DOI: 10.2478/auom-2018-0024
An. cSt. Univ. Ovidius Constancta
sciendo
Vol. 26(2),2018, 175–192

PRECONDITIONED DIRICHLET-DIRICHLET METHODS FOR OPTIMAL CONTROL OF ELLIPTIC PDE

Daniel Loghin

Abstract

The discretization of optimal control of elliptic partial differential equations problems yields optimality conditions in the form of large sparse linear systems with block structure. Correspondingly, when the solution method is a Dirichlet-Dirichlet non-overlapping domain decomposition method, we need to solve interface problems which inherit the block structure. It is therefore natural to consider block preconditioners acting on the interface variables for the acceleration of Krylov methods with substructuring preconditioners.

In this paper we describe a generic technique which employs a preconditioner block structure based on the fractional Sobolev norms corresponding to the domains of the boundary operators arising in the matrix interface problem, some of which may include a dependence on the control regularization parameter. We illustrate our approach on standard linear elliptic control problems. We present analysis which shows that the resulting iterative method converges independently of the size of the problem. We include numerical results which indicate that performance is also independent of the control regularization parameter and exhibits only a mild dependence on the number of the subdomains.

Key Words: Domain decomposition, Steklov-Poincaré operator, PDE-constrained optimization, preconditioned Krylov subspace methods.

2010 Mathematics Subject Classification: 65K10, 65N55, 65F10, 90C51, 49N90.

Received: March, 2017.

Revised: April, 2017.

Accepted: July, 2017.

1 Introduction

Optimal control problems with PDE constraints represent an important class of optimization problem with applications covering virtually every field in science and engineering. It is typical that these are challenging problems, particularly when nonlinearity is present or when the constraints are not of equality type. Another equally important challenge is the size of the resulting linear systems; often solutions can only be achieved iteratively, using some of the standard methodology (multigrid, Krylov subspace solvers, domain decomposition). We review briefly below some existing work in this regard.

1.1 Preconditioned Krylov methods

Given the block structure of the linear system, it is natural to devise block preconditioners. Such approaches are included in [26], [25], [20], [28], [21], [19], [24]. While they offer various generic options of block preconditioners, in all cases it is the distributed elliptic control problem with constant coefficients that is invariably considered. Some exceptions are represented by [13] (abstract formulation with state and control constraints), [12] (gradient inequality constraints), [16] (boundary control).

Block preconditioners are also extended to some other systems: Stokes flow [30], [18], Navier-Stokes flow [4], [5], although a different approach is used (the reduced space method), which involves elimination of some of the variables; the problem is also a nonlinear boundary control problem.

1.2 Multigrid methods

The use of multigrid methods for PDE constrained optimization problems is outlined in the SIAM review paper [6]. The main approach for multigrid to work is to reduce the problem to a positive-definite formulation, so that the standard formulation is applicable. The reduction is not always achievable, so the technique is not guaranteed to be available. Instead, one could work with a technique called one-shot multigrid, first introduced by Ta'asan [29], which is essentially a monolithic method that applies multigrid directly to the system of equations arising from first order optimality conditions (also known as the Karush-Kuhn-Tucker system), without decoupling or solving smaller problems. The technique requires a so-called collective smoothing approach, which may be problem dependent. Note that another occurence of multigrid methods for PDE-constrained optimization is as black-box solvers for block matrix preconditioners inside a preconditioned Krylov method approach (see, e.g. [19]).

1.3 Domain decomposition methods (DD)

The literature describing domain decomposition methods for PDE-constrained optimization is more limited. In fact, even for the Stokes problem (constrained Laplacian), the methods described are technical and often not optimal. Main contributions use off-the-shelf solvers (Schwarz approaches [22], [23], PETSc [4], [10, Ch. 16]). The only attempts to tackle the model problem using advanced DD methodology are the balancing Neumann-Neumann approaches in [8], [9], [3], followed a decade later by a short FETI-DP contribution for linear elasticity control [11]. For a recent contribution with a substructuring approach for topology optimization see [14].

2 Problem formulation

We start by introducing some notation, followed by the problem formulation of our optimal control problem.

2.1 Notation and definitions

Throughout the paper we will use the following notation and standard results. Given an open simply-connected domain U in $\mathbb{R}^2$, its boundary will be denoted by ∂U. We denote by $C_0^\infty(U)$ the space of infinitely differentiable functions defined on U with compact support in U. We will also denote by $L^2(U)$ the Lebesgue space of square-integrable functions defined on U endowed with inner-product $(\cdot,\cdot)$ and by $H^m(U)$ the Sobolev space of order m equipped with norm $\|\cdot\|_{m,U}$ and semi-norm $|\cdot|_{m,U}$ with the convention $H^0(U) = L^2(U)$. The Sobolev spaces of real index $0 \le s \le m$ are defined as interpolation spaces of index $\theta = 1 - s/m$ for the pair $[H^m(U), L^2(U)]$

$$H^s(\Omega) := [H^m(U), L^2(U)]_\theta \qquad \theta = 1 - s/m.$$

For any s, the space $H_0^s(U)$ denotes the completion of $C_0^\infty(U)$ in $H^s(U)$ (see e.g. [15, p 60]). In particular, we shall be interested in the interpolation space

$$H^{1/2}(U) = [H^1(U), L^2(U)]_{1/2}.$$

for which there holds $H_0^{1/2}(U) \equiv H^{1/2}(U)$. Another space of interest is $H_{00}^{1/2}(U)$ which is a subspace of $H_0^{1/2}(U)$ and is defined as the interpolation space of index $1/2$ for the pair $[H_0^1(U), L^2(U)]$

$$H_{00}^{1/2}(U) = [H_0^1(U), H^0(U)]_{1/2}.$$

Norms on $H^{1/2}(U), H_{00}^{1/2}(U)$ will be denoted by the same notation $|\cdot|_{1/2,U}$ or $\|\cdot\|_{1/2,U}$, with the assumption that it is evident from the context which space is under consideration.

