Comput Visual Sci (2010) 13:249–264
DOI 10.1007/s00791-010-0142-4

REGULAR ARTICLE

Domain decomposition and model reduction for the numerical solution of PDE constrained optimization problems with localized optimization variables

Harbir Antil · Matthias Heinkenschloss · Ronald H. W. Hoppe · Danny C. Sorensen

Received: 28 April 2010 / Accepted: 3 August 2010 / Published online: 25 November 2010

Abstract We introduce 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 on an optimal control problem and on a shape optimization problem.

Keywords Optimal control · Shape optimization · Domain decomposition · Model reduction

Communicated by Gabriel Wittum.

The research of RH was supported in part by NSF grants DMS-0511624, DMS-0707602, DMS-0810176, DMS-0811153 and by the German National Science Foundation (DFG) within the Priority Program SPP 1253. The research of MH was supported in part by NSF grant DMS-0915238 and AFOSR grant FA9550-09-1-0225.

H. Antil · M. Heinkenschloss (✉) · D. C. Sorensen
Department of Computational and Applied Mathematics, MS-134, Rice University, 6100 Main Street, Houston, TX 77005-1892, USA
e-mail: heinken@rice.edu

H. Antil
e-mail: Harbir.Antil@rice.edu

D. C. Sorensen
e-mail: sorensen@rice.edu

R. H. W. Hoppe
Department of Mathematics, University of Houston, 4800 Calhoun, Houston, TX 77204-3008, USA
e-mail: rohop@math.uh.edu

R. H. W. Hoppe
Institute of Mathematics, University of Augsburg, 86159 Augsburg, Germany

1 Introduction

We investigate the numerical solution of optimization problems governed by time dependent advection diffusion partial differential equations (PDEs) in which the optimization variables are located in a small spatial region Ω_2 of the entire spatial domain Ω on which the PDE is posed. This scenario arises, for example, in shape optimization when only a small portion of the shape can be modified or in parameter identification problems where the parameters are associated with spatially localized material properties.

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 PDE on the small spatial region interacts 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 [7–9,20] 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, [7,8] discusses the coupling of the Navier-Stokes equations to the linear Oseen equations. In [9] the 3D Navier-Stokes equations are coupled with a 1D model for the flow in blood vessels. Section 3.3 of the review paper [20] discusses the coupling of distributed parameter models with lumped parameter models for the modeling of blood flow. The papers [17,18,23,24] use dimension reduction techniques (see [4] for a recent overview). The papers [17,18] 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 [23,24] 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 Sect. 5 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 paper are of the form

$$\text{minimize} \int_0^T \ell(\mathbf{y}(t),t,\theta)\,dt, \tag{1a}$$

subject to

$$\mathbf{M}(\theta)\frac{d}{dt}\mathbf{y}(t) + \mathbf{A}(\theta)\mathbf{y}(t) = \mathbf{B}(\theta)\mathbf{u}(t), \quad t\in(0,T), \tag{1b}$$

$$\mathbf{M}(\theta)\mathbf{y}(0) = \mathbf{y}_0, \tag{1c}$$

$$\theta \in \Theta. \tag{1d}$$

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 (1) for two applications in Sect. 5. 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 (1) by a reduced order problem

$$\text{minimize} \int_0^T \ell(\widehat{\mathbf{y}}(t),t,\theta)\,dt, \tag{2a}$$

subject to

$$\widehat{\mathbf{M}}(\theta)\frac{d}{dt}\widehat{\mathbf{y}}(t) + \widehat{\mathbf{A}}(\theta)\widehat{\mathbf{y}}(t) = \widehat{\mathbf{B}}(\theta)\mathbf{u}(t), \quad t\in(0,T), \tag{2b}$$

$$\widehat{\mathbf{M}}(\theta)\mathbf{y}(0) = \widehat{\mathbf{y}}_0, \tag{2c}$$

$$\theta \in \Theta, \tag{2d}$$

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 (1) is well approximated by the solution $\widehat{\theta}_*$ of (2).

Our approach uses domain decomposition techniques to divide the optimality system corresponding to (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 (2). We provide a-priori estimates for the error between the solution θ_* of (1) and the solution $\widehat{\theta}_*$ of (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 3 applies balanced truncation to reduce a linear quadratic optimal control problem. Although this optimization problem is simpler than (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 (1) is presented and analyzed in Sect. 4. In Sect. 5 we discuss two problems which lead to (1) and the application of our approach for the reduction of these problems.

Throughout this paper 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.

2 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 [3,4,6,10,19]. 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

$$\mathcal{M}\mathbf{y}'(t) = \mathcal{A}\mathbf{y}(t) + \mathcal{B}\mathbf{u}(t), \quad t \in (0,T) \tag{3a}$$
$$\mathbf{z}(t) = \mathcal{C}\mathbf{y}(t) + \mathcal{D}_s\mathbf{u}(t), \quad t \in (0,T) \tag{3b}$$
$$\mathbf{y}(0) = \mathbf{y}_0, \tag{3c}$$
$$-\mathcal{M}\boldsymbol{\lambda}'(t) = \mathcal{A}^T\boldsymbol{\lambda}(t) + \mathcal{C}^T\mathbf{w}(t), \quad t \in (0,T) \tag{3d}$$
$$\mathbf{q}(t) = \mathcal{B}^T\boldsymbol{\lambda}(t) + \mathcal{D}_a\mathbf{w}(t), \quad t \in (0,T) \tag{3e}$$
$$\boldsymbol{\lambda}(T) = 0, \tag{3f}$$

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 equation (3) by setting $\mathbf{y} = \mathcal{V}\widehat{\mathbf{y}}$ and projecting (imposing a Galerkin condition) so that

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

Applying an analogous projection to (3d) and (3e) with $\boldsymbol{\lambda}$ replaced by $\mathcal{W}\widehat{\boldsymbol{\lambda}}$, we obtain a reduced order system of order n given by

$$\widehat{\mathbf{y}}'(t) = \widehat{\mathcal{A}}\widehat{\mathbf{y}}(t) + \widehat{\mathcal{B}}\mathbf{u}(t), \quad t \in (0,T) \tag{4a}$$
$$\widehat{\mathbf{z}}(t) = \widehat{\mathcal{C}}\widehat{\mathbf{y}}(t) + \mathcal{D}_s\mathbf{u}(t), \quad t \in (0,T) \tag{4b}$$
$$\widehat{\mathbf{y}}(0) = \widehat{\mathbf{y}}_0, \tag{4c}$$
$$-\widehat{\boldsymbol{\lambda}}'(t) = \widehat{\mathcal{A}}^T\widehat{\boldsymbol{\lambda}}(t) + \widehat{\mathcal{C}}^T\mathbf{w}(t), \quad t \in (0,T) \tag{4d}$$
$$\widehat{\mathbf{q}}(t) = \widehat{\mathcal{B}}^T\widehat{\boldsymbol{\lambda}}(t) + \mathcal{D}_a\mathbf{w}(t), \quad t \in (0,T) \tag{4e}$$
$$\widehat{\boldsymbol{\lambda}}(T) = 0, \tag{4f}$$

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 (3), we factor $\mathcal{M} = \mathcal{R}\mathcal{R}^T$, multiply (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 (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

$$\mathcal{A}\mathcal{P}\mathcal{M} + \mathcal{M}\mathcal{P}\mathcal{A}^T + \mathcal{B}\mathcal{B}^T = 0, \tag{5a}$$
$$\mathcal{A}^T\mathcal{Q}\mathcal{M} + \mathcal{M}\mathcal{Q}\mathcal{A} + \mathcal{C}^T\mathcal{C} = 0. \tag{5b}$$

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

$$\mathbf{U}^T\mathcal{M}\mathbf{L} = \mathbf{Z}\mathbf{S}\mathbf{Y}^T \quad \text{the SVD}, \tag{6a}$$
$$\mathcal{V} = \mathbf{U}\mathbf{Z}_n\mathbf{S}_n^{-1/2}, \tag{6b}$$
$$\mathcal{W} = \mathbf{L}\mathbf{Y}_n\mathbf{S}_n^{-1/2}. \tag{6c}$$

Here, $\mathbf{S}_n = diag(\sigma_1, \sigma_2, \ldots, \sigma_n)$ 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,

$$\begin{aligned} 0 &= \mathcal{W}^T(\mathcal{A}\mathcal{P}\mathcal{M} + \mathcal{M}\mathcal{P}\mathcal{A}^T + \mathcal{B}\mathcal{B}^T)\mathcal{W} \\ &= \widehat{\mathcal{A}}\mathbf{S}_n + \mathbf{S}_n\widehat{\mathcal{A}}^T + \widehat{\mathcal{B}}\widehat{\mathcal{B}}^T, \end{aligned} \tag{7a}$$
$$\begin{aligned} 0 &= \mathcal{V}^T(\mathcal{A}^T\mathcal{Q}\mathcal{M} + \mathcal{M}\mathcal{Q}\mathcal{A} + \mathcal{C}^T\mathcal{C})\mathcal{V} \\ &= \widehat{\mathcal{A}}^T\mathbf{S}_n + \mathbf{S}_n\widehat{\mathcal{A}} + \widehat{\mathcal{C}}^T\widehat{\mathcal{C}}. \end{aligned} \tag{7b}$$

The terminology "balanced" refers to the fact that the controllability and observability Gramians $\mathbf{S}_n$ 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., [3,10,26]) that $\widehat{\mathcal{A}}$ must be stable. Furthermore if $\mathbf{y}_0 = \mathbf{0}$, then for any given inputs $\mathbf{u}$, $\mathbf{w}$ we have

$$\|\mathbf{z} - \widehat{\mathbf{z}}\|_{L^2} \le 2\|\mathbf{u}\|_{L^2}(\sigma_{n+1} + \cdots + \sigma_N), \tag{8a}$$
$$\|\mathbf{q} - \widehat{\mathbf{q}}\|_{L^2} \le 2\|\mathbf{w}\|_{L^2}(\sigma_{n+1} + \cdots + \sigma_N). \tag{8b}$$

Remark 1 One can derive error bounds for inhomogeneous initial values $\mathbf{y}_0$. These require a slight modification of the problem set-up in which the original $\mathcal{B}$ is augmented, see [13]. 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.

3 Balanced truncation model reduction and optimal control

Before we consider the optimization problem (1), we consider a simpler problem, a linear quadratic optimal control problem

$$\min J(\mathbf{u}) \equiv \frac{1}{2}\int_0^T \|\mathbf{C}\mathbf{y}(t) + \mathbf{D}\mathbf{u}(t) - \mathbf{d}(t)\|^2 dt, \tag{9}$$

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

$$\mathbf{M}\mathbf{y}'(t) = \mathbf{A}\mathbf{y}(t) + \mathbf{B}\mathbf{u}(t), \quad t \in (0, T), \tag{10a}$$
$$\mathbf{y}(0) = \mathbf{y}_0. \tag{10b}$$

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

$$\mathbf{v}^T\mathbf{A}\mathbf{v} \le -\alpha\mathbf{v}^T\mathbf{M}\mathbf{v}, \qquad \forall \mathbf{v} \in \mathbb{R}^N. \tag{11}$$

Note that (11) implies that all eigenvalues of the pair $(\mathbf{A}, \mathbf{M})$ have negative real part.

