Superlinear convergence of the GMRES for PDE-constrained optimization problems

by O. Axelsson[1], J. Karátson[2]

Abstract

Optimal control problems for PDEs arise in many important applications. A main step in the solution process is the solution of the arising linear system, where the crucial point is usually finding a proper preconditioner. We propose both proper block diagonal and more involved preconditioners, and derive mesh independent superlinear convergence of the preconditioned GMRES iterations based on a compact perturbation property of the underlying operators.

1 Introduction

Optimal control problems for PDEs, where we want to steer the solution of the modelled process close to some desired target solution by use of a control function, arise in many important applications. Such problems have been dealt with in several publications, such as [2, 3, 10, 16, 22], see also the references therein. Earlier publications have mostly dealt with problems when the control and observation domains coincide, however, in recent papers they may be allowed to be different. The general approaches are the discretize-then-optimize or optimize-then-discretize processes: recent research shows that one should use discretization schemes for which both approaches coincide. A main step in the solution process is the solution of the arising linear system, where the crucial point is usually finding a proper preconditioner.

We propose both proper block diagonal and more involved preconditioners. Mesh independent superlinear convergence is derived for the preconditioned GMRES iterations, based on a compact perturbation property of the underlying operators. These are new contributions to the topic, since previous results for such problems only studied linear convergence properties. The paper begins with the required preliminaries, then the new results are presented in detail for a time-independent distributed control problem, finally some related problems are mentioned in the last section.

2 Preliminaries

We elaborate our preconditioning approach for a time-independent distributed control problem, described below, where the control and observation domains are different. Further related problems will be mentioned in section 4.

[1]Institute of Geonics AS CR, Ostrava, Czech Republic
[2]Department of Applied Analysis & MTA-ELTE Numerical Analysis and Large Networks Research Group, ELTE University; Department of Analysis, Technical University; Budapest, Hungary

2.1 Formulation of the problem

We consider a time-independent distributed control problem, with target solution $\overline{y}$ and control function u, using H^1-regularization, as described in [10]. Let $\Omega \subset \mathbf{R}^d$ be a bounded domain, and Ω_1, Ω_2 given subsets of Ω: the observation region Ω_1 and the control region Ω_2. Minimize

$$J(y,u) := \frac{1}{2}\|y-\overline{y}\|^2_{L^2(\Omega_1)} + \frac{\beta}{2}\|u\|^2_{H^1(\Omega_2)} \tag{2.1}$$

subject to the PDE constraint

$$\begin{cases} -\Delta y = \begin{cases} u & \text{on } \Omega_2 \\ 0 & \text{on } \Omega \setminus \Omega_2 \end{cases} \\ y\big|_{\partial\Omega} = g. \end{cases} \tag{2.2}$$

Here g is a fixed boundary term that admits a Dirichlet lift $\tilde{g} \in H^1(\Omega)$, and $\beta > 0$ is a regularization constant.

This leads to the following system of PDEs in weak form for the state and control variables and the Lagrange multiplier:

find $y \in \tilde{g} + H^1_0(\Omega)$, $u \in H^1(\Omega_2)$, $\lambda \in H^1_0(\Omega)$ such that

$$\begin{aligned} &\int_{\Omega_1} y\mu - \int_{\Omega} \nabla\lambda\cdot\nabla\mu = \int_{\Omega_1}\overline{y}\mu \qquad (\forall \mu \in H^1_0(\Omega)), \\ &\beta\int_{\Omega_2}(\nabla u\cdot\nabla v + uv) + \int_{\Omega_2}\lambda v = 0 \qquad (\forall v \in H^1(\Omega_2)), \\ &\int_{\Omega}\nabla y\cdot\nabla z - \int_{\Omega_2} uz = 0 \qquad (\forall z \in H^1_0(\Omega)). \end{aligned} \tag{2.3}$$

The system can be homogenized, using the splitting $y = y_0 + \tilde{g}$ where $y_0 \in H^1_0(\Omega)$. Therefore, in what follows, we may assume that $g = 0$, and hence $y \in H^1_0(\Omega)$.

The finite element solution is then carried out in a usual way: we introduce suitable finite element subspaces

$$Y_h \subset H^1_0(\Omega), \quad U_h \subset H^1(\Omega_2), \quad \Lambda_h \subset H^1_0(\Omega)$$

and replace the solution and test functions in (2.3) with functions only in the above subspaces. Let us fix proper bases in the subspaces and denote by $\mathbf{y}$, $\mathbf{u}$ and $\boldsymbol{\lambda}$ the coefficient vectors of these finite element solutions. Then we obtain a systems of equations in the following form:

$$\begin{aligned} \mathbf{M}_y\mathbf{y} - \mathbf{K}\boldsymbol{\lambda} &= \overline{\mathbf{y}} \\ \beta(\mathbf{M_u} + \mathbf{K_u})\mathbf{u} + \mathbf{M}^T\boldsymbol{\lambda} &= \mathbf{0} \\ \mathbf{K}\mathbf{y} - \mathbf{M}\mathbf{u} &= \mathbf{0}, \end{aligned} \tag{2.4}$$

Here $\mathbf{M}_y$ and $\mathbf{M}_u$ are the mass matrices corresponding to the subdomains Ω_1 and Ω_2 (i.e. that are used to approximate y and u), and similarly, $\mathbf{K}$ and $\mathbf{K_u}$ are the stiffness matrices corresponding to Ω and Ω_2, respectively, further, the rectangular mass matrix

$\mathbf{M}$ corresponds to function pairs from $\Omega \times \Omega_2$. We note that $\boldsymbol{\lambda}$ and $\mathbf{y}$ have the same dimension, they both represent functions on Ω, whereas $\mathbf{u}$ only corresponds to nodepoints in Ω_2. We also note that the last r.h.s is $\mathbf{0}$ due to $g = 0$. In the general case $g \neq 0$ we would have some $\mathbf{g} \neq 0$ on the last r.h.s, i.e. non-homogenity would only affect the r.h.s. and our results would remain valid.

After rearrangement, we obtain in matrix form that

$$\begin{pmatrix} \mathbf{K} & -\mathbf{M} & \mathbf{0} \\ \mathbf{0} & \beta(\mathbf{M_u} + \mathbf{K_u}) & \mathbf{M}^T \\ -\mathbf{M}_y & \mathbf{0} & \mathbf{K} \end{pmatrix} \begin{pmatrix} \mathbf{y} \\ \mathbf{u} \\ \boldsymbol{\lambda} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \overline{\mathbf{y}} \end{pmatrix} \tag{2.5}$$

Problem (2.3) has a unique solution, as well as system (2.4). See [2, 3, 10] for more details on the problem.

Our goal is to define an efficient preconditioned iterative solution method for the above linear system, and to derive a mesh independent superlinear convergence rate. Previous work of the authors includes such superlinear estimates on coercive or complex-valued equations [4, 5, 6, 9]. The present paper includes its extension to indefinite real-valued systems.

