To appear in *Journal of Computational and Applied Mathematics*

A Spatial Domain Decomposition Method for Parabolic Optimal Control Problems ⋆

Matthias Heinkenschloss [a,*]

[a] *Department of Computational and Applied Mathematics, MS-134, Rice University, 6100 Main Street, Houston, TX 77005-1892*

Michael Herty [b]

[b] *Fachbereich Mathematik, Technische Universität Kaiserslautern, Postfach 30 49, D-67653 Kaiserslautern, Germany*

Abstract

We present a non-overlapping spatial domain decomposition method for the solution of linear–quadratic parabolic optimal control problems. The spatial domain is decomposed into non-overlapping subdomains. The original parabolic optimal control problem is decomposed into smaller problems posed on space-time cylinder subdomains with auxiliary state and adjoint variables imposed as Dirichlet boundary conditions on the space-time interface boundary. The subdomain problems are coupled through Robin transmission conditions. This leads to a Schur complement equation in which the unknowns are the auxiliary state adjoint variables on the space-time interface boundary. The Schur complement operator is the sum of space-time subdomain Schur complement operators. The application of these subdomain Schur complement operators is equivalent to the solution of an subdomain parabolic optimal control problem. The subdomain Schur complement operators are shown to be invertible and the application of their inverses is equivalent to the solution of a related subdomain parabolic optimal control problem. We introduce a new family of Neumann-Neumann type preconditioners for the Schur complement system including several different coarse grid corrections. We compare the numerical performance of our preconditioners with an alternative approach recently introduced by Benamou.

Key words: Parabolic optimal control problems, preconditioners, domain decomposition, Neumann-Neumann methods
1991 MSC: 65K10, 49M05, 65N55, 65F10, 90C06, 90C20

1 Introduction

This paper introduces a new spatial domain decomposition method for the solution of linear-quadratic parabolic optimal control problems. Such problems arise directly in many applications [1–3], but also as subproblems in Newton or sequential quadratic programming methods for the solution of nonlinear parabolic optimal control problems, such as those described in [4–9]. The numerical solution of such problems is difficult because of the large storage requirements arising out of the strong coupling of states, adjoints and controls. The domain decomposition methods introduced in this paper introduce optimization-level parallelism into the solution approach and reduce the amount of permanent storage required.

Domain decomposition methods have been applied previously to linear-quadratic time dependent optimal control problems. They split into time domain decomposition methods [10–13] and spatial domain decomposition methods [14–16]. Like [14–16], the approach introduced in this paper is also based on a decomposition of the spatial domain. The resulting subproblems are smaller linear-quadratic parabolic optimal control problems posed on a spatial-subdomain-time cylinder. The difference between the approaches [14–16] and the approach introduced here lies in the way the subdomain problems are coupled and in the solution method for the coupled subdomain problems.

Our spatial domain decomposition method for linear-quadratic parabolic optimal control problems is based on the so-called Neumann-Neumann domain decomposition methods. Of the domain decomposition method for elliptic partial differential equations, Neumann-Neumann methods are among the most successful ones. Their derivation and discussions of their convergence properties can be found in the books [17–19] and the references given therein. Recently, Neumann-Neumann methods were generalized to solve linear-quadratic elliptic optimal control problems. The results in [21,22,20] have shown that their performance on linear-quadratic elliptic optimal control model problems is comparable to their good performance for single elliptic partial differential equations. This paper extends Neumann-Neumann methods to the solution of linear-quadratic parabolic optimal control problems. We discuss the various possible extensions and give numerical results on their performance. Further we compare Neumann-Neumann preconditioned methods with the approach of Benamou in [15]. We present numerical results on the number of iterations for both approaches and show the dependency on the mesh size and on the regularization parameter for the control.

⋆ This work was supported in part by NSF Grants ACI-0121360 and CNS-0435425 and DAAD Grant D/04/23833.

* Corresponding author
Email addresses: heinken@rice.edu (Matthias Heinkenschloss), herty@rhrk.uni-kl.de (Michael Herty).

To illustrate our ideas, we consider the model problem

$$\begin{aligned}\text{minimize} \quad & \frac{\alpha_1}{2}\int_0^T\int_\Omega (y(x,t)-\hat{y}(x,t))^2 dxdt + \frac{\alpha_2}{2}\int_\Omega (y(x,T)-\hat{y}_T(x))^2 dx \\ & + \frac{\alpha_3}{2}\int_0^T\int_\Omega u^2(x,t)dxdt, \qquad (1a)\end{aligned}$$

$$\begin{aligned}\text{subject to} \quad & \partial_t y(x,t) - \mu\Delta y(x,t) + \mathbf{a}(x)\cdot\nabla y(x,t) \\ & + c(x)y(x,t) = f(x,t)+u(x,t) && \text{in } \Omega\times(0,T), && (1b)\\ & y(x,t) = 0 && \text{on } \partial\Omega\times(0,T), && (1c)\\ & y(x,0) = y_0(x) && \text{in } \Omega, && (1d)\end{aligned}$$

where $\hat{y}, \hat{y}_T, \mathbf{a}, c, f$ are given functions and $\alpha_1, \alpha_2 \geq 0$, $\alpha_3 > 0$, $\mu > 0$ are given parameters. The problem (1) has to be solved for y and u. Detailed model problem assumptions will be introduced in Section 2.

2 The model problem

We collect some well known results that serve as the foundation of the subsequent sections. In particular, we specify the setting for the model problem (1), recall a result on the existence and uniqueness of its solution, and review the well-known necessary and sufficient optimality conditions.

Let $\Omega \subset \mathbb{R}^d$, $d = 1,2,3$, be an open, bounded set with Lipschitz boundary (if $d = 2$ or 3). We consider the state space

$$\mathcal{Y} = W(0,T) = \left\{y \,:\, y \in L^2(0,T;H_0^1(\Omega)), y' \in L^2(0,T;H^{-1}(\Omega))\right\}$$

and the control space

$$\mathcal{U} = L^2(0,T;L^2(\Omega)).$$

We assume that the problem data satisfy $y_0, \hat{y}_T \in L^2(\Omega)$, $\hat{y} \in L^2(\Omega\times(0,T))$, $\mathbf{a} \in W^{1,\infty}(\Omega)$, $c \in L^\infty(\Omega)$, $f \in L^2(\Omega\times(0,T))$ be given functions and let $\alpha_1, \alpha_2 \geq 0$, $\alpha_3 > 0$, $\mu > 0$ be given parameters. We define the bilinear forms $a : H_0^1(\Omega)\times H_0^1(\Omega) \to \mathbb{R}$ and $b : L^2(\Omega)\times H_0^1(\Omega) \to \mathbb{R}$ as

$$\begin{aligned} a(y,\phi) &= \int_\Omega \mu\nabla y(x)\nabla\phi(x) + \mathbf{a}(x)\cdot\nabla y(x)\phi(x) + c(x)y(x)\phi(x)dx, \\ b(u,\phi) &= -\int_\Omega u(x)\phi(x)dx, \end{aligned}$$

respectively, and we use $\langle\cdot,\cdot\rangle_{L^2(\Omega)}$ and $\|\cdot\|_{L^2(\Omega)}$ to denote the inner product and the norm in $L^2(\Omega)$.

We are interested in the solution $y \in Y$, $u \in U$ of the optimal control problem

$$\text{minimize} \quad \frac{\alpha_1}{2}\int_0^T \|y(t)-\hat{y}(t)\|^2_{L^2(\Omega)}dt + \frac{\alpha_2}{2}\|y(T)-\hat{y}_T\|^2_{L^2(\Omega)} + \frac{\alpha_3}{2}\int_0^T \|u(t)\|^2_{L^2(\Omega)}dt, \tag{2a}$$

$$\text{subject to} \quad \langle y'(t),\phi\rangle_{L^2(\Omega)} + a(y(t),\phi) + b(u(t),\phi) = \langle f(t),\phi\rangle_{L^2(\Omega)} \quad \forall \phi \in H^1_0(\Omega), \tag{2b}$$

$$y(0) = y_0. \tag{2c}$$

Theorem 2.1 *The optimal control problem* (2) *has a unique solution* $(u_*, y_*) \in \mathcal{U}\times\mathcal{Y}$*, which, together with the adjoint variable* $p_* \in \mathcal{Y}$*, is characterized by the necessary and sufficient optimality conditions*

$$-\langle p'(t),\psi\rangle_{L^2(\Omega)} + a(\psi,p(t)) = -\alpha_1\langle y(t)-\hat{y}(t),\psi\rangle_{L^2(\Omega)}, \tag{3a}$$

$$p(T) = -\alpha_2(y(T)-\hat{y}_T), \tag{3b}$$

$$\alpha_3\langle u(t),\mu\rangle_{L^2(\Omega)} + b(\mu,p(t)) = 0, \tag{3c}$$

$$\langle y'(t),\phi\rangle_{L^2(\Omega)} + a(y(t),\phi) + b(u(t),\phi) = \langle f(t),\phi\rangle_{L^2(\Omega)}, \tag{3d}$$

$$y(0) = y_0. \tag{3e}$$

for all $\psi,\phi \in H^1_0(\Omega)$*,* $\mu \in L^2(\Omega)$*.*

Proof: Since

$$a(y,\phi) = \int_\Omega \mu\nabla y(x)\nabla\phi(x) + \tfrac{1}{2}\mathbf{a}(x)\cdot\nabla y(x)\phi(x) - \tfrac{1}{2}\mathbf{a}(x)\cdot\nabla\phi(x)y(x) + (c(x) - \tfrac{1}{2}\nabla\cdot\mathbf{a}(x))y(x)\phi(x)dx$$

for all $y,\phi \in H^1_0(\Omega)$, we have

$$a(y,y) + \lambda\|y\|^2_{L^2(\Omega)} \geq \int_\Omega \mu(\nabla y(x)\nabla y(x) + y(x)y(x))dx$$

for all $y \in H^1_0(\Omega)$ and all $\lambda \geq \mu + \|c - \frac{1}{2}\nabla\cdot\mathbf{a}\|_{L^\infty(\Omega)}$. Furthermore, there exists a constant $\kappa > 0$, depending on $\mu, \mathbf{a}, c$ such that $a(y,\phi) \leq \kappa\|y\|_{H^1(\Omega)}\|\phi\|_{H^1(\Omega)}$ for all $y,\phi \in H^1(\Omega)$. The statement of the theorem follows from, e.g., [3, p. 114,116]. □

The adjoint equation (3a) is the weak form of the

$$-\partial_t p(x,t) - \mu\Delta p(x,t) - \mathbf{a}(x)\cdot\nabla p(x,t) + (c(x) - \nabla\cdot\mathbf{a}(x))p(x,t) = -\alpha_1(y(x,t)-\widehat{y}(x,t)) \quad \text{in } \Omega\times(0,T), \tag{4a}$$

$$p(x,t) = 0 \quad \text{on } \partial\Omega\times(0,T), \tag{4b}$$

$$p(x,T) = -\alpha_2(y(x,T)-\widehat{y}_T(x)) \quad \text{in } \Omega. \tag{4c}$$

Equation (3c) states that

$$\alpha_3 u - p = 0 \tag{5}$$

a.e. in $\Omega\times(0,T)$.

2.1 Semi-discretization

Let $V^h \subset H_0^1(\Omega)$, $U^h \subset L^2(\Omega)$ be finite dimensional subspaces with bases $\phi_1, \ldots, \phi_m$ and $\mu_1, \ldots, \mu_n$, respectively. We approximate the states and controls by $y_h \in H^1(0,T;V^h)$, $u_h \in L^2(0,T;U^h)$ defined as

$$y_h(t) = \sum_{l=1}^{m} y_k(t)\phi_k, \quad u_h(t) = \sum_{l=1}^{n} u_k(t)\mu_k. \tag{6}$$

We define $\mathbf{A} \in \mathbb{R}^{m\times m}$, $\mathbf{B} \in \mathbb{R}^{m\times n}$, $\mathbf{M} \in \mathbb{R}^{m\times m}$, $\mathbf{Q} \in \mathbb{R}^{n\times n}$, $\mathbf{f} \in L^2(0,T;\mathbb{R}^m)$, $\mathbf{c} \in L^2(0,T;\mathbb{R}^m)$ and $\mathbf{d} \in \mathbb{R}^m$ as follows:

$$\mathbf{A}_{jk} = a(\phi_k, \phi_j), \quad \mathbf{M}_{jk} = \langle \phi_k, \phi_j \rangle_{L^2(\Omega)},$$

$$\mathbf{c}_j(t) = -\alpha_1 \langle \hat{y}(t), \phi_j \rangle_{L^2(\Omega)}, \quad \mathbf{f}_j(t) = \langle f(t), \phi_j \rangle_{L^2(\Omega)}, \quad \mathbf{d}_j = -\alpha_2 \langle \hat{y}_T, \phi_j \rangle_{L^2(\Omega)}$$

for $j,k = 1,\ldots,m$, and

$$\mathbf{B}_{jk} = b(\mu_k, \phi_j), \quad \mathbf{Q}_{jk} = \langle \mu_k, \mu_j \rangle_{L^2(\Omega)}$$

for $j,k = 1,\ldots,n$. We set $\mathbf{y}(t) = (y_1(t),\ldots,y_m(t))^T$ and $\mathbf{u}(t) = (u_1(t),\ldots,u_n(t))^T$ where y_i, u_i, p_i are the functions in (6).

We now replace y,u by y_h, u_h defined in (6) and require (2b) to hold for $\phi = \phi_k$, $k = 1,\ldots,m$. This finite element semi-discretization of the optimal control problem (2) leads to a large-scale linear quadratic problem of the form

$$\begin{aligned}
\text{minimize } & \int_0^T \frac{\alpha_1}{2}\mathbf{y}(t)^T\mathbf{M}\mathbf{y}(t) + \mathbf{c}(t)^T\mathbf{y}(t)dt + \frac{\alpha_2}{2}\mathbf{y}(T)^T\mathbf{M}\mathbf{y}(T) + \mathbf{d}^T\mathbf{y}(T) \\
& + \int_0^T \frac{\alpha_3}{2}\mathbf{u}(t)^T\mathbf{Q}\mathbf{u}(t)dt, && \text{(7a)} \\
\text{subject to } & \mathbf{M}\mathbf{y}'(t) + \mathbf{A}\mathbf{y}(t) + \mathbf{B}\mathbf{u}(t) = \mathbf{f}(t), \quad t \in (0,T), && \text{(7b)} \\
& \mathbf{y}(0) = \mathbf{y}_0. && \text{(7c)}
\end{aligned}$$

The necessary and sufficient optimality conditions for (7) are given by

$$\begin{aligned}
-\mathbf{M}\mathbf{p}'(t) + \mathbf{A}^T\mathbf{p}(t) &= -\alpha_1\mathbf{M}\mathbf{y}(t) - \mathbf{c}(t), && \text{(8a)} \\
\mathbf{M}\mathbf{p}(T) &= -\alpha_2\mathbf{M}\mathbf{y}(T) - \mathbf{d}, && \text{(8b)} \\
\alpha_3\mathbf{Q}\mathbf{u}(t) + \mathbf{B}^T\mathbf{p}(t) &= 0, && \text{(8c)} \\
\mathbf{M}\mathbf{y}'(t) + \mathbf{A}\mathbf{y}(t) + \mathbf{B}\mathbf{u}(t) &= \mathbf{f}(t), && \text{(8d)} \\
\mathbf{y}(0) &= \mathbf{y}_0. && \text{(8e)}
\end{aligned}$$

The system (8) is equivalent to the semi–discretization of (3) obtained by replacing y,u,p in (3) by y_h, u_h, p_h defined in (6) and requiring the equations (3) to hold for all $\psi = \phi_k$, $k = 1,\ldots,m$, $\mu = \mu_k$, $k = 1,\ldots,n$, and $\phi = \phi_k$, $k = 1,\ldots,m$.