We want to reduce this optimization problem using balanced truncation model reduction and establish bounds for the error between the solution $\mathbf{u}_*$ of (9), (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 (1) in a simpler setting involving less notation.

The necessary and sufficient optimality conditions for (9), (10) are given by

$$\mathbf{M}\mathbf{y}'(t) = \mathbf{A}\mathbf{y}(t) + \mathbf{B}\mathbf{u}(t), \quad t \in (0, T) \tag{12a}$$
$$\mathbf{z}(t) = \mathbf{C}\mathbf{y}(t) + \mathbf{D}\mathbf{u}(t) - \mathbf{d}(t), \quad t \in (0, T) \tag{12b}$$
$$\mathbf{y}(0) = \mathbf{y}_0, \tag{12c}$$
$$-\mathbf{M}\boldsymbol{\lambda}'(t) = \mathbf{A}^T\boldsymbol{\lambda}(t) + \mathbf{C}^T\mathbf{z}(t), \quad t \in (0, T) \tag{12d}$$
$$\mathbf{q}(t) = \mathbf{B}^T\boldsymbol{\lambda}(t) + \mathbf{D}^T\mathbf{z}(t), \quad t \in (0, T) \tag{12e}$$
$$\boldsymbol{\lambda}(T) = 0, \tag{12f}$$
$$\mathbf{q}(t) = \mathbf{0}, \quad t \in (0, T). \tag{12g}$$

The optimality system (12) is written in a slightly unconventional way to highlight its connection with the system (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

$$\widehat{\mathbf{y}}'(t) = \widehat{\mathbf{A}}\widehat{\mathbf{y}}(t) + \widehat{\mathbf{B}}\mathbf{u}(t), \quad t \in (0, T) \tag{13a}$$
$$\widehat{\mathbf{z}}(t) = \widehat{\mathbf{C}}\widehat{\mathbf{y}}(t) + \mathbf{D}\mathbf{u}(t) - \mathbf{d}(t), \quad t \in (0, T) \tag{13b}$$
$$\widehat{\mathbf{y}}(0) = \widehat{\mathbf{y}}_0, \tag{13c}$$
$$-\widehat{\boldsymbol{\lambda}}'(t) = \widehat{\mathbf{A}}^T\widehat{\boldsymbol{\lambda}}(t) + \widehat{\mathbf{C}}^T\widehat{\mathbf{z}}(t), \quad t \in (0, T) \tag{13d}$$
$$\widehat{\mathbf{q}}(t) = \widehat{\mathbf{B}}^T\widehat{\boldsymbol{\lambda}}(t) + \mathbf{D}^T\widehat{\mathbf{z}}(t), \quad t \in (0, T) \tag{13e}$$
$$\widehat{\boldsymbol{\lambda}}(T) = 0, \tag{13f}$$
$$\widehat{\mathbf{q}}(t) = \mathbf{0}, \quad t \in (0, T), \tag{13g}$$

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

$$\mathbf{y}_0 = \mathbf{0} \tag{14}$$

cf., Remark 1. This can always be achieved by representing the solution of (10) as $\mathbf{y} = \mathbf{y}_u + \mathbf{y}_h$, where $\mathbf{y}_h$ solves (10) with $\mathbf{u} \equiv \mathbf{0}$ and $\mathbf{y}_u$ solves (10) with $\mathbf{y}_0 = \mathbf{0}$ and then writing the optimal control problem (9), (10) as a problem in $\mathbf{y}_u$.

We note that the reduced optimality system (13) is the optimality system for the reduced optimal control problem

$$\min \widehat{J}(\mathbf{u}) \equiv \frac{1}{2}\int_0^T \|\widehat{\mathbf{C}}\widehat{\mathbf{y}}(t) + \mathbf{D}\mathbf{u}(t) - \mathbf{d}(t)\|^2 dt \tag{15}$$

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

$$\widehat{\mathbf{y}}'(t) = \widehat{\mathbf{A}}\widehat{\mathbf{y}}(t) + \widehat{\mathbf{B}}\mathbf{u}(t), \quad t \in (0, T), \tag{16a}$$
$$\widehat{\mathbf{y}}(0) = \widehat{\mathbf{y}}_0. \tag{16b}$$

Next we provide an estimate for the error between the solution $\mathbf{u}_*$ of (9), (10) and the solution $\widehat{\mathbf{u}}_*$ of (15), (16). We assume that J is a strictly convex quadratic function. More precisely, we assume the existence of $\kappa > 0$ such that

$$\langle \mathbf{u} - \mathbf{w}, \nabla J(\mathbf{u}) - \nabla J(\mathbf{w})\rangle_{L^2} \ge \kappa \|\mathbf{u} - \mathbf{w}\|_{L^2}^2 \tag{17}$$

for all $\mathbf{u}, \mathbf{w} \in L^2$. If $\mathbf{u}_*$ solves (9), (10) and $\widehat{\mathbf{u}}_*$ solves (15), (16), then

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

and (17) implies

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

Hence

$$\|\mathbf{u}_* - \widehat{\mathbf{u}}_*\|_{L^2} \le \kappa^{-1}\|\nabla \widehat{J}(\widehat{\mathbf{u}}_*) - \nabla J(\widehat{\mathbf{u}}_*)\|_{L^2}. \tag{18}$$

Thus, to estimate the error we need to estimate the error in the gradients between the original problem (9), (10) and of the reduced (15), (16).

To emphasize the dependence of the solution of (12a–12f) and of (13a–13f) 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 (12a–12f) and $\widehat{\mathbf{y}}(\mathbf{u})$, $\widehat{\mathbf{z}}(\mathbf{u})$, $\widehat{\boldsymbol{\lambda}}(\mathbf{u})$, $\widehat{\mathbf{q}}(\mathbf{u})$ satisfy (13a–13f), 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 (8) for balanced truncation model reduction directly, since (3d, 3e) and (4d, 4e) both depend on the same input $\mathbf{w}$, whereas (12d, 12e) has input $\mathbf{z}$ and (13d, 13e) has input $\widehat{\mathbf{z}}$.

We consider the auxiliary adjoint equation

$$-\mathbf{M}\widetilde{\boldsymbol{\lambda}}'(t) = \mathbf{A}^T\widetilde{\boldsymbol{\lambda}}(t) + \mathbf{C}^T\widehat{\mathbf{z}}(t), \quad t \in (0,T) \tag{19a}$$

$$\widetilde{\mathbf{q}}(t) = \mathbf{B}^T\widetilde{\boldsymbol{\lambda}}(t) + \mathbf{D}^T\widehat{\mathbf{z}}(t), \quad t \in (0,T) \tag{19b}$$

$$\widetilde{\boldsymbol{\lambda}}(T) = 0. \tag{19c}$$

Lemma 1 *Let (11) be satisfied. For any $\mathbf{z}, \widehat{\mathbf{z}} \in L^2$ the outputs $\mathbf{q}$ and $\widetilde{\mathbf{q}}$ of (13d)–(13f) and (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

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

Lemma 4 in the "Appendix" gives

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

The desired inequality follows since

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

□

The error estimate (8) for balanced truncation model reduction implies

$$\|\mathbf{z} - \widehat{\mathbf{z}}\|_{L^2} \le 2\|\mathbf{u}\|_{L^2}(\sigma_{n+1} + \cdots + \sigma_N), \tag{20a}$$

$$\|\widehat{\mathbf{q}} - \widetilde{\mathbf{q}}\|_{L^2} \le 2\|\widehat{\mathbf{z}}\|_{L^2}(\sigma_{n+1} + \cdots + \sigma_N) \tag{20b}$$

for all $\mathbf{u} \in L^2$ and all $\widehat{\mathbf{z}} \in L^2$. We can now use Lemma 1 and the balanced truncation model reduction error estimates (20) to derive a bound for the error between the solutions $\mathbf{u}_*$ of (9), (10) and $\widehat{\mathbf{u}}_*$ of (15), (16).

Theorem 1 *Let (11) 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*

$$\begin{aligned}&\|\nabla J(\mathbf{u}) - \nabla \widehat{J}(\mathbf{u})\|_{L^2}\\ &\quad \le 2\left(c\|\mathbf{u}\|_{L^2} + \|\widehat{\mathbf{z}}(\mathbf{u})\|_{L^2}\right)(\sigma_{n+1} + \cdots + \sigma_N),\end{aligned}$$

where c is the constant specified in Lemma 1.

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 (12a–12f), let $\widehat{\mathbf{y}}(\mathbf{u})$, $\widehat{\mathbf{z}}(\mathbf{u})$, $\widehat{\boldsymbol{\lambda}}(\mathbf{u})$, $\widehat{\mathbf{q}}(\mathbf{u})$ satisfy (13a–13f), and let $\widetilde{\boldsymbol{\lambda}}(\mathbf{u})$, $\widetilde{\mathbf{q}}(\mathbf{u})$ satisfy (19).

We have $\nabla J(\mathbf{u}) = \mathbf{q}(\mathbf{u})$, $\nabla \widehat{J}(\mathbf{u}) = \widehat{\mathbf{q}}(\mathbf{u})$. Lemma 1 and the balanced truncation model reduction error estimates (20) imply

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

□

Inequality (18) and Theorem 1 imply the following estimate for the error in the optimal controls.

Corollary 1 *Let (11) be satisfied and let $\kappa > 0$ be a constant such that (17) holds. Furthermore, let $\mathbf{u}_*$ solve (9), (10) and let $\widehat{\mathbf{u}}_*$ be the solution of (15), (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*

$$\begin{aligned}&\|\mathbf{u}_* - \widehat{\mathbf{u}}_*\|_{L^2}\\ &\quad \le \frac{2}{\kappa}\left(c\|\widehat{\mathbf{u}}_*\|_{L^2} + \|\widehat{\mathbf{z}}_*\|_{L^2}\right)(\sigma_{n+1} + \cdots + \sigma_N),\end{aligned}$$

where c is the constant specified in Lemma 1.

