

Reduced order modeling based shape optimization of surface acoustic wave driven microfluidic biochips[☆]

Harbir Antil^a, Matthias Heinkenschloss^b, Ronald H.W. Hoppe^{a,c,*},
Christopher Linsenmann^c, Achim Wixforth^d

^a *Department of Mathematics, University of Houston, Houston, TX 77204-3008, USA*

^b *Comput. and Appl. Mathematics, Rice University, Houston, TX 77005-1892, USA*

^c *Institute of Mathematics, University of Augsburg, D-86159 Augsburg, Germany*

^d *Institute of Physics, University of Augsburg, D-86159 Augsburg, Germany*

[☆] The work of the authors has been supported in part by NSF grants DMS-0511624, DMS-0707602, DMS-0810176, DMS-0811153, DMS-0914788, by AFOSR grant FA9550-09-1-0225, and by the German National Science Foundation (DFG) within the Priority Program SPP 1253.

* Corresponding author at: Department of Mathematics, University of Houston, 651 P.G. Hoffman, Houston, TX 77204-3008, USA.
Tel.: +1 7137433452; fax: +1 7137433505.

E-mail address: rohop@math.uh.edu (R.H.W. Hoppe).

1. Introduction

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 [28,31]. 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 (left)).

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 [15,21,33–35]. 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 (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.

This paper is concerned with the numerical solution of a shape optimization problem in which pressure driven capillary barriers between the microchannels and the reservoirs are determined such that a very precise and bubble-free filling of the reservoirs can be guaranteed. This leads to an optimization problem governed by a system of time dependent partial differential equations (PDEs), the so-called state PDEs. The numerical solution of such problems using gradient based methods involves a coupled system of PDEs consisting of the forward in time system of state PDEs and the backward in time system of adjoint PDEs. Because of the size of this coupled system of state and adjoint PDEs, the numerical solution is very costly, both in terms of computing time and memory requirements. We will apply model reduction to replace the original state equation by a much smaller system of ordinary differential equations (ODEs) in such a way that the solution of the optimization problem governed by the reduced order system has approximately the same solution as the original optimization problem. This is achieved by a combination of balanced truncation model reduction and domain decomposition and a careful exploitation of the problem structure.

The paper is organized as follows: in Section 2, we introduce the mathematical model describing the operational behavior of microfluidic biochips with emphasis on the SAW induced fluid flow as described by the compressible Navier–Stokes equations and on the homogenization approach for a proper separation of the time scales. Section 3 is devoted to the optimal design of microfluidic biochips based on the optimization of objective functionals subject to the equations describing the acoustic streaming. For model reduction purposes, we present the basic idea of balanced truncation model reduction and its combination with the domain decomposition methodology. Finally, in Section 4 we provide a detailed documentation of simulation results that illustrate both the validity of our model as well as the feasibility of the model reduction based optimization.

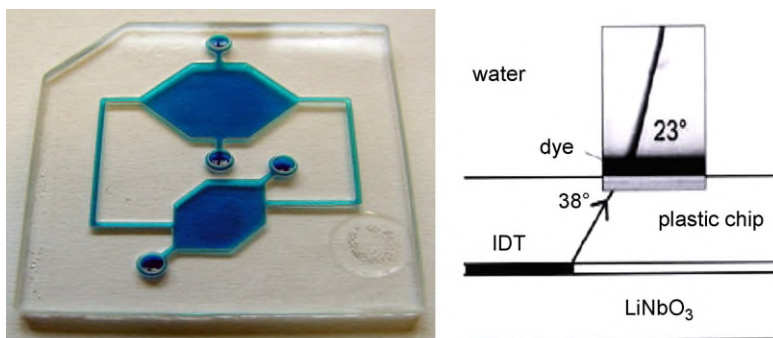


Fig. 1. Microfluidic biochip (left) and sharp jet created by surface acoustic waves (right).

2. Modeling and simulation of piezoelectrically actuated acoustic streaming in microfluidic biochips

The operational behavior of SAW driven microfluidic biochips is modeled by a coupled system consisting of the linearized equations of piezoelectricity and the compressible Navier–Stokes equations (see, e.g. [1,2]). In the sequel, we will focus on the SAW induced fluid flow in the fluidic network on top of the biochip. For a detailed description of the piezoelectrics, we refer to [16].

We denote by $\Omega(t) \subset \mathbb{R}^2$, $t \in [0, T]$, the time dependent domain occupied by the fluid with boundary $\Gamma(t) = \bar{\Gamma}_D(t) \cup \bar{\Gamma}_N(t)$, $\Gamma_D(t) \cap \Gamma_N(t) = \emptyset$. Here, $\Gamma_D(t)$ is that part of the boundary where the SAW enter the fluid filled microchannels. We denote by \mathbf{v} and p the velocity and the pressure, and we refer to ρ , η , and ξ as the density of the fluid, the standard and bulk viscosities. The pair (\mathbf{v}, p) satisfies the following initial-boundary value problem

$$\rho \left(\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} \right) = \nabla \cdot \boldsymbol{\sigma} \quad \text{in } \Omega(t), t \in (0, T], \quad (1a)$$

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 \quad \text{in } \Omega(t), t \in (0, T], \quad (1b)$$

$$\mathbf{v}(\cdot + \mathbf{u}(\cdot, t), t) = \frac{\partial \mathbf{u}}{\partial t}(\cdot, t) \quad \text{on } \Gamma_D(t), t \in (0, T], \quad (1c)$$

$$\boldsymbol{\sigma} \mathbf{n} = 0 \quad \text{on } \Gamma_N(t), t \in (0, T], \quad (1d)$$

$$\mathbf{v}(\cdot, 0) = \mathbf{v}_0, \quad p(\cdot, 0) = p_0 \quad \text{in } \Omega(0), \quad (1e)$$

where $\boldsymbol{\sigma} = (\sigma_{ij})_{i,j=1}^2$, $\sigma_{ij} := -p\delta_{ij} + 2\eta\varepsilon_{ij}(\mathbf{v}) + \delta_{ij}(\xi - 2\eta/3)\nabla \cdot \mathbf{v}$ and \mathbf{u} in (1c) 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 (see [16]).

Since the deflection of the walls of the microchannels by the SAW is approximately 10^{-9} m compared to lengths, widths, and heights of the microchannels in the range of micrometer to millimeter, we will henceforth neglect the time dependence of Ω .

As mentioned in Section 1, the SAW induced fluid flow exhibits two different time scales. When the SAW enter the fluid filled microchannels, sharp jets are created within nanoseconds (cf. Fig. 1 (right)). The SAW propagate along the channels and experience a significant damping which results in an almost stationary flow pattern (acoustic streaming). This relaxation process happens on a time scale of milliseconds. The multiscale character of the problem can be appropriately taken care of by a homogenization approach. Following [1,25], 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

$$\rho = \rho_0 + \varepsilon\rho' + \varepsilon^2\rho'' + O(\varepsilon^3),$$

$$\mathbf{v} = \mathbf{v}_0 + \varepsilon\mathbf{v}' + \varepsilon^2\mathbf{v}'' + O(\varepsilon^3),$$

$$p = p_0 + \varepsilon p' + \varepsilon^2 p'' + O(\varepsilon^3).$$

We assume $\mathbf{v}_0 \equiv 0$ and $p_0 \equiv 0$ (fluid at rest, if no SAW actuation) and choose ρ_0 as the density of the carrier fluid. Then, collecting all terms of order $O(\varepsilon)$ and setting $\rho_1 = \varepsilon\rho'$, $\mathbf{v}_1 := \varepsilon\mathbf{v}'$, $p_1 := \varepsilon p'$, we find that the triple $(\rho_1, \mathbf{v}_1, p_1)$ satisfies the following initial-boundary value problem for linear compressible Navier–Stokes equations with time periodic boundary conditions on Γ_D

$$\rho_0 \frac{\partial \mathbf{v}_1}{\partial t} - \nabla \cdot \boldsymbol{\sigma}_1 = \mathbf{0} \quad \text{in } \Omega \times (0, T_1], \quad (2a)$$

$$\frac{\partial \rho_1}{\partial t} + \rho_0 \nabla \cdot \mathbf{v}_1 = 0 \quad \text{in } \Omega \times (0, T_1], \quad (2b)$$

$$\mathbf{v}_1 = \mathbf{g}_1 \quad \text{on } \Gamma_D \times (0, T_1], \quad (2c)$$

$$\boldsymbol{\sigma}_1 \mathbf{n} = 0 \quad \text{on } \Gamma_N \times (0, T_1], \quad (2d)$$

$$\mathbf{v}_1(\cdot, 0) = \mathbf{0}, \quad p_1(\cdot, 0) = 0 \quad \text{in } \Omega, \quad (2e)$$

where $T_1 := 2\pi/\omega$ with ω being the angular frequency of the time harmonic SAW excitation, $\mathbf{g}_1 := \partial\mathbf{u}/\partial t$ in (2c), and

$$\boldsymbol{\sigma}_1 = ((\sigma_1)_{ij})_{i,j=1}^2, \quad (\sigma_1)_{ij} := -p_1\delta_{ij} + 2\eta\varepsilon_{ij}(\mathbf{v}_1) + \delta_{ij}(\xi - 2\eta/3)\nabla \cdot \mathbf{v}_1.$$

Moreover, p_1 and ρ_1 are related by the constitutive equation

$$p_1 = c_0^2\rho_1 \quad \text{in } \Omega \times (0, T_1]. \tag{3}$$

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.

