KKT PRECONDITIONERS FOR PDE-CONSTRAINED OPTIMIZATION WITH THE HELMHOLTZ EQUATION*

DREW P. KOURI†, DENIS RIDZAL‡, AND RAY TUMINARO§

Abstract. This paper considers preconditioners for the linear systems that arise from optimal control and inverse problems involving the Helmholtz equation. Specifically, we explore an all-at-once approach. The main contribution centers on the analysis of two block preconditioners. Variations of these preconditioners have been proposed and analyzed in prior works for optimal control problems where the underlying partial differential equation is a Laplace-like operator. In this paper, we extend some of the prior convergence results to Helmholtz-based optimization applications. Our analysis examines situations where control variables and observations are restricted to subregions of the computational domain. We prove that solver convergence rates do not deteriorate as the mesh is refined or as the wavenumber increases. More specifically, for one of the preconditioners we prove accelerated convergence as the wavenumber increases. Additionally, in situations where the control and observation subregions are disjoint, we observe that solver convergence rates have a weak dependence on the regularization parameter. We give a partial analysis of this behavior. We illustrate the performance of the preconditioners on control problems motivated by acoustic testing.

1. Introduction and Motivation. This paper investigates solution techniques for the first-order optimality Karush-Kuhn-Tucker (KKT) systems that arise from linear-quadratic optimization problems involving the Helmholtz equation. Efficient solution of these systems is motivated by optimal control applications, e.g., direct-field acoustic testing [22], remote sensing applications, e.g., source inversion in seismology [2], etc. As is well known, the discretized Helmholtz equation is often difficult to solve using iterative methods [13], although several successful techniques have now been proposed, e.g., [3, 9, 10, 11, 12, 28]. We examine block preconditioning methods for the more complex saddle-point system associated with the KKT conditions, in which the Helmholtz operator is one of the system blocks. A key requirement for the block preconditioners is that they leverage suitable solvers for the Helmholtz block. While our focus is on KKT solvers involving the Helmholtz equation, the proposed preconditioners and the accompanying analyses may be useful in solving optimization problems with other partial differential equations (PDEs).

To orient the initial discussion, consider a discretized linear-quadratic optimization problem and its KKT system,

$$\begin{cases} \min\limits_{\mathbf{u},\mathbf{z}} \ \frac{1}{2}\mathbf{u}^*\mathbf{C}\mathbf{u} - \mathbf{u}^*\mathbf{w} + \frac{\beta}{2}\mathbf{z}^*\mathbf{R}\mathbf{z} \\ \text{subject to} \quad \mathbf{K}\mathbf{u} + \mathbf{L}\mathbf{z} = 0 \end{cases} \qquad \text{and} \qquad \begin{pmatrix} \mathbf{C} & 0 & \mathbf{K}^* \\ 0 & \beta\mathbf{R} & \mathbf{L}^* \\ \mathbf{K} & \mathbf{L} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ \mathbf{z} \\ \boldsymbol{\lambda} \end{pmatrix} = \begin{pmatrix} \mathbf{w} \\ 0 \\ 0 \end{pmatrix},$$

where $\mathbf{K} \in \mathbb{C}^{m\times m}$ is an invertible PDE matrix; $\mathbf{L} \in \mathbb{C}^{m\times n}$ is a (generally rectangular)

*DPK and DR were supported by the U.S. Air Force Office of Scientific Research, Optimization Program. DR was additionally supported by the Sandia National Laboratories' LDRD Program. RT was supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Applied Mathematics program. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under grant DE-NA-0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

†Sandia National Labs, Albuquerque, NM 87123 (dpkouri@sandia.gov)

‡Sandia National Labs, Albuquerque, NM 87123 (dridzal@sandia.gov)

§Sandia National Labs, Livermore, CA 94551 (rstumin@sandia.gov)

control-to-state matrix; $\mathbf{C} \in \mathbb{C}^{m\times m}$ is a (generally singular) positive semidefinite observation matrix; and $\mathbf{R} \in \mathbb{R}^{n\times n}$ is a positive definite control penalty matrix or regularization matrix. For the remainder of the paper we refer to $\beta > 0$ as the regularization parameter. Moreover, $\mathbf{u} \in \mathbb{C}^m$ denotes the state variables, $\mathbf{z} \in \mathbb{C}^n$ the control variables, $\mathbf{w} \in \mathbb{C}^m$ the observations, and $\boldsymbol{\lambda} \in \mathbb{C}^m$ the Lagrange multipliers. We distinguish between two cases, the *general case* and the *idealized case*, as follows:

- *General case*: No further assumptions are made on $\mathbf{K}, \mathbf{L}, \mathbf{C}$ and $\mathbf{R}$.
- *Idealized case*: Let $\mathbf{M} \in \mathbb{R}^{m\times m}$ be a typical (positive definite) mass matrix that accompanies the PDE discretization. Set $n = m$ and $\mathbf{L} = \mathbf{C} = \mathbf{R} = \mathbf{M}$.

We note that the general case naturally accommodates restrictions of control variables and observations to subregions of the physical domain. The idealized case does not; moreover, it requires that the discretizations of the state and control spaces be identical, which may not be true even if the control variables are defined on the entire computational domain. These limitations are impractical in many applications; for instance, in direct-field acoustic testing, the control variables correspond to the output amplitudes of a set of speakers, which are restricted to a small subregion of the physical domain, and the observation region is another small subregion surrounded by the speakers. See Figure 5.1 for an illustration.

Our work builds on a series of related preconditioning papers [4, 8, 23, 26, 27, 29] for PDE-constrained optimization with the Poisson equation. With the notable exception of [8, 23, 26], these papers focus on the idealized case as opposed to the general case. We briefly summarize our contributions before reviewing them in the context of prior works. First, we analyze the problem where $\mathbf{K}$ is a discretization of the Helmholtz operator, by extending some of the prior convergence results. The analysis is done for the general case, where controls and observations may be defined in subregions of the computational domain. Under some assumptions on the governing Helmholtz equation, we prove accelerated convergence of the preconditioner as the wavenumber increases. Second, we examine the setting where the control and observation subregions are disjoint, and show that the solver convergence is weakly dependent on the regularization parameter β, deteriorating substantially only when β is very small. Third, we analyze a preconditioner that is motivated by the idealized case. In this setting, we prove β independence and, under some problem assumptions, wavenumber independence.

We highlight a few features of related existing methods, in the context of optimal control problems governed by the Poisson equation. A significant drawback of most of the earlier-developed preconditioners is that they are only valid for the idealized case. For example, the proofs for the eigenvalue bounds on the preconditioned Schur complement in [29] require that the mass matrix have a Cholesky factorization. As already noted, [8, 23, 26] consider preconditioning extensions that fit the general case, which includes the uses of a perturbed mass matrix that will be discussed further. In [4], a preconditioner is proposed for the KKT system associated with the idealized case, which is *exact* requiring only that the PDE matrix be Hermitian. That is, this method converges in three iterations when exact solvers are used for the preconditioner's sub-block solves. While the exact property is highly desirable, the method requires imaginary arithmetic even when not required by the application. Further, the method's extension from the idealized case to the general case is not clear and the preconditioner will no longer be exact in this scenario. This preconditioning approach is discussed further in Section 4.2. The general-case KKT preconditioner considered in the current manuscript is not exact. When evaluating preconditioners

for PDE-constrained optimization, we are interested in how the solver's convergence rate depends on mesh refinement and on the penalty parameter β. The preconditioner in [29] is analyzed and shown to be optimal for the idealized case when $\mathbf{K}$ corresponds to a Poisson-like operator. That is, the convergence rate does not deteriorate as the PDE mesh is refined, though the convergence rate does deteriorate as the penalty parameter β becomes smaller. An adaptation of this preconditioner is presented in [27]. This adaptation is less sensitive to the penalty parameter β, however this improvement is still oriented towards the idealized case. On the other hand, [26] considers a range of time-dependent optimal control problems that yield KKT systems belonging to the general case. Different adaptations of the preconditioner in [27] are considered with an associated analysis for several scenarios. In our manuscript, we leverage a perturbation idea presented in [26] (explained in Sections 4.1 and 5.3), and later used in [8, 23], to adapt the preconditioner to the scenario where $\mathbf{C}$ is singular, e.g., when observations are defined in a subregion. In addition to extending convergence results to Helmholtz-based PDEs, our analysis reveals additional insights for the general case (even when $\mathbf{K}$ is not necessarily a Helmholtz operator).

The remainder of the paper is organized as follows. First, we briefly review the Helmholtz optimal control problem, laying the necessary groundwork for proving the conditioning results that follow. We then describe in more detail the KKT preconditioners focusing first on the general case. While the general-case preconditioner is motivated by [29], our analysis differs substantially, requiring lower bounds on the singular values (rather than eigenvalues), stability estimates for the Helmholtz equation, as well as understanding resonance and wavenumber dependence. The main theoretical drawback in the general case is that there is some convergence degradation based on the penalty parameter, β. We follow this by considering the *exact* preconditioning method of [4], but focusing on the scenario where the matrix is not Hermitian, due to impedance boundary conditions or damping. Our analysis highlights that the associated convergence rates are only weakly dependent on β, though the results are limited to the idealized case. Numerical results for direct field acoustic testing are presented to complement the theoretical discussion. We conclude with a further discussion of the general case to better understand the nature of the preconditioned eigen-spectrum and why convergence degradation is less pronounced when the control and observation regions are disjoint.

2. Notation. Throughout, Ω will be an open bounded and connected subset of $\mathbb{R}^d$ ($d = 1, 2, 3$) with Lipschitz boundary $\partial\Omega$ that is partitioned into three non-overlapping segments Γ_i, $i = 1, 2, 3$. That is, $\Gamma_1 \cup \Gamma_2 \cup \Gamma_3 = \partial\Omega$ and $\Gamma_i \cap \Gamma_j$ has measure zero for $i \neq j$. Additionally, W and Z will denote Hilbert spaces: W is the space of observations and Z is the space of controls. For a set $D \subseteq \mathbb{R}^k$ with $k \in \mathbb{N}$, we denote the space of square-integrable, complex-valued functions defined on D by $L^2(D)$ and we denote the Sobolev space of functions in $L^2(D)$ whose weak derivatives are also in $L^2(D)$ by $H^1(D)$. Moreover, we denote the space of essentially bounded functions defined on D by $L^\infty(D)$. When $\Gamma_3 \neq \emptyset$, we denote by $H^1_{\Gamma_3}(\Omega)$ the subset of functions in $H^1(\Omega)$ whose trace on Γ_3 is zero. If $\Gamma_3 \neq \emptyset$, we will denote $U := H^1_{\Gamma_3}(\Omega)$. Otherwise, $U := H^1(\Omega)$.

3. Optimal Control of the Helmholtz Equation. We consider the generic Helmholtz optimal control problem

$$\min_{u \in U,\, z \in Z} \ \frac{1}{2}\|\mathcal{W}(u) - w\|_W^2 + \frac{\beta}{2}\|z\|_Z^2 \tag{3.1a}$$

subject to

$$-\Delta u - \left(\kappa^2 + \imath\gamma\kappa\right) u = \mathcal{F}(z) \quad \text{in } \Omega \tag{3.1b}$$
$$\partial_\nu u = \mathcal{B}_1(z) \quad \text{on } \Gamma_1 \tag{3.1c}$$
$$\partial_\nu u - \imath\delta\kappa u = \mathcal{B}_2(z) \quad \text{on } \Gamma_2 \tag{3.1d}$$
$$u = 0 \qquad \text{on } \Gamma_3, \tag{3.1e}$$

where $\imath = \sqrt{-1}$, $\omega = 2\pi f$ is the angular frequency for $f > 0$, $c \in L^\infty(\Omega)$ with $0 < c_0 \le c(x) \le c_1 < \infty$ almost everywhere (a.e.) is the speed of sound in the medium occupying Ω, $\kappa = \omega/c$ is the wavenumber, $\gamma \in L^\infty(\Omega)$ with $0 \le \gamma_0 \le \gamma(x) \le \gamma_1 < \infty$ a.e. is the damping factor, and $\delta \in L^\infty(\Gamma_2)$ with $0 < \delta_0 \le \delta(x) \le \delta_1 < \infty$ a.e. is the impedance factor on Γ_2. The solution to (3.1b-3.1e), $u \in U$, is the acoustic pressure associated with the control variable $z \in Z$. In addition, $\mathcal{W} : L^2(\Omega) \to W$ is a continuous linear operator defining the observations of the acoustic pressure, $w \in W$ is the target data, $\beta > 0$ is a control penalty parameter, $\mathcal{F} : Z \to L^2(\Omega)$ is a continuous linear operator defining the control action on the interior of Ω, and $\mathcal{B}_i : Z \to L^2(\Gamma_i)$ is a continuous linear operator defining the control action on Γ_i, $i = 1, 2$.