Remark 2.2 *We use linear finite element methods. For advection dominated problems with coarse grids (relative to the ratio of diffusion to advection), standard linear finite element methods lead to spurious oscillations in the computed solution. In this case stabilization methods or other modifications are needed, see, e.g., [23,24]. Stabilization methods can be incorporated easily into our discretization as described, e.g., in [25,23,24].*

3 Domain decomposition Schur complement formulation of the model problem

Our domain decomposition approach is formulated for the semi–discrete problem (7). We preview it using the original problem formulation (1). We subdivide the domain Ω into non-overlapping subdomains $\Omega_1,\ldots,\Omega_s$. The interface between the subdomains is denoted by Γ. We introduce auxiliary states and adjoints y_Γ, p_Γ defined on $\Gamma\times(0,T)$. Now we restrict the optimality conditions (1b–d), (4), (5) to subdomain Ω_i and impose the Dirichlet conditions that the states and adjoints match y_Γ, p_Γ, respectively, on $\Gamma\cap\partial\Omega_i\times(0,T)$. Let y_i,u_i,p_i be the state, control, and adjoint components of the solutions of these subproblems. They can be viewed as functions of the interface variables y_Γ, p_Γ. The subdomain states y_i, controls u_i, and adjoints p_i are the restrictions of the solution to the original optimality conditions (1b–d), (4), (5), if the subdomain states and adjoints satisfy certain transmission conditions at the subdomain interfaces. Following [26,27] we require that

$$\begin{aligned}
\mu\tfrac{\partial}{\partial\mathbf{n}_i}y_i(x,t)-(\tfrac{1}{2}\mathbf{a}(x)\mathbf{n}_i)y_i(x,t) &= -\Big(\mu\tfrac{\partial}{\partial\mathbf{n}_j}y_j(x,t)-(\tfrac{1}{2}\mathbf{a}(x)\mathbf{n}_j)y_j(x,t)\Big),\\
\mu\tfrac{\partial}{\partial\mathbf{n}_i}p_i(x,t)+(\tfrac{1}{2}\mathbf{a}(x)\mathbf{n}_i)p_i(x,t) &= -\Big(\mu\tfrac{\partial}{\partial\mathbf{n}_j}p_j(x,t)+(\tfrac{1}{2}\mathbf{a}(x)\mathbf{n}_j)p_j(x,t)\Big)
\end{aligned}\tag{9}$$

on $\partial\Omega_i\cap\partial\Omega_j\times(0,T)$ for adjacent subdomains Ω_i,Ω_j. Here $\mathbf{n}_i$ denotes the unit outward normal for subdomain Ω_i. Since the subdomain states y_i, controls u_i, and adjoints p_i are functions of interface variables y_Γ, p_Γ, the collection of transmission conditions (9) leads to an operator equation in y_Γ, p_Γ. This equation will be solved iteratively, using Krylov subspace methods. The properties of the system operator are used to derive preconditioners.

3.1 Domain decomposition in space

We discretize (2) using conforming linear finite elements. Given a triangulation $\{T_l\}$ of Ω, the space V^h used in the discretization of the states is given by

$$V^h=\left\{v\in H_0^1(\Omega)\ :\ v|_{T_l}\in P^1(T_l)\text{ for all }k\right\}.$$

We divide Ω into nonoverlapping subdomains Ω_i, $i=1,\ldots,s$, such that each T_l belongs to exactly one $\overline{\Omega_i}$. We define $\Gamma_i=\partial\Omega_i\setminus\partial\Omega$ and $\Gamma=\cup_{i=1}^s\Gamma_i$.

Let $\{x_j\}_{j=1}^m$ be the set of vertices of $\{T_l\}$ that lie inside Ω and let $\{\phi_j\}_{j=1}^m$ be the piecewise linear nodal basis for V^h. Let m_I^i be the number of vertices in Ω_i, let m_Γ^i be the number

of vertices on $\Gamma \cap \partial\Omega_i$, and let m_Γ be the number of vertices on the subdomain interfaces Γ. Hence the number of discretized state variables for a given time t is given by $m = m_\Gamma + \sum_{i=1}^{s} m_I^i$.

To approximate the control we define the discrete spaces

$$U_i^h = \left\{ u \in C^0(\Omega_i) \ : \ u \text{ is linear on } \Omega_i \cap T_l \text{ for all } T_l \subset \overline{\Omega}_i \right\},$$

which we identify with a subspace of $L^2(\Omega)$ by extending functions $u_i \in U_i^h$ by zero onto Ω. The space of semi–discrete control is $L^2(0,T,U^h)$ where

$$U^h = \cup_{i=1}^{s} U_i^h \subset L^2(\Omega).$$

Note that our controls are continuous on each Ω_i, $i = 1,\ldots,s$, and linear on each $\Omega_i \cap T_l$, but that are not assumed to be continuous at $\partial\Omega_i \cap \partial\Omega_j$, $i \neq j$. Other control discretizations might introduce discrete controls defined on a small band of with h around the interface Γ. Since $u \in L^2(\Omega \times (0,T))$ such controls would not be meaningful as the mesh size h is reduced. See [21,22] for more discussion.

Let $\{\mu_j^i\}_{j=1}^{n^i}$ be the piecewise linear nodal basis for U_i^h, where n^i is the number of vertices in $\overline{\Omega}_i$. We identify μ_j^i with a function in $L^2(\Omega)$ by extending μ_j^i by zero outside $\overline{\Omega}_i$. We set

$$\mu_1 = \mu_1^1, \ldots, \mu_{n^i} = \mu_{n^1}^1,\ \mu_{n^1+1} = \mu_1^2, \ldots, \mu_{n^1+n^2} = \mu_{n^2}^2, \ldots.$$

The number of discretized control variables for a given time t is given by $n = \sum_{i=1}^{s} n^i$.

3.2 Decomposition of the semi–discretized model problem

3.2.1 The decomposed optimality conditions

In the presence of advection terms, the spatial domain decomposition requires a careful choice of bilinear forms for local subproblems. See [19, Sec 11.5.1] for a discussion. Following [26] we define the local bilinear forms

$$a_i(y,\phi) = \int_{\Omega_i} \mu \nabla y(x) \nabla \phi(x) + \mathbf{a}(x) \nabla y(x) \phi(x) + c(x) y(x) \phi(x) dx - \int_{\Gamma_i} \tfrac{1}{2} \mathbf{a}(x) \mathbf{n}_i y(x) \phi(x) dx, \tag{10}$$

$i = 1,\ldots,s$, where $\mathbf{n}_i$ is the unit outward normal for the ith subdomain. We have

$$\sum_{i=1}^{s} \int_{\Gamma_i} \mathbf{a}(x) \mathbf{n}_i y(x) \phi(x) dx = \sum_{i=1}^{s} \sum_{j \neq i} \int_{\Omega_i \cap \Omega_j} \mathbf{a}(x) \mathbf{n}_i y(x) \phi(x) dx = 0, \tag{11}$$

since each interface boundary segment $\Omega_i \cap \Omega_j$ appears twice in the above sum and $\mathbf{n}_i = -\mathbf{n}_j$ on $\Omega_i \cap \Omega_j$. Consequently

$$\sum_{i=1}^{s} a_i(y,\phi) = a(y,\phi) \quad \forall y,\phi \in H_0^1(\Omega). \tag{12}$$

Integration by parts yields

$$a_i(y,\phi) = \int_{\Omega_i} \mu \nabla y(x) \nabla \phi(x) + \tfrac{1}{2}\mathbf{a}(x)\nabla y(x)\phi(x) - \tfrac{1}{2}\mathbf{a}(x)\nabla\phi(x)y(x) + (c(x) - \tfrac{1}{2}\nabla\mathbf{a}(x))y(x)\phi(x)dx$$

for all $y,\phi \in H_0^1(\Omega)$, $i = 1,\ldots,s$. Hence

$$a_i(y,y) + \lambda_i \|y\|^2_{L^2(\Omega_i)} \geq \int_{\Omega_i} \mu(\nabla y(x)\nabla y(x) + y(x)y(x))dx \tag{13}$$

for all $y \in H_0^1(\Omega)$ and all $\lambda_i \geq \mu + \|c - \frac{1}{2}\nabla\mathbf{a}\|_{L^\infty(\Omega_I)}$, $i = 1,\ldots,s$. The validity of (13) motivates the choice (10) of the local bilinear form, instead of the naive choice

$$\begin{aligned}
&\int_{\Omega_i} \mu\nabla y(x)\nabla\phi(x) + \mathbf{a}(x)\nabla y(x)\phi(x) + c(x)y(x)\phi(x)dx \\
&= \int_{\Omega_i} \mu\nabla y(x)\nabla\phi(x) + \tfrac{1}{2}\mathbf{a}(x)\nabla y(x)\phi(x) - \tfrac{1}{2}\mathbf{a}(x)\nabla\phi(x)y(x) + (c(x) - \tfrac{1}{2}\nabla\mathbf{a}(x))y(x)\phi(x)dx \\
&\quad + \tfrac{1}{2}\int_{\partial\Omega_i} \mathbf{a}(x)\mathbf{n}_i y(x)\phi(x)dx,
\end{aligned}$$

which due to the boundary integral may not allow an estimate of the form (13).

For $i = 1,\ldots,s$, we define the submatrices $\mathbf{A}^i_{II} \in \mathbb{R}^{m^i_I \times m^i_I}$, $\mathbf{A}^i_{\Gamma I} \in \mathbb{R}^{m^i_\Gamma \times m^i_I}$, $\mathbf{A}^i_{I\Gamma} \in \mathbb{R}^{m^i_I \times m^i_\Gamma}$, and $\mathbf{A}^i_{\Gamma\Gamma} \in \mathbb{R}^{m^i_\Gamma \times m^i_\Gamma}$, where, as before, m^i_I is the number of nodes in Ω_i and m^i_Γ is the number of nodes in $\Gamma \cap \partial\Omega_i$, as follows. Let i_k be the global node number of the kth node in Ω_i and let γ_k be the global node number of the kth node in $\Gamma \cap \partial\Omega_i$. We set

$$\begin{aligned}
(\mathbf{A}^i_{II})_{jk} &= a_i(\phi_{i_k},\phi_{i_j}), \quad x_{i_j}, x_{i_k} \in \Omega_i, \\
(\mathbf{A}^i_{I\Gamma})_{jk} &= a_i(\phi_{\gamma_k},\phi_{i_j}), \quad x_{i_j} \in \Omega_i,\ x_{\gamma_k} \in \Gamma \cap \partial\Omega_i, \\
(\mathbf{A}^i_{\Gamma I})_{jk} &= a_i(\phi_{i_k},\phi_{\gamma_j}), \quad x_{\gamma_j} \in \Gamma \cap \partial\Omega_i,\ x_{i_k} \in \Omega_i, \\
(\mathbf{A}^i_{\Gamma\Gamma})_{jk} &= a_i(\phi_{\gamma_k},\phi_{\gamma_j}), \quad x_{\gamma_j}, x_{\gamma_k} \in \Gamma \cap \partial\Omega_i,
\end{aligned}$$

and $\mathbf{A}_{\Gamma\Gamma} = \sum_{i=1}^s (I^i_\Gamma)^T \mathbf{A}^i_{\Gamma\Gamma} I^i_\Gamma$, where I^i_Γ is a matrix of size $m^i_\Gamma \times m_\Gamma$ with entries given by zero or one which maps a vector of coefficient unknowns on the interface boundary Γ to a subvector with coefficient unknowns associated with the interface boundary $\Gamma \cap \partial\Omega_i$ of the ith subdomain. Note that the modification (10) of the local bilinear form a_i only changes

$\mathbf{A}^i_{\Gamma\Gamma}$. Because of the identity (12), the stiffness matrix can be written as

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}^1_{II} & & & \mathbf{A}^1_{I\Gamma} I^1_\Gamma \\ & \ddots & & \vdots \\ & & \mathbf{A}^s_{II} & \mathbf{A}^s_{I\Gamma} I^s_\Gamma \\ (I^1_\Gamma)^T \mathbf{A}^1_{\Gamma I} & \dots & (I^s_\Gamma)^T \mathbf{A}^s_{\Gamma I} & \mathbf{A}_{\Gamma\Gamma} \end{pmatrix},$$

after a suitable reordering of rows and columns. Similar decompositions can be introduced for $\mathbf{M}$, $\mathbf{c}(t)$, and $\mathbf{d}$. For example, for $i = 1, \dots, s$, we define

$$\begin{aligned} (\mathbf{d}^i_I)_j &= -\alpha_2 \langle \hat{y}_T, \phi_{i_j} \rangle_{L^2(\Omega_i)}, \quad x_{i_j} \in \Omega_i, \\ (\mathbf{d}^i_\Gamma)_j &= -\alpha_2 \langle \hat{y}_T, \phi_{\gamma_j} \rangle_{L^2(\Omega_i)}, \quad x_{\gamma_j} \in \Gamma \cap \partial\Omega_i, \end{aligned}$$

and $\mathbf{d}_\Gamma = \sum_{i=1}^s (I^i_\Gamma)^T \mathbf{d}^i_\Gamma$. After a suitable reordering, the vector $\mathbf{d}$ can be written as

$$\mathbf{d} = \begin{pmatrix} \mathbf{d}^1_I \\ \vdots \\ \mathbf{d}^s_I \\ \mathbf{d}_\Gamma \end{pmatrix}.$$

The vectors $\mathbf{y}_0$, $\mathbf{y}(t)$, and $\mathbf{p}(t)$ are partitioned correspondingly. For example, $\mathbf{y}^i_I(t)$ denotes the subvector of $\mathbf{y}(t)$ with indices k such that $x_k \in \Omega_i$, $\mathbf{y}_\Gamma(t)$ denotes the subvector of $\mathbf{y}(t)$ with indices k such that $x_k \in \Gamma$, and $\mathbf{y}^i_\Gamma(t)$ denotes the subvector of $\mathbf{y}(t)$ with indices k such that $x_k \in \Gamma \cap \partial\Omega_i$. Defining functions analogously to (6), $\mathbf{y}^i_I(t)$ represents a function in $H^1(0,T;V_{i,0})$, $\mathbf{y}_\Gamma(t)$ represents a function in $H^1(0,T;V_\Gamma)$, and $\mathbf{y}^i_\Gamma(t)$ represents a function in $H^1(0,T;V_{i,\Gamma_i})$. The subvectors $\mathbf{y}^i_I(t)$, $\mathbf{p}^i_\Gamma(t)$ and $\mathbf{p}_\Gamma(t)$ are interpreted analogously.