2.2 Superlinear convergence of the GMRES

In what follows, we will need the solution of linear systems

$$Au = b \tag{2.6}$$

with a given nonsingular matrix $A \in \mathbf{R}^{n\times n}$. When A is large and sparse, one generally uses a Krylov type iterative method, see e.g. [1, 11, 21]. In this paper we are interested in superlinear convergence rates of the iteration. Here we summarize briefly the required background.

For the symmetric positive-definite case, the well-known superlinear estimate of the standard CG method is obtained as follows, see e.g. [1]. Let us consider the decomposition

$$A = I + E, \tag{2.7}$$

where I is the identity matrix, and let $\lambda_j(E) =: \mu_j$. Let us define the polynomial $P_k(\lambda) := \prod_{j=1}^{k} \left(1 - \frac{\lambda}{\lambda_j}\right)$, where $\lambda_j := \lambda_j(A)$ are ordered according to $|\lambda_j - 1|$, i.e. such that $|\mu_1| \geq |\mu_2| \geq ... \geq |\mu_n|$. Since $P_k(\lambda_i) = 0 \quad (i = 1, \ldots, k)$, and using that $|\mu_j - \mu_i| \leq 2|\mu_j|$ $(i \geq k+1, \quad 1 \leq j \leq k)$ and $\frac{1}{\lambda_j} \leq \|A^{-1}\|$, one obtains

$$\max_{\lambda\in\sigma(A)} |P_k(\lambda)| = \max_{i\geq k+1} |P_k(\lambda_i)| = \max_{i\geq k+1} \prod_{j=1}^{k} \frac{|\mu_j - \mu_i|}{\lambda_j} \leq \left(2\|A^{-1}\|\right)^k \prod_{j=1}^{k} |\mu_j| \tag{2.8}$$

where $\mu_j = \lambda_j - 1$. Using the minimax property of the CG method, (2.8) and the arithmetic-geometric means inequality, and returning to the notation $\lambda_j(E) = \mu_j$, we finally obtain that

$$\left(\frac{\|e_k\|_A}{\|e_0\|_A}\right)^{1/k} \leq \frac{2\|A^{-1}\|}{k} \sum_{j=1}^{k} \left|\lambda_j(E)\right| \qquad (k = 1, 2, ..., n). \tag{2.9}$$

In the present paper the matrix is nonsymmetric, for which also several Krylov algorithms exist, in particular, GMRES and its variants are most widely used. There exist similar efficient superlinear convergence estimates for the GMRES, based on the decomposition (2.7). In fact, the sharpest one has been proved in [17], using products of singular values and the residual error vectors $r_k := Au_k - b$, on the Hilbert space level for an invertible operator $A \in B(H)$. One has

$$\frac{\|r_k\|}{\|r_0\|} \leq \prod_{j=1}^{k} s_j(E) s_j(A^{-1}) \qquad (k = 1, 2, ...) \tag{2.10}$$

where the singular values for a general bounded operator are defined as the distances from the best approximations with rank less than j. Hence, clearly, $s_j(A^{-1}) \leq \|A^{-1}\|$ for all j, thus the right hand side (r.h.s.) above is bounded by $\left(\prod_{j=1}^{k} s_j(E)\right) \|A^{-1}\|^k$. Using the inequality between the geometric and arithmetic means, we obtain the following estimate:

$$\left(\frac{\|r_k\|}{\|r_0\|}\right)^{1/k} \leq \frac{\|A^{-1}\|}{k} \sum_{j=1}^{k} s_j(E) \qquad (k = 1, 2, ...), \tag{2.11}$$

where the r.h.s. is a sequence decresing towards zero.

3 Numerical solution and mesh-independent superlinear convergence

3.1 Discretization and block matrix formulations

We consider a finite element discretization of problem (2.3) as described in subsection 2.1 . The convergence of the finite element solutions to the exact one is ensured by the standard approximation property: denoting $V_h := Y_h \times U_h \times \Lambda_h$ for all considered $h > 0$, and letting n be the dimension of V_h,

$$\text{for any } \underline{x} \in \mathcal{H}, \quad \text{dist}(\underline{x}, V_h) := \min\{\|\underline{x} - \underline{w}_h\| : \ \underline{w}_h \in V_h\} \to 0 \qquad (\text{as } \ n \to \infty). \tag{3.1}$$

Let us denote by $\mathcal{A}_h$ the global stiffness matrix of system (2.5):

$$\mathcal{A}_h := \begin{pmatrix} \mathbf{K} & -\mathbf{M} & \mathbf{0} \\ \mathbf{0} & \beta(\mathbf{M_u} + \mathbf{K_u}) & \mathbf{M}^T \\ -\mathbf{M}_y & \mathbf{0} & \mathbf{K} \end{pmatrix} \tag{3.2}$$

and let us also use compressed notations for the solution vector and the r.h.s. as

$$\mathbf{c} := \begin{pmatrix} \mathbf{y} \\ \mathbf{u} \\ \boldsymbol{\lambda} \end{pmatrix}, \qquad \mathbf{b} := \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \overline{\mathbf{y}} \end{pmatrix}, \tag{3.3}$$

i.e. the system (2.5) which we wish to solve is

$$\mathcal{A}_h \mathbf{c} = \mathbf{b}. \tag{3.4}$$

We will denote the total DOF by n, i.e. the size of above system is $n \times n$.

Let us define the block diagonal and the split part, respectively:

$$\mathcal{S}_h := \begin{pmatrix} \mathbf{K} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \beta(\mathbf{M_u} + \mathbf{K_u}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{K} \end{pmatrix}, \qquad \mathcal{Q}_h := \mathcal{A}_h - \mathcal{S}_h = \begin{pmatrix} \mathbf{0} & -\mathbf{M} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{M}^T \\ -\mathbf{M}_y & \mathbf{0} & \mathbf{0} \end{pmatrix}. \tag{3.5}$$

By the definition of the used stiffness and mass matrices, we have the following relation between the above matrices and the underlying inner product $\langle ., . \rangle_{\mathcal{H}}$ and operators Q, L. Let

$$y, z \in Y_h, \quad u, v \in U_h, \quad \lambda, \mu \in \Lambda_h$$

be given functions and let $\mathbf{y}$, $\mathbf{z}$, $\mathbf{u}$, $\mathbf{v}$, $\boldsymbol{\mu}$ and $\boldsymbol{\lambda}$ be their coefficient vectors, respectively. Following (3.14) and (3.3), let

$$\underline{x} := \begin{pmatrix} y \\ u \\ \lambda \end{pmatrix}, \quad \underline{w} := \begin{pmatrix} z \\ v \\ \mu \end{pmatrix}, \quad \mathbf{c} := \begin{pmatrix} \mathbf{y} \\ \mathbf{u} \\ \boldsymbol{\lambda} \end{pmatrix} \quad \text{and} \quad \mathbf{d} := \begin{pmatrix} \mathbf{y} \\ \mathbf{u} \\ \boldsymbol{\lambda} \end{pmatrix}. \tag{3.6}$$

Then we have

$$\langle \underline{x}, \underline{w} \rangle_{\mathcal{H}} = \mathcal{S}_h \mathbf{c} \cdot \mathbf{d}, \qquad \langle Q\underline{x}, \underline{w} \rangle_{\mathcal{H}} = \mathcal{Q}_h \mathbf{c} \cdot \mathbf{d} \qquad \text{and} \qquad \langle L\underline{x}, \underline{w} \rangle_{\mathcal{H}} = \mathcal{A}_h \mathbf{c} \cdot \mathbf{d} \tag{3.7}$$

where $\cdot$ denotes the ordinary inner product on $\mathbf{R}^n$. Accordingly, the natural inner product on $\mathbf{R}^n$ for our problem is the $\mathcal{S}_h$-inner product.