2.2 Optimal control problems with elliptic constraints

Let Ω denote an open bounded domain in $\mathbb{R}^2$ with Lipschitz boundary $\partial\Omega$ and let $\Omega_d \subseteq \Omega$. Consider the following linear distributed control problem

$$\text{PDECO}: \begin{cases} \text{Find } (y,u) \in H_0^1(\Omega) \times L^2(\Omega_d) \text{ such that} \\ \quad \dfrac{1}{2}\|y - y_d\|^2_{L^2(\Omega_d)} + \dfrac{\alpha^2}{2}\|u\|^2_{L^2(\Omega)} = \min! \\ \text{subject to} \\ \qquad \begin{aligned} Ły &= f + u && \text{in } \Omega, \\ y &= 0 && \text{on } \partial\Omega, \end{aligned} \end{cases}$$

where the unknowns are the state y and the control u while the problem data is given by the desired state $y_d \in L^2(\Omega)$ and the PDE data $f \in L^2(\Omega)$, with α^2 a regularization parameter. $\mathcal{L}$ is a general second-order elliptic operator; for the purpose of a simpler exposition, we will assume throughout that $Ł = -\Delta$.

Let now $V = H_0^1(\Omega), Q = L^2(\Omega)$ and consider the weak formulation of the above control problem, with the constant term $\|y_d\|^2_{L^2(\Omega_d)}$ ignored:

$$\text{WF}: \begin{cases} \text{Find } (y,u) \in V \times Q \text{ such that} \\ \quad \dfrac{1}{2}m(y,y) - c(y) + \dfrac{\alpha^2}{2}q(u,u) = \min! \\ \text{subject to} \\ \qquad a(y,v) - b(u,v) = \ell(v) \quad \forall v \in V, \end{cases}$$

where

$$m(y,z) = (y,z)_{L^2(\Omega_d)}\,,\ q(w,v) = (w,v)_{L^2(\Omega)}\,,\ b(w,v) = (w,v)_{L^2(\Omega)}\,,$$

$$\ell(v) = (f,v)_{L^2(\Omega)}\,,\ c(z) = (y_d,z)_{L^2(\Omega)}\,,\ a(z,v) = (\nabla z, \nabla v)_{L^2(\Omega)}\,.$$

Let now $V_h \subset V$, $Q_h \subset Q$ denote suitable finite dimensional spaces. With the

above notation, the discrete weak formulation reads

$$\text{DWF}: \begin{cases} \text{Find } (y_h, u_h) \in V_h \times Q_h \text{ such that} \\ \quad \frac{1}{2} m(y_h, y_h) - c(y_h) + \frac{\alpha^2}{2} q(u_h, u_h) = \min! \\ \quad \text{subject to} \\ \quad a(y_h, v_h) - b(u_h, v_h) = \ell(v_h) \quad \forall v_h \in V_h, \end{cases}$$

which in turn yields the quadratic programming problem

$$\text{QP}: \begin{cases} \text{Find } \mathbf{y} \in \mathbb{R}^n, \mathbf{u} \in \mathbb{R}^m \text{ such that} \\ \quad \frac{1}{2} \mathbf{y}^T M \mathbf{y} - \mathbf{c}^T \mathbf{y} + \frac{\alpha^2}{2} \mathbf{u}^T Q \mathbf{u} = \min! \\ \text{subject to} \\ \quad A\mathbf{y} - B\mathbf{u} = \mathbf{b}, \end{cases}$$

where $A, M \in \mathbb{R}^{n \times n}, B \in \mathbb{R}^{n \times m}, Q \in \mathbb{R}^{m \times m}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^n$. The first order conditions for the above QP problem yield the following linear system of equations:

$$\begin{bmatrix} M & 0 & A^T \\ 0 & \alpha^2 Q & -B^T \\ A & -B & 0 \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \mathbf{u} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} \mathbf{c} \\ \mathbf{0} \\ \mathbf{b} \end{bmatrix}$$

When $\dim V_h = \dim Q_h = n$ and $\Omega_d = \Omega$, the matrices M, Q, B are equal and invertible $M = Q = B \in \mathbb{R}^{n \times n}$ and the linear system can be reduced to positive-definite form

$$\begin{bmatrix} \alpha A & -M \\ M & \alpha A^T \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \hat{\mathbf{p}} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{b}} \\ \mathbf{c} \end{bmatrix},$$

with $\hat{\mathbf{p}} = \mathbf{p}/\alpha, \hat{\mathbf{b}} = \alpha \mathbf{b}$. We discuss next an approach that allows for the design of optimal domain decomposition solvers.