For our analysis, we consider the weak form of (3.1b-3.1e). We define the sesquilinear forms $a : U \times U \to \mathbb{C}$, $b : U \times U \to \mathbb{C}$, and $\ell : Z \times U \to \mathbb{C}$ as

$$\begin{aligned} a(u, v) &= \int_\Omega \left[\nabla u(x) \cdot \nabla \overline{v}(x) - \kappa^2 u(x)\overline{v}(x)\right] \mathrm{d}x \\ b(u, v) &= \int_\Omega \gamma\kappa u(x)\overline{v}(x)\,\mathrm{d}x + \int_{\Gamma_2} \delta\kappa u(x)\overline{v}(x)\,\mathrm{d}x \\ \ell(z; v) &= \int_\Omega [\mathcal{F}(z)](x)\overline{v}(x)\,\mathrm{d}x + \int_{\Gamma_1} [\mathcal{B}_1(z)](x)\overline{v}(x)\,\mathrm{d}x + \int_{\Gamma_2} [\mathcal{B}_2(z)](x)\overline{v}(x)\,\mathrm{d}x, \end{aligned} \tag{3.2}$$

where $\overline{v}$ is the complex conjugate of v. The variational formulation of the Helmholtz equation (3.1b-3.1e) is then to find $u \in U$ that satisfies

$$a(u, v) - \imath b(u, v) = \ell(z; v) \quad \forall\, v \in U. \tag{3.3}$$

Note that the sesquilinear form a satisfies the Gårding inequality

$$\mathrm{Re}\,(a(u, u) - \imath b(u, u)) + 2\kappa^2 \|u\|^2_{L^2(\Omega)} \ge \min\{\kappa^2, 1\} \|u\|^2_U \quad \forall\, u \in U \tag{3.4}$$

and therefore, the Fredholm alternative ensures the existence of solutions if one can prove uniqueness. To avoid complication, we make the following well-posedness assumption for (3.3).

Assumption 3.1 (Well-Posedness of the Helmholtz Equation)**.** *The continuous linear operator $K : U \to U^*$ defined by $Ku = a(u, \cdot) - \imath b(u, \cdot)$ for all $u \in U$ has a bounded inverse denoted K^{-1}.*

Remark 3.2 (Resonance)**.** *In general, the Helmholtz operator K may be singular. For example, in the absence of damping and impedance (i.e., $\gamma \equiv 0$ and $\Gamma_2 = \emptyset$), K is singular if κ^2 is an eigenvalue of the Laplacian on Ω with appropriately prescribed boundary conditions (i.e., if the speed of sound c is constant, then ω is a resonance frequency). Assumption 3.1 excludes these cases.*

If Assumption 3.1 holds, then an application of the direct method of the calculus of variations ensures that (3.1) has a solution, which is unique since $\beta > 0$, cf. [18, Th. 1.43]. Moreover, the Banach-Nečas-Babuška theorem ensures that the inf-sup

condition holds [15, Lem. 6.94]. We denote the associated inf-sup constant by $C_0(\omega) > 0$, i.e.,

$$C_0(\omega) := \inf_{u \in U\setminus\{0\}} \sup_{v \in U\setminus\{0\}} \frac{|a(u,v) - \imath b(u,v)|}{\|u\|_U \|v\|_U}.$$

Using the inf-sup condition, it is straightforward to show that if $u \in U$ solves (3.3) for a fixed $z \in Z$, then

$$\|u\|_U \leq C_0(\omega)^{-1} \|\ell(z;\cdot)\|_{U^*} \leq \tilde{C}(\omega)\|z\|_Z, \tag{3.5}$$

where $\tilde{C}(\omega)$ is proportional to $C_0(\omega)^{-1}$. We note that the stability estimate (3.5) can be further refined under additional geometric assumptions, cf. [14, 17].

Remark 3.3 (Existence and Stability of Solutions when $\Gamma_2 = \partial\Omega$)**.** *When $\Gamma_1 = \Gamma_3 = \emptyset$, $\gamma \equiv 0$, $\delta \equiv 1$ and $c \equiv c_0$ is constant, the variational problem,*

$$\textit{find } u \in U \textit{ such that} \quad a(u,v) - \imath b(u,v) = f(v) \quad \forall\, v \in U,$$

has a unique solution for all $f \in U^*$ that satisfies the stability bound

$$\|\nabla u\|_{L^2(\Omega)} + \omega \|u\|_{L^2(\Omega)} \leq C(\omega)\|f\|_{U^*}, \tag{3.6}$$

where $C(\omega)$ is a positive constant that depends on the wavenumber [24, Prop. 8.1.3]. In addition, if Ω is either star-shaped with a smooth boundary or convex, then $C(\omega) \equiv C$ is independent of ω and the stability constant $\tilde{C}(\omega)$ in (3.5) is bounded from above by a constant times $\min\{1,\omega\}^{-1}$. In this case, we have the following bound on the $L^2(\Omega)$-norm of u,

$$\|u\|_{L^2(\Omega)} \leq \frac{C}{\omega}\|f\|_{U^*}.$$

3.1. Discretization. We discretize (3.3) using a Galerkin method. Let U_m denote a finite-dimensional subspace of U with real-valued basis $\{\varphi_1, \ldots, \varphi_m\} \subset U_m$. We approximate $u \in U$ as a linear combination of the basis functions $\{\varphi_i\}$ with scalar coefficients $\mathbf{u} = (\mathbf{u}_1, \ldots, \mathbf{u}_m)^\top \in \mathbb{C}^m$. Similarly, let Z_n be a finite-dimensional subspace of Z with real-valued basis $\{\psi_1, \ldots, \psi_n\} \subset Z_n$. We discretize the control variable z as a linear combination of the basis functions $\{\psi_i\}$ with scalar coefficients $\mathbf{z} = (\mathbf{z}_1, \ldots, \mathbf{z}_n)^\top \in \mathbb{C}^n$. To simplify the presentation, we introduce the following matrices and vector: $\mathbf{M}$, $\mathbf{A}$, $\mathbf{B} \in \mathbb{R}^{m\times m}$, $\mathbf{C} \in \mathbb{C}^{m\times m}$, $\mathbf{L} \in \mathbb{C}^{n\times m}$, $\mathbf{R} \in \mathbb{R}^{n\times n}$, and $\mathbf{w} \in \mathbb{C}^m$ with respective elements

$$\begin{aligned}
&\mathbf{M}_{ij} = \int_\Omega \varphi_i(x)\varphi_j(x)\,\mathrm{d}x, && i,\, j = 1,\ldots,m,\\
&\mathbf{A}_{ij} = a(\varphi_i,\varphi_j), && i,\, j = 1,\ldots,m,\\
&\mathbf{B}_{ij} = b(\varphi_i,\varphi_j), && i,\, j = 1,\ldots,m,\\
&\mathbf{C}_{ij} = \langle \mathcal{W}(\varphi_i), \overline{\mathcal{W}(\varphi_j)}\rangle_W, && i,\, j = 1,\ldots,m,\\
&\mathbf{L}_{ij} = -\ell(\psi_j,\varphi_i), && i = 1,\ldots,m, \quad j = 1,\ldots,n,\\
&\mathbf{R}_{ij} = \langle \psi_i,\psi_j\rangle_Z, && i,\, j = 1,\ldots,n,\\
&\mathbf{w}_i = \langle \mathcal{W}(\varphi_i), \overline{w}\rangle_W, && i = 1,\ldots,m.
\end{aligned}$$

We note that $\mathbf{A}$ is symmetric indefinite, $\mathbf{B}$ is symmetric positive semidefinite, $\mathbf{M}$ and $\mathbf{R}$ are symmetric positive definite, and $\mathbf{C}$ is Hermitian positive semidefinite. With these matrices and vectors, we define our discretized optimization problem as

$$\min_{\mathbf{u},\mathbf{z}} \ \frac{1}{2}\mathbf{u}^*\mathbf{C}\mathbf{u} - \mathbf{u}^*\mathbf{w} + \frac{\beta}{2}\mathbf{z}^*\mathbf{R}\mathbf{z} \tag{3.7a}$$

$$\text{subject to} \quad (\mathbf{A} - \imath\mathbf{B})\mathbf{u} + \mathbf{L}\mathbf{z} = 0. \tag{3.7b}$$

We denote the discretized Helmholtz operator by $\mathbf{K} := (\mathbf{A} - \imath\mathbf{B}) \in \mathbb{C}^{m\times m}$. The KKT system associated with (3.7) is

$$\begin{pmatrix} \mathbf{C} & 0 & \mathbf{K}^* \\ 0 & \beta\mathbf{R} & \mathbf{L}^* \\ \mathbf{K} & \mathbf{L} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ \mathbf{z} \\ \boldsymbol{\lambda} \end{pmatrix} = \begin{pmatrix} \mathbf{w} \\ 0 \\ 0 \end{pmatrix} \tag{3.8}$$

where $\boldsymbol{\lambda} = (\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_m)^\top \in \mathbb{C}^m$ denote the Lagrange multipliers associated with (3.7b). Here, $\mathbf{L}$ is typically rectangular, $\mathbf{K}$ is indefinite and $\mathbf{C}$ is not necessarily invertible.

Recall that the $L^2(\Omega)$ and Z inner products for elements of U_m and Z_n, respectively, can be evaluated using the mass matrices $\mathbf{M}$ and $\mathbf{R}$. We denote the associated finite-dimensional norms by $\|\cdot\|_{\mathbf{M}}$ and $\|\cdot\|_{\mathbf{R}}$. For the subsequent analysis, we make the following stability assumption on the approximation subspaces U_m and Z_n.

Assumption 3.4 (Stability of Discretization)**.** *The discretized Helmholtz operator* $\mathbf{K}$ *is invertible and therefore* (3.7b) *has a unique solution* $\mathbf{U}\mathbf{z} \in \mathbb{C}^m$ *for all* $\mathbf{z} \in \mathbb{C}^n$*. Moreover, the discrete solution* $\mathbf{U}\mathbf{z}$ *satisfies the stability condition*

$$\|\mathbf{U}\mathbf{z}\|_{\mathbf{M}} \le \tilde{c}(\omega)\|\mathbf{z}\|_{\mathbf{R}} \quad \forall\, \mathbf{z} \in \mathbb{C}^n, \tag{3.9}$$

where $\tilde{c}(\omega)$ *is a positive constant. Note that the solution operator* $\mathbf{U}$ *is the matrix* $\mathbf{U} = -\mathbf{K}^{-1}\mathbf{L}$.

Remark 3.5 (Discretization Independent Stability)**.** *The stability constant* $\tilde{c}(\omega)$ *in* (3.9) *is often independent of discretization parameters such as the mesh size in finite elements. For example, using Lagrange finite elements with mesh size* h *to define* U_m*, one can show that the discretized Helmholtz operator* (3.7b) *is uniformly stable for small enough* h*. That is, there exists* $c_0(\omega) > 0$ *independent of* h *such that*

$$\sup_{w\in U_m\setminus\{0\}} \frac{|a(v,w) - \imath b(v,w)|}{\|w\|_U} \ge c_0(\omega)\|v\|_U \quad \forall\, v \in U_m.$$

See [5] for more details. In this setting, $\tilde{c}(\omega)$ *is proportional to* $c_0(\omega)^{-1}$*. Moreover, in some cases, one can analytically determine the functional dependence of* $c_0(\omega)$ *on the frequency* ω*; see, e.g., [6, 7, 14, 19, 20, 24] for such results.*

4. KKT Preconditioners. In this section, we extend the preconditioners in [29] to accommodate the discretized Helmholtz optimization problem (3.7). All preconditioners described in this section have the form

$$\boldsymbol{\mathcal{P}}_\alpha = \begin{pmatrix} (\mathbf{C} + \alpha\mathbf{M}) & 0 & 0 \\ 0 & \beta\mathbf{R} & 0 \\ 0 & 0 & \mathbf{Q}(\mathbf{C} + \alpha\mathbf{M})^{-1}\mathbf{Q}^* \end{pmatrix}, \tag{4.1}$$

for fixed $\alpha \ge 0$, where $\mathbf{Q} \in \mathbb{C}^{m\times m}$ is an invertible matrix. We require $\alpha > 0$ when $\mathbf{C}$ is singular. Moreover, we denote by $\mathbf{S}_\alpha = (\mathbf{K}(\mathbf{C} + \alpha\mathbf{M})^{-1}\mathbf{K}^* + \frac{1}{\beta}\mathbf{L}\mathbf{R}^{-1}\mathbf{L}^*)$ the Schur

complement associated with the perturbed KKT system matrix

$$\boldsymbol{\mathcal{A}}_\alpha := \begin{pmatrix} \mathbf{C}+\alpha\mathbf{M} & 0 & \mathbf{K}^* \\ 0 & \beta\mathbf{R} & \mathbf{L}^* \\ \mathbf{K} & \mathbf{L} & 0 \end{pmatrix} = \boldsymbol{\mathcal{A}}_0 + \begin{pmatrix} \alpha\mathbf{M} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{4.2}$$

where $\boldsymbol{\mathcal{A}}_0$ is the original KKT system matrix (3.8) and by $\widehat{\mathbf{S}}_\alpha = \mathbf{Q}(\mathbf{C}+\alpha\mathbf{M})^{-1}\mathbf{Q}^*$ the approximate Schur complement. When $\alpha = 0$, we will drop the subscripts. The following proposition shows that the spectrum of the preconditioned KKT matrix in (3.8) is determined by the spectrum of the preconditioned Schur complement. This is a generalization of Proposition 3.2 in [29]. In the subsequent subsections, we will provide bounds on the preconditioned Schur compliment for two choices of $\mathbf{Q}$.

