Mesh independent superlinear PCG rates via compact-equivalent operators

by O. Axelsson[1] and J. Karátson[2]

Abstract

The subject of the paper is the mesh independent convergence of the preconditioned conjugate gradient method for nonsymmetric elliptic problems. The approach of equivalent operators is involved, in which one uses the discretization of another suitable elliptic operator as preconditioning matrix. By introducing the notion of compact-equivalent operators, it is proved that for a wide class of elliptic problems the superlinear convergence of the obtained PCGM is mesh independent under FEM discretizations, that is, the rate of superlinear convergence is given in the form of a sequence which is mesh independent and is determined only by the elliptic operators.

Key words: nonsymmetric elliptic problems, conjugate gradient method, preconditioning, equivalent operators, superlinear convergence, mesh independence.

AMS Classification: 65J10, 65F10, 65N15.

1   Introduction

The conjugate gradient method is a widespread way of solving large linear algebraic systems, such as those arising from discretized elliptic problems, in particular when combined with a suitable preconditioning. Since its first presentation in [19] the convergence of the CGM has been well established, as summarized in [2]. The convergence theory of the CGM often involves linear operators in Hilbert space, see various both classical and recent results [12, 13, 18, 24, 27, 28] and the authors' papers [4, 5, 6]. A basic reason to use Hilbert space theory is to derive mesh independence of the convergence estimates, by which it can be shown that the preconditioned CGM can be competitive with multigrid methods [12].

The theory of equivalent operators in Hilbert space has proved an efficient clear framework for the convergence study of the preconditioned CGM for elliptic problems [12, 16, 23]. Thereby one uses the discretization of a suitable linear elliptic operator as preconditioning matrix, see also [8, 10, 28]. As a main result, mesh independence of linear convergence rates is rigorously characterized in [12, 23]. We note that in [16], for proper boundary conditions, when the preconditioned operator is a compact perturbation of the identity then convergence is expected to be faster than any linear rate.

Our goal is to complete the above results by showing that for a class of elliptic problems, the superlinear convergence of the PCGM is mesh independent under FEM discretizations. This means that a bound on the rate of superlinear convergence is given in the form of a sequence which is mesh independent and is determined only by the elliptic operators. To describe the suitable class of problems, we introduce the notion of compact-equivalent operators which expresses that preconditioning one of them with the other yields a compact perturbation of the identity. This notion and the convergence result give a refinement of the case of equivalent operators: roughly speaking, if the two operators (the original and preconditioner) are equivalent

[1]Department of Information Technology, Uppsala University, Sweden & Institute of Geonics AS CR, Ostrava, Czech Republic; owea@it.uu.se
[2]Department of Applied Analysis, ELTE University, H-1117 Budapest, Hungary; karatson@cs.elte.hu

then the corresponding PCGM provides mesh independent linear convergence, whereas if the two operators are compact-equivalent then the PCGM provides mesh independent superlinear convergence.

Our present results are extensions of the earlier ones [6, 22], where such mesh independence was proved for the GCG-LS method for elliptic Dirichlet problems, however, with severe restrictions: except some special cases, both the original and preconditioning operators had to contain constant coefficients. Now we show that two elliptic operators, with homogeneous Dirichlet conditions on the same portion of the boundary, are compact-equivalent if and only if their principal parts coincide up to a constant factor. Within this class, the proof of the mesh independence result then contains no restrictions except standard smoothness and coercivity assumptions on the operators.

The paper is organized as follows: the required background is given in Section 2, compact-equivalent operators are introduced and characterized in Section 3, and the mesh independence result is proved in Section 4.

2 Background

2.1 Conjugate gradient algorithms

Let us consider a linear system

$$Bu = f \tag{1}$$

with a given nonsingular matrix $B \in \mathbf{R}^{n \times n}$, $f \in \mathbf{R}^n$ and solution u. Let $\langle .,. \rangle$ be a given inner product on $\mathbf{R}^n$ and, denoting by B^* the adjoint of B w.r.t. this inner product, assume that $B + B^* > 0$, i.e., is positive definite.

If B is self-adjoint, then the standard CGM reads as follows [2, 27]: let $u_0 \in \mathbf{R}^n$ be arbitrary, $d_0 := Bu_0 - f$; for given u_k and d_k, with $\hat{r}_k := Bu_k - f$, we let

$$u_{k+1} = u_k - \alpha_k d_k, \text{ where } \alpha_k = \frac{\langle \hat{r}_k, d_k \rangle}{\langle Bd_k, d_k \rangle}; \quad d_{k+1} = \hat{r}_{k+1} + \beta_k d_k, \text{ where } \beta_k = \frac{\|\hat{r}_{k+1}\|^2}{\|\hat{r}_k\|^2}. \tag{2}$$

Then, using the error vector $e_k = u_k - u$ and its energy norm $\|e_k\|_B = \langle Be_k, e_k \rangle^{1/2}$, respectively, and with the decomposition $B = I + C$ (where I is the identity matrix), the following celebrated estimate holds [2, 27]:

$$\left(\frac{\|e_k\|_B}{\|e_0\|_B} \right)^{1/k} \leq \frac{2}{k} \|B^{-1}\| \sum_{j=1}^{k} |\lambda_j(C)| \qquad (k = 1, 2, ..., n), \tag{3}$$

which shows superlinear convergence if the eigenvalues $|\lambda_1(C)| \geq |\lambda_2(C)| \geq ...$ approach zero.

Since this result is basic for the whole paper, and for completeness, we present a derivation of (3) following [2]. The optimality of the CGM implies

$$\frac{\|e_k\|_B}{\|e_0\|_B} \leq \min_{P_k \in \pi_k^1} \max_{\lambda \in \sigma(B)\}} |P_k(\lambda)|$$

where π_k^1 denotes the set of polynomials P_k of degree k with $P_k(0) = 1$. Let $\lambda_j := \lambda_j(B)$ and $\mu_j := \lambda_j(C) (= \lambda_j - 1)$. Then the polynomials $P_k(\lambda) := \prod_{j=1}^{k} \left(1 - \frac{\lambda}{\lambda_j}\right)$ satisfy $P_k(\lambda_i) = 0$

2

$(i = 1, .., k)$ and

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

hence, using the arithmetic-geometric inequality,

$$\max_{\lambda \in \sigma(B)\}} |P_k(\lambda)|^{1/k} \leq \frac{2}{k} \sum_{j=1}^{k} \frac{|\mu_j|}{|1 + \mu_j|} \leq \frac{2}{k} \left(\sup \frac{1}{|\lambda_j|} \right) \sum_{j=1}^{k} |\mu_j|$$

which yields (3).

For nonsymmetric B, several CG algorithms exist (see e.g. [2, 11]). The GCG-LS method [1, 2] is defined directly for the equation (1) and produces a similar estimate as (3) if B is normal. Mesh independent bounds in [6, 22] for (3) for some elliptic problems have been given using the GCG-LS method. Alternatively, one can consider the normal equation and apply a symmetric CG algorithm, which we will do in this paper. For clearness, let us hereby consider a nonsymmetric linear system

$$Au = b \tag{4}$$

with given $A \in \mathbf{R}^{n \times n}$, $b \in \mathbf{R}^n$, such that

$$A + A^* > 0. \tag{5}$$

Let us apply the iteration (2) for equation $A^*Au = A^*b$, i.e. with $B = A^*A$ and $f = A^*b$. Then, with notations $s_k = \hat{r}_k$ and $r_k = A^{-T}\hat{r}_k$, we obtain the following algorithmic form: let $u_0 \in \mathbf{R}^n$ be arbitrary, $r_0 := Au_0 - b$, $s_0 := d_0 := A^*r_0$; for given d_k, u_k, r_k and s_k, we let

$$\begin{cases} z_k = Ad_k, \\[2mm] \alpha_k = \dfrac{\langle r_k, z_k \rangle}{\|z_k\|^2} \,, \quad u_{k+1} = u_k - \alpha_k d_k \,, \quad r_{k+1} = r_k - \alpha_k z_k \,; \\[2mm] s_{k+1} = A^*r_{k+1}, \\[2mm] \beta_k = \dfrac{\|s_{k+1}\|^2}{\|s_k\|^2} \,, \quad d_{k+1} = s_{k+1} + \beta_k d_k. \end{cases} \tag{6}$$

Let us consider the decomposition

$$A = I + K.$$

Then, using the relations $B = I + (K^* + K + K^*K)$, $\|e_k\|_B = \|Ae_k\| = \|r_k\|$ and $\|B^{-1}\| \leq \nu^{-1}$ where $\nu := \min_{x \in \mathbf{R}^n} \frac{\|Ax\|^2}{\|x\|^2}$, estimate (3) can be reformulated as

$$\left(\frac{\|r_k\|}{\|r_0\|} \right)^{1/k} \leq \frac{2}{k\nu} \sum_{i=1}^{k} \Big(|\lambda_i(K^* + K)| + \lambda_i(K^*K) \Big) \qquad (k = 1, 2, ..., n). \tag{7}$$

The goal of this paper is to derive a mesh independent bound for (7) when (4) comes from a preconditioned discretized elliptic PDE using suitable equivalent operators.

2.2 Singular values of compact operators

Let H be a real Hilbert space. We shall consider compact operators, i.e., operators C such that the image (Cv_i) of any bounded sequence (v_i) contains a convergent subsequence.

Definition 2.1 (i) We call $\lambda_i(F)$ $(i = 1, 2, \ldots)$ 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 $|\lambda_1(F)| \geq |\lambda_2(F)| \geq \ldots$

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

$$s_i(C) := \lambda_i(C^*C)^{1/2}, \qquad (i = 1, 2, \ldots)$$

where $\lambda_i(C^*C)$ are the ordered eigenvalues of C^*C. In particular, if C is self-adjoint then $s_i(C) = |\lambda_i(C)|$.