For $i = 1, \dots, s$ we define

$$b_i : L^2(\Omega_i) \times H^1(\Omega_i) \to \mathbb{R}, \quad b_i(u,\phi) = -\int_{\Omega_i} u_i(x)\phi_i(x)dx.$$

and the submatrices $\mathbf{B}^i_{II} \in \mathbb{R}^{m^i_I \times n^i}$, $\mathbf{B}^i_{\Gamma I} \in \mathbb{R}^{m^i_\Gamma \times n^i}$ with entries

$$\begin{aligned} (\mathbf{B}^i_{II})_{jk} &= b_i(\mu^i_k, \phi_{i_j}), \quad x_{i_j} \in \Omega_i, x_k \in \overline{\Omega_i}, \\ (\mathbf{B}^i_{\Gamma I})_{jk} &= b_i(\mu^i_k, \phi_{i_j}), \quad x_{i_j} \in \partial\Omega_i \setminus \partial\Omega, x_k \in \overline{\Omega_i}. \end{aligned}$$

After a suitable reordering of rows and columns, the matrix $\mathbf{B}$ can be written as

$$\mathbf{B} = \begin{pmatrix} \mathbf{B}_{II}^1 & & \\ & \ddots & \\ & & \mathbf{B}_{II}^s \\ (I_\Gamma^1)^T \mathbf{B}_{\Gamma I}^1 & \dots & (I_\Gamma^s)^T \mathbf{B}_{\Gamma I}^s \end{pmatrix}.$$

Note that in our particular control discretization, all basis functions μ_k^i for the discretised control u_h have support in only one subdomain Ω_i. Consequently, there is no $\mathbf{B}_{\Gamma\Gamma}^i$. The matrix $\mathbf{Q}$ and the vector $\mathbf{u}(t)$ can be decomposed analogously into

$$\mathbf{Q} = \begin{pmatrix} \mathbf{Q}_{II}^1 & & \\ & \ddots & \\ & & \mathbf{Q}_{II}^s \end{pmatrix}, \quad \mathbf{u}(t) = \begin{pmatrix} \mathbf{u}_I^1(t) \\ \vdots \\ \mathbf{u}_I^s(t) \end{pmatrix}.$$

The function $\mathbf{u}_I^i(t)$ represents a function in $L^2(0,T;U_i)$. Due to our control discretization in space, there is no $\mathbf{u}_\Gamma(t)$. Such an interface control would be semi–discrete version of $u|_\Gamma$, which is not defined since $u \in L^2(\Omega \times (0,T))$.

As before, let

$$I_\Gamma^i \in \mathbb{N}^{m_\Gamma^i \times m_\Gamma} \tag{14a}$$

be the matrix with zero or one entries that extracts out of a vector $\mathbf{v}_\Gamma \in \mathbb{R}^{m_\Gamma}$ the subvector $\mathbf{v}_\Gamma^i \in \mathbb{R}^{m_\Gamma^i}$ whose components correspond to vertices $x_k \in \Gamma \cap \partial\Omega_i$ and let

$$\mathbf{I}_\Gamma^i = \begin{pmatrix} I_\Gamma^i & \\ & I_\Gamma^i \end{pmatrix}. \tag{14b}$$

For given t, we can partition the semi–discrete states, adjoints, and controls into $\mathbf{y}_I^i(t)$, $i = 1,\dots,s$, $\mathbf{y}_\Gamma(t)$, $\mathbf{p}_I^i(t)$, $i = 1,\dots,s$, $\mathbf{p}_\Gamma(t)$, and $\mathbf{u}_I^i(t)$, $i = 1,\dots,s$, respectively. We define $\mathbf{y}_\Gamma^i(t) = I_\Gamma^i \mathbf{y}_\Gamma(t)$, $\mathbf{p}_\Gamma^i(t) = I_\Gamma^i \mathbf{p}_\Gamma(t)$, $i = 1,\dots,s$. The optimality conditions (8) can now be

decomposed into the systems

$$
\begin{aligned}
&-\mathbf{M}_{II}^i \frac{d}{dt}\mathbf{p}_I^i(t) + (\mathbf{A}_{II}^i)^T \mathbf{p}_I^i(t) + \alpha_1 \mathbf{M}_{II}^i \mathbf{y}_I^i(t) \\
&-\mathbf{M}_{I\Gamma}^i \frac{d}{dt}\mathbf{p}_\Gamma^i(t) + (\mathbf{A}_{\Gamma I}^i)^T \mathbf{p}_\Gamma^i(t) + \alpha_1 \mathbf{M}_{I\Gamma}^i \mathbf{y}_\Gamma^i(t) = -\mathbf{c}_I^i(t), \qquad t \in (0,T) \qquad (15a)\\
&\mathbf{M}_{II}^i \mathbf{p}_I^i(T) + \mathbf{M}_{I\Gamma}^i \mathbf{p}_\Gamma^i(T) + \alpha_2 \mathbf{M}_{II}^i \mathbf{y}_I^i(T) + \alpha_2 \mathbf{M}_{I\Gamma}^i \mathbf{y}_\Gamma^i(T) = -\mathbf{d}_I^i, \qquad (15b)\\
&\alpha_3 \mathbf{Q}_{II} \mathbf{u}_I^i(t) + (\mathbf{B}_{II}^i)^T \mathbf{p}_I^i(t) + (\mathbf{B}_{\Gamma I}^i)^T \mathbf{p}_\Gamma^i(t) = 0, \qquad t \in (0,T) \qquad (15c)\\
&\mathbf{M}_{II}^i \frac{d}{dt}\mathbf{y}_I^i(t) + \mathbf{A}_{II}^i \mathbf{y}_I^i(t) \\
&+\mathbf{M}_{I\Gamma}^i \frac{d}{dt}\mathbf{y}_\Gamma^i(t) + \mathbf{A}_{I\Gamma}^i \mathbf{y}_\Gamma^i(t) + \mathbf{B}_{II}^i \mathbf{u}_I^i(t) = \mathbf{f}_I^i(t), \qquad t \in (0,T) \qquad (15d)\\
&\mathbf{y}_I^i(0) = (\mathbf{y}_0)_I^i. \qquad (15e)
\end{aligned}
$$

for $i = 1, \ldots, s$, and into the interface coupling condition

$$
\begin{aligned}
&\sum_{i=1}^{s} (I_\Gamma^i)^T \Big(\mathbf{M}_{\Gamma I}^i \mathbf{p}_I^i(T) + \mathbf{M}_{\Gamma\Gamma}^i \mathbf{p}_\Gamma^i(T) + \alpha_2 \mathbf{M}_{\Gamma I}^i \mathbf{y}_I^i(T) + \alpha_2 \mathbf{M}_{\Gamma\Gamma}^i \mathbf{y}_\Gamma^i(T) \Big) = -\mathbf{d}_\Gamma, \qquad (16a)\\
&\sum_{i=1}^{s} (I_\Gamma^i)^T \Big(-\mathbf{M}_{\Gamma I}^i \frac{d}{dt}\mathbf{p}_I^i(t) + (\mathbf{A}_{I\Gamma}^i)^T \mathbf{p}_I^i(t) - \mathbf{M}_{\Gamma\Gamma}^i \frac{d}{dt}\mathbf{p}_\Gamma^i(t) + (\mathbf{A}_{\Gamma\Gamma}^i)^T \mathbf{p}_\Gamma^i(t) \\
&\qquad +\alpha_1 \mathbf{M}_{\Gamma I}^i \mathbf{y}_I^i(t) + \alpha_1 \mathbf{M}_{\Gamma\Gamma}^i \mathbf{y}_\Gamma^i(t) \Big) = -\mathbf{c}_\Gamma(t), \quad t \in (0,T), \qquad (16b)\\
&\sum_{i=1}^{s} (I_\Gamma^i)^T \Big(\mathbf{M}_{\Gamma I}^i \frac{d}{dt}\mathbf{y}_I^i(t) + \mathbf{A}_{\Gamma I}^i \mathbf{y}_I^i(t) + \mathbf{M}_{\Gamma\Gamma}^i \frac{d}{dt}\mathbf{y}_\Gamma^i(t) \\
&\qquad +\mathbf{A}_{\Gamma\Gamma}^i \mathbf{y}_\Gamma^i(t) + \mathbf{B}_{\Gamma I}^i \mathbf{u}_I^i(t) \Big) = \mathbf{f}_\Gamma(t), \qquad t \in (0,T), \qquad (16c)
\end{aligned}
$$

$$
\mathbf{y}_\Gamma^i(0) = I_\Gamma^i(\mathbf{y}(0))_\Gamma, \quad i = 1, \ldots, s. \qquad (16d)
$$

We now view the solution $\mathbf{y}_I^i, \mathbf{u}_I^i, \mathbf{p}_I^i$ of (15) as a function of $\mathbf{y}_\Gamma, \mathbf{p}_\Gamma$. The equation (16) then represents a system of equations in $\mathbf{y}_\Gamma, \mathbf{p}_\Gamma$. Before, we give a precise statement of this system, we give an interpretation of the system (15) as the optimality conditions of an optimal control problem.

Theorem 3.1 *Let $\mathbf{y}_\Gamma^i, \mathbf{p}_\Gamma^i$ be given. The system* (15) *are the necessary and sufficient optimality conditions for the following subdomain optimal control problem in the variables $\mathbf{y}_I^i, \mathbf{u}_I^i$ and adjoint variable $\mathbf{p}_I^i$.*

$$
\begin{aligned}
\textit{Minimize} \quad & \int_0^T \frac{\alpha_1}{2} \mathbf{y}_I^i(t)^T \mathbf{M}_{II}^i \mathbf{y}_I^i(t) \\
& \quad + \left(\mathbf{c}_I^i(t) - \mathbf{M}_{I\Gamma}^i \frac{d}{dt} \mathbf{p}_\Gamma^i(t) + (\mathbf{A}_{\Gamma I}^i)^T \mathbf{p}_\Gamma^i(t) + \alpha_1 \mathbf{M}_{I\Gamma}^i \mathbf{y}_\Gamma^i(t) \right)^T \mathbf{y}_I^i(t) dt \\
& + \frac{\alpha_2}{2} \mathbf{y}_I^i(T)^T \mathbf{M}_{II}^i \mathbf{y}_I^i(T) + \left(\mathbf{d}_I^i + \mathbf{M}_{I\Gamma}^i \mathbf{p}_\Gamma^i(T) + \alpha_2 \mathbf{M}_{I\Gamma}^i \mathbf{y}_\Gamma^i(T) \right)^T \mathbf{y}_I^i(T) \\
& + \int_0^T \frac{\alpha_3}{2} \mathbf{u}_I^i(t)^T \mathbf{Q}_{II}^i \mathbf{u}_I^i(t) + \mathbf{p}_\Gamma^i(t)^T (\mathbf{B}_{\Gamma I}^i) \mathbf{u}_I^i(t) dt, \qquad (17a) \\
\textit{subject to} \quad & \mathbf{M}_{II}^i \frac{d}{dt} \mathbf{y}_I^i(t) + \mathbf{A}_{II}^i \mathbf{y}_I^i(t) + \mathbf{B}_{II}^i \mathbf{u}_I^i(t) \\
& = \mathbf{f}_I^i(t) - \mathbf{M}_{I\Gamma}^i \frac{d}{dt} \mathbf{y}_\Gamma^i(t) - \mathbf{A}_{I\Gamma}^i \mathbf{y}_\Gamma^i(t) \quad t \in (0,T) \qquad (17b) \\
& \mathbf{y}_I^i(0) = (\mathbf{y}_0)_I^i. \qquad (17c)
\end{aligned}
$$

Proof: The proof is standard [28] and is omitted. □

In Appendix A.1, we will give an interpretation of the subdomain problems (15) as well as of the interface coupling conditions (16) in terms of the original problem formulation. In particular, we will show that the subdomain problems (15) can formally be interpreted as the optimality conditions of a semi–discretized version of a subdomain optimal control problem that is a restriction of the original optimal control problem (1) to the subdomain Ω_i with Dirichlet conditions for the state on the subdomain interface Γ_i and with an addition to the objective function that arises from the transmission condition (9) for the state. Moreover, we will show that the transmission conditions (16b), (16c) can formally be interpreted as discretizations of the transmission conditions (9) for the state and the adjoint.

3.2.2 Schur–operator equations

We now return to the solution of the decomposed system of optimality conditions (15), (16). The solutions $\mathbf{y}_I^i, \mathbf{u}_I^i, \mathbf{p}_I^i$ of (15) can be viewed as an affine linear function of the interface variables $\mathbf{y}_\Gamma$, $\mathbf{p}_\Gamma$. If we take this view, (16) is a system of linear equations in $\mathbf{y}_\Gamma$, $\mathbf{p}_\Gamma$. This motivates the following definitions. We define the linear map

$$
\mathbf{S}_i : \left(H^1(0,T;\mathbb{R}^{m_\Gamma^i}) \right)^2 \to \mathbb{R}^{m_\Gamma^i} \times L^2(0,T;\mathbb{R}^{m_\Gamma^i}) \times L^2(0,T;\mathbb{R}^{m_\Gamma^i}) \qquad (18a)
$$

by

$$\begin{aligned}&\mathbf{S}_i(\mathbf{y}_\Gamma,\mathbf{p}_\Gamma)\\&=\begin{pmatrix}\mathbf{M}^i_{\Gamma I}\mathbf{p}^i_I(T)+\mathbf{M}^i_{\Gamma\Gamma}\mathbf{p}^i_\Gamma(T)+\alpha_2\mathbf{M}^i_{\Gamma I}\mathbf{y}^i_I(T)+\alpha_2\mathbf{M}^i_{\Gamma\Gamma}\mathbf{y}^i_\Gamma(T)\\-\mathbf{M}^i_{\Gamma I}\frac{d}{dt}\mathbf{p}^i_I+(\mathbf{A}^i_{I\Gamma})^T\mathbf{p}^i_I-\mathbf{M}^i_{\Gamma\Gamma}\frac{d}{dt}\mathbf{p}^i_\Gamma+(\mathbf{A}^i_{\Gamma\Gamma})^T\mathbf{p}^i_\Gamma+\alpha_1\mathbf{M}^i_{\Gamma I}\mathbf{y}^i_I+\alpha_1\mathbf{M}^i_{\Gamma\Gamma}\mathbf{y}^i_\Gamma\\\mathbf{M}^i_{\Gamma I}\frac{d}{dt}\mathbf{y}^i_I+\mathbf{A}^i_{\Gamma I}\mathbf{y}^i_I+\mathbf{M}^i_{\Gamma\Gamma}\frac{d}{dt}\mathbf{y}^i_\Gamma+\mathbf{A}^i_{\Gamma\Gamma}\mathbf{y}^i_\Gamma+\mathbf{B}^i_{\Gamma I}\mathbf{u}^i_I\end{pmatrix},\end{aligned}\tag{18b}$$