Proposition 4.1. *Let the preconditioner $\boldsymbol{\mathcal{P}}_\alpha$ be given by (4.1) for some invertible matrix $\mathbf{Q} \in \mathbb{C}^{m\times m}$ and $\alpha \geq 0$ and let μ be an eigenvalue of $\boldsymbol{\mathcal{P}}_\alpha^{-1}\boldsymbol{\mathcal{A}}_0$. We require $\alpha > 0$ if $\mathbf{C}$ is singular. Then*

$$\begin{aligned} \textit{either} \quad & \frac{\lambda}{\alpha+\lambda} \leq \mu \leq 1, \\ \textit{or} \quad & \frac{1}{2}(1+\sqrt{1+4\sigma_1}) - \frac{\alpha}{\alpha+\lambda} \leq \mu \leq \frac{1}{2}(1+\sqrt{1+4\sigma_m}), \\ \textit{or} \quad & \frac{1}{2}(1-\sqrt{1+4\sigma_m}) - \frac{\alpha}{\alpha+\lambda} \leq \mu \leq \frac{1}{2}(1-\sqrt{1+4\sigma_1}), \end{aligned} \tag{4.3}$$

where $0 \leq \sigma_1 \leq \cdots \leq \sigma_m$ are the eigenvalues of the preconditioned Schur complement

$$\widehat{\mathbf{S}}_\alpha^{-1}\mathbf{S}_\alpha = (\mathbf{Q}(\mathbf{C}+\alpha\mathbf{M})^{-1}\mathbf{Q}^*)^{-1}(\mathbf{K}(\mathbf{C}+\alpha\mathbf{M})^{-1}\mathbf{K}^* + \tfrac{1}{\beta}\mathbf{L}\mathbf{R}^{-1}\mathbf{L}^*) \tag{4.4}$$

and λ denotes the minimum eigenvalue of $\mathbf{M}^{-1}\mathbf{C}$.

Proof. Since $\mathbf{C}$ is positive semidefinite and $\mathbf{M}$ is positive definite, we have that $(\mathbf{C}+\alpha\mathbf{M})$ is nonsingular if $\alpha > 0$. Moreover, if $\mathbf{C}$ is nonsingular, then $(\mathbf{C}+\alpha\mathbf{M})$ is invertible for any $\alpha \geq 0$. To analyze the preconditioned KKT matrix, we note that

$$\boldsymbol{\mathcal{P}}_\alpha^{-1}\boldsymbol{\mathcal{A}}_0 = \boldsymbol{\mathcal{P}}_\alpha^{-1}\boldsymbol{\mathcal{A}}_\alpha - \boldsymbol{\mathcal{P}}_\alpha^{-1}\begin{pmatrix} \alpha\mathbf{M} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{4.5}$$

The spectrum of the first term on the right hand side of (4.5) satisfies similar bounds as those in Proposition 3.2 of [29] and the same proof mechanics as in [29] apply with little modification. We note that the presence of a rectangular matrix, $\mathbf{L}$, does not alter the original proof, which focuses on square matrices. We refer the reader to [29] for the details. We can rewrite the second term on the right hand side of (4.5) as

$$\boldsymbol{\mathcal{P}}_\alpha^{-1}\begin{pmatrix} \alpha\mathbf{M} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} (\mathbf{I}+\alpha^{-1}\mathbf{M}^{-1}\mathbf{C})^{-1} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

which has eigenvalues 0 and $\alpha/(\hat{\lambda}+\alpha)$ where $\hat{\lambda} \geq 0$ denotes an eigenvalue of $\mathbf{M}^{-1}\mathbf{C}$. The desired bounds follow from Horn's conjecture [21]. That is, the minimum (maximum) eigenvalue of a sum of matrices are always greater (less) than the sum of the respective minimum (maximum) eigenvalues of the summands. □

Remark 4.2 (Singular Observation Matrices)**.** *When* $\mathbf{C}$ *is singular, the bounds in Proposition 4.1 may not be tight. In this case,* λ *in Proposition 4.1 is zero and the first set of bounds on* μ *is the interval* $[0, 1]$*, whereas the shift on the lower limit in the remaining bounds is* $\alpha/(\alpha+\lambda) = 1$*. Therefore, the intervals in* (4.3) *may include zero. We discuss this further in Section 5.3 for the special case of lumped mass matrices.*

4.1. General Case. We consider the case that $\mathbf{C}$, $\mathbf{L}$ and $\mathbf{R}$ are arbitrary. To precondition the resulting KKT system matrix, we follow [29] and choose

$$\mathbf{Q} = \mathbf{K}.$$

The preconditioned Schur complement in this setting is

$$\widehat{\mathbf{S}}_\alpha^{-1}\mathbf{S}_\alpha = \mathbf{I} + \tfrac{1}{\beta}(\mathbf{K}(\mathbf{C}+\alpha\mathbf{M})^{-1}\mathbf{K}^*)^{-1}\mathbf{L}\mathbf{R}^{-1}\mathbf{L}^*. \tag{4.6}$$

Throughout, we assume that $\alpha = 0$ if $\mathbf{C}$ is invertible and $\alpha > 0$ otherwise. The following result provides bounds on the spectrum of the preconditioned Schur complement.

Proposition 4.3. *Let Assumption 3.4 hold. Then the eigenvalues* σ_i*,* $i = 1, \ldots, m$*, of the preconditioned Schur complement satisfy the bounds*

$$1 \le \sigma_i \le 1 + \frac{1}{\beta}(C_{\text{obs}}^2 + \alpha)\tilde{c}(\omega)^2 \quad i = 1, \ldots, m. \tag{4.7}$$

Proof. The second term in the preconditioned Schur complement (4.6) is spectrally equivalent to $(\mathbf{C}+\alpha\mathbf{M})^{\frac{1}{2}}\mathbf{U}\mathbf{R}^{-1}\mathbf{U}^*(\mathbf{C}+\alpha\mathbf{M})^{\frac{1}{2}}$, and therefore we must bound the singular values of $(\mathbf{C}+\alpha\mathbf{M})^{\frac{1}{2}}\mathbf{U}\mathbf{R}^{-\frac{1}{2}}$. Fix $\mathbf{z} \in \mathbb{C}^n$. Then Assumption 3.4 ensures that

$$\|\mathbf{M}^{\frac{1}{2}}\mathbf{U}\mathbf{R}^{-\frac{1}{2}}\mathbf{z}\|_2 = \|\mathbf{U}\mathbf{R}^{-\frac{1}{2}}\mathbf{z}\|_{\mathbf{M}} \le \tilde{c}(\omega)\|\mathbf{R}^{-\frac{1}{2}}\mathbf{z}\|_{\mathbf{R}} = \tilde{c}(\omega)\|\mathbf{z}\|_2. \tag{4.8}$$

Since (4.8) holds for any $\mathbf{z}$, we have that $\|\mathbf{M}^{\frac{1}{2}}\mathbf{U}\mathbf{R}^{-\frac{1}{2}}\|_2 \le \tilde{c}(\omega)$ and the bounds

$$0 \le \sigma_i(\mathbf{M}^{\frac{1}{2}}\mathbf{U}\mathbf{R}^{-\frac{1}{2}}) \;\le\; \tilde{c}(\omega)$$

follow from the properties of the singular values. Additionally, we have that

$$\|(\mathbf{C}+\alpha\mathbf{M})^{\frac{1}{2}}\mathbf{U}\mathbf{R}^{-\frac{1}{2}}\mathbf{z}\|_2^2 = \mathbf{u}^*(\mathbf{C}+\alpha\mathbf{M})\mathbf{u},$$

where $\mathbf{u} = \mathbf{U}\mathbf{R}^{-\frac{1}{2}}\mathbf{z}$. Since $\mathcal{W}$ is a continuous linear operator from $L^2(\Omega)$ into W, we have that the definition of $\mathbf{C}$ ensures

$$\mathbf{u}^*\mathbf{C}\mathbf{u} = \|\mathcal{W}(u)\|_W^2 \le C_{\text{obs}}^2\|u\|_{L^2(\Omega)}^2 = C_{\text{obs}}^2\mathbf{u}^*\mathbf{M}\mathbf{u},$$

where u denotes the expansion in the $\{\varphi_1, \ldots, \varphi_m\}$ basis associated with the coefficient vector $\mathbf{u}$. This combined with (4.8) yields the bounds

$$0 \le \sigma_i((\mathbf{C}+\alpha\mathbf{M})^{\frac{1}{2}}\mathbf{U}\mathbf{R}^{-\frac{1}{2}}) \;\le\; (C_{\text{obs}}^2+\alpha)^{\frac{1}{2}}\tilde{c}(\omega),$$

and the desired bound (4.7) follows. □

Remark 4.4 (Full Observation)**.** *When* $\mathbf{C} = \mathbf{M}$*, we have that* $C_{\text{obs}} = 1$ *and we can set* $\alpha = 0$*. In this setting, the upper bound in* (4.7) *simplifies to*

$$1 + \frac{1}{\beta}\tilde{c}(\omega)^2.$$

Remark 4.5 (β Dependence)**.** *As with the optimal control of the Poisson equation, this choice of preconditioner results in relatively strong β-dependence. In particular, as $\beta \downarrow 0$, the intervals defined in Proposition 4.1 grow.*

Remark 4.6 (Wavenumber Independence)**.** *When U_m is defined using Lagrange finite elements (see Remark 3.5) and when the discrete stability constant $\tilde{c}(\omega)$ behaves like the reciprocal of ω (see [5, 6] for such examples), the bounds in Proposition 4.3 become tighter (i.e., the preconditioned eigenvalues are more clustered) as ω increases. Therefore, as the wavenumber increases and the mesh is refined accordingly, we expect the convergence of the KKT preconditioner to improve, so long as the discretized Helmholtz operator $\mathbf{K}$ is inverted exactly. This is in contrast to solving the forward Helmholtz problem, where preconditioners tend to degrade with increasing wavenumber. Clearly if the stability constant $\tilde{c}(\omega)$ is a wavenumber-independent constant (which is often the case for small ω), then the intervals in Proposition 4.3 remain well-behaved as ω approaches zero. Thus, we would expect the convergence behavior of the KKT preconditioner to be relatively insensitive to discretization and wavenumber changes in the small wavenumber case.*

4.2. Idealized Case. We now consider block diagonal preconditioners for the case where $\mathbf{C} = \mathbf{L} = \mathbf{R} = \mathbf{M}$. This occurs when $W = Z = L^2(\Omega)$, $\mathcal{W} = \mathcal{F}$ are identity operators on $L^2(\Omega)$, and $\mathcal{B}_i \equiv 0$, $i = 1, 2$. This models the idealized situation where observations and controls are located on the entire domain Ω. In this case, we set $\alpha = 0$ and the Schur complement simplifies to $\mathbf{S} = \mathbf{K}\mathbf{M}^{-1}\mathbf{K}^* + \frac{1}{\beta}\mathbf{M}$. To precondition the KKT matrix in (3.8), we now modify our choice of $\mathbf{Q}$ from the general case and choose

$$\mathbf{Q} = \mathbf{K} - \frac{\imath}{\sqrt{\beta}}\mathbf{M}.$$

This choice was originally proposed in [4]. Note that this choice of $\mathbf{Q}$ represents a damping of the original Helmholtz operator. This damping is closely related to the shifted Laplace preconditioners in [11, 12], which are used to facilitate the solution of the forward Helmholtz equation. Unfortunately, this shift introduces imaginary arithmetic into the solution procedure. This is more problematic in the Laplace setting than for Helmholtz applications where many situations will already require imaginary arithmetic to address things like impedance boundary conditions. Expanding the approximate Schur complement $\widehat{\mathbf{S}}$, we see that

$$\widehat{\mathbf{S}} = \mathbf{Q}\mathbf{M}^{-1}\mathbf{Q}^* = \mathbf{K}\mathbf{M}^{-1}\mathbf{K}^* + \frac{\imath}{\sqrt{\beta}}(\mathbf{K} - \mathbf{K}^*) + \frac{1}{\beta}\mathbf{M} = \mathbf{S} + \frac{2}{\sqrt{\beta}}\mathbf{B}. \tag{4.9}$$

That is, $\widehat{\mathbf{S}}$ is a positive semidefinite perturbation of the Schur complement $\mathbf{S}$. In particular, if $\mathbf{B} = 0$ (i.e., $\gamma \equiv 0$ and $\Gamma_2 = \emptyset$), then $\mathbf{K} = \mathbf{A}$ is symmetric and the preconditioner is exact. In this case, Proposition 4.1 ensures that GMRES will converge in at most three iterations. Notice that this result relies on a cancellation of terms associated with $\mathbf{C} = \mathbf{L} = \mathbf{R} = \mathbf{M}$. That is, when $\mathbf{Q} = \mathbf{K}$ the Schur complement and its approximation differ by a $\frac{1}{\beta}\mathbf{M}$ term, which is eliminated by choosing the imaginary shift. It is worth noting that [26, 27] consider a real shift of $\mathbf{K}$, which avoids imaginary arithmetic. The approximate Schur complement is no longer exact in the Hermitian case, but the authors argue that the difference between the Schur complement and its approximation is smaller than when no shift is applied. The main limitation for the shift methods occurs when considering general control

and observation models. For example, in the general control case with $\mathbf{C} = \mathbf{M}$, the Schur compliment is $\mathbf{S} = \mathbf{K}\mathbf{M}^{-1}\mathbf{K}^* + \frac{1}{\beta}\mathbf{L}\mathbf{R}^{-1}\mathbf{L}^*$. Here, the $\frac{1}{\beta}\mathbf{L}\mathbf{R}^{-1}\mathbf{L}^*$ term does not reduce to $\frac{1}{\beta}\mathbf{M}$ and therefore, the approximate Schur complement in (4.9) is unlikely to perform better than the unshifted preconditioner with $\mathbf{Q} = \mathbf{K}$.