Note that the size of $\sigma_{n+1} + \cdots + \sigma_N$ 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 The optimization problem

We now return to the optimization problem (1). The Lagrangian associated with this problem is

$$\begin{aligned}&L(\mathbf{y}, \boldsymbol{\lambda}, \theta)\\ &\quad = \int_0^T \ell(\mathbf{y}(t), t, \theta)dt\\ &\qquad + \int_0^T \boldsymbol{\lambda}(t)^T\left(\mathbf{M}(\theta)\mathbf{y}'(t) + \mathbf{A}(\theta)\mathbf{y}(t) - \mathbf{B}(\theta)\mathbf{u}(t)\right)dt.\end{aligned}$$

Since Θ is a closed convex set, the first order necessary optimality conditions for (1) are given by $\theta \in \Theta$,

$$\mathbf{M}(\theta)\frac{d}{dt}\mathbf{y}(t) + \mathbf{A}(\theta)\mathbf{y}(t) = \mathbf{B}(\theta)\mathbf{u}(t), \tag{21a}$$

$$-\mathbf{M}(\theta)^T\frac{d}{dt}\boldsymbol{\lambda}'(t) + \mathbf{A}(\theta)^T\boldsymbol{\lambda}(t) = -\nabla_{\mathbf{y}}\ell(\mathbf{y}(t), t, \theta), \tag{21b}$$

for $t \in (0, T)$,

$$\int_0^T D_\theta \ell(\mathbf{y}(t), t, \theta)(\widetilde{\theta} - \theta)dt + \int_0^T \boldsymbol{\lambda}(t)^T \Big[\big(D_\theta \mathbf{M}(\theta)(\widetilde{\theta} - \theta)\big) \frac{d}{dt}\mathbf{y}(t) + \big(D_\theta \mathbf{A}(\theta)(\widetilde{\theta} - \theta)\big)\mathbf{y}(t) - \big(D_\theta \mathbf{B}(\theta)(\widetilde{\theta} - \theta)\big)\mathbf{u}(t) \Big] dt \geq 0 \tag{21c}$$

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

4.1 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 (1a) is of the form

$$\ell(\mathbf{y}(t), t, \theta) = \tfrac{1}{2}\|\mathbf{C}_I^{(1)}\mathbf{y}_I^{(1)}(t) - \mathbf{d}_I^{(1)}(t)\|^2 + \widetilde{\ell}(\mathbf{y}_\Gamma(t), \mathbf{y}_I^{(2)}(t), t, \theta). \tag{22}$$

In the following section we will use domain decomposition to decompose the optimality conditions (21) 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. 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., [22, Chap. 4] and [25, Chap. 1]. Our notation follows that of [22,25]. 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 equation (1b) can be written as

$$\mathbf{M}_{II}^{(1)}\frac{d}{dt}\mathbf{y}_I^{(1)} + \mathbf{M}_{I\Gamma}^{(1)}\frac{d}{dt}\mathbf{y}_\Gamma + \mathbf{A}_{II}^{(1)}\mathbf{y}_I^{(1)} + \mathbf{A}_{I\Gamma}^{(1)}\mathbf{y}_\Gamma = \mathbf{B}_I^{(1)}\mathbf{u}_I^{(1)}, \tag{23a}$$

$$\mathbf{M}_{II}^{(2)}(\theta)\frac{d}{dt}\mathbf{y}_I^{(2)} + \mathbf{M}_{I\Gamma}^{(2)}(\theta)\frac{d}{dt}\mathbf{y}_\Gamma + \mathbf{A}_{II}^{(2)}(\theta)\mathbf{y}_I^{(2)} + \mathbf{A}_{I\Gamma}^{(2)}(\theta)\mathbf{y}_\Gamma = \mathbf{B}_I^{(2)}(\theta)\mathbf{u}_I^{(2)}, \tag{23b}$$

$$\mathbf{M}_{\Gamma I}^{(1)}\frac{d}{dt}\mathbf{y}_I^{(1)} + \mathbf{M}_{\Gamma\Gamma}(\theta)\frac{d}{dt}\mathbf{y}_\Gamma + \mathbf{M}_{\Gamma I}^{(2)}\frac{d}{dt}\mathbf{y}_I^{(2)} + \mathbf{A}_{\Gamma I}^{(1)}\mathbf{y}_I^{(1)} + \mathbf{A}_{\Gamma\Gamma}(\theta)\mathbf{y}_\Gamma + \mathbf{A}_{\Gamma I}^{(2)}\mathbf{y}_I^{(2)} = \mathbf{B}^{\Gamma}(\theta)\mathbf{u}_\Gamma. \tag{23c}$$

The optimality conditions (21) can now be written as (23a–23c) and the adjoint equations

$$-\mathbf{M}_{II}^{(1)}\frac{d}{dt}\boldsymbol{\lambda}_I^{(1)} - \mathbf{M}_{I\Gamma}^{(1)}\frac{d}{dt}\boldsymbol{\lambda}_\Gamma + \left(\mathbf{A}_{II}^{(1)}\right)^T\boldsymbol{\lambda}_I^{(1)} + \left(\mathbf{A}_{\Gamma I}^{(1)}\right)^T\boldsymbol{\lambda}_\Gamma = -\left(\mathbf{C}_I^{(1)}\right)^T\left(\mathbf{C}_I^{(1)}\mathbf{y}_I^{(1)} - \mathbf{d}_I^{(1)}\right), \tag{23d}$$

$$-\mathbf{M}_{II}^{(2)}(\theta)\frac{d}{dt}\boldsymbol{\lambda}_I^{(2)} - \mathbf{M}_{I\Gamma}^{(2)}(\theta)\frac{d}{dt}\boldsymbol{\lambda}_\Gamma + \left(\mathbf{A}_{II}^{(2)}(\theta)\right)^T\boldsymbol{\lambda}_I^{(2)} + \left(\mathbf{A}_{\Gamma I}^{(2)}(\theta)\right)^T\boldsymbol{\lambda}_\Gamma = -\nabla_{\mathbf{y}_I^{(2)}}\widetilde{\ell}(\mathbf{y}_\Gamma, \mathbf{y}_I^{(2)}, t, \theta), \tag{23e}$$

$$-\mathbf{M}_{\Gamma I}^{(1)}\frac{d}{dt}\boldsymbol{\lambda}_I^{(1)} - \mathbf{M}_{\Gamma\Gamma}(\theta)\frac{d}{dt}\boldsymbol{\lambda}_\Gamma - \mathbf{M}_{\Gamma I}^{(2)}\frac{d}{dt}\boldsymbol{\lambda}_I^{(2)} + \left(\mathbf{A}_{I\Gamma}^{(1)}\right)^T\boldsymbol{\lambda}_I^{(1)} + \left(\mathbf{A}_{\Gamma\Gamma}(\theta)\right)^T\boldsymbol{\lambda}_\Gamma + \left(\mathbf{A}_{I\Gamma}^{(2)}\right)^T\boldsymbol{\lambda}_I^{(2)} = -\nabla_{\mathbf{y}_\Gamma}\widetilde{\ell}(\mathbf{y}_\Gamma, \mathbf{y}_I^{(2)}, t, \theta), \tag{23f}$$

and

$$\int_0^T D_\theta \widetilde{\ell}(\mathbf{y}_\Gamma, \mathbf{y}_I^{(2)}, t, \theta)(\widetilde{\theta} - \theta)dt + \int_0^T \begin{pmatrix} \boldsymbol{\lambda}_\Gamma \\ \boldsymbol{\lambda}_I^{(2)} \end{pmatrix}^T \Big[\Big(D_\theta \mathbf{M}^{(2)}(\theta)(\widetilde{\theta} - \theta)\Big) \frac{d}{dt}\begin{pmatrix} \mathbf{y}_\Gamma \\ \mathbf{y}_I^{(2)} \end{pmatrix} + \Big(D_\theta \mathbf{A}^{(2)}(\theta)(\widetilde{\theta} - \theta)\Big)\begin{pmatrix} \mathbf{y}_\Gamma \\ \mathbf{y}_I^{(2)} \end{pmatrix} - \Big(D_\theta \mathbf{B}^{(2)}(\theta)(\widetilde{\theta} - \theta)\Big)\begin{pmatrix} \mathbf{u}_\Gamma \\ \mathbf{u}_I^{(2)} \end{pmatrix} \Big] dt \geq 0 \tag{23g}$$

for all $\widetilde{\theta} \in \Theta$, where we have set

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

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

4.2 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 (23) interact with the other components of the system and we have to make sure that the resulting subsystem is of the form (3) to which balanced truncation can be applied. This is the reason why we have assumed that the integrand ℓ in the objective function (1a) is of the form (22).

If we inspect (23) to see how $\mathbf{y}^{(1)}_I$ and $\boldsymbol{\lambda}^{(1)}_I$ interact with the other components of the system, we are led to

$$\begin{aligned}
\mathbf{M}^{(1)}_{II}\frac{d}{dt}\mathbf{y}^{(1)}_I &= -\mathbf{A}^{(1)}_{II}\mathbf{y}^{(1)}_I - \mathbf{M}^{(1)}_{I\Gamma}\frac{d}{dt}\mathbf{y}_\Gamma \\
&\quad +\mathbf{B}^{(1)}_I\mathbf{u}^{(1)}_I - \mathbf{A}^{(1)}_{I\Gamma}\mathbf{y}_\Gamma && \text{(24a)}\\
\mathbf{z}^{(1)}_I &= -\mathbf{C}^{(1)}_I\mathbf{y}^{(1)}_I + \mathbf{d}^{(1)}_I, && \text{(24b)}\\
\mathbf{z}_\Gamma &= -\mathbf{M}^{(1)}_{\Gamma I}\frac{d}{dt}\mathbf{y}^{(1)}_I - \mathbf{A}^{(1)}_{\Gamma I}\mathbf{y}^{(1)}_I, && \text{(24c)}\\
-\mathbf{M}^{(1)}_{II}\frac{d}{dt}\boldsymbol{\lambda}^{(1)}_I &= -\left(\mathbf{A}^{(1)}_{II}\right)^T\boldsymbol{\lambda}^{(1)}_I + \mathbf{M}^{(1)}_{I\Gamma}\frac{d}{dt}\boldsymbol{\lambda}_\Gamma \\
&\quad -\left(\mathbf{C}^{(1)}_I\right)^T\mathbf{w}^{(1)}_I - \left(\mathbf{A}^{(1)}_{\Gamma I}\right)^T\boldsymbol{\lambda}_\Gamma && \text{(24d)}\\
\mathbf{q}^{(1)}_I &= \left(\mathbf{B}^{(1)}_I\right)^T\boldsymbol{\lambda}^{(1)}_I, && \text{(24e)}\\
\mathbf{q}_\Gamma &= \mathbf{M}^{(1)}_{\Gamma I}\frac{d}{dt}\boldsymbol{\lambda}^{(1)}_I - \left(\mathbf{A}^{(1)}_{I\Gamma}\right)^T\boldsymbol{\lambda}^{(1)}_I. && \text{(24f)}
\end{aligned}$$

In fact (24a) and (24d) are identical to (23a) and (23d), respectively, if $\mathbf{w}^{(1)}_I = -\mathbf{z}^{(1)}_I = \mathbf{C}^{(1)}_I\mathbf{y}^{(1)}_I - \mathbf{d}^{(1)}_I$. The output (24b) enters into (23d) and the output (24c) enters into (23c). Similarly, the output (24f) enters into (23f). The output (24e) is included as an auxiliary variable. It does not enter into any of the equations in (23), but is included to establish the connection with the generic system (3).

