure nodes) and 4800 time steps.

Part II

Stochastic problems

Chapter 6

Stochastic collocation on sparse grids

This chapter introduces the sparse grid stochastic collocation method, which enables uncertainty quantification for problems governed by partial differential equations with random input data. The key question is: How does the uncertainty in the data effect the output of the simulation? This question is related to problems which frequently arise in applications: How large is the effect of measurement errors on a predicted quantity? How robust is a particular design with respect to changes in the environment? What production tolerances can be accepted in order to achieve a certain rate of failure? The stochastic collocation method relies on a parametrization of the random input. It requires the numerical solutions of a set of deterministic problems in order to create an interpolation of the solution in the random parameter domain. Reduced-order modeling is introduced to this framework to accelerate the computation of the statistics of a scalar quantity of interest.

A Karhunen-Loève expansion is often used to parameterize correlated random fields. The terms in the expansion consist of deterministic functions in space multiplied by random coefficients, see section 6.1. While the spatial

functions can usually be computed by means of analytical or numerical methods, the determination of the coefficients usually involves additional modeling.

Having a suitable Karhunen-Loève expansion of the random input available, the solution of the stochastic boundary value problem can be interpreted as a function of the random coefficients. Assuming smooth dependence of the solution on the values of the random coefficients, it is reasonable to approximate the solution by an interpolation using multivariate Lagrangian polynomials with global support. The essence of the collocation method, see section 6.2, is to solve the boundary value problem individually for each of the interpolation points, which are also called collocation points in this context. Collocation has turned out to be attractive in the stochastic setting, because the method combines advantages of its most prominent competitors, namely Monte-Carlo methods and stochastic Galerkin methods. In particular, independent runs of a conventional deterministic code are sufficient to solve the stochastic system – like in a Monte Carlo simulation. Still, fast convergence can be achieved for solutions which depend smoothly on the random variables – like in a stochastic Galerkin method (Le Maître and Knio, 2010).

The stochastic collocation method requires the choice of a set of collocation points. Sparse grids have become widely accepted for this purpose as they diminish the curse of dimensionality that is imposed on standard tensor product grids. An introduction to sparse grids is given in section 6.3.

Sparse grids are able to reduce the number of unknowns compared to a tensor product grid. Still, the number of necessary deterministic solutions can be quite large. In the setting of stochastic collocation, reduced-order modeling can be viewed as either a more sophisticated replacement of the polynomial interpolation or as a cheap surrogate for the deterministic simulations at the collocation points. Section 6.4 presents the use of POD-Galerkin modeling for the purpose of accelerating the sparse-grid collocation method. An overview of alternative approaches is provided, additionally.

For further reading about stochastic collocation, a few publications are listed in the following. The basic sparse grid algorithm was published by Smolyak (1963). In Xiu and Hesthaven (2005) the sparse grid idea is applied to stochastic collocation for elliptic partial differential equations with random coefficients. Babuška et al. (2007) apply stochastic collocation to elliptic partial differential equations with random coefficients and forcing terms using a tensor product grid to discretize the stochastic domain. In Nobile et al. (2008a) the stochastic collocation method is presented in conjunction with isotropic sparse grids and in Nobile et al. (2008b) the method is extended to anisotropic sparse grids. Ma and Zabaras (2009) propose an adaptive sparse grid collocation scheme that handles discontinuities in the stochastic domain.

6.1 Karhunen-Loève expansion

The Karhunen-Loève expansion can be viewed as a generalization of the proper orthogonal decomposition, introduced earlier in this thesis, to the context of stochastic processes. In the following, the method is presented with reference to Ghanem and Spanos (1991). More details about the underlying probability theory can be found e.g. in Rao and Swift (2006).

Let $(\Theta, \Sigma, \mathcal{P})$ be a complete probability space, where Θ is a sample space, Σ is a σ-field on Θ, and $\mathcal{P}$ is a probability measure. The random field $\alpha(y, \theta) : \Omega \times \Theta \to \mathbb{R}$ with bounded domain Ω is introduced. The dth moment of $\alpha(y, \theta)$ is given by

$$E[\alpha(y, \theta)^d] = \int_\Theta \alpha(y, \theta)^d \, \mathrm{d}\mathcal{P}(\theta).$$

It is assumed that $\alpha(y, \theta)$ is a centered second-order random field, i.e. its first moment is zero and its second moment is finite. Consequently, $\alpha(y, \cdot)$ is a random variable with zero expected value and finite second moment on $(\Theta, \Sigma, \mathcal{P})$ for any $y \in \Omega$. Furthermore, $\alpha(\cdot, \theta)$ is a realization of the random field for any $\theta \in \Theta$.

To enable the formulation of the Karhunen Loève expansion, let the autocovariance function of $\alpha(y, \theta)$ be denoted by

$$C(y_1, y_2) = \int_\Theta \alpha(y_1, \theta)\alpha(y_2, \theta) \, \mathrm{d}\mathcal{P}(\theta),$$

which is bounded, symmetric and positive definite. By the theorem of Mercer (1909), there exists a decomposition

$$C(y_1, y_2) = \sum_{i=1}^{\infty} \lambda_i f_i(y_1) f_i(y_2),$$

where λ_i and $f_i(y)$ for $i = 1, 2, \ldots$ are the eigenvalues and eigenfunctions of the covariance kernel. Hence, they are given by the solutions of the integral equation

$$\int_\Omega C(y_1, y_2) f(y_2) \, \mathrm{d}y_2 = \lambda f(y_1).$$

The eigenfunctions are orthogonal and can be normalized, which results in

$$\int_\Omega f_i(y) f_j(y) \, \mathrm{d}y = \delta_{ij}, \qquad i, j = 1, 2, \ldots.$$

The expansion theorem (Courant and Hilbert, 1953) provides the expansion

$$\alpha(y,\theta) = \sum_{i=1}^{\infty} \sqrt{\lambda_i} f_i(y) \xi_i(\theta),$$

with random variables

$$\xi_i(\theta) = \frac{1}{\sqrt{\lambda_i}} \int_\Omega \alpha(y,\theta) f_i(y) \, \mathrm{d}y, \qquad i = 1, 2, \ldots.$$

In the present context, the expansion is called Karhunen-Loève expansion. It can be shown (Ghanem and Spanos, 1991) that the random variables are centered,

$$E[\xi_i(\theta)] = \int_\Theta \xi_i(\theta) \, \mathrm{d}\mathcal{P}(\theta) = 0, \qquad i = 1, 2, \ldots$$

and mutually uncorrelated with unit variance,

$$E[\xi_i(\theta)\xi_j(\theta)] = \int_\Theta \xi_i(\theta)\xi_j(\theta) \, \mathrm{d}\mathcal{P}(\theta) = \delta_{ij}, \qquad i, j = 1, 2, \ldots.$$

The Karhunen-Loève expansion exhibits a number of favorable properties. Most prominently, a truncated Karhunen-Loève expansion with K terms minimizes the mean square error with respect to the original random field among all K-term expansions. A proof and additional properties can be found in Ghanem and Spanos (1991).

The rate of decay of the eigenvalues determines the convergence of a truncated Karhunen-Loève expansion. It is linked to the correlation length of the stochastic process in the following way: A fast decay of the eigenvalues can be expected if the correlation length is in the order of magnitude of the length of the domain. In this case, a Karhunen-Loève expansion provides a reasonable tool for the numerical analysis of the considered problem. On the other hand, a slow decay of the eigenvalues can be expected if the correlation length is small compared to the length of the domain. In this case a large number of terms is necessary in order to accurately represent the random field.

For some types of covariance functions, e.g. exponential functions, it is possible to derive the eigenvalues and eigenfunctions analytically. Otherwise, they have to be approximated numerically. Derivations and algorithms can be found in Schieche (2012a) and references therein. Often modeling decisions have to be taken in order to determine the distribution functions of the random variables. A notable exception is the case of Gaussian random fields, where the respective random variables are independent and standard normally distributed (Le Maître and Knio, 2010, sec. 2.1.4).

6.2 Stochastic collocation

Assume a random field is modeled as a function of a finite number of independent random variables via a truncated Karhunen-Loève expansion and given as an input to a PDE problem. Let an integral of the solution be requested as an output. The procedure to compute the output for one realization of the input can be summarized as follows:

- Create a Karhunen-Loève representation of a realization of the random field.
- Solve the PDE problem with the realization of the random field as an input.
- Compute the integral of the solution.

If the quantity of interest needs to be computed for a large number of realizations, or if integrations over the sample space have to be performed, it is beneficial to replace the procedure by a simpler model, e.g. a sparse grid interpolation. The cost of this method is determined by the numerical solutions of a decoupled set of deterministic problems at predefined collocation points in the image domain of the Karhunen-Loève coefficients. In the following, a summary of the stochastic collocation method is given, based on Xiu and Hesthaven (2005), Babuška et al. (2007), and Le Maître and Knio (2010).

The stochastic collocation method is formulated for problems whose input is characterized by a number of mutually independent stochastic parameters with zero mean and unit variance. The Karhunen-Loève expansion has been introduced as a method to transform a spatially correlated random field into a linear combination of deterministic fields with random coefficients. Still, these random coefficients are not independent in general. A common approach is to introduce independence as a modeling assumption. See Xiu and Hesthaven (2005) for a discussion on the treatment of dependent random variables.

There are model problems which naturally involve a finite number of independent random coefficients. An example is the thermal block with random conductivity studied in Peherstorfer et al. (2013), where the physical domain is divided into blocks and the conductivity in each block is treated as a random parameter. Such types of problems are naturally suited for the stochastic collocation method, because they do not involve the truncation and modeling errors that occur in the presence of a Karhunen-Loève expansion of a non-Gaussian field.

The stochastic collocation method is introduced for a general set of equations, where $\mathcal{A}$ is some possibly non-linear differential operator, f is a source

term, $\mathcal{B}$ is a boundary operator and g some boundary data. The general stochastic boundary value problem is to find a stochastic function $q(\vec{x}, \theta)$ such that $\mathcal{P}$-almost everywhere

$$\begin{aligned} \mathcal{A}(\vec{x}, \vec{\xi}(\theta); q(\vec{x}, \theta)) &= f(\vec{x}, \vec{\xi}(\theta)), & (\vec{x}, \theta) &\in \Omega \times \Theta, \\ \mathcal{B}(\vec{x}, \vec{\xi}(\theta); q(\vec{x}, \theta)) &= g(\vec{x}, \vec{\xi}(\theta)), & (\vec{x}, \theta) &\in \partial\Omega \times \Theta \end{aligned}$$

for some random vector $\vec{\xi}(\theta) = (\xi^1(\theta), \ldots, \xi^K(\theta))^T : \Theta \to \mathbb{R}^K$ which contains random variables of zero mean and unit variance. By the Doob-Dynkin lemma, the solution of the stochastic boundary value problem can be interpreted as a function of the stochastic parameters, so that $q(\vec{x}, \theta) = q(\vec{x}, \vec{\xi}(\theta))$. Let $\Xi^k = \xi^k(\Theta)$ denote the images of the random coefficients for all $k = 1, \ldots, K$ and define $\Xi = \prod_{k=1}^{K} \Xi^k$. Assume that the random vector possesses the joint probability density function $\rho : \Xi \to \mathbb{R}^+$. Now the stochastic boundary value problem can be transferred to the image space (Xiu and Hesthaven, 2005): Find $q(\vec{x}, \vec{\xi})$ such that ρ-almost everywhere

$$\begin{aligned} \mathcal{A}(\vec{x}, \vec{\xi}; q(\vec{x}, \vec{\xi})) &= f(\vec{x}, \vec{\xi}), & (\vec{x}, \vec{\xi}) &\in \Omega \times \Xi \\ \mathcal{B}(\vec{x}, \vec{\xi}; q(\vec{x}, \vec{\xi})) &= g(\vec{x}, \vec{\xi}), & (\vec{x}, \vec{\xi}) &\in \partial\Omega \times \Xi. \end{aligned}$$

The stochastic collocation method focuses on this formulation of the equations.

In the following, the solution $q(\vec{x}, \vec{\xi})$ is represented by a polynomial interpolation. To this end, assume there exists a set of interpolation points $\vec{\xi}_1, \ldots, \vec{\xi}_N \in \Xi$ together with a set of multivariate Lagrangian polynomials $\mathcal{L}_1(\vec{\xi}), \ldots, \mathcal{L}_N(\vec{\xi})$ defined for $\vec{\xi} \in \Xi$, so that $\mathcal{L}_i(\vec{\xi}_j) = \delta_{ij}$ for all $i, j = 1, \ldots, N$. Let $\hat{q}(\vec{x}, \vec{\xi})$ denote the Lagrangian interpolation of the stochastic solution:

$$\hat{q}(\vec{x}, \vec{\xi}) = \sum_{n=1}^{N} \mathcal{L}_n(\vec{\xi}) q(\vec{x}, \vec{\xi}_n).$$

The collocation method requires the equations to be fulfilled at the interpolation points, which are called collocation points in this context:

$$\begin{aligned} \mathcal{A}(\vec{x}, \vec{\xi}_n; q(\vec{x}, \vec{\xi}_n)) &= f(\vec{x}, \vec{\xi}_n), & \vec{x} &\in \Omega, & n &= 1, \ldots, N, \\ \mathcal{B}(\vec{x}, \vec{\xi}_n; q(\vec{x}, \vec{\xi}_n)) &= g(\vec{x}, \vec{\xi}_n), & \vec{x} &\in \partial\Omega, & n &= 1, \ldots, N. \end{aligned}$$

This set of decoupled deterministic equations can be solved with a suitable numerical scheme. Once the numerical solutions are available, the Lagrangian representation of the solution can be evaluated for any point in Ξ. Also, if

integrals over Ξ are to be computed, these can be obtained by using the analytic expressions of the Lagrangian polynomials. For instance, the expected value of the solution can be approximated via

$$E[q(\vec{x},\theta)] \approx E[\hat{q}(\vec{x},\vec{\xi})] = \int_{\Xi} \hat{q}(\vec{x},\vec{\xi})\rho(\vec{\xi})\,\mathrm{d}\Xi.$$

If $\rho(\vec{\xi})$ is the product of probability densities of uniform distributions, then the integral reduces to a weighted sum of integrals over the Lagrangian polynomials.

6.3 Sparse grids

The stochastic collocation method requires the choice of a set of distinct collocation points $\vec{\xi}_1, \ldots, \vec{\xi}_N$. In the context of sparse grid collocation based on global polynomials, the collocation points define a set of Lagrangian polynomials $\mathcal{L}_1(\vec{\xi}), \ldots, \mathcal{L}_N(\vec{\xi})$ used to construct a multivariate polynomial approximation $\hat{q}(\vec{x},\vec{\xi})$ of the solution $q(\vec{x},\vec{\xi})$. Smolyak's algorithm has become a popular means of generating collocation points on a sparse grid (Smolyak, 1963). In the following, the Smolyak sparse-grid interpolation is presented for a domain Ξ that is given by the K-dimensional hypercube $[-1,1]^K$. The one-dimensional case is introduced as a foundation for the multi-dimensional case. The material is based on Barthelmann et al. (2000), Xiu and Hesthaven (2005) and Schieche (2012a).