Some useful properties of compact operators are listed below:

Proposition 2.1 *Let C be a compact operator in H. Then*

(a) for any $k \in \mathbf{N}^+$ and any orthonormal vectors $u_1, \ldots, u_k \in H$,

$$\sum_{i=1}^{k} |\langle Cu_i, u_i \rangle| \leq \sum_{i=1}^{k} s_i(C).$$

(b) If B is bounded linear operator in H, then

$$s_i(BC) \leq \|B\| \, s_i(C) \qquad (i = 1, 2, \ldots).$$

(c) (Variational characterization of the eigenvalues). If C is also self-adjoint, then

$$|\lambda_i(C)| = \min_{H_{i-1} \subset H} \max_{\substack{u \perp H_{i-1} \\ u \neq 0}} \frac{|\langle Cu, u \rangle|}{\|u\|^2},$$

where H_{i-1} stands for an arbitrary $(i-1)$-dimensional subspace.

(d) If a sequence $(u_i) \subset H$ satisfies $\langle u_i, u_j \rangle = \langle Cu_i, u_j \rangle = 0$ $(i \neq j)$, then

$$\inf_i |\langle Cu_i, u_i \rangle| / \|u_i\|^2 = 0.$$

PROOF. Statements (a) and (b) are the consequences of [14, Chap. VI, Corollary 3.3 and Proposition 1.3, resp.], for statement (c) see [15, Theorem III.9.1]. To prove (d), assume the contrary that the infimum equals $\delta > 0$. We may assume that $\langle Cu_i, u_i \rangle$ has constant sign (otherwise we consider such a subsequence only). Then the orthonormal sequence $v_i := u_i / \|u_i\|$ satisfies for all $i \neq j$

$$2\delta \leq |\langle Cv_i, v_i \rangle + \langle Cv_j, v_j \rangle| = |\langle C(v_i - v_j), v_i - v_j \rangle| \leq \|C(v_i - v_j)\| \, \|v_i - v_j\| = \sqrt{2}\|C(v_i - v_j)\|,$$

hence the image (Cv_i) of the bounded sequence (v_i) contains no convergent subsequence, i.e. C is not compact. ∎

3 Compact-equivalent operators in Hilbert space

In this section we introduce and characterize compact-equivalent operators. Roughly speaking, the compact-equivalence of the unbounded operators N and L expresses that $N^{-1}L$ is a compact perturbation of constant times the identity. To avoid difficulties with domains and ranges, our definition will use a weak form of the operators in a suitable energy space H_S. In particular, no regularity is required in the case of elliptic operators.

The fact that a compact perturbation of constant times identity is a bounded operator implies that compact-equivalent operators are equivalent in the sense of [12]. Hence, when we characterize compact-equivalent elliptic operators (under standard smoothness and coercivity assumptions), we can a priori assume that they have homogeneous Dirichlet conditions on the same portion of the boundary [23]. Within this class, compact-equivalence will hold if and only if the principal parts of the operators coincide up to some constant.

3.1 Basic definitions

In what follows, let H be a real Hilbert space. Let S be a (generally unbounded) linear symmetric operator in H which is coercive, i.e., there exists $p > 0$ such that $\langle Su, u \rangle \geq p\|u\|^2$ $(u \in D(S))$. Then the energy space H_S is the completion of $D(S)$ under the inner product $\langle u, v \rangle_S = \langle Su, v \rangle$, and the coercivity implies $H_S \subset H$. The corresponding S-norm is denoted by $\|u\|_S$, and the space of bounded linear operators on H_S by $B(H_S)$.

Definition 3.1 Let S be a linear symmetric coercive operator in H. We say that a linear operator L in H is *S-bounded and S-coercive*, and write $L \in BC_S(H)$, if the following properties hold:

(i) $D(L) \subset H_S$ and $D(L)$ is dense in H_S in the S-norm;

(ii) there exists $M > 0$ such that

$$|\langle Lu, v \rangle| \leq M\|u\|_S\|v\|_S \qquad (u, v \in D(L));$$

(iii) there exists $m > 0$ such that

$$\langle Lu, u \rangle \geq m\|u\|_S^2 \qquad (u \in D(L)).$$

Definition 3.2 For any $L \in BC_S(H)$, let $L_S \in B(H_S)$ be defined by

$$\langle L_S u, v \rangle_S = \langle Lu, v \rangle \qquad (u, v \in D(L)).$$

Remark 3.1 (a) The above definition makes sense since L_S is the bounded linear operator on H_S that represents the unique extension to H_S of the densely defined S-bounded bilinear form $u, v \mapsto \langle Lu, v \rangle$.

(b) L_S is coercive on H_S.

(c) If in particular $R(L) \subset R(S)$ (where $R(.)$ denotes the range), then $L_S|_{D(L)} = S^{-1}L$.

Remark 3.2 Definition 3.2 uses the idea of weak form of operators from [23]. Namely, if H_S is a subspace of $H^1(\Omega)$ consisting of functions vanishing on a fixed portion of the boundary, then L_S coincides with the weak operator L_w using (2.15) in [23].

5

Now let us consider an operator equation

$$Lu = g \qquad (8)$$

where $L \in BC_S(H)$ and $g \in H$.

Definition 3.3 We call $u \in H_S$ the *weak solution* of equation (8) if

$$\langle L_S u, v \rangle_S = \langle g, v \rangle \qquad (v \in H_S). \qquad (9)$$

Remark 3.3 (a) For all $g \in H$ the weak solution of (8) exists and is unique. This follows in the usual way from the Lax-Milgram theorem, since $v \mapsto \langle g, v \rangle$ is a bounded linear functional on H_S by the coercivity of S.

(b) If $u \in D(L)$, then u satisfies (8) (i.e. it is a strong solution) if and only if it is a weak solution.

3.2 Compact-equivalent operators

Definition 3.4 Let L and N be S-bounded and S-coercive operators in H. We call L and N *compact-equivalent in H_S* if

$$L_S = \mu N_S + Q_S \qquad (10)$$

for some constant $\mu > 0$ and compact operator $Q_S \in B(H_S)$.

Remark 3.4 (i) It follows in a straightforward way that the property compact-equivalence is an equivalence relation.

(ii) In the special case $R(L) \subset R(N)$, compact-equivalence of L and N means that $N^{-1}L$ is a compact perturbation of constant times the identity in the space H_S. Indeed, it is easy to see that here $N^{-1}L = N_S^{-1}L_S|_{D(L)}$, and by definition the latter is the perturbation of μI with the operator $N_S^{-1}Q_S|_{D(L)}$, which is compact since N_S^{-1} is bounded. (In the general case the 'weakly preconditioned' form $N_S^{-1}L_S$ is also a compact perturbation.)

Now we characterize compact-equivalence for elliptic operators. Let $H = L^2(\Omega)$ and let us define the operators

$$N_1 u \equiv -\mathrm{div}\,(A_1 \nabla u) + \mathbf{b}_1 \cdot \nabla u + c_1 u \qquad \text{for } u_{|\Gamma_D} = 0,\ \tfrac{\partial u}{\partial \nu_{A_1}} + \alpha_1 u_{|\Gamma_N} = 0,$$
$$N_2 u \equiv -\mathrm{div}\,(A_2 \nabla u) + \mathbf{b}_2 \cdot \nabla u + c_2 u \qquad \text{for } u_{|\Gamma_D} = 0,\ \tfrac{\partial u}{\partial \nu_{A_2}} + \alpha_2 u_{|\Gamma_N} = 0$$

where $\frac{\partial u}{\partial \nu_{A_i}} = A_i \nu \cdot \nabla u$ denotes the weighted normal derivative. (The formal domain of N_i to be used in Definition 3.2 consists of those $u \in H^2(\Omega)$ that satisfy the above boundary conditions, however, this is nowhere used elsewhere.) The following properties hold, where $i = 1, 2$:

Assumptions 3.2

(i) $\Omega \subset \mathbf{R}^d$ is a bounded piecewise C^1 domain; Γ_D, Γ_N are disjoint open measurable subparts of $\partial\Omega$ such that $\partial\Omega = \overline{\Gamma}_D \cup \overline{\Gamma}_N$;

(ii) $A_i \in C^1(\overline{\Omega}, \mathbf{R}^{d \times d})$ and for all $x \in \overline{\Omega}$ the matrix $A_i(x)$ is symmetric; $\mathbf{b}_i \in C^1(\overline{\Omega})^d$, $c_i \in L^\infty(\Omega)$, $\alpha_i \in L^\infty(\Gamma_N)$;

(iii) we have the coercivity properties $\min\limits_{\lambda\in\sigma(A_i(x))}\lambda \geq p > 0$ with p independent of x, $\hat{c}_i :=$ $c_i - \frac{1}{2}\operatorname{div}\mathbf{b}_i \geq 0$ in Ω and $\hat{\alpha}_i := \alpha_i + \frac{1}{2}(\mathbf{b}_i\cdot\nu) \geq 0$ on Γ_N;

(iv) either $\Gamma_D \neq \emptyset$, or $\hat{c}_i$ or $\hat{\alpha}_i$ has a positive lower bound.

For the study of such operators we define the space

$$H_D^1(\Omega) := \{u \in H^1(\Omega) : u_{|\Gamma_D} = 0\} \quad \text{with} \quad \langle u, v\rangle_S := \int_\Omega (G\,\nabla u\cdot\nabla v + huv) + \int_{\Gamma_N}\beta uv\,d\sigma \quad (11)$$

where G has the same properties as A_i above in (ii)-(iii), and $h \in L^\infty(\Omega)$, $h \geq 0$ if $\Gamma_D \neq \emptyset$ and $h \geq \delta_0 > 0$ if $\Gamma_D = \emptyset$, further, $\beta \in L^\infty(\Gamma_N)$ and $\beta \geq 0$. Then $H_D^1(\Omega)$ is the energy space H_S of the operator $Su := -\operatorname{div}(G\,\nabla u) + hu$ on $D(S) := \{u \in H^2(\Omega) : u_{|\Gamma_D} = 0, \frac{\partial u}{\partial\nu_G}_{|\Gamma_N} = 0\}$. It is easy to check the properties in Definition 3.1 from the above assumptions, which means that $N_1, N_2 \in C_S(L^2(\Omega))$.