If

$$\mathbf{M}^{(1)}_{I\Gamma} = \mathbf{0} \quad \text{and} \quad \mathbf{M}^{(1)}_{\Gamma I} = \mathbf{0}, \tag{25}$$

then (24) is given by

$$\begin{aligned}
\mathbf{M}^{(1)}_{II}\frac{d}{dt}\mathbf{y}^{(1)}_I &= -\mathbf{A}^{(1)}_{II}\mathbf{y}^{(1)}_I + \left(\mathbf{B}^{(1)}_I \mid -\mathbf{A}^{(1)}_{I\Gamma}\right)\begin{pmatrix}\mathbf{u}^{(1)}_I \\ \mathbf{y}_\Gamma\end{pmatrix} && \text{(26a)}\\
\begin{pmatrix}\mathbf{z}^{(1)}_I \\ \mathbf{z}_\Gamma\end{pmatrix} &= \begin{pmatrix}-\mathbf{C}^{(1)}_I \\ -\mathbf{A}^{(1)}_{\Gamma I}\end{pmatrix}\mathbf{y}^{(1)}_I + \begin{pmatrix}\mathbf{I} \\ \mathbf{0}\end{pmatrix}\mathbf{d}^{(1)}_I, && \text{(26b)}\\
-\mathbf{M}^{(1)}_{II}\frac{d}{dt}\boldsymbol{\lambda}^{(1)}_I &= -(\mathbf{A}^{(1)}_{II})^T\boldsymbol{\lambda}^{(1)}_I + \begin{pmatrix}-\mathbf{C}^{(1)}_I \\ -\mathbf{A}^{(1)}_{\Gamma I}\end{pmatrix}^T\begin{pmatrix}\mathbf{w}^{(1)}_I \\ \boldsymbol{\lambda}_\Gamma\end{pmatrix} && \text{(26c)}\\
\begin{pmatrix}\mathbf{q}^{(1)}_I \\ \mathbf{q}_\Gamma\end{pmatrix} &= \left(\mathbf{B}^{(1)}_I \mid -\mathbf{A}^{(1)}_{I\Gamma}\right)^T\boldsymbol{\lambda}^{(1)}_I. && \text{(26d)}
\end{aligned}$$

This system is exactly of the form (3) that is needed for balanced truncation. We assume that

$$\mathbf{v}^T\mathbf{A}\mathbf{v} \le -\alpha\mathbf{v}^T\mathbf{M}\mathbf{v}, \qquad \forall \mathbf{v} \in \mathbb{R}^N. \tag{27}$$

Note that assumption (27) implies

$$\mathbf{v}^T\mathbf{A}^{(1)}_{II}\mathbf{v} \le -\alpha\mathbf{v}^T\mathbf{M}^{(1)}_{II}\mathbf{v}, \qquad \forall \mathbf{v} \in \mathbb{R}^{N^{(1)}_I}. \tag{28}$$

As a consequence of (28) all eigenvalues of the pair $(\mathbf{A}^{(1)}_{II}, \mathbf{M}^{(1)}_{II})$ have negative real part and, hence, balanced truncation model reduction can be applied to (26) which leads to the following reduced subsystem

$$\begin{aligned}
\frac{d}{dt}\widehat{\mathbf{y}}^{(1)}_I &= -\widehat{\mathbf{A}}^{(1)}_{II}\widehat{\mathbf{y}}^{(1)}_I - \widehat{\mathbf{A}}^{(1)}_{I\Gamma}\mathbf{y}_\Gamma + \widehat{\mathbf{B}}^{(1)}_I\mathbf{u}^{(1)}_I && \text{(29a)}\\
\widehat{\mathbf{z}}^{(1)}_I &= -\widehat{\mathbf{C}}^{(1)}_I\widehat{\mathbf{y}}^{(1)}_I + \mathbf{d}^{(1)}_I, && \text{(29b)}\\
\widehat{\mathbf{z}}_\Gamma &= -\widehat{\mathbf{A}}^{(1)}_{\Gamma I}\widehat{\mathbf{y}}^{(1)}_I, && \text{(29c)}\\
-\frac{d}{dt}\widehat{\boldsymbol{\lambda}}^{(1)}_I &= -\left(\widehat{\mathbf{A}}^{(1)}_{II}\right)^T\widehat{\boldsymbol{\lambda}}^{(1)}_I - \left(\widehat{\mathbf{A}}^{(1)}_{\Gamma I}\right)^T\boldsymbol{\lambda}_\Gamma - \left(\widehat{\mathbf{C}}^{(1)}_I\right)^T\mathbf{w}^{(1)}_I && \text{(29d)}\\
\widehat{\mathbf{q}}^{(1)}_I &= \left(\widehat{\mathbf{B}}^{(1)}_I\right)^T\widehat{\boldsymbol{\lambda}}^{(1)}_I, && \text{(29e)}\\
\widehat{\mathbf{q}}_\Gamma &= -\left(\widehat{\mathbf{A}}^{(1)}_{I\Gamma}\right)^T\widehat{\boldsymbol{\lambda}}^{(1)}_I. && \text{(29f)}
\end{aligned}$$

We assume that

$$\mathbf{y}^{(1)}_{I,0} = \mathbf{0}, \tag{30}$$

cf., Remark 1.

Balanced truncation generates a reduced order model (29) such that the error between the input-to-output maps of (24) and (29) can be estimated by

$$\left\|\begin{pmatrix}\mathbf{z}^{(1)}_I \\ \mathbf{z}_\Gamma\end{pmatrix} - \begin{pmatrix}\widehat{\mathbf{z}}^{(1)}_I \\ \widehat{\mathbf{z}}_\Gamma\end{pmatrix}\right\|_{L^2} \le 2\left\|\begin{pmatrix}\mathbf{u}^{(1)}_I \\ \mathbf{y}_\Gamma\end{pmatrix}\right\|_{L^2}\tau, \tag{31a}$$
$$\left\|\begin{pmatrix}\mathbf{q}^{(1)}_I \\ \mathbf{q}_\Gamma\end{pmatrix} - \begin{pmatrix}\widehat{\mathbf{q}}^{(1)}_I \\ \widehat{\mathbf{q}}_\Gamma\end{pmatrix}\right\|_{L^2} \le 2\left\|\begin{pmatrix}\mathbf{w}^{(1)}_I \\ \boldsymbol{\lambda}_\Gamma\end{pmatrix}\right\|_{L^2}\tau, \tag{31b}$$

where

$$\tau = \sigma_{n+1} + \cdots + \sigma_N. \tag{31c}$$

To be consistent with (25) 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 (23) is given by the state equation

$$\frac{d}{dt}\widehat{\mathbf{y}}_I^{(1)} + \widehat{\mathbf{A}}_{II}^{(1)}\widehat{\mathbf{y}}_I^{(1)} + \widehat{\mathbf{A}}_{I\Gamma}^{(1)}\widehat{\mathbf{y}}_\Gamma = \widehat{\mathbf{B}}_I^{(1)}\mathbf{u}_I^{(1)}, \tag{32a}$$

$$\mathbf{M}_{II}^{(2)}\frac{d}{dt}\widehat{\mathbf{y}}_I^{(2)} + \mathbf{A}_{II}^{(2)}\widehat{\mathbf{y}}_I^{(2)} + \mathbf{A}_{I\Gamma}^{(2)}\widehat{\mathbf{y}}_\Gamma = \mathbf{B}_I^{(2)}\mathbf{u}_I^{(2)}, \tag{32b}$$

$$\mathbf{M}_{\Gamma\Gamma}\frac{d}{dt}\widehat{\mathbf{y}}_\Gamma + \widehat{\mathbf{A}}_{\Gamma I}^{(1)}\widehat{\mathbf{y}}_I^{(1)} + \mathbf{A}_{\Gamma\Gamma}\widehat{\mathbf{y}}_\Gamma + \mathbf{A}_{\Gamma I}^{(2)}\widehat{\mathbf{y}}_I^{(2)} = \mathbf{B}_\Gamma\mathbf{u}_\Gamma, \tag{32c}$$

the adjoint equation

$$-\frac{d}{dt}\widehat{\boldsymbol{\lambda}}_I^{(1)} + \left(\widehat{\mathbf{A}}_{II}^{(1)}\right)^T\widehat{\boldsymbol{\lambda}}_I^{(1)} + \left(\widehat{\mathbf{A}}_{\Gamma I}^{(1)}\right)^T\widehat{\boldsymbol{\lambda}}_\Gamma = -\left(\widehat{\mathbf{C}}_I^{(1)}\right)^T\left(\widehat{\mathbf{C}}_I^{(1)}\widehat{\mathbf{y}}_I^{(1)} - \mathbf{d}_I^{(1)}\right), \tag{32d}$$

$$-\mathbf{M}_{II}^{(2)}\frac{d}{dt}\widehat{\boldsymbol{\lambda}}_I^{(2)} + \left(\mathbf{A}_{II}^{(2)}\right)^T\widehat{\boldsymbol{\lambda}}_I^{(2)} + \left(\mathbf{A}_{\Gamma I}^{(2)}\right)^T\widehat{\boldsymbol{\lambda}}_\Gamma = -\nabla_{\widehat{\mathbf{y}}_I^{(2)}}\widetilde{\ell}(\widehat{\mathbf{y}}_\Gamma, \widehat{\mathbf{y}}_I^{(2)}, t, \theta), \tag{32e}$$

$$-\mathbf{M}_{\Gamma\Gamma}\frac{d}{dt}\widehat{\boldsymbol{\lambda}}_\Gamma + \left(\widehat{\mathbf{A}}_{I\Gamma}^{(1)}\right)^T\widehat{\boldsymbol{\lambda}}_I^{(1)} + \mathbf{A}_{\Gamma\Gamma}^T\widehat{\boldsymbol{\lambda}}_\Gamma + \left(\mathbf{A}_{I\Gamma}^{(2)}\right)^T\widehat{\boldsymbol{\lambda}}_I^{(2)} = -\nabla_{\widehat{\mathbf{y}}_\Gamma}\widetilde{\ell}(\widehat{\mathbf{y}}_\Gamma, \widehat{\mathbf{y}}_I^{(2)}, t, \theta), \tag{32f}$$

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

$$\begin{aligned}
&\int_0^T D_\theta\widetilde{\ell}(\widehat{\mathbf{y}}_\Gamma, \widehat{\mathbf{y}}_I^{(2)}, t, \theta)(\widetilde{\theta} - \theta)dt \\
&+ \int_0^T \begin{pmatrix} \widehat{\boldsymbol{\lambda}}_\Gamma \\ \widehat{\boldsymbol{\lambda}}_I^{(2)} \end{pmatrix}^T \left[\left(D_\theta\mathbf{M}^{(2)}(\theta)(\widetilde{\theta} - \theta)\right) \frac{d}{dt}\begin{pmatrix} \widehat{\mathbf{y}}_\Gamma \\ \widehat{\mathbf{y}}_I^{(2)} \end{pmatrix} \right. \\
&+ \left(D_\theta\mathbf{A}^{(2)}(\theta)(\widetilde{\theta} - \theta)\right)\begin{pmatrix} \widehat{\mathbf{y}}_\Gamma \\ \widehat{\mathbf{y}}_I^{(2)} \end{pmatrix} \\
&\left. - \left(D_\theta\mathbf{B}^{(2)}(\theta)(\widetilde{\theta} - \theta)\right)\begin{pmatrix} \mathbf{u}_\Gamma \\ \mathbf{u}_I^{(2)} \end{pmatrix} \right] dt \geq 0
\end{aligned} \tag{32g}$$

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

The reduced order optimality system (32) is the first order necessary optimality system for the reduced order semidiscretized shape optimization problem

$$\begin{aligned}
\text{minimize} \quad &\int_0^T \tfrac{1}{2}\|\widehat{\mathbf{C}}_I^{(1)}\widehat{\mathbf{y}}_I^{(1)}(t) - \mathbf{d}_I^{(1)}(t)\|_2^2 \\
&+\widetilde{\ell}(\widehat{\mathbf{y}}_\Gamma(t), \widehat{\mathbf{y}}_I^{(2)}(t), t, \theta)\, dt,
\end{aligned} \tag{33}$$

subject to (32a–32c) 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 Error analysis

We define the objective functions

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

where $\mathbf{y}_I^{(1)}, \mathbf{y}_I^{(2)}, \mathbf{y}_\Gamma$ solve (23a–23c) and where $\widehat{\mathbf{y}}_I^{(1)}, \widehat{\mathbf{y}}_I^{(2)}, \widehat{\mathbf{y}}_\Gamma$ solve (32a–32c). 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 (1) and (33) 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

$$\nabla J(\theta_*)^T(\theta - \theta_*) \geq 0 \qquad \nabla\widehat{J}(\widehat{\theta}_*)^T(\theta - \widehat{\theta}_*) \geq 0 \tag{34}$$

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

$$(\nabla J(\theta_*) - \nabla\widehat{J}(\widehat{\theta}_*))^T(\widehat{\theta}_* - \theta_*) \geq 0 \tag{35}$$

If we assume the convexity condition

$$(\nabla J(\widehat{\theta}_*) - \nabla J(\theta_*))^T(\widehat{\theta}_* - \theta_*) \geq \kappa\|\widehat{\theta}_* - \theta_*\|^2, \tag{36}$$

then combining (35) and (36) 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

$$\|\theta_* - \widehat{\theta}_*\| \leq \kappa^{-1}\|\nabla\widehat{J}(\widehat{\theta}_*) - \nabla J(\widehat{\theta}_*)\|. \tag{37}$$

As before, assuming (36), an estimate of the error in the solution of (1) and (33) requires an estimate of the error in the gradient of the full and the reduced order optimization problem.

