

DOMAIN DECOMPOSITION PRECONDITIONERS FOR LINEAR-QUADRATIC ELLIPTIC OPTIMAL CONTROL PROBLEMS

MATTHIAS HEINKENSCHLOSS AND HOANG NGUYEN

ABSTRACT. We develop and analyze a class of overlapping domain decomposition (DD) preconditioners for linear-quadratic elliptic optimal control problems. Our preconditioners utilize the structure of the optimal control problems. Their execution requires the parallel solution of subdomain linear-quadratic elliptic optimal control problems, which are essentially smaller subdomain copies of the original problem. This work extends to optimal control problems the application and analysis of overlapping DD preconditioners, which have been used successfully for the solution of single PDEs.

We prove that for a class of problems the performance of the two-level versions of our preconditioners is independent of the mesh size and of the subdomain size.

1. INTRODUCTION

This paper is concerned with the development and analysis of a class of domain decomposition (DD) preconditioners for linear-quadratic elliptic optimal control problems. Such problems arise in many applications (see e.g., [35, 41, 47]) and, perhaps more importantly, they arise as subproblems in Newton-type or sequential quadratic programming (SQP) methods for many nonlinear elliptic control problems (see, e.g., [9, 18, 33, 34, 43]). After a discretization, a linear-quadratic elliptic optimal control problem leads to a large-scale quadratic programming problem whose solution is, under suitable conditions, characterized through the linear system of optimality conditions. This linear system is large-scale and indefinite and it usually has to be solved iteratively. The spectrum of this system is determined, among other things by the mesh size h and the control regularization parameter α (see the formulation (1.2) below). Often one observes that the condition number of the optimality system matrix grows like h^{-2} and like α^{-1} . Good preconditioners are important for the overall performance of solution methods for elliptic control problems. We note that the structure of the optimality system arising in elliptic control problems is different from the structure of the saddle point systems arising in the solution of the Stokes equation or in the solution of elliptic partial differential equations (PDEs) using mixed finite element methods. Thus, the preconditioners developed in that context cannot be used in the solution of optimal control problems, in general.

This paper is concerned with the development of DD preconditioners for the optimality system of linear-quadratic elliptic optimal control problems. Our preconditioners utilize the structure of the optimal control problems. The execution of our preconditioners requires the parallel solution of subdomain linear-quadratic elliptic optimal control problems, which are essentially smaller subdomain copies of the original problem. We extend

Date: November 23, 2004.

1991 Mathematics Subject Classification. Primary 49M05, 65N55; Secondary 49N10, 49M27.

This work was supported in part by NSF Grant ACI-0121360 and by the Los Alamos National Laboratory Computer Science Institute (LACSI) through LANL contract number 03891-99-23 as part of the prime contract (W-7405-ENG-36) between the Department of Energy and the Regents of the University of California.

overlapping DD preconditioners which have been used successfully for the solution of single PDEs to optimal control problems. We prove that the performance of the two-level versions of our preconditioners is independent of the mesh size h and of the subdomain size H . Numerical studies indicate that the performance of our preconditioners for optimal control problems is comparable to the performance of their counterparts applied to single PDEs. Moreover, the performance of our preconditioners seems to be rather insensitive to the size of the control regularization parameter α .

Let V and U be Hilbert spaces with duals V' and U' , respectively, let

$$(1.1a) \quad a : V \times V \rightarrow \mathbb{R}, \quad b : U \times V \rightarrow \mathbb{R},$$

be continuous bilinear forms, let

$$(1.1b) \quad m : V \times V \rightarrow \mathbb{R}, \quad q : U \times U \rightarrow \mathbb{R},$$

be symmetric and continuous bilinear forms, let

$$(1.1c) \quad c, f : V \rightarrow \mathbb{R}, \quad d : U \rightarrow \mathbb{R}$$

be continuous linear functionals and let $\alpha > 0$. We consider the linear-quadratic elliptic optimal control problem

$$(1.2a) \quad \text{minimize} \quad \frac{1}{2}m(y, y) - c(y) + \frac{\alpha}{2}q(u, u) - d(u)$$

$$(1.2b) \quad \text{subject to} \quad a(y, \phi) + b(u, \phi) = f(\phi) \quad \forall \phi \in V.$$

The unknowns are the state $y \in V$ and the control $u \in U$. Assumptions on the (bi)linear forms that ensure existence and uniqueness of a solution of (1.2) will be given in Section 2.1.

The formulation (1.2) covers many problems, including the example problems

$$(1.3a) \quad \text{minimize} \quad \frac{1}{2} \int_{\Omega_o} (y(x) - \hat{y}(x))^2 dx + \frac{\alpha}{2} \int_{\partial\Omega} u^2(x) dx,$$

$$(1.3b) \quad \text{subject to} \quad -\Delta y(x) + \sigma y(x) = \hat{f}(x) \quad \text{in } \Omega,$$

$$(1.3c) \quad \frac{\partial}{\partial n} y(x) = u(x) \quad \text{on } \partial\Omega$$

and

$$(1.4a) \quad \text{minimize} \quad \frac{1}{2} \int_{\Omega_o} (y(x) - \hat{y}(x))^2 dx + \frac{\alpha}{2} \int_{\Omega} u^2(x) dx,$$

$$(1.4b) \quad \text{subject to} \quad -\Delta y(x) = \hat{f}(x) + u(x) \quad \text{in } \Omega,$$

$$(1.4c) \quad y(x) = 0 \quad \text{on } \partial\Omega.$$

In both examples $\Omega_o \subset \Omega$, $\hat{y} \in L^2(\Omega_o)$, $\hat{f} \in L^2(\Omega)$ are given functions, and $\alpha > 0$ is a given parameter. In (1.3), $\sigma \geq 0$ is also given. Example (1.3) is a special case of (1.2) with $V = H^1(\Omega)$, $U = L^2(\partial\Omega)$, $a(y, \phi) = \int_{\Omega} \nabla y \nabla \phi + \sigma y \phi dx$, $m(y, \phi) = \int_{\Omega_o} y \phi dx$, $b(u, \phi) = \int_{\partial\Omega} u \phi dx$, $q(u, \phi) = \int_{\partial\Omega} u \phi dx$, $c(\phi) = \int_{\Omega_o} \hat{y} \phi dx$, $f(\phi) = \int_{\Omega} \hat{f} \phi dx$, and $d = 0$. Example (1.4) is a special case of (1.2) with $V = H_0^1(\Omega)$, $U = L^2(\Omega)$, $a(y, \phi) = \int_{\Omega} \nabla y \nabla \phi dx$, m, c, f defined as in example (1.3) and $b(u, \phi) = \int_{\Omega} u \phi dx$, $q(u, \phi) = \int_{\Omega} u \phi dx$, $d = 0$.

Our DD preconditioners for (1.2) are derived from a framework based on subspace decomposition. Such a framework is well known for elliptic PDEs (see, e.g., [12, 20, 21, 22, 48, 50].) Since the optimality system for the linear-quadratic elliptic optimal control problem (1.2) is highly indefinite, the well-posedness of subspace projection operators and

the equivalence between the optimality system for (1.2) and a system involving combinations of the subspace projection operators need to be examined. This is done in Section 2, where we will provide conditions based on the structure of linear-quadratic elliptic optimal control problems that guarantee the well-posedness of subspace projection operators and allow an equivalent reformulation of the optimality system for (1.2) using the subspace projection operators.

The subspace decomposition framework is applied in Section 3 to derive overlapping DD preconditioners for (1.2). We extend the ideas in [10, 12, 13] to examine the performance of our preconditioners when applied with GMRES. We will show that, under suitable conditions on the optimal control problem and on the finite element mesh and subdomains, the convergence factor of GMRES preconditioned with a two-level version of our overlapping preconditioners is independent of the mesh size h and of the subdomain size H . This result extends to the optimal control context results that are well known for overlapping DD methods applied to individual PDEs. The general idea is to split the operator arising in the optimality system for (1.2) into a symmetric positive definite part and a remainder part, so that the symmetric positive definite part ‘dominates’ the remainder part. This requires a somewhat non-standard non-symmetric formulation of the optimality conditions, which in the discretized case corresponds to an interchange of two row blocks. This will be illustrated in Section 3.3. In this section we will also show that this row block interchange is not necessary in the actual use of our preconditioners, i.e., in practice one can work with the symmetric version of the optimality conditions.

While the general outline of the proof for the convergence factor estimates follows that of [10, 12, 13], many technical details need to be extended to the optimal control setting. This is accomplished in Section 3.2. Our estimates apply to the case of distributed controls (e.g. Example (1.4)). To estimate the performance of our overlapping preconditioners, we need some discretization error bounds for linear-quadratic elliptic optimal control problems. Since these results are somewhat scattered in the literature and sometimes focus on specific linear-quadratic elliptic optimal control problems, we present the necessary error estimates in Section 5, focussing on the setting of this paper.

Other DD methods for linear-quadratic elliptic optimal control problems are given in [1, 2, 3, 4, 5, 6, 42, 36, 37] and DD methods for a class of elliptic parameter identification problems are discussed in [19, 40, 49]. We refer to [37] and [44] for a brief comparison of the different approaches. Our overlapping DD methods, like the DD methods in [1, 2, 3, 36, 37] are based on a decomposition of the optimality system. The subproblems that have to be solved (in parallel) are smaller subdomain versions of the optimality system. Hence, our methods as well as those in [1, 2, 3, 36, 37] require fewer communication per computation compared to the approaches in [4, 5, 6, 42] which require the (parallel) solution of smaller subdomain versions of the governing PDE and its adjoint. The higher computation to communication ratio may be preferable on some computing platforms. The subdomain optimality systems arising in our approach are optimality systems of smaller subdomain copies of the original optimal control problem, allowing code reuse. This is also true for [37], but not for [1, 2, 3]. Finally, while convergence proofs are presented in [1, 2, 3, 6, 42], there are no results in [1, 2, 3, 4, 5, 6, 42, 36, 37] that estimate the performance of the respective DD methods with respect to mesh size h , subdomain size H , or regularization parameter α . We are able to provide such estimates for our overlapping DD methods.

The non-overlapping Neumann-Neumann methods studied in [36, 37] can also be derived from the general framework in Section 2, following the work of [20, 24] for elliptic PDEs. In this sense, the present paper complements [36, 37].

Our work in this paper is also related to [11], where a nonlinear overlapping DD method is applied to the solution of (systems) of nonlinear PDEs. The method in [11] could be applied to the optimality system for an elliptic optimal control problem. If this is done for linear-quadratic problems, the method is identical to the overlapping DD methods in this paper. Hence, this paper provides theoretical justification of the approach in [11] for linear-quadratic optimal control problems.

2. SCHWARZ FRAMEWORK

We describe a general domain decomposition approach for solving linear-quadratic optimal control problems. First, we briefly review conditions for the existence of a unique solution of the abstract control problem (1.2). We then present a subspace decomposition of the optimal control problem, where the space of state, adjoint and control variables are decomposed into local spaces. This is an extension of the subspace methods well-known for analyzing domain decomposition algorithms for PDE problems (see, e.g., [12, 20, 21, 22, 48, 50].) However, since the optimality system for the linear-quadratic elliptic optimal control problem (1.2) is highly indefinite, the well-posedness of subspace projection operators and the equivalence between the optimality system for (1.2) and a system involving combinations of the subspace projection operators need to be examined. In this section we will give conditions based on the structure of linear-quadratic elliptic optimal control problems that guarantee the well-posedness of subspace projection operators and allow the equivalent reformulation of the optimality system for (1.2) using an addition of the subspace projection operators.

2.1. Problem Formulation. We assume that the state equation (1.2b) is surjective, i.e., that for all $l \in V'$ there exist $y \in V$, $u \in U$ such that

$$(2.1a) \quad a(y, \phi) + b(u, \phi) = l(\phi) \quad \forall \phi \in V$$

and that the objective function is strictly convex on the null-space of the constraints, i.e., that there exists a constant $\zeta > 0$ such that

$$(2.1b) \quad m(y, y) + \alpha q(u, u) \geq \zeta(\|y\|_V^2 + \|u\|_U^2)$$

for all $y \in V$, $u \in U$ with

$$(2.1c) \quad a(y, \phi) + b(u, \phi) = 0 \quad \phi \in V.$$

Remark 2.1. *The conditions (2.1) are satisfied if a and q are coercive. This is the case for the first example problem (1.3) if $\sigma > 0$ and for the second example problem (1.4). If $\sigma = 0$, the bilinear form in the first example problem (1.3) is not coercive on $V = H^1(\Omega)$. However, one can show that the first example problem with $\sigma = 0$ satisfies (2.1) [37].*

The next result is standard, see, e.g., [38, 41, 44].

Theorem 2.2. *Let conditions (2.1) be satisfied. Problem (1.2) has a unique solution $y^* \in V$, $u^* \in U$. The pair $y^* \in V$, $u^* \in U$ solves (1.2), if and only if there exists a unique*

adjoint variable $p^* \in V$ such that

$$(2.2a) \quad a(\theta, p^*) + m(y^*, \theta) = c(\theta) \quad \forall \theta \in V,$$

$$(2.2b) \quad \alpha q(u^*, \mu) + b(\mu, p^*) = d(\mu) \quad \forall \mu \in U,$$

$$(2.2c) \quad a(y^*, \phi) + b(u^*, \phi) = f(\phi) \quad \forall \phi \in V.$$

Moreover, there exists a constant $\kappa > 0$ such that

$$(2.3) \quad \|y^*\|_V + \|u^*\|_U + \|p^*\|_V \leq \kappa (\|c\|_{V'} + \|d\|_{U'} + \|f\|_{V'}).$$

We define the space $Z \stackrel{\text{def}}{=} V \times U \times V$, and let $z \stackrel{\text{def}}{=} (y, u, p) \in Z$, $\psi \stackrel{\text{def}}{=} (\theta, \mu, \phi) \in Z$. Assuming that the conditions (2.1) are satisfied, the problem of solving for the optimality conditions (2.2) may be stated as finding the unique $z^* \in Z$ such that

$$(2.4) \quad K(z^*, \psi) = g(\psi) \quad \forall \psi \in Z,$$

where $K : Z \times Z \rightarrow \mathbb{R}$, and $g : Z \rightarrow \mathbb{R}$ are defined as

$$(2.5) \quad \begin{aligned} K((y, u, p), (\theta, \mu, \phi)) &= a(\theta, p) + m(y, \theta) + \alpha q(u, \mu) + b(\mu, p) \\ &\quad + a(y, \phi) + b(u, \phi), \\ g((\theta, \mu, \phi)) &= c(\theta) + d(\mu) + f(\phi). \end{aligned}$$

Remark 2.3. Note that the optimality system (2.2) can also be written as (2.4) with K and g defined by

$$(2.6) \quad \begin{aligned} K((y, u, p), (\theta, \mu, \phi)) &= a(\phi, p) + m(y, \phi) + \alpha q(u, \mu) + b(\mu, p) \\ &\quad + a(y, \theta) + b(u, \theta), \\ g((\theta, \mu, \phi)) &= c(\phi) + d(\mu) + f(\theta). \end{aligned}$$

If we compare (2.5) to (2.6) we see that the roles of θ and ϕ are interchanged. The representation of (2.4) with (2.6) corresponds to a reordering of the equations in (2.2) in the order c - b - a and interchanging θ and ϕ of the original optimality system. After a finite element discretization, (2.4) with (2.5) is obtained from (2.4) with (2.6) by a row permutation (see Section 3.3).

The bilinear form (2.5) is symmetric, i.e., satisfies $K(z, \psi) = K(\psi, z)$ for all $z, \psi \in Z$. The bilinear form (2.6) is not symmetric, but can be split into a symmetric part and a remainder part. This splitting will be important for the convergence analysis for the overlapping methods presented in Section 3.

All results derived in the remainder of this section for (2.5) remain true if (2.5) is replaced by (2.6).

2.2. Subspace Decomposition of the Optimal Control Problem. Assume that the spaces V, U are decomposable into subspaces as

$$(2.7a) \quad V = V_0 + V_1 + \dots + V_N,$$

$$(2.7b) \quad U = U_0 + U_1 + \dots + U_N,$$

with $V_i \subset V$ and $U_i \subset U$, $i = 0, \dots, N$. The space Z may then be decomposed as

$$(2.8) \quad Z = Z_0 + Z_1 + \dots + Z_N,$$

with $Z_i = V_i \times U_i \times V_i$. In a domain decomposition problem with N (possibly overlapping) subdomains, each of the spaces $Z_i, i = 1, \dots, N$, is associated with subdomain i . The special space Z_0 is used as a coarse space in a two level method and would not be needed in one-level methods.