For the weak formulation of (2a)–(2e), we adopt standard notation from Lebesgue and Sobolev space theory [26]. In particular, bold letters will refer to Lebesgue and Sobolev spaces of vector-valued functions. We substitute ρ_1 in (2b) by means of (3) and introduce the function spaces

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

We note that $H^1((0, T_1); \mathbf{H}^{-1}(\Omega)) \cap L^2((0, T_1); \mathbf{H}^1(\Omega))$ is continuously embedded in $C([0, T_1], \mathbf{L}^2(\Omega))$. The weak formulation of (2a)–(2e) amounts to the computation of $(\mathbf{v}_1, p_1) \in \mathbf{V}_{\mathbf{g}_1} \times W$ such that

$$\left\langle \rho_0 \frac{\partial \mathbf{v}_1}{\partial t}, \mathbf{w} \right\rangle + a(\mathbf{v}_1, \mathbf{w}) + b(p_1, \mathbf{w}) = \mathbf{0}, \quad \mathbf{w} \in \mathbf{H}_{0,\Gamma_D}^1(\Omega), \tag{4a}$$

$$\left(\rho_0^{-1} c_0^{-2} \frac{\partial p_1}{\partial t}, q \right)_{0,\Omega_2} - b(q, \mathbf{v}_1) = 0, \quad q \in L^2(\Omega), \tag{4b}$$

$$\mathbf{v}_1(\cdot, 0) = \mathbf{0}, \quad p_1(\cdot, 0) = 0. \tag{4c}$$

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

$$a(\mathbf{v}, \mathbf{w}) := \eta \int_{\Omega_2} \nabla \mathbf{v} : \nabla \mathbf{w} \, dx + \left(\xi + \frac{\eta}{3} \right) \int_{\Omega_2} \nabla \cdot \mathbf{v} \nabla \cdot \mathbf{w} \, dx, \tag{5a}$$

$$b(p, \mathbf{w}) := - \int_{\Omega_2} p \nabla \cdot \mathbf{w} \, dx. \tag{5b}$$

Theorem 1. *For the solution of the variational problem (4a)–(4c) there holds:*

If $\mathbf{g}_1 \in \mathbf{L}^2((0, T); \mathbf{H}_{00}^{1/2}(\Gamma_D))$, then there exists a unique solution $(\mathbf{v}_1, p_1) \in \mathbf{V}_{\mathbf{g}_1} \times W$ of (4a)–(4c) satisfying the stability estimate

$$\|(\mathbf{v}_1, p_1)\|_{\mathbf{V}_{\mathbf{g}_1} \times W} \leq C_{T_1} \|\mathbf{g}_1\|_{\mathbf{L}^2((0,T); \mathbf{H}_{00}^{1/2}(\Gamma_D))}, \tag{6}$$

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

Proof. The existence can be shown by the Galerkin method, whereas the uniqueness and the stability estimate (6) can be derived using the ellipticity of the bilinear form $a(\cdot, \cdot)$ and the fact that the bilinear form $b(\cdot, \cdot)$ satisfies an inf-sup condition. \square

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

$$\rho_0 \frac{\partial \mathbf{v}_2}{\partial t} - \nabla \cdot \boldsymbol{\sigma}_2 = \left\langle -\rho_1 \frac{\partial \mathbf{v}_1}{\partial t} - \rho_0 [\nabla \mathbf{v}_1] \mathbf{v}_1 \right\rangle \quad \text{in } \Omega \times (0, T], \tag{7a}$$

$$\rho_0 \nabla \cdot \mathbf{v}_2 = \langle -\nabla \cdot (\rho_1 \mathbf{v}_1) \rangle \quad \text{in } \Omega \times (0, T], \tag{7b}$$

$$\mathbf{v}_2 = \mathbf{g}_2 \quad \text{on } \Gamma_D \times (0, T], \tag{7c}$$

$$\boldsymbol{\sigma}_2 \mathbf{n} = 0 \quad \text{on } \Gamma_N \times (0, T], \quad (7d)$$

$$\mathbf{v}_2(\cdot, 0) = \mathbf{0}, \quad p_2(\cdot, 0) = 0 \quad \text{in } \Omega, \quad (7e)$$

where $\mathbf{g}_2 := -\langle [\nabla \mathbf{v}_1] \mathbf{u} \rangle$ in (7c) and

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

The density ρ_2 can be obtained via the constitutive equation

$$p_2 = c_0^2 \rho_2 \quad \text{in } \Omega \times (0, T]. \quad (8)$$

The compressible Stokes system (7a)–(7c) is used as a model for the acoustic streaming.

The weak formulation of (7a)–(7c) requires the computation of $(\mathbf{v}_2, p_2) \in \mathbf{V}_{\mathbf{g}_2} \times W$, where

$$\mathbf{V}_{\mathbf{g}_2} := \{\mathbf{v} \in H^1((0, T), \mathbf{H}^{-1}(\Omega)) \cap L^2((0, T), \mathbf{H}^1(\Omega)) \mid \mathbf{v}|_{\Gamma_D} = \mathbf{g}_2\},$$

$$W := H^1((0, T_2); L^2(\Omega)),$$

such that

$$\left\langle \rho_0 \frac{\partial \mathbf{v}_2}{\partial t}, \mathbf{w} \right\rangle + a(\mathbf{v}_2, \mathbf{w}) + b(p_2, \mathbf{w}) = (\mathbf{f}, \mathbf{w})_{0,\Omega} \quad \mathbf{w} \in \mathbf{H}_{0,\Gamma_D}^1(\Omega), \quad (9a)$$

$$b(q, \mathbf{v}_2) = (f, q)_{0,\Omega}, \quad q \in L^2(\Omega), \quad (9b)$$

$$\mathbf{v}_2(\cdot, 0) = \mathbf{0}, \quad p_2(\cdot, 0) = 0. \quad (9c)$$

Here, the bilinear forms $a(\cdot, \cdot)$, $b(\cdot, \cdot)$ are as in (5a), (5b), and the right-hand sides \mathbf{f} , f are given by

$$\mathbf{f} := - \left\langle \rho_1 \frac{\partial \mathbf{v}_1}{\partial t} + \rho_0 [\nabla \mathbf{v}_1] \mathbf{v}_1 \right\rangle, \quad f = -(\rho_0^{-1} \nabla \cdot (\rho_1 \mathbf{v}_1)).$$

Theorem 2. *If $\mathbf{f} \in \mathbf{L}^2(\Omega)$, $f \in L^2(\Omega)$, and $\mathbf{g}_2 \in \mathbf{H}_{00}^{1/2}(\Gamma_D)$, the weak formulation (9a), (9b) 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*

$$\|(\mathbf{v}_2, p_2)\|_{\mathbf{V}_{\mathbf{g}_2} \times W} \leq C_{T_2} (\|\mathbf{f}\|_{0,\Omega} + \|f\|_{0,\Omega} + \|\mathbf{g}_2\|_{\mathbf{H}_{00}^{1/2}(\Gamma_D)}). \quad (10)$$

Proof. The proof follows along the same lines as that of Theorem 1. \square

3. Reduced order modeling based shape optimization

The performance of SAW driven microfluidic biochips can be significantly improved by an optimal design of the walls of the microchannels and the reservoirs as well as by the optimal placement of the IDT in order to obtain a maximum pumping rate. Another design issue is the shape optimization of pressure driven capillary barriers between the microchannels and the reservoirs such that a very precise and bubble-free filling of the reservoirs can be guaranteed. For the efficient solution of these optimal design problems, in previous work [5–7] 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 [7], the amount of computational time is still significant, and there is a need for further reductions. Such reductions can be achieved by replacing the original compressible Stokes system (7a)–(7e) in the optimization problem by a reduced order model. Many approaches for the computation of reduced order models exist, such as Proper Orthogonal Decomposition (POD), Balanced Truncation Model Reduction (BTMR), Krylov subspace methods, or reduced basis methods. See, e.g., [11]. A crucial requirement is that the error between the solution of the original optimization problem and the optimization problem governed by the reduced order equations can be controlled and made small. To achieve this, one needs to use a model reduction approach which provides a bound for the error between the so-called input-to-output maps of

the original full order model and the reduced order model. We use BTMR because it provides an error estimate and therefore our approach provides a-priori estimates for the error in the solutions and lets the user control the size of the error estimate. Alternatively, one could consider other approaches, such as interpolation based methods [9,19]. We comment more on this issue in Remark 4 below.

We use BTMR combined with domain decomposition (DD). BTMR is used because of the existence of error bounds. However, these error bounds only exist for linear time invariant problems. In our case, the design parameters determine the domain Ω . Therefore the system is not linear time invariant. Fortunately, only a small portion of the domain, corresponding to the capillary barriers between the microchannels and the reservoirs are modified. Therefore we use DD to decompose the problem into a part which is governed by a linear time invariant system and a small part corresponding to the subdomain in which the domain can be changed. BTMR is then applied to the large subproblem governed by the linear time invariant system. For the standard time-dependent Stokes system, such a combined domain decomposition and balanced truncation model reduction (DDBTMR) approach has been developed and analyzed in [4] (see also [3] for DDBTMR applied to convection–diffusion problems). Here, we will consider DDBTMR for the compressible Stokes system (7a)–(7e).

3.1. Semi-discretization of the compressible Stokes system

Let $\Omega(\theta) \subset \mathbb{R}^2$ be a bounded domain that depends on design variables $\theta = (\theta_1, \dots, \theta_d)^T \in \Theta$, where $\theta_i, 1 \leq i \leq d$, are the Bézier control points of a Bézier curve representation of the boundary and

$$\Theta := \{\theta_i \in \mathbb{R} \mid \theta_{min}^{(i)} \leq \theta_i \leq \theta_{max}^{(i)}, 1 \leq i \leq d\}.$$

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 $\Gamma_{lat}(\theta)$ which are mutually disjoint such that

$$\partial\Omega(\theta) = \bar{\Gamma}_{in}(\theta) \cup \bar{\Gamma}_{out}(\theta) \cup \bar{\Gamma}_{lat}(\theta).$$

We set

$$\begin{aligned} Q(\theta) &:= \Omega(\theta) \times (0, T), & \Sigma(\theta) &:= \partial\Omega(\theta) \times (0, T), \\ \Sigma_{in}(\theta) &:= \Gamma_{in}(\theta) \times (0, T), & \Sigma_{lat}(\theta) &:= \Gamma_{lat}(\theta) \times (0, T) \end{aligned} \quad T > 0,$$

and consider shape optimization problems of the form

$$\min_{\theta \in \Theta} J(\theta) := \int_0^T \int_{\Omega(\theta)} \ell(\mathbf{v}(\theta), p(\theta), x, t) dx dt, \tag{11}$$

subject to (7a)–(7e), where the integrand ℓ in the objective functional is chosen according to the individual design problem.