Proposition 3.1 *The elliptic operators N_1 and N_2 are compact-equivalent in $H_D^1(\Omega)$ if and only if their principal parts coincide up to some constant $\mu > 0$, i.e. $A_1 = \mu A_2$.*

PROOF.   We have for all $u, v \in H_D^1(\Omega)$

$$\langle (N_i)_S u, v\rangle_S = \int_\Omega\Big(A_i\,\nabla u\cdot\nabla v + (\mathbf{b}_i\cdot\nabla u)v + c_i uv\Big)\,dx + \int_{\Gamma_N}\alpha_i uv\,d\sigma\,.$$

Hence

$$(N_1)_S - \mu(N_2)_S = J_S + Q_S$$

where, using notations $\mathbf{b} := \mathbf{b}_1 - \mu\mathbf{b}_2$, $c := c_1 - \mu c_2$ and $\alpha := \alpha_1 - \mu\alpha_2$, we have

$$\langle J_S u, v\rangle_S = \int_\Omega (A_1 - \mu A_2)\,\nabla u\cdot\nabla v \quad \text{and} \quad \langle Q_S u, v\rangle_S = \int_\Omega\Big((\mathbf{b}\cdot\nabla u)v + cuv\Big)\,dx + \int_{\Gamma_N}\alpha uv\,d\sigma\,. \tag{12}$$

Here Q_S is compact, which is known [16] when N_1 and N_2 have the same boundary conditions. Otherwise we use the equality

$$\int_\Omega(\mathbf{b}\cdot\nabla u)v\,dx = -\int_\Omega u(\mathbf{b}\cdot\nabla v)\,dx - \int_\Omega(\operatorname{div}\mathbf{b})uv\,dx + \int_{\Gamma_N}(\mathbf{b}\cdot\nu)\,uv\,d\sigma \quad (u, v \in H_D^1(\Omega)) \tag{13}$$

whence, using notations $\tilde{c} := c - \operatorname{div}\mathbf{b}$ and $\tilde{\alpha} := \alpha + \mathbf{b}\cdot\nu$,

$$\|Q_S u\|_S = \sup_{\substack{v\in H_D^1(\Omega)\\ \|v\|_S=1}} |\langle Q_S u, v\rangle_S| = \sup_{\substack{v\in H_D^1(\Omega)\\ \|v\|_S=1}} \left| -\int_\Omega u(\mathbf{b}\cdot\nabla v)\,dx + \int_\Omega \tilde{c}uv\,dx + \int_{\Gamma_N}\tilde{\alpha}\,uv\,d\sigma \right|\,.$$

Using the embedding estimates

$$\|v\|_{L^2(\Omega)} \leq C_\Omega\|v\|_S\,, \qquad \|v\|_{L^2(\Gamma_N)} \leq C_{\Gamma_N}\|v\|_S \qquad (v \in H_D^1(\Omega)) \tag{14}$$

(where $C_\Omega, C_{\Gamma_N} > 0$) and $\|\nabla v\|_{L^2(\Omega)} \leq p^{-1/2}\|v\|_S$, and letting $K_1 := p^{-1/2}\|\mathbf{b}\|_{L^\infty(\Omega)} + C_\Omega\|\tilde{c}\|_{L^\infty(\Omega)}$, $K_2 := C_{\Gamma_N}\|\tilde{\alpha}\|_{L^\infty(\Gamma_N)}$, we obtain

$$\|Q_S u\|_S \leq K_1\|u\|_{L^2(\Omega)} + K_2\|u\|_{L^2(\Gamma_N)} \tag{15}$$

whence Q_S is compact.

It remains to prove that if $A_1 \neq \mu A_2$ then J_S is not compact. Using Proposition 2.1 (d), it suffices to find a sequence $(u_i) \subset H_0^1(\Omega) \subset H_D^1(\Omega)$ satisfying

$$\langle u_i, u_j \rangle_S = \langle J_S u_i, u_j \rangle_S = 0 \qquad (i \neq j), \tag{16}$$

$$\inf_i |\langle J_S u_i, u_i \rangle_S| / \|u_i\|_S^2 = \delta > 0. \tag{17}$$

Let $A := A_1 - \mu A_2$. Since A is not identically zero, there is $x_0 \in \Omega$ such that $A_0 := A(x_0) \neq 0$. Here A_0 is symmetric, hence there is $u_0 \in H_0^1(\Omega)$ such that $\int_\Omega A_0 \, \nabla u_0 \cdot \nabla u_0 \neq 0$. Let

$$\varepsilon := \Big| \int_\Omega A_0 \, \nabla u_0 \cdot \nabla u_0 \Big| / \Big(\int_\Omega |\nabla u_0|^2 \Big), \qquad \Omega_{\varepsilon/2} := \{ x \in \Omega : \ \|A(x) - A_0\| < \varepsilon/2 \}$$

which is an open set since A is continuous. Fix $z' \in \Omega$, and for any $z \in \Omega$ and $R > 0$ let $\Omega_{z,R} := \{ x \in \mathbf{R}^d : \ z' + R(x-z) \in \Omega \}$. Let $z_i \in \Omega$, $R_i > 0$ $(i \in \mathbf{N}^+)$ such that $\Omega_i := \Omega_{z_i, R_i} \subset \Omega_{\varepsilon/2}$ and $\overline{\Omega}_i$ are pairwise disjoint sets. We define $u_i \in H_0^1(\Omega)$ by $u_i(x) := u_0(z' + R_i(x - z_i))$ for $x \in \Omega_i$ and $u_i(x) := 0$ for $x \in \Omega \setminus \Omega_i$. Since supp $u_i = \overline{\Omega}_i$ are disjoint, (16) is satisfied. Further, using the fact $\Omega_i \subset \Omega_{\varepsilon/2}$ and a linear transformation $\Omega_i \to \Omega$ in the integral, we obtain

$$\frac{|\langle J_S u_i, u_i \rangle_S|}{\int_{\Omega_i} |\nabla u_i|^2} = \frac{\left| \int_{\Omega_i} A \, \nabla u_i \cdot \nabla u_i \right|}{\int_{\Omega_i} |\nabla u_i|^2} \geq \frac{\left| \int_{\Omega_i} A_0 \, \nabla u_i \cdot \nabla u_i \right|}{\int_{\Omega_i} |\nabla u_i|^2} - \frac{\varepsilon}{2} = \frac{\left| \int_\Omega A_0 \, \nabla u_0 \cdot \nabla u_0 \right|}{\int_\Omega |\nabla u_0|^2} - \frac{\varepsilon}{2} = \frac{\varepsilon}{2}.$$

Since for $u \in H_0^1(\Omega)$ have $\|u\|_S^2 \leq C \cdot \int_\Omega |\nabla u|^2$, the above estimate yields (17) with $\delta = \frac{\varepsilon}{2C} > 0$. ∎

4 Compact-equivalent preconditioning and mesh independent superlinear convergence rates

We prove the mesh independent convergence results for the PCGM in four stages. First we consider symmetric preconditioning operators, which is more straightforward to handle. Then, by suitable modifications of the proof, we turn to arbitrary preconditioning operators (in the studied coercive framework) where the general result is obtained. In both the symmetric and nonsymmetric cases we first consider an abstract Hilbert space level, then derive the corresponding estimates for elliptic problems.

The distinction between the symmetric and nonsymmetric preconditioners also reflects that the symmetric ones are in general much more relevant in practice. For nonsymmetric preconditioners our result has mostly a theoretical interest for its generality, nevertheless, some practical examples will be given here as well.

For simplicity we will consider compact-equivalence with $\mu = 1$ in (10), which is clearly no restriction, since if a preconditioner N_S satisfies $L_S = \mu N_S + Q_S$ then we can consider the preconditioner μN_S instead.

4.1 The abstract operator equation and its discretization

Let us consider the operator equation

$$Lu = g \tag{18}$$

where $L \in BC_S(H)$ and $g \in H$, and let $u \in H_S$ be the weak solution as in Definition 3.3. Equation (18) will be solved numerically using a Galerkin discretization: let

$$V_h = span\{\varphi_1, \ldots, \varphi_n\} \subset H_S,$$

where φ_i are linearly independent, be a given finite-dimensional subspace and

$$\mathbf{L}_h := \left\{ \langle L_S \varphi_i, \varphi_j \rangle_S \right\}_{i,j=1}^n .$$

Finding the discrete solution $u_h \in V_h$ requires solving the $n \times n$ system

$$\mathbf{L}_h \, \mathbf{c} = \mathbf{b} \tag{19}$$

with $\mathbf{b} = \{\langle g, \varphi_j \rangle\}_{j=1}^n$. Since $L \in BC_S(H)$, the symmetric part of $\mathbf{L}_h$ is positive definite, hence system (19) has a unique solution. Moreover, if a sequence of such subspaces V_h satisfies $\inf_{v \in V_h} \|u - v\|_S \to 0$ for all $u \in H_S$, then the coercivity of L_S implies in the standard way [7] that u_h converges to the exact weak solution in H_S-norm.

4.2 Symmetric preconditioning in Hilbert space

We introduce the stiffness matrix of S

$$\mathbf{S}_h = \left\{ \langle \varphi_i, \varphi_j \rangle_S \right\}_{i,j=1}^n \tag{20}$$

as preconditioner for system (19), and wish to solve

$$\mathbf{S}_h^{-1} \mathbf{L}_h \, \mathbf{c} = \tilde{\mathbf{b}} \tag{21}$$

(with $\tilde{\mathbf{b}} = \mathbf{S}_h^{-1} \mathbf{b}$) using the CGM. Let us endow $\mathbf{R}^n$ with the $\mathbf{S}_h$-inner product $\langle \mathbf{c}, \mathbf{d} \rangle_{\mathbf{S}_h} := \mathbf{S}_h \mathbf{c} \cdot \mathbf{d}$. Then the $\mathbf{S}_h$-adjoint of $\mathbf{S}_h^{-1} \mathbf{L}_h$ is $\mathbf{S}_h^{-1} \mathbf{L}_h^T$, hence we apply the CG algorithm (6) with $A = \mathbf{S}_h^{-1} \mathbf{L}_h$ and $A^* = \mathbf{S}_h^{-1} \mathbf{L}_h^T$.

