RICE UNIVERSITY

Domain Decomposition Methods for Linear-Quadratic Elliptic Optimal Control Problems

by

Hoang Q. Nguyen

A Thesis Submitted
in Partial Fulfillment of the
Requirements for the Degree

Doctor of Philosophy

Approved, Thesis Committee:

Dr. Matthias Heinkenschloss, Chairman
Associate Professor of Computational and Applied Mathematics

Dr. William W. Symes
Noah Harding Professor of Computational and Applied Mathematics

---

Dr. Matteo Pasquali
Assistant Professor of Chemical Engineering

Dr. Mark P. Embree
Assistant Professor of Computational and Applied Mathematics

Houston, Texas

July, 2004

Abstract

Domain Decomposition Methods for Linear-Quadratic Elliptic Optimal Control Problems

by

Hoang Q. Nguyen

This thesis is concerned with the development of domain decomposition (DD) based preconditioners for linear-quadratic elliptic optimal control problems (LQ-EOCPs), their analysis, and numerical studies of their performance on model problems. The solution of LQ-EOCPs arises in many applications, either directly or as subproblems in Newton or Sequential Quadratic Programming methods for the solution of nonlinear elliptic optimal control problems. After a finite element discretization, convex LQ-EOCPs lead to large scale symmetric indefinite linear systems. The solution of these large systems is a very time consuming step and must be done iteratively, typically with a preconditioned Krylov subspace method. Developing good preconditioners for these linear systems is an important part of improving the overall performance of the solution method.

The DD preconditioners for LQ-EOCPs studied in this thesis are extensions of overlapping and nonoverlapping Neumann-Neumann DD preconditioners applied to single elliptic partial differential equations (PDEs). In our case, DD is applied on the optimization level. In particular, the proposed preconditioners require the parallel solution of subdomain optimal control problems that are related to restrictions of

the original LQ-EOCP to a subdomain. Numerical results on several test problems have shown that the new preconditioners are effective. Their performance relative to decreases in finite element mesh size or increase in number of subdomains seem to be numerically comparable to that of overlapping and Neumann-Neumann preconditioners for single PDEs. Remarkably, the proposed preconditioners seem to be rather insensitive to control regularization parameters. For overlapping methods, theoretical results are provided to support the numerical observations.

Acknowledgements

First of all, I would like to express deep gratitude to my advisor, Dr. Matthias Heinkenschloss, for all his generous help and guidance on this research. His dedication and support help make this process well worth it in the end.

I would like to thank the members of my thesis committee, Dr. William Symes, Dr. Matteo Pasquali and Dr. Mark Embree for their interest in my research. Special thanks to Dr. Embree for proofreading the thesis.

I would like to thank the CAAM faculty for accepting me into the Department, and for many interesting and challenging classes. I thank my fellow students, including Agata, Denis, Eric, Fernando, Luis, Yunkai, Zenaida, as well as the staff at CAAM, especially Daria, for making the graduate life more pleasant. Also, I would like to thank the many teachers and researchers over the years who have contributed to this thesis in some way.

Most importantly, I thank my parents, Hanh and Lan, and my family for their constant love and support. This thesis is dedicated to the memory of my mother.

The research in this thesis was supported by Contract No. 74837-001-0349 from the Regents of University of California (Los Alamos National Laboratory) to William Marsh Rice University and by NSF grant ACI-0121360.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Linear-Quadratic Optimal Control Problems

This thesis is concerned with the development, analysis and numerical study of a class of domain decomposition methods for the solution of linear-quadratic elliptic optimal control problems. We illustrate our algorithms using the model problems

$$\begin{aligned} \min \quad & \tfrac{1}{2}\int_{\Omega_o}(y(x)-y_d(x))^2\,dx + \tfrac{\alpha}{2}\int_{\Omega}(u(x))^2\,dx \\ \text{s.t.} \quad & \\ & -\Delta y(x) = f(x)+u(x) \qquad \text{in } \Omega, \\ & y(x) = 0 \qquad \text{on } \partial\Omega, \end{aligned} \tag{1.1.1}$$

and

$$\begin{aligned} \min \quad & \tfrac{1}{2}\int_{\Omega_o}(y(x)-y_d(x))^2\,dx + \tfrac{\alpha}{2}\int_{\partial\Omega}(u(x))^2\,dx \\ \text{s.t.} \quad & \\ & -\Delta y(x)+y(x) = f(x) \qquad \text{in } \Omega, \\ & \tfrac{\partial}{\partial n}y(x) = u(x) \qquad \text{on } \partial\Omega, \end{aligned} \tag{1.1.2}$$

where y denotes the state, u denotes the control, Ω_o is a measurable subset of Ω, $y_d \in L^2(\Omega_o)$ is the given desired state, f is a given function and $\alpha > 0$ is the regularization parameter. The model control problem (1.1.1) is an optimization problem in the variables $u \in L^2(\Omega)$, $y \in H_0^1(\Omega)$. In our computational examples $\Omega \subset \mathbb{R}^2$, but the

formulation of our methods extends to $\Omega \subset \mathbb{R}^3$. The model control problem (1.1.2) is an optimization problem in the variables $u \in L^2(\partial\Omega)$, $y \in H^1(\Omega)$.

Mostly, our domain decomposition methods are formulated and analyzed for the abstract problem class

$$\begin{array}{ll} \min & \frac{1}{2}m(y,y) - c(y) + \frac{1}{2}q(u,u) - d(u) \\ \text{s.t.} & a(y,\phi) + b(u,\phi) = f(\phi) \quad \forall \phi \in V, \end{array} \tag{1.1.3}$$

where the optimization variables are $y \in V, u \in U$, V and U being Hilbert spaces,

$$a : V \times V \to \mathbb{R}, \quad b : U \times V \to \mathbb{R},$$

are continuous bilinear forms,

$$m : V \times V \to \mathbb{R}, \quad q : U \times U \to \mathbb{R},$$

are symmetric and continuous bilinear forms, and

$$c : V \to \mathbb{R}, \quad d : U \to \mathbb{R}, \quad f : V \to \mathbb{R},$$

are continuous linear functionals with

$$\begin{array}{rcll} m(v,v) & \geq & 0 & \forall v \in V, \\ q(u,u) & \geq & \zeta \|u\|_U^2 & \forall u \in U, \\ a(v,v) & \geq & \eta \|v\|_V^2 & \forall v \in V, \end{array}$$

for some positive constants ζ and η. The model problems (1.1.1) and (1.1.2) are special cases of (1.1.3) with

$$\begin{aligned} & V = H_0^1(\Omega), \quad U = L^2(\Omega), \\ & a(y,v) = \int_\Omega \nabla y(x) \nabla v(x) dx, \quad b(u,v) = -\int_\Omega u(x)v(x)dx, \\ & m(y,v) = \int_{\Omega_o} y(x)v(x)dx, \quad q(u,v) = \alpha \int_\Omega u(x)v(x)dx, \\ & f(v) = \int_\Omega f(x)v(x)dx, \quad c(v) = \int_\Omega y_d(x)v(x)dx, \quad d(v) = 0, \end{aligned}$$

and

$$
\begin{aligned}
&V = H^1(\Omega), \quad U = L^2(\partial\Omega),\\
&a(y,v) = \int_\Omega \nabla y(x)\nabla v(x) + y(x)v(x)dx, \quad b(u,v) = \int_{\partial\Omega} u(x)v(x)dx,\\
&m(y,v) = \int_{\Omega_o} y(x)v(x)dx, \quad q(u,v) = \alpha\int_{\partial\Omega} u(x)v(x)dx,\\
&f(v) = \int_\Omega f(x)v(x)dx, \quad c(v) = \int_\Omega y_d(x)v(x)dx, \quad d(v) = 0,
\end{aligned}
$$

respectively.

Elliptic linear-quadratic optimal control problems arise in several applications (see, e.g., [GHS93, Lio71, SCH96]), but they also arise as subproblems in the sequential quadratic programming (SQP) method or in (quasi-) Newton methods for the solution of nonlinear elliptic optimal control problems such as those described, e.g., in [GHR97, GHS91, Lit00].

After a discretization of problem (1.1.3), we have the quadratic programming (QP) problem

$$
\begin{aligned}
\min_{\mathbf{y},\mathbf{u}} \quad & \tfrac{1}{2}\mathbf{y}^T\mathbf{M}\mathbf{y} + \tfrac{\alpha}{2}\mathbf{u}^T\mathbf{Q}\mathbf{u} - \mathbf{c}^T\mathbf{y}\\
\text{s.t.} \quad & \mathbf{A}\mathbf{y} + \mathbf{B}\mathbf{u} = \mathbf{b},
\end{aligned}
\tag{1.1.4}
$$

where $\mathbf{A} \in \mathbb{R}^{m\times m}$ is invertible, $\mathbf{M} \in \mathbb{R}^{m\times m}$ is positive semidefinite, $\mathbf{Q} \in \mathbb{R}^{n\times n}$ is positive definite, and $\mathbf{B} \in \mathbb{R}^{m\times n}$, $\mathbf{b}, \mathbf{c} \in \mathbb{R}^m$, $\mathbf{d} \in \mathbb{R}^n$.

The necessary and sufficient optimality conditions for (1.1.4) are given by

$$
\begin{pmatrix} \mathbf{M} & \mathbf{0} & \mathbf{A}^T \\ \mathbf{0} & \alpha\mathbf{Q} & \mathbf{B}^T \\ \mathbf{A} & \mathbf{B} & \mathbf{0} \end{pmatrix}
\begin{pmatrix} \mathbf{y} \\ \mathbf{u} \\ \mathbf{p} \end{pmatrix} =
\begin{pmatrix} \mathbf{c} \\ \mathbf{0} \\ \mathbf{b} \end{pmatrix},
\tag{1.1.5}
$$

where $\mathbf{p}$ is the vector of Lagrange multipliers or adjoint variables. We denote the system matrix in (1.1.5) by $\mathbf{K}$, which is known as the KKT (Karush-Kuhn-Tucker) matrix.

Since $\mathbf{A}$ is invertible, the first and last equation in (1.1.5) can be used to eliminate

$\mathbf{y}$ and $\mathbf{p}$, respectively. This leads to the system

$$\widehat{\mathbf{Q}}\,\mathbf{u} = \widehat{\mathbf{g}}, \tag{1.1.6}$$

where

$$\begin{aligned} \widehat{\mathbf{Q}} &= \alpha\mathbf{Q} + \mathbf{B}^T\mathbf{A}^{-T}\mathbf{M}\mathbf{A}^{-1}\mathbf{B}, \\ \widehat{\mathbf{g}} &= -\mathbf{B}^T\mathbf{A}^{-T}(\mathbf{c} - \mathbf{M}\mathbf{A}^{-1}\mathbf{b}). \end{aligned}$$

The symmetric positive definite matrix $\widehat{\mathbf{Q}}$ is also known as the reduced Hessian. Given $\mathbf{u}$, the discrete states $\mathbf{y}$ and adjoints $\mathbf{p}$ can be computed from

$$\begin{aligned} \mathbf{y} &= \mathbf{A}^{-1}(\mathbf{b} - \mathbf{B}\mathbf{u}), \\ \mathbf{p} &= \mathbf{A}^{-T}[\mathbf{c} - \mathbf{M}\mathbf{A}^{-1}(\mathbf{b} - \mathbf{B}\mathbf{u})]. \end{aligned}$$

For large scale problems, especially those governed by elliptic partial differential equations (PDEs) in $\Omega \subset \mathbb{R}^3$ or by systems of elliptic PDEs, the linear systems (1.1.5) and (1.1.6) have to be solved iteratively. The system matrix in (1.1.5) is symmetric indefinite and (1.1.5) can be solved using SYMMLQ, MINRES [PS75] if symmetric positive definite preconditioners are used, or sQMR [FN94, FN95] if a certain class of symmetric preconditioners is used, or using GMRES [SS86] or other Krylov subspace methods for nonsymmetric systems if general preconditioners are used. The system (1.1.6) can be solved using the preconditioned conjugate gradient (CG) method. See [Gre97, Saa03, TB97, Dem97, DDSV98, BBC+93, GV96] for additional details on Krylov subspace methods.

The application of $\widehat{\mathbf{Q}}$ to a vector $\mathbf{v}$ requires the application of $\mathbf{A}^{-1}$ and $\mathbf{A}^{-T}$ to a vector. For our application, these tasks correspond to the solution of elliptic partial differential equations. If these tasks are performed using iterative methods, only an approximation $\widetilde{\widehat{\mathbf{Q}}\mathbf{v}}$ of $\widehat{\mathbf{Q}}\mathbf{v}$ is available. The quality of the approximation $\widetilde{\widehat{\mathbf{Q}}\mathbf{v}}$ depends on the accuracy with which the systems involving $\mathbf{A}$ and $\mathbf{A}^T$ are solved. Hence, in practice, one does not solve (1.1.6), but

$$(\widehat{\mathbf{Q}} + \mathbf{E})\,\mathbf{u} = \widehat{\mathbf{g}}, \tag{1.1.7}$$

where the perturbation $\mathbf{E}$ depends on the iterative method applied to solve systems involving $\mathbf{A}$ and $\mathbf{A}^T$. The perturbation $\mathbf{E}$ may be nonsymmetric. The size of $\mathbf{E}$ depends on the accuracy with which the systems involving $\mathbf{A}$ and $\mathbf{A}^T$ are solved. The impact of the nonsymmetry of $\mathbf{E}$ on the conjugate residual methods for (1.1.6) is studied in [MS92]. Instead of solving (1.1.6), we prefer to solve (1.1.5). Matrix-vector applications with the KKT matrix in (1.1.5) can be computed exactly (except for floating point arithmetic). Approximate inverses of $\mathbf{A}$ and $\mathbf{A}^T$ may enter in the execution of preconditioners for (1.1.5).

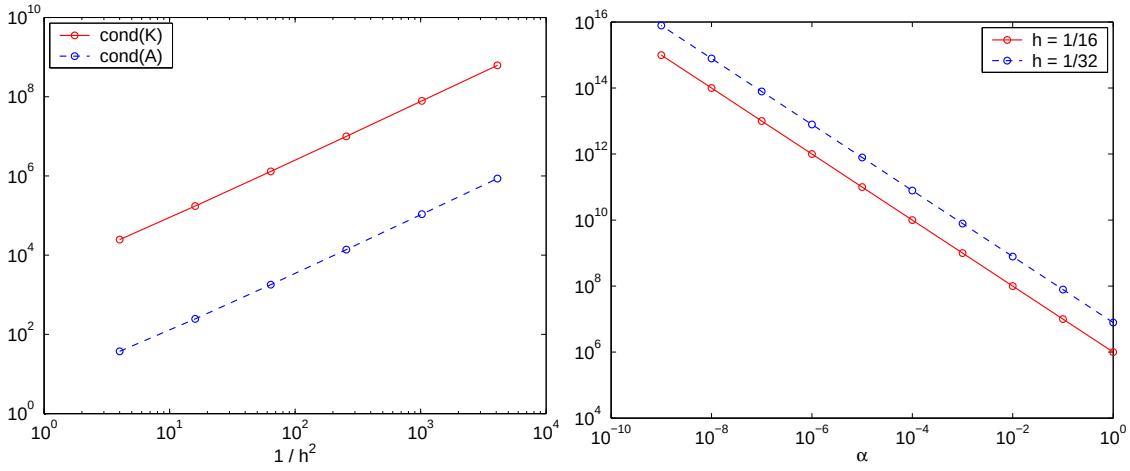

Figure 1.1: Left: Condition number of $\mathbf{K}$ and $\mathbf{A}$ grows as $O(h^{-2})$. Right: Condition number of $\mathbf{K}$ depends linearly on on the regularization parameter α.

The efficient solution of KKT systems (1.1.5) requires the construction of preconditioners. In Figure 1.1, we plot the condition number of the KKT matrix resulting from the finite element discretization of (1.1.1) on $\Omega = (-1, 1)^2$ using piecewise linear finite element on triangles for both state y and control u. The conditioning of the resulting KKT matrix is influenced by two parameters, the mesh size h and the regularization parameter α. It is known that second order elliptic PDEs lead to stiffness matrices $\mathbf{A}$ with condition number that grows as $O(h^{-2})$ for all space dimensions ([AB01, Sec. 5.5].) The KKT matrix $\mathbf{K}$ inherits this property from the $\mathbf{A}$ block as the mesh parameter h decreases (left plot of Figure 1.1, with constant $\alpha = 0.1$).

Another source of ill-conditioning is a small value for the regularization parameter α. Since the problem (1.1.3) is not strictly convex for $\alpha = 0$, we expect the optimality system to be increasingly ill-conditioned as α approaches zero (right plot of Figure 1.1.) It is important to study the behavior of the preconditioned system for varying mesh size h and regularization parameter α.

Preconditioners for (1.1.5) have been developed in [AH03, BH98, BS01, BG99, BG00, HPS02, GMPS92, GG03, GHN01, HA00, KGW00, LV98, MN00, MGW00, MS00]. We note that although (1.1.5) corresponds to a saddle point problem, its structure is quite different from the saddle point problems arising, e.g., from the Stokes problem (see, e.g., [LW03, SW94]) or from mixed finite element discretizations of elliptic PDEs [BF91].

The papers [GMPS92, GHN01, KGW00, LV98] discuss preconditioners that use sparse matrix decompositions of KKT matrices obtained from the matrix in (1.1.5) by modification of blocks. For very large PDE constrained problems the direct factorization is no longer feasible or competitive because of storage requirements and run-time. Our domain decomposition preconditioners do not require direct matrix factorizations. The preconditioners in [AH03, BH98, BS01, BG99, BG00] are based on block decompositions of the matrix in (1.1.5). Their effectiveness depends on how well $\mathbf{A}^{-1}$ and $\mathbf{A}^{-T}$ can be approximated (e.g., by applying iterative solvers to the solution of systems with $\mathbf{A}$ and $\mathbf{A}^T$) and on the availability of an effective preconditioner for the reduced Hessian. We will provide a few more details on this class of preconditioners in Section 1.2. Deriving preconditioners for the reduced Hessian for general problems is difficult. The domain decomposition preconditioners derived in this thesis do not require preconditioners for the reduced Hessian. The preconditioners in [BH98, MGW00] can only be applied to (1.1.5) if $\mathbf{M}$ is invertible. This is the case for our model problems (1.1.1) and (1.1.2) only if $\Omega_o = \Omega$. Invertibility of $\mathbf{M}$ is not required for our preconditioners. Multigrid solvers for (1.1.5) for classes of linear-quadratic PDE constrained problems are discussed in [AH03, HPS02, AH03, MS00].

The preconditioners in this thesis are based on domain decomposition of the elliptic control problem. They are motivated by the successful application of domain decomposition methods to the parallel solution of partial differential equations (see, e.g., the books [SBG96, QV99] and the proceedings of domain decomposition conferences, many of which are available electronically at `http://www.ddm.org/`). The domain decomposition preconditioners introduced and studied in this thesis incorporate the structure of elliptic linear-quadratic optimal control problems (1.1.3). Their execution requires the solution of elliptic linear-quadratic optimal control problems restricted to subdomains. These subdomain solves can be performed in parallel. If individual subdomain solves are performed iteratively, existing preconditioners for (1.1.5) can be incorporated for these tasks. In this sense, our domain decomposition preconditioners complement existing preconditioners for (1.1.5). Before we outline our domain decomposition preconditioners in more detail in Section 1.3, we briefly review prior work on domain decomposition based methods for the solution of elliptic linear-quadratic optimal control problems.

1.2 Domain Decomposition Methods

Domain decomposition based iterative methods for the solution of (1.1.3) can be split into three groups, depending on how much the domain decomposition is integrated with the optimization.

Group 1. The methods in the first group apply preconditioned Krylov subspace methods to (1.1.5) (or (1.1.6)) and use domain decomposition methods for the approximate solution of systems involving $\mathbf{A}$ and $\mathbf{A}^T$. Hence, domain decomposition is only used at the PDE level. The methods in [BG99, BG00] belong in this group; they will be reviewed in more detail below.

Group 2. The methods in the second group use domain decomposition to reformulate the optimization problem (1.1.3) or its discretization (1.1.4). The reformulations

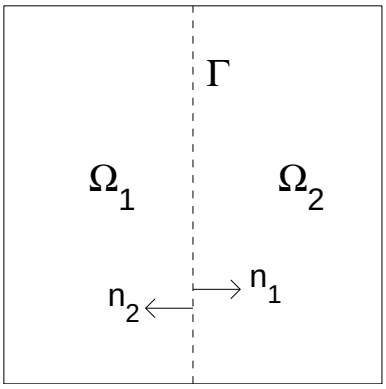

Figure 1.2: A two-subdomain partitioning of Ω into Ω_1 and Ω_2, with common interface Γ and outward normal vectors n_1 and n_2.

change the number of variables visible to the outer iteration. For example, a reformulation may eliminate state variables in the interior of subdomains by viewing them as functions of the controls in that subdomain as well as state variables on the subdomain interfaces. The optimization variables visible to the outer iteration then consist of the control variables and the state variables restricted to the subdomain interfaces. The approaches in [Bou98, DL94, LP98] belong in the second group. We will review them in more detail below. The linear systems that need to solved during the execution of the methods in the second group are related to PDEs restricted to subdomains.

Group 3. The methods in the third group apply domain decomposition to the optimality system. Hence, domain decomposition is integrated with the optimization. The linear systems that need to solved during the execution of the methods in the third group are related to optimal control problems restricted to subdomains. The methods [Ben94, Ben96, Ben98] as well as the methods proposed in this thesis belong to this group.

In all three groups, domain decomposition introduces parallelism into the optimization, but affects the outer iteration differently.

To review existing domain decomposition methods for optimal control problem in more detail, we consider the model problem (1.1.1) or its finite element discretization (1.1.4) . To simplify the presentation, we frequently use a two-subdomain formulation to state and compare the these approaches. The domain $\Omega \subset \mathbb{R}^2$ is partitioned into Ω_1 and Ω_2 as sketched in Figure 1.2, with subdomain interface Γ and outward normal vectors n_1 and n_2 on the interface.

1.2.1 Group 1

We consider (1.1.5) and note that

$$\begin{bmatrix} \mathbf{M} & \mathbf{0} & \mathbf{A}^T \\ \mathbf{0} & \alpha\mathbf{Q} & \mathbf{B}^T \\ \mathbf{A} & \mathbf{B} & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{M}\mathbf{A}^{-1} & \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{I} & \mathbf{B}^T\mathbf{A}^{-T} \\ \mathbf{I} & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \widehat{\mathbf{Q}} & \mathbf{0} \\ \mathbf{0} & -\mathbf{M}\mathbf{A}^{-1}\mathbf{B} & \mathbf{A}^T \end{bmatrix}, \quad (1.2.1)$$

where $\widehat{\mathbf{Q}}$ is the reduced Hessian (cf. (1.1.6)). After a permutation, the matrices on the right hand side of (1.2.1) are upper/lower block triangular. The application of inverses of these block triangular matrices to a given vector requires the application of inverses of $\mathbf{A}$, $\mathbf{A}^T$ and $\widehat{\mathbf{Q}}$. A preconditioner is obtained if we approximate the application of these sub-block inverses. This leads to

$$\begin{bmatrix} \mathbf{M} & \mathbf{0} & \mathbf{A}^T \\ \mathbf{0} & \alpha\mathbf{Q} & \mathbf{B}^T \\ \mathbf{A} & \mathbf{B} & \mathbf{0} \end{bmatrix}^{-1} \approx \begin{bmatrix} \widetilde{\mathbf{A}} & \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \widetilde{\widehat{\mathbf{Q}}} & \mathbf{B}^T \\ \mathbf{0} & -\mathbf{M}\widetilde{\mathbf{A}}^{-1}\mathbf{B} & \widetilde{\mathbf{A}}^T \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{M}\widetilde{\mathbf{A}}^{-1} & \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{I} & \mathbf{B}^T\widetilde{\mathbf{A}}^{-T} \\ \mathbf{I} & \mathbf{0} & \mathbf{0} \end{bmatrix}^{-1}. \quad (1.2.2)$$

The right hand side in (1.2.2) is used to construct (left, right, split) preconditioners for the KKT matrix. This is one of the approaches studied in [BG99, BG00]. Related block factorizations of the KKT matrix (1.1.5) are introduced in [BH98, BS01, HA00]. One may reduce the expense of the application of the preconditioner by dropping the $\mathbf{M}\widetilde{\mathbf{A}}^{-1}$ blocks in (1.2.2) (see [BG99, BG00]).

The quality with which the product on the right hand side in (1.2.2) approximates the inverse of the KKT matrix (1.1.5) depends on how well the inverses of $\widetilde{\mathbf{A}}$, $\widetilde{\mathbf{A}}^T$ and

$\widetilde{\widehat{\mathbf{Q}}}$ approximate those of $\mathbf{A}$, $\mathbf{A}^T$ and $\widehat{\mathbf{Q}}$. The construction of good and computationally inexpensive approximations of the reduced Hessian $\widehat{\mathbf{Q}}$ is difficult in general. In [BG99, BG00] quasi-Newton preconditioners [MN00] or the use of two-step stationary iterative methods for the approximate solution of $\widehat{\mathbf{Q}}\mathbf{v} = \mathbf{r}$ are proposed. In some cases, problem specific approximations of the reduced Hessian can be found [BS01]. In [BG99, BG00] the approximate inverses of $\mathbf{A}$, $\mathbf{A}^T$ correspond to the use of PDE domain decomposition methods applied to the solution of systems with matrices $\mathbf{A}$ or $\mathbf{A}^T$.

1.2.2 Group 2

We consider the model problem (1.1.1). On subdomain Ω_i we consider the differential equation

$$\begin{aligned} -\Delta y_i(x) &= f(x) + u(x) && \text{in } \Omega_i, \\ y_i(x) &= 0 && \text{on } \partial\Omega_i \backslash \Gamma, \\ y_i(x) &= y_\Gamma(x) && \text{on } \Gamma. \end{aligned} \tag{1.2.3}$$

Given $u \in L^2(\Omega_i)$ and $y_\Gamma \in H_{00}^{1/2}(\Gamma)$, the PDE (1.2.3) has a unique solution. Let

$$H_{00}^{1/2}(\Gamma) \times L^2(\Omega_i) \ni (y_\Gamma, u) \mapsto y_i(y_\Gamma, u) \in \left\{ v \in H^1(\Omega_i) \ : \ v = 0 \text{ on } \partial\Omega_i \backslash \Gamma \right\}$$

be the mapping of control and interface data into the solution of (1.2.3). If

$$\frac{\partial}{\partial n_1}(y_1(y_\Gamma, u))(x) = -\frac{\partial}{\partial n_2}(y_2(y_\Gamma, u))(x) \quad \text{on } \Gamma, \tag{1.2.4}$$

then $y \in H_0^1(\Omega)$ defined as $y|_{\Omega_i} = y_i$, $i = 1, 2$ solves the state equation

$$\begin{aligned} -\Delta y(x) &= f(x) + u(x) && \text{in } \Omega, \\ y(x) &= 0 && \text{on } \partial\Omega \end{aligned}$$

(see, e.g., [QV99, Ch. 1]). Thus, the model control problem (1.1.1) can now be equivalently formulated as

$$\begin{aligned} \min \quad & \textstyle\sum_{i=1}^{2} \left(\frac{1}{2} \int_{\Omega_i} [(y_i(y_\Gamma, u))(x) - y_d(x)]^2 \, dx + \frac{\alpha}{2} \int_{\Omega_i} (u(x))^2 \, dx \right) \\ \text{s.t.} \quad & \frac{\partial}{\partial n_1}(y_1(y_\Gamma, u))(x) = -\frac{\partial}{\partial n_2}(y_2(y_\Gamma, u))(x) \quad \text{on } \Gamma. \end{aligned} \tag{1.2.5}$$

The model control problem (1.1.1) is an optimization problem in the variables $u \in L^2(\Omega)$, $y \in H_0^1(\Omega)$. Its reformulation (1.2.5) is an optimization problem in the variables $u \in L^2(\Omega)$, $y_\Gamma \in H_{00}^{1/2}(\Gamma)$.

Using the domain decomposition structure, the stiffness matrix $\mathbf{A}$ corresponding to the state equation (1.1.1) may be written as

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{II}^1 & & \mathbf{A}_{I\Gamma}^1 \\ & \mathbf{A}_{II}^2 & \mathbf{A}_{I\Gamma}^2 \\ \mathbf{A}_{\Gamma I}^1 & \mathbf{A}_{\Gamma I}^2 & \mathbf{A}_{\Gamma\Gamma}^{1+2} \end{pmatrix}$$

where $\mathbf{A}_{II}^i$ contains the stiffness matrix contributions from the interactions of nodel basis functions in the interior of Ω_i, $\mathbf{A}_{I\Gamma}^i$ contains the stiffness matrix contributions from the interactions of nodal basis functions in the interior of Ω_i with those on the interface boundary Γ, $\mathbf{A}_{\Gamma I}^i = (\mathbf{A}_{I\Gamma}^i)^T$, and $\mathbf{A}_{\Gamma\Gamma}^i$ contains the contributions from subdomain Ω_i to the area integral for the entries in $\mathbf{A}_{\Gamma\Gamma}$. (For more details see Section 4.3.1 and [QV99, Ch. 2]). Similarly, the vector representing y can be partitioned into blocks as

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_I^1 \\ \mathbf{y}_I^2 \\ \mathbf{y}_\Gamma \end{pmatrix}, \tag{1.2.6}$$

where $\mathbf{y}_I^i$ corresponds to the nodes in the interior of Ω_i and $\mathbf{y}_\Gamma$ corresponds to the nodes on Γ. Analogous notation is used for the remaining matrices $\mathbf{B}, \mathbf{M}$, etc. Exposing

the domain decomposition structure, the KKT system (1.1.5) is now written as

$$\left(\begin{array}{ccc|cc|ccc}
\mathbf{M}^1_{II} & & \mathbf{M}^1_{I\Gamma} & & & \mathbf{A}^1_{II} & & \mathbf{A}^1_{I\Gamma} \\
& \mathbf{M}^2_{II} & \mathbf{M}^2_{I\Gamma} & & & & \mathbf{A}^2_{II} & \mathbf{A}^2_{I\Gamma} \\
\mathbf{M}^1_{\Gamma I} & \mathbf{M}^2_{\Gamma I} & \mathbf{M}^{1+2}_{\Gamma\Gamma} & & & \mathbf{A}^1_{\Gamma I} & \mathbf{A}^2_{\Gamma I} & \mathbf{A}^{1+2}_{\Gamma\Gamma} \\
\hline
& & & \alpha\mathbf{Q}^1_{II} & & \mathbf{B}^1_{II} & & \mathbf{B}^1_{I\Gamma} \\
& & & & \alpha\mathbf{Q}^2_{II} & & \mathbf{B}^2_{II} & \mathbf{B}^2_{I\Gamma} \\
\hline
\mathbf{A}^1_{II} & & \mathbf{A}^1_{I\Gamma} & \mathbf{B}^1_{II} & & & & \\
& \mathbf{A}^2_{II} & \mathbf{A}^2_{I\Gamma} & & \mathbf{B}^2_{II} & & & \\
\mathbf{A}^1_{\Gamma I} & \mathbf{A}^2_{\Gamma I} & \mathbf{A}^{1+2}_{\Gamma\Gamma} & \mathbf{B}^1_{\Gamma I} & \mathbf{B}^2_{\Gamma I} & & &
\end{array}\right)
\left(\begin{array}{c}
\mathbf{y}^1_I \\ \mathbf{y}^2_I \\ \mathbf{y}_\Gamma \\ \hline \mathbf{u}^1_I \\ \mathbf{u}^2_I \\ \hline \mathbf{p}^1_I \\ \mathbf{p}^2_I \\ \mathbf{p}_\Gamma
\end{array}\right)
=
\left(\begin{array}{c}
\mathbf{c}^1_I \\ \mathbf{c}^2_I \\ \mathbf{c}^{1+2}_\Gamma \\ \hline \mathbf{0} \\ \mathbf{0} \\ \hline \mathbf{b}^1_I \\ \mathbf{b}^2_I \\ \mathbf{b}^{1+2}_\Gamma
\end{array}\right). \tag{1.2.7}$$

The finite element discretization of (1.2.3) leads to

$$\mathbf{A}^i_{II}\mathbf{y}^i_I + \mathbf{A}^i_{I\Gamma}\mathbf{y}_\Gamma + \mathbf{B}^i_{II}\mathbf{u}^i_I = \mathbf{b}^i_I,$$

where $\mathbf{u}^i_I$ represents the restriction of $u|_{\Omega_i}$. Hence

$$\mathbf{y}^i_I = (\mathbf{A}^i_{II})^{-1}\left(\mathbf{b}^i_I - \mathbf{B}^i_{II}\mathbf{u}^i_I - \mathbf{A}^i_{I\Gamma}\mathbf{y}_\Gamma\right). \tag{1.2.8}$$

The interface condition (1.2.4) corresponds to

$$\overline{\mathbf{A}}\mathbf{y}_\Gamma + \overline{\mathbf{B}}_1\mathbf{u}^1_I + \overline{\mathbf{B}}_2\mathbf{u}^2_I = \overline{\mathbf{b}},$$

where

$$\begin{aligned}
\overline{\mathbf{A}} &= \mathbf{A}^{1+2}_{\Gamma\Gamma} - \mathbf{A}^1_{\Gamma I}(\mathbf{A}^1_{II})^{-1}\mathbf{A}^1_{I\Gamma} - \mathbf{A}^2_{\Gamma I}(\mathbf{A}^2_{II})^{-1}\mathbf{A}^2_{I\Gamma}, \\
\overline{\mathbf{B}}_i &= \mathbf{B}^i_{\Gamma I} - \mathbf{A}^i_{\Gamma I}(\mathbf{A}^i_{II})^{-1}\mathbf{B}^i_{II}, \quad i = 1, 2, \\
\overline{\mathbf{b}} &= \mathbf{b}^{1+2}_\Gamma - (\mathbf{A}^1_{II})^{-1}\mathbf{b}^1_I - (\mathbf{A}^2_{II})^{-1}\mathbf{b}^2_I.
\end{aligned}$$

Using the domain decomposition structure and (1.2.8), the discretized objective func-

tion $\frac{1}{2}\mathbf{y}^T\mathbf{M}\mathbf{y} + \frac{\alpha}{2}\mathbf{u}^T\mathbf{Q}\mathbf{u} - \mathbf{c}^T\mathbf{y}$ is written as

$$\begin{aligned}
&\frac{1}{2}\left((\mathbf{y}_I^1)^T\mathbf{M}_{II}^1\mathbf{y}_I^1 + 2(\mathbf{y}_I^1)^T\mathbf{M}_{I\Gamma}^1\mathbf{y}_\Gamma + (\mathbf{y}_I^2)^T\mathbf{M}_{II}^2\mathbf{y}_I^2 + 2(\mathbf{y}_I^2)^T\mathbf{M}_{I\Gamma}^2\mathbf{y}_\Gamma + (\mathbf{y}_\Gamma)^T\mathbf{M}_{\Gamma\Gamma}^{1+2}\mathbf{y}_\Gamma\right)\\
&+\frac{\alpha}{2}(\mathbf{u}_I^1)^T\mathbf{Q}_{II}^1\mathbf{u}_I^1 + \frac{\alpha}{2}(\mathbf{u}_I^2)^T\mathbf{Q}_{II}^2\mathbf{u}_I^2 - (\mathbf{c}_I^1)^T\mathbf{y}_I^1 - (\mathbf{c}_I^2)^T\mathbf{y}_I^2 - (\mathbf{c}_\Gamma^{1+2})^T\mathbf{y}_\Gamma\\
=\ & \frac{1}{2}\mathbf{y}_\Gamma^T\overline{\mathbf{M}}\mathbf{y}_\Gamma + \mathbf{y}_\Gamma^T\overline{\mathbf{N}}\begin{pmatrix}\mathbf{u}_I^1\\ \mathbf{u}_I^2\end{pmatrix} + \frac{1}{2}\begin{pmatrix}\mathbf{u}_I^1\\ \mathbf{u}_I^2\end{pmatrix}^T\overline{\mathbf{Q}}\begin{pmatrix}\mathbf{u}_I^1\\ \mathbf{u}_I^2\end{pmatrix} - \overline{\mathbf{c}}^T\mathbf{y}_\Gamma - \overline{\mathbf{d}}^T\begin{pmatrix}\mathbf{u}_I^1\\ \mathbf{u}_I^2\end{pmatrix}\\
&-(\mathbf{c}_I^1)^T(\mathbf{A}_{II}^1)^{-1}\mathbf{b}_I^1 - (\mathbf{c}_I^2)^T(\mathbf{A}_{II}^2)^{-1}\mathbf{b}_I^2
\end{aligned}$$

where

$$\begin{aligned}
\overline{\mathbf{M}} &= \begin{pmatrix}-(\mathbf{A}_{II}^1)^{-1}\mathbf{A}_{I\Gamma}^1\\ -(\mathbf{A}_{II}^2)^{-1}\mathbf{A}_{I\Gamma}^2\\ \mathbf{I}\end{pmatrix}^T\begin{pmatrix}\mathbf{M}_{II}^1 & & \mathbf{M}_{I\Gamma}^1\\ & \mathbf{M}_{II}^2 & \mathbf{M}_{I\Gamma}^2\\ \mathbf{M}_{\Gamma I}^1 & \mathbf{M}_{\Gamma I}^2 & \mathbf{M}_{\Gamma\Gamma}^{1+2}\end{pmatrix}\begin{pmatrix}-(\mathbf{A}_{II}^1)^{-1}\mathbf{A}_{I\Gamma}^1\\ -(\mathbf{A}_{II}^2)^{-1}\mathbf{A}_{I\Gamma}^2\\ \mathbf{I}\end{pmatrix},\\
\overline{\mathbf{N}} &= \begin{pmatrix}-(\mathbf{A}_{II}^1)^{-1}\mathbf{A}_{I\Gamma}^1\\ -(\mathbf{A}_{II}^2)^{-1}\mathbf{A}_{I\Gamma}^2\\ \mathbf{I}\end{pmatrix}^T\begin{pmatrix}\mathbf{M}_{II}^1 & \\ & \mathbf{M}_{II}^2\\ \mathbf{M}_{\Gamma I}^1 & \mathbf{M}_{\Gamma I}^2\end{pmatrix}\begin{pmatrix}-(\mathbf{A}_{II}^1)^{-1}\mathbf{B}_{II}^1 & \\ & -(\mathbf{A}_{II}^2)^{-1}\mathbf{B}_{II}^2\end{pmatrix},\\
\overline{\mathbf{Q}} &= \begin{pmatrix}\alpha\mathbf{Q}_{II}^1 + (\mathbf{B}_{II}^1)^T(\mathbf{A}_{II}^1)^{-T}\mathbf{M}_{II}^1(\mathbf{A}_{II}^1)^{-1}\mathbf{B}_{II}^1 & \\ & \alpha\mathbf{Q}_{II}^2 + (\mathbf{B}_{II}^2)^T(\mathbf{A}_{II}^2)^{-T}\mathbf{M}_{II}^2(\mathbf{A}_{II}^2)^{-1}\mathbf{B}_{II}^2\end{pmatrix},\\
\overline{\mathbf{c}} &= \mathbf{c}_\Gamma^{1+2} - (\mathbf{A}_{I\Gamma}^1)^T(\mathbf{A}_{II}^1)^{-T}\mathbf{c}_I^1 - (\mathbf{A}_{I\Gamma}^2)^T(\mathbf{A}_{II}^2)^{-T}\mathbf{c}_I^2,\\
\overline{\mathbf{d}} &= \begin{pmatrix}(\mathbf{B}_{II}^1)^T(\mathbf{A}_{II}^1)^{-T}\mathbf{c}_I^1\\ (\mathbf{B}_{II}^2)^T(\mathbf{A}_{II}^2)^{-T}\mathbf{c}_I^2\end{pmatrix}.
\end{aligned}$$

The discretization of (1.2.5) is given by

$$\begin{aligned}
\min\quad & \tfrac{1}{2}\mathbf{y}_\Gamma^T\overline{\mathbf{M}}\mathbf{y}_\Gamma + \mathbf{y}_\Gamma^T\overline{\mathbf{N}}\begin{pmatrix}\mathbf{u}_I^1\\ \mathbf{u}_I^2\end{pmatrix} + \tfrac{1}{2}\begin{pmatrix}\mathbf{u}_I^1\\ \mathbf{u}_I^2\end{pmatrix}^T\overline{\mathbf{Q}}\begin{pmatrix}\mathbf{u}_I^1\\ \mathbf{u}_I^2\end{pmatrix} - \overline{\mathbf{c}}^T\mathbf{y}_\Gamma - \overline{\mathbf{d}}^T\begin{pmatrix}\mathbf{u}_I^1\\ \mathbf{u}_I^2\end{pmatrix}\\
\text{s.t.}\quad & \overline{\mathbf{A}}\mathbf{y}_\Gamma + \overline{\mathbf{B}}_1\mathbf{u}_I^1 + \overline{\mathbf{B}}_2\mathbf{u}_I^2 = \overline{\mathbf{b}},
\end{aligned} \tag{1.2.9}$$

where we have dropped the constant $-(\mathbf{c}_I^1)^T(\mathbf{A}_{II}^1)^{-1}\mathbf{b}_I^1 - (\mathbf{c}_I^2)^T(\mathbf{A}_{II}^2)^{-1}\mathbf{b}_I^2$ from the objective function. The necessary and sufficient optimality conditions for (1.2.9) lead

to the linear system

$$\left(\begin{array}{c|c|c} \overline{\mathbf{M}} & \overline{\mathbf{N}} & \overline{\mathbf{A}}^T \\ \hline \overline{\mathbf{N}}^T & \overline{\mathbf{Q}} & \begin{pmatrix} \overline{\mathbf{B}}_1^T \\ \overline{\mathbf{B}}_2^T \end{pmatrix} \\ \hline \overline{\mathbf{A}} & (\overline{\mathbf{B}}_1 \mid \overline{\mathbf{B}}_2) & \end{array}\right) \left(\begin{array}{c} \mathbf{y}_\Gamma \\ \hline \mathbf{u}_I^1 \\ \mathbf{u}_I^2 \\ \hline \mathbf{p}_\Gamma \end{array}\right) = \left(\begin{array}{c} \overline{\mathbf{c}} \\ \hline \overline{\mathbf{d}} \\ \hline \overline{\mathbf{b}} \end{array}\right). \tag{1.2.10}$$

We denote the system matrix in (1.2.10) by $\overline{\mathbf{K}}$. It can be seen that $\overline{\mathbf{K}}$ is the Schur complement of (1.2.7) with respect to $\mathbf{y}_\Gamma, \mathbf{u}_I^1, \mathbf{u}_I^2, \mathbf{p}_\Gamma$.