For the spatial discretization of the time-dependent Stokes system we may use one of the standard methods such as P2-P1 Taylor Hood elements or discontinuous pressure elements [12,13,17]. We will discuss this in more detail in Subsection 3.3. 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

$$\min_{\theta \in \Theta} J(\theta) := \int_0^T \ell(\mathbf{v}, \mathbf{p}, t, \theta) dt, \tag{12a}$$

where \mathbf{v}, \mathbf{p} solve

$$\mathbf{E}(\theta) \frac{d}{dt} \begin{pmatrix} \mathbf{v}(t) \\ \mathbf{p}(t) \end{pmatrix} + \mathbf{S}(\theta) \begin{pmatrix} \mathbf{v}(t) \\ \mathbf{p}(t) \end{pmatrix} = \begin{pmatrix} \mathbf{g}_1^{(\theta)}(t) \\ \mathbf{g}_2^{(\theta)}(t) \end{pmatrix} \quad t \in (0, T), \tag{12b}$$

$$\mathbf{M}(\theta)\mathbf{v}(0) = \mathbf{v}^{(0)}, \tag{12c}$$

$$-\mathbf{B}(\theta)\mathbf{M}(\theta)^{-1}\mathbf{v}^{(0)} + \mathbf{g}_2^{(\theta)}(0) = \mathbf{0}. \tag{12d}$$

Here, the integrand $\ell(\cdot)$ in (12a) results from the spatial discretization of the integral ℓ in the objective functional (11). The block matrix $\mathbf{E}(\theta)$ and the discrete Stokes operator $\mathbf{S}(\theta)$ in (12b) are given by

$$\mathbf{E}(\theta) := \begin{pmatrix} \mathbf{M}(\theta) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \quad \mathbf{S}(\theta) := \begin{pmatrix} \mathbf{A}(\theta) & \mathbf{B}^T(\theta) \\ \mathbf{B}(\theta) & \mathbf{0} \end{pmatrix}, \tag{13}$$

where $\mathbf{M}(\theta) \in \mathbb{R}^{n \times n}$, $\mathbf{A}(\theta) \in \mathbb{R}^{n \times n}$ and $\mathbf{B}(\theta) \in \mathbb{R}^{m \times n}$ denote 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 (12b) stems from the semi-discretization of the compressibility condition and the boundary condition at the inflow boundary. The vector $\mathbf{v}^{(0)}$ refers to the initial velocity satisfying (12d). 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 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 (12b), (12c) as well as its continuous dependence on the data of the problem is a consequence of the following result [4] which also plays a prominent role with regard to the application of BTMR and the derivation of upper estimates for the modeling error.

Theorem 3. *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 $\text{Ker} \mathbf{B}$, i.e., there exists a constant $\alpha > 0$ such that

$$\mathbf{v}^T \mathbf{A} \mathbf{v} \geq \alpha \|\mathbf{v}\|^2, \quad \mathbf{v} \in \text{Ker} \mathbf{B}. \tag{14}$$

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

Consider the initial value problem

$$\mathbf{E} \frac{d}{dt} \begin{pmatrix} \mathbf{v}(t) \\ \mathbf{p}(t) \end{pmatrix} + \mathbf{S} \begin{pmatrix} \mathbf{v}(t) \\ \mathbf{p}(t) \end{pmatrix} = \begin{pmatrix} \mathbf{g}_1(t) \\ \mathbf{g}_2(t) \end{pmatrix}, \quad t \in (0, T], \tag{15a}$$

$$\mathbf{M}\mathbf{v}(0) = \mathbf{v}^{(0)}, \tag{15b}$$

where \mathbf{E}, \mathbf{S} are as in (13) 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 $-\mathbf{B}\mathbf{M}^{-1}\mathbf{v}^{(0)} + \mathbf{g}_2(0) = \mathbf{0}$. Under the assumptions (i), (ii) and (iii), the initial value problem (15) has a unique solution $(\mathbf{v}, \mathbf{p}) \in C([0, T]; \mathbb{R}^n) \times C([0, T]; \mathbb{R}^m / (\text{Ker} \mathbf{B}^T))$, and there exist constants $C_1 \geq 0, C_2 \geq 0$, depending on $\mathbf{A}, \mathbf{B}, \mathbf{M}$ such that

$$\begin{aligned} \|\mathbf{v}\|_{L^2} &\leq C_1 \|\mathbf{v}^{(0)}\| + C_2 (\|\mathbf{g}_1\|_{L^2} + \|\mathbf{g}_2\|_{L^2}), \\ \|\mathbf{p}\|_{L^2} &\leq C_1 \|\mathbf{v}^{(0)}\| + C_2 \left(\|\mathbf{g}_1\|_{L^2} + \|\mathbf{g}_2\|_{L^2} + \left\| \frac{d}{dt} \mathbf{g}_2 \right\|_{L^2} \right). \end{aligned}$$

Proof. We refer to the proof of Theorem 2.1 in [4]. \square

3.2. Balanced truncation model reduction

Balanced truncation model reduction (BTMR) is a model reduction technique that replaces a large-scale dynamical system by a system of substantially lower dimension while keeping almost the same response characteristics (cf., e.g. [8,11] and the references therein). Originally developed for systems of ordinary differential equations, it has recently

been extended to descriptor systems [23,27,29,32]. In particular, BTMR for semidiscretized Stokes and linearized Navier–Stokes systems has been studied in [4,23,29,32]. We consider

$$\mathbf{M} \frac{d}{dt} \mathbf{v}(t) = -\mathbf{A} \mathbf{v}(t) - \mathbf{B}^T \mathbf{p}(t) + \mathbf{K} \mathbf{u}(t), \quad t \in (0, T), \quad (16a)$$

$$0 = -\mathbf{B} \mathbf{v}(t) + \mathbf{L} \mathbf{u}(t), \quad t \in (0, T), \quad (16b)$$

$$\mathbf{z}(t) = \mathbf{C} \mathbf{v}(t) + \mathbf{F} \mathbf{p}(t) + \mathbf{D} \mathbf{u}(t), \quad t \in (0, T), \quad (16c)$$

$$\mathbf{M} \mathbf{v}(0) = \mathbf{v}^{(0)}, \quad (16d)$$

$$\mathbf{B} \mathbf{M}^{-1} \mathbf{v}^{(0)} = \mathbf{L} \mathbf{u}(0), \quad (16e)$$

and

$$-\mathbf{M} \frac{d}{dt} \boldsymbol{\lambda}(t) = -\mathbf{A}^T \boldsymbol{\lambda}(t) - \mathbf{B}^T \boldsymbol{\kappa}(t) + \mathbf{C}^T \mathbf{w}(t), \quad t \in (0, T), \quad (17a)$$

$$0 = -\mathbf{B} \boldsymbol{\lambda}(t) + \mathbf{F}^T \mathbf{w}(t), \quad t \in (0, T), \quad (17b)$$

$$\mathbf{q}(t) = \mathbf{K}^T \boldsymbol{\lambda}(t) + \mathbf{L}^T \boldsymbol{\kappa}(t) + \mathbf{D}^T \mathbf{w}(t), \quad t \in (0, T), \quad (17c)$$

$$\mathbf{M} \boldsymbol{\lambda}(T) = \boldsymbol{\lambda}^{(T)}, \quad (17d)$$

$$\mathbf{B} \mathbf{M}^{-1} \boldsymbol{\lambda}^{(T)} = \mathbf{F}^T \mathbf{w}(T), \quad (17e)$$

where the matrices $\mathbf{M} \in \mathbb{R}^{n \times n}$, $\mathbf{A} \in \mathbb{R}^{n \times n}$ and $\mathbf{B} \in \mathbb{R}^{m \times n}$ are assumed to satisfy the assumptions of Theorem 3. Moreover, $\mathbf{K} \in \mathbb{R}^{n \times r}$, $\mathbf{L} \in \mathbb{R}^{m \times r}$, $\mathbf{C} \in \mathbb{R}^{s \times n}$, $\mathbf{F} \in \mathbb{R}^{s \times m}$, and $\mathbf{D} \in \mathbb{R}^{s \times r}$. The system (17) is the adjoint system corresponding to (16). We will see later that systems of the type (17) and (16) arise within the optimality system corresponding to (12).