For each $i = 0, \dots, N$, we assume that there exist local continuous bilinear forms

$$(2.9a) \quad a_i, m_i : V_i \times V_i \rightarrow \mathbb{R}, \quad b_i : U_i \times V_i \rightarrow \mathbb{R}, \quad q_i : U_i \times U_i \rightarrow \mathbb{R},$$

that are local approximations to the corresponding global bilinear forms. For $i = 0, \dots, N$ we require that m_i and q_i are symmetric, that for all $l_i \in V_i'$ there exist $y_i \in V_i$, $u_i \in U_i$ with

$$(2.10a) \quad a_i(y_i, \phi_i) + b_i(u_i, \phi_i) = l_i(\phi_i) \quad \forall \phi_i \in V_i$$

and that there exists a constant $\zeta_i > 0$ such that

$$(2.10b) \quad m_i(y_i, y_i) + \alpha q_i(u_i, u_i) \geq \zeta_i (\|y_i\|_{V_i}^2 + \|u_i\|_{U_i}^2)$$

for all $y_i \in V_i$, $u_i \in U_i$ with

$$(2.10c) \quad a_i(y_i, \phi_i) + b_i(u_i, \phi_i) = 0 \quad \phi_i \in V_i.$$

Often the local forms are defined as restrictions of the global forms to the local spaces, i.e.,

$$(2.11) \quad \begin{aligned} a_i(y_i, \phi_i) &= a(y_i, \phi_i), & b_i(u_i, \mu_i) &= b(u_i, \mu_i), \\ m_i(y_i, \phi_i) &= m(y_i, \phi_i), & q_i(u_i, \mu_i) &= q(u_i, \mu_i) \end{aligned}$$

for all $y_i, \phi_i \in V_i$ and all $u_i, \mu_i \in U_i$.

For each $i = 0, \dots, N$, we define a bilinear form $K_i : Z_i \times Z_i \rightarrow \mathbb{R}$ as a local approximation for K :

$$(2.12) \quad \begin{aligned} &K_i((y_i, u_i, p_i), (\theta_i, \mu_i, \phi_i)) \\ &= a_i(\theta_i, p_i) + m_i(y_i, \theta_i) + \alpha q_i(u_i, \mu_i) + b_i(\mu_i, p_i) + a_i(y_i, \phi_i) + b_i(u_i, \phi_i). \end{aligned}$$

Lemma 2.4. *If the local bilinear forms satisfy the requirements (2.10), then for each $z \in Z$ there exists a unique solution $z_i \in Z_i$ of*

$$(2.13) \quad K_i(z_i, \psi_i) = K(z, \psi_i) \quad \forall \psi_i \in Z_i.$$

Proof. Let $z = (y, u, p) \in Z$ be arbitrary. If we set $z_i = (y_i, u_i, p_i) \in Z_i$, then (2.13) can be written as

$$(2.14) \quad \begin{aligned} &a_i(\theta_i, p_i) + m_i(y_i, \theta_i) + \alpha q_i(u_i, \mu_i) + b_i(\mu_i, p_i) + a_i(y_i, \phi_i) + b_i(u_i, \phi_i) \\ &= a(\theta_i, p) + m(y, \theta_i) + \alpha q(u, \mu_i) + b(\mu_i, p) + a(y, \phi_i) + b(u, \phi_i) \end{aligned}$$

for all $(\theta_i, \mu_i, \phi_i) \in V_i \times U_i \times V_i$. Since z is fixed, the right hand side of (2.14) defines a continuous linear functional

$$(2.15) \quad g_z((\theta_i, \mu_i, \phi_i)) = c_z(\theta_i) + d_z(\mu_i) + f_z(\phi_i),$$

on Z_i , where

$$c_z(\theta_i) = a(\theta_i, p) + m(y, \theta_i), \quad d_z(\mu_i) = \alpha q(u, \mu_i) + b(\mu_i, p), \quad f_z(\phi_i) = a(y, \phi_i) + b(u, \phi_i).$$

Equation (2.14) is equivalent to the system

$$(2.16a) \quad a_i(\theta_i, p_i) + m_i(y_i, \theta_i) = c_z(\theta_i) \quad \forall \theta_i \in V_i,$$

$$(2.16b) \quad \alpha q_i(u_i, \mu_i) + b_i(\mu_i, p_i) = d_z(\mu_i) \quad \forall \mu_i \in U_i,$$

$$(2.16c) \quad a_i(y_i, \phi_i) + b_i(u_i, \phi_i) = f_z(\phi_i) \quad \forall \phi_i \in V_i.$$

Using assumptions (2.10) together with Theorem 2.2 we see that (2.16) is the necessary and sufficient optimality conditions for (y_i, u_i) to be the unique solution of the subspace optimal control problem

$$(2.17a) \quad \min_{y_i \in V_i, u_i \in U_i} \frac{1}{2} m_i(y_i, y_i) - c_z(y_i) + \frac{\alpha}{2} q_i(u_i, u_i) - d_z(u_i)$$

$$(2.17b) \quad \text{s.t.} \quad a_i(y_i, \phi_i) + b_i(u_i, \phi_i) = f_z(\phi_i) \quad \forall \phi_i \in V_i.$$

with corresponding adjoint p_i . This means z_i is uniquely defined by z . \square

For each $i = 0, \dots, N$, we define the operator

$$\begin{aligned} T_i : Z &\rightarrow Z, \\ z &\mapsto z_i \end{aligned}$$

where $z_i \in Z_i$ is the unique solution of (2.13). By definition of T_i ,

$$(2.18) \quad K_i(T_i z, \psi_i) = K(z, \psi_i) \quad \forall \psi_i \in Z_i.$$

It is easy to show that T_i is linear. Moreover, if K_i is the restriction of K , it is easy to show that $T_i^2 = T_i$, i.e., T_i is a projection. We note that $T_i z$ may also be stated as the solution of the linear problem

$$(2.19) \quad K_i(T_i z, \psi_i) = g_z(\psi_i) \quad \forall \psi_i \in Z_i,$$

where g_z is defined as $g_z(\psi_i) = K(z, \psi_i)$.

We are interested in transforming the problem (2.4) into an equivalent problem that is better conditioned. We define the operator $T : Z \rightarrow Z$ as

$$(2.20) \quad Tz = \sum_{i=0}^N T_i z.$$

To establish the nonsingularity of T , we impose stronger assumptions on the local bilinear forms that those in (2.10). For $i = 0, \dots, N$ we require that m_i and q_i are symmetric and that there exist $\eta > 0$, $\rho > 0$ such that

$$(2.21a) \quad a_i(y_i, y_i) \geq \eta \|y_i\|_{V_i}^2 \quad \forall y_i \in V_i,$$

$$(2.21b) \quad m_i(y_i, y_i) \geq 0 \quad \forall y_i \in V_i,$$

$$(2.21c) \quad q_i(u_i, u_i) \geq \rho \|u_i\|_{U_i}^2, \quad \forall u_i \in U_i.$$

Assumptions (2.21) imply (2.10). In fact the continuity and coercivity of a_i imply that $a_i(y_i, \phi_i) = l_i(\phi_i)$ for all $\phi_i \in V_i$ has a unique solution y_i^l for given $l_i \in V_i'$. Hence, (2.10a) is satisfied with $y_i = y_i^l$ and $u_i = 0$. The continuity and coercivity of a_i imply that for each $u_i \in U_i$, $a_i(y_i, \phi_i) + b_i(u_i, \phi_i) = 0$ has a unique solution $y_i(u_i)$, which depends Lipschitz continuously on u_i . This together with Assumptions (2.21b,c) imply (2.10b).

Lemma 2.5. *Let bilinear forms satisfy (2.1). If the local bilinear forms satisfy the requirements (2.21) for $i = 0, \dots, N$, then the operator T is nonsingular.*

Proof. Let $z \in Z$ satisfy $Tz = 0$. If we denote $z_i = (y_i, u_i, p_i) = T_i z$, $i = 0, \dots, N$, then

$$(2.22) \quad \sum_{i=0}^N (y_i, u_i, p_i) = (0, 0, 0).$$

For each $i = 0, \dots, N$, $(y_i, u_i, p_i) = T_i z$ is the solution of (2.16). Setting $\theta_i = y_i$, $\mu_i = u_i$, $\phi_i = p_i$ in (2.16), then summing over $i = 0, \dots, N$ and using (2.22) gives

$$(2.23a) \quad \sum_{i=0}^N a_i(y_i, p_i) + m_i(y_i, y_i) = \sum_{i=0}^N c_z(y_i) = c_z\left(\sum_{i=0}^N y_i\right) = 0,$$

$$(2.23b) \quad \sum_{i=0}^N \alpha q_i(u_i, u_i) + b_i(u_i, p_i) = \sum_{i=0}^N d_z(u_i) = d_z\left(\sum_{i=0}^N u_i\right) = 0,$$

$$(2.23c) \quad \sum_{i=0}^N a_i(y_i, p_i) + b_i(u_i, p_i) = \sum_{i=0}^N f_z(p_i) = f_z\left(\sum_{i=0}^N p_i\right) = 0.$$

Adding equation (2.23a) to (2.23b) and subtracting (2.23c) results in

$$\sum_{i=0}^N \alpha q_i(u_i, u_i) + m_i(y_i, y_i) = 0.$$

Each term in the previous sum is nonnegative by assumptions (2.21b,c), so we must have $u_i = 0$ for $i = 0, \dots, N$. Setting $\phi_i = y_i$ in (2.16c) and summing over $i = 0, \dots, N$ gives

$$\sum_{i=0}^N a_i(y_i, y_i) + b_i(u_i, y_i) = \sum_{i=0}^N f_z(y_i) = 0,$$

which implies $y_i = 0$, $i = 0, \dots, N$, by requirement (2.21a). Finally, setting $\theta_i = p_i$ in (2.16a) and summing over i yields

$$\sum_{i=0}^N a_i(p_i, p_i) + m_i(y_i, p_i) = \sum_{i=0}^N c_z(p_i) = 0,$$

which implies $p_i = 0$, $i = 0, \dots, N$. This shows $T_i z = 0$ for $i = 0, \dots, N$. With (2.18) we can deduce $K(z, \psi_i) = 0$ for all $\psi_i \in Z_i$, $i = 0, \dots, N$, and since $Z = Z_0 + \dots + Z_N$, we obtain $K(z, \psi) = 0$ for all $\psi \in Z$. By Theorem 2.2, $z = 0$. \square

The following example shows that the assumption (2.10) on the local bilinear forms, which is sufficient to guarantee existence of the projection operators T_i , is not sufficient to ensure nonsingularity of $\sum_{i=0}^N T_i$.

Example 2.6. Let $V = \mathbb{R}$, $U = \mathbb{R}$. We define bilinear forms $a(v, \phi) = 0$, $m(v, \phi) = v\phi$, $b(u, \mu) = u\mu$, and $q(u, \mu) = 0$. It is easy to verify that the bilinear forms satisfy (2.1). The bilinear form K is associated with the matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

and in this example we use K to also denote the above 3×3 matrix.

We set $V_i = V$, $U_i = U$, $i = 0, 1$ and we define the bilinear forms $a_0 = a$, $m_0 = m$, $b_0 = b$, $q_0 = q$ and $a_1 = a$, $m_1 = m$, $b_1 = -b$, $q_1 = q$. It is easy to verify that (2.10) is satisfied for $i = 0, 1$, but not (2.21). The bilinear forms K_i , $i = 0, 1$ are associated with the matrices

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}$$

which we also denote by K_0 , K_1 .

The matrix representations of the projection operators T_i , $i = 0, 1$, are given by $T_0 = K_0^{-1}K = I$ and $T_1 = K_1^{-1}K = \text{diag}(1, -1, -1)$. Obviously, $T_0 + T_1$ is singular.

We consider the problem

$$(2.24) \quad Tz = r,$$

where

$$(2.25) \quad r \stackrel{\text{def}}{=} \sum_{i=0}^N r_i, \quad r_i \stackrel{\text{def}}{=} T_i z^*, \quad i = 0, \dots, N,$$

and z^* is the solution of (2.4). The local right hand sides $r_i, i = 0, \dots, N$, can be computed without knowing the solution z^* . In fact, (2.4) and (2.13) imply that

$$K_i(T_i z^*, \psi_i) = K(z^*, \psi_i) = g(\psi_i) \quad \forall \psi_i \in Z_i.$$

This means that $r_i = T_i z^* \in Z_i, i = 0, \dots, N$, can be computed by solving the local problem

$$K_i(r_i, \psi_i) = g(\psi_i) \quad \forall \psi_i \in Z_i.$$

Theorem 2.7. *If the bilinear forms satisfy (2.1) and if the local bilinear forms satisfy (2.21) for $i = 0, \dots, N$, then problems (2.4) and (2.24) have the same solution.*

Proof. i. Problem (2.4) has a unique solution by Theorem 2.2. Let z^* satisfy $K(z^*, \psi) = g(\psi)$ for all $\psi \in Z$, then

$$K_i(T_i z^*, \psi_i) = K(z^*, \psi_i) = g(\psi_i) = K_i(r_i, \psi_i) \quad \forall \psi_i \in Z_i.$$

This means $K_i(T_i z^* - r_i, \psi_i) = 0$ for all $\psi_i \in Z_i$, which implies $T_i z^* = r_i$. Therefore, $Tz^* = \sum_{i=0}^N T_i z^* = \sum_{i=0}^N r_i = r$.

ii. Suppose that we have found a solution w that solves $Tw = r$. Then $Tw = Tz^*$, or $T(z^* - w) = 0$. Since T is nonsingular by Lemma 2.5, we have $(z^* - w) = 0$, so w is a solution to (2.4). \square

The transformed problem (2.24) may be solved by using a linear iterative method. The operator T is nonsymmetric, in general, and we solve (2.24) using GMRES [46], QMR [26] or any other method for nonsymmetric systems [45]. In Section 3.3, we show that the operator T has structure that allows the application of the symmetric QMR (sQMR) method [27, 28].

Let $A(\cdot, \cdot)$ be an inner product on Z and let $\|\cdot\|_A$ be the norm on Z induced by $A(\cdot, \cdot)$. We define

$$(2.26) \quad c_T = \inf_{z \neq 0} \frac{A(Tz, z)}{A(z, z)}, \quad C_T = \sup_{z \neq 0} \frac{\|Tz\|_A}{\|z\|_A}.$$

If GMRES with inner product $A(\cdot, \cdot)$ is applied to the solution of (2.24) and if $c_T > 0$, then the residual $r - Tz^{(k)}$ in the k th GMRES iteration obeys

$$(2.27) \quad \|r - Tz^{(k)}\|_A = \left(1 - \frac{c_T^2}{C_T^2}\right)^{k/2} \|r - Tz^{(0)}\|_A$$

(see [25].) In the next section we analyze a specific domain decomposition method, and estimate a lower bound for c_T and an upper bound for C_T in terms of the key problem parameters H (subdomain size), h (mesh element size) and α (regularization parameter).

We conclude this section by remarking that a multiplicative Schwarz and hybrid methods for the optimal control case can also be formulated in a straightforward way (see [48, Sec. 5.1]). However, for these methods results corresponding to Theorem 2.7 still need to be established. It is also possible to allow non-nested spaces V_i, U_i , i.e., spaces with $V_i \not\subset V, U_i \not\subset U$, if one introduces appropriate of interpolation operators (see [48, Sec. 5.1]).