The gradients are given by

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

where $\mathbf{y}_I^{(1)}$, $\mathbf{y}_I^{(2)}$, $\mathbf{y}_\Gamma$, $\boldsymbol{\lambda}_I^{(1)}$, $\boldsymbol{\lambda}_I^{(2)}$, $\boldsymbol{\lambda}_\Gamma$ solve (23a–23f), and

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

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 (32a–32f), respectively. The difference is given by

$$\begin{aligned}
&\left(\nabla J(\theta)-\nabla\widehat{J}(\theta)\right)^T\widetilde{\theta}\\
&\quad=\int_0^T\left(D_\theta\widetilde{\ell}(\mathbf{y}_\Gamma,\mathbf{y}_I^{(2)},t,\theta)-D_\theta\widetilde{\ell}(\widehat{\mathbf{y}}_\Gamma,\widehat{\mathbf{y}}_I^{(2)},t,\theta)\right)\widetilde{\theta}dt\\
&\qquad+\int_0^T\begin{pmatrix}\boldsymbol{\lambda}_\Gamma\\ \boldsymbol{\lambda}_I^{(2)}\end{pmatrix}^T\left\{\left(D_\theta\mathbf{M}^{(2)}(\theta)\widetilde{\theta}\right)\frac{d}{dt}\begin{pmatrix}\mathbf{y}_\Gamma-\widehat{\mathbf{y}}_\Gamma\\ \mathbf{y}_I^{(2)}-\widehat{\mathbf{y}}_I^{(2)}\end{pmatrix}\right.\\
&\qquad\left.+\left(D_\theta\mathbf{A}^{(2)}(\theta)\widetilde{\theta}\right)\begin{pmatrix}\mathbf{y}_\Gamma-\widehat{\mathbf{y}}_\Gamma\\ \mathbf{y}_I^{(2)}-\widehat{\mathbf{y}}_I^{(2)}\end{pmatrix}\right\}dt\\
&\qquad+\int_0^T\begin{pmatrix}\boldsymbol{\lambda}_\Gamma-\widehat{\boldsymbol{\lambda}}_\Gamma\\ \boldsymbol{\lambda}_I^{(2)}-\widehat{\boldsymbol{\lambda}}_I^{(2)}\end{pmatrix}^T\left\{\left(D_\theta\mathbf{M}^{(2)}(\theta)\widetilde{\theta}\right)\frac{d}{dt}\begin{pmatrix}\widehat{\mathbf{y}}_\Gamma\\ \widehat{\mathbf{y}}_I^{(2)}\end{pmatrix}\right.\\
&\qquad+\left(D_\theta\mathbf{A}^{(2)}(\theta)\widetilde{\theta}\right)\begin{pmatrix}\widehat{\mathbf{y}}_\Gamma\\ \widehat{\mathbf{y}}_I^{(2)}\end{pmatrix}\\
&\qquad\left.-\left(D_\theta\mathbf{B}^{(2)}(\theta)\widetilde{\theta}\right)\begin{pmatrix}\mathbf{u}_\Gamma\\ \mathbf{u}_I^{(2)}\end{pmatrix}\right\}dt. \qquad (38)
\end{aligned}$$

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

Lemma 2 *Let (27) be valid. If $\mathbf{y}_I^{(1)}$, $\mathbf{y}_I^{(2)}$, $\mathbf{y}_\Gamma$ solve (23a–23c), and $\widehat{\mathbf{y}}_I^{(1)}$, $\widehat{\mathbf{y}}_I^{(2)}$, $\widehat{\mathbf{y}}_\Gamma$ solve (32a–32c), then*

$$\begin{aligned}
&\left\|\mathbf{C}_I^{(1)}\mathbf{y}_I^{(1)}-\widehat{\mathbf{C}}_I^{(1)}\widehat{\mathbf{y}}_I^{(1)}\right\|_{L^2}\\
&\quad\le\left(2+\frac{4\|\mathbf{M}^{-1}\|\|\mathbf{C}_I^{(1)}\|}{\alpha}\right)\left\|\begin{pmatrix}\mathbf{u}_I^{(1)}\\ \widehat{\mathbf{y}}_\Gamma\end{pmatrix}\right\|_{L^2}\tau \qquad (39a)
\end{aligned}$$

and

$$\left\|\begin{pmatrix}\mathbf{y}_I^{(2)}-\widehat{\mathbf{y}}_I^{(2)}\\ \mathbf{y}_\Gamma-\widehat{\mathbf{y}}_\Gamma\end{pmatrix}\right\|_{L^2}\le\frac{4\|\mathbf{M}^{-1}\|}{\alpha}\left\|\begin{pmatrix}\mathbf{u}_I^{(1)}\\ \widehat{\mathbf{y}}_\Gamma\end{pmatrix}\right\|_{L^2}\tau, \qquad (39b)$$

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

Proof Let $\mathbf{y}_I^{(1)}$, $\mathbf{y}_I^{(2)}$, $\mathbf{y}_\Gamma$ solve (23a–23c), and let $\widehat{\mathbf{y}}_I^{(1)}$, $\widehat{\mathbf{y}}_I^{(2)}$, $\widehat{\mathbf{y}}_\Gamma$ solve (32a–32c). Furthermore, let $\widetilde{\mathbf{y}}_I^{(1)}$ solve

$$\mathbf{M}_{II}^{(1)}\frac{d}{dt}\widetilde{\mathbf{y}}_I^{(1)}(t)+\mathbf{A}_{II}^{(1)}\widetilde{\mathbf{y}}_I^{(1)}(t)+\mathbf{A}_{I\Gamma}^{(1)}\widehat{\mathbf{y}}_\Gamma(t)=\mathbf{B}_I^{(1)}\mathbf{u}_I^{(1)}(t) \qquad (40)$$

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

The balanced truncation error bound (31) implies

$$\left\|\begin{pmatrix}\mathbf{C}_I^{(1)}\widetilde{\mathbf{y}}_I^{(1)}-\widehat{\mathbf{C}}_I^{(1)}\widehat{\mathbf{y}}_I^{(1)}\\ \mathbf{A}_{\Gamma I}^{(1)}\widetilde{\mathbf{y}}_\Gamma-\widehat{\mathbf{A}}_{\Gamma I}^{(1)}\widehat{\mathbf{y}}_\Gamma\end{pmatrix}\right\|_{L^2}\le 2\left\|\begin{pmatrix}\mathbf{u}_I^{(1)}\\ \widehat{\mathbf{y}}_\Gamma\end{pmatrix}\right\|_{L^2}\tau. \qquad (41)$$

The equations (23a–23c), (32a–32c), and (40) give

$$\begin{aligned}
&\mathbf{M}_{II}^{(1)}(\theta)\frac{d}{dt}\left(\mathbf{y}_I^{(1)}-\widetilde{\mathbf{y}}_I^{(1)}\right)\\
&\quad+\mathbf{A}_{II}^{(1)}(\theta)(\mathbf{y}_I^{(1)}-\widetilde{\mathbf{y}}_I^{(1)})+\mathbf{A}_{I\Gamma}^{(1)}(\theta)(\mathbf{y}_\Gamma-\widehat{\mathbf{y}}_\Gamma)=0, \qquad (42a)\\
&\mathbf{M}_{II}^{(2)}(\theta)\frac{d}{dt}(\mathbf{y}_I^{(2)}-\widehat{\mathbf{y}}_I^{(2)})\\
&\quad+\mathbf{A}_{II}^{(2)}(\theta)(\mathbf{y}_I^{(2)}-\widehat{\mathbf{y}}_I^{(2)})+\mathbf{A}_{I\Gamma}^{(2)}(\theta)(\mathbf{y}_\Gamma-\widehat{\mathbf{y}}_\Gamma)=0, \qquad (42b)\\
&\mathbf{M}_{\Gamma\Gamma}(\theta)\frac{d}{dt}(\mathbf{y}_\Gamma-\widehat{\mathbf{y}}_\Gamma)+\mathbf{A}_{\Gamma\Gamma}(\theta)(\mathbf{y}_\Gamma-\widehat{\mathbf{y}}_\Gamma)\\
&\quad+\mathbf{A}_{\Gamma I}^{(1)}(\mathbf{y}_I^{(1)}-\widetilde{\mathbf{y}}_I^{(1)})+\mathbf{A}_{\Gamma I}^{(2)}(\theta)(\mathbf{y}_I^{(2)}-\widehat{\mathbf{y}}_I^{(2)})\\
&\quad=\widehat{\mathbf{A}}_{\Gamma I}^{(1)}\widehat{\mathbf{y}}_I^{(1)}-\mathbf{A}_{\Gamma I}^{(1)}\widetilde{\mathbf{y}}_I^{(1)} \qquad (42c)
\end{aligned}$$

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 5 in the "Appendix" to (42) followed by an application of (41) gives

$$\begin{aligned}
\left\|\begin{pmatrix}\mathbf{y}_I^{(1)}-\widetilde{\mathbf{y}}_I^{(1)}\\ \mathbf{y}_I^{(2)}-\widehat{\mathbf{y}}_I^{(2)}\\ \mathbf{y}_\Gamma-\widehat{\mathbf{y}}_\Gamma\end{pmatrix}\right\|_{L^2}&\le\frac{2\|\mathbf{M}^{-1}\|}{\alpha}\left\|\widehat{\mathbf{A}}_{\Gamma I}^{(1)}\widehat{\mathbf{y}}_I^{(1)}-\mathbf{A}_{\Gamma I}^{(1)}\widetilde{\mathbf{y}}_I^{(1)}\right\|_{L^2}\\
&\le\frac{4\|\mathbf{M}^{-1}\|}{\alpha}\left\|\begin{pmatrix}\mathbf{u}_I^{(1)}\\ \widehat{\mathbf{y}}_\Gamma\end{pmatrix}\right\|_{L^2}\tau. \qquad (43)
\end{aligned}$$

This implies (39b). The estimate (39a) follows from (41) and (43). □

The errors in the adjoints are estimated similarly.