Let $\chi^1, \chi^2, \ldots$ be sets of nodes with $\chi^i = \{Y_1^i, \ldots, Y_{m_i}^i\}$ and $Y_1^i, \ldots, Y_{m_i}^i \in [-1,1]$ for all $i \in \mathbb{N}$. It is assumed that none of the sets contains duplicate nodes: $Y_j^i \neq Y_k^i$ for all $j \neq k$ and $i \in \mathbb{N}$. The one-dimensional Lagrangian polynomials $\mathcal{L}_j^i(\xi) : [-1,1] \to \mathbb{R}$ are introduced for $j = 1, \ldots, m_i$ and $i \in \mathbb{N}$, so that $\mathcal{L}_j^i(Y_k^i) = \delta_{jk}$ for all $j,k = 1, \ldots, m_i$ and $i \in \mathbb{N}$. Lagrangian polynomial approximations of a function $f(\xi) : [-1,1] \to \mathbb{R}$ can be formed like

$$\mathcal{I}^i(f)(\xi) = \sum_{j=1}^{m_i} \mathcal{L}_j^i(\xi) f(Y_j^i), \qquad \xi \in [-1,1], \qquad i \in \mathbb{N}.$$

From now on, it is assumed that $m_i = 2^{i-1} + 1$ for $i \geq 1$ and that the underlying sequences of nodes are the extrema of the Chebyshev polynomials,

$$Y_j^i = \cos\left(\frac{\pi(j-1)}{m_i - 1}\right), \qquad j = 1, \ldots, m_i, \tag{6.1}$$

for $i > 1$ and $Y_1^i = 0$ for $i = 1$. The one-dimensional Lagrangian polynomials are defined by requiring $\mathcal{I}^i(f)(\xi) = f(\xi)$ for $\xi \in [-1, 1]$ and for all f which are polynomials of degree less than m_i. The choice of Chebyshev points over other options has two reasons. Firstly, it is easy to find a nested sequence of nodes so that $\chi^i \subset \chi^j$ holds for all $i < j$. This requirement is helpful because the computational results at the grid points of a lower refinement level can be reused for grids with a higher refinement level. Secondly, the resulting Clenshaw-Curtis quadrature converges fast with the order of the polynomial (Trefethen, 2011), and the computation of the statistical moments relies on this one-dimensional quadrature.

Based on the one-dimensional polynomial approximations, K-dimensional interpolations can be defined using a tensor product approach. Consider a coordinate vector $\vec{\xi} = (\xi^1, \dots, \xi^K)^T \in \mathbb{R}^K$ and a function $f(\vec{\xi}) = f(\xi^1, \dots, \xi^K) : [-1, 1]^K \to \mathbb{R}$. Let $\vec{i} = (i_1, \dots, i_K)^T$ be a multi-index that determines which set of nodes and polynomial to use in each direction. For example, $i_2 = 3$ means that in the direction of ξ^2 the nodes $Y_1^3, \dots, Y_{m_3}^3$ of χ^3 and the respective Lagrangian polynomials $\mathcal{L}_1^3, \dots, \mathcal{L}_{m_3}^3$ are used. The tensor product Lagrangian polynomial approximation of $f(\xi^1, \dots, \xi^K)$ is given by

$$(\mathcal{I}^{i_1} \otimes \cdots \otimes \mathcal{I}^{i_K})(f)(\vec{\xi}) = \sum_{j_1=1}^{m_{i_1}} \cdots \sum_{j_K=1}^{m_{i_K}} f(Y_{j_1}^{i_1}, \dots, Y_{j_K}^{i_K}) \prod_{k=1}^{K} \mathcal{L}_{j_k}^{i_k}(\xi^k).$$

This formulation allows for different sets of underlying one-dimensional nodes in each direction. This enables a dimension-adaptive interpolation, for instance. Assume for now that the same nodes are used in each direction and that $L = i_1 = \cdots = i_K$ denotes the uniform refinement level. Respective tensor product grids for $K = 2$ and refinement levels of $L = 0, \dots, 5$ are shown in Figure 6.1. The 'curse of dimensionality' is the observation that the number of function evaluations, $N = \prod_{k=1}^{K} m_{i_k}$, increases exponentially with the dimension K. Table 6.1 illustrates that the number of collocation points quickly becomes prohibitively expensive even if a moderate number of nodes is used in the underlying one-dimensional interpolations.

Smolyak's algorithm is an attempt to diminish this effect. The Smolyak polynomial interpolation of level $L \geq 0$ can be written as

$$A(L, K)(f)(\vec{\xi}) = \sum_{L-K+1 \leq |\vec{i}| \leq L} (-1)^{L-|\vec{i}|} \cdot \binom{K-1}{L-|\vec{i}|} \cdot (\mathcal{I}^{i_1} \otimes \cdots \otimes \mathcal{I}^{i_K})(f)(\vec{\xi}),$$

where $|\vec{i}| = i_1 + \cdots + i_K$. Let N denote the number of distinct nodes at which the function f has to be evaluated and let $\vec{\xi}_1, \dots, \vec{\xi}_N$ denote these nodes,

where $\vec{\xi}_n = (\xi_n^1, \ldots, \xi_n^K)^T \in \mathbb{R}^K$ for all $n = 1, \ldots, N$. It can be shown that the Smolyak interpolation can be written as

$$A(L, K)(f)(\vec{\xi}) = \sum_{n=1}^{N} f(\vec{\xi}_n)\mathcal{L}_n(\vec{\xi})$$

with suitable Lagrangian polynomials $\mathcal{L}_1(\vec{\xi}), \ldots, \mathcal{L}_N(\vec{\xi}) : [-1, 1]^K \to \mathbb{R}$.

An illustration of the resulting sparse grids, defined by the nodes $\vec{\xi}_1, \ldots, \vec{\xi}_N$, is given in Figure 6.2 for $K = 2$. Note that the number of points does not grow as fast as in the case of the tensor product grid. This effect can also be seen in Table 6.1, where the number of collocation points is presented for a sparse grid based on Smolyak's algorithm with varying refinement level L and dimension K.

Table 6.1: Number of collocations points N depending on the stochastic dimension K and the sparse grid refinement level L for a tensor product grid.

L	0	1	2	3	4	5	6
$K=0$	1	1	1	1	1	1	1
$K=1$	1	3	5	9	17	33	65
$K=2$	1	9	25	81	289	1089	4225
$K=3$	1	27	125	729	4913	35937	274625
$K=4$	1	81	625	6561	83521	$1.2 \cdot 10^6$	$1.8 \cdot 10^7$
$K=5$	1	243	3125	59049	$1.4 \cdot 10^6$	$3.9 \cdot 10^7$	$1.2 \cdot 10^9$
$K=6$	1	729	15625	531441	$2.4 \cdot 10^7$	$1.3 \cdot 10^9$	$7.5 \cdot 10^{10}$

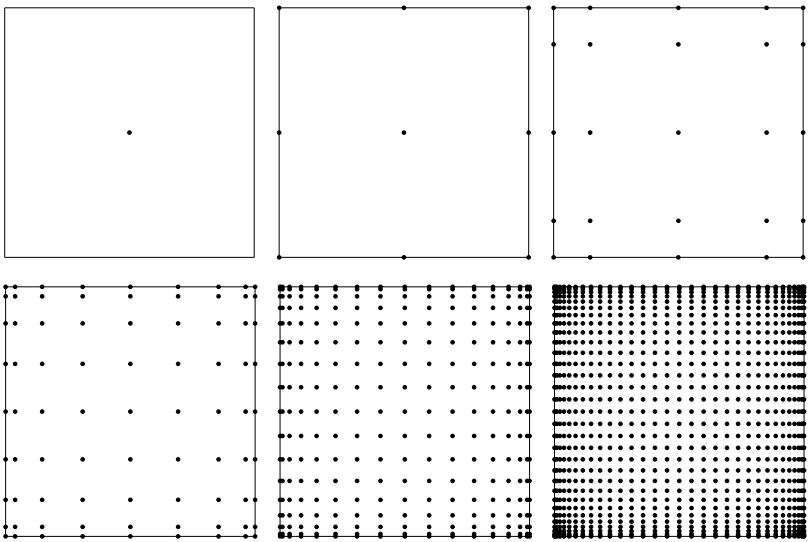

Figure 6.1: Tensor product grids of dimension $K = 2$ based on a nested sequence of Chebyshev points.

Table 6.2: Number of collocations points N depending on the stochastic dimension K and the grid refinement level L for Smolyak's algorithm.

L	0	1	2	3	4	5	6
$K = 0$	1	1	1	1	1	1	1
$K = 1$	1	3	5	9	17	33	65
$K = 2$	1	5	13	29	65	145	321
$K = 3$	1	7	25	69	177	441	1 073
$K = 4$	1	9	41	137	401	1 105	2 929
$K = 5$	1	11	61	241	801	2 433	6 993
$K = 6$	1	13	85	389	1 457	4 865	15 121

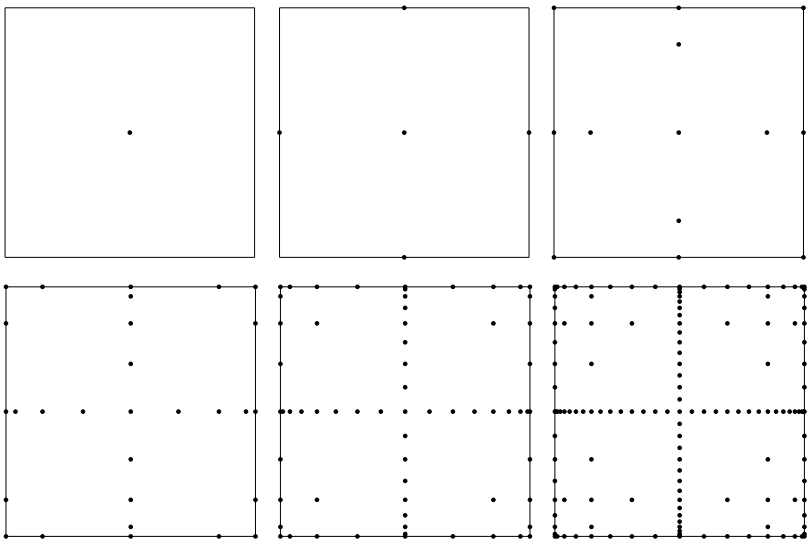

Figure 6.2: Sparse grids of dimension $K = 2$ based on a nested sequence of Chebyshev points.

6.4 Reduced-order modeling

Reduced-order modeling provides a computationally inexpensive input-output map from a point in a parameter domain to a respective quantity of interest. Because the creation of a POD-Galerkin reduced-order model involves the computation of a set of snapshots, such a model can be more efficient than a respective finite element model only if the number of evaluation points is larger than the number of snapshots required to create the model. It seems that the stochastic collocation method is an ideal area of application for reduced-order modeling, as the number of sampling points can be large even if sparse grids are employed. Indeed, a few publications have appeared, where reduced-order modeling is applied to problems that involve uncertainty.

Burkardt et al. (2007) apply POD-Galerkin modeling to a non-linear unsteady stochastic PDE problem with homogeneous Dirichlet boundary conditions. The problem is governed by a heat equation augmented with a cubic non-linearity. A space-time Brownian white noise acts as a source term in the whole domain. In the approach of Burkardt et al., finite-element approximations of the noise term and the non-linear term are still remaining in the POD Galerkin model, which could deteriorate the reduced-order performance.

A similar approach is taken by Gunzburger and Ming (2011) for an incompressible Navier-Stokes flow over a backward-facing step, where a random inflow condition is given by a fixed spatial profile multiplied by time-dependent Brownian white noise. Auxiliary steady computations are performed in addition to the snapshot simulations in order to obtain functions which can be used to implement the stochastic Dirichlet condition via a control function approach. An optimal control problem is presented, where the expected value of the noise is taken as a control and the expected value of the vorticity is taken as a cost function. To solve the problem, reduced-order state and adjoint equations are used as components in a gradient-based optimization routine.

In Boyaval et al. (2009), the reduced basis method (Prud'homme et al., 2002b) is applied to an elliptic boundary value problem with a stochastic Robin boundary condition, which is parametrized using a Karhunen-Loève expansion. An a posteriori error estimator is presented, which are used to assess the errors in the statistical outputs of interest. Statistics are computed with a Monte Carlo method employing evaluations of the reduced-order model.

Chen et al. (2012) present a reduced basis model that builds upon the sparse grid stochastic collocation method for a linear elliptic problem, where the coefficient is parametrized by a finite number of uniformly distributed random variables. The reduced-order model is evaluated at the collocation points of a fine stochastic grid and, subsequently, the fine stochastic grid

is used to compute the final stochastic quantities. In a subsequent report the authors extend the method to problems with non-uniform probability distributions (Chen et al., 2013).

The paper of Elman and Liao (2012) combines the sparse grid stochastic collocation method with the reduced basis method for a variety of applications, including a diffusion problem and a steady incompressible Navier-Stokes problem, with uncertainty in the coefficients and in the boundary. The boundary conditions are implemented using the control function method.

Pasetto et al. (2011) use a POD-Galerkin reduced-order model embedded in a Monte Carlo method to compute the statistics of the hydraulic head in a groundwater flow application. The governing equation is an elliptic partial differential equation with a spatially correlated random source term. The authors observe that their method provides a significant speed-up compared to a conventional Monte Carlo method.

The approach of the present thesis assumes that inhomogeneous Dirichlet conditions are the source of uncertainty. The output of interest is assumed to be a linear functional of the solution of a PDE problem. The goal is to create a stochastic sparse grid interpolation of this output for a dimension $K^\star$ and a Smolyak level $L^\star$ at the computational cost of finite element simulations necessary for the collocation with dimension K and level L. To achieve this, some finite element simulations are replaced by reduced-order simulations. If a fast and accurate reduced-order model can be obtained with snapshots of dimension $K < K^\star$ and/or level $L < L^\star$, then significant savings of computation time can be expected. The following POD-aided stochastic collocation procedure is used:

1. Choose the dimension K and the refinement level L for the computation of the POD snapshots.

2. Create the inhomogeneous Dirichlet boundary conditions for the collocation points of K and L, perform the respective finite element simulations and store the numerical solutions.

3. Create POD basis functions from the numerical solutions.

4. Create a reduced-order model from the POD basis functions. The model takes a realization of $\vec{\xi}$ as input and provides the functional of interest as output.