Figure 1.3 show the dependence of the condition number for the KKT matrix $\overline{\mathbf{K}}$ in the optimality system for (1.2.9) versus finite element mesh size h and regularization parameter α. The problem data are the same as those used in Figure 1.1. In particular, the condition number plots for $\mathbf{K}$ in the respective subplots in Figure 1.3 and in Figure 1.3 are identical. In the right plot of Figure 1.3 we use the mesh size $h = 1/16$. The left plot of Figure 1.3 reveals that for fixed regularization parameter α (in the left plot of Figure 1.3, $\alpha = 10^{-1}$) the Schur complement matrix $\overline{\mathbf{K}}$ is better conditioned than the original KKT matrix $\mathbf{K}$. Still, the condition number of $\overline{\mathbf{K}}$ is large even for moderate h and grows like $1/h^2$ as $h \to 0$. The right plot of Figure 1.3 shows that for fixed mesh size h (in the left plot of Figure 1.3, $h = 1/16$) the conditioning of the Schur complement matrix $\overline{\mathbf{K}}$ is independent of the regularization parameter α. This is due to the fact that in $\overline{\mathbf{Q}}$ the block diagonals are of the form $\alpha\mathbf{Q}_{II}^i + (\mathbf{B}_{II}^i)^T(\mathbf{A}_{II}^i)^{-T}\mathbf{M}_{II}^i(\mathbf{A}_{II}^i)^{-1}\mathbf{B}_{II}^i$, $i = 1, 2$, i.e., for small $\alpha > 0$, the second term dominates. Overall, the condition number of the Schur complement matrix $\overline{\mathbf{K}}$ is large even for moderate values of h and α and preconditioning of $\overline{\mathbf{K}}$ is needed for an efficient solution of (1.2.9).

A domain decomposition reformulation of the type (1.2.5) is used in [DL94]. However, their optimization problem is the identification of the transmissivity from pressure data. This is a nonlinear optimization problem and the interest in [DL94] was the impact of problem formulation on the nonlinear outer optimization iteration.

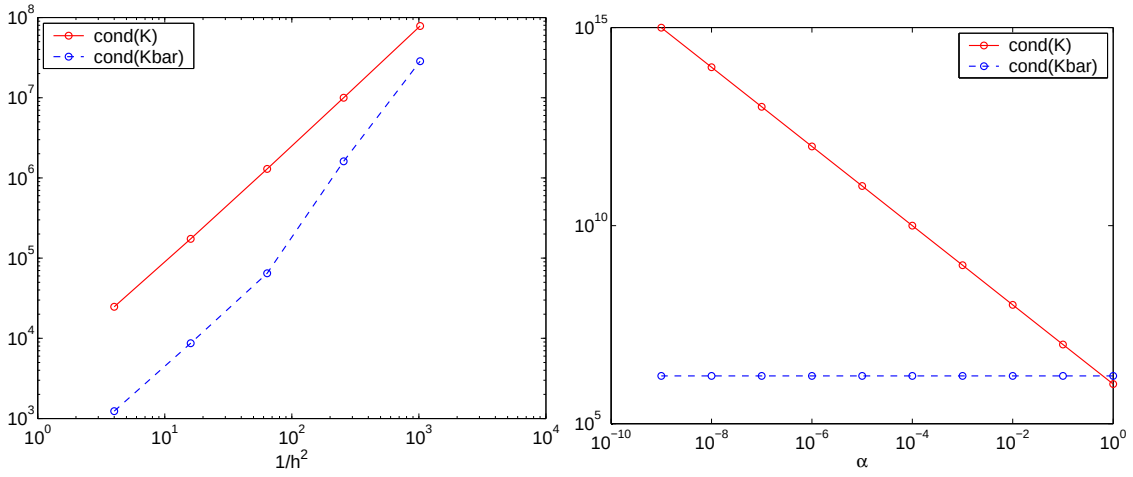

Figure 1.3: Left: Condition numbers of $\mathbf{K}$ and $\overline{\mathbf{K}}$ grow as $O(h^{-2})$. Right: Condition numbers of $\mathbf{K}$ and $\overline{\mathbf{K}}$ depend linearly on the regularization parameter α.

Instead of using (1.2.3), one may consider the map

$$H^{-1/2}(\Gamma) \times L^2(\Omega_i) \ni (z,u) \mapsto y_i(z,u) \in \left\{ v \in H^1(\Omega_i) \ : \ v = 0 \text{ on } \partial\Omega_i \backslash \Gamma \right\}$$

where y_i is the solution of

$$\begin{aligned} -\Delta y_i(x) &= f(x) + u(x) && \text{in } \Omega_i, \\ y_i(x) &= 0 && \text{on } \partial\Omega_i \backslash \Gamma, \\ \tfrac{\partial}{\partial n_i} y_i(x) &= (-1)^i z(x) && \text{on } \Gamma. \end{aligned} \tag{1.2.11}$$

If

$$y_1(z,u)(x) = y_2(z,u)(x) \quad \text{on } \Gamma,$$

then $y \in H_0^1(\Omega)$ defined as $y|_{\Omega_i} = y_i$, $i = 1, 2$ solves the state equation

$$\begin{aligned} -\Delta y(x) &= f(x) + u(x) && \text{in } \Omega, \\ y(x) &= 0 && \text{on } \partial\Omega \end{aligned}$$

Thus, the model control problem (1.1.1) can also be equivalently formulated as

$$\begin{aligned} \min \quad & \textstyle\sum_{i=1}^{2} \left(\frac{1}{2} \int_{\Omega_i} [(y_i(z,u))(x) - y_d(x)]^2 \, dx + \frac{\alpha}{2} \int_{\Omega_i} (u(x))^2 \, dx \right) \\ \text{s.t.} \quad & y_1(z,u)(x) = y_2(z,u)(x) \quad \text{on } \Gamma. \end{aligned} \tag{1.2.12}$$

This decomposition is used in [Bou98], where an augmented Lagrangian approach is used to solve (1.2.12). This decomposition is also related to the overlapping domain decomposition method in [LP98].

Note that if more than two subdomain are used, (1.2.11) is only well defined if each subdomain Ω_i shares a boundary segment with $\partial\Omega$. This is the case for the decompositions used in [Bou98, LP98].

The papers [Bou98, LP98] prove convergence of their methods for solving (1.2.12) or the related optimization problems considered in the respective papers, but the convergence proofs are modifications of the convergence results of gradient-type or augmented Lagrangian-type methods applied to the domain decomposition reformulation of the optimization problem. The given convergence proofs do not reveal the dependence of the convergence on mesh size h and subdomain size H. The numerical results reported in [Bou98, LP98] also do not illustrate the dependence of the convergence on mesh size h and subdomain size H.

1.2.3 Group 3

Benamou ([Ben94, Ben96, Ben98]) introduced a domain decomposition approach that is applied to the optimality system, thus solving the PDE and the optimization problem concurrently.

We consider (1.1.1) and the corresponding necessary and sufficient optimality conditions

$$\begin{aligned} -\Delta y &= f + u && \text{in } \Omega, \\ -\Delta p &= y - y_d && \text{in } \Omega, \\ \alpha u + p &= 0 && \text{in } \Omega, \\ y &= 0 && \text{on } \partial\Omega, \\ p &= 0 && \text{on } \partial\Omega. \end{aligned}$$

By eliminating the control, the model problem is reduced to a coupled system of state

and adjoint PDEs

$$\begin{aligned} -\Delta y &= f - \alpha^{-1} p && \text{in } \Omega, \\ -\Delta p &= y - y_d && \text{in } \Omega. \end{aligned} \tag{1.2.13}$$

The system (1.2.13) is equivalent to

$$\begin{aligned} -\Delta y_i &= f - \alpha^{-1} p_i && \text{in } \Omega_i, \\ -\Delta p_i &= y_i - y_d && \text{in } \Omega_i, \\ y_i &= 0 && \text{on } \partial\Omega_i \backslash \Gamma, \\ p_i &= 0 && \text{on } \partial\Omega_i \backslash \Gamma, \end{aligned} \tag{1.2.14}$$

$i = 1, 2$, together with the Robin transmission conditions

$$\begin{aligned} \tfrac{\partial}{\partial n_1} y_1 + \lambda p_1 &= -\tfrac{\partial}{\partial n_2} y_2 + \lambda p_2 && \text{on } \Gamma, \\ \tfrac{\partial}{\partial n_1} p_1 - \lambda y_1 &= -\tfrac{\partial}{\partial n_2} p_2 - \lambda y_2 && \text{on } \Gamma, \end{aligned} \tag{1.2.15}$$

for suitable $\lambda \in \mathbb{R}$. The system (1.2.14)-(1.2.15) may be solved using the following iteration in which for a given guess $y_i^{(k)}, p_i^{(k)}$, $i = 1, 2$, of the solution of (1.2.14)-(1.2.15), the new guess $y_i^{(k+1)}, p_i^{(k+1)}$, $i = 1, 2$, is computed by solving (in parallel)

$$\begin{aligned} -\Delta y_i^{(k+1)} &= f - \alpha^{-1} p_i^{(k+1)} && \text{in } \Omega_i, \\ -\Delta p_i^{(k+1)} &= y_i^{(k+1)} - y_d && \text{in } \Omega_i, \\ y_i^{(k+1)} &= 0 && \text{on } \partial\Omega_i \backslash \Gamma, \\ p_i^{(k+1)} &= 0 && \text{on } \partial\Omega_i \backslash \Gamma, \end{aligned} \tag{1.2.16}$$

together with the Robin transmission conditions

$$\begin{aligned} \tfrac{\partial}{\partial n_i} y_i^{(k+1)} + \lambda p_i^{(k+1)} &= -\tfrac{\partial}{\partial n_j} y_j^{(k)} + \lambda p_j^{(k)} && \text{on } \Gamma, \\ \tfrac{\partial}{\partial n_i} p_i^{(k+1)} - \lambda y_i^{(k+1)} &= -\tfrac{\partial}{\partial n_j} p_j^{(k)} - \lambda y_j^{(k)} && \text{on } \Gamma, \end{aligned} \tag{1.2.17}$$

$i = 1, 2$, where $j = 2$ if $i = 1$ and $j = 1$ if $i = 2$ ([Ben98]). The weighting parameter λ may depend on the iteration n. This is an extension of the Robin transmission condition for PDEs suggested in [Lio90] (see also [QV99]).

Other Robin-type transmission conditions besides (1.2.15) are possible and lead to slightly different iterative methods ([Ben96, Ben98]). Benamou ([Ben94, Ben96,

Ben98]) proves convergence of the iteration (1.2.16), (1.2.17), but, unfortunately, no convergence rates could be established nor are results available that show the dependence of the iteration on mesh size h and subdomain size H. This corresponds to what seems to be known for the convergence of Robin-Robin-type methods for single PDEs [QV99].

1.3 Proposed Methods

We consider new approaches to using domain decomposition for the solution of linear-quadratic elliptic optimal control problems, where the full optimal control problem is decomposed into smaller control problems. Both overlapping and nonoverlapping methods are considered, which are inspired by methods developed originally for PDE problems. We implement these as one-level and two-level preconditioners for some Krylov iterative methods.

Numerical experiments show that our methods inherit the good convergence properties from domain decomposition methods for PDEs. These include weak dependence on the mesh size parameter h, and, in the two-level case, relative insensitivity to the subdomain size H. In addition, our methods perform well even when the regularization parameter α is set near zero. For our overlapping methods, theoretical results are provided that describe the dependence of the convergence behavior on mesh size h and subdomain size H.

Our two-level overlapping method is proven to have a convergence rate that can be bounded independently of the subdomain size H and element size h, and applies to all regularization parameters $\alpha > 0$. Convergence results for the DD methods for optimal control studied in [Ben94, Ben96, Ben98, Bou98, LP98] do not show dependence of the convergence behavior on H, h, α. Moreover, most previous methods have been formulated only as one-level methods, which suffer from a significant increase in iterations as the number of subdomains is increased.

1.3.1 General Formulation

In Chapter 2, we present a general approach to applying domain decomposition ideas toward solving linear-quadratic optimal control problems. A variational formulation is used in Section 2.1 to describe an abstract control problem over one domain. We verify that if the bilinear forms in the formulation satisfy certain coercivity assumptions then the control problem has a unique solution.

In Section 2.2, we present a subspace decomposition of the optimal control problem, where the space of state, adjoint and control variables are decomposed into subspaces. In a typical implementation, these subspaces are associated with either local subdomains or a coarse grid. This approach is an extension of the well-known subspace methods for analyzing domain decomposition algorithms for PDE problems. For each subspace, we define a few local bilinear forms analogous to the global bilinear forms specifying the one-domain control problem. These forms allow the variational definition of projection operators onto the subspaces. If the local bilinear forms satisfy certain coercivity requirements, then the associated projection operator is shown to be well defined. We combine these local projection operators over all subspaces by extending the additive method from domain decomposition for PDEs. The result is a transformed problem that is shown to be equivalent to the original control problem. We may interpret the linear operator of the transformed problem as a preconditioned version of the original operator. This transformed problem can be solved by a linear iterative method, such as a Krylov subspace method (e.g., GMRES.)

In Section 2.3, we discuss how the convergence rate of the GMRES method can be estimated when applied to the transformed/preconditioned problem. If the linear operator representing this problem has strictly positive eigenvalues for its Hermitian part, then the convergence rate can be bounded by using two parameters: (i) the minimal eigenvalue of the Hermitian part of the operator, and (ii) the norm of the operator. We can analyze the convergence of a specific domain decomposition method by bounding these two parameters in terms of some key problem specifications such

as mesh size, subdomain size and regularization parameter.

1.3.2 Overlapping Methods

In Chapter 3, we discuss an overlapping domain decomposition approach for optimal control. The one-level and two-level overlapping methods are extensions of the well-known additive Schwarz methods for PDE problems. These methods for optimal control are implemented as preconditioners within a Krylov iterative method for solving the KKT optimality system.

In Section 3.1, we describe a model distributed control problem in variational formulation. The one-level and two-level methods are described in Sections 3.2 and 3.3 respectively. We use a standard piecewise linear finite element method for both analysis and implementation. The partitioning and triangulation of the subdomains can be done in the same manner as for a PDE problem. For the one-level method, we may triangulate first with a fine mesh then partition the set of elements into subdomains. For the two-level method, we can simplify the implementation by triangulating first with a coarse mesh to define the subdomains, then refine it into a fine mesh. This allows the coarse grid function space to be nested inside the fine grid space. Each subdomain is extended to a larger overlapping region, with any part outside the original domain being cut off. The minimum overlap amount is either kept fixed, or allowed to stay proportional to the subdomain diameter.

In Section 3.4.1, we analyze the convergence rate of the two-level method in terms of how some key problem parameters can affect this rate. Our approach is to extend the analysis of nonsymmetric PDEs by Cai and Widlund ([CW92], [CW93]) to the optimal control problem. We express the bilinear form describing the optimality conditions as the sum of a symmetric positive definite (s.p.d.) component and a nonsymmetric component. If the s.p.d. component can be made to dominate the nonsymmetric component by varying some problem parameters, then the linear operator of the preconditioned problem would have strictly positive eigenvalues for

its Hermitian part, thus allowing the use of a convenient estimate for the GMRES convergence rate. Our main theoretical result for the two-level overlapping method shows that if the coarse grid is made sufficiently fine then the convergence rate can be bounded independently of the number of subdomains or the discretization level. This qualitative behavior does not depend on the regularization parameter, and is similar to the two-level result for nonsymmetric PDEs.

We also discuss an alternative preconditioning method suggested in [CW92], where only the symmetric positive definite component of the subdomain preconditioner is used. For the optimal control problem, this method performs well only in the case of relatively large regularization parameter.

In Section 3.4.2, we consider the difficulties involved in applying the convergence analysis to the one-level method. Without a coarse grid component, the convergence of the one-level method cannot be bounded independently of the number of subdomains in the general case.

In Section 3.4.3, we consider the effect on the convergence rate if the overlap amount δ is allowed to decrease independently of H. Similar to the PDE case, a small positive δ is enough to retain the good convergence behavior for the two-level method.

We present these overlapping preconditioning methods at the algebraic level in Section 3.5, which is useful for describing the implementation details. In Section 3.6, we present numerical results for both overlapping methods. In general, the one-level method is good for a small number of subdomains while the two-level method is good for many subdomains. The regularization parameter α and the overlap amount δ have a larger effect on convergence for the one-level method.

1.3.3 Nonoverlapping Methods

In Chapter 4, we discuss a nonoverlapping approach to combining domain decomposition with optimal control. We use the Schur complement viewpoint for analysis and

to develop new preconditioners for solving the optimality system. These preconditioners, the one-level Neumann-Neumann and two-level Balancing Neumann-Neumann, are extensions of well-known domain decomposition methods for PDEs.

In Section 4.1, we derive a two-subdomain variational formulation of an optimal control problem, consisting of two decoupled local optimality systems plus an interface condition. Each of the two optimality systems represents a local control problem on a subdomain, with the same boundary value for the state and adjoint variables on the common interface. The interface condition requires the normal derivative of the states and adjoints to be continuous on the common interface. This alternative viewpoint of the optimality conditions is shown to be equivalent to the original one-domain formulation. We also derive a formulation of the two local control problems defined on the subdomains that correspond to the decoupled local optimality systems.

We next formulate a Schur complement system with respect to the state and adjoint variables on the common interface. The PDE concept of harmonic extension from interface to subdomain interior is generalized to the optimal control problem, where the interface values of both states and adjoints are specified. The local Schur complement operator for each subdomain is then defined in terms of these generalized harmonic extensions. The global Schur system operator follows directly as the sum of these local operators.

As in the PDE case, the action of a local Schur operator may be viewed as a mapping of Dirichlet data to Neumann data on the common interface, but now the action includes both state and adjoint variables. In our case, computing this mapping requires solving some local optimal control problems rather than some local PDEs. Conversely, the action of an inverse local Schur operator may be viewed as a mapping of Neumann data to Dirichlet data on the interface. For PDE problems, this inverse Schur operator is undefined when the underlying differential operator is singular (e.g., for the case of an interior subdomain having pure Neumann data on its boundary/interface.) A similar situation can occur in some optimal control problems.

However, the inverse local Schur operators are used only as part of a preconditioner in a Schur complement method, meaning they do not need to be computed exactly. For both PDE and optimal control problems, one possible method for approximating the inverse of a singular Schur operator is to modify slightly the underlying singular differential operator to make it positive definite. For our model distributed control problem, we formulate the local problem that needs to be solved when a local Schur operator is inverted.

We next present the Neumann-Neumann method for two subdomains and show how it may be viewed as a preconditioned iterative method. In Section 4.2, we apply a finite element method to the control problem and consider the Neumann-Neumann method at the discrete level. The somewhat unusual discretization of the control variable on the common interface is discussed.

In Section 4.3, we discuss the Neumann-Neumann method for optimal control at the algebraic level. The local and global Schur complement systems and preconditioners are specified explicitly in matrix notations. We extend these to the case of many subdomains by defining the restriction and extension matrices for the optimality system variables. For many subdomains, the one-level Neumann-Neumann method is inadequate, so we also implement the Balancing Neumann-Neumman method for optimal control, which is an extension of the method for PDE problems ([Man93].)

Numerical results are presented in Section 4.4, where the general convergence properties of the new preconditioners for optimal control are observed to be similar to that for a PDE problem.

Chapter 2

Formulation of the Optimal Control Problem

We describe a general domain decomposition approach for solving linear-quadratic optimal control problems. First, a variational formulation is used to describe an abstract control problem over one domain. If the bilinear forms satisfy certain coercivity assumptions then the problem is shown to have a unique solution. 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. An additive domain decomposition method is presented and is shown to be equivalent to the one-domain problem formulation. We then discuss how the convergence rate of this algorithm for optimal control may be analyzed by adapting ideas from the PDE problem.

2.1 Problem Formulation for One Domain

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

$$a \;:\; V \times V \to \mathbb{R}, \tag{2.1.1a}$$

$$b \;:\; U \times V \to \mathbb{R}, \tag{2.1.1b}$$

be continuous bilinear forms, let

$$m \;:\; V \times V \to \mathbb{R}, \tag{2.1.2a}$$

$$q \;:\; U \times U \to \mathbb{R}, \tag{2.1.2b}$$

be symmetric and continuous bilinear forms, and let

$$c \;:\; V \to \mathbb{R}, \tag{2.1.3a}$$

$$d \;:\; U \to \mathbb{R}, \tag{2.1.3b}$$

$$f \;:\; V \to \mathbb{R}, \tag{2.1.3c}$$

be continuous linear functionals. We assume that there exist constants $\zeta, \eta > 0$ such that

$$m(v,v) \;\geq\; 0 \qquad \forall v \in V, \tag{2.1.4a}$$

$$q(u,u) \;\geq\; \zeta \|u\|_U^2 \qquad \forall u \in U, \tag{2.1.4b}$$

$$a(v,v) \;\geq\; \eta \|v\|_V^2 \qquad \forall v \in V. \tag{2.1.4c}$$

We consider the linear-quadratic elliptic optimal control problem

$$\min_{y \in V, u \in U} \quad \tfrac{1}{2} m(y,y) - c(y) + \tfrac{1}{2} q(u,u) - d(u) \tag{2.1.5a}$$

$$\text{s.t.} \quad a(y,\phi) + b(u,\phi) = f(\phi) \quad \forall \phi \in V. \tag{2.1.5b}$$

The constraint (2.1.5b) is the weak form of an elliptic PDE with state variable y and control u.

The following two theorems establish the existence and uniqueness of the solution of (2.1.5) and provide a characterization of this solution. Both results are standard, but their proofs are included for completeness.

Theorem 2.1.1 *If the conditions* (2.1.4) *are satisfied, then problem* (2.1.5) *has a unique solution* y^*, u^*. *Moreover, there exists* $\kappa > 0$ *such that*

$$\|y^*\|_V + \|u^*\|_U \leq \kappa \left(\|c\|_{V'} + \|d\|_{U'} + \|f\|_{V'} \right). \tag{2.1.6}$$

Proof: There exist operators $A \in L(V, V')$, $B \in L(U, V')$ such that

$$\langle Ay, \phi \rangle_{V',V} = a(y, \phi), \quad \langle Bu, \phi \rangle_{V',V} = b(u, \phi) \quad \forall y \in V, u \in U, \phi \in V.$$

The coercivity assumption (2.1.4) on a implies that A is continuously invertible and $\|A^{-1}\|_{L(V',V)} \leq \eta^{-1}$.

For every $u \in U$ the state equation (2.1.5b) has a unique solution

$$y(u) = -A^{-1}Bu + A^{-1}f.$$

Hence, the constrained linear-quadratic elliptic optimal control problem (2.1.5) is equivalent to the unconstrained quadratic problem

$$\begin{aligned} \min_{u \in U} \quad & \tfrac{1}{2} m(A^{-1}Bu, A^{-1}Bu) + \tfrac{1}{2} q(u,u) - m(A^{-1}f, A^{-1}Bu) + c(A^{-1}Bu) - d(u) \\ & + \tfrac{1}{2} m(A^{-1}f, A^{-1}f) - c(A^{-1}f). \end{aligned} \tag{2.1.7}$$

The coercivity assumptions (2.1.4) on m and h imply that the objective function in (2.1.7) is strictly convex.

Standard arguments (e.g.,[Zei95, Sec. 2.5]) now imply the existence of a unique solution u^* of (2.1.7) which satisfies

$$m(A^{-1}Bu^*, A^{-1}B\mu) + q(u^*, \mu) = m(A^{-1}f, A^{-1}B\mu) - c(A^{-1}B\mu) + d(\mu) \quad \forall \mu \in U. \tag{2.1.8}$$

Setting $\mu = u^*$ and using (2.1.4a), (2.1.4b) results in the bound

$$\begin{aligned} \|u^*\|_U \quad \leq \quad & \zeta^{-1} (\, \|m\| \|A^{-1}\|^2_{L(V',V)} \|B\|_{L(U,V')} \|f\|_{V'} + \\ & \|A^{-1}\|_{L(V',V)} \|B\|_{L(U,V')} \|c\|_{V'} + \|d\|_{U'} \,). \end{aligned} \tag{2.1.9}$$

This inequality and

$$\|y^*\|_V = \|y(u^*)\|_V \leq \|A^{-1}\|_{L(V',V)} \left(\|B\|_{L(U,V')}\|u^*\|_U + \|f\|_{V'}\right)$$

can be combined to give

$$\|y^*\|_V + \|u^*\|_U \quad \leq \quad \kappa_c\|c\|_{V'} + \kappa_d\|d\|_{U'} + \kappa_f\|f\|_{V'},$$

with

$$\begin{aligned}
\kappa_c &= \zeta^{-1}(\|A^{-1}\|_{L(V',V)}\|B\|_{L(U,V')} + 1)\|A^{-1}\|_{L(V',V)}\|B\|_{L(U,V')}, \\
\kappa_d &= \zeta^{-1}(\|A^{-1}\|_{L(V',V)}\|B\|_{L(U,V')} + 1), \\
\kappa_f &= \zeta^{-1}(\|A^{-1}\|_{L(V',V)}\|B\|_{L(U,V')} + 1)\|m\|\|A^{-1}\|^2_{L(V',V)}\|B\|_{L(U,V')} + \|A^{-1}\|_{L(V',V)},
\end{aligned}$$

which implies (2.1.6) with $\kappa = \max\{\kappa_c, \kappa_d, \kappa_f\}$. □

Theorem 2.1.2 *Let the conditions* (2.1.4) *be satisfied. If* $y^* \in V, u^* \in U$ *solve* (2.1.5)*, then there exists a unique adjoint variable* $p^* \in V$ *such that*

$$a(\theta, p^*) + m(y^*, \theta) = c(\theta) \qquad \forall \theta \in V, \tag{2.1.10a}$$

$$q(u^*, \mu) + b(\mu, p^*) = d(\mu) \qquad \forall \mu \in U, \tag{2.1.10b}$$

$$a(y^*, \phi) + b(u^*, \phi) = f(\phi) \qquad \forall \phi \in V. \tag{2.1.10c}$$

On the other hand, if $y^* \in V, u^* \in U, p^* \in V$ *solve* (2.1.10)*, then* $y^* \in V, u^* \in U$ *solve* (2.1.5)*. Moreover, there exists a constant* $\kappa > 0$ *such that*

$$\|y^*\|_V + \|u^*\|_U + \|p^*\|_V \leq \kappa \left(\|c\|_{V'} + \|d\|_{U'} + \|f\|_{V'}\right). \tag{2.1.11}$$

Proof:

Let $y^* \in V, u^* \in U$ solve (2.1.5). Equation (2.1.10c) is the constraint in (2.1.5) evaluated at y^*, u^*. Moreover, the conditions on the (bi)linear forms a, m, c imply

that (2.1.10a) has a unique solution p^*. Hence, it remains to show that (2.1.10b) is valid. The solution $y^* = -A^{-1}Bu^* + A^{-1}f$, u^* satisfies

$$m(y^*, -A^{-1}B\mu) + q(u^*, \mu) = c(-A^{-1}B\mu) + d(\mu) \quad \forall \mu \in U \tag{2.1.12}$$

(cf. (2.1.8)). By the definitions of A, B, $\theta = -A^{-1}B\mu$ solves $a(\theta, \phi) + b(\mu, \phi) = 0$ for all $\phi \in V$. If we set $\phi = p^*$, (2.1.12) can be written as

$$a(-A^{-1}B\mu, p^*) + m(y^*, -A^{-1}B\mu) + q(u^*, \mu) + b(\mu, p^*) = c(-A^{-1}B\mu) + d(\mu) \quad \forall \mu \in U. \tag{2.1.13}$$

Using the definition (2.1.10a) of p^*, we see that (2.1.13) is equivalent to (2.1.10b).

Let $y^* \in V, u^* \in U, p^* \in V$ solve (2.1.10). Equation (2.1.10c) implies

$$y^* = -A^{-1}Bu^* + A^{-1}f.$$

Inserting this representation of y^* into (2.1.10a) , (2.1.10b) , subtracting the resulting equations and setting $\theta = A^{-1}B\mu$ implies

$$\begin{aligned} & m(A^{-1}Bu^*, A^{-1}B\mu) + q(u^*, \mu) - a(A^{-1}B\mu, p^*) + b(\mu, p^*) \\ = \;& m(A^{-1}f, A^{-1}B\mu) - c(A^{-1}B\mu) + d(\mu) \quad \forall \mu \in U. \end{aligned} \tag{2.1.14}$$

By definition of A, B, $z = -A^{-1}B\mu$ solves $a(z, \phi) + b(\mu, \phi) = 0$ for all $\phi \in V$. Hence (2.1.14) implies (2.1.8), which shows that $y^* \in V, u^* \in U$ solve (2.1.5).

Equation (2.1.10a) and the Lax-Milgram lemma imply

$$\|p^*\|_V \;\leq\; \eta^{-1}\|c\|_{V'} + \eta^{-1}\|m\|\|y^*\|_V.$$

The desired inequality (2.1.11) now follows with (2.1.6). □

We define the space $Z = V \times U \times V$, and let $z \equiv (y, u, p) \in Z$, $\psi \equiv (\theta, \mu, \phi) \in Z$. Assuming that the conditions (2.1.4) are satisfied, the problem of solving for the optimality conditions (2.1.10) may be stated as finding the unique $z^* \in Z$ such that

$$K(z^*, \psi) = g(\psi) \qquad \forall \psi \in Z, \tag{2.1.15}$$

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

$$K(z,\psi) = a(\theta,p) + m(y,\theta) + q(u,\mu) + b(\mu,p) + a(y,\phi) + b(u,\phi), \quad (2.1.16a)$$

$$g(\psi) = c(\theta) + d(\mu) + f(\phi). \quad (2.1.16b)$$

2.2 Subspace Decomposition of the Optimal Control Problem

We extend the subspace decomposition method for analyzing PDE problems to our abstract optimal control problem. The approach follows that of Dryja and Widlund and their co-workers ([DW89, DW90, CW92, DSW94]).

The spaces V, U are assumed to be decomposable into subspaces as

$$V = V_0 + V_1 + ... + V_N, \quad (2.2.1a)$$

$$U = U_0 + U_1 + ... + U_N, \quad (2.2.1b)$$

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

$$Z = Z_0 + Z_1 + ... + Z_N, \quad (2.2.2)$$

with $Z_i = V_i \times U_i \times V_i$. This means every $z = (y, u, p) \in Z$ may be written as $(y, u, p) = \sum_{i=0}^{N} (y_i, u_i, p_i)$ for some (not necessarily unique) $(y_i, u_i, p_i) \in Z_i, i = 0, ..., N$. In a domain decomposition problem with N (possibly overlapping) subdomains, each of the spaces $Z_i, i = 1, ..., N$, 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 $i = 0, ..., N$, we assume that there exist local bilinear forms

$$a_i : V_i \times V_i \to \mathbb{R}, \quad (2.2.3a)$$

$$b_i : U_i \times V_i \to \mathbb{R}, \quad (2.2.3b)$$

$$m_i : V_i \times V_i \to \mathbb{R}, \quad (2.2.3c)$$

$$q_i : U_i \times U_i \to \mathbb{R}, \quad (2.2.3d)$$

that are local approximations to the corresponding global bilinear forms. We require the existence of constants $\zeta_i, \eta_i > 0$, $i = 0, ..., N$, such that

$$\begin{aligned} m_i(v_i, v_i) &\geq 0 && \forall v_i \in V_i, && (2.2.4a) \\ q_i(u_i, u_i) &\geq \zeta_i \|u_i\|_U^2 && \forall u_i \in U_i, && (2.2.4b) \\ a_i(v_i, v_i) &\geq \eta_i \|v_i\|_V^2 && \forall v_i \in V_i, && (2.2.4c) \end{aligned}$$

$i = 0, ..., N$. It is often possible that these local forms can be defined by the restriction of the global forms to the local spaces, i.e.,

$$\begin{aligned} a_i(v_i, w_i) &= a(v_i, w_i) && \forall v_i, w_i \in V_i, && (2.2.5a) \\ b_i(v_i, w_i) &= b(v_i, w_i) && \forall v_i \in U_i, w_i \in V_i, && (2.2.5b) \\ m_i(v_i, w_i) &= m(v_i, w_i) && \forall v_i, w_i \in V_i, && (2.2.5c) \\ q_i(v_i, w_i) &= q(v_i, w_i) && \forall v_i, w_i \in U_i. && (2.2.5d) \end{aligned}$$

For some problems, a local definition (2.2.5a) for $a_i(\cdot, \cdot)$ does not satisfy the coercivity requirement (2.2.4c), and has to be modified slightly. For each $i = 0, ..., N$, we define a bilinear form $K_i : Z_i \times Z_i \to \mathbb{R}$ as a local approximation for $K(\cdot, \cdot)$:

$$\begin{aligned} & K_i((y_i, u_i, p_i), (\theta_i, \mu_i, \phi_i)) \\ &= a_i(\theta_i, p_i) + m_i(y_i, \theta_i) + q_i(u_i, \mu_i) + b_i(\mu_i, p_i) + a_i(y_i, \phi_i) + b_i(u_i, \phi_i). \quad (2.2.6) \end{aligned}$$

Lemma 2.2.1 *If the local bilinear forms satisfy the requirements* (2.2.4)*, then for each $z \in Z$ there exists a unique solution z_i of*

$$K_i(z_i, \psi_i) = K(z, \psi_i) \qquad \forall \psi_i \in Z_i. \tag{2.2.7}$$

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 (2.2.7) can be written as

$$\begin{aligned} & a_i(\theta_i, p_i) + m_i(y_i, \theta_i) + q_i(u_i, \mu_i) + b_i(\mu_i, p_i) + a_i(y_i, \phi_i) + b_i(u_i, \phi_i) \\ &= a(\theta_i, p) + m(y, \theta_i) + q(u, \mu_i) + b(\mu_i, p) + a(y, \phi_i) + b(u, \phi_i) \quad (2.2.8) \end{aligned}$$

for all $(\theta_i, \mu_i, \phi_i) \in V_i \times U_i \times V_i$. For a given fixed z, the right hand side of (2.2.8) defines a continuous linear functional on Z_i :

$$g_z((\theta_i, \mu_i, \phi_i)) = c_z(\theta_i) + d_z(\mu_i) + f_z(\phi_i), \tag{2.2.9}$$

with

$$c_z(\theta_i) = a(\theta_i, p) + m(y, \theta_i), \tag{2.2.10a}$$

$$d_z(\mu_i) = q(u, \mu_i) + b(\mu_i, p), \tag{2.2.10b}$$

$$f_z(\phi_i) = a(y, \phi_i) + b(u, \phi_i). \tag{2.2.10c}$$

If we set two out of the three test functions θ_i, μ_i, ϕ_i in (2.2.8) to zero in turn, then (2.2.8) is equivalent to the system

$$a_i(\theta_i, p_i) + m_i(y_i, \theta_i) = c_z(\theta_i) \qquad \forall \theta_i \in V_i, \tag{2.2.11a}$$

$$q_i(u_i, \mu_i) + b_i(\mu_i, p_i) = d_z(\mu_i) \qquad \forall \mu_i \in U_i, \tag{2.2.11b}$$

$$a_i(y_i, \phi_i) + b_i(u_i, \phi_i) = f_z(\phi_i) \qquad \forall \phi_i \in V_i. \tag{2.2.11c}$$

Using assumptions (2.2.4) together with Theorems 2.1.1 and 2.1.2, we see that (2.2.11) gives the necessary and sufficient optimality conditions for (y_i, u_i, p_i) to be the unique solution of the subspace optimal control problem

$$\min_{y_i \in V_i, u_i \in U_i} \quad \tfrac{1}{2} m_i(y_i, y_i) - c_z(y_i) + \tfrac{1}{2} q_i(u_i, u_i) - d_z(u_i) \tag{2.2.12a}$$

$$\text{s.t.} \quad a_i(y_i, \phi_i) + b_i(u_i, \phi_i) = f_z(\phi_i) \quad \forall \phi_i \in V_i. \tag{2.2.12b}$$

This means z_i is uniquely defined by z. □

For each $i = 0, ..., N$, we define the projection operator $T_i : Z \to Z_i$ in terms of K and its local approximation K_i. For any $z \in Z$, let $T_i z \in Z_i$ be the solution of

$$K_i(T_i z, \psi_i) = K(z, \psi_i) \qquad \forall \psi_i \in Z_i, \tag{2.2.13}$$

which exists and is unique by Lemma 2.2.1. We note that $T_i z$ may also be stated as the solution of the linear problem

$$K_i(T_i z, \psi_i) \;=\; g_z(\psi_i) \qquad \forall \psi_i \in Z_i, \tag{2.2.14}$$

where g_z is defined as $g_z(\psi_i) = K(z, \psi_i)$. Since $g_z(\cdot)$ is linear in z, $T_i z$ also depends linearly on z.