Lemma 3 *Let (27) be valid and assume that*

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

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* (23a–23f), *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* (32a–32f), *then*

$$\left\|\begin{pmatrix}\boldsymbol{\lambda}_I^{(2)}-\widehat{\boldsymbol{\lambda}}_I^{(2)}\\ \boldsymbol{\lambda}_\Gamma-\widehat{\boldsymbol{\lambda}}_\Gamma\end{pmatrix}\right\|_{L^2}\le c_\lambda(\sigma_{n+1}+\cdots+\sigma_N),\tag{44}$$

where

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

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 (23a–23f), 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 (32a–32f) 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

$$\begin{aligned}
&-\mathbf{M}_{II}^{(1)}\frac{d}{dt}\widetilde{\boldsymbol{\lambda}}_I^{(1)}(t)+(\mathbf{A}_{II}^{(1)})^T\widetilde{\boldsymbol{\lambda}}_I^{(1)}(t)+(\mathbf{A}_{\Gamma I}^{(1)})^T\widehat{\boldsymbol{\lambda}}_\Gamma(t)\\
&\quad=-(\mathbf{C}_I^{(1)})(\widehat{\mathbf{C}}_I^{(1)}\widehat{\mathbf{y}}_I^{(1)}-\mathbf{d}_I^{(1)})
\end{aligned}\tag{45}$$

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

The balanced truncation error bound (31) implies

$$\begin{aligned}
&\left\|\begin{pmatrix}(\mathbf{B}_I^{(1)})^T\widetilde{\boldsymbol{\lambda}}_I^{(1)} & -(\widehat{\mathbf{B}}_I^{(1)})^T\widehat{\boldsymbol{\lambda}}_I^{(1)}\\ (\mathbf{A}_{I\Gamma}^{(1)})^T\widetilde{\boldsymbol{\lambda}}_I^{(1)} & -(\widehat{\mathbf{A}}_{I\Gamma}^{(1)})^T\widehat{\boldsymbol{\lambda}}_I^{(1)}\end{pmatrix}\right\|_{L^2}\\
&\quad\le 2\left\|\begin{pmatrix}\widehat{\mathbf{C}}_I^{(1)}\widehat{\mathbf{y}}_I^{(1)}-\mathbf{d}_I^{(1)}\\ \widehat{\boldsymbol{\lambda}}_\Gamma\end{pmatrix}\right\|_{L^2}(\sigma_{n+1}+\cdots+\sigma_N).
\end{aligned}\tag{46}$$

The equations (23d–23f), (32d–32f), and (45) imply

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

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 5 gives the estimate

$$\begin{aligned}
&\left\|\begin{pmatrix}\boldsymbol{\lambda}_I^{(1)}-\widetilde{\boldsymbol{\lambda}}_I^{(1)}\\ \boldsymbol{\lambda}_I^{(2)}-\widehat{\boldsymbol{\lambda}}_I^{(2)}\\ \boldsymbol{\lambda}_\Gamma-\widehat{\boldsymbol{\lambda}}_\Gamma\end{pmatrix}\right\|_{L^2}\\
&\quad\le\frac{2\|\mathbf{M}^{-1}\|}{\alpha}\|\mathbf{C}_I^{(1)}\|\|\mathbf{C}_I^{(1)}\mathbf{y}_I^{(1)}-\widehat{\mathbf{C}}_I^{(1)}\widehat{\mathbf{y}}_I^{(1)}\|_{L^2}\\
&\qquad+\frac{2\|\mathbf{M}^{-1}\|}{\alpha}\|(\widehat{\mathbf{A}}_{I\Gamma}^{(1)})^T\widehat{\boldsymbol{\lambda}}_I^{(1)}-(\mathbf{A}_{I\Gamma}^{(1)})^T\widetilde{\boldsymbol{\lambda}}_I^{(1)}\|_{L^2}\\
&\qquad+\frac{2L\|\mathbf{M}^{-1}\|}{\alpha}\left\|\begin{pmatrix}\mathbf{y}_I^{(2)}-\widehat{\mathbf{y}}_I^{(2)}\\ \mathbf{y}_\Gamma-\widehat{\mathbf{y}}_\Gamma\end{pmatrix}\right\|_{L^2}.
\end{aligned}\tag{48}$$

The error estimate follows from (39), (46) and (48). □

Equation (38) and Lemmas 2, 3 imply the following result

Theorem 2 *Let the assumptions of Lemma 3 be valid and assume that*

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

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*

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

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

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

Proof The inequality follows directly from equation (38) and Lemmas 2, 3. □

Corollary 2 *If the assumptions of Theorem 2 and (36) hold, then there exists $c > 0$ dependent on $\mathbf{u}$, $\widehat{\mathbf{y}}$, and $\widehat{\boldsymbol{\lambda}}$ such that*

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

Remark 2 (i) The error estimates in Theorem 2 and Corollary 2 rely on an estimate of the type (31) 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 (27) 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, is necessary for the application of balanced truncation model reduction. Secondly, we use it in connection with Lemma 4. 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 4. If such estimates are used, assumption (27) could be weakened.

5 Numerical examples

5.1 Optimal control of water pollution

This example is motivated by [5], where adaptive finite elements are considered for a steady state version of the optimal control problem described below. See also [1] for a related problem.

The domain Ω is shown in Fig. 1. The boundary specifications in Fig. 1 are those for the advection diffusion equation (50).

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

$$-\mu\Delta\mathbf{V}(x) + \nabla p(x) = \mathbf{0}, \quad \text{in } \Omega \tag{49a}$$
$$\nabla \cdot \mathbf{V}(x) = 0, \quad \text{in } \Omega \tag{49b}$$
$$\mathbf{V}(x) = \mathbf{V}_{in}(x), \quad \text{on } \Gamma_{\text{in}} \tag{49c}$$
$$\mathbf{V}(x) = \mathbf{0}, \quad \text{on } \Gamma_0 \tag{49d}$$
$$-\mu\nabla\mathbf{V}(x)\mathbf{n} + p(x)\mathbf{n} = 0, \quad \text{on } \Gamma_{\text{out}}. \tag{49e}$$

The problem data are chosen as in [5]. In particular, $\mu = 0.1$ and $\mathbf{V}_{in}(x) = (1 - (x_2/0.2)^2, 0)^T$. Furthermore, the inflow boundary is $\Gamma_{\text{in}} = \{(x_1, x_2) \in \overline{\Omega} : x_1 = 0\}$, the outflow boundary is $\Gamma_{\text{out}} = \{(x_1, x_2) \in \overline{\Omega} : x_1 = 1.2\}$, and $\Gamma_0 = \partial\Omega\backslash(\Gamma_{\text{in}} \cup \Gamma_{\text{out}})$.

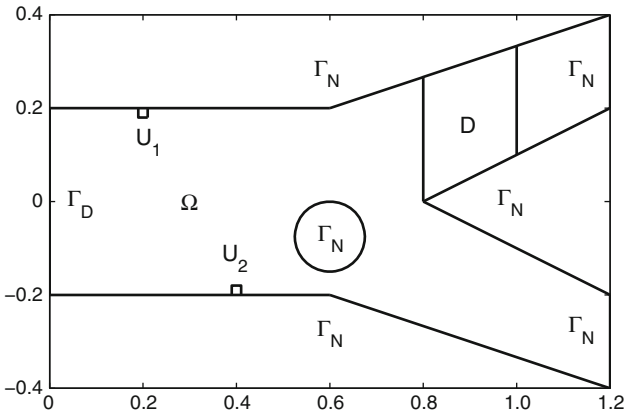


Fig. 1 The domain Ω with boundary conditions for the advection diffusion equation (50)

The optimal control problem is governed by the advection diffusion equation

$$\frac{\partial}{\partial t}y(x,t) - \nabla(k\nabla y(x,t)) + \mathbf{V}(x)\cdot\nabla y(x,t) \tag{50a}$$
$$= u(x,t)\chi_{U_1}(x) + u(x,t)\chi_{U_2}(x) \quad \text{in } \Omega, \tag{50b}$$

with boundary and initial conditions

$$y(x,t) = 0 \quad \text{in } \Gamma_D, \tag{50c}$$
$$\frac{\partial}{\partial n}y(x,t) = 0 \quad \text{in } \Gamma_N, \tag{50d}$$
$$y(x,0) = 0 \quad \text{in } \Omega. \tag{50e}$$

Here χ_S is the characteristic function corresponding to the set S. Furthermore, $k = 0.015$, V is the solution of (49), the boundary segments Γ_D and Γ_N and the control regions U_1 and U_2 are shown in Fig. 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 Fig. 1 and $d \equiv 0.5$.

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 [12] with $m = 4$ shifts to solve the controllability and observability Lyapunov equations (5). For the model reduction, we select those Hankel singular values σ_n, with $\sigma_n \geq 10^{-4}\sigma_1$. Table 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.

Table 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

Grid number	m	k	N	n
1	168	9	1, 545	9
2	283	16	2,673	9
3	618	29	6,036	9

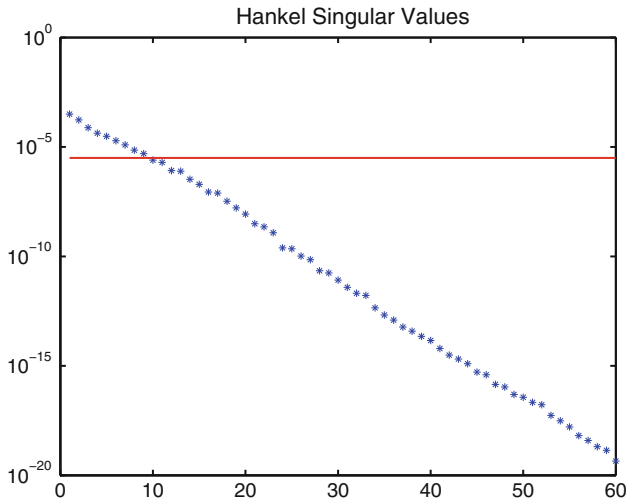


Fig. 2 The largest Hankel singular values and the threshold $10^{-4}\sigma_1$

Figure 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.

For the numerical solution of the optimal control problem (9), (10) and of its reduced version (15), (16) we use the Crank–Nicolson method in time with time step size 10^{-2}. The resulting problem is solved using the Conjugate Gradient method with initial guess $\mathbf{u} = \mathbf{0}$. The Conjugate Gradient is stopped if the initial residual is reduced by a factor 10^{-4}. Figure 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 1. For the fine grid problem, the error between full and the reduced order model solutions is $\|u_* - \widehat{u}_*\|^2_{L^2} = 6.2 \cdot 10^{-3}$.

The convergence histories of the Conjugate Gradient algorithm applied to the full order and reduced order optimal control problem are shown in Fig. 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 1.