5. Evaluate the reduced-order model at the collocation points of $K^\star$ and $L^\star$, and store the resulting values of the functional.

6. Create a sparse grid interpolation of dimension $K^\star$ and level $L^\star$ using the data generated with the reduced-order model.

This procedure is closely related to the approach of Chen et al. (2012). The main difference is the choice and computation of the basis functions, as the current approach uses a POD of a fixed set of snapshots and the approach of Chen et al. uses a greedy procedure. A comparison is presented in Table 6.3. Both approaches rely on finite element simulations at a set of snapshot collocation points. In the approach of Chen et al., at first a larger set of training points is defined, which consists of all points of a given stochastic collocation plus a set of points distributed randomly in the stochastic domain. From the set of training points, snapshot points are chosen iteratively by the greedy procedure, using an a posteriori error bound. In the POD method, however, there is no distinction between training points and snapshot points. Instead, finite element simulations are performed at all collocation points of a chosen stochastic refinement level. A POD of this snapshot set is performed, and the leading singular vectors are used to build the reduced-order model. In short, the reduced basis method puts effort in the selection of a small number of snapshot points to exclude redundant information, while the POD method takes a larger number of snapshots and applies a POD to exclude redundant information. As a necessary ingredient, the reduced basis method requires an a posteriori error bound or error estimator.

Table 6.3: Comparison of the present POD-Galerkin approach with the reduced basis approach of Chen et al. (2012).

	POD-Galerkin	reduced basis method
training points	collocation points	collocation and random points
snapshot points	training points	subset of training points
reduced space	POD of snapshots	orthogonalized snapshots

More details about the reduced-order modeling in the context of sparse grid stochastic collocation are given in the next chapter for the particular example of a Boussinesq flow in a cavity. In particular, the interface relating the stochastic coordinates with the reduced-order coefficients are explained in detail for the considered test case.

Chapter 7

Natural convection in a cavity with stochastic temperature Dirichlet boundary condition

This chapter contains an extended and revised version of Ullmann and Lang (2014, to appear). It presents the application of the stochastic sparse grid collocation method to a thermally driven flow in a cavity. The source of uncertainty is the temperature at one of the solid walls surrounding the cavity. A Karhunen-Loève expansion is used to parametrize the stochastic boundary data. The stochastic parameter domain is discretized using sparse grid collocation. The resulting deterministic problems are discretized with finite elements in the physical domain. POD-Galerkin modeling is introduced in order to replace the finite element model with a computationally cheap surrogate model. All numerical computations are performed with Matlab. It is demonstrated that POD-Galerkin reduced-order modeling significantly diminishes the total computation time needed to approximate the statistics of an integral output of interest.

7.1 Governing equations

The problem of Ganapathysubramanian and Zabaras (2007) and Schieche and Lang (2010) is used as a test case. It consists of a steady convective flow in the unit square $\Omega = (0,1) \times (0,1)$ with stochastic boundary conditions. Note that the references employ an unsteady problem formulation, but actually solve the equations for a steady state, which is then used for further processing. Based on the observations made in the references, it is assumed that a steady solution exists and is unique for the considered parameter choices. For computational efficiency, however, a steady solution is computed directly by employing a steady problem formulation.

First, the boundary conditions of the problem are considered, because the uncertainty enters the problem via the temperature prescribed at the boundary. To this end, the boundary Γ of the computational domain Ω is split into the parts Γ_{D}^{q} and Γ_{N}^{q}, as sketched in Fig. 7.1. The spatial coordinate is denoted by $\vec{x} = (x, y)^T$. To characterize the stochastic aspect of the problem, let Θ be the sample space of random events and let elements of Θ be denoted by θ. Further assume $\alpha : (0,1) \times \Theta \to \mathbb{R}$ is a centered second-order random field which is spatially correlated with the exponential covariance function

$$C(y_1, y_2) = e^{-|y_1 - y_2|}.$$

The random field is coupled to the flow problem via a function $q_{\mathrm{D}} : \bar{\Omega} \times \Theta \to \mathbb{R}$, which represents an extension of the temperature Dirichlet boundary data into the domain. It is assumed that q_{D} is continuous in x and fulfills

$$q_{\mathrm{D}}(\vec{x}, \theta) = \begin{cases} 0.5 & \text{if } x = 0, \\ -0.5 + \alpha(y, \theta) & \text{if } x = 1. \end{cases} \tag{7.1}$$

Let the velocity vector be denoted by $\vec{u} = (u, v)^T : \bar{\Omega} \times \Theta \to \mathbb{R}^2$, the pressure by $p : \bar{\Omega} \times \Theta \to \mathbb{R}$ and the temperature by $q : \bar{\Omega} \times \Theta \to \mathbb{R}$. The following boundary conditions are imposed:

$$\vec{u}(\vec{x}, \theta) = \vec{0}, \quad (\vec{x}, \theta) \in \Gamma \times \Theta, \tag{7.2}$$

$$q(\vec{x}, \theta) - q_{\mathrm{D}}(\vec{x}, \theta) = 0, \quad (\vec{x}, \theta) \in \Gamma_{\mathrm{D}}^{q} \times \Theta, \tag{7.3}$$

$$\nabla q(\vec{x}, \theta) \cdot \vec{n} = 0, \quad (\vec{x}, \theta) \in \Gamma_{\mathrm{N}}^{q} \times \Theta. \tag{7.4}$$

The vector $\vec{n}$ denotes the boundary unit normal vector pointing outward of the domain.

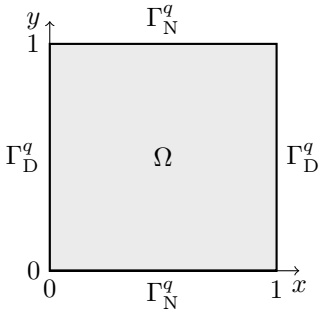

Figure 7.1: Sketch of the domain Ω with the locations of the temperature boundaries Γ^q_{D} and Γ^q_{N}.

In the interior domain, the problem is governed by the steady dimensionless incompressible Navier-Stokes equations with the Boussinesq approximation for the temperature forcing term. All parameters are fixed and incorporated explicitly in the equations, which are consequently given by

$$\vec{u} \cdot \nabla \vec{u} + \nabla p - \nabla \cdot (\nabla \vec{u} + (\nabla \vec{u})^T) + 5000 \vec{g} q = \vec{0}, \quad (\vec{x}, \theta) \in \Omega \times \Theta, \tag{7.5}$$

$$\vec{u} \cdot \nabla q - \nabla \cdot \nabla q = 0, \quad (\vec{x}, \theta) \in \Omega \times \Theta, \tag{7.6}$$

$$\nabla \cdot \vec{u} = 0, \quad (\vec{x}, \theta) \in \Omega \times \Theta \tag{7.7}$$

with the direction of gravity given by $\vec{g} = (0, -1)^T$.

The Nusselt number $Nu : \Theta \to \mathbb{R}$ at the left boundary is defined as the output of the simulation:

$$Nu(q(\vec{x}, \theta)) = -\int_0^1 \partial_x q(\vec{x}, \theta)|_{\vec{x}=(0,y)^T} \, \mathrm{d}y. \tag{7.8}$$

The further computations aim at approximating the statistics of this quantity.

7.2 Karhunen-Loève expansion

A truncated Karhunen-Loève expansion of the stochastic process determining the temperature boundary condition is given by

$$\alpha(y, \theta) \approx \alpha(y, \vec{\xi}(\theta)) = \sum_{i=1}^{K} \sqrt{\lambda_i} \xi^i(\theta) f_i(y). \tag{7.9}$$

Some general properties about Karhunen-Loève expansions have been given in section 6.1. In particular, the random vector $\vec{\xi}(\theta) = (\xi^1(\theta), \ldots, \xi^K(\theta))^T$ contains mutually uncorrelated random variables with zero mean and unit variance.

For the case of exponential covariance it is possible to derive the eigenvalues $\lambda_1, \ldots, \lambda_K$ and the eigenfunctions $f_1, \ldots, f_K$ analytically. By applying the procedure described in Ghanem and Spanos (1991) to the current case, one obtains the following results: If $\omega_1 < \omega_3 < \ldots$ are the strictly positive solutions of

$$\cos(\omega/2) - \omega \sin(\omega/2) = 0$$

and $\omega_2 < \omega_4 < \ldots$ are the strictly positive solutions of

$$\omega \cos(\omega/2) + \sin(\omega/2) = 0,$$

then

$$\lambda_n = \frac{2}{\omega_n^2 + 1}, \qquad n = 1, 2, \ldots$$

and

$$f_n(y) = \begin{cases} \frac{1}{2} + \frac{\cos(\omega_n y)}{\sqrt{\frac{1}{2} + \frac{\sin(\omega_n)}{2\omega_n}}} & n = 1, 3, \ldots, \\ \frac{1}{2} + \frac{\sin(\omega_n y)}{\sqrt{\frac{1}{2} - \frac{\sin(\omega_n)}{2\omega_n}}} & n = 2, 4, \ldots. \end{cases}$$

To determine the probability distributions of the random variables, additional knowledge about the random field $\alpha(y, \theta)$ is required besides its mean and covariance function. To circumvent this issue, $\vec{\xi}(\theta) = (\xi^1(\theta), \ldots, \xi^K(\theta))^T$ is used to denote a *modeled* random vector from now on. The modeled random variables are assumed to be mutual independent and uniformly distributed with interval $[-\sqrt{3}, \sqrt{3}]$, which implies zero mean and unit variance. A realization of the stochastic boundary condition can now be computed via

$$q_{\mathrm{D}}(\vec{x}, \vec{\xi}(\theta)) = \begin{cases} 0.5 & \text{if } x = 0, \\ -0.5 + \sum_{i=1}^{K} \sqrt{\lambda_i} \xi^i(\theta) f_i(y) & \text{if } x = 1, \end{cases} \tag{7.10}$$

where a realization of the random vector can be computed with a suitable pseudo-random number generator. Examples of respective Karhunen-Loève expansions and the underlying functions are presented in Figure 7.2.

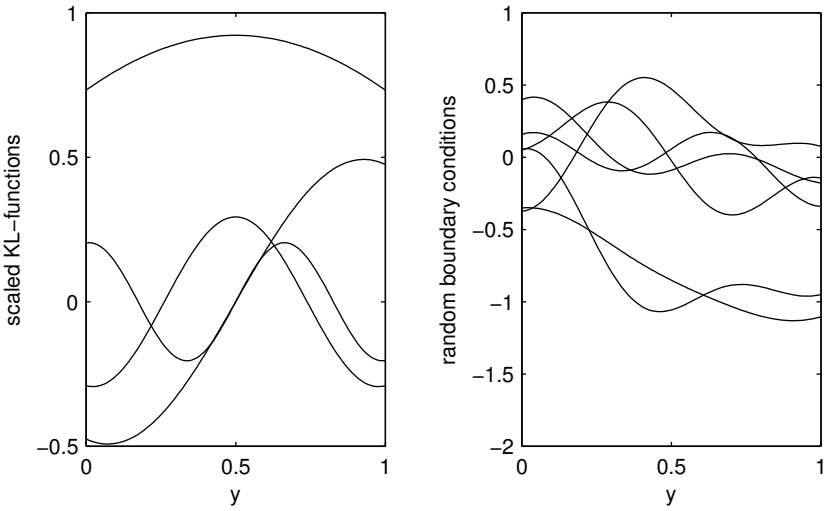

Figure 7.2: Scaled Karhunen-Loève eigenfunctions $\sqrt{\lambda_i} f_i(y)$ for $i = 1, \ldots, 4$ (left) and 6 random realizations of the boundary condition q_{D} using $K = 4$ (right).

7.3 Sparse grid stochastic collocation

The stochastic boundary condition has been modeled as a function of the random vector $\vec{\xi}(\theta)$, which consists of K independent, uniformly distributed random parameters. The sparse grid stochastic collocation method is based on a formulation of the problem in the image space Ξ of the random vector, see section 6.2. Elements of Ξ are denoted by $\vec{\xi} = (\xi^1, \ldots, \xi^K)^T$. The respective governing equations are

$$\vec{u} \cdot \nabla \vec{u} + \nabla p - \nabla \cdot (\nabla \vec{u} + (\nabla \vec{u})^T) + 5000 \vec{g} q = \vec{0}, \quad (\vec{x}, \vec{\xi}) \in \Omega \times \Xi, \tag{7.11}$$

$$\vec{u} \cdot \nabla q - \nabla \cdot \nabla q = 0, \quad (\vec{x}, \vec{\xi}) \in \Omega \times \Xi, \tag{7.12}$$

$$\nabla \cdot \vec{u} = 0, \quad (\vec{x}, \vec{\xi}) \in \Omega \times \Xi, \tag{7.13}$$

$$\vec{u}(\vec{x}, \vec{\xi}) = \vec{0}, \quad (\vec{x}, \vec{\xi}) \in \Gamma \times \Xi, \tag{7.14}$$

$$q(\vec{x}, \vec{\xi}) - q_{\mathrm{D}}(\vec{x}, \vec{\xi}) = 0, \quad (\vec{x}, \vec{\xi}) \in \Gamma_{\mathrm{D}}^q \times \Xi, \tag{7.15}$$

$$\nabla q(\vec{x}, \vec{\xi}) \cdot \vec{n} = 0, \quad (\vec{x}, \vec{\xi}) \in \Gamma_{\mathrm{N}}^q \times \Xi \tag{7.16}$$

with unknowns $\vec{u} : \bar{\Omega} \times \Xi \to \mathbb{R}^2$, $p : \bar{\Omega} \times \Xi \to \mathbb{R}$ and $q : \bar{\Omega} \times \Xi \to \mathbb{R}$ and with

$$q_{\mathrm{D}}(\vec{x}, \vec{\xi}) = \begin{cases} 0.5 & \text{if } x = 0, \\ -0.5 + \sum_{i=1}^{K} \sqrt{\lambda_i} \xi^i f_i(y) & \text{if } x = 1. \end{cases} \tag{7.17}$$

To discretize the image space, collocation points $\vec{\xi}_1, \ldots, \vec{\xi}_N \in \Xi$ are provided by Smolyak's algorithm applied to a nested sequence of Chebyshev points. Details of the algorithm are shown in section 6.3. The Chebyshev points are chosen in an interval of $[-\sqrt{3}, \sqrt{3}]$, which complies with the support of the probability density of the random variables. The number N of collocation points depends on the dimension K of Ξ and the Smolyak refinement level L. With the help of the respective multivariate Lagrangian polynomials $\mathcal{L}_1(\vec{\xi}), \ldots, \mathcal{L}_N(\vec{\xi})$, a sparse grid approximation of the temperature field depending on the image domain of the random vector is given by

$$q(\vec{x}, \vec{\xi}) \approx \hat{q}(\vec{x}, \vec{\xi}) = \sum_{n=1}^{N} \mathcal{L}_n(\vec{\xi}) q(\vec{x}, \vec{\xi}_n). \tag{7.18}$$

It can be deduced from (7.8) and (7.18) that the Nusselt number computed from the sparse grid approximation of the temperature field is given by

$$Nu(q(\vec{x}, \vec{\xi})) \approx Nu(\hat{q}(\vec{x}, \vec{\xi})) = \sum_{n=1}^{N} \mathcal{L}_n(\vec{\xi}) Nu(q(\vec{x}, \vec{\xi}_n)). \tag{7.19}$$

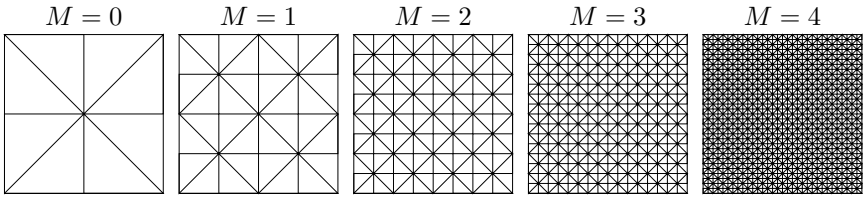

Figure 7.3: Uniformly refined meshes.