Let us now assume that L and S are compact-equivalent with $\mu = 1$. In this special case (10) holds with $N_S = I$:

$$L_S = I + Q_S . \tag{22}$$

Hence, letting

$$\mathbf{Q}_h = \left\{ \langle Q_S \varphi_i, \varphi_j \rangle_S \right\}_{i,j=1}^n ,$$

system (21) takes the form

$$\left(\mathbf{I}_h + \mathbf{S}_h^{-1} \mathbf{Q}_h \right) \mathbf{c} = \tilde{\mathbf{b}} \tag{23}$$

where $\mathbf{I}_h$ is the $n \times n$ identity matrix. Using (7), the CG algorithm (6) thus provides the estimate

$$\left(\frac{\|r_k\|_{\mathbf{S}_h}}{\|r_0\|_{\mathbf{S}_h}} \right)^{1/k} \leq \frac{2}{k \nu_h} \sum_{i=1}^k \left(|\lambda_i(\mathbf{S}_h^{-1} \mathbf{Q}_h^T + \mathbf{S}_h^{-1} \mathbf{Q}_h)| + \lambda_i(\mathbf{S}_h^{-1} \mathbf{Q}_h^T \mathbf{S}_h^{-1} \mathbf{Q}_h) \right) \tag{24}$$

$(k = 1, 2, ..., n)$, where

$$\nu_h = \min_{\substack{\mathbf{c} \in \mathbf{R}^n \\ \mathbf{c} \neq \mathbf{0}}} \frac{\|\mathbf{S}_h^{-1} \mathbf{L}_h \mathbf{c}\|_{\mathbf{S}_h}^2}{\|\mathbf{c}\|_{\mathbf{S}_h}^2} . \tag{25}$$

Our goal is to give a bound on (24) that is independent of the subspace V_h.

Proposition 4.1 *Let L be S-bounded and S-coercive. Let $\mathbf{S}_h$, $\mathbf{Q}_h$ be defined as above and let $s_i(Q_S)$ and $\lambda_i(Q_S^* + Q_S)$ $(i = 1, 2, \dots)$ denote the corresponding singular values resp. ordered eigenvalues where Q_S, defined in (22), is compact on H_S. Then the following relations hold:*

(a)

$$\sum_{i=1}^k \lambda_i(\mathbf{S}_h^{-1} \mathbf{Q}_h^T \mathbf{S}_h^{-1} \mathbf{Q}_h) \leq \sum_{i=1}^k s_i(Q_S)^2 \qquad (k = 1, \dots, n),$$

9

(b) $$\sum_{i=1}^{k}|\lambda_i(\mathbf{S}_h^{-1}\mathbf{Q}_h^T + \mathbf{S}_h^{-1}\mathbf{Q}_h)| \le \sum_{i=1}^{k}|\lambda_i(Q_S^* + Q_S)| \qquad (k = 1,\dots,n),$$

(c) $$\nu_h \ge m^2 \quad \textit{for } \nu_h \textit{ in } (25), \quad \textit{where} \quad m := \inf_{\substack{u\in D(L)\\ u\neq 0}} \frac{\langle Lu, u\rangle}{\|u\|_S^2}.$$

PROOF. (a) Let $\lambda_i := \lambda_i(\mathbf{S}_h^{-1}\mathbf{Q}_h^T \mathbf{S}_h^{-1}\mathbf{Q}_h)$ $(i = 1,...,n)$ and let $\mathbf{c}^i = (c_1^i,\dots,c_n^i) \in \mathbf{R}^n$ be corresponding eigenvectors such that

$$\mathbf{S}_h\,\mathbf{c}^i \cdot \mathbf{c}^l = \delta_{il} \qquad (i,l = 1,...,n), \tag{26}$$

where $\cdot$ denotes the ordinary inner product on $\mathbf{R}^n$. Then

$$\mathbf{S}_h^{-1}\mathbf{Q}_h\,\mathbf{c}^i \cdot \mathbf{Q}_h\,\mathbf{c}^i = \lambda_i \qquad (i = 1,...,n). \tag{27}$$

Let $\mathbf{d}^i := \mathbf{S}_h^{-1}\mathbf{Q}_h\,\mathbf{c}^i$ for all i, that is

$$\mathbf{S}_h\,\mathbf{d}^i = \mathbf{Q}_h\,\mathbf{c}^i \tag{28}$$

which turns (27) into

$$\mathbf{S}_h\mathbf{d}^i \cdot \mathbf{d}^i = \lambda_i. \tag{29}$$

Now let $u_i = \sum_{j=1}^{n} c_j^i\varphi_j \in V_h$ and $z_i = \sum_{j=1}^{n} d_j^i\varphi_j \in V_h$ $(i = 1,...,n)$. Then (29) yields

$$\|z_i\|_S^2 = \lambda_i. \tag{30}$$

Further, for all $v = \sum_{j=1}^{n} p_j\varphi_j \in V_h$, with notation $\mathbf{p} = (p_1,\dots,p_n) \in \mathbf{R}^n$, (28) yields $\mathbf{S}_h\,\mathbf{d}^i \cdot \mathbf{p} = \mathbf{Q}_h\,\mathbf{c}^i \cdot \mathbf{p}$, which implies

$$\langle z_i, v\rangle_S = \langle Q_S u_i, v\rangle_S \qquad (v \in V_h),$$

i.e. z_i is the orthogonal projection of $Q_S u_i \in H_S$ into V_h. Therefore $\|z_i\|_S \le \|Q_S u_i\|_S$, and (30) provides

$$\sum_{i=1}^{k}\lambda_i \le \sum_{i=1}^{k}\|Q_S u_i\|_S^2 = \sum_{i=1}^{k}\langle Q_S^* Q_S u_i, u_i\rangle_S. \tag{31}$$

Here $\langle u_i, u_l\rangle_S = \mathbf{S}_h\,\mathbf{c}^i \cdot \mathbf{c}^l$ for all $i,l = 1,...,n$, hence by (26) the vectors u_i are orthonormal in H_S. Therefore Proposition 2.1 (a) for the operator $C = Q_S^* Q_S$ in the space H_S yields the desired estimate.

(b) The proof is similar to that of (a). Now let $\lambda_i := \lambda_i(\mathbf{S}_h^{-1}\mathbf{Q}_h^T + \mathbf{S}_h^{-1}\mathbf{Q}_h)$ and let $\mathbf{c}^i = (c_1^i,\dots,c_n^i) \in \mathbf{R}^n$ be corresponding eigenvectors with property (26). Then

$$(\mathbf{Q}_h^T + \mathbf{Q}_h)\,\mathbf{c}^i = \lambda_i\,\mathbf{S}_h\,\mathbf{c}^i \qquad (i = 1,...,n)$$

and (26) yields

$$\lambda_i = (\mathbf{Q}_h^T + \mathbf{Q}_h)\,\mathbf{c}^i \cdot \mathbf{c}^i = 2\,\mathbf{Q}_h\,\mathbf{c}^i \cdot \mathbf{c}^i.$$

For $u_i = \sum_{j=1}^{n} c_j^i\varphi_j \in V_h$ we thus obtain

$$\sum_{i=1}^{k}|\lambda_i| = 2\sum_{i=1}^{k}|\langle Q_S u_i, u_i\rangle_S| = \sum_{i=1}^{k}|\langle (Q_S^* + Q_S)u_i, u_i\rangle_S|, \tag{32}$$

10

and Proposition 2.1 (a) for the operator $C = Q_S^* + Q_S$ in the space H_S yields the desired estimate.

(c) We have

$$
\min_{\substack{\mathbf{c}\in\mathbf{R}^n \\ \mathbf{c}\neq\mathbf{0}}} \frac{\|\mathbf{S}_h^{-1}\mathbf{L}_h\mathbf{c}\|_{\mathbf{S}_h}}{\|\mathbf{c}\|_{\mathbf{S}_h}} = \min_{\substack{\mathbf{c}\in\mathbf{R}^n \\ \mathbf{c}\neq\mathbf{0}}} \frac{\|\mathbf{S}_h^{-1}\mathbf{L}_h\mathbf{c}\|_{\mathbf{S}_h}\|\mathbf{c}\|_{\mathbf{S}_h}}{\|\mathbf{c}\|_{\mathbf{S}_h}^2} \geq \min_{\substack{\mathbf{c}\in\mathbf{R}^n \\ \mathbf{c}\neq\mathbf{0}}} \frac{\langle\mathbf{S}_h^{-1}\mathbf{L}_h\mathbf{c},\mathbf{c}\rangle_{\mathbf{S}_h}}{\|\mathbf{c}\|_{\mathbf{S}_h}^2} = \min_{\substack{\mathbf{c}\in\mathbf{R}^n \\ \mathbf{c}\neq\mathbf{0}}} \frac{\mathbf{L}_h\,\mathbf{c}\cdot\mathbf{c}}{\mathbf{S}_h\,\mathbf{c}\cdot\mathbf{c}}
$$

$$
= \min_{\substack{u\in V_h \\ u\neq 0}} \frac{\langle L_S u, u\rangle_S}{\|u\|_S^2} \geq \inf_{\substack{u\in H_S \\ u\neq 0}} \frac{\langle L_S u, u\rangle_S}{\|u\|_S^2} = \inf_{\substack{u\in D(L) \\ u\neq 0}} \frac{\langle L_S u, u\rangle_S}{\|u\|_S^2} = \inf_{\substack{u\in D(L) \\ u\neq 0}} \frac{\langle Lu, u\rangle}{\|u\|_S^2} = m
$$

where the density of $D(L)$ in H_S has been used. ∎

In virtue of (24) and Proposition 4.1, we have proved

Theorem 4.1 *Let L be S-bounded and S-coercive, and let L and S be compact-equivalent with $\mu = 1$. Let the compact operator Q_S be as in (22). Then for any subspace $V_h = span\{\varphi_1,\ldots,\varphi_n\} \subset H_S$, the CG algorithm (6) with $\mathbf{S}_h$-inner product, applied for the $n \times n$ preconditioned system (21), yields*

$$
\left(\frac{\|r_k\|_{\mathbf{S}_h}}{\|r_0\|_{\mathbf{S}_h}}\right)^{1/k} \leq \varepsilon_k \qquad (k = 1, 2, ..., n) \tag{33}
$$

$$
where \quad \varepsilon_k = \frac{2}{km^2} \sum_{i=1}^{k} \Big(|\lambda_i(Q_S^* + Q_S)| + s_i(Q_S)^2\Big) \;\to\; 0 \qquad (as\ k\to\infty) \tag{34}
$$