When $\mathbf{B}$ is not identically zero, we have the following bounds on the eigenvalues of the preconditioned Schur complement.

Proposition 4.7. *Let Assumption 3.4 hold. Then the eigenvalues σ_i, $i = 1, \ldots, m$, of the preconditioned Schur complement satisfy the bounds*

$$\frac{1 + \beta\tilde{c}(\omega)^{-2}}{1 + \beta\tilde{c}(\omega)^{-2} + 2\sqrt{\beta}\omega c_0^{-1}(\gamma_1 + C_{\Gamma_2}\delta_1)} \leq \sigma_i \leq 1 \quad i = 1, \ldots, m, \tag{4.10}$$

where $C_{\Gamma_2} > 0$ is a positive constant that is independent of ω and β.

Proof. Let $(\sigma, \mathbf{v})$ be an eigenpair of the preconditioned Schur complement $\widehat{\mathbf{S}}^{-1}\mathbf{S}$, then we have that $\sigma > 0$ and

$$(1 - \sigma)\mathbf{S}\mathbf{v} = \sigma\frac{2}{\sqrt{\beta}}\mathbf{B}\mathbf{v}. \tag{4.11}$$

Since $\mathbf{S}$ is positive definite and $\mathbf{B}$ is positive semidefinite, we have that $0 < \sigma \leq 1$ and $\sigma = 1$ if and only if $\mathbf{B}\mathbf{v} = 0$. Consider the case when $\sigma < 1$. Premultiplying (4.11) by $\mathbf{v}^*/(\mathbf{v}^*\mathbf{M}\mathbf{v})$ gives

$$(1 - \sigma)\left(\frac{\mathbf{v}^*\mathbf{K}^*\mathbf{M}^{-1}\mathbf{K}\mathbf{v}}{\mathbf{v}^*\mathbf{M}\mathbf{v}} + \frac{1}{\beta}\right) = \sigma\frac{2}{\sqrt{\beta}}\frac{\mathbf{v}^*\mathbf{B}\mathbf{v}}{\mathbf{v}^*\mathbf{M}\mathbf{v}}. \tag{4.12}$$

Let v be the linear combination of $\{\varphi_1, \ldots, \varphi_m\}$ with coefficients given by the vector $\mathbf{v}$. Then the definition of $\mathbf{B}$ and the assumptions on the problem data (γ, κ, δ) ensure that

$$\begin{aligned}\mathbf{v}^*\mathbf{B}\mathbf{v} = b(v, v) &= \int_\Omega \gamma\kappa v(x)\overline{v}(x)\,\mathrm{d}x + \int_{\Gamma_2} \delta\kappa v(x)\overline{v}(x)\,\mathrm{d}x \\ &\leq \frac{\omega}{c_0}(\gamma_1\|v\|^2_{L^2(\Omega)} + \delta_1\|v\|^2_{L^2(\Gamma_2)}) = \frac{\omega}{c_0}(\gamma_1 + C_{\Gamma_2}\delta_1)\mathbf{v}^*\mathbf{M}\mathbf{v},\end{aligned}$$

where $C_{\Gamma_2} > 0$ denotes the maximum generalized eigenvalue of the pair of matrices $(\mathbf{M}_{\Gamma_2}, \mathbf{M})$ where $\mathbf{M}_{\Gamma_2} \in \mathbb{C}^{m\times m}$ denotes the mass matrix associated with the boundary Γ_2. We now bound the ratio

$$\frac{\mathbf{v}^*\mathbf{K}^*\mathbf{M}^{-1}\mathbf{K}\mathbf{v}}{\mathbf{v}^*\mathbf{M}\mathbf{v}} \tag{4.13}$$

from below. Let $\mathbf{u} = \mathbf{v}/\sqrt{\mathbf{v}^*\mathbf{M}\mathbf{v}}$, then

$$\mathbf{u}^*\mathbf{K}^*\mathbf{M}^{-1}\mathbf{K}\mathbf{u} = \mathbf{u}^*(\mathbf{M}^{-\frac{1}{2}}\mathbf{K})^*(\mathbf{M}^{-\frac{1}{2}}\mathbf{K})\mathbf{u} = \|\mathbf{M}^{-\frac{1}{2}}\mathbf{K}\mathbf{u}\|_2^2.$$

By defining $\mathbf{y} = \mathbf{M}^{-\frac{1}{2}}\mathbf{K}\mathbf{u}$, we see that $\mathbf{u} = \mathbf{K}^{-1}\mathbf{M}^{\frac{1}{2}}\mathbf{y} = -\mathbf{U}\mathbf{M}^{-\frac{1}{2}}\mathbf{y}$ and it follows from Assumption 3.4 that

$$1 = \|\mathbf{u}\|_\mathbf{M} = \|\mathbf{U}\mathbf{M}^{-\frac{1}{2}}\mathbf{y}\|_\mathbf{M} \leq \tilde{c}(\omega)\|\mathbf{M}^{-\frac{1}{2}}\mathbf{y}\|_\mathbf{M} = \tilde{c}(\omega)\|\mathbf{y}\|_2.$$

In particular, we have that

$$\|\mathbf{M}^{-\frac{1}{2}}\mathbf{K}\mathbf{u}\|_2^2 \geq \tilde{c}(\omega)^{-2}.$$

In fact, this holds for all vectors $\mathbf{u} \in \mathbb{C}^m$ satisfying $\mathbf{u}^*\mathbf{M}\mathbf{u} = 1$. As a result, the ratio (4.13) is bounded from below by $\tilde{c}(\omega)^{-2}$. By applying these upper and lower bounds to (4.12), we obtain

$$(1-\sigma)\left(\tilde{c}(\omega)^{-2} + \frac{1}{\beta}\right) \leq \sigma \frac{2}{\sqrt{\beta}} \frac{\omega}{c_0}(\gamma_1 + C_{\Gamma_2}\delta_1) \tag{4.14}$$

and the desired result follows. □

Remark 4.8 (β Independence)**.** *We note that as $\beta \downarrow 0$, the lower bound in* (4.10) *converges to one and as $\beta \uparrow \infty$, the lower bound converges again to one. This suggests that this choice of preconditioner will perform relatively independent of β. In fact, the value $\beta = \tilde{c}(\omega)^2$ minimizes the lower bound as a function of β, giving*

$$\frac{1}{1 + \tilde{c}(\omega)\omega c_0^{-1}(\gamma_1 + C_{\Gamma_2}\delta_1)} \leq \sigma_i \leq 1. \tag{4.15}$$

This minimum lower bound represents the worst case from a preconditioning perspective.

Remark 4.9 (Wavenumber Independence)**.** *When $\tilde{c}(\omega)$ is proportional to ω^{-1} as in [5, 6], the lower bound in* (4.10) *tends to one as $\omega \uparrow \infty$. It is worth emphasizing in this case that when $\beta = \tilde{c}(\omega)^2$, the lower bound* (4.15) *is independent of ω. This yields the largest range of possible eigenvalues σ_i and results in the worst expected numerical performance for the preconditioner.*

Remark 4.10 (Relation to Previous Work)**.** *In [26], the authors analyze the approximate Schur complement $\widehat{\mathbf{S}}$ with*

$$\mathbf{Q} = \mathbf{K} + \frac{1}{\sqrt{\beta}}\mathbf{M}$$

in the context of elliptic and parabolic control, respectively. Unfortunately, this choice of $\mathbf{Q}$ does not apply to Helmholtz owing to the fact that $\mathbf{K}^{-1}\mathbf{M}$ has complex eigenvalues, in general. Even in the case that $\gamma \equiv 0$ and $\Gamma_2 = \emptyset$, the Helmholtz equation is real valued, but the matrix $\mathbf{K}$ is symmetric indefinite.

5. Numerical Results. We perform numerical studies to corroborate the theory. The studies are motivated by direct-field acoustic testing (DFAT). We consider a simplified DFAT problem in two dimensions, where the goal is to generate a specific acoustic pressure w in an observation region Ω_w by applying acoustic control z distributed continuously in control region Ω_z. In other words, we minimize

$$\frac{1}{2}\int_{\Omega_w} (u(x) - w(x))\overline{(u(x) - w(x))}\mathrm{d}x + \frac{\beta}{2}\int_{\Omega_z} z(x)\overline{z(x)}\mathrm{d}x$$

subject to the equations (3.1b)–(3.1e). The desired acoustic pressure has the form of a plane wave, $w(x) = \exp\left(i\kappa\cos(\pi/4)x_1 + i\kappa\sin(\pi/4)x_2\right)$. We study the *general case* supported by the theory in Section 4.1, focusing on disjoint observations and controls, i.e., $\Omega_w \cap \Omega_z = \emptyset$. We also study the *idealized case* supported by the theory in Section 4.2 corresponding to $\Omega_w = \Omega_z = \Omega$ and $\mathbf{C} = \mathbf{L} = \mathbf{R} = \mathbf{M}$, which is more of academic value. The computational domain is a square, defined by $\Omega = (-5, 5)^2$. In the general case, Ω_w is a disc of radius 2 centered at the origin, and Ω_z is an annulus centered at the origin, with inner radius 2.5 and outer radius 2.6, see Figure 5.1.

We discretize (3.1b)–(3.1e) in space using Q1 elements. We use the same discretization for the state and control spaces, in order to include the idealized case in our study. We note that the discrete state and control spaces are typically different in practice, resulting in rectangular $\mathbf{L}$ matrices even if $\Omega_w = \Omega_z = \Omega$. Computational meshes on Ω are generated using Cubit [1] with up to five levels of refinement (including the coarsest mesh, depicted in Figure 5.1), needed to resolve high-frequency waves. The number of elements, denoted by N_e, are $N_e \in \{7\,424, 29\,696, 118\,784, 475\,136, 1\,900\,544\}$. We study (3.1b)–(3.1e) with wavenumbers $\kappa \in \{5, 10, 20, 40, 80\}$, in three configurations based on the damping coefficient and boundary conditions: *undamped/radiating*, where $\gamma = 0$, $\delta = 1$, specifically $\partial_\nu u - \imath\kappa u = 0$ on $\partial\Omega$; *damped/reflective*, where $\gamma = 1$, $\delta = 0$, specifically $\partial_\nu u = 0$ on $\partial\Omega$; and *damped/radiating*, where $\gamma = 1$, $\delta = 1$, with $\partial_\nu u - \imath\kappa u = 0$ on $\partial\Omega$. We note that the Robin condition $\partial_\nu u - \imath\kappa u = 0$ is an approximation of the Sommerfeld radiation condition. The undamped/reflective configuration ($\gamma = 0$, $\delta = 0$) is omitted as it may result in resonance, violating our theoretical assumptions.