We are interested in transforming the problem (2.1.15) into an equivalent problem that is better conditioned. We define the operator $T : Z \to Z$ as

$$Tz = \sum_{i=0}^{N} T_i z. \tag{2.2.15}$$

Lemma 2.2.2 *If the local bilinear forms satisfy the requirements* (2.2.4) *for* $i = 0, ..., N$, *then the operator* T *is nonsingular.*

Proof: Let $z \in Z$ and suppose that $Tz = 0$. We need to show this implies $z = 0$. Let $z_i = (y_i, u_i, p_i) = T_i z$ for $i = 0, ..., N$. Then

$$\sum_{i=0}^{N} (y_i, u_i, p_i) \;=\; (0, 0, 0). \tag{2.2.16}$$

For each $i = 0, ..., N$, the optimality conditions are

$$a_i(\theta_i, p_i) + m_i(y_i, \theta_i) \;=\; c_z(\theta_i) \qquad \forall \theta_i \in V_i, \tag{2.2.17a}$$

$$q_i(u_i, \mu_i) + b_i(\mu_i, p_i) \;=\; d_z(\mu_i) \qquad \forall \mu_i \in U_i, \tag{2.2.17b}$$

$$a_i(y_i, \phi_i) + b_i(u_i, \phi_i) \;=\; f_z(\phi_i) \qquad \forall \phi_i \in V_i. \tag{2.2.17c}$$

Setting $\theta_i = y_i$, $\mu_i = u_i$, $\phi_i = p_i$, then summing over $i = 0, ..., N$ and using (2.2.16),

$$\sum_{i=0}^{N} a_i(y_i, p_i) + m_i(y_i, y_i) \;=\; \sum_{i=0}^{N} c_z(y_i) \;=\; c_z(\sum_{i=0}^{N} y_i) \;=\; 0, \tag{2.2.18a}$$

$$\sum_{i=0}^{N} q_i(u_i, u_i) + b_i(u_i, p_i) \;=\; \sum_{i=0}^{N} d_z(u_i) \;=\; d_z(\sum_{i=0}^{N} u_i) \;=\; 0, \tag{2.2.18b}$$

$$\sum_{i=0}^{N} a_i(y_i, p_i) + b_i(u_i, p_i) \;=\; \sum_{i=0}^{N} f_z(p_i) \;=\; f_z(\sum_{i=0}^{N} p_i) \;=\; 0. \tag{2.2.18c}$$

Adding equation (2.2.18a) to (2.2.18b) and subtracting (2.2.18c) results in

$$\sum_{i=0}^{N} q_i(u_i, u_i) + m_i(y_i, y_i) \;\; = \;\; 0. \tag{2.2.19a}$$

Each term in the sum is nonnegative by assumptions (2.2.4a,b), so we must have $u_i = 0$ for $i = 0, ..., N$. Setting $\phi_i = y_i$ in (2.2.17c) and summing over $i = 0, ..., N$ gives

$$\sum_{i=0}^{N} a_i(y_i, y_i) + b_i(u_i, y_i) \;\; = \;\; \sum_{i=0}^{N} f_z(y_i) \; = \; 0, \tag{2.2.20a}$$

which implies $y_i = 0$, $i = 0, ..., N$, by requirement (2.2.4c). Finally, setting $\theta_i = p_i$ in (2.2.17a) and summing over i yields

$$\sum_{i=0}^{N} a_i(p_i, p_i) + m_i(y_i, p_i) \;\; = \;\; \sum_{i=0}^{N} c_z(p_i) \; = \; 0, \tag{2.2.21a}$$

which implies $p_i = 0$, $i = 0, ..., N$. This shows $T_i z = 0$ for $i = 0, ..., N$, which implies $z = 0$.

□

We consider the problem of finding z^* such that

$$Tz^* = g_T. \tag{2.2.22}$$

The right hand side g_T is chosen so that z^* is the same as the solution of problem (2.1.15), i.e.,

$$g_T = \sum_{i=0}^{N} g_i, \tag{2.2.23}$$

where $g_i = T_i z^* \in Z_i, i = 0, ..., N$. The local right hand sides g_i, $i = 0, ..., N$, can be computed without knowing the solution z^*. In fact, (2.1.15) and (2.2.7) imply that

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

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

$$K_i(g_i, \psi_i) = g(\psi_i) \qquad \forall \psi_i \in Z_i. \tag{2.2.24}$$

Lemma 2.2.3 *If the local bilinear forms satisfy the requirements* (2.2.4) *for* $i = 0, ..., N$, *then problems* (2.1.15) *and* (2.2.22) *have the same solution.*

Proof: (i) Problem (2.1.15) has a unique solution by Theorem 2.1.1. Let z^* satisfy $K(z^*, \psi) = g(\psi)$ for all $\psi \in Z$, then

$$\begin{aligned} K_i(T_i z^*, \psi_i) &= K(z^*, \psi_i) && \forall \psi_i \in Z_i, \\ &= g(\psi_i) && \forall \psi_i \in Z_i, \\ &= K_i(g_i, \psi_i) && \forall \psi_i \in Z_i. \end{aligned}$$

This means

$$K_i(T_i z^* - g_i, \psi_i) = 0 \qquad \forall \psi_i \in Z_i,$$

which implies $T_i z^* = g_i$. Therefore, $T z^* = \sum_{i=0}^{N} T_i z^* = \sum_{i=0}^{N} g_i = g_T$.

(ii) Suppose that we have found a solution w that solves $Tw = g_T$. Then $Tw = Tz^*$, or $T(z^* - w) = 0$. Since T is nonsingular by Lemma 2.2.2, we have $(z^* - w) = 0$, so w is a solution to (2.1.15).

□

The transformed problem (2.2.22) may be solved by using a linear iterative method. The operator T is nonsymmetric, in general, and we solve (2.2.22) using GMRES [SS86] or QMR [FN92]. The operator T may have structure that allows the application of the symmetric QMR (sQMR) method [FN94, FN95].

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

$$c_T = \inf_{z \neq 0} \frac{A(Tz, z)}{A(z, z)}, \tag{2.2.25}$$

$$C_T = \sup_{z \neq 0} \frac{\|Tz\|_A}{\|z\|_A}. \tag{2.2.26}$$

If GMRES with inner product $A(\cdot,\cdot)$ is applied to the solution of (2.2.22) and if $c_T > 0$, then the residual $r^{(k)} = g_T - Tz^{(k)}$ in the kth GMRES iteration obeys

$$\|r^{(k)}\|_A = \left(1 - \frac{c_T^2}{C_T^2}\right)^{k/2} \|r^{(0)}\|_A \tag{2.2.27}$$

(see [EES83].) To analyze a specific domain decomposition method, we 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).

2.3 Abstract Convergence Analysis

For the PDE problem, domain decomposition convergence theories have evolved into a general abstract Schwarz framework suitable for analyzing a variety of algorithms ([DW89, DW90, DW92, CW92, Xu92a, DSW94, DW95, GO95, SBG96].) In the PDE case, a particular domain decomposition algorithm may be defined within the Schwarz framework by specifying:

(i) The family of subspaces of a given finite element space V^h. (This typically corresponds to one subspace per subdomain plus a coarse space. For example, the subspaces $V_i^h, i = 0, ..., N$, that allow a finite element space V^h to be decomposed as $V^h = \sum_{i=0}^N V_i^h$.)

(ii) The projection operators onto the subspaces that are used to compute the correction steps. (For example, the $T_i, i = 0, ..., N$, operators defined in terms of the global bilinear form specifying the problem and the local bilinear forms on the subspaces. These may be either exact or approximate projections.)

(iii) A polynomial function of the projection operators, which specifies the order that the subspace corrections are performed. (For example, an additive operator is defined by $T = \sum_{i=0}^N T_i$.)

To analyze the convergence rate of a specific DD algorithm, one would first verify that the chosen subspaces (i) and projection operators (ii) satisfy a few assumptions

made by the framework ([DW95], [SBG96, Chap. 5].) For a chosen combined operator T specified by (iii), the spectral properties of T are then analyzed under the verified assumptions.

For the case of symmetric positive definite elliptic PDEs, the spectral properties of the additive operator T have been well analyzed for various overlapping and non-overlapping algorithms. The iterative method of choice is the conjugate gradient method, which has a theoretical convergence bound expressible in terms of the largest and smallest eigenvalues of T.

For nonsymmetric or indefinite PDE problems, the additive operator T is non-symmetric. Among many options, we may use the preconditioned GMRES iterative method [SS86], which is equivalent to the Generalized Conjugate Residual method [EES83], or we may use QMR [FN92]. The operator T may have structure that allow the application of the symmetric QMR (sQMR) method [FN94, FN95].

We stated earlier that if GMRES with inner product $A(\cdot,\cdot)$ is applied to the solution of (2.2.22), then the residual $r^{(k)} = g_T - Tz^{(k)}$ in the kth GMRES iteration obeys (2.2.27). To estimate the parameters c_T, C_T in (2.2.27) for the optimal control problem, we extend the approach developed for nonsymmetric PDEs by Cai and Widlund [CW92, CW93].

The bilinear form $K(\cdot,\cdot)$ defined in (2.1.16) associated with the optimal control problem (2.1.5) may be split into a symmetric positive definite part $A(\cdot,\cdot)$ and an indefinite part $N(\cdot,\cdot)$

$$K(z,\psi) \;=\; A(z,\psi) + N(z,\psi), \tag{2.3.1}$$

where

$$z, \psi \in Z \equiv V \times U \times V,$$

and V, U are the state and the control space introduced in Section 2.1. The symmetric positive definite part $A(\cdot,\cdot)$ in the splitting (2.3.1) of K serves as the inner product in GMRES.

We use the notation

$$z = (y, u, p), \quad \psi = (\theta, \mu, \phi), \quad z_i = (y_i, u_i, p_i), \quad \psi_i = (\theta_i, \mu_i, \phi_i).$$

One splitting of (2.3.1) is given by

$$A((y,u,p),(\theta,\mu,\phi)) = a(y,\theta) + q(u,\mu) + a(p,\phi), \tag{2.3.2a}$$

$$N((y,u,p),(\theta,\mu,\phi)) = m(y,\phi) + b(\mu,p) + b(u,\theta). \tag{2.3.2b}$$

Similarly, the local forms $K_i(\cdot,\cdot), i = 0, ..., N$, may be split as

$$K_i(z_i,\psi_i) = A_i(z_i,\psi_i) + N_i(z_i,\psi_i), \tag{2.3.3a}$$

$$A_i((y_i,u_i,p_i),(\theta_i,\mu_i,\phi_i)) = a_i(y_i,\theta_i) + q_i(u_i,\mu_i) + a_i(p_i,\phi_i), \tag{2.3.3b}$$

$$N_i((y_i,u_i,p_i),(\theta_i,\mu_i,\phi_i)) = m_i(y_i,\phi_i) + b_i(\mu_i,p_i) + b_i(u_i,\theta_i). \tag{2.3.3c}$$

We consider the case where the local bilinear forms are the restriction of the global forms to the local spaces, meaning $K_i(z_i,\psi_i) = K(z_i,\psi_i)$ for all $z_i, \psi_i \in Z_i$. From the definition of T_i in (2.2.13),

$$A(T_i z, \psi_i) + N(T_i z, \psi_i) = A(z,\psi_i) + N(z,\psi_i) \qquad \forall \psi_i \in Z_i. \tag{2.3.4}$$

Setting $\psi_i = T_i z$, we have

$$A(T_i z, T_i z) + N(T_i z, T_i z) = A(z, T_i z) + N(z, T_i z), \tag{2.3.5}$$

or

$$A(T_i z, z) = A(T_i z, T_i z) + N(T_i z - z, T_i z). \tag{2.3.6}$$

Summing up over $i = 0, ..., N$, we have

$$\sum_{i=0}^{N} A(T_i z, z) = \sum_{i=0}^{N} A(T_i z, T_i z) + \sum_{i=0}^{N} N(T_i z - z, T_i z). \tag{2.3.7}$$

The first sum on the right hand side is non-negative while the second sum may be positive, negative or zero. We may write

$$\sum_{i=0}^{N} A(T_i z, z) \geq \sum_{i=0}^{N} A(T_i z, T_i z) - |\sum_{i=0}^{N} N(T_i z - z, T_i z)|, \tag{2.3.8}$$

or for $z \neq 0$,

$$\frac{A(Tz,z)}{A(z,z)} \;\geq\; \frac{\sum_{i=0}^{N} A(T_i z, T_i z)}{A(z,z)} - \frac{|\sum_{i=0}^{N} N(T_i z - z, T_i z)|}{A(z,z)}. \tag{2.3.9}$$

Let $A(\cdot,\cdot)$ be an inner product on Z. This is true for $A(\cdot,\cdot)$ defined in (2.3.2a) under the assumptions (2.1.4b), (2.1.4c). Let T^* denote the adjoint of T with respect to this inner product. The minimal eigenvalue of the Hermitian part of T can then be characterized by

$$\inf_{z\neq 0} \frac{A(\frac{1}{2}(T+T^*)z,z)}{A(z,z)} \;=\; \inf_{z\neq 0} \frac{A(Tz,z)}{A(z,z)}. \tag{2.3.10}$$

Inequality (2.3.9) implies that the eigenvalues of the Hermitian part of T and, hence, c_T defined in (2.2.25) will be all positive if the second term on the right hand side of (2.3.9) can be made small enough relative to the first term. If this is the case, the bound (2.2.27) for the GMRES method is applicable.

To analyze a specific domain decomposition method, we estimate a lower bound for c_T defined in (2.2.25) and an upper bound for C_T defined in (2.2.26) in terms of the key problem parameters H (subdomain size), h (mesh element size) and α (regularization parameter.)

Chapter 3

Overlapping Methods

The original domain decomposition method by Schwarz in 1870 was stated as an alternating method for two overlapping subdomains (see [SBG96, Sec. 1.1].) Since then, the overlapping method has been studied by many researchers for various PDE problems (see [Lio88].) Some recent theoretical studies on the one-level version by P. Lions may be found in [Lio88], [Lio89]. The two-level version was introduced by Dryja and Widlund for symmetric positive definite PDEs in [DW89], and extended to nonsymmetric problems by Cai and Widlund in [CW92].

3.1 Problem Formulation

We consider the model elliptic control problem

$$\begin{array}{ll} \min\limits_{y,u} & \frac{1}{2}\int_\Omega (y(x)-y_d(x))^2\,dx + \frac{\alpha}{2}\int_\Omega (u(x))^2\,dx \\ \text{s.t.} & \begin{cases} -\Delta y(x) &= f(x)+u(x) \qquad \text{in } \Omega, \\ y(x) &= 0 \qquad\qquad\quad \text{on } \partial\Omega. \end{cases} \end{array} \tag{3.1.1}$$

We define the spaces $V \equiv H_0^1(\Omega)$, $U \equiv L^2(\Omega)$, and write problem (3.1.1) in weak formulation

$$\min_{y\in V, u\in U} \quad \tfrac{1}{2}m(y,y) - c(y) + \tfrac{1}{2}\alpha q(u,u) \tag{3.1.2a}$$

$$\text{s.t.} \quad a(y,\phi) + b(u,\phi) = f(\phi) \quad \forall \phi \in V, \tag{3.1.2b}$$

where

$$m : V \times V \to \mathbb{R} \qquad m(y,p) = \int_\Omega y(x)p(x)\,dx, \tag{3.1.3a}$$

$$a : V \times V \to \mathbb{R} \qquad a(y,p) = \int_\Omega \nabla y(x) \cdot \nabla p(x)\,dx, \tag{3.1.3b}$$

$$q : U \times U \to \mathbb{R} \qquad q(u,v) = \int_\Omega u(x)v(x)\,dx, \tag{3.1.3c}$$

$$b : U \times V \to \mathbb{R} \qquad b(u,p) = -\int_\Omega u(x)p(x)\,dx, \tag{3.1.3d}$$

$$c : V \to \mathbb{R} \qquad c(y) = \int_\Omega y(x)y_d(x)\,dx, \tag{3.1.3e}$$

$$f : V \to \mathbb{R} \qquad f(y) = \int_\Omega f(x)y(x)\,dx. \tag{3.1.3f}$$

The optimality conditions are

$$a(\theta,p) + m(y,\theta) = m(y_d,\theta) \qquad \forall \theta \in V, \tag{3.1.4a}$$

$$\alpha q(u,\mu) + b(\mu,p) = 0 \qquad \forall \mu \in U, \tag{3.1.4b}$$

$$a(y,\phi) + b(u,\phi) = f(\phi) \qquad \forall \phi \in V, \tag{3.1.4c}$$

which are the weak forms of the following equations:

$$\text{Adjoint equation:} \quad \begin{cases} -\Delta p(x) = -(y(x) - y_d(x)) & \text{in } \Omega, \\ p(x) = 0 & \text{on } \partial\Omega, \end{cases}$$

$$\text{Gradient equation:} \quad \begin{cases} -p(x) + \alpha u(x) = 0 & \text{in } \Omega, \end{cases}$$

$$\text{State equation:} \quad \begin{cases} -\Delta y(x) = f(x) + u(x) & \text{in } \Omega, \\ y(x) = 0 & \text{on } \partial\Omega. \end{cases}$$

We simplify the optimality conditions by using the gradient equation to eliminate the control $u = \alpha^{-1}p$. The reduced optimality conditions may be written as a system

of coupled PDEs:

$$\begin{aligned} -\Delta p(x) &= -(y(x) - y_d(x)) && \text{in } \Omega, && (3.1.5a)\\ p(x) &= 0 && \text{on } \partial\Omega, && (3.1.5b)\\ -\Delta y(x) &= f(x) + \alpha^{-1} p(x) && \text{in } \Omega, && (3.1.5c)\\ y(x) &= 0 && \text{on } \partial\Omega, && (3.1.5d) \end{aligned}$$

or in weak form:

$$\begin{aligned} a(\theta, p) + m(y, \theta) &= m(y_d, \theta) && \forall \theta \in V, && (3.1.6a)\\ a(y, \phi) - \alpha^{-1} m(p, \phi) &= f(\phi) && \forall \phi \in V. && (3.1.6b) \end{aligned}$$

We define the product space

$$Z \stackrel{\text{def}}{=} H_0^1(\Omega) \times H_0^1(\Omega), \tag{3.1.7}$$

and let $K : Z \times Z \to \mathbb{R}$ denote a bilinear form corresponding to the optimality system (3.1.6)

$$\begin{aligned} K(z, \psi) &= K((y, p), (\theta, \phi)) \\ &= a(y, \phi) + a(\theta, p) + m(y, \theta) - \alpha^{-1} m(p, \phi). && (3.1.8) \end{aligned}$$

Our control problem is to find $z \in Z$ such that

$$K(z, \psi) = g(\psi) \qquad \forall \psi \in Z, \tag{3.1.9}$$

where $g : Z \to \mathbb{R}$ is defined as

$$g(\psi) = g((\theta, \phi)) = m(y_d, \theta) + f(\phi). \tag{3.1.10}$$

3.2 One-level Method

We apply the finite element method for solving (3.1.9) by discretizing the domain $\Omega \subset \mathbb{R}^d\,(d = 2, 3)$, with a family of triangulations $\{\mathcal{T}^h\}$. For each triangular/tetrahedral

element τ in a triangulation $\mathcal{T}^h$ of the family, let h_τ denote the length of the longest side of τ. The characteristic parameter h of each triangulation $\mathcal{T}^h$ is defined as the maximum side length among all elements in the triangulation

$$h = \max\{h_\tau : \tau \in \mathcal{T}^h\}. \tag{3.2.1}$$

Let ρ_τ denote the diameter of the largest circle/sphere inscribed in an element τ. The family $\{\mathcal{T}^h\}$ is assumed to be regular (e.g., [QV94, Sec. 3.4]), meaning there exists a constant $c_1 \geq 1$ such that for all triangulations $\mathcal{T}^h \in \{\mathcal{T}^h\}$,

$$\max\{h_\tau/\rho_\tau : \tau \in \mathcal{T}^h\} \leq c_1. \tag{3.2.2}$$

We further assume that the family $\{\mathcal{T}^h\}$ is quasi-uniform (e.g.,[QV94, Sec. 3.5]), so there exists a constant $c_2 > 0$ such that for all triangulations $\mathcal{T}^h \in \{\mathcal{T}^h\}$,

$$\min\{h_\tau/h : \tau \in \mathcal{T}^h\} \geq c_2. \tag{3.2.3}$$

Let $V^h(\Omega)$ be the space of continuous piecewise linear functions defined on a triangulation $\mathcal{T}^h$, and with value zero on the outside boundary $\partial\Omega$

$$V^h = \{v^h : v^h \in C^0(\Omega), v^h|_\tau \text{ linear } \forall \tau \in \mathcal{T}^h, v^h = 0 \text{ on } \partial\Omega\}. \tag{3.2.4}$$

We partition the set of elements in Ω into N subdomains Ω_i, $i = 1, ..., N$, with maximum diameter less than or equal to H. It is often desirable to have the partitioned subdomains roughly equal-sized, for load-balancing reasons in parallel implementations. For the pair of state and adjoint variables, we define the finite element product space

$$Z^h \stackrel{\text{def}}{=} V^h \times V^h. \tag{3.2.5}$$

Our discretized problem is to find $z^h \in Z^h$ such that

$$K(z^h, \psi^h) = g(\psi^h) \qquad \forall \psi^h \in Z^h. \tag{3.2.6}$$

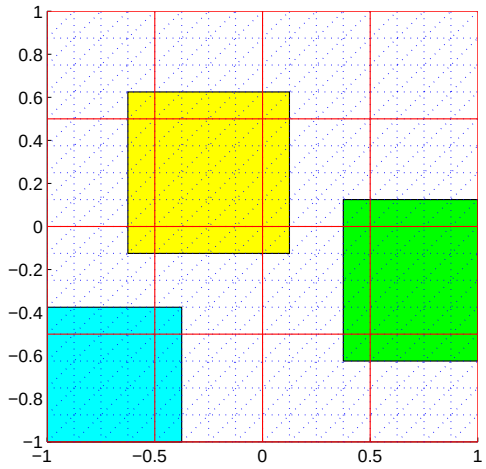

Figure 3.1: A sample triangulation partitioned into 16 subdomains. The three shaded regions are examples of extended subdomains.

In an overlapping method for problem (3.2.6), we extend each subdomain Ω_i to a larger region $\widehat{\Omega}_i$, with an overlap such that

$$\text{distance}(\partial\widehat{\Omega}_i \cap \Omega, \partial\Omega_i \cap \Omega) \geq \delta, \quad \forall i \tag{3.2.7}$$

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

$$\delta \;\geq\; \beta H \tag{3.2.8}$$

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 for a 2-D example.)

We define a finite element subspace $V_i^h \subset V^h$ associated with each extended subdomain $\widehat{\Omega}_i, i = 1, ..., N$, by

$$V_i^h \;=\; \{v \in V^h : v(x) = 0 \text{ on } \Omega \backslash \widehat{\Omega}_i\}. \tag{3.2.9}$$

The associated local product spaces are

$$Z_i^h \;\stackrel{\text{def}}{=}\; V_i^h \times V_i^h \qquad i = 1, ..., N. \tag{3.2.10}$$

We note that the finite element space V^h is decomposable as

$$V^h = V_1^h + V_2^h + ... + V_N^h \tag{3.2.11}$$

so that each $v^h \in V^h$ can be written as $v^h = \sum_{i=1}^N v_i^h$, for some $v_i^h \in V_i^h, i = 1, ..., N$. Similarly, the product space Z^h may be decomposed as

$$Z^h = Z_1^h + Z_2^h + ... + Z_N^h. \tag{3.2.12}$$

On each Z_i^h, we define the bilinear form $K_i : Z_i^h \times Z_i^h \to \mathbb{R}$ as a local restriction of $K(\cdot, \cdot)$:

$$K_i((y_i, p_i), (\theta_i, \phi_i)) = a(y_i, \theta_i) + a(p_i, \phi_i) + m(y_i, \phi_i) - \alpha^{-1} m(p_i, \theta_i). \tag{3.2.13}$$

For each $i = 1, ..., N$, we define $T_i : Z^h \to Z_i^h$ as the operator such that $T_i z^h$ is the solution of the finite element equations

$$K_i(T_i z^h, \psi_i^h) = K(z^h, \psi_i^h) \qquad \forall \psi_i^h \in Z_i^h. \tag{3.2.14}$$

We define the discrete one-level additive operator

$$T = T_1 + T_2 + ... + T_N, \tag{3.2.15}$$

and replace the original problem (3.2.6) by the problem of finding z^h such that

$$T z^h = g_T. \tag{3.2.16}$$

The right hand side g_T is chosen so that the two problems have the same solution:

$$g_T = \sum_{i=1}^N g_T^{(i)}, \tag{3.2.17}$$

where $g_T^{(i)} = T_i z^h$ can be computed without knowing the solution z^h by solving a local finite element problem

$$K_i(g_T^{(i)}, \psi_i^h) = g(\psi_i^h) \qquad \forall \psi_i^h \in Z_i^h. \tag{3.2.18}$$

The T_i operator defined by (3.2.14) may be considered a discrete version of the abstract operator T_i from Section 2.2, but with the control variables eliminated. To show that it is a projection operator onto Z_i^h, we note that for any $z^h \in Z^h$,

$$\begin{aligned} K(T_i(T_i z^h), \psi_i^h) &= K_i(T_i(T_i z^h), \psi_i^h) \qquad \forall \psi_i^h \in Z_i^h \\ &= K(T_i z^h, \psi_i^h) \qquad \forall \psi_i^h \in Z_i^h, \end{aligned}$$

(using (3.2.14)), or

$$K(T_i(T_i z^h) - T_i z^h, \psi_i^h) = 0 \qquad \forall z^h \in Z^h, \forall \psi_i^h \in Z_i^h, i = 1, ..., N,$$

which implies $T_i^2 = T_i$.

We may view these T_i as error projectors onto Z_i^h in a simple iterative method. Suppose that we have $z^{(k)}$ as the current approximation to the solution z^h of problem (3.2.16), then the error is $e^{(k)} = z^h - z^{(k)}$. The next approximation given by the Richardson iterative method is

$$z^{(k+1)} = z^{(k)} + g_T - T z^{(k)}.$$

Subtracting both sides from z^h and substituting $T z^h$ for g_T, we have

$$\begin{aligned} z^h - z^{(k+1)} &= z^h - z^{(k)} - T z^h + T z^{(k)} \\ e^{(k+1)} &= e^{(k)} - T e^{(k)} \\ &= (I - T) e^{(k)}, \end{aligned}$$

so the error propagation operator is $I - T$. At each iteration of this method, the current error $e^{(k)}$ is projected onto each of the subspaces Z_i^h as $T_i e^{(k)}$, which can be used as a subspace correction step. These subspace corrections are combined into one global correction step as $T e^{(k)} = \sum_i T_i e^{(k)}$, which is used to update $z^{(k)}$, thus changing the error by $-T e^{(k)}$.

In practice, the transformed problem (3.2.16) is solved by a Krylov subspace method, which may be viewed as an accelerated version of a simple iterative method. In Section 3.5, we will discuss an equivalent algebraic formulation and other implementation issues.

3.3 Two-level Method

In the one-level method, the correction step computed at each iteration is determined by solving only local subdomain problems. For second order elliptic PDEs, it is known that the local subdomain solves may not be effective at removing certain components of the global error. As discussed in [SBG96, Sec. 2.1], since the local spaces $V_i^h, i = 1, ..., N$, have value zero on the boundary of the subdomain $\widehat{\Omega}_i$, they cannot contain any nonzero constant function over $\widehat{\Omega}_i$. Consider the problem of projecting the global error onto a subspace V_i^h associated with an interior subdomain $\widehat{\Omega}_i$ (where the Dirichlet boundary condition on $\partial\Omega$ does not apply.) If the error e over $\widehat{\Omega}_i$ contains a large constant shift, then the projection operator T_i cannot compute a good subspace correction step $T_i e$ in one iteration. This situation also applies to the optimal control problem since these subspaces are used for the state and adjoint variables. Moreover, in the one-level method, the effect of a subdomain correction step is propagated at each iteration only to the overlapped subdomain neighbors. Therefore, many iterations may be required to remove the troublesome error component for the case of many subdomains.

A two-level method can solve these problems by using a coarse grid over the entire domain Ω for transporting global information to all subdomains at each iteration. The subspace associated with this coarse grid is able to represent functions that are constant over an interior subdomain, thus allowing the computation of a good correction step when the error has a large constant component. Our two-level method for optimal control is an extension of the the two-level overlapping method for PDEs introduced by Dryja and Widlund [DW89].

In the two-level method, we use two families of triangulation to discretize the domain Ω, with the coarse-level family denoted by $\{\mathcal{T}^H\}$ and the fine-level family by $\{\mathcal{T}^h\}$. The parameters H and h denote the characteristic maximum element size of a triangulation in the families, as described below. For simplicity, we define the two levels of triangulations so that the coarse grid finite element space can be nested

inside the fine grid space. The following description of the triangulation process is for a 2-dimensional problem, but is extendible to three dimensions. We consider only piecewise linear triangular elements, and assume that the domain Ω is polygonal.

For the coarse-level triangulation, we first partition Ω into N nonoverlapping polygonal subdomains, denoted as Ω_i, $i = 1, ..., N$, so that $\overline{\Omega} = \cup_{i=1}^{N} \overline{\Omega}_i$. To simplify the triangulation, we may let each subdomain be either a triangle or a quadrilateral. If subdomain Ω_i is triangular then it defines a coarse element in the triangulation $\mathcal{T}^H$. Otherwise, a quadrilateral Ω_i is subdivided into two triangular elements in $\mathcal{T}^H$. For each triangle τ in a coarse triangulation $\mathcal{T}^H$, let H_τ denote the length of the longest side of τ. The characteristic parameter H of each $\mathcal{T}^H$ is defined as the maximum side length among all triangles:

$$H = \max\{H_\tau : \tau \in \mathcal{T}^H\}. \tag{3.3.1}$$

(H may be considered an approximation of the maximum subdomain diameter.) Let ρ_τ denote the diameter of the largest circle/sphere inscribed in τ. The family of coarse triangulation $\{\mathcal{T}^H\}$ is assumed to be regular, i.e., there exists a constant $c_3 \geq 1$ such that

$$\max\{H_\tau/\rho_\tau : \tau \in \mathcal{T}^H\} \leq c_3. \tag{3.3.2}$$

We also assume the family $\{\mathcal{T}^H\}$ is quasi-uniform, so there exist a constant $c_4 > 0$ such that for all $\mathcal{T}^H \in \{\mathcal{T}^H\}$,

$$\min\{H_\tau/H : \tau \in \mathcal{T}^H\} \geq c_4. \tag{3.3.3}$$

We construct a fine-level $\mathcal{T}^h$ by dividing each coarse triangle in $\mathcal{T}^H$ into smaller triangles so that $\mathcal{T}^h$ defines a finite element triangulation over the entire domain Ω (i.e., the triangulation is conforming between adjacent subdomains; see Figure 3.2.) As in the one-level method, the parameter h denotes the maximum side length of any fine-level triangles. We again assume that the family of fine-level triangulations $\{\mathcal{T}^h\}$ is quasi-uniform, and define an associated finite element space for each triangulation

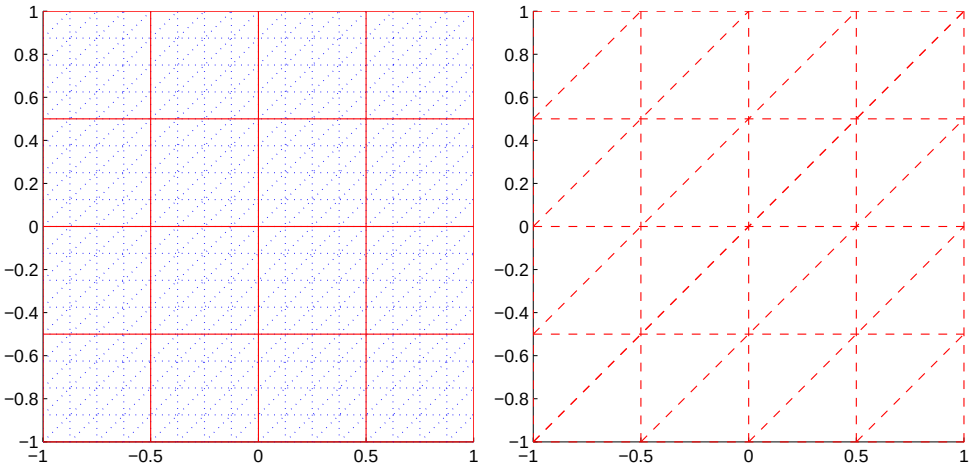

Figure 3.2: Left: A sample fine mesh triangulation with 16 subdomains. Right: The corresponding coarse mesh triangulation.

as

$$V^h = \{v^h : v^h \in C^0(\Omega), v^h|_\tau \text{ linear } \forall \tau \in \mathcal{T}^h, v^h = 0 \text{ on } \partial\Omega\}. \tag{3.3.4}$$

The coarse-level finite element space may be defined as

$$\begin{aligned} V^H &= \{v^H : v^H \in C^0(\Omega), v^H|_{\tau_H} \text{ linear } \forall \tau_H \in \mathcal{T}^H, v^H = 0 \text{ on } \partial\Omega\}, \\ &= \{v^H : v^H \in V^h(\Omega), v^H|_{\tau_H} \text{ linear } \forall \tau_H \in \mathcal{T}^H\}, \end{aligned} \tag{3.3.5}$$

which means $V^H \subset V^h$. It is possible to construct non-nested triangulations so that the coarse space V^H is not a subset of V^h. In this case, we would need an interpolation operator from V^H to V^h in order to exchange information between the two grids [SBG96, Sec. 2.8].

The product space for the coarse-level problem is

$$Z^H \stackrel{\text{def}}{=} V^H \times V^H. \tag{3.3.6}$$

For convenience, we also define $V_0 \stackrel{\text{def}}{=} V^H$ and $Z_0 \stackrel{\text{def}}{=} Z^H$. The coarse space operator $T_0 : Z^h \to Z^H$ is defined such that $T_0 z^h$ is the solution of the coarse finite element equations

$$K(T_0 z^h, \psi^H) = K(z^h, \psi^H) \qquad \forall \psi^H \in Z^H. \tag{3.3.7}$$

We define the two-level additive operator as

$$T = T_0 + T_1 + ... + T_N, \tag{3.3.8}$$

where $T_i, i = 1, ..., N$ are the local operators previously defined for the one-level method. The variational problem (3.2.6) is now transformed into

$$Tz^h = g_T, \tag{3.3.9}$$

with right hand

$$g_T = \sum_{i=0}^{N} g_T^{(i)}. \tag{3.3.10}$$

and

$$K(g_T^{(0)}, \psi^H) = g(\psi^H) \qquad \forall \psi^H \in Z^H. \tag{3.3.11}$$

It is possible to define other two-level operators by combining the local operators $T_i, i = 0, ..., N$ differently. For example, the multiplicative operator or some hybrid between the additive and multiplicative method can be extended from the PDE case to the optimal control case [SBG96, Sec. 2.3].

3.4 Convergence Analysis

3.4.1 Two-Level Method

We analyze the convergence properties of a two-level overlapping method for solving the optimal control problem (3.1.1), or to find $z \in Z$ such that

$$K(z, \psi) = g(\psi) \qquad \forall \psi \in Z. \tag{3.4.1}$$

After a finite element discretization, we have the problem of finding $z^h \in Z^h$ such that

$$K(z^h, \psi^h) = g(\psi^h) \qquad \forall \psi^h \in Z^h. \tag{3.4.2}$$

We assume that the amount of overlap between regions is uniformly generous and is kept proportional to H. The maximum diameter of each extended region is assumed to be less than or equal to $C_\delta H$, for some constant $C_\delta > 1$. To use some of the theoretical error estimates, certain regularity assumptions on the PDE problems are necessary. We will assume that the domain Ω is either a bounded polygonal region in $\mathbb{R}^2$, or a bounded polyhedral region in $\mathbb{R}^3$. For simplicity, we will consider only piecewise linear finite elements.

To simplify the labeling of non-critical constants, we often use C without any subscript to denote a generic constant that never depends on the parameters H, h or α, but can change from line to line.

The bilinear form $K(\cdot,\cdot)$ may be split into two components

$$K(z,\psi) = A(z,\psi) + N(z,\psi) \qquad \forall z,\psi \in Z \tag{3.4.3}$$

with symmetric positive definite part

$$A((y,p),(\theta,\phi)) = a(y,\theta) + a(p,\phi) \tag{3.4.4}$$

and nonsymmetric part

$$N((y,p),(\theta,\phi)) = m(y,\phi) - \alpha^{-1}\, m(p,\theta). \tag{3.4.5}$$

We also define $M : Z \times Z \to \mathbb{R}$ as

$$M((y,p),(\theta,\phi)) = m(y,\theta) + m(p,\phi) \tag{3.4.6}$$

and define two norms on Z as

$$\|z\|_M^2 \;=\; M(z,z) \;=\; m(y,y) + m(p,p), \tag{3.4.7}$$

$$\|z\|_A^2 \;=\; A(z,z) \;=\; a(y,y) + a(p,p). \tag{3.4.8}$$

By a Poincaré inequality [GT01, eqn. 7.44], there exists a constant $C_\Omega > 0$ such that

$$\|y\|_{L^2(\Omega)} \;\le\; C_\Omega\, |y|_{H^1(\Omega)} \qquad \forall y \in H_0^1(\Omega), \tag{3.4.9}$$

where for a domain $\Omega \subset \mathbb{R}^d$ with measure $|\Omega|$,

$$C_\Omega = \left(\frac{|\Omega|}{|B_d|} \right)^{1/d}, \tag{3.4.10}$$

with $|B_d|$ being the volume of the unit ball in $\mathbb{R}^d$. Therefore, we have

$$\|z\|_M \leq C_\Omega \|z\|_A \qquad \forall z \in Z. \tag{3.4.11}$$

For a function $v \in V$, we use the following standard notations for norms and semi-norms defined over a domain $\mathcal{D} \subset \mathbb{R}^d$:

$$\begin{aligned}
\|v\|_{L^2(\mathcal{D})} &= \left(\int_{\mathcal{D}} |v(x)|^2 \, dx \right)^{1/2}, \\
\|v\|_{L^\infty(\mathcal{D})} &= \sup\{|v(x)| : x \in \mathcal{D}\}, \\
|v|_{H^1(\mathcal{D})} &= \|\nabla v\|_{L^2(\mathcal{D})}, \\
\|v\|_{H^1(\mathcal{D})} &= (\|v\|_{L^2(\mathcal{D})} + \|\nabla v\|_{L^2(\mathcal{D})})^{1/2}.
\end{aligned}$$

To analyze the convergence properties, we will need a number of intermediate results, which will be stated as Lemmas. They are extensions of results by Cai and Widlund ([Cai90], [CW92], [CW93], [SBG96, Chaper 5]) and Schatz ([Sch74]) to the optimal control problem.