BTMR can be applied to (17) and (16) by eliminating the variables \mathbf{p} and $\boldsymbol{\kappa}$ via projection. We choose $\mathbf{v}(t) = \mathbf{v}_H(t) + \mathbf{v}_P(t)$, where

$$\mathbf{v}_P(t) = \mathbf{M}^{-1} \mathbf{B}^T (\mathbf{B} \mathbf{M}^{-1} \mathbf{B}^T)^{-1} \mathbf{L} \mathbf{u}(t) \quad (18)$$

is a particular solution of (16b) and $\mathbf{v}_H(t)$ satisfies $\mathbf{0} = \mathbf{B} \mathbf{v}_H(t)$. If we insert $\mathbf{v}(t) = \mathbf{v}_H(t) + \mathbf{v}_P(t)$, (18) into (16a–c), we obtain

$$\mathbf{M} \frac{d \mathbf{v}_H}{dt}(t) = -\mathbf{A} \mathbf{v}_H(t) - \mathbf{B}^T \mathbf{p}(t) + \tilde{\mathbf{B}} \mathbf{u}(t) - \mathbf{B}^T (\mathbf{B} \mathbf{M}^{-1} \mathbf{B}^T)^{-1} \mathbf{L} \frac{d \mathbf{u}}{dt}(t), \quad (19a)$$

$$\mathbf{0} = \mathbf{B} \mathbf{v}_H(t), \quad (19b)$$

$$\mathbf{z}(t) = \mathbf{C} \mathbf{v}_H(t) + \mathbf{F} \mathbf{p}(t) + (\mathbf{D} + \mathbf{C} \mathbf{M}^{-1} \mathbf{B}^T (\mathbf{B} \mathbf{M}^{-1} \mathbf{B}^T)^{-1} \mathbf{L}) \mathbf{u}(t), \quad (19c)$$

where $\tilde{\mathbf{B}} = \mathbf{K} - \mathbf{A} \mathbf{M}^{-1} \mathbf{B}^T (\mathbf{B} \mathbf{M}^{-1} \mathbf{B}^T)^{-1} \mathbf{L}$. Eq. (19a and b) imply that

$$\mathbf{p}(t) = (\mathbf{B} \mathbf{M}^{-1} \mathbf{B}^T)^{-1} \left(-\mathbf{B} \mathbf{M}^{-1} \mathbf{A} \mathbf{v}_H(t) + \mathbf{B} \mathbf{M}^{-1} \tilde{\mathbf{B}} \mathbf{u}(t) - \mathbf{L} \frac{d \mathbf{u}}{dt}(t) \right), \quad (20)$$

and $\boldsymbol{\Pi}^T \mathbf{v}_H(t) = \mathbf{v}_H(t)$, where $\boldsymbol{\Pi} = \mathbf{I} - \mathbf{B}^T (\mathbf{B} \mathbf{M}^{-1} \mathbf{B}^T)^{-1} \mathbf{B} \mathbf{M}^{-1}$ is an oblique projector that satisfies $\boldsymbol{\Pi}^2 = \boldsymbol{\Pi}$, $\boldsymbol{\Pi} \mathbf{M} = \mathbf{M} \boldsymbol{\Pi}^T$, $\text{null}(\boldsymbol{\Pi}) = \text{range}(\mathbf{B}^T)$ and $\text{range}(\boldsymbol{\Pi}) = \text{null}(\mathbf{B} \mathbf{M}^{-1})$. Inserting (20) into (19a and c), it follows that

$$\boldsymbol{\Pi} \mathbf{M} \boldsymbol{\Pi}^T \frac{d}{dt} \mathbf{v}_H(t) = -\boldsymbol{\Pi} \mathbf{A} \boldsymbol{\Pi}^T \mathbf{v}_H(t) + \boldsymbol{\Pi} \tilde{\mathbf{B}} \mathbf{u}(t), \quad (21a)$$

$$\mathbf{z}(t) = \tilde{\mathbf{C}} \boldsymbol{\Pi}^T \mathbf{v}_H(t) + \tilde{\mathbf{D}} \mathbf{u}(t) - \mathbf{F} (\mathbf{B} \mathbf{M}^{-1} \mathbf{B}^T)^{-1} \mathbf{L} \frac{d}{dt} \mathbf{u}(t), \quad (21b)$$

where $\tilde{\mathbf{C}} = \mathbf{C} - \mathbf{F} (\mathbf{B} \mathbf{M}^{-1} \mathbf{B}^T)^{-1} \mathbf{B} \mathbf{M}^{-1} \mathbf{A}$ and $\tilde{\mathbf{D}} = \mathbf{D} + \mathbf{C} \mathbf{M}^{-1} \mathbf{B}^T (\mathbf{B} \mathbf{M}^{-1} \mathbf{B}^T)^{-1} \mathbf{L} + \mathbf{F} (\mathbf{B} \mathbf{M}^{-1} \mathbf{B}^T)^{-1} \mathbf{B} \mathbf{M}^{-1} \tilde{\mathbf{B}}$. Setting $\mathbf{v}^{(0)} = \boldsymbol{\Pi} \mathbf{v}^{(0)} + (\mathbf{I} - \boldsymbol{\Pi}) \mathbf{v}^{(0)}$ and using (16e), we obtain the initial condition

$$\boldsymbol{\Pi} \mathbf{M} \mathbf{v}_H(0) = \boldsymbol{\Pi} \mathbf{M} \boldsymbol{\Pi}^T \mathbf{v}_H(0) = \boldsymbol{\Pi} \mathbf{v}^{(0)} (=:\mathbf{v}_H^{(0)}). \quad (21c)$$

In the same way, we transform (17). Setting $\lambda = \lambda_H(t) + \lambda_P(t)$, where $\lambda_P(t) = \mathbf{M}^{-1} \mathbf{B}^T (\mathbf{B} \mathbf{M}^{-1} \mathbf{B}^T)^{-1} \mathbf{F}^T \mathbf{w}(t)$ and where λ_H solves

$$-\mathbf{\Pi} \mathbf{M} \mathbf{\Pi}^T \frac{d}{dt} \lambda_H(t) = -\mathbf{\Pi} \mathbf{A}^T \mathbf{\Pi}^T \lambda_H(t) + \mathbf{\Pi} \tilde{\mathbf{C}}^T \mathbf{w}(t), \quad (22a)$$

$$\mathbf{q}(t) = \tilde{\mathbf{B}}^T \mathbf{\Pi}^T \lambda_H(t) + \tilde{\mathbf{D}}^T \mathbf{w}(t) + \mathbf{L}^T (\mathbf{B} \mathbf{M}^{-1} \mathbf{B}^T)^{-1} \mathbf{F}^T \frac{d\mathbf{w}}{dt}(t), \quad (22b)$$

$$\mathbf{M} \lambda_H(T) = \mathbf{\Pi} \lambda^{(T)}. \quad (22c)$$

The derivation of (22) from (17) is analogous to the derivation of (21) from (16) and therefore we have omitted the details. We note that the transformed system (22) is the adjoint system of (21).

BTMR generates projection matrices $\mathbf{V}, \mathbf{W} \in \mathbb{R}^{n \times k}$ with $k \ll n$ such that

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

The reduced order model for (21) is obtained by replacing $\mathbf{v}_H(t)$ in (21) by $\mathbf{V} \hat{\mathbf{v}}(t)$ and multiplying the resulting equation by \mathbf{W}^T . This gives

$$\frac{d}{dt} \hat{\mathbf{v}}(t) = -\mathbf{W}^T \mathbf{A} \mathbf{V} \hat{\mathbf{v}}(t) + \mathbf{W}^T \tilde{\mathbf{B}} \mathbf{u}(t), \quad (23a)$$

$$\hat{\mathbf{z}}(t) = \tilde{\mathbf{C}} \mathbf{V} \hat{\mathbf{v}}(t) + \tilde{\mathbf{D}} \mathbf{u}(t) - \mathbf{F} (\mathbf{B} \mathbf{M}^{-1} \mathbf{B}^T)^{-1} \mathbf{L} \frac{d}{dt} \mathbf{u}(t), \quad (23b)$$

$$\hat{\mathbf{v}}(t) = \mathbf{W}^T \mathbf{\Pi} \mathbf{v}^{(0)}. \quad (23c)$$

Similarly, the reduced order model for (22) is obtained by replacing $\lambda_H(t)$ in (22) by $\mathbf{W} \hat{\lambda}(t)$ and multiplying the resulting equation by \mathbf{V}^T . This gives

$$-\frac{d}{dt} \hat{\lambda}(t) = -\mathbf{V}^T \mathbf{A}^T \mathbf{W} \hat{\lambda}(t) + \mathbf{V}^T \tilde{\mathbf{C}}^T \mathbf{w}(t), \quad (24a)$$

$$\hat{\mathbf{g}}(t) = \tilde{\mathbf{B}}^T \mathbf{W} \hat{\lambda}(t) + \tilde{\mathbf{D}}^T \mathbf{w}(t) + \mathbf{L}^T (\mathbf{B} \mathbf{M}^{-1} \mathbf{B}^T)^{-1} \mathbf{F}^T \frac{d}{dt} \mathbf{w}(t), \quad (24b)$$

$$\hat{\lambda}(t) = \mathbf{V}^T \mathbf{\Pi} \lambda^{(T)}. \quad (24c)$$

If $\mathbf{v}_H(0) = \lambda_H(T) = \mathbf{0}$, then for any given inputs \mathbf{u}, \mathbf{w} we have

$$\|\mathbf{z} - \hat{\mathbf{z}}\|_{L^2} \leq 2 \|\mathbf{u}\|_{L^2} (\sigma_{k+1} + \dots + \sigma_n), \quad (25a)$$

$$\|\mathbf{q} - \hat{\mathbf{g}}\|_{L^2} \leq 2 \|\mathbf{w}\|_{L^2} (\sigma_{k+1} + \dots + \sigma_n), \quad (25b)$$

where $\sigma_1 \geq \dots \geq \sigma_k > \sigma_{k+1} + \dots + \sigma_n \geq 0$ are the so-called Hankel singular values that are computed when generating the projection matrices \mathbf{V}, \mathbf{W} .

Remark 4.

- (i) To compute the projection matrices \mathbf{V}, \mathbf{W} one must compute approximations to the so-called controllability and observability Gramians, which are solutions of associated matrix Lyapunov equations. For Stokes systems this is described in [23,32]. We use the multishift ADI algorithm described in [23], which is itself an extension of the algorithm in [20] to descriptor systems. The computational cost depends on the size $m + n$ of the Stokes matrix in 13, as well as the number s of observations and the number r of inputs. Computation of approximate controllability and observability Gramians in low rank factored form is the most expensive part. As we will discuss in Section 4, in our application the number of finite element nodes on the interface between subdomains Ω_1 and $\Omega_2(\theta)$ lead to auxiliary controls and outputs and better ways to compute the projection matrices \mathbf{V}, \mathbf{W} for problems with larger number numbers of controls and observations would reduce the cost of the reduced order model computation. In our application, the one-time cost of reduced order model computation is amortized over several optimization iterations and overall leads to significant computational savings.

