PDE-constrained optimization with local control and boundary observations: Robust preconditioners

Ole Løseth Elvetun*and Bjørn Fredrik Nielsen†

December 5, 2014

Abstract

We consider PDE-constrained optimization problems with control functions defined on a subregion of the domain of the state equation. The main purpose of this paper is to define and analyze robust preconditioners for KKT systems associated with such optimization tasks. That is, preconditioners that lead to iteration bounds, for the MINRES scheme, that are independent of the regularization parameter α and the mesh size h.

Our analysis addresses elliptic control problems, subject to Tikhonov regularization, and covers cases with boundary observations only and locally defined control functions. A number of numerical experiments are presented.

Keywords: PDE-constrained optimization, preconditioning, minimal residual method.

AMS subject classifications: 49K20, 65F08, 65N21, 65F15.

1 Introduction

Parameter robust preconditioners for KKT systems arising in connection with PDE-constrained optimization have been successfully constructed [10, 11, 12, 13]. Nevertheless, these methods typically assume that observation data is available throughout the entire domain of the state equation, and that the control function also is defined on this domain. In [6] it is explained how one may handle problems with boundary observations only. The purpose of this text is to further explore this issue. More specifically, to investigate

*Department of Mathematical Sciences and Technology, Norwegian University of Life Sciences, Norway. Email: `ole.elvetun@nmbu.no`

†Department of Mathematical Sciences and Technology, Norwegian University of Life Sciences, Norway; Simula Research Laboratory; Center for Cardiological Innovation, Oslo University Hospital. Email: `bjorn.f.nielsen@nmbu.no`

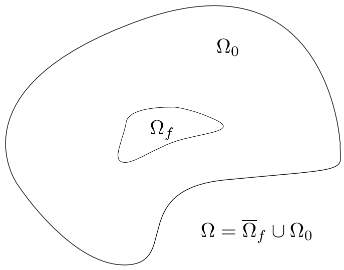


Figure 1: An example of a domain Ω with subdomains Ω_f and Ω_0.

how to design parameter robust preconditioners for problems with locally defined control functions and with boundary observations only.

Our work is motivated by the fact that many inverse problems, arising in the engineering sciences and in medical imaging, involve locally defined controls and limited observation data. This is, for example, the case for the inverse problem of electrocardiography (ECG).

2 Model problem

Consider the problem:

$$\min_{\mathbf{f},u} \left\{ \frac{1}{2} \|u-d\|^2_{L^2(\partial\Omega)} + \frac{1}{2}\alpha \|\mathbf{f}\|^2 \right\} \tag{1}$$

subject to

$$\Delta u - u = \begin{cases} \mathbf{f} & \text{in } \Omega_f, \\ 0 & \text{in } \Omega_0 = \Omega \setminus \overline{\Omega}_f, \end{cases} \tag{2}$$

$$\frac{\partial u}{\partial \mathbf{n}} = 0 \quad \text{on } \partial\Omega, \tag{3}$$

where $\mathbf{n}$ denotes the outward directed normal vector, of unit length, of $\partial\Omega$. We are thus aiming at using a L^2-boundary observation d, of u, to identify a source $\mathbf{f}$ defined on the subregion Ω_f of the domain Ω of the state equation. Note that Ω_0 represents the region $\Omega \setminus \overline{\Omega}_f$, see Figure 1. We assume that Ω_f and Ω are bounded and open domains, with Lipschitz boundaries, and that $\partial\Omega_f \cap \partial\Omega = \emptyset$.

Since the state u belongs to $H^1(\Omega)$, it is natural to seek a control $\mathbf{f}$ in the dual space $H^1(\Omega_f)'$, and the state equations (2)-(3) take the form

$$\int_\Omega \nabla u \cdot \nabla w + uw \, dx = -\langle \mathbf{f}, Rw \rangle \quad \forall w \in H^1(\Omega),$$

where $R : H^1(\Omega) \to H^1(\Omega_f)$ denotes the restriction operator, which, for the sake of simple notation, will be omitted.

Riesz' representation theorem implies that any $\mathbf{f} \in H^1(\Omega_f)'$ can be uniquely represented by a function $f \in H^1(\Omega_f)$. Hence, our optimization problem can be expressed as

$$\min_{f \in H^1(\Omega_f),\, u \in H^1(\Omega)} \left\{ \frac{1}{2} \|u - d\|^2_{L^2(\partial\Omega)} + \frac{1}{2}\alpha \|f\|^2_{H^1(\Omega_f)} \right\} \tag{4}$$

subject to

$$\int_\Omega \nabla u \cdot \nabla w + uw \, dx = -\int_{\Omega_f} \nabla f \cdot \nabla w + fw \, dx \quad \forall w \in H^1(\Omega). \tag{5}$$

Please note that (4)-(5) is similar to the inverse problem of electrocardiography, provided that the ST-shift in the transmembrane potential of the heart is used as the unknown source/control. But the inverse ECG problem involves conductivity tensors and the state equation does not contain any zero order terms.

3 Alternative formulation

We will now show that one can replace the state space $H^1(\Omega)$ in (4)-(5) with a function space consisting of functions satisfying, in a suitable weak sense, $\Delta\phi - \phi = 0$ in Ω_0. In Section 4 we employ this fact, and properly weighted Sobolev norms, to prove that the Brezzi conditions hold with α-independent constants. Thereafter, this insight is used to remove one of the unknowns from the KKT system and to develop parameter robust preconditioners.

From (5) it follows that the solution u of the state equation satisfies

$$(u, \psi)_{H^1(\Omega)} = 0 \quad \forall \psi \in S,$$

where

$$S = \left\{ \psi \in H^1(\Omega) \mid \psi|_{\overline{\Omega}_f} = 0 \right\},$$

i.e.

$$u \in U = S^\perp.$$

We will now briefly argue that $H^1(\Omega)$ in (4)-(5) can be replaced by U.

- Assume that u satisfies (5). Then, we know that $u \in U$ and, since U is a subspace of $H^1(\Omega)$, it follows that

$$(u, w)_{H^1(\Omega)} = -(f, w)_{H^1(\Omega_f)} \quad \forall w \in U. \tag{6}$$

- Let $w \in H^1(\Omega)$ be arbitrary and recall the orthogonal decomposition

$$w = q + q^\perp, \quad q \in S \text{ and } q^\perp \in U.$$

Assume that $u \in U$ satisfies (6). Then,

$$\begin{aligned}(u,w)_{H^1(\Omega)} &= (u,q)_{H^1(\Omega)} + (u,q^\perp)_{H^1(\Omega)} \\ &= (u,q^\perp)_{H^1(\Omega)} \\ &= -(f,q^\perp)_{H^1(\Omega_f)} \\ &= -(f,q^\perp)_{H^1(\Omega_f)} - (f,q)_{H^1(\Omega_f)} \\ &= -(f,w)_{H^1(\Omega_f)},\end{aligned}$$

where the second last equality follows from the fact that $q \in S$, i.e. $q|_{\overline{\Omega}_f} = 0$. Since $w \in H^1(\Omega)$ was arbitrary, we conclude that: If $u \in U$ satisfies (6), then u also satisfies (5).

It follows that we may rephrase (4)-(5) as follows:

$$\min_{f \in H^1(\Omega_f),\, u \in U} \left\{ \frac{1}{2} \|u-d\|^2_{L^2(\partial\Omega)} + \frac{1}{2}\alpha \|f\|^2_{H^1(\Omega_f)} \right\} \tag{7}$$

subject to

$$\int_\Omega \nabla u \cdot \nabla w + uw \, dx = -\int_{\Omega_f} \nabla f \cdot \nabla w + fw \, dx \quad \forall w \in U. \tag{8}$$

3.1 Helmholtz-harmonic extensions

We will now explain why functions in U can be regarded as Helmholtz-harmonic extensions, to the entire domain Ω, of functions defined on Ω_f. Below it will become evident why we use the term "Helmholtz-harmonic".

Let $\phi \in U = S^\perp$ be arbitrary. Then,

$$(\phi,\psi)_{H^1(\Omega)} = 0 \quad \forall \psi \in S,$$

or, since all $\psi \in S$ satisfy $\psi|_{\overline{\Omega}_f} = 0$,

$$\int_{\Omega_0} \nabla\phi \cdot \nabla\psi + \phi\psi \, dx = 0 \quad \forall \psi \in S,$$

where we recall that $\Omega_0 = \Omega \setminus \overline{\Omega}_f$, see Figure 1. The functions in S may be regarded as zero-extensions of functions belonging to $\{q \in H^1(\Omega_0) \mid q|_{\partial\Omega_f} = 0\}$. Thus, provided that the boundaries of Ω_f and Ω are Lipschitz, we may conclude that

$$\tilde{\phi} = \phi|_{\Omega_0}$$

is the weak solution of

$$\Delta\tilde{\phi} - \tilde{\phi} = 0 \quad \text{in } \Omega_0, \tag{9}$$

$$\tilde{\phi} = \phi \quad \text{on } \partial\Omega_f, \tag{10}$$

$$\frac{\partial\tilde{\phi}}{\partial\mathbf{n}} = 0 \quad \text{on } \partial\Omega. \tag{11}$$

We therefore refer to the functions in U as Helmholtz-harmonic on Ω_0.