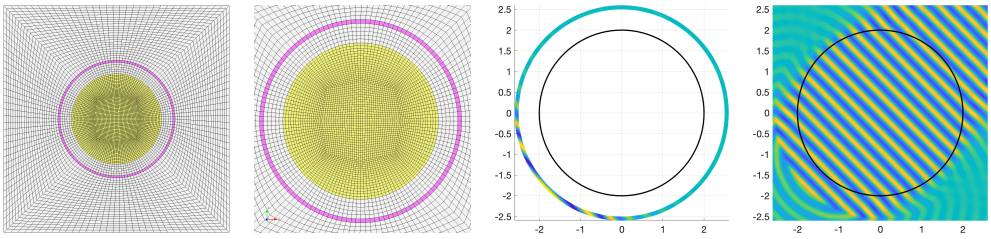

FIG. 5.1. *Left: Computational domain with the observation region Ω_w (yellow) and the control region Ω_z (magenta), meshed at the coarsest level. Center/Left: Zoom of the mesh for the observation and control regions. Center/Right: Real part of the optimal control in the general case for $\kappa = 20, N_e = 475\,136, \beta = 10^{-5}$, and the undamped/absorbing configuration. The boundary of the observation region is shown in black. Right: Zoom of the real part of the optimal state for the given problem parameters. The generated acoustic pressure is a close match to the desired plane wave.*

To solve the KKT systems we use SYMMLQ [25], preconditioned with an inverse of $\boldsymbol{\mathcal{P}}_\alpha$ defined in (4.1). For the general case we set

$$\alpha = \min(10^{-4}, \beta)$$

and $\mathbf{Q} = \mathbf{K}$. For the idealized case, where $\Omega_w = \Omega_z = \Omega$ and $\mathbf{C} = \mathbf{L} = \mathbf{R} = \mathbf{M}$, we use $\alpha = 0$ and $\mathbf{Q} = \mathbf{K} - \frac{\imath}{\sqrt{\beta}}\mathbf{M}$. We performed all studies with flexible GMRES (FGMRES) [31] as well. While this solver reduced the number of iterations by approximately 20% compared to SYMMLQ in some cases, FGMRES could not consistently reach the desired relative residual tolerances due to loss of numerical accuracy for $\beta \leq 10^{-5}$. For this reason, we omit FGMRES results. It is worth noting that for small β the condition number of the unpreconditioned KKT matrix, $\boldsymbol{\mathcal{A}}_0$, is very large. Matlab's `condest` gives a condition number estimate of $1.9 \cdot 10^{10}$ for $\kappa = 5, N_e = 7\,424$, and $\beta = 10^{-6}$ in the undamped/radiating configuration.

We choose the initial guess for the linear solver to be the zero vector. The solver terminates at iterate $\mathbf{x}_\star$ if the relative residual $\|\mathbf{b} - \boldsymbol{\mathcal{A}}_0\mathbf{x}_\star\|_2/\|\mathbf{b}\|_2$ satisfies

$$\|\mathbf{b} - \boldsymbol{\mathcal{A}}_0\mathbf{x}_\star\|_2/\|\mathbf{b}\|_2 \leq \min\{10^{-4}, 10^{-1}\beta\kappa^2/(1+\kappa^2)\},$$

where $\mathbf{b}^* = (\mathbf{w}^*\ 0^*\ 0^*)$. This relative residual tolerance is chosen for two reasons. First, requiring that it be smaller than 10^{-4} is sufficient to match the accuracy of

the studied PDE discretizations given a large β and a small ω (i.e., small κ). Second, for small β and large ω (i.e., large κ) it is also important that the relative residual tolerance depend on the values of β and ω, as follows. Assumption 3.4 ensures that the state equation is uniquely solvable and therefore (3.1) can be tranformed into a strongly convex optimization problem for the control only. We denote the reduced objective function by J. Let $\mathbf{z}_\star$ denote the unique minimizer of J, i.e., $\nabla J(\mathbf{z}_\star) = 0$. Owing to strong convexity and the Cauchy-Schwarz inequality, we have

$$\beta\|\mathbf{z}-\mathbf{z}_\star\|_2^2 \leq \langle \nabla J(\mathbf{z}) - \nabla J(\mathbf{z}_\star), \mathbf{z}-\mathbf{z}_\star\rangle_2 \quad \Longrightarrow \quad \|\mathbf{z}-\mathbf{z}_\star\|_2 \leq \|\nabla J(\mathbf{z})\|_2/\beta.$$

Therefore to obtain a solution $\mathbf{z}$ that satisfies $\|\mathbf{z}-\mathbf{z}_\star\|_2 \leq \epsilon$ for some $\epsilon > 0$, we must choose the gradient tolerance $\|\nabla J(\mathbf{z})\|_2 \leq \epsilon\beta$. The relationship to the residual of the discrete KKT system can be obtained by considering the gradient equation in (3.8). Let $\mathbf{x}^* = (\mathbf{u}^*\ \mathbf{z}^*\ \boldsymbol{\lambda}^*)$, $\mathbf{Uz} = -\mathbf{K}^{-1}\mathbf{Lz}$ be the solution to the state equation, and $\boldsymbol{\Lambda}(\mathbf{z}) = -\mathbf{K}^{-*}(\mathbf{CUz}-\mathbf{w})$ be the solution to the adjoint equation for the control $\mathbf{z}$. By adding and subtracting terms and applying the triangle inequality, we can bound the control error as

$$\begin{aligned}\beta\|\mathbf{z}-\mathbf{z}_\star\|_2 &\leq \|\beta\mathbf{Rz}+\mathbf{L}^*\boldsymbol{\Lambda}(\mathbf{z})\|_2 \leq \|\beta\mathbf{Rz}+\mathbf{L}^*\boldsymbol{\lambda}\|_2 + \|\mathbf{U}^*(\mathbf{K}^*\boldsymbol{\lambda}+\mathbf{CUz}-\mathbf{w})\|_2 \\ &\leq \|\beta\mathbf{Rz}+\mathbf{L}^*\boldsymbol{\lambda}\|_2 + \|\mathbf{U}\|_2\|\mathbf{K}^*\boldsymbol{\lambda}+\mathbf{Cu}-\mathbf{w}\|_2 + \|\mathbf{U}^*\mathbf{CK}^{-1}\|_2\|\mathbf{Ku}+\mathbf{Lz}\|_2.\end{aligned}$$

From this, we see that the error in the controls is further bounded by

$$\|\mathbf{z}-\mathbf{z}_\star\|_2 \leq \frac{1}{\beta}\left(\sqrt{1+\|\mathbf{U}\|_2^2+\|\mathbf{U}^*\mathbf{CK}^{-1}\|_2^2}\right)\|\mathbf{b}-\boldsymbol{\mathcal{A}}_0\mathbf{x}\|_2.$$

Roughly speaking, the constant multiplying the residual in the control error is of order $(1+\tilde{c}(\omega)^2)/\beta$. Therefore, in an attempt to achieve a control error of order ϵ, we can employ the residual tolerance $\|\mathbf{b}-\boldsymbol{\mathcal{A}}_0\mathbf{x}\|_2 \leq \epsilon\beta/(1+\tilde{c}(\omega)^2)$. When $\tilde{c}(\omega)$ is proportional to $1/\omega$ as in Remark 4.6 and since κ is proportional to ω, we obtain the relative residual tolerance $\epsilon\beta\kappa^2/(\kappa^2+1)$. With this tolerance choice, we observe numerically that the relative error in the optimal controls remains roughly constant.

5.1. Results for the General Case. Here we analyze the numerical results corresponding to the general case defined above, i.e., the case of disjoint observations and controls. As an aside, the solution of a sample problem with $\kappa = 20, N_e = 475\,136, \beta = 10^{-5}$, and the undamped/radiating configuration is depicted in Figure 5.1.

Tables 5.1, 5.2 and 5.3 show the iteration counts as reported by Matlab's implementation of SYMMLQ. We note that the number of applications of the KKT operator $\boldsymbol{\mathcal{A}}_0$ and the number of applications of the preconditioner $\boldsymbol{\mathcal{P}}_\alpha^{-1}$ may be slightly higher for each reported iteration count. In all cases studied here, the total number of preconditioner applications equals the iteration count plus two. We make four observations. First, the convergence is relatively mesh independent. Second, there is some β dependence in the convergence results. This dependence is relatively mild for $\beta \geq 10^{-2}$ and more pronounced for $\beta \leq 10^{-5}$. For very small β, we recall that the linear system is extremely ill-conditioned, exacerbating the effects of numerical error in Krylov solvers. We note that the presence of damping nearly halves iteration numbers (compare Table 5.1 with Tables 5.2 and 5.3). Third, there is a clear convergence improvement with increasing ω, see, e.g., the diagonal of the $\beta = 10^{-4}$ chart in Table 5.1. These observations confirm Proposition 4.3, Remark 4.5, and Remark 4.6.

Fourth, we emphasize that the iteration numbers are relatively modest even for very small β. The eigenvalues of the preconditioned operator, $\boldsymbol{\mathcal{P}}_\alpha^{-1}\boldsymbol{\mathcal{A}}_0$, for the smallest

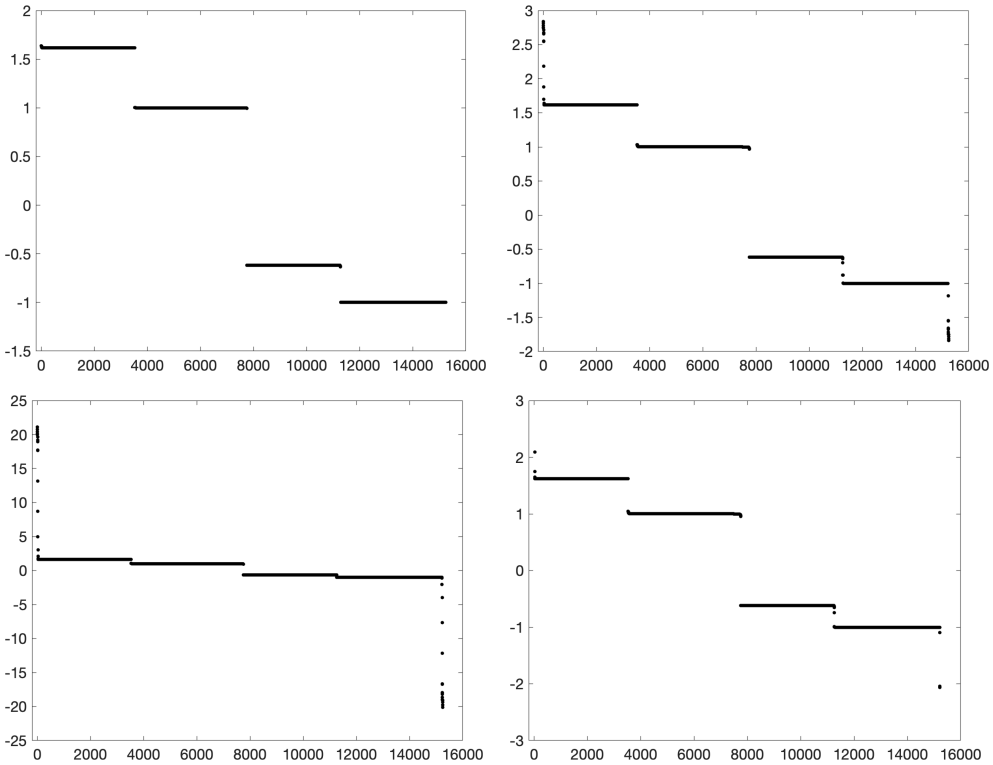

FIG. 5.2. *Top left: Eigenvalues of* $\boldsymbol{\mathcal{P}}_\alpha^{-1}\boldsymbol{\mathcal{A}}_0$*, for the general case in undamped/radiating configuration, with* $\beta = 10^{-1}$*,* $\kappa = 5$*, and* $N_e = 7\,424$*. Top right:* $\beta = 10^{-3}$*. Bottom left:* $\beta = 10^{-5}$*. Bottom right:* $\beta = 10^{-5}$*, zoomed.*

mesh, $N_e = 7\,424$, and smallest wavenumber, $\kappa = 5$, are depicted in Figure 5.2. We observe four clusters of eigenvalues, located away from zero, for any β. Some stretching of the spectrum is evident as β decreases, however the eigenvalues remain far from zero. The smallest absolute value of any eigenvalue is larger than $1/2$. We attempt to explain these results in Section 5.3.

5.2. Results for the Idealized Case. Here we analyze the numerical results corresponding to the idealized case, i.e., $\Omega_w = \Omega_z = \Omega$ and $\mathbf{C} = \mathbf{L} = \mathbf{R} = \mathbf{M}$. We focus on the iteration numbers in Tables 5.4, 5.5 and 5.6, and make four observations. First, solver convergence is relatively mesh independent. Second, the convergence rate dependence on β is mild, and there is a general convergence improvement with very small β. Third, one can see a slow reduction in the number of required iterations as the wavenumber increases for a fixed mesh size and a fixed β when $\beta \geq 10^{-1}$. This is most easily seen within the last row of each table. For some *intermediate* β values, we see some rise in iterations followed by a reduction as the wavenumber increases. This provides evidence that asymptotically convergence rates improve with increasing wavenumber. However, this improvement is not monotone and not yet apparent in the $\beta \leq 10^{-5}$ data. The previous three observations confirm Proposition 4.7, Remark 4.8, and Remark 4.9. Fourth, we note that the presence of damping has a modest negative effect on iteration numbers for small β.