where $\mathbf{y}^i_I,\mathbf{u}^i_I,\mathbf{p}^i_I$ is the solution of (15) (or, equivalently, of (17)) with $(\mathbf{y}_0)^i_I=\mathbf{0}$, $\mathbf{f}^i_I=\mathbf{0}$, $\mathbf{c}^i_I=\mathbf{0}$, $\mathbf{d}^i_I=\mathbf{0}$. Furthermore, let $\widehat{\mathbf{y}}_\Gamma\in H^1(0,1;\mathbb{R}^{m_\Gamma})$ satisfy $\widehat{\mathbf{y}}_\Gamma(0)=(\mathbf{y}(0))_\Gamma$ and let $\widehat{\mathbf{p}}_\Gamma=0$. We define

$$\mathbf{r}^i\in\mathbb{R}^{m^i_\Gamma}\times L^2(0,T;\mathbb{R}^{m^i_\Gamma})\times L^2(0,T;\mathbb{R}^{m^i_\Gamma})\tag{19a}$$

by

$$\mathbf{r}^i=\begin{pmatrix}-\mathbf{d}^i_\Gamma-\alpha_2\mathbf{M}^i_{\Gamma I}\widehat{\mathbf{y}}^i_I(T)\\-\mathbf{c}^i_\Gamma+\mathbf{M}^i_{\Gamma I}\frac{d}{dt}\widehat{\mathbf{p}}^i_I-(\mathbf{A}^i_{I\Gamma})^T\widehat{\mathbf{p}}^i_I-\alpha_1\mathbf{M}^i_{\Gamma I}\widehat{\mathbf{y}}^i_I\\\mathbf{f}^i_\Gamma-\mathbf{M}^i_{\Gamma I}\frac{d}{dt}\widehat{\mathbf{y}}^i_I-\mathbf{A}^i_{\Gamma I}\widehat{\mathbf{y}}^i_I-\mathbf{B}^i_{\Gamma I}\widehat{\mathbf{u}}^i_I\end{pmatrix},\tag{19b}$$

where $\widehat{\mathbf{y}}^i_I,\widehat{\mathbf{u}}^i_I,\widehat{\mathbf{p}}^i_I$ is the solution of (15) (or, equivalently, of (17)) with $\mathbf{y}_\Gamma=\widehat{\mathbf{y}}_\Gamma$ and $\mathbf{p}_\Gamma=\widehat{\mathbf{p}}_\Gamma=0$.

Pointwise application of matrix $\mathbf{I}^i_\Gamma$ defined in (14) induces an operator $(H^1(0,T;\mathbb{R}^{m_\Gamma}))^2\to(H^1(0,T;\mathbb{R}^{m^i_\Gamma}))^2$ and pointwise application of matrix $(\mathbf{I}^i_\Gamma)^T$ induces an operator $\mathbb{R}^{m^i_\Gamma}\times(L^2(0,T;\mathbb{R}^{m^i_\Gamma}))^2\to\mathbb{R}^{m_\Gamma}\times(L^2(0,T;\mathbb{R}^{m_\Gamma}))^2$. These operator will also be denoted by $\mathbf{I}^i_\Gamma$ and $(\mathbf{I}^i_\Gamma)^T$, respectively. The system (15), (16b) can now be written as an operator equation

$$\sum_{i=1}^s(\mathbf{I}^i_\Gamma)^T\mathbf{S}_i\mathbf{I}^i_\Gamma(\mathbf{y}_\Gamma,\mathbf{p}_\Gamma)=\sum_{i=1}^s(\mathbf{I}^i_\Gamma)^T\mathbf{r}^i\tag{20}$$

in the unknowns $\mathbf{y}_\Gamma,\mathbf{p}_\Gamma\in H^1(0,T;\mathbb{R}^{m_\Gamma})$. If the solution $\mathbf{y}_\Gamma,\mathbf{p}_\Gamma$ of (20) is computed, then the remaining components $\mathbf{y}^i_I$, $\mathbf{u}^i_I$, $\mathbf{p}^i_I$ $i=1,\dots,s$, of $\mathbf{y},\mathbf{u},\mathbf{p}$ can be computed by solving (15) (or, equivalently, of (17)).

In the next theorem, we will show how to apply the inverse of the subdomain operator $\mathbf{S}_i$, $i=1,\dots,s$. For this result it is useful to introduce the notation

$$\mathbf{A}^i=\begin{pmatrix}\mathbf{A}^i_{II}&\mathbf{A}^i_{I\Gamma}\\\mathbf{A}^i_{\Gamma I}&\mathbf{A}^i_{\Gamma\Gamma}\end{pmatrix},\quad\mathbf{M}^i=\begin{pmatrix}\mathbf{M}^i_{II}&\mathbf{M}^i_{I\Gamma}\\\mathbf{M}^i_{\Gamma I}&\mathbf{M}^i_{\Gamma\Gamma}\end{pmatrix},\quad\mathbf{B}^i=\begin{pmatrix}\mathbf{B}^i_{II}\\\mathbf{B}^i_{\Gamma I}\end{pmatrix}\tag{21a}$$

and

$$\mathbf{y}^i=\begin{pmatrix}\mathbf{y}^i_I\\\mathbf{y}^i_\Gamma\end{pmatrix},\quad\mathbf{p}^i=\begin{pmatrix}\mathbf{p}^i_I\\\mathbf{p}^i_\Gamma\end{pmatrix}\in\mathbb{R}^{m^i},\tag{21b}$$

where $m^i = m_I^i + m_\Gamma^i$. Furthermore, let

$$I^i \in \mathbb{R}^{m_\Gamma^i \times m^i} \tag{22}$$

be the matrix with zero or one entries that extracts out of a vector $\mathbf{v}^i \in \mathbb{R}^{m^i}$ the subvector $\mathbf{v}_\Gamma^i \in \mathbb{R}^{m_\Gamma^i}$ whose components correspond to vertices $x_k \in \Gamma \cap \partial\Omega_i$.

Theorem 3.2 *Let* $\mathbf{r}^i = (\mathbf{r}_1^i, \mathbf{r}_2^i, \mathbf{r}_3^i) \in \mathbb{R}^{m_\Gamma^i} \times L^2(0,T;\mathbb{R}^{m_\Gamma^i}) \times L^2(0,T;\mathbb{R}^{m_\Gamma^i})$ *be given. The solution* $\mathbf{y}_\Gamma^i, \mathbf{p}_\Gamma^i \in H^1(0,T;\mathbb{R}^{m_\Gamma^i})$ *of*

$$\mathbf{S}_i(\mathbf{y}_\Gamma^i, \mathbf{p}_\Gamma^i) = \mathbf{r}^i$$

is given by

$$\mathbf{y}_\Gamma^i(t) = I^i \mathbf{y}^i(t), \quad \mathbf{p}_\Gamma^i(t) = I^i \mathbf{p}^i(t),$$

where $\mathbf{y}^i, \mathbf{u}^i, \mathbf{p}^i$ *solve*

$$-\mathbf{M}^i \frac{d}{dt}\mathbf{p}^i(t) + (\mathbf{A}^i)^T \mathbf{p}^i(t) + \alpha_1 \mathbf{M}^i \mathbf{y}_I^i(t) = \begin{pmatrix} \mathbf{0} \\ \mathbf{r}_2^i(t) \end{pmatrix}, \qquad t \in (0,T) \tag{23a}$$

$$\mathbf{M}^i \mathbf{p}^i(T) + \alpha_2 \mathbf{M}^i \mathbf{y}^i(T) = \begin{pmatrix} \mathbf{0} \\ \mathbf{r}_1^i \end{pmatrix}, \tag{23b}$$

$$\alpha_3 \mathbf{Q}_{II} \mathbf{u}_I^i(t) + (\mathbf{B}^i)^T \mathbf{p}^i(t) = \mathbf{0}, \qquad t \in (0,T) \tag{23c}$$

$$\mathbf{M}^i \frac{d}{dt}\mathbf{y}^i(t) + \mathbf{A}^i \mathbf{y}^i(t) + \mathbf{B}^i \mathbf{u}^i(t) = \begin{pmatrix} \mathbf{0} \\ \mathbf{r}_3^i(t) \end{pmatrix}, \qquad t \in (0,T) \tag{23d}$$

$$\mathbf{y}^i(0) = \mathbf{0}. \tag{23e}$$

The equations (23) *are the system of necessary and sufficient optimality conditions for the optimal control problem*

$$\begin{aligned} \textit{Minimize} \quad & \int_0^T \frac{\alpha_1}{2} \mathbf{y}^i(t)^T \mathbf{M}^i \mathbf{y}^i(t) - (\mathbf{0}^T, \mathbf{r}_2^i(t)^T)\mathbf{y}^i(t)dt \\ & + \frac{\alpha_2}{2} \mathbf{y}^i(T)^T \mathbf{M}^i \mathbf{y}^i(T) - (\mathbf{0}^T, \mathbf{r}_1^i)^T \mathbf{y}^i(T) + \int_0^T \frac{\alpha_3}{2} \mathbf{u}_I^i(t)^T \mathbf{Q}_{II}^i \mathbf{u}_I^i(t)dt, \end{aligned} \tag{24a}$$

$$\textit{subject to} \quad \mathbf{M}^i \frac{d}{dt}\mathbf{y}^i(t) + \mathbf{A}^i \mathbf{y}^i(t) + \mathbf{B}^i \mathbf{u}_I^i(t) = \begin{pmatrix} \mathbf{0} \\ \mathbf{r}_3^i(t) \end{pmatrix}, \quad t \in (0,T) \tag{24b}$$

$$\mathbf{y}^i(0) = \mathbf{0}. \tag{24c}$$

Proof: By definition (18) of $\mathbf{S}_i$ we see that $\mathbf{y}_\Gamma^i, \mathbf{p}_\Gamma^i$ with $\mathbf{y}_\Gamma^i(0) = \mathbf{0}$, satisfies $\mathbf{S}_i(\mathbf{y}_\Gamma^i, \mathbf{p}_\Gamma^i) = \mathbf{r}^i$

if and only if

$$
\begin{aligned}
&-\mathbf{M}_{II}^i \frac{d}{dt}\mathbf{p}_I^i(t) + (\mathbf{A}_{II}^i)^T \mathbf{p}_I^i(t) + \alpha_1 \mathbf{M}_{II}^i \mathbf{y}_I^i(t) \\
&-\mathbf{M}_{I\Gamma}^i \frac{d}{dt}\mathbf{p}_\Gamma^i(t) + (\mathbf{A}_{\Gamma I}^i)^T \mathbf{p}_\Gamma^i(t) + \alpha_1 \mathbf{M}_{I\Gamma}^i \mathbf{y}_\Gamma^i(t) = \mathbf{0}, \qquad t \in (0,T) \qquad (25a)\\
&\mathbf{M}_{I\Gamma}^i \mathbf{p}_\Gamma^i + \mathbf{M}_{II}^i \mathbf{p}_I^i(T) + \alpha_2 \mathbf{M}_{II}^i \mathbf{y}_I^i(T) + \alpha_2 \mathbf{M}_{I\Gamma}^i \mathbf{y}_\Gamma^i(T) = \mathbf{0}, \qquad (25b)\\
&\alpha_3 \mathbf{Q}_{II} \mathbf{u}_I^i(t) + (\mathbf{B}_{II}^i)^T \mathbf{p}_I^i(t) + (\mathbf{B}_{\Gamma I}^i)^T \mathbf{p}_\Gamma^i(t) = \mathbf{0}, \qquad t \in (0,T) \qquad (25c)\\
&\mathbf{M}_{II}^i \frac{d}{dt}\mathbf{y}_I^i(t) + \mathbf{A}_{II}^i \mathbf{y}_I^i(t) \\
&+\mathbf{M}_{I\Gamma}^i \frac{d}{dt}\mathbf{y}_\Gamma^i(t) + \mathbf{A}_{I\Gamma}^i \mathbf{y}_\Gamma^i(t) + \mathbf{B}_{II}^i \mathbf{u}_I^i(t) = \mathbf{0}, \qquad t \in (0,T) \qquad (25d)\\
&\mathbf{y}_I^i(0) = \mathbf{0}, \qquad (25e)
\end{aligned}
$$

$$
\begin{aligned}
&-\mathbf{M}_{\Gamma I}^i \frac{d}{dt}\mathbf{p}_I^i(t) + \mathbf{A}_{\Gamma I}^i \mathbf{p}_I^i(t) - \mathbf{M}_{\Gamma\Gamma}^i \frac{d}{dt}\mathbf{p}_\Gamma^i(t) + (\mathbf{A}_{\Gamma\Gamma}^i)^T \mathbf{p}_\Gamma^i(t) \\
&+\alpha_1 \mathbf{M}_{\Gamma I}^i \mathbf{y}_I^i(t) + \alpha_1 \mathbf{M}_{\Gamma\Gamma}^i \mathbf{y}_\Gamma^i(t) = \mathbf{r}_2^i(t), \qquad t \in (0,T) \qquad (25f)\\
&\mathbf{M}_{\Gamma I}^i \frac{d}{dt}\mathbf{y}_I^i(t) + \mathbf{A}_{\Gamma I}^i \mathbf{y}_I^i(t) \\
&+\mathbf{M}_{\Gamma\Gamma}^i \frac{d}{dt}\mathbf{y}_\Gamma^i(t) + \mathbf{A}_{\Gamma\Gamma}^i \mathbf{y}_\Gamma^i(t) + \mathbf{B}_{\Gamma I}^i \mathbf{u}_I^i(t) = \mathbf{r}_3^i(t), \qquad t \in (0,T) \qquad (25g)\\
&\mathbf{y}_\Gamma^i(0) = \mathbf{0}, \qquad (25h)\\
&\mathbf{M}_{\Gamma I}^i \mathbf{p}_I^i(T) + \mathbf{M}_{\Gamma\Gamma}^i \mathbf{p}_\Gamma^i(T) + \alpha_2 \mathbf{M}_{\Gamma\Gamma}^i \mathbf{y}_\Gamma^i(T) + \alpha_2 \mathbf{M}_{\Gamma I}^i \mathbf{y}_I^i(T) = \mathbf{r}_1^i. \qquad (25i)
\end{aligned}
$$

The equations (25) can be written in the more compact notation (23).

The interpretation of (23) as the necessary and sufficient optimality conditions for (24) can be proven using standard techniques and we omit the proof. □

In Appendix A.2 we interpret the system (23) as the semi–discretization of a system of partial differential equations.