2.3 An auxiliary problem: reaction-diffusion system

Motivated by the above discussion, we consider the block linear system

$$K\mathbf{x} = \begin{bmatrix} \alpha A & -M \\ M & \alpha A \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \end{bmatrix}. \tag{1}$$

We associate with this problem the system of PDE

$$\begin{cases} -\alpha\Delta u_1 - u_2 &= f_1 \quad \text{in } \Omega, \\ u_1 - \alpha\Delta u_2 &= f_2 \quad \text{in } \Omega, \\ u_1 = u_2 &= 0 \quad \text{on } \partial\Omega. \end{cases}$$

Equivalently,

$$\text{RD} : \begin{cases} \mathcal{L}\mathbf{u} := -\alpha\Delta\mathbf{u} + S\mathbf{u} &= \mathbf{f} \quad \text{in } \Omega, \\ \mathbf{u} &= \mathbf{0} \quad \text{on } \partial\Omega, \end{cases} \qquad S = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

The weak formulation of the above problem reads

$$\begin{cases} \text{Find } \mathbf{u} \in V \times V \text{ such that for all } \mathbf{v} \in V \times V \\ B(\mathbf{u}, \mathbf{v}) = (\mathbf{f}, \mathbf{v}), \end{cases}$$

where $V = H_0^1(\Omega)$ and

$$B(\mathbf{w}, \mathbf{v}) = \int_\Omega \left(\alpha\nabla\mathbf{w} : \nabla\mathbf{v} + \mathbf{w}^T S\mathbf{v}\right) \mathrm{d}\Omega.$$

In the next section we consider a non-overlapping domain decomposition method for the solution of the linear system arising from the discretization of the above weak formulation.

3 Domain decomposition formulation.

In the following we describe a standard domain decomposition (DD) strategy of our PDE system and consider the resulting discrete formulation.

3.1 Non-overlapping DD for reaction-diffusion systems.

Consider a non-overlapping decomposition of Ω:

$$\bar{\Omega} = \bigcup_{i=1}^{N} \bar{\Omega}_i, \qquad \Omega_i \cap \Omega_j \equiv \emptyset \quad (i \neq j),$$

and let $\Gamma \subset \mathbb{R}^{d-1}$ denote the set of internal boundaries associated with the above partition of Ω

$$\Gamma = \bigcup_{i=1}^{N} \Gamma_i \quad (\Gamma_i := \partial\Omega_i \setminus \partial\Omega).$$

We will use the notation $\mathbf{v}_i := \mathbf{v}\,|_{\Omega_i}$.

Let $\mathbf{u}_i\,|_{\Gamma_i} = \boldsymbol{\lambda}_i$. Our model problem decouples as

$$\begin{cases} Ł\mathbf{u}_i &= \mathbf{f} & \text{in } \Omega_i, \\ \mathbf{u}_i &= \mathbf{0} & \text{on } \partial\Omega \cap \partial\Omega_i, \\ \mathbf{u}_i &= \boldsymbol{\lambda}_i & \text{on } \Gamma_i. \end{cases}$$

Moreover, letting $\mathbf{u}_i = \mathbf{u}_i^{\{1\}} + \mathbf{u}_i^{\{2\}}$, the above system will split as

$$\begin{cases} Ł\mathbf{u}_i^{\{1\}} &= \mathbf{f} & \text{in } \Omega_i, \\ \mathbf{u}_i^{\{1\}} &= \mathbf{0} & \text{on } \partial\Omega \cap \partial\Omega_i, \\ \mathbf{u}_i^{\{1\}} &= \mathbf{0} & \text{on } \Gamma_i. \end{cases} \qquad \begin{cases} Ł\mathbf{u}_i^{\{2\}} &= \mathbf{0} & \text{in } \Omega_i, \\ \mathbf{u}_i^{\{2\}} &= \mathbf{0} & \text{on } \partial\Omega \cap \partial\Omega_i, \\ \mathbf{u}_i^{\{2\}} &= \boldsymbol{\lambda}_i & \text{on } \Gamma_i. \end{cases}$$

Multiplying by $\mathbf{v}_i \in [H_D^1(\Omega_i)]^2 = \left\{\mathbf{w}_i \in [H^1(\Omega_i)]^2 : \mathbf{w}_i|_{\partial\Omega} = 0\right\}$ and integrating we get

$$B(\mathbf{u}_i^{\{1\}}, \mathbf{v}_i) = (\mathbf{f}, \mathbf{v}_i) + \int_{\Gamma_i} \mathbf{n}_i \cdot \nabla \mathbf{u}_i^{\{1\}} \mathbf{v}_i \mathrm{d}s(\Gamma_i)$$

$$B(\mathbf{u}_i^{\{2\}}, \mathbf{v}_i) = \int_{\Gamma_i} \mathbf{n}_i \cdot \nabla \mathbf{u}_i^{\{2\}} \mathbf{v}_i \mathrm{d}s(\Gamma_i).$$

Adding and then summing over i we find

$$B(\mathbf{u}, \mathbf{v}) = (\mathbf{f}, \mathbf{v}) + \sum_{i=1}^{N} \int_{\Gamma_i} \mathbf{n}_i \cdot \nabla \mathbf{u}_i^{\{1\}} \mathbf{v}_i \mathrm{d}s(\Gamma_i) + \sum_{i=1}^{N} \int_{\Gamma_i} \mathbf{n}_i \cdot \nabla \mathbf{u}_i^{\{2\}} \mathbf{v}_i \mathrm{d}s(\Gamma_i)$$

which yields the *Steklov-Poincaré equation* for $\boldsymbol{\lambda}_i$:

$$\sum_{i=1}^{N} \int_{\Gamma_i} \mathbf{n}_i \cdot \nabla \mathbf{u}_i^{\{2\}} \mathbf{v}_i \mathrm{d}s(\Gamma_i) = -\sum_{i=1}^{N} \int_{\Gamma_i} \mathbf{n}_i \cdot \nabla \mathbf{u}_i^{\{1\}} \mathbf{v}_i \mathrm{d}s(\Gamma_i).$$