The BT reduced order model preserves stability. While the Stokes matrix \mathbf{S} in the full order state equations is symmetric, the matrix $\mathbf{V}^T \mathbf{A} \mathbf{W}$ in the reduced order state equations (23a) is no longer symmetric. However, this has negligible effect on the performance of the reduced order model.

- (ii) The error bound (25) depends on the sum of the truncated singular values $\sigma_{k+1} + \dots + \sigma_n$. For large scale problems only approximations of the first $n_1 \ll n$ singular values $\sigma_1, \dots, \sigma_{n_1}$ are computed. We determine the truncation index k as the smallest integer such that $\sigma_{k+1} < \text{tol}$, see Section 4. Therefore, we cannot compute $\sigma_{k+1} + \dots + \sigma_n$, but can only monitor the upper bound $\sigma_{k+1} + \dots + \sigma_{n_1} + (n - n_1)\sigma_{n_1}$ or the more pessimistic upper bound $(n - k)\sigma_{k+1}$. If the Hankel singular values decay rapidly, as they do in our application, these upper bounds seem sufficient.

Decay rates of Hankel singular values have been studied in [10]. The only critical parameter for our system is the viscosity. For decreasing viscosity (increasing Reynolds number), the poles approach the imaginary axis. However, for the microfluidic flows under consideration, the Reynolds number is close to one so that stability does not become an issue.

- (iii) The bound (25) holds when $\mathbf{v}_H(0) = \lambda_H(T) = \mathbf{0}$. Inhomogeneous initial conditions can be handled as discussed in [22].
- (iv) We will use the bounds (25) to derive estimates between the solution of the original, full order optimization problem and the reduced order optimization problem. Any other model reduction methods for which such a bound is available could in principle be used as well. In particular, reduced order model based on interpolation and Krylov subspaces, which are nicely surveyed in [9], are of interest. These aim at minimizing the so-called \mathcal{H}_2 error of the system, which is slightly different than (25). The extension of the analysis in [4] to optimal \mathcal{H}_2 reduced order models, the efficient computation of these reduced order model for Stokes type systems, and numerical comparisons with BTMR are possible future research tasks.

3.3. Domain decomposition

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

$$\bar{\Omega}(\theta) = \bar{\Omega}_1 \cup \bar{\Omega}_2(\theta), \quad \Omega_1 \cap \Omega_2(\theta) = \emptyset, \quad \Gamma := \bar{\Omega}_1 \cap \bar{\Omega}_2(\theta), \quad (26)$$

where Γ stands for the interface between the subdomains. In our application, $\Omega_2(\theta)$ corresponds to the small region between the microchannels and the reservoirs. This is the only subdomain that depends on the design variables θ . The subdomain Ω_1 is independent of θ . We assume that the objective functional can be split accordingly

$$J(\theta) := J_1(\mathbf{v}, p) + J_2(\mathbf{v}(\theta), p(\theta), \theta). \quad (27)$$

Here, $J_1(\mathbf{v}, p)$ is given in terms of observation operators $C : \mathbf{L}^2((0, T); \mathbf{V}) \rightarrow \mathbf{L}^2((0, T); (L^2(\Omega_1))^q)$, $F : \mathbf{L}^2((0, T); L_0^2(\Omega)) \rightarrow \mathbf{L}^2((0, T); (L^2(\Omega_1))^q)$ and a feedthrough operator $D : \mathbf{L}^2((0, T); \mathbf{L}^2(\Omega)) \rightarrow \mathbf{L}^2((0, T); (L^2(\Omega_1))^q)$, $q \in \mathbb{N}$. For a given function $\mathbf{d} \in \mathbf{L}^2((0, T); (L^2(\Omega_1))^q)$, we define

$$J_1(\mathbf{v}, p) := \int_0^T \int_{\Omega_1} |C\mathbf{v} + Fp + D\mathbf{u} - \mathbf{d}|^2 dx dt. \quad (28)$$

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 (12b)–(12d) has a unique solution $(\mathbf{v}, \mathbf{p}) \in \mathbf{L}^2((0, T); \mathbb{R}^n) \times \mathbf{L}^2((0, T); \mathbb{R}^m / (\text{Ker} \mathbf{B}^T))$. 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 (12b)–(12d) 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 is made for both subdomains.

In our semidiscrete problem 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

$$\mathbf{x}(t) := (\mathbf{v}_1, \mathbf{p}_1, \mathbf{v}_2, \mathbf{p}_2, \mathbf{v}_\Gamma, \mathbf{p}_0)^T, \quad t \in [0, T].$$

With this discretization and partitioning of variables, the matrices $\mathbf{A}(\theta)$ and $\mathbf{B}(\theta)$ can be partitioned as follows

$$\mathbf{A}(\theta) = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{0} & \mathbf{A}_{1\Gamma} \\ \mathbf{0} & \mathbf{A}_{22}(\theta) & \mathbf{A}_{2\Gamma}(\theta) \\ \mathbf{A}_{1\Gamma}^T & \mathbf{A}_{2\Gamma}^T(\theta) & \mathbf{A}_{\Gamma\Gamma}(\theta) \end{pmatrix}, \quad \mathbf{B}(\theta) = \begin{pmatrix} \mathbf{B}_{11} & \mathbf{0} & \mathbf{B}_{1\Gamma} \\ \mathbf{0} & \mathbf{B}_{22}(\theta) & \mathbf{B}_{2\Gamma}(\theta) \\ 0 & 0 & \mathbf{B}_0(\theta) \end{pmatrix}.$$

Likewise, the matrices $\mathbf{K}(\theta)$, $\mathbf{L}(\theta)$ and the lumped mass matrix $\mathbf{M}(\theta)$ admit the decompositions

$$\begin{aligned} \mathbf{K}(\theta) &= (\mathbf{K}_1, \mathbf{K}_2(\theta), \mathbf{K}_\Gamma(\theta))^T, & \mathbf{L}(\theta) &= (\mathbf{L}_1, \mathbf{L}_2(\theta), \mathbf{L}_0(\theta))^T, \\ \mathbf{M}(\theta) &= \mathbf{blockdiag}(\mathbf{M}_1, \mathbf{M}_2(\theta), \mathbf{M}_\Gamma(\theta)). \end{aligned}$$

We set

$$\mathbf{E}(\theta) = \begin{pmatrix} \mathbf{E}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_2(\theta) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{E}_\Gamma(\theta) \end{pmatrix} \quad \mathbf{S}(\theta) = \begin{pmatrix} \mathbf{S}_1 & \mathbf{0} & \mathbf{S}_{1\Gamma} \\ \mathbf{0} & \mathbf{S}_2(\theta) & \mathbf{S}_{2\Gamma}(\theta) \\ \mathbf{S}_{1\Gamma}^T & \mathbf{S}_{2,\Gamma}^T(\theta) & \mathbf{S}_\Gamma(\theta) \end{pmatrix},$$

where the submatrices are partitioned accordingly, and

$$\mathbf{N}(\theta) = (\mathbf{K}_1 | \mathbf{L}_1 | \mathbf{K}_2(\theta) | \mathbf{L}_2(\theta) | \mathbf{K}_\Gamma(\theta) | \mathbf{L}_0(\theta))^T.$$

We further denote by $\mathbf{C}_1 \in \mathbb{R}^{q \times n_1}$, $\mathbf{F}_1 \in \mathbb{R}^{q \times m_1}$, $\mathbf{D}_1 \in \mathbb{R}^{q \times r}$, $\mathbf{d}(t) \in \mathbb{R}^q$, $t \in (0, T)$, the matrices and the vector and by $\ell(\mathbf{v}_2, \mathbf{v}_\Gamma, \mathbf{p}_2, \mathbf{p}_0, t, \theta)$ the functional resulting from the semi-discretization of the inner integrals in J_2 . We set

$$J(\theta) := J_1(\mathbf{v}_1, \mathbf{p}_1, \mathbf{p}_0) + J_2(\mathbf{v}_2(\theta), \mathbf{v}_\Gamma(\theta), \mathbf{p}_2(\theta), \mathbf{p}_0(\theta), \theta), \quad (29)$$

where J_1 and J_2 are given by

$$J_1(\mathbf{v}_1, \mathbf{p}_1, \mathbf{p}_0) = \frac{1}{2} \int_0^T |\mathbf{C}_1 \mathbf{v}_1(t) + \mathbf{F}_1 \mathbf{p}_1(t) + \mathbf{F}_0 \mathbf{p}_0(t) + \mathbf{D}_1 \mathbf{u}(t) - \mathbf{d}(t)|^2 dt,$$

$$J_2(\mathbf{v}_2, \mathbf{v}_\Gamma, \mathbf{p}_2, \mathbf{p}_0, \theta) = \int_0^T \ell(\mathbf{v}_2, \mathbf{v}_\Gamma, \mathbf{p}_2, \mathbf{p}_0, t, \theta) dt.$$

The semi-discretized, domain decomposed shape optimization problem can be formulated according to

$$\min_{\theta \in \Theta} J(\theta) \quad (30a)$$

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

$$\mathbf{P}(\theta) \mathbf{x}(t) := \mathbf{E}(\theta) \frac{d}{dt} \mathbf{x}(t) + \mathbf{S}(\theta) \mathbf{x}(t) = \mathbf{N}(\theta) \mathbf{u}(t) \quad t \in (0, T], \quad (30b)$$

$$\mathbf{M}(\theta) \mathbf{v}(0) = \mathbf{v}^{(0)}(\theta). \quad (30c)$$

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 (30a)–(30c) is given by

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\mu}, \theta) := J(\mathbf{v}, \mathbf{p}, \theta) + \int_0^T \boldsymbol{\mu}(t)^T (\mathbf{P}(\theta) \mathbf{x}(t) - \mathbf{N}(\theta) \mathbf{u}(t)) dt, \quad (31)$$

and the optimality conditions read

$$\begin{aligned} \nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \boldsymbol{\mu}, \theta) &= 0, \quad \nabla_{\boldsymbol{\mu}} \mathcal{L}(\mathbf{x}, \boldsymbol{\mu}, \theta) = 0, \\ \nabla_{\theta} \mathcal{L}(\mathbf{x}, \boldsymbol{\mu}, \theta)^T (\tilde{\theta} - \theta) &\geq 0 \quad \forall \tilde{\theta} \in \Theta. \end{aligned} \quad (32)$$

