OPTIMIZATION AND MODEL REDUCTION OF TIME DEPENDENT PDE-CONSTRAINED OPTIMIZATION PROBLEMS: APPLICATIONS TO SURFACE ACOUSTIC WAVE DRIVEN MICROFLUIDIC BIOCHIPS

A Dissertation

Presented to

the Faculty of the Department of Mathematics

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

By

Harbir Antil

December 2009

OPTIMIZATION AND MODEL REDUCTION OF TIME DEPENDENT PDE-CONSTRAINED OPTIMIZATION PROBLEMS: APPLICATIONS TO SURFACE ACOUSTIC WAVE DRIVEN MICROFLUIDIC BIOCHIPS

Harbir Antil

APPROVED:

Prof. Ronald H. W. Hoppe,
Chairman

Prof. Roland Glowinski

Prof. Yuri Kuznetsov

Prof. Matthias Heinkenschloss

Dean, College of Natural Sciences and Mathematics

Acknowledgements

First of all, I would like to thank my advisor Dr. Ronald Hoppe for his guidance throughout this journey. He was my mentor during my Master's and he introduced me to the field of scientific computing, specifically PDE-constrained optimization. Without his insight and deep scientific knowledge, it won't have been possible to achieve this goal. He has been a great source of inspiration and never ceased to encourage me throughout my Ph.D. From him, I learnt the true meaning of perseverance. I am thankful to him, and consider myself fortunate to work under him.

I am very grateful to Dr. Matthias Heinkenschloss for introducing me to the field of domain decomposition and model reduction. This final year of my Ph.D., I have learnt a great deal from him, covering both theoretical and practical aspects. He has given a new direction to my work, and I have been deeply influenced by his research.

Dr. Christopher Linsenmann my colleague and close friend has helped me from the very beginning. This journey would have been different without his support.

Special thanks to Dr. Heinkenschloss, Dr. Hoppe, Dr. Linsenmann, and Natasha Sharma for proof-reading this dissertation.

Several people contributed to the achievement of this work. Recognition goes to all my colleagues and friends at the Department of Mathematics, University of Houston, and Mathematical Sciences Foundation, Delhi.

Finally, I am grateful to my brother and sister-in-law and all other family members for their love and constant moral support.

OPTIMIZATION AND MODEL REDUCTION OF TIME DEPENDENT PDE-CONSTRAINED OPTIMIZATION PROBLEMS: APPLICATIONS TO SURFACE ACOUSTIC WAVE DRIVEN MICROFLUIDIC BIOCHIPS

An Abstract of a Dissertation

Presented to

the Faculty of the Department of Mathematics

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

By

Harbir Antil

December 2009

Abstract

The optimal design of structures and systems described by partial differential equations (PDEs) often gives rise to large-scale optimization problems, in particular if the underlying system of PDEs represents a multiscale, multiphysics problem. Therefore, reduced order modeling techniques such as balanced truncation model reduction (BTMR), proper orthogonal decomposition (POD), or reduced basis methods (RB) are used to significantly decrease the computational complexity while maintaining the desired accuracy of the approximation. We are interested in such shape optimization problems where the design issue is restricted to a relatively small portion of the computational domain and in optimal control problems where the nonlinearity is local in nature. In these cases, it appears to be natural to rely on a full order model only in that specific part of the domain and to use a reduced order model elsewhere. A convenient methodology to realize this idea is a suitable combination of domain decomposition techniques and BTMR. We will consider such an approach for optimal control and shape optimization problems governed by advection-diffusion equations and derive explicit error bounds for the modeling error.

As an application in life sciences, we will be concerned with the optimal design of capillary barriers as part of a network of microchannels and reservoirs on surface acoustic wave driven microfluidic biochips. Here, the state equations represent a multiscale multiphysics problem consisting of the linearized equations of piezoelectricity and the compressible Navier-Stokes equations. The multiscale character is due to the occurrence of fluid flow on different time scales. A standard homogenization approach by means of a state parameter results in a first-order time periodic linearized compressible Navier-Stokes equations and a second-order compressible Stokes system. The second-order compressible Stokes system provides

an appropriate model for the optimal design of the capillary barriers.

Another application considered is the shape optimization of an aorto-coronaric bypass. Finally, in order to address environmental issues, we present an optimal control problem where our aim is to reduce the water pollution in a region of choice.

Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

Microfluidics is the science dealing with the behavior, precise control, and manipulation of fluids in the sub-millimeter scale. We are all familiar with the revolution brought to us by the advancement in microelectronics in our day to day life by providing, smaller, cheaper, and highly efficient devices. One should expect microfluidic sciences to follow the same path.

In life sciences, a popular concept is "labs-on-a-chip" which is defined as chip-based miniature laboratories that can be controlled electronically. Microfluidic biochips represent an important example (cf. Fig. 1.1 (left)). The miniaturized chip laboratories are able to do complex tasks within a few micrometers for which usually a full-size laboratory is required. Often only a very tiny amount of sample is available, e.g., in forensics and in

gene expression profiling analysis.

Microfluidic biochips are used in pharmaceutical, medical, and forensic applications for high throughput screening, genotyping, and sequencing in genomics, protein profiling in proteomics, and cytometry in cell analysis [92, 101]. They provide a much better sensitivity and a greater flexibility than traditional approaches. More importantly, they give rise to a significant speed-up of the hybridization processes and allow the in-situ investigation of these processes at an extremely high time resolution. This can be achieved by integrating the fluidics on top of the chip by means of a lithographically produced network of channels and reservoirs (cf. Fig. 1.1 (left)).

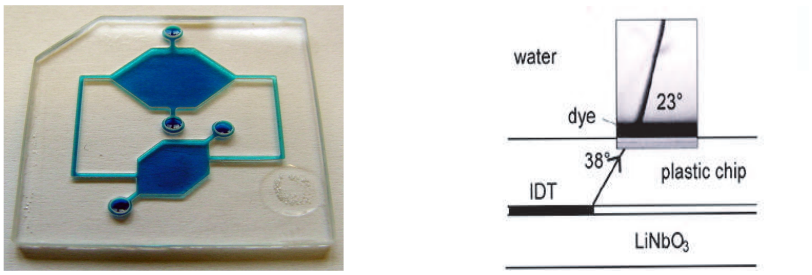


Figure 1.1: Microfluidic biochip (left) and sharp jet created by surface acoustic waves (right)

The idea is to inject a DNA or protein containing probe and to transport it in the fluid to a reservoir where a chemical analysis is performed. The fluid flow can be taken care of by external pumps which, however, do not guarantee a very precise control of the fluid flow and are subject to wear. A new generation of biochips is based on a surface acoustic waves (SAW)-driven fluid flow [47, 58, 118, 119, 122]. Surface acoustic waves are generated by interdigital transducers (IDT), well-known from Micro-Electro-Mechanical Systems (MEMS). They propagate through the base of the device with amplitudes in the range of nanometers and enter the fluid-filled microchannels creating sharp jets (cf. Fig. 1.1 (right)). This happens within nanoseconds. In the microchannels, the SAW get

significantly damped so that an almost stationary fluid pattern emerges which is called acoustic streaming. This relaxation process occurs on a time scale of milliseconds. We are thus faced with a multiscale, multiphysics problem whose mathematical modeling and numerical simulation represents a significant challenge. It is also a challenging problem with regard to various optimization issues such as the optimal design of the microchannels in order to achieve a maximum pumping rate. Another one is the design of pressure driven capillary barriers between the channels and the reservoirs to guarantee a precise filling of the reservoirs with the probes (cf. Fig. 1.2). This amounts to the solution of a shape optimization problem where the mathematical model for the acoustic streaming consists of the linearized equations of piezoelectricity and the compressible Navier-Stokes equations.

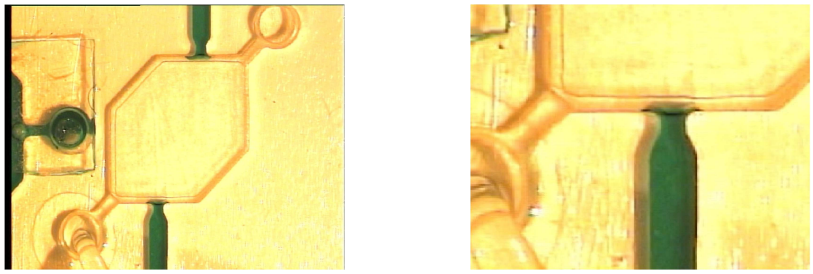


Figure 1.2: Capillary barriers

The multiscale character of the problem can be appropriately taken care of by a homogenization approach. Following [10, 75], after homogenization we obtain a linearized compressible Navier-Stokes equation and a compressible Stokes system. We will discuss this in more detail in Chapters 2 and 6.

For the efficient solution of the optimal design problems, we have developed an adaptive multilevel interior-point method of barrier type featuring a predictor-corrector continuation method with an adaptive choice of the barrier parameter along the barrier path. The prediction step relies on a nested-iteration type tangent continuation, and the correction

step is a Newton-multigrid method for the KKT system. Despite the fact that this approach leads to a considerable reduction in the computational work compared to more standard optimization strategies, the amount of computational time is still significant, and there is a need for further reductions. Such reductions can be achieved by model reduction based optimization methods using reduced order models for the underlying state equations generated, e.g., by Proper Orthogonal Decomposition (POD), Balanced Truncation Model Reduction (BTMR), Krylov subspace methods, or reduced basis methods.

Among these techniques, we have chosen BTMR combined with a domain decomposition approach referred to as DDBTMR, since for many of the above mentioned optimal design problems the region, where the optimal design has to be implemented, is small compared to the rest of the microfluidic network. Our approach uses domain decomposition applied to the optimality system to isolate the subsystem that explicitly depends on the optimization variables from the remaining linear optimality system. We derive estimates for the error between the solution of the original optimization problem and the solution of the reduced problem and these estimates are confirmed by numerical results. We first consider shape optimization problems governed by the Stokes system and then extend DDBTMR to microfluidic biochips where the governing equations are compressible Stokes equations.

1.1 Biochip in action

A microfluidic biochip (cf. Fig. 1.3 (left)) consists of a lithographically produced network of microchannels located on top of a substrate (glass or plastic plate coated by a piezoelectric material such as $LiNbO_3$). An IDT, which is attached to a chip holder (cf. Fig 1.4 (left)) is placed on top of the substrate. The chip holder holds an RF-input connection for receiving the high frequency signal produced by the signal generator (cf. Fig. 1.4 (right)). This

high frequency signal (around 100 MHz) causes the IDT to excite and launch SAWs. The SAW move like miniature earthquakes and transport fluids and solid matter across the microchannels (cf. Fig.1.3 (right)).

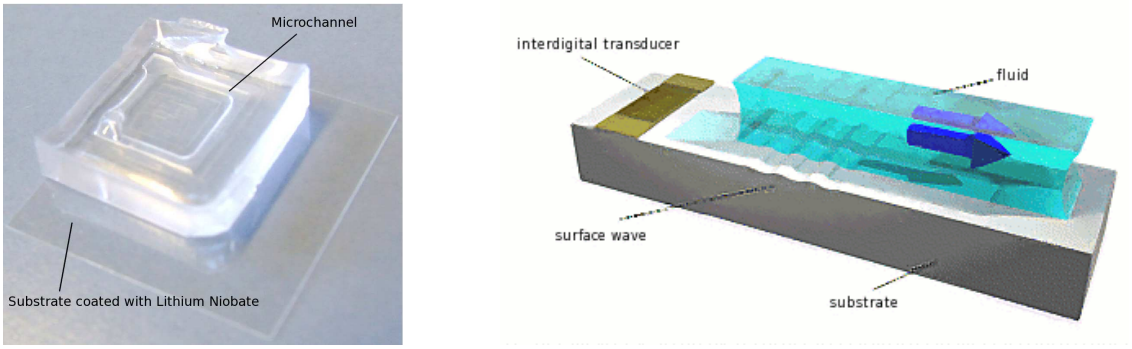


Figure 1.3: Microfluidic biochip placed on a substrate (left) and fluid flow in part of a microchannel steered by SAW generated using IDT (right)

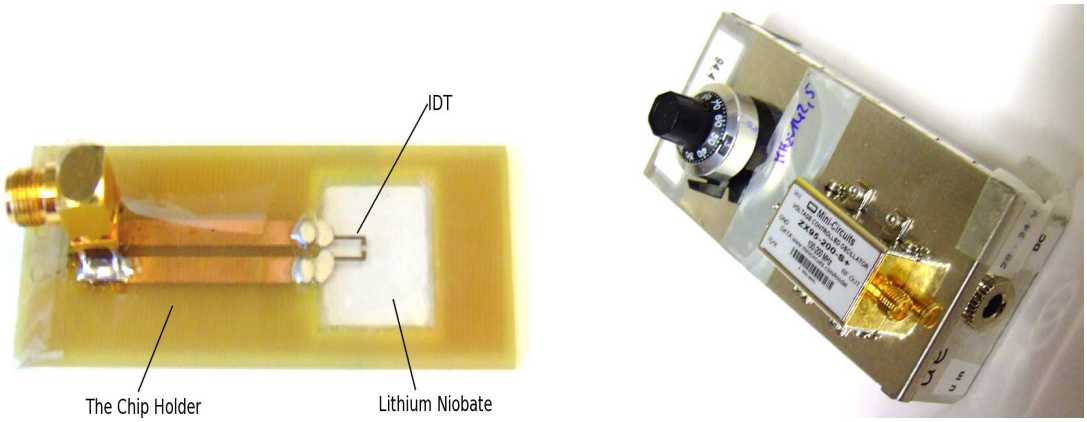


Figure 1.4: IDT on a substrate (left) and RF signal generator (right)

1.2 Outline of the thesis

This thesis is organized as follows:

In **Chapter 2**, we will present a mathematical model for the acoustic streaming that can be derived by a homogenization approach based on the coupling of the equations of piezoelectricity and the compressible Navier-Stokes equations.

Chapter 3 is devoted to shape optimization by means of a so-called all-at-once approach featuring the simultaneous optimization and numerical solution of the state equations (stationary Stokes system in this case).

The all-at-once approach is realized by a barrier type interior-point method in terms of a path-following predictor-corrector continuation scheme with an adaptive choice of the continuation steplength along the barrier path. The barrier method is implemented within a multilevel framework using a nested iteration type predictor and a Newton multigrid technique as a corrector. The results obtained are compared with classical barrier methods such as the long-step path-following method and Mehrotra's predictor-corrector method. Some parts of this chapter have been taken from [16] and [17].

Chapter 4 introduces a technique for the dimension reduction of a class of PDE constrained optimization problems governed by linear time dependent advection-diffusion equations for which the optimization variables are related to spatially localized quantities. Our approach uses domain decomposition applied to the optimality system to isolate the subsystem that explicitly depends on the optimization variables from the remaining linear optimality subsystem. We apply balanced truncation model reduction to the linear optimality subsystem. The resulting coupled reduced optimality system can be interpreted as the optimality system of a reduced optimization problem. We derive estimates for the error between the solution of the original optimization problem and the solution of the reduced problem. The approach is demonstrated numerically with an optimal control problem and a shape optimization problem. This chapter is based on [14].

In **Chapter 5**, the technique introduced in Chapter 4 is extended to optimization problems governed by the Stokes equations. Here, we are interested in shape optimization problems where the design issue is restricted to a relatively small portion of the computational domain. In this case, it appears to be natural to rely on a full order model only in that specific part of the domain and to use a reduced order model elsewhere. A convenient methodology to realize this idea consists in a suitable combination of domain decomposition techniques and balanced truncation model reduction. We will consider such an approach for shape optimization problems associated with the time-dependent Stokes system and derive explicit error bounds for the modeling error.

Although conceptually the approach in this case is same as in Chapter 4, the extension requires several important changes. These are due to incompressibility constraints that affect the model reduction, the domain decomposition and the coupling of both, as well as the error analysis. As a numerical example, we will be concerned with the optimal design of capillary barriers as part of a network of microchannels and reservoirs on domain motivated by microfluidic biochip, but with Stokes equations as the state system. This chapter contains the results from [15].

Chapter 6: The mathematical model for microfluidic biochips represents a multiphysics problem consisting of the piezoelectric equations coupled with multiscale compressible Navier-Stokes equations that have to be treated by an appropriate homogenization. We discussed the modeling approach in Chapter 2, here we present algorithmic tools for the numerical simulation and address optimal design issues. In particular, the optimal design of specific parts of the biochips leads to large-scale optimization problems. In order to reduce the computational complexity, we present a combination of domain decomposition and balanced truncation model reduction which allows explicit error bounds for the error between the reduced order and the fine-scale optimization problem, similar to Stokes case

in Chapter 5. It is shown that this approach gives rise to a significant reduction of the problem size while maintaining the accuracy of the approximation. This chapter is based on [13].

CHAPTER 2

Modeling microfluidic biochips

In this chapter we will present a mathematical model for the acoustic streaming that can be derived by a homogenization approach based on the coupling of the equations of piezoelectricity and the compressible Navier-Stokes equations. In Section 2.1 we will first review the piezoelectric equations, and then take care of the multiscale character of the microfluidic biochip problem using an homogenization approach, which gives us a linearized compressible Navier-Stokes equation and a compressible Stokes system. For further details we refer to [10, 13, 48, 49, 74, 75]. Section 2.2 is devoted to the general discussion of a shape optimization problem. We will be more specific in the later chapters.

2.1 Mathematical modeling

In this section, we will develop and analyze a mathematical model describing the operational behavior of SAW driven microfluidic biochips. The model consists of the equations of piezoelectricity unilaterally coupled with the compressible Navier-Stokes equations. In particular, in Subsection 2.1.1 we will be concerned with the piezoelectric equations, whereas Subsection 2.1.2 will be devoted to the compressible Navier-Stokes equations. Using techniques from heterogeneous multiscale methods [37, 38], we will derive a compressible Stokes system which serves as a model for the acoustic streaming.

Throughout this section, we will use standard notation from Lebesgue and Sobolev space theory. In particular, for a bounded polygonal or polyhedral domain $\Omega \subset \mathbb{R}^d, d \in \{2,3\}$, with boundary $\Gamma = \partial\Omega$, we denote by $L^2(\Omega)$ and $\mathbf{L}^2(\Omega)$ the Hilbert spaces of scalar and vector-valued Lebesgue integrable functions on Ω with inner products $(\cdot,\cdot)_{0,\Omega}$ and norms $\|\cdot\|_{0,\Omega}$. Likewise, $H^1(\Omega)$ and $\mathbf{H}^1(\Omega)$ refer to the Sobolev spaces with inner products $(\cdot,\cdot)_{1,\Omega}$ and norms $\|\cdot\|_{1,\Omega}$, whereas $H^{1/2}(\Gamma')$ and $\mathbf{H}^{1/2}(\Gamma'), \Gamma' \subset \Gamma$, stand for the associated trace spaces. We further refer to $H^{-1}(\Omega)$ and $\mathbf{H}^{-1}(\Omega)$ as the dual spaces of $H_0^1(\Omega)$ and $\mathbf{H}_0^1(\Omega)$, respectively. For further notation account we refer to Appendix C.

2.1.1 The piezoelectric equations

In piezoelectric materials, the stress tensor $\boldsymbol{\sigma}$ depends linearly on the electric field $\mathbf{E}$ according to a generalized Hooke's law

$$\boldsymbol{\sigma}(\mathbf{u},\mathbf{E}) \;=\; \mathbf{c}\,\boldsymbol{\varepsilon}(\mathbf{u}) \;-\; \mathbf{e}\,\mathbf{E}\,. \tag{2.1}$$

Here, $\boldsymbol{\varepsilon}(\mathbf{u}) := (\nabla\mathbf{u} + (\nabla\mathbf{u})^T)/2$ refers to the linearized strain tensor with $\mathbf{u}$ denoting the mechanical displacement, whereas $\mathbf{c}$ and $\mathbf{e}$ stand for the symmetric fourth-order elasticity

tensor and the symmetric third-order piezoelectric tensor, respectively. Hence, the application of an electric field will cause a displacement of the material. Piezoelectric materials also show the reverse effect to generate an electric field when subjected to mechanical stress. These properties are called the piezoelectric effect and the inverse piezoelectric effect. The origin of the piezoelectric effect is related to an asymmetry in the unit cell of a piezoelectric crystal and can be observed only in materials with a polar axis (cf., e.g., [40, 82]).

The frequency of the electromagnetic wave is small compared to the frequency of the generated acoustic wave so that a coupling will be neglected. In particular, the electric field will be considered as quasistatic. Moreover, it is irrotational and hence, according to $\mathbf{E} = -\nabla\Phi$ it can be expressed as the gradient of an electric potential Φ. Since piezoelectric materials are nearly perfect insulators, the only remaining quantity of interest in Maxwell's equations is the dielectric displacement $\mathbf{D}$ which is related to the electric field by the constitutive equation

$$\mathbf{D} \;=\; \boldsymbol{\epsilon}\,\mathbf{E} \;+\; \mathbf{P}\;, \tag{2.2}$$

where $\boldsymbol{\epsilon}$ is the electric permittivity of the material and $\mathbf{P}$ stands for the polarization. In piezoelectric materials, the polarization $\mathbf{P}$ due to external strain is linear, i.e., there holds

$$\mathbf{P} \;=\; \mathbf{e}\,\boldsymbol{\varepsilon}(\mathbf{u})\;. \tag{2.3}$$

We assume that the piezoelectric material with density $\rho_p > 0$ occupies some rectangular domain Ω_1 with boundary $\Gamma_1 = \partial\Omega_1$ and exterior unit normal $\mathbf{n}_1$ and decompose boundary in two different ways such that

$$\begin{aligned}\Gamma_1 \quad &= \overline{\Gamma}_{E,D} \cup \overline{\Gamma}_{E,N}, \quad \Gamma_{E,D} \cap \Gamma_{E,N} = \emptyset,\\ \Gamma_1 \quad &= \overline{\Gamma}_{p,D} \cup \overline{\Gamma}_{p,N}, \quad \Gamma_{p,D} \cap \Gamma_{p,N} = \emptyset,\end{aligned}$$

where $\Gamma_{E,D} \subset \Gamma_1$ is a $(d-1)$ dimension manifold and $\Gamma_{E,N} := \Gamma_1 \setminus \overline{\Gamma}_{E,D}$. Given boundary data $\Phi_{E,D}$ on $\Gamma_{E,D}$, the pair $(\mathbf{u}, \Phi)$ satisfies the following initial-boundary value problem

for the piezoelectric equations

$$\rho_p \frac{\partial^2 \mathbf{u}}{\partial t^2} - \nabla \cdot \boldsymbol{\sigma}(\mathbf{u}, \mathbf{E}) = 0 \qquad \text{in } \Omega_1 \times (0, T_1), \tag{2.4a}$$

$$\nabla \cdot \mathbf{D}(\mathbf{u}, \mathbf{E}) = 0 \qquad \text{in } \Omega_1 \times (0, T_1), \tag{2.4b}$$

$$\mathbf{u} = 0 \quad \text{on } \Gamma_{p,D} \times (0, T_1) \quad , \quad \mathbf{n}_1 \cdot \boldsymbol{\sigma} = \boldsymbol{\sigma}_{\mathbf{n}_1} \qquad \text{on } \Gamma_{p,N} \times (0, T_1), \tag{2.4c}$$

$$\Phi = \Phi_{E,D} \quad \text{on } \Gamma_{E,D} \times (0, T_1) \quad , \quad \mathbf{n}_1 \cdot \mathbf{D} = D_{\mathbf{n}_1} \qquad \text{on } \Gamma_{E,N} \times (0, T_1) \ , \tag{2.4d}$$

$$\mathbf{u}(\cdot, 0) = 0 \quad , \quad \frac{\partial \mathbf{u}}{\partial t}(\cdot, 0) = 0 \qquad \text{in } \Omega_1, \tag{2.4e}$$

which have to be completed by the constitutive equations (2.1),(2.2) and (2.3).

Assuming time periodic excitations $\Phi_{E,D}(\cdot, t) = \mathrm{Re}\left(\hat{\Phi}_{E,D} \exp(-\, i\omega t)\right)$ such that $\hat{\Phi}_{E,D} \in H^{1/2}(\Gamma_{E,D})$, we are looking for time harmonic solutions

$$\mathbf{u}(\cdot, t) \ = \ \mathrm{Re}\,(\, \mathbf{u}(\cdot) \exp(-\, i\omega t)\,) \quad , \quad \Phi(\cdot, t) \ = \ \mathrm{Re}\,(\, \Phi(\cdot) \exp(-\, i\omega t)\,) \ .$$

This leads to a saddle point problem for a Helmholtz-type equation which in its weak form amounts to the computation of $(\mathbf{u}, \Phi) \in \mathbf{V} \times W$, where $\mathbf{V} := H^1_{0,\Gamma_{p,D}}(\Omega_1)^d$ and $W := \{\varphi \in H^1(\Omega_1) \mid \varphi_{\Gamma_{E,D}} = \hat{\Phi}_{E,D}\}$, such that for all $\mathbf{v} \in \mathbf{V}$ and $\psi \in W_0 := H^1_{0,\Gamma_{E,D}}(\Omega_1)$

$$a(\mathbf{u}, \mathbf{v}) + b(\Phi, \mathbf{v}) - \omega^2 \rho_p (\mathbf{u}, \mathbf{v})_{0,\Omega} = \ell_1(\mathbf{v}), \tag{2.5a}$$

$$b(\psi, \mathbf{u}) - c(\Phi, \psi) \qquad\qquad = \ell_2(\psi). \tag{2.5b}$$

Here, $H^1_{0,\Gamma_{p,D}}(\Omega_1)^d := \{\mathbf{v} \in H^1(\Omega_1)^d \mid \mathbf{v}|_{\Gamma_{p,D}} = 0\}$, $H^1_{0,\Gamma_{E,D}}(\Omega_1) := \{\psi \in H^1(\Omega_1) \mid \psi|_{\Gamma_{E,D}} = 0\}$, and the sesquilinear forms $a(\cdot,\cdot), b(\cdot,\cdot), c(\cdot,\cdot)$ and the functionals $\ell_1 \in \mathbf{V}^*, \ell_2 \in W^*$ are given by

$$a(\mathbf{v}, \mathbf{w}) \ := \ \int_{\Omega_1} \mathbf{c}\, \boldsymbol{\varepsilon}(\mathbf{v}) : \boldsymbol{\varepsilon}(\bar{\mathbf{w}})\ dx \quad , \quad b(\varphi, \mathbf{v}) \ := \ \int_{\Omega_1} \mathbf{e}\, \nabla\varphi : \boldsymbol{\varepsilon}(\bar{\mathbf{v}})\ dx \ ,$$

$$c(\varphi, \psi) \ := \ \int_{\Omega_1} \boldsymbol{\epsilon}\, \nabla\varphi \cdot \nabla\bar{\psi}\ dx \ ,$$

$$\ell_1(\mathbf{v}) \ := \ \langle \boldsymbol{\sigma}_{\mathbf{n}_1}, \mathbf{v} \rangle_{p,N} \quad , \quad \ell_2(\psi) \ := \ \langle D_{\mathbf{n}_1}, \psi \rangle_{E,N} \ ,$$

with $\langle\cdot,\cdot\rangle_{p,N}, \langle\cdot,\cdot\rangle_{E,N}$ denoting the dual pairings between the associated trace spaces and their dual spaces, respectively.

We denote by $\mathbf{A} : \mathbf{V} \to \mathbf{V}^*, \mathbf{B} : W \to \mathbf{V}^*$, and $\mathbf{C} : W \to W^*$ the operators associated with the sesquilinear forms and by $\mathbf{I}$ the injection $\mathbf{I} : \mathbf{V} \to \mathbf{V}^*$. Then, an equivalent formulation of (2.5a),(2.5b) is

$$(\mathbf{A} - \omega^2\rho_p\mathbf{I})\mathbf{u} + \mathbf{B}\Phi = \mathbf{f}, \tag{2.6a}$$

$$\mathbf{B}^*\mathbf{u} - \mathbf{C}\Phi = f. \tag{2.6b}$$

Here, the right-hand sides $\mathbf{f} \in \mathbf{V}^*$ and $f \in W^*$ are given by

$$\mathbf{f} \;:=\; \ell_1 \;-\; \mathbf{B}\tilde{\Phi}_{E,D} \quad , \quad f \;:=\; \ell_2 \;+\; \mathbf{C}\tilde{\Phi}_{E,D} \;, \tag{2.7}$$

where $\tilde{\Phi}_{E,D}$ stands for the extension of the Dirichlet data onto W.
In particular, the operator $\mathbf{A}$ is symmetric and $\mathbf{V}$-elliptic, and the operator $\mathbf{C}$ is symmetric and W-elliptic. The symmetry of $\mathbf{A}$ results from the symmetry of the elasticity tensor $\mathbf{c}$, whereas the $\mathbf{V}$-ellipticity is a direct consequence of the positive definiteness of $\mathbf{c}$ and Korn's inequality. Likewise, the symmetry of $\mathbf{C}$ follows from the symmetry of the dielectric permittivity $\boldsymbol{\epsilon}$ and the W-ellipticity can be deduced from the positive definiteness of $\boldsymbol{\epsilon}$.
Elimination of Φ from (2.6a),(2.6b) results in the Schur complement system

$$\mathbf{S}\mathbf{u} \;-\; \omega^2\rho_p\,\mathbf{u} \;=\; \mathbf{g}\;. \tag{2.8}$$

Here, the Schur complement operator $\mathbf{S} : \mathbf{V} \to \mathbf{V}^*$ is defined according to

$$\mathbf{S} \;:=\; \mathbf{A} \,+\, \mathbf{B}\mathbf{C}^{-1}\mathbf{B}^*\;, \tag{2.9}$$

whereas the right-hand side $\mathbf{g}$ is given by

$$\mathbf{g} \;:=\; \mathbf{f} \;+\; \mathbf{B}\mathbf{C}^{-1}f\;. \tag{2.10}$$

Theorem 2.1.1 *For the Schur complement* $\mathbf{S}$ *given by* (2.9) *and the Schur complement system* (2.8) *there holds:*

(i) The spectrum of $\mathbf{S}$ *consists of a sequence of countably many real eigenvalues* $0 < \zeta_1^2 < \zeta_2^2 < \dots$ *tending to infinity, i.e.,* $\lim_{j\to\infty} \zeta_j^2 = \infty$.

(ii) If $\omega^2\rho_p$ *is not an eigenvalue of* $\mathbf{S}$*, for every* $\mathbf{g} \in \mathbf{V}^*$, (2.8) *admits a unique solution* $\mathbf{u} \in \mathbf{V}$ *depending continuously on* $\mathbf{g}$.

(iii) If $\omega^2\rho_p \in \mathbb{R}$ *is an eigenvalue of* $\mathbf{S}$, (2.8) *is solvable if and only if* $\mathbf{g} \in Ker(\mathbf{S}-\omega^2\rho_p\,\mathbf{I})^0$ *where*

$$Ker(\mathbf{S} - \omega^2\rho_p\ \mathbf{I})^0\ :=\ \{\mathbf{v}^* \in \mathbf{V}^* \mid \langle \mathbf{v}^*, \mathbf{v}\rangle = 0\ ,\ \mathbf{v} \in Ker(\mathbf{S} - \omega^2\rho_p\ \mathbf{I})\}\ .$$

Proof: The Schur complement system (2.8) can be rewritten according to

$$\mathbf{Su}\ -\ \omega^2\rho_p\,\mathbf{u}\ =\ -\omega^2\rho_p\,\mathbf{S}\Big(\mathbf{S}_R^{-1} - \omega^{-2}\rho_p^{-1}\,\mathbf{I}\Big)\,\mathbf{u}\ =\ \mathbf{g}\ ,$$

where $\mathbf{S}_R^{-1} : \mathbf{L}^2(\Omega_1) \to \mathbf{V}$ is given by $\mathbf{S}_R^{-1}\mathbf{v} = \mathbf{S}^{-1}\mathbf{v}$, $\mathbf{v} \in \mathbf{L}^2(\Omega_1)$. It is easy to see that $\mathbf{S}_R^{-1}$ is a compact self-adjoint endomorphism on $\mathbf{L}^2(\Omega_1)$ and hence, the assertions $(i), (ii)$, and (iii) follow from the Hilbert-Schmidt theory and the Fredholm alternative (cf., e.g., [123]). □

2.1.2 The compressible Navier-Stokes equations

Since compressible effects dominate the SAW-induced fluid flow, it has to be described by the compressible Navier-Stokes equations. We denote by $\Omega_2(t), t \in [0, T_2]$, the time dependent domain occupied by the fluid with boundary $\Gamma_2(t) = \overline{\Gamma(t)}_{2,D} \cup \overline{\Gamma(t)}_{2,N}$, $\Gamma_{2,D}(t) \cap$

$\Gamma_{2,N}(t) = \emptyset$. Here, $\Gamma_{2,D}(t)$ is that part of the boundary where the SAWs enter the fluid-filled microchannels. We assume that the coupling between the piezoelectric and the Navier-Stokes equations is unilateral and occurs by means of the deflection of the walls of the microchannels caused by the SAWs. We denote by $\mathbf{v}$ and p the velocity and the pressure, and we refer to ρ_f, η, and ξ as the density of the fluid and the standard and bulk viscosities. Then, the pair $(\mathbf{v}, p)$ satisfies the following initial-boundary value problem

$$\rho_f\big(\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v}\cdot\nabla\mathbf{v}\big) = \nabla\cdot\boldsymbol{\sigma} \qquad \text{in } \Omega_2(t),\ t\in(0,T_2], \tag{2.11a}$$

$$\frac{\partial \rho_f}{\partial t} + \nabla\cdot(\rho_f\mathbf{v}) = 0 \qquad \text{in } \Omega_2(t),\ t\in(0,T_2], \tag{2.11b}$$

$$\mathbf{v}(\cdot + \mathbf{u}(\cdot,t),t) = \frac{\partial \mathbf{u}}{\partial t}(\cdot,t) \qquad \text{on } \Gamma_{2,D}(t),\ t\in(0,T_2], \tag{2.11c}$$

$$\boldsymbol{\sigma}\mathbf{n} = 0 \qquad \text{on } \Gamma_{2,N}(t),\ t\in(0,T_2], \tag{2.11d}$$

$$\mathbf{v}(\cdot,0) = \mathbf{v}_0, \quad p(\cdot,0) = p_0 \qquad \text{in } \Omega_2(0), \tag{2.11e}$$

where

$$\boldsymbol{\sigma} = (\sigma_{ij})_{i,j=1}^{d}\ ,\ \sigma_{ij} := -p\ \delta_{ij} + 2\eta\varepsilon_{ij}(\mathbf{v}) + \delta_{ij}(\xi - 2\eta/3)\nabla\cdot\mathbf{v},$$

and $\mathbf{u}$ in (2.11c) stands for the deflection of the walls of the microchannels caused by the SAW. We note that $\mathbf{u}$ can be computed by the solution of the linearized equations of piezoelectricity as described in previous subsection, for more details see [49]. In this dissertation we assume that $\mathbf{u}$ is given as the time periodic boundary data. We will discuss this in more detail in Chapter 6.

Since the deflection of the walls of the microchannels by the SAWs is approximately $10^{-9}\,m$ compared to lengths, widths, and heights of the microchannels in the range of μm to mm, in the sequel we will neglect the time dependence of Ω_2.

The SAW-induced fluid flow exhibits two different time scales. When the SAWs enter the fluid-filled microchannels, sharp jets and vortices are created within nanoseconds (cf.

Fig. 1.1 and 6.1). The SAWs propagate along the channels and experience a significant damping which results in an almost stationary flow pattern, called acoustic streaming. This relaxation process happens on a time scale of milliseconds. The multiscale character can be appropriately taken care of by a homogenization approach. Following [10, 75], we introduce a scale parameter $0 < \varepsilon \ll 1$ which represents the maximum deflection of the walls of the microchannels, and we consider the asymptotic expansions

$$\begin{aligned} \rho_f &= \rho_{f,0} + \varepsilon\, \rho_f' + \varepsilon^2\, \rho_f'' + O(\varepsilon^3)\,, \\ \mathbf{v} &= \mathbf{v}_0 + \varepsilon\, \mathbf{v}' + \varepsilon^2\, \mathbf{v}'' + O(\varepsilon^3)\,, \\ p &= p_0 + \varepsilon\, p' + \varepsilon^2\, p'' + O(\varepsilon^3)\,. \end{aligned}$$

Collecting all terms of order $O(\varepsilon)$, assuming $\mathbf{v}_0 \equiv \mathbf{0}$ (fluid at rest, if no SAW actuation), and setting $\rho_{f,1} = \varepsilon\rho_f', \mathbf{v}_1 := \varepsilon\mathbf{v}', p_1 := \varepsilon p'$, we find that the triple $(\rho_{f,1}, \mathbf{v}_1, p_1)$ satisfies the linear system (for notation convenience we assume that this problem has time scale $[0, T_1]$)

$$\rho_{f,0}\frac{\partial \mathbf{v}_1}{\partial t} - \nabla\cdot\boldsymbol{\sigma}_1 = \mathbf{0} \qquad \text{in } \Omega_2\times(0,T_1], \tag{2.12a}$$

$$\frac{\partial \rho_{f,1}}{\partial t} + \rho_{f,0}\nabla\cdot\mathbf{v}_1 = 0 \qquad \text{in } \Omega_2\times(0,T_1], \tag{2.12b}$$

$$\mathbf{v}_1 = \mathbf{g}_1 \qquad \text{on } \Gamma_{2,D}\times(0,T_1], \tag{2.12c}$$

$$\boldsymbol{\sigma}_1\mathbf{n} = 0 \qquad \text{on } \Gamma_{2,N}\times(0,T_1], \tag{2.12d}$$

$$\mathbf{v}_1(\cdot,0) = \mathbf{0}, \quad p_1(\cdot,0) = 0 \qquad \text{in } \Omega_2, \tag{2.12e}$$

where $\boldsymbol{\sigma}_1 = ((\sigma_1)_{ij})_{i,j=1}^d$, $(\sigma_1)_{ij} := -p_1\,\delta_{ij} + 2\eta\varepsilon_{ij}(\mathbf{v}_1) + \delta_{ij}(\xi - 2\eta/3)\nabla\cdot\mathbf{v}_1$, $\mathbf{g}_1 := \partial\mathbf{u}/\partial t$ and where p_1 and $\rho_{f,1}$ are related by the constitutive equation

$$p_1 = c_0^2\,\rho_{f,1} \quad \text{in } \Omega_2\times(0,T_1]\,. \tag{2.13}$$

Here, c_0 stands for the small signal sound speed in the fluid. The system describes the propagation and damping of the acoustic waves in the microchannels.

We substitute $\rho_{f,1}$ in (2.12b) by means of (2.13) and introduce the function spaces

$$\begin{aligned}\mathbf{V}_{\mathbf{g}_1} &:= \{\mathbf{v} \in L^2((0,T_1);\mathbf{H}^1(\Omega_2)) \cap H^1((0,T_1);\mathbf{H}^{-1}(\Omega_2)) \mid \mathbf{v}|_{\Gamma_{2,D}} = \mathbf{g}_1\} \ , \\ W &:= H^1((0,T_1);L^2(\Omega_2)) \ .\end{aligned}$$

We note that $H^1((0,T_1);\mathbf{H}^{-1}(\Omega_2)) \cap L^2((0,T_1);\mathbf{H}^1(\Omega_2))$ is continuously embedded in $C([0,T_1],\mathbf{L^2}(\Omega_2))$. The weak formulation of (2.12a)-(2.12d) amounts to the computation of $(\mathbf{v}_1,p_1) \in \mathbf{V}_{\mathbf{g}_1} \times W$ such that for all $(\mathbf{w},q) \in \mathbf{H}^1_{0,\Gamma_{2,D}}(\Omega_2) \times L^2(\Omega_2)$ there holds

$$\langle \rho_{f,0} \frac{\partial \mathbf{v}_1}{\partial t}, \mathbf{w}\rangle + a(\mathbf{v}_1,\mathbf{w}) + b(p_1,\mathbf{w}) = \mathbf{0} \ , \qquad \mathbf{w} \in \mathbf{H}^1_{0,\Gamma_{2,D}}(\Omega_2), \tag{2.14a}$$

$$(\rho_{f,0}^{-1} c_0^{-2} \frac{\partial p_1}{\partial t}, q)_{0,\Omega_2} - b(q,\mathbf{v}_1) = 0 \ , \qquad q \in L^2(\Omega_2), \tag{2.14b}$$

$$\mathbf{v}_1(\cdot,0) = \mathbf{0}, \quad p_1(\cdot,0) = 0. \tag{2.14c}$$

Here, $\langle\cdot,\cdot\rangle$ stands for the respective dual pairing, and the bilinear forms $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$ are given by

$$a(\mathbf{v},\mathbf{w}) := \eta \int_{\Omega_2} \nabla\mathbf{v} : \nabla\mathbf{w}\, dx \ + \ (\xi + \frac{\eta}{3}) \int_{\Omega_2} \nabla\cdot\mathbf{v}\, \nabla\cdot\mathbf{w}\, dx \ , \tag{2.15a}$$

$$b(p,\mathbf{w}) := \ -\int_{\Omega_2} p\, \nabla\cdot\mathbf{w}\, dx \ . \tag{2.15b}$$

For time periodic excitations $\mathbf{g}_1$, we also consider the time periodic problem:

Find $(\mathbf{v}_1,p_1) \in \mathbf{V}_{\mathbf{g}_1} \times W$ such that for all $(\mathbf{w},q) \in \mathbf{H}^1_{0,\Gamma_{2,D}}(\Omega_2) \times L^2(\Omega_2)$ there holds

$$\langle \rho_{f,0} \frac{\partial \mathbf{v}_1}{\partial t}, \mathbf{w}\rangle + a(\mathbf{v}_1,\mathbf{w}) + b(p_1,\mathbf{w}) = \mathbf{0}, \tag{2.16a}$$

$$(\rho_{f,0}^{-1} c_0^{-2} \frac{\partial p_1}{\partial t}, q)_{0,\Omega_2} - b(q,\mathbf{v}_1) = 0, \tag{2.16b}$$

$$\mathbf{v}_1(\cdot,0) = \mathbf{v}_1(\cdot,T), \quad p_1(\cdot,0) = p_1(\cdot,T). \tag{2.16c}$$

Theorem 2.1.2 *For the solution of the variational problems* (2.14a)-(2.14c) *and* (2.16a)-(2.16c) *there holds:*

(i) If $\mathbf{g}_1 \in \mathbf{L}^2((0,T);\mathbf{H}^{1/2}_{00}(\Gamma_{2,D}))$ *then there exists a unique solution* $(\mathbf{v}_1, p_1) \in \mathbf{V}_{\mathbf{g}_1} \times W$ *of* (2.14a)-(2.14c) *satisfying the stability estimate*

$$\|(\mathbf{v}_1, p_1)\|_{\mathbf{V}_{\mathbf{g}_1}\times W} \;\leq\; C_{T_1}\, \|\mathbf{g}_1\|_{\mathbf{L}^2((0,T_1);\mathbf{H}^{1/2}_{00}(\Gamma_{2,D}))}\;, \tag{2.17}$$

where $C_{T_1} > 0$ *is a constant depending on* T_1.

(ii) If the forcing term $\mathbf{g}_1$ *is time periodic, then there exists a unique solution* $(\mathbf{v}_1, p_1)$ *of* (2.16a)-(2.16c).

Proof: Taking advantage of the ellipticity of the bilinear form $a(\cdot,\cdot)$, i.e.,

$$a(\mathbf{v},\mathbf{v}) \;\geq\; \alpha\, \|\mathbf{v}\|^2_{1,\Omega_2} \quad , \quad \alpha > 0\;,$$

and the fact that the bilinear form $b(\cdot,\cdot)$ satisfies the inf-sup condition

$$\inf_{q\in L^2_0(\Omega_2)\setminus\{0\}} \sup_{\mathbf{v}\in\mathbf{H}^1_{0,\Gamma_{2,D}}(\Omega_2)\setminus\{\mathbf{0}\}} \frac{b(q,\mathbf{v})}{\|q\|_{0,\Omega_2}\|\mathbf{v}\|_{1,\Omega_2}} \;\geq\; \beta \;>\; 0\;,$$

the existence of a solution $(\mathbf{v}_1, p_1) \in \mathbf{V}_{\mathbf{g}_1} \times W$ of (2.14a)-(2.14c) satisfying (2.17) can be shown by standard arguments based on the Galerkin method (cf., e.g., [95]). The uniqueness is an immediate consequence of (2.17). For the proof of (ii) we refer to Theorem 3.12 in [75]. □

Collecting all terms of order $O(\varepsilon^2)$, neglecting the time derivative with respect to the pressure, and performing the time-averaging

$$\langle w\rangle := T_1^{-1}\int_{t_0}^{t_0+T_1} w\, dt,$$

we arrive at the compressible Stokes system (for notation convenience we assume that this problem has a time scale $[0,T]$)

$$\rho_{f,0}\frac{\partial \mathbf{v}_2}{\partial t} - \nabla\cdot\boldsymbol{\sigma}_2 = \langle -\rho_{f,1}\frac{\partial \mathbf{v}_1}{\partial t} - \rho_{f,0}[\nabla \mathbf{v}_1]\mathbf{v}_1\rangle \qquad \text{in } \Omega_2\times(0,T], \tag{2.18a}$$

$$\rho_{f,0}\nabla\cdot\mathbf{v}_2 = \langle -\nabla\cdot(\rho_{f,1}\mathbf{v}_1)\rangle \qquad \text{in } \Omega_2\times(0,T], \tag{2.18b}$$

$$\mathbf{v}_2 = \mathbf{g}_2 \qquad \text{on } \Gamma_{2,D}\times(0,T], \tag{2.18c}$$

$$\boldsymbol{\sigma}_2\mathbf{n} = 0 \qquad \text{on } \Gamma_{2,N}\times(0,T], \tag{2.18d}$$

$$\mathbf{v}_2(\cdot,0) = \mathbf{0},\ p_2(\cdot,0) = 0 \qquad \text{in } \Omega_2, \tag{2.18e}$$

where $\mathbf{g}_2 := -\langle[\nabla\mathbf{v}_1]\mathbf{u}\rangle$ in (2.18c) and

$$\boldsymbol{\sigma}_2 = ((\sigma_2)_{ij})_{i,j=1}^d\ ,\ (\sigma_2)_{ij} := -p_2\ \delta_{ij} + 2\eta\varepsilon_{ij}(\mathbf{v}_2) + \delta_{ij}(\xi - 2\eta/3)\nabla\cdot\mathbf{v}_2.$$

The density $\rho_{f,2}$ can be obtained via the constitutive equation

$$p_2 \ = \ c_0^2\ \rho_{f,2} \quad \text{in } \Omega_2\times(0,T]\ . \tag{2.19}$$

The compressible Stokes system (2.18a)-(2.18d) is used as a model for the acoustic streaming.

The weak formulation of (2.18a)-(2.18d) requires the computation of $(\mathbf{v}_2,p_2)\in\mathbf{V}_{\mathbf{g}_2}\times W$, where

$$\begin{aligned}\mathbf{V}_{\mathbf{g}_2} &:= \{\mathbf{v}\in H^1((0,T),\mathbf{H}^{-1}(\Omega_2))\cap L^2((0,T),\mathbf{H}^1(\Omega_2))\mid \mathbf{v}|_{\Gamma_D} = \mathbf{g}_2\}\ ,\\ W &:= H^1((0,T);L^2(\Omega_2)),\end{aligned}$$

such that

$$\langle\rho_{f,0}\frac{\partial\mathbf{v}_2}{\partial t},\mathbf{w}\rangle + a(\mathbf{v}_2,\mathbf{w}) + b(p_2,\mathbf{w}) = (\mathbf{f},\mathbf{w})_{0,\Omega_2}, \qquad \mathbf{w}\in\mathbf{H}^1_{0,\Gamma_{2,D}}(\Omega_2), \tag{2.20a}$$

$$b(q,\mathbf{v}_2) = (f,q)_{0,\Omega_2}, \qquad q\in L^2(\Omega_2), \tag{2.20b}$$

$$\mathbf{v}_2(\cdot,0) = \mathbf{0}, \quad p_2(\cdot,0) = 0. \tag{2.20c}$$

Here, the bilinear forms $a(\cdot,\cdot), b(\cdot,\cdot)$ are as in (2.15a),(2.15b), and the right-hand sides $\mathbf{f}, f$ are given by

$$\mathbf{f} \;:=\; -\langle \rho_{f,1}\,\frac{\partial \mathbf{v}_1}{\partial t} + \rho_{f,0}\,[\nabla \mathbf{v}_1]\mathbf{v}_1\rangle \quad , \quad f \;:=\; -\langle \rho_{f,0}^{-1}\,\nabla\cdot(\rho_{f,1}\,\mathbf{v}_1)\rangle \;.$$

Theorem 2.1.3 *If* $\mathbf{f} \in \mathbf{L}^2(\Omega_2), f \in L^2(\Omega_2)$, *and* $\mathbf{g}_2 \in \mathbf{H}^{1/2}_{00}(\Gamma_{2,D})$, *the weak formulation* (2.20a),(2.20b) *of the compressible Stokes system admits a unique solution* $(\mathbf{v}_2, p_2) \in \mathbf{V}_{\mathbf{g}_2} \times W$. *Moreover, there exists a constant* $C_T > 0$ *depending on* T *such that*

$$\|(\mathbf{v}_2,p_2)\|_{\mathbf{V}_{\mathbf{g}_2}\times W} \;\leq\; C_T\,\Big(\|\mathbf{f}\|_{0,\Omega_2} \;+\; \|f\|_{0,\Omega_2} \;+\; \|\mathbf{g}_2\|_{\mathbf{H}^{1/2}_{00}(\Gamma_{2,D})}\Big). \tag{2.21}$$

Proof: The proof follows along the same lines as that of Theorem 2.1.2. □

2.2 Shape optimization

We have performed shape optimization of the walls of the microchannels and reservoirs using objective functionals of tracking type or representing the pumping rate at selected cross sections or minimizing the vorticity using (2.18a)-(2.18b) as state system.

The optimal design of the shape or topology of structures is an area within the theory of optimization with applications, e.g., in aero- and fluid dynamics, electromagnetics, and mechanics whose importance is reflected by a series of monographs on this topic that have been published during the past decades (cf. [9, 21, 22, 32, 57, 60, 61, 85, 91, 104]).

A typical shape optimization problem associated with a time-dependent PDE or a system thereof as the underlying state equation amounts to the minimization of a shape functional J over bounded domains $\Omega\times[0,T]$ with $\Omega \subset \mathbb{R}^d$ and $T > 0$. The state function $\mathbf{y}$

is assumed to satisfy an initial-boundary value problem as described by means of a partial differential operator L, and there may be further equality and/or inequality constraints on the domain. We also assume that the domain Ω does not vary with time. Then the shape optimization problem over a set of domains is

$$\inf_{\Omega} \; J(\mathbf{y}, \Omega), \quad J(\mathbf{y}, \Omega) \; := \int_0^T \int_{\Omega} \ell(x, t, \mathbf{y}(x,t)) \; dxdt, \tag{2.22a}$$

subject to

$$L\mathbf{y}(x,t) \; = \; \mathbf{f}(x,t) \qquad \text{in } \Omega \times (0,T], \tag{2.22b}$$

$$\mathbf{y}(x,t) \; = \; \mathbf{g}(x,t) \qquad \text{on } \Gamma \times (0,T], \tag{2.22c}$$

$$\mathbf{y}(x,0) \; = \; \mathbf{y}_0(x) \qquad \text{in } \Omega, \tag{2.22d}$$

$$h(x,t) \; \geq \; 0 \qquad \text{in } \Omega \times [0,T]. \tag{2.22e}$$

In order to cope with the inherent difficulty that the minimization is over a certain class of domains instead of a set of functions in an appropriate function space, we have used the classical approach based on a parametrization of the domain by a finite number of design variables: The boundary Γ is represented by a composite Bézier curve using a certain number of Bézier control points $\theta \in \Theta \subset \mathbb{R}^k, k \in \mathbb{N}$, which serve as design variables. We further assume that $\Theta := \{\theta_i \in \mathbb{R} \mid \theta_i^{min} \leq \theta_i \leq \theta_i^{max}, 1 \leq i \leq k\}$ is independent of time and is closed and convex. The equality and/or inequality constraints are expressed by means of the design variables. The above shape optimization amounts to solving

$$\inf_{\theta \in \Theta} \; J(\theta), \quad J(\theta) \; := \int_0^T \int_{\Omega(\theta)} \ell(x, t, \theta, \mathbf{y}(x,t)) \; dxdt, \tag{2.23a}$$

subject to

$$
\begin{aligned}
L\mathbf{y}(x,t) &= \mathbf{f}(x,t) && \text{in } \Omega(\theta)\times(0,T], && (2.23b)\\
\mathbf{y}(x,t) &= \mathbf{g}(x,t) && \text{on } \Gamma(\theta)\times(0,T], && (2.23c)\\
\mathbf{y}(x,0) &= \mathbf{y}_0(x) && \text{in } \Omega(\theta), && (2.23d)\\
\theta &\in \Theta. && && (2.23e)
\end{aligned}
$$

For the finite element approximation of (2.23a)-(2.23c) we choose $\widehat{\theta}$ as reference design variable and refer to $\widehat{\Omega} := \Omega(\widehat{\theta})$ as the associated reference domain. Then, the actual domain $\Omega(\theta)$ can be obtained from the reference domain $\widehat{\Omega}$ by means of a mapping $\Omega(\theta) = \Phi(\widehat{\Omega};\theta)$. The advantage of using the reference domain $\widehat{\Omega}$ is that finite element approximations can be performed with respect to that fixed domain without being forced to remesh for every new set of the design variables.

An alternative approach would be the use of shape calculus [32], i.e., of shape gradients and shape Hessians in case of interior-point methods requiring first- and second-order information. Since stable numerical implementations of shape Hessians were not available at the time the research has been conducted, we decided to use the standard approach described above.

CHAPTER 3

Optimal design of stationary Stokes flow by path-following interior-point methods

Recall that in Chapter 2 we derived the acoustic streaming system i.e., the compressible Stokes system given by (2.18). The system derived generates almost a stationary pattern called acoustic streaming. Keeping in mind our goal i.e., modeling, simulation, and shape optimization of microfluidic biochips, as a first step in this chapter we consider shape optimization problems governed by incompressible stationary Stokes system. The key features of this chapter are all-at-once approach, primal-dual barrier method in terms of a path-following predictor-corrector continuation scheme with adaptive choice of the continuation steplength, extension of barrier method to multilevel framework and comparison with the classical primal-dual path-following methods. This chapter is based on [16] and [17].

Optimal design problems associated with fluid flow problems play a decisive role in a wide variety of engineering applications (cf., e.g., [85] and the references therein). A typical example is to design the geometry of the container of the fluid, e.g., a channel, a reservoir, or a network of channels and reservoirs, in such a way that a desired flow velocity and/or pressure profile is achieved. The solution of the problem amounts to the minimization of an objective functional that depends on the so-called state variables (velocity, pressure) and on the design variables which determine the geometry of the fluid-filled domain. The state variables are supposed to satisfy the underlying fluid mechanical equations, and there are typically further technologically motivated constraints, e.g., bilateral constraints on the design variables which restrict the shape of the fluid-filled domain to that what is technologically feasible.

Shape optimization problems have been extensively studied and are well documented in the literature (cf., e.g., the monographs [9, 21, 22, 29, 32, 60, 61, 85, 91, 99, 104]). The traditional approach relies on a separate treatment of the design objective and the state equation by an iterative cycle that starts from a given design, computes an approximate solution of the state equation for that design, invokes some sensitivity analysis for an update of the design, and continues this way until convergence is achieved. In contrast to this successive approximation, recently so-called 'all-at-once methods' or 'one-shot methods' have attracted considerable attention in PDE constrained optimization whose characteristic feature is that the numerical solution of the state equation is an integral part of the optimization routine. In particular, it has been shown that this novel approach may lead to significant savings of computational time (see e.g. [6, 7, 8, 66, 68, 69, 102]).

In this chapter, we consider the optimal design of stationary fluid flow problems as described by the Stokes system. The objective is to design the geometry of a channel

or a particular geometric feature of a channel such that a desired profile of the velocity and/or the pressure is realized as closely as possible. The design variables are chosen as the Bézier control points of a globally continuous Bézier curve representation of the walls of the channel. The control points are subject to bilateral constraints. For instance, for the shape optimization of a backward facing step (cf. Fig. 3.1), we use k_B number of control points for the lower backward facing step (curved line).

The approach that we are pursuing here is an 'all-at-once method' based on a primal-dual formulation where the Stokes system is coupled by Lagrange multipliers and the constraints on the design variables are taken care of by parameterized logarithmic barrier functions.

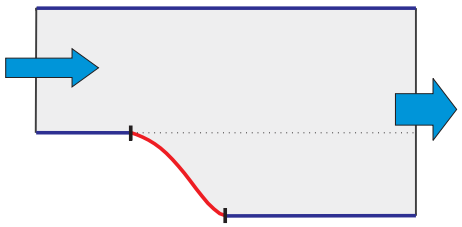


Figure 3.1: Channel with a backward facing step

This leads to a family of minimization subproblems parameterized by the barrier parameter. The optimality conditions result in a parameter dependent nonlinear system whose solution gives rise to the so-called central path (cf., e.g., [46, 120]). A significant challenge is to follow the central path as closely as possible as the barrier parameter goes to zero. Here, we use three path-following strategies. The first one is an adaptive continuation method with tangent continuation as a predictor and Newton's method as a corrector following the ideas from [33], whereas the second and third ones are variants of the long-step target following algorithm and Mehrotra's algorithm known from linear programming (cf., e.g., [121]). We note that path-following algorithms for shape optimization problems

in structural mechanics have been used in [63].

3.1 The shape optimization of parametrized Stokes equations

Let $\Omega(\theta) \subset \mathbb{R}^d$ be a bounded domain that depends on design variables $\theta = (\theta_1, \cdots, \theta_k)^T \in \Theta$, where $\Theta \subset \mathbb{R}^k$ is a given convex set, $\theta_i, 1 \leq i \leq k$, are the Bézier control points of a Bézier curve representation of the boundary and $\Theta := \{\theta_i \in \mathbb{R} \mid \theta_i^{min} \leq \theta_i \leq \theta_i^{max}, 1 \leq i \leq k\}$. We assume that the boundary $\partial\Omega(\theta)$ consists of an inflow boundary $\Gamma_{\text{in}}(\theta)$, an outflow boundary $\Gamma_{\text{out}}(\theta)$, and a lateral boundary $\Gamma_{\text{lat}}(\theta)$ such that $\partial\Omega(\theta) = \overline{\Gamma}_{\text{in}}(\theta) \cup \overline{\Gamma}_{\text{out}}(\theta) \cup \overline{\Gamma}_{\text{lat}}(\theta), \Gamma_{\text{in}}(\theta) \cap \Gamma_{\text{out}}(\theta) \cap \Gamma_{\text{lat}}(\theta) = \emptyset$. Consider shape optimization problems associated with the stationary Stokes system of the form

$$\inf_{\mathbf{v}\in\mathbf{H}^1(\Omega), p\in L^2(\Omega), \theta\in\Theta} J(\mathbf{v}, p, \theta) \tag{3.1a}$$

where

$$J(\mathbf{v}, p, \theta) := \int_{\Omega(\theta)} \ell(\mathbf{v}(\theta), p(\theta), x, \theta)\, dx, \tag{3.1b}$$

and where $\mathbf{v}(\theta), p(\theta)$ solve

$$-\nu\, \Delta\mathbf{v}(x) + \nabla p(x) = \mathbf{f}(x), \qquad x \in \Omega(\theta), \tag{3.1c}$$

$$\nabla \cdot \mathbf{v}(x) = 0, \qquad x \in \Omega(\theta), \tag{3.1d}$$

$$\mathbf{v}(x) = \mathbf{v}_{\text{in}}(x), \qquad x \in \Gamma_{\text{in}}(\theta), \tag{3.1e}$$

$$\mathbf{v}(x, t) = 0, \qquad x \in \Gamma_{\text{lat}}(\theta), \tag{3.1f}$$

$$(\nabla\mathbf{v}(x) - p(x)I)\mathbf{n} = 0\ , \qquad x \in \Gamma_{\text{out}}(\theta). \tag{3.1g}$$

where $\mathbf{n}$ and ν denoting the exterior unit normal vector and viscosity of the fluid, respectively and $\mathbf{f}$ denotes the forcing term. It is well-known that the weak formulation of (3.1c-g) admits a unique solution (cf., e.g., [80]).