Let $\boldsymbol{\eta} \in \Lambda := [H_{00}^{1/2}(\Gamma)]^2$ and let $\mathbf{v}_i$ be the solution to the problem

$$\begin{cases} Ł\mathbf{v}_i &= 0 & \text{in } \Omega_i, \\ \mathbf{v}_i &= 0 & \text{on } \partial\Omega_i \setminus \Gamma_i, \\ \mathbf{v}_i &= \boldsymbol{\eta_i} & \text{on } \Gamma_i. \end{cases}$$

We call $\mathbf{v}_i$ the Ł-extension of the boundary data $\boldsymbol{\eta}_i$ to Ω_i:

$$\mathbf{v}_i = \mathcal{G}_i \boldsymbol{\eta}_i.$$

The *Steklov-Poincaré matrix operator* $\mathcal{S} : \Lambda \to \Lambda'$ is defined via

$$(\mathcal{S}\boldsymbol{\eta}, \boldsymbol{\mu})_{\cdot} = \sum_{i=1}^{N} \int_{\Gamma_i} \mathbf{n}_i \cdot \nabla(\mathcal{G}_i \boldsymbol{\eta}_i)\, \boldsymbol{\mu}_i \mathrm{d}s(\Gamma_i) =: \sum_{i=1}^{N} (\mathcal{S}_i \boldsymbol{\eta}_i, \boldsymbol{\mu}_i)_{\cdot}$$

We note here that $\mathcal{G}_i$ are matrix Green's functions for $\boldsymbol{\mathcal{L}}$ restricted to Ω_i. The following result describes the continuity and coercivity properties of $\mathcal{S}$.

Lemma 3.1. *There exist constants α_1, α_2 such that for all $\boldsymbol{\eta}, \boldsymbol{\mu} \in \Lambda$*

$$\alpha_1 \|\boldsymbol{\eta}\|_\Lambda^2 \leq (\mathcal{S}\boldsymbol{\eta}, \boldsymbol{\eta}), \qquad (\mathcal{S}\boldsymbol{\eta}, \boldsymbol{\mu}) \leq \alpha_2 \|\boldsymbol{\eta}\|_\Lambda \|\boldsymbol{\mu}\|_\Lambda. \tag{2}$$

Proof. The proof is straightforward to adapt from [1]. □

Our domain decomposition problem is summarised below.

$$\text{(i)} \begin{cases} Ł\mathbf{u}_i^{\{1\}} & = \mathbf{f} \ \text{ in } \Omega_i, \\ \mathbf{u}_i^{\{1\}} & = \mathbf{0} \ \text{ on } \partial\Omega_i, \end{cases}$$

$$\text{(ii)} \begin{cases} \mathcal{S}\boldsymbol{\lambda} = -\displaystyle\sum_{i=1}^{N} \mathbf{n}_i \cdot \nabla \mathbf{u}_i^{\{1\}} \ \text{ on } \Gamma, \end{cases}$$

$$\text{(iii)} \begin{cases} Ł u_i^{\{2\}} & = \mathbf{0} \ \text{ in } \Omega_i, \\ \mathbf{u}_i^{\{2\}} & = \boldsymbol{\lambda}_i \ \text{ on } \Gamma_i. \\ \mathbf{u}_i^{\{2\}} & = \mathbf{0} \ \text{ on } \partial\Omega_i \setminus \Gamma_i. \end{cases}$$

The resulting solution is

$$\mathbf{u}|_{\Omega_i} = \mathbf{u}_i^{\{1\}} + \mathbf{u}_i^{\{2\}}.$$

3.2 Matrix formulation

Consider now the linear system (1) permuted to the block form

$$K\mathbf{u} = \begin{bmatrix} K_{II} & K_{IB} \\ K_{BI} & K_{BB} \end{bmatrix} \begin{bmatrix} \mathbf{u}_I \\ \mathbf{u}_B \end{bmatrix} = \begin{bmatrix} \mathbf{f}_I \\ \mathbf{f}_B \end{bmatrix}$$

where

$$K_{II} = \bigoplus_{i=1}^{N} \begin{bmatrix} \alpha A_{I_i I_i} & -M_{I_i I_i} \\ M_{I_i I_i} & \alpha A_{I_i I_i} \end{bmatrix}$$

and $\mathbf{u}_I \in \mathbb{R}^{n_I}, \mathbf{u}_B \in \mathbb{R}^{n_B}$ with

$$n_I = \sum_{i=1}^{N} n_{I_i}, \quad n_I + n_B = n.$$

Let

$$S = K_{BB} - K_{BI} K_{II}^{-1} K_{IB}.$$

Steps (i–iii) in our DD approach above represent a Schur complement approach:

$$\text{(i) } K_{II}\mathbf{u}_I^{\{1\}} = \mathbf{f}_I,$$

$$\text{(ii) } S\mathbf{u}_B = \mathbf{f}_B - K_{BI}\mathbf{u}_I^{\{1\}}.$$

$$\text{(iii) } K_{II}\mathbf{u}_I^{\{2\}} = -K_{IB}\mathbf{u}_B.$$

where S is the discrete representation of $\mathcal{S}$. We note that steps (i) and (iii) involve N decoupled linear systems which can be solved in parallel. The crux of the problem is step (ii) which in general cannot be solved via a direct solver with optimal complexity, as S is a dense matrix of size increasing with N as well as n. However, if we approach this problem in an iterative context, sparse approximations to S can be devised and employed successfully. We describe next a useful preconditioner for S based on the properties inherited from the Steklov-Poincaré operator. We first review briefly the concept of discrete interpolation norms; these are norms which arise naturally when considering fractional Sobolev spaces, as is the case here (cf. Eqn (2)).