Standard stability estimates and the trace theorem imply that

$$\|\phi\|_{H^1(\Omega_0)} = \|\tilde{\phi}\|_{H^1(\Omega_0)} \leq c\|\phi\|_{H^{1/2}(\partial\Omega_f)} \leq C\|\phi\|_{H^1(\Omega_f)}. \tag{12}$$

Throughout this text, c and C denote generic positive constants that are independent of the regularization parameter α and the grid parameter h.

Lemma 3.1.1 *There exists a positive constant c such that*

$$\|\phi\|_{H^1(\Omega_f)} \leq \|\phi\|_{H^1(\Omega)} \leq c\|\phi\|_{H^1(\Omega_f)} \quad \forall \phi \in U.$$

Proof

The first inequality follows from the assumption that $\Omega_f \subset \Omega$. The second inequality is a consequence of (12).

■

4 KKT system

The Lagrangian associated with (7)-(8) reads

$$\begin{aligned} \mathcal{L}(f,u,w) \quad &= \quad \left\{ \frac{1}{2}\|u-d\|^2_{L^2(\partial\Omega)} + \frac{1}{2}\alpha\|f\|^2_{H^1(\Omega_f)} \right\} \\ &\quad +(f,w)_{H^1(\Omega_f)} + (u,w)_{H^1(\Omega)}, \end{aligned}$$

with $f \in H^1(\Omega_f)$, $u \in U$ and $w \in U$. And, from the first order optimality conditions

$$\frac{\partial \mathcal{L}}{\partial f} = 0, \ \frac{\partial \mathcal{L}}{\partial u} = 0, \ \frac{\partial \mathcal{L}}{\partial w} = 0,$$

we obtain the optimality system: Determine $(f,u,w) \in H^1(\Omega_f) \times U \times U$ such that

$$\alpha(f,\psi)_{H^1(\Omega_f)} + (\psi,w)_{H^1(\Omega_f)} = 0 \quad \forall \psi \in H^1(\Omega_f), \tag{13}$$

$$(u-d,\phi)_{L^2(\partial\Omega)} + (\phi,w)_{H^1(\Omega)} = 0 \quad \forall \phi \in U, \tag{14}$$

$$(f,\varphi)_{H^1(\Omega_f)} + (u,\varphi)_{H^1(\Omega)} = 0 \quad \forall \varphi \in U. \tag{15}$$

This system can be written on the form

$$\underbrace{\begin{bmatrix} \alpha\tilde{A}_f & 0 & A_f' \\ 0 & M_\partial & A' \\ A_f & A & 0 \end{bmatrix}}_{\mathcal{A}_\alpha^{3\times 3}} \begin{bmatrix} f \\ u \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ \tilde{M}_\partial d \\ 0 \end{bmatrix}, \tag{16}$$

where

$$\begin{aligned}
A &: U \to U', & u &\mapsto (u, \cdot)_{H^1(\Omega)}, & (17)\\
A_f &: H^1(\Omega_f) \to U', & f &\mapsto (f, \cdot)_{H^1(\Omega_f)}, & (18)\\
\tilde{A}_f &: H^1(\Omega_f) \to H^1(\Omega_f)', & f &\mapsto (f, \cdot)_{H^1(\Omega_f)}, & (19)\\
M_\partial &: U \to U', & u &\mapsto (u, \cdot)_{L^2(\partial\Omega)}, & (20)\\
\tilde{M}_\partial &: L^2(\partial\Omega) \to U', & d &\mapsto (d, \cdot)_{L^2(\partial\Omega)}. & (21)
\end{aligned}$$

The notation "′" is used to denote dual operators and dual spaces.

4.1 Weighted norms

For $\alpha > 0$, standard techniques can be employed to show that the Brezzi conditions hold for (16). In the standard L^2- and H^1-norms, however, the constants in the Brezzi conditions depend on α: Typically, the constant appearing in the coercivity condition is of order $O(\alpha)$. Consequently, we can not easily obtain an α-robust preconditioner with these norms. To remedy this, we can follow the procedure in [12] and introduce weighted Hilbert spaces, which are constructed in such a manner that the constants appearing in the Brezzi conditions are independent of the regularization parameter α.

For the control, state and multiplier spaces, we will work with the weighted norms

$$\begin{aligned}
\|f\|_{F_\alpha}^2 &= \alpha \|f\|_{H^1(\Omega_f)}^2,\\
\|u\|_{U_\alpha}^2 &= \alpha \|u\|_{H^1(\Omega)}^2 + \|u\|_{L^2(\partial\Omega)}^2,\\
\|w\|_{U_{\alpha^{-1}}}^2 &= \frac{1}{\alpha} \|w\|_{H^1(\Omega)}^2.
\end{aligned}$$

Note that we must have $\alpha > 0$ for these norms to make sense.

4.2 Inf-sup condition

The weighted norms give us the possibility to consider the operator $\mathcal{A}_\alpha^{3\times 3}$ in (16) as a mapping

$$\mathcal{A}_\alpha^{3\times 3} : F_\alpha \times U_\alpha \times U_{\alpha^{-1}} \to F_\alpha' \times U_\alpha' \times U_{\alpha^{-1}}'.$$

The analysis of saddle-point problems is standard and consists of three steps: Boundedness, coercivity on the kernel of the state equation, and the inf-sup condition. That the first two conditions are fulfilled, with α independent constants, follows from the results published in [9]. Their analysis will therefore be omitted. We are left with proving that the inf-sup condition holds, which we state in the following lemma:

Lemma 4.2.1 *There exists a constant $c > 0$, independent of $\alpha > 0$, such that*

$$\inf_{w \in U_{\alpha^{-1}}} \sup_{(f,u) \in F_\alpha \times U_\alpha} \frac{(f,w)_{H^1(\Omega_f)} + (u,w)_{H^1(\Omega)}}{\|(f,u)\|_{F_\alpha \times U_\alpha} \|w\|_{U_{\alpha^{-1}}}} \geq c.$$

Proof

Let $w \in U_{\alpha^{-1}} \setminus \{0\}$ be arbitrary. With $f = w|_{\Omega_f}$ and $u = 0$ we find that

$$\begin{aligned}
\sup_{(f,u) \in F_\alpha \times U_\alpha} \frac{(f,w)_{H^1(\Omega_f)} + (u,w)_{H^1(\Omega)}}{\|(f,u)\|_{F_\alpha \times U_\alpha} \|w\|_{U_{\alpha^{-1}}}} &\geq \frac{(w,w)_{H^1(\Omega_f)}}{\|w\|_{F_\alpha} \|w\|_{U_{\alpha^{-1}}}} \\
&= \frac{\|w\|^2_{H^1(\Omega_f)}}{\sqrt{\alpha}\|w\|_{H^1(\Omega_f)} \sqrt{\alpha^{-1}} \|w\|_{H^1(\Omega)}} \\
&\geq \frac{1}{\tilde{c}} \frac{\|w\|^2_{H^1(\Omega_f)}}{\|w\|_{H^1(\Omega_f)} \|w\|_{H^1(\Omega_f)}} \\
&= \frac{1}{\tilde{c}} = c,
\end{aligned}$$

where the last inequality follows from Lemma 3.1.1.

■

We conclude that both $\|\mathcal{A}_\alpha^{3\times3}\|$ and $\|[\mathcal{A}_\alpha^{3\times3}]^{-1}\|$ are bounded independently of $\alpha > 0$.

5 Reducing the size of the KKT system

The main reason why the inf-sup condition in Lemma 4.2.1 holds, with an α independent constant, is the fact that $H^1(\Omega_f)$ and U are isomorphic. We will now see how this property also can be used to remove the Lagrange multiplier from the KKT system.

First, however, we will formalize the isomorphism between $H^1(\Omega_f)$ and U. Define the extension operator $E : H^1(\Omega_f) \to U$ as

$$E\phi = \begin{cases} \phi & \text{in } \Omega_f, \\ \tilde{\phi} & \text{in } \Omega_0, \end{cases} \tag{22}$$

where $\tilde{\phi}$ is the weak solution of (9)-(11). From standard theory for elliptic PDEs and Lemma 3.1.1, it follows that this operator is an isomorphism between $H^1(\Omega_f)$ and U.