3. OVERLAPPING DOMAIN DECOMPOSITION METHODS

3.1. Space Decomposition. Let $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, be a polygonal/polyhedral domain with boundary $\partial\Omega$ and let $\mathcal{D} = \Omega$ or $\mathcal{D} = \partial\Omega$. We assume that the state and the control space satisfy

$$V \subset H^1(\Omega), \quad U = L^2(\mathcal{D}).$$

We apply the finite element method for solving (1.2) or, equivalently, (2.2). Let $\{\mathcal{T}^h\}$ be a family of quasi-uniform meshes. For $\tau \in \mathcal{T}^h$, let h_τ denote the length of the longest side of τ and define

$$(3.1) \quad h = \max\{h_\tau : \tau \in \mathcal{T}^h\}.$$

We discretize states and controls using continuous piecewise linear functions. Let

$$\widehat{V}^h = \{v^h \in H^1(\Omega) : v^h|_\tau \text{ linear } \forall \tau \in \mathcal{T}^h\}.$$

The spaces of discretized states and controls are given by

$$(3.2) \quad V^h = \{v^h \in V : v^h \in \mathcal{P}_1(\tau) \forall \tau \in \mathcal{T}^h\},$$

where \mathcal{P}_1 denotes the space of piecewise linear polynomials, and

$$(3.3) \quad U^h = \left\{ u^h \in U : u^h = v^h|_{\mathcal{D}} \text{ for some } v^h \in \widehat{V}^h \right\}$$

respectively. Note that the space V^h may incorporate homogeneous Dirichlet boundary information; the space \widehat{V}^h does not. We set

$$Z^h = V^h \times U^h \times V^h.$$

Our discretization of (1.2) is given by

$$(3.4a) \quad \text{minimize} \quad \frac{1}{2}m(y^h, y^h) - c(y^h) + \frac{\alpha}{2}q(u^h, u^h) - d(u^h)$$

$$(3.4b) \quad \text{subject to} \quad a(y^h, \phi^h) + b(u^h, \phi^h) = f(\phi^h) \quad \forall \phi^h \in V^h.$$

The unknowns are $y^h \in V^h$ and $u^h \in U^h$. The necessary and sufficient optimality conditions for (3.4) are given by

$$(3.5) \quad K(z^h, \psi^h) = g(\psi^h) \quad \forall \psi^h \in Z^h,$$

where K and g are defined in (2.6), i.e., are given by

$$(3.6a) \quad K(z, \psi) = a(\phi, p) + m(y, \phi) + \alpha q(u, \mu) + b(\mu, p) + a(y, \theta) + b(u, \theta),$$

$$(3.6b) \quad g(\psi) = c(\phi) + d(\mu) + f(\theta)$$

where $z = (y, u, p)$ and $\psi = (\theta, \mu, \phi)$. The choice (3.6) over (2.5) for the (bi)linear forms K, g will be motivated in Section 3.2 below. Note that (3.5) can also be obtained by discretization of (2.4) with (3.6).

We partition the set of elements in Ω into N subdomains Ω_i , $i = 1, \dots, N$, with maximum diameter less than or equal to H . Each subdomain Ω_i is extended to a larger region $\widehat{\Omega}_i$, with an overlap such that

$$(3.7) \quad \text{distance}(\partial\widehat{\Omega}_i \cap \Omega, \partial\Omega_i \cap \Omega) \geq \delta, \quad \forall i$$

for some $\delta > 0$. In general, we assume that the amount of overlap is kept proportional to the subdomain diameter, i.e.,

$$(3.8) \quad \delta \geq \beta H$$

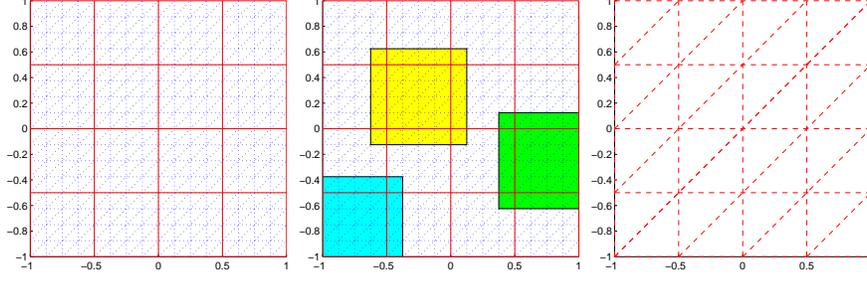


FIGURE 3.1. A sample triangulation of $\Omega = (-1, 1)^2$ partitioned into 16 subdomains (left plot). The three shaded regions are examples of extended subdomains (middle plot). A coarse mesh (right plot).

for some constant $\beta > 0$. Each extended region is assumed to not cut through any fine mesh element, and any part of $\widehat{\Omega}_i$ that is outside of the original domain Ω is cut off (see Figure 3.1). Corresponding to $\widehat{\Omega}_i$ we define $\widehat{\mathcal{D}}_i, i = 1, \dots, N$, as follows

$$\begin{aligned} \widehat{\mathcal{D}}_i &= \mathcal{D} \cap \partial \widehat{\Omega}_i & \text{if } \mathcal{D} = \partial \Omega, \\ \widehat{\mathcal{D}}_i &= \widehat{\Omega}_i & \text{if } \mathcal{D} = \Omega. \end{aligned}$$

We define state and control subspaces associated with each extended subdomain $\widehat{\Omega}_i, i = 1, \dots, N$, by

$$(3.9) \quad V_i^h = \left\{ v^h \in V^h : v^h = 0 \text{ on } \Omega \setminus \widehat{\Omega}_i \right\},$$

$$(3.10) \quad U_i^h = \left\{ u^h \in U^h : u^h = v_h|_{\widehat{\mathcal{D}}_i} \text{ for some } v^h \in \widehat{V}^h \right\}.$$

In the case of boundary controls, the set $U_i^h = \emptyset$ for all extended subdomains $\widehat{\Omega}_i$ for which the interior (relative to \mathcal{D}) of $\mathcal{D} \cap \overline{\widehat{\Omega}_i}$ is empty. We say that $\widehat{\Omega}_i, i = 1, \dots, N$, is a *control subdomain* if

$$(3.11) \quad \text{int}(\mathcal{D} \cap \overline{\widehat{\Omega}_i}) = \emptyset.$$

Here the interior is taken relative to \mathcal{D} .

The associated local product spaces are

$$(3.12) \quad Z_i^h \stackrel{\text{def}}{=} V_i^h \times U_i^h \times V_i^h \quad i = 1, \dots, N.$$

Since $V^h = V_1^h + V_2^h + \dots + V_N^h$ and $U^h = U_1^h + U_2^h + \dots + U_N^h$ we have

$$(3.13) \quad Z^h = Z_1^h + Z_2^h + \dots + Z_N^h.$$

On each $Z_i^h, i = 1, \dots, N$, we define the bilinear form $K_i : Z_i^h \times Z_i^h \rightarrow \mathbb{R}$ as a local restriction of K defined in (3.6),

$$(3.14) \quad \begin{aligned} &K_i((y_i, u_i, p_i), (\theta_i, \mu_i, \phi_i)) \\ &= a(\phi_i, p_i) + m(y_i, \phi_i) + \alpha q(u_i, \mu_i) + b(\mu_i, p_i) + a(y_i, \theta_i) + b(u_i, \theta_i). \end{aligned}$$

If $\widehat{\Omega}_i$ is not a control subdomain, the terms involving u_i, μ_i are dropped from (3.14).

In the two-level method, we use two families of meshes to discretize the domain Ω . The coarse-level family is denoted by $\{\mathcal{T}^H\}$ and the fine-level family by $\{\mathcal{T}^h\}$. For the coarse-level triangulation, we first partition Ω into N nonoverlapping elements, denoted as $\Omega_i, i = 1, \dots, N$, and then, if necessary, subdivide $\Omega_i, i = 1, \dots, N$, to obtain a coarse

mesh \mathcal{T}^H (see the right plot in Figure 3.1). The coarse mesh is subdivided further to obtain the fine mesh \mathcal{T}^h . Let H_τ denote the length of the longest side of $\tau \in \mathcal{T}^H$ and define

$$(3.15) \quad H = \max\{H_\tau : \tau \in \mathcal{T}^H\}.$$

We assume that both families $\{\mathcal{T}^H\}$ and $\{\mathcal{T}^h\}$ are quasi-uniform.

The coarse-level finite element space are defined as

$$(3.16) \quad V^H = \{v^H \in V : v^H \in \mathcal{P}_1(\tau) \forall \tau \in \mathcal{T}^H\},$$

$$(3.17) \quad U^H = \left\{u^H \in U : u^H = v^H|_{\mathcal{D}} \text{ for some } v^H \in \widehat{V}^H\right\},$$

where, as before, $\widehat{V}^H = \{v^h \in H^1(\Omega) : v^h \in \mathcal{P}_1(\tau) \forall \tau \in \mathcal{T}^H\}$, \mathcal{P}_1 denotes the space of piecewise linear polynomials, and

$$Z^H \stackrel{\text{def}}{=} V^H \times U^H \times V^H.$$

For convenience, we also define $V_0 \stackrel{\text{def}}{=} V^H$, $U_0 \stackrel{\text{def}}{=} U^H$ and $Z_0 \stackrel{\text{def}}{=} Z^H$. By our construction of the fine and coarse meshes, we have $V^H \subset V^h$ and $U^H \subset U^h$. It is possible to construct non-nested triangulations so that the coarse spaces V^H, U^H are not subsets of V^h, U^h , respectively. In this case, we would need interpolation operators in order to exchange information between meshes ([48, Sec. 2.8].)

The local bilinear form K_0 corresponding to the coarse space is defined as $K_0 = K$. In particular, the local bilinear forms a_0, m_0, q_0 for the coarse grid are given by a, m, q .

We make the following assumptions.

A1 We assume that the local bilinear forms $a_i, m_i, q_i, i = 1, \dots, N$, which are defined as restrictions of the bilinear forms a, m, q , satisfy (2.21).

A2 We assume that the bilinear forms a, m, q satisfy (2.21).

Note that since $a_i, m_i, q_i, i = 1, \dots, N$, are defined as restrictions of the bilinear forms a, m, q , Assumption A2 implies Assumption A1. Assumption A1 is required for the one-level method, while Assumption A2 is required for the two-level method,

Remark 3.1. *In the example problem (1.3) with $\sigma = 0$, we have $V = H^1(\Omega)$ and $a(y, \phi) = \int_{\Omega} \nabla y \nabla \phi dx$. The local spaces (3.9) satisfy*

$$V_i^h \subset \left\{v \in H^1(\widehat{\Omega}_i) : v = 0 \text{ on } \partial\widehat{\Omega}_i \setminus \partial\Omega\right\}.$$

If the relative interior of $\partial\widehat{\Omega}_i \setminus \partial\Omega$ is nonempty, then the Poincare inequality implies that the local bilinear form

$$a_i(y_i, \phi_i) = \int_{\widehat{\Omega}_i} \nabla y_i \nabla \phi_i dx$$

is coercive. The positive semidefiniteness of m_i and the coercivity of q_i follow immediately from the definition of m and q in example problem (1.3). Hence, assumption A1 is satisfied for the example problem (1.3). The bilinear form a , however, violates the assumption A2.

For example problem (1.3) with $\sigma > 0$ and for example problem (1.4) assumptions A1 and A2 are satisfied.

The transformed system (2.24) is now computed for the discretized problem, i.e., in (2.24) we replace Z by Z^h , Z_i by $Z_i^h, i = 1, \dots, N$, and Z_0 by Z^H . That is, for $i = 0, \dots, N$, we define the projections

$$(3.18) \quad T_i : \begin{array}{l} Z^h \rightarrow Z^h, \\ z^h \mapsto z_i^h, \end{array}$$

where $z_i^h \in Z_i^h, i = 1, \dots, N$, is the solution of

$$(3.19) \quad K_i(z_i^h, \psi_i^h) = K(z^h, \psi_i^h) \quad \forall \psi_i^h \in Z_i^h$$

and $z_0^h \in Z^H$ is the solution of

$$(3.20) \quad K(z_0^h, \psi^H) = K(z^h, \psi^H) \quad \forall \psi^H \in Z^H.$$

The right hand sides $r_i \in Z_i^h, i = 1, \dots, N$, and $r_0 \in Z^H$ are defined through

$$(3.21) \quad K_i(r_i, \psi_i^h) = g(\psi_i^h) \quad \forall \psi_i^h \in Z_i^h.$$

and

$$(3.22) \quad K(r_0, \psi^H) = g(\psi^H) \quad \forall \psi^H \in Z^H,$$

respectively.

The one-level additive operator is

$$T = T_1 + \dots + T_N.$$

Under assumption A1, Theorem 2.7 shows that (2.4), (3.6) is equivalent to

$$(3.23) \quad (T_1 + \dots + T_N)z^h = r_1 + \dots + r_N.$$

The two-level additive operator is

$$T = T_0 + T_1 + \dots + T_N.$$

Under assumption A2, Theorem 2.7 shows that (2.4), (3.6) is equivalent to

$$(3.24) \quad (T_0 + T_1 + \dots + T_N)z^h = r_0 + r_1 + \dots + r_N.$$

3.2. Convergence Analysis for the Two-Level Method. In this section we present a convergence analysis for overlapping domain decomposition methods for elliptic linear-quadratic optimal control problems with distributed controls. Our convergence analysis is based on the convergence theory by Cai and Widlund for non-symmetric or indefinite elliptic PDEs. See [10, 12, 13] and [48, Chaper 5]. We split K into a symmetric positive definite part A and the remainder $N = K - A$. The symmetric positive definite part A generates the inner product that is used in GMRES. To derive convergence estimates for the GMRES residual, we need to show that A dominates N . Many technical details in our convergence proof use the fact that K corresponds to the optimality system for an elliptic linear-quadratic optimal control problem and differ substantially from those presented in [10, 12, 13] and [48, Chaper 5].

The bilinear form K defined in (3.6) may be split into two components

$$(3.25) \quad K(z, \psi) = A(z, \psi) + N(z, \psi) \quad \forall z, \psi \in Z$$

with symmetric positive definite part

$$(3.26) \quad A((y, u, p), (\theta, \mu, \phi)) = a(y, \theta) + \alpha q(u, \mu) + a(p, \phi)$$

and

$$(3.27) \quad N((y, u, p), (\theta, \mu, \phi)) = m(y, \phi) + b(\mu, p) + b(u, \theta).$$

It is this splitting that motivates the choice (3.6) over (2.5).

We make the following assumption.

A3 There exist constants $0 < c_a \leq C_a$, $0 < c_q \leq C_q$, $0 < C_m, C_b$, such that

$$(3.28a) \quad c_a \|y\|_{H^1(\Omega)}^2 \leq a(y, y), \quad a(y, \phi) \leq C_a \|y\|_{H^1(\Omega)} \|\phi\|_{H^1(\Omega)},$$

$$(3.28b) \quad c_q \|u\|_{L^2(\Omega)}^2 \leq q(u, u), \quad q(u, \mu) \leq C_q \|u\|_{L^2(\Omega)} \|\mu\|_{L^2(\Omega)},$$

$$(3.28c) \quad 0 \leq m(y, y), \quad m(y, \phi) \leq C_m \|y\|_{L^2(\Omega)} \|\phi\|_{L^2(\Omega)},$$

$$(3.28d) \quad b(u, \phi) \leq C_b \|u\|_{L^2(\Omega)} \|\phi\|_{L^2(\Omega)},$$

for all $y, \phi \in V$ and all $u, \mu \in U$ and

$$(3.28e) \quad b(u_i, \phi_i) \leq C_b \|u_i\|_{L^2(\hat{\Omega}_i)} \|\phi\|_{L^2(\hat{\Omega}_i)},$$

for all $\phi_i \in V_i^h$ and all $u_i \in U_i^h$.

These assumptions are satisfied for example problem (1.4). Assumption (3.28c) excludes problems with boundary controls, such as the example problem (1.3). We will comment on this restriction in the conclusion section.

Since our local bilinear forms $a_i, m_i, q_i, i = 1, \dots, N$, are defined as restrictions of the bilinear forms a, m, q , (3.28) implies Assumptions A1, A2.

To prove convergence of our domain decomposition method, we also need the following regularity assumption.

A4 Suppose that for every $l_1, l_3 \in L^2(\Omega) \times L^2(\Omega)$ the solution $w \in Z$ of the adjoint problem

$$(3.29) \quad K(\psi, w) = \langle (l_1, 0, l_3), \psi \rangle_{L^2(\Omega) \times L^2(\Omega) \times L^2(\Omega)} \quad \forall \psi \in Z$$

satisfies $w \in H^2(\Omega) \times H^1(\Omega) \times H^2(\Omega)$ and that there exists $C > 0$ such that

$$(3.30) \quad \|w\|_{H^2(\Omega) \times H^1(\Omega) \times H^2(\Omega)} \leq C \|(l_1, 0, l_3)\|_M \quad \forall l_1, l_3 \in L^2(\Omega).$$

Recall that (3.6) is nonsymmetric, i.e., A4 is not an assumption on the regularity of the solution of the optimal control problem.

Under assumption (3.28), the bilinear form (3.26) defines an inner product on Z^h and we set

$$\|z\|_A^2 = A(z, z).$$

We apply GMRES with the A -inner product to solve (3.24). The GMRES residuals obey (2.27), where c_T, C_T are defined in (2.26). The following theorem is the main result of this section.