Since $\mathcal{A}_h$ is regular, we note that it satisfies an inf-sup condition:

$$\inf_{\substack{\mathbf{c} \in \mathbf{R}^n \\ \mathbf{c} \neq \mathbf{0}}} \sup_{\substack{\mathbf{d} \in \mathbf{R}^n \\ \mathbf{d} \neq \mathbf{0}}} \frac{\mathcal{A}_h \, \mathbf{c} \cdot \mathbf{d}}{\|\mathbf{c}\|_{\mathcal{S}_h} \|\mathbf{d}\|_{\mathcal{S}_h}} =: \, m_h \, > 0 \tag{3.8}$$

where, on the other hand, m_h might in general depend on h.

3.2 Iterative solution and block diagonal preconditioning

We will use the block diagonal matrix $\mathcal{S}_h$ as preconditioner. Since $\mathcal{A}_h = \mathcal{S}_h + \mathcal{Q}_h$, we obtain $\mathcal{S}_h^{-1}\mathcal{A}_h = \mathcal{I}_h + \mathcal{S}_h^{-1}\mathcal{Q}_h$, where $\mathcal{I}_h$ denotes the identity matrix. Hence the preconditioned form of (3.4) becomes

$$(\mathcal{I}_h + \mathcal{S}_h^{-1}\mathcal{Q}_h)\, \mathbf{c} = \tilde{\mathbf{b}} \tag{3.9}$$

where $\tilde{\mathbf{b}} := \mathcal{S}_h^{-1}\mathbf{b}$. We apply a preconditioned GMRES method to solve (3.9). The preconditioner is based on the idea of equivalent operators [6, 12]. Let us introduce the uniformly positive elliptic operator

$$S \begin{pmatrix} y \\ u \\ \lambda \end{pmatrix} := \begin{pmatrix} -\Delta y \\ \beta(-\Delta u + u) \\ -\Delta \lambda \end{pmatrix} \qquad \text{for } \; y_{|\partial\Omega} = \lambda_{|\partial\Omega} = 0 \tag{3.10}$$

in the product space $\mathcal{H}$, where $\beta > 0$ is the constant used in (2.3). Then the stiffness matrix of S coincides with the diagonal preconditioner $\mathcal{S}_h$ introduced in (3.5). The auxiliary problems with $\mathcal{S}_h$ are thus discretizations of uncoupled positive definite elliptic equations with constant coefficients, and hence can be solved with an optimal order of the number of operations [15, 20]. Consequently, if we prove mesh independent rate of convergence, then the overall number of operations is also of optimal order.

As seen above, the preconditioned system takes the form (3.9), i.e. we have a counterpart of (2.7). Applying the GMRES algorithm for the matrix $A = \mathcal{S}_h^{-1}\mathcal{A}_h$ (with inverse $(\mathcal{S}_h^{-1}\mathcal{A}_h)^{-1} = \mathcal{A}_h^{-1}\mathcal{S}_h$) and inner product $\langle \mathbf{c}, \mathbf{d}\rangle_{\mathcal{S}_h} := \mathcal{S}_h\,\mathbf{c}\cdot\mathbf{d}$, we obtain the following counterpart of estimate (2.11):

$$\left(\frac{\|r_k\|_{\mathcal{S}_h}}{\|r_0\|_{\mathcal{S}_h}}\right)^{1/k} \le \frac{\|\mathcal{A}_h^{-1}\mathcal{S}_h\|_{\mathcal{S}_h}}{k} \sum_{i=1}^{k} s_i(\mathcal{S}_h^{-1}\mathcal{Q}_h) \qquad (k = 1, 2, ..., n). \tag{3.11}$$

Our goal is to give a bound on (3.11) that is independent of the subspaces $Y_h,\ U_h, \Lambda_h$. This will be shown by a suitable modification of our results in [4, 5].

3.3 Hilbert space background

We introduce the Hilbert space

$$\mathcal{H} := H_0^1(\Omega) \times H^1(\Omega_2) \times H_0^1(\Omega)$$

with inner product

$$\left\langle \begin{pmatrix} y \\ u \\ \lambda \end{pmatrix}, \begin{pmatrix} z \\ v \\ \mu \end{pmatrix} \right\rangle_{\mathcal{H}} := \langle y, z\rangle_{H_0^1(\Omega)} + \langle u, v\rangle_{H^1(\Omega_2)} + \langle \lambda, \mu\rangle_{H_0^1(\Omega)},$$

where

$$\langle y, z\rangle_{H_0^1(\Omega)} := \int_\Omega \nabla y \cdot \nabla z, \qquad \langle u, v\rangle_{H^1(\Omega_2)} := \beta \int_{\Omega_2} (\nabla u \cdot \nabla v + uv)$$

with $\beta > 0$ defined in (2.3). Define $b \in H_0^1(\Omega)$ by

$$\langle b, \mu\rangle_{H_0^1(\Omega)} := -\int_{\Omega_1} \overline{y}\mu \qquad (\forall \mu \in H_0^1(\Omega))$$

(i.e. b is the Riesz representant of the integral functional), and also the bounded linear operators $Q_1 : H_0^1(\Omega) \to H_0^1(\Omega)$ and $Q_2 : H^1(\Omega_2) \to H_0^1(\Omega)$ via

$$\langle Q_1 y, \mu\rangle_{H_0^1(\Omega)} := \int_{\Omega_1} y\mu \quad (y, \mu \in H_0^1(\Omega)), \qquad \langle Q_2 u, z\rangle_{H_0^1(\Omega)} := \int_{\Omega_2} uz \quad (u \in H^1(\Omega_2),\ z \in H_0^1(\Omega)).$$

Then system (2.3) can be rewritten as follows:

$$\begin{aligned} &\langle y, z\rangle_{H_0^1(\Omega)} - \langle Q_2 u, z\rangle_{H_0^1(\Omega)} = 0 \qquad (\forall z \in H_0^1(\Omega)), \\ &\langle u, v\rangle_{H^1(\Omega_2)} + \langle \lambda, Q_2 v\rangle_{H_0^1(\Omega)} = 0 \qquad (\forall v \in H^1(\Omega_2)), \\ &\langle \lambda, \mu\rangle_{H_0^1(\Omega)} - \langle Q_1 y, \mu\rangle_{H_0^1(\Omega)} = \langle b, \mu\rangle_{H_0^1(\Omega)} \qquad (\forall \mu \in H_0^1(\Omega)). \end{aligned} \tag{3.12}$$