For the finite element approximation of (3.1) we choose $\widehat{\theta} \in \Theta$ as a reference design and refer to $\widehat{\Omega} := \Omega(\widehat{\theta})$ as the associated reference domain. Then, the actual domain $\Omega(\theta)$ can be obtained from the reference domain $\widehat{\Omega}$ by means of a mapping $\Omega(\theta) = \Phi(\widehat{\Omega}; \theta)$. The advantage of using the reference domain $\widehat{\Omega}$ is that finite element approximations can be performed with respect to that fixed domain without being forced to remesh for every new set of the design variables.

For the discretization of the stationary Stokes system we use one of the many standard methods [51], such as the classical P2-P1 Taylor Hood element, or methods with discontinuous pressure discretizations. We assume that the simplicial triangulation $\mathcal{T}_h$ of the spatial domain $\Omega(\theta)$ is geometrically conforming and aligns with $\Gamma_{\text{in}}(\theta), \Gamma_{\text{lat}}(\theta)$ and $\Gamma_{\text{out}}(\theta)$. This leads to the discrete optimization problem

$$\inf_{\mathbf{v}\in\mathbb{R}^n, p\in\mathbb{R}^m, \theta\in\Theta} J(\mathbf{v}, p, \theta) \tag{3.2a}$$

where

$$J(\theta) := \ \boldsymbol{\ell}(\mathbf{v}(\theta), p(\theta), x, \theta), \tag{3.2b}$$

and where $\mathbf{v}(\theta), p(\theta)$ solve

$$\mathbf{S}(\theta) \begin{pmatrix} \mathbf{v} \\ p \end{pmatrix} = \begin{pmatrix} \mathbf{g}_1(\theta) \\ \mathbf{g}_2(\theta) \end{pmatrix} . \tag{3.2c}$$

For ease of notation, we have kept the same symbol for the velocity $\mathbf{v}$ and pressure p as it was there in the continuous setting. Here, $\boldsymbol{\ell}(\cdot)$ in (3.2b) results from the spatial discretization of the integral of the objective functional in (3.1a). The discrete Stokes

operator $\mathbf{S}(\theta)$ in (3.2c) is given by

$$\mathbf{S}(\theta) := \begin{pmatrix} \mathbf{A}(\theta) & \mathbf{B}^T(\theta) \\ \mathbf{B}(\theta) & \mathbf{0} \end{pmatrix}, \tag{3.3}$$

where $\mathbf{A}(\theta) \in \mathbb{R}^{n\times n}$ and $\mathbf{B}(\theta) \in \mathbb{R}^{m\times n}$ are the stiffness matrix, and the matrix representation of the discrete divergence operator. The vector $\mathbf{g}_2(\theta) \in \mathbb{R}^m$ in (3.2c) stems from the semi-discretization of the incompressibility condition due to the boundary condition at the inflow boundary. We note that the data of the discrete problem depend on the design variable θ due to the dependence of the spatial domain on θ.

Assume that $\theta \in \Theta$ satisfies the so-called box constraints i.e.,

$$\Theta := \{\theta_i \in \mathbb{R} \mid \theta_i^{min} \leq \theta_i \leq \theta_i^{max}, 1 \leq i \leq k\}. \tag{3.4}$$

Then due to the nonlinear dependence on the design variables, (3.2) represents an inequality constrained nonlinear programming problem. It will be numerically solved by path-following primal-dual interior-point methods as described in the next section.

3.2 Path-following interior-point methods

Introduce Lagrange multipliers $\boldsymbol{\lambda}_{\mathbf{v}} \in \mathbb{R}^n$, $\lambda_p \in \mathbb{R}^m$ for the PDE constraints (3.2c) and couple the inequality constraints (3.4) by logarithmic barrier functions with a barrier parameter $\beta = 1/\mu > 0$, $\mu \to \infty$,. This leads to the saddle point problem

$$\inf_{\mathbf{y},\theta} \sup_{\boldsymbol{\lambda}} \mathcal{L}^{(\mu)}(\mathbf{y}, \boldsymbol{\lambda}, \theta), \tag{3.5}$$

where $\mathbf{y} := (\mathbf{v}, p)$, $\boldsymbol{\lambda} := (\boldsymbol{\lambda}_{\mathbf{v}}, \lambda_p)$ and $\mathcal{L}^{(\mu)}$ stands for the Lagrangian

$$\mathcal{L}^{(\mu)}(\mathbf{y}, \boldsymbol{\lambda}, \theta) = \mathcal{B}^{(\mu)}(\mathbf{y}, \theta) + \boldsymbol{\lambda}^T\big(\mathbf{S}(\theta)\mathbf{y} - \mathbf{g}\big), \tag{3.6}$$

and $\mathcal{B}^{(\mu)}(\mathbf{y}, \theta)$ is the so-called barrier function as given by

$$\mathcal{B}^{(\mu)}(\mathbf{y}, \theta) \; := J(\mathbf{y}, \theta) \; - \; \frac{1}{\mu} \sum_{i=1}^{m} \left[\ln(\theta_i - \theta_i^{min}) + \ln(\theta_i^{max} - \theta_i) \right]. \tag{3.7}$$

(for details cf., e.g., [120]) and $\mathbf{g} = (\mathbf{g}_1, \mathbf{g}_2)^T$.

The central path $\mu \longmapsto \mathbf{x}(\mu) := \big(\mathbf{y}(\mu), \boldsymbol{\lambda}(\mu), \theta(\mu)\big)^T$ is given as the solution of the nonlinear system

$$\mathcal{F}(\mathbf{x}(\mu), \mu) \; = \; \begin{pmatrix} \mathcal{L}_{\mathbf{y}}^{(\mu)}(\mathbf{y}, \boldsymbol{\lambda}, \theta) \\ \mathcal{L}_{\boldsymbol{\lambda}}^{(\mu)}(\mathbf{y}, \boldsymbol{\lambda}, \theta) \\ \mathcal{L}_{\theta}^{(\mu)}(\mathbf{y}, \boldsymbol{\lambda}, \theta) \end{pmatrix} \; = \; \mathbf{0}, \tag{3.8}$$

where the subindices refer to the derivatives of the Lagrangian with respect to the primal, the dual, and the design variables. The choice of the barrier parameter strongly influences the performance of the interior-point method. There are static strategies with the Fiacco-McCormick approach as the most prominent one (cf. [44]), where the barrier parameter is fixed until an approximate solution of (3.5) has been obtained, and there is a variety of dynamic update strategies (cf. [20, 39, 50, 88, 111, 113, 114]). Convergence properties of the Fiacco-McCormick approach have been studied in [27] and [115], whereas a convergence analysis of dynamic update strategies has been addressed in [20, 39, 88, 113].

We consider the solution of (3.8) by an adaptive continuation method based on the affine invariant convergence theory of Newton-type methods and nonlinear variants of the long-step and Mehrotra's path-following method (see [84]).

3.2.1 Adaptive continuation method

The adaptive continuation method is a predictor-corrector method with an adaptively determined continuation step size in the prediction step and Newton's method as a corrector.

It relies on the affine invariant convergence theory of Newton and Newton-type methods (cf., e.g., [33]) and ensures that the iterates stay within a neighborhood (contraction tube) of the central path so that convergence to a local minimum of the original minimization problem can be achieved (cf. Fig. 3.2).

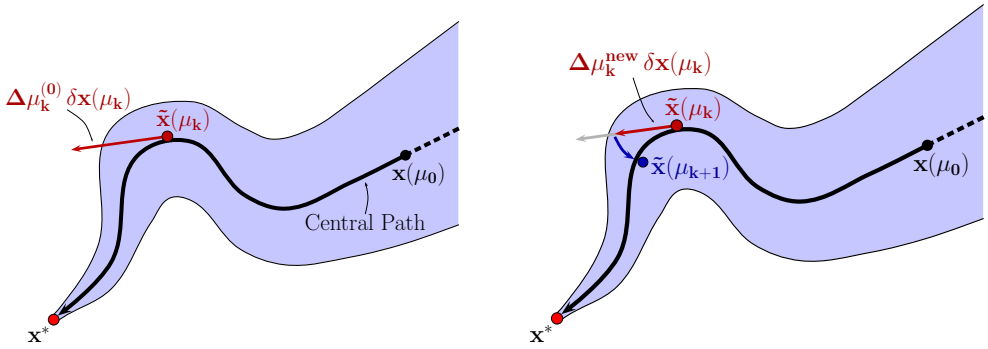


Figure 3.2: Adaptive continuation method: Prediction step (left) and correction step (right)

Prediction step: The predictor step relies on tangent continuation along the trajectory of the Davidenko equation

$$\mathcal{F}_{\mathbf{x}}(\mathbf{x}(\mu),\mu)\;\mathbf{x}'(\mu) \;=\; -\mathcal{F}_{\mu}(\mathbf{x}(\mu),\mu) \tag{3.9}$$

and amounts to the implementation of an explicit Euler step: Given some approximation $\widetilde{\mathbf{x}}(\mu_k)$ at $\mu_k > 0$, compute $\widetilde{\mathbf{x}}^{(j_0)}(\mu_{k+1})$, where $\mu_{k+1} = \mu_k + \Delta\mu_k^{(j)}$, according to

$$\mathcal{F}_{\mathbf{x}}(\widetilde{\mathbf{x}}(\mu_k),\mu_k)\;\delta\mathbf{x}(\mu_k) \;=\; -\;\mathcal{F}_{\mu}(\widetilde{\mathbf{x}}(\mu_k),\mu_k)\;, \tag{3.10a}$$

$$\widetilde{\mathbf{x}}^{(j_0)}(\mu_{k+1}) \;=\; \widetilde{\mathbf{x}}(\mu_k) \;+\; \Delta\mu_k^{(j)}\;\delta\mathbf{x}(\mu_k)\;, \tag{3.10b}$$

starting with $j = 0$ ($j \geq 1$ only if required by the correction step (see below)). We use $\Delta\mu_0^{(0)} = \Delta\mu_0$ for some given initial step size $\Delta\mu_0$, whereas for $k \geq 1$ the predicted step size $\Delta\mu_k^{(0)}$ is chosen by

$$\Delta\mu_k^{(0)} \;:=\; \Big(\frac{\|\Delta\mathbf{x}^{(j_0)}(\mu_k)\|}{\|\widetilde{\mathbf{x}}(\mu_k) - \widetilde{\mathbf{x}}^{(j_0)}(\mu_k)\|}\;\frac{\sqrt{2}-1}{2\Pi(\mu_k)}\Big)^{1/2}\;\Delta\mu_{k-1}\;, \tag{3.11}$$

where $\Delta\mu_{k-1}$ is the computed continuation step size, $\Delta\mathbf{x}^{(j_0)}(\mu_k)$ is the first Newton correction (see below), and $\Pi(\mu_k) < 1$ is the contraction factor associated with a successful previous continuation step.

Correction step: As a corrector, we use Newton's method applied to $\mathcal{F}(\mathbf{x}(\mu_{k+1}), \mu_{k+1}) = 0$ with $\widetilde{\mathbf{x}}^{(j_0)}(\mu_{k+1})$ from (3.10) as a start vector. In particular, for $\ell \geq 0$ (Newton iteration index) and $j_\ell \geq 0$ (j being the steplength correction index) we compute $\Delta\mathbf{x}^{(j_\ell)}(\mu_{k+1})$ according to

$$\mathcal{F}_{\mathbf{x}}(\widetilde{\mathbf{x}}^{(j_\ell)}(\mu_{k+1}), \mu_{k+1})\ \Delta\mathbf{x}^{(j_\ell)}(\mu_{k+1}) = -\ \mathcal{F}(\widetilde{\mathbf{x}}^{(j_\ell)}(\mu_{k+1}), \mu_{k+1})\ , \tag{3.12}$$

update $\widetilde{\mathbf{x}}^{(j_{\ell+1})}(\mu_{k+1}) := \widetilde{\mathbf{x}}^{(j_\ell)}(\mu_{k+1}) + \Delta\mathbf{x}^{(j_\ell)}(\mu_{k+1})$ and compute $\overline{\Delta\mathbf{x}}^{(j_\ell)}(\mu_{k+1})$ as the associated simplified Newton correction

$$\mathcal{F}_{\mathbf{x}}(\widetilde{\mathbf{x}}^{(j_\ell)}(\mu_{k+1}), \mu_{k+1})\ \overline{\Delta\mathbf{x}}^{(j_\ell)}(\mu_{k+1}) = -\ \mathcal{F}(\widetilde{\mathbf{x}}^{(j_\ell)}(\mu_{k+1}) + \Delta\mathbf{x}^{(j_\ell)}(\mu_{k+1}), \mu_{k+1})\ . \tag{3.13}$$

We monitor convergence of Newton's method by means of

$$\Pi^{(j_\ell)}(\mu_{k+1})\ :=\ \|\overline{\Delta\mathbf{x}}^{(j_\ell)}(\mu_{k+1})\| / \|\Delta\mathbf{x}^{(j_\ell)}(\mu_{k+1})\|\ .$$

In case of successful convergence, we set $\widetilde{\mathbf{x}}(\mu_{k+1}) := \widetilde{\mathbf{x}}^{(j_\ell)}(\mu_{k+1})$ with ℓ being the current Newton iteration index, accept the current step size $\Delta\mu_k := \Delta\mu_k^{(j)}$ with current steplength correction index j and proceed with the next continuation step. However, if the monotonicity test

$$\Pi^{(j_\ell)}(\mu_{k+1})\ <\ 1 \tag{3.14}$$

fails for some $j_\ell \geq 0$, the predicted steplength $\Delta\mu_k^{(j)}$ has been chosen too large so that the predicted solution $\widetilde{\mathbf{x}}^{(j_0)}(\mu_{k+1})$ is not situated within the Kantorovich neighborhood of $\mathbf{x}(\mu_{k+1})$, i.e., it is outside the contraction tube around the central path (cf. Fig. 3.2 (right)). The corrector step provides a correction of the steplength for the tangent direction $\delta\mathbf{x}(\mu_k)$

such that the new iterate stays within the contraction tube. To do so, the continuation step from (3.10b) has to be repeated with the reduced step size

$$\begin{aligned} \Delta\mu_k^{(j+1)} &:= \Big(\frac{\sqrt{2}-1}{g(\Pi^{(j_\ell)})}\Big)^{1/2} \Delta\mu_k^{(j)} \ , \\ g(\Pi) &:= \sqrt{\Pi+1}-1 \end{aligned} \tag{3.15}$$

until we either achieve convergence or for some prespecified lower bound $\Delta\mu_{min}$ observe

$$\Delta\mu_k^{(j+1)} \ < \ \Delta\mu_{min} \ .$$

In the latter case, we stop the algorithm and report convergence failure.

A Newton step: The Newton steps are realized by an inexact Newton method featuring right-transforming iterations (cf., e.g., [66, 68, 77]). Consider for example we want to solve the KKT system (3.12). The block structure of this system is given by (for notational convenience, in the following the upper index (μ) in the Lagrangian will be dropped)

$$\overbrace{\begin{pmatrix} \nabla_{\mathbf{y},\mathbf{y}} J(\theta) & \mathbf{S}(\theta) & \mathcal{L}_{\mathbf{y},\theta} \\ \mathbf{S}(\theta) & \mathbf{0} & \mathcal{L}_{\boldsymbol{\lambda},\theta} \\ \mathcal{L}_{\theta,\mathbf{y}} & \mathcal{L}_{\theta,\boldsymbol{\lambda}} & \mathcal{L}_{\theta,\theta} \end{pmatrix}}^{\mathcal{F}_{\mathbf{x}}} \overbrace{\begin{pmatrix} \Delta\mathbf{y} \\ \Delta\boldsymbol{\lambda} \\ \Delta\theta \end{pmatrix}}^{\Delta\mathbf{x}} = -\overbrace{\begin{pmatrix} \nabla_{\mathbf{y}}\mathcal{L} \\ \nabla_{\boldsymbol{\lambda}}\mathcal{L} \\ \nabla_{\theta}\mathcal{L} \end{pmatrix}}^{\mathcal{F}} . \tag{3.16}$$

Introduce the positive auxiliary variables $\mathbf{z} = (\mathbf{z}^{(1)}, \mathbf{z}^{(2)})^T$ with entries

$$z_i^{(1)} := \frac{1}{\mu(\theta_i - \theta_i^{min})}, \qquad z_i^{(2)} := \frac{1}{\mu(\theta_i^{max} - \theta_i)}, \qquad 1 \le i \le k \ , \tag{3.17}$$

which are well-defined as $\mu > 0$ and for $\theta \in \mathrm{int}(\Theta)$. We can rewrite (3.17) as

$$\mathbf{D}_1\mathbf{z}^{(1)} = \frac{1}{\mu}\mathbf{e}, \qquad \mathbf{D}_1 := \mathrm{diag}(\theta_1 - \theta_1^{min}, ..., \theta_k - \theta_k^{min}), \tag{3.18a}$$

$$\mathbf{D}_2\mathbf{z}^{(2)} = \frac{1}{\mu}\mathbf{e}, \qquad \mathbf{D}_2 := \mathrm{diag}(\theta_1^{max} - \theta_1, ..., \theta_k^{max} - \theta_k), \tag{3.18b}$$

where $\mathbf{e} := (1, ..., 1)^T \in \mathbb{R}^k$. The Newton linearization corresponding to (3.18) is given by

$$\mathbf{Z}_1 \Delta\theta + \mathbf{D}_1 \Delta\mathbf{z}^{(1)} = -(\mathbf{D}_1 \mathbf{z}^{(1)} - \frac{1}{\mu}\mathbf{e}) \tag{3.19a}$$

$$-\mathbf{Z}_2 \Delta\theta + \mathbf{D}_2 \Delta\mathbf{z}^{(2)} = -(\mathbf{D}_2 \mathbf{z}^{(2)} - \frac{1}{\mu}\mathbf{e}) \tag{3.19b}$$

where $\mathbf{Z}_i := \mathrm{diag}(\mathbf{z}_1^{(i)}, ..., \mathbf{z}_k^{(i)})$, $1 \leq i \leq 2$. The new KKT system reads as follows

$$\overbrace{\begin{pmatrix} \nabla_{\mathbf{y},\mathbf{y}} J(\theta) & \mathbf{S}(\theta) & \mathcal{L}_{\mathbf{y},\theta} & \mathbf{0} \\ \mathbf{S}(\theta) & \mathbf{0} & \mathcal{L}_{\boldsymbol{\lambda},\theta} & \mathbf{0} \\ \mathcal{L}_{\theta,\mathbf{y}} & \mathcal{L}_{\theta,\boldsymbol{\lambda}} & \mathcal{L}_{\theta,\theta} & \widehat{\mathbf{I}} \\ \mathbf{0} & \mathbf{0} & \widehat{\mathbf{Z}} & \mathbf{D} \end{pmatrix}}^{\mathbf{K}} \begin{pmatrix} \Delta\mathbf{y} \\ \Delta\boldsymbol{\lambda} \\ \Delta\theta \\ \Delta\mathbf{z} \end{pmatrix} = - \begin{pmatrix} \nabla_{\mathbf{y}}\mathcal{L} \\ \nabla_{\boldsymbol{\lambda}}\mathcal{L} \\ \nabla_{\theta}\mathcal{L} \\ \widehat{\mathbf{D}}\mathbf{z} - 1/\mu\ \mathbf{e} \end{pmatrix}. \tag{3.20}$$

where $\mathbf{K}$ is the KKT matrix and

$$\widehat{\mathbf{D}} := \begin{pmatrix} \mathbf{D}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_2 \end{pmatrix}, \quad \widehat{\mathbf{I}} := \begin{pmatrix} -\mathbf{I}_k \\ \mathbf{I}_k \end{pmatrix}^T, \quad \widehat{\mathbf{Z}} := \begin{pmatrix} \mathbf{Z}_1 \\ -\mathbf{Z}_2 \end{pmatrix}.$$

We then apply the condensation and right-transforming iterations to solve (3.20). For complete details, we refer to [11, 77].

The first- and second-order derivatives w.r.t. design variables θ occurring in the KKT system are computed by automatic differentiation (cf., e.g., [55]) based on the automatic differentiation package from the INTLAB toolbox (see [100]).

3.2.2 Long-step path-following method

Long-step path-following method amounts to compute an increment $\Delta\mathbf{x} := (\Delta\mathbf{y}, \Delta\boldsymbol{\lambda}, \Delta\theta, \Delta\mathbf{z})^T$ by solving the KKT (Newton's) system as before, namely,

$$\mathbf{K}\ \Delta\mathbf{x} = -\widetilde{\mathbf{g}}, \tag{3.21}$$

where, $\widetilde{\mathbf{g}} := (\mathbf{g}, \mathbf{D}_1\mathbf{z}^{(1)} - \sigma\mu^{-1}\mathbf{e}, \mathbf{D}_2\mathbf{z}^{(2)} - \sigma\mu^{-1}\mathbf{e})^T$, with an additional centering parameter $\sigma > 0$.

We define $\mathcal{N}_{-\infty}(\gamma), 0 < \gamma \ll 1$, as the following neighborhood of the central path

$$\mathcal{N}_{-\infty}(\gamma) := \left\{(\mathbf{y}, \boldsymbol{\lambda}, \theta, \mathbf{z}) \mid \mathbf{s}_\theta^T\mathbf{z} \geq \mu^{-1}\gamma\right\} \ , \tag{3.22}$$

where $\mathbf{s}(\theta) := (s^{(1)}(\theta), s^{(2)}(\theta))^T, s^{(1)}(\theta) := \theta - \theta^{min}, s^{(2)}(\theta) := \theta^{max} - \theta$.

The long-step path-following algorithm proceeds as follows:

Initialization: Specify $0 < \gamma \ll 1$, bounds $0 < \sigma_{min} < \sigma_{max} < 1$ for the centering parameter, and choose a start iterate

$$\mathbf{x}^{(0)} = (\mathbf{y}^{(0)}, \boldsymbol{\lambda}^{(0)}, \theta^{(0)}, \mathbf{z}^{(0)}) \in \mathcal{N}_{-\infty}(\gamma) \ .$$

Iteration loop: For $j = 0, 1, 2, \ldots$ set

$$\mu^{(j)} := \max\left(\frac{k}{s^{(1)}(\theta^{(j)})^T z^{(1)}(\theta^{(j)})}, \frac{k}{s^{(2)}(\theta^{(j)})^T z^{(2)}(\theta^{(j)})}\right). \tag{3.23}$$

Choose $\sigma^{(j)} \in (\sigma_{min}, \sigma_{max})$, and compute

$$\Delta\mathbf{x}^{(j)} = (\Delta\mathbf{y}^{(j)}, \Delta\boldsymbol{\lambda}^{(j)}, \Delta\theta^{(j)}, \Delta\mathbf{z}^{(j)})$$

as the solution of (3.21). Set the next iterate as

$$\mathbf{x}^{(j+1)} = \mathbf{x}^{(j)} + \eta^{(j)}\Delta\mathbf{x}^{(j)},$$

where, $\eta^{(j)} = \max\{\eta \in (0,1) \mid \mathbf{x}^{(j)} + \eta\Delta\mathbf{x}^{(j)} \in \mathcal{N}_{-\infty}(\gamma)\}$.

Given a tolerance ϵ, the iteration will be terminated, if for some $j^* \geq 1$

$$\Delta J := \left|J(\mathbf{y}^{(j^*)}, \theta^{(j^*)}) - J(\mathbf{y}^{(j^*-1)}, \theta^{(j^*-1)})\right| \ < \ \epsilon.$$

For $j \geq 1$, a possible choice of the centering parameter $\sigma^{(j)}$ is $\sigma^{(j)} := \left(\mu^{(j-1)}/\mu^{(j)}\right)^2$. The solution of (3.21) is computed based on static condensation of the slack variables and the application of right-transforming iterations to the resulting reduced Hessian system (cf., e.g., [66, 77]).

3.2.3 Nonlinear version of Mehrotra's method

Using the same notation as in Subsection 3.2.2, the nonlinear version of Mehrotra's method (cf. [84]) is as follows:

Initialization: Choose a start iterate

$$\mathbf{x}^{(0)} = (\mathbf{y}^{(0)}, \boldsymbol{\lambda}^{(0)}, \theta^{(0)}, \mathbf{z}^{(0)}) \ .$$

Iteration loop: For $j = 0, 1, 2, \ldots$ define $\mu^{(j)}$ as in (3.23) and solve (3.21) with $\sigma = 0$ for affine scaling direction $\Delta\mathbf{x}_{\text{aff}}^{(j)} = \left(\Delta\mathbf{y}_{\text{aff}}^{(j)}, \Delta\boldsymbol{\lambda}_{\text{aff}}^{(j)}, \Delta\theta_{\text{aff}}^{(j)}, \Delta\mathbf{z}_{\text{aff}}^{(j)}\right)^T$.

Choose step lengths for the primal (p) and dual (d) components as

$$\beta_{\text{aff}}^{p} = \max\left\{\beta \in (0,1) \ \middle| \ \theta^{(j)} + \beta\Delta\theta_{\text{aff}}^{(j)} \geq 0\right\},$$
$$\beta_{\text{aff}}^{d} = \max\left\{\beta \in (0,1) \ \middle| \ \mathbf{z}^{(j)} + \beta\Delta\mathbf{z}_{\text{aff}}^{(j)} \geq 0\right\}.$$

Choose centering parameter adaptively by first setting

$$\mu_{\text{aff}} = \max_{1\leq\nu\leq 2}\left\{\frac{k}{\left(s^{(\nu)}(\theta^{(j)}) + \beta_{\text{aff}}^{p}\Delta\theta_{\text{aff}}^{(j)}\right)^T\left(\mathbf{z}^{(\nu)}(\theta^{(j)}) + \beta_{\text{aff}}^{d}\Delta\mathbf{z}_{\text{aff}}^{(\nu),(j)}\right)}\right\},$$

then $\sigma^{(j)} = (\mu^{(j)}/\mu_{aff})^3$.

Solve (3.21) for $(\Delta\mathbf{y}^{(j)}, \Delta\boldsymbol{\lambda}^{(j)}, \Delta\theta^{(j)}, \Delta\mathbf{z}^{(j)})^T$ with right-hand side $\widetilde{\mathbf{g}}$ replaced by

$$\widetilde{\mathbf{g}} = \left(\mathbf{g}, \mathbf{D}_1\mathbf{z}^{(1)} + \Delta\mathbf{D}_1\Delta\mathbf{z}_{\text{aff}}^{(1),(j)} - \sigma^{(j)}{\mu^{(j)}}^{-1}\mathbf{e}, \mathbf{D}_2\mathbf{z}^{(2)} + \Delta\mathbf{D}_2\Delta\mathbf{z}_{\text{aff}}^{(2),(j)} - \sigma^{(j)}{\mu^{(j)}}^{-1}\mathbf{e}\right)^T,$$

where

$$\Delta\mathbf{D}_1 = \operatorname{diag}\left(\theta_i^{(j)} + \beta_{\text{aff}}^p \ \Delta\theta_{i,\text{aff}}^{(j)} - \theta_i^{min}\right), \qquad 1 \le i \le k,$$
$$\Delta\mathbf{D}_2 = \operatorname{diag}\left(\theta_i^{max} - (\theta_i^{(j)} + \beta_{\text{aff}}^p \ \Delta\theta_{i,\text{aff}}^{(j)})\right), \qquad 1 \le i \le k.$$

Again choose the step lengths as

$$\bar{\beta}^p = \max\left\{\beta \in \mathbb{R}^+ \mid \theta^{(j)} + \beta\Delta\theta^{(j)} \ge 0\right\},$$
$$\bar{\beta}^d = \max\left\{\beta \in \mathbb{R}^+ \mid \mathbf{z}^{(j)} + \beta\Delta\mathbf{z}^{(j)} \ge 0\right\},$$
$$\beta^p = \max\left\{0.99\bar{\beta}^p, 1\right\}, \quad \beta^d = \max\left\{0.99\bar{\beta}^d, 1\right\},$$

and compute a new iterate $\mathbf{x}^{(j+1)}$ according to

$$(\mathbf{y}^{(j+1)}, \theta^{(j+1)}) = (\mathbf{y}^{(j)}, \theta^{(j)}) + \beta^p(\Delta\mathbf{y}^{(j)}, \Delta\theta^{(j)}),$$
$$(\boldsymbol{\lambda}^{(j+1)}, \mathbf{z}^{(j+1)}) = (\boldsymbol{\lambda}^{(j)}, \mathbf{z}^{(j)}) + \beta^d(\Delta\boldsymbol{\lambda}^{(j)}, \Delta\mathbf{z}^{(j)}).$$

The termination criterion is the same as in the long-step path-following method.

3.3 Multigrid and interior-point methods

We perform the predictor-corrector scheme (cf. Subsection 3.2.1) in a multilevel framework with respect to a hierarchy of discretizations. We describe the multilevel approach in case of a two-level scheme with the levels $\ell-1$ and ℓ (cf. Fig 3.3). In the general case of more than two levels, the multilevel predictor-corrector continuation method consists of a recursive application of the two-level scheme (cf. [17, 77]). The extension to other path-following methods discussed in Subsections 3.2.2 and 3.2.3 follows on the similar lines.

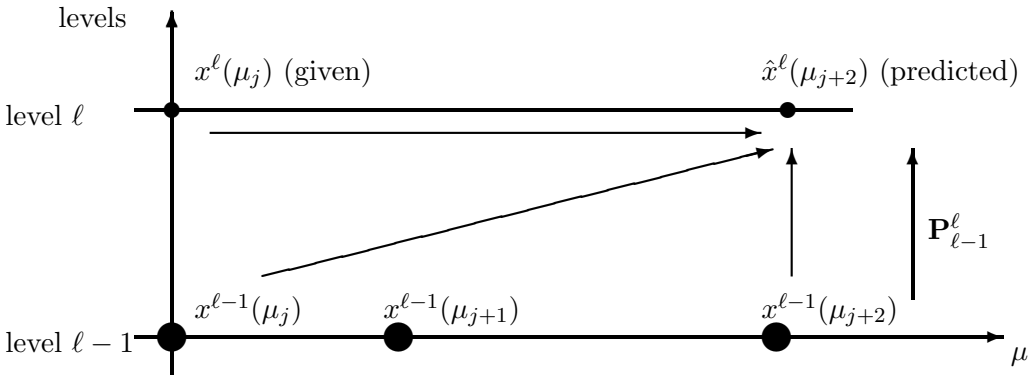


Figure 3.3: Two-level predictor-corrector scheme

3.3.1 Multigrid adaptive continuation method

The solutions on coarse grid level $\ell - 1$ are computed using the adaptively chosen barrier parameters $\mu_j < \mu_{j+1} < \mu_{j+2}$ using the path-following technique as described in Subsection 3.2.1. The fine grid solution $\mathbf{x}^\ell(\mu_j)$ is obtained either by an initialization step or results from a previous application of the multigrid prediction-correction scheme.

Prediction step: We use the the level $\ell - 1$ approximations $x^{\ell-1}(\mu_j)$ and $x^{\ell-1}(\mu_{j+2})$ as well as the level ℓ approximation $x^\ell(\mu_j)$ to obtain a level ℓ prediction at μ_{k+2} using the formula

$$\widehat{\mathbf{x}}^\ell(\mu_{j+2}) = \mathbf{x}^\ell(\mu_j) + \mathbf{P}^\ell_{\ell-1}\left(\mathbf{x}^{\ell-1}(\mu_{j+2}) - \mathbf{x}^{\ell-1}(\mu_j)\right) \tag{3.24}$$

where $\mathbf{P}^\ell_{\ell-1}$ is the prolongation operator from level $\ell - 1$ to ℓ.

Correction step: The approximation $\widehat{\mathbf{x}}^\ell(\mu_{j+2})$ is then corrected by the two-level Newton multigrid scheme incorporating a two-level PDE solver with augmented Lagrangian as the pre- and post-smoother. For acceptance of Newton's method a monotonicity test is performed [33].

3.3.2 Multigrid long-step path-following method

The solutions on coarse grid level $\ell - 1$ are computed using the barrier parameters $\mu_j < \mu_{j+1} < \mu_{j+2}$ using long-step path-following technique as described in Subsection 3.2.2.

Prediction step: The prediction on level ℓ is made using the formula (3.24).

Correction step: The approximation $\widehat{\mathbf{x}}^\ell(\mu_{j+2})$ in step above on level ℓ is corrected using the long-step method with augmented Lagrangian as the pre- and post-smoother.

3.3.3 Multigrid Mehrotra's method

The solutions on coarse grid level $\ell - 1$ are computed using the barrier parameters $\mu_j < \mu_{j+1} < \mu_{j+2}$ using Mehrotra's method as described in Subsection 3.2.3.

Prediction step: The prediction on level ℓ is made using the formula (3.24).

Correction step: The approximation $\widehat{\mathbf{x}}^\ell(\mu_{j+2})$ in step above on level ℓ is corrected using the Mehrotra's method with augmented Lagrangian as the pre- and post-smoother.

3.4 Applications

3.4.1 Channel with a backward facing step

As a benchmark problem, we consider Stokes flow in a channel with a backward facing step (cf. Figure 3.1). The objective is to design the wall of the step in such a way that a desired velocity $\mathbf{v}^d$ is attained.

The computational domain $\Omega \subset \mathbb{R}^2$ is displayed in Figure 3.4. The boundary $\partial\Omega$ is

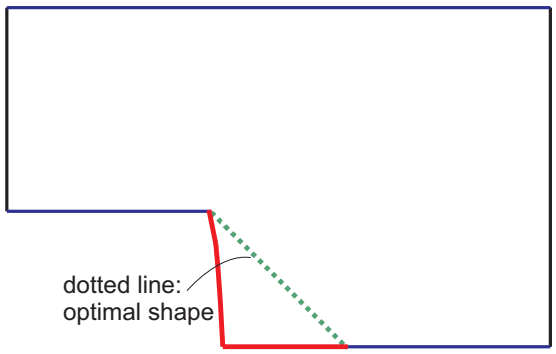


Figure 3.4: Initial and optimal shape of the backward facing step

decomposed into $\Gamma_{\text{in}} = \{0\} \times (-1, 1), \Gamma_{\text{out}} = \{10\} \times (-1.5, 1)$, and $\Gamma_{\text{lat}} = \partial\Omega \setminus (\Gamma_{\text{in}} \cup \Gamma_{\text{out}})$. The data of the problem is chosen as follows. Assume $\mathbf{f} = 0$ in Ω, a Poiseuille velocity profile $\mathbf{v}_{\text{in}}(x_1, x_2) = 6(x_2 + 1)(1 - x_2)$ on Γ_{in}, outflow boundary conditions on Γ_{out}, and no-slip conditions on Γ_{lat}. The objective is to design the shape of the backward step $\Gamma_{2,B}$, with solid line in Figure 3.4 in such a way that a prescribed velocity profile $\mathbf{v}^d$ is achieved in Ω. We use a parametrization of the backward step by means of the Bézier control points $\theta \in \mathbb{R}^k, k = k_B$, of Bézier curve representations of $\Gamma_{2,B}$, where k_B refers to the number of control points for $\Gamma_{2,B}$. The shape optimization problem amounts to the minimization of

$$J(\theta) = \int_{\Omega(\theta)} |\mathbf{v}(x) - \mathbf{v}^d(x)|^2 dx$$

subject to the Stokes equations (3.1c-g) and design parameter constraints

$$\theta^{min} \leq \theta \leq \theta^{max},$$

with viscosity $\nu = 1$. The bounds $\theta^{min}, \theta^{max}$ on the design parameters are chosen such that the design constraints are never active in this example. We use $k_B = 5$ Bézier control points to specify the step boundary. The desired velocity $\mathbf{v}^d$ is computed by specifying the optimal parameter θ^* and solving the state equation on $\Omega(\theta^*)$. The optimal domain $\Omega(\theta^*)$ is shown in Figure 3.4 marked by the dashed line.

We consider a geometrically conforming simplicial triangulation $\mathcal{T}_h(\Omega)$ of the reference domain. The discretization in space is taken care of by P2-P1 Taylor-Hood elements. For $D \subseteq \overline{\Omega}$, we denote by $\mathcal{N}_{\mathbf{v},h}(D), \mathcal{N}_{p,h}(D)$ the set of velocity and pressure nodal points in D.

We use automatic differentiation [54, 100] to compute the derivatives with respect to the design variables θ. The stationary optimization problems are solved using various methods described in Sections 3.2-3.3.

Table 3.1 reflects the convergence history of the iterative process using the adaptive continuation method as described in Subsection 3.2.1 and Figure 3.5 shows the convergence of objective functional J.

k	μ	$\Delta\mu$	ΔJ
0	1.0e+2	3.0e+2	–
1	1.0e+2	3.0e+2	1.70e+0
			6.02e-1
			8.80e-2
			2.31e-1
2	2.7e+2	1.7e+2	1.06e-2
			1.90e-4
			2.05e-5
3	3.6e+2	9.8e+1	1.32e-5
4	3.7e+2	3.9e+2	1.08e-5

Table 3.1: Backward facing step: Adaptive continuation strategy convergence history

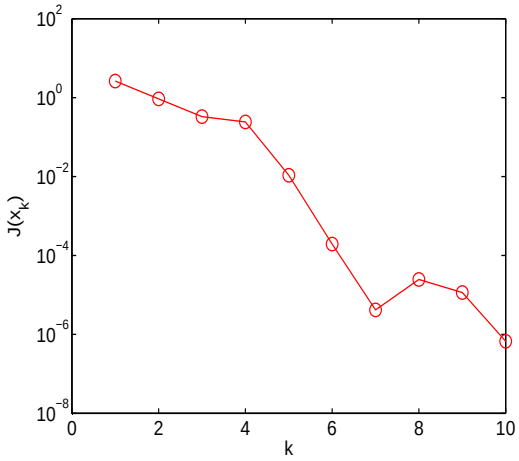


Figure 3.5: Backward facing step: Convergence history of the objective functional

The adaptive continuation method has been compared with the dynamic barrier update strategy from [50], which was also used in [66, 68], and the nonlinear variant of Mehrotra's path-following method. For these three methods, Tables 3.2 and 3.3 contain the number of continuation steps, the final value of the objective functional, and the execution time on a coarse mesh ($h_{max} = 0.3$) and on a finer mesh ($h_{max} = 0.1$), where 'no conv.' means no convergence. The results show that the adaptive continuation method is more efficient

than the dynamic barrier update strategy and more robust than Mehrotra's method for which convergence only occurs, if the mesh is fine enough. On the other hand, for this robustness one has to pay a price in terms of the execution time (cf. Table 3.3).

Method	Cont. Steps	J	Exec. Time
ACM	4	6.68E-07	5 min
DBUS	8	9.32E-07	8 min
MEHR	–	no conv.	–

Table 3.2: Backward facing step: Comparison of adaptive continuation method (ACM), dynamic barrier update strategy (DBUS) from [50] and Mehrotra's method (MEHR) on a coarse mesh ($h_{max} = 0.3$)

Method	Cont. Steps	J	Exec. Time
ACM	3	4.98e-07	59 min
DBUS	11	9.42e-07	142 min
MEHR	8	6.61e-07	34 min

Table 3.3: Backward facing step: Comparison of adaptive continuation method (ACM), dynamic barrier update strategy (DBUS) from [50] and Mehrotra's method (MEHR) on a finer mesh ($h_{max} = 0.1$)

We consider multigrid extension of the above developed interior-point methods i.e., adaptive continuation, long-step and Mehrota's method. Three levels of refinement is used with the number of unknowns on the coarsest and finest level to be 7599 and 116859 respectively. The results obtained using different interior point methods are almost the same. The multilevel code works three times faster as compared to the single grid. We omit the data table in this case and will show it in next example.

The computed optimal value of the design parameter using all the methods is

$$\theta^* = (+1.0, -0.5, -0.55, -0.6, -0.6, -1.5).$$

3.4.2 Capillary barrier

One of the issues in the optimal design of the biochips is to make sure that the reservoir is filled with a very precise amount of the fluid. This is taken care of by a capillary barrier placed between a channel and the reservoir (see Fig. 3.6).

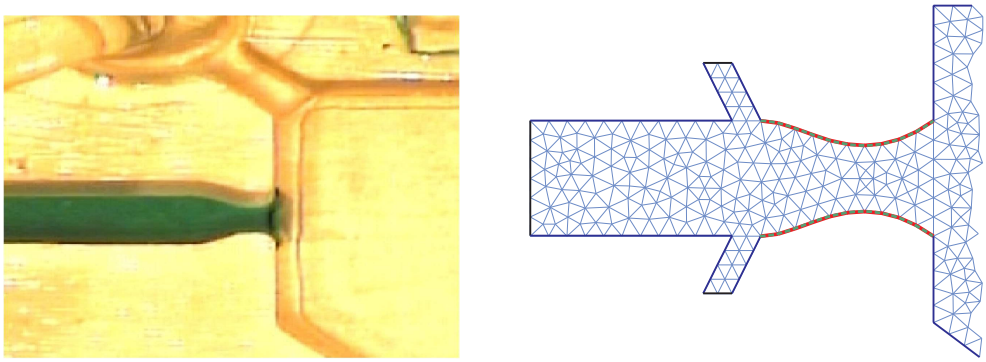


Figure 3.6: Channel with capillary barrier on an SAW driven microfluidic biochip (left) and optimal shape obtained (right)

As computational domain we have chosen part of a channel with a capillary barrier at its end and part of a reservoir connected with the channel by the capillary barrier. The objective is to design the walls of the barriers in such a way that a desired velocity profile $\mathbf{v}^d$ is attained.

The boundary $\partial\Omega$ is decomposed into $\Gamma_{\text{in}} = \{-2\} \times (-1, 1), \Gamma_{\text{out}} = \{5\} \times (-1, 1)$, and $\Gamma_{\text{lat}} = \partial\Omega \setminus (\Gamma_{\text{in}} \cup \Gamma_{\text{out}})$. The data of the problem is chosen as follows. Assume $\mathbf{f} = 0$ in Ω, a Poiseuille velocity profile $\mathbf{v}_{\text{in}}(x_1, x_2) = 6(x_2 + 1)(1 - x_2)$ on Γ_{in}, outflow boundary conditions on Γ_{out}, and no-slip conditions on Γ_{lat}. The objective is to design the shape of the barrier (dotted walls in Figure 3.6(right)) top $\Gamma_{2,T}$ and the bottom $\Gamma_{2,B}$ of $\partial\Omega$ in such a way that a prescribed velocity profile $\mathbf{v}^d$ is achieved in Ω. We use a parametrization of the barrier by means of the Bézier control points $\theta \in \mathbb{R}^k, k = k_T + k_B$, of Bézier curve representations of $\Gamma_{2,T}$ and $\Gamma_{2,B}$, where k_T and k_B refers to the number of control points for

$\Gamma_{2,T}$ and $\Gamma_{2,B}$, respectively. The shape optimization problem amounts to the minimization of

$$J(\theta) = \int_{\Omega(\theta)} |\mathbf{v}(x) - \mathbf{v}^d(x)|^2 dx$$

subject to the Stokes equations (3.1c-g) and design parameter constraints

$$\theta^{min} \leq \theta \leq \theta^{max},$$

with viscosity $\nu = 1$. The bounds $\theta^{min}, \theta^{max}$ on the design parameters are chosen such that the design constraints are never active in this example. We use $k = k_T + k_B = 16$ Bézier control points for the Bézier curve representation of the capillary barrier as design variables. The desired velocity $\mathbf{v}^d$ is computed by specifying the optimal parameter θ^* and solving the state equation on $\Omega(\theta^*)$.

Figure 3.6 (right) displays the computed optimal shape of the barrier together with an underlying finite element mesh. The channel additionally has passive outlet valves (cf. Figure 3.6) that are activated when the barrier operates in stopping mode and back flow occurs. Figure 3.7 (left) provides a visualization of the velocity field for the optimized channel under conditions of flow from the channel into the reservoir. Likewise, Figure 3.7 (right) displays the velocity field for the optimized channel under back flow conditions, i.e., when the capillary barrier operates in stopping mode.

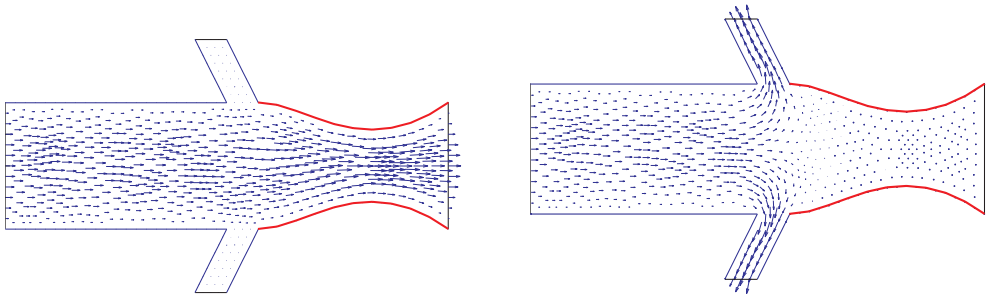


Figure 3.7: Velocity field: Optimal configuration barrier non-stopping mode (left) and stopping mode (right)

We compared the adaptive continuation method with the long-step path-following algorithm and Mehrotra's algorithm, first for a single grid and then consider the extension to the multigrid. As in the previous example, we observed robustness of the adaptive continuation method as it converges on relatively coarse meshes where the two other algorithms fail to convergence. Below we report the convergence history and execution times of all three algorithms for a sufficiently fine finite element mesh with a total of $N_{dof} = 62916$ degrees of freedom with a tolerance $\epsilon = 10^{-3}$ as termination criterion. Tables 3.4, 3.5 and 3.6 display the convergence histories of the three algorithms for a single grid.

k	μ	$\Delta\mu$	ΔJ	Time
0	2.0E+02	5.0E+02	–	
1	6.9E+02	4.9E+02	2.83E+00	
2	1.2E+03	5.3E+02	4.58E-05	747 min

Table 3.4: Capillary barrier: Adaptive continuation method

k	μ	σ	η	ΔJ	Time
0	2.0E+02	–	–	–	
1	2.0E+02	1.0E-03	1	2.35E+00	
2	3.2E+03	1.2E-01	1	4.80E-01	
3	9.2E+03	7.2E-03	1	5.34E-04	360 min

Table 3.5: Capillary barrier: Long-step path-following method

k	μ	σ	β^p, β^d	ΔJ	Time
0	2.0E+02	–	–	–	
1	2.0E+02	1.0E-05	0.99	2.37E+00	
2	3.2E+03	4.1E-02	0.99	4.62E-01	
3	1.2E+03	1.4E-07	0.99	5.04E-04	371 min

Table 3.6: Capillary barrier: Mehrotra's method

Extension to the multigrid is considered of the above developed interior-point methods i.e., adaptive continuation, long-step and Mehrota's method. Four levels of refinement are used with the number of unknowns on corasest and finest level to be 2362 and 141634 respectively. The results obtained using different interior-point methods are almost the

same. The multilevel code works three times faster as compared to the single grid. Table 3.7 shows the results obtained for the multigrid adaptive continuation strategy.

level	k	μ	$\Delta\mu$	ΔJ
1	0	2.0E+02	5.0E+02	
				2.83E+00
	1	6.3E+02	4.3E+02	1.87E-05
	2	1.1E+03	4.9E+02	3.40E-06
	3	1.6E+03	5.1E+02	1.09E-06
	4	2.3E+03	6.8E+02	5.70E-07
	5	3.5E+03	1.1E+03	3.63E-07
	6	5.3E+03	1.9E+03	1.99E-07
	7	8.8E+03	3.5E+03	1.02E-07
	8	1.6E+04	7.3E+03	4.50E-08
2	2	1.1E+03	9.2E+02	
	4	2.3E+03	1.2E+03	
	6	5.3E+03	3.0E+03	
	8	1.6E+04	1.1E+04	
3	4	2.3E+03	2.1E+03	
	8	1.6E+04	1.4E+04	
4	8	1.6E+04	1.6E+04	

Table 3.7: History of the adaptive multilevel predictor-corrector strategy (capillary barrier, 4 levels)

3.5 Concluding remarks

In this chapter, we have provided an 'all-at-once approach' for the optimal design of stationary flow problems described by linear and nonlinear Stokes flow featuring path-following primal-dual interior-point methods by means of an adaptive predictor-corrector type continuation method, a long-step path-following algorithm and a nonlinear version of Mehrotra's algorithm. A multilevel extension of these methods is discussed and implemented. The computation of the first-order derivatives in the KKT systems and the second-order

derivatives in the Hessians is significantly facilitated by automatic differentiation. Numerical examples including a benchmark problem and a real-life design problems demonstrate that the methods can be used in shape optimization with the adaptive continuation method being the most robust algorithm at the expense of slightly higher execution times on finer finite element meshes.

CHAPTER 4

Domain decomposition and model reduction for optimal control and shape optimization of an advection-diffusion system

We consider optimization problems where the optimization variables are located in a small spatial region Ω_2 of the entire spatial domain Ω on which the PDE is posed. Consider, for instance the shape optimization of capillary barriers for the microfluidic biochips, where only a small portion of the shape can be modified, or the parameter identification problems where the parameters are associated with spatially localized material properties.

In this chapter we investigate this for the numerical solution of optimization problems governed by time dependent advection-diffusion partial differential equations (PDEs) and is based on [14]. The approach is extended to the time dependent Stokes or the time dependent linearized Navier-Stokes equations, linearized around a steady state in Chapter 5.

In Chapter 6 we consider the compressible Navier-Stokes equations for the shape optimization of microfluidic biochips. The extension to the Stokes is non-trivial and handling the advection-diffusion case first provides us the insight into the main ideas behind our approach and the error analysis.

Although the optimization parameters are located in a small spatial region Ω_2, standard methods for the numerical optimization of such systems require the repeated solution of the governing PDE (the state equation) and the associated adjoint PDE over the entire domain Ω. It is desirable to reduce the overall problem size by essentially reducing the optimization problem to the small spatial region on which the optimization parameters act. Since the governing PDEs on the small spatial region interact with the solution on the entire domain, it is not feasible to simply truncate the domain, but one has to carefully reduce the problem to preserve the important interactions between the different components of the system. For a class of problems we present a systematic approach based on domain decomposition and balanced truncation model reduction to reduce the subproblems corresponding to the large subdomain $\Omega \setminus \Omega_2$.

There are many examples where domain decomposition and some form of model reduction is used to reduce the computational complexity of the simulation. For example, the papers [42, 43, 45, 93] use physics-based model reduction. A complex system of PDEs is replaced by a simpler model away from the region Ω_2 of interest. Specifically, [42, 43] discusses the coupling of the Navier-Stokes equations to the linear Oseen equations. In [45] the 3D Navier-Stokes equations are coupled with a 1D model for the flow in blood vessels. Section 3.3 of the review paper [93] discusses the coupling of distributed parameter models with lumped parameter models for the modeling of blood flow. The papers [79, 78, 108, 109] use dimension reduction techniques (see [23] for a recent overview). The papers [79, 78]

describe the use of domain decomposition and Proper Orthogonal Decomposition (POD) for the simulation of flows with shocks. Domain decomposition and balanced truncation model reduction is used in [108, 109] for the simulation of PDEs with spatially localized nonlinearities. The approach in these two papers is related to ours, except that we apply it in the optimization context. Moreover, we provide an a priori bound for the error between the solution of the original and the model reduced optimization problem.

We study optimization problems governed by advection-diffusion equations of the type

$$\frac{\partial}{\partial t}\mathbf{y}(x,t) - \nabla(k(x)\nabla\mathbf{y}(x,t)) + V(x)\cdot\nabla\mathbf{y}(x,t)) = \mathbf{f}(x,t)$$

in $\Omega \times (0,T)$, together with suitable boundary and initial conditions. The optimization variables can, for example, be shape parameters that describe the domain Ω or they can be related to the parameters k, V, f in the PDE. In Section 4.4 we discuss an optimal control problem in which the optimization variable is related to the source f and a shape optimization problem in which the optimization variables are shape parameters.

After a discretization in space the optimization problems studied in this chapter are of the form

$$\text{minimize} \int_0^T \ell(\mathbf{y}(t),t,\theta)dt, \tag{4.1a}$$

subject to

$$\mathbf{M}(\theta)\frac{d}{dt}\mathbf{y}(t) + \mathbf{A}(\theta)\mathbf{y}(t) = \mathbf{B}(\theta)\mathbf{u}(t), \qquad t \in (0,T), \tag{4.1b}$$

$$\mathbf{M}(\theta)\mathbf{y}(0) = \mathbf{y}_0, \tag{4.1c}$$

$$\theta \in \Theta. \tag{4.1d}$$

Here $\mathbf{M}(\theta), \mathbf{A}(\theta) \in \mathbb{R}^{N\times N}$ are mass and stiffness matrices that arise from a spatial discretization. Furthermore Θ is a closed convex set of admissible parameters and $\mathbf{B}(\theta) \in$

$\mathbb{R}^{N\times m}$, $\mathbf{u}$ are given inputs which relate to the source f and boundary data in the advection-diffusion equation. We will discuss the derivation of (4.1) for two applications in Section 4.4. Since the optimization variables θ are related to spatially localized quantities (shape parameters, coefficients,..) in the advection-diffusion equation, only few entries of $\mathbf{M}(\theta), \mathbf{A}(\theta), \mathbf{B}(\theta)$ depend on θ.

Our goal is to replace (4.1) by a reduced order problem

$$\text{minimize } \int_0^T \ell(\widehat{\mathbf{y}}(t), t, \theta)dt, \tag{4.2a}$$

subject to

$$\widehat{\mathbf{M}}(\theta)\frac{d}{dt}\widehat{\mathbf{y}}(t) + \widehat{\mathbf{A}}(\theta)\widehat{\mathbf{y}}(t) = \widehat{\mathbf{B}}(\theta)\mathbf{u}(t), \qquad t \in (0,T), \tag{4.2b}$$

$$\widehat{\mathbf{M}}(\theta)\mathbf{y}(0) = \widehat{\mathbf{y}}_0, \tag{4.2c}$$

$$\theta \in \Theta, \tag{4.2d}$$

with matrices $\widehat{\mathbf{M}}(\theta), \widehat{\mathbf{A}}(\theta) \in \mathbb{R}^{n\times n}$, $\widehat{\mathbf{B}}(\theta) \in \mathbb{R}^{n\times m}$, such that $n \ll N$ and such that the solution θ_* of (4.1) is well approximated by the solution $\widehat{\theta}_*$ of (4.2).

Our approach uses domain decomposition techniques to divide the optimality system corresponding to (4.1) into linear subproblems and small nonlinear subproblems. Balanced truncation is applied to the linear subproblems with inputs and outputs determined by the original in- and outputs as well as the interface conditions between the subproblems. The reduced optimality system is identified as the optimality system of a reduced optimization problem (4.2). We provide a priori estimates for the error between the solution θ_* of (4.1) and the solution $\widehat{\theta}_*$ of (4.2). These bounds depend on the balanced truncation error bounds as well as properties of the subsystem that is not reduced.

We expect that this combination of domain decomposition and balanced truncation will lead to a substantial reduction of the original problem, if the nonlinearities are localized, i.e.,

the nonlinear subproblems are small relative to the other subdomains, and if the interfaces between the subproblems are relatively small. This is confirmed by our numerical results.

In the next section we provide a brief review of balanced truncation model reduction. Section 4.2 applies balanced truncation to reduce a linear quadratic optimal control problem. Although this optimization problem is simpler than (4.1) it is relevant for many applications and already provides insight into the main ideas behind our approach and the corresponding error analysis. The integration of domain decomposition and balanced truncation model reduction for the reduction of (4.1) is presented and analyzed in Section 4.3. In Section 4.4 we discuss two problems which lead to (4.1) and the application of our approach for the reduction of these problems.

Throughout this chapter we use $\|\cdot\|$ to denote the Euclidean norm in $\mathbb{R}^N$ or the corresponding matrix norm in $\mathbb{R}^{N\times N}$. Instead of $L^p(0,T;\mathbb{R}^N)$ we simply write L^p.

4.1 Balanced truncation model reduction

Model reduction seeks to replace a large-scale system of differential or difference equations by a system of substantially lower dimension that has nearly the same response characteristics. Balanced reduction is a particular method that preserves asymptotic stability and also provides an error bound on the discrepancy between the outputs of the full and reduced order system [18, 23, 34, 53, 86]. We use balanced truncation model reduction because of the availability of an error bound.

We briefly review balanced truncation model reduction for linear time invariant systems

in state space form

$$\mathcal{M}\mathbf{y}'(t) = \mathcal{A}\mathbf{y}(t) + \mathcal{B}\mathbf{u}(t), \qquad t \in (0,T) \tag{4.3a}$$

$$\mathbf{z}(t) = \mathcal{C}\mathbf{y}(t) + \mathcal{D}_s\mathbf{u}(t), \qquad t \in (0,T) \tag{4.3b}$$

$$\mathbf{y}(0) = \mathbf{y}_0, \tag{4.3c}$$

$$-\mathcal{M}\boldsymbol{\lambda}'(t) = \mathcal{A}^T\boldsymbol{\lambda}(t) + \mathcal{C}^T\mathbf{w}(t), \qquad t \in (0,T) \tag{4.3d}$$

$$\mathbf{q}(t) = \mathcal{B}^T\boldsymbol{\lambda}(t) + \mathcal{D}_a\mathbf{w}(t), \qquad t \in (0,T) \tag{4.3e}$$

$$\boldsymbol{\lambda}(T) = 0, \tag{4.3f}$$

where $\mathcal{M} \in \mathbb{R}^{N\times N}$ is symmetric positive definite, $\mathcal{A} \in \mathbb{R}^{N\times N}$, $\mathcal{B} \in \mathbb{R}^{N\times m}$, $\mathcal{C} \in \mathbb{R}^{k\times N}$, $\mathcal{D}_s \in \mathbb{R}^{k\times m}$, and $\mathcal{D}_a \in \mathbb{R}^{m\times k}$.

Projection methods for model reduction generally produce $N \times n$ matrices $\mathcal{V}, \mathcal{W}$ with $n \ll N$ and with $\mathcal{W}^T\mathcal{M}\mathcal{V} = I_n$. One obtains a reduced form of equations (4.3) by setting $\mathbf{y} = \mathcal{V}\widehat{\mathbf{y}}$ and projecting (imposing a Galerkin condition) so that

$$\mathcal{W}^T[\mathcal{M}\mathcal{V}\frac{d}{dt}\widehat{\mathbf{y}}(t) - \mathcal{A}\mathcal{V}\widehat{\mathbf{y}}(t) - \mathcal{B}\mathbf{u}(t)] = 0, \quad t \in (0,T).$$

Applying an analogous projection to (4.3d,e) with $\boldsymbol{\lambda}$ replaced by $\mathcal{W}\widehat{\boldsymbol{\lambda}}$, we obtain a reduced order system of order n given by

$$\widehat{\mathbf{y}}'(t) = \widehat{\mathcal{A}}\widehat{\mathbf{y}}(t) + \widehat{\mathcal{B}}\mathbf{u}(t), \qquad t \in (0,T) \tag{4.4a}$$

$$\widehat{\mathbf{z}}(t) = \widehat{\mathcal{C}}\widehat{\mathbf{y}}(t) + \mathcal{D}_s\mathbf{u}(t), \qquad t \in (0,T) \tag{4.4b}$$

$$\widehat{\mathbf{y}}(0) = \widehat{\mathbf{y}}_0, \tag{4.4c}$$

$$-\widehat{\boldsymbol{\lambda}}'(t) = \widehat{\mathcal{A}}^T\widehat{\boldsymbol{\lambda}}(t) + \widehat{\mathcal{C}}^T\mathbf{w}(t), \qquad t \in (0,T) \tag{4.4d}$$

$$\widehat{\mathbf{q}}(t) = \widehat{\mathcal{B}}^T\widehat{\boldsymbol{\lambda}}(t) + \mathcal{D}_a\mathbf{w}(t), \qquad t \in (0,T) \tag{4.4e}$$

$$\widehat{\boldsymbol{\lambda}}(T) = 0, \tag{4.4f}$$

with $\widehat{\mathcal{A}} = \mathcal{W}^T\mathcal{A}\mathcal{V}$, $\widehat{\mathcal{B}} = \mathcal{W}^T\mathcal{B}$, $\widehat{\mathcal{C}} = \mathcal{C}\mathcal{V}$, and $\widehat{\mathbf{y}}_0 = \mathcal{W}^T\mathcal{M}\mathbf{y}_0$.

Balanced reduction is a particular technique for constructing the projecting matrices $\mathcal{V}$ and $\mathcal{W}$. Originally, balanced reduction was developed for state space systems with $\mathcal{M} = I$. To apply it to (4.3), we factor $\mathcal{M} = \mathcal{R}\mathcal{R}^T$, multiply (4.3) by $\mathcal{R}^{-1}$, and substitute $\widetilde{\mathbf{y}} = \mathcal{R}^T\mathbf{y}$, $\widetilde{\boldsymbol{\lambda}} = \mathcal{R}^T\boldsymbol{\lambda}$. Then we apply the standard balanced reduction to the resulting system. Afterwards we transform back to the original variables and express all operations in terms of the original system (4.3).