Theorem 3.2. *Let the bilinear forms satisfy A3 and $b = q$, and let $\alpha > 0$ be given. Furthermore, let A4 be satisfied. If GMRES with A -norm defined (3.26) is used to solve (3.24) where T_i and $r_i, i = 0, \dots, N$ are defined in (3.18)–(3.20) and (3.21)–(3.22), respectively, with $K_i = K$ and if H is sufficiently small then $c_T > 0$ and $1 - c_T/C_T$ can be bounded independently of H and h .*

The proof of Theorem 3.2 requires a number of technical lemmas, which will be presented next. To simplify the labeling of non-critical constants throughout this section, we use C without any subscript to denote a generic constant that does not depend on the mesh size h , the subdomain size H , and the regularization parameter α .

We define $M : Z \times Z \rightarrow \mathbb{R}$,

$$(3.31) \quad M((y, u, p), (\theta, \mu, \phi)) = \langle y, \theta \rangle_{L^2(\Omega)} + \langle u, \mu \rangle_{L^2(\Omega)} + \langle p, \phi \rangle_{L^2(\Omega)}$$

and

$$\|z\|_M^2 = M(z, z).$$

By (3.28) there exists $C_A > 0$ (independent of H , h , and α) such that

$$(3.32) \quad \|z\|_M \leq C_A \max\{1, \alpha^{-1}\} \|z\|_A \quad \forall z \in Z.$$

Lemma 3.3. *There exists a constant $C_K > 0$ (independent of H , h , and α) such that*

$$(3.33) \quad |K(z, \psi)| \leq C_K \max\{\alpha, \alpha^{-1}\} \|z\|_A \|\psi\|_A \quad \forall z, \psi \in Z.$$

Proof. The definition (3.6) of K and (3.28) give

$$\begin{aligned} |K(z, \psi)| &= |a(y, \theta) + \alpha q(u, \mu) + a(p, \phi) + m(y, \phi) + b(\mu, p) + b(u, \theta)| \\ &\leq C \left(\|y\|_{H^1} \|\theta\|_{H^1} + \alpha \|u\|_{L^2} \|\mu\|_{L^2} + \|p\|_{H^1} \|\phi\|_{H^1} \right. \\ &\quad \left. + \|y\|_{H^1} \|\phi\|_{H^1} + \|\mu\|_{L^2} \|p\|_{H^1} + \|u\|_{L^2} \|\theta\|_{H^1} \right). \end{aligned}$$

On the other hand, the definition (3.26) of A and (3.28) imply

$$\begin{aligned} \|z\|_A \|\psi\|_A &= \left([a(y, y) + \alpha q(u, u) + a(p, p)] [a(\theta, \theta) + \alpha q(\mu, \mu) + a(\phi, \phi)] \right)^{1/2} \\ &= a(y, y)^{1/2} a(\theta, \theta)^{1/2} + \alpha^{1/2} a(y, y)^{1/2} q(\mu, \mu)^{1/2} + a(y, y)^{1/2} a(\phi, \phi)^{1/2} \\ &\quad + \alpha^{1/2} q(u, u)^{1/2} a(\theta, \theta)^{1/2} + \alpha q(u, u)^{1/2} q(\mu, \mu)^{1/2} + \alpha^{1/2} q(u, u)^{1/2} a(\phi, \phi)^{1/2} \\ &\quad + a(p, p)^{1/2} a(\theta, \theta)^{1/2} + \alpha^{1/2} a(p, p)^{1/2} q(\mu, \mu)^{1/2} + a(p, p)^{1/2} a(\phi, \phi)^{1/2} \\ &\geq C \min\{1, \alpha\} \left(\|y\|_{H^1} \|\theta\|_{H^1} + \|u\|_{L^2} \|\mu\|_{L^2} + \|p\|_{H^1} \|\phi\|_{H^1} \right. \\ &\quad \left. + \|y\|_{H^1} \|\phi\|_{H^1} + \|\mu\|_{L^2} \|p\|_{H^1} + \|u\|_{L^2} \|\theta\|_{H^1} \right). \end{aligned}$$

The desired estimate follow from the previous bounds on $|K(z, \psi)|$ and $\|z\|_A \|\psi\|_A$. \square

The first part of the following lemma corresponds to [12, L. 4], [50, eq. (4.3), L. 7.2]. The constant N_O in the following lemma plays the same role as the spectral radius of the strengthened Cauchy-Schwarz inequality matrix \mathcal{E} that is often used in the analysis of PDE problems ([48, Sec. 5.2].) The second inequality in the following lemma can be proven analogously to the first inequality [44].

Lemma 3.4. *There exists a constant N_O (independent of H , h , and α) such that every decomposition $z^h = \sum_{i=0}^N z_i^h$, $z_i^h \in Z_i^h$, satisfies*

$$\left\| \sum_{i=0}^N z_i^h \right\|_A^2 \leq N_O \sum_{i=0}^N \|z_i^h\|_A^2, \quad \left\| \sum_{i=0}^N z_i^h \right\|_M^2 \leq N_O \sum_{i=0}^N \|z_i^h\|_M^2.$$

The following lemma is proven, e.g., in [21, Sec. 4], [50, Lemma 7.1], [48, Sec. 5.3.1].

Lemma 3.5. *There exists a constant $C_{0a} > 0$ (independent of H , h (and α)) such that for all $v^h \in V^h$, there exists a two-level decomposition $v^h = \sum_{i=0}^N v_i^h$, $v_i^h \in V_i^h$, with*

$$\sum_{i=0}^N a(v_i^h, v_i^h) \leq C_{0a}^2 a(v^h, v^h).$$

The constant $C_{0a} > 0$ is independent of H , h (and α) provided that the amount of overlap δ defined in (3.7) is proportional to H (see (3.8)). Without the latter assumption one can show that the constant C_{0a} in Lemma 3.5 satisfies

$$(3.34) \quad C_{0a}^2 \leq \tilde{C}_{0a}^2 (1 + H/\delta),$$

where \tilde{C}_{0a} is independent of H , h (and α) and δ is the amount of overlap defined in (3.7) (see [23], [48, Thm. 2].)

Lemma 3.6. *There exists a constant $C_{0q} > 0$ (independent of H , h , and α) such that for all $u^h \in U^h$, there exists a two-level decomposition $u^h = \sum_{i=0}^N u_i^h$, $u_i^h \in U_i^h$, with*

$$\sum_{i=0}^N q(u_i^h, u_i^h) \leq C_{0q}^2 q(u^h, u^h).$$

Proof. Let $u^h \in U^h$ be given. We set $u_0^h = 0$ and we let $\tilde{u}_i \in L^2(\Omega_i)$ be such that

$$u^h = \sum_{i=1}^N \tilde{u}_i, \quad \tilde{u}_i(x) = 0, x \notin \bar{\Omega}_i.$$

(One can choose \tilde{u}_i related to $u^h \chi_{\bar{\Omega}_i}$, where $\chi_{\bar{\Omega}_i}$ is the indicator function, but one needs to avoid to include function values of u^h on the boundary $\partial\Omega_i$ more than once.) We define $u_i^h = I_h \tilde{u}_i$. We have

$$u^h = I_h u^h = I_h \sum_{i=1}^N \tilde{u}_i = \sum_{i=1}^N I_h \tilde{u}_i = \sum_{i=1}^N u_i^h.$$

For elements $\tau \in \mathcal{T}_h$ inside Ω_i $u_i^h|_\tau = u^h|_\tau$ and, hence, $\|u_i^h\|_{L^2(\tau)} = \|u^h\|_{L^2(\tau)}$. For elements outside $\bar{\Omega}_i$ that are not adjacent to $\bar{\Omega}_i$, $u_i^h|_\tau = 0$ and, hence, $\|u_i^h\|_{L^2(\tau)} \leq \|u^h\|_{L^2(\tau)}$. For all other elements $\tau \in \mathcal{T}_h$ we have

$$\begin{aligned} \|u_i^h\|_{L^2(\tau)} &\leq |\tau|^{1/2} \|u_i^h\|_{L^\infty(\tau)} \leq |\tau|^{1/2} \|u^h\|_{L^\infty(\tau)} \\ &\leq C |\tau|^{1/2} h^{-d/2} \|u^h\|_{L^2(\tau)} \leq C h^{d/2} h^{-d/2} \|u^h\|_{L^2(\tau)}, \end{aligned}$$

where we have used $|\tau|$ to denote the measure of τ and we have used an inverse inequality [15, Thm. 17.2] to bound the $L^\infty(\tau)$ -norm by the $L^2(\tau)$ -norm. Consequently,

$$\sum_{i=0}^N \|u_i^h\|_{L^2(\tau)}^2 \leq C \|u^h\|_{L^2(\tau)}^2,$$

where we have used the fact that any $\tau \in \mathcal{T}_h$ belongs only to a finite number of overlapping domains $\bar{\Omega}_i$ that is bounded independently of N . Summing over the elements $\tau \in \mathcal{T}_h$ yields

$$\sum_{i=0}^N \|u_i^h\|_{L^2(\Omega)}^2 \leq C \|u^h\|_{L^2(\Omega)}^2,$$

which together with (3.28b) gives the desired result. \square

It is straightforward to extend Lemmas 3.5 and 3.6 to the product space $Z^h = V^h \times U^h \times V^h$ as follows.

Lemma 3.7. *There exists a constant $C_0 > 0$ (independent of H , h , and α) such that for all $z^h \in Z^h$, there exists a two-level decomposition $z^h = \sum_{i=0}^N z_i^h$, $z_i^h \in Z_i^h$, with*

$$\sum_{i=0}^N A(z_i^h, z_i^h) \leq C_0^2 A(z^h, z^h).$$

The next lemma establishes properties of the coarse grid projection. Its proof requires discretization error estimates for optimal control problems. These are briefly summarized in Section 5. We also need the following assumption on the regularity of the adjoint problem as stated in Assumption A4.

Lemma 3.8. *Assume that A4 holds and that $z^h = (y^h, u^h, p^h) \in Z^h$. There exist $C_1, C_2 > 0$ independent of H and h , such that the solution $z^H = (y^H, u^H, p^H) \in Z^H$ of*

$$(3.35) \quad K(z^H, \psi^H) = K(z^h, \psi^H) \quad \forall \psi^H \in Z^H$$

obeys

$$(3.36) \quad \|z^H\|_A \leq C_1 \|z^h\|_A,$$

and

$$(3.37) \quad \|y^h - y^H\|_{L^2(\Omega)} + \|p^h - p^H\|_{L^2(\Omega)} \leq C_2 \max\{\alpha, \alpha^{-1}\} H \|z^h\|_A.$$

Before we present the proof, we remark that by definition (3.18), (3.20) of T_0 , we have

$$z^H = T_0 z^h,$$

where z^h, z^H are given as in the previous lemma. Hence Lemma 3.8 provides estimates for the coarse grid operator.

Proof of Lemma 3.8. First we note that (3.35) has a unique solution. In fact, in the proof of Lemma 2.4 we have shown that (3.35) corresponds to a coarse grid optimal control problem of the type (2.17). By Theorem 2.2 applied to the coarse grid optimal control problem, the coarse grid optimal control problem has a unique solution. Moreover, the coarse grid solution obeys (3.36) (see (5.5)).

For brevity we define $\tilde{z}^h = (y^h, 0, p^h)$ and $\tilde{z}^H = (y^H, 0, p^H)$. We consider the adjoint problems

$$(3.38a) \quad K(\psi, w) = \langle \tilde{z}^h - \tilde{z}^H, \psi \rangle_{L^2(\Omega) \times L^2(\Omega) \times L^2(\Omega)} \quad \forall \psi \in Z,$$

$$(3.38b) \quad K(\psi^h, w^h) = \langle \tilde{z}^h - \tilde{z}^H, \psi^h \rangle_{L^2(\Omega) \times L^2(\Omega) \times L^2(\Omega)} \quad \forall \psi^h \in Z^h,$$

$$(3.38c) \quad K(\psi^H, w^H) = \langle \tilde{z}^h - \tilde{z}^H, \psi^H \rangle_{L^2(\Omega) \times L^2(\Omega) \times L^2(\Omega)} \quad \forall \psi^H \in Z^H.$$

If we set $\psi^h = z^h - z^H$ in (3.38b) and note that $\langle \tilde{z}^h - \tilde{z}^H, z^h - z^H \rangle_{L^2(\Omega) \times L^2(\Omega) \times L^2(\Omega)} = \|\tilde{z}^h - \tilde{z}^H\|_{L^2(\Omega) \times L^2(\Omega) \times L^2(\Omega)}^2$, then

$$(3.39) \quad K(z^h - z^H, w^h) = \|y^h - y^H\|_{L^2(\Omega)}^2 + \|p^h - p^H\|_{L^2(\Omega)}^2.$$

Equation (3.35) with $\psi^H = w^H \in Z^H$ yields

$$(3.40) \quad K(z^h - z^H, w^H) = 0.$$

Subtracting (3.39) and (3.40), we have

$$(3.41) \quad K(\tilde{z}^h - \tilde{z}^H, w^h - w^H) = \|y^h - y^H\|_{L^2(\Omega)}^2 + \|p^h - p^H\|_{L^2(\Omega)}^2.$$

We now use Lemma 3.3 to bound $K(\cdot, \cdot)$

$$(3.42) \quad \begin{aligned} & \|y^h - y^H\|_{L^2(\Omega)}^2 + \|p^h - p^H\|_{L^2(\Omega)}^2 \\ & \leq C_K \max\{\alpha, \alpha^{-1}\} \|z^h - z^H\|_A \|w^h - w^H\|_A. \end{aligned}$$

Let w be the solution of (3.38a). Standard error estimates for linear quadratic optimal control problems (see Theorem 5.3) give

$$\|w^h - w\|_A \leq Ch \|w\|_{H^2(\Omega) \times H^1(\Omega) \times H^2(\Omega)},$$

$$\|w^H - w\|_A \leq CH \|w\|_{H^2(\Omega) \times H^1(\Omega) \times H^2(\Omega)}.$$

Using the previous two inequalities and assumption A4, we obtain

$$\begin{aligned} \|w^h - w^H\|_A & \leq \|w^h - w\|_A + \|w^H - w\|_A \leq CH \|w\|_{H^2(\Omega) \times H^1(\Omega) \times H^2(\Omega)} \\ & \leq CH (\|y^h - y^H\|_{L^2(\Omega)} + \|p^h - p^H\|_{L^2(\Omega)}). \end{aligned}$$

Hence, (3.42) implies

$$\begin{aligned} & \|y^h - y^H\|_{L^2(\Omega)}^2 + \|p^h - p^H\|_{L^2(\Omega)}^2 \\ & \leq C_K C \max\{\alpha, \alpha^{-1}\} H \|z^h - z^H\|_A (\|y^h - y^H\|_{L^2(\Omega)} + \|p^h - p^H\|_{L^2(\Omega)}) \end{aligned}$$

which, together with (3.36), gives (3.37). \square

Lemma 3.9. *If each overlapping region $\widehat{\Omega}_i, i = 1, \dots, N$, has diameter less than or equal to $C_\delta H$, then there exists $C > 0$ independent of H and h such that*

$$\|v_i^h\|_{L^2(\widehat{\Omega}_i)} \leq C C_\delta H \|v_i^h\|_{H^1(\widehat{\Omega}_i)} \quad \forall v_i^h \in V_i^h, \quad i = 1, \dots, N.$$

Proof. Since v_i^h is zero outside of $\widehat{\Omega}_i$, by a Poincaré inequality (e.g., [29, eqn. 7.44])

$$\|v_i^h\|_{L^2(\widehat{\Omega}_i)} \leq C |\widehat{\Omega}_i|^{1/d} \|\nabla v_i^h\|_{L^2(\widehat{\Omega}_i)} \leq C C_\delta H \|\nabla v_i^h\|_{L^2(\widehat{\Omega}_i)}$$

where $|\widehat{\Omega}_i|$ denotes the measure of $\widehat{\Omega}_i \subset \mathbb{R}^d$. This implies the desired inequality. \square

The following lemma says that the bilinear form K is positive definite on the subspace $Z_i^h, i = 1, \dots, N$, if H is sufficiently small.