Due to the special structure of $\mathbf{A}(\theta)$, etc., introduced by the domain decomposition, the optimality conditions (32) can be split into a coupled system of optimality conditions associated with the subdomains Ω_1 , $\Omega_2(\theta)$, and the interface $\Gamma(\theta)$. The optimality subsystem associated with the fixed subdomain Ω_1 is exactly of the form (16), (17), were parts of the inputs $\mathbf{K}\mathbf{u}$, $\mathbf{C}^T \mathbf{w}$ and parts of the outputs \mathbf{z} , \mathbf{q} correspond to the interface conditions between Ω_1 and $\Gamma(\theta)$. The idea is to apply BTMR described in Section 3.2 to this optimality subsystem and to keep the full order model for the subproblems associated with $\Omega_2(\theta)$ and $\Gamma(\theta)$. This leads to a reduced order optimality system that has exactly the same structure as (32), but is of much smaller size. In particular, the reduced order optimality system is the optimality system of an optimization problem of the form (30), but where the large submatrices \mathbf{A}_{11} , etc., and variables \mathbf{v}_1 , which correspond to the fixed subdomain are replaced by small matrices $\hat{\mathbf{A}}_{11}$, etc., and variables $\hat{\mathbf{v}}_1$.

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 on the integrand ℓ , it can be shown that

$$\|\theta^* - \hat{\theta}^*\| \leq C(\sigma_{k+1} + \dots + \sigma_n), \quad (33)$$

where $\sigma_1 \geq \dots \geq \sigma_k > \sigma_{k+1} + \dots + \sigma_n \geq 0$ are the Hankel singular values provided by BTMR for the Ω_1 optimality subsystem. We refer to [4] for details.

4. Numerical results

In actual numerical simulations of acoustic streaming, we consider (2a)–(2e) in dimensionless form according to

$$\frac{VT}{L} \frac{d\mathbf{v}_1}{dt} - \frac{VT^2}{L^3} (\tilde{\nu}_1 \Delta \mathbf{v}_1 + \tilde{\nu}_2 \nabla(\nabla \cdot \mathbf{v}_1)) + \nabla p_1 = 0 \quad \text{in } \Omega \times (0, T_1], \quad (34a)$$

$$\frac{L^3}{c_0^2 VT^3} \frac{dp_1}{dt} + \nabla \cdot \mathbf{v}_1 = 0 \quad \text{in } \Omega \times (0, T_1], \quad (34b)$$

$$\frac{VT}{L} \mathbf{v}_1 = \frac{d\mathbf{u}}{dt} \quad \text{on } \Gamma_D \times (0, T_1], \quad \boldsymbol{\sigma}_1 \mathbf{n} = 0 \quad \text{on } \Gamma_N \times (0, T_1], \quad (34c)$$

$$\mathbf{v}_1(\cdot, 0) = 0, \quad p_1(\cdot, 0) = 0 \quad \text{in } \Omega, \quad (34d)$$

where the parameters and their meanings are shown in Table 1. Similarly, the dimensionless form of (7a)–(7e) is as follows:

$$\frac{VT}{L} \frac{d\mathbf{v}_2}{dt} - \frac{VT^2}{L^3} (\tilde{\nu}_1 \Delta \mathbf{v}_2 + \tilde{\nu}_2 \nabla(\nabla \cdot \mathbf{v}_2)) + \nabla p_2 = \langle \mathbf{f}_v \rangle \quad \text{in } \Omega \times \mathbb{R}_+, \quad (35a)$$

Table 1
Numerical and physical parameters for the numerical simulation of acoustic streaming.

Parameter	Value and units	Description
V	1.0×10^{-1} m/s	Dimensionless velocity scale
L	1.0×10^{-7} m	Dimensionless length scale
T	1.0×10^{-8} s	Dimensionless time scale
f	1.0×10^8 Hz	Frequency of the SAW device
c_0	1.484×10^3 m/s	Small signal sound speed in water
ρ	1.0×10^3 kg/m ³	Density of liquid
u_0	1.0×10^{-9} m	Maximal SAW displacement
C_d	8.06×10^3 1/m	Damping parameter of the LSAW
$\tilde{\nu}_1$	1.002×10^{-6} m ² /s	Kinematic viscosity of water
$\tilde{\nu}_2$	1.002×10^{-6} m ² /s	Kinematic bulk viscosity of water

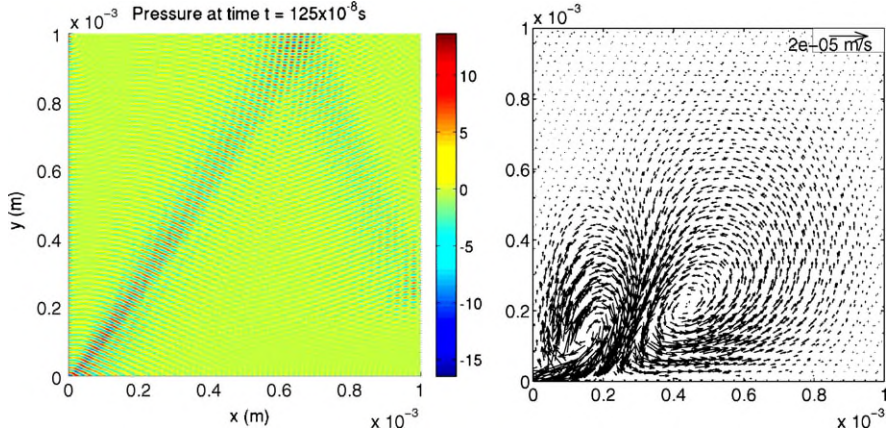


Fig. 2. Left: pressure at $t = 1.25 \mu\text{s}$. Right: velocity field \mathbf{v}_2 in m/s .

$$\nabla \cdot \mathbf{v}_2 = \langle f_p \rangle \quad \text{in } \Omega \times \mathbb{R}_+, \quad (35b)$$

$$\mathbf{v}_2 = \langle (\nabla \mathbf{v}_1) \mathbf{u} \rangle \quad \text{on } \Gamma_D \times \mathbb{R}_+, \quad \boldsymbol{\sigma}_2 \mathbf{n} = \mathbf{0} \quad \text{on } \Gamma_N \times \mathbb{R}_+, \quad (35c)$$

$$\mathbf{v}_2(\cdot, 0) = \mathbf{0}, \quad p_2(\cdot, 0) = 0 \quad \text{in } \Omega, \quad (35d)$$

where

$$\mathbf{f}_v := -\frac{V^2 T^2}{L^2} (\nabla \cdot \mathbf{v}_1) \mathbf{v}_1 + (\nabla \mathbf{v}_1) \mathbf{v}_1, \quad f_p := -\frac{L^2}{T^2 c_0^2} \nabla \cdot (p_1 \mathbf{v}_1).$$

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.

4.1. 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 [25] 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

$$\begin{aligned} u_1(t, x_1) &= 0.6\epsilon \exp(-\hat{C}_d x_1) \sin(2\pi(-\hat{k}x_1 + ft)), \\ u_2(t, x_2) &= -\epsilon \exp(-\hat{C}_d x_1) \cos(2\pi(-\hat{k}x_1 + ft)). \end{aligned}$$

with parameters $\epsilon = u_0/L$, $\hat{C}_d = C_d L$, $\hat{k} = L/\lambda$, $\hat{f} = fT$, where $\lambda = 24 \mu\text{m}$ is the SAW wavelength (cf. Table 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 $t_{end} = \tau_1 N$.

Fig. 2 displays the computed pressure at $t = 1.25 \mu\text{s}$ (left) and the associated velocity field \mathbf{v}_2 (right). Both coincide well with experimental measurements reported in [25].

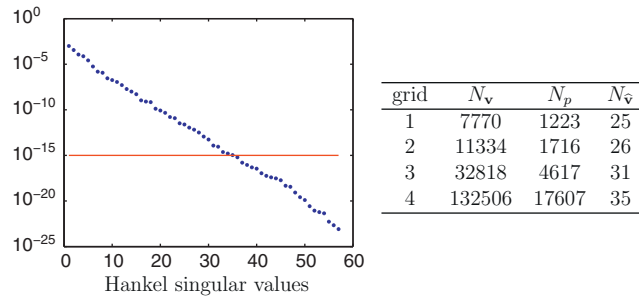


Fig. 3. Left: the largest Hankel singular values and the threshold $10^{-12}\sigma_1$ for the fine grid problem. Right: the degrees of freedom N_v , N_p for the velocity and the pressure in the full order model and the degrees of freedom $N_{\check{v}}$ for velocity in the reduced order model for four different grids.

4.2. BTMR of a microfluidic biochip

A simplified biochip geometry as used for measurements in [14] consists of a rectangular channel occupying the domain $\Omega := (0, 10) \times (0, 1) \cup (0, 10) \cup (9, 10) \cup (0, 1) \times (0, 10) \cup (9, 10) \times (0, 10)$ (in mm^2). The IDT is placed at $x_1 = 0$ aligned with the top wall of the biochip. Since the width of the IDT is 6 mm, we set $\Gamma_{in} = \{0\} \times (9.4, 10)$. The function \mathbf{u} describing the SAW excitation on Γ_{in} is chosen according to

$$\begin{aligned} u_1(t, x_1) &= 0.6\epsilon \sin(2\pi(-\hat{k}x_1 + ft)), \\ u_2(t, x_1) &= -\epsilon \cos(2\pi(-\hat{k}x_1 + ft)), \end{aligned}$$

where the constants are the same as in Subsection 4.1. We further choose $\Gamma_{out} = \{10\} \times (0, 1)\text{mm}^2$ which ensures the uniqueness of the pressure for the acoustic streaming subproblem. On $\Gamma_{lat} = \partial\Omega \setminus (\Gamma_{in} \cup \Gamma_{out})$, the velocity is set to be zero. The geometry and the IDT specifications are exactly the same as in the experimental measurements performed in [14].