Lemma 3.4.1 *There exists a constant $C_K > 0$ (independent of H, h and α) such that*

$$|K(z, \psi)| \leq C_K \max\{1, \alpha^{-1}\} \|z\|_A \, \|\psi\|_A \qquad \forall z, \psi \in Z. \tag{3.4.12}$$

Proof: We note that

$$\begin{aligned}
|K(z,\psi)| &= |a(y,\theta) + a(p,\phi) + m(y,\phi) - \alpha^{-1} m(p,\theta)| \\
&\leq a(y,y)^{1/2} a(\theta,\theta)^{1/2} + a(p,p)^{1/2} a(\phi,\phi)^{1/2} + m(y,y)^{1/2} m(\phi,\phi)^{1/2} \\
&\quad + \alpha^{-1} m(p,p)^{1/2} m(\theta,\theta)^{1/2} \\
&\leq a(y,y)^{1/2} a(\theta,\theta)^{1/2} + a(p,p)^{1/2} a(\phi,\phi)^{1/2} + C_\Omega^2 a(y,y)^{1/2} a(\phi,\phi)^{1/2} \\
&\quad + \alpha^{-1} C_\Omega^2 a(p,p)^{1/2} a(\theta,\theta)^{1/2},
\end{aligned}$$

and

$$\begin{aligned}\|z\|_A \|\psi\|_A &= ([a(y,y)+a(p,p)][a(\theta,\theta)+a(\phi,\phi)])^{1/2} \\ &\geq a(y,y)^{1/2}a(\theta,\theta)^{1/2} + a(y,y)^{1/2}a(\phi,\phi)^{1/2} \\ &\quad + a(p,p)^{1/2}a(\theta,\theta)^{1/2} + a(p,p)^{1/2}a(\phi,\phi)^{1/2},\end{aligned}$$

so a sufficient condition for inequality (3.4.12) is $C_K \geq \max\{1, C_\Omega^2\}$.

□

Lemma 3.4.2 *There exists a constant $C_N > 0$ (independent of H, h and α) such that*

$$|N(z,\psi)| \leq C_N \max\{1,\alpha^{-1}\}\|z\|_A \|\psi\|_M \qquad \forall z,\psi \in Z, \tag{3.4.13}$$

$$|N(z,\psi)| \leq C_N \max\{1,\alpha^{-1}\}\|\psi\|_A \|z\|_M \qquad \forall z,\psi \in Z, \tag{3.4.14}$$

$$|N(z,z)| \leq \tfrac{1}{2}\max\{1,\alpha^{-1}\}\|z\|_M^2 \qquad \forall z \in Z. \tag{3.4.15}$$

Proof: We note that

$$\begin{aligned}|N(z,\psi)| &= |m(y,\phi) - \alpha^{-1}m(p,\theta)| \\ &\leq |m(y,\phi)| + \alpha^{-1}|m(p,\theta)| \\ &\leq m(y,y)^{1/2}m(\phi,\phi)^{1/2} + \alpha^{-1}m(p,p)^{1/2}m(\theta,\theta)^{1/2}.\end{aligned}$$

Since

$$\begin{aligned}\|z\|_A \|\psi\|_M &= ([a(y,y)+a(p,p)][m(\theta,\theta)+m(\phi,\phi)])^{1/2} \\ &\geq C_\Omega^{-1}([m(y,y)+m(p,p)][m(\theta,\theta)+m(\phi,\phi)])^{1/2},\end{aligned}$$

and

$$\begin{aligned}\|\psi\|_A \|z\|_M &= ([a(\theta,\theta)+a(\phi,\phi)][m(y,y)+m(p,p)])^{1/2} \\ &\geq C_\Omega^{-1}([m(\theta,\theta)+m(\phi,\phi)][m(y,y)+m(p,p)])^{1/2},\end{aligned}$$

a sufficient condition for (3.4.13) and (3.4.14) is $C_N \geq C_\Omega$. Inequality (3.4.15) follows from

$$\begin{aligned} |N(z,z)| &= |m(y,p) - \alpha^{-1} m(p,y)| \\ &\leq |\alpha^{-1} - 1|\, |m(y,p)| \\ &\leq |\alpha^{-1} - 1|\, m(y,y)^{1/2} m(p,p)^{1/2} \\ &\leq \tfrac{1}{2} |\alpha^{-1} - 1|\, \|z\|_M^2. \end{aligned}$$

□

Lemma 3.4.3 *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*

$$\| \sum_{i=0}^N z_i^h \|_A^2 \leq N_O \sum_{i=0}^N \| z_i^h \|_A^2 \tag{3.4.16}$$

and

$$\| \sum_{i=0}^N z_i^h \|_M^2 \leq N_O \sum_{i=0}^N \| z_i^h \|_M^2. \tag{3.4.17}$$

Proof: We note that

$$\begin{aligned} \| \sum_{i=0}^N z_i^h \|_A^2 &= A(\sum_{i=0}^N z_i^h, \sum_{i=0}^N z_i^h) \\ &= \sum_{i=0}^N A(z_i^h, z_i^h) + \sum_{i=0}^N \sum_{j=0, j\neq i}^N A(z_i^h, z_j^h) \\ &\leq (N+1) \sum_{i=0}^N A(z_i^h, z_i^h) \end{aligned}$$

(since $A(z_i^h, z_j^h) + A(z_j^h, z_i^h) \leq A(z_i^h, z_i^h) + A(z_j^h, z_j^h)$ for $i \neq j$). Therefore, a simple bound for N_O is $N+1$. Because $A(z_i^h, z_j^h) = 0$ for those pairs of subdomains $\widehat{\Omega}_i$

and $\widehat{\Omega}_j$ that do not overlap, a smaller bound for N_O may be derived by counting the number of overlapping neighbors, with the coarse grid being viewed as overlapping all other subdomains. We still can improve on this bound by using a coloring argument. Let N_c denote the minimal number of colors required to color the subdomains so that no pair of subdomains that overlap has the same color. We also give the coarse component z_0^h its own color, so the components $z_i^h, i = 0, ..., N$ can be partitioned into N_c+1 disjoint sets. Let $\mathcal{J}_k, k = 1, ..., N_c+1$ denote the index sets of subdomains having the color k, then for each k

$$A(z_i^h, z_j^h) \;=\; 0 \qquad i, j \in \mathcal{J}_k, i \neq j,$$

and

$$A(\sum_{i\in\mathcal{J}_k} z_i^h, \sum_{i\in\mathcal{J}_k} z_i^h) \;=\; \sum_{i\in\mathcal{J}_k} A(z_i^h, z_i^h). \tag{3.4.18}$$

For each color k, let $w_k = \sum_{i\in\mathcal{J}_k} z_i^h$, then

$$\begin{aligned} A(\sum_{i=0}^{N} z_i^h, \sum_{i=0}^{N} z_i^h) &= A(\sum_{k=1}^{N_c+1}\sum_{i\in\mathcal{J}_k} z_i^h, \sum_{k=1}^{N_c+1}\sum_{i\in\mathcal{J}_k} z_i^h) \\ &= A(\sum_{k=1}^{N_c+1} w_k, \sum_{k=1}^{N_c+1} w_k) \\ &= \sum_{k=1}^{N_c+1}\sum_{l=1}^{N_c+1} A(w_k, w_l) \\ &= \sum_{k=1}^{N_c+1} A(w_k, w_k) + \sum_{k=1}^{N_c+1}\sum_{l=1,l\neq k}^{N_c+1} A(w_k, w_l). \end{aligned}$$

Since $A(w_k, w_l) + A(w_l, w_k) \leq A(w_k, w_k) + A(w_l, w_l)$ for $k \neq l$, we have

$$\begin{aligned} A(\sum_{i=0}^{N} z_i^h, \sum_{i=0}^{N} z_i^h) &\leq (N_c+1)\sum_{k=1}^{N_c+1} A(w_k, w_k) \\ &= (N_c+1)\sum_{k=1}^{N_c+1}\sum_{i\in\mathcal{J}_k} A(z_i^h, z_i^h) \\ &= (N_c+1)\sum_{i=0}^{N} A(z_i^h, z_i^h), \end{aligned}$$

which allows the bound $N_O \leq N_c + 1$. The bound for the $\|\cdot\|_M$ norm can be derived the same way.

□

The constant N_O 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 [SBG96, Sec. 5.2]. Here, we use the same N_O for the two different norms, with N_O being bounded independently of H, h or α.

Since the family of finite element triangulation $\{\mathcal{T}^h\}$ is assumed to be quasi-uniform, we may use the following inverse inequality to help prove the next Lemma. For the piecewise linear finite element space V^h, there exists a constant $C > 0$ independent of h such that

$$|u^h|_{H^1(\tau)} \leq Ch^{-1}\|u^h\|_{L^2(\tau)} \qquad \forall \tau \in \mathcal{T}^h, \forall u^h \in V^h \tag{3.4.19}$$

(e.g., [QV94, Sec. 3.5], [BS02, Sec. 4.5]).

The following Lemma applies to the single PDE problem and is used in the convergence analysis of domain decomposition methods for symmetric positive definite problems. We note that this Lemma is independent of any particular T_i projection operator.

Lemma 3.4.4 *There exists a constant $C_0 > 0$ (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*

$$\sum_{i=0}^{N} a(v_i^h, v_i^h) \quad \leq \quad C_0^2\, a(v^h, v^h). \tag{3.4.20}$$

The constant C_0 is sometimes called the *partition constant*. The proof is given in [DW89, Sec. 4], [Xu92a, Lemma 7.1], [SBG96, Sec. 5.3.1]. Due to its importance in comparing the one-level and two-level methods, and its extension to the small-overlap case (Section 3.4.3), we include a detailed explanation here. The basic idea

is to construct first a coarse component $v_0^h \in V^H$ by extracting some low energy (smooth) part of v^h. For example, v_0^h may be defined by an L^2 projection of v^h onto the coarse space V^H. The local components $v_i^h, i = 1, ..., N$ are then constructed as some partition of unity of the remaining $v^h - v_0^h$. This allows the $a(\cdot, \cdot)$ energy of each component to be kept low and bounded independently of the number of subdomains.

Proof: Let $Q_H : V^h \to V^H$ be defined such that for any $v^h \in V^h$, $Q_H v^h$ satisfies

$$(Q_H v^h, w^H)_{L^2(\Omega)} = (v^h, w^H)_{L^2(\Omega)} \qquad \forall w^H \in V^H. \tag{3.4.21}$$

We now show that a sufficient decomposition of v may be chosen as

$$\begin{aligned} v_0^h &= Q_H v^h, \\ v_i^h &= I^h[\theta_i(v^h - v_0^h)], \quad i = 1, ..., N, \end{aligned}$$

where the functions $\theta_i \in C_0(\Omega), i = 1, ..., N$, with $\theta_i(x) = 0$ on $\Omega \backslash \widehat{\Omega}_i$, define a partition of unity, i.e., with $0 \leq \theta_i(x) \leq 1$, and $\sum_{i=1}^N \theta_i(x) = 1$. The operator $I^h : C_0(\Omega) \to V^h(\Omega)$ interpolates a function onto the finite element space. We note that $\theta_i(x) = 1$ on the region of $\widehat{\Omega}_i$ that does not overlap any other extended subdomain, and $\theta_i(x) = 0$ outside of $\widehat{\Omega}_i$. In between these two regions, we can construct θ_i so that it decreases approximately linearly. Assuming that the overlap between each region and is kept proportional to H, the maximum gradient can be bounded from above as

$$\|\nabla \theta_i\|_{L^\infty(\widehat{\Omega}_i)} \leq C/H. \tag{3.4.22}$$

Considering one element τ at a time and letting $\bar{\theta}_i$ denote the average of θ_i over τ, we can bound the maximum deviation of θ_i from its average as

$$\|\theta_i - \bar{\theta}_i\|_{L^\infty(\tau)} \leq h \|\nabla \theta_i\|_{L^\infty(\tau)} \leq Ch/H. \tag{3.4.23}$$

The energy of local component v_i^h over an arbitrary element τ is bounded as

$$\begin{aligned} |v_i^h|_{H^1(\tau)} &= |\, I^h[\bar{\theta}_i(v^h - v_0^h) + (\theta_i - \bar{\theta}_i)(v^h - v_0^h)]\, |_{H^1(\tau)} \\ &\leq |\bar{\theta}_i(v^h - v_0^h)|_{H^1(\tau)} + |I^h[(\theta_i - \bar{\theta}_i)(v^h - v_0^h)]|_{H^1(\tau)}. \end{aligned}$$

The first term on the right can be bounded by $|v^h - v_0^h|_{H^1(\tau)}$, since $\|\bar{\theta}_i\|_{L^\infty(\tau)} \leq 1$. The second term on the right may be bounded by the inverse inequality (3.4.19), yielding

$$|v_i^h|_{H^1(\tau)} \leq |v^h - v_0^h|_{H^1(\tau)} + Ch^{-1}\|I^h[(\theta_i - \bar{\theta}_i)(v^h - v_0^h)]\|_{L^2(\tau)}.$$

Using (3.4.23) in the second term, we have

$$|v_i^h|_{H^1(\tau)} \leq |v^h - v_0^h|_{H^1(\tau)} + CH^{-1}\|v^h - v_0^h\|_{L^2(\tau)},$$

or

$$|v_i^h|^2_{H^1(\tau)} \leq 2|v^h - v_0^h|^2_{H^1(\tau)} + CH^{-2}\|v^h - v_0^h\|^2_{L^2(\tau)}.$$

We sum this up over $i = 1, ..., N$, while noting that the number of components v_i^h that is nonzero over any τ is bounded above independently of N (an estimate for this bound is N_c, the minimal number of colors needed to color the overlapping subdomains)

$$\sum_{i=1}^{N} |v_i^h|^2_{H^1(\tau)} \leq C|v^h - v_0^h|^2_{H^1(\tau)} + CH^{-2}\|v^h - v_0^h\|^2_{L^2(\tau)}.$$

Summing over all elements τ in the domain Ω, we have

$$\sum_{i=1}^{N} |v_i^h|^2_{H^1(\Omega)} \leq C|v^h - v_0^h|^2_{H^1(\Omega)} + CH^{-2}\|v^h - v_0^h\|^2_{L^2(\Omega)}. \tag{3.4.24}$$

Because v_0^h is constructed by an L^2 projection (3.4.21), the remainder component $v^h - v_0^h$ can be shown to have the properties

$$\|v^h - v_0^h\|^2_{L^2(\Omega)} \leq CH^2|v^h|^2_{H^1(\Omega)}$$

and

$$|v^h - v_0^h|^2_{H^1(\Omega)} \leq C|v^h|^2_{H^1(\Omega)}$$

e.g., [BX91]. These may be combined with (3.4.24) to obtain the two-level bound

$$\sum_{i=0}^{N} |v_i^h|^2_{H^1(\Omega)} \leq C_0^2|v^h|^2_{H^1(\Omega)},$$

where C_0 is independent of H and h.

□

It is straightforward to extend Lemma 3.4.4 to the product space $Z^h = V^h \times V^h$ as follows.

Lemma 3.4.5 *There exists a constant $C_0 > 0$ (independent of H and h) 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*

$$\sum_{i=0}^{N} A(z_i^h, z_i^h) \;\leq\; C_0^2\, A(z^h, z^h). \tag{3.4.25}$$

Proof: For any $y^h \in V^h$ and $p^h \in V^h$, there exist decompositions $y^h = \sum_{i=0}^N y_i^h, y_i^h \in V_i^h$ and $p^h = \sum_{i=0}^N p_i^h, p_i^h \in V_i^h$, such that

$$\begin{aligned}
\sum_{i=0}^{N} a(y_i^h, y_i^h) &\leq C_0^2\, a(y^h, y^h), \\
\sum_{i=0}^{N} a(p_i^h, p_i^h) &\leq C_0^2\, a(p^h, p^h).
\end{aligned}$$

A sufficient decomposition of $z^h = (y^h, p^h) \in Z^h$ is simply $z^h = \sum_{i=0}^N z_i^h = \sum_{i=0}^N (y_i^h, p_i^h)$, which satisfies (3.4.25). □

Next, we will derive some basic error estimates for the optimal control problem that will be used later for bounding the error in the coarse component.

We define two families of operators parameterized by h: $\mathcal{A}_h \in \mathcal{L}(V^h, (V^h)')$ and $\mathcal{M}_h \in \mathcal{L}(V^h, (V^h)')$ as

$$\langle \mathcal{A}_h v^h, \phi^h \rangle_{V',V} \;=\; a(v^h, \phi^h), \tag{3.4.26a}$$

$$\langle \mathcal{M}_h v^h, \phi^h \rangle_{V',V} \;=\; m(v^h, \phi^h), \tag{3.4.26b}$$

where $\langle \cdot, \cdot \rangle_{V',V}$ denotes the duality pairing between V' and V. From the definition of $a(\cdot,\cdot)$ and $m(\cdot,\cdot)$ in (3.1.3), we note that

$$\begin{aligned} \langle \mathcal{A}_h v^h, v^h \rangle_{V',V} &\geq \eta \|v^h\|_V^2 \qquad \forall v^h \in V^h, \quad (3.4.27) \\ \langle \mathcal{M}_h v^h, v^h \rangle_{V',V} &\geq 0 \qquad\quad \forall v^h \in V^h, \quad (3.4.28) \end{aligned}$$

which means $\mathcal{A}_h$ is continuously invertible. The operator $\mathcal{K}_h \in \mathcal{L}(Z^h, (Z^h)')$ associated with the discretized problem (3.4.2) is

$$\mathcal{K}_h = \begin{pmatrix} \mathcal{A}_h & -\alpha^{-1}\mathcal{M}_h \\ \mathcal{M}_h & \mathcal{A}_h \end{pmatrix}. \quad (3.4.29)$$

Lemma 3.4.6 *Let Z^h be equipped with the A-norm. The operator $\mathcal{K}_h$ defined by* (3.4.29) *and* (3.4.26) *is invertible, and there exists a $\kappa > 0$ such that*

$$\|\mathcal{K}_h^{-1}\|_{\mathcal{L}((Z^h)',Z^h)} \leq \kappa \qquad \forall\, h > 0 \quad (3.4.30)$$

Proof: Since a and m are continuous bilinear forms, the families of operators $\mathcal{A}_h$ and $\mathcal{M}_h$ are uniformly bounded. By (3.4.27), the inverse of $\mathcal{A}_h$ exists and is uniformly bounded. It is easy to verify that the inverse of $\mathcal{K}_h$ is

$$\mathcal{K}_h^{-1} = \begin{pmatrix} \mathcal{J}_h^{-1} & \alpha^{-1}\mathcal{A}_h^{-1}\mathcal{M}_h\mathcal{J}_h^{-1} \\ -\mathcal{A}_h^{-1}\mathcal{M}_h\mathcal{J}_h^{-1} & \mathcal{J}_h^{-1} \end{pmatrix}, \quad (3.4.31)$$

where $\mathcal{J}_h = \mathcal{A}_h + \alpha^{-1}\mathcal{M}_h\mathcal{A}_h^{-1}M_h$. Using (3.4.27), we see that

$$\langle \mathcal{J}_h v^h, v^h \rangle_{V',V} \geq \eta \|v^h\|_V^2 \qquad \forall v^h \in V^h,$$

so $\mathcal{J}_h$ and $\mathcal{J}_h^{-1}$ are uniformly bounded. Each "block entry" on the right side of (3.4.31) corresponds to an operator that is uniformly bounded, so $\mathcal{K}_h^{-1}$ is uniformly bounded. □

The next Lemma provides some error estimates for the finite element discretization of the optimal control problem.

Lemma 3.4.7 *Let $\mathcal{K} \in \mathcal{L}(Z, Z')$ be the operator associated with the problem* (3.4.1)*, and let $\mathcal{K}_h \in \mathcal{L}(Z^h, (Z^h)')$ be the operator associated with the discretized problem* (3.4.2)*. Let $l \in Z'$ and assume that the solution $z \in Z$ of*

$$\mathcal{K}z = l \tag{3.4.32}$$

satisfies $z = (y, p) \in (H^2(\Omega) \cap H_0^1(\Omega))^2$. If $z^h \in Z^h$ is the solution of the discretized problem

$$\mathcal{K}_h z^h = l_h, \tag{3.4.33}$$

where $l_h \in (Z^h)'$ is defined by $\langle l_h, \psi^h \rangle_{Z',Z} = \langle l, \psi^h \rangle_{Z',Z}$ for all $\psi^h \in Z^h$, then there exists a constant $C > 0$ independent of h such that

$$\|z - z^h\|_A \;\leq\; Ch\|z\|_{(H^2(\Omega))^2}. \tag{3.4.34}$$

Proof: We derive (3.4.34) first by applying some standard finite element error bounds on the state and adjoint variables. Let $\Pi_h : C^0(\overline{\Omega}) \to V^h$ be the finite element space interpolation operator. Using [Cia02, Thm. 3.2.1], an H^1 error estimates for the interpolant $\Pi_h v$ is

$$\|v - \Pi_h v\|_{H^1(\Omega)} \;\leq\; Ch\|v\|_{H^2(\Omega)} \qquad \forall v \in H^2(\Omega). \tag{3.4.35}$$

We define the restriction operator $\mathcal{R}_h : Z \to Z^h$ as

$$\mathcal{R}_h(z) = (\Pi_h y, \Pi_h p), \tag{3.4.36}$$

which has the error estimate

$$\begin{aligned} \|z - \mathcal{R}_h(z)\|_A &\leq Ch \max\{\|y\|_{H^2(\Omega)}, \|p\|_{H^2(\Omega)}\} \\ &\leq Ch\|z\|_{(H^2(\Omega))^2} \end{aligned} \tag{3.4.37}$$

Subtracting $\mathcal{K}_h\mathcal{R}_h(z)$ from both sides of (3.4.33), we have

$$\mathcal{K}_h(z^h - \mathcal{R}_h(z)) = l_h - \mathcal{K}_h\mathcal{R}_h(z), \tag{3.4.38}$$

which yields the error estimate

$$\begin{aligned}\|z^h - \mathcal{R}_h(z)\|_A &\leq \|\mathcal{K}_h^{-1}\|_{\mathcal{L}((Z^h)',Z^h)}\|l_h - \mathcal{K}_h\mathcal{R}_h(z)\|_{(Z^h)'} \\ &\leq \kappa\|l_h - \mathcal{K}_h\mathcal{R}_h(z)\|_{(Z^h)'}, \end{aligned} \tag{3.4.39}$$

using Lemma 3.4.6. The last factor on the right can be bounded as

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

Using (3.4.32),

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

The estimate (3.4.39) is now

$$\begin{aligned}\|z^h - \mathcal{R}_h(z)\|_A &\leq \kappa\|\mathcal{K}\|_{\mathcal{L}(Z,Z')}\|z - \mathcal{R}_h(z)\|_A, \\ &\leq C\|z - \mathcal{R}_h(z)\|_A, \end{aligned} \tag{3.4.40}$$

for some $C > 0$ independent of h. We derive (3.4.34) by combining (3.4.40) and (3.4.37) as

$$\begin{aligned}\|z - z^h\|_A &\leq \|z - \mathcal{R}_h(z)\|_A + \|z^h - \mathcal{R}_h(z)\|_A, \\ &\leq C\|z - \mathcal{R}_h(z)\|_A, \\ &\leq Ch\|z\|_{(H^2(\Omega))^2}.\end{aligned}$$

□

The next Lemma provides some norm bounds on the coarse component in terms of the H^1 norms of the fine grid component. To derive these results, we need to make an assumption on the regularity of the problem, then apply the Aubin-Nitsche method ([Cia02, Sec. 3.2], [Hac92, Sec. 8.4], [AB01, Sec. 5.4]).

Assumption A1: For every $l \in (L^2(\Omega))^2$, the solution w of the adjoint problem

$$K(\psi, w) = (l, \psi)_{(L^2(\Omega))^2} \qquad \forall \psi \in Z \tag{3.4.41}$$

satisfies $w \in (H^2(\Omega) \cap H_0^1(\Omega))^2$, and

$$\|w\|_{(H^2(\Omega))^2} \leq C\|l\|_{(L^2(\Omega))^2} \tag{3.4.42}$$

for some $C > 0$ independent of l.

Lemma 3.4.8 *Let $z^h \in Z^h$. The solution z^H of the problem*

$$K(z^H, \psi^H) = K(z^h, \psi^H) \qquad \forall \psi^H \in Z^H \tag{3.4.43}$$

satisfies

$$\|z^H\|_A \leq C_1\|z^h\|_A, \tag{3.4.44}$$

and

$$\|z^h - z^H\|_M \leq C_2 H\|z^h\|_A \tag{3.4.45}$$

for some positive constants C_1, C_2 independent of H and h.

Proof: We note that problem (3.4.43) has a unique solution corresponding to a coarse grid optimal control problem (see Lemma 2.2.1). We may write the problem as

$$\mathcal{K}_H z^H = \mathcal{K}_h z^h \tag{3.4.46}$$

where $\mathcal{K}_H \in \mathcal{L}(Z^H, (Z^H)')$ is defined analogously to $\mathcal{K}_h$ in (3.4.29) and (3.4.26). That is,

$$\mathcal{K}_H = \begin{pmatrix} \mathcal{A}_H & -\alpha^{-1}\mathcal{M}_H \\ \mathcal{M}_H & \mathcal{A}_H \end{pmatrix},$$

which is invertible. This means

$$z^H = \mathcal{K}_H^{-1}\mathcal{K}_h z^h,$$

and since $\mathcal{K}_h$ and $\mathcal{K}_H^{-1}$ are uniformly bounded,

$$\|z^H\|_A \leq C\|z^h\|_A. \tag{3.4.47}$$

for some $C > 0$ independent of h.

To derive (3.4.45), we consider the adjoint problems

$$K(\psi, w) = (z^h - z^H, \psi) \qquad \forall \psi \in Z, \tag{3.4.48}$$

$$K(\psi^h, w^h) = (z^h - z^H, \psi^h) \qquad \forall \psi^h \in Z^h, \tag{3.4.49}$$

$$K(\psi^H, w^H) = (z^h - z^H, \psi^H) \qquad \forall \psi^H \in Z^H, \tag{3.4.50}$$

with $w \in Z$, $w^h \in Z^h$, $w^H \in Z^H$. Setting $\psi^h = z^h - z^H$ in (3.4.49), we have

$$K(z^h - z^H, w^h) = \|z^h - z^H\|^2_{(L^2(\Omega))^2}. \tag{3.4.51}$$

Problem (3.4.43) is equivalent to

$$K(z^h - z^H, \psi^H) = 0 \qquad \forall \psi^H \in Z^H,$$

which also holds for $\psi^H = w^H \in Z^H$:

$$K(z^h - z^H, w^H) = 0. \tag{3.4.52}$$

Subtracting (3.4.51) and (3.4.52), we have

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

We now use Lemma 3.4.1 to bound $K(\cdot,\cdot)$

$$\|z^h - z^H\|^2_{(L^2(\Omega))^2} \leq C_K \max\{1, \alpha^{-1}\}\|z^h - z^H\|_{(H^1(\Omega))^2}\|w^h - w^H\|_{(H^1(\Omega))^2}. \quad (3.4.53)$$

Let w be a solution of (3.4.48), with finite element approximations w^h and w^H. The error may be bounded by Lemma 3.4.7 as

$$\begin{aligned} \|w^h - w\|_A &\leq Ch\|w\|_{(H^2(\Omega))^2}, \\ \|w^H - w\|_A &\leq CH\|w\|_{(H^2(\Omega))^2}. \end{aligned}$$

This allows the bound

$$\begin{aligned} \|w^h - w^H\|_A &\leq \|w^h - w\|_A + \|w^H - w\|_A \\ &\leq CH\|w\|_{(H^2(\Omega))^2}, \\ &\leq CH\|z^h - z^H\|_{(L^2(\Omega))^2}, \end{aligned}$$

where we use the assumption $A1$ to bound the last step. Inequality (3.4.53) is now

$$\|z^h - z^H\|^2_{(L^2(\Omega))^2} \leq C_K \max\{1, \alpha^{-1}\}\|z^h - z^H\|_{(H^1(\Omega))^2} CH\|z^h - z^H\|_{(L^2(\Omega))^2}.$$

Canceling one L^2 norm factor from each side yields

$$\|z^h - z^H\|_{(L^2(\Omega))^2} \leq CH\|z^h - z^H\|_{(H^1(\Omega))^2}. \quad (3.4.54)$$

Using (3.4.47) in (3.4.54), we have

$$\|z^h - z^H\|_M \leq CH\|z^h\|_A.$$

□

We note that by the definition of T_0, the bounds (3.4.44) and (3.4.45) are equivalent to

$$\|T_0 z^h\|_A \leq C_1\|z^h\|_A, \quad (3.4.55)$$

and

$$\|T_0 z^h - z^h\|_M \;\leq\; C_2 H \|z^h\|_A. \tag{3.4.56}$$

The following Lemma states a similar L^2 norm bound in terms of the H^1 norm for the local components $z_i^h, i > 0$.

Lemma 3.4.9 *Assuming that each overlapping region $\widehat{\Omega}_i, i = 1, ..., N$ has diameter less than or equal to $C_\delta H$, then*

$$\|z_i^h\|_M \;\leq\; \tfrac{1}{2} C_\delta H \|z_i^h\|_A \qquad \forall z_i \in Z_i^h,\ i = 1, ..., N. \tag{3.4.57}$$

Proof: Let $(y_i^h, p_i^h) = z_i^h$, then since z_i^h is zero outside of $\widehat{\Omega}_i$, by a Poincaré inequality, there exists a constant $C_{\widehat{\Omega}_i} > 0$ such that

$$\begin{aligned} \|z_i^h\|_M^2 &= \|y_i^h\|^2_{L^2(\widehat{\Omega}_i)} + \|p_i^h\|^2_{L^2(\widehat{\Omega}_i)} \\ &\leq C^2_{\widehat{\Omega}_i} (|y_i^h|^2_{H^1(\widehat{\Omega}_i)} + |p_i^h|^2_{H^1(\widehat{\Omega}_i)}), \end{aligned}$$

where (using [GT01, eqn. 7.44] with $|B_d|$ = volume of unit ball in $\mathbb{R}^d$)

$$C_{\widehat{\Omega}_i} \;=\; (|\widehat{\Omega}_i|/|B_d|)^{1/d} \;\leq\; C_\delta H/2.$$

Therefore,

$$\begin{aligned} \|z_i^h\|_M^2 &\leq \tfrac{1}{4} C_\delta^2 H^2 (|y_i^h|^2_{H^1(\widehat{\Omega}_i)} + |p_i^h|^2_{H^1(\widehat{\Omega}_i)}) \\ &\leq \tfrac{1}{4} C_\delta^2 H^2 \|z_i^h\|_A^2. \end{aligned}$$

□

The following Lemma says that the bilinear form $K(\cdot, \cdot)$ is positive definite on the subspaces $Z_i^h, i = 1, ..., N$, if H is sufficiently small.

Lemma 3.4.10 *Assume that each overlapping region* $\widehat{\Omega}_i, i = 1, ..., N$ *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 for all local spaces* $Z_i, i = 1, ..., N$

$$K(z_i^h, z_i^h) \;\geq\; C_{H_0} A(z_i^h, z_i^h) \qquad \forall z_i^h \in Z_i^h. \tag{3.4.58}$$

Proof: Let $z_i^h = (y_i^h, p_i^h) \in Z_i^h$, then

$$\begin{aligned}
K(z_i^h, z_i^h) &= a(y_i^h, y_i^h) + a(p_i^h, p_i^h) + m(y_i^h, p_i^h) - \alpha^{-1} m(p_i^h, y_i^h) \\
&\geq a(y_i^h, y_i^h) + a(p_i^h, p_i^h) - |\alpha^{-1} - 1|\, |m(y_i^h, p_i^h)| \\
&\geq a(y_i^h, y_i^h) + a(p_i^h, p_i^h) - \max\{1, \alpha^{-1}\}\, \tfrac{1}{2}[m(y_i^h, y_i^h) + m(p_i^h, p_i^h)] \\
&\geq a(y_i^h, y_i^h) + a(p_i^h, p_i^h) - \max\{1, \alpha^{-1}\}\, \tfrac{1}{8} C_\delta^2 H^2 [a(y_i^h, y_i^h) + a(p_i^h, p_i^h)] \\
&= (1 - \tfrac{1}{8} \max\{1, \alpha^{-1}\} C_\delta^2 H^2) A(z_i^h, z_i^h),
\end{aligned}$$

with Lemma 3.4.9 being used to bound $m(\cdot, \cdot)$ in terms of $a(\cdot, \cdot)$. If we set H_0 so that the factor in front of $A(z_i^h, z_i^h)$ is positive, e.g.,

$$H_0 = \left(\frac{8(1 - \epsilon)}{C_\delta^2 \max\{1, \alpha^{-1}\}} \right)^{1/2}$$

for some $\epsilon \in (0, 1)$, then for $H \leq H_0$ inequality (3.4.58) is satisfied with $C_{H_0} = (1 - \frac{1}{8} \max\{1, \alpha^{-1}\} C_\delta^2 H_0^2)$.

□

Lemma 3.4.11 *Let* H_0 *be defined as in Lemma 3.4.10. There exists a constant* $C_{H_1} > 0$ *such that if* $H \leq H_0$ *then*

$$\sum_{i=0}^{N} A(T_i z^h, T_i z^h) \;\leq\; C_{H_1} A(z^h, z^h) \qquad \forall z^h \in Z^h. \tag{3.4.59}$$

Proof: We let $H \leq H_0$ so that $K(\cdot, \cdot)$ is positive definite. For the coarse component, inequality (3.4.55) gives

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

For the local components, we use (3.4.58) and the definition of $T_i, i = 1, ..., N$,

$$\begin{aligned} A(T_i z^h, T_i z^h) &\leq C_{H_0}^{-1} K(T_i z^h, T_i z^h) \\ &= C_{H_0}^{-1} K(z^h, T_i z^h). \end{aligned}$$

We sum over $i = 1, ..., N$, then use Lemma 3.4.1 and Lemma 3.4.3

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

This implies

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

so we may use $C_{H_1} = C_{H_0}^{-2} C_K^2 \max\{1, \alpha^{-1}\}^2 N_O + C_1^2$.