Lemma 3.10. *Assume that each overlapping region $\widehat{\Omega}_i, i = 1, \dots, N$ has diameter less than or equal to $C_\delta H$. There exist constants $H_0 > 0, C_{H_0} > 0$ such that if $H \leq H_0$, then*

$$(3.43) \quad K(z_i^h, z_i^h) \geq C_{H_0} A(z_i^h, z_i^h) \quad \forall z_i^h \in Z_i^h, \quad i = 1, \dots, N.$$

Proof. Let $z_i^h = (y_i^h, u_i^h, p_i) \in Z_i^h$. The definition (3.6) of K , (3.28) and the first inequality in Lemma 3.9 imply

$$\begin{aligned} K(z_i^h, z_i^h) &= a(y_i^h, y_i^h) + \alpha q(u_i^h, u_i^h) + a(p_i^h, p_i^h) + m(y_i^h, p_i^h) + b(u_i^h, p_i^h) + b(u_i^h, y_i^h), \\ &\geq a(y_i^h, y_i^h) + \alpha q(u_i^h, u_i^h) + a(p_i^h, p_i^h) \\ &\quad - C \left(\|y_i^h\|_{L^2(\widehat{\Omega}_i)} \|p_i^h\|_{L^2(\widehat{\Omega}_i)} + \|u_i^h\|_{L^2(\widehat{\Omega}_i)} \|p_i^h\|_{L^2(\widehat{\Omega}_i)} + \|u_i^h\|_{L^2(\widehat{\Omega}_i)} \|y_i^h\|_{L^2(\widehat{\Omega}_i)} \right) \\ &\geq a(y_i^h, y_i^h) + \alpha q(u_i^h, u_i^h) + a(p_i^h, p_i^h) \\ &\quad - C \left(\frac{1}{2} \|y_i^h\|_{L^2(\widehat{\Omega}_i)}^2 + \frac{1}{2} \|p_i^h\|_{L^2(\widehat{\Omega}_i)}^2 + \frac{\alpha c_q}{4C} \|u_i^h\|_{L^2(\widehat{\Omega}_i)}^2 + \frac{C}{\alpha c_q} \|p_i^h\|_{L^2(\widehat{\Omega}_i)}^2 \right. \\ &\quad \left. + \frac{\alpha c_q}{4C} \|u_i^h\|_{L^2(\widehat{\Omega}_i)}^2 + \frac{C}{\alpha c_q} \|y_i^h\|_{L^2(\widehat{\Omega}_i)}^2 \right) \\ &\geq a(y_i^h, y_i^h) - C(1 + \alpha^{-1}) C_\delta H \|y_i^h\|_{H^1(\widehat{\Omega}_i)}^2 + \frac{\alpha}{2} q(u_i^h, u_i^h) \\ &\quad + a(p_i^h, p_i^h) - C(1 + \alpha^{-1}) C_\delta H \|p_i^h\|_{H^1(\widehat{\Omega}_i)}^2. \end{aligned}$$

The assertion now follows from the definition (3.26) of A and (3.28). \square

Lemma 3.11. *Let H_0 be defined as in Lemma 3.10. There exists a constant $C_{H_1} > 0$ such that if $H \leq H_0$ then*

$$\sum_{i=0}^N A(T_i z^h, T_i z^h) \leq C_{H_1} A(z^h, z^h) \quad \forall z^h \in Z^h.$$

Proof. Recall, that by definition (3.18), (3.20) of T_0 , the function z^H defined by (3.35) satisfies $z^H = T_0 z^h$. Hence, by inequality (3.36),

$$A(T_0 z^h, T_0 z^h) \leq C_1^2 A(z^h, z^h).$$

Let $H < H_0$. Lemma 3.10 and the definition (3.18), (3.19) of T_i imply

$$A(T_i z^h, T_i z^h) \leq C_{H_0}^{-1} K(T_i z^h, T_i z^h) = C_{H_0}^{-1} K(z^h, T_i z^h) \quad \forall z^h \in Z^h,$$

$i = 1, \dots, N$. We sum over $i = 1, \dots, N$, then use Lemma 3.3 and Lemma 3.4

$$\begin{aligned} \sum_{i=1}^N A(T_i z^h, T_i z^h) &\leq C_{H_0}^{-1} K(z^h, \sum_{i=1}^N T_i z^h) \\ &\leq C_{H_0}^{-1} C_K \max\{\alpha, \alpha^{-1}\} \|z^h\|_A \left\| \sum_{i=1}^N T_i z^h \right\|_A \\ &\leq C_{H_0}^{-1} C_K \max\{\alpha, \alpha^{-1}\} A(z^h, z^h)^{1/2} N_O^{1/2} \left(\sum_{i=1}^N A(T_i z^h, T_i z^h) \right)^{1/2} \end{aligned}$$

This implies

$$\sum_{i=1}^N A(T_i z^h, T_i z^h) \leq C_{H_0}^{-2} C_K^2 \max\{\alpha, \alpha^{-1}\}^2 N_O A(z^h, z^h).$$

The desired inequality follows if we set $C_{H_1} = C_{H_0}^{-2} C_K^2 \max\{\alpha, \alpha^{-1}\}^2 N_O + C_1^2$. \square

Lemma 3.12. *There exist constants $H_2 > 0$, $C_{H_2} > 0$ such that if $H \leq H_2$ then*

$$\sum_{i=0}^N A(T_i z^h, T_i z^h) \geq C_{H_2} A(z^h, z^h) \quad \forall z^h \in Z^h.$$

Proof. Let $z^h \in Z^h$ be arbitrary. By Lemma 3.7 there exists a representation $z^h = \sum_{i=0}^N z_i^h$, $z_i^h \in Z_i^h$ with $(\sum_{i=0}^N \|z_i^h\|_A^2)^{1/2} \leq C_0 \|z^h\|_A$. We derive an upper bound for $K(z^h, z^h)$ by using this decomposition of z^h , the definition (3.18) of T_i and Lemma 3.3,

$$K(z^h, z^h) = \sum_{i=0}^N K(z^h, z_i^h) = \sum_{i=0}^N K(T_i z^h, z_i^h) \leq C_K \max\{\alpha, \alpha^{-1}\} \sum_{i=0}^N \|T_i z^h\|_A \|z_i^h\|_A.$$

Applying the Cauchy-Schwarz inequality and Lemma 3.7,

$$\begin{aligned} K(z^h, z^h) &\leq C_K \max\{\alpha, \alpha^{-1}\} \left(\sum_{i=0}^N \|T_i z^h\|_A^2 \right)^{1/2} \left(\sum_{i=0}^N \|z_i^h\|_A^2 \right)^{1/2} \\ (3.44) \quad &\leq C_K C_0 \max\{\alpha, \alpha^{-1}\} \left(\sum_{i=0}^N \|T_i z^h\|_A^2 \right)^{1/2} \|z^h\|_A. \end{aligned}$$

To derive a lower bound for $K(z^h, z^h)$ we proceed as in the proof of Lemma 3.10 to obtain

$$\begin{aligned} K(z^h, z^h) &\geq a(y^h, y^h) + \frac{\alpha}{2} q(u^h, u^h) + a(p^h, p^h) \\ (3.45) \quad &\quad - C \left((1 + \alpha^{-1}) \|y^h\|_{L^2(\Omega)}^2 + (1 + \alpha^{-1}) \|p^h\|_{L^2(\Omega)}^2 \right). \end{aligned}$$

Let $z^H = (y^H, u^H, p^H)$ be defined by (3.35) and recall, that by definition (3.18), (3.20) of T_0 , we have $z^H = T_0 z^h$. Applying the triangle inequality and using Lemma 3.8, (3.32),

we obtain the lower bound

$$\begin{aligned}
K(z^h, z^h) &\geq a(y^h, y^h) + \frac{\alpha}{2}q(u^h, u^h) + a(p^h, p^h) \\
&\quad - C\left((1 + \alpha^{-1})\|y^h - y^H\|_{L^2(\Omega)}^2 + (1 + \alpha^{-1})\|p^h - p^H\|_{L^2(\Omega)}^2\right) \\
&\quad - C(1 + \alpha^{-1})\|z^H\|_M^2 \\
&\geq a(y^h, y^h) + \frac{\alpha}{2}q(u^h, u^h) + a(p^h, p^h) \\
(3.46) \quad &\geq C \max\{\alpha, \alpha^{-1}\}^2 (1 + \alpha^{-1}) H^2 \|z^h\|_A^2 - C \max\{1, \alpha^{-1}\} (1 + \alpha^{-1}) \|z^h\|_A \|T_0 z^h\|_A.
\end{aligned}$$

Combining the upper bound (3.44) and the lower bound (3.46), we have

$$\begin{aligned}
(3.47) \quad &\left(C_K C_0 \max\{\alpha, \alpha^{-1}\} + C \max\{1, \alpha^{-1}\} (1 + \alpha^{-1})\right) \left(\sum_{i=0}^N \|T_i z^h\|_A^2\right)^{1/2} \|z^h\|_A \\
&\geq (1/2 - C \max\{1, \alpha\}^2 (1 + \alpha^{-1}) H^2) \|z^h\|_A^2.
\end{aligned}$$

If H_2 is chosen such that $C \max\{1, \alpha\}^2 (1 + \alpha^{-1}) H_2^2 < \min\{1, \alpha/2\}$, then (3.47) implies the desired result with

$$C_{H_2} = \frac{\left(1/2 - C \max\{1, \alpha\}^2 (1 + \alpha^{-1}) H_2^2\right)^2}{\left(C_K C_0 \max\{\alpha, \alpha^{-1}\} + C \max\{1, \alpha^{-1}\} (1 + \alpha^{-1})\right)^2}.$$

□

The following Lemma bounds the contribution by the local components ($i > 0$) to the nonsymmetric part of $K(\cdot, \cdot)$.

Lemma 3.13. *Let H_0 be defined as in Lemma 3.11. There exists a constant $C_{H_3} > 0$ such that if $H \leq H_0$ then*

$$(3.48) \quad \left| \sum_{i=1}^N N(T_i z^h - z^h, T_i z^h) \right| \leq C_{H_3} \max\{1, \alpha^{-1}\} H A(z^h, z^h) \quad \forall z^h \in Z^h.$$

Proof. Let $T_i z^h = (y_i^h, u_i^h, p_i^h)$. We note that

$$(3.49) \quad \left| \sum_{i=1}^N N(T_i z^h - z^h, T_i z^h) \right| \leq \left| \sum_{i=1}^N N(T_i z^h, T_i z^h) \right| + \left| N(z^h, \sum_{i=1}^N T_i z^h) \right|.$$

The definition (3.27) of N and (3.28) imply

$$\begin{aligned}
&|N(T_i z^h, T_i z^h)| \\
&= |m(y_i^h, p_i^h) + b(u_i^h, p_i^h) + b(u_i^h, y_i^h)| \\
&\leq C \left(\|y_i^h\|_{L^2(\hat{\Omega}_i)} \|p_i^h\|_{L^2(\hat{\Omega}_i)} + \|p_i^h\|_{L^2(\hat{\Omega}_i)} \|u_i^h\|_{L^2(\hat{\Omega}_i)} + \|u_i^h\|_{L^2(\hat{\Omega}_i)} \|y_i^h\|_{L^2(\hat{\Omega}_i)} \right) \\
&\leq C \left(\|y_i^h\|_{L^2(\hat{\Omega}_i)}^2 + \|p_i^h\|_{L^2(\hat{\Omega}_i)}^2 + H^{-1} \|y_i^h\|_{L^2(\hat{\Omega}_i)}^2 + H^{-1} \|p_i^h\|_{L^2(\hat{\Omega}_i)}^2 + H \|u_i^h\|_{L^2(\hat{\Omega}_i)}^2 \right).
\end{aligned}$$

With Lemma 3.9 and (3.28) we obtain

$$\begin{aligned}
|N(T_i z^h, T_i z^h)| &\leq C C_\delta^2 H \left(\|y_i^h\|_{H^1(\hat{\Omega}_i)}^2 + \|p_i^h\|_{H^1(\hat{\Omega}_i)}^2 + \|u_i^h\|_{L^2(\hat{\Omega}_i)}^2 \right) \\
&\leq C C_\delta^2 \max\{1, \alpha^{-1}\} H A(T_i z^h, T_i z^h).
\end{aligned}$$

Lemma 3.11 now implies

$$(3.50) \quad \left| \sum_{i=1}^N N(T_i z^h, T_i z^h) \right| \leq CC_\delta^2 C_{H_1} \max\{1, \alpha^{-1}\} HA(z^h, z^h),$$

provided $H \leq H_0$.

Let $z^h = (y^h, u^h, p^h)$ and $T_i z^h = (y_i^h, u_i^h, p_i^h)$. The definition (3.27) of N implies

$$(3.51) \quad N(z^h, \sum_{i=1}^N T_i z^h) = m(y^h, \sum_{i=1}^N p_i^h) + b(\sum_{i=1}^N u_i^h, p^h) + b(u^h, \sum_{i=1}^N y_i^h).$$

Assumption (3.28) and Lemmas 3.4, 3.9 imply

$$(3.52) \quad \begin{aligned} m(y^h, \sum_{i=1}^N p_i^h) &\leq C \|y^h\|_{L^2(\Omega)} \left\| \sum_{i=1}^N p_i^h \right\|_{L^2(\Omega)}, \\ &\leq CN_O^{1/2} \|y^h\|_{L^2(\Omega)} \left(\sum_{i=1}^N \|p_i^h\|_{L^2(\widehat{\Omega}_i)}^2 \right)^{1/2} \\ &\leq CN_O^{1/2} C_\delta H \|y^h\|_{L^2(\Omega)} \left(\sum_{i=1}^N \|p_i^h\|_{H^1(\widehat{\Omega}_i)}^2 \right)^{1/2}. \end{aligned}$$

Similarly,

$$(3.53) \quad \begin{aligned} b(u^h, \sum_{i=1}^N y_i^h) &\leq C \|u^h\|_{L^2(\Omega)} \left\| \sum_{i=1}^N y_i^h \right\|_{L^2(\Omega)}, \\ &\leq CN_O^{1/2} \|u^h\|_{L^2(\Omega)} \left(\sum_{i=1}^N \|y_i^h\|_{L^2(\widehat{\Omega}_i)}^2 \right)^{1/2} \\ &\leq CN_O^{1/2} C_\delta H \|u^h\|_{L^2(\Omega)} \left(\sum_{i=1}^N \|y_i^h\|_{H^1(\widehat{\Omega}_i)}^2 \right)^{1/2}. \end{aligned}$$

The estimate of the term $b(\sum_{i=1}^N u_i^h, p^h)$ is little more involved. Let $\tilde{p}_i^h \in V_i^h$ be the $L^2(\widehat{\Omega}_i)$ projection of p^h , i.e.,

$$\int_{\widehat{\Omega}_i} \tilde{p}_i^h \phi_i^h = \int_{\widehat{\Omega}_i} p^h \phi_i^h \quad \forall \phi_i^h \in V_i^h.$$

It is known, see, e.g., [8], [39, L. 1.8.1], that

$$(3.54) \quad \|\tilde{p}_i^h - p_i^h\|_{L^2(\Omega)} \leq Ch \|p_i^h\|_{H^1(\Omega)}$$

and

$$(3.55) \quad \|\tilde{p}_i^h\|_{H^1(\Omega)} \leq C \|p_i^h\|_{H^1(\Omega)}.$$

Using assumption (3.28), (3.54), $h \leq H$, Lemma 3.9 and (3.55), we obtain

$$\begin{aligned}
b\left(\sum_{i=1}^N u_i^h, p^h\right) &= \sum_{i=1}^N b(u_i^h, p^h - \tilde{p}_i^h) + b(u_i^h, \tilde{p}_i^h), \\
&\leq C \sum_{i=1}^N \|u_i^h\|_{L^2(\widehat{\Omega}_i)} (\|p^h - \tilde{p}_i^h\|_{L^2(\widehat{\Omega}_i)} + \|\tilde{p}_i^h\|_{L^2(\widehat{\Omega}_i)}) \\
&\leq CH \sum_{i=1}^N \|u_i^h\|_{L^2(\widehat{\Omega}_i)} (\|p^h\|_{H^1(\widehat{\Omega}_i)} + \|\tilde{p}_i^h\|_{H^1(\widehat{\Omega}_i)}) \\
&\leq CH \sum_{i=1}^N \|u_i^h\|_{L^2(\widehat{\Omega}_i)} \|p^h\|_{H^1(\widehat{\Omega}_i)} \\
&\leq CH \left(\sum_{i=1}^N \|u_i^h\|_{L^2(\widehat{\Omega}_i)}^2 \right)^{1/2} \left(\sum_{i=1}^N \|p^h\|_{H^1(\widehat{\Omega}_i)}^2 \right)^{1/2}.
\end{aligned}$$

Since the number of subdomains that overlap a given subdomain $\widehat{\Omega}_i$ is bounded independently of N , we have

$$\sum_{i=1}^N \|p^h\|_{H^1(\widehat{\Omega}_i)}^2 \leq C \|p^h\|_{H^1(\Omega)}^2.$$

Hence, we obtain

$$(3.56) \quad b\left(\sum_{i=1}^N u_i^h, p^h\right) \leq CH \left(\sum_{i=1}^N \|u_i^h\|_{L^2(\widehat{\Omega}_i)}^2 \right)^{1/2} \|p^h\|_{H^1(\Omega)}^2.$$

If we combine (3.51)–(3.56), recall that $z^h = (y^h, u^h, p^h)$, $T_i z^h = (y_i^h, u_i^h, p_i^h)$, and use Lemma (3.11), then

$$\begin{aligned}
&N(z^h, \sum_{i=1}^N T_i z^h) \\
&\leq CH \left(\|y^h\|_{L^2(\Omega)} \left(\sum_{i=1}^N \|p_i^h\|_{H^1(\widehat{\Omega}_i)}^2 \right)^{1/2} \right. \\
&\quad \left. + \|u^h\|_{L^2(\Omega)} \left(\sum_{i=1}^N \|y_i^h\|_{H^1(\widehat{\Omega}_i)}^2 \right)^{1/2} + \|p^h\|_{H^1(\Omega)}^2 \left(\sum_{i=1}^N \|u_i^h\|_{L^2(\widehat{\Omega}_i)}^2 \right)^{1/2} \right) \\
&\leq C \max\{1, \alpha^{-1}\} H \|z^h\|_A \left(\sum_{i=1}^N A(T_i z^h, T_i z^h) \right)^{1/2} \\
(3.57) &\leq C \max\{1, \alpha^{-1}\} H \|z^h\|_A^2,
\end{aligned}$$

provided $H < H_0$.