System (3.12) can be formulated in a more concise way. Let us define the operator

$$Q := \begin{pmatrix} 0 & -Q_2 & 0 \\ 0 & 0 & Q_2^* \\ -Q_1 & 0 & 0 \end{pmatrix} \tag{3.13}$$

and denote

$$\underline{x} := \begin{pmatrix} y \\ u \\ \lambda \end{pmatrix}, \quad \underline{w} := \begin{pmatrix} z \\ v \\ \mu \end{pmatrix} \quad \text{and} \quad \underline{b} := \begin{pmatrix} 0 \\ 0 \\ b \end{pmatrix} \tag{3.14}$$

in $\mathcal{H}$. Then (3.12) is equivalent to

$$\langle \underline{x}, \underline{w} \rangle_{\mathcal{H}} + \langle Q\underline{x}, \underline{w} \rangle_{\mathcal{H}} = \langle \underline{b}, \underline{w} \rangle_{\mathcal{H}} \qquad (\forall \underline{w} \in \mathcal{H})$$

or simply the operator equation

$$(I + Q)\,\underline{x} = \underline{b} \tag{3.15}$$

in $\mathcal{H}$. Using notation

$$L := I + Q,$$

we may just write

$$L\underline{x} = \underline{b}.$$

Since L is a compact perturbation of the identity, the well-posedness of the above equation implies using Fredholm theory that L is invertible, in particular the inf-sup condition holds:

$$\inf_{\substack{\underline{x}\in\mathcal{H} \\ \underline{x}\neq 0}} \sup_{\substack{\underline{w}\in\mathcal{H} \\ \underline{w}\neq 0}} \frac{\langle L\underline{x}, \underline{w} \rangle_{\mathcal{H}}}{\|\underline{x}\|_{\mathcal{H}} \|\underline{w}\|_{\mathcal{H}}} =: m > 0. \tag{3.16}$$

Our estimates will involve compact operators in a real Hilbert space H, see, e.g., [13, Chap. VI], and the following notions:

Definition 3.1 (i) We call $\lambda_j(F)$ $(j = 1, 2, \dots)$ the *ordered eigenvalues* of a compact self-adjoint linear operator F in H if each of them is repeated as many times as its multiplicity and $|\lambda_1(F)| \geq |\lambda_2(F)| \geq ...$

(ii) The *singular values* of a compact operator C in H are

$$s_j(C) := \lambda_j(C^*C)^{1/2} \qquad (j = 1, 2, \dots)$$

where $\lambda_j(C^*C)$ are the ordered eigenvalues of C^*C.

A basic property of compact operators is that $s_j(C) \to 0$ as $j \to \infty$.

Now we verify that the operators in our decomposition of the problem are compact.

Proposition 3.1 *The operators Q_1 and Q_2 in* (3.12) *are compact.*

PROOF. It is well-known that the Riesz representant of the L^2 inner product in a Sobolev space defines a compact operator, see, e.g., [14] (in fact, it is the inverse of the Laplacian or its shifted version). The operators Q_1 and Q_2 define the Riesz representants of L^2 inner products on Ω_1 resp. Ω_2, i.e. the above-mentioned compact operator is only composed with a restriction operator from Ω to Ω_1 or Ω_2 in $L^2(\Omega)$. Since this restriction is obviously bounded, it preserves compactness. ∎

This proposition readily yields the same for the corresponding operator matrix:

Corollary 3.1 *Operator Q in* (3.13) *is compact.*

We will also need the following result for the inf-sup condition:

Proposition 3.2 *[9] Let $L \in B(\mathcal{H})$ be an invertible operator in a Hilbert space $\mathcal{H}$, that is,*

$$m := \inf_{\substack{u\in\mathcal{H}\\ u\neq 0}} \sup_{\substack{v\in\mathcal{H}\\ v\neq 0}} \frac{|\langle Lu, v\rangle|}{\|u\|\|v\|} > 0, \tag{3.17}$$

and let the decomposition $L = I + Q$ hold for some compact operator Q. Let $(V_n)_{n\in\mathbf{N}^+}$ be a sequence of closed subspaces of $\mathcal{H}$ such that the approximation property (3.1) *holds. Then the sequence of real numbers*

$$m_n := \inf_{\substack{u_n\in V_n\\ u_n\neq 0}} \sup_{\substack{v_n\in V_n\\ v_n\neq 0}} \frac{|\langle Lu_n, v_n\rangle|}{\|u_n\|\|v_n\|} \qquad (n \in \mathbf{N}^+)$$

satisfies $\liminf m_n \geq m$.

3.4 The superlinear convergence result

Proposition 3.3 *Let $\mathcal{S}_h$ and $\mathcal{Q}_h$ be defined as in* (3.5)*, and let $s_i(Q)$ $(i = 1, 2, \dots)$ denote the ordered singular values of the operator Q defined in* (3.13)*. Then the following relations hold:*

(*a*) $$s_i(\mathcal{S}_h^{-1}\mathcal{Q}_h) \leq s_i(Q) \qquad (k = 1, \dots, n),$$

(*b*) $$\|\mathcal{A}_h^{-1}\mathcal{S}_h\|_{\mathcal{S}_h} \leq \frac{1}{m_0},$$

for some constant $m_0 > 0$ independent of h.

PROOF. (a) The first estimate is a special case of our result in [9], but such that we now have a better constant in the bound due to the symmetric preconditioner. Namely, by [9, Prop. 5.4], if $\mathcal{N}_h$ is the stiffness matrix of an operator N in $\mathcal{H}$ that satisfies

$$\inf_{\substack{u_h\in V_h\\ u\neq 0}} \sup_{\substack{v_h\in V_h\\ v\neq 0}} \frac{|\langle Nu_h, v_h\rangle_{\mathcal{H}}|}{\|u\|_{\mathcal{H}}\|v\|_{\mathcal{H}}} =: m_1 > 0,$$

then

$$\lambda_i(\mathcal{S}_h^{-1}\mathcal{Q}_h^T\mathcal{N}_h^{-T}\mathcal{S}_h\mathcal{N}_h^{-1}\mathcal{Q}_h) \le \frac{1}{m_1^2}\, s_i(Q_S)^2 \qquad (j=1,2,\ldots,n). \tag{3.18}$$

Now we can set $N = I$ (the identity operator), in which case $m_1 = 1$, further, we have $\mathcal{N}_h = \mathcal{N}_h^T = \mathcal{S}_h$. Hence (3.18) becomes

$$\lambda_i(\mathcal{S}_h^{-1}\mathcal{Q}_h^T\,\mathcal{S}_h^{-1}\mathcal{Q}_h) \le s_i(Q_S)^2 \qquad (j=1,2,\ldots,n).$$

Taking square roots, this is the same as we wanted to prove.