$\beta = 10^{1}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	6				
29 696	6	6			
118 784	7	6	6		
475 136	7	6	6	6	
1 900 544	8	7	6	6	6

$\beta = 10^{0}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	6				
29 696	8	6			
118 784	8	7	6		
475 136	8	8	7	6	
1 900 544	8	8	7	7	6

$\beta = 10^{-1}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	8				
29 696	8	8			
118 784	13	8	8		
475 136	13	8	8	8	
1 900 544	13	11	8	8	7

$\beta = 10^{-2}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	19				
29 696	20	13			
118 784	21	16	8		
475 136	21	17	12	8	
1 900 544	21	17	13	8	8

$\beta = 10^{-3}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	32				
29 696	33	24			
118 784	36	24	17		
475 136	37	24	19	13	
1 900 544	38	28	19	16	11

$\beta = 10^{-4}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	63				
29 696	64	55			
118 784	66	57	40		
475 136	74	53	41	27	
1 900 544	76	60	43	28	19

$\beta = 10^{-5}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	99				
29 696	99	104			
118 784	112	104	88		
475 136	121	110	93	66	
1 900 544	130	116	96	67	44

$\beta = 10^{-6}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	170				
29 696	187	182			
118 784	180	186	167		
475 136	182	201	166	136	
1 900 544	202	208	179	145	109

TABLE 5.1
General case, undamped/radiating configuration: Preconditioned SYMMLQ iteration numbers.

$\beta = 10^{1}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	4				
29 696	4	4			
118 784	6	4	4		
475 136	6	6	4	4	
1 900 544	7	6	6	6	6

$\beta = 10^{0}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	4				
29 696	6	4			
118 784	7	6	4		
475 136	8	7	6	4	
1 900 544	8	8	7	6	6

$\beta = 10^{-1}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	8				
29 696	8	4			
118 784	8	8	6		
475 136	8	8	7	7	
1 900 544	8	8	8	7	7

$\beta = 10^{-2}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	8				
29 696	10	8			
118 784	13	8	8		
475 136	13	8	8	8	
1 900 544	15	12	8	8	8

$\beta = 10^{-3}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	20				
29 696	20	13			
118 784	23	17	11		
475 136	24	17	13	8	
1 900 544	24	17	13	8	8

$\beta = 10^{-4}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	45				
29 696	45	29			
118 784	45	30	21		
475 136	45	31	21	17	
1 900 544	45	34	24	17	13

$\beta = 10^{-5}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	73				
29 696	71	67			
118 784	76	67	44		
475 136	79	65	47	30	
1 900 544	79	65	51	30	23

$\beta = 10^{-6}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	125				
29 696	127	117			
118 784	127	119	95		
475 136	131	121	93	71	
1 900 544	131	121	95	73	49

TABLE 5.2
General case, damped/reflective configuration: Preconditioned SYMMLQ iteration numbers.

$\beta = 10^{1}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	4				
29 696	4	4			
118 784	6	4	4		
475 136	6	6	4	4	
1 900 544	7	6	6	6	6

$\beta = 10^{0}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	4				
29 696	6	4			
118 784	7	6	4		
475 136	8	7	6	4	
1 900 544	8	8	7	6	6

$\beta = 10^{-1}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	8				
29 696	8	4			
118 784	8	8	6		
475 136	8	8	7	7	
1 900 544	8	8	8	7	7

$\beta = 10^{-2}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	8				
29 696	10	8			
118 784	13	8	8		
475 136	13	8	8	8	
1 900 544	15	11	8	8	8

$\beta = 10^{-3}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	18				
29 696	20	13			
118 784	22	17	11		
475 136	23	17	13	8	
1 900 544	24	17	13	8	8

$\beta = 10^{-4}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	43				
29 696	43	26			
118 784	43	28	21		
475 136	43	31	21	17	
1 900 544	43	33	24	17	13

$\beta = 10^{-5}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	69				
29 696	69	63			
118 784	73	61	42		
475 136	73	61	49	28	
1 900 544	75	61	49	30	23

$\beta = 10^{-6}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	115				
29 696	112	107			
118 784	121	108	89		
475 136	123	109	89	69	
1 900 544	123	109	89	73	52

TABLE 5.3
General case, damped/radiating configuration: Preconditioned SYMMLQ iteration numbers.

$\beta = 10^{1}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	10				
29 696	10	10			
118 784	12	10	8		
475 136	12	10	8	8	
1 900 544	12	10	8	8	8

$\beta = 10^{0}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	14				
29 696	14	12			
118 784	14	14	12		
475 136	14	14	12	10	
1 900 544	16	14	12	10	10

$\beta = 10^{-1}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	14				
29 696	14	16			
118 784	14	16	16		
475 136	16	16	16	14	
1 900 544	16	16	16	14	12

$\beta = 10^{-2}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	12				
29 696	12	14			
118 784	12	14	16		
475 136	12	16	16	18	
1 900 544	12	16	18	18	16

$\beta = 10^{-3}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	4				
29 696	4	10			
118 784	4	10	14		
475 136	4	10	14	18	
1 900 544	4	10	14	18	18

$\beta = 10^{-4}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	2				
29 696	2	4			
118 784	2	4	8		
475 136	2	4	8	16	
1 900 544	2	4	8	16	20

$\beta = 10^{-5}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	2				
29 696	2	2			
118 784	2	2	2		
475 136	2	2	2	6	
1 900 544	2	2	2	6	16

$\beta = 10^{-6}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	2				
29 696	2	2			
118 784	2	2	2		
475 136	2	2	2	2	
1 900 544	2	2	2	2	4

TABLE 5.4
Idealized case, undamped/radiating configuration: Preconditioned SYMMLQ iteration numbers.

$\beta = 10^{1}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	10				
29 696	10	8			
118 784	10	8	8		
475 136	10	10	8	8	
1 900 544	12	10	8	8	8

$\beta = 10^{0}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	12				
29 696	14	12			
118 784	14	12	10		
475 136	14	12	10	10	
1 900 544	16	12	12	10	8

$\beta = 10^{-1}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	16				
29 696	18	16			
118 784	18	16	14		
475 136	20	18	14	12	
1 900 544	20	18	14	12	10

$\beta = 10^{-2}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	16				
29 696	16	18			
118 784	18	20	18		
475 136	18	20	20	16	
1 900 544	20	22	20	16	14

$\beta = 10^{-3}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	12				
29 696	12	16			
118 784	12	16	20		
475 136	14	16	20	20	
1 900 544	14	18	22	22	20

$\beta = 10^{-4}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	10				
29 696	10	12			
118 784	10	12	16		
475 136	12	14	16	20	
1 900 544	12	14	18	22	24

$\beta = 10^{-5}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	8				
29 696	10	10			
118 784	10	10	12		
475 136	10	12	14	16	
1 900 544	10	12	14	16	20

$\beta = 10^{-6}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	8				
29 696	8	10			
118 784	8	10	10		
475 136	8	10	12	14	
1 900 544	10	10	12	14	16

TABLE 5.5
Idealized case, damped/reflective configuration: Preconditioned SYMMLQ iteration numbers.

$\beta = 10^{1}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	8				
29 696	10	8			
118 784	10	8	8		
475 136	10	8	8	8	
1 900 544	10	10	8	8	6

$\beta = 10^{0}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	12				
29 696	12	10			
118 784	14	12	10		
475 136	14	12	10	8	
1 900 544	14	12	10	8	8

$\beta = 10^{-1}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	16				
29 696	16	14			
118 784	18	16	12		
475 136	18	16	14	12	
1 900 544	20	18	14	12	10

$\beta = 10^{-2}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	16				
29 696	18	18			
118 784	18	18	18		
475 136	20	20	18	14	
1 900 544	20	22	20	16	14

$\beta = 10^{-3}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	12				
29 696	14	16			
118 784	14	18	20		
475 136	16	18	20	20	
1 900 544	16	20	22	22	18

$\beta = 10^{-4}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	10				
29 696	10	12			
118 784	10	12	18		
475 136	12	14	18	22	
1 900 544	12	14	20	24	24

$\beta = 10^{-5}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	8				
29 696	10	10			
118 784	10	10	12		
475 136	10	12	14	18	
1 900 544	10	12	14	18	24

$\beta = 10^{-6}$

$N_e \backslash \kappa$	5	10	20	40	80
7 424	8				
29 696	8	10			
118 784	8	10	10		
475 136	8	10	12	14	
1 900 544	10	10	12	14	16

TABLE 5.6
Idealized case, damped/radiating configuration: Preconditioned SYMMLQ iteration numbers.

5.3. Discussion of Disjoint Observations/Controls with Mass Lumping. As noted in Remark 4.5, convergence rates may degrade with decreasing β. To understand why this degradation may not be excessive when the observation and control regions are disjoint, i.e., $\Omega_w \cap \Omega_z = \emptyset$, we refine the analysis presented in [29].

Consider the preconditioned linear system $\boldsymbol{\mathcal{P}}_\alpha^{-1}\boldsymbol{\mathcal{A}}_\alpha$ where $\boldsymbol{\mathcal{P}}_\alpha$ is defined by (4.1) with $\mathbf{Q} = \mathbf{K}$. Recall, the α perturbation (as suggested in [26]) avoids the singularity of $\mathbf{C}$. Thus, the preconditioner can be applied and Remark 4.5 holds. As discussed in the proof of Proposition 4.1, the eigenvalues of this preconditioned system lie in three clusters: $\mu = 1$ and μ near $(1 \pm \sqrt{5})/2$. Our goal in this discussion is to gain further insight regarding the distribution of eigenvalues within these clusters. To achieve this, we abuse notation slightly and redefine the previous mass matrix definitions ($\mathbf{C}$, $\mathbf{R}$, and $\mathbf{M}$) to simplify the analysis that follows. In particular, all mass matrices are now assumed to be diagonal (e.g., lumping is used). Further, the only nonzeros in $\mathbf{M}$ correspond to $\mathbf{C}$ entries that are zero. That is, perturbation terms associated with $\alpha\mathbf{M}$ are zero for rows in $\mathbf{C}$ that correspond to observation variables. This change in the perturbation definition should be negligeable as these α-terms would have been quite small relative to the associated nonzero $\mathbf{C}$ entries. Additionally, we assume that $\mathbf{L}^\top = \begin{pmatrix} 0 & \mathbf{R}^\top \end{pmatrix}$. That is, the same discretization is used for the state and control space in that the basis functions for control are a subset of those used for the state discretization. Following [29], we investigate the eigenvalues of $\boldsymbol{\mathcal{P}}_\alpha^{-\frac{1}{2}}\boldsymbol{\mathcal{A}}_\alpha\boldsymbol{\mathcal{P}}_\alpha^{-\frac{1}{2}}$ given by

$$\begin{pmatrix} \mathbf{I} & 0 & \widehat{\mathbf{K}}^* \\ 0 & \mathbf{I} & \widehat{\mathbf{L}}^* \\ \widehat{\mathbf{K}} & \widehat{\mathbf{L}} & 0 \end{pmatrix} \quad \text{where} \quad \begin{cases} \widehat{\mathbf{K}} = \widehat{\mathbf{S}}^{-\frac{1}{2}}\mathbf{K}(\mathbf{C} + \alpha\mathbf{M})^{-\frac{1}{2}} \\ \widehat{\mathbf{L}} = \frac{1}{\sqrt{\beta}}\widehat{\mathbf{S}}^{-\frac{1}{2}}\mathbf{L}\mathbf{R}^{-\frac{1}{2}} \end{cases}.$$

It is easy to see that any eigenvector of the form $\mathbf{v}^* = (\mathbf{v}_\mathbf{u}^* \ \mathbf{v}_\mathbf{z}^* \ \mathbf{v}_{\boldsymbol{\lambda}}^*)$ corresponds to an eigenvalue μ equal to one when $\mathbf{v}_{\boldsymbol{\lambda}} = 0$ and $\widehat{\mathbf{K}}\mathbf{v}_\mathbf{u} + \widehat{\mathbf{L}}\mathbf{v}_\mathbf{z} = 0$. As $(\widehat{\mathbf{K}} \ \ \widehat{\mathbf{L}})$ has dimensions $m \times (m+n)$, there are generally n linearly independent such vectors.

To analyze $\mu \neq 1$, we first expand $\boldsymbol{\mathcal{P}}_\alpha^{-\frac{1}{2}}\boldsymbol{\mathcal{A}}_\alpha\boldsymbol{\mathcal{P}}_\alpha^{-\frac{1}{2}}\mathbf{v} = \mu\mathbf{v}$ as

$$\mathbf{v}_\mathbf{u} + \widehat{\mathbf{K}}^*\mathbf{v}_{\boldsymbol{\lambda}} = \mu\mathbf{v}_\mathbf{u}, \quad \mathbf{v}_\mathbf{z} + \widehat{\mathbf{L}}^*\mathbf{v}_{\boldsymbol{\lambda}} = \mu\mathbf{v}_\mathbf{z}, \quad \text{and} \quad \widehat{\mathbf{K}}\mathbf{v}_\mathbf{u} + \widehat{\mathbf{L}}\mathbf{v}_\mathbf{z} = \mu\mathbf{v}_{\boldsymbol{\lambda}}.$$

Eliminating $\mathbf{v}_\mathbf{u}$ and $\mathbf{v}_\mathbf{z}$ from the third equation yields

$$\widehat{\mathbf{K}}\widehat{\mathbf{K}}^*\mathbf{v}_{\boldsymbol{\lambda}} + \widehat{\mathbf{L}}\widehat{\mathbf{L}}^*\mathbf{v}_{\boldsymbol{\lambda}} = \mu(\mu-1)\mathbf{v}_{\boldsymbol{\lambda}} \tag{5.1}$$

$$\mathbf{v}_{\boldsymbol{\lambda}} + \frac{1}{\beta}\widehat{\mathbf{S}}^{-\frac{1}{2}}\mathbf{L}\mathbf{R}^{-1}\mathbf{L}^*\widehat{\mathbf{S}}^{-\frac{1}{2}}\mathbf{v}_{\boldsymbol{\lambda}} = \mu(\mu-1)\mathbf{v}_{\boldsymbol{\lambda}}, \tag{5.2}$$

which has been multiplied by $\mu - 1$. Pre-multiplication by $\mathbf{v}_{\boldsymbol{\lambda}}^*$ gives a quadratic equation and leads to the bounds in Section 4.1, or similarly in [29].