The desired estimate now follows from (3.49), (3.50), (3.57). \square

The contribution from the coarse component to the nonsymmetric part can be bounded similarly.

Lemma 3.14. *Assume that $b = q$. There exists $C > 0$ such that*

$$|N(T_0 z^h - z^h, T_0 z^h)| \leq C \max\{\alpha, \alpha^{-2}\} H A(z^h, z^h) \quad \forall z^h \in Z^h.$$

Proof. Let $z^h = (y^h, u^h, p^h)$ and $T_0 z^h = z^H = (y^H, u^H, p^H)$. By definition (3.18), (3.20) of T_0 ,

$$K(z^H, \psi^H) = K(z^h, \psi^H) \quad \forall \psi^H \in Z^H.$$

Inserting the definition (3.6) of K , into the previous identity, we obtain

$$\alpha q(u^H, \mu^H) + b(\mu^H, p^H) = \alpha q(u^h, \mu^H) + b(\mu^H, p^h) \quad \forall \mu^H \in U^H.$$

With $b = q$ and $\mu^H = y^H$, this implies

$$(3.58) \quad b(u^H - u^h, y^H) = \alpha^{-1} b(y^H, p^h - p^H).$$

The definition (3.27) of N and (3.58) imply

$$\begin{aligned} & |N(T_0 z^h - z^h, T_0 z^h)| \\ &= |m(y^H - y^h, p^H) + b(u^H, p^H - p^h) + b(u^H - u^h, y^H)| \\ &= |m(y^H - y^h, p^H) + b(u^H, p^H - p^h) + \alpha^{-1} b(y^H, p^h - p^H)| \\ &\leq C \max\{1, \alpha^{-1}\} \left(\|y^H - y^h\|_{L^2(\Omega)} \|p^H\|_{L^2(\Omega)} \right. \\ &\quad \left. + \|p^H - p^h\|_{L^2(\Omega)} \|u^H\|_{L^2(\Omega)} + \|p^H - p^h\|_{L^2(\Omega)} \|y^H\|_{L^2(\Omega)} \right). \end{aligned}$$

We use Lemma 3.8 to obtain

$$|N(T_0 z^h - z, T_0 z^h)| \leq C \max\{1, \alpha^{-1}\} \max\{\alpha, \alpha^{-1}\} H \|z^h\|_A^2.$$

□

Now we are able to prove our main convergence result, Theorem 3.2.

Proof of Theorem 3.2. We show that both c_T (the minimal eigenvalue of the Hermitian part of T) and C_T (the norm of T) can be bounded independently of H and h for H sufficiently small.

First we provide a bound for c_T . From the definition of T_i in (3.18),

$$A(T_i z^h, \psi_i^h) + N(T_i z^h, \psi_i^h) = K(T_i z^h, \psi_i^h) = K(z^h, \psi_i^h) = A(z^h, \psi_i^h) + N(z^h, \psi_i^h)$$

for all $\psi_i^h \in Z_i^h$. Setting $\psi_i^h = T_i z^h$, we have $A(T_i z^h, T_i z^h) + N(T_i z^h, T_i z^h) = A(z^h, T_i z^h) + N(z^h, T_i z^h)$, or

$$A(T_i z^h, z^h) = A(T_i z^h, T_i z^h) + N(T_i z^h - z^h, T_i z^h).$$

Summing up over $i = 0, \dots, N$, we have

$$\begin{aligned} \sum_{i=0}^N A(T_i z^h, z^h) &= \sum_{i=0}^N A(T_i z^h, T_i z^h) + \sum_{i=0}^N N(T_i z^h - z^h, T_i z^h), \\ (3.59) \quad &\geq \sum_{i=0}^N A(T_i z^h, T_i z^h) - \left| \sum_{i=0}^N N(T_i z^h - z^h, T_i z^h) \right|. \end{aligned}$$

We let $H < \min\{H_0, H_2\}$ and then bound the first term on the right from below using Lemma 3.12

$$\sum_{i=0}^N A(T_i z^h, T_i z^h) \geq C_{H_2} A(z^h, z^h),$$

and bound the second term on the right from above by Lemmas 3.13 and 3.14

$$\begin{aligned} \left| \sum_{i=1}^N N(T_i z^h - z^h, T_i z^h) \right| &\leq C_{H_3} \max\{1, \alpha^{-1}\} H A(z^h, z^h), \\ |N(T_0 z^h - z^h, T_0 z^h)| &\leq C \max\{\alpha, \alpha^{-2}\} H A(z^h, z^h). \end{aligned}$$

Therefore,

$$(3.60) \quad A(Tz^h, z^h) \geq (C_{H_2} - C_{H_3} \max\{1, \alpha^{-1}\} H - C \max\{\alpha, \alpha^{-2}\} H) A(z^h, z^h)$$

where the constants are independent of H or h . For H sufficiently small, $C_{H_2} - C_{H_3} \max\{1, \alpha^{-1}\} H - C \max\{\alpha, \alpha^{-2}\} H$ is positive, so that

$$c_T = \inf_{z \neq 0} \frac{A(Tz, z)}{A(z, z)} > 0,$$

and remains bounded away from zero if H is decreased further.

To bound the norm of T , we use Lemma 3.4 and Lemma 3.11.

$$\begin{aligned} C_T^2 &= \sup_{z \neq 0} \frac{A(Tz, Tz)}{A(z, z)} \\ &\leq \sup_{z \neq 0} \frac{N_O \sum_{i=0}^N A(T_i z, T_i z)}{A(z, z)} \\ &\leq N_O C_{H_1}, \end{aligned}$$

with N_O and C_{H_1} being bounded independently of H and h . □

3.3. Algebraic Viewpoint. The overlapping methods for optimal control may be formulated at the algebraic level, which is useful for explaining many implementation issues.

Let $\{\phi_j\}_{j=1}^n, \{\mu_j\}_{j=1}^m$ be bases of V^h and U^h , respectively. The discretized states, controls and adjoints can be written as

$$y^h(x) = \sum_{j=1}^n y_j \phi_j(x), \quad u^h(x) = \sum_{j=1}^m u_j \mu_j(x), \quad p^h(x) = \sum_{j=1}^n p_j \phi_j(x),$$

respectively. We use bold face for the corresponding vectors of coefficients representing y^h, u^h, p^h , e.g., $\mathbf{y} = (y_1, \dots, y_n)^T$.

If we define matrices $\mathbf{A}, \mathbf{M} \in \mathbb{R}^{n \times n}$, $\mathbf{B} \in \mathbb{R}^{n \times m}$, $\mathbf{Q} \in \mathbb{R}^{m \times m}$ and vectors $\mathbf{b}, \mathbf{c} \in \mathbb{R}^n$, $\mathbf{d} \in \mathbb{R}^m$ with entries

$$\begin{aligned} \mathbf{A}_{jk} &= a(\phi_k, \phi_j), & \mathbf{M}_{jk} &= m(\phi_k, \phi_j), & \mathbf{B}_{jk} &= b(\mu_k, \phi_j), & \mathbf{Q}_{jk} &= q(\mu_k, \mu_j), \\ & & \mathbf{b}_j &= f(\phi_j), & \mathbf{c}_j &= c(\phi_j), & \mathbf{d}_j &= d(\mu_j), \end{aligned}$$

then the equivalent algebraic formulation of (3.4) is

$$(3.61a) \quad \min \quad \frac{1}{2} \mathbf{y}^T \mathbf{M} \mathbf{y} + \frac{\alpha}{2} \mathbf{u}^T \mathbf{Q} \mathbf{u} - \mathbf{c}^T \mathbf{y} - \mathbf{d}^T \mathbf{u},$$

$$(3.61b) \quad \text{s.t.} \quad \mathbf{A} \mathbf{y} + \mathbf{B} \mathbf{u} = \mathbf{b}.$$

If there exists a feasible point and if the Hessian of the objective function is positive definite on the null-space of the constraints (both assumptions are satisfied for a discretization of the example problems (1.3) and (1.3) using conforming piecewise linear elements as described

as described in Section 3.1 (see also [37]), then the necessary and sufficient optimality conditions for (3.61) are given by

$$(3.62) \quad \begin{pmatrix} \mathbf{A} & \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \alpha\mathbf{Q} & \mathbf{B}^T \\ \mathbf{M} & \mathbf{0} & \mathbf{A}^T \end{pmatrix} \begin{pmatrix} \mathbf{y} \\ \mathbf{u} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} \mathbf{b} \\ \mathbf{d} \\ \mathbf{c} \end{pmatrix}.$$

The system (3.62) is the matrix representation of (3.5), (3.6). We also use the notation

$$(3.63) \quad \mathbf{Kz} = \mathbf{g}$$

instead of (3.62). Note that

$$(3.64) \quad \begin{pmatrix} \mathbf{A} & \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \alpha\mathbf{Q} & \mathbf{B}^T \\ \mathbf{M} & \mathbf{0} & \mathbf{A}^T \end{pmatrix} = \begin{pmatrix} \mathbf{A} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \alpha\mathbf{Q} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{A}^T \end{pmatrix} + \begin{pmatrix} \mathbf{0} & \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{B}^T \\ \mathbf{M} & \mathbf{0} & \mathbf{0} \end{pmatrix}$$

corresponds to the splitting (3.25).

Let

$$(3.65) \quad \mathbf{K}_i = \begin{pmatrix} \mathbf{A}_i & \mathbf{B}_i & \mathbf{0} \\ \mathbf{0} & \alpha\mathbf{Q}_i & \mathbf{B}_i^T \\ \mathbf{M}_i & \mathbf{0} & \mathbf{A}_i^T \end{pmatrix} \in \mathbb{R}^{(n_i+m_i+n_i) \times (n_i+m_i+n_i)}, \quad i = 1, \dots, N,$$

be the submatrix of (3.62) associated with the extended subdomain $\widehat{\Omega}_i$, $i = 1, \dots, N$, i.e., \mathbf{K}_i is the KKT matrix associated with the space Z_i^h . Furthermore, let

$$(3.66) \quad \mathbf{K}_0 = \begin{pmatrix} \mathbf{A}_0 & \mathbf{B}_0 & \mathbf{0} \\ \mathbf{0} & \alpha\mathbf{Q}_0 & \mathbf{B}_0^T \\ \mathbf{M}_0 & \mathbf{0} & \mathbf{A}_0^T \end{pmatrix} \in \mathbb{R}^{(n_0+m_0+n_0) \times (n_0+m_0+n_0)}$$

be the KKT matrix associated with the coarse space Z^H . By

$$\mathbf{R}_i = \begin{pmatrix} \mathbf{R}_i^y & & \\ & \mathbf{R}_i^u & \\ & & \mathbf{R}_i^p \end{pmatrix} \in \mathbb{R}^{(n_i+m_i+n_i) \times (n+m+n)}, \quad i = 1, \dots, N,$$

we denote the restriction matrix that maps the global $(\mathbf{y}, \mathbf{u}, \mathbf{p})$ vector to the local vector $(\mathbf{y}_i, \mathbf{u}_i, \mathbf{p}_i)$ corresponding to the extended subdomain $\widehat{\Omega}_i$. In the case $i = 0$, $\mathbf{R}_0 \in \mathbb{R}^{(n_0+m_0+n_0) \times (n+m+n)}$ is the interpolation operator from the coarse grid to the fine grid. The block diagonals of \mathbf{R}_0 are computed as described in [48, p. 62].

For a boundary control problem such as (1.3) with extended subdomains $\widehat{\Omega}_i$ that do not share a boundary segment with Ω , the middle row and column block of \mathbf{K}_i and the middle row block of \mathbf{R}_i have to be deleted. The same issue may also arise when distributed control is executed only on a subset Ω_c of Ω . To simplify the presentation, we do not distinguish between these cases. The discussion in this section can be easily extended to include control problems with extended subdomains $\widehat{\Omega}_i$ on which no control is exercised.

The algebraic representations of (3.19), (3.20) are given by

$$\mathbf{K}_i \mathbf{z}_i = \mathbf{R}_i \mathbf{Kz}, \quad i = 0, \dots, N.$$

Hence,

$$\mathbf{T}_i = \mathbf{R}_i^T \mathbf{K}_i^{-1} \mathbf{R}_i \mathbf{K} \in \mathbb{R}^{(n+m+n) \times (n+m+n)}, \quad i = 0, \dots, N.$$

Similarly,

$$\mathbf{r}_i = \mathbf{R}_i^T \mathbf{K}_i^{-1} \mathbf{R}_i \mathbf{g} \in \mathbb{R}^{n+m+n}, \quad i = 0, \dots, N.$$

The algebraic representations of the transformed system (3.24) is given by

$$(3.67) \quad \sum_{i=0}^N \mathbf{T}_i \mathbf{z} = \sum_{i=0}^N \mathbf{r}_i.$$

If we define

$$(3.68) \quad \mathbf{P} = \sum_{i=0}^N \mathbf{R}_i^T \mathbf{K}_i^{-1} \mathbf{R}_i,$$

then we see that (3.67) is just the preconditioned equation

$$(3.69) \quad \mathbf{PKz} = \mathbf{Pg}$$

corresponding to (3.63).

It is interesting to observe that the system (3.62) is nonsymmetric. This choice is motivated by the existence of a splitting (3.25), (3.64). On the other hand, the nonsymmetric form of (3.62) and the resulting nonsymmetry in the preconditioner \mathbf{P} seems to preclude the application of iterative methods for symmetric indefinite systems that allow symmetric indefinite preconditioners such as the symmetric QMR (sQMR) method [27, 28]. This is not the case, as the following discussion will show.