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

4.3 Symmetric preconditioning for discretized elliptic problems

4.3.1 General elliptic equations

Let us consider an elliptic problem

$$
\begin{cases} Lu \equiv -\mathrm{div}\,(A\,\nabla u) + \mathbf{b}\cdot\nabla u + cu = g \\ u_{|\Gamma_D} = 0, \qquad \frac{\partial u}{\partial \nu_A} + \alpha u_{|\Gamma_N} = 0 \end{cases} \tag{35}
$$

where L satisfies Assumptions 3.2 and $g \in L^2(\Omega)$. We define $H_D^1(\Omega) = \{u \in H^1(\Omega) : u_{|\Gamma_D} = 0\}$, then Assumptions 3.2 ensure that problem (35) has a unique weak solution $u \in H_D^1(\Omega)$. Now let $V_h = span\{\varphi_1,\ldots,\varphi_n\} \subset H_D^1(\Omega)$ be a given FEM subspace. We seek the FEM solution $u_h \in V_h$, which requires solving the $n \times n$ system

$$
\mathbf{L}_h\,\mathbf{c} = \mathbf{b} \tag{36}
$$

where

$$
(\mathbf{L}_h)_{i,j} = \int_\Omega \Big(A\,\nabla\varphi_i\cdot\nabla\varphi_j + (\mathbf{b}\cdot\nabla\varphi_i)\varphi_j + c\varphi_i\varphi_j\Big) + \int_{\Gamma_N} \alpha\varphi_i\varphi_j\,d\sigma
$$

and $\mathbf{b}_j = \int_\Omega g\varphi_j$ ($j = 1,...,n$). Following subsection 4.2, we define a preconditioner for system (36) as the discretization of a suitable symmetric elliptic operator. Let

$$
Su := -\mathrm{div}\,(A\,\nabla u) + hu \qquad for\ u \in H^2(\Omega):\ u_{|\Gamma_D} = 0,\ \frac{\partial u}{\partial \nu_A} + \beta u_{|\Gamma_N} = 0, \tag{37}
$$

where $h \in L^\infty(\Omega)$ and $h \geq 0$ if $\Gamma_D \neq \emptyset$ and $h \geq \delta_0 > 0$ if $\Gamma_D = \emptyset$, further, $\beta \in L^\infty(\Gamma_N)$ and $\beta \geq 0$. The corresponding inner product on $H_D^1(\Omega)$ is

$$\langle u, v \rangle_S := \int_\Omega (A \, \nabla u \cdot \nabla v + huv) + \int_{\Gamma_N} \beta uv \, d\sigma \,. \tag{38}$$

We introduce the matrix

$$\mathbf{S}_h = \left\{ \langle \varphi_i, \varphi_j \rangle_S \right\}_{i,j=1}^n \tag{39}$$

as preconditioner for system (36), and then solve system (21) using the CG algorithm (6) with the $\mathbf{S}_h$-inner product and with $A = \mathbf{S}_h^{-1} \mathbf{L}_h$ and $A^* = \mathbf{S}_h^{-1} \mathbf{L}_h^T$.

Theorem 4.2 *Let $V_h \subset H_D^1(\Omega)$ be an arbitrary FEM subspace and consider the FEM discretization (36) of problem (35), using the stiffness matrix $\mathbf{S}_h$ as preconditioner. Then the superlinear convergence of the preconditioned CGM is mesh independent in the sense of Theorem 4.1, i.e., we have*

$$\left(\frac{\|r_k\|_{\mathbf{S}_h}}{\|r_0\|_{\mathbf{S}_h}} \right)^{1/k} \leq \varepsilon_k \qquad (k = 1, 2, ..., n) \tag{40}$$

for the mesh independent sequence $\varepsilon_k \to 0$ from (34).

PROOF. The coercivity and boundedness assumptions on the coefficients of L and S imply in a standard way that L is S-bounded and S-coercive. Proposition 3.1 yields that L and S are compact-equivalent in $H_D^1(\Omega)$ if the latter is endowed with the inner product (38). Therefore Theorem 4.1 is valid with the compact operator Q_S defined via

$$\langle Q_S u, v \rangle_S = \int_\Omega \left((\mathbf{b} \cdot \nabla u)v + (c - h)uv \right) + \int_{\Gamma_N} (\alpha - \beta)uv \, d\sigma \qquad (u, v \in H_D^1(\Omega)), \tag{41}$$

which satisfies (22). ∎

We note that the above result is an extension of [6], where the mesh independence property has been proved for Dirichlet boundary conditions when either S is the symmetric part of L, or both L and S have constant coefficients.

Remark 4.1 Finding the correction terms in algorithm (6) with the present choice $A = \mathbf{S}_h^{-1} \mathbf{L}_h$ and $A^* = \mathbf{S}_h^{-1} \mathbf{L}_h^T$ are equivalent to the auxiliary problems

$$\text{find } z_k \in V_h : \qquad \langle z_k, v \rangle_S = \langle L_S d_k, v \rangle_S \qquad (v \in V_h),$$
$$\text{find } s_{k+1} \in V_h : \qquad \langle s_{k+1}, v \rangle_S = \langle L_S^* d_k, v \rangle_S \qquad (v \in V_h),$$

i.e., z_k and s_{k+1} are the FEM solutions in V_h of the symmetric elliptic problems of the form $Sz_k = Ld_k$ and $Ss_{k+1} = L^* r_{k+1}$ with the boundary conditions of (37).

Proposition 4.2 *Under the conditions of Theorem 4.2, the sequence ε_k in (40) satisfies*

$$\varepsilon_k \leq \frac{4s}{k} \sum_{i=1}^k \frac{1}{\mu_i}, \tag{42}$$

where μ_i $(i \in \mathbf{N}^+)$ are the solutions of eigenvalue problem

$$Su = \mu u, \quad u_{|\Gamma_D} = 0, \quad r \left(\frac{\partial u}{\partial \nu_A} + \beta u \right)_{|\Gamma_N} = \mu u \tag{43}$$

and $s, r > 0$ are constants defined below. When the asymptotics $\mu_i = O(i^{2/d})$ holds (in particular, for Dirichlet boundary conditions),

$$\varepsilon_k \le O\Big(\frac{\log k}{k}\Big) \quad \text{if} \ \ d = 2 \quad \text{and} \quad \varepsilon_k \le O\Big(\frac{1}{k^{2/d}}\Big) \quad \text{if} \ \ d \ge 3. \tag{44}$$

PROOF. From (41) and (13) for $v = u$, letting $d = c - h$ and $\gamma = \alpha - \beta$, we obtain

$$\langle Q_S u, u\rangle_S = \int_\Omega \Big(d - \frac{1}{2}(\text{div } \mathbf{b})\Big)u^2 \ + \int_{\Gamma_N} \Big(\gamma + \frac{1}{2}(\mathbf{b} \cdot \nu)\Big)u^2 \, d\sigma \le C_1 \|u\|_{L^2(\Omega)}^2 + C_2 \|u\|_{L^2(\Gamma_N)}^2.$$

We have $|\langle (Q_S^* + Q_S)u, u\rangle_S| = 2|\langle Q_S u, u\rangle_S|$, hence the variational characterization of the eigenvalues yields

$$|\lambda_i(Q_S^* + Q_S)| = \min_{H_{i-1} \subset H_S} \max_{\substack{u \perp H_{i-1} \\ u \neq 0}} \frac{|\langle (Q_S^* + Q_S)u, u\rangle_S|}{\|u\|_S^2} \le 2 \min_{H_{i-1} \subset H_S} \max_{\substack{u \perp H_{i-1} \\ u \neq 0}} \frac{C_1 \|u\|_{L^2(\Omega)}^2 + C_2 \|u\|_{L^2(\Gamma_N)}^2}{\|u\|_S^2},$$

where H_{i-1} stands for an arbitrary $(i-1)$-dimensional subspace. On the other hand, here Q_S falls into the type (12), hence (15) implies

$$\|Q_S u\|_S^2 \le 2K_1^2 \|u\|_{L^2(\Omega)}^2 + 2K_2^2 \|u\|_{L^2(\Gamma_N)}^2.$$

Since $s_i(Q_S)^2 = \lambda_i(Q_S^* Q_S)$ and $\langle Q_S^* Q_S u, u\rangle_S = \|Q_S u\|_S^2$, we obtain as above that

$$s_i(Q_S)^2 = \min_{H_{i-1} \subset H_S} \max_{\substack{u \perp H_{i-1} \\ u \neq 0}} \frac{\langle Q_S^* Q_S u, u\rangle_S}{\|u\|_S^2} \le \min_{H_{i-1} \subset H_S} \max_{\substack{u \perp H_{i-1} \\ u \neq 0}} \frac{2K_1^2 \|u\|_{L^2(\Omega)}^2 + 2K_2^2 \|u\|_{L^2(\Gamma_N)}^2}{\|u\|_S^2}.$$

Altogether, letting $\ \ s := \frac{C_1 + K_1^2}{m^2}, \quad r := \frac{C_1 + K_1^2}{C_2 + K_2^2}$, formula (34) implies

$$\varepsilon_k \le \frac{4s}{k} \sum_{i=1}^k \hat\mu_i \qquad \text{where} \ \ \hat\mu_i = \min_{H_{i-1} \subset H_S} \max_{\substack{u \perp H_{i-1} \\ u \neq 0}} \frac{\|u\|_{L^2(\Omega)}^2 + \frac{1}{r}\|u\|_{L^2(\Gamma_N)}^2}{\|u\|_S^2},$$

in which the fraction equals $1/\mu$ for (43), hence the equality $\hat\mu_i = \frac{1}{\mu_i}$ follows from the variational characterization of the eigenvalues.