□

Lemma 3.4.12 *There exist constants $H_2 > 0$, $C_{H_2} > 0$ such that if $H \leq H_2$ then*

$$\sum_{i=0}^{N} A(T_i z^h, T_i z^h) \geq C_{H_2} A(z^h, z^h) \qquad \forall z^h \in Z^h. \tag{3.4.60}$$

Proof: Lemma 3.4.5 allows us to express any $z^h \in Z^h$ as a two-level decomposition $z^h = \sum_{i=0}^{N} z_i^h, z_i^h \in Z_i^h$ that has the "nice" property of low energy sum. We derive an upper bound for $K(z^h, z^h)$ by using this decomposition of z^h, the definition of T_i and

Lemma 3.4.1,

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

Applying the Cauchy-Schwarz inequality and Lemma 3.4.5,

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

A lower bound for $K(z^h, z^h)$ may be derived using Lemmas 3.4.2 and 3.4.8

$$\begin{aligned} K(z^h, z^h) &= A(z^h, z^h) + N(z^h, z^h) \\ &\geq A(z^h, z^h) - |N(z^h, z^h)| \\ &\geq \|z^h\|_A^2 - \tfrac{1}{2} \max\{1, \alpha^{-1}\} \|z^h\|_M^2 \\ &\geq \|z^h\|_A^2 - \tfrac{1}{2} \max\{1, \alpha^{-1}\} \|z^h - T_0 z^h\|_M^2 - \tfrac{1}{2} \max\{1, \alpha^{-1}\} \|T_0 z^h\|_M^2 \\ &\geq \|z^h\|_A^2 - \tfrac{1}{2} \max\{1, \alpha^{-1}\} C_2^2 H^2 \|z^h\|_A^2 - \tfrac{1}{2} \max\{1, \alpha^{-1}\} C_\Omega^2 \|T_0 z^h\|_A^2 \\ &\geq (1 - \tfrac{1}{2} \max\{1, \alpha^{-1}\} C_2^2 H^2) \|z^h\|_A^2 - \tfrac{1}{2} \max\{1, \alpha^{-1}\} C_\Omega^2 C_1 \|z^h\|_A \|T_0 z^h\|_A. \end{aligned}$$

Combining the upper and lower bounds and canceling $\|z^h\|_A$ on each side,

$$\begin{aligned} &C_K \max\{1, \alpha^{-1}\} C_0 (\textstyle\sum_{i=0}^{N} \|T_i z^h\|_A^2)^{1/2} \\ &\qquad \geq (1 - \tfrac{1}{2} \max\{1, \alpha^{-1}\} C_2^2 H^2) \|z^h\|_A - \tfrac{1}{2} \max\{1, \alpha^{-1}\} C_\Omega^2 C_1 \|T_0 z^h\|_A, \end{aligned}$$

or

$$\begin{aligned} &(C_K \max\{1, \alpha^{-1}\} C_0 + \tfrac{1}{2} \max\{1, \alpha^{-1}\} C_\Omega^2 C_1) (\textstyle\sum_{i=0}^{N} \|T_i z^h\|_A^2)^{1/2} \\ &\qquad \geq (1 - \tfrac{1}{2} \max\{1, \alpha^{-1}\} C_2^2 H^2) \|z\|_A. \end{aligned}$$

Let $H_2 = ((1-\epsilon)2\max\{1,\alpha^{-1}\}^{-1}C_2^{-2})^{1/(2)}$, for some $\epsilon \in (0,1)$. If $H \leq H_2$ then squaring both sides gives

$$\begin{aligned}(C_K \max\{1,\alpha^{-1}\}C_0 + \tfrac{1}{2}\max\{1,\alpha^{-1}\}C_\Omega^2 C_1)^2(\textstyle\sum_{i=0}^N \|T_i z^h\|_A^2) \\ \geq\ (1-\tfrac{1}{2}\max\{1,\alpha^{-1}\}C_2^2 H^2)^2\|z^h\|_A^2.\end{aligned}$$

This implies, for $H \leq H_2$,

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

with

$$C_{H_2} = \frac{(1-\frac{1}{2}\max\{1,\alpha^{-1}\}C_2^2 H_2^2)^2}{(\max\{1,\alpha^{-1}\})^2(C_K C_0 + \frac{1}{2}C_\Omega^2 C_1)^2}.$$

□

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

Lemma 3.4.13 *There exist constants $H_3 > 0$, $C_{H_3} > 0$ such that if $H \leq H_3$ then*

$$|\sum_{i=1}^N N(T_i z^h - z^h, T_i z^h)| \ \leq\ C_{H_3} H A(z^h, z^h). \tag{3.4.61}$$

Proof: Let H_0 be defined as in Lemma 3.4.10, and assume that $H \leq H_0$. We note that

$$|\sum_{i=1}^N N(T_i z^h - z^h, T_i z^h)| \ \leq\ |\sum_{i=1}^N N(T_i z^h, T_i z^h)| + |N(z^h, \sum_{i=1}^N T_i z^h)|. \tag{3.4.62}$$

The first term on the right is bounded by Lemmas 3.4.2, 3.4.9 and 3.4.11

$$\begin{aligned}|\sum_{i=1}^N N(T_i z^h, T_i z^h)| &\leq \tfrac{1}{2}\max\{1,\alpha^{-1}\}\sum_{i=1}^N \|T_i z^h\|_M^2 \\ &\leq \tfrac{1}{8}\max\{1,\alpha^{-1}\}C_\delta^2 H^2 \sum_{i=1}^N \|T_i z^h\|_A^2 \\ &\leq \tfrac{1}{8}\max\{1,\alpha^{-1}\}C_\delta^2 C_{H_1} H^2 A(z^h, z^h) \end{aligned} \tag{3.4.63}$$

The second term on the right of (3.4.62) is bounded using Lemmas 3.4.2, 3.4.3 and 3.4.9

$$\begin{aligned}
|N(z^h, \sum_{i=1}^{N} T_i z^h)| &\leq C_N \max\{1, \alpha^{-1}\} \|z^h\|_A \| \sum_{i=1}^{N} T_i z^h \|_M \\
&\leq C_N \max\{1, \alpha^{-1}\} \|z^h\|_A N_O^{1/2} (\sum_{i=1}^{N} \|T_i z^h\|_M^2)^{1/2} \\
&\leq C_N \max\{1, \alpha^{-1}\} \|z^h\|_A N_O^{1/2} (\sum_{i=1}^{N} \tfrac{1}{4} C_\delta^2 H^2 \|T_i z^h\|_A^2)^{1/2} \\
&\leq \tfrac{1}{2} C_N \max\{1, \alpha^{-1}\} C_\delta N_O^{1/2} H \|z^h\|_A (\sum_{i=1}^{N} \|T_i z^h\|_A^2)^{1/2}.
\end{aligned}$$

The last factor on the right can be bounded by Lemma 3.4.11, so that

$$|N(z^h, \sum_{i=1}^{N} T_i z^h)| \leq \tfrac{1}{2} C_N \max\{1, \alpha^{-1}\} C_\delta C_{H_1}^{1/2} N_O^{1/2} H A(z^h, z^h). \quad (3.4.64)$$

The H^2 factor in (3.4.63) goes to zero faster than H in (3.4.64), so if H_3 is chosen small enough so that the first term in (3.4.62) is smaller than the second, i.e.,

$$\tfrac{1}{8} \max\{1, \alpha^{-1}\} C_\delta^2 C_{H_1} H_3^2 \leq \tfrac{1}{2} C_N \max\{1, \alpha^{-1}\} C_\delta C_{H_1}^{1/2} N_O^{1/2} H_3,$$

or

$$H_3 = \min\{H_0, 4 C_N C_\delta^{-1} C_{H_1}^{-1/2} N_O^{1/2}\},$$

then (3.4.61) holds for $H \leq H_3$, with $C_{H_3} = C_N \max\{1, \alpha^{-1}\} C_\delta C_{H_1}^{1/2} N_O^{1/2}$. □

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

Lemma 3.4.14 *There exists a constant $C_{H_{3C}} > 0$ such that*

$$|N(T_0 z^h - z^h, T_0 z^h)| \leq C_{H_{3C}} H A(z^h, z^h). \quad (3.4.65)$$

Proof: We use Lemma 3.4.2 followed by Lemma 3.4.8,

$$\begin{aligned} |N(T_0 z^h - z^h, T_0 z^h)| &\leq C_N \max\{1, \alpha^{-1}\} \|T_0 z^h\|_A \, \|T_0 z^h - z^h\|_M, \\ &\leq C_N \max\{1, \alpha^{-1}\} C_1 \|z^h\|_A \, C_2 H \|z^h\|_A, \\ &\leq C_{H_{3C}} H A(z^h, z^h), \end{aligned}$$

with $C_{H_{3C}} = C_N \max\{1, \alpha^{-1}\} C_1 C_2$.

□

We are now ready to combine the previous results to prove our main convergence theorem. As a reminder, the GMRES convergence rate can be specified by two parameters:

$$\begin{aligned} c_T &= \inf_{z \neq 0} \frac{A(Tz, z)}{A(z, z)}, \\ C_T &= \sup_{z \neq 0} \frac{\|Tz\|_A}{\|z\|_A}, \end{aligned}$$

assuming that $c_T > 0$ (where c_T is the minimal eigenvalue of the Hermitian part of T, and C_T is the A-norm of T.) We will show that both c_T and C_T can be bounded independently of H and h for H sufficiently small.

Theorem 3.4.1 *For the two-level additive overlapping method, using exact subdomain solves and uniform overlap proportional to the subdomain diameter H, if H is sufficiently small then the convergence rate can be bounded independently of H and h.*

Proof: From the definition of T_i,

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

for all $\psi_i^h \in Z_i^h$. Setting $\psi_i^h = T_i z^h$ and summing over $i = 0, ..., N$, we have

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

We let H be at least as small as $\min\{H_2, H_3\}$, then bound the first term on the right from below using Lemma 3.4.12

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

and bound the second term on the right from above by Lemmas 3.4.13 and 3.4.14

$$\begin{aligned}|\sum_{i=1}^{N} N(T_i z^h - z^h, T_i z^h)| &\leq C_{H_3} H A(z^h, z^h), \\ |N(T_0 - z^h, T_0 z^h)| &\leq C_{H_{3C}} H A(z^h, z^h).\end{aligned}$$

Therefore,

$$A(T z^h, z^h) \geq (C_{H_2} - C_{H_3} H - C_{H_{3C}} H) A(z^h, z^h) \tag{3.4.66}$$

where the constants are independent of H or h. For H sufficiently small, $(C_{H_2} - C_{H_3} H - C_{H_{3C}} H)$ is positive, so that

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

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

To bound the norm of T, we use Lemma 3.4.3 and Lemma 3.4.11.

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

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

We note that this convergence property holds for all choices of the regularization parameter $\alpha > 0$. Even though the constants $C_{H_1}, C_{H_2}, C_{H_3}, C_{H_{3C}}$ depend on α, they only affect how slow the worst case can be, not the qualitative (mesh independence) behavior of the two-level method when H is sufficiently small.

For the nonsymmetric PDE considered by Cai and Widlund [CW92], there is an alternative way to precondition the problem, where only the symmetric positive definite part of the bilinear form is used in the local operators $T_i, i > 0$. Both their theoretical and numerical results indicate that this second method (their Algorithm 2) has similar behavior to the first method (Algorithm 1.) This approach of using only the symmetric positive definite part in the preconditioner is discussed further in [XC92] and [Xu92b].

Our theory also predicts a similarity between the two methods as long as α is not too small. In the alternative method, the definition of T_i in (2.3.4) is modified to leave out the nonsymmetric term $N(T_i z, \psi_i)$, so that

$$A(T_i z^h, \psi_i^h) \;=\; A(z^h, \psi_i^h) + N(z^h, \psi_i^h) \qquad \forall \psi_i^h \in Z_i^h. \tag{3.4.67}$$

The effect can be seen in Lemma 3.4.13, where the absolute sum $|\sum_{i=1}^{N} N(T_i z^h, T_i z^h)|$ of inequalities (3.4.62) and (3.4.63) now does not exist. Since this sum is bounded by an H^2 factor, its contribution to the right hand side of (3.4.62) diminishes faster than the second term as H decreases. This suggests for large α the effect of dropping this term is not significant. Our numerical experiments show that when $\alpha > 10^{-2}$, the two methods give nearly identical results, while for $\alpha < 10^{-6}$ the original method is much better.

3.4.2 One-Level Method

We now consider the difficulties that would arise if the convergence analysis of the previous section is applied to the one-level method. The crucial difference between the two methods can be seen in Lemma 3.4.4 (and consequently Lemma 3.4.5.) To achieve the bound

$$\sum_{i=0}^{N} a(v_i^h, v_i^h) \;\leq\; C_0^2\, a(v^h, v^h) \qquad \forall v^h \in V^h \tag{3.4.68}$$

(for some constant $C_0 > 0$ independent of H and h), a global component v_0^h is required in order to reduce the energy of the remaining components enough to bound the sum independently of the number of subdomains N. In general, it is not possible to bound the local component energy $a(v_i^h, v_i^h), i > 0$, by $a(v^h|_{\widehat{\Omega}_i}, v^h|_{\widehat{\Omega}_i})$, the energy of v^h restricted to the overlapping region $\widehat{\Omega}_i$. This is due to those functions v^h that are constant over $\widehat{\Omega}_i$, which would have zero energy and cannot bound the positive energy of $a(v_i^h, v_i^h)$. It is possible to bound each local energy in terms of the energy over the entire domain Ω

$$a(v_i^h, v_i^h) \;\leq\; C\, a(v^h, v^h), \tag{3.4.69}$$

but this would yield the disappointing result

$$\sum_{i=1}^{N} a(v_i^h, v_i^h) \;\leq\; CN\, a(v^h, v^h). \tag{3.4.70}$$

Since N is $O(H^{-2})$ in 2-D, instead of Lemma 3.4.5 we now have the bound

$$\sum_{i=1}^{N} A(z_i^h, z_i^h) \;\leq\; \tilde{C}_0^2 H^{-2}\, A(z^h, z^h) \tag{3.4.71}$$

for some constant $\tilde{C}_0 > 0$ independent of H and h.

If we follow through the analysis of the previous section for the one-level method, then Lemmas 3.4.8 and 3.4.14 can be omitted. Other Lemmas either remain unchanged or need only a change in the starting index of the sum (begin at $i = 1$ instead of $i = 0$).

Lemma 3.4.11 is essentially the same, but without the component T_0:

$$\sum_{i=1}^{N} A(T_i z^h, T_i z^h) \;\; \leq \;\; \tilde{C}_{H_1} A(z^h, z^h) \tag{3.4.72}$$

for some $\tilde{C}_{H_1} > 0$, if $H < H_1$. This implies the norm of T for the one-level method can be bounded similarly to the two-level method (i.e., independent of H and h.)

Lemma 3.4.12 no longer applies because it depends on Lemma 3.4.5 to derive an upper bound on $K(z^h, z^h)$. If we use the bound (3.4.71) to continue the analysis, then the result is the disappointing bound

$$\sum_{i=1}^{N} A(T_i z^h, T_i z^h) \;\; \geq \;\; \tilde{C}_2 H^2 A(z^h, z^h) \tag{3.4.73}$$

for some $\tilde{C}_2 > 0$.

For the nonsymmetric part, Lemma 3.4.13 remains unchanged, so we have the following one-level bound instead of (3.4.66)

$$A(T z^h, z^h) \;\; \geq \;\; (\tilde{C}_2 H^2 - C_{H_3} H) A(z^h, z^h).$$

This means the minimal eigenvalue of the Hermitian part of T cannot be bounded away from zero as H is decreased.

3.4.3 Small Overlap

The convergence analysis for the two-level method in Section 3.4 assumes that the amount of overlap is kept proportional to the subdomain diameter H (i.e., $\delta = \beta H$ for some constant $\beta > 0$.) For PDE problems, numerical experiments show that the two-level overlapping method performs well even when the overlap δ is kept fixed at a small amount (e.g., $\delta \approx h$.) This desirable property is also observed for our optimal control problem. In this section, we consider the effect on the convergence rate when the overlap δ is allowed to decrease independently of H.

The two convergence parameters c_T and C_T from the proof of Theorem 3.4.1 are bounded in terms of a few constants that are independent of H and h, but may be dependent on the overlap δ. We consider how each of these constants may change when δ approaches zero (and C_δ approaches 1.)

The parameter C_T (the A-norm of T) is bounded by constants N_O and C_{H_1}. From the proof of Lemma 3.4.3, it is easy to see that the bound $N_O \leq N_c + 1$ is unaffected by small δ. The constant C_{H_1} depends on the constant C_{H_0}, but we see from Lemmas 3.4.10 and 3.4.11 that neither constant is affected significantly by δ approaching zero. Thus, the norm of T remains approximately unchanged in this situation.

The parameter c_T (minimal eigenvalue of the Hermitian part of T) depends on C_{H_2}, C_{H_3} and $C_{H_{3C}}$. Considering C_{H_3} in Lemma 3.4.13 and $C_{H_{3C}}$ in Lemma 3.4.14, we see that these constants are well behaved as δ goes to zero.

The remaining constant C_{H_2} is the only one with a significant dependence on δ, because it depends on the partition constant C_0 of Lemma 3.4.4. As H decreases, c_T approaches C_{H_2} from above, so we are interested in finding a lower bound for C_{H_2} in terms of δ. From the estimate for C_{H_2} in the proof of Lemma 3.4.12, if H is sufficiently small (e.g., $H < H_2$) then we can bound $C_{H_2} \geq CC_0^{-2}$ for some $C > 0$ independent of δ.

The partition constant C_0 depends on how the finite element space V^h is decomposed into subspaces $V_0^h, V_1^h, ..., V_N^h$. As seen in Lemma 3.4.4, C_0 reflects how these subspaces interact, so it depends on the amount of overlap between the subdomains. In the proof of Lemma 3.4.4, we assume that the overlap between each region is kept proportional to H, so the gradient of any component of the partition of unity can be bounded above by

$$\|\nabla \theta_i\|_{L^\infty(\widehat{\Omega}_i)} \leq C/H. \tag{3.4.74}$$

With a constant small overlap of δ, this bound is now replaced by

$$\|\nabla \theta_i\|_{L^\infty(\widehat{\Omega}_i)} \leq C/\delta. \tag{3.4.75}$$

By continuing the analysis with this substitution, the constant C_0 has been shown by Dryja and Widlund [DW94, Thm. 3] to have the bound

$$C_0^2 \leq C(1 + H/\delta) \tag{3.4.76}$$

for some C independent of δ (see also [SBG96, Thm. 2].)

Since we use the same space V^h and the same two-level decomposition for the state and adjoint variables, this estimate also applies to the control problem. Using (3.4.66), we may bound the minimal eigenvalue of the symmetric part of the two-level T as

$$c_T \;\geq\; C(1 + H/\delta)^{-1} - C_{H_3} H - C_{H_{3C}} H^{\gamma}. \tag{3.4.77}$$

For H sufficiently small, c_T remains bounded away from zero for a positive, fixed overlap δ, implying the convergence does not deteriorate for many subdomains. This bound also agrees with the observation that a larger overlap tends to improve the convergence rate.

The bound (3.4.77) suggests that convergence can be slow if the overlap is set to zero. The difficulty is due to the local subspaces $V_i^h, i > 0$, being associated with nonoverlapping subdomains, so that they can no longer represent any function that is nonzero on the interface between adjacent subdomains. This means none of the global error on the interface can be projected onto these subspaces to be corrected locally. Numerical experiments confirm that convergence for the two-level method is poor in this case.

3.5 Algebraic Viewpoint

The overlapping methods for optimal control may be formulated at the algebraic level, which is useful for explaining many implementation issues. We consider the model

control problem (3.1.1) and its equivalent weak formulation (3.1.2):

$$\min_{y\in V,u\in U} \quad \tfrac{1}{2}m(y,y) - c(y) + \frac{\alpha}{2}q(u,u) \tag{3.5.1a}$$

$$\text{s.t.} \quad a(y,\phi) + b(u,\phi) = f(\phi) \quad \forall \phi \in V, \tag{3.5.1b}$$

with $V = H_0^1(\Omega)$ and $U = L^2(\Omega)$. After discretization with a finite element method, we have the problem

$$\min_{y\in V^h,u\in U^h} \quad \tfrac{1}{2}m(y,y) - c(y) + \frac{\alpha}{2}q(u,u) \tag{3.5.2a}$$

$$\text{s.t.} \quad a(y,\phi) + b(u,\phi) = f(\phi) \quad \forall \phi \in V^h, \tag{3.5.2b}$$

with finite element spaces

$$\begin{aligned} V^h &= \{v^h : v^h \in C^0(\Omega), v^h|_{\tau_j} \text{ linear}, v^h = 0 \text{ on } \partial\Omega\}, \\ U^h &= V^h. \end{aligned}$$

If V^h has the basis $\{\phi_j\}_{j=1}^n$, then every $v \in V^h$ can be written as

$$v(x) = \sum_{j=1}^{n} v_j \phi_j(x) \tag{3.5.3}$$

for some $v_j \in \mathbb{R}, j = 1, ..., n$. We use bold face for the corresponding vector of coefficients representing v:

$$\mathbf{v} = (v_1, ..., v_n)^T. \tag{3.5.4}$$

Every bilinear form on $V^h \times V^h$ or $U^h \times V^h$ can be associated with a matrix. For example, there exists a matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ corresponding to the bilinear form $a(\cdot,\cdot)$ and operator $A \in \mathcal{L}(V^h, (V^h)')$ such that

$$a(u,v) = \langle Au, v\rangle = \mathbf{v}^T\mathbf{A}\mathbf{u} \qquad \forall u, v \in V^h. \tag{3.5.5}$$

The equivalent algebraic formulation of (3.5.2) is

$$\min_{\mathbf{y},\mathbf{u}} \quad \tfrac{1}{2}\mathbf{y}^T\mathbf{M}\mathbf{y} + \frac{\alpha}{2}\mathbf{u}^T\mathbf{Q}\mathbf{u} - \mathbf{c}^T\mathbf{y} \tag{3.5.6a}$$

$$\text{s.t.} \quad \mathbf{A}\mathbf{y} + \mathbf{B}\mathbf{u} = \mathbf{b}, \tag{3.5.6b}$$

with matrix and vector entries defined by

$$\begin{aligned}
\mathbf{M}_{j,k} &= \int_\Omega \phi_j(x)\,\phi_k(x)\,dx \\
\mathbf{Q}_{j,k} &= \int_\Omega \phi_j(x)\,\phi_k(x)\,dx \\
\mathbf{A}_{j,k} &= \int_\Omega \nabla\phi_j(x)\cdot\nabla\phi_k(x)\,dx \\
\mathbf{B}_{j,k} &= -\int_\Omega \phi_j(x)\,\phi_k(x)\,dx \\
\mathbf{c}_j &= \int_\Omega y_d(x)\,\phi_j(x)\,dx \\
\mathbf{b}_j &= \int_\Omega f(x)\,\phi_j(x)\,dx.
\end{aligned}$$

The optimality conditions may be derived by defining the discrete Lagrangian as

$$L(\mathbf{y},\mathbf{u},\mathbf{p}) = \frac{1}{2}\mathbf{y}^T\mathbf{M}\mathbf{y} + \frac{\alpha}{2}\mathbf{u}^T\mathbf{Q}\mathbf{u} - \mathbf{c}^T\mathbf{y} + \mathbf{p}^T(\mathbf{A}\mathbf{y} + \mathbf{B}\mathbf{u} - \mathbf{b}). \tag{3.5.7}$$

By setting to zero the partial derivative of L with respect to each argument, we arrive at the first order necessary optimality conditions, known as the KKT system

$$\begin{pmatrix} \mathbf{M} & \mathbf{0} & \mathbf{A}^T \\ \mathbf{0} & \alpha\mathbf{Q} & \mathbf{B}^T \\ \mathbf{A} & \mathbf{B} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{y} \\ \mathbf{u} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} \mathbf{c} \\ \mathbf{0} \\ \mathbf{b} \end{pmatrix}. \tag{3.5.8}$$

Since the objective function of (3.5.6) is strictly convex for positive α, this condition is also sufficient.

The control $\mathbf{u} = \alpha^{-1}\mathbf{p}$ is eliminated from this system by using the gradient equation. After a block row interchange, this results in the following system of coupled (discretized) PDEs

$$\begin{pmatrix} \mathbf{A} & -\alpha^{-1}\mathbf{M} \\ \mathbf{M} & \mathbf{A}^T \end{pmatrix} \begin{pmatrix} \mathbf{y} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} \mathbf{b} \\ \mathbf{c} \end{pmatrix}. \tag{3.5.9}$$

We write this system in abbreviated form as

$$\mathbf{K}\,\mathbf{z} = \mathbf{g}. \tag{3.5.10}$$

Since $\mathbf{A}^T = \mathbf{A}$ in this case, the system matrix may be written as the sum of a symmetric positive definite component and a nonsymmetric component:

$$\begin{pmatrix} \mathbf{A} & -\alpha^{-1}\mathbf{M} \\ \mathbf{M} & \mathbf{A}^T \end{pmatrix} = \begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{A} \end{pmatrix} + \begin{pmatrix} \mathbf{0} & -\alpha^{-1}\mathbf{M} \\ \mathbf{M} & \mathbf{0} \end{pmatrix}. \tag{3.5.11}$$

This splitting is the algebraic equivalent of the bilinear form splitting (3.4.3) used in the convergence analysis.

The reduced KKT system (3.5.10) is solved by a preconditioned Krylov subspace method, with the overlapping methods implemented as preconditioners. The one-level and two-level preconditioners are built by combining some basic components, each of which represents a subspace correction operator. For our optimal control problem, each subspace correction corresponds to solving a smaller control problem, either on an extended subdomain or on the coarse grid. We define the basic component for each subdomain $\widehat{\Omega}_i, i = 1, ..., N$ as

$$\mathbf{P}_i \;=\; \widehat{\mathbf{R}}_i^T \widehat{\mathbf{K}}_i^{-1} \widehat{\mathbf{R}}_i, \tag{3.5.12}$$

where the restriction matrix $\widehat{\mathbf{R}}_i$ maps the global vector $(\mathbf{y}, \mathbf{p}) \in \mathbb{R}^{(n+n)}$ to a local vector $(\mathbf{y}_i, \mathbf{p}_i) \in \mathbb{R}^{(n_i+n_i)}$ for the extended subdomain $\widehat{\Omega}_i$:

$$\widehat{\mathbf{R}}_i = \begin{pmatrix} \widehat{\mathbf{R}}_i^y & \\ & \widehat{\mathbf{R}}_i^p \end{pmatrix}. \tag{3.5.13}$$

The transpose of this matrix represents the zero-extension map of a local vector to a full global vector. The matrix $\widehat{\mathbf{K}}_i$ is the local reduced KKT system matrix

$$\widehat{\mathbf{K}}_i = \begin{pmatrix} \widehat{\mathbf{A}}_i & -\alpha^{-1}\widehat{\mathbf{M}}_i \\ \widehat{\mathbf{M}}_i & \widehat{\mathbf{A}}_i^T \end{pmatrix} \tag{3.5.14}$$

of a local optimal control problem defined on the extended subdomain $\widehat{\Omega}_i$:

$$\begin{array}{ll} \min\limits_{y_i \in V_i, u_i \in U_i} & \frac{1}{2}\int_{\widehat{\Omega}_i} (y_i(x))^2\,dx + \frac{\alpha}{2}\int_{\widehat{\Omega}_i} (u_i(x))^2\,dx - \int_{\widehat{\Omega}_i} \hat{y}(x) y_i(x)\,dx \\ \text{s.t.} & \begin{cases} -\Delta y_i(x) &= \hat{f}(x) + u_i(x) \qquad \text{in } \widehat{\Omega}_i, \\ y_i(x) &= 0 \qquad\qquad\qquad\quad \text{on } \partial\widehat{\Omega}_i, \end{cases} \end{array} \tag{3.5.15}$$

with $V_i = H_0^1(\widehat{\Omega}_i)$, $U_i = L^2(\widehat{\Omega}_i)$. The functions $\hat{y}$ and $\hat{f}$ represent some right hand side data for the KKT system. The equivalent algebraic formulation is

$$\min_{\mathbf{y}_i,\mathbf{u}_i} \quad \tfrac{1}{2}\mathbf{y}_i^T\widehat{\mathbf{M}}_i\mathbf{y}_i + \frac{\alpha}{2}\mathbf{u}_i^T\widehat{\mathbf{Q}}_i\mathbf{u}_i - \hat{\mathbf{c}}^T\mathbf{y}_i \tag{3.5.16a}$$

$$\text{s.t.} \quad \widehat{\mathbf{A}}_i\mathbf{y}_i + \widehat{\mathbf{B}}_i\mathbf{u}_i = \hat{\mathbf{b}}. \tag{3.5.16b}$$

for some data $\hat{\mathbf{c}}$ and $\hat{\mathbf{b}}$.

The one-level additive preconditioner is built from the components $\mathbf{P}_i, i = 1, ..., N$, by simply adding them together:

$$\begin{aligned} \mathbf{P}_{A1} &= \mathbf{P}_1 + \mathbf{P}_2 + ... + \mathbf{P}_N \\ &= \sum_{i=1}^{N} \widehat{\mathbf{R}}_i^T\widehat{\mathbf{K}}_i^{-1}\widehat{\mathbf{R}}_i. \end{aligned} \tag{3.5.17}$$

Applying the preconditioner means transforming the original problem (3.5.10) into the equivalent problem

$$\mathbf{P}_{A1}\mathbf{K}\mathbf{z} = \mathbf{P}_{A1}\mathbf{g}, \tag{3.5.18}$$

which is better conditioned. It is possible to define other preconditioners by combining the basic components $\mathbf{P}_i$ in a different way. In a multiplicative method, the components are applied one at a time in substeps, with the residual recomputed for the next substep by using the latest available solution estimate.

For the two-level method, we define an additional component of the preconditioner corresponding to the coarse grid correction:

$$\mathbf{P}_0 = \widehat{\mathbf{R}}_0^T\widehat{\mathbf{K}}_0^{-1}\widehat{\mathbf{R}}_0, \tag{3.5.19}$$

where the matrix $\widehat{\mathbf{K}}_0$ is the reduced KKT optimality system matrix of an optimal control problem defined on the coarse grid

$$\widehat{\mathbf{K}}_0 = \begin{pmatrix} \widehat{\mathbf{A}}_0 & -\alpha^{-1}\widehat{\mathbf{M}}_0 \\ \widehat{\mathbf{M}}_0 & \widehat{\mathbf{A}}_0^T \end{pmatrix}. \tag{3.5.20}$$

The matrix $\widehat{\mathbf{R}}_0$ represents a linear interpolation operator between the coarse and fine grid spaces. Let

$$V^h = \text{span } \{\phi_j; j = 1, ..., n\}, \tag{3.5.21}$$

$$V^H = \text{span } \{\psi_k; k = 1, ..., n_0\}, \tag{3.5.22}$$

then since the coarse grid is nested in the fine grid, we have $V^H \subset V^h$. Therefore, the coarse grid basis functions can be expressed as a linear combinations of the fine grid basis functions

$$\psi_k = \sum_j (\widehat{\mathbf{R}}_0^y)_{kj} \phi_j, \tag{3.5.23}$$

where $\widehat{\mathbf{R}}_0^y \in \mathbb{R}^{n_0 \times n}$ interpolates a fine grid state variable vector to a coarse grid vector. The adjoint variables have the same interpolation matrix $\widehat{\mathbf{R}}_0^p = \widehat{\mathbf{R}}_0^y$, so

$$\widehat{\mathbf{R}}_0 = \begin{pmatrix} \widehat{\mathbf{R}}_0^y & \\ & \widehat{\mathbf{R}}_0^p \end{pmatrix}. \tag{3.5.24}$$

There are a few possible implementations of a two-level preconditioner, depending on how we choose the order and interaction between the coarse and fine grid solvers. These choices are analogous to the different overlapping methods for PDEs: additive, multiplicative or some hybrid [SBG96, Sec. 2.3]. Our two-level choice is the additive method, with preconditioner

$$\begin{aligned} \mathbf{P}_{A2} &= \mathbf{P}_0 + \mathbf{P}_1 + ... + \mathbf{P}_N \\ &= \widehat{\mathbf{R}}_0^T \widehat{\mathbf{K}}_0^{-1} \widehat{\mathbf{R}}_0 + \sum_{i=1}^{N} \widehat{\mathbf{R}}_i^T \widehat{\mathbf{K}}_i^{-1} \widehat{\mathbf{R}}_i. \end{aligned} \tag{3.5.25}$$

The original problem (3.5.10) is now transformed into

$$\mathbf{P}_{A2} \mathbf{K} \mathbf{z} = \mathbf{P}_{A2} \mathbf{g}. \tag{3.5.26}$$

Comparing these algebraic equations to the variational formulations of Sections 3.2 and 3.3, we note that the original variational problem (3.2.6)

$$K(z^h, w^h) = g(w^h) \qquad \forall w^h \in Z^h$$

corresponds to the algebraic system $\mathbf{Kz} = \mathbf{g}$. If we interpret the equivalent transformed problem (3.3.9)

$$Tz^h = g_T$$

as a preconditioned version of (3.2.6) (i.e., as having algebraic representation (3.5.26)), then the two-level operator $T = T_0 + T_1 + ... + T_N$ would correspond to the preconditioned system matrix $\mathbf{P}_{A2}\mathbf{K}$. Also, since $\mathbf{Kz} = \mathbf{g}$ is transformed into

$$(\mathbf{P}_0\mathbf{K} + \mathbf{P}_1\mathbf{K} + ... + \mathbf{P}_N\mathbf{K})\mathbf{z} = (\mathbf{P}_0 + ... + \mathbf{P}_N)\mathbf{g}, \tag{3.5.27}$$

the local projection operators $T_i, i = 0, ..., N$, may be considered to have the matrix representation $\mathbf{P}_i\mathbf{K}$.

Considering the KKT system (3.5.9), we note that it is possible to permute the block rows of the system to yield the equivalent symmetric problem

$$\begin{pmatrix} \mathbf{M} & \mathbf{A}^T \\ \mathbf{A} & -\alpha^{-1}\mathbf{M} \end{pmatrix} \begin{pmatrix} \mathbf{y} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} \mathbf{c} \\ \mathbf{b} \end{pmatrix}. \tag{3.5.28}$$

The system (3.5.28) can be represented in terms of the first KKT system (3.5.10) as

$$\mathbf{\Pi}\,\mathbf{K}\,\mathbf{z} = \mathbf{\Pi}\,\mathbf{g}, \qquad \mathbf{\Pi}\,\mathbf{K} = \begin{pmatrix} \mathbf{M} & \mathbf{A}^T \\ \mathbf{A} & -\alpha^{-1}\mathbf{M} \end{pmatrix}, \qquad \mathbf{\Pi}\,\mathbf{g} = \begin{pmatrix} \mathbf{c} \\ \mathbf{b} \end{pmatrix},$$

where $\mathbf{\Pi} \in \mathbb{R}^{(n+n)\times(n+n)}$ is the permutation matrix that interchanges the first n and last n components of a vector. Similarly, we may define local permutation matrices $\mathbf{\Pi}_i \in \mathbb{R}^{(n_i+n_i)\times(n_i+n_i)}$, $i = 0, ..., N$, that interchanges the first n_i and last n_i components of a local vector. The permuted version of the local KKT matrix (3.5.14) is

$$\mathbf{\Pi}_i\widehat{\mathbf{K}}_i = \begin{pmatrix} \widehat{\mathbf{M}}_i & \widehat{\mathbf{A}}_i^T \\ \widehat{\mathbf{A}}_i & -\alpha^{-1}\widehat{\mathbf{M}}_i \end{pmatrix}. \tag{3.5.29}$$

We note that the permutation and restriction matrices have following properties:

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

which allow the preconditioned system matrix $\mathbf{P}_{A2}\mathbf{K}$ to be written as

$$\begin{aligned}\mathbf{P}_{A2}\mathbf{K} &= \sum_{i=0}^{N} \widehat{\mathbf{R}}_i^T \widehat{\mathbf{K}}_i^{-1} \widehat{\mathbf{R}}_i \, \mathbf{K} \\ &= \sum_{i=0}^{N} \widehat{\mathbf{R}}_i^T \widehat{\mathbf{K}}_i^{-1} \mathbf{\Pi}_i \widehat{\mathbf{R}}_i \mathbf{\Pi} \, \mathbf{K} \\ &= \sum_{i=0}^{N} \widehat{\mathbf{R}}_i^T (\mathbf{\Pi}_i \widehat{\mathbf{K}}_i)^{-1} \widehat{\mathbf{R}}_i \mathbf{\Pi} \, \mathbf{K}.\end{aligned}$$

Instead of solving $\mathbf{P}_{A2}\mathbf{K}\mathbf{z} = \mathbf{P}_{A2}\mathbf{g}$, we may solve the equivalent permuted system

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

A GMRES method yields the same result for either system, and our convergence theory developed in Section 3.4 applies to both cases.

We may compare these various formulations in terms of how the error can be reduced in a simple iterative method. Let $\mathbf{z}^*$ denote the true solution ($\mathbf{K}\mathbf{z}^* = \mathbf{g}$), and let $\mathbf{z}^{(k)}$ denote the approximate solution at iteration k. The next iterate is updated by applying a preconditioner $\mathbf{P}$ approximating $\mathbf{K}^{-1}$ (e.g., $\mathbf{P}_{A1}$ or $\mathbf{P}_{A2}$) to the residual vector to generate a correction step:

$$\mathbf{z}^{(k+1)} = \mathbf{z}^{(k)} + \mathbf{P}(\mathbf{g} - \mathbf{K}\mathbf{z}^{(k)}) \tag{3.5.30}$$

(this may be considered a preconditioned Richardson method.) The error $\mathbf{e}^{(k)} = \mathbf{z}^* - \mathbf{z}^{(k)}$ is updated to

$$\begin{aligned}\mathbf{e}^{(k+1)} &= \mathbf{z}^* - \mathbf{z}^{(k+1)} \\ &= \mathbf{z}^* - \mathbf{z}^{(k)} - \mathbf{P}(\mathbf{g} - \mathbf{K}\mathbf{z}^{(k)}) \\ &= \mathbf{e}^{(k)} - \mathbf{P}(\mathbf{K}\mathbf{z}^* - \mathbf{K}\mathbf{z}^{(k)}) \\ &= \mathbf{e}^{(k)} - \mathbf{P}\mathbf{K}\mathbf{e}^{(k)} \\ &= (\mathbf{I} - \mathbf{P}\mathbf{K})\mathbf{e}^{(k)},\end{aligned}$$

so we may view $\mathbf{I} - \mathbf{PK}$ as an error propagation matrix for this iterative method. In the original problem formulation without preconditioning, this matrix would be $\mathbf{I} - \mathbf{K}$. Thus, the transformed problem has the potential for improved convergence, for example when $\mathbf{P}$ is a good approximation for $\mathbf{K}^{-1}$. We note that if $\rho(\mathbf{I} - \mathbf{PK}) < 1$, then the iterative method (3.5.30) would converge to $\mathbf{z}^*$. In practice, we use a preconditioned Krylov subspace method to solve the linear system rather than directly applying the steps (3.5.30).

3.6 Numerical Results

The overlapping methods are tested on the model control problem

$$\min_{y,u} \quad \frac{1}{2}\int_\Omega (y(x) - y_d(x))^2\,dx + \frac{\alpha}{2}\int_\Omega u^2(x)\,dx \tag{3.6.1}$$

subject to

$$\begin{aligned} -\Delta y(x) &= f(x) + u(x) \qquad \text{in } \Omega, \\ y(x) &= 0 \qquad\qquad\quad \text{on } \partial\Omega. \end{aligned} \tag{3.6.2}$$

The code was implemented in Matlab, using the PDE Toolbox for finite element assembly. The domain $\Omega = (-1,1) \times (-1,1)$ is partitioned into N equal-sized square non-overlapping subdomains $\Omega_i, i = 1, ..., N$ in a checkerboard pattern (left plot of Figure 3.2). The side length of each subdomain is denoted by H. The domain is triangulated with a conforming regular mesh, with each element having width h. Each subdomain Ω_i is then extended to a larger overlapping region $\widehat{\Omega}_i$ by a uniform width around the perimeter. The amount of overlap is an integer multiple of the mesh parameter h, so that the extension does not cut through any fine mesh element. Also, any part of the extension that is outside of the original domain Ω is cut off.

We use the non-overlapping subdomain partitioning of the one-level method to help define a coarse mesh. Each original square subdomain is divided into two triangles so that piecewise linear elements may be used (right plot of Figure 3.2).

The desired state is $y_d(x) = \sin(\pi x_1)\sin(\pi x_2)$ and the right hand side data is $f(x) = 1$. Other given data yield similar trends for the iteration counts as those reported below. For each case, we use a zero initial guess and a stopping tolerance of 10^{-4} times the initial residual norm (or right hand side norm.) The following reported iterations are for the outer loop, while using exact solves for the subdomain problems. We use full (unrestarted) GMRES in all cases. Since the convergence theory was developed for the A-norm, we modified the Matlab version of GMRES to use this norm and the $A(\cdot,\cdot)$ inner product. Our numerical experiments show that the convergence trends are similar if the standard L^2 norm is used, even though the iteration counts are different.

Table 3.1 summarizes the GMRES iteration counts for both overlapping preconditioners using a small fixed overlap of h (fine mesh element size) and regularization parameters $\alpha = 1, 10^{-4}$ and 10^{-8}. Tables 3.3 and 3.5 provide the same views, but using overlap of $2h$ and $4h$ respectively.

We also consider the behavior of the two methods when the overlap is allowed to stay proportional to H (subdomain size.) Table 3.7 summarizes the results for an overlap that is kept at $H/4$, while Table 3.8 shows the iterations for an overlap of $H/2$. The results seem to agree with the theory that for small enough H the iteration counts can be bounded independently of H and h.

To reduce information overload, we view the results for the three cases of $\alpha = 1, \alpha = 10^{-4}$ and $\alpha = 10^{-8}$ as representing typical results for "large" α, "medium" α and "small" α, respectively. Other in-between values of α yield iteration counts that are in line with expectations for intermediate results. We note that within these tables, moving along a diagonal (top left to bottom right) corresponds to viewing results for a fixed H/h ratio.

For comparison, we solve a PDE problem with the overlapping methods, using the state equation (3.6.2) with $u = 0$. We again use GMRES and the same stopping criteria on the A-norm. The results for various overlap amounts are listed in tables

immediately following the corresponding results for the control problem.

We observe the following general convergence results for the one-level method:

- for large α and fixed δ, the iterations grow as $O(1/H)$;
- for small α and fixed δ, the iterations are nearly independent of H, h and H/h;
- increasing δ while fixing other parameters (α, H, h) improves the convergence;
- for δ proportional to H, convergence is nearly independent of h and H/h.

For the two-level method, we observe that:

- convergence is relatively insensitive to α (cf. one-level);
- increasing δ has relatively small effect;
- for fixed δ and all α, h, the convergence is good for small H;
- for δ proportional to H, convergence is nearly independent of H and h for small H.

Table 3.1: GMRES iterations, overlap = h. Left: One-level. Right: Two-level.