(b) From (3.8) we have

$$\inf_{\substack{\mathbf{c}\in\mathbf{R}^n\\ \mathbf{c}\neq\mathbf{0}}} \frac{\|\mathcal{S}_h^{-1}\mathcal{A}_h\mathbf{c}\|_{\mathcal{S}_h}}{\|\mathbf{c}\|_{\mathcal{S}_h}} = \inf_{\substack{\mathbf{c}\in\mathbf{R}^n\\ \mathbf{c}\neq\mathbf{0}}} \sup_{\substack{\mathbf{d}\in\mathbf{R}^n\\ \mathbf{d}\neq\mathbf{0}}} \frac{\langle\mathcal{S}_h^{-1}\mathcal{A}_h\mathbf{c},\mathbf{d}\rangle_{\mathcal{S}_h}}{\|\mathbf{c}\|_{\mathcal{S}_h}\|\mathbf{d}\|_{\mathcal{S}_h}} = \inf_{\substack{\mathbf{c}\in\mathbf{R}^n\\ \mathbf{c}\neq\mathbf{0}}} \sup_{\substack{\mathbf{d}\in\mathbf{R}^n\\ \mathbf{d}\neq\mathbf{0}}} \frac{\mathcal{A}_h\,\mathbf{c}\cdot\mathbf{d}}{\|\mathbf{c}\|_{\mathcal{S}_h}\|\mathbf{d}\|_{\mathcal{S}_h}} =: m_h > 0$$

from (3.8). Using (3.7),

$$\inf_{\substack{\underline{x}\in V_h\\ \underline{x}\neq 0}} \sup_{\substack{\underline{w}\in V_h\\ \underline{w}\neq 0}} \frac{\langle L\underline{x},\,\underline{w}\rangle_{\mathcal{H}}}{\|\underline{x}\|_{\mathcal{H}}\|\underline{w}\|_{\mathcal{H}}} = \inf_{\substack{\mathbf{c}\in\mathbf{R}^n\\ \mathbf{c}\neq\mathbf{0}}} \sup_{\substack{\mathbf{d}\in\mathbf{R}^n\\ \mathbf{d}\neq\mathbf{0}}} \frac{\mathcal{A}_h\,\mathbf{c}\cdot\mathbf{d}}{\|\mathbf{c}\|_{\mathcal{S}_h}\|\mathbf{d}\|_{\mathcal{S}_h}} = m_h > 0.$$

On the other hand, (3.16) holds on the whole space $\mathcal{H}$:

$$\inf_{\substack{\underline{x}\in\mathcal{H}\\ \underline{x}\neq 0}} \sup_{\substack{\underline{w}\in\mathcal{H}\\ \underline{w}\neq 0}} \frac{\langle L\underline{x},\,\underline{w}\rangle_{\mathcal{H}}}{\|\underline{x}\|_{\mathcal{H}}\|\underline{w}\|_{\mathcal{H}}} =: m > 0.$$

However, Proposition 3.2 yields

$$\liminf m_h \ge m\,(>0)$$

as the dimension n of V_h tends to ∞. This implies that m_h is bounded away from zero, i.e. there exists $m_0 > 0$ such that

$$\inf_{\substack{\mathbf{c}\in\mathbf{R}^n\\ \mathbf{c}\neq\mathbf{0}}} \frac{\|\mathcal{S}_h^{-1}\mathcal{A}_h\mathbf{c}\|_{\mathcal{S}_h}}{\|\mathbf{c}\|_{\mathcal{S}_h}} \ge m_0$$

independently of h. Hence finally

$$\|\mathcal{A}_h^{-1}\mathcal{S}_h\|_{\mathcal{S}_h} = \|(\mathcal{S}_h^{-1}\mathcal{A}_h)^{-1}\|_{\mathcal{S}_h} = \sup_{\substack{\mathbf{c}\in\mathbf{R}^n\\ \mathbf{c}\neq\mathbf{0}}} \frac{\|\mathbf{c}\|_{\mathcal{S}_h}}{\|\mathcal{S}_h^{-1}\mathcal{A}_h\mathbf{c}\|_{\mathcal{S}_h}} \le \frac{1}{m_0}. \qquad \blacksquare$$

In virtue of (3.11) and Proposition 3.3, we have proved

Theorem 3.1 *Under the setting of Proposition 3.3, for any subspace $V_h := Y_h \times U_h \times \Lambda_h \subset \mathcal{H}$, the GMRES iteration for the $n \times n$ preconditioned system (3.9) provides the mesh independent superlinear convergence estimate*

$$\left(\frac{\|r_k\|_{\mathcal{S}_h}}{\|r_0\|_{\mathcal{S}_h}}\right)^{1/k} \le\ \varepsilon_k \qquad (k=1,2,...,n), \tag{3.19}$$

$$\textit{where}\quad \varepsilon_k = \frac{1}{km_0}\ \sum_{i=1}^{k} s_i(Q)\ \to 0 \qquad (\textit{as } k\to\infty) \tag{3.20}$$

and $(\varepsilon_k)_{k\in\mathbf{N}^+}$ is a sequence independent of n and V_h.

4 Some generalizations

4.1 Block preconditioners of PRESB type

Instead of the block diagonal preconditioner used in the previous sections, one can apply a more general block preconditioner of "preconditioned square block matrix" (PRESB) type, extending the method in [7].

For this, one first rewrites system (2.5) by eliminating the variable $\mathbf{u}$. Namely, substituting $\mathbf{u} = -\frac{1}{\beta}(\mathbf{M_u} + \mathbf{K_u})^{-1}\mathbf{M}^T\boldsymbol{\lambda}$, system (2.5) can be reduced to the 2 by 2 system

$$\begin{pmatrix} \mathbf{K} & \frac{1}{\beta}\mathbf{M}(\mathbf{M_u} + \mathbf{K_u})^{-1}\mathbf{M}^T \\ -\mathbf{M}_y & \mathbf{K} \end{pmatrix} \begin{pmatrix} \mathbf{y} \\ \boldsymbol{\lambda} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ -\overline{\mathbf{y}} \end{pmatrix}. \tag{4.1}$$

Here one introduces the scaled vector $\hat{\boldsymbol{\lambda}} := \frac{1}{\sqrt{\beta}}\boldsymbol{\lambda}$ and multiplies the second equation with $-\frac{1}{\sqrt{\beta}}\boldsymbol{\lambda}$ to get

$$\widehat{\mathcal{A}}_h \begin{pmatrix} \mathbf{y} \\ \hat{\boldsymbol{\lambda}} \end{pmatrix} \equiv \begin{pmatrix} \mathbf{K} & \widehat{\mathbf{M}} \\ \widehat{\mathbf{M}}_y & -\mathbf{K} \end{pmatrix} \begin{pmatrix} \mathbf{y} \\ \hat{\boldsymbol{\lambda}} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \widehat{\mathbf{y}} \end{pmatrix},$$

where $\widehat{\mathbf{M}}_y := \frac{1}{\sqrt{\beta}}\mathbf{M}_y$, $\widehat{\mathbf{M}} := \frac{1}{\sqrt{\beta}}\mathbf{M}(\mathbf{M_u} + \mathbf{K_u})^{-1}\mathbf{M}^T$ and $\widehat{\mathbf{y}} := \frac{1}{\sqrt{\beta}}\mathbf{y}$.