3.3 Solution algorithm

In the previous section we have shown that the semi–discrete optimal control problem (7) is equivalent to the linear operator equation (20), i.e.,

$$
\sum_{i=1}^{s} (\mathbf{I}_\Gamma^i)^T \mathbf{S}_i \mathbf{I}_\Gamma^i (\mathbf{y}_\Gamma, \mathbf{p}_\Gamma) = \sum_{i=1}^{s} (\mathbf{I}_\Gamma^i)^T \mathbf{r}_i \qquad (26)
$$

After suitable discretization in time we obtain a symmetric operator $\mathbf{S}_i^{\Delta t}$ as shown in [29]. Hence, we solve (26) using preconditioned sQMR method [30]. The inverse of the system

operator $\sum_{i=1}^{s}(\mathbf{I}_\Gamma^i)^T\mathbf{S}_i\mathbf{I}_\Gamma^i$ is approximated by a weighted sum of inverses of the subdomain operators $\mathbf{S}_i$. This choice is motivated by Neumann-Neumann domain decomposition preconditioners that have been used sucessfully for the solution of elliptic PDEs (see [17–19] and the references given therein) as well as elliptic linear-quadratic optimal control problems [21,22]. We let $D_\Gamma^i \in \mathbb{R}^{m_\Gamma^i \times m_\Gamma^i}$ be positive definite diagonal matrices such that

$$\sum_{i=1}^{s}(\mathbf{I}_\Gamma^i)^T D_\Gamma^i = \mathbf{I}.$$

In our case the entry D_{kk}^i is equal to one over the number of subdomains containing the interface node x_k. We set

$$\mathbf{D}_\Gamma^i = \begin{pmatrix} D_\Gamma^i & & \\ & \ddots & \\ & & D_\Gamma^i \end{pmatrix}. \tag{27}$$

Pointwise application of the matrix $\mathbf{D}_\Gamma^i$ induces an operator $H^1(0,T;\mathbb{R}^{m_\Gamma}))^2\ [L^2(0,T;\mathbb{R}^{m_\Gamma}))^2] \to H^1(0,T;\mathbb{R}^{m_\Gamma}))^2\ [L^2(0,T;\mathbb{R}^{m_\Gamma}))^2]$. This operator will also be denoted by $\mathbf{D}_\Gamma^i$.

The (symmetric) Neumann-Neumann (NN) preconditioner for $\sum_{i=1}^{s}(\mathbf{I}_\Gamma^i)^T\mathbf{S}_i\mathbf{I}_\Gamma^i$ is now given by

$$\mathbf{P}_{NN} = \sum_{i=1}^{s}(\mathbf{I}_\Gamma^i)^T\mathbf{D}_\Gamma^i\mathbf{S}_i^{-1}\mathbf{D}_\Gamma^i\mathbf{I}_\Gamma^i. \tag{28}$$

It is well-known that the NN preconditioner $\mathbf{P}_{NN}$ detoriates for large numbers of subdomains. This behaviour is observed for the NN preconditioner for elliptic equations (see, e.g., [17–19]) as well as for elliptic linear-quadratic control problems ([21,22]). The numerical tests reported on in Section 5 show the same qualitative behavior of preconditioner (28) for parabolic problems. To remedy this, we introduce a coarse space Schur complement $\mathbf{S}_0$ and use the symmetric preconditioner

$$\mathbf{P}_{NNc} = (\mathbf{I} - \mathbf{I}_0^T\mathbf{S}_0^{-1}\mathbf{I}_0\mathbf{S})\left(\sum_{i=1}^{s}(\mathbf{I}_\Gamma^i)^T\mathbf{D}_\Gamma^i\mathbf{S}_i^{-1}\mathbf{D}_\Gamma^i\mathbf{I}_\Gamma^i\right)(\mathbf{I} - \mathbf{S}\mathbf{I}_0^T\mathbf{S}_0^{-1}\mathbf{I}_0) + \mathbf{I}_0^T\mathbf{S}_0^{-1}\mathbf{I}_0 \tag{29}$$

(cf. [18, Sec. 4.3.3], [19, Sec. 6.2.2], [21,22]). Here, $\mathbf{S}$ is the fine Schur operator defined by

$$\mathbf{S}(\mathbf{y}_\Gamma,\mathbf{p}_\Gamma) = \sum_{i=1}^{s}(\mathbf{I}_\Gamma^i)^T\mathbf{S}_i\mathbf{I}_\Gamma^i(\mathbf{y}_\Gamma,\mathbf{p}_\Gamma), \tag{30}$$

$\mathbf{I}$ is the identity, $\mathbf{S}_0$ is the coarse grid Schur operator and $\mathbf{I}_0$ is a restriction operator. To define the coarse grid Schur operator $\mathbf{S}_0$ we consider problem (2). We introduce a coarse semi–discretization as in Section 3.1 with $m = 2s$ basis functions where s is the number of subdomains. The coarse Schur operator has the same structure as the Schur operator (18) and (20). However, the evaluation of the subdomain coarse Schur operators $\mathbf{S}_{0,i}$ now requires the solution of the optimal control problem (17) where the finite dimensional subspaces for $\mathbf{y}_I^i, \mathbf{p}_I^i$ are of dimension one. The restriction operator is defined as

map $\mathbf{I}_0 : H^1(0,T;\mathbb{R}^{m_\Gamma})^2 \to H^1(0,T;\mathbb{R}^s)^2$. It returns for each subdomain two functions: at each time t the weighted sum of the states and the weighted sum of the adjoints on the common nodes. The weights are choosen as reciprocal number of subdomains sharing the node. In the case $\Omega_i \in \mathbb{R}^1$ the restriction operator can also be choosen as the identity map $\mathbf{I}_0 = \mathbf{Id} : H^1(0,T;\mathbb{R}^{m_\Gamma})^2 \to H^1(0,T;\mathbb{R}^{m_\Gamma})^2$.

To apply $\mathbf{S}_0^{-1}$ we solve $\mathbf{S}_0^{-1}\mathbf{v}_{0,\Gamma}$ using sQMR. Since $\mathbf{S}_0$ has the same structure as $\mathbf{S}$ we introduce a Neumann-Neumann preconditioner for $\mathbf{S}_0^{-1}$ given by

$$\mathbf{P}_{NN,0} = \sum_{i=1}^{s} (\mathbf{I}_\Gamma^i)^T \mathbf{D}_\Gamma^i \mathbf{S}_{0,i}^{-1} \mathbf{D}_\Gamma^i \mathbf{I}_\Gamma^i \tag{31}$$

Here $\mathbf{S}_{0,i}^{-1}$ are subdomain Schur complement operators on the coarse grid. In summary, the preconditioned symmetric balanced Neumann-Neumann preconditioner $\mathbf{P}_{NNc,0}$ for (26) is given by (29) combined with preconditioned evaluation of the coarse grid schur operator using the preconditioner $\mathbf{P}_{NN,0}$, (31).

We discuss the evaluation of a subdomain Schur operators and their inverses. In Theorem 3.2 we have shown that the evaluation of matrix-vector products of the form $\mathbf{S}_i^{-1}\mathbf{v}_\Gamma^i$ is equivalent to solving an optimal control subdomain problem (24). Further, the evaluation of matrix-vector products of the form $\mathbf{S}_i\mathbf{v}_\Gamma^i$ requires the solution of an optimal control subdomain problem given by (17) with homogenous data, i.e., $(\mathbf{y}_0)_I^i = \mathbf{f}_I^i(t) = \mathbf{d}_I^i(t) = \mathbf{c}_I^i(t) = \mathbf{0}$ for all t. The same is true for the coarse grid operators $\mathbf{S}_{0,i}$ and $\mathbf{S}_{0,i}^{-1}$. The subdomain optimal control problems are solved using the conjugate gradient method applied to the reduced formulation of the respective subdomain optimal control problem. The reduced form of the optimal control problems (24) or (17) is the one in which the state is viewed as function of the control and the optimal control problem is posed as a minimization problem in the controls only.

We now introduce the full discretization and a last modification of our preconditioner. To solve (26) we discretize using the backward Euler method on an equidistant time grid

$$t_k = k\Delta t, \quad k = 0,\ldots,K$$

with time step size $\Delta t = T/K$. We use the same notation for the operators as before. Especially, $\mathbf{S}$ is now an operator defined on $\mathbb{R}^{2(m_\Gamma \times K)}$ and the coarse schur operator $\mathbf{S}_0$ is defined on $\mathbb{R}^{2(s\times K)}$. Details can be found in [29].

In addition to the coarse grid Schur operator $\mathbf{S}_0$ discussed previously, we also experiment with a coarse grid Schur operator that also involves a coarse time grid. We define a second time equidistant time grid $\overline{t_k} = k\overline{\Delta t}$ for $k = 0,\ldots,K_c$ with $K_c \ll K$. The coarse time grid Schur operator $\mathbf{S}_0^t$ is defined by equations (18) and (20) for $m = 2s$ base functions in space and on the time grid $0,\ldots,K_c$. Hence $\mathbf{S}_0^t : \mathbb{R}^{2(s\times K_c)} \to \mathbb{R}^{2(s\times K_c)}$ and $\mathbf{S}_0^t \equiv \mathbf{S}_0$ iff $K_c = K$. To derive the preconditioner we need to introduce $\mathbf{I}_0^t$ as restriction operator in space and time. To be more precise: $\mathbf{I}_0^t : \mathbb{R}^{2(m_\Gamma\times K+1)} \to \mathbb{R}^{2(s\times K_c)}$. For $\Omega \in \mathbb{R}^1$, $\mathbf{I}_0^t$ returns for each $\overline{t_k}$ two values: the average of all state values and the average of all adjoint values with

$\bar{t}_k \leq t_j < \bar{t}_{k+1}$. The final preconditioner $\mathbf{P}^t_{NN}$ for solving (26) is then given by (29) wherein $\mathbf{S}_0$ and $\mathbf{I}_0$ are replaced by $\mathbf{S}^t_0$ and $\mathbf{I}^t_0$ respectively. Analogously we derive $\mathbf{P}^t_{NNc,0}$.

We specify the tolerances for the numerical results. The conjugate gradient method applied to the subdomain optimal control problems corresponding to the local Schur complements and their inverses is stopped when the norm of the gradient is less than 10^{-10}. The preconditioned sQMR computing $\mathbf{S}_0^{-1}\mathbf{v}_\Gamma$ is stopped when the preconditioned residual is less than 10^{-4}. The preconditioned sQMR applied to the linear operator equation (26) is stopped if the preconditioned residual is less than 10^{-4}.

The following Table 1 compares the costs per preconditioned sQMR iteration solving (26) in terms of evaluations of coarse and fine grid schur operators for a domain decomposition approach with s subdomains. Since the coarse grid Schur operator acts on a much smaller domain, the evaluation of $\mathbf{S}_{0,i}$ and $\mathbf{S}_{0,i}^{-1}$ is cheaper than those of $\mathbf{S}_i$ and $\mathbf{S}_i^{-1}$.

Preconditioner \ Number of evaluations of	$\mathbf{S}_i$	$\mathbf{S}_i^{-1}$	$\mathbf{S}_0^{-1}$	$\mathbf{S}_{0,i}$	$\mathbf{S}_{0,i}^{-1}$
Neumann-Neumann (P_{NN})	s	s	0	0	0
NN with coarse space (P_{NNc})	$2s$	s	2	0	0
NN with iterative solution of coarse space Schur complement eqn. $(P_{NNc,0})$	$2s$	s	0	$2s\ell$	$2s\ell$

Table 1
Number of evaluations per each computation of the preconditioner in a preconditioned sQMR iteration solving (26) with s subdomains. Here ℓ is the number of preconditioned sQMR iterations needed for the solution of $\mathbf{S}_0^{-1}\mathbf{v}_\Gamma$.

4 An iterative method based on skew symmetric Robin transmission conditions

In [14] Benamou proposes a spatial domain decomposition methods for parabolic optimal control problems. His approach [14] is based on [31] and is sketched here for the model problem (1) with $\mu = 1$ and $\mathbf{a} = \mathbf{0}$, $c = 0$. If advection is present, the interface conditions, equations (32c) and (32g) below, have to be modified using, e.g., the ideas in [25]. We use the decomposition of the spatial domain Ω introduced in Section 3.

The domain decomposition method proceeds as follows. Let y_i^k, u_i^k, p_i^k be approximations of states, controls, and adjoints in subdomain i computed in iteration k. The new approxi-

mations in iteration $k+1$ are computed as the solution of

$$\begin{aligned}
\partial_t y_i(x,t) - \Delta y_i(x,t) &= f(x,t) + u_i(x,t) && \text{in } \Omega_i \times (0,T), && (32a)\\
y_i(x,t) &= 0 && \text{on } (\partial\Omega_i \cap \partial\Omega) \times (0,T), && (32b)\\
\frac{\partial}{\partial n_i} y_i(x,t) + \beta p_i(x,t) &= \frac{\partial}{\partial n_i} y_j^k(x,t) + \beta p_j^k(x,t) && \text{on } (\partial\Omega_i \cap \partial\Omega_j) \times (0,T), && (32c)\\
y_i(x,0) &= y_0(x) && \text{in } \Omega_i, && (32d)\\
-\partial_t p_i(x,t) - \Delta p_i(x,t) &= -\alpha_1 (y_i(x,t) - \widehat{y}(x,t)) && \text{in } \Omega_i \times (0,T), && (32e)\\
p_i(x,t) &= 0 && \text{on } (\partial\Omega_i \cap \partial\Omega) \times (0,T), && (32f)\\
\frac{\partial}{\partial n_i} p_i(x,t) - \beta y_i(x,t) &= \frac{\partial}{\partial n_i} p_j^k(x,t) - \beta y_j^k(x,t) && \text{on } (\partial\Omega_i \cap \partial\Omega_j) \times (0,T), && (32g)\\
p_i(x,T) &= -\alpha_2 (y_i(x,T) - \widehat{y}_T(x)) && \text{in } \Omega_i, && (32h)\\
p_i(x,t) + \alpha_3 u_i(x,t) &= 0 && \text{on } \Omega_i \times (0,T). && (32i)
\end{aligned}$$

In (32), $\beta > 0$ is a given parameter. In [14] the choice $\beta = O(1/h)$ is recommended, where h describes the spatial mesh size. Note that the subdomains Ω_i and Ω_j are coupled through the skew symmetric Robin transmission conditions (32c), (32g).

We now discuss a few implementation details not provided in [14]. For the implementation of the method it will be convenient to introduce

$$z_{ji}^k = \frac{\partial}{\partial n_i} y_j^k(x,t) + \beta p_j^k(x,t), \quad q_{ji}^k = \frac{\partial}{\partial n_i} p_j^k(x,t) - \beta y_j^k(x,t). \tag{33}$$

The introduction of these variables is motivated by [25]. The system (32) is now written as

$$\begin{aligned}
\partial_t y_i(x,t) - \Delta y_i(x,t) &= f(x,t) + u_i(x,t) && \text{in } \Omega_i \times (0,T), && (34a)\\
y_i(x,t) &= 0 && \text{on } (\partial\Omega_i \cap \partial\Omega) \times (0,T), && (34b)\\
\frac{\partial}{\partial n_i} y_i(x,t) + \beta p_i(x,t) &= z_{ji}^k(x,t) && \text{on } (\partial\Omega_i \cap \partial\Omega_j) \times (0,T), && (34c)\\
y_i(x,0) &= y_0(x) && \text{in } \Omega_i, && (34d)\\
-\partial_t p_i(x,t) - \Delta p_i(x,t) &= -\alpha_1 (y_i(x,t) - \widehat{y}(x,t)) && \text{in } \Omega_i \times (0,T), && (34e)\\
p_i(x,t) &= 0 && \text{on } (\partial\Omega_i \cap \partial\Omega) \times (0,T), && (34f)\\
\frac{\partial}{\partial n_i} p_i(x,t) - \beta y_i(x,t) &= q_{ji}^k(x,t) && \text{on } (\partial\Omega_i \cap \partial\Omega_j) \times (0,T), && (34g)\\
p_i(x,T) &= -\alpha_2 (y_i(x,T) - \widehat{y}_T(x)) && \text{in } \Omega_i, && (34h)\\
p_i(x,t) + \alpha_3 u_i(x,t) &= 0 && \text{on } \Omega_i \times (0,T). && (34i)
\end{aligned}$$

If $y_i^{k+1}, u_i^{k+1}, p_i^{k+1}$ denotes the solution of (34), then (33), (34c), and (34g) imply

$$z_{ij}^{k+1} = 2\beta p_j^{k+1}(x,t) - z_{ji}^k, \quad q_{ij}^{k+1} = -2\beta y_j^{k+1}(x,t) - q_{ji}^k. \tag{35}$$

A semi–discretization of the system (34) using finite elements is straight forward. Given an

intial guess for y_i^0, p_i^0, $i = 1,\dots,s$, i.e, for z_{ij}^0, q_{ij}^0, $i,j = 1,\dots,s$, updates $z_{ij}^{k+1}, q_{ij}^{k+1}$ can be computed using (35).