As a first task, we solve the acoustic subproblem (34). We iterate from $t = 0$ to $t_{end} = 1.0\mu\text{s}$ with time step $\tau_1 = 0.1$. The computed pressure at t_{end} is in excellent agreement with the measurements in [14]. The solution (\mathbf{v}_1, p_1) is used to generate the right-hand side and the boundary data for the acoustic streaming subsystem (35) which is solved from $t = 0$ to $t_f = 0.1$ ms.

The second task is to apply BTMR to the subsystem (35) 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

$$\mathbf{z}(t) = \int_{\Omega_{\text{obs}}} \nabla \times \mathbf{v}_2 dx, \quad \Omega_{\text{obs}} = (1.5, 2.5) \times (9, 10)\text{mm}^2 \quad (36)$$

in some part Ω_{obs} of the domain. The semidiscretization of this problem described in Section 3.1 leads to a system (16), where the inputs \mathbf{u} in (16a) correspond to the inputs \mathbf{u} in (35c) and the outputs \mathbf{z} in (16c) correspond to (36). This is a simulation problem and BTMR as described in [23] 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.

We apply BTMR to semidiscretizations of (35) on several grids, where 1 refers to the coarsest grid and 4 refers to the finest grid. Fig. 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 [14]. 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 Fig. 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 Fig. 3 shows the numbers N_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 $N_{\check{v}}$ in the reduced order model (23). In particular, we see that BTMR is very effective and dramatically reduced the size of the system.

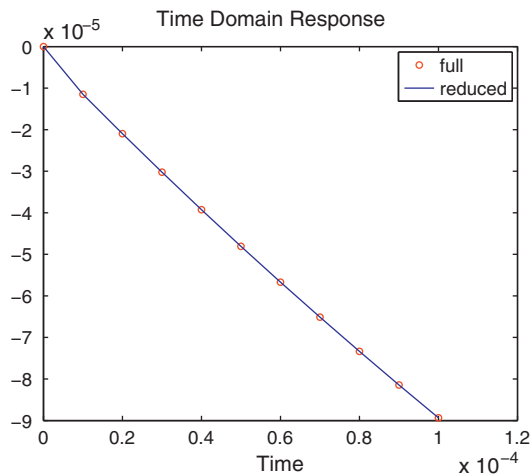


Fig. 4. Time response for the full (circles) and for the reduced order model (solid line).

To illustrate the BTMR error bound (25a) we show the time domain response of the output \mathbf{z} for the full order model and $\hat{\mathbf{z}}$ for the reduced order model in Fig. 4. As predicted by the theory, the reduced order model approximates the full order model very accurately.

4.3. DDBTMR applied to shape optimization of capillary barriers

We consider fluid flow described by subsystem (35) 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 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 Fig. 5. It is decomposed into subdomains $\Omega_1 = \Omega \setminus \Omega_2$, and $\Omega_2 = (1.5, 2.5) \times (9, 10)\text{mm}^2$. The boundary $\partial\Omega$ is split into $\Gamma_{in} = \{0\} \times (9.4, 10)$, $\Gamma_{out} = \{10\} \times (0, 1)\text{mm}^2$, and $\Gamma_{lat} = \partial\Omega \setminus (\Gamma_{in} \cup \Gamma_{out})$. We assume that an IDT of width 6mm is placed at Γ_{in} and that the input velocity profile (u_1, u_2) is the same as in Subsection 4.2. The forces (\mathbf{f}_v, f_p) in $\Omega \times (0, T)$ are computed by solving the acoustic subproblem (34) for $0 = t_0 \leq t \leq t_{end} = 1\mu\text{s}$ with step size $\tau_1 = 0.1$ and is equal to the right hand side in (35a-b). We further choose a constant velocity profile $\mathbf{v}_{in}(x_1, x_2)$ on $\Gamma_{in} \times (0, T)$ as given by (35c), outflow boundary conditions on $\Gamma_{out} \times (0, T)$, and no-slip conditions on $\Gamma_{lat} \times (0, T)$. 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

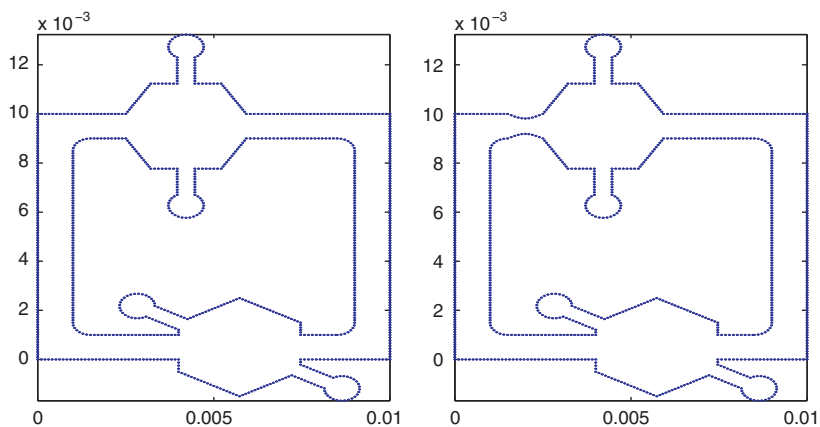


Fig. 5. The reference domain Ω_{ref} (left, in (m)) and the optimal domain (right, in (m)).

vorticity is minimized in Ω_{obs} (two bulb shaped structures associated with the lower reservoir in Fig. 5). The subdomain Ω_2 is parameterized by representing the top and bottom boundary by Bézier curves with d_T and d_B control points, respectively. This leads to a parametrization $\Omega_2(\theta)$ of Ω_2 with parameters $\theta \in \mathbb{R}^{d_T+d_B}$. We use $d_T = d_B = 6$.

The shape optimization problem amounts to the minimization of

$$J(\theta) = \int_0^{t_{\text{end}}} \int_{\Omega_{\text{obs}}} |\nabla \times \mathbf{v}(x, t)|^2 dxdt + \int_0^{t_{\text{end}}} \int_{\Omega_2(\theta)} |\mathbf{v}(x, t) - \mathbf{v}^d(x, t)|^2 dxdt, \quad (37)$$

subject to subsystem (35) and design parameter constraints

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

where $t_{\text{end}} = 0.1$ ms. The bounds θ^{\min} , θ^{\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 Fig. 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 semidiscretization in space is performed as described in Subsections 3.1 and 3.3. Let $N_{\mathbf{v}}^{(1)}$, $N_{\mathbf{v}}^{(2)}$, $N_{\mathbf{v}}^{\Gamma}$ be the number of velocity degrees of freedom in the subdomains $\bar{\Omega}_1 \setminus \Gamma$, $\bar{\Omega}_2 \setminus \Gamma$ 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 $\bar{\Omega}_1$, $\bar{\Omega}_2$ 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 are solved using a projected BFGS method with Armijo line search [24]. The optimization algorithm is terminated when the norm of the projected gradient is less than $\epsilon = 2 \times 10^{-8}$. We use automatic differentiation [18,30] 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 [23]. For BTMR, we truncate the Hankel singular values by selecting the smallest index k for which $\sigma_{k+1} < 10^{-12}\sigma_1$. We apply DDBTMR to semidiscretizations generated using four grids. Fig. 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 Fig. 6 (left). The table in Fig. 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 $N_{\hat{\mathbf{v}}}^{(1)}$, $N_{\hat{\mathbf{v}}}$ 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 $N_{\hat{\mathbf{v}}}^{(1)} = 766$. 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 $N_{\hat{\mathbf{v}}} = 914 + 766 = 1680$ 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 semidiscretization of $\nabla \times \mathbf{v}(x, t)$, $x \in \Omega_{\text{obs}}$. See (37). 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,

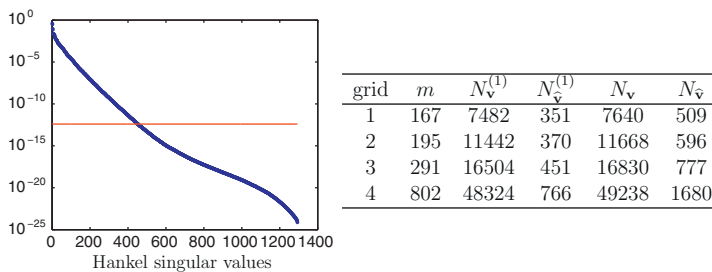


Fig. 6. Left: the largest Hankel singular values computed for the fine grid problem and the threshold $10^{-12}\sigma_1$. Right: the number of m of observations, the numbers $N_{\mathbf{v}}^{(1)}$, $N_{\mathbf{v}}$ of velocity degrees of freedom in subdomain Ω_1 and in Ω , respectively, for the full order model, and the number $N_{\hat{\mathbf{v}}}^{(1)}$, $N_{\hat{\mathbf{v}}}$ of velocity degrees of freedom in subdomain Ω_1 and in Ω , respectively, for the reduced order model.

Table 2

Optimal shape parameters θ^* and $\hat{\theta}^*$ computed by minimizing the full and the reduced order model.

θ^*	(9.8833, 9.7467, 9.7572, 9.8671, 9.1336, 9.2015, 9.1971, 9.1310) $\times 10^{-3}$
$\hat{\theta}^*$	(9.8694, 9.7374, 9.7525, 9.8628, 9.1498, 9.2044, 9.1895, 9.1204) $\times 10^{-3}$

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 $N_{\hat{\varphi}}$.