To compute the balanced reduction, we first have to compute the controllability and observability Gramians $\mathcal{P}, \mathcal{Q}$, respectively. Under the assumptions of stability, controllability and observability, the matrices $\mathcal{P}, \mathcal{Q}$ are both symmetric and positive definite and they solve the Lyapunov equations

$$\mathcal{A}\mathcal{P}\mathcal{M} + \mathcal{M}\mathcal{P}\mathcal{A}^T + \mathcal{B}\mathcal{B}^T = 0, \tag{4.5a}$$

$$\mathcal{A}^T\mathcal{Q}\mathcal{M} + \mathcal{M}\mathcal{Q}\mathcal{A} + \mathcal{C}^T\mathcal{C} = 0. \tag{4.5b}$$

There are direct methods for the small dense case and iterative methods for the large sparse setting to compute $\mathcal{P} = \mathbf{U}\mathbf{U}^T$ and $\mathcal{Q} = \mathbf{L}\mathbf{L}^T$ in factored form. In the large-scale setting the factorization is typically a low rank approximation.

The balancing transformation is constructed by the SVD

$$\mathbf{U}^T\mathcal{M}\mathbf{L} = \mathbf{Z}\mathbf{S}\mathbf{Y}^T, \tag{4.6a}$$

$$\mathcal{V} = \mathbf{U}\mathbf{Z}_n\mathbf{S}_n^{-1/2}, \tag{4.6b}$$

$$\mathcal{W} = \mathbf{L}\mathbf{Y}_n\mathbf{S}_n^{-1/2}. \tag{4.6c}$$

Here, $\mathbf{S}_n = \operatorname{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n)$ with $\mathbf{S} = \mathbf{S}_N$. The σ_j are in decreasing order and n is selected to be the smallest positive integer such that $\sigma_{n+1} < \tau\sigma_1$ where $\tau > 0$ is a

prespecified constant. The matrices $\mathbf{Z}_n, \mathbf{Y}_n$ consist of the corresponding leading n columns of $\mathbf{Z}, \mathbf{Y}$.

It is easily verified that $\mathcal{P}\mathcal{M}\mathcal{W} = \mathcal{V}\mathbf{S}_n$ and that $\mathcal{Q}\mathcal{M}\mathcal{V} = \mathcal{W}\mathbf{S}_n$. Hence,

$$\begin{aligned} 0 &= \mathcal{W}^T(\mathcal{A}\mathcal{P}\mathcal{M} + \mathcal{M}\mathcal{P}\mathcal{A}^T + \mathcal{B}\mathcal{B}^T)\mathcal{W} \\ &= \widehat{\mathcal{A}}\mathbf{S}_n + \mathbf{S}_n\widehat{\mathcal{A}}^T + \widehat{\mathcal{B}}\widehat{\mathcal{B}}^T, \end{aligned} \tag{4.7a}$$

$$\begin{aligned} 0 &= \mathcal{V}^T(\mathcal{A}^T\mathcal{Q}\mathcal{M} + \mathcal{M}\mathcal{Q}\mathcal{A} + \mathcal{C}^T\mathcal{C})\mathcal{V} \\ &= \widehat{\mathcal{A}}^T\mathbf{S}_n + \mathbf{S}_n\widehat{\mathcal{A}} + \widehat{\mathcal{C}}^T\widehat{\mathcal{C}}. \end{aligned} \tag{4.7b}$$

The terminology "balanced" refers to the fact that the controllability and observability Gramians of the reduced systems are both diagonal and equal. This is true for every possible order n of the truncation.

It is well known (see, e.g., [18, 53, 124]) that $\widehat{\mathcal{A}}$ must be stable. Furthermore if $\mathbf{y}_0 = \mathbf{0}$, and $\sigma_n > \sigma_{n+1}$, then for any given inputs $\mathbf{u}$, $\mathbf{w}$ we have

$$\|\mathbf{z} - \widehat{\mathbf{z}}\|_{L^2} \leq 2\|\mathbf{u}\|_{L^2}(\sigma_{n+1} + \ldots + \sigma_N), \tag{4.8a}$$

$$\|\mathbf{q} - \widehat{\mathbf{q}}\|_{L^2} \leq 2\|\mathbf{w}\|_{L^2}(\sigma_{n+1} + \ldots + \sigma_N). \tag{4.8b}$$

Remark 4.1.1 *One can derive error bounds for inhomogeneous initial values $\mathbf{y}_0$. These require a slight modification of the problem setup in which the original $\mathcal{B}$ is augmented. Since we are interested in the handling of local nonlinearities and our examples have homogeneous initial values $\mathbf{y}_0 = \mathbf{0}$, we omit this extension.*

4.2 Balanced truncation model reduction and optimal control

Before we consider the optimization problem (4.1), we consider a simpler problem, a linear quadratic optimal control problem

$$\min J(\mathbf{u}) \equiv \frac{1}{2}\int_0^T \|\mathbf{C}\mathbf{y}(t) + \mathbf{D}\mathbf{u}(t) - \mathbf{d}(t)\|^2 dt, \tag{4.9}$$

where $\mathbf{y}(t) = \mathbf{y}(\mathbf{u}; t)$ is the solution of

$$\mathbf{M}\mathbf{y}'(t) = \mathbf{A}\mathbf{y}(t) + \mathbf{B}\mathbf{u}(t), \qquad t \in (0, T), \tag{4.10a}$$

$$\mathbf{y}(0) = \mathbf{y}_0. \tag{4.10b}$$

Here $\mathbf{M} \in \mathbb{R}^{N\times N}$ is symmetric positive definite, $\mathbf{A} \in \mathbb{R}^{N\times N}$, $\mathbf{B} \in \mathbb{R}^{N\times m}$, $\mathbf{C} \in \mathbb{R}^{k\times N}$, $\mathbf{D} \in \mathbb{R}^{k\times m}$, and $\mathbf{d} \in L^2(0, T)$ is a given function. We assume that there exists $\alpha > 0$ such that

$$\mathbf{v}^T\mathbf{A}\mathbf{v} \le -\alpha\mathbf{v}^T\mathbf{M}\mathbf{v}, \qquad \forall \mathbf{v} \in \mathbb{R}^N. \tag{4.11}$$

Note that (4.11) implies that all eigenvalues of the pair $(\mathbf{A}, \mathbf{M})$ have negative real part.

Remark 4.2.1 *We can show that under suitable assumptions, the finite element discretization of advection dominated problems satisfies* (4.11)*. We refer to Chapter 6 in [94] and Chapter 3 in [73] for complete details.*

We want to reduce this optimization problem using balanced truncation model reduction and establish bounds for the error between the solution $\mathbf{u}_*$ of (4.9), (4.10) and the solution $\widehat{\mathbf{u}}_*$ of the reduced optimal control problem. This will provide some insight into the process that will be applied for the reduction of the optimization problem (4.1) in a simpler setting involving less notation.

The necessary and sufficient optimality conditions for (4.9), (4.10) are given by

$$\mathbf{M}\mathbf{y}'(t) = \mathbf{A}\mathbf{y}(t) + \mathbf{B}\mathbf{u}(t), \qquad t \in (0,T) \tag{4.12a}$$

$$\mathbf{z}(t) = \mathbf{C}\mathbf{y}(t) + \mathbf{D}\mathbf{u}(t) - \mathbf{d}(t), \qquad t \in (0,T) \tag{4.12b}$$

$$\mathbf{y}(0) = \mathbf{y}_0, \tag{4.12c}$$

$$-\mathbf{M}\boldsymbol{\lambda}'(t) = \mathbf{A}^T\boldsymbol{\lambda}(t) + \mathbf{C}^T\mathbf{z}(t), \qquad t \in (0,T) \tag{4.12d}$$

$$\mathbf{q}(t) = \mathbf{B}^T\boldsymbol{\lambda}(t) + \mathbf{D}^T\mathbf{z}(t), \qquad t \in (0,T) \tag{4.12e}$$

$$\boldsymbol{\lambda}(T) = 0, \tag{4.12f}$$

$$\mathbf{q}(t) = \mathbf{0}, \qquad t \in (0,T). \tag{4.12g}$$

The optimality system (4.12) is written in a slightly unconventional way to highlight its connection with the system (4.3) to which balanced truncation model reduction can be applied.

We use balanced truncation model reduction to compute $\mathcal{W}, \mathcal{V}$ and the reduced optimality system

$$\widehat{\mathbf{y}}'(t) = \widehat{\mathbf{A}}\widehat{\mathbf{y}}(t) + \widehat{\mathbf{B}}\mathbf{u}(t), \qquad t \in (0,T) \tag{4.13a}$$

$$\widehat{\mathbf{z}}(t) = \widehat{\mathbf{C}}\widehat{\mathbf{y}}(t) + \mathbf{D}\mathbf{u}(t) - \mathbf{d}(t), \qquad t \in (0,T) \tag{4.13b}$$

$$\widehat{\mathbf{y}}(0) = \widehat{\mathbf{y}}_0, \tag{4.13c}$$

$$-\widehat{\boldsymbol{\lambda}}'(t) = \widehat{\mathbf{A}}^T\widehat{\boldsymbol{\lambda}}(t) + \widehat{\mathbf{C}}^T\widehat{\mathbf{z}}(t), \qquad t \in (0,T) \tag{4.13d}$$

$$\widehat{\mathbf{q}}(t) = \widehat{\mathbf{B}}^T\widehat{\boldsymbol{\lambda}}(t) + \mathbf{D}^T\widehat{\mathbf{z}}(t), \qquad t \in (0,T) \tag{4.13e}$$

$$\widehat{\boldsymbol{\lambda}}(T) = 0, \tag{4.13f}$$

$$\widehat{\mathbf{q}}(t) = \mathbf{0}, \qquad t \in (0,T), \tag{4.13g}$$

with $\widehat{\mathbf{A}} = \mathcal{W}^T\mathbf{A}\mathcal{V}$, $\widehat{\mathbf{B}} = \mathcal{W}^T\mathbf{B}$, $\widehat{\mathbf{C}} = \mathbf{C}\mathcal{V}$, and $\widehat{\mathbf{y}}_0 = \mathcal{W}^T\mathbf{M}\mathbf{y}_0$. We assume that

$$\mathbf{y}_0 = \mathbf{0}, \tag{4.14}$$

cf., Remark 4.1.1.

We note that the reduced optimality system (4.13) is the optimality system for the reduced optimal control problem

$$\min \widehat{J}(\mathbf{u}) \equiv \frac{1}{2}\int_0^T \|\widehat{\mathbf{C}}\widehat{\mathbf{y}}(t) + \mathbf{D}\mathbf{u}(t) - \mathbf{d}(t)\|^2 dt \tag{4.15}$$

where $\widehat{\mathbf{y}}(t) = \widehat{\mathbf{y}}(\mathbf{u}; t)$ solves

$$\widehat{\mathbf{y}}'(t) = \widehat{\mathbf{A}}\widehat{\mathbf{y}}(t) + \widehat{\mathbf{B}}\mathbf{u}(t), \qquad t \in (0, T), \tag{4.16a}$$

$$\widehat{\mathbf{y}}(0) = \widehat{\mathbf{y}}_0. \tag{4.16b}$$

Next we provide an estimate for the error between the solution $\mathbf{u}_*$ of (4.9), (4.10) and the solution $\widehat{\mathbf{u}}_*$ of (4.15), (4.16). We assume that J is a strictly convex quadratic function. More precisely, we assume the existence of $\kappa > 0$ such that

$$\langle \mathbf{u} - \mathbf{w}, \nabla J(\mathbf{u}) - \nabla J(\mathbf{w})\rangle_{L^2} \geq \kappa \|\mathbf{u} - \mathbf{w}\|_{L^2}^2 \tag{4.17}$$

for all $\mathbf{u}, \mathbf{w} \in L^2$. If $\mathbf{u}_*$ solves (4.9), (4.10) and $\widehat{\mathbf{u}}_*$ solves (4.15), (4.16), then

$$\nabla J(\mathbf{u}_*) = \nabla \widehat{J}(\widehat{\mathbf{u}}_*) = 0$$

and (4.17) implies

$$\begin{aligned}
\|\mathbf{u}_* - \widehat{\mathbf{u}}_*\|_{L^2} \|\nabla \widehat{J}(\widehat{\mathbf{u}}_*) - \nabla J(\widehat{\mathbf{u}}_*)\|_{L^2} &= \|\mathbf{u}_* - \widehat{\mathbf{u}}_*\|_{L^2} \|\nabla J(\mathbf{u}_*) - \nabla J(\widehat{\mathbf{u}}_*)\|_{L^2} \\
&\geq \langle \mathbf{u}_* - \widehat{\mathbf{u}}_*, \nabla J(\mathbf{u}_*) - \nabla J(\widehat{\mathbf{u}}_*)\rangle_{L^2} \\
&\geq \kappa \|\mathbf{u}_* - \widehat{\mathbf{u}}_*\|_{L^2}^2.
\end{aligned}$$

Hence

$$\|\mathbf{u}_* - \widehat{\mathbf{u}}_*\|_{L^2} \leq \kappa^{-1} \|\nabla \widehat{J}(\widehat{\mathbf{u}}_*) - \nabla J(\widehat{\mathbf{u}}_*)\|_{L^2}. \tag{4.18}$$

Thus, to estimate the error we need to estimate the error in the gradients between the original problem (4.9), (4.10) and of the reduced (4.15), (4.16).

To emphasize the dependence of the solution of (4.12a-f) and of (4.13a-f) on the inputs $\mathbf{u}$, we often write $\mathbf{y}(\mathbf{u})$, $\mathbf{z}(\mathbf{u})$, $\boldsymbol{\lambda}(\mathbf{u})$, $\mathbf{q}(\mathbf{u})$ and $\widehat{\mathbf{y}}(\mathbf{u})$, $\widehat{\mathbf{z}}(\mathbf{u})$, $\widehat{\boldsymbol{\lambda}}(\mathbf{u})$, $\widehat{\mathbf{q}}(\mathbf{u})$. If for given $\mathbf{u}$ the functions $\mathbf{y}(\mathbf{u})$, $\mathbf{z}(\mathbf{u})$, $\boldsymbol{\lambda}(\mathbf{u})$, $\mathbf{q}(\mathbf{u})$ satisfy (4.12a-f) and $\widehat{\mathbf{y}}(\mathbf{u})$, $\widehat{\mathbf{z}}(\mathbf{u})$, $\widehat{\boldsymbol{\lambda}}(\mathbf{u})$, $\widehat{\mathbf{q}}(\mathbf{u})$ satisfy (4.13a-f), then

$$\nabla J(\mathbf{u}) = \mathbf{q}(\mathbf{u}), \quad \nabla \widehat{J}(\mathbf{u}) = \widehat{\mathbf{q}}(\mathbf{u}).$$

To estimate the error $\|\mathbf{q}(\widehat{\mathbf{u}}_*) - \widehat{\mathbf{q}}(\widehat{\mathbf{u}}_*)\|_{L^2}$ we cannot use the error estimate (4.8) for balanced truncation model reduction directly, since (4.3d,e) and (4.4d,e) both depend on the same input $\mathbf{w}$, whereas (4.12d,e) has input $\mathbf{z}$ and (4.13d,e) has input $\widehat{\mathbf{z}}$.

We consider the auxiliary adjoint equation

$$-\mathbf{M}\widetilde{\boldsymbol{\lambda}}'(t) = \mathbf{A}^T\widetilde{\boldsymbol{\lambda}}(t) + \mathbf{C}^T\widehat{\mathbf{z}}(t), \qquad t \in (0,T) \tag{4.19a}$$

$$\widetilde{\mathbf{q}}(t) = \mathbf{B}^T\widetilde{\boldsymbol{\lambda}}(t) + \mathbf{D}^T\widehat{\mathbf{z}}(t), \qquad t \in (0,T) \tag{4.19b}$$

$$\widetilde{\boldsymbol{\lambda}}(T) = 0. \tag{4.19c}$$

Lemma 4.2.2 *Let* (4.11) *be satisfied. For any* $\mathbf{z}, \widehat{\mathbf{z}} \in L^2$ *the outputs* $\mathbf{q}$ *and* $\widetilde{\mathbf{q}}$ *of* (4.13d-f) *and* (4.19)*, respectively, satisfy*

$$\|\widetilde{\mathbf{q}} - \mathbf{q}\|_{L^2} \le c\|\widehat{\mathbf{z}} - \mathbf{z}\|_{L^2},$$

where $c = \alpha^{-1}2\|\mathbf{C}\mathbf{M}^{-1/2}\|\|\mathbf{M}^{-1/2}\mathbf{B}\| + \|\mathbf{D}\|$.

Proof: Since $\mathbf{M}$ is symmetric positive definite, $\mathbf{M}^{1/2}$ exists and is symmetric positive definite. The scaled adjoints $\mathbf{M}^{1/2}(\widetilde{\boldsymbol{\lambda}} - \boldsymbol{\lambda})$ satisfy

$$\begin{aligned} -\mathbf{M}^{1/2}(\widetilde{\boldsymbol{\lambda}} - \boldsymbol{\lambda})'(t) =& \mathbf{M}^{-1/2}\mathbf{A}^T\mathbf{M}^{-1/2}\mathbf{M}^{1/2}(\widetilde{\boldsymbol{\lambda}} - \boldsymbol{\lambda})(t) + \mathbf{M}^{-1/2}\mathbf{C}^T(\widehat{\mathbf{z}} - \mathbf{z})(t), \\ \mathbf{M}^{1/2}(\widetilde{\boldsymbol{\lambda}} - \boldsymbol{\lambda})(T) =& 0. \end{aligned}$$

Lemma A.1.1 in the Appendix A gives

$$\|\mathbf{M}^{1/2}(\widetilde{\boldsymbol{\lambda}} - \boldsymbol{\lambda})\|_{L^2} \leq \frac{2\|\mathbf{C}\mathbf{M}^{-1/2}\|}{\alpha}\|\widehat{\mathbf{z}} - \mathbf{z}\|_{L^2}.$$

The desired inequality follows since

$$\widetilde{\mathbf{q}} - \mathbf{q} = \mathbf{B}^T\mathbf{M}^{-1/2}\mathbf{M}^{1/2}(\widetilde{\boldsymbol{\lambda}} - \boldsymbol{\lambda}) + \mathbf{D}^T(\widehat{\mathbf{z}} - \mathbf{z}).$$

□

We assume that the Hankel singular values satisfy

$$\sigma_1 \geq \ldots \geq \sigma_n > \sigma_{n+1} \geq \ldots \geq \sigma_N \ . \tag{4.20}$$

The error estimate (4.8) for balanced truncation model reduction implies

$$\|\mathbf{z} - \widehat{\mathbf{z}}\|_{L^2} \leq 2\|\mathbf{u}\|_{L^2}(\sigma_{n+1} + \ldots + \sigma_N), \tag{4.21a}$$

$$\|\widehat{\mathbf{q}} - \widetilde{\mathbf{q}}\|_{L^2} \leq 2\|\widehat{\mathbf{z}}\|_{L^2}(\sigma_{n+1} + \ldots + \sigma_N) \tag{4.21b}$$

for all $\mathbf{u} \in L^2$ and all $\widehat{\mathbf{z}} \in L^2$. We can now use Lemma 4.2.2 and the balanced truncation model reduction error estimates (4.21) to derive a bound for the error between the solutions $\mathbf{u}_*$ of (4.9), (4.10) and $\widehat{\mathbf{u}}_*$ of (4.15), (4.16).

Theorem 4.2.3 *Let* (4.11) *and* (4.20) *be satisfied. For any* $\mathbf{u} \in L^2$ *let* $\widehat{\mathbf{y}}(\mathbf{u})$ *be the corresponding reduced state and* $\widehat{\mathbf{z}}(\mathbf{u}) = \widehat{\mathbf{C}}\widehat{\mathbf{y}}(\mathbf{u}) + \mathbf{D}\mathbf{u} - \mathbf{d}$. *The error in the gradients obeys*

$$\|\nabla J(\mathbf{u}) - \nabla \widehat{J}(\mathbf{u})\|_{L^2} \leq 2\,(c\|\mathbf{u}\|_{L^2} + \|\widehat{\mathbf{z}}(\mathbf{u})\|_{L^2})\,(\sigma_{n+1} + \ldots + \sigma_N),$$

where c *is the constant specified in Lemma 4.2.2.*

Proof: For arbitrary $\mathbf{u} \in L^2$ let the functions $\mathbf{y}(\mathbf{u})$, $\mathbf{z}(\mathbf{u})$, $\boldsymbol{\lambda}(\mathbf{u})$, $\mathbf{q}(\mathbf{u})$ satisfy (4.12a-f), let $\widehat{\mathbf{y}}(\mathbf{u})$, $\widehat{\mathbf{z}}(\mathbf{u})$, $\widehat{\boldsymbol{\lambda}}(\mathbf{u})$, $\widehat{\mathbf{q}}(\mathbf{u})$ satisfy (4.13a-f), and let $\widetilde{\boldsymbol{\lambda}}(\mathbf{u})$, $\widetilde{\mathbf{q}}(\mathbf{u})$ satisfy (4.19).

We have $\nabla J(\mathbf{u}) = \mathbf{q}(\mathbf{u})$, $\nabla \widehat{J}(\mathbf{u}) = \widehat{\mathbf{q}}(\mathbf{u})$. Lemma 4.2.2 and the balanced truncation model reduction error estimates (4.21) imply

$$\begin{aligned}\|\mathbf{q}(\mathbf{u}) - \widehat{\mathbf{q}}(\mathbf{u})\|_{L^2} &\leq \|\mathbf{q}(\mathbf{u}) - \widetilde{\mathbf{q}}(\mathbf{u})\|_{L^2} + \|\widetilde{\mathbf{q}}(\mathbf{u}) - \widehat{\mathbf{q}}(\mathbf{u})\|_{L^2} \\ &\leq c\|\widehat{\mathbf{z}}(\mathbf{u}) - \mathbf{z}(\mathbf{u})\|_{L^2} + 2\|\widehat{\mathbf{z}}(\mathbf{u})\|_{L^2}(\sigma_{n+1} + \ldots + \sigma_N) \\ &\leq 2\left(c\|\mathbf{u}\|_{L^2} + \|\widehat{\mathbf{z}}(\mathbf{u})\|_{L^2}\right)(\sigma_{n+1} + \ldots + \sigma_N).\end{aligned}$$

□

Inequality (4.18) and Theorem 4.2.3 imply the following estimate for the error in the optimal controls.

Corollary 4.2.4 *Let* (4.11) *and* (4.20) *be satisfied and let* $\kappa > 0$ *be a constant such that* (4.17) *holds. Furthermore, let* $\mathbf{u}_*$ *solve* (4.9), (4.10) *and let* $\widehat{\mathbf{u}}_*$ *be the solution of* (4.15), (4.16) *with corresponding state* $\widehat{\mathbf{y}}_*$ *and* $\widehat{\mathbf{z}}_* = \widehat{\mathbf{C}}\widehat{\mathbf{y}}_* + \mathbf{D}\mathbf{u}_* - \mathbf{d}$. *The error between the solutions satisfies*

$$\|\mathbf{u}_* - \widehat{\mathbf{u}}_*\|_{L^2} \leq \frac{2}{\kappa}\left(c\|\widehat{\mathbf{u}}_*\|_{L^2} + \|\widehat{\mathbf{z}}_*\|_{L^2}\right)(\sigma_{n+1} + \ldots + \sigma_N),$$

where c *is the constant specified in Lemma 4.2.2.*

Note that the size of $\sigma_{n+1} + \ldots + \sigma_N$ can be controlled by the user during the computation of the reduced order models. Moreover, $\|\widehat{\mathbf{u}}_*\|_{L^2}$ and $\|\widehat{\mathbf{z}}_*\|_{L^2}$ can be computed.

4.3 The optimization problem

We now return to the optimization problem (4.1). The Lagrangian associated with this problem is

$$L(\mathbf{y},\boldsymbol{\lambda},\theta)=\int_0^T \ell(\mathbf{y}(t),t,\theta)dt+\int_0^T \boldsymbol{\lambda}(t)^T\Big(\mathbf{M}(\theta)\mathbf{y}'(t)+\mathbf{A}(\theta)\mathbf{y}(t)-\mathbf{B}(\theta)\mathbf{u}(t)\Big)\,dt.$$

Since Θ is a closed convex set, the first-order necessary optimality conditions for (4.1) with $\theta\in\Theta$, are given by the state-adjoint system

$$\mathbf{M}(\theta)\frac{d}{dt}\mathbf{y}(t)+\mathbf{A}(\theta)\mathbf{y}(t)=\mathbf{B}(\theta)\mathbf{u}(t), \qquad t\in(0,T), \tag{4.22a}$$

$$-\mathbf{M}(\theta)^T\frac{d}{dt}\boldsymbol{\lambda}'(t)+\mathbf{A}(\theta)^T\boldsymbol{\lambda}(t)=-\nabla_{\mathbf{y}}\ell(\mathbf{y}(t),t,\theta), \qquad t\in(0,T), \tag{4.22b}$$

and the inequality

$$\begin{aligned}&\int_0^T D_\theta\ell(\mathbf{y}(t),t,\theta)(\widetilde{\theta}-\theta)dt\\ &+\int_0^T \boldsymbol{\lambda}(t)^T\Big[\big(D_\theta\mathbf{M}(\theta)(\widetilde{\theta}-\theta)\big)\frac{d}{dt}\mathbf{y}(t)+\big(D_\theta\mathbf{A}(\theta)(\widetilde{\theta}-\theta)\big)\mathbf{y}(t)\\ &\qquad-\big(D_\theta\mathbf{B}(\theta)(\widetilde{\theta}-\theta)\big)\mathbf{u}(t)\Big]\,dt\geq 0\end{aligned} \tag{4.22c}$$

for all $\widetilde{\theta}\in\Theta$, and $\mathbf{y}(0)=\mathbf{y}_0$, $\boldsymbol{\lambda}(T)=\mathbf{0}$.

4.3.1 Domain decomposition

We assume that $\Omega(\theta)$ is decomposed into a subdomain Ω_1 independent of θ and a subdomain $\Omega_2(\theta)$ that depends on θ. More precisely, we assume

$$\overline{\Omega(\theta)}=\overline{\Omega_1}\cup\overline{\Omega_2(\theta)},\quad \Omega_1\cap\Omega_2(\theta)=\emptyset.$$

Moreover, we assume that the integrand ℓ in the objective function (4.1a) is of the form

$$\ell(\mathbf{y}(t),t,\theta)=\ \tfrac{1}{2}\|\mathbf{C}_I^{(1)}\mathbf{y}_I^{(1)}(t)-\mathbf{d}_I^{(1)}(t)\|^2+\widetilde{\ell}(\mathbf{y}_\Gamma(t),\mathbf{y}_I^{(2)}(t),t,\theta). \tag{4.23}$$

In the following section we will use domain decomposition to decompose the optimality conditions (4.22) into three components, one corresponding to the fixed subdomain Ω_1, one corresponding to the variable subdomain $\Omega_2(\theta)$, and one corresponding to the interface between Ω_1 and $\Omega_2(\theta)$. The decomposed problems will be used to identify linear quadratic subproblems corresponding to the fixed domain Ω_1, which will be reduced using balanced truncation model reduction.

We note that both subdomains Ω_1 and $\Omega_2(\theta)$ could be subdivided further. This additional structure can be used in the implementation of the balanced truncation and the optimization algorithm for the solution of the reduced shape optimization problem. However, the division of $\Omega(\theta)$ into Ω_1 and $\Omega_2(\theta)$ is enough to study the essential features of our approach.

We use a standard nonoverlapping domain decomposition approach (substructuring) to decompose the optimality system. See, e.g., [103, Ch. 4] and [112, Ch. 1]. Our notation follows that of [103, 112]. The finite element stiffness matrix can be decomposed into

$$\mathbf{A}(\theta) = \begin{pmatrix} \mathbf{A}_{II}^{(1)} & \mathbf{A}_{I\Gamma}^{(1)} & \mathbf{0} \\ \mathbf{A}_{\Gamma I}^{(1)} & \mathbf{A}_{\Gamma\Gamma}(\theta) & \mathbf{A}_{\Gamma I}^{(2)}(\theta) \\ \mathbf{0} & \mathbf{A}_{I\Gamma}^{(2)}(\theta) & \mathbf{A}_{II}^{(2)}(\theta) \end{pmatrix}$$

where

$$\mathbf{A}_{\Gamma\Gamma}(\theta) = \mathbf{A}_{\Gamma\Gamma}^{(1)} + \mathbf{A}_{\Gamma\Gamma}^{(2)}(\theta).$$

The matrices $\mathbf{M}$, $\mathbf{B}$ admit similar representations and the vectors $\mathbf{y}, \mathbf{u}$ can be structured accordingly.

In the following we frequently omit the argument t and, for example, simply write $\mathbf{y}_I^{(1)}$ instead of $\mathbf{y}_I^{(1)}(t)$.

Using the domain decomposition structure, the state equations (4.1b) can be written as

$$\mathbf{M}_{II}^{(1)}\frac{d}{dt}\mathbf{y}_I^{(1)} + \mathbf{M}_{I\Gamma}^{(1)}\frac{d}{dt}\mathbf{y}_\Gamma + \mathbf{A}_{II}^{(1)}\mathbf{y}_I^{(1)} + \mathbf{A}_{I\Gamma}^{(1)}\mathbf{y}_\Gamma = \mathbf{B}_I^{(1)}\mathbf{u}_I^{(1)}, \tag{4.24a}$$

$$\mathbf{M}_{II}^{(2)}(\theta)\frac{d}{dt}\mathbf{y}_I^{(2)} + \mathbf{M}_{I\Gamma}^{(2)}(\theta)\frac{d}{dt}\mathbf{y}_\Gamma + \mathbf{A}_{II}^{(2)}(\theta)\mathbf{y}_I^{(2)} + \mathbf{A}_{I\Gamma}^{(2)}(\theta)\mathbf{y}_\Gamma = \mathbf{B}_I^{(2)}(\theta)\mathbf{u}_I^{(2)}, \tag{4.24b}$$

$$\begin{aligned}\mathbf{M}_{\Gamma I}^{(1)}\frac{d}{dt}\mathbf{y}_I^{(1)} + \mathbf{M}_{\Gamma\Gamma}(\theta)\frac{d}{dt}\mathbf{y}_\Gamma + \mathbf{M}_{\Gamma I}^{(2)}\frac{d}{dt}\mathbf{y}_I^{(2)} \\ + \mathbf{A}_{\Gamma I}^{(1)}\mathbf{y}_I^{(1)} + \mathbf{A}_{\Gamma\Gamma}(\theta)\mathbf{y}_\Gamma + \mathbf{A}_{\Gamma I}^{(2)}\mathbf{y}_I^{(2)} = \mathbf{B}^\Gamma(\theta)\mathbf{u}_\Gamma.\end{aligned} \tag{4.24c}$$

The optimality conditions (4.22) can now be written as (4.24a-c) and the adjoint equations

$$-\mathbf{M}_{II}^{(1)}\frac{d}{dt}\boldsymbol{\lambda}_I^{(1)} - \mathbf{M}_{I\Gamma}^{(1)}\frac{d}{dt}\boldsymbol{\lambda}_\Gamma + \big(\mathbf{A}_{II}^{(1)}\big)^T\boldsymbol{\lambda}_I^{(1)} + \big(\mathbf{A}_{\Gamma I}^{(1)}\big)^T\boldsymbol{\lambda}_\Gamma = -\big(\mathbf{C}_I^{(1)}\big)^T\big(\mathbf{C}_I^{(1)}\mathbf{y}_I^{(1)} - \mathbf{d}_I^{(1)}\big), \tag{4.24d}$$

$$\begin{aligned}-\mathbf{M}_{II}^{(2)}(\theta)\frac{d}{dt}\boldsymbol{\lambda}_I^{(2)} - \mathbf{M}_{I\Gamma}^{(2)}(\theta)\frac{d}{dt}\boldsymbol{\lambda}_\Gamma + \big(\mathbf{A}_{II}^{(2)}(\theta)\big)^T\boldsymbol{\lambda}_I^{(2)} \\ + \big(\mathbf{A}_{\Gamma I}^{(2)}(\theta)\big)^T\boldsymbol{\lambda}_\Gamma = -\nabla_{\mathbf{y}_I^{(2)}}\widetilde{\ell}(\mathbf{y}_\Gamma, \mathbf{y}_I^{(2)}, t, \theta),\end{aligned} \tag{4.24e}$$

$$\begin{aligned}-\mathbf{M}_{\Gamma I}^{(1)}\frac{d}{dt}\boldsymbol{\lambda}_I^{(1)} - \mathbf{M}_{\Gamma\Gamma}(\theta)\frac{d}{dt}\boldsymbol{\lambda}_\Gamma - \mathbf{M}_{\Gamma I}^{(2)}\frac{d}{dt}\boldsymbol{\lambda}_I^{(2)} \\ + \big(\mathbf{A}_{I\Gamma}^{(1)}\big)^T\boldsymbol{\lambda}_I^{(1)} + \big(\mathbf{A}_{\Gamma\Gamma}(\theta)\big)^T\boldsymbol{\lambda}_\Gamma + \big(\mathbf{A}_{I\Gamma}^{(2)}\big)^T\boldsymbol{\lambda}_I^{(2)} = -\nabla_{\mathbf{y}_\Gamma}\widetilde{\ell}(\mathbf{y}_\Gamma, \mathbf{y}_I^{(2)}, t, \theta),\end{aligned} \tag{4.24f}$$

and

$$\begin{aligned}&\int_0^T D_\theta\widetilde{\ell}(\mathbf{y}_\Gamma, \mathbf{y}_I^{(2)}, t, \theta)(\widetilde{\theta} - \theta)dt \\ &+ \int_0^T \begin{pmatrix}\boldsymbol{\lambda}_\Gamma \\ \boldsymbol{\lambda}_I^{(2)}\end{pmatrix}^T \Big[\big(D_\theta\mathbf{M}^{(2)}(\theta)(\widetilde{\theta} - \theta)\big)\,\frac{d}{dt}\begin{pmatrix}\mathbf{y}_\Gamma \\ \mathbf{y}_I^{(2)}\end{pmatrix} + \big(D_\theta\mathbf{A}^{(2)}(\theta)(\widetilde{\theta} - \theta)\big)\begin{pmatrix}\mathbf{y}_\Gamma \\ \mathbf{y}_I^{(2)}\end{pmatrix} \\ &\qquad - \big(D_\theta\mathbf{B}^{(2)}(\theta)(\widetilde{\theta} - \theta)\big)\begin{pmatrix}\mathbf{u}_\Gamma \\ \mathbf{u}_I^{(2)}\end{pmatrix}\Big]\,dt \geq 0\end{aligned} \tag{4.24g}$$

for all $\widetilde{\theta} \in \Theta$.

We have set

$$\mathbf{M}^{(2)}(\theta) = \begin{pmatrix} \mathbf{M}_{\Gamma\Gamma}(\theta) & \mathbf{M}^{(2)}_{\Gamma I}(\theta) \\ \mathbf{M}^{(2)}_{I\Gamma}(\theta) & \mathbf{M}^{(2)}_{II}(\theta) \end{pmatrix}, \quad \mathbf{A}^{(2)}(\theta) = \begin{pmatrix} \mathbf{A}_{\Gamma\Gamma}(\theta) & \mathbf{A}^{(2)}_{\Gamma I}(\theta) \\ \mathbf{A}^{(2)}_{I\Gamma}(\theta) & \mathbf{A}^{(2)}_{II}(\theta) \end{pmatrix},$$
$$\mathbf{B}^{(2)}(\theta) = \begin{pmatrix} \mathbf{B}_{\Gamma}(\theta) \\ \mathbf{B}^{(2)}_{I}(\theta) \end{pmatrix}.$$

We apply balanced truncation model reduction to the optimality subsystem that corresponds to the fixed subdomain Ω_1.

4.3.2 Balanced truncation model reduction of the fixed subdomain problem

We will apply balanced truncation model reduction to the optimality subsystem that corresponds to the fixed subdomain Ω_1. To accomplish this we need to identify how $\mathbf{y}^{(1)}_I$ and $\boldsymbol{\lambda}^{(1)}_I$ in (4.24) interact with the other components of the system and we have to make sure that the resulting subsystem is of the form (4.3) to which balanced truncation can be applied. This is the reason why we have assumed that the integrand ℓ in the objective function (4.1a) is of the form (4.23).

If we inspect (4.24) to see how $\mathbf{y}^{(1)}_I$ and $\boldsymbol{\lambda}^{(1)}_I$ interact with the other components of the system, we are led to

$$\mathbf{M}^{(1)}_{II}\frac{d}{dt}\mathbf{y}^{(1)}_I = -\mathbf{A}^{(1)}_{II}\mathbf{y}^{(1)}_I - \mathbf{M}^{(1)}_{I\Gamma}\frac{d}{dt}\mathbf{y}_\Gamma + \mathbf{B}^{(1)}_I\mathbf{u}^{(1)}_I - \mathbf{A}^{(1)}_{I\Gamma}\mathbf{y}_\Gamma \tag{4.25a}$$
$$\mathbf{z}^{(1)}_I = -\mathbf{C}^{(1)}_I\mathbf{y}^{(1)}_I + \mathbf{d}^{(1)}_I, \tag{4.25b}$$
$$\mathbf{z}_\Gamma = -\mathbf{M}^{(1)}_{\Gamma I}\frac{d}{dt}\mathbf{y}^{(1)}_I - \mathbf{A}^{(1)}_{\Gamma I}\mathbf{y}^{(1)}_I, \tag{4.25c}$$

$$-\mathbf{M}_{II}^{(1)}\frac{d}{dt}\boldsymbol{\lambda}_I^{(1)} = -\left(\mathbf{A}_{II}^{(1)}\right)^T\boldsymbol{\lambda}_I^{(1)} + \mathbf{M}_{I\Gamma}^{(1)}\frac{d}{dt}\boldsymbol{\lambda}_\Gamma - \left(\mathbf{C}_I^{(1)}\right)^T\mathbf{w}_I^{(1)} - \left(\mathbf{A}_{\Gamma I}^{(1)}\right)^T\boldsymbol{\lambda}_\Gamma \tag{4.25d}$$

$$\mathbf{q}_I^{(1)} = \left(\mathbf{B}_I^{(1)}\right)^T\boldsymbol{\lambda}_I^{(1)}, \tag{4.25e}$$

$$\mathbf{q}_\Gamma = \mathbf{M}_{\Gamma I}^{(1)}\frac{d}{dt}\boldsymbol{\lambda}_I^{(1)} - \left(\mathbf{A}_{I\Gamma}^{(1)}\right)^T\boldsymbol{\lambda}_I^{(1)}. \tag{4.25f}$$

In fact (4.25a) and (4.25d) are identical to (4.24a) and (4.24d), respectively, if $\mathbf{w}_I^{(1)} = -\mathbf{z}_I^{(1)} = \mathbf{C}_I^{(1)}\mathbf{y}_I^{(1)} - \mathbf{d}_I^{(1)}$. The output (4.25b) enters into (4.24d) and the output (4.25c) enters into (4.24c). Similarly, the output (4.25f) enters into (4.24f). The output (4.25e) is included as an auxiliary variable. It does not enter into any of the equations in (4.24), but is included to establish the connection with the generic system (4.3).

If

$$\mathbf{M}_{I\Gamma}^{(1)} = \mathbf{0} \quad \text{and} \quad \mathbf{M}_{\Gamma I}^{(1)} = \mathbf{0}, \tag{4.26}$$

then (4.25) is given by

$$\mathbf{M}_{II}^{(1)}\frac{d}{dt}\mathbf{y}_I^{(1)} = -\mathbf{A}_{II}^{(1)}\mathbf{y}_I^{(1)} + \left(\mathbf{B}_I^{(1)} | -\mathbf{A}_{I\Gamma}^{(1)}\right)\begin{pmatrix}\mathbf{u}_I^{(1)}\\ \mathbf{y}_\Gamma\end{pmatrix} \tag{4.27a}$$

$$\begin{pmatrix}\mathbf{z}_I^{(1)}\\ \mathbf{z}_\Gamma\end{pmatrix} = \begin{pmatrix}-\mathbf{C}_I^{(1)}\\ -\mathbf{A}_{\Gamma I}^{(1)}\end{pmatrix}\mathbf{y}_I^{(1)} + \begin{pmatrix}\mathbf{I}\\ \mathbf{0}\end{pmatrix}\mathbf{d}_I^{(1)}, \tag{4.27b}$$

$$-\mathbf{M}_{II}^{(1)}\frac{d}{dt}\boldsymbol{\lambda}_I^{(1)} = -(\mathbf{A}_{II}^{(1)})^T\boldsymbol{\lambda}_I^{(1)} + \begin{pmatrix}-\mathbf{C}_I^{(1)}\\ -\mathbf{A}_{\Gamma I}^{(1)}\end{pmatrix}^T\begin{pmatrix}\mathbf{w}_I^{(1)}\\ \boldsymbol{\lambda}_\Gamma\end{pmatrix} \tag{4.27c}$$

$$\begin{pmatrix}\mathbf{q}_I^{(1)}\\ \mathbf{q}_\Gamma\end{pmatrix} = \left(\mathbf{B}_I^{(1)} | -\mathbf{A}_{I\Gamma}^{(1)}\right)^T\boldsymbol{\lambda}_I^{(1)}. \tag{4.27d}$$

This system is exactly of the form (4.3) that is needed for balanced truncation. We assume that

$$\mathbf{v}^T\mathbf{A}\mathbf{v} \le -\alpha\mathbf{v}^T\mathbf{M}\mathbf{v}, \qquad \forall\mathbf{v}\in\mathbb{R}^N. \tag{4.28}$$

Note that assumption (4.28) implies

$$\mathbf{v}^T \mathbf{A}_{II}^{(1)} \mathbf{v} \leq -\alpha \mathbf{v}^T \mathbf{M}_{II}^{(1)} \mathbf{v}, \qquad \forall \mathbf{v} \in \mathbb{R}^{N_I^{(1)}}. \tag{4.29}$$

As a consequence of (4.29) all eigenvalues of the pair $(\mathbf{A}_{II}^{(1)}, \mathbf{M}_{II}^{(1)})$ have negative real part and, hence, balanced truncation model reduction can be applied to (4.27) which leads to the following reduced subsystem

$$\frac{d}{dt}\widehat{\mathbf{y}}_I^{(1)} = -\widehat{\mathbf{A}}_{II}^{(1)}\widehat{\mathbf{y}}_I^{(1)} - \widehat{\mathbf{A}}_{I\Gamma}^{(1)}\mathbf{y}_\Gamma + \widehat{\mathbf{B}}_I^{(1)}\mathbf{u}_I^{(1)} \tag{4.30a}$$

$$\widehat{\mathbf{z}}_I^{(1)} = -\widehat{\mathbf{C}}_I^{(1)}\widehat{\mathbf{y}}_I^{(1)} + \mathbf{d}_I^{(1)}, \tag{4.30b}$$

$$\widehat{\mathbf{z}}_\Gamma = -\widehat{\mathbf{A}}_{\Gamma I}^{(1)}\widehat{\mathbf{y}}_I^{(1)}, \tag{4.30c}$$

$$-\frac{d}{dt}\widehat{\boldsymbol{\lambda}}_I^{(1)} = -\big(\widehat{\mathbf{A}}_{II}^{(1)}\big)^T\widehat{\boldsymbol{\lambda}}_I^{(1)} - \big(\widehat{\mathbf{A}}_{\Gamma I}^{(1)}\big)^T\boldsymbol{\lambda}_\Gamma - \big(\widehat{\mathbf{C}}_I^{(1)}\big)^T\mathbf{w}_I^{(1)} \tag{4.30d}$$

$$\widehat{\mathbf{q}}_I^{(1)} = \big(\widehat{\mathbf{B}}_I^{(1)}\big)^T\widehat{\boldsymbol{\lambda}}_I^{(1)}, \tag{4.30e}$$

$$\widehat{\mathbf{q}}_\Gamma = -\big(\widehat{\mathbf{A}}_{I\Gamma}^{(1)}\big)^T\widehat{\boldsymbol{\lambda}}_I^{(1)}. \tag{4.30f}$$

We assume that

$$\mathbf{y}_{I,0}^{(1)} = \mathbf{0}, \tag{4.31}$$

cf., Remark 4.1.1.

Balanced truncation generates a reduced order model (4.30) such that the error between the input-to-output maps of (4.25) and (4.30) can be estimated by

$$\left\| \begin{pmatrix} \mathbf{z}_I^{(1)} \\ \mathbf{z}_\Gamma \end{pmatrix} - \begin{pmatrix} \widehat{\mathbf{z}}_I^{(1)} \\ \widehat{\mathbf{z}}_\Gamma \end{pmatrix} \right\|_{L^2} \leq 2 \left\| \begin{pmatrix} \mathbf{u}_I^{(1)} \\ \mathbf{y}_\Gamma \end{pmatrix} \right\|_{L^2} \tau, \tag{4.32a}$$

$$\left\| \begin{pmatrix} \mathbf{q}_I^{(1)} \\ \mathbf{q}_\Gamma \end{pmatrix} - \begin{pmatrix} \widehat{\mathbf{q}}_I^{(1)} \\ \widehat{\mathbf{q}}_\Gamma \end{pmatrix} \right\|_{L^2} \leq 2 \left\| \begin{pmatrix} \mathbf{w}_I^{(1)} \\ \boldsymbol{\lambda}_\Gamma \end{pmatrix} \right\|_{L^2} \tau, \tag{4.32b}$$

where

$$\tau = \sigma_{n+1} + \ldots + \sigma_N. \tag{4.32c}$$

To be consistent with (4.26) we also assume that $\mathbf{M}_{I\Gamma}^{(2)} = \mathbf{0}$ and $\mathbf{M}_{\Gamma I}^{(2)} = \mathbf{0}$. The reduced order optimality system corresponding to (4.24) is given by the state equation

$$\frac{d}{dt}\widehat{\mathbf{y}}_I^{(1)} + \widehat{\mathbf{A}}_{II}^{(1)}\widehat{\mathbf{y}}_I^{(1)} + \widehat{\mathbf{A}}_{I\Gamma}^{(1)}\widehat{\mathbf{y}}_\Gamma = \widehat{\mathbf{B}}_I^{(1)}\mathbf{u}_I^{(1)}, \tag{4.33a}$$

$$\mathbf{M}_{II}^{(2)}\frac{d}{dt}\widehat{\mathbf{y}}_I^{(2)} + \mathbf{A}_{II}^{(2)}\widehat{\mathbf{y}}_I^{(2)} + \mathbf{A}_{I\Gamma}^{(2)}\widehat{\mathbf{y}}_\Gamma = \mathbf{B}_I^{(2)}\mathbf{u}_I^{(2)}, \tag{4.33b}$$

$$\mathbf{M}_{\Gamma\Gamma}\frac{d}{dt}\widehat{\mathbf{y}}_\Gamma + \widehat{\mathbf{A}}_{\Gamma I}^{(1)}\widehat{\mathbf{y}}_I^{(1)} + \mathbf{A}_{\Gamma\Gamma}\widehat{\mathbf{y}}_\Gamma + \mathbf{A}_{\Gamma I}^{(2)}\widehat{\mathbf{y}}_I^{(2)} = \mathbf{B}_\Gamma\mathbf{u}_\Gamma, \tag{4.33c}$$

the adjoint equation

$$-\frac{d}{dt}\widehat{\boldsymbol{\lambda}}_I^{(1)} + \big(\widehat{\mathbf{A}}_{II}^{(1)}\big)^T\widehat{\boldsymbol{\lambda}}_I^{(1)} + \big(\widehat{\mathbf{A}}_{\Gamma I}^{(1)}\big)^T\widehat{\boldsymbol{\lambda}}_\Gamma = -\big(\widehat{\mathbf{C}}_I^{(1)}\big)^T\big(\widehat{\mathbf{C}}_I^{(1)}\widehat{\mathbf{y}}_I^{(1)} - \mathbf{d}_I^{(1)}\big), \tag{4.33d}$$

$$-\mathbf{M}_{II}^{(2)}\frac{d}{dt}\widehat{\boldsymbol{\lambda}}_I^{(2)} + \big(\mathbf{A}_{II}^{(2)}\big)^T\widehat{\boldsymbol{\lambda}}_I^{(2)} + \big(\mathbf{A}_{\Gamma I}^{(2)}\big)^T\widehat{\boldsymbol{\lambda}}_\Gamma = -\nabla_{\widehat{\mathbf{y}}_I^{(2)}}\widetilde{\ell}(\widehat{\mathbf{y}}_\Gamma, \widehat{\mathbf{y}}_I^{(2)}, t, \theta), \tag{4.33e}$$

$$-\mathbf{M}_{\Gamma\Gamma}\frac{d}{dt}\widehat{\boldsymbol{\lambda}}_\Gamma + \big(\widehat{\mathbf{A}}_{I\Gamma}^{(1)}\big)^T\widehat{\boldsymbol{\lambda}}_I^{(1)} + \mathbf{A}_{\Gamma\Gamma}^T\widehat{\boldsymbol{\lambda}}_\Gamma + \big(\mathbf{A}_{I\Gamma}^{(2)}\big)^T\widehat{\boldsymbol{\lambda}}_I^{(2)} = -\nabla_{\widehat{\mathbf{y}}_\Gamma}\widetilde{\ell}(\widehat{\mathbf{y}}_\Gamma, \widehat{\mathbf{y}}_I^{(2)}, t, \theta), \tag{4.33f}$$

where $\mathbf{M}_{II}^{(2)} = \mathbf{M}_{II}^{(2)}(\theta)$, $\mathbf{M}_{\Gamma\Gamma} = \mathbf{M}_{\Gamma\Gamma}(\theta)$, $\mathbf{A}_{II}^{(2)} = \mathbf{A}_{II}^{(2)}(\theta)$, $\mathbf{A}_{\Gamma\Gamma} = \mathbf{A}_{\Gamma\Gamma}(\theta)$, $\mathbf{A}_{I\Gamma}^{(2)} = \mathbf{A}_{I\Gamma}^{(2)}(\theta)$, $\mathbf{A}_{\Gamma I}^{(2)} = \mathbf{A}_{\Gamma I}^{(2)}(\theta)$, $\mathbf{B}_I^{(2)} = \mathbf{B}_I^{(2)}(\theta)$, $\mathbf{B}_\Gamma = \mathbf{B}_\Gamma(\theta)$, and by

$$\begin{aligned}
&\int_0^T D_\theta\widetilde{\ell}(\widehat{\mathbf{y}}_\Gamma, \widehat{\mathbf{y}}_I^{(2)}, t, \theta)(\widetilde{\theta} - \theta)dt \\
&+ \int_0^T \begin{pmatrix} \widehat{\boldsymbol{\lambda}}_\Gamma \\ \widehat{\boldsymbol{\lambda}}_I^{(2)} \end{pmatrix}^T \Big[\big(D_\theta\mathbf{M}^{(2)}(\theta)(\widetilde{\theta} - \theta)\big)\,\frac{d}{dt}\begin{pmatrix} \widehat{\mathbf{y}}_\Gamma \\ \widehat{\mathbf{y}}_I^{(2)} \end{pmatrix} + \big(D_\theta\mathbf{A}^{(2)}(\theta)(\widetilde{\theta} - \theta)\big)\begin{pmatrix} \widehat{\mathbf{y}}_\Gamma \\ \widehat{\mathbf{y}}_I^{(2)} \end{pmatrix} \\
&\qquad - \big(D_\theta\mathbf{B}^{(2)}(\theta)(\widetilde{\theta} - \theta)\big)\begin{pmatrix} \mathbf{u}_\Gamma \\ \mathbf{u}_I^{(2)} \end{pmatrix}\Big]\,dt \geq 0
\end{aligned} \tag{4.33g}$$

for all $\widetilde{\theta} \in \Theta$.

The reduced order optimality system (4.33) is the first-order necessary optimality system for the reduced order semidiscretized shape optimization problem

$$\text{minimize } \int_0^T \ \tfrac{1}{2}\|\widehat{\mathbf{C}}_I^{(1)}\widehat{\mathbf{y}}_I^{(1)}(t) - \mathbf{d}_I^{(1)}(t)\|_2^2 + \widetilde{\ell}(\widehat{\mathbf{y}}_\Gamma(t), \widehat{\mathbf{y}}_I^{(2)}(t), t, \theta)\,dt, \tag{4.34}$$

subject to (4.33a-c) with initial conditions $\widehat{\mathbf{y}}_I^{(1)}(0) = \widehat{\mathbf{y}}_{I,0}^{(1)}$, $\widehat{\mathbf{y}}_I^{(2)}(0) = \mathbf{y}_{I,0}^{(2)}$, $\widehat{\mathbf{y}}_\Gamma(0) = \mathbf{y}_{\Gamma,0}$ and parameter constraints $\theta \in \Theta$.

4.3.3 Error analysis

We define the objective functions

$$J(\theta) = \int_0^T \tfrac{1}{2}\|\mathbf{C}_I^{(1)}\mathbf{y}_I^{(1)}(t) - \mathbf{d}_I^{(1)}(t)\|_2^2 + \widetilde{\ell}(\mathbf{y}_\Gamma(t), \mathbf{y}_I^{(2)}(t), t, \theta)\, dt,$$
$$\widehat{J}(\theta) = \int_0^T \tfrac{1}{2}\|\widehat{\mathbf{C}}_I^{(1)}\widehat{\mathbf{y}}_I^{(1)}(t) - \mathbf{d}_I^{(1)}(t)\|_2^2 + \widetilde{\ell}(\widehat{\mathbf{y}}_\Gamma(t), \widehat{\mathbf{y}}_I^{(2)}(t), t, \theta)\, dt,$$

where $\mathbf{y}_I^{(1)}, \mathbf{y}_I^{(2)}, \mathbf{y}_\Gamma$ solve (4.24a-c) and where $\widehat{\mathbf{y}}_I^{(1)}, \widehat{\mathbf{y}}_I^{(2)}, \widehat{\mathbf{y}}_\Gamma$ solve (4.33a-c). Using these objective functions, which treat the states $\mathbf{y}_I^{(1)}, \mathbf{y}_I^{(2)}, \mathbf{y}_\Gamma$ and $\widehat{\mathbf{y}}_I^{(1)}, \widehat{\mathbf{y}}_I^{(2)}, \widehat{\mathbf{y}}_\Gamma$ as implicit functions of $\theta \in \Theta$, the optimization problems (4.1) and (4.34) can be written as

$$\min_{\theta\in\Theta} J(\theta) \quad \text{and} \quad \min_{\theta\in\Theta} \widehat{J}(\theta)$$

respectively. Recall that Θ is a closed convex set. If $\theta_* \in \Theta$ and $\widehat{\theta}_* \in \Theta$ are solutions of these problems, then

$$\nabla J(\theta_*)^T(\theta - \theta_*) \geq 0, \qquad \nabla \widehat{J}(\widehat{\theta}_*)^T(\theta - \widehat{\theta}_*) \geq 0 \tag{4.35}$$

for all $\theta \in \Theta$. This implies

$$(\nabla J(\theta_*) - \nabla \widehat{J}(\widehat{\theta}_*))^T(\widehat{\theta}_* - \theta_*) \geq 0 \tag{4.36}$$

If we assume the convexity condition

$$(\nabla J(\widehat{\theta}_*) - \nabla J(\theta_*))^T(\widehat{\theta}_* - \theta_*) \geq \kappa\|\widehat{\theta}_* - \theta_*\|^2, \tag{4.37}$$

then combining (4.36) and (4.37) leads to

$$(\nabla J((\widehat{\theta}_*) - \nabla \widehat{J}(\widehat{\theta}_*))^T(\widehat{\theta}_* - \theta_*) \geq \kappa\|\widehat{\theta}_* - \theta_*\|^2.$$

Hence, we have the error estimate

$$\|\theta_* - \widehat{\theta}_*\| \leq \kappa^{-1} \|\nabla \widehat{J}(\widehat{\theta}_*) - \nabla J(\widehat{\theta}_*)\|. \tag{4.38}$$

As before, assuming (4.37), an estimate of the error in the solution of (4.1) and (4.34) requires an estimate of the error in the gradient of the full and the reduced order optimization problem.

The gradients are given by

$$\begin{aligned}
&\nabla J(\theta)^T \widetilde{\theta} \\
&= \int_0^T D_\theta \widetilde{\ell}(\mathbf{y}_\Gamma(t), \mathbf{y}_I^{(2)}(t), t, \theta) \widetilde{\theta} dt \\
&\quad + \int_0^T \begin{pmatrix} \boldsymbol{\lambda}_\Gamma(t) \\ \boldsymbol{\lambda}_I^{(2)}(t) \end{pmatrix}^T \Big\{ \big(D_\theta \mathbf{M}^{(2)}(\theta)\widetilde{\theta}\big) \, \frac{d}{dt} \begin{pmatrix} \mathbf{y}_\Gamma(t) \\ \mathbf{y}_I^{(2)}(t) \end{pmatrix} \\
&\qquad + \big(D_\theta \mathbf{A}^{(2)}(\theta)\widetilde{\theta}\big) \begin{pmatrix} \mathbf{y}_\Gamma(t) \\ \mathbf{y}_I^{(2)}(t) \end{pmatrix} - \big(D_\theta \mathbf{B}^{(2)}(\theta)\widetilde{\theta}\big) \begin{pmatrix} \mathbf{u}_\Gamma(t) \\ \mathbf{u}_I^{(2)}(t) \end{pmatrix} \Big\} dt
\end{aligned}$$

where $\mathbf{y}_I^{(1)}$, $\mathbf{y}_I^{(2)}$, $\mathbf{y}_\Gamma$, $\boldsymbol{\lambda}_I^{(1)}$, $\boldsymbol{\lambda}_I^{(2)}$, $\boldsymbol{\lambda}_\Gamma$ solve (4.24a-f), and

$$\begin{aligned}
&\nabla \widehat{J}(\theta)^T \widetilde{\theta} \\
&= \int_0^T D_\theta \widetilde{\ell}(\widehat{\mathbf{y}}_\Gamma(t), \widehat{\mathbf{y}}_I^{(2)}(t), t, \theta) \widetilde{\theta} dt \\
&\quad + \int_0^T \begin{pmatrix} \widehat{\boldsymbol{\lambda}}_\Gamma(t) \\ \widehat{\boldsymbol{\lambda}}_I^{(2)}(t) \end{pmatrix}^T \Big\{ \big(D_\theta \mathbf{M}^{(2)}(\theta)\widetilde{\theta}\big) \, \frac{d}{dt} \begin{pmatrix} \widehat{\mathbf{y}}_\Gamma(t) \\ \widehat{\mathbf{y}}_I^{(2)}(t) \end{pmatrix} \\
&\qquad + \big(D_\theta \mathbf{A}^{(2)}(\theta)\widetilde{\theta}\big) \begin{pmatrix} \widehat{\mathbf{y}}_\Gamma(t) \\ \widehat{\mathbf{y}}_I^{(2)}(t) \end{pmatrix} - \big(D_\theta \mathbf{B}^{(2)}(\theta)\widetilde{\theta}\big) \begin{pmatrix} \mathbf{u}_\Gamma(t) \\ \mathbf{u}_I^{(2)}(t) \end{pmatrix} \Big\} dt
\end{aligned}$$

where $\widehat{\mathbf{y}}_I^{(1)}$, $\widehat{\mathbf{y}}_I^{(2)}$, $\widehat{\mathbf{y}}_\Gamma$, $\widehat{\boldsymbol{\lambda}}_I^{(1)}$, $\widehat{\boldsymbol{\lambda}}_I^{(2)}$, $\widehat{\boldsymbol{\lambda}}_\Gamma$ solve (4.33a-f), respectively.

