Available online at www.sciencedirect.com
ScienceDirect
Comput. Methods Appl. Mech. Engrg. 196 (2007) 1612–1621
Computer methods in applied mechanics and engineering
www.elsevier.com/locate/cma

Optimal left and right additive Schwarz preconditioning for minimal residual methods with Euclidean and energy norms

Marcus Sarkis [a,b,1], Daniel B. Szyld [c,*,2]

[a] *Instituto Nacional de Matemática Pura e Aplicada, Rio de Janeiro, Brazil*
[b] *Worcester Polytechnic Institute, Worcester, MA 01609, USA*
[c] *Department of Mathematics, Temple University, Philadelphia, PA 19122, USA*

Received 11 July 2005; received in revised form 30 December 2005; accepted 5 March 2006

Abstract

For the solution of non-symmetric or indefinite linear systems arising from discretizations of elliptic problems, two-level additive Schwarz preconditioners are known to be optimal in the sense that convergence bounds for the preconditioned problem are independent of the mesh and the number of subdomains. These bounds are based on some kind of *energy norm*. However, in practice, iterative methods which minimize the Euclidean norm of the residual are used, despite the fact that the usual bounds are non-optimal, i.e., the quantities appearing in the bounds may depend on the mesh size; see [X.-C. Cai, J. Zou, Some observations on the l^2 convergence of the additive Schwarz preconditioned GMRES method, Numer. Linear Algebra Appl. 9 (2002) 379–397]. In this paper, iterative methods are presented which minimize the same energy norm in which the optimal Schwarz bounds are derived, thus maintaining the Schwarz optimality. As a consequence, bounds for the Euclidean norm minimization are also derived, thus providing a theoretical justification for the practical use of Euclidean norm minimization methods preconditioned with additive Schwarz. Both left and right preconditioners are considered, and relations between them are derived. Numerical experiments illustrate the theoretical developments.

Keywords: Additive Schwarz preconditioning; Krylov subspace iterative methods; Minimal residuals; GMRES; Indefinite and non-symmetric elliptic problems; Energy norm minimization

1. Introduction

We consider minimal residual methods for the solution of non-symmetric or indefinite large systems of linear equations of the form

$$Bx = f, \tag{1}$$

where B is the discretization of a partial differential operator; see Section 2 for a description of the class of operators we consider. GMRES [24] is a popular Krylov subspace method for the iterative solution of non-symmetric linear systems, where at each step the norm of the residual is minimized over nested affine spaces of increasing dimension. The norm used in this minimization is usually taken to be the l_2 norm, i.e., the Euclidean norm associated with the standard inner product $(x, y) = x^{\mathrm{T}}y$. For references on discussions of other inner products in this context, see Section 4.

Additive Schwarz (AS) refers to a class of extensively used preconditioners for (1); we describe them in Section 2. There are two main components to their appeal. First, they are easily parallelizable, since several smaller linear systems need to be solved: one system for each of the subdomains, usually corresponding to the restriction of the differential operator to that subdomain. These are called local problems. Second, if a coarse problem is introduced,

* Corresponding author.
E-mail addresses: msarkis@fluid.impa.br (M. Sarkis), szyld@math.temple.edu (D.B. Szyld).
[1] Research supported in part by CNPQ (Brazil) under grant 305539/2003-8 and by the US National Science Foundation under grant CGR 9984404.
[2] Research supported in part by the US Department of Energy under grant DE-FG02-05ER25672.

0045-7825/$ - see front matter
doi:10.1016/j.cma.2006.03.027

they are optimal in the sense that bounds on the convergence rate of the preconditioned iterative method are independent (or slowly dependent) on the finite element mesh size and the number of subproblems; see, e.g., [22,28,30]. These bounds are given using some kind of energy norm (or equivalent Sobolev $\mathscr{H}^1$ norm), i.e., the norm induced by the A-inner product $(x,y)_A = x^{\mathrm{T}}Ay$, for some appropriate symmetric positive definite matrix A. Usually A is taken to be the symmetric part of B, i.e., $(B+B^{\mathrm{T}})/2$, if it is positive definite (i.e., if B is positive real), or some other symmetric positive definite matrix related to B; see further Section 2 for the operators we consider here.