Additional insight is revealed by examining $\frac{1}{\beta}\widehat{\mathbf{S}}^{-\frac{1}{2}}\mathbf{L}\mathbf{R}^{-1}\mathbf{L}^*\widehat{\mathbf{S}}^{-\frac{1}{2}}$. As its rank is n, there are $m - n$ linearly independent $\mathbf{v}_{\boldsymbol{\lambda}}$ where this term is zero, implying $m - n$ eigenvalues are identically equal to $(1 + \sqrt{5})/2$ and another $m - n$ eigenvalues are equal to $(1 - \sqrt{5})/2$. Now, pre-multiply (5.2) by $\widehat{\mathbf{S}}^{\frac{1}{2}}$ to obtain

$$\mathbf{T}\bar{\mathbf{v}}_{\boldsymbol{\lambda}} = \widehat{\mu}\ \widehat{\mathbf{S}}\bar{\mathbf{v}}_{\boldsymbol{\lambda}} \tag{5.3}$$

where $\bar{\mathbf{v}}_{\boldsymbol{\lambda}} = \widehat{\mathbf{S}}^{-\frac{1}{2}}\mathbf{v}_{\boldsymbol{\lambda}}$, $\mathbf{T} = \frac{1}{\beta}\mathbf{L}\mathbf{R}^{-1}\mathbf{L}^*$, and $\widehat{\mu} = \mu^2 - \mu - 1$. Computationally, we observe that most non-zero $\widehat{\mu}$ are relatively small, and so most of the eigenvalues μ

are clustered near $(1 \pm \sqrt{5})/2$. Here, we provide some imprecise arguments as to why most $\widehat{\mu}$ may be small. Apply a simple transformation to (5.3) such that

$$(\mathbf{C}+\alpha\mathbf{M})^{\frac{1}{2}}\mathbf{T}(\mathbf{C}+\alpha\mathbf{M})^{\frac{1}{2}}\widehat{\mathbf{v}}_{\boldsymbol{\lambda}} = \widehat{\mu}(\mathbf{C}+\alpha\mathbf{M})^{\frac{1}{2}}\mathbf{K}(\mathbf{C}+\alpha\mathbf{M})^{-1}\mathbf{K}^{*}(\mathbf{C}+\alpha\mathbf{M})^{\frac{1}{2}}\widehat{\mathbf{v}}_{\boldsymbol{\lambda}}$$

where $\bar{\mathbf{v}}_{\boldsymbol{\lambda}} = (\mathbf{C}+\alpha\mathbf{M})^{\frac{1}{2}}\widehat{\mathbf{v}}_{\boldsymbol{\lambda}}$ and we have used the definition of $\widehat{\mathbf{S}}$. If we assume that the restriction of $\mathbf{M}$ to the control region coincides with $\mathbf{R}$, then the left hand side matrix is diagonal with n nonzeros equal to α/β. If the right hand side matrix contained no α-dependence, then the eigenvalues would be proportional to α/β and so a small α could mitigate a small β. Of course, the right hand side matrix depends on α, but we make the following observations that give some expectation that this dependence may be weak: (1) the eigenvalues of $(\mathbf{C}+\alpha\mathbf{M})^{\frac{1}{2}}\mathbf{K}(\mathbf{C}+\alpha\mathbf{M})^{-\frac{1}{2}}$ are independent of α; (2) the entries of $(\mathbf{C}+\alpha\mathbf{M})^{\frac{1}{2}}\mathbf{K}(\mathbf{C}+\alpha\mathbf{M})^{-\frac{1}{2}}$ only differ from those of $\mathbf{K}$ near the observation interface; (3) any $\widehat{\mathbf{v}}_{\boldsymbol{\lambda}}$ must essentially satisfy the homogeneous PDE associated with $(\mathbf{C}+\alpha\mathbf{M})^{\frac{1}{2}}\widehat{\mathbf{S}}(\mathbf{C}+\alpha\mathbf{M})^{\frac{1}{2}}$ away from the control region; (4) any $\widehat{\mathbf{v}}_{\boldsymbol{\lambda}}$ must resemble an eigenfunction of the PDE associated with $\widehat{\mathbf{S}}$ within the interior of the control region. For example, if the eigenvectors of $\widehat{\mathbf{S}}$ are sine functions, then the $\widehat{\mathbf{v}}_{\boldsymbol{\lambda}}$ will resemble different sine functions within the different regions. While the interface conditions affect the frequency of these eigenfunctions and how they are stitched together, α only affects the observation interface while it is the eigenfunctions within the control region that determine $\widehat{\mu}$ (where the left hand side matrix is nonzero). The influence of the observation interface on these control region eigen-modes depends on many complex factors, but it is intuitive that this influence will be less pronounced when the observation and control regions are distant from each other.

The main disadvantage with the previous discussion is that the optimization problem is perturbed. Instead, we are interested in how the preconditioner performs on the original matrix. Consider

$$\boldsymbol{\mathcal{A}}_0 = \begin{pmatrix} 0 & 0 & 0 & \mathbf{K}_{\emptyset}^{*} \\ 0 & \mathbf{C}_{\mathrm{o}} & 0 & \mathbf{K}_{\mathrm{o}}^{*} \\ 0 & 0 & \beta\mathbf{R} & \mathbf{L}^{*} \\ \mathbf{K}_{\emptyset} & \mathbf{K}_{\mathrm{o}} & \mathbf{L} & 0 \end{pmatrix} \quad , \quad \boldsymbol{\mathcal{P}}_{\alpha} = \begin{pmatrix} \alpha\mathbf{M}_{\emptyset} & 0 & 0 & 0 \\ 0 & \mathbf{C}_{\mathrm{o}} & 0 & 0 \\ 0 & 0 & \beta\mathbf{R} & 0 \\ 0 & 0 & 0 & \widehat{\mathbf{S}} \end{pmatrix},$$

and

$$\widehat{\mathbf{S}} = \alpha\mathbf{K}_{\emptyset}\mathbf{M}_{\emptyset}^{-1}\mathbf{K}_{\emptyset}^{*} + \mathbf{K}_{\mathrm{o}}\mathbf{C}_{\mathrm{o}}^{-1}\mathbf{K}_{\mathrm{o}}^{*}.$$

Here, we adapt the notation to investigate the splitting of $\mathbf{K}$ into $\mathbf{K}_{\mathrm{o}}$ and $\mathbf{K}_{\emptyset}$, induced by the degree-of-freedom partitioning based on the observation region and its complement, respectively. This also induces a splitting of $\mathbf{C}+\alpha\mathbf{M}$ into a $\mathbf{C}_{\mathrm{o}}$ part and an $\alpha\mathbf{M}_{\emptyset}$ part. The preconditioned system is then

$$\boldsymbol{\mathcal{P}}_{\alpha}^{-\frac{1}{2}}\boldsymbol{\mathcal{A}}_0\boldsymbol{\mathcal{P}}_{\alpha}^{-\frac{1}{2}} = \begin{pmatrix} 0 & 0 & 0 & \widehat{\mathbf{K}}_{\emptyset}^{*} \\ 0 & \mathbf{I} & 0 & \widehat{\mathbf{K}}_{\mathrm{o}}^{*} \\ 0 & 0 & \mathbf{I} & \widehat{\mathbf{L}}^{*} \\ \widehat{\mathbf{K}}_{\emptyset} & \widehat{\mathbf{K}}_{\mathrm{o}} & \widehat{\mathbf{L}} & 0 \end{pmatrix}$$

where

$$\widehat{\mathbf{K}}_{\emptyset} = \frac{1}{\sqrt{\alpha}}\widehat{\mathbf{S}}^{-\frac{1}{2}}\mathbf{K}_{\emptyset}\mathbf{M}_{\emptyset}^{-\frac{1}{2}} \quad , \quad \widehat{\mathbf{K}}_{\mathrm{o}} = \widehat{\mathbf{S}}^{-\frac{1}{2}}\mathbf{K}_{\mathrm{o}}\mathbf{C}_{\mathrm{o}}^{-\frac{1}{2}} \quad , \quad \widehat{\mathbf{L}} = \frac{1}{\sqrt{\beta}}\widehat{\mathbf{S}}^{-\frac{1}{2}}\mathbf{L}\mathbf{R}^{-\frac{1}{2}}.$$

It follows that

$$\widehat{\mathbf{K}}_{\text{ø}}\widehat{\mathbf{K}}_{\text{ø}}^{*} + \widehat{\mathbf{K}}_{\text{o}}\widehat{\mathbf{K}}_{\text{o}}^{*} = \mathbf{I}. \tag{5.4}$$

Consider a generic eigenvector with the form $\mathbf{v}^{*} = (\mathbf{v}_{\text{ø}}^{*}\ \mathbf{v}_{\text{o}}^{*}\ \mathbf{v}_{\mathbf{z}}^{*}\ \mathbf{v}_{\boldsymbol{\lambda}}^{*})$, where $\mathbf{v}_{\mathbf{u}}$ is now split into two subvectors $\mathbf{v}_{\text{ø}}^{*}$ and $\mathbf{v}_{\text{o}}^{*}$ induced by the domain partitioning into the observation region and its complement. It is easy to show that there are $n_{\text{ø}} - n$ eigenvalues equal to one with eigenvectors satisfying $\mathbf{v}_{\text{o}} = \mathbf{v}_{\mathbf{z}} = 0$, $\widehat{\mathbf{K}}_{\text{o}}^{*}\mathbf{v}_{\boldsymbol{\lambda}} = 0$, $\widehat{\mathbf{L}}^{*}\mathbf{v}_{\boldsymbol{\lambda}} = 0$, and $\mathbf{v}_{\text{ø}} = \widehat{\mathbf{K}}_{\text{ø}}^{*}\mathbf{v}_{\boldsymbol{\lambda}}$. This follows from (5.4), which implies that $\mathbf{v}_{\boldsymbol{\lambda}} = \widehat{\mathbf{K}}_{\text{ø}}\widehat{\mathbf{K}}_{\text{ø}}^{*}\mathbf{v}_{\boldsymbol{\lambda}}$ when $\widehat{\mathbf{K}}_{\text{o}}^{*}\mathbf{v}_{\boldsymbol{\lambda}} = 0$. There should be at least $n_{\text{ø}} - n$ such vectors based on the dimensions of $\widehat{\mathbf{K}}_{\text{o}}^{*}$ and $\widehat{\mathbf{L}}^{*}$. Here, the total number of observation (and non-observation) variables is n_{o} ($n_{\text{ø}}$) and $n \leq n_{\text{ø}}$ follows from the control region being disjoint from the observation region. Further, it is trivial to verify that there are $n_{\text{ø}} - n$ eigenvalues equal to minus one where the associated eigenvectors are identical to those of the plus one eigenvalues with the exception that the fourth component is changed from $\mathbf{v}_{\boldsymbol{\lambda}}$ to $-\mathbf{v}_{\boldsymbol{\lambda}}$.