The difference between the full and the reduced gradients is given by

$$\begin{aligned}
&\Big(\nabla J(\theta) - \nabla \widehat{J}(\theta)\Big)^T \widetilde{\theta} \\
&= \int_0^T \Big(D_\theta \widetilde{\ell}(\mathbf{y}_\Gamma, \mathbf{y}_I^{(2)}, t, \theta) - D_\theta \widetilde{\ell}(\widehat{\mathbf{y}}_\Gamma, \widehat{\mathbf{y}}_I^{(2)}, t, \theta) \Big) \widetilde{\theta} dt \\
&\quad + \int_0^T \begin{pmatrix} \boldsymbol{\lambda}_\Gamma \\ \boldsymbol{\lambda}_I^{(2)} \end{pmatrix}^T \Big\{ \big(D_\theta \mathbf{M}^{(2)}(\theta)\widetilde{\theta} \big) \frac{d}{dt} \begin{pmatrix} \mathbf{y}_\Gamma - \widehat{\mathbf{y}}_\Gamma \\ \mathbf{y}_I^{(2)} - \widehat{\mathbf{y}}_I^{(2)} \end{pmatrix} + \big(D_\theta \mathbf{A}^{(2)}(\theta)\widetilde{\theta} \big) \begin{pmatrix} \mathbf{y}_\Gamma - \widehat{\mathbf{y}}_\Gamma \\ \mathbf{y}_I^{(2)} - \widehat{\mathbf{y}}_I^{(2)} \end{pmatrix} \Big\} dt \\
&\quad + \int_0^T \begin{pmatrix} \boldsymbol{\lambda}_\Gamma - \widehat{\boldsymbol{\lambda}}_\Gamma \\ \boldsymbol{\lambda}_I^{(2)} - \widehat{\boldsymbol{\lambda}}_I^{(2)} \end{pmatrix}^T \Big\{ \big(D_\theta \mathbf{M}^{(2)}(\theta)\widetilde{\theta} \big) \frac{d}{dt} \begin{pmatrix} \widehat{\mathbf{y}}_\Gamma \\ \widehat{\mathbf{y}}_I^{(2)} \end{pmatrix} + \big(D_\theta \mathbf{A}^{(2)}(\theta)\widetilde{\theta} \big) \begin{pmatrix} \widehat{\mathbf{y}}_\Gamma \\ \widehat{\mathbf{y}}_I^{(2)} \end{pmatrix} \\
&\qquad\qquad - \big(D_\theta \mathbf{B}^{(2)}(\theta)\widetilde{\theta} \big) \begin{pmatrix} \mathbf{u}_\Gamma \\ \mathbf{u}_I^{(2)} \end{pmatrix} \Big\} dt. \qquad (4.39)
\end{aligned}$$

We begin with an estimate of the error in the states.

Lemma 4.3.1 *Let* (4.20) *and* (4.28) *be valid. If* $\mathbf{y}_I^{(1)}$, $\mathbf{y}_I^{(2)}$, $\mathbf{y}_\Gamma$ *solve (4.24a-c), and* $\widehat{\mathbf{y}}_I^{(1)}$, $\widehat{\mathbf{y}}_I^{(2)}$, $\widehat{\mathbf{y}}_\Gamma$ *solve (4.33a-c), then*

$$\Big\| \mathbf{C}_I^{(1)} \mathbf{y}_I^{(1)} - \widehat{\mathbf{C}}_I^{(1)} \widehat{\mathbf{y}}_I^{(1)} \Big\|_{L^2} \leq \left(2 + \frac{4 \|\mathbf{M}^{-1}\| \|\mathbf{C}_I^{(1)}\|}{\alpha} \right) \left\| \begin{pmatrix} \mathbf{u}_I^{(1)} \\ \widehat{\mathbf{y}}_\Gamma \end{pmatrix} \right\|_{L^2} \tau \qquad (4.40a)$$

and

$$\left\| \begin{pmatrix} \mathbf{y}_I^{(2)} - \widehat{\mathbf{y}}_I^{(2)} \\ \mathbf{y}_\Gamma - \widehat{\mathbf{y}}_\Gamma \end{pmatrix} \right\|_{L^2} \leq \frac{4 \|\mathbf{M}^{-1}\|}{\alpha} \left\| \begin{pmatrix} \mathbf{u}_I^{(1)} \\ \widehat{\mathbf{y}}_\Gamma \end{pmatrix} \right\|_{L^2} \tau, \qquad (4.40b)$$

where $\tau = \sigma_{n+1} + \ldots + \sigma_N$.

Proof: Let $\mathbf{y}_I^{(1)}$, $\mathbf{y}_I^{(2)}$, $\mathbf{y}_\Gamma$ solve (4.24a-c), and let $\widehat{\mathbf{y}}_I^{(1)}$, $\widehat{\mathbf{y}}_I^{(2)}$, $\widehat{\mathbf{y}}_\Gamma$ solve (4.33a-c). Furthermore, let $\widetilde{\mathbf{y}}_I^{(1)}$ solve

$$\mathbf{M}_{II}^{(1)} \frac{d}{dt} \widetilde{\mathbf{y}}_I^{(1)}(t) + \mathbf{A}_{II}^{(1)} \widetilde{\mathbf{y}}_I^{(1)}(t) + \mathbf{A}_{I\Gamma}^{(1)} \widehat{\mathbf{y}}_\Gamma(t) = \mathbf{B}_I^{(1)} \mathbf{u}_I^{(1)}(t) \qquad (4.41)$$

with initial condition $\widetilde{\mathbf{y}}_I^{(1)}(0) = \mathbf{y}_{I,0}^{(1)}$.

The balanced truncation error bound (4.32) implies

$$\left\| \begin{pmatrix} \mathbf{C}_I^{(1)}\widetilde{\mathbf{y}}_I^{(1)} - \widehat{\mathbf{C}}_I^{(1)}\widehat{\mathbf{y}}_I^{(1)} \\ \mathbf{A}_{\Gamma I}^{(1)}\widetilde{\mathbf{y}}_\Gamma - \widehat{\mathbf{A}}_{\Gamma I}^{(1)}\widehat{\mathbf{y}}_\Gamma \end{pmatrix} \right\|_{L^2} \leq 2 \left\| \begin{pmatrix} \mathbf{u}_I^{(1)} \\ \widehat{\mathbf{y}}_\Gamma \end{pmatrix} \right\|_{L^2} \tau. \tag{4.42}$$

The equations (4.24a-c), (4.33a-c), and (4.41) give

$$\mathbf{M}_{II}^{(1)}(\theta)\frac{d}{dt}(\mathbf{y}_I^{(1)} - \widetilde{\mathbf{y}}_I^{(1)}) + \mathbf{A}_{II}^{(1)}(\theta)(\mathbf{y}_I^{(1)} - \widetilde{\mathbf{y}}_I^{(1)}) + \mathbf{A}_{I\Gamma}^{(1)}(\theta)(\mathbf{y}_\Gamma - \widehat{\mathbf{y}}_\Gamma) = 0, \tag{4.43a}$$

$$\mathbf{M}_{II}^{(2)}(\theta)\frac{d}{dt}(\mathbf{y}_I^{(2)} - \widehat{\mathbf{y}}_I^{(2)}) + \mathbf{A}_{II}^{(2)}(\theta)(\mathbf{y}_I^{(2)} - \widehat{\mathbf{y}}_I^{(2)}) + \mathbf{A}_{I\Gamma}^{(2)}(\theta)(\mathbf{y}_\Gamma - \widehat{\mathbf{y}}_\Gamma) = 0, \tag{4.43b}$$

$$\begin{aligned} &\mathbf{M}_{\Gamma\Gamma}(\theta)\frac{d}{dt}(\mathbf{y}_\Gamma - \widehat{\mathbf{y}}_\Gamma) + \mathbf{A}_{\Gamma\Gamma}(\theta)(\mathbf{y}_\Gamma - \widehat{\mathbf{y}}_\Gamma) \\ &+\mathbf{A}_{\Gamma I}^{(1)}(\mathbf{y}_I^{(1)} - \widetilde{\mathbf{y}}_I^{(1)}) + \mathbf{A}_{\Gamma I}^{(2)}(\theta)(\mathbf{y}_I^{(2)} - \widehat{\mathbf{y}}_I^{(2)}) = \widehat{\mathbf{A}}_{\Gamma I}^{(1)}\widehat{\mathbf{y}}_I^{(1)} - \mathbf{A}_{\Gamma I}^{(1)}\widetilde{\mathbf{y}}_I^{(1)} \end{aligned} \tag{4.43c}$$

with initial conditions $\mathbf{y}_I^{(1)}(0) - \widetilde{\mathbf{y}}_I^{(1)}(0) = 0$, $\mathbf{y}_I^{(2)}(0) - \widehat{\mathbf{y}}_I^{(2)}(0) = 0$, $\mathbf{y}_\Gamma(0) - \widehat{\mathbf{y}}_\Gamma(0) = 0$.

Application of Lemma A.1.2 in the Appendix A to (4.43) followed by an application of (4.42) gives

$$\left\| \begin{pmatrix} \mathbf{y}_I^{(1)} - \widetilde{\mathbf{y}}_I^{(1)} \\ \mathbf{y}_I^{(2)} - \widehat{\mathbf{y}}_I^{(2)} \\ \mathbf{y}_\Gamma - \widehat{\mathbf{y}}_\Gamma \end{pmatrix} \right\|_{L^2} \leq \frac{2\|\mathbf{M}^{-1}\|}{\alpha} \left\| \widehat{\mathbf{A}}_{\Gamma I}^{(1)}\widehat{\mathbf{y}}_I^{(1)} - \mathbf{A}_{\Gamma I}^{(1)}\widetilde{\mathbf{y}}_I^{(1)} \right\|_{L^2} \leq \frac{4\|\mathbf{M}^{-1}\|}{\alpha} \left\| \begin{pmatrix} \mathbf{u}_I^{(1)} \\ \widehat{\mathbf{y}}_\Gamma \end{pmatrix} \right\|_{L^2} \tau. \tag{4.44}$$

This implies (4.40b). The estimate (4.40a) follows from (4.42) and (4.44). □

The errors in the adjoints are estimated similarly.

Lemma 4.3.2 *Let* (4.20) *and* (4.28) *be valid and assume that*

$$\|\nabla_{\mathbf{y}_\Gamma}\widetilde{\ell}(\mathbf{y}_I^{(2)},\mathbf{y}_\Gamma,t,\theta)-\nabla_{\mathbf{y}_\Gamma}\widetilde{\ell}(\widetilde{\mathbf{y}}_I^{(2)},\widetilde{\mathbf{y}}_\Gamma,t,\theta)\|\le L\left(\|\mathbf{y}_I^{(2)}-\widetilde{\mathbf{y}}_I^{(2)}\|^2+\|\mathbf{y}_\Gamma-\widetilde{\mathbf{y}}_\Gamma\|^2\right)^{1/2},$$
$$\|\nabla_{\mathbf{y}_I^{(2)}}\widetilde{\ell}(\mathbf{y}_I^{(2)},\mathbf{y}_\Gamma,t,\theta)-\nabla_{\mathbf{y}_I^{(2)}}\widetilde{\ell}(\widetilde{\mathbf{y}}_I^{(2)},\widetilde{\mathbf{y}}_\Gamma,t,\theta)\|\le L\left(\|\mathbf{y}_I^{(2)}-\widetilde{\mathbf{y}}_I^{(2)}\|^2+\|\mathbf{y}_\Gamma-\widetilde{\mathbf{y}}_\Gamma\|^2\right)^{1/2}$$

for all $\mathbf{y}_I^{(2)}-\widetilde{\mathbf{y}}_I^{(2)}\in\mathbb{R}^{N_I^{(2)}}$, $\mathbf{y}_\Gamma-\widetilde{\mathbf{y}}_\Gamma\in\mathbb{R}^{N_\Gamma}$, $\theta\in\Theta$. *If* $\mathbf{y}_I^{(1)}$, $\mathbf{y}_I^{(2)}$, $\mathbf{y}_\Gamma$, $\boldsymbol{\lambda}_I^{(1)}$, $\boldsymbol{\lambda}_I^{(2)}$, $\boldsymbol{\lambda}_\Gamma$ *solve (4.24a-f), and* $\widehat{\mathbf{y}}_I^{(1)}$, $\widehat{\mathbf{y}}_I^{(2)}$, $\widehat{\mathbf{y}}_\Gamma$, $\widehat{\boldsymbol{\lambda}}_I^{(1)}$, $\widehat{\boldsymbol{\lambda}}_I^{(2)}$, $\widehat{\boldsymbol{\lambda}}_\Gamma$ *solve (4.33a-f), then*

$$\left\|\begin{pmatrix}\boldsymbol{\lambda}_I^{(2)}-\widehat{\boldsymbol{\lambda}}_I^{(2)}\\ \boldsymbol{\lambda}_\Gamma-\widehat{\boldsymbol{\lambda}}_\Gamma\end{pmatrix}\right\|_{L^2}\le c_\lambda(\sigma_{n+1}+\ldots+\sigma_N),\tag{4.45}$$

where

$$c_\lambda=\frac{4\|\mathbf{M}^{-1}\|}{\alpha}\left\|\begin{pmatrix}\widehat{\mathbf{C}}_I^{(1)}\widehat{\mathbf{y}}_I^{(1)}-\mathbf{d}_I^{(1)}\\ \widehat{\boldsymbol{\lambda}}_\Gamma\end{pmatrix}\right\|_{L^2}$$
$$+\left(\frac{2\|\mathbf{C}_I^{(1)}\|\|\mathbf{M}^{-1}\|}{\alpha}\Big(2+\frac{4\|\mathbf{C}_I^{(1)}\|\|\mathbf{M}^{-1}\|}{\alpha}\Big)+\frac{8L\|\mathbf{M}^{-1}\|^2}{\alpha^2}\right)\left\|\begin{pmatrix}\mathbf{u}_I^{(1)}\\ \widehat{\mathbf{y}}_\Gamma\end{pmatrix}\right\|_{L^2}.$$

Proof: Let $\mathbf{y}_I^{(1)}$, $\mathbf{y}_I^{(2)}$, $\mathbf{y}_\Gamma$, $\boldsymbol{\lambda}_I^{(1)}$, $\boldsymbol{\lambda}_I^{(2)}$, $\boldsymbol{\lambda}_\Gamma$ solve (4.24a-f), and $\widehat{\mathbf{y}}_I^{(1)}$, $\widehat{\mathbf{y}}_I^{(2)}$, $\widehat{\mathbf{y}}_\Gamma$, $\widehat{\boldsymbol{\lambda}}_I^{(1)}$, $\widehat{\boldsymbol{\lambda}}_I^{(2)}$, $\widehat{\boldsymbol{\lambda}}_\Gamma$ solve (4.33a-f) and set

$$\widehat{\mathbf{z}}_I^{(1)}=\widehat{\mathbf{C}}_I^{(1)}\widehat{\mathbf{y}}_I^{(1)}.$$

Furthermore, let $\widetilde{\boldsymbol{\lambda}}_I^{(1)}$ solve

$$-\mathbf{M}_{II}^{(1)}\frac{d}{dt}\widetilde{\boldsymbol{\lambda}}_I^{(1)}(t)+(\mathbf{A}_{II}^{(1)})^T\widetilde{\boldsymbol{\lambda}}_I^{(1)}(t)+(\mathbf{A}_{\Gamma I}^{(1)})^T\widehat{\boldsymbol{\lambda}}_\Gamma(t)=-(\mathbf{C}_I^{(1)})(\widehat{\mathbf{C}}_I^{(1)}\widehat{\mathbf{y}}_I^{(1)}-\mathbf{d}_I^{(1)})\tag{4.46}$$

with the final condition $\widetilde{\boldsymbol{\lambda}}_I^{(1)}(T)=0$.

The balanced truncation error bound (4.32) implies

$$\left\|\begin{pmatrix}(\mathbf{B}_I^{(1)})^T\widetilde{\boldsymbol{\lambda}}_I^{(1)}-(\widehat{\mathbf{B}}_I^{(1)})^T\widehat{\boldsymbol{\lambda}}_I^{(1)}\\ (\mathbf{A}_{I\Gamma}^{(1)})^T\widetilde{\boldsymbol{\lambda}}_I^{(1)}-(\widehat{\mathbf{A}}_{I\Gamma}^{(1)})^T\widehat{\boldsymbol{\lambda}}_I^{(1)}\end{pmatrix}\right\|_{L^2}\le 2\left\|\begin{pmatrix}\widehat{\mathbf{C}}_I^{(1)}\widehat{\mathbf{y}}_I^{(1)}-\mathbf{d}_I^{(1)}\\ \widehat{\boldsymbol{\lambda}}_\Gamma\end{pmatrix}\right\|_{L^2}(\sigma_{n+1}+\ldots+\sigma_N).\tag{4.47}$$

The equations (4.24d-f), (4.33d-f), and (4.46) imply

$$\begin{aligned}
-\mathbf{M}_{II}^{(1)}\frac{d}{dt}(\boldsymbol{\lambda}_I^{(1)}-\widetilde{\boldsymbol{\lambda}}_I^{(1)})+(\mathbf{A}_{II}^{(1)})^T(\boldsymbol{\lambda}_I^{(1)}-\widetilde{\boldsymbol{\lambda}}_I^{(1)})&\\
+(\mathbf{A}_{\Gamma I}^{(1)})^T(\boldsymbol{\lambda}_\Gamma-\widehat{\boldsymbol{\lambda}}_\Gamma)=&-(\mathbf{C}_I^{(1)})(\mathbf{C}_I^{(1)}\mathbf{y}_I^{(1)}-\widehat{\mathbf{C}}_I^{(1)}\widehat{\mathbf{y}}_I^{(1)}),\\
-\mathbf{M}_{II}^{(2)}(\theta)\frac{d}{dt}(\boldsymbol{\lambda}_I^{(2)}-\widehat{\boldsymbol{\lambda}}_I^{(2)})+(\mathbf{A}_{II}^{(2)}(\theta))^T(\boldsymbol{\lambda}_I^{(2)}-\widehat{\boldsymbol{\lambda}}_I^{(2)})&\\
+(\mathbf{A}_{\Gamma I}^{(2)}(\theta))^T(\boldsymbol{\lambda}_\Gamma-\widehat{\boldsymbol{\lambda}}_\Gamma)=&-(\nabla_{\mathbf{y}_I^{(2)}}\ell(\mathbf{y}_I^{(2)},\mathbf{y}_\Gamma,\theta,t)\\
&-\nabla_{\widehat{\mathbf{y}}_I^{(2)}}\ell(\widehat{\mathbf{y}}_I^{(2)},\widehat{\mathbf{y}}_\Gamma,\theta,t)),\\
-\mathbf{M}_{\Gamma\Gamma}(\theta)\frac{d}{dt}(\boldsymbol{\lambda}_\Gamma-\widehat{\boldsymbol{\lambda}}_\Gamma)+(\mathbf{A}_{\Gamma\Gamma}(\theta))^T(\boldsymbol{\lambda}_\Gamma-\widehat{\boldsymbol{\lambda}}_\Gamma)&\\
+(\mathbf{A}_{I\Gamma}^{(1)})^T(\boldsymbol{\lambda}_I^{(1)}-\widetilde{\boldsymbol{\lambda}}_I^{(1)})+(\mathbf{A}_{I\Gamma}^{(2)}(\theta))^T(\boldsymbol{\lambda}_I^{(2)}-\widehat{\boldsymbol{\lambda}}_I^{(2)})=&(\widehat{\mathbf{A}}_{I\Gamma}^{(1)})^T\widehat{\boldsymbol{\lambda}}_I^{(1)}-(\mathbf{A}_{I\Gamma}^{(1)})^T\widetilde{\boldsymbol{\lambda}}_I^{(1)}\\
&-(\nabla_{\mathbf{y}_\Gamma}\ell(\mathbf{y}_I^{(2)},\mathbf{y}_\Gamma,\theta,t)\\
&-\nabla_{\widehat{\mathbf{y}}_\Gamma}\ell(\widehat{\mathbf{y}}_I^{(2)},\widehat{\mathbf{y}}_\Gamma,\theta,t))
\end{aligned}$$

with final conditions $\boldsymbol{\lambda}_I^{(1)}(T)=\widetilde{\boldsymbol{\lambda}}_I^{(1)}(T)=0$, $\boldsymbol{\lambda}_I^{(2)}(T)=\widehat{\boldsymbol{\lambda}}_I^{(2)}(T)=0$, and $\boldsymbol{\lambda}_\Gamma(T)=\widehat{\boldsymbol{\lambda}}_\Gamma(T)=0$. Lemma A.1.2 gives the estimate

$$\begin{aligned}
\left\|\begin{pmatrix}\boldsymbol{\lambda}_I^{(1)}-\widetilde{\boldsymbol{\lambda}}_I^{(1)}\\ \boldsymbol{\lambda}_I^{(2)}-\widehat{\boldsymbol{\lambda}}_I^{(2)}\\ \boldsymbol{\lambda}_\Gamma-\widehat{\boldsymbol{\lambda}}_\Gamma\end{pmatrix}\right\|_{L^2}&\leq\frac{2\|\mathbf{M}^{-1}\|}{\alpha}\|\mathbf{C}_I^{(1)}\|\|\mathbf{C}_I^{(1)}\mathbf{y}_I^{(1)}-\widehat{\mathbf{C}}_I^{(1)}\widehat{\mathbf{y}}_I^{(1)}\|_{L^2}\\
&+\frac{2\|\mathbf{M}^{-1}\|}{\alpha}\|(\widehat{\mathbf{A}}_{I\Gamma}^{(1)})^T\widehat{\boldsymbol{\lambda}}_I^{(1)}-(\mathbf{A}_{I\Gamma}^{(1)})^T\widetilde{\boldsymbol{\lambda}}_I^{(1)}\|_{L^2}\\
&+\frac{2L\|\mathbf{M}^{-1}\|}{\alpha}\left\|\begin{pmatrix}\mathbf{y}_I^{(2)}-\widehat{\mathbf{y}}_I^{(2)}\\ \mathbf{y}_\Gamma-\widehat{\mathbf{y}}_\Gamma\end{pmatrix}\right\|_{L^2}.
\end{aligned}\tag{4.49}$$

The error estimate follows from (4.49), (4.40) and (4.47). □

Equation (4.39) and Lemmas 4.3.1, 4.3.2 imply the following result:

Theorem 4.3.3 *Let the assumptions of Lemma 4.3.2 be valid and assume that*

$$\|\nabla_\theta \widetilde{\ell}(\mathbf{y}_I^{(2)}, \mathbf{y}_\Gamma, t, \theta) - \nabla_\theta \widetilde{\ell}(\widetilde{\mathbf{y}}_I^{(2)}, \widetilde{\mathbf{y}}_\Gamma, t, \theta)\| \leq L \left(\|\mathbf{y}_I^{(2)} - \widetilde{\mathbf{y}}_I^{(2)}\|^2 + \|\mathbf{y}_\Gamma - \widetilde{\mathbf{y}}_\Gamma\|^2 \right)^{1/2}$$

for all $\mathbf{y}_I^{(2)} - \widetilde{\mathbf{y}}_I^{(2)} \in \mathbb{R}^{N_I^{(2)}}$, $\mathbf{y}_\Gamma - \widetilde{\mathbf{y}}_\Gamma \in \mathbb{R}^{N_\Gamma}$, $\theta \in \Theta$, *and*

$$\max \left\{ \|D_\theta \mathbf{M}^{(2)}(\theta)\widetilde{\theta}\|, \|D_\theta \mathbf{A}^{(2)}(\theta)\widetilde{\theta}\|, \|D_\theta \mathbf{B}^{(2)}(\theta)\widetilde{\theta}\| \right\} \leq \gamma$$

for all $\|\widetilde{\theta}\| \leq 1$ *and all* $\theta \in \Theta$. *Then, there exists* $c > 0$ *dependent on* $\mathbf{u}$, $\widehat{\mathbf{y}}$, *and* $\widehat{\boldsymbol{\lambda}}$ *such that*

$$\|\nabla J(\theta) - \nabla \widehat{J}(\theta)\|_{L^2} \leq \frac{c}{\alpha}(\sigma_{n+1} + ... + \sigma_N).$$

Proof: The inequality follows directly from equation (4.39) and Lemmas 4.3.1, 4.3.2. □

Corollary 4.3.4 *If the assumptions of Theorem 4.3.3 and* (4.37) *hold, then there exists* $c > 0$ *dependent on* $\mathbf{u}$, $\widehat{\mathbf{y}}$, *and* $\widehat{\boldsymbol{\lambda}}$ *such that*

$$\|\theta_* - \widehat{\theta}_*\| \leq \frac{c}{\alpha\kappa}(\sigma_{n+1} + ... + \sigma_N).$$

Remark 4.3.5 *(i) The error estimates in Theorem 4.3.3 and Corollary 4.3.4 rely on an estimate of the type* (4.32) *of the errors between the input-output operators of the full state and adjoint systems and the reduced state and adjoint systems. Balanced truncation model reduction provides such a bound. Any other model reduction technique for which such a bound is available can be used as well.*

(ii) The assumption (4.28) *is used in two ways. First, it implies that all eigenvalues of the pair* $(\mathbf{A}_{II}^{(1)}, \mathbf{M}_{II}^{(1)})$ *have negative real part and, consequently, it is necessary for the application of balanced truncation model reduction. Secondly, we use it in connection with Lemma A.1.1. We could, for example, use Gronwall type estimates to derive different*

bounds for the solution of a dynamical system in terms of the right hand side of the dynamical system. These bounds can be easily substituted for the bound in Lemma A.1.1. If such estimates are used, assumption (4.28) *could be weakened.*

4.4 Numerical examples

4.4.1 Optimal control of water pollution

This example is motivated by [31], where adaptive finite elements are considered for a steady state version of the optimal control problem described below. See also [4] for a related problem.

The domain Ω is shown in Figure 4.1. The boundary specifications in Figure 4.1 are those for the advection-diffusion equation (4.51).

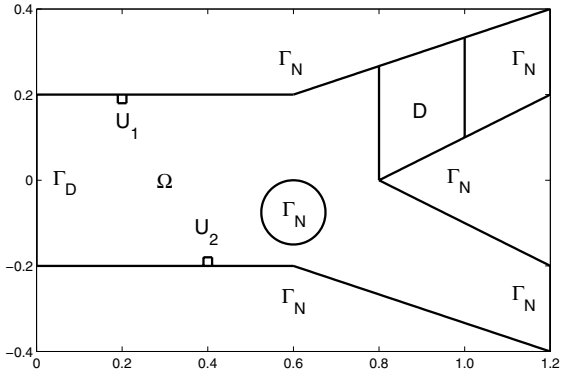


Figure 4.1: The domain Ω with boundary conditions for the advection-diffusion equation (4.51)

The advection $\mathbf{V}$ is the solution of the steady Stokes equation

$$-\mu\Delta\mathbf{V}(x) + \nabla p(x) = \mathbf{0}, \qquad \text{in } \Omega \tag{4.50a}$$

$$\nabla \cdot \mathbf{V}(x) = 0, \qquad \text{in } \Omega \tag{4.50b}$$

$$\mathbf{V}(x) = \mathbf{V}_{in}(x), \qquad \text{on } \Gamma_{\text{in}} \tag{4.50c}$$

$$\mathbf{V}(x) = \mathbf{0}, \qquad \text{on } \Gamma_0 \tag{4.50d}$$

$$(\mu\nabla\mathbf{V}(x)\mathbf{n} - p(x)I)\ \mathbf{n} = \mathbf{0}, \qquad \text{on } \Gamma_{\text{out}}. \tag{4.50e}$$

The problem data are chosen as in [31]. In particular, $\mu = 0.1$ and $\mathbf{V}_{in}(x) = (1 - (x_2/0.2)^2, 0)^T$. Furthermore, the inflow boundary is $\Gamma_{\text{in}} = \left\{(x_1, x_2) \in \overline{\Omega} \ : \ x_1 = 0\right\}$, the outflow boundary is $\Gamma_{\text{out}} = \left\{(x_1, x_2) \in \overline{\Omega} \ : \ x_1 = 1.2\right\}$, and $\Gamma_0 = \partial\Omega \setminus (\Gamma_{\text{in}} \cup \Gamma_{\text{out}})$.

The optimal control problem is governed by the advection-diffusion equation

$$\frac{\partial}{\partial t}y(x,t) - \nabla(k\nabla y(x,t)) + \mathbf{V}(x) \cdot \nabla y(x,t)) = u(x,t)\chi_{U_1}(x) + u(x,t)\chi_{U_2}(x) \qquad \text{in } \Omega, \tag{4.51a}$$

with boundary and initial conditions

$$y(x,t) = 0 \qquad \text{in } \Gamma_D, \tag{4.51b}$$

$$\frac{\partial}{\partial n}y(x,t) = 0 \qquad \text{in } \Gamma_N, \tag{4.51c}$$

$$y(x,0) = 0 \qquad \text{in } \Omega. \tag{4.51d}$$

Here χ_S is the characteristic function corresponding to the set S. Furthermore, $k = 0.015$, V is the solution of (4.50), the boundary segments Γ_D and Γ_N and the control regions U_1 and U_2 are shown in Figure 4.1. In our experiments, the final time is $T = 4$.

The objective function is

$$\frac{1}{2}\int_0^T \int_D (y(x,t) - d(x,t))^2 dx\, dt + \frac{10^{-4}}{2}\int_0^T \int_{U_1 \cup U_2} u^2(x,t) dx\, dt,$$

where D is the observation region shown in Figure 4.1.

For the spatial discretization we use piecewise linear finite elements on three different triangulations with decreasing mesh sizes. We use the modified low-rank Smith method in [56] with $m = 4$ shifts to solve the controllability and observability Lyapunov equations (4.5). For the model reduction, we select those Hankel singular values σ_n, with $\sigma_n \geq 10^{-4}\sigma_1$. Table 4.1 displays the size of the reduced and the full order problems for the three grid sizes. The size of the reduced order model is insensitive to the size of the discretization.

grid number	m	k	N	n
1	168	9	1545	9
2	283	16	2673	9
3	618	29	6036	9

Table 4.1: The number m of observations, the number k of controls, the size N of the full order system, and the size n of the reduced order system for three discretizations.

Figure 4.2 shows the largest Hankel singular values for the fine grid discretization, together with the threshold $10^{-4}\sigma_1$ indicated by the solid line.

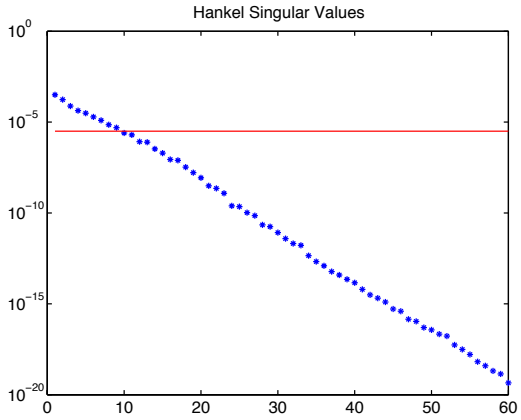


Figure 4.2: The largest Hankel singular values and the threshold $10^{-4}\sigma_1$

For the numerical solution of (4.9), (4.10) we use the Crank-Nicolson method in time with time step size 10^{-2}. The resulting problem is solved using the conjugate gradient method. The conjugate gradient is stopped if the initial residual is reduced by a factor 10^{-4}. Figure 4.3 shows the integrals $\int_{U_1} u^2(x,t)dx$ and $\int_{U_2} u^2(x,t)dx$ of the optimal controls computed using the full and the reduced order model on the fine grid problem. The full and the reduced order model solutions are in excellent agreement as expected by Corollary 4.2.4. For the fine grid problem, the error between full and the reduced order model solutions is $\|u_* - \widehat{u}_*\|^2_{L^2} = 6.2 \cdot 10^{-3}$.

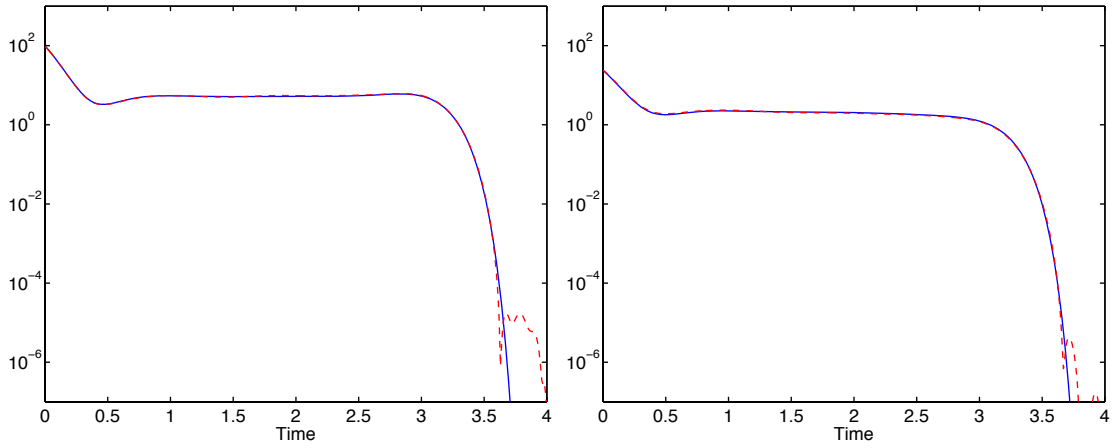


Figure 4.3: The left plot shows the integrals $\int_{U_1} u_*^2(x,t)dx$ and $\int_{U_1} \widehat{u}_*^2(x,t)dx$ of the optimal controls computed using the full (solid line) and the reduced order model (dashed line). The right plot shows the integrals $\int_{U_2} u_*^2(x,t)dx$ and $\int_{U_2} \widehat{u}_*^2(x,t)dx$ of the optimal controls computed using the full (solid line) and the reduced order model (dashed line). The full and reduced order model solutions are in excellent agreement.

The convergence histories of the conjugate gradient algorithm applied to the full order and reduced order optimal control problem are shown in Figure 4.4. The convergence behavior of the conjugate gradient algorithm applied to the full and the reduced order problems is nearly identical. Although there is no rigorous theoretical justification for this

behavior, it is not surprising, given the gradient error bounds derived in Theorem 4.2.3.

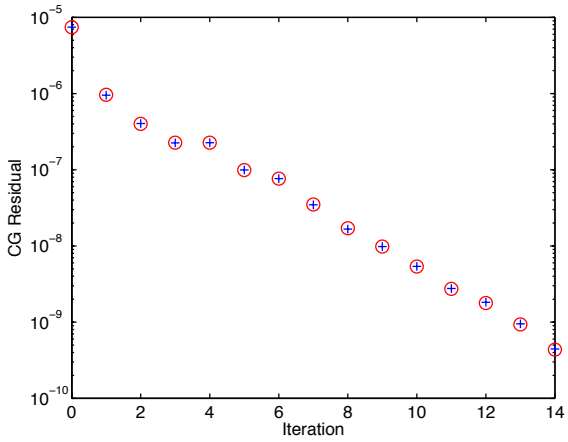


Figure 4.4: The convergence histories of the conjugate gradient algorithm applied to the full (+) and the reduced (o) order optimal control problems

4.4.2 Shape optimization

Our second example is a shape optimization problem governed by the heat equation. The domain Ω is of the type shown in Figure 4.5 with a circular hole Ω_H. It is decomposed into subdomains $\Omega_1 = \Omega_A \cup \Omega_B$ and $\Omega_2 = \Omega_C \setminus \Omega_H$. The boundary $\partial\Omega$ is decomposed into $\Gamma_L, \Gamma_R, \Gamma_T, \Gamma_B$, and $\Gamma_H = \partial\Omega_H$. The interface between Ω_1 and Ω_2 is given by $\Gamma_I = (\overline{\Omega}_A \cap \overline{\Omega}_C) \cup (\overline{\Omega}_B \cap \overline{\Omega}_C)$.

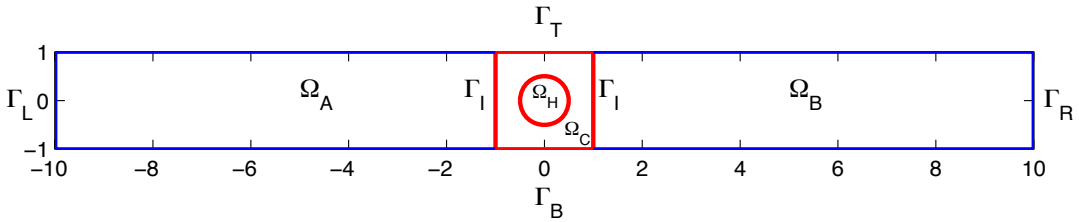


Figure 4.5: Reference domain Ω_{ref}

Assuming a heat source $\mathbf{f}$ in $\Omega_2 \times (0,T)$, no heat flux through $\partial\Omega$ at any time, and zero initial temperature, the objective is to design the shape of the top $\Gamma_{2,T}$ and the bottom $\Gamma_{2,B}$ of $\partial\Omega_2$ in such a way that a prescribed temperature distribution $\mathbf{y}^d$ is achieved in $\Omega_2 \times (0,T)$ and on $(\Gamma_L \cup \Gamma_R) \times (0,T)$. We use a parametrization $\Omega_2(\theta)$ of Ω_2 by means of the Bézier control points $\theta \in \mathbb{R}^k, k = k_T + k_B$, of Bézier curve representations of $\Gamma_{2,T}$ and $\Gamma_{2,B}$, where k_T and k_B refer to the number of control points for $\Gamma_{2,T}$ and $\Gamma_{2,B}$, respectively. The shape optimization problem amounts to the minimization of

$$J(\theta) = \int_0^T \int_{\Gamma_L \cup \Gamma_R} |y - y^d|^2 ds dt + \int_0^T \int_{\Omega_2(\theta)} |y - y^d|^2 dx dt$$

subject to the differential equation

$$\begin{aligned} y_t(x,t) - \Delta y(x,t) + y(x,t) &= f(x,t) && \text{in } \Omega(\theta) \times (0,T), \\ n \cdot \nabla y(x,t) &= 0 && \text{on } \partial\Omega(\theta) \times (0,T), \\ y(x,0) &= 0 && \text{in } \Omega(\theta), \end{aligned}$$

and design parameter constraints

$$\theta^{min} \leq \theta \leq \theta^{max}.$$

We set $f = 100$ in $\Omega_2(\theta) \times (0,T)$ and $f = 0$ else. Furthermore, we specify $T = 4$. The bounds $\theta^{min}, \theta^{max}$ on the design parameters are chosen such that the design constraints are never active in this example. We use $k_T = 3, k_B = 3$ Bézier control points to specify the top and the bottom boundary of the variable subdomain $\Omega_2(\theta)$. The desired temperature y^d is computed by specifying the optimal parameter θ_* (specified in Table 4.3 below) and solving the state equation on $\Omega(\theta_*)$. The optimal domain $\Omega(\theta_*)$ is shown in Figure 4.6.

For the semi-discretization in space we use conforming piecewise linear finite elements with respect to a simplicial triangulation of the computational domain $\Omega(\theta)$ that aligns with

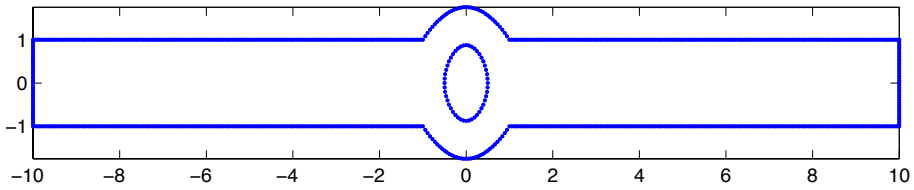


Figure 4.6: Optimal domain

its decomposition into the subdomains Ω_1 and Ω_2. For $D \subseteq \bar{\Omega}$, we denote by $\mathcal{N}_h(D)$ the set of nodal points in D. We use the domain decomposition methodology as described in the previous section and set $N_{dof}^{(\nu)} = \text{card}(\mathcal{N}_h(\bar{\Omega}_\nu \setminus \Gamma_I))$, $\nu = 1, 2$, and $N_{dof}^{\Gamma_I} := \text{card}(\mathcal{N}_h(\Gamma_I))$ so that $N_{dof} = N_{dof}^{(1)} + N_{dof}^{(2)} + N_{dof}^{\Gamma_I}$ is the total number of degrees of freedom.

The matrices $\mathbf{A}, \mathbf{M}$ in the semidiscretized optimization problem (4.1) are given as usual. If ϕ_i are the piecewise linear basis functions associated with the triangulation of $\Omega(\theta)$, then, for example,

$$\mathbf{A}(\theta)_{ij} = \int_{\Omega(\theta)} (\nabla \phi_j^T \nabla \phi_i + \phi_j \phi_i) dx.$$

The matrix $\mathbf{B}(\theta) \in \mathbb{R}^{N_{dof} \times 1}$ corresponds to the right-hand side f and is given by $\mathbf{B}(\theta)_i = \int_{\Omega_2(\theta)} \phi_i dx$ such that with $\mathbf{u} = 100$, $\int_{\Omega(\theta)} f(x,t)\phi_i dx = \mathbf{B}(\theta)\mathbf{u}$ (recall that $f = 100$ in $\Omega_2(\theta) \times (0,T)$ and $f = 0$ else). If the boundary data in the heat equation were nonzero, they would also be incorporated into $\mathbf{B}(\theta)$ by adding another column. For example, $n \cdot \nabla y(x,t) = g_1(x)g_2(t)$ on $\partial\Omega(\theta) \times (0,T)$ would lead to a second column of $\mathbf{B}(\theta)_{i,2} = \int_{\partial\Omega(\theta)} \phi_i g_1(x) dx$.

The observation matrix $\mathbf{C}_I^{(1)}$ in (4.23) is associated with the term $\int_0^T \int_{\Gamma_L \cup \Gamma_R} |y - y^d|^2 ds dt$ in the objective function. If ϕ_i, $i = 1, \ldots, k_1$, are the basis functions associated with the nodes on $\Gamma_L \cup \Gamma_R$, then we compute the entries of $\mathbf{C}_I^{(1)} \in \mathbb{R}^{k_1 \times N_{dof}^{(1)}}$ as $(\mathbf{C}_I^{(1)})_{i,j} = \int_{\Omega_1} \phi_i(x)\phi_j(x) dx$ for $i = 1, \ldots, k_1$, and $j = 1, ..., N_{dof}^{(1)}$.

We use automatic differentiation [54, 100] to compute the derivatives with respect to the

design variables θ. The semi-discretized optimization problems are solved using a projected BFGS method with Armijo line search [71]. The optimization algorithm is terminated when the norm of projected gradient is less than $\epsilon = 10^{-4}$.

As before, we use the modified low-rank Smith method in [56] with $m = 4$ shifts to solve the controllability and observability Lyapunov equations (4.5). Figure 4.7 shows the largest Hankel singular values. For the model reduction, we select those Hankel singular values σ_j, with $\sigma_j \geq 10^{-4}\sigma_1$. The threshold $10^{-4}\sigma_1$ is indicated by the solid line in Figure 4.7.

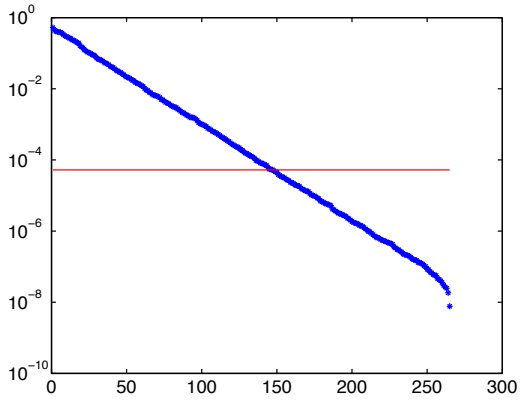


Figure 4.7: The largest Hankel singular values and the threshold $10^{-4}\sigma_1$

Table 4.2 displays the sizes for the full and the reduced order problems.

	$N_{dof}^{(1)}$	N_{dof}
Reduced	147	581
Full	4280	4714

Table 4.2: Sizes of the full and the reduced order problems

The optimal shape parameters θ_* and $\widehat{\theta}_*$ computed by minimizing the full and the reduced

order model, respectively, are shown in Table 4.3. The error $\|\theta^* - \widehat{\theta}^*\|_2 = 2.325 \cdot 10^{-4}$ is proportional to the threshold applied to the truncation of the Hankel singular values, as predicted by Corollary 4.3.4.

θ_*	(1.00, 2.0000, 2.0000, -2.0000, -2.0000, -1.00)
$\widehat{\theta}_*$	(1.00, 1.9999, 2.0001, -2.0001, -1.9998, -1.00)

Table 4.3: Optimal shape parameters θ_* and $\widehat{\theta}_*$ (rounded to 5 digits) computed by minimizing the full and the reduced order model, respectively

The convergence histories of the projected BFGS algorithm applied to the full and the reduced order problems are shown in Figure 4.8. Except for the final iterations, the convergence behavior of the optimization algorithm applied to the full and the reduced order problems is nearly identical. Although there is no rigorous theoretical justification for this behavior, it is not surprising, given the gradient error bounds derived in Theorem 4.3.3.

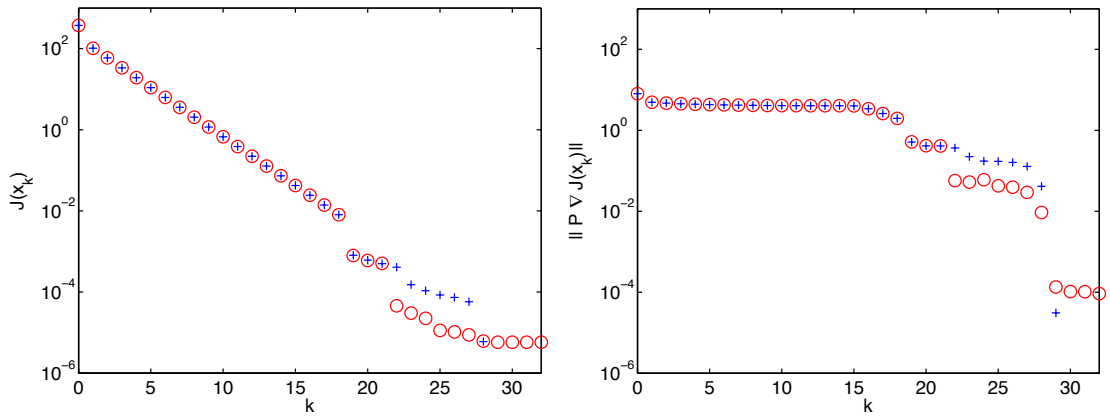


Figure 4.8: The convergence histories of the projected BFGS algorithm applied to the full and the reduced order problems. The left figure shows the convergence history of the objective functionals for the full (+) and reduced (o) order model. The right figure shows the convergence history of the projected gradients for the full (+) and reduced (o) order model.

4.5 Concluding remarks

In this chapter, we have integrated domain decomposition and balanced truncation model reduction for the numerical solution of a class of PDE-constrained optimization problems which are governed by linear time dependent advection-diffusion equations and for which the optimization variables are related to spatially localized quantities. Our approach leads to a reduced optimization problem with the same structure as the original one, but of potentially much smaller dimension. We have derived an estimate for the error between the solution of the original optimization problem and the solution of the reduced problem. The estimate is largely determined by the balanced truncation error estimate.

CHAPTER 5

Domain decomposition and balanced truncation model reduction for shape optimization of the Stokes system

In this chapter we extend the technique introduced in Chapter 4 to study the numerical solution of a class of shape optimization problems governed by the time dependent Stokes or the time dependent linearized Navier-Stokes equations, linearized around a steady state, in which only a small part of the overall domain is modified. The numerical solution of such optimization problems using gradient based optimization methods requires the solution of coupled systems of partial differential equations (PDEs) involving the forward in time governing equation and the backward in time adjoint equation. The solution of this coupled system can be very expensive, both in terms of computing time and memory. This chapter is based on [15].

Our approach to reduce the computational complexity of the numerical optimization is an integration of domain decomposition and model reduction. Domain decomposition in space is used to decouple the small subproblem that corresponds to the subdomain whose shape is modified by the optimization from the fixed subdomain problem. Balanced truncation model reduction is used to replace the subproblem corresponding to the fixed subdomain by a substantially smaller problem. Domain decomposition identifies the proper connectivities between the subproblems, which are used in the balanced truncation. In principle any model reduction technique can be used, but balanced truncation provides an error bound for the quality of the reduced order subsystem. This error bound will be used to derive an error bound for the coupled shape optimization problem.

Chapter 4 was devoted to the numerical solution of optimal control and shape optimization problems associated with linear time dependent advection-diffusion equations by DDBTMR. This chapter extends the approach and analysis presented in Chapter 4 to shape optimization problems governed by the time dependent Stokes or the time dependent linearized Navier-Stokes equations, linearized around a steady state. Although conceptually the approach in this chapter is similar to the one in Chapter 4, the extension to the Stokes system requires several important changes. These are due to the presence of the incompressibility constraints and affect the model reduction, the domain decomposition, the coupling of both, and the analysis.

The chapter is organized as follows: Section 5.1 is devoted to an appropriate setup of the problem. Balanced truncation model reduction (BTMR) for the semi-discretized Stokes system is reviewed in Section 5.2. Section 5.3 introduces the domain decomposition (DD) methodology, including the specification of the optimality systems for the respective subdomain and interface problems. This is followed by the application of BTMR to the

domain decomposed optimality system in Section 5.4. Section 5.5 is concerned with an a priori estimate of the modelling error which, under certain assumptions, is shown to be largely determined by the BTMR error bound. Two numerical examples concerning the application of DDBTMR to the shape optimization problems are presented in Section 5.6. The first example is motivated by the shape optimization of a capillary barrier in a surface acoustic wave-driven microfluidic biochip (c.f., Subsection 5.6.1). The shape optimization of an aorto-coronaric bypass is considered in Subsection 5.6.2. Finally, Section 5.7 contains some concluding remarks as well as an outlook to possible extensions. While problems governed by the Stokes system are used to demonstrate our approach, it can be applied to problems governed by the Oseen equation or linearized Navier-Stokes equations, linearized around a steady state.

5.1 Shape optimization of the time dependent Stokes system

Let $\Omega(\theta) \subset \mathbb{R}^2$ be a bounded domain that depends on design variables $\theta = (\theta_1, \cdots, \theta_d)^T \in \Theta$, where $\Theta \subset \mathbb{R}^d$ is a given convex set, e.g., $\theta_i, 1 \leq i \leq d$, are the Bézier control points of a Bézier curve representation of the boundary and $\Theta := \{\theta_i \in \mathbb{R} \mid \theta_i^{min} \leq \theta_i \leq \theta_i^{max}, 1 \leq i \leq d\}$. We assume that the boundary $\partial\Omega(\theta)$ consists of an inflow boundary $\Gamma_{in}(\theta)$, an outflow boundary $\Gamma_{out}(\theta)$, and a lateral boundary $\Gamma_{lat}(\theta)$ such that $\partial\Omega(\theta) = \overline{\Gamma}_{in}(\theta) \cup \overline{\Gamma}_{out}(\theta) \cup \overline{\Gamma}_{lat}(\theta), \Gamma_{in}(\theta) \cap \Gamma_{out}(\theta) \cap \Gamma_{lat}(\theta) = \emptyset$. We set $Q(\theta) := \Omega(\theta) \times (0,T), \Sigma(\theta) := \partial\Omega(\theta) \times (0,T), \Sigma_{in}(\theta) := \Gamma_{in}(\theta) \times (0,T), \Sigma_{lat}(\theta) := \Gamma_{lat}(\theta) \times (0,T), T > 0$, and consider shape optimization problems associated with the time dependent Stokes system of the form

$$\inf_{\theta \in \Theta} J(\theta) \tag{5.1a}$$

where

$$J(\theta) := \int_0^T \int_{\Omega(\theta)} \ell(\mathbf{v}(\theta), p(\theta), x, t)\, dx\, dt, \tag{5.1b}$$

and where $\mathbf{v}(\theta), p(\theta)$ solve

$$\frac{\partial}{\partial t}\mathbf{v}(x,t) - \nu\, \Delta \mathbf{v}(x,t) + \nabla p(x,t) = \mathbf{f}(x,t)\ , \qquad (x,t) \in Q(\theta), \tag{5.1c}$$

$$\nabla \cdot \mathbf{v}(x,t) = 0\ , \qquad (x,t) \in Q(\theta), \tag{5.1d}$$

$$\mathbf{v}(x,t) = \mathbf{v}^{in}(x,t)\ , \qquad (x,t) \in \Sigma_{in}(\theta), \tag{5.1e}$$

$$\mathbf{v}(x,t) = \mathbf{0}\ , \qquad (x,t) \in \Sigma_{lat}(\theta), \tag{5.1f}$$

$$(\nabla \mathbf{v}(x,t) - p(x,t)I)\mathbf{n} = 0\ , \qquad (x,t) \in \Sigma_{out}(\theta), \tag{5.1g}$$

$$\mathbf{v}(x,0) = \mathbf{v}^{(0)}(x)\ , \qquad x \in \Omega(\theta). \tag{5.1h}$$

Here, $\mathbf{v} = \mathbf{v}(x,t) = (v_1(x,t), v_2(x,t))^T$ and $p = p(x,t)$ stand for the velocity and the pressure, $\mathbf{f} = \mathbf{f}(x,t)$ is a given forcing term, $\mathbf{v}^{in}$ denotes a prescribed normal velocity on $\Sigma_{in}(\theta)$, $\mathbf{v}^{(0)} = \mathbf{v}^{(0)}(x), x \in \Omega(\theta)$, is the velocity distribution at initial time $t = 0$, satisfying $\nabla \cdot \mathbf{v}^{(0)} = 0$, $\nu > 0$ refers to the viscosity of the fluid, and $\mathbf{t}, \mathbf{n}$ are the unit tangential and unit exterior normal vector on $\partial\Omega(\theta)$. Moreover, the integrand $\ell(\cdot)$ in the objective functional J is a given function of the velocity, the pressure, and the independent variables x, t.

For the spatial discretization of the time dependent Stokes system we use one of the many standard methods [51], such as the classical P2-P1 Taylor Hood element, or methods with discontinuous pressure discretizations. We will discuss this in more detail in Section 5.3 and for details of numerical implementation we refer to Appendix B. We assume that the simplicial triangulation $\mathcal{T}_h$ of the spatial domain $\Omega(\theta)$ is geometrically conforming and aligns with $\Gamma_{in}(\theta), \Gamma_{lat}(\theta)$ and $\Gamma_{out}(\theta)$. This leads to the semi-discrete optimization problem

$$\inf_{\theta\in\Theta} J(\theta) \tag{5.2a}$$

where

$$J(\theta) := \int_0^T \boldsymbol{\ell}(\mathbf{v}(\theta), \mathbf{p}(\theta), x, t, \theta)\ dt, \tag{5.2b}$$

and where $\mathbf{v}(\theta), \mathbf{p}(\theta)$ solve

$$\mathbf{E}(\theta)\ \frac{d}{dt}\begin{pmatrix} \mathbf{v}(t) \\ \mathbf{p}(t) \end{pmatrix} + \mathbf{S}(\theta)\begin{pmatrix} \mathbf{v}(t) \\ \mathbf{p}(t) \end{pmatrix} = \begin{pmatrix} \mathbf{g}_1(\theta)(t) \\ \mathbf{g}_2(\theta)(t) \end{pmatrix}\ ,\ t \in (0, T], \tag{5.2c}$$

$$\mathbf{M}(\theta)\mathbf{v}(0) = \mathbf{v}^{(0)}(\theta), \tag{5.2d}$$

$$-\mathbf{B}(\theta)\mathbf{M}^{-1}\mathbf{v}^{(0)}(\theta) + \mathbf{g}_2(\theta)(0) = \mathbf{0}. \tag{5.2e}$$

Here, the integrand $\boldsymbol{\ell}(\cdot)$ in (5.2b) results from the spatial discretization of the inner integral of the objective functional in (5.1b). The block matrix $\mathbf{E}(\theta)$ and the discrete Stokes operator $\mathbf{S}(\theta)$ in (5.2c) are given by

$$\mathbf{E}(\theta) := \begin{pmatrix} \mathbf{M}(\theta) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}\quad ,\quad \mathbf{S}(\theta) := \begin{pmatrix} \mathbf{A}(\theta) & \mathbf{B}^T(\theta) \\ \mathbf{B}(\theta) & \mathbf{0} \end{pmatrix}, \tag{5.3}$$

where $\mathbf{M}(\theta) \in \mathbb{R}^{n\times n}, \mathbf{A}(\theta) \in \mathbb{R}^{n\times n}$ and $\mathbf{B}(\theta) \in \mathbb{R}^{m\times n}$ are the lumped mass matrix, the stiffness matrix, and the matrix representation of the discrete divergence operator. The vector $\mathbf{g}_2(\theta)(t) \in \mathbb{R}^m$ in (5.2c) stems from the semi-discretization of the incompressibility condition due to the boundary condition at the inflow boundary and $\mathbf{v}^{(0)}(\theta)$ is the initial velocity satisfying the discrete incompressibility condition (5.2e). We note that the data of the semi-discrete problem depend on the design variable θ due to the dependence of the spatial domain on θ.

The Oseen equation and the linearized Navier-Stokes equations, linearized around a steady state, also lead to systems of the type (5.2c)–(5.3). The existence and uniqueness

of a solution $(\mathbf{v},\mathbf{p}) \in L^2((0,T);\mathbb{R}^n) \times L^2((0,T);\mathbb{R}^m/(\text{Ker } \mathbf{B}^T))$ of the semi-discretized equations (5.2c),(5.2d) as well as its continuous dependence on the data of the problem is a consequence of the following result which will also play a prominent role with regard to the application of BTMR and the derivation of upper estimates for the modeling error. The following result applies to the semi discretized Stokes system, but also to class of problems governed by the Oseen equations or the linearized Navier-Stokes equations.

Theorem 5.1.1 *Let* $\mathbf{A},\mathbf{M} \in \mathbb{R}^{n\times n}, \mathbf{B} \in \mathbb{R}^{m\times n}, m < n$, *be matrices with the following properties:*

(i) $\mathbf{M}$ *is symmetric positive definite.*

(ii) $\mathbf{A}$ *is positive definite (not necessarily symmetric) on Ker* $\mathbf{B}$*, i.e., there exists a constant* $\alpha > 0$ *such that*

$$\mathbf{v}^T\mathbf{A}\mathbf{v} \geq \alpha\ \mathbf{v}^T\mathbf{M}\mathbf{v} \quad , \quad \mathbf{v} \in Ker\, \mathbf{B}. \tag{5.4}$$

(iii) $\mathbf{B}$ *has full row rank* m*.*

Consider the initial value problem

$$\mathbf{E}\frac{d}{dt}\begin{pmatrix}\mathbf{v}(t)\\ \mathbf{p}(t)\end{pmatrix} + \mathbf{S}\begin{pmatrix}\mathbf{v}(t)\\ \mathbf{p}(t)\end{pmatrix} = \begin{pmatrix}\mathbf{g}_1(t)\\ \mathbf{g}_2(t)\end{pmatrix} \quad , \quad t \in (0,T], \tag{5.5a}$$

$$\mathbf{M}\mathbf{v}(0) = \mathbf{v}^{(0)}, \tag{5.5b}$$

where $\mathbf{E},\mathbf{S}$ *are as in* (5.3) *and* $\mathbf{g}_1 \in C([0,T];\mathbb{R}^n), \mathbf{g}_2, d\mathbf{g}_2/dt \in C([0,T];\mathbb{R}^m)$ *and* $\mathbf{v}^{(0)} \in \mathbb{R}^n$ *satisfies*

$$-\mathbf{B}\mathbf{M}^{-1}\mathbf{v}^{(0)} \ + \ \mathbf{g}_2(0) \ = \ 0. \tag{5.6}$$

Under the assumptions (i),(ii) and (iii), the initial value problem (5.5a),(5.5b) *has a unique solution* $(\mathbf{v},\mathbf{p}) \in C([0,T];\mathbb{R}^n) \times C([0,T];\mathbb{R}^m/(Ker\ \mathbf{B}^T))$*, and there exist constants* $C_1 - C_8 \geq 0$*, depending on* $\mathbf{A},\mathbf{B},\mathbf{M}$ *such that*

$$\|\mathbf{v}\|_{L^2} \leq C_1 \left(\|\mathbf{v}^{(0)}\| + C_2\|\mathbf{g}_2(0)\| \right) + C_3\|\mathbf{g}_1\|_{L^2} + C_4\|\mathbf{g}_2\|_{L^2},$$
$$\|\mathbf{p}\|_{L^2} \leq C_1C_5 \left(\|\mathbf{v}^{(0)}\| + C_2\|\mathbf{g}_2(0)\| \right) + C_6\|\mathbf{g}_1\|_{L^2} + C_7\|\mathbf{g}_2\|_{L^2} + C_8 \left\| \frac{d}{dt}\mathbf{g}_2 \right\|_{L^2},$$

where

$$C_1 = \frac{\sqrt{2}\left\|\overline{\mathbf{M}}^{-1/2}\right\| \left\|\overline{\mathbf{M}}^{1/2}\right\|}{\sqrt{\alpha}}\ , \qquad C_2 = \left\|\mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\right\|\ , \qquad C_3 = \frac{2\left\|\overline{\mathbf{M}}^{-1}\right\|}{\alpha}\ ,$$
$$C_4 = \frac{2\left\|\overline{\mathbf{M}}^{-1}\right\|}{\alpha} \left\|\mathbf{A}\mathbf{M}^{-1}\mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\right\| + \left\|\mathbf{M}^{-1}\mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\right\|\ ,$$
$$C_5 = \left\|(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{B}\mathbf{M}^{-1}\right\| \|\mathbf{A}\|\ ,$$
$$C_6 = \left\|(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{B}\mathbf{M}^{-1}\right\| \left(\frac{2}{\alpha}\left\|\overline{\mathbf{M}}^{-1}\right\| \|\mathbf{A}\| + 1 \right)\ ,$$
$$C_7 = \left\|(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{B}\mathbf{M}^{-1}\right\|\ \left\|\mathbf{A}\mathbf{M}^{-1}\mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\right\| \left(\frac{2}{\alpha}\left\|\overline{\mathbf{M}}^{-1}\right\| \|\mathbf{A}\| + 1 \right)\ ,$$
$$C_8 = \left\|(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\right\|.$$

Proof: We set $\mathbf{\Pi} := \mathbf{I} - \mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{B}\mathbf{M}^{-1}$. Since $\mathbf{\Pi}^2 = \mathbf{\Pi}$, $\mathbf{\Pi}\mathbf{M} = \mathbf{M}\mathbf{\Pi}^T$, $\text{null}(\mathbf{\Pi}) = \text{range}(\mathbf{B}^T)$ and $\text{range}(\mathbf{\Pi}) = \text{null}(\mathbf{B}\mathbf{M}^{-1})$, i.e., $\mathbf{\Pi}$ is an oblique projector.