Since both the state function, u, and the dual function, w, in the KKT system (13)-(15) belong to U, we may express them on the form

$$\begin{aligned} u &= E\hat{u}, \\ w &= E\hat{w}, \end{aligned}$$

where $\hat{u}, \hat{w} \in H^1(\Omega_f)$. Hence, equations (13)-(15) can be reformulated as

$$\begin{aligned} \alpha(f,\psi)_{H^1(\Omega_f)} + (\psi,\hat{w})_{H^1(\Omega_f)} &= 0 \quad \forall \psi \in H^1(\Omega_f), & (23)\\ (E\hat{u}-d, E\phi)_{L^2(\partial\Omega)} + (E\phi, E\hat{w})_{H^1(\Omega)} &= 0 \quad \forall \phi \in H^1(\Omega_f), & (24)\\ (f,\varphi)_{H^1(\Omega_f)} + (E\hat{u}, E\varphi)_{H^1(\Omega)} &= 0 \quad \forall \varphi \in H^1(\Omega_f). & (25) \end{aligned}$$

On this form, the relation between the control, f, and the dual, $\hat{w}$, becomes clear. In fact, from (23) it follows that

$$\hat{w} = -\alpha f.$$

Consequently, we can replace $\hat{w}$ with $-\alpha f$ in (24) to obtain the equations

$$\begin{aligned} -\frac{1}{\alpha}(E\hat{u}-d, E\phi)_{L^2(\partial\Omega)} + (Ef, E\phi)_{H^1(\Omega)} &= 0 \quad \forall \phi \in H^1(\Omega_f), & (26)\\ (E\hat{u}, E\varphi)_{H^1(\Omega)} + (f,\varphi)_{H^1(\Omega_f)} &= 0 \quad \forall \varphi \in H^1(\Omega_f). & (27) \end{aligned}$$

We can then write this system on the block form

$$\underbrace{\begin{bmatrix} -\frac{1}{\alpha}\mathsf{M}_\partial & \mathsf{A} \\ \mathsf{A} & \mathsf{A}_f \end{bmatrix}}_{\mathcal{A}_\alpha^{2\times 2}} \begin{bmatrix} \hat{u} \\ f \end{bmatrix} = \begin{bmatrix} -\frac{1}{\alpha}\tilde{\mathsf{M}}_\partial d \\ 0 \end{bmatrix}, \tag{28}$$

where

$$\begin{aligned} \mathsf{M}_\partial &: H^1(\Omega_f) \to H^1(\Omega_f)', & \hat{u} &\mapsto (E\hat{u}, E\,\cdot\,)_{L^2(\partial\Omega)}, & (29)\\ \tilde{\mathsf{M}}_\partial &: L^2(\partial\Omega) \to H^1(\Omega_f)', & d &\mapsto (d, E\,\cdot\,)_{L^2(\partial\Omega)}, & (30)\\ \mathsf{A} &: H^1(\Omega_f) \to H^1(\Omega_f)', & \hat{u} &\mapsto (E\hat{u}, E\,\cdot\,)_{H^1(\Omega)}, & (31)\\ \mathsf{A}_f &: H^1(\Omega_f) \to H^1(\Omega_f)', & f &\mapsto (f, \,\cdot\,)_{H^1(\Omega_f)}. & (32) \end{aligned}$$

We conclude that (28) has a unique solution since (16) has a unique solution and the extension operator $E : H^1(\Omega_f) \to U$ is isomorphic. (Also for other KKT-systems, arising in connection with PDE-constrained optimization, it is sometimes possible to reduce the problem to a 2×2 block system, see e.g. [13, 4].)

6 Analysis of the reduced system

For the original system (16) we have concluded that, in properly weighted Hilbert spaces, we have Brezzi constants which are independent of the regularization parameter α. It is not self-evident that similar properties can be proved for the reduced system (28).

The form of (28), and the fact that $Eu \in U$, motivate us to introduce the weighted norm

$$\|u\|^2_{U_{1+\alpha^{-1}}} = \|Eu\|^2_{H^1(\Omega)} + \frac{1}{\alpha}\|Eu\|^2_{L^2(\partial\Omega)}, \; u \in H^1(\Omega_f),$$

for the state function. Note that we, for the sake of simplicity, no longer use the notation $\hat{u}$ for functions in $H^1(\Omega_f)$. To further increase readability, we define the product space

$$V_\alpha = U_{1+\alpha^{-1}} \times H^1(\Omega_f). \tag{33}$$

We can now define the bilinear form

$$a(\,\cdot\,;\,\cdot\,) : V_\alpha \times V_\alpha \to \mathbb{R},$$

associated with (28), as

$$\begin{aligned} a(u, f; \phi, \varphi) = &- \frac{1}{\alpha}(Eu, E\phi)_{L^2(\partial\Omega)} + (Ef, E\phi)_{H^1(\Omega)} \\ &+ (Eu, E\varphi)_{H^1(\Omega)} + (f, \varphi)_{H^1(\Omega_f)}. \end{aligned}$$

According to Babuška theory [1], (28) is well-posed if and only if $a(\,\cdot\,,\,\cdot\,)$ is continuous and weakly coercive. Since we in Section 5 concluded that (28) is well-posed, it follows that $a(\,\cdot\,,\,\cdot\,)$ fulfills these two conditions. To obtain an α-robust preconditioner, however, we must show that the constants appearing in the continuity and coercivity bounds are independent of α, provided that proper weighted norms are applied.

Lemma 6.0.2 *There exists a constant $C_1 > 0$, independent of $\alpha > 0$, such that*

$$|a(u, f; \phi, \varphi)| \leq C_1 \|(u, f)\|_{V_\alpha} \|(\phi, \varphi)\|_{V_\alpha},$$

where V_α is defined in (33).

Proof

Cauchy-Schwartz' inequality implies that

$$\begin{aligned} |a(u, f; \phi, \varphi)| \leq &\frac{1}{\sqrt{\alpha}}\|Eu\|_{L^2(\partial\Omega)} \frac{1}{\sqrt{\alpha}}\|E\phi\|_{L^2(\partial\Omega)} \\ &+ \|Ef\|_{H^1(\Omega)}\|E\phi\|_{H^1(\Omega)} \\ &+ \|Eu\|_{H^1(\Omega)}\|E\varphi\|_{H^1(\Omega)} \\ &+ \|f\|_{H^1(\Omega_f)}\|\varphi\|_{H^1(\Omega_f)} \\ \leq &\, \tilde{c} \left[\frac{1}{\sqrt{\alpha}}\|Eu\|_{L^2(\partial\Omega)} \frac{1}{\sqrt{\alpha}}\|E\phi\|_{L^2(\partial\Omega)} \right. \\ &+ \|f\|_{H^1(\Omega_f)}\|E\phi\|_{H^1(\Omega)} \\ &+ \|Eu\|_{H^1(\Omega)}\|\varphi\|_{H^1(\Omega_f)} \\ &\left. + \|f\|_{H^1(\Omega_f)}\|\varphi\|_{H^1(\Omega_f)} \right] \\ \leq &\, 4\tilde{c}\|(u, f)\|_{V_\alpha}\|(\phi, \varphi)\|_{V_\alpha}, \end{aligned}$$

where $\tilde{c} = \max\{\|E\|, 1\}$. The result follows with $C_1 = 4\tilde{c}$.
■

The weak coercivity of the bilinear operator $a(\,\cdot\,,\,\cdot\,)$ is defined in terms of two inf-sup conditions. Babuška theory asserts that the constants in the inf-sup conditions coincide when the system is well-posed, provided that only reflexive Banach spaces are involved. This constant will be independent of α, as the following lemma expresses.

Lemma 6.0.3 (Weak coercivity) *There exists a constant $C_2 > 0$, independent of $\alpha > 0$, such that*

$$\inf_{(\phi,\varphi)\in V_\alpha} \sup_{(u,f)\in V_\alpha} \frac{a(u,f;\phi,\varphi)}{\|(u,f)\|_{V_\alpha}\|(\phi,\varphi)\|_{V_\alpha}} \geq C_2,$$

and

$$\inf_{(u,f)\in V_\alpha} \sup_{(\phi,\varphi)\in V_\alpha} \frac{a(u,f;\phi,\varphi)}{\|(u,f)\|_{V_\alpha}\|(\phi,\varphi)\|_{V_\alpha}} \geq C_2.$$

Proof

Let $(\phi,\varphi) \in V_\alpha \setminus \{(0,0)\}$ be arbitrary. With $u = -\phi$ and $f = \phi + \varphi$ we get

$$\begin{aligned}
\sup_{(u,f)\in V_\alpha} \frac{a(u,f;\phi,\varphi)}{\|(u,f)\|_{V_\alpha}\|(\phi,\varphi)\|_{V_\alpha}} &\geq \frac{\frac{1}{\alpha}\|E\phi\|^2_{L^2(\partial\Omega)} + \|E\phi\|^2_{H^1(\Omega)} + \|\varphi\|^2_{H^1(\Omega_f)} + (\phi,\varphi)_{H^1(\Omega_f)}}{\|(\phi,\phi+\varphi)\|_{V_\alpha}\|(\phi,\varphi)\|_{V_\alpha}} \\
&\geq \frac{\|\phi\|^2_{U_{1+\alpha^{-1}}} + \|\varphi\|^2_{H^1(\Omega_f)} - \|E\phi\|_{H^1(\Omega)}\|\varphi\|_{H^1(\Omega_f)}}{\|(\phi,\phi+\varphi)\|_{V_\alpha}\|(\phi,\varphi)\|_{V_\alpha}} \\
&\geq \frac{\|\phi\|^2_{U_{1+\alpha^{-1}}} + \|\varphi\|^2_{H^1(\Omega_f)} - \frac{1}{2}\|E\phi\|^2_{H^1(\Omega)} - \frac{1}{2}\|\varphi\|^2_{H^1(\Omega_f)}}{\|(\phi,\phi+\varphi)\|_{V_\alpha}\|(\phi,\varphi)\|_{V_\alpha}} \\
&\geq \frac{\|\phi\|^2_{U_{1+\alpha^{-1}}} + \|\varphi\|^2_{H^1(\Omega_f)} - \frac{1}{2}\|\phi\|^2_{U_{1+\alpha^{-1}}} - \frac{1}{2}\|\varphi\|^2_{H^1(\Omega_f)}}{\|(\phi,\phi+\varphi)\|_{V_\alpha}\|(\phi,\varphi)\|_{V_\alpha}} \\
&\geq \frac{1}{2\sqrt{3}} \frac{\|(\phi,\varphi)\|^2_{V_\alpha}}{\|(\phi,\varphi)\|_{V_\alpha}\|(\phi,\varphi)\|_{V_\alpha}} \\
&= \frac{1}{2\sqrt{3}},
\end{aligned}$$