3.3 Discrete interpolation norms

Let $X_h = (V_h, \|\cdot\|_X), Y_h = (V_h, \|\cdot\|_Y)$ denote two finite-dimensional spaces with

$$V_h = \text{span}\,\{\phi_i\}_{1\le i\le n}$$

and

$$\|u_h\|_X = \|\mathbf{u}\|_{H_X}, \quad \|u_h\|_Y = \|\mathbf{u}\|_{H_Y},$$

where

$$(H_X)_{ij} = (\phi_i, \phi_j)_X, \quad (H_Y)_{ij} = (\phi_i, \phi_j)_Y.$$

We define for all $\theta \in [0,1]$ [2]

$$\|\mathbf{u}\|_{H_\theta}^2 = \|\mathbf{u}\|_{H_Y}^2 + \|E^{1-\theta}\mathbf{u}\|_{H_Y}^2.$$

where $E = J^{1/2} := (H_Y^{-1}H_X)^{1/2}$. Then

$$H_\theta = H_Y + (E^{1-\theta})^T H_Y E^{1-\theta} = H_Y + H_Y J^{1-\theta}.$$

The above derivation follows the definition of interpolation norms of Lions and Magenes [15]. It is shown in [2] that the norm induced by H_θ and the Sobolev norm $\|\cdot\|_\theta$ are equivalent over the interpolation space $[X_h, Y_h]_\theta$. Given that the properties in Lemma 3.1 involve fractional Sobolev spaces, we are led to consider the following discrete interpolation norms.

Let $U \subset \mathbb{R}^d$. Let $X = H_0^1(U), Y = L^2(U)$ and consider the interpolation space $H_{00}^{1/2}(U) = [H_0^1(U), L^2(U)]_{1/2}$. Let $V_h = \text{span}\,\{\phi_i\} \subset H_0^1(U)$ and let

$$(L_0)_{ij} = (\phi_i, \phi_j)\,, \quad (L_1)_{ij} = (\nabla\phi_i, \nabla\phi_j)$$

Then $J = L_0^{-1}L_1$ and

$$H_{1/2} = L_0 + L_0 J^{1/2} = L_0 + L_0(L_0^{-1}L_1)^{1/2}.$$

We show below that the availability of the discrete interpolation norm induced by $H_{1/2}$ allows for the construction of optimal preconditioners for our optimal control problem.

3.4 Generalized Lanczos approximations

The computation of $H_{1/2}^{-1}\mathbf{r}$, where $\mathbf{r} \in \mathbb{R}^{n_B}$, which arises in the implementation of our preconditioner (see next subsection) can be achieved via a sparse procedure by using a generalized Lanczos method. This approach was outlined in [1]; we include a brief summary below. Given matrices L_0, L_1, the generalised Lanczos algorithm applied to the stencil $[L_1^{-1}, L_0^{-1}]$ constructs $L_1^{-1}-$orthogonal vectors $\mathbf{v}_i$ (Lanczos vectors) such that

$$L_0^{-1}V_k = L_1^{-1}V_k T_k + \beta_{k+1}L_1^{-1}\mathbf{v}_{k+1}\mathbf{e}_k^T, \quad V_k^T L_1^{-1} V_k = I_k$$

where T_k is a symmetric positive definite tridiagonal matrix of small size k. Given a vector $\mathbf{r}$, the computation $H_{1/2}^{-1}\mathbf{r}$ can be approximated by

$$H_{1/2}^{-1}\mathbf{r} \approx \mathbf{z} := L_0^{-1}V_k(I_k + T_k^{1/2})^{-1}\mathbf{e}_1\|\mathbf{r}\|_{L_1^{-1}}. \tag{3}$$

The above expression involves sparse computations (inverses of nearly tridiagonal matrices L_0, L_1) with the exception of the matrix square-root computation on a small matrix of size $k\times k$ and the multiplication by $V_k \in \mathbb{R}^{n_B\times k}$, both of which require dense algebra; however, this does not affect the complexity of order $O(n_B)$ of the overall procedure.

3.5 Interface preconditioners

Consider the equivalent problem

$$(KP^{-1})(P\mathbf{u}) = \mathbf{f}.$$

where P is the block preconditioner

$$P = \begin{bmatrix} K_{II} & K_{IB} \\ & P_{BB} \end{bmatrix}.$$

Note that if $P_{BB} = S$

$$KP^{-1} = \begin{bmatrix} I & \\ K_{BI}K_{II}^{-1} & I \end{bmatrix}.$$

and convergence is guaranteed in 2 iterations [17]. In this sense, this choice of preconditioner is optimal. As S corresponds to an operator acting on the interface Γ, we will refer to P_{BB} as an interface preconditioner.
Since Γ is a one-dimensional manifold in Ω, we adapt the above definition of $H_{1/2}$ as follows. Let $S_h = \operatorname{span}\{\phi_i\} \subset H_0^1(\Gamma)$ where ϕ_i are the restrictions to Γ of the basis elements used to represent a piecewise polynomial function on Ω. Define

$$(L_0)_{ij} = (\phi_i, \phi_j)_{L^2(\Gamma)}, \quad (L_1)_{ij} = (\nabla_\Gamma \phi_i, \nabla_\Gamma \phi_j)_{L^2(\Gamma)},$$

where ∇_Γ is the tangential gradient operator:

$$\nabla_\Gamma v = \nabla v - \mathbf{n}(\mathbf{n} \cdot \nabla v).$$

Then the corresponding discrete interpolation norm of index $1/2$ is

$$H_{1/2} = L_0 + L_0(L_0^{-1}L_1)^{1/2}.$$

A result similar to that of Lemma 3.1 holds for S. Since the lemma refers to the product space $\Lambda = [H^{1/2}(\Gamma)]^2$, we introduce the following block-diagonal matrix:

$$H := \begin{bmatrix} H_{1/2} & \\ & H_{1/2} \end{bmatrix}.$$

Thus, H can be viewed as a matrix inducing a vector discrete fractional Sobolev norm of index $1/2$.