Cai and Zou [12] pointed out that when using AS with GMRES minimizing the l_2 norm of the residual, the optimality results of AS may be lost. They show explicit examples in which the quantities used in the GMRES convergence bounds depend on the mesh size. Nevertheless, this AS/GMRES method is widely used, e.g., it is standard in PETSc [4]; see also [22,30]. In this paper, we present a version of GMRES where the minimization is done using some energy norm. In this form, we preserve the optimality of the preconditioner. Thus, both the bounds for the minimal residual method and those providing the independence of the mesh are in *the same* energy norm. By using the same energy norm in the minimization as that used to obtained the optimal bounds, one avoids the possible pitfalls of the mesh dependence in the bounds highlighted by Cai and Zou [12]. In particular we mention that while Cai and Zou [12] found that certain operators cease to be positive real (in the l_2 norm), we show that they become positive real in the energy norm; cf. the discussion in [18, p. 32].

The iterative methods using the energy norm are more expensive at each step, and thus we do not advocate their use in practice in all cases; see Remark 7.1 for cases when it might be computationally advantageous to use the energy norm minimization methods. As it turns out, the analysis of the energy norm based methods do provide the theoretical justification for the use of the Euclidean norm based methods; see Section 6. We show that asymptotically, for a fixed mesh, the two behave in the same manner. Therefore, we say that the standard AS/GMRES is asymptotically optimal. We show experimentally that for many problems the asymptotic regime occurs rather rapidly, and thus, the number of iterations to achieve a desired small tolerance is the same using either method; see Section 7.

We consider both left and right preconditioning. We show relations between these two situations, both in the Euclidean and the energy norm; see Remark 3.1 and Proposition 5.1. These relations provide us with optimality results in both left and right preconditioning.

2. Additive Schwarz methods for a class of non-symmetric problems

In this section, we follow the description of a class of non-symmetric problems from [30, Chapter 11]; see also [10,11,22,28, Section 5.4].

Let $\Omega \subset \mathbb{R}^d$ be a region of interest which is polygonal and an open bounded domain, and let $\mathscr{T}_h(\Omega)$ be a regular shaped and quasi-uniform triangulation of Ω. Let V be the traditional finite element space formed by piecewise linear and continuous functions vanishing on the boundary of Ω; for details about finite elements formulations, see, e.g., [7,8]. Consider the following discrete partial differential equation. Find $u \in V$ such that

$$b(u,v) = f(v) \quad \text{for all } v \in V,$$

where

$$b(u,v) = a(u,v) + s(u,v) + c(u,v), \tag{2}$$

$$a(u,v) = \int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}x, \tag{3}$$

$$s(u,v) = \int_\Omega (b\cdot\nabla u)v + (\nabla\cdot bu)v\,\mathrm{d}x, \quad b\in\mathbb{R}^d,$$

$$c(u,v) = \int_\Omega cuv\,\mathrm{d}x, \quad \text{and} \quad f(v) = \int_\Omega fv\,\mathrm{d}x.$$

We note that $a(\cdot,\cdot)$ is positive definite, $s(\cdot,\cdot)$ is antisymmetric, and $c(\cdot,\cdot)$ is an L_2 inner product with a weight function $c \in L^\infty$ smooth enough. Hence, if the mesh size is small enough, this problem has a unique solution [30].