(a) $\alpha = 1$

GMRES Iterations, $\boldsymbol{A}$-norm						
One-level, overlap=h, $\alpha = 1.0$						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	7	10	12	16	21	31
1/2		12	17	24	33	49
1/4			21	30	43	64
1/8				41	57	85
1/16					83	116

GMRES Iterations, $\boldsymbol{A}$-norm						
Two-level, overlap=h, $\alpha = 1.0$						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	8	11	12	15	19	25
1/2		12	14	15	19	25
1/4			13	14	16	20
1/8				14	14	16
1/16					14	14

(b) $\alpha = 10^{-4}$

GMRES Iterations, $\boldsymbol{A}$-norm						
One-level, overlap=h, $\alpha = 10^{-4}$						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	6	8	10	13	17	23
1/2		12	12	18	25	35
1/4			14	22	32	43
1/8				31	42	63
1/16					62	85

GMRES Iterations, $\boldsymbol{A}$-norm						
Two-level, overlap=h, $\alpha = 10^{-4}$						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	7	12	12	16	21	26
1/2		15	16	18	23	32
1/4			16	18	21	26
1/8				18	17	19
1/16					16	16

(c) $\alpha = 10^{-8}$

GMRES Iterations, $\boldsymbol{A}$-norm						
One-level, overlap=h, $\alpha = 10^{-8}$						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	7	10	9	9	9	9
1/2		15	14	12	12	14
1/4			14	14	13	14
1/8				15	13	14
1/16					15	15

GMRES Iterations, $\boldsymbol{A}$-norm						
Two-level, overlap=h, $\alpha = 10^{-8}$						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	8	13	14	14	15	15
1/2		18	20	19	20	18
1/4			16	18	18	19
1/8				19	20	21
1/16					20	21

Table 3.2: PDE only, GMRES iterations, overlap = h. Left: One-level. Right: Two-level.

One-level, overlap=h, PDE only						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	3	6	8	10	13	18
1/2		8	12	15	20	28
1/4			14	18	26	38
1/8				25	35	57
1/16					50	74

Two-level, overlap=h, PDE only						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	4	8	10	12	15	20
1/2		11	13	14	18	24
1/4			13	14	16	20
1/8				14	15	20
1/16					16	17

Table 3.3: GMRES iterations, overlap = $2h$. Left: One-level. Right: Two-level.

(a) $\alpha = 1$

GMRES Iterations, $\boldsymbol{A}$-norm						
One-level, overlap=$2h$, $\alpha = 1.0$						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	1	9	10	13	17	23
1/2		8	13	17	25	35
1/4			14	22	31	46
1/8				26	41	59
1/16					51	83

GMRES Iterations, $\boldsymbol{A}$-norm						
Two-level, overlap=$2h$, $\alpha = 1.0$						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	2	10	11	13	16	20
1/2		9	13	14	16	20
1/4			12	15	15	16
1/8				15	15	15
1/16					17	15

(b) $\alpha = 10^{-4}$

GMRES Iterations, $\boldsymbol{A}$-norm						
One-level, overlap=$2h$, $\alpha = 10^{-4}$						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	2	7	8	10	13	18
1/2		8	10	12	17	25
1/4			10	15	22	32
1/8				18	29	43
1/16					37	62

GMRES Iterations, $\boldsymbol{A}$-norm						
Two-level, overlap=$2h$, $\alpha = 10^{-4}$						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	2	11	11	13	17	23
1/2		10	14	16	18	24
1/4			10	16	18	22
1/8				15	18	18
1/16					22	17

(c) $\alpha = 10^{-8}$

GMRES Iterations, $\boldsymbol{A}$-norm						
One-level, overlap=$2h$, $\alpha = 10^{-8}$						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	3	9	9	9	8	8
1/2		11	11	10	10	10
1/4			10	11	10	11
1/8				11	11	11
1/16					12	11

GMRES Iterations, $\boldsymbol{A}$-norm						
Two-level, overlap=$2h$, $\alpha = 10^{-8}$						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	5	12	13	14	13	13
1/2		14	15	16	16	16
1/4			12	14	16	16
1/8				13	16	17
1/16					13	17

Table 3.4: PDE only, GMRES iterations, overlap = $2h$. Left: One-level. Right: Two-level.

One-level, overlap=$2h$, PDE only						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	1	6	7	9	11	14
1/2		7	10	12	15	22
1/4			9	15	19	28
1/8				16	25	37
1/16					31	55

Two-level, overlap=$2h$, PDE only						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	2	8	10	11	13	17
1/2		8	13	14	15	20
1/4			9	14	14	16
1/8				13	15	18
1/16					19	18

Table 3.5: GMRES iterations, overlap = $4h$. Left: One-level. Right: Two-level.

(a) $\alpha = 1$

GMRES Iterations, $\boldsymbol{A}$-norm						
One-level, overlap=$4h$, $\alpha = 1.0$						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	1	1	9	11	13	17
1/2		6	9	13	18	26
1/4			10	14	23	32
1/8				16	27	43
1/16					31	53

GMRES Iterations, $\boldsymbol{A}$-norm						
Two-level, overlap=$4h$, $\alpha = 1.0$						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	2	2	11	12	14	17
1/2		6	10	14	16	17
1/4			11	12	15	16
1/8				15	17	15
1/16					22	18

(b) $\alpha = 10^{-4}$

GMRES Iterations, $\boldsymbol{A}$-norm						
One-level, overlap=$4h$, $\alpha = 10^{-4}$						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	2	1	7	8	10	14
1/2		6	7	10	12	18
1/4			10	10	15	22
1/8				13	18	31
1/16					22	38

GMRES Iterations, $\boldsymbol{A}$-norm						
Two-level, overlap=$4h$, $\alpha = 10^{-4}$						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	2	2	11	12	14	17
1/2		8	10	15	17	19
1/4			11	11	17	19
1/8				13	16	19
1/16					21	24

(c) $\alpha = 10^{-8}$

GMRES Iterations, $\boldsymbol{A}$-norm						
One-level, overlap=$4h$, $\alpha = 10^{-8}$						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	3	3	9	8	7	7
1/2		7	8	10	9	8
1/4			11	9	10	8
1/8				11	9	10
1/16					11	9

GMRES Iterations, $\boldsymbol{A}$-norm						
Two-level, overlap=$4h$, $\alpha = 10^{-8}$						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	5	5	11	12	13	12
1/2		8	10	14	15	15
1/4			12	11	14	16
1/8				13	12	15
1/16					14	13

Table 3.6: PDE only, GMRES iterations, overlap = $4h$. Left: One-level. Right: Two-level.

One-level, overlap=$4h$, PDE only						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	1	1	6	7	9	11
1/2		5	8	10	13	16
1/4			9	10	16	20
1/8				11	17	26
1/16					19	35

Two-level, overlap=$4h$, PDE only						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	2	2	9	10	11	14
1/2		6	8	13	14	16
1/4			9	10	15	16
1/8				11	14	16
1/16					17	20

Table 3.7: GMRES iterations, overlap = $H/4$. Left: One-level, Right: Two-level.

(a) $\alpha = 1$

GMRES Iterations, $\boldsymbol{A}$-norm						
One-level, overlap=$\boldsymbol{H/4}$, $\boldsymbol{\alpha = 1.0}$						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1		10	10	11	11	12
1/2			17	17	18	19
1/4				30	31	32
1/8					57	59
1/16						116

GMRES Iterations, $\boldsymbol{A}$-norm						
Two-level, overlap=$\boldsymbol{H/4}$, $\boldsymbol{\alpha = 1.0}$						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1		11	11	12	13	13
1/2			14	14	16	16
1/4				14	15	16
1/8					14	15
1/16						14

(b) $\alpha = 10^{-4}$

GMRES Iterations, $\boldsymbol{A}$-norm						
One-level, overlap=$\boldsymbol{H/4}$, $\boldsymbol{\alpha = 10^{-4}}$						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1		8	8	8	9	9
1/2			12	12	12	13
1/4				22	22	22
1/8					42	43
1/16						85

GMRES Iterations, $\boldsymbol{A}$-norm						
Two-level, overlap=$\boldsymbol{H/4}$, $\boldsymbol{\alpha = 10^{-4}}$						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1		12	11	12	12	13
1/2			16	16	17	17
1/4				18	18	19
1/8					17	18
1/16						16

(c) $\alpha = 10^{-8}$

GMRES Iterations, $\boldsymbol{A}$-norm						
One-level, overlap=$\boldsymbol{H/4}$, $\boldsymbol{\alpha = 10^{-8}}$						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1		10	9	8	8	7
1/2			14	10	9	8
1/4				14	10	8
1/8					13	11
1/16						15

GMRES Iterations, $\boldsymbol{A}$-norm						
Two-level, overlap=$\boldsymbol{H/4}$, $\boldsymbol{\alpha = 10^{-8}}$						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1		13	13	12	12	12
1/2			20	16	15	15
1/4				18	16	16
1/8					20	17
1/16						21

Table 3.8: GMRES iterations, overlap = $H/2$. Left: One-level. Right: Two-level.

(a) $\alpha = 1$

GMRES Iterations, $\boldsymbol{A}$-norm						
One-level, overlap=$\boldsymbol{H/2}$, $\boldsymbol{\alpha = 1.0}$						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	7	9	9	9	10	8
1/2		12	13	13	14	15
1/4			21	22	23	24
1/8				41	41	43
1/16					83	83

GMRES Iterations, $\boldsymbol{A}$-norm						
Two-level, overlap=$\boldsymbol{H/2}$, $\boldsymbol{\alpha = 1.0}$						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	8	10	11	11	12	13
1/2		12	13	14	15	16
1/4			13	15	15	16
1/8				14	15	15
1/16					14	15

(b) $\alpha = 10^{-4}$

GMRES Iterations, $\boldsymbol{A}$-norm						
One-level, overlap=$\boldsymbol{H/2}$, $\boldsymbol{\alpha = 10^{-4}}$						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	6	7	7	7	8	9
1/2		12	10	10	11	12
1/4			14	15	15	16
1/8				31	29	31
1/16					62	62

GMRES Iterations, $\boldsymbol{A}$-norm						
Two-level, overlap=$\boldsymbol{H/2}$, $\boldsymbol{\alpha = 10^{-4}}$						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	7	11	11	11	11	12
1/2		15	14	15	16	17
1/4			16	16	17	18
1/8				18	18	19
1/16					16	17

(c) $\alpha = 10^{-8}$

GMRES Iterations, $\boldsymbol{A}$-norm						
One-level, overlap=$\boldsymbol{H/2}$, $\boldsymbol{\alpha = 10^{-8}}$						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	7	9	9	8	8	8
1/2		15	11	10	10	9
1/4			14	11	10	10
1/8				15	11	10
1/16					15	11

GMRES Iterations, $\boldsymbol{A}$-norm						
Two-level, overlap=$\boldsymbol{H/2}$, $\boldsymbol{\alpha = 10^{-8}}$						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	8	12	11	11	11	12
1/2		18	15	14	14	14
1/4			16	14	14	15
1/8				19	16	15
1/16					20	17

Chapter 4

Neumann-Neumann Methods

We now consider the nonoverlapping approach to combining domain decomposition with elliptic optimal control. In the engineering field of structural analysis, the nonoverlapping methods for PDEs are called substructuring methods, which are implemented as direct algorithms. The iterative versions of these direct methods have been developed more recently, and are known as iterative substructuring or Schur complement methods. As noted in [SBG96, p.101], the iterative versions may be attractive as the more efficient alternatives that preserve the existing software data structures and concepts. We will focus on two of the best known methods for solving Schur complement systems, the one-level Neumann-Neumann and two-level Balancing Neumann-Neumann methods, and adapt them to the optimal control problem.

In the last decade, the Neumann-Neumann methods and the overlapping Schwarz methods are considered among the most efficient and well-tested domain decomposition algorithms for solving large scale PDE problems (see [PW02].) The question of whether or not to overlap does not seem to have a definite answer in the domain decomposition literature. The DD survey paper by Chan and Mathew [CM94] suggested that overlapping methods are generally easier to understand and are often more robust, but require extra work on the overlapped regions. They have trouble in problems with discontinuous coefficients, while the nonoverlapping methods can

handle large jumps in coefficients across subdomains.

The fundamental relationship between overlapping and nonoverlapping methods is discussed in [BW89] and [CG92], where it is shown that any overlapping method can be viewed as equivalent to a Schur complement method that uses a particular interface preconditioner. This preconditioner is implicitly applied by the overlapping method through extra computation on the overlapped region.

4.1 Schur Complement for Optimality System

We consider a non-overlapping domain decomposition approach to solving the model problem

$$\begin{array}{ll} \min\limits_{y,u} & \frac{1}{2}\int_\Omega (y(x)-y_d(x))^2\,dx + \frac{\alpha}{2}\int_\Omega (u(x))^2\,dx \\ \text{s.t.} & \begin{cases} -\Delta y(x) + y(x) &= f(x)+u(x) \qquad \text{in } \Omega, \\ y(x) &= 0 \qquad \text{on } \partial\Omega. \end{cases} \end{array} \tag{4.1.1}$$

We define the spaces $V \stackrel{\text{def}}{=} H_0^1(\Omega)$, $U \stackrel{\text{def}}{=} L^2(\Omega)$, and write problem (4.1.1) in weak formulation as

$$\min_{y\in V, u\in U} \quad \tfrac{1}{2}m(y,y) - c(y) + \tfrac{1}{2}q(u,u) \tag{4.1.2a}$$

$$\text{s.t.} \quad a(y,\phi) + b(u,\phi) = f(\phi) \quad \forall \phi \in V, \tag{4.1.2b}$$

with

$$m : V\times V \to \mathbb{R} \qquad m(y,p) = \int_\Omega y(x)p(x)\,dx, \tag{4.1.3a}$$

$$a : V\times V \to \mathbb{R} \qquad a(y,p) = \int_\Omega (\nabla y(x)\cdot\nabla p(x) + y(x)p(x))\,dx, \tag{4.1.3b}$$

$$q : U\times U \to \mathbb{R} \qquad q(u,v) = \alpha\int_\Omega u(x)v(x)\,dx, \tag{4.1.3c}$$

$$b : U\times V \to \mathbb{R} \qquad b(u,p) = -\int_\Omega u(x)p(x)\,dx, \tag{4.1.3d}$$

$$c : V \to \mathbb{R} \qquad c(y) = \int_\Omega y(x)y_d(x)\,dx, \tag{4.1.3e}$$

$$f : V \to \mathbb{R} \qquad f(y) = \int_\Omega f(x)y(x)\,dx. \tag{4.1.3f}$$

The optimality conditions are

$$\begin{aligned} a(\phi,p)+m(y,\phi) &= m(y_d,\phi) && \forall \phi \in V, && (4.1.4a)\\ q(u,\mu)+b(\mu,p) &= 0 && \forall \mu \in U, && (4.1.4b)\\ a(y,\varphi)+b(u,\varphi) &= f(\varphi) && \forall \varphi \in V. && (4.1.4c) \end{aligned}$$

To simplify the presentation, we consider the two-subdomain case first, which will be generalized to many subdomains later. We partition the open domain Ω into two disjoint open subdomains Ω_1 and Ω_2, such that $\overline{\Omega_1} \cup \overline{\Omega_2} = \overline{\Omega}$ and $\Omega_1 \cap \Omega_2 = \emptyset$. The interface boundary between the subdomains is denoted by Γ, so that $\Gamma = \overline{\Omega_1} \cap \overline{\Omega_2}$. The unit outward normal on the boundary of Ω is denoted n, while n_i is used to denote the unit outward normal on Γ for subdomain Ω_i (see Figure 1.2).

For each subdomain $\Omega_i, i = 1, 2$, we define the associated function spaces

$$\begin{aligned} V_i &= \{v \in H^1(\Omega_i) : v(x) = 0 \text{ on } \partial\Omega_i \cap \partial\Omega\},\\ U_i &= L^2(\Omega_i),\\ V_i^0 &= \{v \in V_i : v(x) = 0 \text{ on } \Gamma\} \,=\, H_0^1(\Omega_i), \end{aligned}$$

and also define local linear and bilinear forms on these subdomain spaces:

$$\begin{aligned} &a_i : V_i \times V_i \to \mathbb{R} && a_i(y,\psi) = \textstyle\int_{\Omega_i} (\nabla y(x) \cdot \nabla \psi(x) + y(x)\,\psi(x))\,dx,\\ &b_i : U_i \times V_i \to \mathbb{R} && b_i(u,\psi) = -\textstyle\int_{\Omega_i} u(x)\,\psi(x)\,dx,\\ &m_i : V_i \times V_i \to \mathbb{R} && m_i(y,\psi) = \textstyle\int_{\Omega_i} y(x)\,\psi(x)\,dx,\\ &q_i : U_i \times U_i \to \mathbb{R} && q_i(u,\mu) = \alpha \textstyle\int_{\Omega_i} u(x)\,\mu(x)\,dx,\\ &f_i : V_i \to \mathbb{R} && f_i(\psi) = \textstyle\int_{\Omega_i} f(x)\,\psi(x)\,dx. \end{aligned}$$

Let $\mathrm{Tr}_\Gamma\, v$ denote the trace of v on the common interface Γ, and define a function space on Γ as

$$V_\Gamma \;=\; \{v_\Gamma \in H^{1/2}(\Gamma) : v_\Gamma = \mathrm{Tr}_\Gamma(v) \text{ for some } v \in V\} \;=\; H_{00}^{1/2}(\Gamma).$$

For each subdomain $i = 1, 2$, we define continuous linear extension operators from the interface to the subdomain interior

$$\mathcal{R}_i^s, \mathcal{R}_i^a : V_\Gamma \times V_\Gamma \to V_i \qquad (4.1.5a)$$

such that

$$\begin{aligned} \mathrm{Tr}_\Gamma(\mathcal{R}_i^s(v_\Gamma, q_\Gamma)) &= v_\Gamma \qquad \forall v_\Gamma, q_\Gamma \in V_\Gamma, && \text{(4.1.5b)} \\ \mathrm{Tr}_\Gamma(\mathcal{R}_i^a(v_\Gamma, q_\Gamma)) &= q_\Gamma \qquad \forall v_\Gamma, q_\Gamma \in V_\Gamma. && \text{(4.1.5c)} \end{aligned}$$

For convenience, we will also use the short-hand notation

$$\begin{aligned} \mathcal{R}_i^s v_\Gamma &= \mathcal{R}_i^s(v_\Gamma, 0) && \text{(4.1.5d)} \\ \mathcal{R}_i^a q_\Gamma &= \mathcal{R}_i^a(0, q_\Gamma) && \text{(4.1.5e)} \end{aligned}$$

when we fix one argument to zero.

We now consider the following optimality conditions for the two-subdomain formulation of problem (4.1.2). For $i = 1, 2$,

$$\begin{aligned} a_i(\phi, p_i) + m_i(y_i, \phi) &= m_i(y_d, \phi) \qquad \forall \phi \in V_i^0, && \text{(4.1.6a)} \\ b_i(\mu, p_i) + q_i(u_i, \mu) &= 0 \qquad \forall \mu \in U_i, && \text{(4.1.6b)} \\ a_i(y_i, \psi) + b_i(u_i, \psi) &= f_i(\psi) \qquad \forall \psi \in V_i^0, && \text{(4.1.6c)} \\ y_i &= y_\Gamma \qquad \text{on } \Gamma, && \text{(4.1.6d)} \\ p_i &= p_\Gamma \qquad \text{on } \Gamma, && \text{(4.1.6e)} \end{aligned}$$

and

$$\begin{aligned} &\sum_{i=1}^{2} a_i(y_i, \mathcal{R}_i^a(v_\Gamma, q_\Gamma)) + b_i(u_i, \mathcal{R}_i^a(v_\Gamma, q_\Gamma)) + a_i(\mathcal{R}_i^s(v_\Gamma, q_\Gamma), p_i) + m_i(y_i, \mathcal{R}_i^s(v_\Gamma, q_\Gamma)) \\ = &\sum_{i=1}^{2} f_i(\mathcal{R}_i^a(v_\Gamma, q_\Gamma)) + m_i(y_d, \mathcal{R}_i^s(v_\Gamma, q_\Gamma)) \qquad \forall v_\Gamma, q_\Gamma \in V_\Gamma. \qquad \text{(4.1.7)} \end{aligned}$$

The optimality system (4.1.6) is the weak form of the PDE system

$$\begin{aligned}
-\Delta y_i(x) + y_i(x) &= f(x) + u_i(x) && \text{in } \Omega_i, \\
y_i(x) &= 0 && \text{on } \partial\Omega \cap \partial\Omega_i, \\
y_i(x) &= y_\Gamma(x) && \text{on } \Gamma, \\
-\Delta p_i(x) + p_i(x) &= -(y_i(x) - y_d(x)) && \text{in } \Omega_i, \\
p_i(x) &= 0 && \text{on } \partial\Omega \cap \partial\Omega_i, \\
p_i(x) &= p_\Gamma(x) && \text{on } \Gamma, \\
\alpha u_i(x) - p_i(x) &= 0 && \text{in } \Omega_i,
\end{aligned} \tag{4.1.8}$$

and the interface condition (4.1.7) is the weak form of the transmission conditions

$$\begin{aligned}
\tfrac{\partial}{\partial n_1} y_1(x) &= -\tfrac{\partial}{\partial n_2} y_2(x) && \text{on } \Gamma, \\
\tfrac{\partial}{\partial n_1} p_1(x) &= -\tfrac{\partial}{\partial n_2} p_2(x) && \text{on } \Gamma,
\end{aligned} \tag{4.1.9}$$

which enforce the continuity of the normal derivative of the state and adjoint variables across the interface. The following Lemma states the equivalence between the one-domain and two-subdomain optimality conditions.

Lemma 4.1.1 *Solving for the one-domain optimality conditions* (4.1.4) *is equivalent to finding* $(y_\Gamma, p_\Gamma) \in V_\Gamma \times V_\Gamma$ *such that the subdomain solutions* $(y_i, u_i, p_i) \in V_i \times U_i \times V_i$ *given by* (4.1.6) *also satisfy the interface conditions* (4.1.7)*.*

Proof: The proof is analogous to that of [QV99, Lemma 1.2.1] for a PDE problem. (i) Suppose that $(y, u, p) \in V \times U \times V$ solves (4.1.4). Then setting $y_i = y|_{\Omega_i}, u_i = u|_{\Omega_i}, p_i = p|_{\Omega_i}, i = 1, 2,\ y_\Gamma = \mathrm{Tr}_\Gamma\, y,\ p_\Gamma = \mathrm{Tr}_\Gamma\, p$ would satisfy (4.1.6). For each $(v_\Gamma, q_\Gamma) \in V_\Gamma^2$, the two functions $\mathcal{R}^s(v_\Gamma, q_\Gamma)$ and $\mathcal{R}^a(v_\Gamma, q_\Gamma)$ defined as

$$\mathcal{R}^a(v_\Gamma, q_\Gamma) \stackrel{\text{def}}{=} \begin{cases} \mathcal{R}_1^a(v_\Gamma, q_\Gamma) & \text{in } \Omega_1 \\ \mathcal{R}_2^a(v_\Gamma, q_\Gamma) & \text{in } \Omega_2 \end{cases}$$

$$\mathcal{R}^s(v_\Gamma, q_\Gamma) \stackrel{\text{def}}{=} \begin{cases} \mathcal{R}_1^s(v_\Gamma, q_\Gamma) & \text{in } \Omega_1 \\ \mathcal{R}_2^s(v_\Gamma, q_\Gamma) & \text{in } \Omega_2 \end{cases}$$

both belong to V, so the state and adjoint equations (4.1.4a,c) are satisfied:

$$\begin{aligned} a(\mathcal{R}^a(v_\Gamma, q_\Gamma), p) + m(y, \mathcal{R}^a(v_\Gamma, q_\Gamma)) &= m(y_d, \mathcal{R}^a(v_\Gamma, q_\Gamma)), \\ a(y, \mathcal{R}^s(v_\Gamma, q_\Gamma)) + b(u, \mathcal{R}^s(v_\Gamma, q_\Gamma)) &= f(\mathcal{R}^a(v_\Gamma, q_\Gamma)), \end{aligned}$$

which together imply the transmission condition (4.1.7).

(ii) Suppose that $(y_i, u_i, p_i) \in V_i \times U_i \times V_i$ satisfies (4.1.6) and (4.1.7). Let $(y, u, p) \in V \times U \times V$ be defined by $y|_{\Omega_i} = y_i, u|_{\Omega_i} = u_i, p|_{\Omega_i} = p_i$, $i = 1, 2$. Then the continuity conditions (4.1.6d,e) imply that $y \in V$ and $p \in V$. For any $(\phi, \psi) \in V \times V$, let $(\phi_\Gamma, \psi_\Gamma) = (\mathrm{Tr}_\Gamma\, \phi, \mathrm{Tr}_\Gamma\, \psi)$. Then $(\phi|_{\Omega_i} - \mathcal{R}_i^s(\phi_\Gamma, \psi_\Gamma)) \in V_i^0$ and $(\psi|_{\Omega_i} - \mathcal{R}_i^a(\phi_\Gamma, \psi_\Gamma)) \in V_i^0$. We can then satisfy the global adjoint equation

$$\begin{aligned} a(\phi, p) + m(y, \phi) &= \sum_{i=1}^{2} [a_i(\phi|_{\Omega_i} - \mathcal{R}_i^s(\phi_\Gamma, \psi_\Gamma), p_i) + a_i(\mathcal{R}_i^s(\phi_\Gamma, \psi_\Gamma), p_i)] \\ &\quad + \sum_{i=1}^{2} [m_i(y_i, \phi|_{\Omega_i} - \mathcal{R}_i^s(\phi_\Gamma, \psi_\Gamma)) + m_i(y_i, \mathcal{R}_i^s(\phi_\Gamma, \psi_\Gamma))] \\ &= \sum_{i=1}^{2} [m_i(y_d, \phi|_{\Omega_i} - \mathcal{R}_i^s(\phi_\Gamma, \psi_\Gamma)) + m_i(y_d, \mathcal{R}_i^s(\phi_\Gamma, \psi_\Gamma))] \\ &= m(y_d, \phi), \end{aligned}$$

and global state equation

$$\begin{aligned} a(y, \psi) + b(u, \psi) &= \sum_{i=1}^{2} [a(y_i, \psi_\Gamma - \mathcal{R}_i^a(\phi_\Gamma, \psi_\Gamma)) + a(y_i, \mathcal{R}_i^a(\phi_\Gamma, \psi_\Gamma))] \\ &\quad + \sum_{i=1}^{2} [b(u_i, \psi_\Gamma - \mathcal{R}_i^a(\phi_\Gamma, \psi_\Gamma)) + b(u_i, \mathcal{R}_i^a(\phi_\Gamma, \psi_\Gamma)] \\ &= \sum_{i=1}^{2} [f(\psi_\Gamma - \mathcal{R}_i^a(\phi_\Gamma, \psi_\Gamma)) + f(\mathcal{R}_i^a(\phi_\Gamma, \psi_\Gamma))] \\ &= f(\psi). \end{aligned}$$

The global gradient equation can be treated similarly. For any $\mu \in U$, let $\mu_i = \mu|_{\Omega_i}$. Then

$$b(u, \mu) + q(\mu, p) = \sum_{i=1}^{2} [b_i(u_i, \mu_i) + q_i(\mu_i, p_i)] = 0.$$

□

Our goal is to find a formulation of two local optimal control problems that would correspond to the local optimality conditions given by (4.1.6) for $i = 1, 2$. Considering the following state equation over subdomain Ω_i,

$$\begin{aligned} -\Delta y_i(x) + y_i(x) &= f(x) + u_i(x) && \text{in } \Omega_i, \\ y_i(x) &= 0 && \text{on } \partial\Omega_i \cap \partial\Omega, \end{aligned}$$

we note that the normal derivative of y on the interface can be expressed in weak form as

$$\begin{aligned} \int_\Gamma \frac{\partial}{\partial n_i} y_i(x)\psi(x)\,dx &= \int_{\Omega_i} (\nabla y_i(x)\nabla\psi(x) + y_i(x)\psi(x))\,dx - \int_{\Omega_i} u(x)\psi(x)\,dx \\ &\quad - \int_{\Omega_i} f(x)\psi(x)\,dx \qquad \forall \psi \in V_i \\ &= a_i(y_i, \psi) + b_i(u_i, \psi) - f_i(\psi) \qquad \forall \psi \in V_i. \end{aligned} \tag{4.1.10}$$

We consider the following local problem

$$\begin{aligned} \min_{y_i, u_i} \quad & \left(\tfrac{1}{2}\int_{\Omega_i} (y_i(x) - y_d(x))^2\,dx + \tfrac{\alpha}{2}\int_{\Omega_i} (u_i(x))^2\,dx + \int_\Gamma \tfrac{\partial}{\partial n_i} y_i(x)\, p_\Gamma(x)\,dx \right) \\ \text{s.t.} \quad & \begin{cases} -\Delta y_i(x) + y_i(x) = f(x) + u_i(x) & \text{in } \Omega_i, \\ y_i(x) = 0 & \text{on } \partial\Omega_i \cap \partial\Omega, \\ y_i(x) = y_\Gamma(x) & \text{on } \Gamma. \end{cases} \end{aligned} \tag{4.1.11}$$

Using (4.1.10), we set the last term in the objective function as

$$\int_\Gamma \frac{\partial}{\partial n_i} y_i(x)\, p_\Gamma(x)\,dx = a_i(y_i, \mathcal{R}_i^a p_\Gamma) + b_i(u_i, \mathcal{R}_i^a p_\Gamma) - f_i(\mathcal{R}_i^a p_\Gamma),$$

so problem (4.1.11) has the weak formulation

$$\begin{aligned} \min_{y_i, u_i} \quad & \tfrac{1}{2} m_i(y_i - y_d, y_i - y_d) + \tfrac{1}{2} q_i(u_i, u_i) \\ & + a_i(y_i, \mathcal{R}_i^a p_\Gamma) + b_i(u_i, \mathcal{R}_i^a p_\Gamma) - f_i(\mathcal{R}_i^a p_\Gamma) \qquad (4.1.12a) \\ \text{s.t.} \quad & \begin{cases} a_i(y_i, \psi) + b_i(u_i, \psi) = f_i(\psi) & \forall \psi \in V_i^0 \\ y_i = y_\Gamma & \text{on } \Gamma \end{cases} \qquad (4.1.12b) \end{aligned}$$

Lemma 4.1.2 *The local optimal control problem* (4.1.12) *has necessary and sufficient optimality conditions given by the system* (4.1.6).

Proof: Problem (4.1.12) is a convex linear-quadratic control problem, so it has $y_i \in V_i$, $u_i \in U_i$ as a solution if and only if the state equation (4.1.12b) is satisfied, and

$$m_i(y_i - y_d, z_i) + q_i(u_i, v_i) + a_i(z_i, \mathcal{R}_i^a p_\Gamma) + b_i(v_i, \mathcal{R}_i^a p_\Gamma) = 0 \tag{4.1.13}$$

for all $z_i \in V_i^0$, $v_i \in U_i$ that satisfy the homogeneous linear state equation

$$a_i(z_i, \psi) + b_i(v_i, \psi) \;=\; 0 \qquad \forall \psi \in V_i^0. \tag{4.1.14}$$

(i) Suppose that $y_i \in V_i$, $u_i \in U_i$ is a solution of (4.1.12). Let p_i solve (4.1.6a), with $p_i = p_\Gamma$ on Γ, and define $p_i^0 = p_i - \mathcal{R}_i^a p_\Gamma \in V_i^0$. We use equation (4.1.14) with $\psi = p_i^0$,

$$a_i(z_i, p_i^0) + b_i(v_i, p_i^0) = 0, \tag{4.1.15}$$

and combine it with (4.1.13) to get

$$\begin{aligned} 0 &= m_i(y_i - y_d, z_i) + q_i(u_i, v_i) + a_i(z_i, \mathcal{R}_i^a p_\Gamma) + b_i(v_i, \mathcal{R}_i^a p_\Gamma) + a_i(z_i, p_i^0) + b_i(v_i, p_i^0) \\ &= m_i(y_i - y_d, z_i) + q_i(u_i, v_i) + a_i(z_i, p_i) + b_i(v_i, p_i) \\ &= q_i(u_i, v_i) + b_i(v_i, p_i) \end{aligned}$$

with $v_i \in U_i$. Therefore, condition (4.1.6) is satisfied.

(ii) Suppose that $y_i, p_i \in V_i$, $u_i \in U_i$ solve the optimality conditions (4.1.6). We define $p_i^0 = p_i - \mathcal{R}_i^a p_\Gamma \in V_i^0$, and let $z_i \in V_i^0, v_i \in U_i$ satisfy the homogeneous state equation

$$a_i(z_i, \psi) + b_i(v_i, \psi) = 0 \qquad \forall \psi \in V_i^0. \tag{4.1.16}$$

We use condition (4.1.6a) with $\phi = z_i$

$$a_i(z_i, p_i) + m_i(y_i - y_d, z_i) = 0, \tag{4.1.17}$$

and use condition (4.1.6b) with $\mu = v_i$

$$b_i(v_i, p_i) + q_i(u_i, v_i) = 0. \tag{4.1.18}$$

These are combined with (4.1.14) using $\psi = p_i^0$ to get

$$\begin{aligned} 0 &= m_i(y_i - y_d, z_i) + q_i(u_i, v_i) + a_i(z_i, p_i) + b_i(v_i, p_i) - a_i(z_i, p_i^0) - b_i(v_i, p_i^0) \\ &= m_i(y_i - y_d, z_i) + q_i(u_i, v_i) + a_i(z_i, \mathcal{R}_i^a p_\Gamma) + b_i(v_i, \mathcal{R}_i^a p_\Gamma). \end{aligned}$$

Therefore, (4.1.13) is satisfied and $y_i \in V_i$, $u_i \in U_i$ solve the control problem (4.1.12) □

We can formulate a Schur complement system by viewing the interface condition (4.1.7) as a linear equation in (y_Γ, p_Γ). To derive a formulation, we write the solution (y_i, u_i, p_i) of the local optimality system (4.1.6) as the sum of two components

$$(y_i, u_i, p_i) = (\tilde{y}_i, \tilde{u}_i, \tilde{p}_i) + (\hat{y}_i, \hat{u}_i, \hat{p}_i), \tag{4.1.19}$$

where $(\tilde{y}_i, \tilde{u}_i, \tilde{p}_i)$ solves

$$\begin{aligned} a_i(\psi, \tilde{p}_i) + m_i(\tilde{y}_i, \psi) &= 0 \qquad \forall \psi \in V_i^0, \\ b_i(\mu, \tilde{p}_i) + q_i(\tilde{u}_i, \mu) &= 0 \qquad \forall \mu \in U_i, \\ a_i(\tilde{y}_i, \psi) + b_i(\tilde{u}_i, \psi) &= 0 \qquad \forall \psi \in V_i^0, \\ \tilde{y}_i &= y_\Gamma \qquad \text{on } \Gamma, \\ \tilde{p}_i &= p_\Gamma \qquad \text{on } \Gamma, \end{aligned} \tag{4.1.20}$$

and $(\hat{y}_i, \hat{u}_i, \hat{p}_i)$ solves

$$\begin{aligned} a_i(\psi, \hat{p}_i) + m_i(\hat{y}_i, \psi) &= m_i(y_d, \psi) && \forall \psi \in V_i^0, \\ b_i(\mu, \hat{p}_i) + q_i(\hat{u}_i, \mu) &= 0 && \forall \mu \in U_i, \\ a_i(\hat{y}_i, \psi) + b_i(\hat{u}_i, \psi) &= f_i(\psi) && \forall \psi \in V_i^0, \\ \hat{y}_i &= 0 && \text{on } \Gamma, \\ \hat{p}_i &= 0 && \text{on } \Gamma. \end{aligned} \tag{4.1.21}$$

The interface condition (4.1.7) can be written in terms of these components

$$\begin{aligned} &\sum_{i=1}^{2} [\, a_i(\tilde{y}_i + \hat{y}_i, \mathcal{R}_i^a(v_\Gamma, q_\Gamma)) + b_i(\tilde{u}_i + \hat{u}_i, \mathcal{R}_i^a(v_\Gamma, q_\Gamma)) + a_i(\mathcal{R}_i^s(v_\Gamma, q_\Gamma), \tilde{p}_i + \hat{p}_i) \\ &\qquad + m_i(\tilde{y}_i + \hat{y}_i, \mathcal{R}_i^s(v_\Gamma, q_\Gamma))\,] \\ &= \sum_{i=1}^{2} [\, f_i(\mathcal{R}_i^a(v_\Gamma, q_\Gamma)) + m_i(y_d, \mathcal{R}_i^s(v_\Gamma, q_\Gamma))\,] \qquad \forall v_\Gamma, q_\Gamma \in V_\Gamma, \end{aligned} \tag{4.1.22}$$

which may be rearranged as

$$\begin{aligned} &\sum_{i=1}^{2} [\, a_i(\tilde{y}_i, \mathcal{R}_i^a(v_\Gamma, q_\Gamma)) + b_i(\tilde{u}_i, \mathcal{R}_i^a(v_\Gamma, q_\Gamma)) + a_i(\mathcal{R}_i^s(v_\Gamma, q_\Gamma), \tilde{p}_i) + m_i(\tilde{y}_i, \mathcal{R}_i^s(v_\Gamma, q_\Gamma))\,] \\ &= \sum_{i=1}^{2} [\, f_i(\mathcal{R}_i^a(v_\Gamma, q_\Gamma)) + m_i(y_d, \mathcal{R}_i^s(v_\Gamma, q_\Gamma))\,] - \sum_{i=1}^{2} [\, a_i(\hat{y}_i, \mathcal{R}_i^a(v_\Gamma, q_\Gamma)) + b_i(\hat{u}_i, \mathcal{R}_i^a(v_\Gamma, q_\Gamma)) \\ &\qquad + a_i(\mathcal{R}_i^s(v_\Gamma, q_\Gamma), \hat{p}_i) + m_i(\hat{y}_i, \mathcal{R}_i^s(v_\Gamma, q_\Gamma))\,] \qquad \forall v_\Gamma, q_\Gamma \in V_\Gamma. \end{aligned} \tag{4.1.23}$$

We note that $\tilde{y}_i, \tilde{u}_i, \tilde{p}_i$ and the left hand side of (4.1.23) are linear in (y_Γ, p_Γ), while $\hat{y}_i, \hat{u}_i, \hat{p}_i$ and the right hand side are independent of (y_Γ, p_Γ). We express the components $(\tilde{y}, \tilde{u}, \tilde{p})$ in terms of the interface variable (y_Γ, p_Γ) by defining the linear operator

$$\mathcal{H}_i : V_\Gamma \times V_\Gamma \to V_i \times U_i \times V_i \tag{4.1.24}$$

as

$$\mathcal{H}_i(y_\Gamma, p_\Gamma) = \begin{pmatrix} \mathcal{H}_i^y(y_\Gamma, p_\Gamma) \\ \mathcal{H}_i^u(y_\Gamma, p_\Gamma) \\ \mathcal{H}_i^p(y_\Gamma, p_\Gamma) \end{pmatrix} = \begin{pmatrix} \tilde{y}_i \\ \tilde{u}_i \\ \tilde{p}_i \end{pmatrix}. \tag{4.1.25}$$

This $\mathcal{H}_i$ operator is a generalization of the harmonic extension from domain decomposition methods for linear elliptic PDEs.