We split $\mathbf{v}(t) = \mathbf{v}_H(t) + \mathbf{v}_P(t)$ and $\mathbf{v}^{(0)} = \mathbf{\Pi}\mathbf{v}^{(0)} + \mathbf{v}_P^{(0)}$, where

$$\mathbf{v}_H(t) \in \text{Ker}\mathbf{B} \quad , \quad \mathbf{v}_P(t) := \mathbf{M}^{-1}\mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{g}_2(t), \tag{5.7}$$
$$\mathbf{v}_P^{(0)} = \mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{g}_2(0).$$

We note that $\mathbf{v}_P(t)$ and $\mathbf{v}_P^{(0)}$ are particular solutions of the second equation in (5.5a) and

of (5.6), respectively. Then, the initial value problem (5.5a),(5.5b) can be transformed to

$$\mathbf{\Pi M \Pi}^T \frac{d}{dt}\mathbf{v}_H(t) = -\mathbf{\Pi A \Pi}^T \mathbf{v}_H(t) + \mathbf{\Pi}\widetilde{\mathbf{g}}(t) \quad , \quad t \in (0,T], \tag{5.8a}$$

$$\mathbf{\Pi M \Pi}^T \mathbf{v}_H(0) = \mathbf{v}_H^{(0)}, \tag{5.8b}$$

where $\mathbf{v}_H^{(0)} = \mathbf{\Pi}\ \mathbf{v}^{(0)}$ and $\widetilde{\mathbf{g}} \in \mathbb{R}^n$ is given by

$$\widetilde{\mathbf{g}}(t) := \mathbf{g}_1(t) - \mathbf{A}\mathbf{M}^{-1}\mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{g}_2(t). \tag{5.9}$$

Moreover, $\mathbf{p}(t) \in \mathbb{R}^m/(\text{Ker } \mathbf{B}^T)$ can be recovered according to

$$\mathbf{p}(t) = (\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\Big(\mathbf{B}\mathbf{M}^{-1}\Big(-\mathbf{A}\mathbf{v}_H(t) + \widetilde{\mathbf{g}}(t)\Big) - \frac{d}{dt}\mathbf{g}_2(t)\Big). \tag{5.10}$$

In view of (i),(ii), the matrices $\overline{\mathbf{M}} := \mathbf{\Pi M \Pi}^T$ and $\overline{\mathbf{A}} := \mathbf{\Pi A \Pi}^T$ are symmetric positive definite on Ker $\mathbf{B}$ and satisfy

$$-\mathbf{v}^T\overline{\mathbf{A}}\mathbf{v} \leq -\alpha\mathbf{v}^T\overline{\mathbf{M}}\mathbf{v} \quad , \quad \mathbf{v} \in \text{Ker } \mathbf{B}.$$

Then, Lemma A.1.2 in Appendix A implies

$$\|\mathbf{v}_H\|_{L^2} \leq \frac{\sqrt{2}\left\|\overline{\mathbf{M}}^{-1/2}\right\|\left\|\overline{\mathbf{M}}^{1/2}\right\|}{\sqrt{\alpha}}\ \|\mathbf{v}_H^{(0)}\| + \frac{2\left\|\overline{\mathbf{M}}^{-1}\right\|}{\alpha}\ \|\widetilde{\mathbf{g}}\|_{L^2}. \tag{5.11}$$

We conclude due to (5.7),(5.10) and (5.11). □

Remark 5.1.2 *The semi-discretized Stokes system (5.5a) satisfies* (5.4). *We refer to the proof of Theorem 7.1 in [62].*

5.2 Balanced truncation model reduction for Stokes-type systems

Balanced truncation model reduction is a particular model reduction technique that seeks to replace a large-scale system of differential or difference equations by a system of substantially lower dimension that has nearly the same response characteristics, that preserves asymptotic stability and that provides an error bound on the discrepancy between the outputs of the full and reduced order system [18, 23, 34, 53, 86, 124]. Originally, balanced truncation model reduction was developed for systems of ordinary differential equations (ODEs). Recently it has been extended to descriptor systems. An overview of balanced truncation model reduction for descriptor systems can be found in [83]. Balanced truncation model reduction for semi-discretized Stokes and linearized Navier-Stokes systems is studied in [62, 98, 107]. We summarize the basic ideas for a system that is closely related to the optimality system arising in control and shape optimization problems governed by the semi-discretized Stokes or linearized Navier-Stokes equations. Our presentation follows [62]. We consider

$$\mathbf{M}\frac{d}{dt}\mathbf{v}(t) = -\mathbf{A}\mathbf{v}(t) - \mathbf{B}^T\mathbf{p}(t) + \mathbf{K}\mathbf{u}(t), \qquad t \in (0,T), \tag{5.12a}$$

$$0 = -\mathbf{B}\mathbf{v}(t) + \mathbf{L}\mathbf{u}(t), \qquad t \in (0,T) \tag{5.12b}$$

$$\mathbf{z}(t) = \mathbf{C}\mathbf{v}(t) + \mathbf{F}\mathbf{p}(t) + \mathbf{D}\mathbf{u}(t), \qquad t \in (0,T), \tag{5.12c}$$

$$\mathbf{M}\mathbf{v}(0) = \mathbf{v}^{(0)}, \tag{5.12d}$$

$$\mathbf{B}\mathbf{M}^{-1}\mathbf{v}^{(0)} = \mathbf{L}\mathbf{u}(0), \tag{5.12e}$$

and

$$-\mathbf{M}\frac{d}{dt}\boldsymbol{\lambda}(t) = -\mathbf{A}^T\boldsymbol{\lambda}(t) - \mathbf{B}^T\boldsymbol{\kappa}(t) + \mathbf{C}^T\mathbf{w}(t), \qquad t \in (0,T), \tag{5.13a}$$

$$0 = -\mathbf{B}\boldsymbol{\lambda}(t) + \mathbf{F}^T\mathbf{w}(t), \qquad t \in (0,T), \tag{5.13b}$$

$$\mathbf{q}(t) = \mathbf{K}^T\boldsymbol{\lambda}(t) + \mathbf{L}^T\boldsymbol{\kappa}(t) + \mathbf{D}^T\mathbf{w}(t), \qquad t \in (0,T), \tag{5.13c}$$

$$\mathbf{M}\boldsymbol{\lambda}(T) = \boldsymbol{\lambda}^{(T)}, \tag{5.13d}$$

$$\mathbf{B}\mathbf{M}^{-1}\boldsymbol{\lambda}^{(T)} = \mathbf{F}^T\mathbf{w}(T), \tag{5.13e}$$

where $\mathbf{M} \in \mathbb{R}^{n_v \times n_v}$ is a symmetric positive definite matrix, $\mathbf{A} \in \mathbb{R}^{n_v \times n_v}$, $\mathbf{B} \in \mathbb{R}^{n_p \times n_v}$, $n_p < n_v$, is a matrix with rank n_p, $\mathbf{K} \in \mathbb{R}^{n_v \times n_g}$, $\mathbf{L} \in \mathbb{R}^{n_p \times n_g}$, $\mathbf{C} \in \mathbb{R}^{n_z \times n_v}$, $\mathbf{F} \in \mathbb{R}^{n_z \times n_p}$, and $\mathbf{D} \in \mathbb{R}^{n_z \times n_g}$. The terms $\mathbf{D}\mathbf{u}(t)$ and $\mathbf{D}^T\mathbf{w}(t)$ are 'feed through terms' in the output equations. The system (5.13) is the adjoint system corresponding to (5.12). Conditions (5.12e) and (5.13e) ensure the compatibility of the inputs $\mathbf{u}$ and $\mathbf{w}$ with the initial and final values [28].

In addition to the assumptions above, we assume that the generalized eigenvalues of the pair $(\mathbf{A}, \mathbf{M})$ have positive real part. This assumption is needed to apply balanced truncation model reduction.

The numerical method discussed in [62] for computing reduced order models using balanced truncation is applied to the system (5.12,5.13) directly. However, it is derived by eliminating the variables $\mathbf{p}$ and $\boldsymbol{\kappa}$ via projection. This leads to dynamical systems governed by ODEs to which standard balanced truncation can be applied. The application of balanced truncation to the projected system of ODEs is then translated into an approach that applies directly to (5.12,5.13) . Since the transformation of (5.12,5.13) into a a system of ODEs is also important for the later application of balanced truncation in optimization contexts, we summarize the main steps. Details can be found in [62].

As in the proof of Theorem 5.1.1 we choose $\mathbf{v}(t) = \mathbf{v}_{\mathrm{H}}(t) + \mathbf{v}_{\mathrm{P}}(t)$, where

$$\mathbf{v}_{\mathrm{P}}(t) = \mathbf{M}^{-1}\mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{L}\mathbf{u}(t) \tag{5.14}$$

is a particular solution of (5.12b) and $\mathbf{v}_{\mathrm{H}}(t)$ satisfies $\mathbf{0} = \mathbf{B}\mathbf{v}_{\mathrm{H}}(t)$. If we insert $\mathbf{v}(t) = \mathbf{v}_{\mathrm{H}}(t) + \mathbf{v}_{\mathrm{P}}(t)$, (5.14) into (5.12a-c), we obtain

$$\begin{aligned}\mathbf{M}\frac{d}{dt}\mathbf{v}_{\mathrm{H}}(t) = & -\mathbf{A}\mathbf{v}_{\mathrm{H}}(t) - \mathbf{B}^T\mathbf{p}(t) \\ & + \left(\mathbf{K} - \mathbf{A}\mathbf{M}^{-1}\mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{L}\right)\mathbf{u}(t) \\ & - \mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{L}\frac{d}{dt}\mathbf{u}(t), \end{aligned} \tag{5.15a}$$

$$\mathbf{0} = \mathbf{B}\mathbf{v}_{\mathrm{H}}(t), \tag{5.15b}$$

$$\mathbf{z}(t) = \mathbf{C}\mathbf{v}_{\mathrm{H}}(t) + \mathbf{F}\mathbf{p}(t) + \left(\mathbf{D} + \mathbf{C}\mathbf{M}^{-1}\mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{L}\right)\mathbf{u}(t). \tag{5.15c}$$

Equations (5.15a,b) imply that

$$\begin{aligned}\mathbf{p}(t) = & -(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{B}\mathbf{M}^{-1}\mathbf{A}\mathbf{v}_{\mathrm{H}}(t) \\ & + (\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{B}\mathbf{M}^{-1}\left(\mathbf{K} - \mathbf{A}\mathbf{M}^{-1}\mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{L}\right)\mathbf{u}(t) \\ & - (\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{L}\frac{d}{dt}\mathbf{u}(t) \end{aligned} \tag{5.16}$$

and $\mathbf{\Pi}^T\mathbf{v}_{\mathrm{H}}(t) = \mathbf{v}_{\mathrm{H}}(t)$, where

$$\mathbf{\Pi} = \mathbf{I} - \mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{B}\mathbf{M}^{-1}. \tag{5.17}$$

Note that $\mathbf{\Pi}^2 = \mathbf{\Pi}$, $\mathbf{\Pi}\mathbf{M} = \mathbf{M}\mathbf{\Pi}^T$, $\mathrm{null}(\mathbf{\Pi}) = \mathrm{range}(\mathbf{B}^T)$ and $\mathrm{range}(\mathbf{\Pi}) = \mathrm{null}(\mathbf{B}\mathbf{M}^{-1})$, i.e., $\mathbf{\Pi}$ is an oblique projector.

Next, we insert (5.16) into (5.15a,c), use the identity $\mathbf{\Pi}^T\mathbf{v}_{\mathrm{H}}(t) = \mathbf{v}_{\mathrm{H}}(t)$, and multiply the resulting equation (5.15a) by $\mathbf{\Pi}$. Since $\mathbf{\Pi}\mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{L} = \mathbf{0}$ this leads to

$$\mathbf{\Pi}\mathbf{M}\mathbf{\Pi}^T\frac{d}{dt}\mathbf{v}_{\mathrm{H}}(t) = -\mathbf{\Pi}\mathbf{A}\mathbf{\Pi}^T\mathbf{v}_{\mathrm{H}}(t) + \mathbf{\Pi}\widetilde{\mathbf{B}}\mathbf{u}(t), \tag{5.18a}$$

$$\mathbf{z}(t) = \widetilde{\mathbf{C}}\mathbf{\Pi}^T\mathbf{v}_{\mathrm{H}}(t) + \widetilde{\mathbf{D}}\mathbf{u}(t) - \mathbf{F}(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{L}\frac{d}{dt}\mathbf{u}(t), \tag{5.18b}$$

where

$$\widetilde{\mathbf{B}} = \mathbf{K} - \mathbf{A}\mathbf{M}^{-1}\mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{L},$$
$$\widetilde{\mathbf{C}} = \mathbf{C} - \mathbf{F}(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{B}\mathbf{M}^{-1}\mathbf{A},$$
$$\widetilde{\mathbf{D}} = \mathbf{D} + \mathbf{C}\mathbf{M}^{-1}\mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{L} + \mathbf{F}(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{B}\mathbf{M}^{-1}\widetilde{\mathbf{B}}.$$

To obtain the initial condition for $\mathbf{v}_{\mathrm{H}}$ we set $\mathbf{v}^{(0)} = \mathbf{\Pi}\mathbf{v}^{(0)} + (\mathbf{I} - \mathbf{\Pi})\mathbf{v}^{(0)}$ and use (5.12e)

$$\begin{aligned}\mathbf{v}^{(0)} &= \mathbf{\Pi}\mathbf{v}^{(0)} + \mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{B}\mathbf{M}^{-1}\mathbf{v}^{(0)} \\ &= \mathbf{\Pi}\mathbf{v}^{(0)} + \mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{L}\mathbf{u}(0).\end{aligned}$$

Furthermore, we have

$$\begin{aligned}\mathbf{M}\mathbf{v}(0) &= \mathbf{M}\mathbf{v}_{\mathrm{H}}(0) + \mathbf{M}\mathbf{v}_{\mathrm{P}}(0) = \mathbf{M}\mathbf{\Pi}^T\mathbf{v}_{\mathrm{H}}(0) + \mathbf{M}\mathbf{v}_{\mathrm{P}}(0) \\ &= \mathbf{\Pi}\mathbf{M}\mathbf{v}_{\mathrm{H}}(0) + \mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{L}\mathbf{u}(0).\end{aligned}$$

This leads to

$$\mathbf{\Pi}\mathbf{M}\mathbf{v}_{\mathrm{H}}(0) = \mathbf{\Pi}\mathbf{M}\mathbf{\Pi}^T\mathbf{v}_{\mathrm{H}}(0) = \mathbf{\Pi}\mathbf{v}^{(0)} \ (=: \mathbf{v}_H^{(0)}). \tag{5.18c}$$

We can proceed in the same way to transform (5.13). We set $\boldsymbol{\lambda} = \boldsymbol{\lambda}_{\mathrm{H}}(t) + \boldsymbol{\lambda}_{\mathrm{P}}(t)$ where $\boldsymbol{\lambda}_{\mathrm{P}}(t) = \mathbf{M}^{-1}\mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{F}^T\mathbf{w}(t)$. The equations (5.13) can be transformed into

$$\begin{aligned}-\mathbf{M}\frac{d}{dt}\boldsymbol{\lambda}_{\mathrm{H}}(t) = &-\mathbf{A}^T\boldsymbol{\lambda}_{\mathrm{H}}(t) - \mathbf{B}^T\boldsymbol{\kappa}(t) \\ &+ \left(\mathbf{C}^T - \mathbf{A}^T\mathbf{M}^{-1}\mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{F}^T\right)\mathbf{w}(t) \\ &+ \mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{F}^T\frac{d}{dt}\mathbf{w}(t) \qquad (5.19a)\end{aligned}$$
$$\mathbf{0} = \mathbf{B}\boldsymbol{\lambda}_{\mathrm{H}}(t), \qquad (5.19b)$$
$$\mathbf{q}(t) = \mathbf{K}^T\boldsymbol{\lambda}_{\mathrm{H}}(t) + \mathbf{L}^T\boldsymbol{\kappa}(t) + \left(\mathbf{D}^T + \mathbf{K}^T\mathbf{M}^{-1}\mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{F}^T\right)\mathbf{w}(t). \qquad (5.19c)$$

Equations (5.19a,b) imply that

$$\begin{aligned}\boldsymbol{\kappa}(t) = &-(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{B}\mathbf{M}^{-1}\mathbf{A}^T\boldsymbol{\lambda}_{\mathrm{H}}(t)\\ &+(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{B}\mathbf{M}^{-1}\left(\mathbf{C}^T-\mathbf{A}^T\mathbf{M}^{-1}\mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{F}^T\right)\mathbf{w}(t)\\ &+(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{F}^T\frac{d}{dt}\mathbf{w}(t)\end{aligned} \tag{5.20}$$

and $\boldsymbol{\Pi}^T\boldsymbol{\lambda}_{\mathrm{H}}(t)=\boldsymbol{\lambda}_{\mathrm{H}}(t)$, where $\boldsymbol{\Pi}$ is given as before.

Next, we insert (5.20) into (5.19a,c), use the identity $\boldsymbol{\Pi}^T\boldsymbol{\lambda}_{\mathrm{H}}(t)=\boldsymbol{\lambda}_{\mathrm{H}}(t)$, and multiply the resulting equation (5.19a) by $\boldsymbol{\Pi}$. Since $\boldsymbol{\Pi}\mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{F}^T=\mathbf{0}$ this leads to

$$-\boldsymbol{\Pi}\mathbf{M}\boldsymbol{\Pi}^T\frac{d}{dt}\boldsymbol{\lambda}_{\mathrm{H}}(t) = -\boldsymbol{\Pi}\mathbf{A}^T\boldsymbol{\Pi}^T\boldsymbol{\lambda}_{\mathrm{H}}(t)+\boldsymbol{\Pi}\widetilde{\mathbf{C}}^T\mathbf{w}(t), \tag{5.21a}$$

$$\mathbf{q}(t) = \widetilde{\mathbf{B}}^T\boldsymbol{\Pi}^T\boldsymbol{\lambda}_{\mathrm{H}}(t)+\widetilde{\mathbf{D}}^T\mathbf{w}(t)+\mathbf{L}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{F}^T\frac{d}{dt}\mathbf{w}(t), \tag{5.21b}$$

$$\mathbf{M}\boldsymbol{\lambda}_{\mathrm{H}}(T) = \boldsymbol{\Pi}\boldsymbol{\lambda}^{(T)}, \tag{5.21c}$$

where $\widetilde{\mathbf{B}}$, $\widetilde{\mathbf{C}}$ and $\widetilde{\mathbf{D}}$ are given as before.

For model reduction purposes we view $\mathbf{u}$ and $\frac{d}{dt}\mathbf{u}$ as inputs into (5.18) and $\mathbf{w}$ and $\frac{d}{dt}\mathbf{w}$ as inputs into (5.21). The terms involving $\mathbf{u}$ and $\mathbf{w}$ in (5.18b) and (5.21b) are 'feed through' terms, since inputs are directly fed to the outputs $\mathbf{z}$ and $\mathbf{q}$ respectively. These terms are not reduced. Note that the transformed system (5.21) is the adjoint system corresponding to (5.18).

The systems (5.18) and (5.21) are almost in the form to which standard balanced truncation model reduction can be applied. Since $\boldsymbol{\Pi}$ has a non-trivial null-space, the dynamical systems in (5.18) and (5.21) have to be solved for $\mathbf{v}_{\mathrm{H}}$ with $\boldsymbol{\Pi}^T\mathbf{v}_{\mathrm{H}}=\mathbf{v}_{\mathrm{H}}$ and $\boldsymbol{\lambda}_{\mathrm{H}}$ with $\boldsymbol{\Pi}^T\boldsymbol{\lambda}_{\mathrm{H}}=\boldsymbol{\lambda}_{\mathrm{H}}$. This can be made explicit by expressing

$$\boldsymbol{\Pi} = \boldsymbol{\Theta}_l\boldsymbol{\Theta}_r^T \tag{5.22a}$$

with $\boldsymbol{\Theta}_l, \boldsymbol{\Theta}_r \in \mathbb{R}^{n_v \times (n_v - n_p)}$ satisfying

$$\boldsymbol{\Theta}_l^T \boldsymbol{\Theta}_r = \mathbf{I}. \tag{5.22b}$$

Substituting this decomposition into (5.18) shows that $\widetilde{\mathbf{v}}_{\mathrm{H}} = \boldsymbol{\Theta}_l^T \mathbf{v}_{\mathrm{H}} \in \mathbb{R}^{n_v - n_p}$ must satisfy

$$\boldsymbol{\Theta}_r^T \mathbf{M} \boldsymbol{\Theta}_r \frac{d}{dt} \widetilde{\mathbf{v}}_{\mathrm{H}}(t) = -\boldsymbol{\Theta}_r^T \mathbf{A} \boldsymbol{\Theta}_r \widetilde{\mathbf{v}}_{\mathrm{H}}(t) + \boldsymbol{\Theta}_r^T \widetilde{\mathbf{B}} \mathbf{u}(t), \tag{5.23a}$$

$$\mathbf{z}(t) = \widetilde{\mathbf{C}} \boldsymbol{\Theta}_r \frac{d}{dt} \widetilde{\mathbf{v}}_{\mathrm{H}}(t) + \widetilde{\mathbf{D}} \mathbf{u}(t) - \mathbf{F} (\mathbf{B} \mathbf{M}^{-1} \mathbf{B}^T)^{-1} \mathbf{L} \frac{d}{dt} \mathbf{u}(t). \tag{5.23b}$$

An analogous substitution is applied in (5.21). Standard balanced truncation model reduction can now be applied to the system (5.23) and the corresponding adjoint system derived from (5.21). The projection matrices computed by balanced truncation for (5.23) and the corresponding adjoint system derived from (5.21) can then be transformed into projection matrices for the systems (5.18) and (5.21).

Balanced truncation model reduction generates projection matrices $\mathbf{V}, \mathbf{W} \in \mathbb{R}^{n_v \times k}$ with $k \ll n_v$ such that

$$\mathbf{V} = \boldsymbol{\Pi}^T \mathbf{V}, \mathbf{W} = \boldsymbol{\Pi}^T \mathbf{W}, \text{ and } \mathbf{W}^T \mathbf{M} \mathbf{V} = I.$$

The reduced order model for (5.18) is obtained by replacing $\mathbf{v}_{\mathrm{H}}(t)$ in (5.18) by $\mathbf{V} \widehat{\mathbf{v}}(t)$ and multiplying the resulting equation by $\mathbf{W}^T$. This gives

$$\frac{d}{dt} \widehat{\mathbf{v}}(t) = -\mathbf{W}^T \mathbf{A} \mathbf{V} \widehat{\mathbf{v}}(t) + \mathbf{W}^T \widetilde{\mathbf{B}} \mathbf{u}(t), \tag{5.24a}$$

$$\widehat{\mathbf{z}}(t) = \widetilde{\mathbf{C}} \mathbf{V} \widehat{\mathbf{v}}(t) + \widetilde{\mathbf{D}} \mathbf{u}(t) - \mathbf{F} (\mathbf{B} \mathbf{M}^{-1} \mathbf{B}^T)^{-1} \mathbf{L} \frac{d}{dt} \mathbf{u}(t), \tag{5.24b}$$

$$\widehat{\mathbf{v}}(t) = \mathbf{W}^T \boldsymbol{\Pi} \mathbf{v}^{(0)}. \tag{5.24c}$$

Similarly, the reduced order model for (5.21) is obtained by replacing $\boldsymbol{\lambda}_{\mathrm{H}}(t)$ in (5.21)

by $\mathbf{W}\widehat{\boldsymbol{\lambda}}(t)$ and multiplying the resulting equation by $\mathbf{V}^T$. This gives

$$-\frac{d}{dt}\widehat{\boldsymbol{\lambda}}(t) = -\mathbf{V}^T\mathbf{A}^T\mathbf{W}\widehat{\boldsymbol{\lambda}}(t) + \mathbf{V}^T\widetilde{\mathbf{C}}^T\mathbf{w}(t), \tag{5.25a}$$

$$\widehat{\mathbf{q}}(t) = \widetilde{\mathbf{B}}^T\mathbf{W}\widehat{\boldsymbol{\lambda}}(t) + \widetilde{\mathbf{D}}^T\mathbf{w}(t) + \mathbf{L}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{F}^T\frac{d}{dt}\mathbf{w}(t), \tag{5.25b}$$

$$\widehat{\boldsymbol{\lambda}}(t) = \mathbf{V}^T\boldsymbol{\Pi}\boldsymbol{\lambda}^{(T)}. \tag{5.25c}$$

We can show that $\mathbf{W}^T\mathbf{A}\mathbf{V}$ is stable see [62, Sec. 7] for details. Furthermore if $\mathbf{v}_{\mathrm{H}}(0) = \boldsymbol{\lambda}_{\mathrm{H}}(T) = \mathbf{0}$, and

$$\sigma_1 \geq \ldots \geq \sigma_k > \sigma_{k+1} \geq \ldots \geq \sigma_n \ , \tag{5.26}$$

then for any given inputs $\mathbf{u}$, $\mathbf{w}$ we have

$$\|\mathbf{z} - \widehat{\mathbf{z}}\|_{L^2} \leq 2\|\mathbf{u}\|_{L^2}(\sigma_{k+1} + \ldots + \sigma_n), \tag{5.27a}$$

$$\|\mathbf{q} - \widehat{\mathbf{q}}\|_{L^2} \leq 2\|\mathbf{w}\|_{L^2}(\sigma_{k+1} + \ldots + \sigma_n). \tag{5.27b}$$

Remark 5.2.1 *Inhomogeneous initial conditions can be handled by modifying the balanced truncation model reduction as discussed in [19].*

5.3 Domain decomposition

We consider a decomposition of $\Omega(\theta)$ into subdomains $\Omega_1, \Omega_2(\theta)$ such that

$$\overline{\Omega}(\theta) = \overline{\Omega}_1 \cup \overline{\Omega}_2(\theta) \quad , \quad \Omega_1 \cap \Omega_2(\theta) = \emptyset \quad , \quad \Gamma := \overline{\Omega}_1 \cap \overline{\Omega}_2(\theta), \tag{5.28}$$

where Γ stands for the interfaces between the subdomains. The domain decomposition is motivated by such PDE-constrained optimization problems where the optimal design issues focus on a relatively small portion of the domain, namely the subdomain $\Omega_2(\theta)$. Consequently, only that subdomain is supposed to depend on the design variables θ, whereas Ω_1

is independent of θ. In practice, the subdomains Ω_1 and $\Omega_2(\theta)$ can be further subdivided. Multiple subdomains can be incorporated into our approach, but to keep the presentation simple we consider the two subdomain case.

We assume that the objective functional can be split accordingly

$$J(\theta) := J_1(\mathbf{v}, p) + J_2(\mathbf{v}(\theta), p(\theta), \theta). \tag{5.29}$$

Here, $J_1(\mathbf{v}, p)$ is given in terms of observation operators $C : \mathbf{L}^2((0,T);\mathbf{V}) \to L^2((0,T);(L^2(\Omega_1)^q))$, $F : L^2((0,T);L^2_0(\Omega)) \to L^2((0,T);(L^2(\Omega_1))^q)$ and a feedthrough operator $D : \mathbf{L}^2((0,T);\mathbf{L}^2(\Omega)) \to L^2((0,T);(L^2(\Omega_1))^q)$, $q \in \mathbb{N}$. For a given function $\mathbf{d} \in L^2((0,T);(L^2(\Omega_1))^q)$, we define

$$J_1(\mathbf{v}, p) := \int_0^T \int_{\Omega_1} |C\mathbf{v} + Fp + D\mathbf{u} - \mathbf{d}|^2 \, dx \, dt. \tag{5.30}$$

On the other hand, $J_2(\mathbf{v}, p, \theta)$ is supposed to be as in (5.1b) with $\Omega(\theta)$ replaced by $\Omega_2(\theta)$.

We consider geometrically conforming simplicial triangulations $\mathcal{T}_h(\Omega(\theta))$ that align with the decomposition in the sense that their restrictions to $\Omega_1, \Omega_2(\theta)$ represent geometrically conforming triangulations $\mathcal{T}_h(\Omega_1), \mathcal{T}_h(\Omega_2(\theta))$. The semi-discretization in space of the Stokes equation in the domain decomposition context requires some care. See, e.g., [3, 25, 72, 90, 96, 97, 110, 112]. For semi-discretization in space, we may use stable discontinuous pressure elements such as nonconforming P2-P0 or P1-P0 elements [30] or spectral elements [90, 112]. The subsequent analysis also applies, if we use continuous pressure elements such as the Taylor-Hood P2-P1 element or the mini-element [24, 26], provided the incompressibility condition on the interface $\Gamma(\theta)$ is not discretized and hence, we do not explicitly consider the semi-discrete pressure on the interface $\Gamma(\theta)$ (cf., e.g., [97]).

The discretization needs to be such that the coupled problem is solvable, i.e., the local subproblems corresponding to the subdomains Ω_1 and $\Omega_2(\theta)$ as well as those corresponding

to the interface are solvable. The global problem (5.2c)–(5.2e) has a unique solution $(\mathbf{v}, \mathbf{p}) \in L^2((0,T);\mathbb{R}^n) \times L^2((0,T);\mathbb{R}^m/(\text{Ker } \mathbf{B}^T))$. Some of the local problems associated with the subdomain Ω_1 or $\Omega_2(\theta)$ correspond to Stokes subdomain problems with Dirichlet boundary conditions only. Consequently, for these subproblems the pressure is only unique up to a constant. To ensure that the subdomain solution is the restriction of the solution of (5.2c)–(5.2e) to the subdomain, we split the subdomain pressures into a constant and a subdomain pressure with zero spatial average. The latter is determined uniquely as the solution of the subdomain problem, whereas the constant is determined through the coupled problems. This split is not necessary for subdomains with an outflow condition, where the local pressure is unique. However, to simplify the presentation, we assume that the split has to be made for both subdomains.

The velocities are discretized using

$$v_h(x,t) = \sum_{j=1}^{n} \mathbf{v}_j(t)\phi_j(x),$$

where $\phi_j(t)$, $j = 1,\ldots,n_1$ have support in $\overline{\Omega}_1$, $\phi_j(t)$, $j = n_1+1,\ldots,n_1+n_2$ have support in $\overline{\Omega}_2$, and $\phi_j(t)$, $j = n_1+n_2+1,\ldots,n = n_1+n_2+n_\Gamma$ are the remaining basis functions, which are associated with the interface. The semi-discretized pressure $p_h(x,t)$ is the sum of subdomain pressures $p_{h,i}(x,t)$, $i = 1,2$ with zero average on the subdomain, $\int_{\Omega_j} p_{h,i}(x,t)dx = 0$, $i = 1,2$, and constant pressures $p_{0,i}(t)$, $i = 1,2$, for each subdomain. We have

$$p_h(x,t) = \sum_{j=1}^{2} p_{0,j}(t)\chi_{\overline{\Omega}_j}(x) + \sum_{j=1}^{m-2} p_j(t)\psi_j(x),$$

where $\chi_{\mathcal{S}}$ denotes the characteristic function of a set $\mathcal{S} \subset \Omega$, $\psi_j(t)$, $j = 1,\ldots,m_1$ are basis functions that have support in $\overline{\Omega}_1$, and $\psi_j(t)$, $j = m_1+1,\ldots,m-2 = m_1+m_2$ have

support in $\overline{\Omega}_2$. We require that

$$\int_{\Omega_j} \sum_{j=1}^{m_1} p_j(t)\psi_j(x)dx = \int_{\Omega_j} \sum_{j=1}^{m_2} p_{m_1+j}(t)\psi_{m_1+j}(x)dx = 0.$$

Thus, we have velocities $\mathbf{v}_1(t) \in \mathbb{R}^{n_1}$, $\mathbf{v}_2(t) \in \mathbb{R}^{n_2}$, $\mathbf{v}_\Gamma(t) \in \mathbb{R}^{n_\Gamma}$ associated with Ω_1, $\Omega_2(\theta)$, and $\Gamma(\theta)$, respectively. We set $\mathbf{v}(t) = (\mathbf{v}_1(t), \mathbf{v}_2(t), \mathbf{v}_\Gamma(t))^T$. The pressures associated with Ω_1, $\Omega_2(\theta)$ are $\mathbf{p}_1(t) \in \mathbb{R}^{m_1}$, $\mathbf{p}_2(t) \in \mathbb{R}^{m_2}$. Additionally, we have constants $p_{0,1}(t), p_{0,2}(t) \in \mathbb{R}$. We set $\mathbf{p}_0(t) = (p_{0,1}(t), p_{0,2}(t))^T$ and $\mathbf{p}(t) = (\mathbf{p}_1(t), \mathbf{p}_2(t), \mathbf{p}_0(t))^T$. Finally, we define the state variables

$$\mathbf{x}(t) := (\mathbf{v}_1, \mathbf{p}_1, \mathbf{v}_2, \mathbf{p}_2, \mathbf{v}_\Gamma, \mathbf{p}_0)^T \quad , \quad t \in [0, T]. \tag{5.31}$$

The numerical implementation of domain decomposition for the Stokes equations or linearized Navier-Stokes system is discussed in Appendix B.

With this discretization and partitioning of variables, the matrices $\mathbf{A}(\theta)$ and $\mathbf{B}(\theta)$ can be partitioned as follows

$$\mathbf{A}(\theta) = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{0} & \mathbf{A}_{1\Gamma} \\ \mathbf{0} & \mathbf{A}_{22}(\theta) & \mathbf{A}_{2\Gamma}(\theta) \\ \mathbf{A}_{1\Gamma}^T & \mathbf{A}_{2\Gamma}^T(\theta) & \mathbf{A}_{\Gamma\Gamma}(\theta) \end{pmatrix} \quad , \quad \mathbf{B}(\theta) = \begin{pmatrix} \mathbf{B}_{11} & \mathbf{0} & \mathbf{B}_{1\Gamma} \\ \mathbf{0} & \mathbf{B}_{22}(\theta) & \mathbf{B}_{2\Gamma}(\theta) \\ 0 & 0 & \mathbf{B}_0(\theta) \end{pmatrix}. \tag{5.32}$$

Here, $\mathbf{A}_{11} \in \mathbb{R}^{n_1 \times n_1}$, $\mathbf{A}_{22}(\theta) \in \mathbb{R}^{n_2 \times n_2}$, $\mathbf{A}_{\Gamma\Gamma}(\theta) \in \mathbb{R}^{n_\Gamma \times n_\Gamma}$, $\mathbf{A}_{i\Gamma}(\theta) \in \mathbb{R}^{n_i \times n_\Gamma}$, $1 \le i \le 2$, and $\mathbf{B}_{11} \in \mathbb{R}^{m_1 \times n_1}$, $\mathbf{B}_{22}(\theta) \in \mathbb{R}^{m_2 \times n_2}$, $\mathbf{B}_{i\Gamma}(\theta) \in \mathbb{R}^{m_i \times n_\Gamma}$, $1 \le i \le 2$, $\mathbf{B}_0(\theta) \in \mathbb{R}^{2 \times n_\Gamma}$. Likewise, the matrices $\mathbf{K}(\theta), \mathbf{L}(\theta)$ and the lumped mass matrix $\mathbf{M}(\theta)$ admit the decompositions

$$\mathbf{K}(\theta) = (\mathbf{K}_1, \mathbf{K}_2(\theta), \mathbf{K}_\Gamma(\theta))^T \qquad \mathbf{L}(\theta) = (\mathbf{L}_1, \mathbf{L}_2(\theta), \mathbf{L}_0(\theta))^T, \tag{5.33a}$$

$$\mathbf{M}(\theta) = \mathbf{blockdiag}(\mathbf{M}_1, \mathbf{M}_2(\theta), \mathbf{M}_\Gamma(\theta)), \tag{5.33b}$$

where $\mathbf{K}_i(\theta) \in \mathbb{R}^{n_i \times k}, \mathbf{L}_i(\theta) \in \mathbb{R}^{m_i \times k}$, $1 \le i \le 2$, $\mathbf{K}_\Gamma(\theta) \in \mathbb{R}^{n_\Gamma \times k}$, $\mathbf{L}_\Gamma(\theta) \in \mathbb{R}^{2 \times k}$ and

$\mathbf{M}_1 \in \mathbb{R}^{n_1 \times n_1}, \mathbf{M}_2(\theta) \in \mathbb{R}^{n_2 \times n_2}$, $\mathbf{M}_\Gamma(\theta) \in \mathbb{R}^{n_\Gamma \times n_\Gamma}$. We set

$$\mathbf{E}(\theta) = \begin{pmatrix} \mathbf{E}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_2(\theta) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{E}_\Gamma(\theta) \end{pmatrix} \quad , \quad \mathbf{S}(\theta) = \begin{pmatrix} \mathbf{S}_1 & \mathbf{0} & \mathbf{S}_{1\Gamma} \\ \mathbf{0} & \mathbf{S}_2(\theta) & \mathbf{S}_{2\Gamma}(\theta) \\ \mathbf{S}_{1\Gamma}^T & \mathbf{S}_{2,\Gamma}^T(\theta) & \mathbf{S}_\Gamma(\theta) \end{pmatrix}, \tag{5.34}$$

where

$$\mathbf{E}_1 = \begin{pmatrix} \mathbf{M}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \ , \ \mathbf{E}_2(\theta) = \begin{pmatrix} \mathbf{M}_2(\theta) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \ , \ \mathbf{E}_\Gamma(\theta) = \begin{pmatrix} \mathbf{M}_\Gamma(\theta) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \tag{5.35a}$$

$$\mathbf{S}_1 = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{B}_{11}^T \\ \mathbf{B}_{11} & \mathbf{0} \end{pmatrix} \ , \ \mathbf{S}_2(\theta) = \begin{pmatrix} \mathbf{A}_{22}(\theta) & \mathbf{B}_{22}^T(\theta) \\ \mathbf{B}_{22}(\theta) & \mathbf{0} \end{pmatrix}, \tag{5.35b}$$

$$\mathbf{S}_\Gamma(\theta) = \begin{pmatrix} \mathbf{A}_{\Gamma\Gamma}(\theta) & \mathbf{B}_0^T(\theta) \\ \mathbf{B}_0(\theta) & \mathbf{0} \end{pmatrix} \ , \ \mathbf{S}_{i\Gamma}(\theta) = \begin{pmatrix} \mathbf{A}_{i\Gamma}(\theta) & \mathbf{0} \\ \mathbf{B}_{i\Gamma}(\theta) & \mathbf{0} \end{pmatrix} \ , \quad 1 \le i \le 2, \tag{5.35c}$$

and

$$\mathbf{N}(\theta) = (\mathbf{K}_1 \mid \mathbf{L}_1 \mid \mathbf{K}_2(\theta) \mid \mathbf{L}_2(\theta) \mid \mathbf{K}_\Gamma(\theta) \mid \mathbf{L}_0(\theta))^T. \tag{5.36}$$

We further denote by $\mathbf{C}_1 \in \mathbb{R}^{q \times n_1}, \mathbf{F}_1 \in \mathbb{R}^{q \times m_1}, \mathbf{D}_1 \in \mathbb{R}^{q \times n_1}, \mathbf{d}(t) \in \mathbb{R}^q, t \in (0,T)$, the matrices and the vector and by $\boldsymbol{\ell}(\mathbf{v}_2, \mathbf{v}_\Gamma, \mathbf{p}_2, \mathbf{p}_0, t, \theta)$ the functional resulting from the semi-discretization of the inner integrals in J_2. We set

$$J(\theta) := J_1(\mathbf{v}_1, \mathbf{p}_1, \mathbf{p}_0) + J_2(\mathbf{v}_2(\theta), \mathbf{v}_\Gamma(\theta), \mathbf{p}_2(\theta), \mathbf{p}_0(\theta), \theta) \tag{5.37}$$

where J_1 and J_2 are given by

$$J_1(\mathbf{v}_1, \mathbf{p}_1, \mathbf{p}_0) = \frac{1}{2} \int_0^T |\mathbf{C}_1\mathbf{v}_1(t) + \mathbf{F}_1\mathbf{p}_1(t) + \mathbf{F}_0\mathbf{p}_0(t) + \mathbf{D}_1\mathbf{u}(t) - \mathbf{d}(t)|^2 \, dt, \tag{5.38a}$$

$$J_2(\mathbf{v}_2, \mathbf{v}_\Gamma, \mathbf{p}_2, \mathbf{p}_0, \theta) = \int_0^T \boldsymbol{\ell}(\mathbf{v}_2, \mathbf{v}_\Gamma, \mathbf{p}_2, \mathbf{p}_0, t, \theta) \, dt. \tag{5.38b}$$

The semi-discretized, domain decomposed shape optimization problem can be formulated according to

$$\inf_{\theta\in\Theta} J(\theta) \tag{5.39a}$$

where $\mathbf{x} = (\mathbf{v}_1, \mathbf{p}_1, \mathbf{v}_2(\theta), \mathbf{p}_2(\theta), \mathbf{v}_\Gamma(\theta), \mathbf{p}_0(\theta))^T$ solves

$$\mathbf{P}(\theta)\mathbf{x}(t) := \mathbf{E}(\theta)\ \frac{d}{dt}\mathbf{x}(t) + \mathbf{S}(\theta)\mathbf{x}(t) = \mathbf{N}(\theta)\mathbf{u}(t)\ ,\ t \in (0,T], \tag{5.39b}$$

$$\mathbf{M}(\theta)\mathbf{v}(0) = \mathbf{v}^{(0)}(\theta). \tag{5.39c}$$

Remark 5.3.1 *If the Stokes equations are replaced by the Oseen equations or the linearized Navier-Stokes equations, linearized around a steady state, we also arrive at a semi-discretized, domain decomposed shape optimization problem that is essentially of the type* (5.39)*. In this case, the matrix* $\mathbf{A}(\theta)$ *and consequently, the matrix* $\mathbf{S}(\theta)$*, are no longer symmetric. However, this nonsymmetry can be easily incorporated and the discussion in the current and the following sections can be easily extended to classes of problems governed by the Oseen equation and the linearized Navier-Stokes equations.*

Introducing Lagrange multipliers $\boldsymbol{\lambda}(t) \in \mathbb{R}^n, \boldsymbol{\kappa}(t) \in \mathbb{R}^m, t \in [0,T]$, that are partitioned accordingly, and setting

$$\boldsymbol{\mu}(t) = (\boldsymbol{\lambda}_1(t), \boldsymbol{\kappa}_1(t), \boldsymbol{\lambda}_2(t), \boldsymbol{\kappa}_2(t), \boldsymbol{\lambda}_\Gamma(t), \boldsymbol{\kappa}_0(t))^T,$$

the Lagrangian associated with (5.39a)-(5.39c) is given by

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\mu}, \theta) := J(\mathbf{v}, \mathbf{p}, \theta) + \int_0^T \boldsymbol{\mu}(t)^T(\mathbf{P}(\theta)\mathbf{x}(t) - \mathbf{N}(\theta)\mathbf{u}(t))\ dt, \tag{5.40}$$

and the optimality conditions read

$$\nabla_{\mathbf{x}}\mathcal{L}(\mathbf{x}, \boldsymbol{\mu}, \theta) = 0\ ,\ \nabla_{\boldsymbol{\mu}}\mathcal{L}(\mathbf{x}, \boldsymbol{\mu}, \theta) = 0\ ,\ \nabla_\theta\mathcal{L}(\mathbf{x}, \boldsymbol{\mu}, \theta)^T(\tilde{\theta} - \theta) \geq 0\ ,\ \tilde{\theta} \in \Theta. \tag{5.41}$$

It is obvious that due to the special structure of the decomposed optimization problems, the optimality conditions (5.41) can be split into a coupled system of optimality conditions associated with the subdomains Ω_1, $\Omega_2(\theta)$, and the interface $\Gamma(\theta)$.

(i) Optimality system associated with subdomain Ω_1:

$$\mathbf{E}_1 \frac{d}{dt} \begin{pmatrix} \mathbf{v}_1(t) \\ \mathbf{p}_1(t) \end{pmatrix} = -\mathbf{S}_1 \begin{pmatrix} \mathbf{v}_1(t) \\ \mathbf{p}_1(t) \end{pmatrix} - \mathbf{S}_{1\Gamma} \begin{pmatrix} \mathbf{v}_\Gamma(t) \\ \mathbf{p}_0(t) \end{pmatrix} + \begin{pmatrix} \mathbf{K}_1 \\ \mathbf{L}_1 \end{pmatrix} \mathbf{u}(t), \tag{5.42a}$$

$$\mathbf{z}_1(t) = \mathbf{C}_1\mathbf{v}_1(t) + \mathbf{F}_1\mathbf{p}_1(t) + \mathbf{F}_0\mathbf{p}_0(t) + \mathbf{D}_1\mathbf{u}(t) - \mathbf{d}(t), \tag{5.42b}$$

$$\mathbf{M}_1\mathbf{v}_1(0) = \mathbf{v}_1^{(0)}, \tag{5.42c}$$

$$\mathbf{L}_1\mathbf{u}(0) = \mathbf{B}_{11}\mathbf{M}_1^{-1}\mathbf{v}_1^{(0)} + \mathbf{B}_{1\Gamma}\mathbf{M}_\Gamma^{-1}\mathbf{v}_\Gamma^{(0)}(\theta), \tag{5.42d}$$

and

$$-\mathbf{E}_1 \frac{d}{dt} \begin{pmatrix} \boldsymbol{\lambda}_1(t) \\ \boldsymbol{\kappa}_1(t) \end{pmatrix} = -\mathbf{S}_1 \begin{pmatrix} \boldsymbol{\lambda}_1(t) \\ \boldsymbol{\kappa}_1(t) \end{pmatrix} - \mathbf{S}_{1\Gamma} \begin{pmatrix} \boldsymbol{\lambda}_\Gamma(t) \\ \boldsymbol{\kappa}_0(t) \end{pmatrix} - \begin{pmatrix} \mathbf{C}_1^T \\ \mathbf{F}_1^T \end{pmatrix} \mathbf{z}_1(t), \tag{5.43a}$$

$$\mathbf{M}_1\boldsymbol{\lambda}_1(T) = \boldsymbol{\lambda}_1^{(T)}, \tag{5.43b}$$

$$\mathbf{F}_1^T\mathbf{z}_1(T) = -\mathbf{B}_{11}\mathbf{M}_1^{-1}\boldsymbol{\lambda}_1^{(T)} - \mathbf{B}_{1\Gamma}\mathbf{M}_\Gamma(\theta)^{-1}\boldsymbol{\lambda}_\Gamma^{(T)}(\theta). \tag{5.43c}$$

(ii) Optimality system associated with subdomain $\Omega_2(\theta)$:

$$\mathbf{E}_2(\theta) \frac{d}{dt} \begin{pmatrix} \mathbf{v}_2(t) \\ \mathbf{p}_2(t) \end{pmatrix} = -\mathbf{S}_2(\theta) \begin{pmatrix} \mathbf{v}_2(t) \\ \mathbf{p}_2(t) \end{pmatrix} - \mathbf{S}_{2\Gamma}(\theta) \begin{pmatrix} \mathbf{v}_\Gamma(t) \\ \mathbf{p}_0(t) \end{pmatrix} + \begin{pmatrix} \mathbf{K}_2(\theta) \\ \mathbf{L}_2(\theta) \end{pmatrix} \mathbf{u}(t), \tag{5.44a}$$

$$\mathbf{M}_2(\theta)\mathbf{v}_2(0) = \mathbf{v}_2^{(0)}(\theta), \tag{5.44b}$$

$$\mathbf{L}_2(\theta)\mathbf{u}(0) = \mathbf{B}_{22}(\theta)\mathbf{M}_2(\theta)^{-1}\mathbf{v}_2^{(0)}(\theta) + \mathbf{B}_{2\Gamma}(\theta)\mathbf{M}_\Gamma(\theta)^{-1}\mathbf{v}_\Gamma^{(0)}(\theta), \tag{5.44c}$$

and

$$-\mathbf{E}_2(\theta)\frac{d}{dt}\begin{pmatrix}\boldsymbol{\lambda}_2(t)\\ \boldsymbol{\kappa}_2(t)\end{pmatrix} = -\mathbf{S}_2(\theta)\begin{pmatrix}\boldsymbol{\lambda}_2(t)\\ \boldsymbol{\kappa}_2(t)\end{pmatrix} - \mathbf{S}_{2\Gamma}(\theta)\begin{pmatrix}\boldsymbol{\lambda}_\Gamma(t)\\ \boldsymbol{\kappa}_0(t)\end{pmatrix} - \begin{pmatrix}\nabla_{\mathbf{v}_2}\boldsymbol{\ell}(\mathbf{v}_2,\mathbf{p}_2,\mathbf{v}_\Gamma,\mathbf{p}_0,t,\theta)\\ \nabla_{\mathbf{p}_2}\boldsymbol{\ell}(\mathbf{v}_2,\mathbf{p}_2,\mathbf{v}_\Gamma,\mathbf{p}_0,t,\theta)\end{pmatrix}, \tag{5.45a}$$

$$\mathbf{M}_2(\theta)\boldsymbol{\lambda}_2(T) = \boldsymbol{\lambda}_2^{(T)}(\theta), \tag{5.45b}$$

$$\nabla_{\mathbf{p}_2}\boldsymbol{\ell}(\mathbf{v}_2,\mathbf{p}_2,\mathbf{v}_\Gamma,\mathbf{p}_0,t,\theta) = -\mathbf{B}_{22}(\theta)\mathbf{M}_2(\theta)^{-1}\boldsymbol{\lambda}_2^{(T)}(\theta) - \mathbf{B}_{2\Gamma}(\theta)\mathbf{M}_\Gamma(\theta)^{-1}\boldsymbol{\lambda}_\Gamma^{(T)}(\theta). \tag{5.45c}$$

(iii) Optimality system associated with the interface $\Gamma(\theta)$:

$$\mathbf{E}_\Gamma(\theta)\frac{d}{dt}\begin{pmatrix}\mathbf{v}_\Gamma(t)\\ \mathbf{p}_0(t)\end{pmatrix} = -\mathbf{S}_\Gamma(\theta)\begin{pmatrix}\mathbf{v}_\Gamma(t)\\ \mathbf{p}_0(t)\end{pmatrix} - \mathbf{S}_{1\Gamma}^T\begin{pmatrix}\mathbf{v}_1(t)\\ \mathbf{p}_1(t)\end{pmatrix} - \mathbf{S}_{2\Gamma}^T(\theta)\begin{pmatrix}\mathbf{v}_2(t)\\ \mathbf{p}_2(t)\end{pmatrix} + \begin{pmatrix}\mathbf{K}_\Gamma(\theta)\\ \mathbf{L}_0(\theta)\end{pmatrix}\mathbf{u}(t), \tag{5.46a}$$

$$\mathbf{M}_\Gamma(\theta)\mathbf{v}_\Gamma(0) = \mathbf{v}_\Gamma^{(0)}(\theta), \tag{5.46b}$$

$$\mathbf{L}_0(\theta)\mathbf{u}(0) = \mathbf{B}_0(\theta)\mathbf{M}_\Gamma(\theta)^{-1}\mathbf{v}_\Gamma^{(0)}(\theta), \tag{5.46c}$$

and

$$-\mathbf{E}_\Gamma(\theta)\begin{pmatrix}\boldsymbol{\lambda}_\Gamma(t)\\ \boldsymbol{\kappa}_0(t)\end{pmatrix} = -\mathbf{S}_\Gamma(\theta)\begin{pmatrix}\boldsymbol{\lambda}_\Gamma(t)\\ \boldsymbol{\kappa}_0(t)\end{pmatrix} - \mathbf{S}_{1\Gamma}^T\begin{pmatrix}\boldsymbol{\lambda}_1(t)\\ \boldsymbol{\kappa}_1(t)\end{pmatrix} - \mathbf{S}_{2\Gamma}^T(\theta)\begin{pmatrix}\boldsymbol{\lambda}_2(t)\\ \boldsymbol{\kappa}_2(t)\end{pmatrix} - \begin{pmatrix}\nabla_{\mathbf{v}_\Gamma}\boldsymbol{\ell}(\mathbf{v}_2,\mathbf{p}_2,\mathbf{v}_\Gamma,\mathbf{p}_0,t,\theta)\\ \nabla_{\mathbf{p}_0}\boldsymbol{\ell}(\mathbf{v}_2,\mathbf{p}_2,\mathbf{v}_\Gamma,\mathbf{p}_0,t,\theta)\end{pmatrix} - \begin{pmatrix}\mathbf{0}\\ \mathbf{F}_0^T\end{pmatrix}\mathbf{z}_1, \tag{5.47a}$$

$$\mathbf{M}_\Gamma(\theta)\boldsymbol{\lambda}_\Gamma(T) = \boldsymbol{\lambda}_\Gamma^{(T)}(\theta), \tag{5.47b}$$

$$\nabla_{\mathbf{p}_0}\boldsymbol{\ell}(\mathbf{v}_2,\mathbf{p}_2,\mathbf{v}_\Gamma,\mathbf{p}_0,t,\theta)+\mathbf{F}_0^T\mathbf{z}_1 = -\mathbf{B}_0(\theta)\mathbf{M}_\Gamma(\theta)^{-1}\boldsymbol{\lambda}_\Gamma^{(T)}(\theta). \tag{5.47c}$$

The equations (5.42)-(5.47) have to be complemented by the variational inequality

$$\int_0^T \nabla_\theta \boldsymbol{\ell}(\mathbf{v}_2, \mathbf{p}_2, \mathbf{v}_\Gamma, \mathbf{p}_0, t, \theta)^T(\tilde{\theta} - \theta)\, dt \tag{5.48}$$
$$+ \int_0^T \begin{pmatrix} \boldsymbol{\mu}_2(t) \\ \boldsymbol{\mu}_\Gamma(t) \end{pmatrix}^T \begin{pmatrix} (D_\theta \mathbf{P}_2(\theta)(\tilde{\theta} - \theta))\, \mathbf{x}_2(t) - (D_\theta \mathbf{N}_2(\theta)(\tilde{\theta} - \theta))\, \mathbf{u}(t) \\ (D_\theta \mathbf{P}_\Gamma(\theta)(\tilde{\theta} - \theta))\mathbf{x}_\Gamma(t) - (D_\theta \mathbf{N}_\Gamma(\theta)(\tilde{\theta} - \theta))\, \mathbf{u}(t) \end{pmatrix} dt \geq 0$$

for all $\tilde{\theta} \in \Theta$.

Remark 5.3.2 *Since we are faced with implicit Hessenberg index 2 differential-algebraic systems, the final values* $\boldsymbol{\lambda}_1^{(T)}, \boldsymbol{\lambda}_2^{(T)}(\theta)$ *and* $\boldsymbol{\lambda}_\Gamma^{(T)}(\theta)$ *are in general nonzero and have to be computed as outlined in [28]. It seems that for most examples considered in the flow control literature (see, e.g., [1, 57, 85]) the problem structure is such that* $\boldsymbol{\lambda}_1^{(T)} = \mathbf{0}$, $\boldsymbol{\lambda}_2^{(T)}(\theta) = \mathbf{0}$ *and* $\boldsymbol{\lambda}_\Gamma^{(T)}(\theta) = \mathbf{0}$. *For a flow control problem in which the adjoint has a nonzero final time value see, e.g., [106].*

5.4 Balanced truncation model reduction of the domain decomposed optimality system

We construct a reduced order model for the optimality system (5.42)-(5.48) by applying balanced truncation only to the optimality system (5.42),(5.43) associated with the fixed subdomain Ω_1. To do this, we have to examine (5.42)-(5.48) to see how the subsystems (5.42),(5.43) interact with the remaining subsystems.