In order to instantiate the sparse grid interpolation it is necessary to compute $q(\vec{x}, \vec{\xi}_1), \ldots, q(\vec{x}, \vec{\xi}_N)$, which amounts to solving (7.11)–(7.17) for fixed positions $\vec{\xi}_1, \ldots, \vec{\xi}_N$ in Ξ. This transfers the $(2 + K)$-dimensional problem to a set of N independent two-dimensional problems, which can be solved with the finite element method. After $Nu(q(\vec{x}, \vec{\xi}_1), \ldots, Nu(q(\vec{x}, \vec{\xi}_N)$ have been computed and stored, the Nusselt number can be evaluated quickly for any given realization of the random parameter vector $\vec{\xi}(\theta)$ by using the sparse interpolation (7.19). Furthermore, integrals over the stochastic domain can be evaluated via exact integration of the Lagrangian polynomials.

7.4 Finite element model

The finite element simulation of the deterministic problems at the collocation points is based on the finite element model described section 5.2. The main difference is the absence of a time derivative in the present case. The space discretization is based on a regular triangular mesh to which uniform refinement is applied to study the mesh dependent convergence used for the numerical simulation. In this context, the variable M is used to denote the refinement level. The mesh refinement is sketched in Figure 7.3.

Following the usual finite element modeling steps, at first a weak form of the Boussinesq problem is derived: Find $p \in L^2$, $\vec{u} \in H^1_{\vec{u},0}$, $q - q_{\mathrm{D}} \in H^1_{q,0}$, for which

$$
\begin{aligned}
&(\vec{u} \cdot \nabla \vec{u}, \vec{\psi^u}) + (\nabla \vec{u} + (\nabla \vec{u})^T, \nabla \vec{\psi^u}) \\
&\qquad -(p, \nabla \cdot \vec{\psi^u}) + (5000 \vec{g} q, \vec{\psi^u}) = 0 \qquad \forall \vec{\psi^u} \in H^1_{\vec{u},0}, \qquad (7.20)\\
&\qquad (\vec{u} \cdot \nabla q, \psi^q) + (\nabla q, \nabla \psi^q) = 0 \qquad \forall \psi^q \in H^1_{q,0}, \qquad (7.21)\\
&\qquad (\nabla \cdot \vec{u}, \psi^p) = 0 \qquad \forall \psi^p \in L^2. \qquad (7.22)
\end{aligned}
$$

The unknown fields are discretized in space with Taylor-Hood finite elements. In particular continuous, piecewise linear basis functions are used for the pressure and continuous, piecewise quadratic basis functions for all other components of the solution. The respective finite element spaces are denoted by $L^2_{p,h} \subset L^2$, $H^1_{\vec{u},0,h} \subset H^1_{\vec{u},0}$ and $H^1_{q,0,h} \subset H^1_{q,0}$. The velocity boundary is a pure Dirichlet boundary, which leaves only pressure derivatives in the equations and, thus, the pressure field is determined up to an arbitrary constant. To determine a unique pressure field numerically, the value of the pressure is set to a fixed value at one point in the domain. This can be implemented on the discrete level and is not detailed, here. A discrete continuous extension q^h_{D} of the temperature Dirichlet data is defined via

$$q^h_{\mathrm{D}}(\vec{x}, \vec{\xi}) = f^h_0(\vec{x}) + \sum_{i=1}^{K} \sqrt{\lambda_i} \xi^i f^h_i(\vec{x}), \tag{7.23}$$

where $f^h_1, \ldots, f^h_K \in H^1_{q,h}$ are finite element continuous extensions of $f_1, \ldots, f_K$, and $f^h_0 \in H^1_{q,h}$ is used to implement the deterministic component of the boundary condition.

Substituting the finite element spaces in (7.20)–(7.22) and replacing the unknowns by their finite element representations leads to the following discrete weak form: Find $p^h \in L^2_{p,h}$, $\vec{u}^h \in H^1_{\vec{u},0,h}$, $q^h - q^h_{\mathrm{D}} \in H^1_{q,0,h}$, for which

$$\begin{aligned}
&(\vec{u}^h \cdot \nabla \vec{u}^h, \vec{\psi^u}) + (\nabla \vec{u}^h + (\nabla \vec{u}^h)^T, \nabla \vec{\psi^u}) \\
&\qquad -(p^h, \nabla \cdot \vec{\psi^u}) + (5000 \vec{g} q^h, \vec{\psi^u}) = 0 \quad \forall \vec{\psi^u} \in H^1_{\vec{u},0,h}, && (7.24)\\
&\qquad (\vec{u}^h \cdot \nabla q^h, \psi^q) + (\nabla q^h, \nabla \psi^q) = 0 \quad \forall \psi^q \in H^1_{q,h}, && (7.25)\\
&\qquad (\nabla \cdot \vec{u}^h, \psi^p) = 0 \quad \forall \psi^p \in L^2_{p,h}. && (7.26)
\end{aligned}$$

Using the finite element basis functions as test functions leads to the system of non-linear algebraic equations

$$\begin{aligned}
\mathcal{N}^{\mathrm{F}}_{\vec{u}}(\vec{U})\vec{U} + \mathcal{C}^{\mathrm{F}}_{\vec{u}}\, P + \mathcal{K}^{\mathrm{F}}_{\vec{u}}\, \vec{U} + \mathcal{B}^{\mathrm{F}}_{\vec{u}} Q &= \vec{0},\\
\mathcal{N}^{\mathrm{F}}_{q}(\vec{U})Q + \mathcal{K}^{\mathrm{F}}_{q}\, Q &= 0,\\
\mathcal{D}^{\mathrm{F}}_{\vec{u}}\, \vec{U} &= 0.
\end{aligned}$$

The vectors P, $\vec{U}$ and Q contain the nodal values of the finite element solution. The discretized differential operators and suitable discrete Dirichlet boundary conditions are defined in subsection 5.2.1. The trust-region dogleg method (Powell, 1970) of the Matlab® R2012a Optimization Toolbox is used to solve the system of equations. A few pseudo-time-steps with the backward Euler

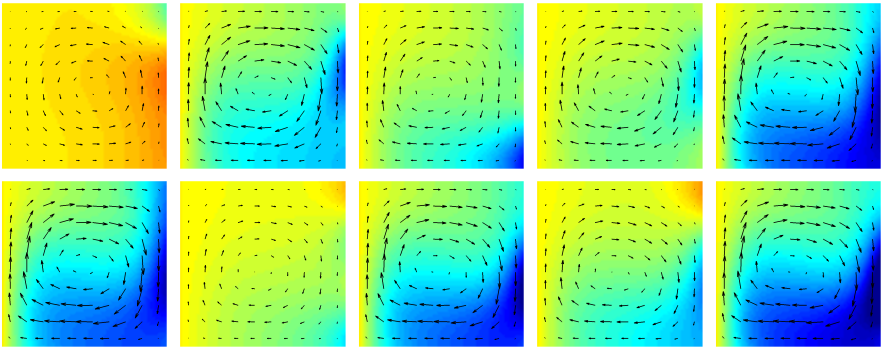

Figure 7.4: Temperature fields (colors) and velocity fields (vectors) resulting from random simulations with $K = 4$ and $M = 4$. All plots use the same color and vector scalings.

method are sufficient to obtain initial values for which the solver converges. The cost of the finite element solution is determined by the LU-decomposition of the Jacobian matrix.

A few solutions with random boundary conditions and $K = 4$ are shown in Fig. 7.4. Notice the absence of thin boundary layers, which justifies the choice of a uniform grid.

7.5 Reduced-order model

A considerable amount of computational work may be needed for the setup of the sparse grid interpolation if the discretizations of the stochastic and the physical space are sufficiently fine. A possible remedy is the use of reduced-order models that are obtained by a Galerkin projection of the deterministic equations on the space spanned by a set of POD basis functions. Just like the sparse grid collocation method, the considered reduced-order models provide a computationally inexpensive mapping from a realization of the random vector to the respective Nusselt number. While the sparse grid collocation is based purely on polynomial interpolations, POD-Galerkin models include knowledge about the governing equations. In the following, the POD approximations of the velocity and the temperature field are introduced and the reduced-order models are formulated.

At first the velocity field is considered. Suppose finite element solutions $\vec{u}_1^h, \ldots, \vec{u}_N^h \in H^1_{\vec{u},0,h}$ at the respective collocation points $\vec{\xi}_1, \ldots, \vec{\xi}_N$ are available. A POD of the snapshots provides POD basis functions $\vec{\phi}_1^{\vec{u}}, \ldots, \vec{\phi}_{R'^{\vec{u}}}^{\vec{u}} \in H^1_{\vec{u},0,h}$. A POD representation $\vec{u}^R \in H^1_{\vec{u},0,R}$ of the velocity field is given by

$$\vec{u}^R(\vec{x}, \vec{\xi}) = \sum_{i=1}^{R'^{\vec{u}}} \vec{\phi}_i^{\vec{u}}(\vec{x}) a^i(\vec{\xi}), \tag{7.27}$$

where $H^1_{\vec{u},0,R} \subset H^1_{\vec{u},0,h}$ is the space spanned by the first $R^{\vec{u}}$ velocity POD basis functions and $R'^{\vec{u}} = R^{\vec{u}}$. The POD basis functions and, consequently, the POD representation of the velocity field fulfill homogeneous Dirichlet conditions and are discretely divergence-free by construction, as they are linear combinations of snapshots.

For the reduced-order approximation of the temperature field, the performance of the control function method and the modified basis function method is compared. Let the finite element temperature solutions for the collocation points $\vec{\xi}_1, \ldots, \vec{\xi}_N$ be given by $q_1^h, \ldots, q_N^h$. Each method is able to generate functions $\phi_1^q, \ldots, \phi_{R^q}^q \in H^1_{q,0,h}$ and $\phi_{R^q+1}^q, \ldots, \phi_{R'^q}^q \in H^1_{q,h}$ with $R'^q = R^q + 1 + K$. Let a general reduced-order representation of the temperature field be given by

$$q^R(\vec{x}, \vec{\xi}) = \sum_{i=1}^{R'^q} \phi_i^q(\vec{x}) b^i(\vec{\xi}), \qquad \vec{\xi} \in \Xi, \qquad \vec{x} \in \bar{\Omega}, \tag{7.28}$$

with coefficient functions $b^1, \ldots, b^{R'^q} : \Xi \to \mathbb{R}$. It is possible to find coefficients $\bar{b}^{R+1}, \ldots, \bar{b}^{R'^q} : \Xi \to \mathbb{R}$ so that the linear combination

$$q_{\mathrm{D}}^R(\vec{x}, \vec{\xi}) = \sum_{i=R^q+1}^{R'^q} \phi_i^q(\vec{x}) \bar{b}^i(\vec{\xi}), \qquad \vec{\xi} \in \Xi, \qquad \vec{x} \in \bar{\Omega} \tag{7.29}$$

results in $q_{\mathrm{D}}^h - q_{\mathrm{D}}^R \in H^1_{q,0,h}$ for any $\vec{\xi} \in \Xi$. If $b^r(\vec{\xi}) = \bar{b}^r(\vec{\xi})$ for $r = R^q+1, \ldots, R'^q$ and $\vec{\xi} \in \Xi$ is used in (7.28), then q^R fulfills the discrete Dirichlet conditions of the snapshot simulation. In case of the control function method it is possible, for example, to use

$$\begin{aligned} \phi_{R^q+1}^q &= f_0^h, & b^{R^q+1} &= 1, \\ \phi_{R^q+1+k}^q &= \sqrt{\lambda_k} f_k^h, & b^{R^q+1+k} &= \xi^k, & k &= 1, \ldots, K. \end{aligned}$$

Fig. 7.5 shows the functions $\phi_1^q, \ldots, \phi_5^q$ and $\phi_{R^q+1}^q, \ldots, \phi_{R^q+5}^q$ resulting from the control function method applied to the temperature snapshots at

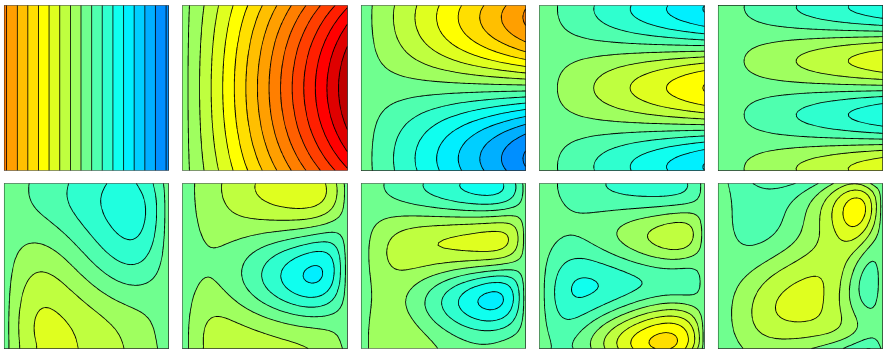

Figure 7.5: Control function method based on functions which are linear in x. Top row: $\phi^q_{R^q+1}, \ldots, \phi^q_{R^q+5}$. Bottom row: $\phi^q_1, \ldots, \phi^q_5$, obtained from the temperature snapshots at the collocation points of $K = 4$ and $L = 6$. The color scales are the same within the top row and within the bottom row.

the collocation points of $K = 4$ and $L = 6$. The functions $\phi^q_{R^q+1}, \ldots, \phi^q_{R^q+5}$ (top row) were chosen a priori with respect to the Karhunen-Loève expansion of the boundary condition and with linearity in x-direction. The functions $\phi^q_1, \ldots, \phi^q_5$ fulfill homogeneous Dirichlet conditions.