where we have used that $\|(\phi, \phi+\varphi)\|_{V_\alpha} \leq \sqrt{3}\|(\phi,\varphi)\|_{V_\alpha}$, which is a consequence of the triangle inequality and that $2ab \leq (a^2+b^2)$:

$$
\begin{aligned}
\|(\phi, \phi+\varphi)\|_{V_\alpha}^2 &= \|E\phi\|_{H^1(\Omega)}^2 + \frac{1}{\alpha}\|E\phi\|_{L^2(\partial\Omega)}^2 + \|\phi+\varphi\|_{H^1(\Omega_f)}^2 \\
&\leq \|E\phi\|_{H^1(\Omega)}^2 + \frac{1}{\alpha}\|E\phi\|_{L^2(\partial\Omega)}^2 \\
&\quad + \|\phi\|_{H^1(\Omega_f)}^2 + 2\|\phi\|_{H^1(\Omega_f)}\|\varphi\|_{H^1(\Omega_f)} + \|\varphi\|_{H^1(\Omega_f)}^2 \\
&\leq \|(\phi,\varphi)\|_{V_\alpha}^2 + \|\phi\|_{H^1(\Omega_f)}^2 + \|\phi\|_{H^1(\Omega_f)}^2 + \|\varphi\|_{H^1(\Omega_f)}^2 \\
&\leq 3\|(\phi,\varphi)\|_{V_\alpha}^2.
\end{aligned}
$$

The first inf-sup condition thus holds with $C_2 = \frac{1}{2\sqrt{3}}$. Since (28) is well-posed, the second inf-sup condition immediately follows, with the same constant, from standard Babuška theory.

■

We have now verified that the bilinear form $a(\cdot,\cdot) : V_\alpha \times V_\alpha \to \mathbb{R}$ is continuous and weakly coercive, with constants that are independent of α. We can then, from the Babuška theory, conclude that

Theorem 6.0.1 *The operator $\mathcal{A}_\alpha^{2\times2} : V_\alpha \to V_\alpha'$, defined in* (28)*, is an isomorphism. That is, $\mathcal{A}_\alpha^{2\times2}$ is bounded and continuously invertible for $\alpha > 0$, in the sense that*

$$\|\mathcal{A}_\alpha^{2\times2}\|_{\mathcal{L}(V_\alpha,V_\alpha')} \leq C_1 \quad \text{and} \quad \|[\mathcal{A}_\alpha^{2\times2}]^{-1}\|_{\mathcal{L}(V_\alpha',V_\alpha)} \leq C_2^{-1},$$

where both C_1 and C_2 are independent of $\alpha > 0$.

Proof

The result follows from Lemma 6.0.2, Lemma 6.0.3 and standard Babuška theory.

■

7 Preconditioners

For a mapping $\mathcal{A}_\alpha$, of the form (16) or (28), from a Hilbert space H onto its dual space H', Krylov subspace methods cannot be applied to solve

$$\mathcal{A}_\alpha x = b.$$

However, assuming that an operator $\mathcal{B}_\alpha : H' \to H$ is available, Krylov subspace methods can be employed to solve

$$\mathcal{B}_\alpha \mathcal{A}_\alpha x = \mathcal{B}_\alpha b, \tag{34}$$

since $\mathcal{B}_\alpha \mathcal{A}_\alpha$ is a mapping from H to H.

In [2, 7, 4] the authors discuss that, to obtain an efficient preconditioner $\mathcal{B}_\alpha$, this mapping should be an isomorphism, with h and α-independent bounds for both $\|\mathcal{B}_\alpha\|$ and $\|\mathcal{B}_\alpha^{-1}\|$. With these ideas in mind, we can propose preconditioners for both the 3×3 and 2×2 block systems analyzed above.

7.1 3×3 system

To construct a suitable preconditioner

$$\mathcal{B}_\alpha^{3\times 3} : F'_\alpha \times U'_\alpha \times U'_{\alpha^{-1}} \to F_\alpha \times U_\alpha \times U_{\alpha^{-1}}$$

for $\mathcal{A}_\alpha^{3\times 3}$, defined in (16), let us recall the scaled norms

$$\begin{aligned} \|f\|_{F_\alpha}^2 &= \alpha \|f\|_{H^1(\Omega_f)}^2, \\ \|u\|_{U_\alpha}^2 &= \alpha \|u\|_{H^1(\Omega)}^2 + \|u\|_{L^2(\partial\Omega)}^2, \\ \|w\|_{U_{\alpha^{-1}}}^2 &= \frac{1}{\alpha} \|w\|_{H^1(\Omega)}^2. \end{aligned}$$

We suggest to use the inverse of the Riesz map of the space

$$W_\alpha = F_\alpha \times U_\alpha \times U_{\alpha^{-1}}$$

as preconditioner, i.e.

$$\mathcal{B}_\alpha^{3\times 3} = \begin{bmatrix} \alpha \tilde{A}_f & 0 & 0 \\ 0 & \alpha A + M_\partial & 0 \\ 0 & 0 & \frac{1}{\alpha} A \end{bmatrix}^{-1} : W'_\alpha \to W_\alpha, \tag{35}$$

see (17)-(20). Clearly, $\mathcal{B}_\alpha^{3\times 3}$ is an isomorphism with α independent bounds for $\|\mathcal{B}_\alpha\|$ and $\|\mathcal{B}_\alpha^{-1}\|$. Moreover, provided that sound discretization techniques are employed, this property will be inherited by the associated discretized operator.

In Section 4 we concluded that both $\|\mathcal{A}_\alpha^{3\times 3}\|$ and $\|[\mathcal{A}_\alpha^{3\times 3}]^{-1}\|$ are bounded independently of $\alpha > 0$. Consequently, also $\|\mathcal{B}_\alpha^{3\times 3}\mathcal{A}_\alpha^{3\times 3}\|$ and $\|[\mathcal{B}_\alpha^{3\times 3}\mathcal{A}_\alpha^{3\times 3}]^{-1}\|$ are well behaved, regardless of the size of $\alpha > 0$. That is, $\mathcal{B}_\alpha^{3\times 3}$ yields a regularization robust preconditioner:

$$\mathcal{B}_\alpha^{3\times 3} \mathcal{A}_\alpha^{3\times 3} x = \mathcal{B}_\alpha^{3\times 3} b. \tag{36}$$

7.2 2×2 system

In the analysis of the operator $\mathcal{A}_\alpha^{2\times 2}$, defined in (28), we used the weighted norm

$$\|u\|_{U_{1+\alpha^{-1}}}^2 = \|Eu\|_{H^1(\Omega)}^2 + \frac{1}{\alpha} \|Eu\|_{L^2(\partial\Omega)}^2, \; u \in H^1(\Omega_f).$$

Thus, a natural preconditioner reads

$$\mathcal{B}_\alpha^{2\times 2} = \begin{bmatrix} \mathsf{A} + \frac{1}{\alpha}\mathsf{M}_\partial & 0 \\ 0 & \mathsf{A}_f \end{bmatrix}^{-1} : V_\alpha' \to V_\alpha, \tag{37}$$

since this is the inverse of the Riesz map of V_α, see (33) and (29)-(32). Sound discretization techniques will, as for the 3×3 system, provide an α-robust preconditioner:

$$\mathcal{B}_\alpha^{2\times 2}\mathcal{A}_\alpha^{2\times 2}x = \mathcal{B}_\alpha^{2\times 2}b. \tag{38}$$

8 Numerical experiments

8.1 The extension operator

In properly weighted Hilbert spaces, the systems (16) and (28) are well-behaved, independently of $\alpha > 0$. To solve these systems numerically, however, we have to represent the subspace $U \subset H^1(\Omega)$, or alternatively, compute the action of the extension operator E, defined in (22). We now address how this can be accomplished.

Let $\{\phi_i\}_{i=1}^N$ be a basis for $V_h \subset H^1(\Omega_f)$, where V_h is a standard scalar FE space. We then get

$$(Eu_h, E\phi_i)_{H^1(\Omega)} = (u_h, \phi_i)_{H^1(\Omega_f)} + (Eu_h, E\phi_i)_{H^1(\Omega_0)}, \tag{39}$$

since the extension leaves the function unchanged throughout Ω_f. Furthermore, with $u_h = \sum_{j=1}^N a_j\phi_j$,

$$(Eu_h, E\phi_i)_{H^1(\Omega_0)} = \sum_{j=1}^N a_j(E\phi_j, E\phi_i)_{H^1(\Omega_0)}.$$

We note from (9)-(11) that $\tilde{\phi}_j = (E\phi_j)|_{\Omega_0}$ is uniquely determined from $\phi_j|_{\partial\Omega_f}$. And,

$$\tilde{\phi}_j = (E\phi_j)|_{\Omega_0} = 0 \text{ if } \phi_j|_{\partial\Omega_f} = 0.$$

Consequently, only the basis functions associated with nodes positioned at the boundary $\partial\Omega_f$ of Ω_f will have non-zero extensions. These extensions are determined by solving (9)-(11). More specifically, one elliptic boundary problem must be solved for each node positioned at $\partial\Omega_f$. This may become CPU demanding if the number of nodes at this interface is large, but the process is easy to parallelize. A more thorough discussion of this issue is present in Section 10.