Proposition 3.2. *There exist constants $\tilde{\alpha}_1, \tilde{\alpha}_2$ such that for all $\boldsymbol{\beta}, \boldsymbol{\rho} \in \mathbb{R}^{n_B}$*

$$\tilde{\alpha}_1 \|\boldsymbol{\beta}\|_H^2 \leq \boldsymbol{\beta}^T S \boldsymbol{\beta}, \qquad \boldsymbol{\beta}^T S \boldsymbol{\rho} \leq \tilde{\alpha}_2 \|\boldsymbol{\beta}\|_H \|\boldsymbol{\rho}\|_H.$$

Proof. Define $\Lambda_h = S_h \times S_h$ and let $\{\boldsymbol{\phi}_i\}_{1 \leq i \leq n_B}$ denote a product basis generated from the basis of S_h. Let $\boldsymbol{\eta}_h, \boldsymbol{\mu}_h \in \Lambda_h$ have representations

$$\boldsymbol{\eta}_h = \sum_{i=1}^{n_B} \boldsymbol{\beta}_i \phi_i, \quad \boldsymbol{\mu}_h = \sum_{i=1}^{n_B} \boldsymbol{\rho}_i \phi_i.$$

Note now that the discrete version of (2) holds with the same constants:

$$\alpha_1 \|\boldsymbol{\eta}_h\|_\Lambda^2 \leq (\mathcal{S}\boldsymbol{\eta}_h, \boldsymbol{\eta}_h), \qquad (\mathcal{S}\boldsymbol{\eta}_h, \boldsymbol{\mu}_h) \leq \alpha_2 \|\boldsymbol{\eta}_h\|_\Lambda \|\boldsymbol{\mu}_h\|_\Lambda.$$

The result then follows from the equivalence of $\|\cdot\|_\Lambda$ and $\|\cdot\|_H$ on $S_h \times S_h$, and by noting that

$$(\mathcal{S}\boldsymbol{\eta}_h, \boldsymbol{\eta}_h) = \boldsymbol{\beta}^T S \boldsymbol{\beta}, \quad (\mathcal{S}\boldsymbol{\eta}_h, \boldsymbol{\mu}_h) = \boldsymbol{\beta}^T S \boldsymbol{\rho}.$$

□

The relevance of this result is directly related to the convergence of the preconditioned GMRES method. The following result is adapted from [7] (see also [27]).

Proposition 3.3. *The residuals of the GMRES algorithm in the H^{-1}-inner product applied to a linear system with coefficient matrix S and right preconditioner H satisfy*

$$\frac{\|\mathbf{r}^k\|_{H^{-1}}}{\|\mathbf{r}^0\|_{H^{-1}}} \leq \left(1 - \frac{\tilde{\alpha}_1^2}{\tilde{\alpha}_2^2}\right)^{k/2}.$$

We therefore expect that preconditioned GMRES with right preconditioner P applied to linear system (1) will converge in a number of steps independent of the size of the problem. We illustrate this in the numerics section below.

4 Numerical experiments

Consider our model problem

$$\text{PDECO}: \begin{cases} \text{Find } (y,u) \in H_0^1(\Omega) \times L^2(\Omega) \text{ such that} \\ \dfrac{1}{2}\|y - y_d\|^2_{L^2(\Omega_d)} + \dfrac{\alpha^2}{2}\|u\|^2_{L^2(\Omega)} = \min! \\ \text{subject to} \\ \qquad -\Delta y + cy = f + u \quad \text{in } \Omega, \\ \qquad\qquad y = 0 \qquad\quad \text{on } \partial\Omega. \end{cases}$$

We experiment with the following test problems, both posed on $\Omega = (0,1)^2$.
Test problem 1 (Nguyen, Heikenschloss 2005, [8]):

$$f = 1, \quad c = 1, \quad y_d(x_1, x_2) = \sin \pi x_1 \sin \pi x_2.$$

Test problem 2 (Pearson, Wathen 2012, [20]):

$$f = 0, \quad c = 0, \quad y_d = \begin{cases} 1 & \text{in } \left(0, \frac{1}{2}\right)^2 =: \Omega_1, \\ 0 & \text{in } \Omega \setminus \bar{\Omega}_1. \end{cases}$$

We used a mixed finite element method for discretizing the weak formulation of the equivalent reaction-diffusion system RD (cf. end of section 1) corresponding to V_h being the space of piecewise linear polynomials defined on an isotropic mesh of triangles. We solved the linear system $K\mathbf{x} = \mathbf{f}$ using GMRES or flexible GMRES (fGMRES), as required, preconditioned by the block-triangular preconditioner

$$P = \begin{bmatrix} K_{II} & K_{IB} \\ & P_{BB} \end{bmatrix}.$$

We worked with two choices of P_{BB}:

- exact implementation: $P_{BB} = H$, where H is the discrete fractional Sobolev norm described in section 2.
- Lanczos approximation: $P_{BB} = H_{\text{Lan}}$, where H_{Lan} denotes the Lanczos approximation (3) of the inverses of the diagonal blocks of H to vectors arising in the preconditioning step of fGMRES.