This leads to

$$\mathbf{E}_1 \frac{d}{dt}\begin{pmatrix} \mathbf{v}_1(t) \\ \mathbf{p}_1(t) \end{pmatrix} = -\mathbf{S}_1 \begin{pmatrix} \mathbf{v}_1(t) \\ \mathbf{p}_1(t) \end{pmatrix} - \mathbf{S}_{1\Gamma} \begin{pmatrix} \mathbf{v}_\Gamma(t) \\ \mathbf{p}_0(t) \end{pmatrix} + \begin{pmatrix} \mathbf{K}_1 \\ \mathbf{L}_1 \end{pmatrix} \mathbf{u}(t), \tag{5.49a}$$

$$\mathbf{z}_1(t) = \mathbf{C}_1\mathbf{v}_1(t) + \mathbf{F}_1\mathbf{p}_1(t) + \mathbf{F}_0\mathbf{p}_0(t) + \mathbf{D}_1\mathbf{u}(t) - \mathbf{d}(t), \tag{5.49b}$$

$$\begin{pmatrix} \mathbf{z}_{v,\Gamma}(t) \\ \mathbf{z}_{p,\Gamma}(t) \end{pmatrix} = -\mathbf{S}_{1\Gamma}^T \begin{pmatrix} \mathbf{v}_1(t) \\ \mathbf{p}_1(t) \end{pmatrix}, \tag{5.49c}$$

$$\mathbf{M}_1\mathbf{v}_1(0) = \mathbf{v}_1^{(0)}, \tag{5.49d}$$

$$\mathbf{L}_1\mathbf{u}(0) = \mathbf{B}_{11}\mathbf{M}_1^{-1}\mathbf{v}_1^{(0)} + \mathbf{B}_{1\Gamma}\mathbf{M}_\Gamma(\theta)^{-1}\mathbf{v}_\Gamma^{(0)}(\theta), \tag{5.49e}$$

and

$$-\mathbf{E}_1 \frac{d}{dt}\begin{pmatrix} \boldsymbol{\lambda}_1(t) \\ \boldsymbol{\kappa}_1(t) \end{pmatrix} = -\mathbf{S}_1 \begin{pmatrix} \boldsymbol{\lambda}_1(t) \\ \boldsymbol{\kappa}_1(t) \end{pmatrix} - \mathbf{S}_{1\Gamma} \begin{pmatrix} \boldsymbol{\lambda}_\Gamma(t) \\ \boldsymbol{\kappa}_0(t) \end{pmatrix} - \begin{pmatrix} \mathbf{C}_1^T \\ \mathbf{F}_1^T \end{pmatrix} \mathbf{z}_1(t), \tag{5.50a}$$

$$\mathbf{q}_1(t) = \mathbf{K}_1^T\boldsymbol{\lambda}_1(t) + \mathbf{L}_1^T\boldsymbol{\kappa}_1(t) - \mathbf{D}_1^T\mathbf{z}_1(t), \tag{5.50b}$$

$$\begin{pmatrix} \mathbf{q}_{v,\Gamma}(t) \\ \mathbf{q}_{p,\Gamma}(t) \end{pmatrix} = -\mathbf{S}_{1\Gamma}^T \begin{pmatrix} \boldsymbol{\lambda}_1(t) \\ \boldsymbol{\kappa}_1(t) \end{pmatrix} - \begin{pmatrix} \mathbf{0} \\ \mathbf{F}_0^T \end{pmatrix} \mathbf{z}_1, \tag{5.50c}$$

$$\mathbf{M}_1\boldsymbol{\lambda}_1(T) = \boldsymbol{\lambda}_1^{(T)}, \tag{5.50d}$$

$$\mathbf{F}_1^T\mathbf{z}_1(T) = -\mathbf{B}_{11}\mathbf{M}_1^{-1}\boldsymbol{\lambda}_1^{(T)} - \mathbf{B}_{1\Gamma}\mathbf{M}_\Gamma(\theta)^{-1}\boldsymbol{\lambda}_\Gamma^{(T)}(\theta). \tag{5.50e}$$

The outputs (5.49c) and (5.50c) are inputs into the subsystems (5.46) and (5.47), respectively. The terms $\mathbf{v}_\Gamma$, $\mathbf{p}_0$, $\boldsymbol{\lambda}_\Gamma$, $\boldsymbol{\kappa}_0$ are auxiliary inputs into the subsystems (5.49a) and (5.50a). The output (5.50b) does not feed into any of the subsystems in (5.42)-(5.48), but is added to emphasize the fact that (5.49) and (5.50) is exactly of the form (5.12) and (5.13). An important observation is that due to the fact that pressures are discontinuous on the boundary, the second row block in $\mathbf{S}_{1\Gamma}^T$ is zero (cf., (5.32) and (5.35c)). Consequently,

$$\mathbf{z}_{p,\Gamma}(t) \equiv 0 \quad \text{and} \quad \mathbf{q}_{p,\Gamma}(t) = -\mathbf{F}_0^T\mathbf{z}_1(t). \tag{5.51}$$

The subsystems (5.49a-c) and (5.50a-c) can be written as

$$\mathbf{E}_1 \frac{d}{dt} \begin{pmatrix} \mathbf{v}_1(t) \\ \mathbf{p}_1(t) \end{pmatrix} = -\mathbf{S}_1 \begin{pmatrix} \mathbf{v}_1(t) \\ \mathbf{p}_1(t) \end{pmatrix} + \begin{pmatrix} -\mathbf{A}_{1\Gamma} & \mathbf{0} & \mathbf{K}_1 \\ -\mathbf{B}_{1\Gamma} & \mathbf{0} & \mathbf{L}_1 \end{pmatrix} \begin{pmatrix} \mathbf{v}_\Gamma(t) \\ \mathbf{p}_0(t) \\ \mathbf{u}(t) \end{pmatrix}, \tag{5.52a}$$

$$\begin{pmatrix} \mathbf{z}_{v,\Gamma}(t) \\ \mathbf{z}_{p,\Gamma}(t) \\ \mathbf{z}_1(t) \end{pmatrix} = \begin{pmatrix} -\mathbf{A}_{1\Gamma}^T & -\mathbf{B}_{1\Gamma}^T \\ \mathbf{0} & \mathbf{0} \\ \mathbf{C}_1 & \mathbf{F}_1 \end{pmatrix} \begin{pmatrix} \mathbf{v}_1(t) \\ \mathbf{p}_1(t) \end{pmatrix} + \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{F}_0\mathbf{p}_0(t) + \mathbf{D}_1\mathbf{u}(t) - \mathbf{d}(t) \end{pmatrix} \tag{5.52b}$$

and

$$-\mathbf{E}_1 \frac{d}{dt} \begin{pmatrix} \boldsymbol{\lambda}_1(t) \\ \boldsymbol{\kappa}_1(t) \end{pmatrix} = -\mathbf{S}_1 \begin{pmatrix} \boldsymbol{\lambda}_1(t) \\ \boldsymbol{\kappa}_1(t) \end{pmatrix} + \begin{pmatrix} -\mathbf{A}_{1\Gamma} & \mathbf{0} & \mathbf{C}_1^T \\ -\mathbf{B}_{1\Gamma} & \mathbf{0} & \mathbf{F}_1^T \end{pmatrix} \begin{pmatrix} \boldsymbol{\lambda}_\Gamma(t) \\ \boldsymbol{\kappa}_0(t) \\ -\mathbf{z}_1(t) \end{pmatrix}, \tag{5.53a}$$

$$\begin{pmatrix} \mathbf{q}_{v,\Gamma}(t) \\ \mathbf{q}_{p,\Gamma}(t) \\ \mathbf{q}_1(t) \end{pmatrix} = \begin{pmatrix} -\mathbf{A}_{1\Gamma}^T & -\mathbf{B}_{1\Gamma}^T \\ \mathbf{0} & \mathbf{0} \\ \mathbf{K}_1^T & \mathbf{L}_1^T \end{pmatrix} \begin{pmatrix} \boldsymbol{\lambda}_1(t) \\ \boldsymbol{\kappa}_1(t) \end{pmatrix} + \begin{pmatrix} \mathbf{0} \\ \mathbf{F}_0^T \\ \mathbf{D}_1^T \end{pmatrix} (-\mathbf{z}_1(t)). \tag{5.53b}$$

To be able to apply the balanced truncation model reduction technique outlined in Section 5.2 we assume that $\mathbf{B}_{11} \in \mathbb{R}^{m_1 \times n_1}$ has rank m_1, that $\mathbf{M}_{11} \in \mathbb{R}^{n_1 \times n_1}$ is symmetric positive definite, and that the generalized eigenvalues of $(\mathbf{A}_{11}, \mathbf{M}_{11})$ have positive real part. These assumptions are satisfied with a proper spatial decomposition of the problem.

If we apply the techniques introduced in Section 5.2 we obtain the following reduced optimality system.

(i) Reduced optimality system associated with the subdomain Ω_1:

$$\frac{d}{dt}\widehat{\mathbf{v}}_1(t) = -\mathbf{W}^T\mathbf{A}_{11}\mathbf{V}\widehat{\mathbf{v}}_1(t) + \mathbf{W}^T\widetilde{\mathbf{B}}_1\begin{pmatrix}\widehat{\mathbf{v}}_\Gamma(t)\\ \widehat{\mathbf{p}}_0(t)\\ \mathbf{u}(t)\end{pmatrix}, \tag{5.54a}$$

$$\begin{pmatrix}\widehat{\mathbf{z}}_{v,\Gamma}(t)\\ \widehat{\mathbf{z}}_{p,\Gamma}(t)\\ \widehat{\mathbf{z}}_1(t)\end{pmatrix} = \widetilde{\mathbf{C}}_1\mathbf{V}\widehat{\mathbf{v}}_1(t) + \widetilde{\mathbf{D}}_1\begin{pmatrix}\widehat{\mathbf{v}}_\Gamma(t)\\ \widehat{\mathbf{p}}_0(t)\\ \mathbf{u}(t)\end{pmatrix} - \widetilde{\mathbf{H}}_1\frac{d}{dt}\begin{pmatrix}\widehat{\mathbf{v}}_\Gamma(t)\\ \widehat{\mathbf{p}}_0(t)\\ \mathbf{u}(t)\end{pmatrix}, \tag{5.54b}$$

$$\widehat{\mathbf{v}}(0) = \mathbf{W}^T\mathbf{\Pi}_1\mathbf{v}_1^{(0)} \tag{5.54c}$$

and

$$-\frac{d}{dt}\widehat{\boldsymbol{\lambda}}_1(t) = -\mathbf{V}^T\mathbf{A}_{11}\mathbf{W}\widehat{\boldsymbol{\lambda}}_1(t) + \mathbf{V}^T\widetilde{\mathbf{C}}_1^T\begin{pmatrix}\widehat{\boldsymbol{\lambda}}_\Gamma(t)\\ \widehat{\boldsymbol{\kappa}}_0(t)\\ -\widehat{\mathbf{z}}_1(t)\end{pmatrix}, \tag{5.55a}$$

$$\begin{pmatrix}\widehat{\mathbf{q}}_{v,\Gamma}(t)\\ \widehat{\mathbf{q}}_{p,\Gamma}(t)\\ \widehat{\mathbf{q}}_1(t)\end{pmatrix} = \widetilde{\mathbf{B}}_1^T\mathbf{W}\widehat{\boldsymbol{\lambda}}_1(t) + \widetilde{\mathbf{D}}_1^T\begin{pmatrix}\widehat{\boldsymbol{\lambda}}_\Gamma(t)\\ \widehat{\boldsymbol{\kappa}}_0(t)\\ -\widehat{\mathbf{z}}_1(t)\end{pmatrix} + \widetilde{\mathbf{H}}_1^T\frac{d}{dt}\begin{pmatrix}\widehat{\boldsymbol{\lambda}}_\Gamma(t)\\ \widehat{\boldsymbol{\kappa}}_0(t)\\ -\widehat{\mathbf{z}}_1(t)\end{pmatrix}, \tag{5.55b}$$

$$\widehat{\boldsymbol{\lambda}}_1(T) = \mathbf{V}^T\mathbf{\Pi}_1\boldsymbol{\lambda}_1^{(T)}. \tag{5.55c}$$

Here $\mathbf{\Pi}_1 = \mathbf{I} - \mathbf{B}_{11}^T(\mathbf{B}_{11}\mathbf{M}_{11}^{-1}\mathbf{B}_{11}^T)^{-1}\mathbf{B}_{11}\mathbf{M}_{11}^{-1}$ and

$$\widetilde{\mathbf{B}}_1 = (-\mathbf{A}_{1\Gamma}\,|\,\mathbf{0}\,|\,\mathbf{K}_1) - \mathbf{A}_{11}\mathbf{M}_{11}^{-1}\mathbf{B}_{11}^T(\mathbf{B}_{11}\mathbf{M}_{11}^{-1}\mathbf{B}_{11}^T)^{-1}\,(-\mathbf{B}_{1\Gamma}\,|\,\mathbf{0}\,|\,\mathbf{L}_1)$$

$$\widetilde{\mathbf{C}}_1 = \begin{pmatrix}-\mathbf{A}_{1\Gamma}^T\\ \mathbf{0}\\ \mathbf{C}_1\end{pmatrix} - \begin{pmatrix}-\mathbf{B}_{1\Gamma}^T\\ \mathbf{0}\\ \mathbf{F}_1\end{pmatrix}(\mathbf{B}_{11}\mathbf{M}_{11}^{-1}\mathbf{B}_{11}^T)^{-1}\mathbf{B}_{11}\mathbf{M}_{11}^{-1}\mathbf{A}_{11},$$

$$\widetilde{\mathbf{D}}_1 = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_0 & \mathbf{D}_1 \end{pmatrix} + \begin{pmatrix} -\mathbf{A}_{1\Gamma}^T \\ \mathbf{0} \\ \mathbf{C}_1 \end{pmatrix} \mathbf{M}_{11}^{-1}\mathbf{B}_{11}^T(\mathbf{B}_{11}\mathbf{M}_{11}^{-1}\mathbf{B}_{11}^T)^{-1}\left(-\mathbf{B}_{1\Gamma} \,|\, \mathbf{0} \,|\, \mathbf{L}_1\right)$$
$$+ \begin{pmatrix} -\mathbf{B}_{1\Gamma}^T \\ \mathbf{0} \\ \mathbf{F}_1 \end{pmatrix} (\mathbf{B}_{11}\mathbf{M}_{11}^{-1}\mathbf{B}_{11}^T)^{-1}\mathbf{B}_{11}\mathbf{M}_{11}^{-1}\widetilde{\mathbf{B}}_1,$$

$$\widetilde{\mathbf{H}}_1 = \begin{pmatrix} -\mathbf{B}_{1\Gamma}^T \\ \mathbf{0} \\ \mathbf{F}_1 \end{pmatrix} (\mathbf{B}_{11}\mathbf{M}_{11}^{-1}\mathbf{B}_{11}^T)^{-1}\left(-\mathbf{B}_{1\Gamma} \,|\, \mathbf{0} \,|\, \mathbf{L}_1\right).$$

Note that the structure of $\widetilde{\mathbf{B}}_1$, $\widetilde{\mathbf{C}}_1$, $\widetilde{\mathbf{D}}_1$, and $\widetilde{\mathbf{H}}_1$ imply

$$\widehat{\mathbf{z}}_{p,\Gamma}(t) \equiv 0 \quad \text{and} \quad \widehat{\mathbf{q}}_{p,\Gamma}(t) \equiv 0. \tag{5.56}$$

The reduced optimality system associated with the subdomain Ω_1 is coupled to the following optimality subsystems.

(ii) Optimality system associated with the subdomain $\Omega_2(\theta)$:

$$\mathbf{E}_2(\theta)\frac{d}{dt}\begin{pmatrix} \widehat{\mathbf{v}}_2(t) \\ \widehat{\mathbf{p}}_2(t) \end{pmatrix} = -\mathbf{S}_2(\theta)\begin{pmatrix} \widehat{\mathbf{v}}_2(t) \\ \widehat{\mathbf{p}}_2(t) \end{pmatrix} - \mathbf{S}_{2\Gamma}(\theta)\begin{pmatrix} \widehat{\mathbf{v}}_\Gamma(t) \\ \widehat{\mathbf{p}}_0(t) \end{pmatrix} + \begin{pmatrix} \mathbf{K}_2(\theta) \\ \mathbf{L}_2(\theta) \end{pmatrix}\mathbf{u}(t), \tag{5.57a}$$

$$\mathbf{M}_2(\theta)\widehat{\mathbf{v}}_2(0) = \mathbf{v}_2^{(0)}(\theta), \tag{5.57b}$$

$$\mathbf{L}_2(\theta)\mathbf{u}(0) = \mathbf{B}_{22}(\theta)\mathbf{M}_2(\theta)^{-1}\mathbf{v}_2^{(0)}(\theta) + \mathbf{B}_{2\Gamma}(\theta)\mathbf{M}_\Gamma(\theta)^{-1}\mathbf{v}_\Gamma^{(0)}(\theta), \tag{5.57c}$$

and

$$-\mathbf{E}_2(\theta)\frac{d}{dt}\begin{pmatrix}\widehat{\boldsymbol{\lambda}}_2(t)\\ \widehat{\boldsymbol{\kappa}}_2(t)\end{pmatrix} = -\mathbf{S}_2(\theta)\begin{pmatrix}\widehat{\boldsymbol{\lambda}}_2(t)\\ \widehat{\boldsymbol{\kappa}}_2(t)\end{pmatrix} - \mathbf{S}_{2\Gamma}(\theta)\begin{pmatrix}\widehat{\boldsymbol{\lambda}}_\Gamma(t)\\ \widehat{\boldsymbol{\kappa}}_0(t)\end{pmatrix} - \begin{pmatrix}\nabla_{\widehat{\mathbf{v}}_2}\boldsymbol{\ell}(\widehat{\mathbf{v}}_2,\widehat{\mathbf{p}}_2,\widehat{\mathbf{v}}_\Gamma,\widehat{\mathbf{p}}_0,t,\theta)\\ \nabla_{\widehat{\mathbf{p}}_2}\boldsymbol{\ell}(\widehat{\mathbf{v}}_2,\widehat{\mathbf{p}}_2,\widehat{\mathbf{v}}_\Gamma,\widehat{\mathbf{p}}_0,t,\theta)\end{pmatrix}, \tag{5.58a}$$

$$\mathbf{M}_2(\theta)\widehat{\boldsymbol{\lambda}}_2(T) = \boldsymbol{\lambda}_2^{(T)}(\theta). \tag{5.58b}$$

(iii) Optimality system associated with the interface $\Gamma(\theta)$:

$$\mathbf{E}_\Gamma(\theta)\frac{d}{dt}\begin{pmatrix}\widehat{\mathbf{v}}_\Gamma(t)\\ \widehat{\mathbf{p}}_0(t)\end{pmatrix} = -\mathbf{S}_\Gamma(\theta)\begin{pmatrix}\widehat{\mathbf{v}}_\Gamma(t)\\ \widehat{\mathbf{p}}_0(t)\end{pmatrix} + \begin{pmatrix}\widehat{\mathbf{z}}_{v,\Gamma}(t)\\ \widehat{\mathbf{z}}_{p,\Gamma}(t)\end{pmatrix} - \mathbf{S}_{2\Gamma}^T(\theta)\begin{pmatrix}\widehat{\mathbf{v}}_2(t)\\ \widehat{\mathbf{p}}_2(t)\end{pmatrix} + \begin{pmatrix}\mathbf{K}_\Gamma(\theta)\\ \mathbf{L}_\Gamma(\theta)\end{pmatrix}\mathbf{u}(t), \tag{5.59a}$$

$$\mathbf{M}_\Gamma(\theta)\widehat{\mathbf{v}}_\Gamma(0) = \mathbf{v}_\Gamma^{(0)}(\theta), \tag{5.59b}$$

and

$$-\mathbf{E}_\Gamma(\theta)\begin{pmatrix}\widehat{\boldsymbol{\lambda}}_\Gamma(t)\\ \widehat{\boldsymbol{\kappa}}_0(t)\end{pmatrix} = -\mathbf{S}_\Gamma(\theta)\begin{pmatrix}\widehat{\boldsymbol{\lambda}}_\Gamma(t)\\ \widehat{\boldsymbol{\kappa}}_0(t)\end{pmatrix} + \begin{pmatrix}\widehat{\mathbf{q}}_{v,\Gamma}(t)\\ \widehat{\mathbf{q}}_{p,\Gamma}(t)\end{pmatrix} - \mathbf{S}_{2\Gamma}^T(\theta)\begin{pmatrix}\widehat{\boldsymbol{\lambda}}_2(t)\\ \widehat{\boldsymbol{\kappa}}_2(t)\end{pmatrix} - \begin{pmatrix}\nabla_{\widehat{\mathbf{v}}_\Gamma}\boldsymbol{\ell}(\widehat{\mathbf{v}}_2,\widehat{\mathbf{p}}_2,\widehat{\mathbf{v}}_\Gamma,\widehat{\mathbf{p}}_0,t,\theta)\\ \nabla_{\widehat{\mathbf{p}}_0}\boldsymbol{\ell}(\widehat{\mathbf{v}}_2,\widehat{\mathbf{p}}_2,\widehat{\mathbf{v}}_\Gamma,\widehat{\mathbf{p}}_0,t,\theta)\end{pmatrix}, \tag{5.60a}$$

$$\mathbf{M}_\Gamma(\theta)\widehat{\boldsymbol{\lambda}}_\Gamma(T) = \boldsymbol{\lambda}_\Gamma^{(T)}(\theta). \tag{5.60b}$$

The equations have to be complemented by the variational inequality

$$\int_0^T \nabla_\theta\boldsymbol{\ell}(\widehat{\mathbf{v}}_2,\widehat{\mathbf{p}}_2,\widehat{\mathbf{v}}_\Gamma,\widehat{\mathbf{p}}_0,t,\theta)^T(\tilde{\theta}-\theta)\,dt \tag{5.61}$$

$$+\int_0^T \begin{pmatrix}\widehat{\boldsymbol{\mu}}_2(t)\\ \widehat{\boldsymbol{\lambda}}_\Gamma(t)\end{pmatrix}^T \begin{pmatrix}(D_\theta\mathbf{P}_2(\theta)(\tilde{\theta}-\theta))\,\widehat{\mathbf{x}}_2(t) - (D_\theta\mathbf{N}_2(\theta)(\tilde{\theta}-\theta))\,\mathbf{u}(t)\\ (D_\theta\mathbf{P}_\Gamma(\theta)(\tilde{\theta}-\theta))\,\widehat{\mathbf{x}}_\Gamma(t) - (D_\theta\mathbf{N}_\Gamma(\theta)(\tilde{\theta}-\theta))\,\mathbf{u}(t)\end{pmatrix} dt \geq 0\ ,\ \tilde{\theta}\in\Theta.$$

We have applied domain decomposition and balanced truncation model reduction to derive the reduced order optimality system (5.54)-(5.61) from the full order optimality system (5.42)-(5.48). This raises the question whether the reduced order optimality system (5.54)-(5.61) is the optimality system for a reduced order optimization problem. This is important, since numerically we solve the shape optimization problem using gradient-based optimization methods rather than explicitly solving the optimality system.

Theorem 5.4.1 *The reduced order optimality system* (5.54)-(5.61) *represents the first order necessary optimality conditions for the shape optimization problem*

$$\min \widehat{J}(\theta) \tag{5.62a}$$

$$s.t.\ \theta \in \Theta$$

where $\widehat{J}(\theta) = \widehat{J}_1(\widehat{\mathbf{v}}_1, \widehat{\mathbf{v}}_\Gamma, \widehat{\mathbf{p}}_0) + \widehat{J}_2(\widehat{\mathbf{v}}_2, \widehat{\mathbf{p}}_2, \widehat{\mathbf{v}}_\Gamma, \widehat{\mathbf{p}}_0, \theta)$,

$$\widehat{J}_1(\widehat{\mathbf{v}}_1, \widehat{\mathbf{v}}_\Gamma) = \frac{1}{2} \int_0^T |\widehat{\mathbf{z}}_1|^2 \ dt, \tag{5.62b}$$

$$\widehat{J}_2(\widehat{\mathbf{v}}_2, \widehat{\mathbf{p}}_2, \widehat{\mathbf{v}}_\Gamma, \widehat{\mathbf{p}}_0, \theta) = \int_0^T \ell(\widehat{\mathbf{v}}_2, \widehat{\mathbf{p}}_2, \widehat{\mathbf{v}}_\Gamma, \widehat{\mathbf{p}}_0, t, \theta) \ dt. \tag{5.62c}$$

and where $\widehat{\mathbf{z}}_1$, $\widehat{\mathbf{v}} = (\widehat{\mathbf{v}}_1, \widehat{\mathbf{v}}_2, \widehat{\mathbf{v}}_\Gamma)^T$ $\widehat{\mathbf{p}} = (\widehat{\mathbf{p}}_2, \widehat{\mathbf{p}}_0)^T$, *are given as the solution of* (5.54), (5.57), (5.59).

The proof uses standard arguments and is omitted.

5.5 A priori estimate of the modeling error

Let $\theta^* \in \Theta$ and $\widehat{\theta}^* \in \Theta$ be local minima of the optimization problem (5.39) and its reduced version (5.62), where the states $\mathbf{v} = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_\Gamma)^T$ and $\mathbf{p} = (\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_0)^T$ solve

(5.42),(5.44),(5.46), and where the reduced states $\widehat{\mathbf{v}} = (\widehat{\mathbf{v}}_1, \widehat{\mathbf{v}}_2, \widehat{\mathbf{v}}_\Gamma)^T$ and $\widehat{\mathbf{p}} = (\widehat{\mathbf{p}}_2, \widehat{\mathbf{p}}_0)^T$ solve (5.54), (5.57), (5.59). Considering the states as implicit functions of the design variables, (5.39) and (5.62), can simply be written as

$$\inf_{\theta\in\Theta} J(\theta) \quad \text{and} \quad \inf_{\theta\in\Theta} \widehat{J}(\theta).$$

We want to derive an upper bound for the modeling error $\|\theta^* - \widehat{\theta}^*\|$ in terms of the Hankel singular values occurring in the BTMR of the optimality system for the fixed subdomain Ω_1. Under the convexity assumption, there exists $\kappa > 0$ such that

$$\Big(\nabla J(\widehat{\theta}^*) - \nabla J(\theta^*)\Big)^T (\widehat{\theta}^* - \theta^*) \geq \kappa \, \|\widehat{\theta}^* - \theta^*\|^2. \tag{5.63}$$

It is easy to see that

$$\|\theta^* - \widehat{\theta}^*\| \leq \kappa^{-1} \, \|\nabla \widehat{J}(\widehat{\theta}^*) - \nabla J(\widehat{\theta}^*)\|, \tag{5.64}$$

see, e.g., [14]. Hence, we need to provide an upper bound for the right-hand side in (5.64). The gradients of the objective functions J and $\widehat{J}$ can be expressed using the Lagrangian in (5.40) and its reduced analogue associated with (5.62). More precisely, we have

$$\begin{aligned}
\Big(\nabla J(\theta) - \nabla \widehat{J}(\theta)\Big)^T \tilde{\theta} = \int_0^T \Big(\nabla_\theta \boldsymbol{\ell}(\mathbf{v}_2, \mathbf{v}_\Gamma, \mathbf{p}_2, \mathbf{p}_0, t, \theta) - \nabla_\theta \boldsymbol{\ell}(\widehat{\mathbf{v}}_2, \widehat{\mathbf{v}}_\Gamma, \widehat{\mathbf{p}}_2, \widehat{\mathbf{p}}_0, t, \theta)\Big)^T \tilde{\theta} \; dt \quad (5.65)\\
+ \int_0^T \begin{pmatrix} \widehat{\boldsymbol{\mu}}_2(t) \\ \widehat{\boldsymbol{\mu}}_\Gamma(t) \end{pmatrix}^T \begin{pmatrix} (D_\theta \mathbf{P}_2(\theta)\tilde{\theta})\,(\mathbf{x}_2 - \widehat{\mathbf{x}}_2)(t) \\ (D_\theta \mathbf{P}_\Gamma(\theta)\tilde{\theta})\,(\mathbf{x}_\Gamma - \widehat{\mathbf{x}}_\Gamma)(t) \end{pmatrix} dt\\
+ \int_0^T \begin{pmatrix} (\boldsymbol{\mu}_2 - \widehat{\boldsymbol{\mu}}_2)(t) \\ (\boldsymbol{\mu}_\Gamma - \widehat{\boldsymbol{\mu}}_\Gamma)(t) \end{pmatrix}^T \begin{pmatrix} (D_\theta \mathbf{P}_2(\theta)\tilde{\theta})\,\mathbf{x}_2(t) - (D_\theta \mathbf{N}_2(\theta)\tilde{\theta})\,\mathbf{u}(t) \\ (D_\theta \mathbf{P}_\Gamma(\theta)\tilde{\theta})\,\mathbf{x}_\Gamma(t) - (D_\theta \mathbf{N}_\Gamma(\theta)\tilde{\theta})\,\mathbf{u}(t) \end{pmatrix} dt
\end{aligned}$$

where $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_\Gamma)^T$, with $\mathbf{x}_i = (\mathbf{v}_i, \mathbf{p}_i)^T$, $i = 1, 2$, $\mathbf{x}_\Gamma = (\mathbf{v}_\Gamma, \mathbf{p}_0)^T$, and $\boldsymbol{\mu} = (\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\mu}_\Gamma)^T$, with $\boldsymbol{\mu}_i = (\boldsymbol{\lambda}_i, \boldsymbol{\kappa}_i)^T$, $i = 1, 2$, $\boldsymbol{\mu}_\Gamma = (\boldsymbol{\lambda}_\Gamma, \boldsymbol{\kappa}_0)^T$ solve (5.42)-(5.47), and where $\widehat{\mathbf{x}} = (\widehat{\mathbf{x}}_1, \widehat{\mathbf{x}}_2, \widehat{\mathbf{x}}_\Gamma)^T$ with $\widehat{\mathbf{x}}_1 = \widehat{\mathbf{v}}_1$, $\widehat{\mathbf{x}}_2 = (\widehat{\mathbf{v}}_2, \widehat{\mathbf{p}}_2)^T$, $\widehat{\mathbf{x}}_\Gamma = (\widehat{\mathbf{v}}_\Gamma, \widehat{\mathbf{p}}_0)^T$ and $\widehat{\boldsymbol{\mu}} = (\widehat{\boldsymbol{\mu}}_1, \widehat{\boldsymbol{\mu}}_2, \widehat{\boldsymbol{\mu}}_\Gamma)^T$ with $\widehat{\boldsymbol{\mu}}_1 = \widehat{\boldsymbol{\lambda}}_1$, $\widehat{\boldsymbol{\mu}}_2 = (\widehat{\boldsymbol{\lambda}}_2, \widehat{\boldsymbol{\kappa}}_2)^T$, $\widehat{\boldsymbol{\mu}}_\Gamma = (\widehat{\boldsymbol{\lambda}}_\Gamma, \widehat{\boldsymbol{\kappa}}_0)^T$ solve (5.54)-(5.60).

In order to estimate (5.65) from above, we have to establish upper bounds for $\mathbf{x}_2 - \widehat{\mathbf{x}}_2, \mathbf{x}_\Gamma - \widehat{\mathbf{x}}_\Gamma$ and $\boldsymbol{\mu}_2 - \widehat{\boldsymbol{\mu}}_2, \boldsymbol{\mu}_\Gamma - \widehat{\boldsymbol{\mu}}_\Gamma$. This will be done in the sequel, where C will denote a generic positive constant not necessarily the same at each occurrence.

We apply the balanced truncation error bound (5.27) to estimate the error due to the reduction of the optimality subsystem corresponding to subdomain Ω_1. The error bound applies when $\mathbf{v}_1^{(0)} = 0$ and $\boldsymbol{\lambda}_1^{(T)} = 0$, which we will assume. This assumption can be relaxed when a modification of balanced truncation is applied. See Remark 5.2.1 in Section 5.2.

In order to provide an estimate of the errors in the adjoint states, we make the following assumption on the matrices $\mathbf{A}(\theta), \mathbf{B}(\theta), \mathbf{M}(\theta)$ defined in (5.32), (5.33b), and submatrices corresponding to subdomain Ω_1. This assumption is satisfied under a proper spatial decomposition of the problem as described in Section 5.3.

($\mathbf{A_1}$) The matrix $\mathbf{B}(\theta) \in \mathbb{R}^{m\times n}$ has rank m, the matrix $\mathbf{M}(\theta) \in \mathbb{R}^{n\times n}$ is symmetric positive definite, and the generalized eigenvalues of $(\mathbf{A}(\theta), \mathbf{M}(\theta))$ have positive real part.
The matrix $\mathbf{B}_{11} \in \mathbb{R}^{m_1\times n_1}$ has rank m_1, the matrix $\mathbf{M}_{11} \in \mathbb{R}^{n_1\times n_1}$ is symmetric positive definite, and the generalized eigenvalues of $(\mathbf{A}_{11}, \mathbf{M}_{11})$ have positive real part.

The first part allows the application of Theorem 5.1.1. The assumption on the submatrices corresponding to subdomain Ω_1 were needed for the application of balanced truncation model reduction to the optimality subsystem (5.52,5.53).

Lemma 5.5.1 *Let* $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_\Gamma)^T$, *where* $\mathbf{x}_i = (\mathbf{v}_i, \mathbf{p}_i)^T$, $1 \le i \le 2$, $\mathbf{x}_\Gamma = (\mathbf{v}_\Gamma, \mathbf{p}_0)^T$, *and* $\widehat{\mathbf{x}} = (\widehat{\mathbf{x}}_1, \widehat{\mathbf{x}}_2, \widehat{\mathbf{x}}_\Gamma)^T$, *where* $\widehat{\mathbf{x}}_1 = \widehat{\mathbf{v}}_1$, $\widehat{\mathbf{x}}_2 = (\widehat{\mathbf{v}}_2, \widehat{\mathbf{p}}_2)^T$, $\widehat{\mathbf{x}}_\Gamma = (\widehat{\mathbf{v}}_\Gamma, \widehat{\mathbf{p}}_0)^T$.

If ($\mathbf{A_1}$), (5.26) *and* $\mathbf{v}_1^{(0)} = 0$ *hold and if* $\mathbf{x}$ *and* $\widehat{\mathbf{x}}$ *satisfy* (5.42), (5.44), (5.46) *and* (5.54),

(5.57), (5.59), *respectively, then*

$$\left\| \begin{pmatrix} \mathbf{v}_2 - \widehat{\mathbf{v}}_2 \\ \mathbf{v}_\Gamma - \widehat{\mathbf{v}}_\Gamma \end{pmatrix} \right\|_{L^2} \leq c_v \Big(\sigma_{k+1} + \cdots + \sigma_n \Big), \tag{5.66a}$$

$$\left\| \begin{pmatrix} \mathbf{p}_2 - \widehat{\mathbf{p}}_2 \\ \mathbf{p}_0 - \widehat{\mathbf{p}}_0 \end{pmatrix} \right\|_{L^2} \leq c_p \Big(\sigma_{k+1} + \cdots + \sigma_n \Big), \tag{5.66b}$$

where

$$c_v = \frac{4}{\alpha} \left\| \overline{\mathbf{M}}^{-1} \right\| \left\| \begin{pmatrix} \mathbf{u} \\ \widehat{\mathbf{x}}_\Gamma \end{pmatrix} \right\|_{L^2},$$

$$c_p = 2 \left\| (\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{B}\mathbf{M}^{-1} \right\| \left(\frac{2}{\alpha} \left\| \overline{\mathbf{M}}^{-1} \right\| \|\mathbf{A}\| + 1 \right) \left\| \begin{pmatrix} \mathbf{u} \\ \widehat{\mathbf{x}}_\Gamma \end{pmatrix} \right\|_{L^2},$$

and

$$\|\mathbf{z}_1 - \widehat{\mathbf{z}}_1\|_{L^2} \leq 2 \Big\{ \frac{2}{\alpha} \left\| \overline{\mathbf{M}}^{-1} \right\| \|\mathbf{C}_1\| + \big(\|\mathbf{F}_1\| + \|\mathbf{F}_0\| \big) \left\| (\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{B}\mathbf{M}^{-1} \right\| \Big(\frac{2}{\alpha} \left\| \overline{\mathbf{M}}^{-1} \right\| \|\mathbf{A}\| + 1 \Big) + 1 \Big\} \left\| \begin{pmatrix} \mathbf{u} \\ \widehat{\mathbf{x}}_\Gamma \end{pmatrix} \right\|_{L^2} \Big(\sigma_{k+1} + \cdots + \sigma_n \Big). \tag{5.66c}$$

Proof: We introduce an auxiliary state $(\widetilde{\mathbf{v}}_1, \widetilde{\mathbf{p}}_1)$ as the solution of

$$\mathbf{E}_1 \frac{d}{dt} \begin{pmatrix} \widetilde{\mathbf{v}}_1(t) \\ \widetilde{\mathbf{p}}_1(t) \end{pmatrix} = -\mathbf{S}_1 \begin{pmatrix} \widetilde{\mathbf{v}}_1(t) \\ \widetilde{\mathbf{p}}_1(t) \end{pmatrix} - \mathbf{S}_{1\Gamma} \begin{pmatrix} \widehat{\mathbf{v}}_\Gamma(t) \\ \widehat{\mathbf{p}}_0(t) \end{pmatrix} + \begin{pmatrix} \mathbf{K}_1 \\ \mathbf{L}_1 \end{pmatrix} \mathbf{u}(t), \tag{5.67a}$$

$$\widetilde{\mathbf{z}}_1(t) = \mathbf{C}_1\widetilde{\mathbf{v}}_1(t) + \mathbf{F}_1\widetilde{\mathbf{p}}_1(t) + \mathbf{F}_0\widehat{\mathbf{p}}_0(t) + \mathbf{D}_1\mathbf{u}(t) - \mathbf{d}(t), \tag{5.67b}$$

$$\begin{pmatrix} \widetilde{\mathbf{z}}_{v,\Gamma}(t) \\ \widetilde{\mathbf{z}}_{p,\Gamma}(t) \end{pmatrix} = -\mathbf{S}_{1\Gamma}^T \begin{pmatrix} \widetilde{\mathbf{v}}_1(t) \\ \widetilde{\mathbf{p}}_1(t) \end{pmatrix}, \tag{5.67c}$$

$$\mathbf{M}_1\widetilde{\mathbf{v}}_1(0) = \mathbf{v}_1^{(0)}, \tag{5.67d}$$

$$\mathbf{L}_1\mathbf{u}(0) = \mathbf{B}_{11}\mathbf{M}_1^{-1}\mathbf{v}_1^{(0)} + \mathbf{B}_{1\Gamma}\mathbf{M}_\Gamma(\theta)^{-1}\mathbf{v}_\Gamma^{(0)}(\theta). \tag{5.67e}$$

Note that because the second row block in $\mathbf{S}_{1\Gamma}^T$ is zero (cf. (5.32) and (5.35c)), we have

$$\widetilde{\mathbf{z}}_{p,\Gamma}(t) \equiv 0. \tag{5.68}$$

This auxiliary system (5.67) is almost identical to (5.49), but has inputs $\widehat{\mathbf{v}}_\Gamma, \widehat{\mathbf{p}}_0$ instead of $\mathbf{v}_\Gamma, \mathbf{p}_0$. Thus the inputs for (5.67) and the reduced system (5.54) are the same and we can apply the balanced truncation error bound (5.27) to this subsystem. The balanced truncation error bound for this subsystem is

$$\left\| \begin{pmatrix} \widetilde{\mathbf{z}}_1 - \widehat{\mathbf{z}}_1 \\ \widetilde{\mathbf{z}}_{v,\Gamma} - \widehat{\mathbf{z}}_{v,\Gamma} \\ \widetilde{\mathbf{z}}_{p,\Gamma} - \widehat{\mathbf{z}}_{p,\Gamma} \end{pmatrix} \right\|_{L^2} \leq 2(\sigma_{k+1} + \ldots + \sigma_n) \left\| \begin{pmatrix} \mathbf{u} \\ \widehat{\mathbf{v}}_\Gamma \\ \widehat{\mathbf{p}}_0 \end{pmatrix} \right\|_{L^2}. \tag{5.69}$$

We set $\mathbf{e_v} = (\mathbf{v}_1 - \widetilde{\mathbf{v}}_1, \mathbf{v}_2 - \widehat{\mathbf{v}}_2, \mathbf{v}_\Gamma - \widehat{\mathbf{v}}_\Gamma)^T$ and $\mathbf{e_p} = (\mathbf{p}_1 - \widetilde{\mathbf{p}}_1, \mathbf{p}_2 - \widehat{\mathbf{p}}_2, \mathbf{p}_0 - \widehat{\mathbf{p}}_0)^T$. It follows from (5.42), (5.44), (5.46), (5.51) and (5.67), (5.57), (5.59), (5.56), (5.68) that $(\mathbf{e_v}, \mathbf{e_p})^T$ satisfies the system

$$\mathbf{E}(\theta)\frac{d}{dt}\begin{pmatrix} \mathbf{e_v}(t) \\ \mathbf{e_p}(t) \end{pmatrix} = -\mathbf{S}(\theta)\begin{pmatrix} \mathbf{e_v}(t) \\ \mathbf{e_p}(t) \end{pmatrix} + \begin{pmatrix} \mathbf{g}_1(t) \\ \mathbf{0} \end{pmatrix}, \ t \in (0, T], \tag{5.70a}$$

$$\mathbf{M}(\theta)\mathbf{e_v}(0) = \mathbf{0}, \tag{5.70b}$$

where

$$\mathbf{g}_1(t) = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \widetilde{\mathbf{z}}_{v,\Gamma} - \widehat{\mathbf{z}}_{v,\Gamma} \end{pmatrix}.$$

Applying Theorem 5.1.1 to (5.70) yields

$$\begin{aligned}
\left\| \begin{pmatrix} \mathbf{v}_1 - \widetilde{\mathbf{v}}_1 \\ \mathbf{v}_2 - \widehat{\mathbf{v}}_2 \\ \mathbf{v}_\Gamma - \widehat{\mathbf{v}}_\Gamma \end{pmatrix} \right\|_{L^2} &\leq \frac{2}{\alpha} \left\| \overline{\mathbf{M}}^{-1} \right\| \|\widetilde{\mathbf{z}}_{v,\Gamma} - \widehat{\mathbf{z}}_{v,\Gamma}\|_{L^2} , \\
\left\| \begin{pmatrix} \mathbf{p}_1 - \widetilde{\mathbf{p}}_1 \\ \mathbf{p}_2 - \widehat{\mathbf{p}}_2 \\ \mathbf{p}_0 - \widehat{\mathbf{p}}_0 \end{pmatrix} \right\|_{L^2} &\leq \left\| (\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{B}\mathbf{M}^{-1} \right\| \left(\frac{2}{\alpha} \left\| \overline{\mathbf{M}}^{-1} \right\| \|\mathbf{A}\| + 1 \right) \|\widetilde{\mathbf{z}}_{v,\Gamma} - \widehat{\mathbf{z}}_{v,\Gamma}\|_{L^2} .
\end{aligned} \tag{5.71}$$

The estimates (5.66a,b) follow from (5.69) and (5.71).

To prove (5.66c) we observe that (5.49b) and (5.67b) imply

$$\begin{aligned}
\|\mathbf{z}_1 - \widehat{\mathbf{z}}_1\|_{L^2} &= \|\mathbf{z}_1 - \widetilde{\mathbf{z}}_1\|_{L^2} + \|\widetilde{\mathbf{z}}_1 - \widehat{\mathbf{z}}_1\|_{L^2} \\
&\leq (\|\mathbf{C}_1\|\|\mathbf{v}_1 - \widetilde{\mathbf{v}}_1\|_{L^2} + \|\mathbf{F}_1\|\|\mathbf{p}_1 - \widetilde{\mathbf{p}}_1\|_{L^2} + \|\mathbf{F}_0\|\|\mathbf{p}_0 - \widehat{\mathbf{p}}_0\|_{L^2}) + \|\widetilde{\mathbf{z}}_1 - \widehat{\mathbf{z}}_1\|_{L^2} .
\end{aligned}$$

Together with (5.69) this implies (5.66c). □

In order to provide an estimate of the errors in the adjoint states, we make the following assumptions.

($\mathbf{A_2}$) $\mathbf{F}_1 = \mathbf{0}$ and $\mathbf{F}_0 = \mathbf{0}$, i.e, the objective function J_1 does not depend explicitly on the pressure.

($\mathbf{A_3}$) There exists a positive constant L_1 such that for all $\mathbf{x}_2, \mathbf{x}_2', \mathbf{x}_\Gamma, \mathbf{x}_\Gamma'$ and all $\theta \in \Theta, t \in$

$[0,T]$ there holds

$$\|\nabla_{\mathbf{v}}\boldsymbol{\ell}(\mathbf{x}_2,\mathbf{x}_\Gamma,t,\theta)-\nabla_{\mathbf{v}}\boldsymbol{\ell}(\mathbf{x}_2',\mathbf{x}_\Gamma',t,\theta)\|\le L_1\Big(\|\delta\mathbf{x}_2\|^2+\|\delta\mathbf{x}_\Gamma\|^2\Big)^{1/2},$$
$$\|\nabla_{\mathbf{p}}\boldsymbol{\ell}(\mathbf{x}_2,\mathbf{x}_\Gamma,t,\theta)-\nabla_{\mathbf{p}}\boldsymbol{\ell}(\mathbf{x}_2',\mathbf{x}_\Gamma',t,\theta)\|\le L_1\Big(\|\delta\mathbf{x}_2\|^2+\|\delta\mathbf{x}_\Gamma\|^2\Big)^{1/2},$$
$$\|\nabla_{\theta}\boldsymbol{\ell}(\mathbf{x}_2,\mathbf{x}_\Gamma,t,\theta)-\nabla_{\theta}\boldsymbol{\ell}(\mathbf{x}_2',\mathbf{x}_\Gamma',t,\theta)\|\le L_1\Big(\|\delta\mathbf{x}_2\|^2+\|\delta\mathbf{x}_\Gamma\|^2\Big)^{1/2},$$

where $\mathbf{v}\in\{\mathbf{v}_2,\mathbf{v}_\Gamma\}, \mathbf{p}\in\{\mathbf{p}_2,\mathbf{p}_0\}$ and $\delta\mathbf{x}_2:=\mathbf{x}_2-\mathbf{x}_2'$, $\delta\mathbf{x}_\Gamma:=\mathbf{x}_\Gamma-\mathbf{x}_\Gamma'$.

($\mathbf{A_4}$) There exists a positive constant γ such that for all $\theta\in\Theta$ and θ' with $\|\theta'\|\le 1$

$$\max\Big(\|D_\theta\mathbf{M}_2(\theta)\theta'\|,\|D_\theta\mathbf{M}_\Gamma(\theta)\theta'\|,\|D_\theta\mathbf{S}_2(\theta)\theta'\|,\|D_\theta\mathbf{S}_\Gamma(\theta)\theta'\|,$$
$$\|D_\theta\mathbf{S}_{2\Gamma}(\theta)\theta'\|,\|D_\theta\mathbf{N}_2(\theta)\theta'\|\|D_\theta\mathbf{N}_\Gamma(\theta)\theta'\|\Big)\le\gamma.$$

Lemma 5.5.2 *Let* $\mathbf{x},\widehat{\mathbf{x}}$ *as in Lemma 5.5.1 and* $\boldsymbol{\mu}=(\boldsymbol{\mu}_1,\boldsymbol{\mu}_2,\boldsymbol{\mu}_\Gamma)^T$*, where*

$$\boldsymbol{\mu}_i=(\boldsymbol{\lambda}_i,\boldsymbol{\kappa}_i)^T\ ,\ 1\le i\le 2\quad,\quad\boldsymbol{\mu}_\Gamma=(\boldsymbol{\lambda}_\Gamma,\boldsymbol{\kappa}_0)^T,$$

and $\widehat{\boldsymbol{\mu}}=(\widehat{\boldsymbol{\mu}}_1,\widehat{\boldsymbol{\mu}}_2,\widehat{\boldsymbol{\mu}}_\Gamma)^T$*, where* $\widehat{\boldsymbol{\mu}}_1=\widehat{\boldsymbol{\lambda}}_1$, $\widehat{\boldsymbol{\mu}}_2=(\widehat{\boldsymbol{\lambda}}_2,\widehat{\boldsymbol{\kappa}}_2)^T$, $\widehat{\boldsymbol{\mu}}_\Gamma=(\widehat{\boldsymbol{\lambda}}_\Gamma,\widehat{\boldsymbol{\kappa}}_0)^T$.

If ($\mathbf{A_1}$) – ($\mathbf{A_3}$), (5.26) *and* $\boldsymbol{\lambda}_1^{(T)}=0$ *hold, and if* $\mathbf{x},\boldsymbol{\mu}$ *and* $\widehat{\mathbf{x}},\widehat{\boldsymbol{\mu}}$, $\widehat{\mathbf{z}}_1$ *solve* (5.42)-(5.47) *and* (5.54)-(5.60)*, respectively, then*

$$\left\|\begin{pmatrix}\boldsymbol{\lambda}_2-\widehat{\boldsymbol{\lambda}}_2\\ \boldsymbol{\lambda}_\Gamma-\widehat{\boldsymbol{\lambda}}_\Gamma\end{pmatrix}\right\|_{L^2}\le c_\lambda\ \Big(\sigma_{k+1}+\cdots+\sigma_n\Big),\tag{5.72a}$$
$$\left\|\begin{pmatrix}\boldsymbol{\kappa}_2-\widehat{\boldsymbol{\kappa}}_2\\ \boldsymbol{\kappa}_0-\widehat{\boldsymbol{\kappa}}_0\end{pmatrix}\right\|_{L^2}\le c_\kappa\ \Big(\sigma_{k+1}+\cdots+\sigma_n\Big),\tag{5.72b}$$

where

$$c_\lambda = \frac{2}{\alpha}\left\|\overline{\mathbf{M}}^{-1}\right\| \quad \left[\left\{\|\mathbf{C}_1\|\big(\frac{4}{\alpha}\left\|\overline{\mathbf{M}}^{-1}\right\|\|\mathbf{C}_1\|+1\big)+L_1\Big(\frac{4}{\alpha}\left\|\overline{\mathbf{M}}^{-1}\right\| + 2\left\|(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{B}\mathbf{M}^{-1}\right\|\big(\frac{2}{\alpha}\left\|\overline{\mathbf{M}}^{-1}\right\|\|\mathbf{A}\|+1\big)\Big)\left\|\begin{pmatrix}\mathbf{u}\\ \widehat{\mathbf{x}}_\Gamma\end{pmatrix}\right\|_{L^2}\right\} + 2\left\|\begin{pmatrix}\widehat{\mathbf{z}}_1\\ \widehat{\boldsymbol{\mu}}_\Gamma\end{pmatrix}\right\|_{L^2}\right],$$

$$c_\kappa = \left\|(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{B}\mathbf{M}^{-1}\right\|\Big(\frac{2}{\alpha}\left\|\overline{\mathbf{M}}^{-1}\right\|\|\mathbf{A}\|+1\Big)$$
$$\left[\left\{\|\mathbf{C}_1\|\big(\frac{4}{\alpha}\left\|\overline{\mathbf{M}}^{-1}\right\|\|\mathbf{C}_1\|+1\big)+L_1\Big(\frac{4}{\alpha}\left\|\overline{\mathbf{M}}^{-1}\right\| + 2\left\|(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{B}\mathbf{M}^{-1}\right\|\big(\frac{2}{\alpha}\left\|\overline{\mathbf{M}}^{-1}\right\|\|\mathbf{A}\|+1\big)\Big)\left\|\begin{pmatrix}\mathbf{u}\\ \widehat{\mathbf{x}}_\Gamma\end{pmatrix}\right\|_{L^2}\right\} + 2\left\|\begin{pmatrix}\widehat{\mathbf{z}}_1\\ \widehat{\boldsymbol{\mu}}_\Gamma\end{pmatrix}\right\|_{L^2}\right].$$

Proof: As in the proof of Lemma 5.5.1, we introduce an auxiliary adjoint state $\widetilde{\boldsymbol{\mu}}_1 = (\widetilde{\boldsymbol{\lambda}}_1, \widetilde{\boldsymbol{\kappa}}_1)^T$ as the solution of

$$-\mathbf{E}_1\frac{d}{dt}\begin{pmatrix}\widetilde{\boldsymbol{\lambda}}_1(t)\\ \widetilde{\boldsymbol{\kappa}}_1(t)\end{pmatrix} = -\mathbf{S}_1\begin{pmatrix}\widetilde{\boldsymbol{\lambda}}_1(t)\\ \widetilde{\boldsymbol{\kappa}}_1(t)\end{pmatrix} - \mathbf{S}_{1\Gamma}\begin{pmatrix}\widehat{\boldsymbol{\lambda}}_\Gamma(t)\\ \widehat{\boldsymbol{\kappa}}_0(t)\end{pmatrix} - \begin{pmatrix}\mathbf{C}_1^T\\ \mathbf{F}_1^T\end{pmatrix}\widehat{\mathbf{z}}_1(t), \tag{5.73a}$$

$$\widetilde{\mathbf{q}}_1(t) = \mathbf{K}_1^T\widetilde{\boldsymbol{\lambda}}_1(t) + \mathbf{L}_1^T\widetilde{\boldsymbol{\kappa}}_1(t) - \mathbf{D}_1^T\widehat{\mathbf{z}}_1(t), \tag{5.73b}$$

$$\begin{pmatrix}\widetilde{\mathbf{q}}_{\lambda,\Gamma}(t)\\ \widetilde{\mathbf{q}}_{\kappa,\Gamma}(t)\end{pmatrix} = -\mathbf{S}_{1\Gamma}^T\begin{pmatrix}\widetilde{\boldsymbol{\lambda}}_1(t)\\ \widetilde{\boldsymbol{\kappa}}_1(t)\end{pmatrix} - \begin{pmatrix}\mathbf{0}\\ \mathbf{F}_0^T\end{pmatrix}\widehat{\mathbf{z}}_1(t), \tag{5.73c}$$

$$\mathbf{M}_1\widetilde{\boldsymbol{\lambda}}_1(T) = \boldsymbol{\lambda}_1^{(T)}, \tag{5.73d}$$

$$\mathbf{F}_1^T\widehat{\mathbf{z}}_1(T) = -\mathbf{B}_{11}\mathbf{M}_1^{-1}\boldsymbol{\lambda}_1^{(T)} - \mathbf{B}_{1\Gamma}\mathbf{M}_\Gamma(\theta)^{-1}\boldsymbol{\lambda}_\Gamma^{(T)}(\theta). \tag{5.73e}$$

Note that due to $\mathbf{F}_1 = \mathbf{0}$ the compatibility condition (5.43c) implies the compatibility condition (5.73e).

Moreover, since the second row block in $\mathbf{S}_{1\Gamma}^T$ is zero (cf. (5.32) and (5.35c)) and $\mathbf{F}_0 = \mathbf{0}$, we have

$$\widetilde{\mathbf{q}}_{\kappa,\Gamma}(t) \equiv 0. \tag{5.74}$$

The inputs for (5.73) and the reduced system (5.55) are the same and we can apply the balanced truncation error bound (5.27) to this subsystem. The balanced truncation error bound for this subsystem is

$$\left\| \begin{pmatrix} \widetilde{\mathbf{q}}_1 - \widehat{\mathbf{q}}_1 \\ \widetilde{\mathbf{q}}_{\lambda,\Gamma} - \widehat{\mathbf{q}}_{\lambda,\Gamma} \\ \widetilde{\mathbf{q}}_{\kappa,\Gamma} - \widehat{\mathbf{q}}_{\kappa,\Gamma} \end{pmatrix} \right\|_{L^2} \leq 2\Big(\sigma_{k+1} + \cdots + \sigma_n\Big) \left\| \begin{pmatrix} \widehat{\mathbf{z}}_1 \\ \widehat{\boldsymbol{\lambda}}_\Gamma \\ \widehat{\boldsymbol{\kappa}}_0 \end{pmatrix} \right\|_{L^2}. \tag{5.75}$$

We set $\mathbf{e}_{\boldsymbol{\lambda}} = (\boldsymbol{\lambda}_1 - \widetilde{\boldsymbol{\lambda}}_1, \boldsymbol{\lambda}_2 - \widehat{\boldsymbol{\lambda}}_2, \boldsymbol{\lambda}_\Gamma - \widehat{\boldsymbol{\lambda}}_\Gamma)^T$ and $\mathbf{e}_{\boldsymbol{\kappa}} = (\boldsymbol{\kappa}_1 - \widetilde{\boldsymbol{\kappa}}_1, \boldsymbol{\kappa}_2 - \widehat{\boldsymbol{\kappa}}_2, \boldsymbol{\kappa}_0 - \widehat{\boldsymbol{\kappa}}_0)^T$. Observing (5.43),(5.45),(5.47), (5.56) as well as (5.73),(5.58),(5.60), (5.74) it follows that

$$-\mathbf{E}_1 \frac{d}{dt} \begin{pmatrix} \mathbf{e}_{\boldsymbol{\lambda}}(t) \\ \mathbf{e}_{\boldsymbol{\kappa}}(t) \end{pmatrix} = -\mathbf{S}(\theta) \begin{pmatrix} \mathbf{e}_{\boldsymbol{\lambda}}(t) \\ \mathbf{e}_{\boldsymbol{\kappa}}(t) \end{pmatrix} + \begin{pmatrix} \mathbf{g}_1(t) \\ \mathbf{g}_2(t) \end{pmatrix}, \ t \in (0, T], \tag{5.76a}$$

$$\mathbf{M}(\theta)\mathbf{e}_{\boldsymbol{\lambda}}(T) = \mathbf{0}, \tag{5.76b}$$

where

$$\mathbf{g}_1(t) = \begin{pmatrix} -\mathbf{C}_1^T(\mathbf{z}_1 - \widehat{\mathbf{z}}_1) \\ -\nabla_{\mathbf{v}_2}\boldsymbol{\ell}(\mathbf{x}_2, \mathbf{x}_\Gamma, t, \theta) + \nabla_{\widehat{\mathbf{v}}_2}\boldsymbol{\ell}(\widehat{\mathbf{x}}_2, \widehat{\mathbf{x}}_\Gamma, t, \theta) \\ \widetilde{\mathbf{q}}_{\lambda,\Gamma} - \widehat{\mathbf{q}}_{\lambda,\Gamma} - \nabla_{\mathbf{v}_\Gamma}\boldsymbol{\ell}(\mathbf{x}_2, \mathbf{x}_\Gamma, t, \theta) + \nabla_{\widehat{\mathbf{v}}_\Gamma}\boldsymbol{\ell}(\widehat{\mathbf{x}}_2, \widehat{\mathbf{x}}_\Gamma, t, \theta) \end{pmatrix},$$

$$\mathbf{g}_2(t) = \begin{pmatrix} -\mathbf{F}_1^T(\mathbf{z}_1 - \widehat{\mathbf{z}}_1) \\ \mathbf{0} \\ -\mathbf{F}_0^T(\mathbf{z}_1 - \widehat{\mathbf{z}}_1) \end{pmatrix}.$$

Applying Theorem 5.1.1 to (5.76) and assumption ($\mathbf{A_3}$) we obtain

$$\left\| \begin{pmatrix} \boldsymbol{\lambda}_1 - \widetilde{\boldsymbol{\lambda}}_1 \\ \boldsymbol{\lambda}_2 - \widehat{\boldsymbol{\lambda}}_2 \\ \boldsymbol{\lambda}_\Gamma - \widehat{\boldsymbol{\lambda}}_\Gamma \end{pmatrix} \right\|_{L^2} \leq \frac{2}{\alpha} \left\| \overline{\mathbf{M}}^{-1} \right\| \Big(\|\mathbf{C}_1\| \|\mathbf{z}_1 - \widehat{\mathbf{z}}_1\|_{L^2} + \left\| \begin{pmatrix} \widetilde{\mathbf{q}}_{\lambda,\Gamma} - \widehat{\mathbf{q}}_{\lambda,\Gamma} \\ \widetilde{\mathbf{q}}_{\kappa,\Gamma} - \widehat{\mathbf{q}}_{\kappa,\Gamma} \end{pmatrix} \right\|_{L^2}$$
$$+ L_1 \left\| \begin{pmatrix} \mathbf{v}_2 - \widehat{\mathbf{v}}_2 \\ \mathbf{v}_\Gamma - \widehat{\mathbf{v}}_\Gamma \end{pmatrix} \right\|_{L^2} + L_1 \left\| \begin{pmatrix} \mathbf{p}_2 - \widehat{\mathbf{p}}_2 \\ \mathbf{p}_0 - \widehat{\mathbf{p}}_0 \end{pmatrix} \right\|_{L^2} \Big). \tag{5.77}$$

Inequality (5.72a) follows using (5.75), (5.77) and Lemma 5.5.1.

Application of Theorem 5.1.1 to (5.76) and using Assumption ($\mathbf{A_3}$) also yields

$$\left\| \begin{pmatrix} \boldsymbol{\kappa}_1 - \widetilde{\boldsymbol{\kappa}}_1 \\ \boldsymbol{\kappa}_2 - \widehat{\boldsymbol{\kappa}}_2 \\ \boldsymbol{\kappa}_0 - \widehat{\boldsymbol{\kappa}}_0 \end{pmatrix} \right\|_{L^2} \leq \left\| (\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}\mathbf{B}\mathbf{M}^{-1} \right\| \Big(\frac{2}{\alpha} \left\| \overline{\mathbf{M}}^{-1} \right\| \|\mathbf{A}\| + 1 \Big)$$
$$\Big(\|\mathbf{C}_1\| \, \|\mathbf{z}_1 - \widehat{\mathbf{z}}_1\|_{L^2} + \left\| \begin{pmatrix} \widetilde{\mathbf{q}}_{\lambda,\Gamma} - \widehat{\mathbf{q}}_{\lambda,\Gamma} \\ \widetilde{\mathbf{q}}_{\kappa,\Gamma} - \widehat{\mathbf{q}}_{\kappa,\Gamma} \end{pmatrix} \right\|_{L^2}$$
$$+ L_1 \left\| \begin{pmatrix} \mathbf{v}_2 - \widehat{\mathbf{v}}_2 \\ \mathbf{v}_\Gamma - \widehat{\mathbf{v}}_\Gamma \end{pmatrix} \right\|_{L^2} + L_1 \left\| \begin{pmatrix} \mathbf{p}_2 - \widehat{\mathbf{p}}_2 \\ \mathbf{p}_0 - \widehat{\mathbf{p}}_0 \end{pmatrix} \right\|_{L^2} \Big)$$
$$+ C \left(\| \frac{d}{dt} \mathbf{F}_1^T (\mathbf{z}_1 - \widehat{\mathbf{z}}_1) \|_{L^2} + \| \frac{d}{dt} \mathbf{F}_0^T (\mathbf{z}_1 - \widehat{\mathbf{z}}_1) \|_{L^2} \right). \tag{5.78}$$

Since $\mathbf{F}_1 = \mathbf{0}$, $\mathbf{F}_0 = \mathbf{0}$, inequality (5.72b) follows using (5.75), (5.78) and Lemma 5.5.1. □

The preceding two lemmas lead to the following bound for the gradients of the objective functions for the full order and the reduced order problem.