When the non-zero extensions have been determined, the matrix contributions associated with $(Eu_h, \phi_i)_{H^1(\Omega)}$ can be assembled by computing the two inner-products in (39). And, it is also straightforward to assemble the "boundary" matrix associated with M_∂, see (28)-(29),

$$(Eu_h, E\phi_i)_{L^2(\partial\Omega)} = \sum_{j=1}^{N} a_j (E\phi_j, E\phi_i)_{L^2(\partial\Omega)}.$$

Remark

Krylov subspace solvers typically require that $\mathcal{A}_\alpha^{2\times 2} p$ is computed, for a given (vector) p. Since the extension E is defined in terms of an elliptic PDE, it should be possible to determine $\mathcal{A}_\alpha^{2\times 2} p$ by employing a multigrid scheme, without computing $E\phi_j$ for all indexes associated with nodes at $\partial\Omega_f$. Similarly, it is also likely that the action of the preconditioner $\mathcal{B}_\alpha^{2\times 2}$ can be directly computed with multigrid schemes, see (37) and (29)-(32). Hence, one would expect that the step of explicitly determining $\{E\phi_j\}$ can be avoided, provided that proper tailored software is available. We, however, used standard software packages, and it turned out to be difficult to avoid this preprocessing task prior to assembling and solving the KKT system.

8.2 Numerical setup

To avoid introducing further notation, we will not define new symbols for the matrices and vectors associated with the operators and functions in (16), (28), (35) and (37). We would like to emphasis that, in this section, all use of these symbols are to the associated discretized versions.

- All simulations were performed using cbc.block; a branch of the FEniCS software [5].
- For all MINRES tests, the preconditioners (35) and (37) were approximated with the Algebraic MultiGrid (AMG) package in PyTrilinos. We used a symmetric Gauss-Seidel smoother, with three smoothing sweeps.
- For the computations of the eigenvalues, presented below, we dumped the matrices to .mat-files and computed the exact preconditioners in Octave.
- In all simulations, we worked on the domains

  $$\Omega = (0,1) \times (0,1),$$
  $$\Omega_f = (.25, .75) \times (.25, .75).$$

  The observation data d was generated by solving (5) with

  $$f(x,y) = 3\cos(\pi x) + y^2, \qquad (40)$$

  and setting $d = u|_{\partial\Omega}$, where u denotes the solution of (5).

- The MINRES iterations were stopped when

$$\frac{(r_k, \mathcal{B}_\alpha^{p\times p} r_k)}{(r_0, \mathcal{B}_\alpha^{p\times p} r_0)} = \frac{(\mathcal{A}_\alpha^{p\times p} x_k - b, \mathcal{B}^{p\times p}[\mathcal{A}_\alpha^{p\times p} x_k - b])}{(\mathcal{A}_\alpha^{p\times p} x_0 - b, \mathcal{B}^{p\times p}[\mathcal{A}_\alpha^{p\times p} x_0 - b])} \leq \epsilon, \tag{41}$$

  where $p = 2, 3$. In other words, we used a standard relative stopping criterion.

8.3 3×3 system

We first consider the numerical solution of (16) with the preconditioner (35). In Table 1 we can not observe any (systematic) growth in the iteration numbers when the regularization parameter α decreases. The small increase in iteration numbers when the mesh parameter $h \to 0$ is most likely linked to the performance of the AMG. This is supported by Table 2, where we do not observe a significant increase of the condition number $\kappa(\mathcal{B}_\alpha^{3\times 3}\mathcal{A}_\alpha^{3\times 3})$ for $h = 2^{-6}$ compared to $h = 2^{-5}$. That is, for small values of α, the condition number equals 8.701 for $h = 2^{-5}$ and it equals 8.705 for $h = 2^{-6}$. Thus, the discretized preconditioner provides iteration counts, for the MINRES method, which are well behaved with respect to the mesh and regularization parameters.

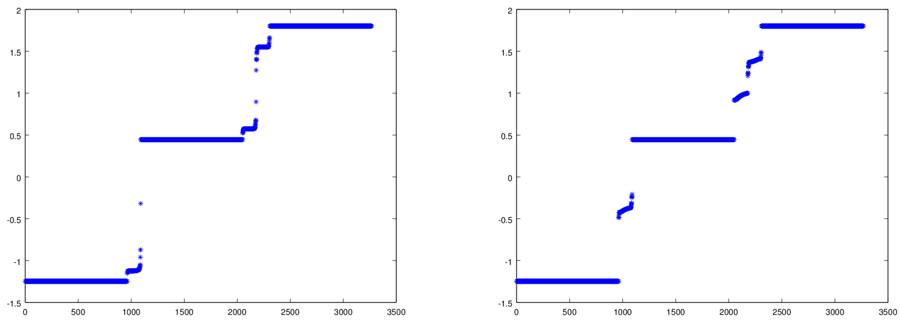
In Figure 2 the spectrum of $\mathcal{B}_\alpha^{3\times 3}\mathcal{A}_\alpha^{3\times 3}$ is displayed for two choices of α. Both spectra are clustered, with three large bands of eigenvalues. The remaining eigenvalues seem to become more clustered for smaller values of α than for $\alpha = 1$.

$\alpha \backslash h$	2^{-5}	2^{-6}	2^{-7}	2^{-8}
1	41	45	47	59
10^{-1}	49	54	55	65
10^{-2}	64	70	72	80
10^{-3}	56	66	68	80
10^{-4}	48	52	57	73

Table 1: The number of MINRES iterations required to solve the KKT system (16), i.e. the 3×3 system, with the preconditioner (35). The stopping criterion was $\epsilon = 10^{-8}$.

8.4 2×2 system

To numerically solve the 3×3 block system (16), with the preconditioner (35), we must implement the extension operator E in (22). Thus, it is practically no additional computational effort to reduce the problem to the 2×2 block system (28), with the preconditioner (37). We will now explore how the MINRES algorithm performs on this reduced system.



(a) $\alpha = 1$.

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

Figure 2: The spectrum of $\mathcal{B}_\alpha^{3\times3}\mathcal{A}_\alpha^{3\times3}$ for two different regularization parameters. These results were computed with a mesh parameter $h = 2^{-6}$.

α	$\lvert\lambda_1\rvert$	$\lvert\lambda_n\rvert$
1	0.31835	1.8019
10^{-1}	0.22199	1.8019
10^{-2}	0.21079	1.8019
10^{-3}	0.20872	1.8019
10^{-4}	0.20751	1.8019
10^{-5}	0.20713	1.8019
10^{-6}	0.20709	1.8019
10^{-7}	0.20708	1.8019
10^{-8}	0.20708	1.8019

(a) Mesh parameter $h = 2^{-5}$.

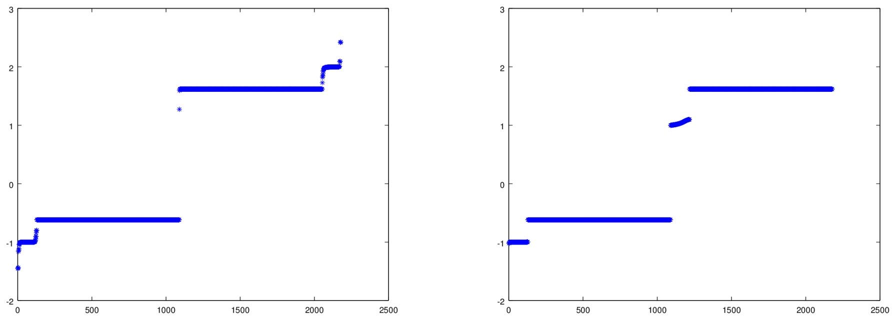
α	$\lvert\lambda_1\rvert$	$\lvert\lambda_n\rvert$
1	0.31831	1.8019
10^{-1}	0.22192	1.8019
10^{-2}	0.21078	1.8019
10^{-3}	0.20886	1.8019
10^{-4}	0.20771	1.8019
10^{-5}	0.20710	1.8019
10^{-6}	0.20700	1.8019
10^{-7}	0.20699	1.8019
10^{-8}	0.20699	1.8019

(b) Mesh parameter $h = 2^{-6}$.

Table 2: The smallest and largest eigenvalues, measured in absolute value, of $\mathcal{B}_\alpha^{3\times3}\mathcal{A}_\alpha^{3\times3}$.