For computational purposes, it is important to note that the system (34) can be interpreted as the optimality conditions for the following optimal control problem in the variables y_i, u_i, w_i.

$$\begin{aligned}\text{Minimize}\quad & \frac{\alpha_1}{2}\int_0^T\int_{\Omega_i}(y_i(x,t)-\hat{y}(x,t))^2dxdt+\frac{\alpha_2}{2}\int_{\Omega_i}(y_i(x,T)-\hat{y}_T(x))^2dx\\ & +\frac{\alpha_3}{2}\int_0^T\int_{\Omega_i}u_i^2(x,t)dxdt+\sum_j\int_0^T\int_{\partial\Omega_i}\Big(\frac{\beta}{2}y_i^2(x,t)+q_{ji}^k(x,t)y_i(x,t)\Big)dxdt,\\ & +\sum_j\int_0^T\int_{\partial\Omega_i}\frac{1}{2\beta}w_i^2(x,t)dxdt, \end{aligned}\tag{36a}$$

subject to

$$\partial_t y_i(x,t)-\Delta y_i(x,t)=\widehat{f}(x,t)+u_i(x,t) \qquad \text{in } \Omega_i\times(0,T), \tag{36b}$$

$$y_i(x,t)=0 \qquad \text{on } (\partial\Omega_i\cap\partial\Omega)\times(0,T), \tag{36c}$$

$$\frac{\partial}{\partial n_i}y_i(x,t)=w_i(x,t)+z_{ji}^k(x,t) \qquad \text{on } (\partial\Omega_i\cap\partial\Omega_j)\times(0,T), \tag{36d}$$

$$y_i(x,0)=y_0(x) \qquad \text{in } \Omega_i. \tag{36e}$$

In fact, the Lagrangian associated with the optimal control problem (36) is given by

$$\begin{aligned} & L(y_i,u_i,w_i,p_i)\\ & =\frac{\alpha_1}{2}\int_0^T\int_{\Omega_i}(y_i(x,t)-\hat{y}(x,t))^2dxdt+\frac{\alpha_2}{2}\int_{\Omega_i}(y_i(x,T)-\hat{y}_T(x))^2dx+\frac{\alpha_3}{2}\int_0^T\int_{\Omega_i}u_i^2(x,t)dxdt\\ & \quad+\sum_{j\neq i}\int_0^T\int_{\partial\Omega_j\cap\partial\Omega_i}\Big(\frac{\beta}{2}y_i^2(x,t)+q_{ji}^k(x,t)y_i(x,t)\Big)dxdt\\ & \quad+\sum_{j\neq i}\int_0^T\int_{\partial\Omega_j\cap\partial\Omega_i}\frac{1}{2\beta}w_i^2(x,t)dxdt\\ & \quad+\int_0^T\int_{\Omega_i}\partial_t y_i(x,t)p_i(x,t)+\nabla y_i(x,t)\nabla p_i(x,t)dxdt\\ & \quad-\int_0^T\int_{\Omega_i}\widehat{f}(x,t)p_i(x,t)+u_i(x,t)p_i(x,t)dxdt\\ & \quad+\int_0^T\sum_{j\neq i}\int_{\partial\Omega_j\cap\partial\Omega_i}\Big(w_i(x,t)+z_{ji}^k(x,t)\Big)p(x,t)dxdt \end{aligned}\tag{37}$$

Setting the derivative of the Lagrangian with respect to p_i to zero gives the state equation (34a)–(34d) with βp_i in (34c) replaced by $-w_i$. Setting the derivative of the Lagrangian with respect to y_i to zero gives the adjoint equation (34e)–(34h). Setting the derivative of the

Lagrangian with respect to u_i to zero gives (34i), and setting derivative of the Lagrangian with respect to w_i to zero gives

$$-w_i(x,t) = \beta p_i(x,t) \quad \text{ on } (\partial\Omega_i \cap \partial\Omega_j) \times (0,T).$$

In our implementation of the iteration (35), we solve (36a). More precisely, we view the solution of (36b)–(36e) as a function of u_i, w_i and use the conjugate gradient method to minimize the convex quadratic objective function (36a), which is viewed as a function in u_i, w_i only.

5 Numerical results

We consider (1) with $\Omega = (0,1)$, $\mu = 1$, $c = 0$, $f = 0$ and $y_0(x) = \sin(2\pi x)$. The desired states $\hat{y}$ and $\hat{y}_T$ are given by the hat functions $\hat{y}(x,t) = \min\{2x, 2(1-x)\}$ and $\hat{y}_T(x) = \min\{2x, 2(1-x)\}$, respectively. The advection $\mathbf{a}$ and the weighting parameters $\alpha_1, \alpha_2, \alpha_3$ will be specified later.

We apply the domain decomposition method to obtain the full discretization of the control problem (7). For the spatial discretization of the problem we use piecewise linear finite elements on an equidistant grid with mesh size $\Delta x = 1/K$ and for the time discretization, we use the backward Euler method with step size $\Delta t = 1/K$. The domain $\Omega = (0,1)$ is subdivided into equidistant subdomains $\Omega_i = ((i-1)H, iH)$, $i = 1, \ldots, s$, $H = 1/s$. Further details can be found in [29]. The full discretization yields symmetric Schur complement operators and the adjoint equation (7) is not equal to the Euler discretization of the adjoint equation (8a).

We note that the case $\Omega \subset \mathbb{R}$ and the backward Euler method are considered for simplicity. The domain decomposition approaches discussed in this paper can handle problems with spatial dimension greater than one and can be applied with other time discretizations.

5.1 Optimal control without advection

In the first example we use $\mathbf{a} = 0$ and $c = 0$. Tables 2 to 5 report on numerical results.

Table 2 shows that the number of preconditioned sQMR iterations is insensitive to the weighting parameters α_1, α_2. The conditioning of the optimal control problem (1) grows as α_1/α_3 and as α_2/α_3. For larger α_1/α_3 and as α_2/α_3 the problem (1) becomes more difficult to solve numerically. The insensitivity of the number of preconditioned sQMR iterations again matches the observations made in [21,22] for Neumann-Neumann methods applied to linear-quadratic elliptic optimal control problems.

Neumann-Neumann ($\mathbf{P}_{NN}$) $\alpha_1 = \alpha_2 = 10$, $\alpha_3 = 1$				Neumann-Neumann ($\mathbf{P}_{NN}$) $\alpha_1 = \alpha_2 = 10^3$, $\alpha_3 = 1$			
s \ K	64	128	256	s \ K	64	128	256
2	1	1	1	2	1	1	1
4	15	16	16	4	8	9	8
8	68	69	75	8	32	39	32
16	>500	445	>500	16	133	140	121

Table 2
Number of preconditioned sQMR iterations needed for the solution of (26) depending on the number of subdomains s and on the discretization size $\Delta x = \Delta t = 1/K$.

Table 3 shows the detoriation of the Neumann-Neumann preconditioner $\mathbf{P}_{NN}$ for large numbers of discretization points and subdomains. The preconditioned balanced Neumann-Neumann preconditioner $\mathbf{P}_{NNc,0}$ does not detoriate as seen in the right part of the table. This matches the observed behaviour in the case of elliptic linear-quadratic optimal control problems.

Neumann-Neumann ($\mathbf{P}_{NN}$)			Neumann-Neumann with coarse grid ($\mathbf{P}_{NNc,0}$)		
s \ K	256	512	s \ K	256	512
2	1	1	2	10	35
4	8	8	4	8	9
8	32	33	8	6	8
16	121	131	16	5	5

Table 3
Number of preconditioned sQMR iterations needed for the solution of (26) depending on the number of subdomains s and on the discretization size $\Delta x = \Delta t = 1/K$. In all computations $\alpha_1 = \alpha_2 = 10^3$, $\alpha_3 = 1$.

Table 4 provides more detailed information about the cost of the iterative solver. Reported are triples $(a;b;c)$, where a denotes the number of (outer) sQMR iterations necessary to solve (26). If a coarse space is used, then the application of $\mathbf{S}_0^{-1}$ to a vector is carried out by applying an (inner) sQMR iteration. In this case the second number b reports the average number of inner sQMR iterations needed to apply $\mathbf{S}_0^{-1}$ to a vector (i.e., b corresponds to ℓ in Table 1). The third number c is the average number cg iterations used to solve the subdomain optimal control problems associated with the application of $\mathbf{S}_i$, $\mathbf{S}_i^{-1}$, $\mathbf{S}_{0,i}$, or $\mathbf{S}_{0,i}^{-1}$. Note that the subdomain optimal control problems associated with the application of $\mathbf{S}_{0,i}$ and $\mathbf{S}_{0,i}^{-1}$ are significantly smaller than those associated with the application of $\mathbf{S}_i$ and $\mathbf{S}_i^{-1}$. Compared to the $\mathbf{P}_{NN}$ preconditioner the number of solved cg problems in $\mathbf{P}_{NNc,0}$ increased by a factor of 10. On the other hand this increase is compensated by the smaller

number of total sQMR iterations and by the number of cg iterations necessary to compute the optimal control on the subdomains. We do not report on the performance of $\mathbf{P}_{NN,c}$. Recall that $\mathbf{P}_{NN,c}$ applies $\mathbf{S}_0^{-1}$ using the sQMR method, whereas $\mathbf{P}_{NNc,0}$ applies $\mathbf{S}_0^{-1}$ using the sQMR method wwith a Neumann-Neumann type preconditioner analogous to (28). We note that, as expected, $\mathbf{P}_{NNc,0}$ requires fewer iterations for applying the preconditioner and, consequently, requires a smaller number (in our computations by a factor of greater than two) of coarse grid subproblems to solve by the conjugate gradient method.

s	K	$\mathbf{P}_{NN}$	$\mathbf{P}_{NNc,0}$
4	32	(7; – ; 22)	(6;3;15)
4	64	(9; – ; 22)	(7;3;16)
8	32	(30; – ; 20)	(4;10;15)
8	64	(68; – ; 12)	(5;10;16)
8	256	(32; – ; 24)	(6;11; 17)
16	128	(140; – ; 22)	(4;44;16)
16	256	(121; – ; 25)	(5;39;16)
16	512	(131; – ; 26)	(5;43;17)

Table 4
Number of iterations for the subproblem solutions in the iterative solver for (26) for varying number of subdomains s and discretization sizes $\Delta x = \Delta t = 1/K$. $(a;b;c)$ is a triple where a is the total number of sQMR iterations needed to solve (26), b is the average numer of sQMR iterations needed by the preconditioner to solve the coarse grid Schur operator equation (i.e., b corresponds to ℓ in Table 1) and c is the average number of cg iterations used to solve the subdomain optimal control problems associated with the application of $\mathbf{S}_i$, $\mathbf{S}_i^{-1}$, $\mathbf{S}_{0,i}$, or $\mathbf{S}_{0,i}^{-1}$. In all computations $\alpha_1 = \alpha_2 = 10^3$, $\alpha_3 = 1$.

Table 5 reports on the effect of coarse time grids by comparing $\mathbf{P}_{NNc,0}$ and $\mathbf{P}^t_{NNc,0}$. The coarse grid Schur complement operator $\mathbf{S}_0$ (used in $\mathbf{P}_{NNc,0}$) is of size $2\,s\,K$, whereas the coarse grid Schur complement operator $\mathbf{S}_0^t$ (used in $\mathbf{P}^t_{NNc,0}$) is of size $2\,s\,K_c$. We observe that when using $\mathbf{P}^t_{NNc,0}$ the number of sQMR iterations slightly increase while the average number of cg iterations per subproblem decreases. This relates to the fact that the subdomain optimization problems for the coarse time grid are of reduced dimension. Another advantage of the coarse time grid approximations is the reduced storage requirement which also improves the computation time.

5.2 Optimal control of an advection–diffusion equation

In the second example we choose a positive advection **a**. For nonzero advection, the Neumann-Neumann preconditioner becomes a Robin-Robin type preconditioner. See Ap-

s	K	K/K_c	$\mathbf{P}_{NNc,0}$	$\mathbf{P}^t_{NNc,0}$
4	16	2	(5;3;13)	(7;3;11)
4	32	2	(6;4;15)	(7;3;14)
4	32	4	(6;4;15)	(8;4;11)
4	64	2	(7;3;16)	(7;3;15)
4	64	4	(7;3;16)	(8;3;14)

Table 5
Number of iterations for the subproblem solutions in the iterative solver for (26) for varying number of subdomains s and discretization sizes $\Delta x = \Delta t = 1/K$. $(a;b;c)$ is a triple where a is the total number of sQMR iterations needed to solve (26), b is the average numer of sQMR iterations needed by the preconditioner to solve the coarse grid Schur operator equation (i.e., b corresponds to ℓ in Table 1) and c is the average number of cg iterations used to solve the subdomain optimal control problems associated with the application of $\mathbf{S}_i$, $\mathbf{S}_i^{-1}$, $\mathbf{S}_{0,i}$, $\mathbf{S}_{0,i}^{-1}$, $\mathbf{S}_{0,i}^t$, or $(\mathbf{S}_{0,i}^t)^{-1}$. In all computations $\alpha_1 = \alpha_2 = 10^3$, $\alpha_3 = 1$.

pendix A. Table 6 shows a strong detoration of the Robin-Robin preconditioner for increasing advection term. In contrast, the iteration numbers for the balanced Robin-Robin preconditioner remain nearly constant.

	$\mathbf{a} = 1$		$\mathbf{a} = 5$		$\mathbf{a} = 10$	
s	$\mathbf{P}_{NN}$	$\mathbf{P}_{NNc,0}$	$\mathbf{P}_{NN}$	$\mathbf{P}_{NNc,0}$	$\mathbf{P}_{NN}$	$\mathbf{P}_{NNc,0}$
4	9	7	13	7	20	7
8	40	5	42	7	57	7

Table 6
Number of preconditioned sQMR iterations needed for the solution of (26) for s subdomains and a discretization size of $\Delta t = \Delta x = 1/K$ with Robin-Robin and balanced Robin-Robin preconditioner. In all computations $\alpha_1 = \alpha_2 = 10^3$, $\alpha_3 = 1$, $\mu = 1$, $K = 64$.