Estimate (44) follows from the asymptotics $\ \mu_i = O(i^{2/d})$ by an elementary calculation. For Dirichlet boundary conditions, this asymptotic behaviour is found in [9]. ∎

Remark 4.2 The asymptotic behaviour $\mu_i = O(i^{2/d})$ is not known for general (other than Dirichlet) boundary conditions up to the authors' knowledge. However, for the simple special case $-\Delta u = \mu u$, $\frac{\partial u}{\partial \nu}|_{\partial\Omega} = \mu u$ where Ω is a disc in $\mathbf{R}^2$, one can easily verify via the sign properties of the Bessel functions that μ_i are asymptotic to the Dirichlet eigenvalues and hence also satisfy $\mu_i = O(i^{2/d})$. This suggests the wider validity of this asymptotic rate.

4.3.2 An example: convection-diffusion equations with Helmholtz preconditioners

As a special case of the preceding paragraph, let us consider the case of a convection-diffusion operator L in (35) and a preconditioning operator S with constant coefficients. Namely, if $A \equiv I$ in (35) then we have the problem

$$\begin{cases} Lu \equiv -\Delta u + \mathbf{b}(x) \cdot \nabla u + c(x)u = g(x) \\ u_{|\Gamma_D} = 0, \qquad \frac{\partial u}{\partial \nu} + \alpha(x)u_{|\Gamma_N} = 0 \end{cases} \tag{45}$$

where for emphasis, the dependence of the coefficients on x has now been indicated unlike before. Let us define the preconditioning operator

$$Su := -\Delta u + hu \qquad \text{for} \;\; u \in H^2(\Omega) : \; u_{|\Gamma_D} = 0, \; \tfrac{\partial u}{\partial \nu} + \beta u_{|\Gamma_N} = 0, \qquad (46)$$

where $h, \beta \in \mathbf{R}$ are constants such that $h \geq 0$ if $\Gamma_D \neq \emptyset$ and $h > 0$ if $\Gamma_D = \emptyset$, further, $\beta \geq 0$.

Then the auxiliary problems with this preconditioning are discrete Helmholtz problems with constant coefficients. For such problems various fast solvers are available (like fast Fourier transform, cyclic reduction or multigrid, see e.g. [17, 25, 26]) which, together with the mesh independence result of Theorem 4.2, turns $\mathbf{S}_h$ into an efficient preconditioner. We point out that this is an extension of [6], where the mesh independence property has been proved for Dirichlet boundary conditions under the strong restriction that the operator L itself has constant coefficients.

4.3.3 Elliptic systems

Analogously to paragraph 4.3.1, we can consider elliptic systems

$$\left.\begin{aligned} L_i u &\equiv -\operatorname{div}(A_i \nabla u_i) + \mathbf{b}_i \cdot \nabla u_i + \sum_{j=1}^{l} V_{ij} u_j = g_i \\ u_{i \,|\Gamma_D} &= 0, \qquad \tfrac{\partial u_i}{\partial \nu_A} + \alpha_i u_{i\,|\Gamma_N} = 0 \end{aligned}\right\} \qquad (i = 1, \ldots, l) \qquad (47)$$

where Ω, A_i and α_i are as in Assumptions 3.2, $\mathbf{b}_i \in C^1(\overline{\Omega})^N$, $g_i \in L^2(\Omega)$, $V_{ij} \in L^\infty(\Omega)$. We assume that $\mathbf{b}_i$ and the matrix $V = \{V_{ij}\}_{i,j=1}^{l}$ satisfy the coercivity property

$$\lambda_{min}(V + V^T) - \max_i \operatorname{div} \mathbf{b}_i \geq 0 \qquad (48)$$

pointwise on Ω, where λ_{min} denotes the smallest eigenvalue, then system (47) has a unique weak solution $u \in H_D^1(\Omega)^l$. Now we choose a FEM subspace $V_h \subset H_D^1(\Omega)^l$ and look for the solution of the corresponding algebraic system $\mathbf{L}_h \mathbf{c} = \mathbf{b}$. We define the preconditioning operator $S = (S_1, \ldots, S_l)$ as the l-tuple of independent operators

$$S_i u_i := -\operatorname{div}(A_i \nabla u) + h_i u \qquad \text{for} \;\; u_i \in H^2(\Omega) : \; u_{i\,|\Gamma_D} = 0, \; \tfrac{\partial u_i}{\partial \nu_A} + \beta_i u_{i\,|\Gamma_N} = 0 \qquad (49)$$

$(i = 1, \ldots, l)$ with the conditions of (37), and let $\mathbf{S}_h$ be the stiffness matrix of S in $H_D^1(\Omega)^l$.

Then, similarly to paragraph 4.3.1, one can verify that the superlinear convergence of the preconditioned CGM is mesh independent in the sense of Theorem 4.1, i.e., (33)–(34) hold.

This result is an extension of [22] where the above preconditioning has been introduced and its efficient parallelizability has been demonstrated; on the other hand, the mesh independence property was proved there for Dirichlet boundary conditions under strong restrictions on the matrix V (antisymmetric, or normal when the operator L itself has constant coefficients).

4.4 Nonsymmetric preconditioning in Hilbert space

Now let N be a general (possibly nonsymmetric) S-bounded and S-coercive operator which is compact-equivalent to L with $\mu = 1$, i.e., (10) becomes

$$L_S = N_S + Q_S. \qquad (50)$$

We introduce the stiffness matrix of N_S

$$\mathbf{N}_h = \left\{ \langle N_S \varphi_i, \varphi_j \rangle_S \right\}_{i,j=1}^{n}$$

14

as preconditioner for system (19), and wish to solve

$$\mathbf{N}_h^{-1}\mathbf{L}_h\,\mathbf{c} = \hat{\mathbf{b}} \tag{51}$$

(with $\hat{\mathbf{b}} = \mathbf{N}_h^{-1}\mathbf{b}$) using the CGM. Since N is nonsymmetric, in order to define an inner product on $\mathbf{R}^n$ we preserve the stiffness matrix of S on V_h, i.e. using (20) we endow $\mathbf{R}^n$ with the $\mathbf{S}_h$-inner product $\langle \mathbf{c}, \mathbf{d} \rangle_{\mathbf{S}_h} := \mathbf{S}_h\,\mathbf{c} \cdot \mathbf{d}$ as earlier. Then the $\mathbf{S}_h$-adjoint of $\mathbf{N}_h^{-1}\mathbf{L}_h$ is $\mathbf{S}_h^{-1}\mathbf{L}_h^T\mathbf{N}_h^{-T}\mathbf{S}_h$, hence we apply the CG algorithm (6) with $A = \mathbf{N}_h^{-1}\mathbf{L}_h$ and $A^* = \mathbf{S}_h^{-1}\mathbf{L}_h^T\mathbf{N}_h^{-T}\mathbf{S}_h$.

Letting

$$\mathbf{Q}_h = \left\{ \langle Q_S\varphi_i, \varphi_j\rangle_S \right\}_{i,j=1}^n,$$

system (21) takes the form

$$(\mathbf{I}_h + \mathbf{N}_h^{-1}\mathbf{Q}_h)\,\mathbf{c} = \hat{\mathbf{b}} \tag{52}$$

where $\mathbf{I}_h$ is the $n \times n$ identity matrix. Using (7), the CG algorithm (6) thus provides

$$\left(\frac{\|r_k\|_{\mathbf{S}_h}}{\|r_0\|_{\mathbf{S}_h}} \right)^{1/k} \leq \frac{2}{k\nu_h}\sum_{i=1}^k\left(\lambda_i(\mathbf{S}_h^{-1}\mathbf{Q}_h^T\mathbf{N}_h^{-T}\mathbf{S}_h + \mathbf{N}_h^{-1}\mathbf{Q}_h) + \lambda_i(\mathbf{S}_h^{-1}\mathbf{Q}_h^T\mathbf{N}_h^{-T}\mathbf{S}_h\mathbf{N}_h^{-1}\mathbf{Q}_h) \right) \tag{53}$$

$(k = 1, 2, ..., n)$ where

$$\nu_h = \min_{\mathbf{c}\in\mathbf{R}^n}\frac{\|\mathbf{N}_h^{-1}\mathbf{L}_h\mathbf{c}\|_{\mathbf{S}_h}^2}{\|\mathbf{c}\|_{\mathbf{S}_h}^2}\,. \tag{54}$$

Again, our goal is to give a bound on (53) that is independent of V_h.

Proposition 4.3 *Let L and N be S-bounded and S-coercive operators, in particular*

$$m := \inf_{\substack{u\in D(L)\\ u\neq 0}}\frac{\langle Lu, u\rangle}{\|u\|_S^2} > 0, \qquad \hat{m} := \inf_{\substack{u\in D(N)\\ u\neq 0}}\frac{\langle Nu, u\rangle}{\|u\|_S^2} > 0, \qquad \hat{M} := \sup_{\substack{u\in D(N)\\ u\neq 0}}\frac{|\langle Nu, v\rangle|}{\|u\|_S\|v\|_S} > 0,$$

and let Q_S be a compact operator on H_S. Let $\mathbf{S}_h$, $\mathbf{N}_h$ and $\mathbf{Q}_h$ be defined as above, and let $s_i(Q_S)$ $(i = 1, 2, \ldots)$ denote the singular values of Q_S. Then the following relations hold:

(a)
$$\sum_{i=1}^k\lambda_i(\mathbf{S}_h^{-1}\mathbf{Q}_h^T\mathbf{N}_h^{-T}\mathbf{S}_h\mathbf{N}_h^{-1}\mathbf{Q}_h) \leq \frac{1}{\hat{m}^2}\sum_{i=1}^k s_i(Q_S)^2 \qquad (k = 1, \ldots, n),$$

(b)
$$\sum_{i=1}^k|\lambda_i(\mathbf{S}_h^{-1}\mathbf{Q}_h^T\mathbf{N}_h^{-T}\mathbf{S}_h + \mathbf{N}_h^{-1}\mathbf{Q}_h)| \leq \frac{2}{\hat{m}}\sum_{i=1}^k s_i(Q_S) \qquad (k = 1, \ldots, n),$$