In Table 3 we observe much smaller iteration numbers than for the 3×3 system, see Table 1. Furthermore, the iteration counts in Table 3 are well behaved with respect to the size of the mesh parameter h, and, if anything, the iteration numbers decreases as α decreases. The latter fact is probably not a generic pattern, but, for this particular model problem, this observation is supported by the eigenvalues reported in Table 4. That is, the computed condition number $\kappa(\mathcal{B}_\alpha^{2\times2}\mathcal{A}_\alpha^{2\times2})$ is non-increasing as $\alpha \to 0$. The spectrum of $\mathcal{B}_\alpha^{2\times2}\mathcal{A}_\alpha^{2\times2}$ is depicted in Figure 3.

According to standard theory, the MINRES scheme requires $O(\kappa(\mathcal{A}))$ iterations to solve the system $\mathcal{A}x = b$. For the 2×2 block system, with $h = 2^{-6}$ and $\alpha = 10^{-4}$, the condition number is $\kappa(\mathcal{B}_\alpha^{2\times2}\mathcal{A}_\alpha^{2\times2}) = 2.618$, while for the 3×3 block system the condition number is $\kappa(\mathcal{B}_\alpha^{3\times3}\mathcal{A}_\alpha^{3\times3}) = 8.675$, which gives the ratio $8.675/2.618 \approx 3.31$. The ratio between the



(a) $\alpha = 1$. (b) $\alpha = 10^{-4}$.

Figure 3: The spectrum of $\mathcal{B}_\alpha^{2\times 2}\mathcal{A}_\alpha^{2\times 2}$ for two different regularization parameters. These results where generated with the mesh parameter $h = 2^{-6}$.

associated iteration counts is $52/17 \approx 3.06$. Similar results hold for the other choices, reported in our tables, of the mesh parameter and the regularization parameter.

$\alpha\backslash h$	2^{-5}	2^{-6}	2^{-7}	2^{-8}
1	31	31	33	37
10^{-1}	30	30	32	35
10^{-2}	28	27	30	34
10^{-3}	21	20	23	27
10^{-4}	18	17	20	24

Table 3: The number of MINRES iterations required to solve the KKT system (28), i.e. the 2×2 system, with the preconditioner (37). The stopping criterion was $\epsilon = 10^{-8}$.

8.5 Comparison with standard preconditioners

Recall the original form (4)-(5) of our optimization problem. The KKT system associated with this problem, without invoking the space U of functions which are Helmholtz-harmonic on Ω_0, will yield an operator

$$\mathcal{C}_\alpha : H^1(\Omega_f) \times H^1(\Omega) \times H^1(\Omega) \rightarrow H^1(\Omega_f)' \times H^1(\Omega)' \times H^1(\Omega)'.$$

Hence, we may simply use the inverse of the Riesz map of $H^1(\Omega_f) \times H^1(\Omega) \times H^1(\Omega)$ to obtain a preconditioned system of the form

$$\mathcal{R}\mathcal{C}_\alpha x = \mathcal{R}b. \tag{42}$$

In this case, no weighting of the involved norms is applied. We will now compare the performance of this methodology with the approaches discussed

α	$\|\lambda_1\|$	$\|\lambda_n\|$
1	0.61803	2.3822
10^{-1}	0.61803	2.2465
10^{-2}	0.61803	1.9123
10^{-3}	0.61803	1.6180
10^{-4}	0.61803	1.6180

(a) Mesh parameter $h = 2^{-5}$.

α	$\|\lambda_1\|$	$\|\lambda_n\|$
1	0.61803	2.4246
10^{-1}	0.61803	2.3213
10^{-2}	0.61803	2.0512
10^{-3}	0.61803	1.6180
10^{-4}	0.61803	1.6180

(b) Mesh parameter $h = 2^{-6}$.

Table 4: The smallest and largest eigenvalues, measured in absolute value, of $\mathcal{B}_\alpha^{2\times2}\mathcal{A}_\alpha^{2\times2}$.

above. We will not undertake a complete "start-to-finish" comparison, since there are many possibilities, particularly with respect to parallelization, to speed up the computation of the extension operator. Instead, we will perform a pure MINRES test, where we measure the wall-time needed by the different schemes.

Tables 5 and 6 contain the speed-up obtained by solving the 3×3 and 2×2 block systems (36) and (38), respectively, instead of applying the Krylov subspace solver to (42). As expected, the speed-up increases when $\alpha \to 0$. For example, with $\alpha = 10^{-6}$ and $h < 10^{-5}$, MINRES solves the 2×2 block system more than 48 times faster than the "standard" preconditioned KKT system (42).

$\alpha\backslash h$	2^{-5}	2^{-6}	2^{-7}	2^{-8}
1	1.33	1.71	1.64	1.49
10^{-1}	1.80	2.10	2.22	2.02
10^{-2}	2.00	2.19	2.35	2.42
10^{-3}	3.16	3.88	3.64	3.54
10^{-4}	4.70	6.26	5.17	4.98
10^{-5}	7.22	9.00	8.74	8.58
10^{-6}	9.22	12.5	11.7	10.7

Table 5: Ratio between the wall-time needed by MINRES to solve (42) and (36) (3×3 block system).

Tables 5 and 6 show that the methods introduced in this paper perform favorable compared with the "standard" preconditioning technique. Nevertheless, the comparison is not truly "objective": The stopping criterion depends on the involved operators and on the regularization parameter α, see (41). Therefore, we also performed an alternative comparison, where we added a prior, given by (40), to the minimization problem. More specifically, we replaced the regularization term in the cost functionals (4) and (7) with

$$\frac{1}{2}\alpha\|f - f_{\text{prior}}\|^2_{H^1(\Omega_f)}, \tag{43}$$

$\alpha\backslash h$	2^{-5}	2^{-6}	2^{-7}	2^{-8}
1	2.40	3.63	3.15	3.74
10^{-1}	3.60	5.25	5.24	5.91
10^{-2}	6.50	8.14	7.75	8.86
10^{-3}	12.7	18.6	15.2	16.3
10^{-4}	15.7	23.8	22.7	23.3
10^{-5}	21.7	40.5	39.0	38.1
10^{-6}	27.7	53.0	48.9	48.1

Table 6: Ratio between the wall-time needed by MINRES to solve (42) and (38) (2×2 block system).

where $f_{\text{prior}}(x, y) = 3\cos(\pi x) + y^2$. Hence, f_{prior} is both used to generate the observation data d and as a prior. The control determined by solving the KKT system will therefore be almost equal to f_{prior}. Our objective is to study how fast the approximate controls f_k, generated by the MINRES algorithm, approaches this function.

In Figure 4, the relative difference

$$\frac{\|f_k - f_{\text{prior}}\|_{H^1(\Omega_f)}}{\|f_{\text{prior}}\|_{H^1(\Omega_f)}}$$

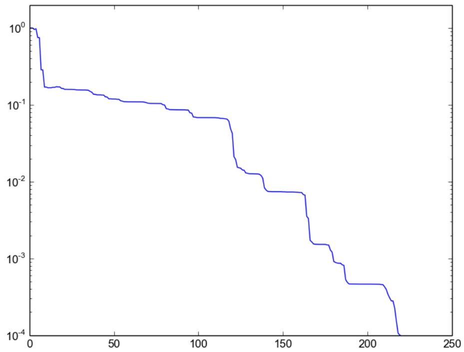
is displayed as a function of the number of MINRES iterations k. More specifically, the MINRES method was applied to (42), the 3×3 block system (36) and the 2×2 block system (38). We observe that the relative difference is reduced to 10^{-4} in approximately 220 iterations by the first scheme, while the two latter preconditioning techniques required 35 and 15 iterations, respectively; see figures 4(a), 4(b) and 4(c). In these experiments we used a zero initial guess for the iteration process, $\alpha = 10^{-4}$ and $h = 1/256$.

9 Further analysis

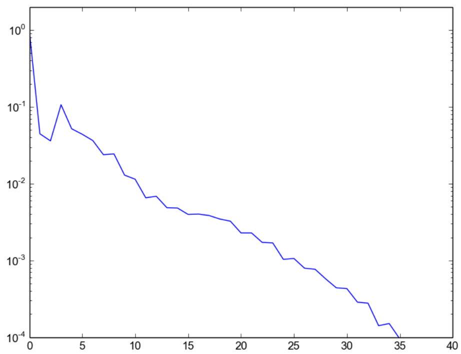
From the analysis presented above, we can conclude that the spectral condition number of the preconditioned operator $\mathcal{B}_\alpha^{2\times2}\mathcal{A}_\alpha^{2\times2}$ is bounded independently of α and h. On the other hand, Figure 3 indicates that the spectrum of this operator may possess further nice properties. The purpose of this section is to investigate this issue from an algebraic point of view. Throughout this section we assume that M_∂, A and A_f are FE operators.