5.3 Comparison with the iteration in Section 4

We now compare the performance of the solution algorithm in Section 3.3 with the performance of the iterative method based on skew symmetric Robin transmission conditions due to Benamou [14], which was sketched in Section 4. For this comparison we set $c = \mathbf{a} = 0$. The subdomain optimal control problems (36) are solved using the conjugate gradient method. The conjugate gradient method is stopped when the norm of the gradient is less than 10^{-10}. The outer iteration (35) is stopped when res $< 10^{-2}$, where

$$\text{res} = \sum_{i=1}^{s} \|y_i - y_j\|_{L^2(\partial\Omega_i \cap \partial\Omega_j)} + \|p_i - p_j\|_{L^2(\partial\Omega_i \cap \partial\Omega_j)}. \tag{38}$$

In Table 7 we report the number of outer iterations (35) as well as the average number of conjugate gradient iterations for the solution of the subdomain optimal control problems (36).

α_1	# iter	# cg
10	164	90
10^2	191	105
10^3	263	151

Table 7
Number of iterations (35) (# iter) and average number of conjugate gradient iterations (#cg) per outer iteration and per subdomain for $\alpha_1 = \alpha_2$, $\alpha_3 = 1$, $s = 4$ subdomains, $K = 64$ and a discretization $\Delta x = \Delta t = 1/K$.

It is not possible to give a precise comparison between the algorithms presented in Sections 3.3 and 4, because of the different stopping criteria used for both. If we assume that for the given stopping criteria both algorithms produce solutions of the same quality, then we may use the number of conjugate gradient steps needed for the solution of the subdomain optimal control problems (A.1), (A.5) in case of the iteration in Sections 3.3 and (36) in case of the iteration in Section 4, respectively, as a rough measure to compare the expense of the two iterations. For the case $\alpha_1 = \alpha_2 = 10^3$, $\alpha_3 = 1$, $s = 4$ subdomains, $K = 64$, the algorithm in Section 4 requires $4 * 263 * 151 = 4 * 39713$ conjugate gradient iterations (cf. last row in Table 7). The algorithm in Section 3.3 with preconditioner $P_{NNc,0}$ requires approximately $4 * 7 * 3 * 16 = 4 * 336$ conjugate gradient iterations (cf. the last row in Table 1 and the second row in Table 4) for the fine grid subproblems and an additional $4 * 7 * 4 * 3 * 16 = 4 * 1344$ conjugate gradient iterations (cf. the last row in Table 1 and the second row in Table 4) for coarse grid subproblems.

6 Conclusion

We presented a domain-decomposition approach for linear quadratic optimal control problems governed by parabolic advection-diffusion equations. The optimality conditions were decomposed using a spatial domain decomposition. This resulted in a Schur complement formulation with unknowns given by the state and adjoint variables restricted to the interfaces of space-time cylinder subdomains. It was shown that the application of the Schur complement operator requires the parallel solution of space-time cylinder subdomain optimal control problems with Dirichlet boundary conditions on the interface. The Schur complement equation was solved using a preconditioned Krylov subspace method (sQMR). The proposed preconditioner is an extension of the Neumann-Neumann preconditioner (or Robin-Robin preconditioner if advection is present) for single partial differential equations to the optimal control context. The application of the one-level version of the preconditioner was shown to require the parallel solution of space-time cylinder subdomain optimal control problems with Neumann (Robin) boundary conditions on the interface. A simple

coarse space preconditioner was added.

The tests indicate that the dependence of the performance of our preconditioned sQMR method is similar to its counterpart applied to elliptic partial differential equations. In particular, the number of sQMR iterations is independent of the size of the subdomain for the balancing Neumann-Neumann (or Robin-Robin, respectively) preconditioner. Last, the preconditioned balanced Robin-Robin preconditioner seems superior to the approach of Benamou by comparing the iteration numbers. Extensions of the approach to several space dimensions are under investigation.

Unfortunately, no theoretical convergence analysis is available yet. Additional research is also required for the construction of less expensive coarse grid Schur complement operators. Furthermore, strategies are needed to dynamically adjust the stopping tolerances for the conjugate gradient method used to solve the space-time cylinder subdomain optimal control problems arising in the application of the Schur complement as well as in the preconditioner to the performance of the outer sQMR iteration. Recent investigations of so-called flexible Krylov subspace methods will be useful for this task.

Acknowledgments

The second author would like to thank the Department of Computational and Applied Mathematics at Rice University, Houston, Texas for their hospitality.

A Interpretation of the subdomain optimality systems

A.1 Interpretation of the decomposed optimality conditions

We give an interpretation of the subdomain problems (15) as well as of the interface coupling conditions (16) in terms of the original problem formulation.

Formally, (15) may be interpreted as a semi–discretization of

$$-\langle p_i'(t),\psi_i\rangle_{L^2(\Omega_i)} + a_i(\psi_i,p_i(t)) = -\alpha_1\langle y_i(t)-\widehat{y}(t),\psi_i\rangle_{L^2(\Omega_i)} \quad \forall\psi_i\in H_0^1(\Omega_i), \tag{A.1a}$$

$$p_i(t)|_{\partial\Omega_i\setminus\partial\Omega} = p_\Gamma|_{\partial\Omega_i\setminus\partial\Omega}, \tag{A.1b}$$

$$p_i(t)|_{\partial\Omega_i\cap\partial\Omega} = 0, \tag{A.1c}$$

$$p_i(T) = -\alpha_2(y_i(T)-\widehat{y}_T)|_{\Omega_i}, \tag{A.1d}$$

$$\alpha_3\langle u_i(t),\mu_i\rangle_{L^2(\Omega_i)} + b_i(\mu_i,p_i(t)) = 0 \quad \forall\mu_i\in L^2(\Omega_i), \tag{A.1e}$$

$$\langle y_i'(t),\phi_i\rangle_{L^2(\Omega_i)} + a_i(y_i(t),\phi_i) + b_i(u_i(t),\phi_i) = \langle f(t),\phi_i\rangle_{L^2(\Omega_i)} \quad \forall\phi_i\in H_0^1(\Omega_i), \tag{A.1f}$$

$$y_i(t)|_{\partial\Omega_i\setminus\partial\Omega} = y_\Gamma|_{\partial\Omega_i\setminus\partial\Omega}, \tag{A.1g}$$

$$y_i(t)|_{\partial\Omega_i\cap\partial\Omega} = 0, \tag{A.1h}$$

$$y_i(0) = y_0|_{\Omega_i}. \tag{A.1i}$$

The system (A.1), in turn, may be interpreted as the weak form of

$$\begin{aligned}
&-\partial_t p_i(x,t) - \mu\Delta p_i(x,t) - \mathbf{a}(x)\nabla p_i(x,t)\\
&\qquad +(c(x)-\nabla\mathbf{a}(x))p_i(x,t) = -\alpha_1(y_i(x,t)-\widehat{y}(x,t)) && \text{in } \Omega_i\times(0,T),\\
&\qquad p_i(x,t) = 0 && \text{on } (\partial\Omega_i\cap\partial\Omega)\times(0,T),\\
&\qquad p_i(x,t) = p_\Gamma(x,t) && \text{on } (\partial\Omega_i\setminus\partial\Omega)\times(0,T),\\
&\qquad p_i(x,T) = -\alpha_2(y_i(x,T)-\widehat{y}_T(x)) && \text{in } \Omega_i,\\
&\qquad -p_i(x,t)+\alpha_3 u_i(x,t) = 0 && \text{on } \Omega_i\times(0,T),\\
&\partial_t y_i(x,t) - \mu\Delta y_i(x,t)\\
&\quad +\mathbf{a}(x)\nabla y_i(x,t) + c(x)y_i(x,t) = f(x,t)+u_i(x,t) && \text{in } \Omega_i\times(0,T),\\
&\qquad y_i(x,t) = 0 && \text{on } (\partial\Omega_i\cap\partial\Omega)\times(0,T),\\
&\qquad y_i(x,t) = y_\Gamma(x,t) && \text{on } (\partial\Omega_i\setminus\partial\Omega)\times(0,T),\\
&\qquad y_i(x,0) = y_0(x) && \text{in } \Omega_i.
\end{aligned}$$

Using the ideas in [22, Sec. 3.1], one can formally interpret (A.1) as the system of optimality conditions for the subdomain optimal control problem

$$\begin{aligned}
\text{minimize}\quad & \frac{\alpha_1}{2}\int_0^T\int_{\Omega_i}(y_i(x,t)-\hat{y}(x,t))^2dxdt\\
&+\int_0^T -\langle y_i(t),\widehat{p}_\Gamma'(t)\rangle_{L^2(\Omega_i)} + a_i(y_i(t),\widehat{p}_\Gamma(t))\\
&\qquad + b_i(u_i(t),\widehat{p}_\Gamma(t)) - \langle f(t),\widehat{p}_\Gamma(t)\rangle_{L^2(\Omega_i)}dt\\
&+\frac{\alpha_2}{2}\int_{\Omega_i}(y_i(x,T)-\hat{y}_T(x))^2dx + \frac{\alpha_3}{2}\int_0^T\int_{\Omega_i}u_i^2(x,t)dxdt
\end{aligned} \tag{A.3a}$$

subject to

$$\langle y_i'(t),\phi_i\rangle_{L^2(\Omega_i)} + a_i(y_i(t),\phi_i) + b_i(u_i(t),\phi_i) = \langle f(t),\phi_i\rangle_{L^2(\Omega_i)} \qquad \forall \phi_i \in H_0^1(\Omega_i), \quad \text{(A.3b)}$$

$$y_i(t)|_{\partial\Omega_i\setminus\partial\Omega} = y_\Gamma|_{\partial\Omega_i\setminus\partial\Omega}, \quad \text{(A.3c)}$$

$$y_i(t)|_{\partial\Omega_i\cap\partial\Omega} = 0, \quad \text{(A.3d)}$$

$$y_i(0) = y_0|_{\Omega_i}, \quad \text{(A.3e)}$$

where $\widehat{p}_\Gamma$ is an extension of p_Γ onto $\Omega_i \times (0,T)$. The subdomain optimal control problem (A.3) is the infinite dimensional analogue of the semi–discretized problem (17).

The constraints (A.3b–e) are the weak form of

$$\begin{aligned}
\partial_t y_i(x,t) - \mu\Delta y_i(x,t) + \mathbf{a}(x)\nabla y_i(x,t) + c(x)y_i(x,t) &= f(x,t) + u(x,t) && \text{in } \Omega\times(0,T),\\
y_i(x,t) &= 0 && \text{on } (\partial\Omega_i\cap\partial\Omega)\times(0,T),\\
y_i(x,t) &= y_\Gamma(x,t) && \text{on } (\partial\Omega_i\setminus\partial\Omega)\times(0,T),\\
y_i(x,0) &= y_0(x) && \text{in } \Omega_i.
\end{aligned}$$

For the interpretation of (A.3a), we note that the definition (10) and application of integration by parts gives

$$\begin{aligned}
&\int_0^T -\langle y_i(t),\widehat{p}_\Gamma'(t)\rangle_{L^2(\Omega_i)} + a_i(y_i(t),\widehat{p}_\Gamma(t)) + b_i(u_i(t),\widehat{p}_\Gamma(t)) - \langle f(t),\widehat{p}_\Gamma(t)\rangle_{L^2(\Omega_i)}dt\\
&= \langle y_0,\widehat{p}_\Gamma(0)\rangle_{L^2(\Omega_i)} - \langle y_i(T),\widehat{p}_\Gamma(T)\rangle_{L^2(\Omega_i)}\\
&\quad + \int_0^T\int_{\Omega_i}\Big(\partial_t y_i(x,t) - \mu\Delta y_i(x,t) + \mathbf{a}(x)\nabla y_i(x,t) + c(x)y_i(x,t) - f(x,t) - u_i(x,t)\Big)\widehat{p}_\Gamma(x,t)dxdt\\
&\quad + \int_0^T\int_{\partial\Omega_i}\Big(\mu\frac{\partial}{\partial\mathbf{n}_i}y_i(x,t) - \tfrac{1}{2}\mathbf{n}_i\mathbf{a}(x)y_i(x,t)\Big)p_\Gamma(x,t)dxdt\\
&= \langle y_0,\widehat{p}_\Gamma(0)\rangle_{L^2(\Omega_i)} - \langle y_i(T),\widehat{p}_\Gamma(T)\rangle_{L^2(\Omega_i)} + \int_0^T\int_{\partial\Omega_i}\Big(\mu\frac{\partial}{\partial\mathbf{n}_i}y_i(x,t) - \tfrac{1}{2}\mathbf{n}_i\mathbf{a}(x)y_i(x,t)\Big)p_\Gamma(x,t)dxdt
\end{aligned}$$

since y_i solves (A.2f–i). Since $\widehat{p}_\Gamma$ is the extension of a function defined on a set of measure zero, the integrals involving $\widehat{p}_\Gamma$ are formally set to zero and the objective (A.3a) can be interpreted as

$$\begin{aligned}
&\frac{\alpha_1}{2}\int_0^T\int_{\Omega_i}(y_i(x,t)-\hat{y}(x,t))^2dxdt + \frac{\alpha_2}{2}\int_{\Omega_i}(y_i(x,T)-\hat{y}_T(x))^2dx\\
&+\frac{\alpha_3}{2}\int_0^T\int_{\Omega_i}u_i^2(x,t)dxdt + \int_0^T\int_{\partial\Omega_i}\Big(\mu\frac{\partial}{\partial n_i}y_i(x,t) - \tfrac{1}{2}\mathbf{n}_i\mathbf{a}(x)y_i(x,t)\Big)p_\Gamma(x,t)dxdt.
\end{aligned}$$

Since (A.3) is an optimization problem over y_i, u_i, we may replace the last integral by

$$\sum_{\partial\Omega_i\cap\partial\Omega_j}\int_0^T\int_{\partial\Omega_i}\Big(\big(\mu\frac{\partial}{\partial n_i}y_i(x,t) - \tfrac{1}{2}\mathbf{n}_i\mathbf{a}(x)y_i(x,t)\big) + \big(\mu\frac{\partial}{\partial n_j}y_j(x,t) - \tfrac{1}{2}\mathbf{n}_j\mathbf{a}(x)y_j(x,t)\big)\Big)p_\Gamma(x,t)dxdt.$$

This will only shift the objective function by a constant. The addition of this constant, however, reveals the connection between the objective function in (A.3) and the transmission

conditions (9) for the states.