Fig. 7.6 shows the functions $\phi^q_1, \ldots, \phi^q_{10}$ resulting from a POD directly applied to the temperature snapshots at the collocation points of $K = 4$ and $L = 6$. The functions are linear combinations of the snapshots. Therefore, $\phi^q_1, \ldots, \phi^q_{10}$ are constant at the left Dirichlet boundary and equal to a constant plus a linear combination of Karhunen-Loève functions at the right boundary.

Fig. 7.7 shows the functions $\phi^q_1, \ldots, \phi^q_{10}$ resulting from a modified POD based on the temperature snapshots at the collocation points of $K = 4$ and $L = 6$. The functions of Fig. 7.7 are computed by linearly combining the functions of Fig. 7.6. While the functions in the top row of Fig. 7.7 fulfill inhomogeneous boundary conditions, the functions in the bottom row fulfill homogeneous Dirichlet conditions.

Having suitable reduced-order subspaces available, the discretized weak form (7.24)–(7.26) can be used to derive the reduced-order model. At first, $\vec{u}^h$ is substituted by $\vec{u}^R$ and q^h by q^R. The discrete continuity equation is automatically fulfilled for $\vec{u}^R$ and, thus, can be neglected. The spaces

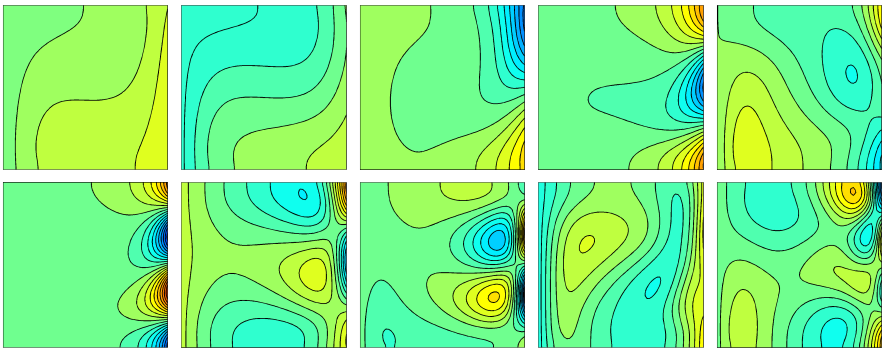

Figure 7.6: POD basis functions of the temperature snapshots at the collocation points of $K = 4$ and $L = 6$. Top row: $\phi_1^q, \ldots, \phi_5^q$. Bottom row: $\phi_6^q, \ldots, \phi_{10}^q$. The color scales are the same as in the bottom row of Fig. 7.5.

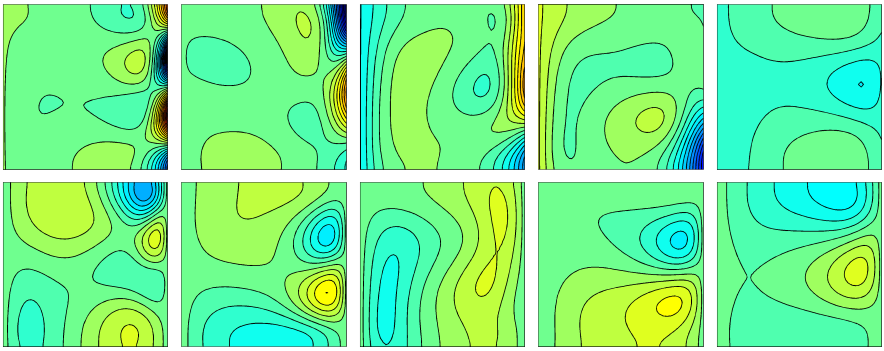

Figure 7.7: Modified basis function method for temperature snapshots at the collocation points of $K = 4$ and $L = 6$. Top row: $\phi_6^q, \ldots, \phi_{10}^q$. Bottom row: $\phi_1^q, \ldots, \phi_5^q$. The color scales are the same as in the bottom row of Fig. 7.5

$H^1_{\vec{u},0,R}$ and $H^1_{q,0,R}$ are chosen as test spaces for the remaining equations. The pressure term becomes zero for all discrete pressures $p^h \in L^2_{p,h}$. This leaves the following reduced-order weak formulation: Find $\vec{u}^R \in H^1_{\vec{u},0,R}$ and $q^R - q^h_{\mathrm{D}} \in H^1_{q,0,R}$, for which

$$\begin{aligned}(\vec{u}^R \cdot \nabla \vec{u}^R, \vec{\phi^u}) + (\nabla \vec{u}^R + (\nabla \vec{u}^R)^T, \nabla \vec{\phi^u}) \\ +(5000\vec{g}q^R, \vec{\phi^u}) = 0 \qquad \forall \vec{\phi^u} \in H^1_{\vec{u},0,R}, \qquad (7.30)\\ (\vec{u}^R \cdot \nabla q^R, \phi^q) + (\nabla q^R, \nabla \phi^q) = 0 \qquad \forall \phi^q \in H^1_{q,0,R}, \qquad (7.31)\end{aligned}$$

After testing against the functions $\vec{\phi}^{\vec{u}}_1, \ldots, \vec{\phi}^{\vec{u}}_{R^{\vec{u}}}$ and $\phi^q_1, \ldots, \phi^q_{R^q}$, respectively, the reduced-order model can be written as

$$\sum_{i=1}^{R'^{\vec{u}}} \sum_{j=1}^{R'^{\vec{u}}} \mathcal{N}^{a,\mathrm{F}}_{rij} a^i a^j + \sum_{i=1}^{R'^{\vec{u}}} \mathcal{K}^{a,\mathrm{F}}_{ri} a^i + \sum_{i=1}^{R'^q} \mathcal{B}^{a,\mathrm{F}}_{ri} b^i = 0, \qquad r = 1, \ldots, R^{\vec{u}}, \qquad (7.32)$$

$$\sum_{i=1}^{R'^{\vec{u}}} \sum_{j=1}^{R'^q} \mathcal{N}^{b,\mathrm{F}}_{rij} a^i b^j + \sum_{i=1}^{R'^q} \mathcal{K}^{b,\mathrm{F}}_{ri} b^i = 0, \qquad r = 1, \ldots, R^q. \qquad (7.33)$$

The unknowns of the equations are $a^1, \ldots, a^{R^{\vec{u}}}$ and $b^1, \ldots, b^{R^{\vec{u}}}$. The values of $b^{R^q+1}, \ldots, b^{R'^q}$ must be provided as input data. The model coefficients are defined in subsection 5.3.3. The reduced-order model is solved by the same iterative non-linear solver as the finite element model. The cost of performing one iteration is determined by the evaluation of the quadratic terms and by the LU-decomposition of the Jacobian matrix.

To obtain a reduced-order equation for the Nusselt number, (7.28) is substituted in (7.8), so that

$$Nu(q^R) = \sum_{i=1}^{R'^q} Nu(\phi^q_i) b^i,$$

where $Nu(\phi^q_r)$ can be evaluated and stored once for each $i = 1, \ldots, R'^q$. Hence, the Nusselt number can be directly computed from any solution of the reduced-order model, without the need to form the temperature field at all.

For the test performed later on, it is important that the reduced-order model can be evaluated for input vectors $\vec{\xi}$ whose length may be incompatible with the dimensions of the reduced-order model. This case occurs, for example, when model evaluations are compared for different stochastic dimensions.

As a remedy, the last entries of $\vec{\xi}$ are removed when $\vec{\xi}$ is too long and zeros are appended when $\vec{\xi}$ is too short.

Now the reduced-order model is ready to be used in the context of the stochastic collocation method via the POD-aided stochastic collocation approach introduced in section 6.4. The basic idea is to replace finite element evaluations by evaluations of the POD reduced-order model for the setup of the sparse-grid interpolation. This approach is different from replacing the interpolation by a reduced-order model.

7.6 Statistics of the Nusselt number

Three aspects are considered in the comparison of the POD-aided stochastic collocation method with the standard stochastic collocation: Firstly, the approximation properties of the respective sparse-grid interpolation of Nu throughout the parameter domain are assessed. Secondly, the probability density function of the Nusselt number is considered, which is computed via a Monte-Carlo simulation using random evaluations of the sparse-grid interpolation. Thirdly, the mean and variance of the Nusselt number are investigated using exact integrations of the underlying Lagrangian polynomials. By default, the control function method with linear functions in x is used to implement the Dirichlet conditions.

In Table 7.1, the total and average computation times are presented for the finite element model (FEM), the reduced-order model (ROM) and the sparse-grid interpolation (SG). The times needed for the creation of the models are displayed alongside. The computation time of the sparse-grid is tested in one instance for the approximation of the probability density function via Monte Carlo sampling with 1 000 000 evaluations of the sparse-grid interpolation at random points (see Sect. 7.6.2). In another instance, the computation of the mean via an exact integration of the multivariate Lagrangian polynomial is considered (see Sect. 7.6.3). All computations are performed with Matlab®. The timings are measured on a single computational thread on an Intel® Xeon® E5-2670 CPU, although the Monte Carlo sampling as well as the independent finite element and reduced-order simulations are well suited for parallel processing.

7.6.1 Sparse grid approximation

The goal of this section is a sparse grid representation of the Nusselt number which is valid throughout Ξ. The results from a standard stochastic colloca-

Table 7.1: Total and average computation times and relevant parameters for the finite element simulation (FEM), the reduced-order simulation (ROM) and the evaluation of the sparse grid (SG).

	K	L	N	M	R	total	average
FEM setup				4		0.46 s	
FEM simulation	4	6	2929	4		13245.52 s	4.5222 s
FEM simulation	4	3	137	4		602.46 s	4.3975 s
ROM setup	4	3	137	4	30	1.67 s	
ROM simulation	4	6	2929	4	30	63.57 s	0.0217 s
SG Nusselt density	4	6	2929			3053.72 s	0.0031 s
SG Nusselt mean	4	6	2929			0.27 s	

tion are compared to the results from a POD-aided stochastic collocation.

Suppose the numerical solutions have been computed with a finite element mesh refinement level of M at the collocation points of the stochastic dimension K and the Smolyak level L. At first, a sparse grid interpolation is created from the available numerical solutions using the standard stochastic collocation. Then, the available numerical solutions are used to build a POD-Galerkin reduced-order model with $R = R^q = R^{\vec{u}}$ basis functions. The reduced-order model is evaluated at the collocation points necessary to build a sparse grid interpolation of stochastic dimension $K^\star = 4$ and sparse grid refinement level $L^\star = 6$.

The standard collocation is to be compared with the POD-aided collocation for varying discretization parameters. The finite element mesh refinement level M is varied from 1 to 4, the stochastic dimension K from 0 to 4, the Smolyak level L from 0 to 6 and, for the reduced-order model, the numbers of POD basis functions $R = 5, 10, 15, 20, 25, 30$ are compared. To estimate the errors of the Nusselt number, a reference solution is used, which is obtained by finite element computations with $M = 5$ at the collocation points of $K = 6$ and $L = 8$. For each of these collocation points the difference between the reference Nusselt number and interpolated Nusselt number can be computed. The maximum absolute difference is taken as an error estimate. A comparison of the estimated errors of the standard and the POD-aided stochastic collocation is given in Fig. 7.8.

The first error plot of Fig. 7.8 presents the case where the stochastic dimen-

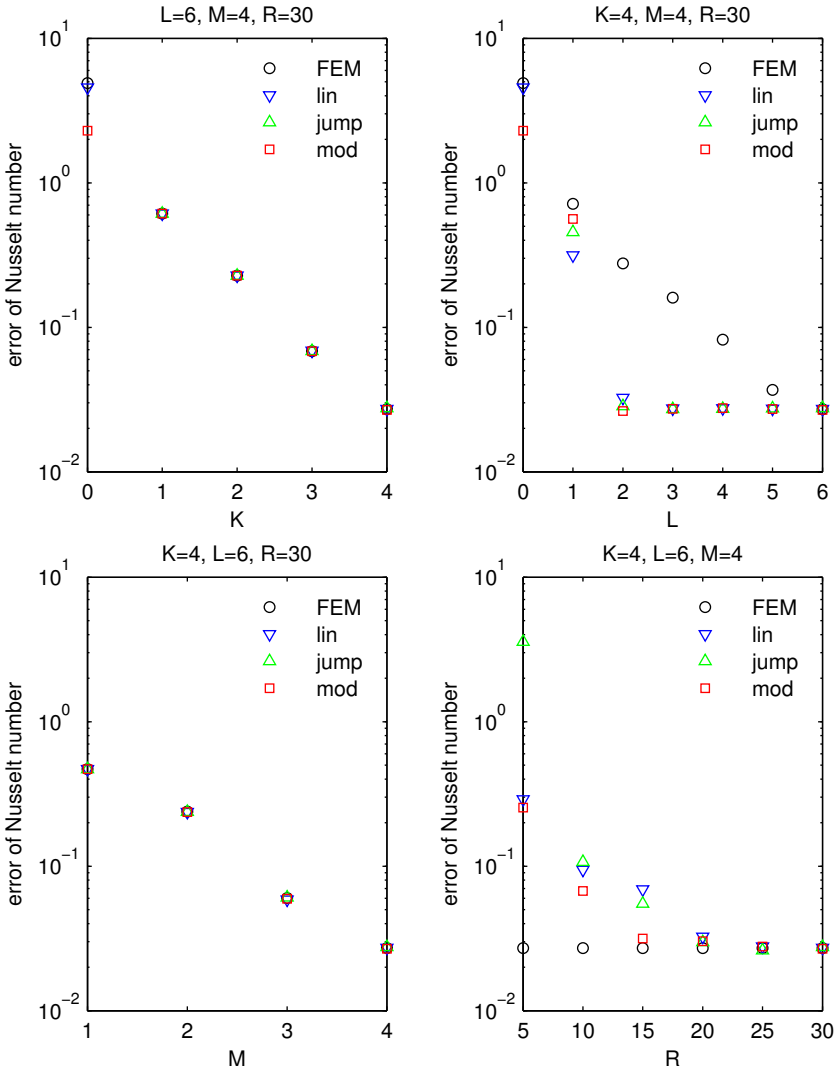

Figure 7.8: Estimates of the maximum absolute error of the Nusselt number throughout the stochastic domain for the sparse grid interpolation (FEM) and the POD-aided sparse grid interpolation based on the control function method with functions which are linear in x (lin), functions which are zero at the interior points (jump), and the modified basis function method (mod).

sion is varied. In particular, a sparse-grid interpolation is built from the finite element snapshots at the collocation points of a sparse grid of dimension K, where $K = 0$ is equivalent to the deterministic case. The error with respect to a reference solution is denoted by black circles. By increasing the dimensions, the error can be decreased. The finite element snapshots corresponding to the dimension K are also used to build reduced-order models using different approaches to implement the Dirichlet conditions. The reduced-order models are used to cheaply obtain an interpolation at a sparse grid of a high dimension. As can be seen in the plot, this does not lead to an improvement over using the snapshots directly.