A member λ of the point spectrum of $\mathcal{B}_\alpha^{2\times2}\mathcal{A}_\alpha^{2\times2}$ must satisfy

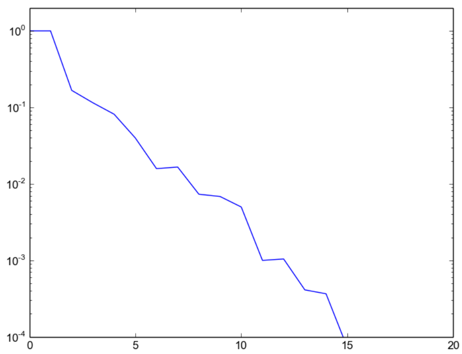
$$\underbrace{\begin{bmatrix} -\frac{1}{\alpha}\mathsf{M}_\partial & \mathsf{A} \\ \mathsf{A} & \mathsf{A}_f \end{bmatrix}}_{\mathcal{A}_\alpha^{2\times2}} \begin{bmatrix} u \\ f \end{bmatrix} = \lambda \underbrace{\begin{bmatrix} \mathsf{A} + \frac{1}{\alpha}\mathsf{M}_\partial & 0 \\ 0 & \mathsf{A}_f \end{bmatrix}}_{(\mathcal{B}_\alpha^{2\times2})^{-1}} \begin{bmatrix} u \\ f \end{bmatrix},$$



(a) "Standard" Riesz preconditioner, see (42).



(b) 3×3 preconditioner, see (35).



(c) 2×2 preconditioner, see (37).

Figure 4: The relative difference $\|f_k - f_{\text{prior}}\|_{H^1(\Omega_f)}/\|f_{\text{prior}}\|_{H^1(\Omega_f)}$ as a function of the number of MINRES iterations k.

or

$$
\begin{aligned}
-\frac{1}{\alpha}\mathsf{M}_{\partial}u + \mathsf{A}f &= \lambda\left(\mathsf{A}u + \frac{1}{\alpha}\mathsf{M}_{\partial}u\right), && (44)\\
\mathsf{A}u + \mathsf{A}_f f &= \lambda \mathsf{A}_f f. && (45)
\end{aligned}
$$

From (45) we find that

$$
f = \frac{1}{\lambda - 1}\mathsf{A}_f^{-1}\mathsf{A}u,\ \lambda \neq 1,
$$

which we may insert into (44) to obtain

$$
-\frac{1}{\alpha}(1+\lambda)\mathsf{M}_{\partial}u + \frac{1}{\lambda-1}\mathsf{A}\mathsf{A}_f^{-1}\mathsf{A}u - \lambda\mathsf{A}u = 0. \tag{46}
$$

Hence, any eigenvalue must satisfy

$$
\frac{1}{\alpha}(1+\lambda)\langle \mathsf{A}^{-1}\mathsf{A}_f\mathsf{A}^{-1}\mathsf{M}_{\partial}u, u\rangle + \frac{1}{1-\lambda}\langle u, u\rangle + \lambda\langle \mathsf{A}^{-1}\mathsf{A}_f u, u\rangle = 0, \tag{47}
$$

where

- $\mathsf{A}^{-1}\mathsf{A}_f\mathsf{A}^{-1}\mathsf{M}_{\partial}$ is semi-positive,
- $\mathsf{A}^{-1}\mathsf{A}_f$ is positive.

Theorem 9.0.1

a) *Let*

$$
\begin{aligned}
\underline{\gamma} &= \lambda_{\min}(\mathsf{A}^{-1}\mathsf{A}_f),\\
\overline{\gamma} &= \lambda_{\max}(\mathsf{A}^{-1}\mathsf{A}_f).
\end{aligned}
$$

Then,

$$
\begin{aligned}
\operatorname{sp}(\mathcal{B}_\alpha^{2\times 2}\mathcal{A}_\alpha^{2\times 2}) \subset & \left[\min\left\{-1, \frac{1}{2} - \sqrt{\frac{1}{4} + \frac{1}{\underline{\gamma}}}\right\}, \max\left\{-1, \frac{1}{2} - \sqrt{\frac{1}{4} + \frac{1}{\overline{\gamma}}}\right\}\right]\\
& \cup \left(1, \frac{1}{2} + \sqrt{\frac{1}{4} + \frac{1}{\underline{\gamma}}}\right].
\end{aligned}
$$

b) *If* $\lambda \in \operatorname{sp}(\mathcal{B}_\alpha^{2\times 2}\mathcal{A}_\alpha^{2\times 2})$ *is an eigenvalue associated with* (u, f)*, where* $u|_{\partial\Omega_f} = 0$*, then*

$$
\lambda = \frac{1 \pm \sqrt{5}}{2} \approx \begin{cases} 1.618, \\ -0.618. \end{cases}
$$

c) *The multiplicity of the eigenvalue* $\lambda = -1$ *equals the dimension of the null-space of the operator*

$$
Q : u \to (u, \phi)_{H^1(\Omega_f)} - (Eu, E\phi)_{H^1(\Omega_0)}, \quad \phi \in V_h,
$$

where V_h *is a FE space for* $H^1(\Omega_f)$*.*

Remarks

For the model problem associated with Figure 3, $\overline{\gamma} = 1.000$ and $\underline{\gamma} = 0.250$. Hence, in this case, invoking a) yields that

$$\mathrm{sp}(\mathcal{B}_\alpha^{2\times 2}\mathcal{A}_\alpha^{2\times 2}) \subset [-1.562, -0.618] \cup (1, 2.562] \quad \forall \alpha > 0.$$

The eigenvalues $\lambda = -0.618, 1.618$, see b), typically have large multiplicity and appear as "long horizontal line segments" in Figure 3. (These numbers coincide with those derived in [8]. We have not fully investigated the connection between the analysis presented in this paper and [8].)

If $\mathrm{supp}(u) \subset \Omega_f$, then $Eu|_{\Omega_0} = 0$ and, hence, $u \neq 0$ can not belong to the null-space of Q, see c). From this we may conclude that the dimension of the kernel of Q, and thus the multiplicity of $\lambda = -1$, is less or equal to the number of nodes at the boundary $\partial\Omega_f$ of Ω_f.

Proof of a)

- Note that, if $u = 0$, then (44) implies that $f = 0$. Hence, there does not exist any eigenfunction (u, f) with $u = 0$. Therefore, in the analysis presented below, we can always assume that $u \neq 0$ in (47).
- Also, for $\lambda = 0$, $(1+\lambda)$, $\frac{1}{1-\lambda} > 0$ and from (47) we find that 0 can not be an eigenvalue.

Positive eigenvalues

- Since $(1+\lambda)$, $\frac{1}{1-\lambda}, \lambda > 0$ for $\lambda \in (0,1)$, (47) yields that the open unit interval $(0,1)$ contains no eigenvalues.
- For $\lambda = 1$, (45) implies that $u = 0$ and it follows from (44) that also $f = 0$. We conclude that 1 does not belong to the spectrum of $\mathcal{B}_\alpha\mathcal{A}_\alpha$.
- For $\lambda > 1$ we find that, see (47),

$$\begin{aligned} \frac{1}{\alpha}(1+\lambda)\langle \mathsf{A}^{-1}\mathsf{A}_f\mathsf{A}^{-1}\mathsf{M}_\partial u, u\rangle \quad &+ \quad \frac{1}{1-\lambda}\langle u, u\rangle + \lambda\langle \mathsf{A}^{-1}\mathsf{A}_f u, u\rangle \\ &\geq \quad \frac{1}{1-\lambda}\langle u, u\rangle + \lambda\underline{\gamma}\langle u, u\rangle \\ &= \quad \left(\frac{1}{1-\lambda} + \lambda\underline{\gamma}\right)\langle u, u\rangle \\ &> \quad 0 \end{aligned}$$

  if $\lambda > \frac{1}{2} + \sqrt{\frac{1}{4} + \frac{1}{\underline{\gamma}}}$. Thus, there are no eigenvalues larger than $\frac{1}{2} + \sqrt{\frac{1}{4} + \frac{1}{\underline{\gamma}}}$.

We conclude that positive eigenvalues must belong to the interval

$$\left(1, \frac{1}{2} + \sqrt{\frac{1}{4} + \frac{1}{\underline{\gamma}}}\right].$$

Negative eigenvalues

- For $\lambda \in (-1, 0)$ it follows that $(1+\lambda)$, $\frac{1}{1-\lambda} > 0$ and $\lambda < 0$. Consequently, see (47),

$$\begin{aligned}\frac{1}{\alpha}(1+\lambda)\langle \mathsf{A}^{-1}\mathsf{A}_f\mathsf{A}^{-1}\mathsf{M}_\partial u, u\rangle &+ \frac{1}{1-\lambda}\langle u, u\rangle + \lambda\langle \mathsf{A}^{-1}\mathsf{A}_f u, u\rangle \\ &\geq \frac{1}{1-\lambda}\langle u, u\rangle + \lambda\overline{\gamma}\langle u, u\rangle \\ &= \left(\frac{1}{1-\lambda} + \lambda\overline{\gamma}\right)\langle u, u\rangle \\ &> 0\end{aligned}$$

  if $\lambda > \frac{1}{2} - \sqrt{\frac{1}{4} + \frac{1}{\overline{\gamma}}}$. Hence, there can not be any eigenvalues in the interval $\left(\max\left\{-1, \frac{1}{2} - \sqrt{\frac{1}{4} + \frac{1}{\overline{\gamma}}}\right\}, 0\right)$.