We define the preconditioner

$$\widehat{\mathcal{S}}_h := \begin{pmatrix} \mathbf{K} + 2\widehat{\mathbf{M}}_y & \widehat{\mathbf{M}}_y \\ \widehat{\mathbf{M}}_y & -\mathbf{K} \end{pmatrix}.$$

As shown in [3], an explicit form of $\widehat{\mathcal{S}}_h^{-1}$ is

$$\widehat{\mathcal{S}}_h^{-1} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{I} & \mathbf{I} \end{pmatrix} \begin{pmatrix} (\mathbf{K} + \widehat{\mathbf{M}}_\mathbf{y})^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{I} & -\widehat{\mathbf{M}}_\mathbf{y} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & -(\mathbf{K} + \widehat{\mathbf{M}}_\mathbf{y})^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{I} & \mathbf{I} \end{pmatrix}.$$

The action of $\widehat{\mathcal{S}}_h^{-1}$ includes two solutions of linear systems with matrix $\mathbf{K} + \widehat{\mathbf{M}}_\mathbf{y}$, which corresponds to FEM solutions of standard elliptic equtions. Hence these auxiliary systems can be solved with an optimal order of the number of operations, and in case of mesh independent rate of convergence, the overall number of operations is also of optimal order as before. Let us summarize the convergence properties.

Superlinear convergence. We have the decomposition $\widehat{\mathcal{A}}_h = \widehat{\mathcal{S}}_h + \widehat{\mathcal{Q}}_h$, where

$$\widehat{\mathcal{Q}}_h := \begin{pmatrix} -2\widehat{\mathbf{M}}_y & \widehat{\mathbf{M}} - \widehat{\mathbf{M}}_y \\ \mathbf{0} & \mathbf{0} \end{pmatrix}.$$

Here, similarly to the 3 by 3 case (3.5), the remainder matrix $\widehat{\mathcal{Q}}_h$ contains only mass matrices, whereas the preconditioner $\widehat{\mathcal{S}}_h$ includes stiffness matrices in both block diagonal terms, i.e. it corresponds to a Sobolev inner product. Hence one can similarly derive that the preconditioned matrix corresponds to a compact perturbation of the identity, and thus we obtain mesh independent superlinear convergence analogously to (3.19).

Linear convergence. In the above results the estimates depend on the parameter $\beta > 0$. If β is small, then superlinear convergence (although valid) is exhibited with large constant multipliers, i.e. it is not a really useful property. On the other hand, one can see that linear convergence can be bounded uniformly w.r.t. β. For this, we estimate the spectrum of $\widehat{\mathcal{S}}_h^{-1}\widehat{\mathcal{A}}_h$ as follows. Let λ be one of its eigenvalues, i.e. let

$$\widehat{\mathcal{A}}_h \begin{pmatrix} \boldsymbol{\xi} \\ \boldsymbol{\eta} \end{pmatrix} = \lambda\, \widehat{\mathcal{S}}_h \begin{pmatrix} \boldsymbol{\xi} \\ \boldsymbol{\eta} \end{pmatrix}$$

for some vector $(\boldsymbol{\xi}, \boldsymbol{\eta})^T \neq (\mathbf{0}, \mathbf{0})^T$. Since $\widehat{\mathcal{A}}_h = \widehat{\mathcal{S}}_h + \widehat{\mathcal{Q}}_h$, we have $(1-\lambda)\widehat{\mathcal{S}}_h\binom{\boldsymbol{\xi}}{\boldsymbol{\eta}} = -\widehat{\mathcal{Q}}_h\binom{\boldsymbol{\xi}}{\boldsymbol{\eta}}$, i.e.

$$(1-\lambda)\begin{pmatrix} \mathbf{K} + 2\widehat{\mathbf{M}}_y & \widehat{\mathbf{M}}_y \\ \widehat{\mathbf{M}}_y & -\mathbf{K} \end{pmatrix}\begin{pmatrix} \boldsymbol{\xi} \\ \boldsymbol{\eta} \end{pmatrix} = \begin{pmatrix} 2\widehat{\mathbf{M}}_y & \widehat{\mathbf{M}}_y - \widehat{\mathbf{M}} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}\begin{pmatrix} \boldsymbol{\xi} \\ \boldsymbol{\eta} \end{pmatrix}.$$

The second row yields $\widehat{\mathbf{M}}_y\boldsymbol{\xi} = \mathbf{K}\boldsymbol{\eta}$. Substituting this in the first equation, we obtain

$$(1-\lambda)\big(\mathbf{K}\boldsymbol{\xi} + (2\mathbf{K} + \widehat{\mathbf{M}}_y)\boldsymbol{\eta}\big) = (2\mathbf{K} + \widehat{\mathbf{M}}_y)\boldsymbol{\eta} - \widehat{\mathbf{M}}\boldsymbol{\eta}.$$

Taking the inner product with $\boldsymbol{\eta}$, and using that $\mathbf{K}\boldsymbol{\xi}\cdot\boldsymbol{\eta} = \mathbf{K}\boldsymbol{\eta}\cdot\boldsymbol{\xi} = \widehat{\mathbf{M}}_y\boldsymbol{\xi}\cdot\boldsymbol{\xi}$, we obtain

$$(1-\lambda)\big(\widehat{\mathbf{M}}_y\boldsymbol{\xi}\cdot\boldsymbol{\xi} + (2\mathbf{K} + \widehat{\mathbf{M}}_y)\boldsymbol{\eta}\cdot\boldsymbol{\eta}\big) = (2\mathbf{K} + \widehat{\mathbf{M}}_y)\boldsymbol{\eta}\cdot\boldsymbol{\eta} - \widehat{\mathbf{M}}\boldsymbol{\eta}\cdot\boldsymbol{\eta},$$

i.e.

$$\widehat{\mathbf{M}}_y\boldsymbol{\xi}\cdot\boldsymbol{\xi} + \widehat{\mathbf{M}}\boldsymbol{\eta}\cdot\boldsymbol{\eta} = \lambda\big(\widehat{\mathbf{M}}_y\boldsymbol{\xi}\cdot\boldsymbol{\xi} + (2\mathbf{K} + \widehat{\mathbf{M}}_y)\boldsymbol{\eta}\cdot\boldsymbol{\eta}\big)$$

or

$$\lambda = \frac{\widehat{\mathbf{M}}_y\boldsymbol{\xi}\cdot\boldsymbol{\xi} + \widehat{\mathbf{M}}\boldsymbol{\eta}\cdot\boldsymbol{\eta}}{\widehat{\mathbf{M}}_y\boldsymbol{\xi}\cdot\boldsymbol{\xi} + (2\mathbf{K} + \widehat{\mathbf{M}}_y)\boldsymbol{\eta}\cdot\boldsymbol{\eta}}\,.$$

Let

$$R(\boldsymbol{\eta}) := \frac{\widehat{\mathbf{M}}\boldsymbol{\eta}\cdot\boldsymbol{\eta}}{(2\mathbf{K} + \widehat{\mathbf{M}}_y)\boldsymbol{\eta}\cdot\boldsymbol{\eta}}, \qquad \theta_{min} := \min_{\boldsymbol{\eta}\neq\mathbf{0}} R(\boldsymbol{\eta}), \qquad \theta_{max} := \max_{\boldsymbol{\eta}\neq\mathbf{0}} R(\boldsymbol{\eta}), \tag{4.2}$$

then we readily obtain

Proposition 4.1 *The eigenvalues of $\widehat{\mathcal{S}}_h^{-1}\widehat{\mathcal{A}}_h$ are real and satisfy*

$$\min\{1, \theta_{min}\} \leq \lambda\big(\widehat{\mathcal{S}}_h^{-1}\widehat{\mathcal{A}}_h\big) \leq \max\{1, \theta_{max}\}$$

with θ_{min} and θ_{max} from (4.2).