Theorem 5.5.3 *If* $(\mathbf{A_1}) - (\mathbf{A_4})$*, and* (5.26) *are valid, then there exists* $c_g > 0$ *such that*

$$\|\nabla J(\theta) - \nabla \widehat{J}(\theta)\| \;\leq\; c_g \,\Big(\sigma_{k+1} \;+\; \cdots \;+\; \sigma_n\Big).$$

where

$$c_g = \left\{ TL_1 + \gamma \left\| \begin{pmatrix} \widehat{\boldsymbol{\mu}}_2 \\ \widehat{\boldsymbol{\mu}}_\Gamma \end{pmatrix} \right\| + \gamma\ (c_\lambda + c_\kappa)\ (\sigma_{k+1} \;+\; \cdots \;+\; \sigma_n) \right\} (c_v + c_p)$$
$$+ \gamma\ (c_\lambda + c_\kappa) \left\{ \left\| \begin{pmatrix} \widehat{\mathbf{x}}_2 \\ \widehat{\mathbf{x}}_\Gamma \end{pmatrix} \right\| + \|\mathbf{u}\| \right\},$$

with $c_v, c_p, c_\lambda, c_\kappa$ *are the constants specified in Lemma 5.5.1 and 5.5.2.*

Proof: The gradients $\nabla J(\theta)$ and $\nabla \widehat{J}(\theta)$ applied to an arbitrary $\widetilde{\theta}$ are given by

$$\nabla J(\theta)^T \widetilde{\theta} = \int_0^T \nabla_\theta \boldsymbol{\ell}(\mathbf{v}_2, \mathbf{p}_2, \mathbf{v}_\Gamma, \mathbf{p}_0, t, \theta)^T \tilde{\theta}\ dt$$
$$+ \int_0^T \begin{pmatrix} \boldsymbol{\mu}_2(t) \\ \boldsymbol{\lambda}_\Gamma(t) \end{pmatrix}^T \begin{pmatrix} (D_\theta \mathbf{P}_2(\theta)\tilde{\theta})\, \mathbf{x}_2(t) - (D_\theta \mathbf{N}_2(\theta)\tilde{\theta})\, \mathbf{u}(t) \\ (D_\theta \mathbf{P}_\Gamma(\theta)\tilde{\theta})\, \mathbf{x}_\Gamma(t) - (D_\theta \mathbf{N}_\Gamma(\theta)\tilde{\theta})\, \mathbf{u}(t) \end{pmatrix} dt,$$

$$\nabla \widehat{J}(\theta)^T \widetilde{\theta} = \int_0^T \nabla_\theta \boldsymbol{\ell}(\widehat{\mathbf{v}}_2, \widehat{\mathbf{p}}_2, \widehat{\mathbf{v}}_\Gamma, \widehat{\mathbf{p}}_0, t, \theta)^T \tilde{\theta}\ dt$$
$$+ \int_0^T \begin{pmatrix} \widehat{\boldsymbol{\mu}}_2(t) \\ \widehat{\boldsymbol{\lambda}}_\Gamma(t) \end{pmatrix}^T \begin{pmatrix} (D_\theta \mathbf{P}_2(\theta)\tilde{\theta})\, \widehat{\mathbf{x}}_2(t) - (D_\theta \mathbf{N}_2(\theta)\tilde{\theta})\, \mathbf{u}(t) \\ (D_\theta \mathbf{P}_\Gamma(\theta)\tilde{\theta})\, \widehat{\mathbf{x}}_\Gamma(t) - (D_\theta \mathbf{N}_\Gamma(\theta)\tilde{\theta})\, \mathbf{u}(t) \end{pmatrix} dt.$$

The estimate now follows from Lemmas 5.5.1 and 5.5.2 and Assumption $(\mathbf{A_3})$. □

Under the convexity assumption (5.63) the bound (5.64) combined with Theorem 5.5.3 gives the desired bound for the error in the solutions computed using the full and the reduced order model.

Corollary 5.5.4 *If the assumptions of Theorem 5.5.3 hold and the convexity assumption* (5.63) *is valid, then there exists* $c_g > 0$ *as specified in Theorem 5.5.3 such that*

$$\|\theta^* - \widehat{\theta}^*\| \;\leq\; \frac{c_g}{\kappa}\Big(\sigma_{k+1} \;+\; \cdots \;+\; \sigma_n\Big).$$

5.6 Numerical examples

5.6.1 Shape optimization of capillary barriers in microfluidic biochips

We saw in Chapter 2 that the induced fluid flow in surface acoustic wave driven microfluidic biochips is taken care of by an homogenization approach so that the resulting flow pattern, called acoustic streaming, can be described by instationary compressible Stokes flow. We refer to [10, 12, 16, 49] for details. The performance of these biochips can be significantly improved by an optimal design of the walls of the microchannels and the capillary barriers between the channels and the reservoirs.

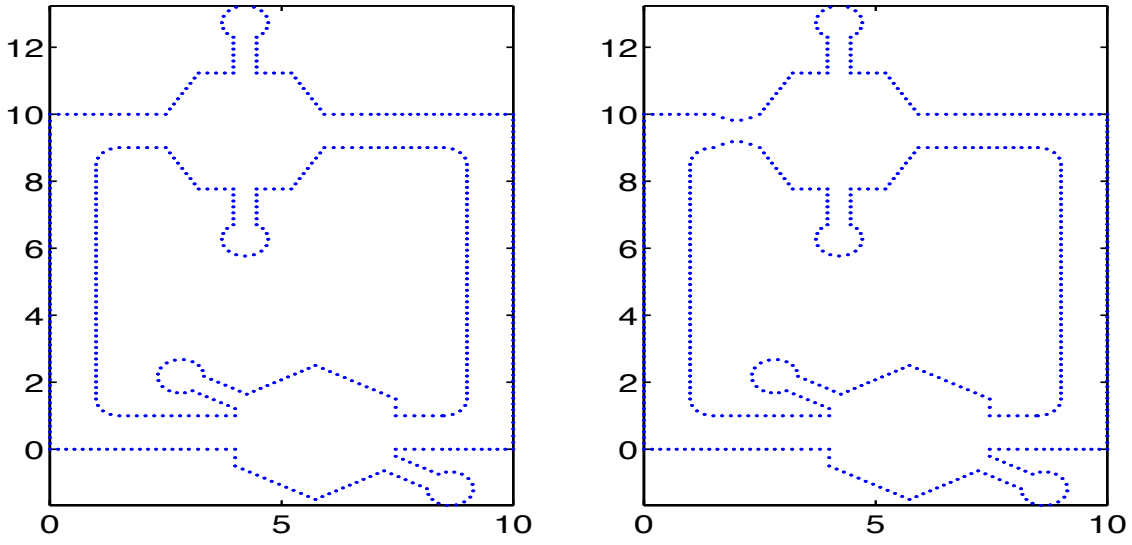


Figure 5.1: The reference domain Ω_{ref} (left) and the optimal domain (right)

As a first step we consider incompressible Stokes flow in a network of microchannels and reservoirs on top of a microfluidic biochip with capillary barriers between the channels and the reservoirs that are designed to guarantee a precise filling of the reservoirs with the DNA or protein probes. The objective is twofold: Firstly, we want to design the walls of the barriers in such a way that a desired velocity profile $\mathbf{v}^d$ is attained and secondly, we want to minimize the vorticity $\nabla \times \mathbf{v}$ in some specific part of the network.

The computational domain $\Omega \subset \mathbb{R}^2$ is displayed in Figure 5.1. It is decomposed into subdomains $\Omega_1 = \Omega \setminus \Omega_2$, and $\Omega_2 = \{1.5, 2.5\} \times \{9, 10\}$. The boundary $\partial\Omega$ is decomposed into $\Gamma_{\text{in}} = \{0\} \times (9, 10), \Gamma_{\text{out}} = \{10\} \times (0, 1)$, and $\Gamma_{\text{lat}} = \partial\Omega \setminus (\Gamma_{\text{in}} \cup \Gamma_{\text{out}})$.
The data of the problem is chosen as follows. Assume $\mathbf{f} = 0$ in $\Omega \times (0, T)$, a Poiseuille velocity profile $\mathbf{v}_{\text{in}}((x_1, x_2), t) = 4(x_2 - 9)(10 - x_2)(1 - \frac{0.8}{15}t)\text{sin(t)}$ on $\Gamma_{\text{in}} \times (0, T)$, outflow boundary conditions on $\Gamma_{\text{out}} \times (0, T)$, and no-slip conditions on $\Gamma_{\text{lat}} \times (0, T)$. The objective is to design the shape of the top $\Gamma_{2,T}$ and the bottom $\Gamma_{2,B}$ of $\partial\Omega_2$ in such a way that a prescribed velocity profile $\mathbf{v}^d$ is achieved in $\Omega_2 \times (0, T)$ and the vorticity is minimized in Ω_{obs} (four bulb-shaped structures in Figure 5.1). We use a parametrization $\Omega_2(\theta)$ of Ω_2 by means of the Bézier control points $\theta \in \mathbb{R}^k, k = k_T + k_B$, of Bézier curve representations of $\Gamma_{2,T}$ and $\Gamma_{2,B}$, where k_T and k_B refer to the number of control points for $\Gamma_{2,T}$ and $\Gamma_{2,B}$, respectively. The shape optimization problem amounts to the minimization of

$$J(\theta) = \int_0^T \int_{\Omega_{\text{obs}}} |\nabla \times \mathbf{v}(x,t)|^2 dxdt + \int_0^T \int_{\Omega_2(\theta)} |\mathbf{v}(x,t) - \mathbf{v}^d(x,t)|^2 dxdt$$

subject to the Stokes equations

$$\begin{aligned}
\mathbf{v}_t(x,t) - \mu\Delta\mathbf{v}(x,t) + \nabla p(x,t) &= \mathbf{f}(x,t), && \text{in } \Omega(\theta)\times(0,T),\\
\nabla\cdot\mathbf{v}(x,t) &= 0, && \text{in } \Omega(\theta)\times(0,T),\\
\mathbf{v}(x,t) &= \mathbf{v}_{\text{in}}(x,t) && \text{on } \Gamma_{\text{in}}\times(0,T),\\
\mathbf{v}(x,t) &= \mathbf{0} && \text{on } \Gamma_{\text{lat}}\times(0,T),\\
(\mu\nabla\mathbf{v}(x,t) - p(x,t)I)\mathbf{n} &= \mathbf{0} && \text{on } \Gamma_{\text{out}}\times(0,T),\\
\mathbf{v}(x,0) &= \mathbf{0} && \text{in } \Omega(\theta)
\end{aligned}$$

and design parameter constraints

$$\theta^{min} \le \theta \le \theta^{max},$$

where $\mu = 1/50$ and $T = 15$. The bounds $\theta^{min}, \theta^{max}$ on the design parameters are chosen such that the design constraints are never active in this example. We use $k_T = 6, k_B = 6$ Bézier control points to specify the top and the bottom boundary of the variable subdomain $\Omega_2(\theta)$ with the respective first and last control points being fixed. The desired velocity $\mathbf{v}^d$ is computed by specifying the optimal parameter θ^* and solving the state equation on $\Omega(\theta^*)$. The optimal domain $\Omega(\theta^*)$ is shown in Figure 5.1.

We consider a geometrically conforming simplicial triangulation $\mathcal{T}_h(\Omega)$ of the reference that aligns with the decomposition of Ω into the subdomains Ω_1 and Ω_2 as well as the respective boundaries. The discretization in space is taken care of by P2-P1 Taylor-Hood elements. For details we refer to Appendix B. For $D \subseteq \overline{\Omega}$, we denote by $\mathcal{N}_{\mathbf{v},h}(D), \mathcal{N}_{p,h}(D)$ the set of velocity and pressure nodal points in D. We use the domain decomposition methodology as before and set $N^{(\nu)}_{\mathbf{v},dof} = \text{card}(\mathcal{N}_{\mathbf{v},h}(\bar{\Omega}_\nu \setminus \Gamma_{I,\mathbf{v}}))$, $\nu = 1, 2$, and $N^{\Gamma_{I,\mathbf{v}}}_{dof} := \text{card}(\mathcal{N}_{\mathbf{v},h}(\Gamma_{I,\mathbf{v}}))$ so that $N_{\mathbf{v},dof} = N^{(1)}_{\mathbf{v},dof} + N^{(2)}_{\mathbf{v},dof} + N^{\Gamma_{I,\mathbf{v}}}_{\mathbf{v},dof}$ is the total number of velocity degrees of freedom. Similarly, $N_{p,dof} = N^{(1)}_{p,dof} + N^{(2)}_{p,dof}$ is the total number of pressure

degrees of freedom.

We use automatic differentiation [54, 100] to compute the derivatives with respect to the design variables θ. The semi-discretized optimization problems are solved using a projected BFGS method with Armijo line search [71]. The optimization algorithm is terminated when the norm of projected gradient is less than $\epsilon = 10^{-4}$. The results in Figure 5.1(right) and Figures 5.2-5.4 and Table 5.2 were generated using the finest grid i.e., grid 4 with $N_{\mathbf{v},dof} = 16806$. We will explain them as we go along.

We use the multishift ADI method [62] to solve the projected Lyapunov equations. We use four shifts in the ADI method which were computed as in [62]. Figure 5.2 shows the largest Hankel singular values. For the model reduction, we select those Hankel singular values σ_j, with $\sigma_j \geq 10^{-3}\sigma_1$. The threshold $10^{-3}\sigma_1$ is indicated by the solid line in Figure 5.2 (left). In this case only twenty-nine Hankel singular values and corresponding singular vectors determine the reduced order model for the velocities in Ω_1.

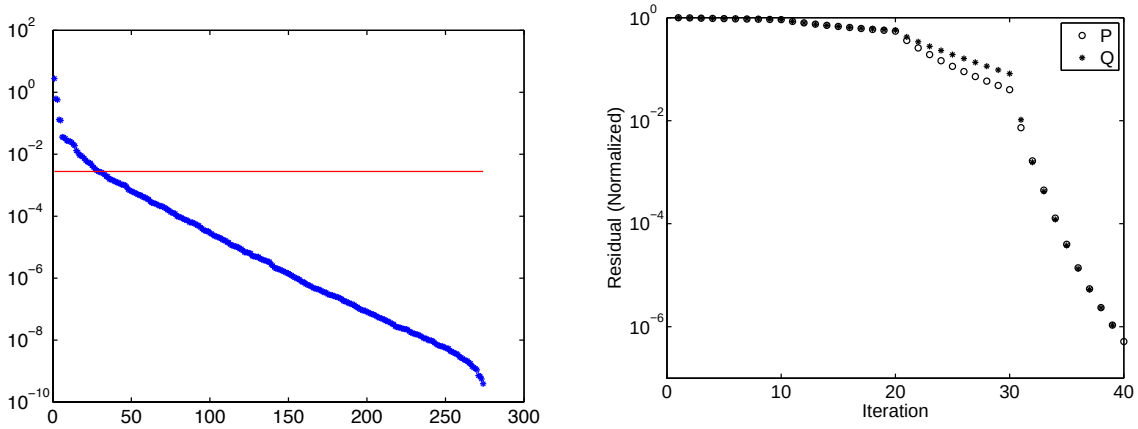


Figure 5.2: The left plot shows the largest Hankel singular values and the threshold $10^{-3}\sigma_1$. The right plot shows normalized residuals [62] generated by the multishift ADI for the approximate solution of the controllability Lyapunov equation (o) and of the observability Lyapunov equation ($*$)

In order to test our model reduction routine we compare full and the reduced order

semi-discretized integrand

$$\ell(\theta_0, t) = \int_{\Omega_{\text{obs}}} |\nabla \times \mathbf{v}(x,t)|^2 dx + \int_{\Omega_2(\theta_0)} |\mathbf{v}(x,t) - \mathbf{v}^d(x,t)|^2 dx$$

as a function of time t for the initial value of the design parameter θ_0. Note that $J(\theta) = \int_0^T \ell(\theta)\, dt$. Figure 5.3 displays the results obtained. The full and reduced order models are both in excellent agreement, which is expected given the theoretical a priori error bound for the balanced truncation model reduction.

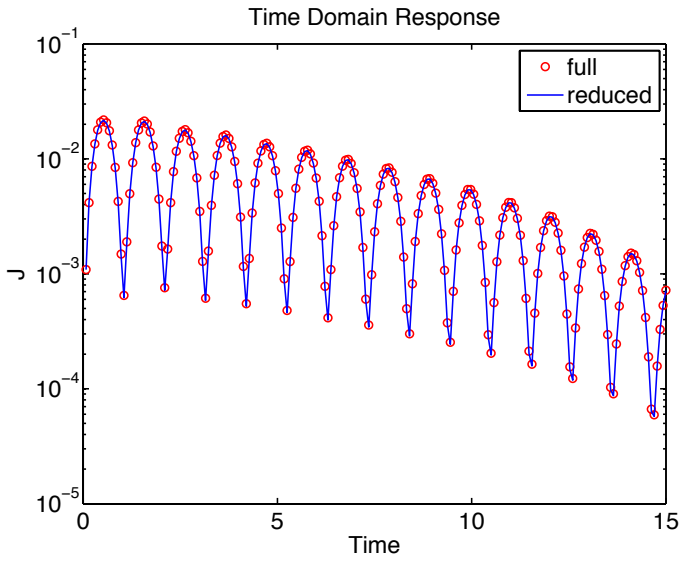


Figure 5.3: Time response for the full (circle) and the reduced (solid line) integrand $\ell(\theta_0)$ for the initial configuration (= reference domain Figure 5.1(left))

Table 5.1 displays the sizes of the reduced and the full order problems (in degrees of freedom (dof)) for an initial coarse grid and three levels of refinement. We observe that the size of the reduced order model is nearly independent of the grid size.

The optimal shape parameters θ^* and $\widehat{\theta}^*$ computed by minimizing the full and the reduced order model, respectively, are shown in Table 5.2. For the finest grid problem, the error between full and the reduced order model solutions is $\|\theta^* - \widehat{\theta}^*\|_2 = 8.0751 \cdot 10^{-3}$.

grid number	m	$N^{(1)}_{\mathbf{v},dof}$	$N^{(1)}_{\widehat{\mathbf{v}},dof}$	$N_{\mathbf{v},dof}$	$N_{\widehat{\mathbf{v}},dof}$
1	149	4752	23	4862	133
2	313	7410	25	7568	183
3	361	11474	26	11700	252
4	537	16472	29	16806	363

Table 5.1: The number m of observations in Ω_1, the numbers $N^{(1)}_{\mathbf{v},dof}, N^{(1)}_{\widehat{\mathbf{v}},dof}$ of velocity dof in Ω_1 (full and reduced order), and the numbers $N_{\mathbf{v},dof}, N_{\widehat{\mathbf{v}},dof}$ of velocity dof in Ω (full and reduced order) for four discretizations

θ^*	(9.8987, 9.7510, 9.7496, 9.8994, 9.0991, 9.2499, 9.2504, 9.0989)
$\widehat{\theta}^*$	(9.9026, 9.7498, 9.7484, 9.9021, 9.0940, 9.2514, 9.2511, 9.0956)

Table 5.2: Optimal shape parameters θ^* and $\widehat{\theta}^*$ (rounded to 5 digits) computed by minimizing the full and the reduced order model

The convergence histories of the projected BFGS algorithm applied to the full and the reduced order problems are shown in Figure 5.4. Except for the final iterations, the convergence behavior of the optimization algorithm applied to the full and the reduced order problems is nearly identical.

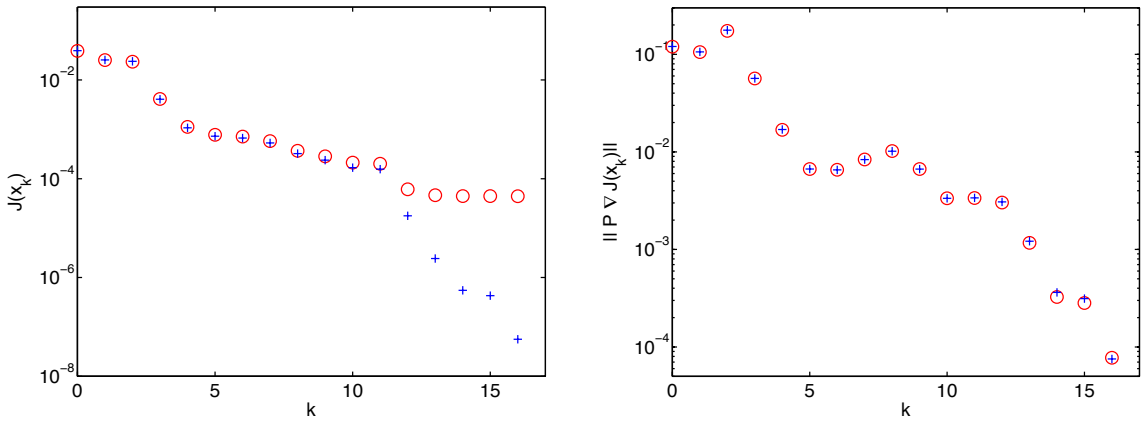


Figure 5.4: The convergence histories of the projected BFGS algorithm applied to the full (+) and the reduced (o) order problems. Left: Objective functionals. Right: Projected gradients.

5.6.2 Shape optimization of aorto-coronaric bypass

In this subsection we consider yet another application of the techniques developed in previous sections to a challenging real life problem. The problem considered is motivated by [2], where an optimal configuration of a 2D aorto-coronaric bypass is obtained and the optimal shape obtained looks like a Taylor patch by minimizing the vorticity in the region of interest.

Consider an arterial bypass bridge configuration shown in Figure 5.5. The dotted region marked (left) and shown separately in Figure 5.5 (right) represents the geometry of the complete anastomosis. The angle between the bypass bridge and the main vessel is called the graft angle. In our optimization problem only the top curved boundary in Figure 5.5 (right) is subject to change during optimization and is the characteristic Bézier curve, characterized by the control points θ.

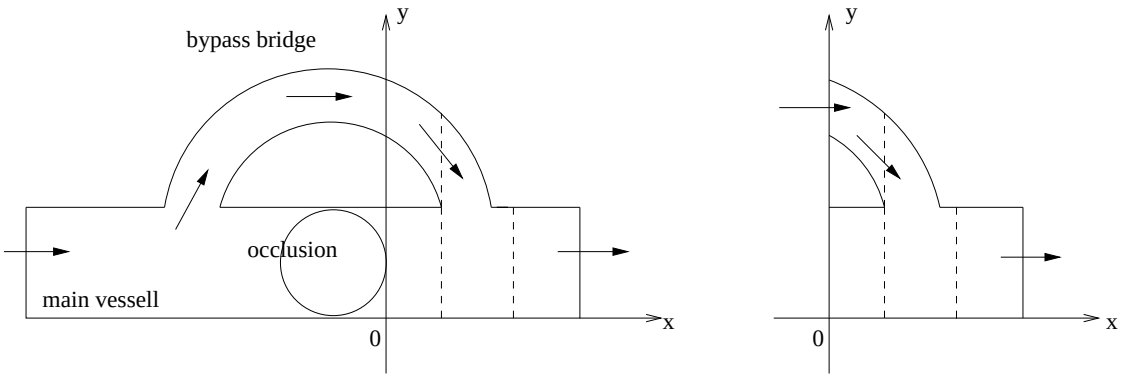


Figure 5.5: Arterial bypass bridge configuration (left), interesting region with respect to shape optimization (right), only top curved boundary (right) is subject to change during optimization.

The computational domain $\Omega \subset \mathbb{R}^2$ is displayed in Figure 5.6. Note we assume that the graft angle is equal to zero. It is decomposed into subdomains $\Omega_1 = \Omega \setminus \Omega_2$, and $\Omega_2 = \{-1.0, 0\} \times \{0, 1.5\}$. The boundary $\partial\Omega$ is decomposed into $\Gamma_{\text{in}} = \{-3\} \times (1, 2), \Gamma_{\text{out}} =$

$\{4\} \times (0,1) \cup \Gamma_{\text{obs}}$, and $\Gamma_{\text{lat}} = \partial\Omega \setminus (\Gamma_{\text{in}} \cup \Gamma_{\text{out}})$, where $\Gamma_{\text{obs}} = \{-3\} \times (0, 0.8)$ is the occlusion wall.

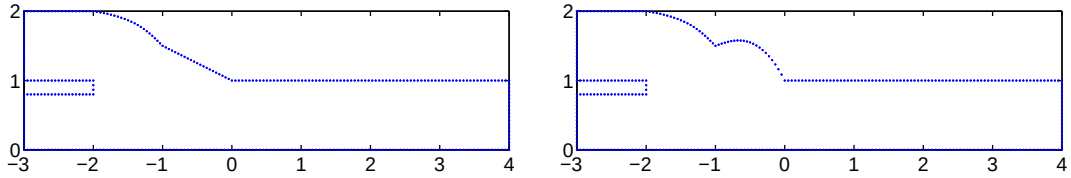


Figure 5.6: Reference (left) and optimal (right) aorto-coronaric bypass

The data of the problem is chosen as in [2]. Assume $\mathbf{f} = 0$ in $\Omega \times (0,T)$, a Poiseuille velocity profile $\mathbf{v}_{\text{in}}((x_1,x_2),t) = 0.475(x_2-1)(2-x_2)(1-t)\text{sin}(3\text{t})$ which is pulsatile in nature with a period $T = 1s$ (heart beat) on $\Gamma_{\text{in}} \times (0,T)$, outflow boundary conditions on $\Gamma_{\text{out}} \times (0,T)$, and no-slip conditions on $\Gamma_{\text{lat}} \times (0,T)$. The input velocity $\mathbf{v}^{\text{in}}$ is chosen so that the Reynolds number Re has the order 10^3, the mean Reynolds number is 1250, the maximum is 2500. For details we refer to [2].

The objective is to design the shape of the top $\Gamma_{2,T}$ with $\Gamma_{2,B}$ fixed of $\partial\Omega_2$ in such a way that a prescribed velocity profile $\mathbf{v}^d$ is achieved in $\Omega_2 \times (0,T)$ and the velocity is equal to zero on Γ_{obs}. We use a parametrization $\Omega_2(\theta)$ of Ω_2 by means of the Bézier control points $\theta \in \mathbb{R}^k, k = k_T + k_B$, of Bézier curve representations of $\Gamma_{2,T}$ and $\Gamma_{2,B}$, where k_T and k_B refer to the number of control points for $\Gamma_{2,T}$ and $\Gamma_{2,B}$, respectively. The shape optimization problem amounts to the minimization of

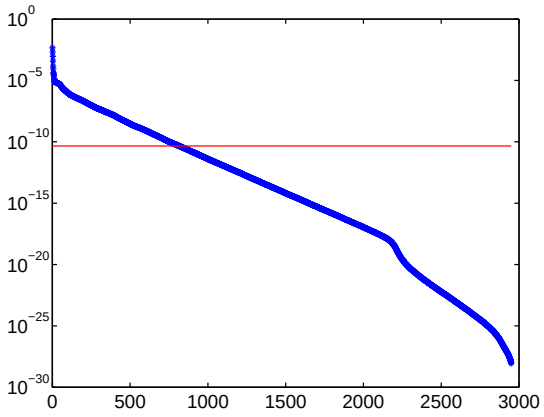
$$J(\theta) = \int_0^T \int_{\Gamma_{\text{obs}}} |\mathbf{v}(x,t)|^2 dxdt + \int_0^T \int_{\Omega_2(\theta)} |\mathbf{v}(x,t) - \mathbf{v}^d(x,t)|^2 dxdt$$

subject to the Stokes equations as in the previous subsection and design parameter constraints

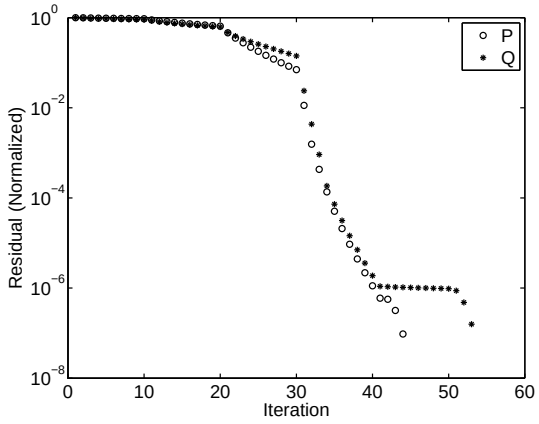
$$\theta^{min} \leq \theta \leq \theta^{max},$$

where $\mu = 4 \cdot 10^{-6} m^2/s$ is the blood kinematic viscosity. The bounds $\theta^{min}, \theta^{max}$ on the design parameters are chosen such that the design constraints are never active in this example. We use $k_T = 5, k_B = 1$ Bézier control points to specify the top and the bottom boundary of the variable subdomain $\Omega_2(\theta)$ with the respective first and last control points being fixed. The desired velocity $\mathbf{v}^d$ is computed by specifying the optimal parameter θ^* (which is chosen according to the optimal shape obtained in [2]) and solving the state equation on $\Omega(\theta^*)$. The optimal domain $\Omega(\theta^*)$ is shown in Figure 5.6.

Again automatic differentiation is used to compute the derivatives with respect to the design variables θ. The semi-discretized optimization problems are solved using a projected BFGS method with Armijo line search [71]. The optimization algorithm is terminated when the norm of projected gradient is less than $\epsilon = 10^{-6}$.



Hankel singular values



Convergence of the multishift ADI

Figure 5.7: The left plot shows the largest Hankel singular values and the threshold $10^{-8}\sigma_1$. The right plot shows normalized residuals [62] generated by the multishift ADI for the approximate solution of the controllability Lyapunov equation (o) and of the observability Lyapunov equation ($*$).

We use the multishift ADI method [62] to solve the projected Lyapunov equations. We use four shifts in the ADI method which were computed as in [62]. Figure 5.7 shows the

largest Hankel singular values. For the model reduction, we select those Hankel singular values σ_j, with $\sigma_j \geq 10^{-8}\sigma_1$. The threshold $10^{-8}\sigma_1$ is indicated by the solid line in Figure 5.7 (left).

Table 5.3 displays the sizes of the reduced and the full order problems (in Degrees of Freedom (dof)).

grid number	m	$N^{(1)}_{\mathbf{v},dof}$	$N^{(1)}_{\widehat{\mathbf{v}},dof}$	$N_{\mathbf{v},dof}$	$N_{\widehat{\mathbf{v}},dof}$
1	251	27554	815	32226	5487

Table 5.3: The number m of observations in Ω_1, the numbers $N^{(1)}_{\mathbf{v},dof}, N^{(1)}_{\widehat{\mathbf{v}},dof}$ of velocity dof in Ω_1 (full and reduced order), and the numbers $N_{\mathbf{v},dof}, N_{\widehat{\mathbf{v}},dof}$ of velocity dof in Ω (full and reduced order)

The optimal shape parameters θ^* and $\widehat{\theta}^*$ computed by minimizing the full and the reduced order model, respectively, are shown in Table 5.4. The error between full and the reduced order model solutions is $\|\theta^* - \widehat{\theta}^*\|_2 = 4.5 \cdot 10^{-3}$.

θ^*	(1.00, 1.2148, 1.2668, 1.3646, 1.00, 0.00)
$\widehat{\theta}^*$	(1.00, 1.2179, 1.2675, 1.3679, 1.00, 0.00)

Table 5.4: Optimal shape parameters θ^* and $\widehat{\theta}^*$ (rounded to 5 digits) computed by minimizing the full and the reduced order model

The convergence histories of the projected BFGS algorithm applied to the full and the reduced order problems are shown in Figure 5.8. The convergence behavior of the optimization algorithm applied to the full and the reduced order problems is nearly identical.

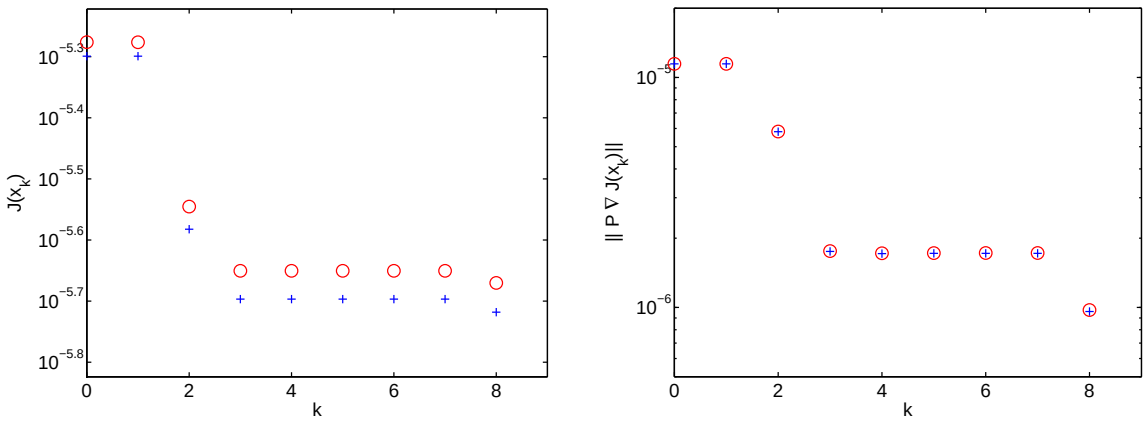


Figure 5.8: The convergence histories of the projected BFGS algorithm applied to the full (+) and the reduced (o) order problems. Left: Objective functionals. Right: Projected gradients.

5.7 Concluding remarks

In this chapter, we have integrated domain decomposition and balanced truncation model reduction for the numerical solution of a class of shape optimization problems governed by the Stokes equations. This approach can be applied when only small part of the overall domain can be modified by the optimization. Our approach leads to a reduced optimization problem with the same structure as the original one, but of potentially much smaller dimension. We have derived an estimate for the error between the solution of the original optimization problem and the solution of the reduced problem. The estimate is largely determined by the balanced truncation error estimate. The approach can be easily extended to shape optimization problems governed by the Oseen equations or the linearized Navier-Stokes equations, linearized around suitable steady flows.

CHAPTER 6

Reduced order modeling based shape optimization of microfluidic biochips

In this chapter we provide a detailed documentation of simulation results for the microfluidic biochips that illustrate both the validity of our model as well as the feasibility of the model reduction based optimization.

Using the notation from previous chapters and denoting by θ^* and $\hat{\theta}^*$ the optimal designs obtained for the full and the reduced order model, respectively, under the assumptions that J is strictly convex, J_1 does not depend explicitly on the pressure p, and some further assumptions, Corollary 5.5.4 gives

$$\|\theta^* - \hat{\theta}^*\| \leq \frac{c_g}{\kappa}\Big(\sigma_{k+1} + \cdots + \sigma_n\Big), \tag{6.1}$$

(Note: The error estimates derived in Chapter 5 hold for biochip systems as well) where $\sigma_1 \geq \ldots \geq \sigma_k > \sigma_{k+1} + \ldots + \sigma_n \geq 0$ are the Hankel singular values provided by BTMR for the Ω_1 optimality subsystem and c_g is the constant from Theorem 5.5.3.

6.1 Numerical results

In actual numerical simulations of acoustic streaming, we consider (2.12a)-(2.12e) in dimensionless form according to

$$\frac{VT}{L}\frac{d\mathbf{v}_1}{dt} - \frac{VT^2}{L^3}(\tilde{\nu}_1\Delta\mathbf{v}_1 + \tilde{\nu}_2\nabla(\nabla\cdot\mathbf{v}_1)) + \nabla p_1 = \mathbf{0} \qquad \text{in } \Omega\times(0,T_1], \tag{6.2a}$$

$$\frac{L^3}{c_0^2VT^3}\frac{dp_1}{dt} + \nabla\cdot\mathbf{v}_1 = 0 \qquad \text{in } \Omega\times(0,T_1], \tag{6.2b}$$

$$\frac{VT}{L}\mathbf{v}_1 = \frac{d\mathbf{u}}{dt} \quad \text{on } \Gamma_D\times(0,T_1] \quad , \quad \boldsymbol{\sigma}_1\mathbf{n} = \mathbf{0} \qquad \text{on } \Gamma_N\times(0,T_1], \tag{6.2c}$$

$$\mathbf{v}_1(\cdot,0) = \mathbf{0} \quad , \quad p_1(\cdot,0) = 0 \qquad \text{in } \Omega, \tag{6.2d}$$

where the parameters and their meanings are shown in Table 6.1. Similarly, the dimensionless form of (2.18a)-(2.18e) is as follows:

$$\frac{VT}{L}\frac{d\mathbf{v}_2}{dt} - \frac{VT^2}{L^3}(\tilde{\nu}_1\Delta\mathbf{v}_2 + \tilde{\nu}_2\nabla(\nabla\cdot\mathbf{v}_2)) + \nabla p_2 = \langle\mathbf{f}_v\rangle \quad \text{in } \Omega\times\mathbb{R}_+, \tag{6.3a}$$

$$\nabla\cdot\mathbf{v}_2 = \langle f_p\rangle \quad \text{in } \Omega\times\mathbb{R}_+, \tag{6.3b}$$

$$\mathbf{v}_2 = \langle(\nabla\mathbf{v}_1)\mathbf{u}\rangle \quad \text{on } \Gamma_D\times\mathbb{R}_+ \quad , \quad \boldsymbol{\sigma}_2\mathbf{n} = \mathbf{0} \qquad \text{on } \Gamma_N\times\mathbb{R}_+, \tag{6.3c}$$

$$\mathbf{v}_2(\cdot,0) = \mathbf{0} \quad , \quad p_2(\cdot,0) = 0 \qquad \text{in } \Omega, \tag{6.3d}$$

where

$$\mathbf{f}_v := -\frac{V^2T^2}{L^2}(\nabla\cdot\mathbf{v}_1)\mathbf{v}_1 + (\nabla\mathbf{v}_1)\mathbf{v}_1 \ , \ f_p := -\frac{L^2}{T^2c_0^2}\nabla\cdot(p_1\mathbf{v}_1).$$

We note that $\langle\mathbf{f}_v\rangle$ represents the time averaged sound velocity in the fluid which is commonly referred to as the effective force.

Parameter	Value and units	Description
V	$1.0 \cdot 10^{-1}$ m/s	Dimensionless velocity scale
L	$1.0 \cdot 10^{-7}$ m	Dimensionless length scale
T	$1.0 \cdot 10^{-8}$ s	Dimensionless time scale
f	$1.0 \cdot 10^{8}$ Hz	Frequency of the SAW device
c_0	$1.484 \cdot 10^{3}$ m/s	Small signal sound speed in water
ρ	$1.0 \cdot 10^{3}$ kg/m^3	Density of liquid
u_0	$1.0 \cdot 10^{-9}$ m	Maximal SAW displacement
C_d	$8.06 \cdot 10^{3}$ 1/m	Damping parameter of the LSAW
$\tilde{\nu}_1$	$1.002 \cdot 10^{-6}$ m^2/s	Kinematic viscosity of water
$\tilde{\nu}_2$	$1.002 \cdot 10^{-6}$ m^2/s	Kinematic bulk viscosity of water

Table 6.1: Numerical and physical parameters for the numerical simulation of acoustic streaming

In Subsection 6.1.1 we model part of a microfluidic biochip with a square domain which acts as a fluid-filled cavity. The main purpose of this subsection is to validate our implementation using the numerical example 6.1.1 from [75] as a benchmark. The purpose of Subsection 6.1.2 is twofold: Firstly, we simulate a simplified biochip geometry used for experimental measurements in [41]. Secondly, we compare the results with those obtained by a reduced order model using BTMR. Finally, in Subsection 6.1.3 we discuss the model reduction based shape optimization of a capillary barrier using DDBTMR combined with a gradient type minimization algorithm.

6.1.1 Fluid-filled cavity

In this subsection we model parts of a microfluidic biochip with a square domain which acts as a fluid-filled cavity. The main purpose of this subsection is to validate our implementation using the numerical Example 6.1.1 from [75] as a benchmark. We consider a fluid-filled square cavity $\Omega = [0, 1 \text{ mm}]^2$ with SAW displacements $\mathbf{u} = (u_1, u_2)^T$ prescribed

at the bottom by

$$u_1(t, x_1) = 0.6\epsilon \exp(-\hat{C}_d x_1) \sin(2\pi(-\hat{k}x_1 + ft)),$$
$$u_2(t, x_2) = -\epsilon \exp(-\hat{C}_d x_1) \cos(2\pi(-\hat{k}x_1 + ft))$$

with parameters $\epsilon = u_0/L$, $\hat{C}_d = C_d L$, $\hat{k} = L/\lambda$, $\hat{f} = fT$, where $\lambda = 24\ \mu m$ is the SAW wavelength (cf. Table 6.1). The velocity $\mathbf{v}_1$ is set to zero on the other three boundaries. The SAW propagates from left to right with exponential attenuation. The fluid is assumed to be initially at rest, i.e., $\mathbf{v}_1 = \mathbf{0}, p_1 = 0$. We use P2-P1 Taylor-Hood finite elements for discretization in space and the Crank-Nicolson scheme for discretization in time.

We iterate until a periodicity condition for the pressure is fulfilled at some time t_{end}: We first compute pressures for k time steps. Then, we choose an offset number of time steps $m = \frac{2\pi}{\omega\tau_1}$ where $\omega = 2\pi f$ and $\tau_1 = 0.1$ denote the angular frequency and the time step size. We vary n from m to k and stop as soon the periodicity condition

$$\left\| \frac{1}{m} \sum_{i=n-m}^{n} p_1^{(i)} - p_1^{(i-m)} \right\|_{L^2} / \|p_1^{(n)}\|_{L^2} \leq \epsilon$$

is satisfied. Otherwise, we go back and increase k. We assume that the iteration stops for some $n = N$ which implies $t_{end} = \tau_1 N$.

Figure 6.1 displays the computed pressure at $t = 1.25\ \mu s$ (left) and the associated velocity field $\mathbf{v}_2$ (right). Both coincide well with experimental measurements reported in [75].

6.1.2 BTMR of a microfluidic biochip

A simplified biochip geometry as used for measurements in [41] is shown in Figure 6.2 (left). The IDT is placed at $x_1 = 0$ aligned with the top wall of the biochip. Since the

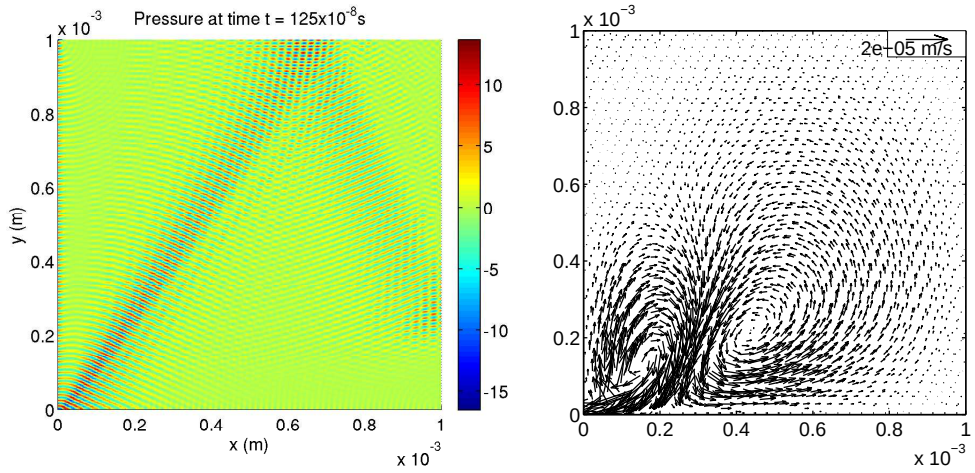


Figure 6.1: Fluid-filled cavity: Pressure at $t = 1.25\ \mu s$ (left) and Velocity field $\mathbf{v}_2$ in m/s (right)

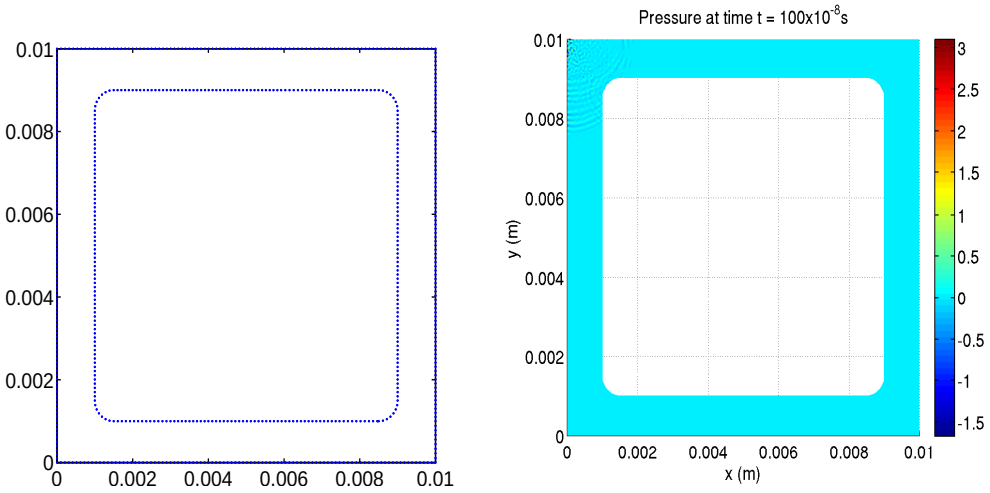


Figure 6.2: Typical biochip geometry (units in m) (left) and pressure at $t_{\text{end}} = 1.0\ \mu s$ (right)

width of the IDT is 0.6 mm, we set $\Gamma_{\text{in}} = \{0\} \times (9.4, 10)\ mm^2$. The function $\mathbf{u}$ describing the SAW excitation on Γ_{in} is chosen according to

$$u_1(t, x_1) = 0.6\epsilon \sin(2\pi(-\hat{k}x_1 + ft)),$$
$$u_2(t, x_1) = -\epsilon \cos(2\pi(-\hat{k}x_1 + ft)),$$

where the constants are the same as in Subsection 6.1.1. We further choose $\Gamma_{\text{out}} = \{10\} \times (0, 1)\ mm^2$ which ensures the uniqueness of the pressure for the acoustic streaming subproblem. On $\Gamma_{\text{lat}} = \partial\Omega \setminus (\Gamma_{\text{in}} \cup \Gamma_{\text{out}})$, the velocity is set to be zero. The geometry and the IDT specifications are exactly the same as in the experimental measurements performed in [41].

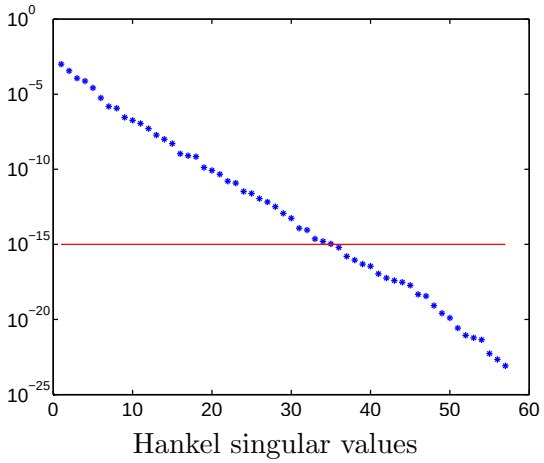
As a first task, we solve the acoustic subproblem (6.2). We iterate from $t = 0$ to $t_{\text{end}} = 1.0\ \mu s$ with time step $\tau_1 = 0.1$. Figure 6.2 (right) shows the computed pressure at t_{end} which is in excellent agreement with the measurements in [41]. The solution $(\mathbf{v}_1, p_1)$ is used to generate the right-hand side and the boundary data for the acoustic streaming subsystem (6.3) which is solved from $t = 0$ to $t_f = 0.1\ ms$.

The second task is to apply BTMR to the subsystem (6.3) observing the vorticity output in some part of the domain. More specifically consider that the output is the mean of the integral of the curl of the velocity

$$\mathbf{z}(t) = \int_{\Omega_{\text{obs}}} \nabla \times \mathbf{v}_2 \ dx \quad , \quad \Omega_{\text{obs}} = (1.5, 2.5) \times (9, 10)\ mm^2 \tag{6.4}$$

in some part Ω_{obs} of the domain. The semidiscretization of this problem leads to a system (5.12), where the inputs $\mathbf{u}$ in (5.12a) correspond to the inputs $\mathbf{u}$ in (6.3c) and the outputs $\mathbf{z}$ in (5.12c) correspond to (6.4). This is a simulation problem and BTMR as described in [62] can be applied directly. No domain decomposition or optimization is involved yet. The purpose of this numerical example is to explore the potential of BTMR for the shape

optimization problem.



grid	$N_{\mathbf{v}}$	N_p	$N_{\widehat{\mathbf{v}}}$
1	7770	1223	25
2	11334	1716	26
3	32818	4617	31
4	132506	17607	35

Figure 6.3: Left: The largest Hankel singular values and the threshold $10^{-12}\sigma_1$ for the fine grid problem. Right: The degrees of freedom $N_{\mathbf{v}}, N_p$ for the velocity and the pressure in the full order model and the degrees of freedom $N_{\widehat{\mathbf{v}}}$ for velocity in the reduced order model for four different grids.

We apply BTMR to semi-discretizations of (6.3) on several grids, where '1' refers to the coarsest grid and '4' refers to the finest grid. Figure 6.3 shows the largest Hankel singular values for problem on the finest grid. We note that the computed velocity obtained from the full order model on the finest mesh is of the order 10^{-4} m/s which is the same as obtained in the experimental results in [41]. For BTMR, we truncate the Hankel singular values by selecting the smallest index k for which $\sigma_{k+1} < 10^{-12}\sigma_1$. The threshold $10^{-12}\sigma_1$ is indicated by the solid line in Figure 6.3 (left). The truncation level $10^{-12}\sigma_1$ is small compared to what one usually sees in the literature for BTMR. This truncation level is motivated by the shape optimization problem discussed in the next subsection and by the scaling of the problem. In this section a coarser level could have been used, but we chose $10^{-12}\sigma_1$ to be comparable with the results in the next subsection.

The table on the right in Figure 6.3 shows the numbers $N_{\mathbf{v}}$ and N_p of velocity and pressure degrees of freedom (dof) for the full order models generated with the four different

grids. The same table also shows the velocity degrees of freedom $N_{\widehat{\mathbf{v}}}$ in the reduced order model (5.24). In particular, we see that BTMR is very effective and dramatically reduced the size of the system.

To illustrate the BTMR error bound (5.27a) we show the time domain response of the output $\mathbf{z}$ for the full order model and $\widehat{\mathbf{z}}$ for the reduced order model in Figure 6.4. As predicted by the theory, the reduced order model approximates the full order model very accurately.

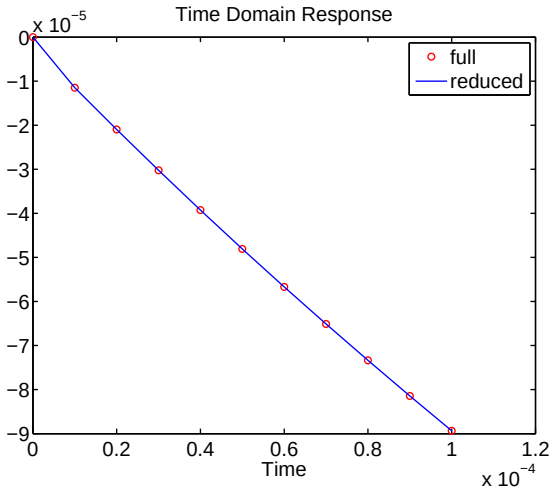


Figure 6.4: Time response for the full (circles) and for the reduced order model (solid line)

6.1.3 DDBTMR applied to shape optimization of capillary barriers

We consider fluid flow described by subsystem (6.3) in a network of microchannels and reservoirs on top of a microfluidic biochip with capillary barriers between the channels and the reservoirs that are designed to guarantee a precise filling of the reservoirs with the DNA or protein probes. The objective is twofold: Firstly, we want to design the capillary barriers in such a way that a desired velocity profile $\mathbf{v}^d$ is attained, and secondly, we want

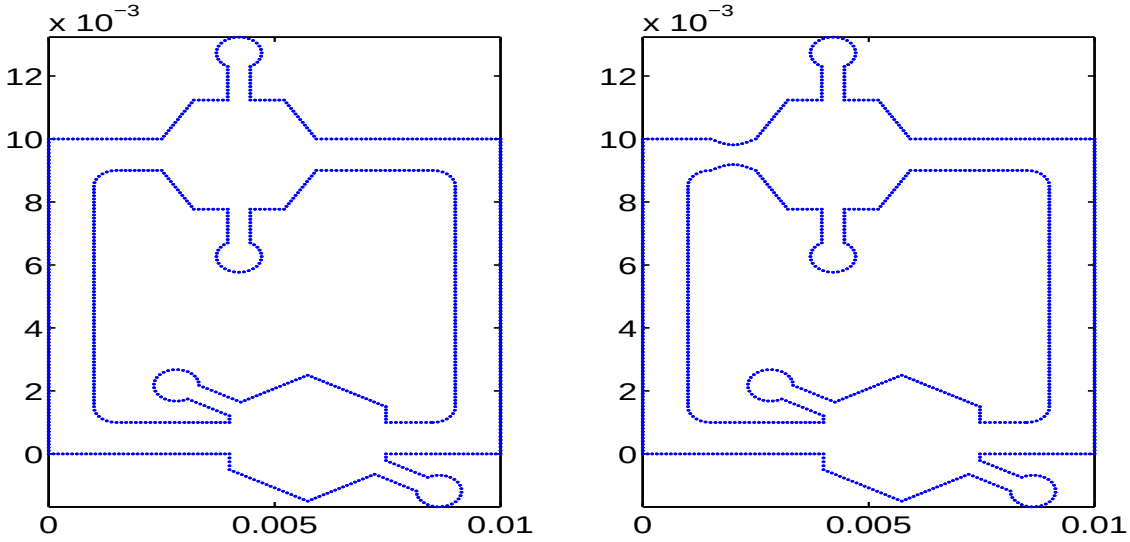


Figure 6.5: The reference domain Ω_{ref} (left) and the optimal domain (right)

to minimize the vorticity $\nabla \times \mathbf{v}$ in some specific part of the network.

The computational domain $\Omega \subset \mathbb{R}^2$ is displayed in Figure 6.5. It is decomposed into subdomains $\Omega_1 = \Omega \setminus \Omega_2$, and $\Omega_2 = (1.5, 2.5) \times (9, 10)\ mm^2$. The boundary $\partial\Omega$ is split into $\Gamma_{\text{in}} = \{0\} \times (9.4, 10), \Gamma_{\text{out}} = \{10\} \times (0, 1)\ mm^2$, and $\Gamma_{\text{lat}} = \partial\Omega \setminus (\Gamma_{\text{in}} \cup \Gamma_{\text{out}})$. We assume that an IDT of width 0.6 mm is placed at Γ_{in} and that the input velocity profile (u_1, u_2) is the same as in Subsection 6.1.2. The forces $(\mathbf{f}_v, f_p)$ in $\Omega \times (0, T)$ are computed by solving the acoustic subproblem (6.2) for $0 = t_0 \leq t \leq t_{end} = 1\ \mu s$ with step size $\tau_1 = 0.1$ and is equal to the right hand side in (6.3a-b). We further choose a constant velocity profile $\mathbf{v}_{\text{in}}(x_1, x_2)$ on $\Gamma_{\text{in}} \times (0, T)$ as given by (6.3c), outflow boundary conditions on $\Gamma_{\text{out}} \times (0, T)$, and no-slip conditions on $\Gamma_{\text{lat}} \times (0, T)$. The objective is to design the shape of the top $\Gamma_{2,T}$ and the bottom $\Gamma_{2,B}$ of $\partial\Omega_2$ in such a way that a prescribed velocity profile $\mathbf{v}^d$ is achieved in $\Omega_2 \times (0, T)$ and that the vorticity is minimized in Ω_{obs} (two bulb-shaped structures associated with the lower reservoir in Figure 6.5). The subdomain Ω_2 is parametrized by Bézier curves representing the top ($\Gamma_{2,T}$) and bottom ($\Gamma_{2,B}$) boundaries with Bézier

control points ℓ_T and ℓ_B, respectively. This leads to a parametrization $\Omega_2(\theta)$ of Ω_2 with parameters $\theta \in \mathbb{R}^{\ell_T+\ell_B}$. We use $\ell_T = \ell_B = 6$.

The shape optimization problem amounts to the minimization of

$$J(\theta) = \int_0^{t_{end}} \int_{\Omega_{\text{obs}}} |\nabla \times \mathbf{v}(x,t)|^2 dxdt + \int_0^{t_{end}} \int_{\Omega_2(\theta)} |\mathbf{v}(x,t) - \mathbf{v}^d(x,t)|^2 dxdt \tag{6.5}$$

subject to subsystem (6.3) and design parameter constraints

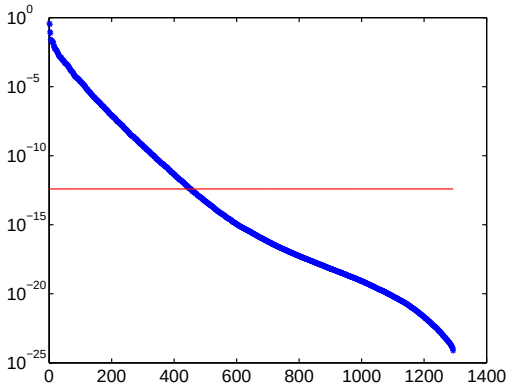
$$\theta^{min} \leq \theta \leq \theta^{max},$$

where $t_{end} = 0.1\ ms$. The bounds $\theta^{min}, \theta^{max}$ on the design parameters are chosen such that the design constraints are never active in this example. The optimal domain $\Omega(\theta^*)$ is shown in Figure 6.5.

We consider a geometrically conforming simplicial triangulation $\mathcal{T}_h(\Omega)$ of Ω that aligns with the decomposition into the subdomains Ω_1 and Ω_2 as well as the respective boundaries. The semi-discretization in space is performed as described in Subsections 5.1, 5.3 and Appendix B. Let $N_{\mathbf{v}}^{(1)}$, $N_{\mathbf{v}}^{(2)}$, $N_{\mathbf{v}}^{\Gamma}$ be the number of velocity degrees of freedom in the subdomains $\overline{\Omega}_1 \setminus \Gamma, \overline{\Omega}_2 \setminus \Gamma$ and in Γ, respectively, and set $N_{\mathbf{v}} = N_{\mathbf{v}}^{(1)} + N_{\mathbf{v}}^{(2)} + N_{\mathbf{v}}^{\Gamma}$. Similarly, let $N_p^{(1)}$, $N_p^{(2)}$ be the number of pressure degrees of freedom in the subdomains $\overline{\Omega}_1, \overline{\Omega}_2$ and let $N_p = N_p^{(1)} + N_p^{(2)}$ be the total number of pressure degrees of freedom.

We solve the semi-discretized optimization problems using a projected BFGS method with Armijo line search [71]. The optimization algorithm is terminated when the norm of the projected gradient is less than $\epsilon = 2 \cdot 10^{-8}$. We use automatic differentiation [54, 100] to compute the derivatives with respect to the design variables θ.

As before, the BTMR of the optimality subsystem is computed using the approach described in [62]. For BTMR, we truncate the Hankel singular values by selecting the



grid	m	$N_{\mathbf{v}}^{(1)}$	$N_{\widehat{\mathbf{v}}}^{(1)}$	$N_{\mathbf{v}}$	$N_{\widehat{\mathbf{v}}}$
1	167	7482	351	7640	509
2	195	11442	370	11668	596
3	291	16504	451	16830	777
4	802	48324	766	49238	1680

Figure 6.6: Left: The largest Hankel singular values computed for the fine grid problem and the threshold $10^{-12}\sigma_1$. Right: The number of observations m, the numbers $N_{\mathbf{v}}^{(1)}$, $N_{\mathbf{v}}$ of velocity degrees of freedom in subdomain Ω_1 and in Ω, respectively, for the full order model, and the number $N_{\widehat{\mathbf{v}}}^{(1)}$, $N_{\widehat{\mathbf{v}}}$ of velocity degrees of freedom in subdomain Ω_1 and in Ω, respectively, for the reduced order model.

smallest index k for which $\sigma_{k+1} < 10^{-12}\sigma_1$. We apply DDBTMR to semi-discretizations generated using four grids. Figure 6.6 shows the largest Hankel singular values computed for the fine grid problem. The threshold $10^{-12}\sigma_1$ is indicated by the solid line in Figure 6.6 (left). The table in in Figure 6.6 shows the sizes $N_{\mathbf{v}}^{(1)}$, $N_{\mathbf{v}}$ of the full order models on the four grids as well as the sizes $N_{\widehat{\mathbf{v}}}^{(1)}$, $N_{\widehat{\mathbf{v}}}$ of the reduced order models in subdomain Ω_1 and in Ω. Note that we apply BTMR only on subdomain Ω_1. For the fine grid, BTMR reduced the size of the Ω_1 subproblem from $N_{\mathbf{v}}^{(1)} = 48324$ to $N_{\widehat{\mathbf{v}}}^{(1)} = 766$. The velocity degrees of freedom in $\Omega_2 \cup \Gamma$ are not reduced. On the fine grid these are $N_{\mathbf{v}}^{(2)} + N_{\mathbf{v}}^{\Gamma} = 914$. Therefore, the reduced order problem has $N_{\widehat{\mathbf{v}}} = 914 + 766 = 1680$ degrees of freedom.

We notice that the reduction by BTMR is not as large as the one reported for the simulation problem in the previous subsection. There are two reasons for this. One reason is that in the simulation problem reported in the previous subsection we had only one observation. Now the observations are determined by the semi-discretization of $\nabla \times \mathbf{v}(x,t)$, $x \in \Omega_{\text{obs}}$, See (6.5). Thus the number m of observations is determined by the degrees of

freedom in Ω_{obs}. The second reason is that the subsystem corresponding to Ω_1 involves auxiliary inputs and outputs that are determined by the interface conditions between subdomains 1 and 2. Thus, the number of inputs and outputs for the subsystem corresponding to Ω_1 is larger than they were in the example problem discussed in the previous subsection. Therefore, the reduction achieved by BTMR on the Ω_1 subsystem is less. Of course, since the $\Omega_2(\theta)$ subsystem is not reduced, this problem size will also determine the size of the coupled reduced order problem, indicated by $N_{\widehat{\mathbf{v}}}$.