- For $\lambda < -1$ we find that $1+\lambda$, $\lambda < 0$ and $\frac{1}{1-\lambda} > 0$. Then, (47) yields that

$$\begin{aligned}\frac{1}{\alpha}(1+\lambda)\langle \mathsf{A}^{-1}\mathsf{A}_f\mathsf{A}^{-1}\mathsf{M}_\partial u, u\rangle &+ \frac{1}{1-\lambda}\langle u, u\rangle + \lambda\langle \mathsf{A}^{-1}\mathsf{A}_f u, u\rangle \\ &\leq \frac{1}{1-\lambda}\langle u, u\rangle + \lambda\underline{\gamma}\langle u, u\rangle \\ &= \left(\frac{1}{1-\lambda} + \lambda\underline{\gamma}\right)\langle u, u\rangle \\ &< 0\end{aligned}$$

  if $\lambda < \frac{1}{2} - \sqrt{\frac{1}{4} + \frac{1}{\underline{\gamma}}}$. Therefore, no eigenvalues are less than $\min\left\{-1, \frac{1}{2} - \sqrt{\frac{1}{4} + \frac{1}{\underline{\gamma}}}\right\}$.

We conclude that negative eigenvalues must belong to the interval

$$\left[\min\left\{-1, \frac{1}{2} - \sqrt{\frac{1}{4} + \frac{1}{\underline{\gamma}}}\right\}, \max\left\{-1, \frac{1}{2} - \sqrt{\frac{1}{4} + \frac{1}{\overline{\gamma}}}\right\}\right].$$

This finishes the proof of a).
■

Proof of b)

If $u|_{\partial\Omega_f} = 0$, then the extension operator E generates a zero extension, i.e. $Eu|_{\Omega_0} = 0$. Consequently,

$$\mathsf{M}_\partial u = 0,$$

see (29). Hence, (46) takes the form

$$\frac{1}{\lambda - 1}\mathsf{A}\mathsf{A}_f^{-1}\mathsf{A}u - \lambda\mathsf{A}u = 0$$

or

$$\frac{1}{\lambda - 1}\mathsf{A}u - \lambda\mathsf{A}_f u = 0. \tag{48}$$

Since $Eu|_{\Omega_0} = 0$ and $Eu|_{\Omega_f} = u$, we find from (31) and (32) that

$$\langle \mathsf{A}u, u\rangle = (Eu, Eu)_{H^1(\Omega)} = (u, u)_{H^1(\Omega_f)} = \langle \mathsf{A}_f u, u\rangle,$$

which, combined with (48), yields that

$$\left(\frac{1}{\lambda - 1} - \lambda\right)\langle \mathsf{A}_f u, u\rangle = 0.$$

Recall that A_f is positive, and therefore λ must satisfy

$$\frac{1}{\lambda - 1} - \lambda = 0$$

or $\lambda = (1 \pm \sqrt{5})/2$.

■

Proof of c)

Let (u, f) be an eigenvector with eigenvalue $\lambda = -1$. Then, (44) becomes

$$\mathsf{A}f = -\mathsf{A}u$$

or

$$f = -u.$$

Inserting this into (45) yields that

$$\mathsf{A}u - \mathsf{A}_f u = \mathsf{A}_f u,$$

which we may express on the form

$$\mathsf{A}u - 2\mathsf{A}_f u = 0.$$

This implies that, see (31) and (32),

$$(Eu, E\phi)_{H^1(\Omega)} - 2(u, \phi)_{H^1(\Omega_f)} = 0 \quad \forall\phi \in V_h.$$

Since $Eu|_{\Omega_f} = u$ and $\Omega_0 = \Omega \setminus \Omega_f$, it follows that

$$(Eu, E\phi)_{H^1(\Omega_0)} - (u, \phi)_{H^1(\Omega_f)} = 0 \quad \forall\phi \in V_h. \tag{49}$$

Evidently, if $u \neq 0$ solves (49) and $f = -u$, then (u, f) will satisfy (44)-(45) with $\lambda = -1$. This completes the argument.

■

10 Summary, discussion and conclusions

We have introduced a robust preconditioner for a PDE-constrained optimization problem with local control and with boundary observations only. This extends previous results, which have mainly focused on optimization tasks for globally defined controls and global observations.

The state equation of our model problem is elliptic, and the robust preconditioning strategy is derived by employing the "natural" Hilbert space for this equation. More specifically, the solution of the state equation is Helmholtz-harmonic on the complement of the support $\Omega_f \subset\subset \Omega$ of the control. Consequently, the Sobolev norm of the functions belonging to this Hilbert space is equivalent to the Sobolev norm associated with Ω_f. Based on this observation, we can define a preconditioner which is robust with respect to the size of the regularization parameter $\alpha > 0$. Furthermore, this approach enables us to significantly reduce the size of the KKT system: All unknowns are only defined on the support Ω_f of the control, and the Lagrange multiplier can be removed from the problem, which yields a 2×2 block-system, instead of a 3×3 system. For our model problem, employing rather fine meshes and with a relatively small regularization parameter, we obtained a reduction of the computing time by a factor in the range $[20, 50]$ - only recording the CPU-time needed to solve the KKT systems.

Prior to solving the 2×2 block system, a Helmholtz-harmonic extension of each FEM basis function, associated with nodes at $\partial\Omega_f$, must be computed. This is, of course, a negative aspect of our methodology. But this process can be fully parallelized with optimal speed-up. Also, if purely serial computations are employed, one will typically have very good initial guesses for an iterative scheme, e.g. CG, for computing these extensions: The Helmholtz-harmonic extensions of neighboring FEM basis functions can be used as initial guesses. Moreover, the benefits of our scheme increases as the size of the support Ω_f of the control decreases, because fewer extensions must be determined and the reduction of the number of unknowns increases.

From an inverse problem perspective, it is certainly not sufficient to solve the KKT system once with one particular choice of the size of the regularization parameter α. On the contrary, since the noise level of the data typically is unknown, a series of KKT systems, with varying degree of regularization, must be solved in order to determine a close-to-optimal value for α, see e.g. [3]. In such a process, it is only necessary to compute the above-mentioned Helmholtz-harmonic extensions once, and the speed-up obtained by solving the 2×2 KKT system, using our α robust preconditioner, will be large. A similar beneficial situation will occur if it is desirable to solve the PDE-constrained optimization problem with a number of different observation data sets, i.e. solving many KKT systems with different data d. This will be the case for a number of inverse problems arising in the engineering sciences. For example, for the inverse problem of electrocardiography (ECG).

By employing tailored multigrid schemes, we also expect that it might be possible to avoid the preprocessing step of explicitly computing the Helmholtz-harmonic extensions. This turned out to be very difficult to achieve by using standard multigrid software packages, and must be regarded as an open problem.

Our investigation only concerns elliptic control problems, and we assume that the unknown control belongs to H^1. If one insists on searching for a L^2-control, which would allow discontinuities, we do not know how to construct a robust preconditioner for the associated KKT system, but one could try to generalize the approach presented in [6]. It is also an open question how to obtain similar results for parabolic state equations or other relevant PDE models.

References

[1] I. Babuška. Error-bounds for finite element method. *Numerische Mathematik*, 16(19):322–333, 1971.

[2] A. Günnel, R. Herzog, and E. Sachs. A note on preconditioners and scalar products in Krylov subspace methods for self-adjoint problems in Hilbert space. *Electronic Transactions on Numerical Analysis*, 41:13–20, 2014.

[3] P. C. Hansen. *Discrete Inverse Problems: Insight and Algorithms.* SIAM, 2010.

[4] R. Herzog and E. Sachs. Superlinear convergence of Krylov subspace methods for self-adjoint problems in Hilbert space. Technische Universität Chemnitz, 2014.

[5] K. A. Mardal and J. B. Haga. Block preconditioning of systems of PDEs. In A. Logg, K. A. Mardal, and G. Wells, editors, *Automated Solution of Differential Equations by the Finite Element Method*, pages 643–654. Springer, 2012.

[6] K. A. Mardal, B. F. Nielsen, and Nordaas. M. Robust preconditioners for PDE-constrained optimization with limited observations. *Submitted*, 2014.

[7] K. A. Mardal and R. Winther. Preconditioning discretizations of systems of partial differential equations. *Numerical Linear Algebra with Applications*, 18(1):1–40, 2011.

[8] M. F Murphy, G. H. Golub, and A. J. Wathen. A note on preconditioning for indefinite linear systems. *SIAM Journal on Scientific Computing*, 21(6):1969–1972, 2000.

[9] B. F. Nielsen and K. A. Mardal. Efficient preconditioners for optimality systems arising in connection with inverse problems. *SIAM Journal on Control and Optimization*, 48(8):5143–5177, 2010.

[10] 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*, 33:1126–1152, 2012.

[11] J. W. Pearson and A. J. Wathen. A new approximation of the Schur complement in preconditioners for PDE-constrained optimization. *Numerical Linear Algebra with Applications*, 19:816–829, 2012.

[12] J. Schöberl and W. Zulehner. Symmetric indefinite preconditioners for saddle point problems with applications to PDE-constrained optimization problems. *SIAM Journal on Matrix Analysis and Applications*, 29(3):752–773, 2007.

[13] W. Zulehner. Nonstandard norms and robust estimates for saddle point problems. *SIAM Journal on Matrix Analysis and Applications*, 32:536–560, 2011.