We off the following interpretation of (16): Let $V_\Gamma := \{v_\Gamma \in H^{1/2}(\Gamma) : v_\Gamma = v|_\Gamma \text{ for some } v \in H_0^1(\Omega)\}$ and similarly we introduce V_{Γ_i}. Given a v_Γ, we use $\widehat{v}_\Gamma \in H_0^1(\Omega)$ to denote its extension. The transmission conditions (16) may be interpreted as a semi–discretizations of

$$\sum_{i=1}^{s} \left(p_i(T) + \alpha_2(y_i(T) - \hat{y}_T)\right)|_{\Gamma_i} = 0, \tag{A.4a}$$

$$\sum_{i=1}^{s} -\langle \partial_t p_i(t), \widehat{v}_\Gamma\rangle_{L^2(\Omega_i)} + a_i(\widehat{v}_\Gamma, p_i(t)) + \alpha_1\langle y_i(t) - \hat{y}(t), \widehat{v}_\Gamma\rangle_{L^2(\Omega_i)} = 0, \tag{A.4b}$$

$$\sum_{i=1}^{s} \langle \partial_t y_i(t), \widehat{q}_\Gamma\rangle_{L^2(\Omega_i)} + a_i(y_i(t), \widehat{q}_\Gamma) + b_i(u_i(t), \widehat{q}_\Gamma) = \langle f, \widehat{v}_\Gamma\rangle_{L^2(\Omega_i)}, \tag{A.4c}$$

for all $v_\Gamma, q_\Gamma \in V_{\Gamma_i}$ respectively. The definition (10) and application of integration by parts gives

$$\begin{aligned}
&-\langle \partial_t p_i(t), \widehat{v}_\Gamma\rangle_{L^2(\Omega_i)} + a_i(\widehat{v}_\Gamma, p_i(t)) + \alpha_1\langle y_i(t) - \hat{y}(t), \widehat{v}_\Gamma\rangle_{L^2(\Omega_i)} \\
&= \int_{\Omega_i} \Big(-\partial_t p_i(x,t) - \mu\Delta p_i(x,t) - \mathbf{a}(x)\nabla p_i(x,t) + (c(x) - \nabla\mathbf{a}(x))p_i(x,t) + \alpha_1(y_i(x,t) - \hat{y}(x,t))\Big)\widehat{v}_\Gamma(x)dx \\
&\quad + \int_{\partial\Omega_i} \Big(\mu\frac{\partial}{\partial\mathbf{n}_i}p_i(x,t) + \tfrac{1}{2}\mathbf{a}(x)\mathbf{n}_i\, p_i(x,t)\Big)v_\Gamma dx \\
&= \int_{\partial\Omega_i} \Big(\mu\frac{\partial}{\partial\mathbf{n}_i}p_i(x,t) + \tfrac{1}{2}\mathbf{a}(x)\mathbf{n}_i\, p_i(x,t)\Big)v_\Gamma dx,
\end{aligned}$$

since p_i solves (A.2a). Hence, (A.4b) may be interpreted as

$$\sum_{i=1}^{s} \int_{\partial\Omega_i} \Big(\mu\frac{\partial}{\partial\mathbf{n}_i}p_i(x,t) + \tfrac{1}{2}\mathbf{a}(x)\mathbf{n}_i\, p_i(x,t)\Big)v_\Gamma dx = 0 \quad \forall v_\Gamma \in V_\Gamma,$$

that is

$$\mu\frac{\partial}{\partial\mathbf{n}_i}p_i(x,t) + \tfrac{1}{2}\mathbf{a}(x)\mathbf{n}_i\, p_i(x,t) = -\Big(\mu\frac{\partial}{\partial\mathbf{n}_j}p_j(x,t) + \tfrac{1}{2}\mathbf{a}(x)\mathbf{n}_j\, p_j(x,t)\Big) \quad x \in \partial\Omega_i \cap \partial\Omega_j.$$

Analogously, (A.4c) may be interpreted as

$$\mu\frac{\partial}{\partial\mathbf{n}_i}y_i(x,t) - \tfrac{1}{2}\mathbf{a}(x)\mathbf{n}_i\, y_i(x,t) = -\Big(\mu\frac{\partial}{\partial\mathbf{n}_j}y_j(x,t) - \tfrac{1}{2}\mathbf{a}(x)\mathbf{n}_j\, y_j(x,t)\Big) \quad x \in \partial\Omega_i \cap \partial\Omega_j.$$

Equation (A.4a) can be seen as the weak form of

$$p_i(x,T) = -\alpha_2(y_i(x,T) - \hat{y}_T(x)) \quad \text{ on } \Gamma_i.$$

A.2 Interpretation of the optimality conditions defining the inverse subdomain Schur complement operator

Formally, the system (23) may be interpreted as the semi–discretization of the following system of differential equations. Let $V := \{v \in H^1(\Omega_i) : v = 0 \text{ on } \partial\Omega_i \cap \partial\Omega\}$. Further, we denote by $Tr_\Gamma(v)$ the trace of v on the interface $\Gamma = \partial\Omega_i \backslash \partial\Omega$ Then, for all $\psi_i \in V$, all $\mu_i \in L^2(\Omega_i)$ and $t \in (0,T)$

$$-\langle p_i'(t), \psi_i\rangle_{L^2(\Omega_i)} + a_i(\psi_i, p_i(t)) = -\alpha_1 \langle y_i(t), \psi_i\rangle_{L^2(\Omega_i)} + \langle r_2(t), Tr_\Gamma(\psi_i)\rangle_{L^2(\Gamma)} \tag{A.5a}$$

$$\alpha_3 \langle u_i(t), \mu_i\rangle_{L^2(\Omega_i)} + b_i(\mu_i, p_i(t)) = 0 \tag{A.5b}$$

$$\langle y_i'(t), \psi_i\rangle_{L^2(\Omega_i)} + a_i(y_i(t), \psi_i) + b_i(u_i(t), \psi_i) = \langle r_3(t), Tr_\Gamma(\psi_i)\rangle_{L^2(\Gamma)} \tag{A.5c}$$

$$\langle p_i(T) + \alpha_2 y_i(T), \psi_i\rangle_{L^2(\Omega_i)} = \langle r_1, Tr_\Gamma(\psi_i)\rangle_{L^2(\Gamma)} \tag{A.5d}$$

The system (A.5) may be interpreted as the weak form of

$$\begin{aligned}
-\partial_t p_i(x,t) - \mu\Delta p_i(x,t) & \\
+(c(x,t) - \nabla\mathbf{a}(x))p(x,t) - \mathbf{a}(x)\nabla p_i(x,t) &= -\alpha_1 y_i(x,t) && \text{in } \Omega_i \times (0,T)\\
p_i(x,t) &= 0 && \text{on } \partial\Omega_i \cap \partial\Omega \times (0,T)\\
\left(\mu\frac{\partial}{\partial \mathbf{n}_i} + \frac{1}{2}\mathbf{a}_i(x)\mathbf{n}_i\right) p_i(x,t) &= r_2(x,t) && \text{on } \Gamma \times (0,T)\\
p_i(x,T) + \alpha_2 y_i(x,T) &= \widehat{r}_1(x) && \text{in } \Omega_i\\
\partial_t y_i(x,t) - \mu\Delta y_i(x,t) + c(x)y_i(x,t) + \mathbf{a}(x)\nabla y_i(x,t) &= u(x,t) && \text{in } \Omega_i \times (0,T)\\
y_i(x,t) &= 0 && \text{on } \partial\Omega_i \cap \partial\Omega \times (0,T)\\
\left(\mu\frac{\partial}{\partial \mathbf{n}_i} - \frac{1}{2}\mathbf{a}_i(x)\mathbf{n}_i\right) y_i(x,t) &= r_3(x,t) && \text{on } \Gamma \times (0,T)\\
y_i(x,0) &= 0 && \text{in } \Omega_i\\
-p_i(x,t) + \alpha_3 u_i(x,t) &= 0 && \text{in } \Omega_i \times (0,T),
\end{aligned}$$

where $\widehat{r}_1$ is the extension of r_1 defined on Γ and extended by zero on Ω_i. Using similar ideas as in the previous sections one can formally interprete the above equations as optimality system for the subdomain optimal control problem

$$\begin{aligned}
\text{minimize } &\frac{\alpha_1}{2}\int_0^T \int_{\Omega_i} y_i^2(x,t)dxdt + \frac{\alpha_2}{2}\int_{\Omega_i} y_i^2(x,T)dx\\
&+ \frac{\alpha_3}{2}\int_0^T \int_{\Omega_i} u_i^2(x,t)dxdt - \int_0^T \int_\Gamma r_2(x,t)y_i(x,t)dxdt - \int_\Gamma r_1(x)y_i(x,T)dx
\end{aligned}$$

subject to

$$\begin{aligned}
\partial_t y_i(x,t) - \mu\Delta y_i(x,t) + c(x)y_i(x,t) + \mathbf{a}(x)\nabla y_i(x,t) &= u(x,t) && \text{in } \Omega_i \times (0,T) \\
y_i(x,t) &= 0 && \text{on } \partial\Omega_i \cap \partial\Omega \times (0,T) \\
\left(\mu\frac{\partial}{\partial \mathbf{n}_i} - \frac{1}{2}\mathbf{a}_i(x)\mathbf{n}_i\right) y_i(x,t) &= r_3(x,t) && \text{on } \Gamma \times (0,T) \\
y_i(x,0) &= 0 && \text{in } \Omega_i.
\end{aligned}$$

References

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

[2] I. Lasiecka, R. Triggiani, Control Theory for Partial Differential Equations. Continuous and Approximation Theories Volume I. Abstract Parabolic Systems, Cambridge University Press, Cambdridge, New York, 1999.

[3] J.-L. Lions, Optimal Control of Systems Governed by Partial Differential Equations, Springer Verlag, Berlin, Heidelberg, New York, 1971.

[4] F. Abergel, R. Temam, On some control problems in fluid mechanics, Theoretical and Computational Fluid Dynamics 1 (1990) 303–325.

[5] M. D. Gunzburger, Perspectives in Flow Control and Optimization, SIAM, Philadelphia, 2003.

[6] M. Heinkenschloss, F. Tröltzsch, Analysis of the Lagrange–SQP–Newton method for the control of a phase field equation, Control and Cybernetics 28 (1999) 177–211.

[7] M. Meyyappan, Computational Modelling in Semiconductor Processing, Artech House, Boston, London, 1995.

[8] R. Pinnau, G. Thömmes, Optimal boundary control of glass cooling processes, Math. Methods Appl. Sci. 27 (11) (2004) 1261–1281.

[9] A. Unger, F. Tröltzsch, Fast solution of optimal control problems in the selective cooling of steel, Zeitschrift für Angewandte Mathematik un Mechanik (ZAMM) 81 (2001) 447–456.

[10] A. Comas, Time domain decomposition methods for second order linear quadratic optimal control problems, Master's thesis, Department of Computational and Applied Mathematics, Rice University, Houston, TX (2004).

[11] J. E. Lagnese, G. Leugering, Time-domain decomposition of optimal control problems for the wave equation, System and Control Letters 48 (2003) 229–242.

[12] J. E. Lagnese, G. Leugering, Domain Decomposition Methods in Optimal Control of Partial Differential Equations, Vol. 148 of International Series of Numerical Mathematics, Birkhäuser Verlag, Basel, 2004.

[13] M. Heinkenschloss, Time–domain decomposition iterative methods for the solution of distributed linear quadratic optimal control problems, Journal of Computational and Applied Mathematics 173 (2005) 169–198.

[14] J.-D. Benamou, A domain decomposition method for control problems, in: P. Bjørstad, M. Espedal, D. Keyes (Eds.), DD9 Proceedings, Domain Decomposition Press, Bergen, Norway, 1998, pp. 266–273.

[15] J. D. Benamou, Domain decomposition, optimal control of systems governed by partial differential equations, and synthesis of feedback laws, J. Optim. Theory Appl. 102 (1) (1999) 15–36.

[16] A. Bounaim, On the optimal control problem of the heat equation: New formulation of the problem using a non-overlapping domain decomposition technique, Tech. rep., Scientific Computing Group, Department of Informatics. University of Oslo (2002).

[17] A. Quarteroni, A. Valli, Domain Decomposition Methods for Partial Differential Equations, Oxford University Press, Oxford, 1999.

[18] B. Smith, P. Bjørstad, W. Gropp, Domain Decomposition. Parallel Multilevel Methods for Elliptic Partial Differential Equations, Cambridge University Press, Cambridge, 1996.

[19] A. Toselli, O. Widlund, Domain Decomposition Methods - Algorithms and Theory, Computational Mathematics, Vol. 34, Springer–Verlag, Berlin, 2004.

[20] R. Bartlett, M. Heinkenschloss, D. Ridzal, B. van Bloemen Waanders, Domain decomposition methods for advection dominated linear–quadratic elliptic optimal control problems, Comp. Meth. Appl. Mech. Engng., submitted, 2005.

[21] M. Heinkenschloss, H. Nguyen, Balancing Neumann-Neumann methods for elliptic optimal control problems, in: R. Kornhuber, R. H. W. Hoppe, J. Periaux, O. Pironneau, O. B. Widlund, J. Xu (Eds.), Domain Decomposition methods in Science and Engineering, Lecture Notes in Computational Science and Engineering, Vol. 40, Springer-Verlag, Heidelberg, 2004, pp. 589–596.

[22] M. Heinkenschloss, H. Nguyen, Neumann-Neumann domain decomposition preconditioners for linear–quadratic elliptic optimal control problems, SIAM J. Sci. Comput., submitted, 2004.

[23] A. Quarteroni, A. Valli, Numerical Approximation of Partial Differential Equations, Springer, Berlin, Heidelberg, New York, 1994.

[24] H. G. Roos, M. Stynes, L. Tobiska, Numerical Methods for Singularly Perturbed Differential Equations, Computational Mathematics, Vol. 24, Springer–Verlag, Berlin, 1996.

[25] G. Lube, L. Müller, F. C. Otto, A non-overlapping domain decomposition method for the advection-diffusion problem, Computing 64 (1) (2000) 49–68.

[26] Y. Achdou, P. L. Tallec, F. Nataf, M. Vidrascu, A domain decomposition preconditioner for an advection-diffusion problem, Comput. Methods Appl. Mech. Engrg. 184 (2-4) (2000) 145–170.

[27] Y. Achdou, F. Nataf, A Robin-Robin preconditioner for an advection-diffusion problem, C. R. Acad. Sci. Paris Sér. I Math. 325 (11) (1997) 1211–1216.

[28] J. Jahn, Introduction to the Theory of Nonlinear Optimization, 2nd Edition, Springer Verlag, Berlin, Heidelberg, New York, 1996.

[29] M. Heinkenschloss, M. Herty, Distributed solution of optimal control problems governed by parabolic equations, in: A. J. Kurdila, P. M. Pardalos, M. Zabarankin (Eds.), Robust Optimization: Directed Design, Nonconvex Optimization and Its Applications, Vol. 81, Springer-Verlag, Heidelberg, 2006, pp. 71–90.

[30] R. W. Freund, N. M. Nachtigal, Software for simplified Lanczos and QMR algorithms, Applied Numerical Mathematics 19 (1995) 319–341.

[31] J.-D. Benamou, A domain decomposition method with coupled transmission conditions for the optimal control of systems governed by elliptic partial differential equations, SIAM J. Numer. Anal. 33 (1996) 2401–2416.