The starting guess $\mathbf{x}^0$ was computed as the solution of the linear system

$$\begin{pmatrix} K_{II} & K_{IB} \\ O & P_{BB} \end{pmatrix} \mathbf{x}^0 = \mathbf{f}.$$

This choice ensures that the initial residual has the block form

$$\mathbf{r}^0 = \begin{pmatrix} \mathbf{0} \\ \mathbf{r}_2^0 \end{pmatrix}$$

so that the Arnoldi basis generated by GMRES has the same zero pattern. This leads to important savings, as the orthogonal basis generated is non-zero only on the constraint interface space. This was taken into account in our implementation of fGMRES, so that only the non-zero part of the Arnoldi basis was stored. We note that this is mathematically equivalent to using fGMRES on the Schur complement problem, but employing a global stopping criterion. The stopping criterion was $\|\mathbf{f} - K\mathbf{x}_m\|_2 / \|\mathbf{f} - K\mathbf{x}_0\|_2 < 10^{-6}$. The size of the Lanczos basis used in the approximation of H_{Lan}^{-1} was in all cases $k = 15$. We experimented with a range of domain subdivisions (N), mesh-sizes (n) and regularization parameters (α^2). The iteration counts are included in Tables 1, 3 for the exact implementation of P and in Tables 2, 4 for the Lanczos approximation. We observe that in all experiments the independence of the size of the problem is present, as indicated by the theory. We also observe that there is a mild dependence on the regularization parameter α^2. This is a parameter that was not considered in our analysis, just as the number

$\alpha^2 =$	1			10^{-2}			10^{-4}		
$N =$	4	16	64	4	16	64	4	16	64
$n =$ 12,675	13	16	21	14	16	22	16	18	21
33,282	13	16	22	14	17	22	16	18	22
132,098	13	16	22	14	17	22	16	18	22

Table 1: GMRES iterations for Problem 1 using $P_{BB} = H$.

$\alpha^2 =$	1			10^{-2}			10^{-4}		
$N =$	4	16	64	4	16	64	4	16	64
$n =$ 12,675	12	14	20	14	17	21	16	18	22
33,282	13	14	19	14	16	21	16	19	23
132,098	14	15	17	15	17	21	17	21	23

Table 2: fGMRES iterations for Problem 1 using $P_{BB} = H_{\text{Lan}}$.

of subdomains was not. However, for this latter parameter also, the number of iterations grows only slowly. We recall here that our method is a one-level method, without the application of a coarse grid solve as many DD approaches

$\alpha^2 =$	1			10^{-2}			10^{-4}		
$N =$	4	16	64	4	16	64	4	16	64
$n =$ 12,675	13	17	22	13	17	24	15	20	26
33,282	13	17	23	13	18	24	15	21	26
132,098	12	17	24	13	18	24	14	21	27

Table 3: GMRES iterations for Problem 2 using $P_{BB} = H$.

$\alpha^2 =$	1			10^{-2}			10^{-4}		
$N =$	4	16	64	4	16	64	4	16	64
$n =$ 12,675	12	16	22	13	18	23	15	20	26
33,282	13	16	22	14	18	23	15	22	27
132,098	14	17	20	14	19	23	15	22	28

Table 4: fGMRES iterations for Problem 2 using $P_{BB} = H_{\text{Lan}}$.

are known to require for scalability. Finally, we remark that the practical version of our preconditioner, given by the generalized Lanczos approximation outperforms in some cases the exact (but inefficient) version of the preconditioner. This is an occurence also noted in the case when the same procedure is applied for the domain decomposition solution of a discrete scalar elliptic problem [1].

5 Conclusion

We introduced a domain decomposition approach based on a reformulation of the first order conditions into a system related to a certain system of PDE (reaction diffusion system). The approach involves applying domain decomposition to this related system, with a key ingredient the interface preconditioner based on discrete fractional Sobolev norms. The resulting solver was analyzed and shown to have performance independent of the size of the problem. We found experimentally that the dependence on the other parameters in the problem is only mild. This makes for a robust and promising approach which we hope to extend to other types of PDE constrained optimization problems.

References

[1] M. Arioli, D. Kourounis, and D. Loghin. Discrete fractional Sobolev norms for domain decomposition preconditioning. *IMA J. Numer. Anal.*, 33(1):318–342, 2013.

[2] M. Arioli and D. Loghin. Discrete interpolation norms with applications. *SIAM J. Num. Anal.*, 47(4):2924–2951, 2009.

[3] Roscoe A. Bartlett, Matthias Heinkenschloss, Denis Ridzal, and Bart G. van Bloemen Waanders. Domain decomposition methods for advection dominated linear-quadratic elliptic optimal control problems. *Comput. Methods Appl. Mech. Engrg.*, 195(44-47):6428–6447, 2006.

[4] George Biros and Omar Ghattas. Parallel Lagrange-Newton-Krylov-Schur methods for PDE-constrained optimization. I. The Krylov-Schur solver. *SIAM J. Sci. Comput.*, 27(2):687–713 (electronic), 2005.

[5] George Biros and Omar Ghattas. Parallel Lagrange-Newton-Krylov-Schur methods for PDE-constrained optimization. II. The Lagrange-Newton solver and its application to optimal control of steady viscous flows. *SIAM J. Sci. Comput.*, 27(2):714–739 (electronic), 2005.

[6] Alfio Borzi and Volker Schulz. Multigrid methods for PDE optimization. *SIAM Rev.*, 51(2):361–395, 2009.

[7] H. C. Elman. *Iterative methods for large sparse non-symmetric systems of linear equations*. PhD thesis, Yale University, New Haven, 1982.