Let A and B be the matrix representations of $v^{\mathrm{T}}Au = a(u,v)$ and $v^{\mathrm{T}}Bu = b(u,v)$, respectively. We mention that these matrix representations depend on the type of boundary conditions, but not on the values of the boundary conditions. Since there is a one-to-one correspondence between functions in the finite element space and nodal values, sometimes we abuse the notation and do not distinguish between them. Let $\|v\|_a = (a(v,v))^{1/2}$, and $\|v\|_A = (v^{\mathrm{T}}Av)^{1/2}$ be the corresponding norms in V and in $\mathbb{R}^n$, respectively. Considering zero Dirichlet boundary conditions and using elementary results we have:

(1) Continuity: there is a constant C, such that

$$|b(u,v)| \leqslant C\|u\|_a\|v\|_a, \quad u,v\in\mathscr{H}_0^1(\Omega).$$

(2) A Gårding inequality: there is a constant C, such that

$$\|u\|_a^2 - C\|u\|_{L^2(\Omega)}^2 \leqslant b(u,u), \quad u\in\mathscr{H}_0^1(\Omega).$$

(3) There is a constant C, such that

$$|s(u,v)| \leqslant C\|u\|_a\|v\|_{L^2(\Omega)}, \quad u,v\in\mathscr{H}_0^1(\Omega),$$

and

$$|c(u,v)| \leqslant C\|u\|_{L^2(\Omega)}\|v\|_{L^2(\Omega)}, \quad u,v\in\mathscr{H}_0^1(\Omega).$$

(4) Regularity (valid for polygonal and smooth domains): there is a constant C, independent of g, where the solution w of the adjoint equation

$$b(\phi,w) = (g,\phi), \quad \phi\in\mathscr{H}_0^1(\Omega)$$

satisfies

$$\|w\|_{\mathscr{H}^{1+\gamma}(\Omega)} \leqslant C\|g\|_{L^2(\Omega)}$$

for some $\gamma > 1/2$.

We introduce a decomposition of V into a sum of $N+1$ subspaces $R_i^{\mathrm{T}} V_i \subset V$, and

$$V = R_0^{\mathrm{T}} V_0 + R_1^{\mathrm{T}} V_1 + \cdots + R_N^{\mathrm{T}} V_N. \tag{4}$$

Here we denote by $R_i^{\mathrm{T}} : V_i \to V$ the extension operator from V_i to V. We note that the decomposition (4) is not necessarily a direct sum of subspaces. Often, the subspaces $R_i^{\mathrm{T}} V_i$, $i = 1, \ldots, N$, are related to a decomposition of the domain Ω into overlapping subregions Ω_i^δ of size O(H) covering Ω. Here δ refers to the amount of overlap between the subregions. The subspace $R_0^{\mathrm{T}} V_0$ is the coarse space. For $u_i, v_i \in V_i$ define

$$b_i(u_i, v_i) = b(R_i^{\mathrm{T}} u_i, R_i^{\mathrm{T}} v_i), \quad a_i(u_i, v_i) = a(R_i^{\mathrm{T}} u_i, R_i^{\mathrm{T}} v_i).$$

Let

$$B_i = R_i B R_i^{\mathrm{T}}, \quad A_i = R_i A R_i^{\mathrm{T}}$$

be the matrix representations of these local bilinear forms. For $i = 0, \ldots, N$, we define $\widetilde{P}_i : V \to V_i$, by

$$b_i(\widetilde{P}_i u, v_i) = b(u, R_i^{\mathrm{T}} v_i), \quad v_i \in V_i,$$

and $\widetilde{Q}_i : V \to V_i$ by

$$a_i(\widetilde{Q}_i u, v_i) = a(u, R_i^{\mathrm{T}} v_i), \quad v_i \in V_i.$$

It is possible to show that the matrices $\widetilde{Q}_i$ are well-defined (since the matrices A_i are invertible) and for H small enough the matrices $\widetilde{P}_i$ are well-defined (since the matrices B_i are invertible for small H); see [11,30]. We now set

$$P_i = R_i^{\mathrm{T}} \widetilde{P}_i = R_i^{\mathrm{T}} B_i^{-1} R_i B, \quad Q_i = R_i^{\mathrm{T}} \widetilde{Q}_i = R_i^{\mathrm{T}} A_i^{-1} R_i B,$$

and we introduce the additive operators

$$P^{(1)} = \sum_{i=0}^{N} P_i = \left(\sum_{i=0}^{N} R_i^{\mathrm{T}} B_i^{-1} R_i \right) B, \tag{5}$$

$$P^{(2)} = P_0 + \sum_{i=1}^{N} Q_i = \left(R_0^{\mathrm{T}} B_0^{-1} R_0 + \sum_{i=1}^{N} R_i^{\mathrm{T}} A_i^{-1} R_i \right) B. \tag{6}$$

The following result can be found, e.g., in [11,30].