In order to observe the uniform behaviour of θ_{min} and θ_{max} as $\beta \to 0$, note that the definition of $\widehat{\mathbf{M}}_y$ and $\widehat{\mathbf{M}}$ implies

$$R(\boldsymbol{\eta}) := \frac{\mathbf{M}(\mathbf{M_u} + \mathbf{K_u})^{-1}\mathbf{M}^T\boldsymbol{\eta}\cdot\boldsymbol{\eta}}{(2\sqrt{\beta}\,\mathbf{K} + \mathbf{M}_y)\boldsymbol{\eta}\cdot\boldsymbol{\eta}} \approx \frac{\mathbf{M}(\mathbf{M_u} + \mathbf{K_u})^{-1}\mathbf{M}^T\boldsymbol{\eta}\cdot\boldsymbol{\eta}}{\mathbf{M}_y\boldsymbol{\eta}\cdot\boldsymbol{\eta}} \qquad \text{as } \beta \to 0.$$

More precisely, we can estimate as follows. We have $(2\sqrt{\beta}\,\mathbf{K} + \mathbf{M}_y)\boldsymbol{\eta}\cdot\boldsymbol{\eta} \geq \mathbf{M}_y\boldsymbol{\eta}\cdot\boldsymbol{\eta}$ in the denominator, hence $R(\boldsymbol{\eta})$ is bounded above uniformly in β. On the other hand, the previously seen equality $\widehat{\mathbf{M}}_y\boldsymbol{\xi} = \mathbf{K}\boldsymbol{\eta}$ implies that $\mathbf{K}\boldsymbol{\eta}$ has zero coordinates where $\widehat{\mathbf{M}}_y\boldsymbol{\xi}$ has, i.e. in the nodes outside Ω_1, hence $\mathbf{K}\boldsymbol{\eta}\cdot\boldsymbol{\eta} = \int_{\Omega_1} |\nabla z_h|^2$ and $\mathbf{M}_y\boldsymbol{\eta}\cdot\boldsymbol{\eta} = \int_{\Omega_1} z_h^2$ (where $z_h \in Y_h$ has coordinate vector $\boldsymbol{\eta}$). Thus the standard condition number estimates yield $\mathbf{K}\boldsymbol{\eta}\cdot\boldsymbol{\eta} \leq O(h^{-2})(\mathbf{M}_y\boldsymbol{\eta}\cdot\boldsymbol{\eta})$. If we choose $\beta = O(h^4)$, then the denominator satisfies $(2\sqrt{\beta}\,\mathbf{K} + \mathbf{M}_y)\boldsymbol{\eta}\cdot\boldsymbol{\eta} = O(h^2)(\mathbf{K}\boldsymbol{\eta}\cdot\boldsymbol{\eta}) + \mathbf{M}_y\boldsymbol{\eta}\cdot\boldsymbol{\eta} \leq const.\ \mathbf{M}_y\boldsymbol{\eta}\cdot\boldsymbol{\eta}$, hence $R(\boldsymbol{\eta})$ is bounded below uniformly in β. Hence, altogether, θ_{min}, θ_{max} and ultimately the spectrum of $\widehat{\mathcal{S}}_h^{-1}\widehat{\mathcal{A}}_h$ are bounded uniformly w.r.t β.

4.2 Boundary control problems

A modification of the distributed control problem (2.1)-(4.3), also studied in [10], is the boundary control problem, in which the same functional (2.1) is minimized subject to the PDE constraint

$$\begin{cases} -\Delta y = f & \text{in } \Omega \\ \frac{\partial y}{\partial n}\big|_{\,\partial\Omega} = u \end{cases} \tag{4.3}$$

where f represents a fixed forcing term and the control function u is applied on the boundary. The FEM solution of this problem leads to a system very similar to (2.5). The mass matrix $\mathbf{M}$ is replaced by an (also rectangular) matrix $\mathbf{N}$ that connects interior and boundary basis functions, further, the mass and stiffness matrices for u act on the boundary, and are denoted by $\mathbf{M}_{\mathbf{u},b}$ and $\mathbf{K}_{\mathbf{u},b}$, respectively. Thus the global system matrix takes the form

$$\mathcal{A}_h := \begin{pmatrix} \mathbf{K} & -\mathbf{N} & \mathbf{0} \\ \mathbf{0} & \beta(\mathbf{M}_{\mathbf{u},b} + \mathbf{K}_{\mathbf{u},b}) & \mathbf{N}^T \\ -\mathbf{M}_y & \mathbf{0} & \mathbf{K} \end{pmatrix}. \tag{4.4}$$

Then our previous results hold for this problem as well with slight changes. In particular, the matrix $\mathbf{N}$ corresponds to the embedding of the boundary space $L^2(\partial\Omega)$ into $H^1(\Omega)$. Hence, in a similar way, we obtain that the preconditioned matrix corresponds to a compact perturbation of the identity. Thus we can again derive mesh independent superlinear convergence of the preconditioned GMRES.

4.3 Box constraints

The functions y and/or u are often assumed to satisfy additional pointwise constraints (box constraints). For instance, for the state variable y, one prescribes

$$y_a \leq y \leq y_b$$

for some given constants y_a and y_b. The corresponding constraint for u is $u_a \leq u \leq u_b$. Box constraints can be dealt with efficiently using a penalty term of so-called Moreau-Yosida type, see [8, 10, 19]. For the distributed control studied in this paper, the objective function (2.1) is modified as

$$J_{MY}(y,u) := J(y,u) + \frac{1}{2\varepsilon}\|\max\{0, y - y_b\}\|^2 + \frac{1}{2\varepsilon}\|\max\{0, y - y_a\}\|^2$$

for the state constrained case (where $\varepsilon > 0$ is a small penalty parameter) and similarly for control constraints. Applying a semi-smooth Newton scheme, one obtains linear systems with small modifications of the system (2.4). After rearrangement as in (2.5), the global system matrix becomes

$$\begin{pmatrix} \mathbf{K} & -\mathbf{M} & \mathbf{0} \\ \mathbf{0} & \beta(\mathbf{M_u} + \mathbf{K_u}) & \mathbf{M}^T \\ -(\mathbf{M}_y + \frac{1}{\varepsilon} G_A \mathbf{M}_y G_A) & \mathbf{0} & \mathbf{K} \end{pmatrix}, \tag{4.5}$$

where G_A is a diagonal matrix with values 0 or 1, depending whether the actual value of $\mathbf{y}$ in that coordinate satisfies or not the box constraint. The new factors G_A at the mass matrix M_y do not change the fact that the term $G_A \mathbf{M}_y G_A$ corresponds to a compact perturbation of the identity, as well as the whole block matrix as before. Hence we obtain mesh independent superlinear convergence again.