If \mathbf{K} is the KKT matrix defined in (3.62) and if $\mathbf{\Pi} \in \mathbb{R}^{(n+m+n) \times (n+m+n)}$ is the permutation matrix that interchanges the first n and last n components of a vector, then

$$(3.70) \quad \mathbf{\Pi K} = \begin{pmatrix} \mathbf{M} & \mathbf{0} & \mathbf{A}^T \\ \mathbf{0} & \alpha \mathbf{Q} & \mathbf{B}^T \\ \mathbf{A} & \mathbf{B} & \mathbf{0} \end{pmatrix}.$$

Similarly, let \mathbf{K}_i , $i = 0, \dots, N$, be defined in (3.65), (3.65). If $\mathbf{\Pi}_i \in \mathbb{R}^{(n_i+m_i+n_i) \times (n_i+m_i+n_i)}$, $i = 0, \dots, N$, then

$$(3.71) \quad \mathbf{\Pi}_i \mathbf{K}_i = \begin{pmatrix} \mathbf{M}_i & \mathbf{0} & \mathbf{A}_i^T \\ \mathbf{0} & \alpha \mathbf{Q}_i & \mathbf{B}_i^T \\ \mathbf{A}_i & \mathbf{B}_i & \mathbf{0} \end{pmatrix} \in \mathbb{R}^{(n_i+m_i+n_i) \times (n_i+m_i+n_i)}.$$

The block diagonal structure of \mathbf{R}_i as well as the fact that the permutations $\mathbf{\Pi}$, $\mathbf{\Pi}_i$ interchange the first and last blocks imply

$$\mathbf{R}_i = \mathbf{\Pi}_i \mathbf{R}_i \mathbf{\Pi}, \quad \mathbf{\Pi}_i = \mathbf{\Pi}_i^T.$$

The preconditioned matrix \mathbf{PK} , where \mathbf{P} is defined in (3.68), can now be written as

$$\mathbf{PK} = \sum_{i=0}^N \mathbf{R}_i^T \mathbf{K}_i^{-1} \mathbf{R}_i \mathbf{K} = \sum_{i=0}^N \mathbf{R}_i^T \mathbf{K}_i^{-1} \mathbf{\Pi}_i \mathbf{R}_i \mathbf{\Pi} \mathbf{K} = \sum_{i=0}^N \mathbf{R}_i^T (\mathbf{\Pi}_i \mathbf{K}_i)^{-1} \mathbf{R}_i \mathbf{\Pi} \mathbf{K}.$$

The matrices $\mathbf{\Pi K}$, $\mathbf{\Pi}_i \mathbf{K}_i$, $i = 0, \dots, N$, are symmetric (cf. (3.70), (3.71)) and indefinite ([30, 36, 37]). Hence, instead of solving the preconditioned system (3.69) we can solve

$$(3.72) \quad \left(\sum_{i=0}^N \mathbf{R}_i^T (\mathbf{\Pi}_i \mathbf{K}_i)^{-1} \mathbf{R}_i \right) (\mathbf{\Pi K}) \mathbf{z} = \left(\sum_{i=0}^N \mathbf{R}_i^T (\mathbf{\Pi}_i \mathbf{K}_i)^{-1} \mathbf{R}_i \right) (\mathbf{\Pi g}),$$

where the matrix $\mathbf{\Pi K}$ is symmetric indefinite and the preconditioner $\sum_{i=0}^N \mathbf{R}_i^T (\mathbf{\Pi}_i \mathbf{K}_i)^{-1} \mathbf{R}_i$ is symmetric. If we use GMRES, then the convergence theory developed in Section 3.2 applies to both cases. The QMR (sQMR) residuals can be linked to the GMRES residuals [26], [45, Sec. 7.3].

4. ADJOINT REGULARITY FOR THE EXAMPLE PROBLEM

Throughout this section we let $\Omega \subset \mathbb{R}^2$ be a convex, open subset with polygonal boundary. We verify that Assumptions A1–A4 are satisfied for the example problem (1.4).

In problem (1.4) with distributed control, we have $V = H_0^1(\Omega)$, $U = L^2(\Omega)$, $a(y, \phi) = \int_{\Omega} \nabla y \nabla \phi dx$, $m(y, \phi) = \int_{\Omega_o} y \phi dx$, $b(u, \phi) = \int_{\Omega} u \phi dx$ and $q(u, \phi) = \int_{\Omega} u \phi dx$. Assumptions A1–A3 are satisfied.

To verify Assumption A4, let $w = (w_y, w_u w_p)$ and $\psi = (\theta, \mu, \phi)$. The adjoint problem (3.29) is equivalent to

$$(4.1a) \quad a(w_p, \phi) + b(w_u, \phi) = \langle l_1, \phi \rangle, \quad \forall \phi \in V,$$

$$(4.1b) \quad \alpha q(\mu, w_u) + b(w_y, \mu) = 0 \quad \forall \mu \in U,$$

$$(4.1c) \quad a(\theta, w_y) + m(\theta, w_p) = \langle l_3, \theta \rangle \quad \forall \theta \in V.$$

Standard estimates analogous to those applied in Theorem 2.2 show that (4.1) has a unique solution $w = (w_y, w_u w_p) \in H^1(\Omega) \times L^2(\Omega) \times H^1(\Omega)$ which satisfies

$$\|w\|_{H^1(\Omega) \times L^2(\Omega) \times H^1(\Omega)} \leq C(\|l_1\|_{L^2(\Omega)} + \|l_3\|_{L^2(\Omega)})$$

for some $C > 0$.

Since $w_u, l_1 \in L^2$ the solution w_p of (4.1a) is in H^2 and obeys

$$\|w_p\|_{H^2(\Omega)} \leq C(\|w_u\|_{L^2(\Omega)} + \|l_1\|_{L^2(\Omega)})$$

for some $C > 0$ [31, Th.3.2.1.2]. Analogously, since $w_p, l_3 \in L^2$, the solution w_y of (4.1c) is in H^2 and obeys

$$\|w_y\|_{H^2(\Omega)} \leq C(\|w_p\|_{L^2(\Omega)} + \|l_3\|_{L^2(\Omega)}).$$

Finally, (4.1b) implies $w_u = w_y \in H^2(\Omega)$. The previous estimates show that Assumption A4 is satisfied.

5. ERROR ESTIMATES FOR ELLIPTIC LINEAR-QUADRATIC OPTIMAL CONTROL PROBLEMS

In this section, we present basic error estimates for elliptic linear-quadratic optimal control problems that are needed in the proof of Lemma 3.8. We focus on the setting of this paper. More results on error estimates for elliptic linear-quadratic optimal control problems may be found in [14, 17, 32] and the references cited therein.

Throughout this section we assume that condition A2 holds, i.e., that the bilinear forms a, m, q satisfy (2.21). Moreover, as in Section 3 we assume that the triangulation is quasi-uniform. We can define operators $\mathcal{A}_{h,k} \in \mathcal{L}(V^h, (V^k)')$, $\mathcal{B}_{h,k} \in \mathcal{L}(U^h, (V^k)')$, $\mathcal{M}_{h,k} \in \mathcal{L}(V^h, (V^k)')$, $\mathcal{Q}_{h,k} \in \mathcal{L}(U^k, (U^k)')$, via

$$(5.1) \quad \begin{aligned} \langle \mathcal{A}_{h,k} v^h, \phi^k \rangle_{V',V} &= a(v^h, \phi^k), & \langle \mathcal{B}_{h,k} u^h, \phi^k \rangle_{V',V} &= b(u^h, \phi^k), \\ \langle \mathcal{M}_{h,k} v^h, \phi^k \rangle_{V',V} &= m(v^h, \phi^k), & \langle \mathcal{Q}_{h,k} u^h, \mu^k \rangle_{U',U} &= q(u^h, \mu^k), \end{aligned}$$

where $\langle \cdot, \cdot \rangle_{V',V}$ the duality pairing between V' and V . If $h = k$, we simply write \mathcal{A}_h, \dots instead of $\mathcal{A}_{h,h}, \dots$. Since a, m, q satisfy (2.21),

$$(5.2) \quad \begin{aligned} \langle \mathcal{A}_h v^h, v^h \rangle_{V',V} &\geq c_a \|v^h\|_V^2, & \langle \mathcal{Q}_h u^h, u^h \rangle_{U',U} &\geq c_q \|u^h\|_U^2, \\ \langle \mathcal{M}_h v^h, v^h \rangle_{V',V} &\geq 0. \end{aligned}$$

for all $v^h \in V^h$, $u^h \in U^h$. In particular, \mathcal{A}_h and \mathcal{Q}_h are continuously invertible.

The operator $\mathcal{K}_{h,k} \in \mathcal{L}(Z^h, (Z^k)')$ associated with the finite element discretization of (3.6) is given by

$$(5.3) \quad \mathcal{K}_{h,k} = \begin{pmatrix} \mathcal{A}_{h,k} & \mathcal{B}_{h,k} & 0 \\ 0 & \alpha \mathcal{Q}_{h,k} & \mathcal{B}_{h,k}^* \\ \mathcal{M}_{h,k} & 0 & \mathcal{A}_{h,k}^* \end{pmatrix}.$$

We also consider $\mathcal{A} \in \mathcal{L}(V, V')$, $\mathcal{B} \in \mathcal{L}(U, V')$, $\mathcal{M} \in \mathcal{L}(V, V')$, $\mathcal{Q} \in \mathcal{L}(U, U')$, $\mathcal{K} \in \mathcal{L}(Z, Z')$, defined analogously to (5.1), (5.3).

Lemma 5.1. *Let Z^h be equipped with the A -norm. If the a, m, b, q are continuous and if a, m, q satisfy (2.21), then the operator $\mathcal{K}_h \in \mathcal{L}(Z^h, (Z^h)')$ defined in (5.3), (5.1) is invertible and there exists $\kappa > 0$ such that*

$$\|\mathcal{K}_h^{-1}\|_{\mathcal{L}((Z^h)', Z^h)} \leq \kappa \quad \forall h.$$

Proof. Since a, b, m are continuous bilinear forms, $\mathcal{A}_h, \mathcal{B}_h, \mathcal{M}_h$ are uniformly bounded. By (5.2), the inverses \mathcal{A}_h and \mathcal{Q}_h exist and are uniformly bounded.

Elementary calculations show that

$$(5.4) \quad \begin{aligned} \mathcal{K}_h &= \begin{pmatrix} 0 & I & 0 \\ \mathcal{B}_h^*(\mathcal{A}_h^*)^{-1} & -\mathcal{B}_h^*(\mathcal{A}_h^*)^{-1}\mathcal{M}_h\mathcal{A}_h^{-1} & I \\ I & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathcal{M}_h & \mathcal{A}^* & 0 \\ \mathcal{A}_h & 0 & 0 \\ 0 & 0 & \widehat{\mathcal{Q}}_h \end{pmatrix} \\ &\times \begin{pmatrix} I & \mathcal{A}_h^{-1}\mathcal{B}_h & 0 \\ 0 & -(\mathcal{A}_h^*)^{-1}\mathcal{M}_h\mathcal{A}_h^{-1}\mathcal{B}_h & I \\ 0 & I & 0 \end{pmatrix}, \end{aligned}$$

where $\widehat{\mathcal{Q}}_h = \alpha \mathcal{Q}_h + \mathcal{B}_h^*(\mathcal{A}_h^*)^{-1}\mathcal{M}_h\mathcal{A}_h^{-1}\mathcal{B}_h$. By (5.2), $\widehat{\mathcal{Q}}_h$ is invertible and its inverse is uniformly bounded. Since $\mathcal{A}_h, \mathcal{B}_h, \mathcal{M}_h, \mathcal{A}_h^{-1}, \widehat{\mathcal{Q}}_h^{-1}$ are uniformly bounded, the operators on the right hand side of (5.4) are invertible and their inverses uniformly bounded. This proves the uniform boundedness of \mathcal{K}_h^{-1} . \square

With (5.1), (5.3), the equation

$$K(z^H, \psi^H) = K(z^h, \psi^h) \quad \forall \psi^H \in Z^H$$

can be written as

$$\mathcal{K}_H z^H = \mathcal{K}_{h,H} z^h.$$

With Lemma 5.1 this implies

$$(5.5) \quad \|z^H\|_A \leq \|\mathcal{K}_H^{-1}\| \|\mathcal{K}_{h,H}\| \|z^h\|_A$$

(cf. 3.36).

Let k be defined as in (3.6) and let $l \in Z'$ be given. We are interested in the error between the solutions z and z^h of

$$(5.6) \quad k(z, \psi) = \langle l, \psi \rangle_{Z', Z} \quad \forall \psi \in Z$$

and

$$(5.7) \quad k(z^h, \psi^h) = \langle l, \psi^h \rangle_{Z', Z}, \quad \forall \psi^h \in Z^h$$

respectively. If we define, $l_h \in (Z^h)'$ by $\langle l_h, \psi^h \rangle_{Z', Z} = \langle l, \psi^h \rangle_{Z', Z}$ for all $\psi^h \in Z^h$, then (5.6), (5.7) can be written as

$$(5.8) \quad \mathcal{K}z = l$$

and

$$(5.9) \quad \mathcal{K}_h z^h = l_h,$$

respectively.

To derive an estimate for the error $\|z^h - z\|_A$, we let $\mathcal{R}_h : Z \rightarrow Z^h$ be a restriction operator and we consider the identity

$$\mathcal{K}_h(z^h - \mathcal{R}_h(z)) = l_h - \mathcal{K}_h \mathcal{R}_h(z).$$

We immediately obtain the estimate

$$(5.10) \quad \|z^h - \mathcal{R}_h(z)\|_A \leq \|\mathcal{K}_h^{-1}\|_{\mathcal{L}((Z^h)', Z^h)} \|l_h - \mathcal{K}_h \mathcal{R}_h(z)\|_{(Z^h)'} \leq \kappa \|l_h - \mathcal{K}_h \mathcal{R}_h(z)\|_A,$$

where we have used Lemma 5.1. An estimate for $\|z^h - z\|_A$ now follows from estimates of $\|l_h - \mathcal{K}_h \mathcal{R}_h(z)\|_{Z'}$ and $\|\mathcal{R}_h(z) - z\|_A$.

Let $z = (y, u, p)$. Let $\Pi_h : C^0(\bar{\Omega}) \rightarrow V^h$ be the V^h interpolant [15, 16]. and let $r_h : L^2(\mathcal{D}) \rightarrow U^h$ be the Clément interpolation operator [15, p. 132], [7, p. 82]. The results in [15, Th. 17.1],[16, Thm. 3.2.1], [15, p. 133], [7, p. 82] guarantee the existence of $C > 0$ (independent of h) such that

$$(5.11a) \quad \|v - \Pi_h v\|_{H^1(\Omega)} \leq Ch \|v\|_{H^2(\Omega)} \quad \forall v \in H^2(\Omega),$$

$$(5.11b) \quad \|u - r_h u\|_{L^2(\mathcal{D})} \leq Ch \|u\|_{H^1(\mathcal{D})} \quad \forall u \in H^1(\mathcal{D}).$$

If we define the restriction operator

$$(5.12) \quad \mathcal{R}_h(z) = (\Pi_h y, r_h u, \Pi_h p)^T$$

and assume

$$z \in H^2(\Omega) \times H^1(\Omega) \times H^2(\Omega),$$

then (5.11) implies

$$(5.13) \quad \|z - \mathcal{R}_h(z)\|_A \leq Ch \max\{\|y\|_{H^2(\Omega)}, \|u\|_{H^1(\Omega)}, \|p\|_{H^2(\Omega)}\}.$$

To estimate $\|l_h - \mathcal{K}_h \mathcal{R}_h(z)\|_{(Z^h)'}$, we use the definition of l_h and \mathcal{K}_h to obtain

$$\begin{aligned} \|l_h - \mathcal{K}_h \mathcal{R}_h(z)\|_{(Z^h)'} &= \sup_{\psi^h \in Z^h \setminus \{0\}} \frac{\langle l_h, \psi^h \rangle_{Z', Z} - \langle \mathcal{K}_h \mathcal{R}_h(z), \psi^h \rangle_{Z', Z}}{\|\psi^h\|_A}, \\ &= \sup_{\psi^h \in Z^h \setminus \{0\}} \frac{\langle l, \psi^h \rangle_{Z', Z} - \langle \mathcal{K} \mathcal{R}_h(z), \psi^h \rangle_{Z', Z}}{\|\psi^h\|_A}. \end{aligned}$$

Since z solves (5.8),

$$\begin{aligned} \|l_h - \mathcal{K}_h \mathcal{R}_h(z)\|_{(Z^h)'} &= \sup_{\psi^h \in Z^h \setminus \{0\}} \frac{\langle l - \mathcal{K}z, \psi^h \rangle_{Z', Z} + \langle \mathcal{K}(z - \mathcal{R}_h(z)), \psi^h \rangle_{Z', Z}}{\|\psi^h\|_A} \\ &= \sup_{\psi^h \in Z^h \setminus \{0\}} \frac{\langle \mathcal{K}(z - \mathcal{R}_h(z)), \psi^h \rangle_{Z', Z}}{\|\psi^h\|_A} \\ &\leq \sup_{\psi^h \in Z^h \setminus \{0\}} \frac{\|\mathcal{K}\|_{\mathcal{L}(Z, Z')} \|z - \mathcal{R}_h(z)\|_A \|\psi^h\|_A}{\|\psi^h\|_A} \\ (5.14) \quad &\leq \|\mathcal{K}\|_{\mathcal{L}(Z, Z')} \|z - \mathcal{R}_h(z)\|_A. \end{aligned}$$

Combining the estimates (5.10), (5.13) and (5.14) gives the following error estimate.