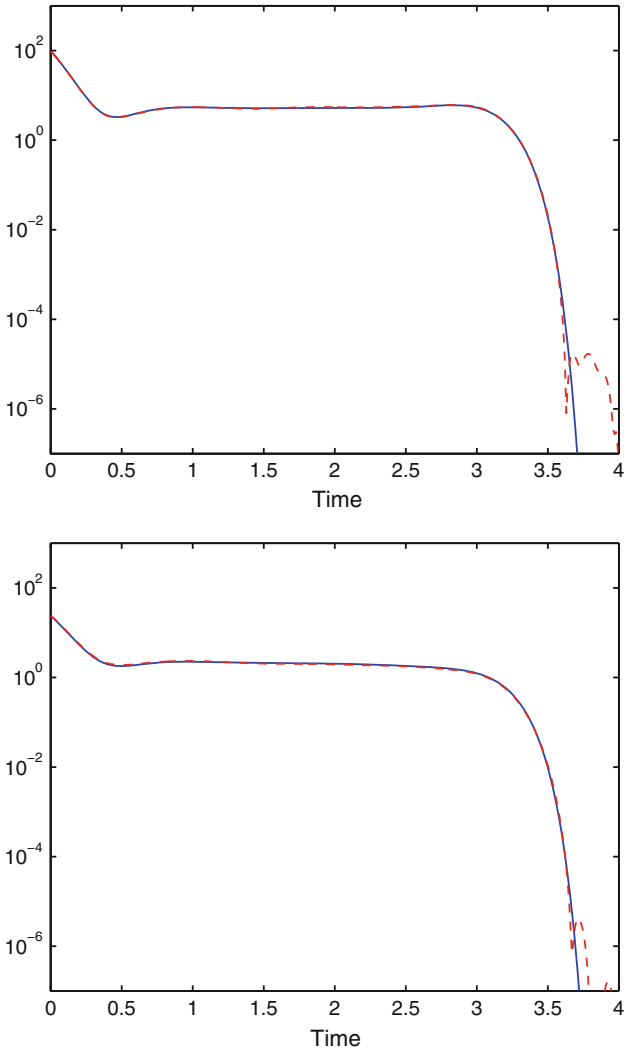


Fig. 3 The *top* plot shows the integrals $\int_{U_1} u_*^2(x,t)dx$ and $\int_{U_1} \widehat{u}_*^2(x,t)dx$ of the optimal controls computed using the full (*solid blue line*) and and the reduced order model (*dashed red line*). The *bottom* plot shows the integrals $\int_{U_2} u_*^2(x,t)dx$ and $\int_{U_2} \widehat{u}_*^2(x,t)dx$ of the optimal controls computed using the full (*solid blue line*) and and the reduced order model (*dashed red line*). The full and reduced order model solutions are in excellent agreement

5.2 Shape optimization

Our second example is a shape optimization problem governed by the heat equation. The domain Ω is of the type shown in Fig. 5 with a circular hole Ω_H. It is decomposed into subdomains $\Omega_1 = \Omega_A \cup \Omega_B$ and $\Omega_2 = \Omega_C \backslash \Omega_H$. The boundary $\partial\Omega$ is decomposed into Γ_L, Γ_R, Γ_T, Γ_B, and $\Gamma_H = \partial\Omega_H$. The interface between Ω_1 and Ω_2 is given by $\Gamma_I = (\overline{\Omega}_A \cap \overline{\Omega}_C) \cup (\overline{\Omega}_B \cap \overline{\Omega}_C)$.

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

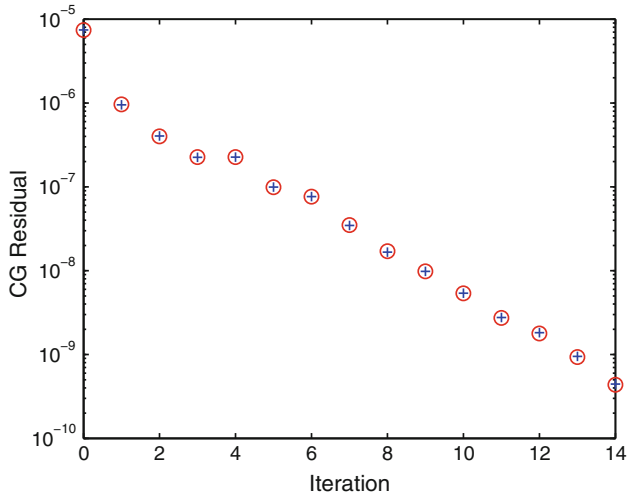


Fig. 4 The convergence histories of the Conjugate Gradient algorithm applied to the full (*blue* +) and the reduced (*red* ◯) order optimal control problems

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

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

and design parameter constraints

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

We set $f = 100$ in $\Omega_2(\theta) \times (0,T)$ and $f = 0$ else. Furthermore $T = 4$. 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 3) below) and solving the state equation on $\Omega(\theta_*)$. The optimal domain $\Omega(\theta_*)$ is shown in Fig. 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 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 $N_{dof}^{(\nu)} = \text{card}(\mathcal{N}_h(\bar{\Omega}_\nu \backslash \Gamma_I))$, $\nu = 1, 2$, 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 (1) are given as usual. If ϕ_i are the piecewise linear basis functions associated with the triangulation of $\Omega(\theta)$, then, for example,

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

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 (22) 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 $j = 1, \ldots, N_{dof}^{(1)}$.

We use automatic differentiation [11,21] 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 [15]. 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 [12] with $m = 4$ shifts to solve the controllability and observability Lyapunov equations (5). Figure 7 shows the largest Hankel singular values. For the model reduction, we

Fig. 5 Reference domain Ω_{ref}

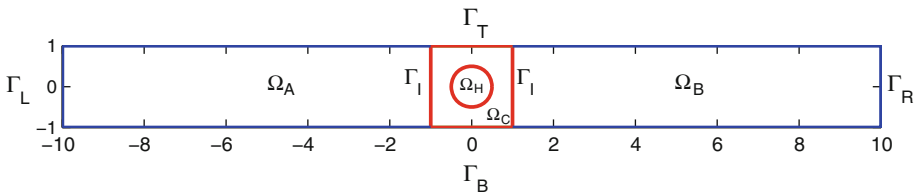
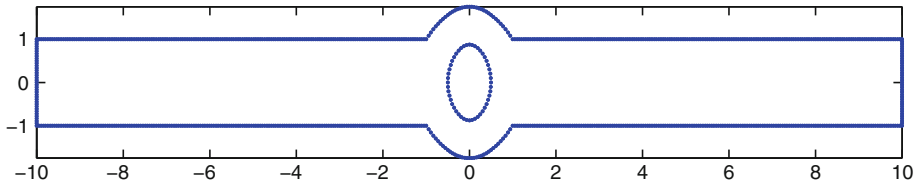


Fig. 6 Optimal domain



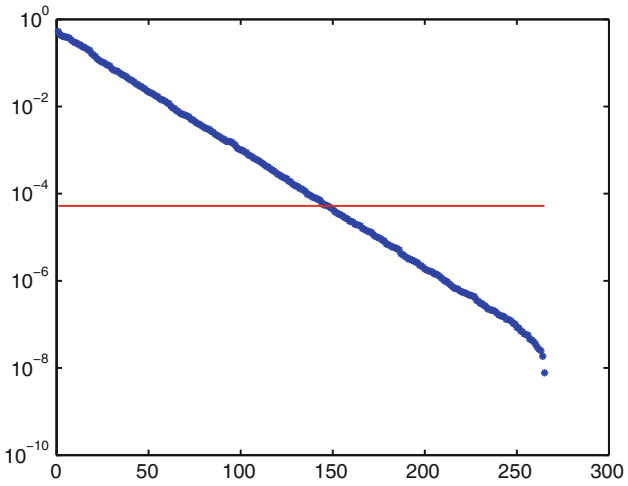


Fig. 7 The largest Hankel singular values and the threshold $10^{-4}\sigma_1$

Table 2 Sizes of the full and the reduced order problems

	$N_{dof}^{(1)}$	N_{dof}
Reduced	147	581
Full	4,280	4,714

Table 3 Optimal shape parameters θ_* and $\widehat{\theta}_*$ (rounded to 5 digits) computed by minimizing the full and the reduced order model, respectively

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)

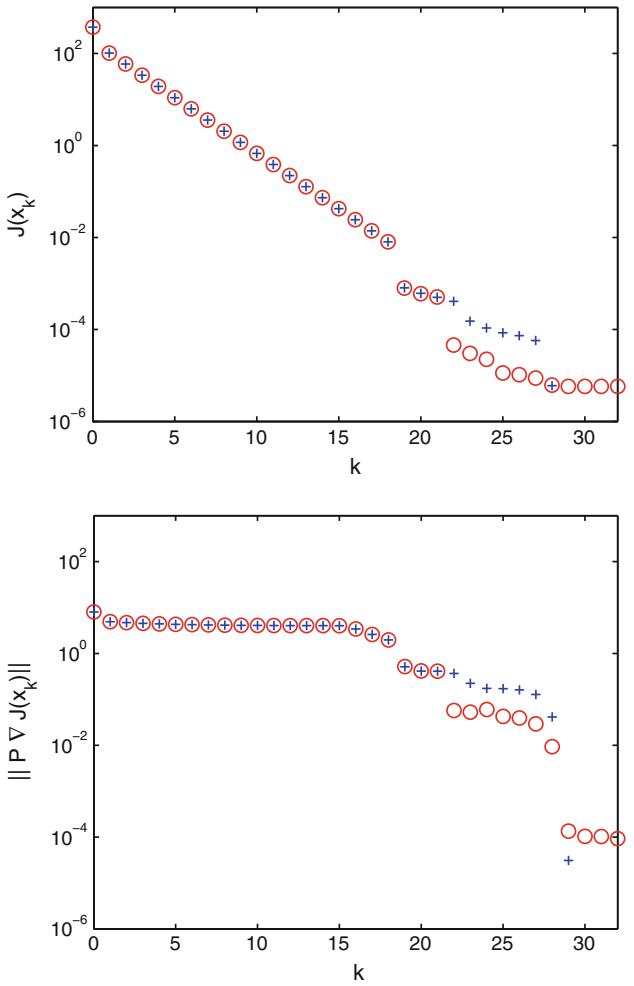


Fig. 8 The convergence histories of the projected BFGS algorithm applied to the full and the reduced order problems. The *top* figure shows the convergence history of the objective functionals for the full (*blue* +) and reduced (*red* ◯) order model. The *bottom* figure shows the convergence history of the projected gradients for the full (*blue* +) and reduced (*red* ◯) order model

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 Fig. 7.

Table 2 displays the sizes for 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 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 2.

The convergence histories of the projected BFGS algorithm applied to the full and the reduced order problems are shown in Fig. 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 2.

6 Conclusions

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.

Our approach can be extended in various ways. In [2] we have extended it to shape optimization problems governed by the incompressible Stokes equations. In this case, the balanced truncation model reduction, the domain decomposition, and their integration needs to be carefully modified due to presence of the incompressibility conditions. It is also possible to admit localized nonlinearities in the PDE, such as those considered in [23,24]. Using model reduction techniques for nonlinear systems such as proper orthognal decomposition (POD) (see e.g., the overview [14]) or extensions of balanced truncation to nonlinear systems [16] one can apply our approach to nonlinear PDEs. However, currently no a-priori error estimates exists for these model reduction techniques and, consequently, no estimate for the error between the solutions of the original optimization problem and of the reduced problem can be obtained.