We note, however, that the superlinear rate is exhibited with large constant multipliers when ε is small. Hence it is worth mentioning that the linear convergence rate is not sensitive to ε. Namely, as shown in [8], for this problem the eigenvalues cluster in two or three intervals: one near the upper bound 1, one in the middle and one near 0. The middle interval is $[\frac{1+\varepsilon}{2+\varepsilon}, 1)$, the upper bound takes values arbitrarily close to unity when $\varepsilon \to 0$. If $\beta = O(\frac{h^2}{\varepsilon})$ then the lower eigenvalues are bounded below by $\frac{\varepsilon}{1+\varepsilon}$ if $\varepsilon < 1$. For very small values of ε, the behaviour is similar to the case when there are several zero eigenvalues [1], i.e. the small eigenvalues have a negligible effect on the solution when a Krylov subspace iteration is used.

4.4 Time-harmonic parabolic optimal control problems

In some problems the control and discrete state functions are time-harmonic, see [7] including an example when the target solution and the control funcion are time-harmonic for a parabolic PDE constraint. This reduces the problem to minimizing $J(y,u) := \frac{1}{2}\|y - \overline{y}\|^2_{L^2(\Omega)} + \frac{\beta}{2}\|u\|^2_{L^2(\Omega)}$ subject to the elliptic PDE constraint

$$\begin{cases} -\Delta y + i\omega y &= u \\ y\big|_{\partial\Omega} &= 0 \end{cases}$$

where y and $\overline{y}$ are real-valued but the control u must be complex-valued. After rearrangement, the global system matrix becomes

$$\mathcal{A}_h := \begin{pmatrix} \mathbf{K} + i\omega\mathbf{M} & -\mathbf{M} \\ \mathbf{M} & \beta(\mathbf{K} + i\omega\mathbf{M}) \end{pmatrix}$$

Introducing the block diagonal preconditioner and the corresponding remainder matrix

$$\mathcal{S}_h := \begin{pmatrix} \mathbf{K} & \mathbf{0} \\ \mathbf{0} & \beta\mathbf{K} \end{pmatrix} \qquad \text{and} \qquad \mathcal{Q}_h := \begin{pmatrix} i\omega\mathbf{M} & -\mathbf{M} \\ \mathbf{M} & i\beta\omega\mathbf{M} \end{pmatrix},$$

respectively, we see that $\mathcal{Q}_h$ contains only mass matrices, whereas the preconditioner $\widehat{\mathcal{S}}_h$ includes stiffness matrices in both block diagonal terms. Then our previous results can

be used with a direct adaptation to the complex case (just replacing the transposed $\mathcal{Q}_h^T$ with the complex adjoint $\mathcal{Q}_h^*$), and we obtain mesh independent superlinear convergence again.

Acknowledgement: The research of O. Axelsson was supported by the Ministry of Education, Youth and Sports from the National Programme of Sustainability (NPU II) project "IT4 Innovations excellence in science LQ1602". The research of J. Karátson was supported by the Hungarian Scientific Research Fund OTKA, No. 112157.

References

[1] Axelsson, O., *Iterative Solution Methods,* Cambridge University Press, 1994.

[2] Axelsson, O., Farouq S., Neytcheva M., Comparison of preconditioned Krylov subspace iteration methods for PDE-constrained optimization problems: Stokes control, *Numerical Algorithms* 73 (3), 631-663.

[3] Axelsson, O., Farouq S., Neytcheva M., A preconditioner for optimal control problems, constrained by Stokes equation with a time-harmonic control, *J. Comput. Appl. Math.* 310 (2017) 5-18.

[4] Axelsson, O., Karátson J., Superlinearly convergent CG methods via equivalent preconditioning for nonsymmetric elliptic operators, *Numer. Math.* 99 (2004), No. 2, 197-223.

[5] Axelsson, O., Karátson J., Mesh independent superlinear PCG rates via compact-equivalent operators, *SIAM J. Numer. Anal.*, 45 (2007), No.4, pp. 1495-1516.

[6] Axelsson, O., Karátson J., Equivalent operator preconditioning for linear elliptic problems, *Numer. Algorithms*, 50 (2009), Issue 3, p. 297-380.

[7] Axelsson, O., Neytcheva M., A comparison of preconditioners for two-by two block matrices with square matrix blocks arising in in optimal control PDE problems, submitted

[8] Axelsson, O., Neytcheva M., Ström, A., An efficient preconditioning method for state box-constrained optimal control problems, submitted

[9] Axelsson, O., Karátson J., Magoulès F., Superlinear convergence under complex shifted Laplace preconditioners for Helmholtz equations, submitted

[10] Barker A.T., Rees T., Stoll M., A Fast Solver for an H1 Regularized PDE-Constrained Optimization Problems, *Comm. Comput. Physics*, 19 (2016), 143-167.

[11] Elman, H.C., Schultz. M.H., Preconditioning by fast direct methods for nonself-adjoint nonseparable elliptic equations, *SIAM J. Numer. Anal.,* 23 (1986), 44–57.

[12] Faber, V., Manteuffel, T., Parter, S.V., On the theory of equivalent operators and application to the numerical solution of uniformly elliptic partial differential equations, *Adv. in Appl. Math.,* 11 (1990), 109-163.

[13] Gohberg, I., Goldberg, S., Kaashoek, M. A., *Classes of Linear Operators*, Vol. I., Operator Theory: Advances and Applications, 49, Birkhäuser Verlag, Basel, 1990.

[14] Goldstein, C. I., Manteuffel, T. A., Parter, S. V., Preconditioning and boundary conditions without H_2 estimates: L_2 condition numbers and the distribution of the singular values, *SIAM J. Numer. Anal.* 30 (1993), no. 2, 343–376.

[15] Hackbusch, W., *Multigrid methods and applications*, Springer Series in Computational Mathematics 4, Springer, Berlin, 1985.

[16] J.-L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer-Verlag, Berlin, 1971.

[17] Moret, I., A note on the superlinear convergence of GMRES, *SIAM J. Numer. Anal.* 34 (1997), 513–516.

[18] Paige, C.; Parlett, B.; van der Vorst, H., Approximate Solutions and Eigenvalue Bounds from Krylov Subspaces, *Numer. Lin. Alg. Appl.* 29, 115-134, 1995.

[19] Pearson, J. W.; Stoll, M.; Wathen, A. J., Preconditioners for state-constrained optimal control problems with Moreau-Yosida penalty function, *Numer. Linear Algebra Appl.* 21 (2014), no. 1, 81-97.

[20] Rossi, T., Toivanen, J., A parallel fast direct solver for block tridiagonal systems with separable matrices of arbitrary dimension, *SIAM J. Sci. Comput.* 20 (1999), no. 5, 1778–1796 (electronic).

[21] Saad, Y., *Iterative Methods for Sparse Linear Systems*, Second Edition, SIAM, 2003.

[22] F. Tröltsch, *Optimal Control of Partial Differential Equations*, Graduate studies in Mathematics, AMS Providence, 2005

[23] Winter, R., Some superlinear convergence results for the conjugate gradient method, *SIAM J. Numer. Anal.*, 17 (1980), 14-17.

[24] Widlund, O., A Lanczos method for a class of non-symmetric systems of linear equations, *SIAM J. Numer. Anal.,* 15 (1978), 801-812.