[8] Matthias Heinkenschloss and Hoang Nguyen. Balancing Neumann-Neumann methods for elliptic optimal control problems. In *Domain decomposition methods in science and engineering*, volume 40 of *Lect. Notes Comput. Sci. Eng.*, pages 589–596. Springer, Berlin, 2005.

[9] Matthias Heinkenschloss and Hoang Nguyen. Neumann-Neumann domain decomposition preconditioners for linear-quadratic elliptic optimal control problems. *SIAM J. Sci. Comput.*, 28(3):1001–1028, 2006.

[10] Michael A. Heroux, Padma Raghavan, and Horst D. Simon, editors. *Parallel processing for scientific computing*, volume 20 of *Software, Environments, and Tools*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2006.

[11] R. Herzog and O. Rheinbach. FETI–DP for optimal control problems. In J. Erhel, M.J. Gander, L. Halpern, G. Pichot, T. Sassi, and O.B. Widlund, editors, *Domain Decomposition Methods in Science and Engineering XXI*, Lecture Notes in Computational Science and Engineering. Springer International Publishing, 2014.

[12] Roland Herzog and Susann Mach. Preconditioned solution of state gradient constrained elliptic optimal control problems. *SIAM J. Numer. Anal.*, 54(2):688–718, 2016.

[13] Roland Herzog and Ekkehard Sachs. Preconditioned conjugate gradient method for optimal control problems with control and state constraints. *SIAM J. Matrix Anal. Appl.*, 31(5):2291–2317, 2010.

[14] M. Kovcvara, D. Loghin, and J. Turner. Constraint interface preconditioning for topology optimization problems. *SIAM J. Sci. Comput.*, 38(1):A128–A145, 2016.

[15] J. L. Lions and E. Magenes. *Problèmes aux limites non homogènes et applications. Vol. 1*. Travaux et Recherches Mathématiques, No. 17. Dunod, Paris, 1968.

[16] T.P. Mathew, M. Sarkis, and C.E. Schaerer. Analysis of block matrix preconditioners for elliptic optimal control problems. *Numerical Linear Algebra with Applications*, 14(4):257–279, 2007. cited By 13.

[17] M. F. Murphy, G. H. Golub, and A. J. Wathen. A note on preconditioning for indefinite linear systems. *SIAM J. Sci. Comp.*, 21:1969–1972, 2000.

[18] John W. Pearson. On the development of parameter-robust preconditioners and commutator arguments for solving Stokes control problems. *Electron. Trans. Numer. Anal.*, 44:53–72, 2015.

[19] John W. Pearson, Martin Stoll, and Andrew J. Wathen. Preconditioners for state-constrained optimal control problems with Moreau-Yosida penalty function. *Numer. Linear Algebra Appl.*, 21(1):81–97, 2014.

[20] John W. Pearson and Andrew J. Wathen. A new approximation of the Schur complement in preconditioners for PDE-constrained optimization. *Numer. Linear Algebra Appl.*, 19(5):816–829, 2012.

[21] John W. Pearson and Andrew J. Wathen. Fast iterative solvers for convection-diffusion control problems. *Electron. Trans. Numer. Anal.*, 40:294–310, 2013.

[22] Ernesto E. Prudencio, Richard Byrd, and Xiao-Chuan Cai. Parallel full space SQP Lagrange-Newton-Krylov-Schwarz algorithms for PDE-constrained optimization problems. *SIAM J. Sci. Comput.*, 27(4):1305–1328 (electronic), 2006.

[23] Ernesto E. Prudencio and Xiao-Chuan Cai. Parallel multilevel restricted Schwarz preconditioners with pollution removing for PDE-constrained optimization. *SIAM J. Sci. Comput.*, 29(3):964–985, 2007.

[24] Yue Qiu, Martin B. van Gijzen, Jan-Willem van Wingerden, Michel Verhaegen, and Cornelis Vuik. Efficient preconditioners for PDE-constrained optimization problem with a multilevel sequentially semiseparable matrix structure. *Electron. Trans. Numer. Anal.*, 44:367–400, 2015.

[25] Tyrone Rees, H. Sue Dollar, and Andrew J. Wathen. Optimal solvers for PDE-constrained optimization. *SIAM J. Sci. Comput.*, 32(1):271–298, 2010.

[26] Tyrone Rees and Martin Stoll. Block-triangular preconditioners for PDE-constrained optimization. *Numer. Linear Algebra Appl.*, 17(6):977–996, 2010.

[27] Y. Saad. *Iterative Methods for Sparse Linear Systems*. PWS Publishing Co., Boston, 1996.

[28] Martin Stoll and Andy Wathen. Preconditioning for partial differential equation constrained optimization with control constraints. *Numer. Linear Algebra Appl.*, 19(1):53–71, 2012.

[29] Shlomo. Ta'asan, Institute for Computer Applications in Science, Engineering., and Langley Research Center. *"One Shot" methods for optimal control of distributed parameter systems I [microform] : finite dimensional control / Shlomo Ta'asan.* Institute for Computer Applications in Science and Engineering : NASA Langley Research Center ; National Technical Information Service, distributor Hampton, Va. : [Springfield, Va], 1991.

[30] Walter Zulehner. Nonstandard norms and robust estimates for saddle point problems. *SIAM J. Matrix Anal. Appl.*, 32(2):536–560, 2011.

Daniel LOGHIN,
School of Mathematics,
University of Birmingham,
Edgbaston, Birmingham B15 2TT, UK
Email: d.loghin@bham.ac.uk