Let V_Γ' denote the dual space of V_Γ. For $i = 1, 2$, we define the local Schur complement operator

$$\mathcal{S}_i : (V_\Gamma)^2 \to (V_\Gamma')^2 \tag{4.1.26}$$

by

$$\begin{aligned} \langle\!\langle \mathcal{S}_i(y_\Gamma, p_\Gamma), (v_\Gamma, q_\Gamma) \rangle\!\rangle &= a_i(\mathcal{H}_i^y(y_\Gamma, p_\Gamma), \mathcal{R}_i^a(v_\Gamma, q_\Gamma)) + b_i(\mathcal{H}_i^u(y_\Gamma, p_\Gamma), \mathcal{R}_i^a(v_\Gamma, q_\Gamma)) \\ &\quad + a_i(\mathcal{R}_i^s(v_\Gamma, q_\Gamma), \mathcal{H}_i^p(y_\Gamma, p_\Gamma)) + m_i(\mathcal{H}_i^y(y_\Gamma, p_\Gamma), \mathcal{R}_i^s(v_\Gamma, q_\Gamma)), \end{aligned} \tag{4.1.27}$$

with $\langle\!\langle \cdot, \cdot \rangle\!\rangle$ denoting the duality pairing between $(V_\Gamma')^2$ and V_Γ^2. For the right hand side of the Schur complement system, we define $g_i \in (V_\Gamma')^2, i = 1, 2$, as

$$\begin{aligned} \langle\!\langle g_i, (v_\Gamma, q_\Gamma) \rangle\!\rangle &= f_i(\mathcal{R}_i^a(v_\Gamma, q_\Gamma)) + m_i(y_d, \mathcal{R}_i^s(v_\Gamma, q_\Gamma)) \\ &\quad - a_i(\hat{y}_i, \mathcal{R}_i^a(v_\Gamma, q_\Gamma)) - b_i(\hat{u}_i, \mathcal{R}_i^a(v_\Gamma, q_\Gamma)) \\ &\quad - a_i(\mathcal{R}_i^s(v_\Gamma, q_\Gamma), \hat{p}_i) - m_i(\hat{y}_i, \mathcal{R}_i^s(v_\Gamma, q_\Gamma)). \end{aligned} \tag{4.1.28}$$

The global Schur complement system is

$$(\mathcal{S}_1 + \mathcal{S}_2)(y_\Gamma, p_\Gamma) = g_1 + g_2 \qquad \text{in } (V_\Gamma')^2, \tag{4.1.29}$$

or

$$\mathcal{S}(y_\Gamma, p_\Gamma) = g, \tag{4.1.30}$$

with $\mathcal{S} = \mathcal{S}_1 + \mathcal{S}_2$ and $g = g_1 + g_2$. By Lemma 4.1.1, the Schur complement system (4.1.30) is equivalent to the optimality conditions (4.1.4).

In domain decomposition for PDEs, the action of a local Schur complement operator may be viewed as a linear mapping of Dirichlet data to Neumann data (weak form) on the common interface. For our control problem, the Schur complement operator $\mathcal{S}_i$ maps the Dirichlet data for both state and adjoint variables to Neumann data. Instead of solving a PDE problem to compute the mapping, we now solve the control problem (4.1.12) when evaluating $\mathcal{S}_i(y_\Gamma, p_\Gamma)$.

Conversely, the action of an inverse local Schur complement operator may be viewed as a mapping of Neumann data to Dirichlet data. This inverse does not exist in the PDE case if the underlying differential operator is singular. For example, the Laplace operator on a subdomain with pure Neumann data on its boundary does not have an inverse. This difficulty can also occur in control problems. However, just as for PDE problems, an inverse Schur operator is used only as part of a preconditioner, which means it does not need to be computed accurately. A common strategy for approximating this inverse is to modify slightly the underlying singular differential operator to make it positive definite.

For our two-subdomain control problem with Dirichlet data on the outside boundary, the inverse Schur problem is well defined.

Lemma 4.1.3 *Let $g_i = (g_i^y, g_i^p) \in (V_\Gamma')^2$. The unique solution $(y_\Gamma, p_\Gamma) \in V_\Gamma^2$ of*

$$\mathcal{S}_i(y_\Gamma, p_\Gamma) = g_i \qquad \text{in } (V_\Gamma')^2 \tag{4.1.31}$$

is given by $(y_\Gamma, p_\Gamma) = (\mathrm{Tr}_\Gamma\, y_i, \mathrm{Tr}_\Gamma\, p_i)$ where (y_i, u_i, p_i) is the unique solution of

$$\begin{aligned} a_i(\psi, p_i) + m_i(y_i, \psi) &= \langle g_i^p, \mathrm{Tr}_\Gamma\, \psi\rangle \qquad \forall \psi \in V_i & \text{(4.1.32a)}\\ b_i(\mu, p_i) + q_i(u_i, \mu) &= 0 \qquad \forall \mu \in U_i & \text{(4.1.32b)}\\ a_i(y_i, \psi) + b_i(u_i, \psi) &= \langle g_i^y, \mathrm{Tr}_\Gamma\, \psi\rangle \qquad \forall \psi \in V_i & \text{(4.1.32c)} \end{aligned}$$

where $\langle\cdot,\cdot\rangle$ denotes the duality pairing between V_Γ' and V_Γ.

Proof: Using the definition (4.1.27) of $\mathcal{S}_i$, equation (4.1.31) becomes

$$\begin{aligned} &a_i(\mathcal{H}_i^y(y_\Gamma, p_\Gamma), \mathcal{R}_i^a(v_\Gamma, q_\Gamma)) + b_i(\mathcal{H}_i^u(y_\Gamma, p_\Gamma), \mathcal{R}_i^a(v_\Gamma, q_\Gamma))\\ &+a_i(\mathcal{R}_i^s(v_\Gamma, q_\Gamma), \mathcal{H}_i^p(y_\Gamma, p_\Gamma)) + m_i(\mathcal{H}_i^y(y_\Gamma, p_\Gamma), \mathcal{R}_i^s(v_\Gamma, q_\Gamma))\\ &\qquad = \langle g_i^y, v_\Gamma\rangle + \langle g_i^p, q_\Gamma\rangle \qquad \forall v_\Gamma, q_\Gamma \in V_\Gamma. \end{aligned} \tag{4.1.33}$$

By using the definition (4.1.25) of $\mathcal{H}_i(y_\Gamma, p_\Gamma)$ in (4.1.33), the equation (4.1.31) is

equivalent to the system

$$\begin{aligned}
&a_i(y_i, \mathcal{R}_i^a(v_\Gamma, q_\Gamma)) + b_i(u_i, p_\Gamma), \mathcal{R}_i^a(v_\Gamma, q_\Gamma)) \\
&\quad + a_i(\mathcal{R}_i^s(v_\Gamma, q_\Gamma), p_i) + m_i(y_i, \mathcal{R}_i^s(v_\Gamma, q_\Gamma)) &&= \langle g_i^y, v_\Gamma\rangle + \langle g_i^p, q_\Gamma\rangle && \forall v_\Gamma, q_\Gamma \in V_\Gamma, \\
&a_i(\psi^0, p_i) + m_i(y_i, \psi^0) &&= 0 && \forall \psi^0 \in V_i^0, \\
&b_i(\mu, p_i) + q_i(u_i, \mu) &&= 0 && \forall \mu \in U_i, \\
&a_i(y_i, \phi^0) + b_i(u_i, \phi^0) &&= 0 && \forall \phi^0 \in V_i^0, \\
&y_i &&= y_\Gamma && \text{on } \Gamma, \\
&p_i &&= p_\Gamma && \text{on } \Gamma.
\end{aligned} \tag{4.1.34}$$

If we set $\psi = \psi^0 + \mathcal{R}_i^s v_\Gamma \in V_i$ and $\phi = \phi^0 + \mathcal{R}_i^a v_\Gamma \in V_i$, then (4.1.34) is equivalent to

$$\begin{aligned}
a_i(\psi, p_i) + m_i(y_i, \psi) &= \langle g_i^y, \operatorname{Tr}_\Gamma \psi\rangle && \forall \psi \in V_i, \\
b_i(\mu, p_i) + q_i(u_i, \mu) &= 0 && \forall \mu \in U_i, \\
a_i(y_i, \phi) + b_i(u_i, \phi) &= \langle g_i^p, \operatorname{Tr}_\Gamma \phi\rangle && \forall \phi \in V_i, \\
y_i &= y_\Gamma && \text{on } \Gamma, \\
p_i &= p_\Gamma && \text{on } \Gamma.
\end{aligned} \tag{4.1.35}$$

The proof will be complete once we prove that (4.1.32) has a unique solution $(y_i, u_i, p_i) \in V_i \times U_i \times V_i$. Suppose that $(y_i^1, u_i^1, p_i^1), (y_i^2, u_i^2, p_i^2) \in V_i \times U_i \times V_i$ are both solutions of (4.1.32). Then $(e_i^y, e_i^u, e_i^p) = (y_i^1 - y_i^2, u_i^1 - u_i^2, p_i^1 - p_i^2) \in V_i \times U_i \times V_i$ satisfies

$$a_i(\psi, e_i^p) + m_i(e_i^y, \psi) = 0 \qquad \forall \psi \in V_i, \tag{4.1.36a}$$

$$b_i(\mu, e_i^p) + q_i(e_i^u, \mu) = 0 \qquad \forall \mu \in U_i, \tag{4.1.36b}$$

$$a_i(e_i^y, \phi) + b_i(e_i^u, \phi) = 0 \qquad \forall \phi \in V_i. \tag{4.1.36c}$$

If we set $\psi = e_i^y$, $\mu = e_i^u$ and $\phi = -e_i^p$ in (4.1.36) and add the three equations, the result is

$$0 = m_i(e_i^y, e_i^y) + q_i(e_i^u, e_i^u) \geq \alpha \|e_i^u\|_{U_i}^2.$$

Therefore, $e_i^u = 0$, which implies $e_i^y = 0$ when combined with (4.1.36c). Using equation (4.1.36a), we then have $e_i^p = 0$. □

The system (4.1.32) can be interpreted as the weak form of

$$\begin{aligned}
-\Delta y_i(x) + y_i(x) &= u_i(x) && \text{in } \Omega_i, \\
y_i(x) &= 0 && \text{on } \partial\Omega_i \cap \partial\Omega, \\
\tfrac{\partial}{\partial n_i} y_i(x) &= g_i^y(x) && \text{on } \Gamma, \\
-\Delta p_i(x) + p_i(x) &= -y_i(x) && \text{in } \Omega_i, \\
p_i(x) &= 0 && \text{on } \partial\Omega_i \cap \partial\Omega, \\
\tfrac{\partial}{\partial n_i} p_i(x) &= g_i^p(x) && \text{on } \Gamma, \\
-p_i(x) + \alpha u_i(x) &= 0 && \text{in } \Omega_i.
\end{aligned} \tag{4.1.37}$$

A domain decomposition Schur complement method is an iterative method with preconditioning on a global Schur system, where the preconditioner is constructed from some or all of the local Schur operators. For PDE problems, a well-known domain decomposition preconditioner is the Neumann-Neumann, where all of the local Schur operators are combined to compute an approximate correction step. We extend these ideas for PDEs to our model optimal control problem by solving iteratively the global Schur optimality equation (4.1.29). Informally, we adjust the common interface variables (y_Γ, p_Γ) until the interface condition (4.1.7) is satisfied. We first state the Neumann-Neumann method for optimal control as a simple iterative method, then show that it may be interpreted as a preconditioning method. This is analogous to the iterative formulation for a PDE given in [QV99, Sec 1.3].

Given some initial (y_Γ^0, p_Γ^0), for each iteration $k \geq 0$:

Step 1. Find $(y_i^{k+1}, u_i^{k+1}, p_i^{k+1})$ (in parallel) for each subdomain $i = 1, 2$, such that

$$\begin{aligned}
a_i(\psi, p_i^{k+1}) + m_i(y_i^{k+1}, \psi) &= m_i(y_d, \psi) && \forall \psi \in V_i^0, \\
b_i(\mu, p_i^{k+1}) + q_i(u_i^{k+1}, \mu) &= 0 && \forall \mu \in U_i, \\
a_i(y_i^{k+1}, \psi) + b_i(u_i^{k+1}, \psi) &= f_i(\psi) && \forall \psi \in V_i^0, \\
y_i^{k+1} &= y_\Gamma^k && \text{on } \Gamma, \\
p_i^{k+1} &= p_\Gamma^k && \text{on } \Gamma
\end{aligned}$$

(i.e., solve the local control problems (4.1.11) using the previous iteration's y_Γ and p_Γ).

Step 2. Compute the residual $(r_y^{k+1}, r_p^{k+1}) \in (V_\Gamma')^2$ of the global Schur optimality equation by combining the residuals from the local equations

$$\begin{aligned}(r_y^{k+1}, r_p^{k+1}) &= (\mathcal{S}_1 + \mathcal{S}_2)(y_\Gamma^k, p_\Gamma^k) - (g_1 + g_2) \\ &= \sum_{i=1}^{2} [\, \mathcal{S}_i(y_\Gamma^k, p_\Gamma^k) - g_i \,].\end{aligned}$$

Step 3. Find (in parallel) the local correction steps for (y_Γ^k, p_Γ^k) needed to reduce the residual (r_y^{k+1}, r_p^{k+1}) to zero locally. This means solving for the correction $(\tilde{y}_{i,\Gamma}^{k+1}, \tilde{p}_{i,\Gamma}^{k+1})$ in the local residual equation $\mathcal{S}_i\,(\tilde{y}_{i,\Gamma}^{k+1}, \tilde{p}_{i,\Gamma}^{k+1}) = (r_y^{k+1}, r_p^{k+1})$, first by finding $(\tilde{y}_i^{k+1}, \tilde{u}_i^{k+1}, \tilde{p}_i^{k+1})$ such that

$$\begin{aligned}a_i(\psi, \tilde{p}_i^{k+1}) + m_i(\tilde{y}_i^{k+1}, \psi) &= \langle r_p^{k+1}, \mathrm{Tr}_\Gamma\, \psi \rangle \qquad && \forall \psi \in V_i, \\ b_i(\mu, \tilde{p}_i^{k+1}) + q_i(\tilde{u}_i^{k+1}, \mu) &= 0 && \forall \mu \in U_i, \\ a_i(\tilde{y}_i^{k+1}, \psi) + b_i(\tilde{u}_i^{k+1}, \psi) &= \langle r_y^{k+1}, \mathrm{Tr}_\Gamma\, \psi \rangle && \forall \psi \in V_i,\end{aligned}$$

and then taking the restriction to the interface

$$(\tilde{y}_{i,\Gamma}^{k+1}, \tilde{p}_{i,\Gamma}^{k+1}) = (\mathrm{Tr}_\Gamma\, \tilde{y}_i^{k+1}, \mathrm{Tr}_\Gamma\, \tilde{p}_i^{k+1}).$$

Step 4. Update (y_Γ^k, p_Γ^k) using a global correction step that is a weighted sum of the local corrections

$$(y_\Gamma^{k+1}, p_\Gamma^{k+1}) = (y_\Gamma^k, p_\Gamma^k) - \theta \sum_{i=1}^{2} \sigma_i\, (\tilde{y}_{i,\Gamma}^{k+1}, \tilde{p}_{i,\Gamma}^{k+1}) \tag{4.1.38}$$

for some positive weights σ_i and positive acceleration parameter θ.

We note that solving for the local correction $(\tilde{y}_{i,\Gamma}^{k+1}, \tilde{p}_{i,\Gamma}^{k+1})$ in step 3 above is the inverse operation of the local Schur operator $\mathcal{S}_i$ (Lemma 4.1.3) :

$$\begin{aligned}(\tilde{y}_{i,\Gamma}^{k+1}, \tilde{p}_{i,\Gamma}^{k+1}) &= (\mathcal{S}_i)^{-1}\, r^{k+1} \\ &= (\mathcal{S}_i)^{-1}\, [\, (\mathcal{S}_1 + \mathcal{S}_2)(y_\Gamma^k, p_\Gamma^k) - (g_1 + g_2)\,].\end{aligned} \tag{4.1.39}$$

The update step (4.1.38) can be written as

$$(y_\Gamma^{k+1}, p_\Gamma^{k+1}) \;=\; (y_\Gamma^k, p_\Gamma^k) + \theta\,(\sigma_1\,(\mathcal{S}_1)^{-1} + \sigma_2\,(\mathcal{S}_2)^{-1})\,[\,g - \mathcal{S}\,(y_\Gamma^k, p_\Gamma^k)\,], \tag{4.1.40}$$

which can be interpreted as a Richardson method for solving $\mathcal{S}(y_\Gamma, p_\Gamma) = g$ with the preconditioner

$$\mathcal{P}_{NN} \;=\; \sigma_1\,(\mathcal{S}_1)^{-1} + \sigma_2\,(\mathcal{S}_2)^{-1} \tag{4.1.41}$$

approximating $\mathcal{S}^{-1}$. This is the Neumann-Neumann preconditioner for our optimal control problem and is analogous to the PDE version in that we approximate the inverse of the global operator by a weighted sum of the local inverses

$$\mathcal{S}^{-1} \;=\; (\mathcal{S}_1 + \mathcal{S}_2)^{-1} \;\approx\; \sigma_1\,(\mathcal{S}_1)^{-1} + \sigma_2\,(\mathcal{S}_2)^{-1}. \tag{4.1.42}$$

4.2 Schur Complement for the Discretized Problem

We apply a finite element method for solving problem (4.1.2) by triangulating the domain Ω with a quasi-uniform family of triangulations $\{\mathcal{T}^h\}$, with h denoting the characteristic maximum element diameter in a mesh. Let V^h be the space of continuous piecewise linear functions defined on a triangulation $\mathcal{T}^h$, and with value zero on the outside boundary $\partial\Omega$:

$$V^h = \{v^h : v^h \in C^0(\Omega), v^h|_\tau \text{ linear } \forall \tau \in \mathcal{T}^h, v^h = 0 \text{ on } \partial\Omega\}. \tag{4.2.1}$$

We partition the set of mesh elements in Ω into two subdomains, Ω_1 and Ω_2, and let Γ denote the common interface separating them (see Figure 4.1). For the state and adjoint space over each subdomain $\Omega_i, i = 1, 2$, let

$$V_i^h \;=\; \{v^h|_{\Omega_i} : v^h \in V^h\}, \tag{4.2.2a}$$

$$V_i^{h,0} \;=\; \{v \in V_i^h : v(x) = 0 \text{ on } \Gamma\}, \tag{4.2.2b}$$

$$V_\Gamma^h \;=\; \{v_{h|_\Gamma} : v^h \in V^h\}. \tag{4.2.2c}$$

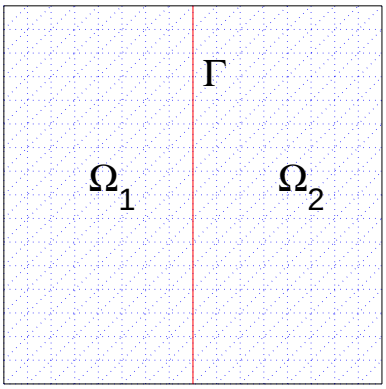

Figure 4.1: A mesh partitioned into two subdomains, with common interface Γ.

For the discretized control, one might consider using a similar approach to decompose V^h by restriction to Ω_1, Ω_2 and Γ. A domain decomposition formulation based on this discretization would introduce "interface controls" defined on a band of width $O(h)$ along Γ (represented by the dotted hat function in the left plot of Figure 4.2). However, since the control u is in $L^2(\Omega)$, it does not make sense to evaluate $u(x)$ on the one-dimensional interface Γ when $\Omega \subset \mathbb{R}^2$. We avoid interface controls by defining the control discretization to be continuous on each subdomain interior, but not on Γ:

$$U_i^h \;=\; \{u^h \in C^0(\Omega_i) : u^h|_{\tau_j} \text{ is linear } \forall \tau_j \in \Omega_i\} \tag{4.2.3}$$

(see the right plot in Figure 4.2). We assume each $U_i^h, i = 1, 2$, has zero extension outside of Ω_i, so that the global control space may be defined as

$$U^h(\Omega) = U_1^h \oplus U_2^h. \tag{4.2.4}$$

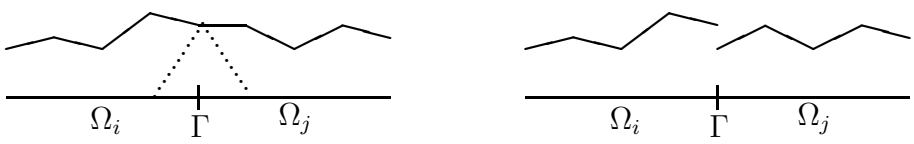

Figure 4.2: Cross-sectional view of the control discretization on the interface

Our discretized problem is

$$\begin{array}{ll} \min\limits_{y^h \in V^h, u^h \in U^h} & \dfrac{1}{2}\displaystyle\int_\Omega (y^h(x) - y_d(x))^2\,dx + \dfrac{\alpha}{2}\int_\Omega (u^h(x))^2\,dx \\ \text{s.t.} & a(y^h, v^h) + b(u^h, v^h) = f(v^h) \quad \forall v^h \in V^h, \end{array} \tag{4.2.5}$$

with optimality conditions

$$a(v^h, p^h) + m(y^h, v^h) = m(y_d, v^h) \qquad \forall v^h \in V^h, \tag{4.2.6a}$$

$$b(u^h, p^h) + q(u^h, w^h) = 0 \qquad \forall w^h \in U^h, \tag{4.2.6b}$$

$$a(y^h, v^h) + b(u^h, v^h) = f(v^h) \qquad \forall v^h \in V^h. \tag{4.2.6c}$$

We also define the discrete extension operators

$$\mathcal{R}_i^{h,s}, \mathcal{R}_i^{h,a} : V_\Gamma^h \times V_\Gamma^h \to V_i^h \tag{4.2.7a}$$

such that

$$\mathrm{Tr}_\Gamma(\mathcal{R}_i^{h,s}(v_\Gamma^h, q_\Gamma^h)) = v_\Gamma^h \qquad \forall v_\Gamma^h, q_\Gamma^h \in V_\Gamma^h, \tag{4.2.7b}$$

$$\mathrm{Tr}_\Gamma(\mathcal{R}_i^{h,a}(v_\Gamma^h, q_\Gamma^h)) = q_\Gamma^h \qquad \forall v_\Gamma^h, q_\Gamma^h \in V_\Gamma^h. \tag{4.2.7c}$$

The superscripts a and s denote the adjoint and state equation, respectively.

The optimality conditions for the two-subdomain problem include the local conditions (for $i = 1, 2$)

$$a_i(v^h, p_i^h) + m_i(y_i^h, v^h) = m_i(y_d, v^h) \qquad \forall v^h \in V_i^{h,0}, \tag{4.2.8a}$$

$$b_i(w^h, p_i^h) + q_i(u_i^h, w^h) = 0 \qquad \forall w^h \in U_i^h, \tag{4.2.8b}$$

$$a_i(y_i^h, v^h) + b_i(u_i^h, v^h) = f_i(v^h) \qquad \forall v^h \in V_i^{h,0}, \tag{4.2.8c}$$

$$y_i^h = y_\Gamma^h \qquad \text{on } \Gamma, \tag{4.2.8d}$$

$$p_i^h = p_\Gamma^h \qquad \text{on } \Gamma, \tag{4.2.8e}$$

and the interface condition

$$\begin{aligned} &\sum_{i=1}^{2} a_i(y_i^h, \mathcal{R}_i^a(v_\Gamma^h, q_\Gamma^h)) + b_i(u_i^h, \mathcal{R}_i^a(v_\Gamma^h, q_\Gamma^h)) + a_i(\mathcal{R}_i^s(v_\Gamma^h, q_\Gamma^h), p_i^h) + m_i(y_i^h, \mathcal{R}_i^s(v_\Gamma^h, q_\Gamma^h)) \\ = &\sum_{i=1}^{2} f_i(\mathcal{R}_i^a(v_\Gamma^h, q_\Gamma^h)) + m_i(y_d, \mathcal{R}_i^s(v_\Gamma^h, q_\Gamma^h)) \qquad \forall v_\Gamma^h, q_\Gamma^h \in V_\Gamma^h. \end{aligned} \tag{4.2.9}$$

We define the discrete harmonic linear extension operators

$$\mathcal{H}_i^h : V_\Gamma^h \times V_\Gamma^h \to V_i^h \times U_i^h \times V_i^h \tag{4.2.10}$$

as

$$\mathcal{H}_i^h(y_\Gamma^h, p_\Gamma^h) = \begin{pmatrix} \mathcal{H}_i^{h,y}(y_\Gamma^h, p_\Gamma^h) \\ \mathcal{H}_i^{h,u}(y_\Gamma^h, p_\Gamma^h) \\ \mathcal{H}_i^{h,p}(y_\Gamma^h, p_\Gamma^h) \end{pmatrix} = \begin{pmatrix} \tilde{y}_i^h \\ \tilde{u}_i^h \\ \tilde{p}_i^h \end{pmatrix} \tag{4.2.11}$$

where $(\tilde{y}_i^h, \tilde{u}_i^h, \tilde{p}_i^h)$ solves

$$\begin{aligned} a_i(v^h, \tilde{p}_i^h) + m_i(\tilde{y}_i^h, v^h) &= 0 \qquad && \forall v^h \in V_i^{h,0}, \\ b_i(w^h, \tilde{p}_i^h) + q_i(\tilde{u}_i^h, w^h) &= 0 \qquad && \forall w^h \in U_i^h, \\ a_i(\tilde{y}_i^h, v^h) + b_i(\tilde{u}_i^h, v^h) &= 0 \qquad && \forall v^h \in V_i^{h,0}, \\ \tilde{y}_i^h &= y_\Gamma^h \qquad && \text{on } \Gamma, \\ \tilde{p}_i^h &= p_\Gamma^h \qquad && \text{on } \Gamma. \end{aligned} \tag{4.2.12}$$

For each subdomain $i = 1, 2$, we define the discrete Schur complement operator

$$\mathcal{S}_i^h : (V_\Gamma^h)^2 \to (V_\Gamma^{h\prime})^2 \tag{4.2.13}$$

by

$$\begin{aligned} &\langle\!\langle \mathcal{S}_i^h(y_\Gamma^h, p_\Gamma^h), (v_\Gamma^h, q_\Gamma^h) \rangle\!\rangle = \\ &\quad a_i(\mathcal{H}_i^{h,y}(y_\Gamma^h, p_\Gamma^h), \mathcal{R}_i^{h,a}(v_\Gamma^h, q_\Gamma^h)) + b_i(\mathcal{H}_i^{h,u}(y_\Gamma^h, p_\Gamma^h), \mathcal{R}_i^{h,a}(v_\Gamma^h, q_\Gamma^h)) \\ &\quad + a_i(\mathcal{R}_i^{h,s}(v_\Gamma^h, q_\Gamma^h), \mathcal{H}_i^{h,p}(y_\Gamma^h, p_\Gamma^h)) + m_i(\mathcal{H}_i^{h,y}(y_\Gamma^h, p_\Gamma^h), \mathcal{R}_i^{h,s}(v_\Gamma^h, q_\Gamma^h)), \end{aligned}$$

with $\langle\!\langle \cdot, \cdot \rangle\!\rangle$ denoting the duality pairing between $(V_\Gamma^{h\prime})^2$ and $(V_\Gamma^h)^2$. For the right hand side of the Schur complement system, we define $g_i^h \in (V_\Gamma^{h\prime})^2$, $i = 1, 2$, as

$$\begin{aligned} \langle\!\langle g_i^h, (v_\Gamma^h, q_\Gamma^h) \rangle\!\rangle &= f_i(\mathcal{R}_i^{h,a}(v_\Gamma^h, q_\Gamma^h)) + m_i(y_d, \mathcal{R}_i^{h,s}(v_\Gamma^h, q_\Gamma^h)) \\ &\quad - a_i(\hat{y}_i, \mathcal{R}_i^{h,a}(v_\Gamma^h, q_\Gamma^h)) - b_i(\hat{u}_i, \mathcal{R}_i^{h,a}(v_\Gamma^h, q_\Gamma^h)) \\ &\quad - a_i(\mathcal{R}_i^{h,s}(v_\Gamma^h, q_\Gamma^h), \hat{p}_i^h) - m_i(\hat{y}_i^h, \mathcal{R}_i^{h,s}(v_\Gamma^h, q_\Gamma^h)). \end{aligned} \tag{4.2.14}$$

The global Schur complement system is

$$(\mathcal{S}_1^h + \mathcal{S}_2^h)(y_\Gamma^h, p_\Gamma^h) = g_1^h + g_2^h \qquad \text{in } (V_\Gamma^{h\prime})^2, \tag{4.2.15}$$

4.3 Algebraic Viewpoint

The Schur complement and Neumann-Neumann methods for optimization may be derived at the algebraic level. We follow the previous convention of using bold face for vectors and matrices. After a finite element discretization, the single-domain optimal control problem (4.1.2) may be written as

$$\begin{array}{ll} \min\limits_{\mathbf{y},\mathbf{u}} & \frac{1}{2}\mathbf{y}^T\mathbf{M}\mathbf{y} + \frac{\alpha}{2}\mathbf{u}^T\mathbf{Q}\mathbf{u} - \mathbf{c}^T\mathbf{y} \\ \text{s.t.} & \mathbf{A}\mathbf{y} + \mathbf{B}\mathbf{u} = \mathbf{b} \end{array} \tag{4.3.1}$$

with optimality condition

$$\begin{pmatrix} \mathbf{M} & \mathbf{0} & \mathbf{A}^T \\ \mathbf{0} & \alpha\mathbf{Q} & \mathbf{B}^T \\ \mathbf{A} & \mathbf{B} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{y} \\ \mathbf{u} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} \mathbf{c} \\ \mathbf{0} \\ \mathbf{b} \end{pmatrix}. \tag{4.3.2}$$

We show how the approach of decomposing the full optimal control problem in the two previous sections may be formulated by decomposing this optimality system. We also derive the Schur complement with respect to this system and the associated Neumann-Neumann preconditioners.

4.3.1 Two Subdomains

Considering the partitioned set of mesh elements in Ω (Figure 4.1), we let $\{\phi^1_{I_j}\}$, $\{\phi^2_{I_j}\}$ and $\{\phi_{\Gamma_j}\}$ denote the sets of nodal basis functions associated with nodes in the interior of Ω_1, interior of Ω_2 and on Γ, respectively. The finite element spaces (4.2.2b,c) may be expressed as

$$V_1^{h,0} = \text{span}\{\phi^1_{I_j}\}, \quad V_2^{h,0} = \text{span}\{\phi^2_{I_j}\}, \quad V_\Gamma^h = \text{span}\{\phi_{\Gamma_j}\},$$

with

$$V^h \;=\; V_1^{h,0} \oplus V_2^{h,0} \oplus V_\Gamma^h. \tag{4.3.3}$$

Each block of the KKT system matrix (4.3.2) can be partitioned according to the location of the nodes within the domain. For example, matrix block $\mathbf{A}$ can be partitioned as

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{II}^1 & & \mathbf{A}_{I\Gamma}^1 \\ & \mathbf{A}_{II}^2 & \mathbf{A}_{I\Gamma}^2 \\ \mathbf{A}_{\Gamma I}^1 & \mathbf{A}_{\Gamma I}^2 & \mathbf{A}_{\Gamma\Gamma}^{1+2} \end{pmatrix}, \tag{4.3.4}$$

where the subscript I denotes interior nodes and Γ denotes interface nodes. The superscript 1 or 2 denotes the subdomain number. Let $x_j, x_k \in \Omega$ be node locations, with associated nodal basis functions ϕ_j, ϕ_k. Then

$$\begin{aligned} (\mathbf{A}_{II}^i)_{j,k} &= \int_{\Omega_i} (\nabla\phi_j \cdot \nabla\phi_k + \phi_j\phi_k)\, dx \qquad x_j, x_k \in \Omega_i, \\ (\mathbf{A}_{I\Gamma}^i)_{j,k} &= \int_{\Omega_i} (\nabla\phi_j \cdot \nabla\phi_k + \phi_j\phi_k)\, dx \qquad x_j \in \Omega_i, x_k \in \Gamma, \\ (\mathbf{A}_{\Gamma\Gamma}^i)_{j,k} &= \int_{\Omega_i} (\nabla\phi_j \cdot \nabla\phi_k + \phi_j\phi_k)\, dx \qquad x_j, x_k \in \Gamma. \end{aligned}$$

We use $\mathbf{A}_{\Gamma\Gamma}^{1+2}$ as an abbreviation for the sum $\mathbf{A}_{\Gamma\Gamma}^1 + \mathbf{A}_{\Gamma\Gamma}^2$, with $\mathbf{A}_{\Gamma\Gamma}^i$ being the contribution from subdomain Ω_i to the area integral for the entries in $\mathbf{A}_{\Gamma\Gamma}$. Similarly, the partitioned block $\mathbf{B}$ has entries

$$\begin{aligned} (\mathbf{B}_{II}^i)_{j,k} &= \int_{\Omega_i} \phi_j\phi_k\, dx \qquad x_j, x_k \in \Omega_i, \\ (\mathbf{B}_{I\Gamma}^i)_{j,k} &= \int_{\Omega_i} \phi_j\phi_k\, dx \qquad x_j \in \Omega_i, x_k \in \Gamma. \end{aligned}$$

Other blocks in the KKT matrix are defined analogously. The vector representing $y \in V^h$ can be partitioned into blocks as

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_I^1 \\ \mathbf{y}_I^2 \\ \mathbf{y}_\Gamma \end{pmatrix}, \tag{4.3.5}$$

with $\mathbf{y}_I^i$ representing $y_i \in V_i^{h,0}$, and $\mathbf{y}_\Gamma$ representing $y_\Gamma \in V_\Gamma^h$.

If each block of the matrix and vectors in the KKT system (4.3.2) is partitioned according to the location of the nodal basis functions, then the system can be expanded

as

$$\left(\begin{array}{ccc|cc|ccc} \mathbf{M}_{II}^1 & & \mathbf{M}_{I\Gamma}^1 & & & \mathbf{A}_{II}^1 & & \mathbf{A}_{I\Gamma}^1 \\ & \mathbf{M}_{II}^2 & \mathbf{M}_{I\Gamma}^2 & & & & \mathbf{A}_{II}^2 & \mathbf{A}_{I\Gamma}^2 \\ \mathbf{M}_{\Gamma I}^1 & \mathbf{M}_{\Gamma I}^2 & \mathbf{M}_{\Gamma\Gamma}^{1+2} & & & \mathbf{A}_{\Gamma I}^1 & \mathbf{A}_{\Gamma I}^2 & \mathbf{A}_{\Gamma\Gamma}^{1+2} \\ \hline & & & \alpha\mathbf{Q}_{II}^1 & & \mathbf{B}_{II}^1 & & \mathbf{B}_{I\Gamma}^1 \\ & & & & \alpha\mathbf{Q}_{II}^2 & & \mathbf{B}_{II}^2 & \mathbf{B}_{I\Gamma}^2 \\ \hline \mathbf{A}_{II}^1 & & \mathbf{A}_{I\Gamma}^1 & \mathbf{B}_{II}^1 & & & & \\ & \mathbf{A}_{II}^2 & \mathbf{A}_{I\Gamma}^2 & & \mathbf{B}_{II}^2 & & & \\ \mathbf{A}_{\Gamma I}^1 & \mathbf{A}_{\Gamma I}^2 & \mathbf{A}_{\Gamma\Gamma}^{1+2} & \mathbf{B}_{\Gamma I}^1 & \mathbf{B}_{\Gamma I}^2 & & & \end{array}\right)\left(\begin{array}{c} \mathbf{y}_I^1 \\ \mathbf{y}_I^2 \\ \mathbf{y}_\Gamma \\ \hline \mathbf{u}_I^1 \\ \mathbf{u}_I^2 \\ \hline \mathbf{p}_I^1 \\ \mathbf{p}_I^2 \\ \mathbf{p}_\Gamma \end{array}\right) = \left(\begin{array}{c} \mathbf{c}_I^1 \\ \mathbf{c}_I^2 \\ \mathbf{c}_\Gamma^{1+2} \\ \hline \mathbf{0} \\ \mathbf{0} \\ \hline \mathbf{b}_I^1 \\ \mathbf{b}_I^2 \\ \mathbf{b}_\Gamma^{1+2} \end{array}\right). \tag{4.3.6}$$

An alternative to grouping these blocks by variable type is to group them by the location in the domain. After performing a symmetric permutation of the rows and columns, the optimality system (4.3.6) above can be written as the following equivalent system

$$\left(\begin{array}{ccc|ccc|cc} \mathbf{M}_{II}^1 & & \mathbf{A}_{II}^1 & & & & \mathbf{M}_{I\Gamma}^1 & \mathbf{A}_{I\Gamma}^1 \\ & \alpha\mathbf{Q}_{II}^1 & \mathbf{B}_{II}^1 & & & & & \mathbf{B}_{I\Gamma}^1 \\ \mathbf{A}_{II}^1 & \mathbf{B}_{II}^1 & & & & & \mathbf{A}_{I\Gamma}^1 & \\ \hline & & & \mathbf{M}_{II}^2 & & \mathbf{A}_{II}^2 & \mathbf{M}_{I\Gamma}^2 & \mathbf{A}_{I\Gamma}^2 \\ & & & & \alpha\mathbf{Q}_{II}^2 & \mathbf{B}_{II}^2 & & \mathbf{B}_{I\Gamma}^2 \\ & & & \mathbf{A}_{II}^2 & \mathbf{B}_{II}^2 & & \mathbf{A}_{I\Gamma}^2 & \\ \hline \mathbf{M}_{\Gamma I}^1 & & \mathbf{A}_{\Gamma I}^1 & \mathbf{M}_{\Gamma I}^2 & & \mathbf{A}_{\Gamma I}^2 & \mathbf{M}_{\Gamma\Gamma}^{1+2} & \mathbf{A}_{\Gamma\Gamma}^{1+2} \\ \mathbf{A}_{\Gamma I}^1 & \mathbf{B}_{\Gamma I}^1 & & \mathbf{A}_{\Gamma I}^2 & \mathbf{B}_{\Gamma I}^2 & & \mathbf{A}_{\Gamma\Gamma}^{1+2} & \end{array}\right)\left(\begin{array}{c} \mathbf{y}_I^1 \\ \mathbf{u}_I^1 \\ \mathbf{p}_I^1 \\ \hline \mathbf{y}_I^2 \\ \mathbf{u}_I^2 \\ \mathbf{p}_I^2 \\ \hline \mathbf{y}_\Gamma \\ \mathbf{p}_\Gamma \end{array}\right) = \left(\begin{array}{c} \mathbf{c}_I^1 \\ \mathbf{0} \\ \mathbf{b}_I^1 \\ \hline \mathbf{c}_I^2 \\ \mathbf{0} \\ \mathbf{b}_I^2 \\ \hline \mathbf{c}_\Gamma^{1+2} \\ \mathbf{b}_\Gamma^{1+2} \end{array}\right). \tag{4.3.7}$$

This permuted optimality system is written in abbreviated form as

$$\left(\begin{array}{ccc} \mathbf{K}_{II}^1 & \mathbf{0} & \mathbf{K}_{I\Gamma}^1 \\ \mathbf{0} & \mathbf{K}_{II}^2 & \mathbf{K}_{I\Gamma}^2 \\ \mathbf{K}_{\Gamma I}^1 & \mathbf{K}_{\Gamma I}^2 & \mathbf{K}_{\Gamma\Gamma}^{1+2} \end{array}\right)\left(\begin{array}{c} \mathbf{z}_I^1 \\ \mathbf{z}_I^2 \\ \mathbf{z}_\Gamma \end{array}\right) = \left(\begin{array}{c} \mathbf{g}_I^1 \\ \mathbf{g}_I^2 \\ \mathbf{g}_\Gamma^{1+2} \end{array}\right), \tag{4.3.8}$$

where $\mathbf{z}$ is used to designate the triplet of state, control and adjoint variables.