(c)
$$\nu_h \geq \frac{m^2}{\hat{M}^2}\,.$$

PROOF. (a) We proceed similarly to Proposition 4.1. Let $\lambda_i := \lambda_i(\mathbf{S}_h^{-1}\mathbf{Q}_h^T\mathbf{N}_h^{-T}\mathbf{S}_h\mathbf{N}_h^{-1}\mathbf{Q}_h)$ $(i = 1, ..., n)$ and let $\mathbf{c}^i = (c_1^i, \ldots, c_n^i) \in \mathbf{R}^n$ be corresponding eigenvectors with property (26). Then

$$\mathbf{S}_h\mathbf{N}_h^{-1}\mathbf{Q}_h\,\mathbf{c}^i\cdot\mathbf{N}_h^{-1}\mathbf{Q}_h\,\mathbf{c}^i = \lambda_i \qquad (i = 1, ..., n). \tag{55}$$

Let $\mathbf{d}^i := \mathbf{N}_h^{-1}\mathbf{Q}_h\,\mathbf{c}^i$ for all i, that is

$$\mathbf{N}_h\,\mathbf{d}^i = \mathbf{Q}_h\,\mathbf{c}^i\,. \tag{56}$$

For this $\mathbf{d}^i$ and λ_i, similarly to Proposition 4.1, we have (29) and, letting $u_i = \sum_{j=1}^{n} c_j^i \varphi_j \ \in V_h$ and $z_i = \sum_{j=1}^{n} d_j^i \varphi_j \ \in V_h$ we obtain (30). Further, for all $v = \sum_{j=1}^{n} p_j \varphi_j \ \in V_h$, with notation $\mathbf{p} = (p_1, \ldots, p_n) \in \mathbf{R}^n$, (56) yields $\mathbf{N}_h \, \mathbf{d}^i \cdot \mathbf{p} = \mathbf{Q}_h \, \mathbf{c}^i \cdot \mathbf{p}$, which means

$$\langle N_S z_i, v \rangle_S = \langle Q_S u_i, v \rangle_S \qquad (v \in V_h).$$

From this we have

$$\|z_i\|_S^2 \leq \frac{1}{\hat{m}} \langle N_S z_i, z_i \rangle_S = \frac{1}{\hat{m}} \langle Q_S u_i, z_i \rangle_S \leq \frac{1}{\hat{m}} \|Q_S u_i\|_S \|z_i\|_S \,,$$

hence $\|z_i\|_S \leq \frac{1}{\hat{m}} \|Q_S u_i\|_S$. Then from (30)

$$\sum_{i=1}^{k} \lambda_i \leq \frac{1}{\hat{m}^2} \sum_{i=1}^{k} \|Q_S u_i\|_S^2 = \frac{1}{\hat{m}^2} \sum_{i=1}^{k} \langle Q_S^* Q_S u_i, u_i \rangle_S, \tag{57}$$

whence the desired estimate follows in the same way as from (31) in Proposition 4.1.

(b)   Now let $\lambda_i := \lambda_i(\mathbf{S}_h^{-1} \mathbf{Q}_h^T \mathbf{N}_h^{-T} \mathbf{S}_h + \mathbf{N}_h^{-1} \mathbf{Q}_h)$ and let $\mathbf{c}^i = (c_1^i, \ldots, c_n^i) \in \mathbf{R}^n$ be corresponding eigenvectors with property (26). Then

$$\lambda_i = \lambda_i \mathbf{S}_h \, \mathbf{c}^i \cdot \mathbf{c}^i = \mathbf{Q}_h^T \, \mathbf{N}_h^{-T} \mathbf{S}_h \, \mathbf{c}^i \cdot \mathbf{c}^i + \mathbf{S}_h \mathbf{N}_h^{-1} \mathbf{Q}_h \, \mathbf{c}^i \cdot \mathbf{c}^i = 2 \, \mathbf{S}_h \mathbf{N}_h^{-1} \mathbf{Q}_h \, \mathbf{c}^i \cdot \mathbf{c}^i = 2 \, \mathbf{Q}_h \, \mathbf{c}^i \cdot \mathbf{e}^i$$

where $\mathbf{e}^i := \mathbf{N}_h^{-T} \mathbf{S}_h \, \mathbf{c}^i$ for all i. Here for all $v = \sum_{j=1}^{n} p_j \varphi_j \ \in V_h$, with notation $\mathbf{p} = (p_1, \ldots, p_n) \in \mathbf{R}^n$, we obtain $\mathbf{e}^i \cdot \mathbf{N}_h \, \mathbf{p} = \mathbf{S}_h \, \mathbf{c}^i \cdot \mathbf{p}$, which means $\langle w_i, N_S v \rangle_S = \langle u_i, v \rangle_S$ for all $v \in V_h$, where $w_i = \sum_{j=1}^{n} e_j^i \varphi_j$ and $u_i = \sum_{j=1}^{n} c_j^i \varphi_j$, or

$$\langle N_S^* w_i, v \rangle_S = \langle u_i, v \rangle_S \qquad (v \in V_h). \tag{58}$$

Denote by P the orthogonal projection of H_S onto V_h. Then (58) yields $u_i = P N_S^* w_i$. Here the linear mapping $(P N_S^*)_{|V_h} : V_h \to V_h$ is one-to-one, since for all $v \in V_h$

$$\langle P N_S^* v, v \rangle_S = \langle N_S^* v, v \rangle_S = \langle N_S v, v \rangle_S \geq \hat{m} \|v\|_S^2. \tag{59}$$

Therefore

$$\mathbf{Q}_h \, \mathbf{c}^i \cdot \mathbf{e}^i = \langle Q_S u_i, w_i \rangle_S = \langle Q_S u_i, (P N_S^*)_{|V_h}^{-1} u_i \rangle_S = \langle u_i, Q_S^* (P N_S^*)_{|V_h}^{-1} u_i \rangle_S.$$

Here the operator $(P N_S^*)_{|V_h}^{-1}$ has a norm-preserving extension $\hat{N}$ from V_h onto H_S (namely, with $\hat{N}|_{(V_h)\perp} := 0$), and from (59) we have $\|\hat{N}\| \leq \frac{1}{\hat{m}}$. Altogether, we obtain

$$\sum_{i=1}^{k} |\lambda_i| = 2 \sum_{i=1}^{k} |\langle Q_S^* (P N_S^*)_{|V_h}^{-1} u_i, u_i \rangle_S| = 2 \sum_{i=1}^{k} |\langle Q_S^* \hat{N} u_i, u_i \rangle_S| \leq 2 \sum_{i=1}^{k} s_i(Q_S^* \hat{N})$$

$$\leq \frac{2}{\hat{m}} \sum_{i=1}^{k} s_i(Q_S^*) = \frac{2}{\hat{m}} \sum_{i=1}^{k} s_i(Q_S)$$

(where, in the inequalities, statments (a) and (b) of Proposition 2.1 have been used, respectively).

(c) Let $\mathbf{c} \in V_h$ be arbitrary, $\mathbf{d} := \mathbf{N}_h^{-1}\mathbf{L}_h\mathbf{c}$. Let $u = \sum\limits_{j=1}^{n} c_j\varphi_j \in V_h$ and $z = \sum\limits_{j=1}^{n} d_j\varphi_j \in V_h$.

Then $m\|u\|_S^2 \leq \langle L_S u, u \rangle_S = \mathbf{L}_h\,\mathbf{c}\cdot\mathbf{c} = \mathbf{N}_h\,\mathbf{d}\cdot\mathbf{c} = \langle N_S z, u \rangle_S \leq \|N_S z\|_S \|u\|_S$, hence

$$m\|u\|_S \leq \|N_S z\|_S$$

and

$$\frac{\|\mathbf{N}_h^{-1}\mathbf{L}_h\mathbf{c}\|_{\mathbf{S}_h}^2}{\|\mathbf{c}\|_{\mathbf{S}_h}^2} = \frac{\mathbf{S}_h\,\mathbf{d}\cdot\mathbf{d}}{\mathbf{S}_h\,\mathbf{c}\cdot\mathbf{c}} = \frac{\|z\|_S^2}{\|u\|_S^2} \geq m^2\,\frac{\|z\|_S^2}{\|N_S z\|_S^2} \geq \frac{m^2}{\hat{M}^2}\,. \qquad \blacksquare$$

In virtue of (53) and Proposition 4.3, we have proved

Theorem 4.3 *Let L and N be S-bounded and S-coercive operators that are compact-equivalent in H_S with $\mu = 1$. Let the compact operator Q_S be as in (50). Then for any subspace $V_h = \mathrm{span}\{\varphi_1, \ldots, \varphi_n\} \subset H_S$, the CG algorithm (6) with $\mathbf{S}_h$-inner product, applied for the $n \times n$ preconditioned system (51), yields*

$$\left(\frac{\|r_k\|_{\mathbf{S}_h}}{\|r_0\|_{\mathbf{S}_h}}\right)^{1/k} \leq \varepsilon_k \qquad (k = 1, 2, ..., n) \tag{60}$$

where $\quad \varepsilon_k = \dfrac{2\hat{M}^2}{km^2}\sum\limits_{i=1}^{k}\Big(\dfrac{2}{\hat{m}}\,s_i(Q_S) + \dfrac{1}{\hat{m}^2}\,s_i(Q_S)^2\Big) \;\to\; 0 \qquad (as\ k \to \infty)$ $\tag{61}$

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

Remark 4.3 When one preconditions L with N, a useful choice for the operator S is the symmetric part of N: i.e., if $D(N) = D(N^*)$ then $S = (N + N^*)/2$, and if $D(N) \neq D(N^*)$ then S is an operator that generates the inner product satisfying $\langle u, v \rangle_S := \frac{1}{2}\big(\langle Nu, v \rangle + \langle u, Nv \rangle\big)$ for $u, v \in D(N)$, see [21]. Then in Proposition 4.3 we have $\langle Nu, u \rangle = \|u\|_S^2$ $(u \in D(N))$, hence $\hat{m} = 1$.