The constant c_g in the estimate (6.1) or the error between the optimal design parameters computed by the full and the reduced order problems, respectively, depend on quantities like α in Theorem 5.1.1(ii), derivatives of $\mathbf{A}(\theta)$ with respect to θ, etc., which in turn depend on the application data given in Table 6.1. In particular using Theorem 5.5.3

$$c_g = \left\{ TL_1 + \gamma \left\| \begin{pmatrix} \widehat{\boldsymbol{\mu}}_2 \\ \widehat{\boldsymbol{\mu}}_\Gamma \end{pmatrix} \right\| + \gamma\ (c_\lambda + c_\kappa)\ (\sigma_{k+1}\ +\ \cdots\ +\ \sigma_n) \right\} (c_v + c_p)$$
$$+ \gamma\ (c_\lambda + c_\kappa) \left\{ \left\| \begin{pmatrix} \widehat{\mathbf{x}}_2 \\ \widehat{\mathbf{x}}_\Gamma \end{pmatrix} \right\| + \|\mathbf{u}\| \right\},$$

with $c_v, c_p, c_\lambda, c_\kappa$ are the constants specified in Lemma 5.5.1 and 5.5.2. Here γ is the bound over the matrix norm of the derivatives of $\mathbf{A}(\theta)$, etc., with respect to θ. Numerical experiments indicate that for the current scaling of the problem, the constant c_g in the estimate (6.1) is large, for example, for the finest grid $\alpha \in (-10^{-4}, -10^{-8})$, $\|\mathbf{A}\| \approx 10^{-2}$, $\|D_\theta \mathbf{A}(\theta)\| \approx 10^{-2}$. Under the current scaling of the problem Lemma 5.5.1 and 5.5.2 implies

$$c_\kappa \approx \frac{\|\mathbf{A}\|^2}{\alpha^2}\ , \qquad c_p \approx \frac{\|\mathbf{A}\|}{\alpha}\ , \qquad \gamma \approx \|D_\theta \mathbf{A}(\theta)\|\ ,$$

then

$$c_g \approx \frac{\|\mathbf{A}\|^3\ \|D_\theta \mathbf{A}(\theta)\|}{\alpha^3} \in (10^4, 10^{16}).$$

Therefore, we require a rather small truncation level of $\sigma_{k+1} < 10^{-12}\sigma_1$ for the Hankel singular values.

The optimal shape parameters θ^* and $\widehat{\theta}^*$ computed by minimizing the full and the reduced order model, respectively, are shown in Table 6.2. For the finest grid, the error between the full and the reduced order model solutions is $\|\theta^* - \widehat{\theta}^*\| = 3.9165 \cdot 10^{-5}$.

θ^*	$(9.8833, 9.7467, 9.7572, 9.8671, 9.1336, 9.2015, 9.1971, 9.1310)\times 10^{-3}$
$\widehat{\theta}^*$	$(9.8694, 9.7374, 9.7525, 9.8628, 9.1498, 9.2044, 9.1895, 9.1204)\times 10^{-3}$

Table 6.2: Optimal shape parameters θ^* and $\widehat{\theta}^*$ computed by minimizing the full and the reduced order model

Conclusions and future work

For shape optimization problems, we have developed an adaptive multilevel interior-point method of barrier type featuring a predictor-corrector continuation method with an adaptive choice of the barrier parameter along the barrier path. The predictor relies on a nested-iteration type tangent continuation, and the corrector is a Newton-multigrid method for the KKT system. Despite the fact that this approach leads to a considerable reduction in the computational work compared to more standard optimization strategies (see Chapter 3), the amount of computational time is still significant, and there is a need for further reductions. This goal can be achieved by model reduction-based optimization methods using reduced order models for the underlying state equations (see Chapters 4-6).

We have integrated domain decomposition and model reduction for systems governed by time dependent advection-diffusion with small localized nonlinearities and the Stokes system in Chapters 4 and 5 respectively. In our case, nonlinearities arise from the dependence on the design variables.

We have proven estimates for the error between the solution of the original and the reduced order problem. These error estimates depend on balanced truncation error estimates.

Reduced order modeling based shape optimization for microfluidic biochips is presented in Chapter 6, where the theoretical error estimates derived in Chapter 5 for the Stokes system still apply in the original form. The results obtained are in good agreement with experimental measurements. Two more challenging real life applications are presented in Chapters 4 and 5. In Chapter 5, we compute the optimal configuration of a 2D aorto-coronaric bypass, and in Chapter 4 we reduced river pollution using our technique applied to an optimal control problem.

Our approach presented in Chapters 4-6 can be extended in various ways. It is possible to admit localized nonlinearities in the PDE, such as those considered in [108, 109]. Using model reduction techniques for nonlinear systems such POD (see e.g., the overview [64]) or extensions of balanced truncation to nonlinear systems [76] one can apply our approach to nonlinear PDEs. However, currently no a priori error estimates exist for these model reduction techniques and, consequently, no estimate for the error between the solution of the original optimization problem and of the reduced problem can be obtained.

One of our future goals is to improve the understanding and the analysis of nonlinear model reduction. Also, we think that we can simultaneously reduce the discretization error and the model reduction error by space-time adaptivity relying on time-like and space-like error indicators and model adaptivity by means of associated weighted residuals. The error indicators can be derived by the goal-oriented dual weighted approach well-known from adaptive finite element methods for PDEs and for PDE constrained optimal control problems [5]. Later on, we would like to replace BTMR by other model reduction techniques such as POD or reduced basis methods. The error estimates in [14, 15] hold

for all parameters and thus make our reduced order models interesting for probabilistic analysis/optimization under uncertainty as well.

APPENDIX A

Auxiliary Lemmas

Lemma A.1.1 *Let $\mathcal{A} \in \mathbb{R}^{N\times N}$ and $\mathcal{B} \in \mathbb{R}^{N\times m}$. If there exists $\alpha > 0$ such that*

$$\mathbf{v}^T \mathcal{A}\mathbf{v} \leq -\alpha \|\mathbf{v}\|^2 \quad \forall \mathbf{v} \in \mathbb{R}^N, \tag{A.1}$$

then the solution of

$$\mathbf{y}'(t) = \mathcal{A}\mathbf{y}(t) + \mathcal{B}\mathbf{u}(t), \quad t \in (0,T), \qquad \mathbf{y}(0) = \mathbf{y}_0 \tag{A.2}$$

obeys

$$\|\mathbf{y}\|_{L^2} \leq \frac{\sqrt{2}}{\sqrt{\alpha}}\|\mathbf{y}_0\| + \frac{2}{\alpha}\|\mathcal{B}\mathbf{u}\|_{L^2} \leq \frac{\sqrt{2}}{\sqrt{\alpha}}\|\mathbf{y}_0\| + \frac{2\|\mathcal{B}\|}{\alpha}\|\mathbf{u}\|_{L^2}.$$

Proof: We multiply the differential equation (A.2) by $\mathbf{y}(t)^T$ to obtain

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\|\mathbf{y}(t)\|^2 &= \mathbf{y}(t)^T\mathcal{A}\mathbf{y}(t) + \mathbf{y}(t)^T\mathcal{B}\mathbf{u}(t)\\ &\leq -\alpha\|\mathbf{y}(t)\|^2 + \mathbf{y}(t)^T\mathcal{B}\mathbf{u}(t).\end{aligned}$$

If we multiply the previous inequality by $\exp(\alpha t)$ we arrive at

$$\frac{d}{dt}\left(e^{\alpha t}\|\mathbf{y}(t)\|^2\right) \leq 2e^{\alpha t}\mathbf{y}(t)^T\mathcal{B}\mathbf{u}(t).$$

Integration from 0 to t gives

$$\|\mathbf{y}(t)\|^2 \leq e^{-\alpha t}\|\mathbf{y}_0\|^2 + \int_0^t 2e^{\alpha(\tau-t)}\mathbf{y}(\tau)^T\mathcal{B}\mathbf{u}(\tau)d\tau$$

and integrating the above inequality from 0 to T yields

$$\begin{aligned}\int_0^T \|\mathbf{y}(t)\|^2 dt &\leq \int_0^T e^{-\alpha t}dt\,\|\mathbf{y}_0\|^2 + \int_0^T\int_0^t 2e^{\alpha(\tau-t)}\mathbf{y}(\tau)^T\mathcal{B}\mathbf{u}(\tau)d\tau dt\\ &\leq \frac{1-e^{-\alpha T}}{\alpha}\|\mathbf{y}_0\|^2 + \int_0^T\int_\tau^T 2e^{\alpha(\tau-t)}dt\;\mathbf{y}(\tau)^T\mathcal{B}\mathbf{u}(\tau)d\tau\\ &= \frac{1-e^{-\alpha T}}{\alpha}\|\mathbf{y}_0\|^2 + \int_0^T \frac{2(1-e^{\alpha(\tau-T)})}{\alpha}\;\mathbf{y}(\tau)^T\mathcal{B}\mathbf{u}(\tau)d\tau\\ &\leq \frac{1}{\alpha}\|\mathbf{y}_0\|^2 + \int_0^T \frac{2}{\alpha}\;\|\mathbf{y}(\tau)\|\,\|\mathcal{B}\mathbf{u}(\tau)\|d\tau\\ &\leq \frac{1}{\alpha}\|\mathbf{y}_0\|^2 + \int_0^T \frac{1}{2}\|\mathbf{y}(\tau)\|^2 + \frac{2}{\alpha^2}\|\mathcal{B}\mathbf{u}(\tau)\|^2 d\tau,\end{aligned}$$

which implies the desired inequality. □

Lemma A.1.2 *Let $\mathcal{M} \in \mathbb{R}^{N\times N}$ be symmetric positive definite, $\mathcal{A} \in \mathbb{R}^{N\times N}$ and $\mathcal{B} \in \mathbb{R}^{N\times m}$. If there exists $\alpha > 0$ such that $\mathbf{v}^T\mathcal{A}\mathbf{v} \leq -\alpha\mathbf{v}^T\mathcal{M}\mathbf{v}$ for all $\mathbf{v} \in \mathbb{R}^N$, then the solution of*

$$\mathcal{M}\mathbf{y}'(t) = \mathcal{A}\mathbf{y}(t) + \mathcal{B}\mathbf{u}(t), \quad t \in (0,T) \tag{A.3}$$

with $\mathbf{y}(0) = \mathbf{y}_0$ *obeys*

$$\|\mathbf{y}\|_{L^2} \leq \frac{\sqrt{2}\|\mathcal{M}^{-1/2}\|\|\mathcal{M}^{1/2}\|}{\sqrt{\alpha}}\|\mathbf{y}_0\| + \frac{2\|\mathcal{M}^{-1}\|}{\alpha}\|\mathcal{B}\mathbf{u}\|_{L^2}.$$

Proof: If we multiply (A.3) by $\mathcal{M}^{-1/2}$ and apply Lemma A.1.1 to the resulting system we obtain the estimate

$$\|\mathcal{M}^{1/2}\mathbf{y}\|_{L^2} \leq \frac{\sqrt{2}}{\sqrt{\alpha}}\|\mathcal{M}^{1/2}\mathbf{y}_0\| + \frac{2}{\alpha}\|\mathcal{M}^{-1/2}\mathcal{B}\mathbf{u}\|_{L^2}.$$

This implies

$$\begin{aligned}
\|\mathbf{y}\|_{L^2} &= \|\mathcal{M}^{-1/2}\mathcal{M}^{1/2}\mathbf{y}\|_{L^2} \\
&\leq \frac{\sqrt{2}\|\mathcal{M}^{-1/2}\|}{\sqrt{\alpha}}\|\mathcal{M}^{1/2}\mathbf{y}_0\| + \frac{2\|\mathcal{M}^{-1/2}\|}{\alpha}\|\mathcal{M}^{-1/2}\mathcal{B}\mathbf{u}\|_{L^2} \\
&\leq \frac{\sqrt{2}\|\mathcal{M}^{-1/2}\|\|\mathcal{M}^{1/2}\|}{\sqrt{\alpha}}\|\mathbf{y}_0\| + \frac{2}{\alpha}\|\mathcal{M}^{-1}\|\|\mathcal{B}\mathbf{u}\|_{L^2}.
\end{aligned}$$

□

APPENDIX B

Implementation of DD Stokes

B.1 Model problems

We consider the following model problem

$$\alpha \mathbf{v} - \Delta \mathbf{v} + \nabla p = \mathbf{f} \qquad \text{in } \Omega, \tag{B.1a}$$

$$\boldsymbol{\nabla} \cdot \mathbf{v} = 0 \qquad \text{in } \Omega, \tag{B.1b}$$

$$(\nabla \mathbf{v} - pI)\mathbf{n} = \mathbf{0} \qquad \text{on } \Gamma_{out}, \tag{B.1c}$$

$$\mathbf{v} = \mathbf{g} \qquad \text{on } \Gamma_{in}, \tag{B.1d}$$

$$\mathbf{v} = \mathbf{0} \qquad \text{on } \partial\Omega \setminus (\Gamma_{in} \cup \Gamma_{out}), \tag{B.1e}$$

where $\alpha \geq 0$. The problem (B.1) with $\alpha > 0$ arises as a subproblem in time stepping methods for the time dependent Stokes equation. In this context, α is determined by the time step size.

The weak form is given by

$$\int_\Omega \alpha \mathbf{v} \cdot \boldsymbol{\psi} \; dx + \int_\Omega \nabla \mathbf{v} : \nabla \boldsymbol{\psi} \; dx - \int_\Omega p \boldsymbol{\nabla} \cdot \boldsymbol{\psi} \; dx = \int_\Omega \mathbf{f} \cdot \boldsymbol{\psi} \; dx \qquad \boldsymbol{\psi} \in \mathbf{V}, \tag{B.2a}$$

$$\int_\Omega \boldsymbol{\nabla} \cdot \mathbf{v} \, \phi \; dx = 0 \qquad \phi \in L^2(\Omega), \tag{B.2b}$$

where $\mathbf{V} = \left\{ \mathbf{v} \in \mathbf{H}^1(\Omega) \; : \; \mathbf{v} = 0 \text{ on } \partial\Omega \setminus \Gamma_{out} \right\}$. The Stokes equation (B.2) has a unique solution $(\mathbf{v}, p) \in \mathbf{H}^1(\Omega) \times L^2(\Omega)$.

We assume that the domain Ω is decomposed into three subdomains as shown in Figure B.1. The interface between subdomains Ω_j and Ω_{j+1} is denoted by $\Gamma_{j,j+1} = \overline{\Omega_j} \cap \overline{\Omega_{j+1}}$.

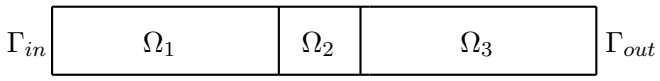


Figure B.1: Example domain

Now we consider the following Stokes equation on subdomain 1.

$$\alpha \mathbf{v} - \Delta \mathbf{v} + \nabla p = \mathbf{f} \qquad \text{in } \Omega_1, \tag{B.3a}$$

$$\boldsymbol{\nabla} \cdot \mathbf{v} = f \qquad \text{in } \Omega_1, \tag{B.3b}$$

$$\mathbf{v} = \mathbf{g} \qquad \text{on } \Gamma_{in}, \tag{B.3c}$$

$$\mathbf{v} = \mathbf{v}_\Gamma \qquad \text{on } \Gamma_{1,2}, \tag{B.3d}$$

$$\mathbf{v} = \mathbf{0} \qquad \text{on } \partial\Omega_1 \setminus (\Gamma_{in} \cup \Gamma_{12}), \tag{B.3e}$$

The weak form is given by

$$\int_{\Omega_1} \alpha \mathbf{v} \cdot \boldsymbol{\psi} \, dx + \int_{\Omega_1} \nabla \mathbf{v} : \nabla \boldsymbol{\psi} \, dx - \int_{\Omega_1} p \boldsymbol{\nabla} \cdot \boldsymbol{\psi} \, dx = \int_{\Omega_1} \mathbf{f} \cdot \boldsymbol{\psi} \, dx \quad \boldsymbol{\psi} \in \mathbf{H}_0^1(\Omega_1), \tag{B.4a}$$

$$\int_{\Omega_1} \boldsymbol{\nabla} \cdot \mathbf{v} \, \phi \, dx = \int_{\Omega_1} f \phi \, dx \qquad \phi \in L^2(\Omega_1). \tag{B.4b}$$

From (B.4b) we obtain (for sufficiently smooth ϕ)

$$\int_{\Omega_1} f \phi \, dx = \int_{\Omega_1} \boldsymbol{\nabla} \cdot \mathbf{v} \phi \, dx = \int_{\partial \Omega_1} \mathbf{v} \cdot n \, \phi \, ds - \int_{\Omega_1} \mathbf{v} \cdot \nabla \phi \, dx,$$

which for $\phi = 1$ gives the compatibility condition

$$\int_{\Omega_1} f \, dx = \int_{\Gamma_{in}} \mathbf{g} \cdot n \, ds + \int_{\Gamma_{12}} \mathbf{v}_\Gamma \cdot n \, ds, \tag{B.5}$$

which must be satisfied to guarantee the existence of a solution. If (B.5) is satisfied, then the Stokes equation (B.3) has a unique solution $(\mathbf{v}, p) \in \mathbf{H}^1(\Omega_1) \times L_0^2(\Omega_1)$.

For the following Stokes equation on subdomain 2 a similar result holds.

$$\alpha \mathbf{v} - \Delta \mathbf{v} + \nabla p = \mathbf{f} \qquad \text{in } \Omega_2, \tag{B.6a}$$

$$\boldsymbol{\nabla} \cdot \mathbf{v} = f \qquad \text{in } \Omega_2, \tag{B.6b}$$

$$\mathbf{v} = \mathbf{v}_\Gamma \qquad \text{on } \Gamma_{12} \cup \Gamma_{23}, \tag{B.6c}$$

$$\mathbf{v} = \mathbf{0} \qquad \text{on } \partial \Omega_2 \setminus (\Gamma_{12} \cup \Gamma_{23}). \tag{B.6d}$$

If the compatibility condition

$$\int_{\Omega_2} f \, dx = \int_{\Gamma_{12} \cup \Gamma_{23}} \mathbf{v}_\Gamma \cdot n \, ds \tag{B.7}$$

is satisfied, then the Stokes equation (B.6) has a unique solution $(\mathbf{v}, p) \in \mathbf{H}^1(\Omega_2) \times L_0^2(\Omega_2)$.

Since the subdomain 3 subproblem

$$\alpha\mathbf{v} - \Delta\mathbf{v} + \nabla p = \mathbf{f} \qquad \text{in } \Omega_3, \tag{B.8a}$$

$$\boldsymbol{\nabla} \cdot \mathbf{v} = f \qquad \text{in } \Omega_3, \tag{B.8b}$$

$$(\nabla\mathbf{v} - pI)\mathbf{n} = \mathbf{0} \qquad \text{on } \Gamma_{out}, \tag{B.8c}$$

$$\mathbf{v} = \mathbf{v}_\Gamma \qquad \text{on } \Gamma_{23}, \tag{B.8d}$$

$$\mathbf{v} = \mathbf{0} \qquad \text{on } \partial\Omega_3 \setminus (\Gamma_{23} \cup \Gamma_{out}) \tag{B.8e}$$

has an outflow condition, no compatibility condition is needed and the Stokes equation (B.8) has a unique solution $(\mathbf{v}, p) \in \mathbf{H}^1(\Omega_3) \times L^2(\Omega_3)$.

Thus the solution of local Stokes problems which only have Dirichlet boundary conditions, such as (B.3) and (B.6) require a consistency condition on the data. Moreover, the pressure for these subproblems is not uniquely defined. Among all pressure solutions (which vary only by a constant), one has to choose the one that corresponds to the restriction of the global pressure onto the respective subdomain. This is typically done by expressing the pressure on the Dirichlet subdomains as the sum of a constant and a pressure with zero average. The constant is determined from a global problem, the pressure with zero average is uniquely determined by the subproblem (provided that the consistency conditions are satisfied).

B.2 Discretization

To simplify the presentation, we assume the inflow velocity is zero, $\mathbf{g} = \mathbf{0}$. Let $\mathcal{T}$ be a triangulation of Ω and let $\{x_j\}$ be the set of nodes (vertices and edge midpoints) in the triangulation. In the Taylor-Hood finite element approximation the spaces of discretized

velocities and pressures are given by

$$\mathbf{V}^h = \left\{ \mathbf{v}^h \in \mathbf{C}(\overline{\Omega}) \ : \ \mathbf{v}^h|_K \in \mathcal{P}_2, \quad \forall K \in \mathcal{T}, \qquad \mathbf{v} = 0 \text{ on } \partial\Omega \setminus \Gamma_{out} \right\}, \tag{B.9a}$$

$$\widetilde{Q}^h = \left\{ q^h \in C(\overline{\Omega}) \ : \ q^h|_K \in \mathcal{P}_1, \quad \forall K \in \mathcal{T} \right\}, \tag{B.9b}$$

respectively. The space of velocities is decomposed in a standard manner.

$$\mathbf{V}_j^h = \left\{ \mathbf{v}^h \in \mathbf{V}^h \ : \ \mathbf{v}^h(x_k) = 0 \text{ for all } x_k \notin \Omega_j \right\}, \qquad j = 1, 2, 3, \tag{B.10a}$$

$$\mathbf{V}_\Gamma^h = \left\{ \mathbf{v}^h \in \mathbf{V}^h \ : \ \mathbf{v}^h(x_k) = 0 \text{ for all } x_k \notin \Gamma_{12} \cup \Gamma_{23} \right\}. \tag{B.10b}$$

Note that

$$\mathbf{V}^h = \mathbf{V}_1^h \oplus \mathbf{V}_2^h \oplus \mathbf{V}_3^h \oplus \mathbf{V}_\Gamma^h.$$

The pressures in subdomains Ω_1 and Ω_2 are represented as the sum of a constant and a pressure with average zero. This leads to the spaces

$$\widetilde{Q}_j^h = \left\{ q^h \in C(\overline{\Omega}_j) \ : \ q^h|_K \in \mathcal{P}_1, \quad \forall K \in \mathcal{T} \right\}, \qquad j = 1, 2, \tag{B.10c}$$

$$Q_j^h = \left\{ q^h \in C(\overline{\Omega}_j) \ : \ q^h|_K \in \mathcal{P}_1, \quad \forall K \in \mathcal{T}, \qquad \int_{\Omega_j} q^h \, dx = 0 \right\}, \quad j = 1, 2, \tag{B.10d}$$

$$Q_3^h = \left\{ q^h \in C(\overline{\Omega}_3) \ : \ q^h|_K \in \mathcal{P}_1, \quad \forall K \in \mathcal{T} \right\}, \tag{B.10e}$$

$$Q_0^h = \left\{ q_1 \chi_{\overline{\Omega}_1} + q_2 \chi_{\overline{\Omega}_2} \ : \ q_1, q_2 \in \mathbb{R} \right\}, \tag{B.10f}$$

where χ_S is the characteristic function corresponding to a set S. Note that

$$\widetilde{Q}^h \subsetneq Q^h \stackrel{\text{def}}{=} Q_1^h \oplus Q_2^h \oplus Q_3^h \oplus Q_0^h.$$

Furthermore, note that if

$$\widetilde{\phi}_i^j, \quad i = 1, \ldots, m_j$$

are piecewise linear basis functions (hat functions) for $\widetilde{Q}_j^h$, then

$$\phi_i^j \stackrel{\text{def}}{=} \widetilde{\phi}_i^j - \frac{1}{|\Omega_j|} \int_{\Omega_j} \widetilde{\phi}_i^j \, dx, \quad i = 1, \ldots, m_j - 1,$$

are basis functions for Q_j^h, where $|\Omega_j|$ is the measure of Ω_j.

The discretization of (B.2) is now given as follows: Find

$$\mathbf{v}^h = \mathbf{v}_1^h + \mathbf{v}_2^h + \mathbf{v}_3^h + \mathbf{v}_\Gamma \in \mathbf{V}_1^h \oplus \mathbf{V}_2^h \oplus \mathbf{V}_3^h \oplus \mathbf{V}_\Gamma^h,$$
$$\mathbf{p}^h = p_1^h + p_2^h + p_3^h + p_0 \in Q_1^h \oplus Q_2^h \oplus Q_3^h \oplus Q_0^h,$$

such that

$$\int_\Omega \alpha \mathbf{v}^h \cdot \boldsymbol{\psi}^h \, dx + \int_\Omega \nabla \mathbf{v}^h : \nabla \boldsymbol{\psi}^h \, dx - \int_{\Omega_1} p^h \boldsymbol{\nabla} \cdot \boldsymbol{\psi} \, dx = \int_\Omega \mathbf{f} \cdot \boldsymbol{\psi} \, dx \tag{B.11a}$$
$$\forall \boldsymbol{\psi} \in \mathbf{V}_1^h \oplus \mathbf{V}_2^h \oplus \mathbf{V}_3^h \oplus \mathbf{V}_\Gamma^h,$$
$$\int_\Omega \boldsymbol{\nabla} \cdot \mathbf{v}^h \, \phi^h \, dx = 0 \tag{B.11b}$$
$$\forall \phi \in Q_1^h \oplus Q_2^h \oplus Q_3^h \oplus Q_0^h.$$

The condition

$$-\int_\Omega \boldsymbol{\nabla} \cdot (\mathbf{v}_1^h + \mathbf{v}_2^h + \mathbf{v}_3^h + \mathbf{v}_\Gamma) \, \phi^h \, dx = 0, \qquad \forall \phi \in Q_1^h \oplus Q_2^h \oplus Q_3^h \oplus Q_0^h$$

leads to

$$\begin{pmatrix} \mathbf{B}_{11} & \mathbf{0} & \mathbf{0} & \mathbf{B}_{1\Gamma} \\ \mathbf{0} & \mathbf{B}_{22} & \mathbf{0} & \mathbf{B}_{2\Gamma} \\ \mathbf{0} & \mathbf{0} & \mathbf{B}_{33} & \mathbf{B}_{3\Gamma} \\ 0 & 0 & 0 & \mathbf{B}_0 \end{pmatrix} \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \mathbf{v}_3 \\ \mathbf{v}_\Gamma \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}.$$

Note that for $\boldsymbol{\psi} \in \mathbf{V}_j^h$ and constant $\chi_{\overline{\Omega}_j}$

$$\int_{\Omega_j} \boldsymbol{\nabla} \cdot \boldsymbol{\psi} \, \chi_{\overline{\Omega}_j} \, dx = \int_{\partial\Omega_j} \boldsymbol{\psi} \cdot n \, \chi_{\overline{\Omega}_j} \, ds - \int_{\Omega_j} \boldsymbol{\psi} \cdot \nabla \chi_{\overline{\Omega}_j} \, dx = 0$$

since $\boldsymbol{\psi} = 0$ on $\partial\Omega_j$, $j = 1, 2, 3$. Hence we have zero blocks in the last row block. The

condition

$$\int_\Omega \alpha(\mathbf{v}_1^h + \mathbf{v}_2^h + \mathbf{v}_3^h + \mathbf{v}_\Gamma) \cdot \boldsymbol{\psi}^h \, dx + \int_\Omega \nabla(\mathbf{v}_1^h + \mathbf{v}_2^h + \mathbf{v}_3^h + \mathbf{v}_\Gamma) : \nabla \boldsymbol{\psi}^h \, dx$$
$$- \int_{\Omega_1} (p_1^h + p_2^h + p_3^h + p_0^h) \, \boldsymbol{\nabla} \cdot \boldsymbol{\psi} \, dx = \int_\Omega \mathbf{f} \cdot \boldsymbol{\psi} \, dx$$

for all $\boldsymbol{\psi} \in \mathbf{V}_1^h \oplus \mathbf{V}_2^h \oplus \mathbf{V}_3^h \oplus \mathbf{V}_\Gamma^h$ leads to

$$\begin{pmatrix} \mathbf{A}_{11} & \mathbf{0} & \mathbf{0} & \mathbf{A}_{1\Gamma} \\ \mathbf{0} & \mathbf{A}_{22} & \mathbf{0} & \mathbf{A}_{2\Gamma} \\ \mathbf{0} & \mathbf{0} & \mathbf{A}_{33} & \mathbf{A}_{3\Gamma} \\ \mathbf{A}_{\Gamma 1} & \mathbf{A}_{\Gamma 2} & \mathbf{A}_{\Gamma 3} & \mathbf{A}_{\Gamma\Gamma} \end{pmatrix} \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \mathbf{v}_3 \\ \mathbf{v}_\Gamma \end{pmatrix} + \begin{pmatrix} \mathbf{B}_{11}^T & \mathbf{0} & \mathbf{0} & 0 \\ \mathbf{0} & \mathbf{B}_{22}^T & \mathbf{0} & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{B}_{33}^T & 0 \\ \mathbf{B}_{1\Gamma}^T & \mathbf{B}_{2\Gamma}^T & \mathbf{B}_{3\Gamma}^T & \mathbf{B}_0^T \end{pmatrix} \begin{pmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \mathbf{p}_3 \\ \mathbf{p}_0 \end{pmatrix} = \begin{pmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \\ \mathbf{f}_3 \\ \mathbf{f}_\Gamma \end{pmatrix}.$$

In the following we use

$$\boldsymbol{\psi}_k^j, \qquad k = 1, \ldots, n_j, \qquad j = 1, 2, 3,$$
$$\boldsymbol{\psi}_k^\Gamma, \qquad k = 1, \ldots, n_\Gamma,$$

to denote the piecewise linear basis functions for the velocities corresponding to Ω_j, $j = 1, 2, 3$, and corresponding to the interface Γ. Furthermore,

$$\widetilde{\phi}_i^j, \quad i = 1, \ldots, m_j, \qquad j = 1, 2, 3,$$

are piecewise linear basis functions (hat functions) for $\widetilde{Q}_1^h$, $\widetilde{Q}_2^h$, $Q_3^h = \widetilde{Q}_3^h$, and

$$\phi_i^j \stackrel{\text{def}}{=} \widetilde{\phi}_i^j - \frac{1}{|\Omega_j|} \int_{\Omega_j} \widetilde{\phi}_i^j \, dx, \quad i = 1, \ldots, m_j - 1, \quad j = 1, 2,$$

are basis functions for Q_j^h, $j = 1, 2$, where $|\Omega_j|$ is the measure of Ω_j, and

$$\phi_i^3 \stackrel{\text{def}}{=} \widetilde{\phi}_i^3 \quad i = 1, \ldots, m_3.$$

For the computation of the matrix $\mathbf{B}_{jj} \in \mathbb{R}^{(m_j-1)\times n_j}$, $j = 1, 2$, with entries

$$(\mathbf{B}_{jj})_{ik} = \int_{\Omega_j} \boldsymbol{\nabla} \cdot \boldsymbol{\psi}_k^j \, \phi_i^j \, dx$$

we note that

$$\begin{aligned}\int_{\Omega_j} \boldsymbol{\nabla} \cdot \boldsymbol{\psi}_k^j \, \phi_i^j \, dx &= \int_{\Omega_j} \boldsymbol{\nabla} \cdot \boldsymbol{\psi}_k^j \left(\widetilde{\phi}_i^j - \frac{1}{|\Omega_j|} \int_{\Omega_j} \widetilde{\phi}_i^j \, dx \right) \, dx \\ &= \int_{\Omega_j} \boldsymbol{\nabla} \cdot \boldsymbol{\psi}_k^j \widetilde{\phi}_i^j \, dx - \int_{\Omega_j} \boldsymbol{\nabla} \cdot \boldsymbol{\psi}_k^j \, dx \, \frac{1}{|\Omega_j|} \int_{\Omega_j} \widetilde{\phi}_i^j \, dx.\end{aligned}$$

Hence if we compute $\widetilde{\mathbf{B}}_{jj} \in \mathbb{R}^{m_j \times n_j}$ with entries

$$(\widetilde{\mathbf{B}}_{jj})_{ik} = \int_{\Omega_j} \boldsymbol{\nabla} \cdot \boldsymbol{\psi}_k^j \, \widetilde{\phi}_i^j \, dx$$

and the averages $c_i \stackrel{\text{def}}{=} \frac{1}{|\Omega_j|} \int_{\Omega_j} \widetilde{\phi}_i^j \, dx$, then, since $\sum_{i=1}^{m_j} \widetilde{\phi}_i^j = 1$, we have

$$\int_{\Omega_j} \boldsymbol{\nabla} \cdot \boldsymbol{\psi}_k^j \, dx = (e^T \widetilde{\mathbf{B}}_{jj})_k,$$

where $e = (1, \ldots, 1)^T \in \mathbb{R}^{m_j}$ and

$$(\mathbf{B}_{jj})_{ik} = (\widetilde{\mathbf{B}}_{jj})_{ik} - (e^T \widetilde{\mathbf{B}}_{jj})_k \, c_i.$$

The same observation applies to the computation of the matrices $\mathbf{B}_{j\Gamma} \in \mathbb{R}^{(m_j-1)\times n_\Gamma}$, $j = 1, 2$.

The matrix $\mathbf{B}_0 \in \mathbb{R}^{2\times n_\Gamma}$, with entries

$$(\mathbf{B}_0)_{jk} = \int_{\Omega_j} \boldsymbol{\nabla} \cdot \boldsymbol{\psi}_k^\Gamma \, dx \quad j = 1, 2$$

can be computed from the matrices $\widetilde{\mathbf{B}}_{j\Gamma} \in \mathbb{R}^{m_j \times n_\Gamma}$, $j = 1, 2$, with entries

$$(\widetilde{\mathbf{B}}_{j\Gamma})_{ik} = \int_{\Omega_j} \boldsymbol{\nabla} \cdot \boldsymbol{\psi}_k^\Gamma \, \widetilde{\phi}_i^j \, dx$$

using

$$(\mathbf{B}_0)_{jk} = (e^T \widetilde{\mathbf{B}}_{j\Gamma})_k.$$

APPENDIX C

Notation and symbols

$\mathbb{N}$	Set of nonzero natural numbers
$\mathbb{R}$	Set of real numbers
∂A	Boundary set of a set $A \subset \mathbb{R}^d$
$A^\top$	Transpose of a matrix $A \in \mathbb{R}^{n\times m}$
$A : B$	For matrices $A, B \in \mathbb{R}^{n\times m}$ the scalar product $A : B := \sum_{i,j} A_{ij}B_{ij}$
SPD	Symmetric and positive definite
BTMR	Balanced truncation model reduction
DD	Domain decomposition
(v, w)	Scalar product of v and w
$L(E, F)$	Space of bounded linear mappings between two normed spaces E and F

E^*	Dual space of a Banach space E, defined as the space of bounded linear functionals on E
$\langle g, f \rangle_{E \times E^*}$	Duality pairing of $f \in E^*$ with $g \in E$
$\|x\|$	Euclidean l^2-norm of $x \in \mathbb{R}^n$, defined as $\|x\| = \left(\sum_{i=1}^{n} {x_i}^2\right)^{1/2}$
Ω	Bounded polygonal or polyhedral domain $\subset \mathbb{R}^d, d \in \{2, 3\}$ with Lipschitz boundary $\Gamma = \partial\Omega$
$L^p(\Omega)$	Lebesgue space of scalar p-integrable functions on Ω
$\mathbf{L}^p(\Omega)$	Vector p-integrable functions on Ω
$L_0^p(\Omega)$	Scalar p-integrable functions with mean zero over Ω
$H^m(\Omega)$	Sobolev space of scalar functions with weak derivatives up to order m in $L^2(\Omega)$
$\mathbf{H}^m(\Omega)$	Sobolev space of vector valued functions; $\mathbf{H}^m(\Omega) := (H^m(\Omega))^d$
$\mathbf{H}_0^m(\Omega)$	Subpace of $\mathbf{H}^m(\Omega)$ with functions of zero trace
$\mathbf{H}^{1/2}(\Gamma)$	Trace space
$\mathbf{H}_{00}^{1/2}(\Gamma)$	Sobolev space defined as a set of functions $\{\mathbf{u} \in \mathbf{H}^{1/2}(\Gamma) \mid \mathcal{E}\mathbf{u} \in \mathbf{H}^{1/2}(\partial\Omega)\}$, where Γ is a proper subset of $\partial\Omega$ and $\mathcal{E}\mathbf{u}$ is the extension by zero of $\mathbf{u}$ to $\partial\Omega$
$\|u\|_{m,\Omega}$	Norm of the Sobolev space $H^m(\Omega)$ or $\mathbf{H}^m(\Omega)$
$\|u\|_{0,\Omega}$	Norm of the space $L^2(\Omega)$ or $\mathbf{L}^2(\Omega)$
$(\cdot, \cdot)_{m,\Omega}$	Inner product of the Sobolev space $H^m(\Omega)$ or $\mathbf{H}^m(\Omega)$
$(\cdot, \cdot)_{0,\Omega}$	Norm of the space $L^2(\Omega)$ or $\mathbf{L}^2(\Omega)$

Bibliography

[1] F. Abergel and R. Temam. On some control problems in fluid mechanics. *Theoretical and Computational Fluid Dynamics*, **1**(6):303–325, 1990.

[2] V. Agoshkov, A. Quarteroni, and G. Rozza. A mathematical approach in the design of arterial bypass using unsteady Stokes equations. *Journal of Scientific Computing*, **28**(2):139–165, 2006.

[3] M. Ainsworth and S. Sherwin. Domain decomposition preconditioners for p and hp finite element approximation of Stokes equations. *Comput. Methods Appl. Mech. Engrg.*, **175**(3-4):243–266, 1999.

[4] V. Akçelik., G. Biros, O. Ghattas, K.R. Long, B. van Bloemen Waanders. A variational finite element method for source inversion for convective-diffusive transport. *Finite Elem. Anal. Des.*, **39**(8):683–705, (2003).

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

[6] G. Biros and O. Ghattas. Parallel Lagrange-Newton-Krylov-Schur methods for PDE-constrained optimization. Part i: The Krylov-Schur solver. *SIAM J. Sci. Comp.*, **27**(2):687–713, 2005.

[7] G. Biros and O. Ghattas. Parallel Lagrange-Newton-Krylov-Schur methods for PDE-constrained optimization. Part ii: The Lagrange-Newton solver and its application to optimal control of steady viscous flows. *SIAM J. Sci. Comp.*, **27**(2), 714–739, 2006.

[8] P. Böhm, R.H.W. Hoppe, G. Mazurkevitch, S.I. Petrova, G. Wachutka, and E. Wolfgang. Optimal structural design of high power electronic devices by topology optimization. In: *Mathematics - Key Technology for the Future. Cooperations between*

Mathematics and Industry (Krebs H, and Jäger W (eds.)), pp. 365–376, Springer, Berlin-Heidelberg-New York, 2003.

[9] G. Allaire. *Shape optimization by the homogenization method*. Springer, Berlin-Heidelberg-New York, 2002.

[10] H. Antil, A. Gantner, R.H.W. Hoppe, D. Köster, K.G. Siebert, and A. Wixforth. Modeling and simulation of piezoelectrically agitated acoustic streaming on microfluidic biochips. In Domain Decomposition Methods in Science and Engineering XVII, (U. Langer et al., eds.), *Lecture Notes in Computational Science and Engineering*, Springer, Berlin-Heidelberg-New York, **60**:305–312, 2008.

[11] H. Antil, R.H.W. Hoppe, and C. Linsenmann. Path-following primal-dual interior-point methods for shape optimization of stationary flow problems. *J. Numer. Math.*, **15**(2):81-100, 2007.

[12] H. Antil, R. Glowinski, R.H.W. Hoppe, C. Linsenmann, T.W. Pan, and A. Wixforth. Modeling, simulation, and optimization of surface acoustic wave driven microfluidic biochips. Submitted to: *J. Comput. Math.*, 2009.

[13] H. Antil, M. Heinkenschloss, R. H. W. Hoppe, C. Linsenmann, and A. Wixforth. Reduced order based shape optimization of surface acoustic wave driven microfluidic biochips. Submitted to: *Mathematics and Computers in Simulation*, 2009.

[14] H. Antil, M. Heinkenschloss, R. H. W. Hoppe, and D.C. Sorensen. Domain decomposition and model reduction for the numerical solution of PDE constrained optimization problems with localized optimization variables. Submitted to: *Comput. Vis. Sci.*, 2009.

[15] H. Antil, M. Heinkenschloss, and R. H. W. Hoppe. Domain decomposition and balanced truncation model reduction for shape optimization of the stokes system. Submitted to: *Optimization Methods and Software*, 2009.

[16] H. Antil, R.H.W. Hoppe, and C. Linsenmann. Optimal design of stationary flow problems by path-following interior-point methods. *Control and Cybernetics*, **37**(4):771–796, 2008.

[17] H. Antil, R.H.W. Hoppe, and C. Linsenmann. Adaptive multilevel interior-point methods in PDE constrained optimization. In Proc. Int. Conf. on Domain Decomposition Methods and Applications (M. Bercovier et al.; eds.), *Lecture Notes in Computer Science and Engineering*, Springer, Berlin-Heidelberg-New York, **70**:15–26, 2009.

[18] A. C. Antoulas. *Approximation of Large-Scale Systems*. SIAM, Philadelphia, 2005.

[19] A. C. Antoulas, M. Heinkenschloss, and T. Reis. On balanced truncation for inhomogeneously initialized systems. *Technical Report TR09*, Department of Computational and Applied Mathematics, Rice University, 2009.

[20] P. Armand, J. Benoist, and D. Orban. Dynamic updates of the barrier parameter in primal-dual methods for nonlinear programming. *Computational Optimization and Applications*, **41**(1):1–25, 2008.

[21] M.P. Bendsøe. *Optimization of structural topology, shape, and material*. Springer, Berlin-Heidelberg-New York, 1995.

[22] M.P. Bendsøe and O. Sigmund. *Topology optimization: Theory, methods and applications*. Springer, Berlin-Heidelberg-New York, 2003.

[23] P. Benner, V. Mehrmann, and D. C. Sorensen, editors. *Dimension reduction of large-scale systems*. Lecture Notes in Computational Science and Engineering, Vol. **45**. Springer-Verlag, Heidelberg, 2005.

[24] D. Braess. *Finite elements: Theory, fast solvers, and applications in elasticity theory*. Cambridge University Press, Cambridge, third edition, 2007.

[25] J. H. Bramble and J. E. Pasciak. A domain decomposition technique for Stokes problems. *Appl. Numer. Math.*, **6**(4):251–261, 1990.

[26] F. Brezzi and M. Fortin. *Mixed and hybrid finite element methods*. Computational Mathematics, Vol. **15**. Springer–Verlag, Berlin, 1991.

[27] R.H. Byrd, J.C. Gilbert, and J. Nocedal. A trust region method based on interior point techniques for nonlinear programming. *Math. Programming*, **89**(1):149-185, 2000.

[28] Y. Cao, S. Li, L. Petzold, and R. Serban. Adjoint sensitivity analysis for differential-algebraic equations: The adjoint DAE system and its numerical solution. *SIAM J. Sci. Comput.*, **24**(3):1076–1089 (electronic), 2002.

[29] A. Cherkaev. *Variational methods for structural optimization.* Springer Verlag, New York, 2000.

[30] M. Crouzeix, and P.A. Raviart. Conforming and nonconforming finite element methods for solving the stationary Stokes equations. I. *Rev. Française Automat. Informat. Recherche Opérationnelle Sér. Rouge*, **7**(R-3):33–75, 1973.

[31] L. Dedé and A. Quarteroni. Optimal control and numerical adaptivity for advection-diffusion equations. *ESAIM: Mathematical Modelling and Numerical Analysis* **39**(5):1019–1040, 2005.

[32] M.C. Delfour and J.P. Zolesio. *Shapes and geometries: Analysis, differential calculus and optimization*. SIAM, Philadelphia, 2001.

[33] P. Deuflhard. *Newton methods for nonlinear problems. Affine invariance and adaptive algorithms*. Springer, Berlin-Heidleberg-New York, 2004.

[34] G.E. Dullerud and F.G. Paganini. *A course in robust control theory.* Texts in Applied Mathematics, Vol. **36**. Springer, Berlin, Heidelberg, New York, 2000.

[35] H.C. Elman. Preconditioners for saddle point problems arising in computational fluid dynamics. *Appl. Numer. Math.*, **43**(1-2):75-89, 2002.

[36] H.C. Elman, D.J. Silvester, and A.J. Wathen. *Finite elements and fast iterative solvers: With applications in incompressible fluid dynamics* . Oxford University Press, USA, 2005.

[37] Weinan E and B. Engquist. The heterogeneous multiscale methods. *Comm. Math. Sciences*, **1**(1):87-132, 2003.

[38] Weinan E and B. Engquist. The heterogoeneous multiscale method for homogenization problems. In Multiscale Methods in Sci. and Engrg., pp. 89-110, *Lect. Notes in Comput Sci. Engrg.*, Vol. **44**, Springer, Berlin-Heidelberg-New York, 2005.

[39] A.S. El-Bakry, R.A. Tapia, T. Tsuchiya, and Y. Zhang. On the formulation and theory of the Newton interior-point method for nonlinear programming. *Journal of Optimization Theory and Applications*, **89**(3):507-541, 1996.

[40] A.C. Eringen and G.A. Maugin. *Electrodynamics of continua volume 1: Foundations and solid media.* Springer, Berlin-Heidelberg-New York, 1990.

[41] M.A. Fallah. *SAW induced acoustic streaming in microchannels of different geometry.* Master's Thesis. Institute of Physics, University of Augsburg, 2008.

[42] L. Fatone, P. Gervasio, and A. Quarteroni. Multimodels for incompressible flows. *J. Math. Fluid Mech.* **2**(2):126–150, 2000.

[43] L. Fatone, P. Gervasio, and A. Quarteroni. Multimodels for incompressible flows: iterative solutions for the Navier-Stokes/Oseen coupling. *M2AN Math. Model. Numer. Anal.* **35**(3):549–574, 2001.

[44] A.V. Fiacco and G.P. McCormick. *Nonlinear programming: Sequential unconstrained minimization techniques.* SIAM, Philadelphia, 1990.

[45] L. Formaggia, J.F. Gerbeau, F. Nobile, and A. Quarteroni. On the coupling of 3D and 1D Navier-Stokes equations for flow problems in compliant vessels. *Comput. Methods Appl. Mech. Engrg.* **191**(6-7):561–582, 2001.

[46] A. Forsgren, P. E. Gill, and M. H. Wright. Interior methods for nonlinear optimization. *SIAM Rev.* **44**(4):525–597, 2002.

[47] T. Franke and A. Wixforth. Microfluidics for miniaturized laboratories on a chip. *ChemPhysChem*, **9**(15):2140-2156, 2008.

[48] A. Gantner. *Mathematical modeling and numerical simulation of piezoelectrical agitated surface acoustic waves*. Ph.D. Thesis, Department of Applied Mathematics, University of Augsburg (2005).

[49] A. Gantner, R.H.W. Hoppe, D. Köster, K.G. Siebert, and A. Wixforth. Numerical simulation of piezoelectrically agitated surface acoustic waves on microfluidic biochips. *Comp. Visual. Sci.*, **10**(3):145-161, 2007.

[50] D.M. Gay, M.I. Overton, and M.H. Wright. Primal-dual interior method for nonconvex nonlinear programming. *Advances in Nonlinear Programming*(Y. Yuan; ed.), Kluwer, Dordrecht, pp. 31-56, 1998.

[51] V. Girault and P.A. Raviart. *Finite element methods for Navier-Stokes equations. Theory and algorithms*. Vol. **5** of Springer Series in Computational Mathematics. Springer-Verlag, Berlin, 1986.

[52] R. Glowinski. *Handbook of numerical analysis: Numerical methods for fluids*. Elsevier, Amsterdam, 2004.

[53] K. Glover. All optimal Hankel-norm approximations of linear multivariable systems and their L^∞-error bounds. *Internat. J. Control*, 39(6):1115–1193, 1984.

[54] A. Griewank and A. Walther. *Evaluating derivatives: Principles and techniques of algorithmic differentiation*. SIAM, Philadelphia, second edition, 2008.

[55] A. Griewank. *Evaluating derivatives: Principles and techniques of algorithmic differentiation*. SIAM, Phildelphia, 2000.

[56] S. Gugercin, D.C. Sorensen, and A.C. Antoulas. A modified low-rank Smith method for large-scale Lyapunov equations. *Numer. Algorithms*, **32**(1):27–55, 2003.

[57] M. D. Gunzburger. *Perspectives in flow control and optimization*. SIAM, Philadelphia, 2003.

[58] Z. Guttenberg, H. Müller, H. Habermüller, A. Geisbauer, J. Pipper, J. Felbel, M. Kilepinski, J. Scriba, and A. Wixforth. Planar chip device for PCR and hybridization with surface acoustic wave pump. *Lab on a Chip*, **5**(3):308–317, 2005.

[59] W. Hackbusch. Multi-grid continuation methods. In *Iterative Solution of Nonlinear Systems* (R. Ansorge et al.; eds.), pp. 20-45, Springer, Berlin-Heidelberg-New York, 1982.

[60] J. Haslinger and P. Neittaanmäki. *Finite element approximation for optimal shape design: Theory and applications*. John Wiley & Sons, Chichester, 1988.

[61] J. Haslinger and R.A.E. Mäkinen. *Introduction to shape optimization: Theory, approximation, and computation*. SIAM, Philadelphia, 2004.

[62] M. Heinkenschloss, D. C. Sorensen, and K. Sun. Balanced truncation model reduction for a class of descriptor systems with application to the Oseen equations. *SIAM Journal on Scientific Computing*, **30**(2):1038–1063, 2008.

[63] J. Herskovits, G. Dias, G. Santos, and C. M. Mota Soares. Shape structural optimization with an interior point nonlinear programming algorithm. *Structural and Multidisciplinary Optimization*, **20**(2):107–115, 2000.

[64] M. Hinze and S. Volkwein. Proper orthogonal decomposition surrogate models for nonlinear dynamical systems: Error estimates and suboptimal control. In P. Benner, V. Mehrmann, and D. C. Sorensen, editors, *Dimension Reduction of Large-Scale Systems*, Lecture Notes in Computational Science and Engineering, Vol. **45**, pages 261–306, Heidelberg, 2005. Springer-Verlag.

[65] R.H.W. Hoppe, C. Linsenmann, and H. Antil. Adaptive path following primal dual interior point methods for shape optimization of linear and nonlinear Stokes flow problems. In *Lecture Notes in Computer Science*, Springer, Berlin-Heidelberg-New York, **4818**:259-266, 2008.

[66] R.H.W. Hoppe, C. Linsenmann, and S.I. Petrova. Primal-dual Newton methods in structural optimization. *Comp. Visual. Sci.*, **9**(2):71-87, 2006.

[67] R.H.W. Hoppe and H.D. Mittelmann. A multi-grid continuation strategy for parameter dependent variational inequalities. *J. Comput. Appl. Math.* **26**(1-2):35-46, (1989).

[68] R.H.W. Hoppe and S.I. Petrova. Primal-dual Newton interior-point methods in shape and topology optimization. *Numer. Linear Algebra Appl.*, **11**:413-429, (2004).

[69] R.H.W. Hoppe, S.I. Petrova, and V. Schulz. A primal-dual Newton-type interior-point method for topology optimization. *Journal of Optimization: Theory and Applications*, **114**(3):545-571, 2002.

[70] F. Ihlenburg. *Finite element analysis of acoustic scattering.* Springer, Berlin-Heidelberg-New York, 1998.

[71] C. T. Kelley. *Iterative methods for optimization.* SIAM, Philadelphia, 1999.

[72] B. N. Khoromskij and G. Wittum. *Numerical solution of elliptic differential equations by reduction to the interface.* Lecture Notes in Computational Science and Engineering, Vol. **36**. Springer Verlag, Berlin, Heidelberg, New York, 2004.

[73] P. Knabner, L. Angermann. *Numerik partieller Differentialgleichungen: eine anwendungsorientierte Einführung.* Springer, 2000.

[74] D. Köster. *Numerical simulation of acoustic streaming on SAW-driven biochips.* Ph.D. Thesis, Department of Applied Mathematics, University of Augsburg (2006).

[75] D. Köster. Numerical simulation of acoustic streaming on SAW-driven biochips. *SIAM J. Comp. Sci.*, **29**:2352-2380, 2007.

[76] S. Lall, J. E. Marsden, and S. Glavaški. A subspace approach to balanced truncation for model reduction of nonlinear control systems. *Internat. J. Robust Nonlinear Control*, **12**(6):519–535, 2002.

[77] C. Linsenmann. *Adaptive multilevel-based shape optimization for a stationary Stokes flow by path-following primal-dual interior point methods*. Ph.D. Thesis, Department of Mathematics, University of Houston, Houston, TX, 2008.

[78] D.J. Lucia, P.S. Beran, and W.A. Silva. Reduced-order modeling: New approaches for computational physics. *Progress in Aerospace Sciences*, **40**(1-2):51–117, 2004.

[79] D.J. Lucia, P.I. King, and P.S. Beran. Domain decomposition for reduced-order modeling of a flow with moving shocks. *AIAA Journal* **40**(11), 2360–2362, 2002.

[80] W. G. Litvinov. *Optimization in elliptic problems with applications to mechanics of deformable bodies and fluid mechanics*. Birkhäuser, Basel, 2000.

[81] B. Maar, and V. Schulz. Interior point methods in topology optimization. *Structural Optimization*, **19**(3):214-224, 2000.

[82] G.A. Maugin. *Continuum mechanics of electromagnetic solids*. North-Holland, Amsterdam, 1987.

[83] V. Mehrmann and T. Stykel. Balanced truncation model reduction for large-scale systems in descriptor form. In P. Benner, V. Mehrmann, and D. C. Sorensen, editors, *Dimension Reduction of Large-Scale Systems*, Lecture Notes in Computational Science and Engineering, Vol. **45**:83–115, Heidelberg, 2005. Springer-Verlag.

[84] S. Mehrotra. On the implementation of a primal-dual interior point method. *SIAM J. on Optimization*, **2**:575–601, 1992.

[85] B. Mohammadi and O. Pironneau. *Applied shape optimization for fluids*. Oxford University Press, Oxford, 2001.

[86] B. C. Moore. Principal component analysis in linear systems: controllability, observability, and model reduction. *IEEE Trans. Automat. Control*, **26**(1):17–32, 1981.

[87] J. Nocedal and M. Overton. Projected Hessian updating algorithms for nonlinearly constrained optimization. *SIAM J. Numer. Anal.*, **22**(5):821–850, 1985.

[88] J. Nocedal, A. Wächter, and R.A. Waltz. Adaptive barrier update strategies for nonlinear interior methods. *Research Report RC 23563, IBM T. J. Watson Research Center, Yorktown*, 2006.

[89] Olympus Medical Systems, *private communication*, 2008.

[90] L. F. Pavarino and O. B. Widlund. Balancing Neumann-Neumann methods for incompressible Stokes equations. *Comm. Pure Appl. Math.*, **55**(3):302–335, 2002.

[91] O. Pironneau. *Optimal shape design for elliptic systems*. Springer, Berlin-Heidelberg-New York, 1984.

[92] J. Pollard and B. Castrodale. Outlook for DNA microarrays: Emerging applications and insights on optimizing microarray studies. *Report. Cambridge Health Institute, Cambridge*, 2003.

[93] A. Quarteroni, M. Tuveri, and A. Veneziani. Computational vascular fluid dynamics: Problems, models and methods. *Computing and Visualization in Science*, **2**(4):163–197, 2000.

[94] A. Quarteroni, A. Valli. *Numerical approximation of partial differential equations*. Springer, Berlin-Heidelberg-New York, 1994.

[95] M. Renardy and R.C. Rogers. *An introduction to partial differential equations*. Springer, Berlin-Heidelberg-New York, 1993.

[96] E. M. Rønquist. A domain decomposition solver for the steady Navier-Stokes equations. In A. Ilin and R. Scott, editors, Proceedings of the Third International Conference on Spectral and High-Order Methods. Houston, Texas, June 1995. *Houston Journal of Mathematics*, **95**, 1996.

[97] E. M. Rønquist. Domain decomposition methods for the steady Stokes equations. In C.-H. Lai, P.E. Bjorstad, M. Cross, and O.B. Widlund, editors, *Eleventh International Conference on Domain Decomposition Methods (London, 1998)*, pages 330–340 (electronic). DDM.org, 1999.

[98] C. W. Rowley. Model reduction for fluids, using balanced proper orthogonal decomposition. *Int. J. on Bifurcation and Chaos*, **15**(3):997–1013, 2005.

[99] G. Rozvany. *Structural design via optimality criteria*. Kluwer, Dordrecht, 1989.

[100] S. M. Rump. INTLAB - INTerval LABoratory. In Tibor Csendes, editor, *Developments in Reliable Computing*, pages 77–104. Kluwer Academic Publishers, Dordrecht, 1999. `http://www.ti3.tu-harburg.de/rump/`.

[101] H.M. Shapiro. *Practical flow cytometry*. Wiley-Liss, New York, 2003.

[102] A. R. Shenoy, M. Heinkenschloss, and E. M. Cliff. Airfoil design by an all-at-once approach. *Int. J. Comput. Fluid Mechanics*, **11**(1):3–25, 1998.

[103] B. Smith, P. Bjørstad, W. Gropp. *Domain decomposition: Parallel multilevel methods for elliptic partial differential equations*. Cambridge University Press, Cambridge, London, New York, 1996.

[104] J. Sokolowski and J.P. Zolesio. *Introduction to shape optimization*. Springer, Berlin-Heidelberg-New York, 1992.

[105] T.M. Squires, and S.R. Quake. Microfluidics: Fluid physics at the nanoliter scale. *Rev. Mod. Phys.*, **77**(3):977-1026, 2005.

[106] R. Stavre. Distributed control of a heat conducting, time-dependent, Navier-Stokes fluid. *Glasg. Math. J.*, **44**(2):191–200, 2002.

[107] T. Stykel. Balanced truncation model reduction for semidiscretized Stokes equation. *Linear Algebra Appl.*, **415**(2-3):262–289, 2006.

[108] K. Sun. *Domain decomposition and model reduction for large-scale dynamical systems*. Ph.D. Thesis, Department of Computational and Applied Mathematics, Rice University, Houston, TX, 2008.

[109] K. Sun, R. Glowinski, M. Heinkenschloss, and D. C. Sorensen. Domain decomposition and model reduction of systems with local nonlinearities. In K. Kunisch, G. Of, and O. Steinbach, editors, *Numerical Mathematics And Advanced Applications. ENUMATH 2007*, pages 389–396, Heidelberg, 2008. Springer-Verlag.

[110] P. Le Tallec and A. Patra. Non-overlapping domain decomposition methods for adaptive *hp* approximations of the Stokes problem with discontinuous pressure fields. *Comput. Methods Appl. Mech. Engrg.*, **145**(3-4):361–379, 1997.

[111] A.L. Tits, A. Wächter, S. Bakhtiari, T.J. Urban, and C.T. Lawrence. A primal-dual interior-point method for nonlinear programming with strong global and local convergence properties. *SIAM J. on Optimization*, **14**:173-199, 2003.

[112] A. Toselli and O. Widlund. *Domain decomposition methods - Algorithms and theory*. Computational Mathematics, Vol. **34**. Springer–Verlag, Berlin, 2004.

[113] M. Ulbrich, S. Ulbrich, and L. Vicente. A globally convergent primal-dual interior point filter method for nonconvex nonlinear programming. *Math. Programming*, **100**(2):379-410, 2004.

[114] R.J. Vanderbei and D.F. Shanno. An interior point algorithm for nonconvex nonlinear programming. *Computational Optimization and Applications*, **13**(1):231-252, 1999.

[115] A. Wächter and L.T. Biegler. Line search filter methods for nonlinear programming: Motivation and global convergence. *SIAM J. on Optimization*, **16**(1):1-31, 2005.

[116] G.M. Whitesides. The origins and the future of microfluidics. *Nature*, **442**(7101):368–373, 2006.

[117] G. Wittum. On the convergence of multigrid iterations with transforming smoothers. Theory with applications to the Navier-Stokes equations. *Numer. Math.*, **57**(1):15-38, 1989.

[118] A. Wixforth. Acoustically driven programmable microfluidics for biological and chemical applications. *JALA*, **11**(6):399-405, 2006.

[119] A. Wixforth, J. Scriba, and G. Gauer. Flatland fluidics. *Mst News*, **5**:42-43, 2002.

[120] M.H. Wright. Interior methods for constrained optimization. *Acta Numerica*, **1**:341-407, 1992.

[121] S. J. Wright. *Primal-dual interior-point methods*. SIAM, Philadelphia, 1997.

[122] L.Y. Yeo and J.R. Friend. Ultrafast microfluidics using surface acoustic waves. *Biomicrofluidics*, **3**:012002-012023, 2009.

[123] E. Zeidler. *Nonlinear functional analysis and its applications. II/A: Linear monotone operators*. Springer, Berlin-Heidelberg-New York, 1990.

[124] K. Zhou, J. C. Doyle, and K. Glover. *Robust and optimal control*. Prentice Hall, Englewood Cliffs, 1996.