The system matrix in (4.3.8) can be factored as

$$\begin{pmatrix} \mathbf{K}_{II}^1 & \mathbf{0} & \mathbf{K}_{I\Gamma}^1 \\ \mathbf{0} & \mathbf{K}_{II}^2 & \mathbf{K}_{I\Gamma}^2 \\ \mathbf{K}_{\Gamma I}^1 & \mathbf{K}_{\Gamma I}^2 & \mathbf{K}_{\Gamma\Gamma}^1 + \mathbf{K}_{\Gamma\Gamma}^2 \end{pmatrix} = \begin{pmatrix} \mathbf{K}_{II}^1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_{II}^2 & \mathbf{0} \\ \mathbf{K}_{\Gamma I}^1 & \mathbf{K}_{\Gamma I}^2 & \mathbf{S}^1 + \mathbf{S}^2 \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{0} & (\mathbf{K}_{II}^1)^{-1}\mathbf{K}_{I\Gamma}^1 \\ \mathbf{0} & \mathbf{I} & (\mathbf{K}_{II}^2)^{-1}\mathbf{K}_{I\Gamma}^2 \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{pmatrix} \tag{4.3.9}$$

where

$$\mathbf{S}^i = \mathbf{K}_{\Gamma\Gamma}^i - \mathbf{K}_{\Gamma I}^i (\mathbf{K}_{II}^i)^{-1} \mathbf{K}_{I\Gamma}^i \tag{4.3.10}$$

is a local Schur complement matrix. The global Schur complement system equivalent to (4.3.8) is formed by summing the local Schur systems

$$\mathbf{S}\,\mathbf{z}_\Gamma \;=\; \mathbf{g} \tag{4.3.11}$$

with

$$\mathbf{S} \;=\; \sum_{i=1}^{2} (\mathbf{K}_{\Gamma\Gamma}^i - \mathbf{K}_{\Gamma I}^i (\mathbf{K}_{II}^i)^{-1} \mathbf{K}_{I\Gamma}^i) \tag{4.3.12a}$$

$$\mathbf{g} \;=\; \sum_{i=1}^{2} (\mathbf{g}_\Gamma^i - \mathbf{K}_{\Gamma I}^i (\mathbf{K}_{II}^i)^{-1} \mathbf{g}_I^i). \tag{4.3.12b}$$

This Schur complement system with respect to the KKT system can be viewed as an alternative expression of the optimality conditions of the discrete control problem (4.3.1).

The Schur complement matrix $\mathbf{S}$ is symmetric indefinite. The inertia of $\mathbf{S}$ is considered in [HN04a], where it is shown that $\mathbf{S}$ has the same number of positive eigenvalues as negative eigenvalues. This number is equal to the number of nodes on the interface Γ (or the length of $\mathbf{y}_\Gamma$). This result also applies to the case of many subdomains, assuming that we do not use controls on the interface.

To solve the system (4.3.11), we apply an iterative method using a type of Neumann-Neumann preconditioner analogous to that from domain decomposition methods for PDEs. The Neumann-Neumann preconditioner for optimal control, $\mathbf{P}_{NN}$, is a weighted sum of the local Schur inverses.

$$\mathbf{S}^{-1} \;\approx\; \mathbf{P}_{NN} \;=\; \sigma_1 (\mathbf{S}^1)^{-1} + \sigma_2 (\mathbf{S}^2)^{-1} \tag{4.3.13}$$

for some positive weights σ_1, σ_2.

To compute the action of a local Schur operator, we apply the definition (4.3.10) to a vector, which involves solving a control problem with Dirichlet data for the states and adjoints on the interface. To apply the inverse $(\mathbf{S}^i)^{-1}$ of a local Schur complement, we solve a control problem with Neumann data on the interface. This can be done without explicitly forming $\mathbf{S}^i$, in a similar manner to the PDE problem:

$$(\mathbf{S}^i)^{-1}\mathbf{g} = (\mathbf{0} \quad \mathbf{I})(\mathbf{K}^i)^{-1} \begin{pmatrix} \mathbf{0} \\ \mathbf{I} \end{pmatrix} \mathbf{g}; \tag{4.3.14}$$

see [SBG96, Sec. 4.2.1].

4.3.2 Many Subdomains

We can generalize the Schur complement method to many subdomains. Suppose that the domain Ω is partitioned into N subdomains, denoted $\Omega_i, i = 1, ..., N$. Let n_y, n_u, n_p denote the total number of state, control and adjoint unknowns, respectively. Similarly, for each subdomain Ω_i, let n_y^i, n_u^i and n_p^i denote the number of state, control and adjoint unknowns associated with the subdomain.

We may use a Boolean matrix $\mathbf{R}_i^y \in \{0,1\}^{n_y^i \times n_y}$ to represent a restriction operator that maps a global vector corresponding to all state variable unknowns $\mathbf{y}$ to a local vector $\mathbf{y}^i$ associated with only subdomain Ω_i. Similarly, let $\mathbf{R}_i^u \in \{0,1\}^{n_u^i \times n_u}$ and $\mathbf{R}_i^p \in \{0,1\}^{n_p^i \times n_p}$ denote the restriction matrices for the control variable $\mathbf{u}$ and adjoint variable $\mathbf{p}$, respectively. The restriction matrix that maps the global $(\mathbf{y}, \mathbf{u}, \mathbf{p})$ vector to a local vector $(\mathbf{y}_i, \mathbf{u}_i, \mathbf{p}_i)$ for subdomain Ω_i is

$$\mathbf{R}_i = \begin{pmatrix} \mathbf{R}_i^y & & \\ & \mathbf{R}_i^u & \\ & & \mathbf{R}_i^p \end{pmatrix}. \tag{4.3.15}$$

The transpose of each of these matrices represents the zero extension map of a local vector to a full global vector. We note that these Boolean matrices are used only as a notational convenience and are not formed explicitly in practice.

Since each finite element node in the domain can be classified as either interior or interface node, the KKT matrix $\mathbf{K}$, solution vector $\mathbf{z}$ and right hand side $\mathbf{g}$ may be partitioned according to the location of the nodes as

$$\begin{pmatrix} \mathbf{K}_{II} & \mathbf{K}_{I\Gamma} \\ \mathbf{K}_{\Gamma I} & \mathbf{K}_{\Gamma\Gamma} \end{pmatrix} \begin{pmatrix} \mathbf{z}_I \\ \mathbf{z}_\Gamma \end{pmatrix} = \begin{pmatrix} \mathbf{g}_I \\ \mathbf{g}_\Gamma \end{pmatrix}. \tag{4.3.16}$$

This global system may be assembled from the local KKT systems for subdomain Ω_i, $i = 1, ..., N$:

$$\begin{pmatrix} \mathbf{K}^i_{II} & \mathbf{K}^i_{I\Gamma} \\ \mathbf{K}^i_{\Gamma I} & \mathbf{K}^i_{\Gamma\Gamma} \end{pmatrix} \begin{pmatrix} \mathbf{z}^i_I \\ \mathbf{z}^i_\Gamma \end{pmatrix} = \begin{pmatrix} \mathbf{g}^i_I \\ \mathbf{g}^i_\Gamma \end{pmatrix}. \tag{4.3.17}$$

The global-local relationship can be expressed as

$$\begin{aligned} \mathbf{K}^i &= \mathbf{R}_i \mathbf{K} \mathbf{R}_i^T \\ \mathbf{z}^i &= \mathbf{R}_i \mathbf{z} \\ \mathbf{g}^i &= \mathbf{R}_i \mathbf{g} \end{aligned}$$

and

$$\begin{aligned} \mathbf{K} &= \sum_{i=1}^{N} \mathbf{R}_i^T \mathbf{K}^i \mathbf{R}_i \\ \mathbf{z} &= \sum_{i=1}^{N} \mathbf{R}_i^T \mathbf{z}^i \\ \mathbf{g} &= \sum_{i=1}^{N} \mathbf{R}_i^T \mathbf{g}^i. \end{aligned}$$

4.3.2.1 One-Level Schur Complement Method

For a Schur complement method operating on the interface space, we define a restriction matrix $\tilde{\mathbf{R}}_i$ similar to $\mathbf{R}_i$, but only involving nodes on the interface. For the state variable, let $\tilde{\mathbf{R}}_i^y$ map a vector of all y values on the interface Γ to a vector of values associated with only Ω_i. Similarly, $\tilde{\mathbf{R}}_i^p$ maps all adjoint values on Γ to those of Ω_i.

The control u does not need to be represented on the interface, so we define

$$\tilde{\mathbf{R}}_i = \begin{pmatrix} \tilde{\mathbf{R}}_i^y & \\ & \tilde{\mathbf{R}}_i^p \end{pmatrix}. \tag{4.3.18}$$

The global Schur complement matrix can be expressed in terms of the local ones as

$$\mathbf{S} \;=\; \sum_{i=1}^{N} \tilde{\mathbf{R}}_i^T \mathbf{S}^{(i)} \tilde{\mathbf{R}}_i. \tag{4.3.19}$$

The Neumann-Neumann preconditioner $\mathbf{P}_{NN}$ approximating $\mathbf{S}^{-1}$ for optimal control can be extended to many subdomains as a weighted sum of all the local Schur inverses

$$\mathbf{P}_{NN} \;=\; \sum_{i=1}^{N} \mathbf{D}_i \, \tilde{\mathbf{R}}_i^T \, \mathbf{S}^{(i)^{-1}} \, \tilde{\mathbf{R}}_i \, \mathbf{D}_i, \tag{4.3.20}$$

for some diagonal scaling matrix $\mathbf{D}_i$ containing positive weights. For each interface node k associated with subdomain Ω_i, the corresponding nodal weight given by matrix entry $(\mathbf{D}_i)_{k,k}$ is computed as the reciprocal of the number of subdomains sharing node k.

4.3.2.2 Two-Level Schur Complement Method

It is well-known that the one-level Neumann-Neumann preconditioner for elliptic PDEs is good for only a small number subdomains. At each iteration, information is passed only between adjacent subdomains since the global correction step is computed only from the local subdomain errors. For large number of subdomains, this method can require many iterations to remove a smooth global error component. This behavior is also observed for the Neumann-Neumann preconditioner (4.3.20) for optimal control. To avoid this problem, we adapt the Balancing Neumann-Neumann method by Mandel [Man93] to the optimal control context (see also [HN04a, HN04b].) Following the description in [SBG96, Sec. 4.3.3], the Balancing Neumann-Neumann

method for solving the system $\mathbf{S}\mathbf{z}_\Gamma = \mathbf{g}$ may be stated as a three-substep method:

$$\mathbf{z}_\Gamma^{(k+1/3)} = \mathbf{z}_\Gamma^{(k)} + \tilde{\mathbf{R}}_0^T \tilde{\mathbf{S}}_0^{-1} \tilde{\mathbf{R}}_0 (\mathbf{g} - \mathbf{S}\mathbf{z}_\Gamma^{(k)}) \tag{4.3.21a}$$

$$\mathbf{z}_\Gamma^{(k+2/3)} = \mathbf{z}_\Gamma^{(k+1/3)} + \sum_{i=1}^{N} \mathbf{D}_i \tilde{\mathbf{R}}_i^T \mathbf{S}^{(i)^{-1}} \tilde{\mathbf{R}}_i \mathbf{D}_i (\mathbf{g} - \mathbf{S}\mathbf{z}_\Gamma^{(k+1/3)}) \tag{4.3.21b}$$

$$\mathbf{z}_\Gamma^{(k+1)} = \mathbf{z}_\Gamma^{(k+2/3)} + \tilde{\mathbf{R}}_0^T \mathbf{S}_0^{-1} \tilde{\mathbf{R}}_0 (\mathbf{g} - \mathbf{S}\mathbf{z}_\Gamma^{(k+2/3)}) \tag{4.3.21c}$$

where

$$\mathbf{S}_0 = \tilde{\mathbf{R}}_0 \mathbf{S} \tilde{\mathbf{R}}_0^T \tag{4.3.22}$$

$$\tilde{\mathbf{R}}_0 = \begin{pmatrix} \tilde{\mathbf{R}}_0^y & \\ & \tilde{\mathbf{R}}_0^p \end{pmatrix}. \tag{4.3.23}$$

The matrix $\mathbf{S}_0$ has dimension $2N \times 2N$ and acts as a coarse Schur operator on a vector containing two unknowns (state and adjoint) per subdomain. The restriction operator $\tilde{\mathbf{R}}_0$ returns for each subdomain two values: the weighted sum of the states and the weighted sum of the adjoints on the common nodes. In all cases, the weights are computed as the reciprocal of the number of subdomains sharing the node.

The three-substep method (4.3.21) may be combined into the Balancing Neumann-Neumann preconditioner as

$$\mathbf{P}_{BNN} = (\mathbf{I} - \tilde{\mathbf{R}}_0^T \mathbf{S}_0^{-1} \tilde{\mathbf{R}}_0 \mathbf{S})(\sum_{i=1}^{N} \mathbf{D}_i \tilde{\mathbf{R}}_i^T \mathbf{S}^{(i)^{-1}} \tilde{\mathbf{R}}_i \mathbf{D}_i)(\mathbf{I} - \mathbf{S}\tilde{\mathbf{R}}_0^T \mathbf{S}_0^{-1} \tilde{\mathbf{R}}_0) + \tilde{\mathbf{R}}_0^T \mathbf{S}_0^{-1} \tilde{\mathbf{R}}_0. \tag{4.3.24}$$

4.4 Numerical Results

Numerical experiments are performed to test the effectiveness of the Neumann-Neumann preconditioners for solving the model control problem (4.1.1). The KKT optimality system is solved using a preconditioned symmetric QMR method as the linear solver [FN94, FN95, FN92]. The code is implemented in Matlab, using the PDE Toolbox for finite element assembly. The domain $\Omega = (-1, 1) \times (-1, 1)$ is partitioned into equal-sized square subdomains in a checkerboard pattern (left plot of

Figure 3.2). The side length of each subdomain is denoted by H. Regular meshes consisting of triangular elements of various widths (denoted as h) are generated. We use $y_d(x) = \sin(\pi x_1)\sin(\pi x_2)$ and $f(x) = 1$ in all cases. The iterations are stopped when the residual 2-norm drops below a tolerance of 10^{-6}. The following reported iterations are for the outer loop, while using exact solves for the subdomain problems.

Table 4.1 summarizes the iteration counts for both preconditioners using right preconditioning.

For comparison, we solved a PDE domain decomposition problem using the PDE version of the one-level NN and two-level BNN preconditioners. The model problem is

$$\begin{aligned} -\Delta y(x) + y(x) &= 1 \qquad \text{in } \Omega, \\ y(x) &= 0 \qquad \text{on } \partial\Omega. \end{aligned} \tag{4.4.1}$$

Table 4.2 shows a similar dependence of these DD preconditioners on H and h.

We also tested the preconditioners on the boundary control problem

$$\min_{y,u} \quad \frac{1}{2}\int_\Omega (y(x) - y_d(x))^2\,dx + \frac{\alpha}{2}\int_{\partial\Omega} u^2(x)\,dx \tag{4.4.2}$$

subject to

$$\begin{aligned} -\Delta y(x) + y(x) &= f(x) \qquad \text{in } \Omega, \\ \tfrac{\partial}{\partial n} y(x) &= u(x) \qquad \text{on } \partial\Omega, \end{aligned} \tag{4.4.3}$$

with $y_d(x) = \sin(\pi x_1)\sin(\pi x_2)$ and $f(x) = 1$ as before. Table 4.3 summarizes the iteration counts for both preconditioners.

The numerical results show that both the Neumann-Neumann and the Balancing Neumann-Neumann preconditioners grow worse slowly when the mesh size h is reduced while other parameters are kept constant. This behavior is consistent with the

PDE version of these preconditioners. As expected, the Neumann-Neumann (NN) preconditioner grows worse quickly as the number of subdomain increases while the Balancing Neumann-Neumann (BNN) preconditioner remains effective. The BNN results seem relatively stable for both problem types over a large range of regularization parameter α. The NN results, however, are dependent on the problem type. In the boundary control problem, the NN method is relatively insensitive to α, while results for the distributed control problem is strongly affected by α.

Table 4.1: SQMR iterations, Distributed Control. Left: One-level Neumann-Neumann. Right: Two-level Balancing Neumann-Neumann.

(a) $\alpha = 1$

SQMR Iterations						
Neumann-Neumann, $\alpha = 1.0$						
Distributed Control						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	2	2	2	1	1	1
1/2		10	10	11	12	15
1/4			35	37	42	49
1/8				115	137	146
1/16					277	353

SQMR Iterations						
Balancing Neumann-Neumann, $\alpha = 1.0$						
Distributed Control						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	2	2	2	1	1	1
1/2		7	7	9	9	10
1/4			7	9	10	13
1/8				7	9	11
1/16					7	8

(b) $\alpha = 10^{-4}$

SQMR Iterations						
Neumann-Neumann, $\alpha = 10^{-4}$						
Distributed Control						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	5	4	3	2	2	1
1/2		6	7	8	9	10
1/4			12	14	17	17
1/8				33	39	42
1/16					104	121

SQMR Iterations						
Balancing Neumann-Neumann, $\alpha = 10^{-4}$						
Distributed Control						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	5	4	3	2	2	1
1/2		5	7	10	11	12
1/4			8	10	13	15
1/8				9	11	15
1/16					7	10

(c) $\alpha = 10^{-8}$

SQMR Iterations						
Neumann-Neumann, $\alpha = 10^{-8}$						
Distributed Control						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	6	6	5	4	3	2
1/2		9	6	7	6	6
1/4			7	7	6	6
1/8				7	6	6
1/16					7	6

SQMR Iterations						
Balancing Neumann-Neumann, $\alpha = 10^{-8}$						
Distributed Control						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	5	6	5	4	3	2
1/2		7	7	6	6	7
1/4			6	6	6	6
1/8				7	6	6
1/16					6	7

Table 4.2: SQMR iterations, PDE only. Left: One-level Neumann-Neumann. Right: Two-level Balancing Neumann-Neumann.

SQMR Iterations						
PDE only, Neumann-Neumann						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	2	2	2	1	1	1
1/2		6	7	7	8	10
1/4			17	20	24	28
1/8				47	57	62
1/16					87	115

SQMR Iterations						
PDE only, Balancing Neumann-Neumann						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	2	2	2	1	1	1
1/2		4	5	6	6	7
1/4			5	7	8	10
1/8				5	7	9
1/16					6	7

Table 4.3: SQMR iterations, Boundary Control. Left: One-level Neumann-Neumann. Right: Two-level Balancing Neumann-Neumann.

(a) $\alpha = 1$

SQMR Iterations						
Neumann-Neumann, $\alpha = 1.0$						
Boundary Control						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	2	2	2	1	1	1
1/2		11	13	15	16	17
1/4			44	54	61	82
1/8				151	186	197
1/16					447	571

SQMR Iterations						
Balancing Neumann-Neumann, $\alpha = 1.0$						
Boundary Control						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	2	2	2	1	1	1
1/2		4	5	6	8	8
1/4			4	7	8	10
1/8				3	7	8
1/16					3	7

(b) $\alpha = 10^{-4}$

SQMR Iterations						
Neumann-Neumann, $\alpha = 10^{-4}$						
Boundary Control						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	3	3	3	2	2	1
1/2		10	13	13	15	16
1/4			41	43	43	46
1/8				115	113	119
1/16					344	334

SQMR Iterations						
Balancing Neumann-Neumann, $\alpha = 10^{-4}$						
Boundary Control						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	3	3	3	2	2	1
1/2		5	7	8	10	11
1/4			6	8	11	12
1/8				5	8	11
1/16					5	8

(c) $\alpha = 10^{-8}$

SQMR Iterations						
Neumann-Neumann, $\alpha = 10^{-8}$						
Boundary Control						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	3	3	3	3	3	3
1/2		10	12	14	15	16
1/4			45	46	54	49
1/8				154	157	156
1/16					413	410

SQMR Iterations						
Balancing Neumann-Neumann, $\alpha = 10^{-8}$						
Boundary Control						
H \ h	1/2	1/4	1/8	1/16	1/32	1/64
1	3	3	3	3	3	3
1/2		6	8	11	14	16
1/4			9	13	18	20
1/8				11	15	22
1/16					11	15

Chapter 5

Summary and Future Work

We have investigated how domain decomposition ideas can be applied to solving quadratic optimal control problems constrained by linear elliptic PDEs. Our general approach is to decompose the original control problem into smaller linear-quadratic control problems that are posed on the subdomains. For the case of many subdomains, we include a coarse-level problem to improve the convergence rate. We considered both overlapping and nonoverlapping domain decomposition, and implemented all methods as preconditioners for Krylov iterative methods on the KKT optimality system. Numerical experiments show that our domain decomposition methods are effective for the model elliptic optimal control problems. The preconditioners for solving the optimality system have convergence properties similar to preconditioners for solving PDE problems.

For overlapping methods, the amount of overlap δ has a similar effect on the control problem as the PDE problem. In general, only a small positive δ is required for reasonable convergence. In the one-level method, increasing δ while keeping other parameters constant results in faster convergence. By comparison, the two-level method is relatively insensitive to δ. The numerical results seem to agree with our convergence theory: if the coarse grid is made fine enough, then the convergence rate of the two-level method can be bounded independently of the mesh parameters. This

qualitative behavior is observed for all choices of the regularization parameters α.

In the nonoverlapping methods, convergence for both the Neumann-Neumann and the Balancing Neumann-Neumann methods depends weakly on the fine mesh parameter h, similar to the PDE case. As the number of subdomains increases, the one-level Neumann-Neumann method deteriorates quickly, as expected, while the Balancing Neumann-Neumann method remains effective. We do not have theoretical bounds on the convergence of these Neumann-Neumann methods. It would be worthwhile to derive the asymptotic relationships between the convergence rate and the parameters H, h and α.

The effect of the regularization parameter α on convergence depends on both the type of control problem and the number of levels. The two-level methods (both overlapping and Balancing Neumann-Neumann) perform well over a large range of α. The one-level method, however, seems to be problem dependent. For the distributed control problem, the normally troublesome case of many subdomains unexpectedly improves as α is reduced. When α is very small, the convergence seems to be independent of the number of subdomains or the mesh discretization level. This effect was not observed for the boundary control problem, where the convergence is relatively insensitive to α.

Our focus so far has been on the ideas and theories behind the extension of domain decomposition methods to optimal control problems. Future work should include the parallel implementation of these methods for large-scale problems. It would be instructive to do a numerical comparison on parallel computers between the proposed methods and some previous methods. In theory, our methods allow the most amount of work to be carried out in parallel, since the subproblems are optimal control problems and not just PDEs. We can also consider more general elliptic optimal control problems to extend these methods to a wider variety of applications.

Bibliography

[AB01] O. Axelsson and V. A. Barker. *Finite Element Solution of Boundary Value Problems*. Classics in Applied Mathematics, Vol. 35. SIAM, Philadelphia, 2001.

[AH03] U. M. Ascher and E. Haber. A multigrid method for distributed parameter estimation problems. *Electron. Trans. Numer. Anal.*, 15:1–17 (electronic), 2003. Tenth Copper Mountain Conference on Multigrid Methods (Copper Mountain, CO, 2001).

[BBC+93] R. Barrett, M. Berry, T. F. Chan, J. Demmel J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine, and H. van der Vorst. *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods*. SIAM, Philadelphia, 1993. Hypertext version is available at http://www.netlib.org/linalg/ html_templates/Templates.html.

[Ben94] J.-D. Benamou. A domain decomposition method with coupled transmission conditions for the optimal control of systems governed by elliptic partial differential equations. Technical Report No. 2246, INRIA, Domaine de Voluceau, 78153 Rocquencourt, France, 1994. Electronically available as http://www.inria.fr/rrrt/rr-2246.html.

[Ben96] J.-D. Benamou. A domain decomposition method with coupled transmission conditions for the optimal control of systems governed by elliptic partial differential equations. *SIAM J. Numer. Anal.*, 33:2401–2416, 1996.

[Ben98] J.-D. Benamou. A domain decomposition method for control problems. In P. Bjørstad, M. Espedal, and D. Keyes, editors, *DD9 Proceedings*, pages 266–273, Bergen, Norway, 1998. Domain Decomposition Press. Available electronically from http://www.ddm.org/DD9/index.html.

[BF91] F. Brezzi and M. Fortin. *Mixed and Hybrid Finite Element Methods*. Computational Mathematics, Vol. 15. Springer–Verlag, Berlin, 1991.

[BG99] G. Biros and O. Ghattas. Parallel preconditioners for KKT systems arising in optimal control of viscous incompressible flows. In *Proceedings of Par-*

allel CFD '99, Williamsburg, VA, May 23–26, 1999, Amsterdam, London, New-York, 1999. North Holland. http://www.cs.cmu.edu/∼oghattas.

[BG00] G. Biros and O. Ghattas. Parallel Lagrange–Newton–Krylov–Schur methods for PDE–constrained optimization. Part I: The Krylov–Schur solver. Technical report, Computational Mechanics Lab, Department of Civil and Environmental Engineering, Carnegie Mellon University, 2000. http://www.cs.cmu.edu/∼oghattas.

[BH98] A. Battermann and M. Heinkenschloss. Preconditioners for Karush–Kuhn–Tucker systems arising in the optimal control of distributed systems. In W. Desch, F. Kappel, and K. Kunisch, editors, *Optimal Control of Partial Differential Equations*, Int. Series of Numer. Math. Vol. 126, pages 15–32, Basel, Boston, Berlin, 1998. Birkhäuser Verlag.

[Bou98] A. Bounaim. A Lagrangian approach to a DDM for an optimal control problem. In P. Bjørstad, M. Espedal, and D. Keyes, editors, *DD9 Proceedings*, pages 283–289, Bergen, Norway, 1998. Domain Decomposition Press. Available electronically from http://www.ddm.org/DD9/index.html.

[BS01] A. Battermann and E. W. Sachs. Block preconditioners for KKT systems in PDE-governed optimal control problems. In K.-H. Hoffmann, R. H. W. Hoppe, and V. Schulz, editors, *Fast solution of discretized optimization problems (Berlin, 2000)*, volume 138 of *Internat. Ser. Numer. Math.*, pages 1–18, Basel, 2001. Birkhäuser.

[BS02] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*. Springer Verlag, Berlin, Heidelberg, New York, second edition, 2002.

[BW89] P. E. Bjørstad and O. B. Widlund. To overlap or not to overlap: A note on a domain decomposition method for elliptic problems. *SIAM J. Sci. Stat. Comput.*, 10:1053–1061, 1989.

[BX91] J. H. Bramble and J. Xu. Some estimates for a weighted L^2 projection. *Math. Comput.*, 56:463–476, 1991.

[Cai90] X. C. Cai. An additive Schwarz algorithm for nonselfadjoint elliptic equations. In T. Chan, R. Glowinski, J. Périaux, and O. Widlund, editors, *Third International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 232–244. SIAM, Philadelphia, PA, 1990.

[CG92] T. F. Chan and D. Goovaerts. On the relationship between overlapping and nonoverlapping domain decomposition methods. *SIAM J. Matrix Anal. Appl.*, 13:663–670, 1992.

[Cia02] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems*. SIAM, Philadelphia, 2002.

[CM94] T. F. Chan and T. P. Mathew. Domain decomposition algorithms. In *Acta Numerica 1994*, pages 61–143. Cambridge University Press, 1994.

[CW92] X.-C. Cai and O. B. Widlund. Domain decomposition algorithms for indefinite elliptic problems. *SIAM J. Sci. Statist. Comput.*, 13(1):243–258, January 1992.

[CW93] X.-C. Cai and O. B. Widlund. Multiplicative Schwarz algorithms for some nonsymmetric and indefinite problems. *SIAM J. Numer. Anal.*, 30(4):936–952, August 1993.

[DDSV98] J. J. Dongarra, I. S. Duff, D. C. Sorensen, and H. A. Van der Vorst. *Numerical Linear Algebra for High-Performance Computers*. SIAM, Philadelphia, 1998.

[Dem97] J. W. Demmel. *Applied Numerical Linear Algebra*. SIAM, Philadelphia, 1997.

[DL94] J. E. Dennis and R. M. Lewis. A comparison of nonlinear programming approaches to an elliptic inverse problem and a new domain decomposition approach. Technical Report TR94–33, Department of Computational and Applied Mathematics, Rice University, P.O.Box 1892, Houston, TX 77251–1892, 1994.

[DSW94] 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(3):1662–1694, 1994.

[DW89] M. Dryja and O. B. Widlund. Some domain decomposition algorithms for elliptic problems. In D. R. Kinkaid and L. J. Hayes, editors, *Iterative Methods for Large Linear Systems*, pages 273–291, San Diego, CA, 1989. Academic Press.

[DW90] M. Dryja and O. B. Widlund. Towards a unified theory of domain decomposition algorithms for elliptic problems. In T. F. Chan, R. Glowinski, J. Periaux, and O. B. Widlund, editors, *Third International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 3–21, Philadelphia, 1990. SIAM.

[DW92] M. Dryja and O. B. Widlund. Additive Schwarz methods for elliptic finite element problems in three dimensions. In D. E. Keyes, T. F. Chan, G. A. Meurant, J. S. Scroggs, and R. G. Voight, editors, *Fifth International*

Symposium on Domain Decomposition Methods for Partial Differential Equations, pages 3–18, Philadelphia, 1992. SIAM.

[DW94] M. Dryja and O. B. Widlund. Domain decomposition algorithms with small overlap. *SIAM J. Sci.Comput.*, 15(3):604–620, May 1994.

[DW95] M. Dryja and O. B. Widlund. Schwarz methods of Neumann-Neumann type for three-dimensional elliptic finite element methods. *Comm. Pure Appl. Math.*, 48:121–155, 1995.

[EES83] S. C. Eisenstat, H. C. Elman, and M. H. Schultz. Variational iterative methods for nonsymmetric systems of linear equations. *SIAM J. Numer. Anal.*, 20:345–357, 1983.

[FN92] R. Freund and N. M. Nachtigal. QMR: a quasi–minimal residual method for non–Hermitian linear systems. *Numerische Mathematik*, 60:315–340, 1992.

[FN94] R. W. Freund and N. M. Nachtigal. A new Krylov-subspace method for symmetric indefinite linear systems. In W. F. Ames, editor, *Proceedings of the 14th IMACS World Congress on Computational and Applied Mathematics*, pages 1253–1256. IMACS, 1994. http://cm.bell-labs.com/cm/cs/doc/nam.html.

[FN95] R. W. Freund and N. M. Nachtigal. Software for simplified Lanczos and QMR algorithms. *Applied Numerical Mathematics*, 19:319–341, 1995.

[GG03] G. H. Golub and C. Greif. On solving block-structured indefinite linear systems. *SIAM J. Sci. Comput.*, 24(6):2076–2092, 2003.

[GHN01] N. I. M. Gould, M. E. Hribar, and J. Nocedal. On the solution of equality constrained quadratic programming problems arising in optimization. *SIAM J. Sci. Comput.*, 23(4):1376–1395, 2001.

[GHR97] 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*, 77:243–268, 1997. `http://link.springer.de/link/service/journals/00211/bibs/7077002/70770243.htm`.

[GHS91] 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.*, pages 123–131, 1991.

[GHS93] M. D. Gunzburger, L. S. Hou, and T. P. Svobotny. Heating and cooling control of temperature distributions along boundaries of flow domains. *J. of Mathematical Systems, Estimation, and Control*, 3:147–172, 1993.

[GMPS92] P. E. Gill, W. Murray, D. B. Ponceleón, and M. A. Saunders. Preconditioners for indefinite systems arising in optimization. *SIAM J. Matrix Anal. Appl.*, 13:292–311, 1992.

[GO95] M. Griebel and P. Oswald. On the abstract theory of additive and multiplicative Schwarz algorithms. *Numer. Math.*, 70:163–180, 1995.

[Gre97] A. Greenbaum. *Iterative Methods for the Solution of Linear Systems*. SIAM, Philadelphia, 1997.

[GT01] D. Gilbarg and N.S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Springer-Verlag, Berlin, second (Classics in Mathematics) edition, 2001.

[GV96] G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, London, third edition, 1996.

[HA00] E. Haber and U. Ascher. Preconditioned all–at–once methods for large sparse parameter estimation problems. *Inverse Problems*, 17:1847–1864, 2000.

[Hac92] W. Hackbusch. *Elliptic Differential Equations: Theory and Numerical Treatment*. Springer–Verlag, Berlin, 1992.

[HN04a] M. Heinkenschloss and H. Nguyen. Balancing Neumann-Neumann methods for elliptic optimal control problems. In R. Kornhuber, R. H. W. Hoppe, J. Periaux, O. Pironneau, O. B. Widlund, and J. Xu, editors, *Proceedings of the 15th International Conference on Domain Decomposition*, Lecture Notes in Computational Science and Engineering, Vol., pages 589–596, Heidelberg, 2004. Springer-Verlag.

[HN04b] M. Heinkenschloss and H. Nguyen. Neumann-Neumann domain decomposition preconditioners for linear-quadratic elliptic optimal control problems. Technical Report TR04–01, Department of Computational and Applied Mathematics, Rice University, 2004.

[HPS02] R. H. W. Hoppe, S. I. Petrova, and V. Schulz. Primal-dual Newton-type interior-point method for topology optimization. *J. Optim. Theory Appl.*, 114(3):545–571, 2002.

[KGW00] C. Keller, N. I. M. Gould, and A. J. Wathen. Constrained preconditioning for indefinite linear systems. *SIAM J. Matrix Analysis and Applications*, 21:1300–1317, 2000.

[Lio71] J.-L. Lions. *Optimal Control of Systems Governed by Partial Differential Equations*. Springer Verlag, Berlin, Heidelberg, New York, 1971.

[Lio88] P. L. Lions. On the Schwarz alternating method. I. In R. Glowinski, G. H. Golub, G. A. Meurant, and J. Périaux, editors, *First International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 1–42, Philadelphia, PA, 1988. SIAM.

[Lio89] P. L. Lions. On the Schwarz alternating method. II. In T. Chan, R. Glowinski, J. Périaux, and O. Widlund, editors, *Domain Decomposition Methods*, pages 47–70, Philadelphia, PA, 1989. SIAM.

[Lio90] P. L. Lions. On the Schwarz alternating method. III: a variant for nonoverlapping subdomains. In T. Chan, R. Glowinski, J. Périaux, and O. Widlund, editors, *Third International Symposium on Domain Decomposition Methods for Partial Differential Equations , held in Houston, Texas, March 20-22, 1989*, pages 202–223, Philadelphia, PA, 1990. SIAM.

[Lit00] 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.

[LP98] J.-L. Lions and O. Pironneau. Sur le contrôle parallèle des systèmes distribués. *Comptes Rendus de l'Académie des Sciences. Série I. Mathématique*, 327(12):993–998, 1998.

[LV98] L. Lukšan and J. Vlček. Indefinitely preconditioned inexact Newton method for large sparse equality constrained non-linear programming problems. *Num. Lin. Alg. Appl.*, 5(3):219–247, 1998.

[LW03] D. Loghin and A. J. Wathen. Schur complement preconditioning for elliptic systems of partial differential equations. *Numerical Linear Algebra with Applications*, 10:423–443, 2003.

[Man93] J. Mandel. Balancing domain decomposition. *Comm. Numer. Meth. Engrg.*, 9:233–241, 1993.

[MGW00] M. F. Murphy, G. H. Golub, and A. J. Wathen. A note on preconditioning for indefinite linear systems. *SIAM J. Sci. Comput.*, 22:1969–1972, 2000.

[MN00] J. L. Morales and J. Nocedal. Automatic preconditioning by limited memory quasi-Newton updating. *SIAM J. Optim.*, 10(4):1079–1096 (electronic), 2000.

[MS92] J. C. Meza and W. W. Symes. Conjugate residual methods for almost symmetric linear systems. *J. Optim. Theory Appl.*, 72:415–440, 1992.

[MS00] B. Maar and V. Schulz. Interior–point multigrid methods for topology optimization. *Structural Optimization*, 19:214–224, 2000.

[PS75] C. C. Paige and M. A. Saunders. Solution of sparse indefinite systems of linear equations. *SIAM J. Numer. Anal.*, 12:617–629, 1975.

[PW02] L. F. Pavarino and O. B. Widlund. Balancing Neumann-Neumann methods for incompressible Stokes equations. *Comm. Pure Appl. Math.*, 55(3):302–335, 2002.

[QV94] A. Quarteroni and A. Valli. *Numerical Approximation of Partial Differential Equations*. Springer, Berlin, Heidelberg, New York, 1994.

[QV99] A. Quarteroni and A. Valli. *Domain Decomposition Methods for Partial Differential Equations*. Oxford University Press, Oxford, 1999.

[Saa03] Y. Saad. *Iterative Methods for Sparse Linear Systems*. SIAM, Philadelphia, second edition, 2003.

[SBG96] 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.

[Sch74] A. H. Schatz. An observation concerning Ritz-Galerkin methods with indefinite bilinear forms. *Math. Comp.*, 28(128):959–962, 1974.

[SCH96] A. Shenoy, E. M. Cliff, and M. Heinkenschloss. Thermal–fluid control via finite–dimensional approximations. In *Proceedings from the 31st AIAA Thermophysics Conference, AIAA Paper 96–1910*, New Orleans, LA, 1996.

[SS86] Y. Saad and M. H. Schultz. GMRES a generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Stat. Comp.*, 7:856–869, 1986.

[SW94] D. Sylvester and A. Wathen. Fast iterative solution of stabilized Stokes systems part II: using general block preconditioners. *SIAM J. Numer. Anal.*, 31:1352–1367, 1994.

[TB97] L. N. Trefethen and D. Bau. *Numerical Linear Algebra*. SIAM, Philadelphia, 1997.

[XC92] J. Xu and X.-C. Cai. A preconditioned GMRES method for nonsymmetric or indefinite problems. *Math. Comp.*, 59:311–319, 1992.

[Xu92a] J. Xu. Iterative methods by space decomposition and subspace correction. *SIAM Review*, 34:581–613, 1992.

[Xu92b] J. Xu. A new class of iterative methods for nonselfadjoint or indefinite problems. *SIAM J. Numer. Anal.*, 29(2):303–319, 1992.

[Zei95] E. Zeidler. *Applied Functional Analysis. Main Principles and Their Applications*. Applied Mathematical Sciences, Vol. 109. Springer Verlag, Berlin, Heidelberg, New-York, 1995.