In the second plot of Fig. 7.8 the error is presented for varying Smolyak level L, where $L = 0$ is equivalent to a fully deterministic simulation. It can be observed that the error of the POD-aided collocation method drops much faster than the error of the standard method when the level is increased. Therefore, by using finite element snapshots at the collocation points of a low level, e.g. $L = 2$, it is possible via the reduced-order model to create a sparse-grid interpolation that is about as accurate as a standard collocation with $L = 5$ or $L = 6$. It seems that the advantage of the POD-aided collocation does not significantly depend on which method is used to implement the Dirichlet conditions.

In Fig. 7.8, the third plot demonstrates the accuracy with respect to the mesh refinement level M. With this plot it is ensured that the resolution of the finite element mesh is adequate. In view of computation times, an over-resolving finite element mesh would give an advantage to the reduced-order model. One can see that by decreasing the mesh resolution from $M = 4$ to $M = 3$ the error of the Nusselt increases significantly, so the finite element mesh with $M = 4$ is not too fine.

The last plot of Fig. 7.8 shows the dependence of the POD-aided stochastic collocation on the number of basis functions. The results of the standard collocation method, which do not depend on R, are shown for comparison. It can be observed that a reduced-order model with 25 to 30 basis functions is sufficient to build a sparse grid whose error is very close to the error of the sparse grid obtained directly from the finite element snapshots. Moreover, the error of the modified basis function method decreases faster than the error of the control function methods which indicates a slight advantage of the former method in this setting. For small R, the control function method does not perform well when used with functions which are zero at all non-Dirichlet points. This can be explained by the fact that the respective control functions and the basis functions contain steep gradients at the boundary, while the exact solution is smooth. A larger number of functions is necessary

to represent a smooth solution close to the boundary, which is essential for computing the Nusselt number.

It can be observed that choosing the refinement parameters $K = 4$, $L = 6$ and $M = 4$ balances the errors introduced by the Karhunen-Loève expansion, the sparse-grid interpolation and the finite element simulation, respectively. The error introduced by a reduced-order model with 30 basis functions is small enough as to be dominated by the other error components. Therefore, this choice of parameters is used for further investigations concerning statistical quantities.

7.6.2 Probability density

In this section, the applicability of the POD-aided collocation method to the computation of the probability density function of the Nusselt number via the Monte Carlo method is assessed. Based on the observations of subsection 7.6.1, finite element snapshots of $K = 4$ and $L = 2$ are used to create a reduced-order model with 30 POD basis functions, employing the control function approach with linear functions in x. The reduced-order model is evaluated at $K^* = 4$ and $L^* = 6$ to create a respective sparse grid interpolation. For the standard collocation, an interpolation is created directly from the finite element solutions of $K = 4$ and $L = 6$. All finite element computations are performed on a mesh with refinement level $M = 4$. Figure 7.9 shows the probability density functions approximated with a Monte Carlo simulation using 1 000 000 samples and a bin width of 0.1. It can be observed that the probability density function computed with the POD-aided collocation is visually indistinguishable from the one obtained with the standard method.

The computation time required to evaluate the stochastic interpolation at a given point in the stochastic domain depends on the number of collocation points. To see whether a smaller number of collocation points is sufficient for an accurate computation of the probability density function, the results of the POD-aided stochastic collocation method for different stochastic dimensions $K^\star$ and Smolyak levels $L^\star$ are compared in Fig. 7.10, again using the control function approach with linear functions in x. Snapshots of $M = 4$, $K = 4$ and $L = 2$ are used to build a reduced-order model with $R = 30$ and sparse grid interpolations of varying $L^\star$ and $K^\star$ are created. For the case of varying stochastic dimension, at least $K^\star = 2$ is needed to capture the correct slope at $Nu > 0.5$ and at least $K^\star = 3$ is needed to capture the peak. For the case of varying Smolyak level, at least $L^\star = 5$ is needed to approximate the peak correctly.

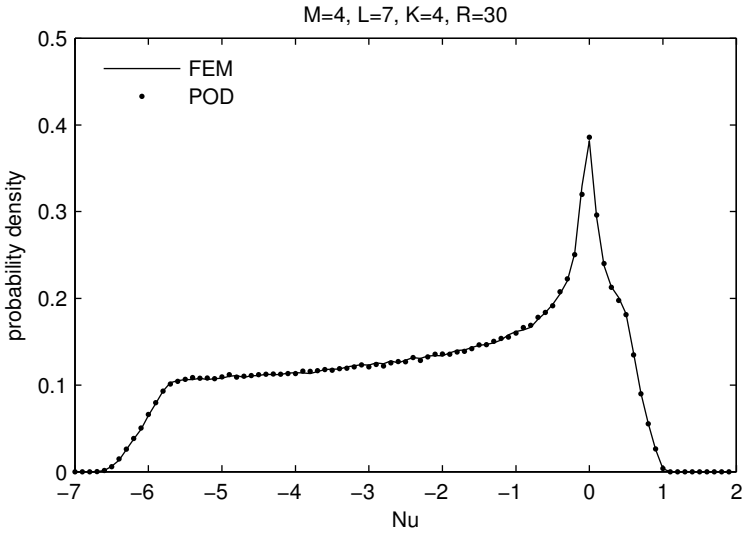

Figure 7.9: Probability density function of the Nusselt number obtained from a Monte Carlo simulation. The line represents the results obtained directly from the finite element simulations with $K = 4$ and $L = 6$ while the dots present the results obtained from a reduced-order model with $R = 30$ that was created from finite element solutions at the collocation points of $K = 4$ and $L = 2$ and evaluated at $K^\star = 4$ and $L^\star = 6$.

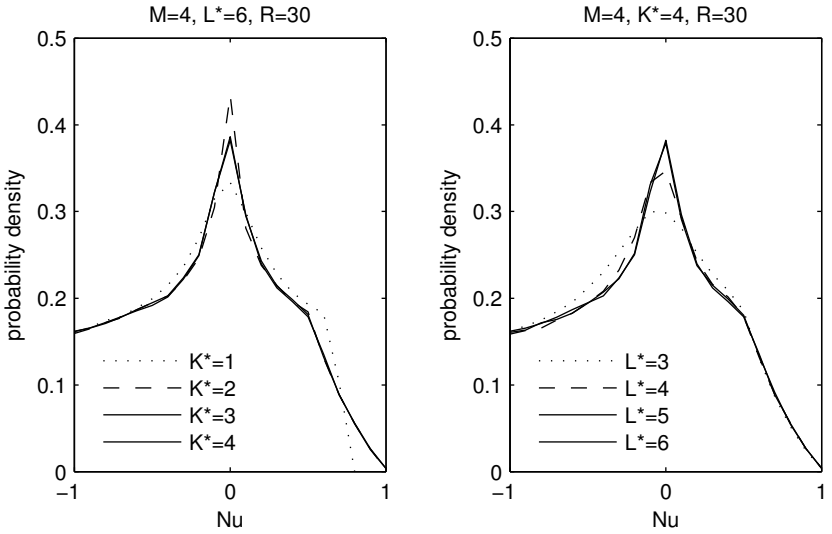

Figure 7.10: Probability density functions of the Nusselt number obtained from a Monte Carlo simulation using a POD-aided stochastic interpolation of varying $K^\star$ and $L^\star$ built from snapshots at the collocation points of $K = 4$ and $L = 2$.

7.6.3 Mean and variance

Often one is not interested in the detailed probability density function of a stochastic quantity, but rather in statistics like the mean and variance. An exact integration of the multivariate Lagrangian polynomials can be used to compute the moments of the Nusselt number. The computation of the integrals is based on one-dimensional Clenshaw-Curtis quadrature weights (Clenshaw and Curtis, 1960), which can be efficiently computed with a routine based on an inverse FFT (Waldvogel, 2006). To obtain the second moments, the square of the Nusselt number is computed at the collocation points and the same quadrature weights as in the case of the mean Nusselt number are used subsequently.

The resulting plots of the mean and variance of the Nusselt number depending on the Smolyak level L are presented in Fig. 7.11 for the standard collocation. The other parameters are set to $K = 4$, $M = 6$ and $R = 50$. Snapshots of level L are used to build a reduced-order model, which is subsequently evaluated at $K^\star = 4$ and $L^\star = 6$ to build the sparse-grid interpolation. Reference values of the mean and variance obtained with $K = 4$, $L = 6$ and $M = 7$ are used to estimate the errors. Note that with this choice, at $L = 6$ the error estimator does not capture the effects of K and L anymore. The second plot of Fig. 7.11 contains two values which are significantly lower than the finite element error. They are most probably a result of errors which canceled by chance, and are therefore not representative. For the computation of the mean Nusselt number it can be observed that the sparse grid refinement level can be reduced by 1 or 2 when POD-Galerkin modeling is employed.

7.7 Conclusions

The applicability of POD-Galerkin reduced-order modeling to the context of stochastic collocation on sparse grids was assessed. For the test case of the computation of the Nusselt number in a natural convective flow setting, the behavior of the different error contributions was studied in a systematic way with the help of a reference solution. By employing a reduced-order model, the number of finite element simulations necessary to create an accurate sparse grid interpolation could be significantly decreased. For example (see Fig. 7.8), by building a reduced-order model from the finite element solutions at the 137 collocation points of Smolyak level $L = 3$ and evaluating the model at the 2929 collocation points of Smolyak level $L^\star = 6$, it was possible to obtain a sparse grid interpolation with a similar accuracy as an interpolation created directly

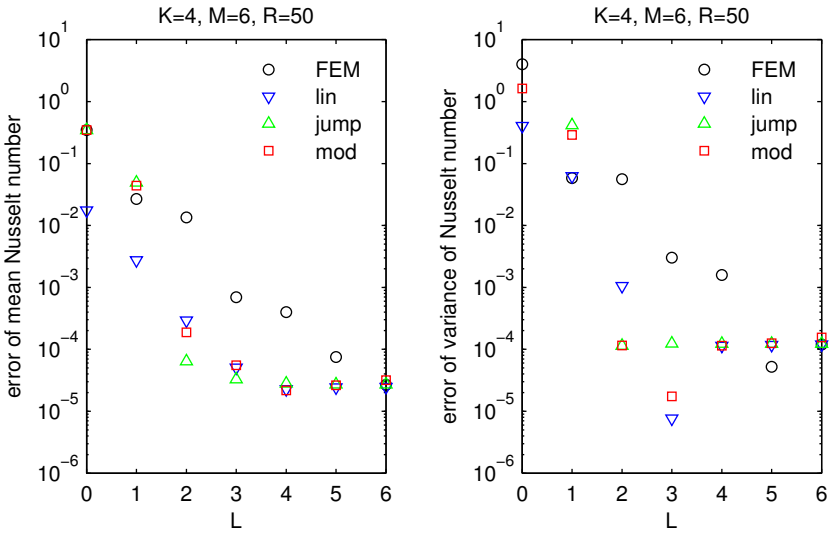

Figure 7.11: Absolute errors of the mean (left) and variance (right) of the Nusselt number depending on the Smolyak level L for the sparse grid interpolation (FEM) and the POD-aided sparse grid interpolation based on the control function method with functions which are linear in x (lin), functions which are zero at the interior points (jump), and the modified basis function method (mod). The errors are obtained with respect to reference solutions with $K = 4$, $L = 6$ and $M = 7$.

from 2929 finite element solutions of level $L = 6$.

To assess the efficiency of the method, it is necessary to inspect the computation times (see Table 7.1): It took about 3 h 41 min to perform the 2929 finite element computations of level $L = 6$ and about 10 min to perform the 137 finite element computations of level $L = 3$. In comparison, 2929 evaluations of the reduced-order model took only 64 s. The additional computational overhead of the reduced-order model, i.e. the creation of the basis functions from the snapshots and the assembly of the system matrices, was less than 2 s. Therefore, the total computation could be decreased by almost a factor of 20 by using reduced-order modeling. The resulting sparse grids had the same number of collocation points, independently of whether they were created using finite element solutions or reduced-order solutions, and so their evaluation times were identical.

Chapter 8

Summary and outlook

POD-Galerkin reduced-order models have been developed and validated for deterministic unsteady problems of increasing complexity, from heat conduction to thermally driven incompressible flow. Alternative derivations of the models have been presented, starting either from a weak formulation or from a spatial semi-discretization, or from a full space-time discretization. For each problem, the different derivations result in the same reduced-order model as long as the time and space dependencies are handled with the discretization schemes of the snapshot simulations.

In the derivation of the reduced-order models for incompressible flows, the discrete divergence-freeness of the velocity POD basis leads to a cancellation of the pressure term and an immediate fulfillment of the continuity equation. With the aid of an auxiliary pressure reduced-order model, pressure-dependent quantities could be computed from the velocity reduced-order coefficients. The pressure model combines the ideas of Akhtar et al. (2009), who suggested a Galerkin projection of a pressure Poisson equation onto a pressure POD basis, with the notion of a discrete pressure Poisson equation according to Gresho and Sani (2000). Although a pressure Poisson equation is absent in the underlying finite element model, the reduced-order pressure converges toward the finite element solution when the dimension of the pressure POD basis is increased.

A strategy has been proposed for the computation of statistical quantities for PDE problems with uncertain data. The procedure combines sparse-grid stochastic collocation with POD-Galerkin reduced-order modeling. At the collocation points of a sparse grid in the stochastic domain, computationally *expensive* evaluations of a finite element model are replaced by computa-

tionally *inexpensive* evaluations of a reduced-order model. Thereafter, the statistics of an integral output of interest are computed using the sparse grid interpolant. The method was demonstrated for the thermally driven flow in a cavity with uncertain temperature Dirichlet conditions. Although the creation of the POD basis required multiple evaluations of the finite element model, a total saving of computational time could be observed.

In order to implement Dirichlet conditions given by parametrized random fields in the reduced-order models, the method of Gunzburger et al. (2007) was generalized and termed *modified basis function method*. As an advantage over the conventional *control function method*, the modified basis function method does not require a predefined lifting function. The method actually provides a suitable lifting function automatically by means of a linear combination of snapshots. When applied to the stochastic thermally driven flow problem, the new method performed similarly compared to the control function method equipped with reasonable lifting functions.

All presented reduced-order models rely on a fixed spatial mesh. While, in principle, geometric variations can be incorporated in the current framework via a transformation to some reference domain, it is not directly possible to utilize snapshots with different numbers of mesh nodes, for instance. Because many applications benefit from adaptive meshing, POD-Galerkin reduced-order modeling based on snapshots from different spatial meshes provides a promising future research direction.