Now consider an eigenvalue, μ, not equal to $1, -1$, or 0. Reconsidering our generic eigenvector, it follows that

$$\mathbf{v}_{\text{ø}} = \frac{1}{\mu}\widehat{\mathbf{K}}_{\text{ø}}^{*}\mathbf{v}_{\boldsymbol{\lambda}}, \quad \mathbf{v}_{\text{o}} = \frac{1}{\mu - 1}\widehat{\mathbf{K}}_{\text{o}}^{*}\mathbf{v}_{\boldsymbol{\lambda}}, \quad \mathbf{v}_{\mathbf{z}} = \frac{1}{\mu - 1}\widehat{\mathbf{L}}^{*}\mathbf{v}_{\boldsymbol{\lambda}},$$

and

$$\widehat{\mathbf{K}}_{\text{ø}}\mathbf{v}_{\text{ø}} + \widehat{\mathbf{K}}_{\text{o}}\mathbf{v}_{\text{o}} + \widehat{\mathbf{L}}\mathbf{v}_{\mathbf{z}} = \mu\mathbf{v}_{\boldsymbol{\lambda}}.$$

Substituting the first expressions into the last one and multiplying by $\mu(\mu - 1)$, yields

$$((\mu - 1)\widehat{\mathbf{K}}_{\text{ø}}\widehat{\mathbf{K}}_{\text{ø}}^{*} + \mu\widehat{\mathbf{K}}_{\text{o}}\widehat{\mathbf{K}}_{\text{o}}^{*} + \mu\widehat{\mathbf{L}}\widehat{\mathbf{L}}^{*})\mathbf{v}_{\boldsymbol{\lambda}} = \mu^{2}(\mu - 1)\mathbf{v}_{\boldsymbol{\lambda}} \tag{5.5}$$

$$(\mu(\widehat{\mathbf{K}}_{\text{ø}}\widehat{\mathbf{K}}_{\text{ø}}^{*} + \widehat{\mathbf{K}}_{\text{o}}\widehat{\mathbf{K}}_{\text{o}}^{*}) - \widehat{\mathbf{K}}_{\text{ø}}\widehat{\mathbf{K}}_{\text{ø}}^{*} + \mu\widehat{\mathbf{L}}\widehat{\mathbf{L}}^{*})\mathbf{v}_{\boldsymbol{\lambda}} = \mu^{2}(\mu - 1)\mathbf{v}_{\boldsymbol{\lambda}} \tag{5.6}$$

$$(\mu - \widehat{\mathbf{K}}_{\text{ø}}\widehat{\mathbf{K}}_{\text{ø}}^{*} + \mu\widehat{\mathbf{L}}\widehat{\mathbf{L}}^{*})\mathbf{v}_{\boldsymbol{\lambda}} = \mu^{2}(\mu - 1)\mathbf{v}_{\boldsymbol{\lambda}} \tag{5.7}$$

$$(-\frac{1}{\mu}\widehat{\mathbf{K}}_{\text{ø}}\widehat{\mathbf{K}}_{\text{ø}}^{*} + \widehat{\mathbf{L}}\widehat{\mathbf{L}}^{*})\mathbf{v}_{\boldsymbol{\lambda}} = (\mu^{2} - \mu - 1)\mathbf{v}_{\boldsymbol{\lambda}}. \tag{5.8}$$

Starting from (5.7) additionally leads to

$$(\mu - (\mathbf{I} - \widehat{\mathbf{K}}_{\text{o}}\widehat{\mathbf{K}}_{\text{o}}^{*}) + \mu\widehat{\mathbf{L}}\widehat{\mathbf{L}}^{*})\mathbf{v}_{\boldsymbol{\lambda}} = \mu^{2}(\mu - 1)\mathbf{v}_{\boldsymbol{\lambda}} \tag{5.9}$$

$$(\widehat{\mathbf{K}}_{\text{o}}\widehat{\mathbf{K}}_{\text{o}}^{*} + \mu\widehat{\mathbf{L}}\widehat{\mathbf{L}}^{*})\mathbf{v}_{\boldsymbol{\lambda}} = (\mu + 1)(\mu - 1)^{2}\mathbf{v}_{\boldsymbol{\lambda}}. \tag{5.10}$$

There exist n_{o} vectors $\mathbf{v}_{\boldsymbol{\lambda}}$ such that $\widehat{\mathbf{K}}_{\text{ø}}^{*}\mathbf{v}_{\boldsymbol{\lambda}} = 0$ and so for these vectors the first term in (5.8) drops out. When this first term is not present, (5.8) is identical to (5.3) after performing the same $\widehat{\mathbf{S}}^{\frac{1}{2}}$ similarity transformation used to obtain (5.3). Thus, one can conclude that there are n_{o} eigenvalues clustered near $(1 + \sqrt{5})/2$ and another n_{o} eigenvalues clustered near $(1 - \sqrt{5})/2$ when the control and observation regions are disjoint.

Starting instead from (5.10), there exist $n_{\text{ø}} - n$ vectors $\mathbf{v}_{\boldsymbol{\lambda}}$ such that the two left hand side terms drop out, which correspond to the $n_{\text{ø}} - n$ eigenvalues with $\mu = 1$ and with $\mu = -1$ already discussed. There are an additional n vectors such that only the first term is zero but not the second. When this first term is not present, (5.10) is similar to (5.3) (again after a $\widehat{\mathbf{S}}^{\frac{1}{2}}$ transformation) with the exception that there is a μ term on the left side and the polynomial expression on the right side is

different. Applying similar arguments, one can conclude that there are n additional eigenvalues near (but not equal) to minus one and an additional $2n$ eigenvalues near (but not equal) to one. Again, these arguments hinge on showing that the left hand side matrix of (5.3) is relatively *small* compared to the right hand side matrix, which has been observed but not proved rigorously. As there are a total of $2n_{\text{o}} + 2n_{\emptyset} + n$ eigenvalues, all have been taken into account: $n_{\emptyset}-n$ identical to one, $n_{\emptyset}-n$ identical to minus one, $2n$ near one, n near minus one, n_{o} near $(1+\sqrt{5})/2$, and n_{o} near $(1-\sqrt{5})/2$. Notice that these four clusters are as illustrated in Figure 5.2 and that these results do not contradict Proposition 4.3 which provides bounds for three clusters as the third cluster in Proposition 4.3 includes both eigenvalues near minus one and near $(1-\sqrt{5})/2$.

5.4. Reproducibility. The studies are performed in Matlab R2019a, with the matrix blocks generated using Trilinos [16], specifically the PDE-OPT suite for PDE-constrained optimization contained in the Rapid Optimization Library [30]. The results can be reproduced by compiling and executing the code in the

`/packages/rol/example/PDE-OPT/published/Helmholtz_KouriRidzalTuminaro2020`

directory of Trilinos, to generate matrix files, followed by running the Matlab script `study.m` with a chosen set of problem parameters. The mesh files needed for the studies are generated using Cubit. Only the files for the two coarsest meshes are included; additionally, the Cubit script `helmholtz-mesh.jou` is given to generate finer meshes. The file `README.md` in the above directory provides further instructions. The C++ code, the Matlab scripts and the Cubit script have been tested with the SHA-1 tag `9884f67` of the `https://github.com/trilinos/Trilinos` repository (`develop` branch) and Cubit version 15.2.

6. Conclusions. We have extended two KKT preconditioning strategies originally developed for PDE constraints associated with symmetric positive definite operators to the indefinite Helmholtz case. We have shown both theoretically and practically that these strategies are only modestly sensitive to wavenumber changes. In fact, the convergence properties often improve as the wavenumber is increased. In the general case, we have illustrated that our proposed preconditioner is quite effective for a broad range of regularization parameters β, though its convergence behavior degrades for very small values of β. Additionally, in the idealized case, where controls and observations are defined in the entire domain and the controls and states are discretized using the same basis functions, we have further shown that our proposed preconditioner has a very weak dependence on the regularization parameter β.

Acknowledgment. We would like to thank Paul Tsuji (Lawrence Livermore National Lab) and Tim Walsh (Sandia National Labs), who provided valuable insight in the initial technical discussions, including algorithmic aspects and practical experiments. We would also like to thank Susanne Bradley and Chen Greif (University of British Columbia). Susanne pointed us to [4] and discussed the imaginary shift preconditioner for the idealized case in her presentation *Preconditioners for Double Saddle Point Systems* at the Preconditioning 2019 Conference in Minneapolis, MN.

REFERENCES

[1] *CUBIT 15.6 user documentation*, Tech. Rep. SAND2020-94156W, Sandia National Labs, 2020.

[2] E. Auger, L. D'Auria, M. Martini, B. Chouet, and P. Dawson, *Real-time monitoring and massive inversion of source parameters of very long period seismic signals: An application to Stromboli Volcano, Italy*, Geophysical Research Letters, 33 (2006).
[3] A. Bayliss, C. I. Goldstein, and E. Turkel, *An iterative method for the Helmholtz equation*, J. Comput. Phys., 49 (1983), pp. 443–457.
[4] Y. Choi, C. Farhat, W. Murray, and M. Saunders, *A practical factorization of a Schur complement for PDE-constrained distributed optimal control*, Journal of Scientific Computing, 65 (2015), pp. 576–597.
[5] P. Ciarlet, *T-coercivity: Application to the discretization of Helmholtz-like problems*, Computers & Mathematics with Applications, 64 (2012), pp. 22 – 34.
[6] L. Demkowicz, *Asymptotic convergence in finite and boundary element methods: Part 1: Theoretical results*, Computers & Mathematics with Applications, 27 (1994), pp. 69 – 84.
[7] J. Douglas, J. Santos, D. Sheen, and L. Bennethum, *Frequency domain treatment of one-dimensional scalar waves*, Mathematical Models and Methods in Applied Sciences, 3 (1993), pp. 171–194.
[8] O. L. Elvetun and B. F. Nielsen, *PDE-constrained optimization with local control and boundary observations: Robust preconditioners*, SIAM Journal on Scientific Computing, 38 (2016), pp. A3461–A3491.
[9] B. Engquist and L. Ying, *Sweeping preconditioner for the Helmholtz equation: Hierarchical matrix representation*, Comm. Pure Appl. Math., 64 (2011), pp. 697–735.
[10] ———, *Sweeping preconditioner for the Helmholtz equation: Moving perfectly matched layers*, Multiscale Model. Simul., 9 (2011), pp. 686–710.
[11] Y. A. Erlangga, C. Vuik, and C. W. Oosterlee, *On a class of preconditioners for solving the Helmholtz equation*, Applied Numerical Mathematics, 50 (2004), pp. 409–425.
[12] ———, *A comparison of multigrid and incomplete LU shifted Laplace preconditioners for the inhomogeneous Helmholtz equation*, Applied Numerical Mathematics, 56 (2006), pp. 648–666.
[13] O. G. Ernst and M. J. Gander, *Why it is difficult to solve Helmholtz problems with classical iterative methods*, in Numerical Analysis of Multiscale Problems, I. Graham, T. Hou, O. Lakkis, and R. Scheichl, eds., vol. 83, Springer-Verlag, 2011, pp. 325–361.
[14] S. Esterhazy and J. Melenk, *On stability of discretizations of the Helmholtz equation*, in Numerical Analysis of Multiscale Problems, Springer Berlin Heidelberg, 2012, pp. 285–324.
[15] W. Hackbusch, *Elliptic Differential Equations: Theory and Numerical Treatment*, Springer Series in Computational Mathematics, Springer Berlin Heidelberg, 2017.
[16] M. A. Heroux, R. A. Bartlett, V. E. Howle, R. J. Hoekstra, J. J. Hu, T. G. Kolda, R. B. Lehoucq, K. R. Long, R. P. Pawlowski, E. T. Phipps, A. G. Salinger, H. K. Thornquist, R. S. Tuminaro, J. M. Willenbring, A. Williams, and K. S. Stanley, *An overview of the Trilinos project*, ACM Transactions on Mathematical Software, 31 (2005), pp. 397–423.
[17] U. Hetmaniuk, *Stability estimates for a class of Helmholtz problems*, Communications in Mathematical Sciences, 5 (2007), pp. 665–678.
[18] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich, *Optimization with PDE constraints*, vol. 23, Springer Science & Business Media, 2008.
[19] F. Ihlenburg, *Finite element analysis of acoustic scattering*, Springer-Verlag, 1998.
[20] F. Ihlenburg and I. Babuska, *Finite element solution of the Helmholtz equation with high wave number, part I: the h-version of the FEM*, Computers and Mathematics with Applications, 30 (1995), pp. 9–37.
[21] A. Knutson and T. Tao, *Honeycombs and sums of Hermitian matrices*, Notices Amer. Math. Soc, 48 (2001).
[22] P. Larkin, *Developments in Direct-Field Acoustic Testing*, Sound & vibration, 48 (2014), pp. 6–10.
[23] K.-A. Mardal, B. F. Nielsen, and M. Nordaas, *Robust preconditioners for pde-constrained optimization with limited observations*, BIT Numerical Mathematics, 57 (2017), pp. 405–431.
[24] J. M. Melenk, *On Generalized Finite Element Methods*, PhD dissertation, University of Maryland, 1995.
[25] C. C. Paige and M. A. Saunders, *Solution of sparse indefinite systems of linear equations*, SIAM Journal on Numerical Analysis, 12 (1975), pp. 617–629.
[26] J. W. Pearson, M. Stoll, and A. J. Wathen, *Regularization-robust preconditioners for time-dependent PDE-constrained optimization problems*, SIAM Journal on Matrix Analysis and Applications, 33 (2012), pp. 1126–1152.
[27] J. W. Pearson and A. J. Wathen, *A new approximation of the Schur complement in precon-*

ditioners for PDE-constrained optimization, Numerical Linear Algebra with Applications, 19 (2012), pp. 816–829.

[28] J. Poulson, B. Engquist, S. Li, and L. Ying, *A parallel sweeping preconditioner for heterogeneous 3D Helmholtz equations*, ArXiv e-prints, (2012).

[29] T. Reese, H. Dollar, and A. Wathen, *Optimal preconditioners for PDE-constrained optimization*, SIAM Journal on Scientific Computing, 32 (2010), pp. 271–298.

[30] D. Ridzal, D. Kouri, and G. von Winckel, *Rapid Optimization Library.* Trilinos Users' Group Meeting, Sandia National Labs, SAND2017-12025PE, 2017.

[31] Y. Saad, *A flexible inner-outer preconditioned GMRES algorithm*, SIAM Journal on Scientific Computing, 14 (1993), pp. 461–469.