The constant C in the estimate (33) for the error between the optimal design parameters computed by the full and the reduced order problems, respectively, depends on quantities like α in Theorem 3(ii), derivatives of $\mathbf{A}(\theta)$ with respect to θ , etc., which in turn depend on the application data given in Table 1. Numerical experiments indicate that for the current scaling of the problem, the constant C in the estimate (33) is large. 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 $\hat{\theta}^*$ computed by minimizing the full and the reduced order model, respectively, are shown in Table 2. For the finest grid, the error between the full and the reduced order model solutions is $\|\theta^* - \hat{\theta}^*\| = 3.9165 \cdot 10^{-5}$.

5. Conclusion

We have applied model reduction to the simulation and shape optimization of a microfluidic biochip. After application of a homogenization approach for a proper separation of the time scales, the fluid in the biochip is modeled by the compressible Stokes equations. The shape variation of parts of the biochip introduces a nonlinearity into the problem, which is spatially localized. To obtain a reduced order model of the problem with the property that the error between the solution of the original shape optimization problem and the shape optimization problem governed by the reduced order equations can be controlled and made small, we integrate domain decomposition and balanced truncation model reduction. Domain decomposition is used to separate the nonlinear subsystem from the linear one, and to state the interface conditions between the two. Balanced truncation model reduction is used to reduce the linear subsystem. We use BTMR because it provides an error estimate for the linear subproblem. We extend this error estimate to obtain the desired error estimate for the solution of the shape optimization problem obtained with the reduced order model.

Numerical experiments show the validity of our approach. In particular, the one-time cost of reduced order model computation is amortized over several optimization iterations and model reduction overall leads to significant computational savings, while delivering a solution to the shape optimization problem with guaranteed error bounds.

There are several directions for future research. The computational cost of our model reduction approach depends on the number s of observations and the number r of inputs. In our application the number of finite element nodes on the interface between subdomains Ω_1 and $\Omega_2(\theta)$ lead to auxiliary controls and outputs. Faster approaches to compute the projection matrices for problems with larger number numbers of controls and observations would reduce the cost of the reduced order model computation.

In principle other model reduction methods than balanced truncation can be used. Reduced order models based on interpolation and Krylov subspaces are interesting candidates. These aim at minimizing the so-called \mathcal{H}_2 error of the system, which is slightly different than the error bound given by balanced truncation (see [9,19]). The extension of the analysis in [4] to optimal \mathcal{H}_2 reduced order models, the efficient computation of these reduced order model for Stokes type systems, and numerical comparisons with BTMR are possible future research tasks.

References

- [1] H. Antil, A. Gantner, R.H.W. Hoppe, D. Köster, K.G. Siebert, 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, vol. 60, Springer, Berlin–Heidelberg–New York, 2008, pp. 305–312.
- [2] H. Antil, R. Glowinski, R.H.W. Hoppe, C. Linsenmann, T.W. Pan, A. Wixforth, Modeling, simulation, and optimization of surface acoustic wave driven microfluidic biochips, *J. Comput. Math.* 28 (2010) 149–169.
- [3] H. Antil, M. Heinkenschloss, R.H.W. Hoppe, D.C. Sorensen, Domain decomposition and model reduction for the numerical solution of PDE constrained optimization problems with localized optimization variables, *Comput. Vis. Sci.* (2010).

- [4] H. Antil, M. Heinkenschloss, R.H.W. Hoppe, Domain decomposition and balanced truncation model reduction for shape optimization of the Stokes system. *Optim. Methods Software*, doi:10.1080/10556781003767904, (2010).
- [5] H. Antil, R.H.W. Hoppe, C. Linsenmann, Path-following primal-dual interior-point methods for shape optimization of stationary flow problems, *J. Numer. Math.* 11 (2007) 81–100.
- [6] H. Antil, R.H.W. Hoppe, C. Linsenmann, Adaptive path following primal dual interior point methods for shape optimization of linear and nonlinear Stokes flow problems, in: *Lecture Notes in Computer Science*, vol. 4818, Springer, Berlin–Heidelberg–New York, 2008, pp. 259–266.
- [7] H. Antil, R.H.W. Hoppe, C. Linsenmann, Optimal design of stationary flow problems by path-following interior-point methods, *Control Cybern.* 37 (2008) 771–796.
- [8] A.C. Antoulas, *Approximation of Large-scale Systems*, SIAM, Philadelphia, 2005.
- [9] A.C. Antoulas, C.A. Beattie, S. Gugercin, Interpolatory model reduction of large-scale dynamical systems, in: J. Mohammadpour, K. Grigoriadis (Eds.), *Efficient Modeling and Control of Large-scale Systems*, Springer, Berlin–Heidelberg–New York, 2010, pp. 3–58.
- [10] A.C. Antoulas, D.C. Sorensen, Y. Zhou, On the decay rate of Hankel singular values and related issues, *Syst. Control Lett.* 46 (2002) 323–342.
- [11] P. Benner, V. Mehrmann, D.C. Sorensen (Eds.), *Dimension Reduction of Large-scale Systems*. *Lecture Notes in Computational Science and Engineering*, vol. 45, Springer, Berlin–Heidelberg–New York, 2005.
- [12] D. Braess, *Finite Elements. Theory, Fast Solvers and Applications in Elasticity Theory*, 3rd ed., Cambridge University Press, Cambridge, 2007.
- [13] F. Brezzi, M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer, Berlin–Heidelberg–New York, 1991.
- [14] M.A. Fallah, *SAW Induced Acoustic Streaming in Microchannels of Different Geometry*, Master's thesis, Institute of Physics, University of Augsburg, 2008.
- [15] T. Franke, A. Wixforth, Microfluidics for miniaturized laboratories on a chip, *ChemPhysChem* 9 (2008) 2140–2156.
- [16] A. Gantner, R.H.W. Hoppe, D. Köster, K.G. Siebert, A. Wixforth, Numerical simulation of piezoelectrically agitated surface acoustic waves on microfluidic biochips, *Comput. Vis. Sci.* 10 (2007) 145–161.
- [17] V. Girault, P.-A. Raviart, *Finite Element Methods for Navier–Stokes Equations. Theory and Algorithms*, Springer, Berlin–Heidelberg–New York, 1986.
- [18] A. Griewank, A. Walther, *Evaluating Derivatives. Principles and Techniques of Algorithmic Differentiation*, 2nd ed., SIAM, Philadelphia, 2008.
- [19] S. Gugercin, A.C. Antoulas, C. Beattie, H_2 model reduction for large-scale linear dynamical systems, *SIAM J. Matrix Anal. Appl.* 30 (2008) 609–638.
- [20] S. Gugercin, D.C. Sorensen, A.C. Antoulas, A modified low-rank Smith method for large-scale Lyapunov equations, *Numer. Algorithms* 32 (2003) 27–55.
- [21] Z. Guttenberg, H. Muller, H. Habermuller, A. Geisbauer, J. Pipper, J. Felbel, M. Kieplinski, J. Scriba, A. Wixforth, Planar chip device for PCR and hybridization with surface acoustic wave pump, *Lab Chip* 5 (2005) 308–317.
- [22] M. Heinkenschloss, T. Reis, A.C. Antoulas, *Balanced Truncation Model Reduction for Systems with Inhomogeneous Initial Conditions*, Technical Report TR09–29, Department of Computational and Applied Mathematics, Rice University, 2009.
- [23] M. Heinkenschloss, D.C. Sorensen, K. Sun, Balanced truncation model reduction for a class of descriptor systems with application to the Oseen equations, *SIAM J. Sci. Comput.* 30 (2008) 1038–1063.
- [24] C.T. Kelley, *Iterative Methods for Optimization*, SIAM, Philadelphia, 1999.
- [25] D. Köster, Numerical simulation of acoustic streaming on SAW-driven biochips, *SIAM J. Comp. Sci.* 29 (2007) 2352–2380.
- [26] J.L. Lions, E. Magenes, *Non-homogeneous Boundary Value Problems and Applications*, vol. I, II, Springer, Berlin–Heidelberg–New York, 1972.
- [27] V. Mehrmann, T. Stykel, Balanced truncation model reduction for large-scale systems in descriptor form, in: *Dimension Reduction of Large-scale Systems*, P. Benner, V. Mehrmann, D.C. Sorensen (Eds.), *Lecture Notes in Computational Science, Engineering*, vol. 45, Springer, Berlin–Heidelberg–New York, 2005, pp. 83–115.
- [28] J. Pollard, B. Castrodale, *Outlook for DNA Microarrays: Emerging Applications and Insights on Optimizing Microarray Studies*, Report, Cambridge Health Institute, Cambridge, 2003.
- [29] C.W. Rowley, Model reduction for fluids, using balanced proper orthogonal decomposition, *Int. J. Bifurcat. Chaos* 15 (2005) 997–1013.
- [30] S.M. Rump, *INTLAB - INTerval LABoratory*, in: Tibor Csendes (Ed.), *Developments in Reliable Computing*, Kluwer Academic Publishers, Dordrecht, 1999, pp. 77–104, <http://www.ti3.tu-harburg.de/rump/>.
- [31] H.M. Shapiro, *Practical Flow Cytometry*, Wiley-Liss, New York, 2003.
- [32] T. Stykel, Balanced truncation model reduction for semidiscretized Stokes equation, *Linear Algebra Appl.* 415 (2006) 262–289.
- [33] A. Wixforth, Acoustically driven programmable microfluidics for biological and chemical applications, *JALA* 11 (2006) 399–405.
- [34] A. Wixforth, J. Scriba, G. Gauer, Flatland fluidics, *mst news* 5 (2002) 42–43.
- [35] L.Y. Yeo, J.R. Friend, Ultrafast microfluidics using surface acoustic waves, *Biomicrofluidics* 3 (2009) 012002–012023.