Theorem 5.2. *Let a, m, b, q be continuous and let a, m, q satisfy (2.21). Let $\{\mathcal{T}_h\}_h$ and $\{\{\tau \cap \mathcal{D} : \tau \in \mathcal{T}_h\}\}_h$ be regular families of triangulations. If the solution $z = (y, u, p)$ of (5.8) satisfies $y, p \in H^2(\Omega)$ and $u \in H^1(\mathcal{D})$, then there exists $C > 0$ independent of h such that the error between the solution z of (5.8) and the solution $z^h \in Z^h$ of (5.9) obeys*

$$\|z - z^h\|_A \leq Ch \max\{\|y\|_{H^2(\Omega)}, \|u\|_{H^1(\mathcal{D})}, \|p\|_{H^2(\Omega)}\} \quad \forall h.$$

Again, let k be defined as in (3.6) and let $l \in Z'$ be given. In the proof of Lemma 3.8 we need an estimate of the error between the solutions w and w^h of

$$(5.15) \quad k(\psi, w) = \langle l, \psi \rangle_{Z', Z} \quad \forall \psi \in Z$$

and

$$(5.16) \quad k(\psi^h, w^h) = \langle l, \psi^h \rangle_{Z', Z}, \quad \forall \psi^h \in Z^h$$

respectively. If we define, $l_h \in (Z^h)'$ by $\langle l_h, \psi^h \rangle_{Z', Z} = \langle l, \psi^h \rangle_{Z', Z}$ for all $\psi^h \in Z^h$, then (5.15), (5.16) can be written as

$$(5.17) \quad \mathcal{K}^* w = l$$

and

$$(5.18) \quad \mathcal{K}_h^* w^h = l_h,$$

respectively. Using $\|\mathcal{K}^*\| = \|\mathcal{K}\|$, $\|(\mathcal{K}^*)^{-1}\| = \|\mathcal{K}^{-1}\|$ and the same techniques applied to prove Theorem 5.2 we can establish the following result.

Theorem 5.3. *Let a, m, b, q be continuous and let a, m, q satisfy (2.21). Let $\{\mathcal{T}_h\}_h$ and $\{\{\tau \cap \mathcal{D} : \tau \in \mathcal{T}_h\}\}_h$ be regular families of triangulations. If the solution $w = (y, u, p)$ of (5.17) satisfies $y, p \in H^2(\Omega)$ and $u \in H^1(\mathcal{D})$, then there exists $C > 0$ independent of h such that the error between the solution w of (5.17) and the solution $w^h \in Z^h$ of (5.18) obeys*

$$\|w - w^h\|_A \leq Ch \max\{\|y\|_{H^2(\Omega)}, \|u\|_{H^1(\mathcal{D})}, \|p\|_{H^2(\Omega)}\} \quad \forall h.$$

6. CONCLUSION

We have extended the framework based on subspace decomposition, which is well known for elliptic PDEs, to derive domain decomposition methods for linear–quadratic elliptic optimal control problems. We have shown that the subdomain problems that arise in our preconditioners are essentially smaller copies of the original optimal control problem, hence allowing code reuse.

The subspace decomposition framework was then applied to derive overlapping DD preconditioners for linear–quadratic elliptic optimal control problems. We have shown that the convergence factor of GMRES preconditioned with a two-level version of our overlapping preconditioners is independent of the mesh size h and of the subdomain size H , provided that the coarse grid is sufficiently small, relative to the control regularization α . This result extends to the optimal control context results that are well known for overlapping DD methods applied to individual PDEs.

Our numerical results in [44] indicate that the convergence behavior of our two-level overlapping DD preconditioners for linear–quadratic elliptic optimal control problems is comparable the performance of their counterparts applied to single elliptic PDEs. However, the numerical results in [44] also indicate that our convergence theorem is too pessimistic for small α . In fact, for the test problem with distributed controls considered in [44], we have not observed any dependence of the size of the coarse grid on α . A sharper theoretical

analysis of the convergence dependence on the size of the control regularization α is part of our future research.

In [44] the conditions $\alpha q(u^*, \mu) + b(\mu, p^*) = d(\mu)$ for all $\mu \in U$ (cf. (2.2b)) and $q = b$, $d = 0$ were used to express the control u^* in terms of p^* and eliminate the former from the optimality system. An overlapping domain decomposition method corresponding to the one discussed in this paper was applied in [44] to the resulting 2×2 system. An advantage, from the theoretical point of view, of the approach in [44] is that implicitly $u = \alpha^{-1}p$ and, hence u inherits all regularity properties from the adjoint p . This is not necessarily the case in our approach and makes the proofs of Lemmas 3.13, 3.14 more involved. However, we believe that the approach followed in his paper is more suitable for the optimization context, since it allows one to formally extend the domain decomposition methods in this paper to problems with point-wise control constraints in several ways.

REFERENCES

- [1] J.-D. Benamou, *A domain decomposition method with coupled transmission conditions for the optimal control of systems governed by elliptic partial differential equations*, Tech. Report No. 2246, INRIA, Domaine de Voluceau, 78153 Rocquencourt, France, 1994, electronically available as <http://www.inria.fr/rrrt/r-2246.html>.
- [2] ———, *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.
- [3] ———, *A domain decomposition method for control problems*, DD9 Proceedings (Bergen, Norway) (P. Bjørstad, M. Espedal, and D. Keyes, eds.), Domain Decomposition Press, 1998, Available electronically from <http://www.ddm.org/DD9/index.html>, pp. 266–273.
- [4] G. Biros and O. Ghattas, *Parallel preconditioners for KKT systems arising in optimal control of viscous incompressible flows*, Proceedings of Parallel CFD '99, Williamsburg, VA, May 23–26, 1999 (Amsterdam, London, New-York), North Holland, 1999, <http://www.cs.cmu.edu/~oghattas>.
- [5] ———, *Parallel Lagrange–Newton–Krylov–Schur methods for PDE-constrained optimization. part I: The Krylov–Schur solver*, Tech. report, Computational Mechanics Lab, Department of Civil and Environmental Engineering, Carnegie Mellon University, 2000, <http://www.cs.cmu.edu/~oghattas>.
- [6] A. Bounaim, *A Lagrangian approach to a DDM for an optimal control problem*, DD9 Proceedings (Bergen, Norway) (P. Bjørstad, M. Espedal, and D. Keyes, eds.), Domain Decomposition Press, 1998, Available electronically from <http://www.ddm.org/DD9/index.html>, pp. 283–289.
- [7] D. Braess, *Finite elements. theory, fast solvers, and applications in solid mechanics*, Cambridge University Press, Cambridge, New York, 1997.
- [8] J. H. Bramble and J. Xu, *Some estimates for a weighted L^2 projection*, *Math. Comput.* **56** (1991), 463–476.
- [9] M. Burger and R. Pinnau, *Fast optimal design of semiconductor devices*, *SIAM J. Appl. Math.* **64** (2003), no. 1, 108–126 (electronic).
- [10] X. C. Cai, *An additive Schwarz algorithm for nonselfadjoint elliptic equations*, Third International Symposium on Domain Decomposition Methods for Partial Differential Equations (T. Chan, R. Glowinski, J. Périaux, and O. Widlund, eds.), SIAM, Philadelphia, PA, 1990, pp. 232–244.
- [11] X.-C. Cai and D. E. Keyes, *Nonlinear preconditioned inexact Newton algorithms*, *SIAM J. Scientific Computing* **24** (2002), 183–200.
- [12] X.-C. Cai and O. B. Widlund, *Domain decomposition algorithms for indefinite elliptic problems*, *SIAM J. Sci. Statist. Comput.* **13** (1992), no. 1, 243–258.
- [13] ———, *Multiplicative Schwarz algorithms for some nonsymmetric and indefinite problems*, *SIAM J. Numer. Anal.* **30** (1993), no. 4, 936–952.
- [14] E. Casas and F. Tröltzsch, *Error estimates for linear-quadratic elliptic control problems*, Analysis and optimization of differential systems (Constanta, 2002), Kluwer Acad. Publ., Boston, MA, 2003, pp. 89–100.
- [15] P. G. Ciarlet, *Basic error estimates for elliptic problems*, Handbook of Numerical Analysis, Vol.2: Finite Element Methods (Part 1) (P.G. Ciarlet and J.L. Lions, eds.), Elsevier/North-Holland, Amsterdam, New York, Oxford, Tokyo, 1991.
- [16] ———, *The finite element method for elliptic problems*, SIAM, Philadelphia, 2002.

- [17] S. S. Collis and M. Heinkenschloss, *Analysis of the streamline upwind/pevov galerkin method applied to the solution of optimal control problems*, Tech. Report TR02–01, Department of Computational and Applied Mathematics, Rice University, Houston, TX 77005–1892, 2002, <http://www.caam.rice.edu/~heinken>.
- [18] M. Delgado, J. A. Montero, and A. Suárez, *Study of the optimal harvesting control and the optimality system for an elliptic problem*, SIAM J. Control Optim. **42** (2003), no. 5, 1559–1577 (electronic).
- [19] J. E. Dennis and R. M. Lewis, *A comparison of nonlinear programming approaches to an elliptic inverse problem and a new domain decomposition approach*, Tech. Report TR94–33, Department of Computational and Applied Mathematics, Rice University, P.O.Box 1892, Houston, TX 77251–1892, 1994.
- [20] M. Dryja, B. F. Smith, and O. B. Widlund, *Schwarz analysis of iterative substructuring algorithms for elliptic problems in three dimensions*, SIAM J. Numer. Anal. **31** (1994), no. 3, 1662–1694.
- [21] M. Dryja and O. B. Widlund, *Some domain decomposition algorithms for elliptic problems*, Iterative Methods for Large Linear Systems (San Diego, CA) (D. R. Kinkaid and L. J. Hayes, eds.), Academic Press, 1989, pp. 273–291.
- [22] ———, *Towards a unified theory of domain decomposition algorithms for elliptic problems*, Third International Symposium on Domain Decomposition Methods for Partial Differential Equations (Philadelphia) (T. F. Chan, R. Glowinski, J. Periaux, and O. B. Widlund, eds.), SIAM, 1990, pp. 3–21.
- [23] ———, *Domain decomposition algorithms with small overlap*, SIAM J. Sci.Comput. **15** (1994), no. 3, 604–620.
- [24] ———, *Schwarz methods of Neumann-Neumann type for three-dimensional elliptic finite element methods*, Comm. Pure Appl. Math. **48** (1995), 121–155.
- [25] S. C. Eisenstat, H. C. Elman, and M. H. Schultz, *Variational iterative methods for nonsymmetric systems of linear equations*, SIAM J. Numer. Anal. **20** (1983), 345–357.
- [26] R. Freund and N. M. Nachtigal, *QMR: a quasi-minimal residual method for non-Hermitian linear systems*, Numerische Mathematik **60** (1992), 315–340.
- [27] R. W. Freund and N. M. Nachtigal, *A new Krylov-subspace method for symmetric indefinite linear systems*, Proceedings of the 14th IMACS World Congress on Computational and Applied Mathematics (W. F. Ames, ed.), IMACS, 1994, pp. 1253–1256.
- [28] ———, *Software for simplified Lanczos and QMR algorithms*, Applied Numerical Mathematics **19** (1995), 319–341.
- [29] D. Gilbarg and N.S. Trudinger, *Elliptic partial differential equations of second order*, second (Classics in Mathematics) ed., Springer-Verlag, Berlin, 2001.
- [30] N. I. M. Gould, *On practical conditions for the existence and uniqueness of solutions to the general equality quadratic programming problem*, Math. Programming **32** (1985), no. 1, 90–99.
- [31] P. Grisvard, *Elliptic problems in nonsmooth domains*, Pitman, Boston, London, Melbourne, 1985.
- [32] M. D. Gunzburger and L. S. Hou, *Finite dimensional approximation of a class of constrained nonlinear optimal control problems*, SIAM J. Control Optimization **34** (1996), 1001–1043.
- [33] M. D. Gunzburger, L. S. Hou, and S. S. Ravindran, *Analysis and approximation of optimal control problems for a simplified Ginzburg-Landau model of superconductivity*, Numerische Mathematik **77** (1997), 243–268.
- [34] M. D. Gunzburger, L. S. Hou, and T. P. Svobotny, *Finite element approximations of an optimal control problem associated with the scalar Ginzburg-Landau equation*, Comput. Math. Appl. (1991), 123–131.
- [35] ———, *Heating and cooling control of temperature distributions along boundaries of flow domains*, J. of Mathematical Systems, Estimation, and Control **3** (1993), 147–172.
- [36] M. Heinkenschloss and H. Nguyen, *Balancing Neumann-Neumann methods for elliptic optimal control problems*, Domain Decomposition methods in Science and Engineering (Heidelberg) (R. Kornhuber, R. H. W. Hoppe, J. Periaux, O. Pironneau, O. B. Widlund, and J. Xu, eds.), Lecture Notes in Computational Science and Engineering, Vol. 40, Springer-Verlag, 2004, pp. 589–596.
- [37] ———, *Neumann-Neumann domain decomposition preconditioners for linear quadratic elliptic optimal control problems*, Tech. Report TR04–01, Department of Computational and Applied Mathematics, Rice University, 2004.
- [38] J. Jahn, *Introduction to the theory of nonlinear optimization*, second ed., Springer Verlag, Berlin, Heidelberg, New York, 1996.
- [39] B. N. Khoromskij and G. Wittum, *Numerical solution of elliptic differential equations by reduction to the interface*, Lecture Notes in Computational Science and Engineering, Vol. 36, Springer Verlag, Berlin, Heidelberg, New York, 2004.
- [40] K. Kunisch and X.-C. Tai, *Nonoverlapping domain decomposition methods for inverse problems*, Proceedings of the 9th International Conference on Domain Decomposition Methods in Bergen, Norway, ddm.org, 1998, pp. 517–524.

- [41] J.-L. Lions, *Optimal control of systems governed by partial differential equations*, Springer Verlag, Berlin, Heidelberg, New York, 1971.
- [42] J.-L. Lions and O. Pironneau, *Sur le contrôle parallèle des systèmes distribués*, Comptes Rendus de l'Académie des Sciences. Série I. Mathématique **327** (1998), no. 12, 993–998.
- [43] W. G. Litvinov, *Optimization in elliptic problems with applications to mechanics of deformable bodies and fluid mechanics*, Operator Theory Advances and Applications, Vol. 119, Birkhäuser Verlag, Basel, Boston, Berlin, 2000.
- [44] H. Q. Nguyen, *Domain decomposition methods for linear-quadratic elliptic optimal control problems*, Ph.D. thesis, Department of Computational and Applied Mathematics, Rice University, Houston, TX, August 2004, Available as CAAM TR04-16.
- [45] Y. Saad, *Iterative methods for sparse linear systems*, PWS Publishing Company, Boston, New-York, Singapore, Toronto, 1996.
- [46] Y. Saad and M. H. Schultz, *GMRES a generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Stat. Comp. **7** (1986), 856–869.
- [47] A. Shenoy, E. M. Cliff, and M. Heinkenschloss, *Thermal–fluid control via finite–dimensional approximations*, Proceedings from the 31st AIAA Thermophysics Conference, AIAA Paper 96–1910 (New Orleans, LA), 1996.
- [48] B. Smith, P. Bjørstad, and W. Gropp, *Domain decomposition. parallel multilevel methods for elliptic partial differential equations*, Cambridge University Press, Cambridge, London, New York, 1996.
- [49] X.-C. Tai, J. Frøyen, M. S. Espedal, and T. F. Chan, *Overlapping domain decomposition and multigrid methods for inverse problems*, Domain decomposition methods, 10 (Boulder, CO, 1997), Contemp. Math., vol. 218, Amer. Math. Soc., Providence, RI, 1998, pp. 523–529.
- [50] J. Xu, *Iterative methods by space decomposition and subspace correction*, SIAM Review **34** (1992), 581–613.

DEPARTMENT OF COMPUTATIONAL AND APPLIED MATHEMATICS, MS-134, RICE UNIVERSITY, 6100
MAIN STREET, HOUSTON, TX 77005-1892
E-mail address: heinken@rice.edu

DEPARTMENT OF COMPUTATIONAL AND APPLIED MATHEMATICS, MS-134, RICE UNIVERSITY, 6100
MAIN STREET, HOUSTON, TX 77005-1892
E-mail address: hnguy@rice.edu