In the presented POD-aided sparse grid collocation method, the snapshots are sampled at predefined collocation points in the stochastic domain. Perhaps, better snapshot locations can be found by applying a greedy method, which enriches the snapshot set iteratively by employing a reduced-order error estimator. The derivation of efficiently computable and sharp error estimators for incompressible flows problems under uncertainty is a subject of current research.

The POD-aided stochastic collocation method presented in this thesis differs from the approach taken by Schieche (2012a,b): The method of Schieche combines a finite element primal model with a reduced-order adjoint model. While adaptivity is used to save computational time in the construction of the sparse grid interpolant, reduced-order modeling is used to alleviate the computational overhead resulting from the adaptivity. In contrast, the method of the present thesis saves computational time in the construction of the sparse grid interpolant by employing reduced-order modeling directly. A combination of both methods promises to further accelerate the computation of statistical quantities while at the same time providing an a posteriori error estimator.

Appendix A

Coefficients of the velocity reduced-order model

$$
\begin{aligned}
&(\vec{\phi}_r^{\vec{u}}, \vec{\phi}_i^{\vec{u}}) \\
&\qquad = (\phi_r^u, \phi_i^u) + (\phi_r^v, \phi_i^v) \\
&\qquad = \left(\sum_{k=1}^{M^u} \Phi_r^{u,k} \psi_k^u, \sum_{l=1}^{M'^u} \Phi_i^{u,l} \psi_l^u \right) + \left(\sum_{k=1}^{M^v} \Phi_r^{v,k} \psi_k^v, \sum_{l=1}^{M'^v} \Phi_i^{v,l} \psi_l^v \right) \\
&\qquad = \sum_{k=1}^{M^u} \sum_{l=1}^{M'^u} \Phi_r^{u,k} (\psi_k^u, \psi_l^u) \Phi_i^{u,l} + \sum_{k=1}^{M^v} \sum_{l=1}^{M'^v} \Phi_r^{v,k} (\psi_k^v, \psi_l^v) \Phi_i^{v,l} \\
&\qquad = (\Phi_r^{u,\mathrm{F}})^T \mathcal{M}_u^{\mathrm{F}} \Phi_i^u + (\Phi_r^{v,\mathrm{F}})^T \mathcal{M}_v^{\mathrm{F}} \Phi_i^v \\
&\qquad = (\vec{\Phi}_r^{\vec{u},\mathrm{F}})^T \mathcal{M}_{\vec{u}}^{\mathrm{F}} \vec{\Phi}_i^{\vec{u}},
\end{aligned}
$$

$$
\begin{aligned}
&(\vec{\phi}_r^{\vec{u}}, \vec{\phi}_i^{\vec{u}} \cdot \nabla \vec{\phi}_j^{\vec{u}}) \\
&= (\phi_r^u, \vec{\phi}_i^{\vec{u}} \cdot \nabla \phi_j^u) + (\phi_r^v, \vec{\phi}_i^{\vec{u}} \cdot \nabla \phi_j^v) \\
&= (\phi_r^u, \phi_i^u \partial_x \phi_j^u + \phi_i^v \partial_y \phi_j^u) + (\phi_r^v, \phi_i^u \partial_x \phi_j^v + \phi_i^v \partial_y \phi_j^v) \\
&= \Big(\sum_{k=1}^{M^u} \Phi_r^{u,k} \psi_k^u, \Big(\sum_{l=1}^{M'^u} \Phi_i^{u,l} \psi_l^u \Big) \partial_x \Big(\sum_{m=1}^{M'^u} \Phi_j^{u,m} \psi_m^u \Big) \\
&\quad + \Big(\sum_{l=1}^{M'^v} \Phi_i^{v,l} \psi_l^v \Big) \partial_y \Big(\sum_{m=1}^{M'^u} \Phi_j^{u,m} \psi_m^u \Big) \Big) \\
&\qquad + \Big(\sum_{k=1}^{M^v} \Phi_r^{v,k} \psi_k^v, \Big(\sum_{l=1}^{M'^u} \Phi_i^{u,l} \psi_l^u \Big) \partial_x \Big(\sum_{m=1}^{M'^v} \Phi_j^{v,m} \psi_m^v \Big) \\
&\quad + \Big(\sum_{l=1}^{M'^v} \Phi_i^{v,l} \psi_l^v \Big) \partial_y \Big(\sum_{m=1}^{M'^v} \Phi_j^{v,m} \psi_m^v \Big) \Big) \\
&= \sum_{k=1}^{M^u} \sum_{m=1}^{M'^u} \Phi_r^{u,k} \Big(\psi_k^u, \sum_{l=1}^{M'^u} \Phi_i^{u,l} \psi_l^u \partial_x \psi_m^u + \sum_{l=1}^{M'^v} \Phi_i^{v,l} \psi_l^v \partial_y \psi_m^u \Big) \Phi_j^{u,m} \\
&\qquad + \sum_{k=1}^{M^v} \sum_{m=1}^{M'^v} \Phi_r^{v,k} \Big(\psi_k^v, \sum_{l=1}^{M'^u} \Phi_i^{u,l} \psi_l^u \partial_x \psi_m^v + \sum_{l=1}^{M'^v} \Phi_i^{v,l} \psi_l^v \partial_y \psi_m^v \Big) \Phi_j^{v,m} \\
&= (\Phi_r^{u,\mathrm{F}})^T \mathcal{N}_u^{\mathrm{F}}(\vec{\Phi}_i^{\vec{u}}) \Phi_j^u + (\Phi_r^{v,\mathrm{F}})^T \mathcal{N}_u^{\mathrm{F}}(\vec{\Phi}_i^{\vec{u}}) \Phi_j^v \\
&= (\vec{\Phi}_r^{\vec{u},\mathrm{F}})^T \mathcal{N}_{\vec{u}}^{\mathrm{F}}(\vec{\Phi}_i^{\vec{u}}) \vec{\Phi}_j^{\vec{u}}.
\end{aligned}
$$

$$
\begin{aligned}
&\frac{1}{Re}(\nabla\vec{\phi}_r^{\vec{u}}, \nabla\vec{\phi}_i^{\vec{u}} + \nabla\cdot\vec{\phi}_i^{\vec{u}}) \\
&= \frac{1}{Re}(\nabla\phi_r^u, \nabla\phi_i^u + \partial_x\vec{\phi}_i^{\vec{u}}) + \frac{1}{Re}(\nabla\phi_r^v, \nabla\phi_i^v + \partial_y\vec{\phi}_i^{\vec{u}}) \\
&= \frac{2}{Re}(\partial_x\phi_r^u, \partial_x\phi_i^u) + \frac{1}{Re}(\partial_y\phi_r^u, \partial_y\phi_i^u) + \frac{1}{Re}(\partial_y\phi_r^u, \partial_x\phi_i^v) \\
&\quad + \frac{1}{Re}(\partial_x\phi_r^v, \partial_y\phi_i^u) + \frac{1}{Re}(\partial_x\phi_r^v, \partial_x\phi_i^v) + \frac{2}{Re}(\partial_y\phi_r^v, \partial_y\phi_i^v) \\
&= \frac{2}{Re}\Big(\partial_x\Big(\sum_{k=1}^{M^u}\Phi_r^{u,k}\psi_k^u\Big), \partial_x\Big(\sum_{l=1}^{M'^u}\Phi_i^{u,l}\psi_l^u\Big)\Big) \\
&\quad + \frac{1}{Re}\Big(\partial_y\Big(\sum_{k=1}^{M^u}\Phi_r^{u,k}\psi_k^u\Big), \partial_y\Big(\sum_{l=1}^{M'^u}\Phi_i^{u,l}\psi_l^u\Big)\Big) \\
&\quad + \frac{1}{Re}\Big(\partial_y\Big(\sum_{k=1}^{M^u}\Phi_r^{u,k}\psi_k^u\Big), \partial_x\Big(\sum_{l=1}^{M'^v}\Phi_i^{v,l}\psi_l^v\Big)\Big) \\
&\quad + \frac{1}{Re}\Big(\partial_x\Big(\sum_{k=1}^{M^v}\Phi_r^{v,k}\psi_k^v\Big), \partial_y\Big(\sum_{l=1}^{M'^u}\Phi_i^{u,l}\psi_l^u\Big)\Big) \\
&\quad + \frac{1}{Re}\Big(\partial_x\Big(\sum_{k=1}^{M^v}\Phi_r^{v,k}\psi_k^v\Big), \partial_x\Big(\sum_{l=1}^{M'^v}\Phi_i^{v,l}\psi_l^v\Big)\Big) \\
&\quad + \frac{2}{Re}\Big(\partial_y\Big(\sum_{k=1}^{M^v}\Phi_r^{v,k}\psi_k^v\Big), \partial_y\Big(\sum_{l=1}^{M'^v}\Phi_i^{v,l}\psi_l^v\Big)\Big) \\
&= \frac{1}{Re}\sum_{k=1}^{M^u}\sum_{l=1}^{M'^u}\Phi_r^{u,k}\Big(2(\partial_x\psi_k^u, \partial_x\psi_l^u) + (\partial_y\psi_k^u, \partial_y\psi_l^u)\Big)\Phi_i^{u,l} \\
&\quad + \frac{1}{Re}\sum_{k=1}^{M^u}\sum_{l=1}^{M'^v}\Phi_r^{u,k}(\partial_y\psi_k^u, \partial_x\psi_l^v)\Phi_i^{v,l} \\
&\quad + \frac{1}{Re}\sum_{k=1}^{M^v}\sum_{l=1}^{M'^u}\Phi_r^{v,k}(\partial_x\psi_k^v, \partial_y\psi_l^u)\Phi_i^{u,l} \\
&\quad + \frac{1}{Re}\sum_{k=1}^{M^v}\sum_{l=1}^{M'^v}\Phi_r^{v,k}\Big((\partial_x\psi_k^v, \partial_y\psi_l^v) + 2(\partial_y\psi_k^v, \partial_y\psi_l^v)\Big)\Phi_i^{v,l} \\
&= (\Phi_r^{u,\mathrm{F}})^T\mathcal{K}_{uu}^{\mathrm{F}}\Phi_i^u + (\Phi_r^{u,\mathrm{F}})^T\mathcal{K}_{uv}^{\mathrm{F}}\Phi_i^v + (\Phi_r^{v,\mathrm{F}})^T\mathcal{K}_{vu}^{\mathrm{F}}\Phi_i^u + (\Phi_r^{v,\mathrm{F}})^T\mathcal{K}_{vv}^{\mathrm{F}}\Phi_i^v \\
&= (\vec{\Phi}_r^{\vec{u},\mathrm{F}})^T\mathcal{K}_{\vec{u}}^{\mathrm{F}}\,\vec{\Phi}_i^{\vec{u}}.
\end{aligned}
$$

Bibliography

I. Akhtar, A. Nayfeh, and C. Ribbens. On the stability and extension of reduced-order Galerkin models in incompressible flows. *Theor. Comput. Fluid Dyn.*, 23(3):213–237, 2009.

I. Babuška. The finite element method with penalty. *Math. Comp.*, 27(122): 221–228, 1973.

I. Babuška, F. Nobile, and R. Tempone. A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM J. Numer. Anal.*, 45(3):1005–1034, 2007.

I. Babuška, J. R. Whiteman, and T. Strouboulis. *Finite elements. An introduction to the method and error estimation.* Oxford University Press, 2011.

J. W. Barrett and C. M. Elliott. Finite element approximation of the Dirichlet problem using the boundary penalty method. *Numer. Math.*, 49(4):343–366, 1986.

V. Barthelmann, E. Novak, and K. Ritter. High dimensional polynomial interpolation on sparse grids. *Adv. Comput. Math.*, 12(4):273–288, 2000.

M. Bergmann, L. Cordier, and J.-P. Brancher. Optimal rotary control of the cylinder wake using proper orthogonal decomposition reduced-order model. *Phys. Fluids*, 17(9):097101:1–097101:21, 2005.

S. Boyaval, C. Le Bris, Y. Maday, N. C. Nguyen, and A. T. Patera. A reduced basis approach for variational problems with stochastic parameters: application to heat conduction with variable Robin coefficient. *Comput. Methods Appl. Mech. Engrg.*, 198(41-44):3187–3206, 2009.

J. Burkardt, M. Gunzburger, and H.-C. Lee. POD and CVT-based reduced-order modeling of Navier-Stokes flows. *Comput. Methods Appl. Mech. Engrg.*, 196(1-3):337–355, 2006a.

J. Burkardt, M. Gunzburger, and H.-C. Lee. Centroidal Voronoi tessellation-based reduced-order modeling of complex systems. *SIAM J. Sci. Comput.*, 28(2):459–484, 2006b.

J. Burkardt, M. D. Gunzburger, and C. Webster. Reduced order modeling of some nonlinear stochastic partial differential equations. *Int. J. Numer. Anal. Model.*, 4(3-4):368–391, 2007.

P. Chen, A. Quarteroni, and G. Rozza. Comparison between reduced basis and stochastic collocation methods for elliptic problems. MATHICSE Report 34.2012, Mathematics Institute of Computational Science and Engineering, Ecole Polytechnique Fédérale de Lausanne, 2012.

P. Chen, A. Quarteroni, and G. Rozza. A weighted reduced basis method for elliptic partial differential equations with random input data. MATHICSE Report 04.2013, Mathematics Institute of Computational Science and Engineering, Ecole Polytechnique Fédérale de Lausanne, 2013.

C. Clenshaw and A. Curtis. A method for numerical integration on an automatic computer. *Numer. Math.*, 2(1):197–205, 1960.

L. Cordier, B. A. E. Majd, and J. Favier. Calibration of POD reduced-order models using Tikhonov regularization. *Int. J. Num. Meth Fluids*, 63(2): 269–296, 2010.

R. Courant and D. Hilbert. *Methods of mathematical physics. Vol. I.* Interscience Publishers Inc., New York, 1953.

J. Crank and P. Nicolson. A practical method for numerical evaluation of solutions of partial differential equations of the heat-conduction type. *Proc. Cambridge Philos. Soc.*, 43:50–67, 1947.

A. Deane, I. Kevrekidis, G. Karniadakis, and S. Orszag. Low-dimensional models for complex geometry flows: Application to grooved channels and circular cylinders. *Phys. Fluids A*, 3(10):2337–2354, 1991.

J. L. Eftang and E. M. Rønquist. Evaluation of flux integral outputs for the reduced basis method. *Math. Models Methods Appl. Sci.*, 20(3):351–374, 2010.

H. C. Elman and Q. Liao. Reduced basis collocation methods for partial differential equations with random coefficients. Technical Report CS-TR-5011, University of Maryland, Computer Science Department, 2012.

B. Erdmann, J. Lang, and R. Roitzsch. KARDOS user's guide. Tech. rep. ZR 02–42, Konrad-Zuse-Zentrum Berlin, 2002.

G. Evans and S. Paolucci. The thermoconvective instability of plane Poiseuille flow heated from below: a proposed benchmark solution for open boundary flows. *Int. J. Numer. Meth. Fluids*, 11(7):1001–1013, 1990.