4.5 Nonsymmetric preconditioning for discretized elliptic problems

This section contains our most general result for elliptic operators: in the studied coercive framework, preconditioning with an arbitrary operator N that is compact-equivalent with L provides mesh independent superlinear convergence. Although this property has mostly a theoretical importance, some practical examples are given here as well. Let us first consider the elliptic problem (35):

$$\begin{cases} Lu \equiv -\mathrm{div}\,(A\,\nabla u) + \mathbf{b}\cdot\nabla u + cu = g \\[2mm] u_{|\Gamma_D} = 0, \qquad \frac{\partial u}{\partial \nu_A} + \alpha u_{|\Gamma_N} = 0 \end{cases} \tag{62}$$

and let us now define the nonsymmetric preconditioning operator

$$Nu := -\mathrm{div}\,(A\,\nabla u) + \mathbf{w}\cdot\nabla u + zu \qquad \text{for } u \in H^2(\Omega): u_{|\Gamma_D} = 0, \tfrac{\partial u}{\partial \nu_A} + \eta u_{|\Gamma_N} = 0, \tag{63}$$

where L and N satisfy Assumptions 3.2 in the obvious sense, further, $g \in L^2(\Omega)$. Accordingly, the preconditioner for the discretized problem (36) is the nonsymmetric stiffness matrix

$$(\mathbf{N}_h)_{i,j} = \int_\Omega \Big(A\,\nabla\varphi_i\cdot\nabla\varphi_j + (\mathbf{w}\cdot\nabla\varphi_i)\,\varphi_j + z\varphi_i\varphi_j\Big) + \int_{\Gamma_N} \eta\varphi_i\varphi_j\,d\sigma\,.$$

We use the same energy space as in the symmetric case, i.e. $H_S = H_D^1(\Omega)$ with inner product (38). We then solve the preconditioned system using the CG algorithm (6) with the $\mathbf{S}_h$-inner product and with $A = \mathbf{N}_h^{-1}\mathbf{L}_h$ and $A^* = \mathbf{S}_h^{-1}\mathbf{L}_h^T\mathbf{N}_h^{-T}\mathbf{S}_h$.

Theorem 4.4 *Let $V_h \subset H_D^1(\Omega)$ be an arbitrary FEM subspace and consider the FEM discretization (36) of problem (35), using the stiffness matrix $\mathbf{N}_h$ as preconditioner. Then the superlinear convergence of the preconditioned CGM is mesh independent in the sense of Theorem 4.3, i.e., (60)–(61) hold.*

PROOF. Similar to that of Theorem 4.2, now Theorem 4.3 is applied in $H_D^1(\Omega)$. ■

Example. Let us consider problem (45), i.e. when in (62) we have

$$Lu = -\Delta u + \mathbf{b}(x) \cdot \nabla u + c(x)u,$$

where for emphasis, the dependence of the coefficients on x has now been indicated. Assume that we are in 2D and, say, $b_1(x)$ has larger values than $b_2(x)$, where $\mathbf{b}(x) = (b_1(x), b_2(x))$ and $x = (x_1, x_2)$. Then one can propose the preconditioning operator

$$Nu = -\Delta u + w_1 \frac{\partial u}{\partial x_1} + zu \qquad \text{for} \ \ u \in H^2(\Omega): \ u_{|\Gamma_D} = 0, \ \frac{\partial u}{\partial \nu} + \eta u_{|\Gamma_N} = 0, \qquad (64)$$

where $w_1, z, \eta \in \mathbf{R}$ are constants such that $z \geq 0$ if $\Gamma_D \neq \emptyset$ and $z > 0$ if $\Gamma_D = \emptyset$, further, $\eta \geq 0$. For convection-dominated problems (i.e. when $|\mathbf{b}|$ is large), the presence of the nonsymmetric term $w_1 \frac{\partial u}{\partial x_1}$ may turn N into a much better approximationion of L than a symmetric preconditioner like (46). Nevertheless, since this term is one-dimensional, the solution of the auxiliary problems remains considerably simpler than the original one, e.g. via local 1D Green's functions [3]. The above operator N has been proposed in [6], where the mesh independence result of the PCGM has been proved for Dirichlet boundary conditions under the strong restriction that the operator L itself has constant coefficients.

Analogously to the symmetric case in paragraph 4.3.3, the above results can be extended to systems in a straightforward way. Namely, let us consider system (47) and introduce the preconditioning operator N as an l-tuple of decoupled operators N_i, where each N_i is of the type (63). Then the superlinear convergence of the preconditioned CGM is mesh independent in the sense of Theorem 4.3, i.e., (60)–(61) hold. Since N_i are decoupled, the resulting algorithm is parallelizable. This turns it into an efficient method if, for instance, each N_i is like (64), or the problem itself is in 1D which occurs e.g. after using the method of splitting in meteorological models with several components.

5 Remarks on singular perturbation problems

For singular perturbation problems such as

$$L_\varepsilon u \equiv -\varepsilon \Delta u + \mathbf{b} \cdot \nabla u + cu = f$$

(plus boundary conditions), where $\varepsilon > 0$ but $\varepsilon << \|\mathbf{b}\|$, one cannot neglect the first order term when forming a preconditioner. Such problems are characterized by thin boundary and/or interior layers and the diffusion term plays a noticable role only in the layer. This property

is not exploited in preconditioners like (64). A possible approach to handle such problems is therefore to use the following defect-correction method:

$$L_{\delta(x)}(u_{k+1} - u_k) = f - L_{\varepsilon}u \qquad (k \in \mathbf{N}^+)$$

where u_0 is given, and in practice only one or two steps need to be performed. Here

$$L_{\delta(x)}u := -\delta(x)\Delta u + \mathbf{b} \cdot \nabla u + cu$$

where $\delta(x) = 0$ outside the layers and increases continuously along each characteristic line (defined by the velocity vector $\mathbf{b}$) from zero to ε in the layers. The widths of the layers are typically chosen as $\varepsilon \log(1/\varepsilon)$. To solve the correction equation by iteration, one can form a preconditioner S by using the operator $\mathbf{b} \cdot \nabla u + hu$ outside the layers and $-\delta(x)\Delta u + \mathbf{b} \cdot \nabla u + hu$ in the layers for some properly chosen function $h \geq 0$. The analysis of the problem will not be considered further in the present paper.

References

[1] AXELSSON, O., A generalized conjugate gradient least square method, *Numer. Math.* 51 (1987), 209-227.

[2] AXELSSON, O., *Iterative Solution Methods,* Cambridge University Press, 1994.

[3] AXELSSON, O., GOLOLOBOV, S. V., A combined method of local Green's functions and central difference method for singularly perturbed convection-diffusion problems, *J. Comput. Appl. Math.* 161 (2003), no. 2, 245–257.

[4] AXELSSON, O., KARÁTSON J., On the rate of convergence of the conjugate gradient method for linear operators in Hilbert space, *Numer. Funct. Anal.* **23** (2002), No. 3-4, 285-302.

[5] AXELSSON, O., KARÁTSON J., Symmetric part preconditioning for the conjugate gradient method in Hilbert space, *Numer. Funct. Anal.* **24** (2003), No. 5-6, 455-474.

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

[7] CIARLET, P. G., *The Finite Element Method for Elliptic Problems,* North-Holland, Amsterdam, 1978

[8] CONCUS, P., GOLUB, G.H., A generalized conjugate method for non-symmetric systems of linear equations, in: *Lect. Notes Math. Syst.* 134 (eds. Glowinski, R., Lions, J.-L.), pp. 56-65, Springer, 1976.

[9] COURANT, R, HILBERT, D., *Methods of Mathematical Physics II.*, Wiley Classics Library, J. Wiley & Sons, 1989.

[10] 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.

[11] FABER, V., MANTEUFFEL, T., PARTER, S.V., Necessary and sufficient conditions for the existence of a conjugate gradient method, *SIAM J. Numer. Anal.* 21 (1984), no. 2, 352–362.

[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] FORTUNA, Z., Some convergence properties of the conjugate gradient method in Hilbert space, *SIAM J. Numer. Anal.* 16 (1979), no. 3, 380–384.

[14] GOHBERG, I., GOLDBERG, S., KAASHOEK, M. A., *Classes of Linear Operators*, Vol. I., Operator Theory: Advances and Applications, 49, Birkhäuser Verlag, Basel, 1990.

[15] GOHBERG, I., GOLDBERG, S., *Basic Operator Theory*, Birkhäuser, Boston, Mass., 1981.

[16] 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.

[17] HACKBUSCH, W., *Multigrid Methods and Applications*, Springer Series in Computational Mathematics 4, Springer, Berlin, 1985.

[18] HAYES, R.M., Iterative methods of solving linear problems in Hilbert space, *Nat. Bur. Standards Appl. Math. Ser*, 39 (1954), 71-104.

[19] HESTENES, M.R., STIEFEL, E., Methods of conjugate gradients for solving linear systems, *J. Res. Nat. Bur. Standards*, Sect. B, 49 (1952) No.6., 409-436.

[20] KARÁTSON J., Mesh independent superlinear convergence estimates of the conjugate gradient method for some equivalent self-adjoint operators *Appl. Math.* (Prague) 50 (2005), No. 3, 277-290.

[21] KARÁTSON J., Superlinear PCG algorithms: symmetric part preconditioning and boundary conditions, Preprint 2006-10, ELTE Dept. Appl. Anal. Comp. Math., http://www.cs.elte.hu/applanal/preprints; submitted

[22] KARÁTSON J., KURICS T., Superlinearly convergent PCG algorithms for some non-symmetric elliptic systems, Preprint 2006-09, ELTE Dept. Appl. Anal. Comp. Math., http://www.cs.elte.hu/applanal/preprints; submitted

[23] MANTEUFFEL, T., PARTER, S. V., Preconditioning and boundary conditions, *SIAM J. Numer. Anal.* 27 (1990), no. 3, 656–694.

[24] NEVANLINNA, O., *Convergence of Iterations for Linear Equations,* Birkhäuser, Basel, 1993.

[25] 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).

[26] SWARZTRAUBER, P. N., The methods of cyclic reduction, Fourier analysis and the FACR algorithm for the discrete solution of Poisson's equation on a rectangle, *SIAM Rev.* 19 (1977), no. 3, 490–501.

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

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