Appendix

Lemma 4 *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*

$$\mathbf{v}^T\mathcal{A}\mathbf{v} \le -\alpha\|\mathbf{v}\|^2 \quad \forall \mathbf{v} \in \mathbb{R}^N, \tag{51}$$

then the solution of

$$\mathbf{y}'(t) = \mathcal{A}\mathbf{y}(t) + \mathcal{B}\mathbf{u}(t), \quad t \in (0,T), \qquad \mathbf{y}(0) = \mathbf{y}_0 \tag{52}$$

obeys

$$\|\mathbf{y}\|_{L^2} \le \frac{\sqrt{2}}{\sqrt{\alpha}}\|\mathbf{y}_0\| + \frac{2}{\alpha}\|\mathcal{B}\mathbf{u}\|_{L^2} \le \frac{\sqrt{2}}{\sqrt{\alpha}}\|\mathbf{y}_0\| + \frac{2\|\mathcal{B}\|}{\alpha}\|\mathbf{u}\|_{L^2}.$$

Proof We multiply the differential equation (52) by $\mathbf{y}(t)^T$ to obtain

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\|\mathbf{y}(t)\|^2 &= \mathbf{y}(t)^T\mathcal{A}\mathbf{y}(t) + \mathbf{y}(t)^T\mathcal{B}\mathbf{u}(t)\\ &\le -\alpha\|\mathbf{y}(t)\|^2 + \mathbf{y}(t)^T\mathcal{B}\mathbf{u}(t).\end{aligned}$$

If we multiply the previous inequality by $\exp(\alpha t)$ we arrive at

$$\frac{d}{dt}\left(e^{\alpha t}\|\mathbf{y}(t)\|^2\right) \le 2e^{\alpha t}\mathbf{y}(t)^T\mathcal{B}\mathbf{u}(t).$$

Integration from 0 to t gives

$$\|\mathbf{y}(t)\|^2 \le e^{-\alpha t}\|\mathbf{y}_0\|^2 + \int_0^t 2e^{\alpha(\tau-t)}\mathbf{y}(\tau)^T\mathcal{B}\mathbf{u}(\tau)d\tau$$

and integration of this resulting equation from 0 to T yields

$$\begin{aligned}&\int_0^T \|\mathbf{y}(t)\|^2 dt\\ &\le \int_0^T e^{-\alpha t}dt\,\|\mathbf{y}_0\|^2 + \int_0^T\int_0^t 2e^{\alpha(\tau-t)}\mathbf{y}(\tau)^T\mathcal{B}\mathbf{u}(\tau)d\tau dt\\ &\le \frac{1-e^{-\alpha T}}{\alpha}\|\mathbf{y}_0\|^2 + \int_0^T\int_\tau^T 2e^{\alpha(\tau-t)}dt\;\mathbf{y}(\tau)^T\mathcal{B}\mathbf{u}(\tau)d\tau\\ &= \frac{1-e^{-\alpha T}}{\alpha}\|\mathbf{y}_0\|^2 + \int_0^T \frac{2(1-e^{\alpha(\tau-T)})}{\alpha}\,\mathbf{y}(\tau)^T\mathcal{B}\mathbf{u}(\tau)d\tau\\ &\le \frac{1}{\alpha}\|\mathbf{y}_0\|^2 + \int_0^T \frac{2}{\alpha}\,\|\mathbf{y}(\tau)\|\,\|\mathcal{B}\mathbf{u}(\tau)\|d\tau\\ &\le \frac{1}{\alpha}\|\mathbf{y}_0\|^2 + \int_0^T \frac{1}{2}\|\mathbf{y}(\tau)\|^2 + \frac{2}{\alpha^2}\|\mathcal{B}\mathbf{u}(\tau)\|^2 d\tau,\end{aligned}$$

which implies the desired inequality. □

Lemma 5 *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 $\mathbf{v}^T\mathcal{A}\mathbf{v} \le -\alpha\mathbf{v}^T\mathcal{M}\mathbf{v}$ for all $\mathbf{v} \in \mathbb{R}^N$, then the solution of*

$$\mathcal{M}\mathbf{y}'(t) = \mathcal{A}\mathbf{y}(t) + \mathcal{B}\mathbf{u}(t), \quad t \in (0,T) \tag{53}$$

with $\mathbf{y}(0) = \mathbf{y}_0$ obeys

$$\|\mathbf{y}\|_{L^2} \le \frac{\sqrt{2}\|\mathcal{M}^{-1/2}\|\|\mathcal{M}^{1/2}\|}{\sqrt{\alpha}}\|\mathbf{y}_0\| + \frac{2\|\mathcal{M}^{-1}\|}{\alpha}\|\mathcal{B}\mathbf{u}\|_{L^2}.$$

Proof If we multiply (53) by $\mathcal{M}^{-1/2}$ and apply Lemma 4) to the resulting system we obtain the estimate

$$\|\mathcal{M}^{1/2}\mathbf{y}\|_{L^2} \le \frac{\sqrt{2}}{\sqrt{\alpha}}\|\mathcal{M}^{1/2}\mathbf{y}_0\| + \frac{2}{\alpha}\|\mathcal{M}^{-1/2}\mathcal{B}\mathbf{u}\|_{L^2}.$$

This implies

$$\begin{aligned}\|\mathbf{y}\|_{L^2} &= \|\mathcal{M}^{-1/2}\mathcal{M}^{1/2}\mathbf{y}\|_{L^2}\\ &\le \frac{\sqrt{2}\|\mathcal{M}^{-1/2}\|}{\sqrt{\alpha}}\|\mathcal{M}^{1/2}\mathbf{y}_0\| + \frac{2\|\mathcal{M}^{-1/2}\|}{\alpha}\|\mathcal{M}^{-1/2}\mathcal{B}\mathbf{u}\|_{L^2}\\ &\le \frac{\sqrt{2}\|\mathcal{M}^{-1/2}\|\|\mathcal{M}^{1/2}\|}{\sqrt{\alpha}}\|\mathbf{y}_0\| + \frac{2}{\alpha}\|\mathcal{M}^{-1}\|\|\mathcal{B}\mathbf{u}\|_{L^2}.\end{aligned}$$

□

References

1. Akçelik, V., Biros, G., Ghattas, O., Long, K.R., van Bloemen Waanders, B.: A variational finite element method for source inversion for convective-diffusive transport. Finite Elem. Anal. Des. **39**(8), 683–705 (2003)

2. Antil, H., Heinkenschloss, M., Hoppe, R.H.W.: Domain Decomposition and Balanced Truncation Model Reduction for Shape Optimization of the Stokes System. Technical Report TR09–24, Department of Computational and Applied Mathematics, Rice University (2009). Optim. Meth. Softw. (in press)
3. Antoulas, A.C.: Approximation of Large-Scale Dynamical Systems, Advances in Design and Control, vol. 6. Society for Industrial and Applied Mathematics (SIAM), Philadelphia (2005)
4. Benner, P., Mehrmann, V., Sorensen, D.C. (eds.): Dimension Reduction of Large-Scale Systems. Lecture Notes in Computational Science and Engineering, vol. 45. Springer, Heidelberg (2005)
5. Dedé, L., Quarteroni, A.: Optimal control and numerical adaptivity for advection-diffusion equations. ESAIM: Math. Model. Numer. Anal. **39**, 1019–1040 (2005)
6. Dullerud, G.E., Paganini, F.: A Course in Robust Control Theory. Texts in Applied Mathematics, vol. 36. Springer, Berlin (2000)
7. Fatone, L., Gervasio, P., Quarteroni, A.: Multimodels for incompressible flows. J. Math. Fluid Mech. **2**(2), 126–150 (2000)
8. Fatone, L., Gervasio, P., Quarteroni, A.: Multimodels for incompressible flows: iterative solutions for the Navier–Stokes/Oseen coupling. M2AN Math. Model. Numer. Anal. **35**(3), 549–574 (2001)
9. Formaggia, L., Gerbeau, J.F., Nobile, F., Quarteroni, A.: On the coupling of 3D and 1D Navier-Stokes equations for flow problems in compliant vessels. Comput. Methods Appl. Mech. Eng. **191** (6–7), 561–582 (2001)
10. Glover, K.: All optimal Hankel-norm approximations of linear multivariable systems and their L^∞-error bounds. Int. J. Control **39**(6), 1115–1193 (1984)
11. Griewank, A., Walther, A.: Evaluating Derivatives, 2nd edn. Society for Industrial and Applied Mathematics (SIAM), Philadelphia (2008). Principles and techniques of algorithmic differentiation
12. Gugercin, S., Sorensen, D.C., Antoulas, A.C.: A modified low-rank Smith method for large-scale Lyapunov equations. Numer. Algorithms **32**(1), 27–55 (2003)
13. Heinkenschloss, M., Reis, T., Antoulas, A.C.: Balanced Truncation Model Reduction for Systems with Inhomogeneous Initial Conditions. Technical Report TR09-29, Department of Computational and Applied Mathematics, Rice University (2009)
14. Hinze, M., Volkwein, S.: Proper orthogonal decomposition surrogate models for nonlinear dynamical systems: Error estimates and suboptimal control. In: Benner, P., Mehrmann, V., Sorensen, D.C. (eds.) Dimension Reduction of Large-Scale Systems, Lecture Notes in Computational Science and Engineering, vol. 45, pp. 261–306. Springer, Heidelberg (2005)
15. Kelley, C.T.: Iterative Methods for Optimization. SIAM, Philadelphia (1999)
16. Lall, S., Marsden, J.E., Glavaški, S.: A subspace approach to balanced truncation for model reduction of nonlinear control systems. Int. J. Robust Nonlinear Control **12**(6), 519–535 (2002)
17. Lucia, D.J., Beran, P.S., Silva, W.A.: Reduced-order modeling: new approaches for computational physics. Prog. Aerosp. Sci. **40** (1–2), 51–117 (2004)
18. Lucia, D.J., King, P.I., Beran, P.S.: Domain decomposition for reduced-order modeling of a flow with moving shocks. AIAA J. **40**, 2360–2362 (2002)
19. Moore, B.C.: Principal component analysis in linear systems: controllability, observability, and model reduction. IEEE Trans. Automat. Control **26**(1), 17–32 (1981)
20. Quarteroni, A., Tuveri, M., Veneziani, A.: Computational vascular fluid dynamics: problems, models and methods. Comput. Vis. Sci. **2**(4), 163–197 (2000)
21. Rump, S.M.: INTLAB—INTerval LABoratory. In: Csendes, T. (ed.) Developments in Reliable Computing, pp. 77–104. Kluwer, Dordrecht (1999). http://www.ti3.tu-harburg.de/rump/
22. Smith, B., Bjørstad, P., Gropp, W.: Domain Decomposition. Parallel Multilevel Methods for Elliptic Partial Differential Equations. Cambridge University Press, Cambridge (1996)
23. Sun, K.: Domain Decomposition and Model Reduction for Large-Scale Dynamical Systems. Ph.D. thesis, Department of Computational and Applied Mathematics, Rice University, Houston (2008)
24. Sun, K., Glowinski, R., Heinkenschloss, M., Sorensen, D.C. : Domain decomposition and model reduction of systems with local nonlinearities. In: Kunisch, K., Of, G., Steinbach, O. (eds.) Numerical Mathematics and Advanced Applications. ENUMATH 2007., pp. 389–396. Springer, Heidelberg (2008)
25. Toselli, A., Widlund, O.: Domain Decomposition Methods—Algorithms and Theory. Computational Mathematics, vol. 34. Springer, Berlin (2004)
26. Zhou, K., Doyle, J.C., Glover, K.: Robust and Optimal Control. Prentice Hall, Englewood Cliffs (1996)