B. Ganapathysubramanian and N. Zabaras. Sparse grid collocation schemes for stochastic natural convection problems. *J. Comput. Phys.*, 225:652–685, 2007.

R. Ghanem and P. Spanos. *Stochastic finite elements: a spectral approach.* Dover publications, 1991.

B. Gottermeier and J. Lang. Adaptive two-step peer methods for thermally coupled incompressible flow. In J. C. F. Pereira, A. Sequeira, and J. M. C. Pereira, editors, *Proceedings of the V European Conference on Computational Fluid Dynamics ECCOMAS CFD 2010*, Lisbon, Portugal, June 2010.

W. R. Graham, J. Peraire, and K. Y. Tang. Optimal control of vortex shedding using low-order models. I. Open-loop model development. *Int. J. Numer. Meth. Eng.*, 44(7):945–972, 1999.

M. A. Grepl and A. T. Patera. A posteriori error bounds for reduced-basis approximations of parametrized parabolic partial differential equations. *M2AN Math. Model. Numer. Anal.*, 39(1):157–181, 2005.

M. A. Grepl, Y. Maday, N. C. Nguyen, and A. T. Patera. Efficient reduced-basis treatment of nonaffine and nonlinear partial differential equations. *ESAIM: Mathematical Modelling and Numerical Analysis*, 41:575–605, 5 2007. ISSN 1290-3841.

P. M. Gresho and R. L. Sani. *Incompressible Flow and the Finite Element method.* John Wiley & Sons, 2000.

M. D. Gunzburger. *Finite element methods for viscous incompressible flows.* Computer Science and Scientific Computing. Academic Press Inc., Boston, MA, 1989.

M. D. Gunzburger and S. L. Hou. Treating inhomogeneous essential boundary conditions in finite element methods and the calculation of boundary stresses. *SIAM J. Numer. Anal.*, 29(2):390–424, 1992.

M. D. Gunzburger and J. Ming. Optimal control of stochastic flow over a backward-facing step using reduced-order modeling. *SIAM J. Sci. Comput.*, 33(5):2641–2663, 2011.

M. D. Gunzburger, J. S. Peterson, and J. N. Shadid. Reduced-order modeling of time-dependent PDEs with multiple parameters in the boundary data. *Comput. Methods Appl. Mech. Engrg.*, 196(4-6):1030–1047, 2007.

B. Haasdonk and M. Ohlberger. Reduced basis method for finite volume approximations of parametrized linear evolution equations. *ESAIM-Math. Model. Num. Anal.*, 42(2):277–302, 2008.

P. Holmes, J. Lumley, and G. Berkooz. *Turbulence, Coherent Structures, Dynamical Systems and Symmetry.* Cambridge University Press, 1996.

K. Ito and S. S. Ravindran. A reduced basis method for control problems governed by PDEs. In *Control and estimation of distributed parameter systems (Vorau, 1996)*, volume 126 of *Internat. Ser. Numer. Math.*, pages 153–168. Birkhäuser, 1998.

A. Janon, M. Nodet, and C. Prieur. Certified reduced-basis solutions of viscous Burgers equation parametrized by initial and boundary values. *ESAIM-Math. Model. Num. Anal.*, 47(2):317–348, 2013.

V. John. Reference values for drag and lift of a two-dimensional time-dependent flow around a cylinder. *Int. J. Numer. Meth. Eng.*, 44(7):777–788, 2004.

M. Kahlbacher. Proper orthogonal decomposition for parameter estimation of bilinear elliptic problems. Master's thesis, University of Graz, 2006.

I. Kalashnikova and M. F. Barone. On the stability and convergence of a Galerkin reduced order model (ROM) of compressible flow with solid wall and far-field boundary treatment. *Int. J. Numer. Meth. Eng.*, 83(10):1345–1375, 2010.

I. Kalashnikova and M. F. Barone. Efficient non-linear proper orthogonal decomposition/Galerkin reduced order models with stable penalty enforcement of boundary conditions. *Int. J. Numer. Meth. Eng.*, 90(11):1337–1362, 2012.

P. Knabner and L. Angermann. *Numerical methods for elliptic and parabolic partial differential equations*, volume 44 of *Texts in Applied Mathematics*. Springer-Verlag, New York, 2003.

D. Knezevic, N. Nguyen, and A. T. Patera. Reduced basis approximation and a posteriori error estimation for the parametrized unsteady boussinesq equation. *Math. Models Methods Appl. Sci.*, 21(7):1415–1442, 2011.

K. Kunisch and S. Volkwein. Control of the Burgers equation by a reduced-order approach using proper orthogonal decomposition. *J. Optimiz. Theory App.*, 102(2):345–371, 1999.

K. Kunisch and S. Volkwein. Galerkin proper orthogonal decomposition methods for parabolic problems. *Numer. Math.*, 90(1):117–148, 2001. ISSN 0029-599X.

K. Kunisch and S. Volkwein. Galerkin proper orthogonal decomposition methods for a general equation in fluid dynamics. *SIAM J. Numer. Anal.*, 40 (2):492–515, 2002a. ISSN 0036-1429.

K. Kunisch and S. Volkwein. Crank-Nicolson Galerkin proper orthogonal decomposition approximations for a general equation in fluid dynamics. In *18th GAMM Seminar on Multigrid and related methods for optimization problems*, pages 97–114, 2002b.

J. Lang. Adaptive incompressible flow computations with linearly implicit time discretization and stabilized finite elements. In K. Papailiou, D. Tsahalis, J. Periaux, C. Hirsch, and M. Pandolfi, editors, *Computational Fluid Dynamics '98*, 1998.

S. Larsson and V. Thomée. *Partial differential equations with numerical methods*, volume 45 of *Texts in Applied Mathematics*. Springer-Verlag, Berlin, 2003.

O. P. Le Maître and O. M. Knio. *Spectral methods for uncertainty quantification*. Scientific Computation. Springer, New York, 2010.

S. J. Leon, Å. Björck, and W. Gander. Gram–Schmidt orthogonalization: 100 years and more. *Numer. Lin. Alg. Appl.*, 20(3):492–532, 2013.

X. Ma and G. E. Karniadakis. A low-dimensional model for simulating three-dimensional cylinder flow. *J. Fluid Mech.*, 458:181–190, 2002.

X. Ma and N. Zabaras. An adaptive hierarchical sparse grid collocation algorithm for the solution of stochastic differential equations. *J. Comput. Phys.*, 228(8):3084–3113, 2009.

J. Mercer. Functions of positive and negative type, and their connection with the theory of integral equations. *Phil. Trans. R. Soc. Lond. A*, 83(441–458): 415–446, 1909.

C. Meyer. *Matrix analysis and applied linear algebra*. Society for Industrial and Applied Mathematics (SIAM), 2000.

N. C. Nguyen, K. Veroy, and A. T. Patera. Certified real-time solution of parametrized partial differential equations. In *Handbook of Materials Modeling*. Springer Netherlands, 2005.

B. Noack, K. Afanasiev, M. Morzynski, G. Tadmor, and F. Thiele. A hierarchy of low-dimensional models for the transient and post-transient cylinder wake. *J. Fluid Mech.*, 497:335–363, 2003.

B. Noack, P. Papas, and P. A. Monkewitz. The need for a pressure-term representation in empirical Galerkin models of incompressible shear flows. *J. Fluid Mech.*, 523:339–365, 2005.

F. Nobile, R. Tempone, and C. G. Webster. A sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM J. Numer. Anal.*, 46(5):2309–2345, 2008a.

F. Nobile, R. Tempone, and C. G. Webster. An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM J. Numer. Anal.*, 46(5):2411–2442, 2008b.

A. K. Noor and J. M. Peters. Reduced basis technique for nonlinear analysis of structures. *AIAA Journal*, 18(4):455–462, 1980.

C. Norberg. Fluctuating lift on a circular cylinder: review and new measurements. *J. Fluid. Struct.*, 17(1):57–96, 2003.

D. Pasetto, A. Guadagnini, and M. Putti. POD-based Monte Carlo approach for the solution of regional scale groundwater flow driven by randomly distributed recharge. *Adv. Water Res.*, 34(11):1450–1463, 2011.

B. Peherstorfer, S. Zimmer, and H.-J. Bungartz. Model reduction with the reduced basis method and sparse grids. In J. Garke and M. Griebel, editors, *Sparse grids and applications*, pages 223–242. Springer, Heidelberg, 2013.

J. S. Peterson. The reduced basis method for incompressible viscous flow calculations. *SIAM J. Sci. Statist. Comput.*, 10(4):777–786, 1989.

O. Posdziech and R. Grundmann. A systematic approach to the numerical calculation of fundamental quantities of the two-dimensional flow over a circular cylinder. *J. Fluid. Struct.*, 23(3):479–499, 2007.

M. Powell. A Fortran subroutine for solving systems of nonlinear algebraic equations. In P. Rabinowitz, editor, *Numerical methods for nonlinear algebraic equations*, chapter 7, pages 115–161. Gordon and Breach, 1970.

C. Prud'homme, D. V. Rovas, K. Veroy, L. Machiels, Y. Maday, A. T. Patera, and G. Turinici. Reliable real-time solution of parametrized partial differential equations: Reduced-basis output bound methods. *J. Fluid. Eng.–T. ASME*, 124(1):70–80, 2002a.

C. Prud'homme, D. V. Rovas, K. Veroy, and A. T. Patera. A mathematical and computational framework for reliable real-time solution of parametrized partial differential equations. *M2AN Math. Model. Numer. Anal.*, 36(5): 747–771, 2002b.

M. M. Rao and R. J. Swift. *Probability theory with applications*, volume 582 of *Mathematics and Its Applications*. Springer, New York, second edition, 2006.

S. Ravindran. Error analysis for Galerkin POD approximation of the nonstationary Boussinesq equations. *Num. Meth. Part. Diff. Eq.*, 27(6):1639–1665, 2011.

D. Rempfer. Investigations of boundary layer transition via Galerkin projection on empirical eigenfunctions. *Phys. Fluids*, 8(1):175 – 188, 1996.

G. Rozza and K. Veroy. On the stability of the reduced basis method for Stokes equations in parametrized domains. *Comput. Methods Appl. Mech. Engrg.*, 196(7):1244–1260, 2007.

E. W. Sachs and S. Volkwein. POD-Galerkin approximations in PDE-constrained optimization. *GAMM-Mitt.*, 33(2):194–208, 2010.

B. Schieche. *Unsteady adaptive stochastic collocation methods on sparse grids*. PhD thesis, Technische Universität Darmstadt, 2012a.

B. Schieche. Adaptive stochastic collocation on sparse grids. *PAMM*, 12(1): 653–654, 2012b.

B. Schieche and J. Lang. Stochastic analysis of nusselt numbers for natural convection with uncertain boundary conditions. Preprint 2608, Technische Universität Darmstadt, Department of Mathematics, 2010.

B. Schieche and J. Lang. Uncertainty quantification for thermo-convective Poiseuille flow using stochastic collocation. *Int. J. Comp. Sci. Eng.*, 2013. to appear.

E. Schmidt. Zur Theorie der linearen und nichtlinearen Integralgleichungen. *Math. Ann.*, 63(4):433–476, 1907.

S. Sirisup and G. E. Karniadakis. Stability and accuracy of periodic flow solutions obtained by a POD-penalty method. *Phys. D*, 202(3-4):218–237, 2005.

S. Sirisup, G. Karniadakis, D. Xiu, and I. Kevrekidis. Equation-free/galerkin-free POD-assisted computation of incompressible flows. *J. Comput. Phys.*, 207:568–587, 2005.

L. Sirovich. Turbulence and the dynamics of coherent structures. parts I, II and III. *Quart. Appl. Math.*, 45:561–590, 1987.

J. Smagorinsky. General circulation experiments with the primitive equations, I, The basic experiment. *Monthly Weather Review*, 91:99–164, 1963.

S. Smolyak. Quadrature and interpolation formulas for tensor products of certain classes of functions. *Soviet Math. Dokl.*, 4:240–243, 1963.

G. W. Stewart. On the early history of the singular value decomposition. *SIAM Rev.*, 35(4):551–566, 1993.

C. Taylor and P. Hood. A numerical solution of the Navier-Stokes equations using the finite element technique. *Int. J. Comput. Fluids*, 1(1):73–100, 1973.

L. N. Trefethen. Six myths of polynomial interpolation and quadrature. *Math. Today (Southend-on-Sea)*, 47(4):184–188, 2011.

L. N. Trefethen and D. Bau, III. *Numerical linear algebra*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.

S. Ullmann and J. Lang. A POD-Galerkin reduced model with updated coefficients for Smagorinsky LES. In J. C. F. Pereira, A. Sequeira, and J. M. C. Pereira, editors, *Proceedings of the V European Conference on Computational Fluid Dynamics ECCOMAS CFD 2010*, Lisbon, Portugal, June 2010.

S. Ullmann and J. Lang. POD and CVT Galerkin reduced-order modeling of the flow around a cylinder. *PAMM*, 12(1):697–698, 2012.

S. Ullmann and J. Lang. POD-Galerkin modeling and sparse-grid collocation for a natural convection problem with stochastic boundary conditions. In J. Garke and D. Pflüger, editors, *Sparse Grids and Applications – Munich 2012*, Lecture Notes in Computational Science and Engineering. Springer, 2014, to appear.

S. Ullmann, S. Löbig, and J. Lang. Adaptive large eddy simulation and reduced-order modeling. In J. Janicka, A. Sadiki, M. Schäfer, and C. Heeger, editors, *Flow and Combustion in Advanced Gas Turbine Combustors*. Springer, 2013.

S. Volkwein. Optimal control of a phase-field model using proper orthogonal decomposition. *Z. Angew. Math. Mech.*, 81(2):83–97, 2001.

S. Volkwein. Proper orthogonal decomposition: Application in optimization and control, 2008. Lecture Notes, CEA-EDF-INRIA Summer School on Model Order Reduction and Reduced Basis Methods, Saint-Lambert-des-Bois, France, 23 June–4 July 2008.

J. Waldvogel. Fast construction of the Fejér and Clenshaw-Curtis quadrature rules. *BIT Num. Math.*, 46(1):195–202, 2006.

D. Xiu and J. S. Hesthaven. High-order collocation methods for differential equations with random inputs. *SIAM J. Sci. Comput.*, 27(3):1118–1139, 2005.

Curriculum Vitae

10/2003–6/2006	Studium Computational Science (B. Sc.), Technische Universität Chemnitz
9/2006–8/2008	Studium Applied Mathematics (M. Sc.), Delft University of Technology
10/2008–	Wissenschaftlicher Mitarbeiter, Fachbereich Mathematik, Technische Universität Darmstadt
2/2014	Promotion in Mathematik, Technische Universität Darmstadt