Theorem 2.1. *There exist constants* $H_0 > 0$, $c(H_0) > 0$, $C(H_0) > 0$, *and* $C_0(\delta)$, *such that if* $H \leqslant H_0$, *then for* $i = 1, 2$, *and* $u \in V$,

$$\frac{a(u, P^{(i)} u)}{a(u, u)} \geqslant c_p, \tag{7}$$

and

$$\|P^{(i)} u\|_a \leqslant C_p \|u\|_a, \tag{8}$$

where $C_p = C(H_0)$ *and* $c_p = C_0(\delta)^{-2} c(H_0)$.

We mention that similar bounds also hold for hybrid versions of the preconditioners [30].

3. Preconditioned GMRES

In this section, we first review the standard preconditioned GMRES [24] (minimizing the Euclidean norm of the residual), which we use later as a model for other versions. We begin with left preconditioned GMRES.

It follows from the form of the preconditioners (5) and (6) that we can write generically $P^{(i)} = M^{-1} B$. The first factor is indeed non-singular, so this notation is consistent; see [6,17,21]. The left preconditioned problem is therefore given by

$$M^{-1} B x = M^{-1} f. \tag{9}$$

Let x_0 be an initial approximation, $r_0 = f - B x_0$ the corresponding initial residual, and $s_0 = M^{-1} r_0$. The left preconditioned GMRES minimizes the residual norm

$$\|M^{-1} f - M^{-1} B x\|_2 = \|M^{-1} r_0 - M^{-1} B (x - x_0)\|_2, \tag{10}$$

among all vectors x from the affine subspace

$$x_0 + \mathscr{K}_m^L = x_0 + \mathrm{span}\{s_0, M^{-1} B s_0, \ldots, (M^{-1} B)^{m-1} s_0\},$$

where $\mathscr{K}_m^L$ is the Krylov subspace generated by $M^{-1} B$ and s_0.

Let $Z_m = [z_1, \ldots, z_m]$ be a matrix whose columns are an orthonormal basis of $\mathscr{K}_m^L$, such that the Arnoldi relation

$$M^{-1} B Z_m = Z_{m+1} \overline{H}_m^L \tag{11}$$

holds, where $\overline{H}_m^L$ is $(m+1) \times m$ upper Hessenberg and $z_1 = s_0 / \beta$. Let $\overline{H}_m = \overline{H}_m^L$. It follows that since we are looking for $x - x_0 = Z_m y$ for some $y \in \mathbb{R}^m$, minimizing (10) is equivalent to finding the minimizer of the smaller problem

$$y_m = \underset{y \in \mathbb{R}^m}{\mathrm{argmin}} \|\beta e_1 - \overline{H}_m y\|_2, \tag{12}$$

and setting $x_m = x_0 + Z_m y_m$; see, e.g., [5,23], for further algorithmic details.

We present next an algorithm to compute the mth approximation x_m with left preconditioned GMRES. This algorithm corresponds to full GMRES; for restarted GMRES one sets $x_0 := x_m$ and restarts the iteration.

Algorithm 3.1

1. Compute $r_0 = f - B x_0$, $s_0 = M^{-1} r_0$, $\beta = (s_0, s_0)^{1/2}$, and $z_1 = s_0 / \beta$
2. For $j = 1, \ldots, m$, Do:
3. Compute $w := B\, z_j$, and $z := M^{-1} w$
4. For $i = 1, \ldots, j$, Do:
5. $h_{i,j} := (z, z_i)$
6. $z := z - h_{i,j} z_i$
7. EndDo
8. Compute $h_{j+1,j} = (z, z)^{1/2}$ and $z_{j+1} = z / h_{j+1,j}$
9. EndDo
10. Define $Z_m := [z_1, \ldots, z_m]$, $\overline{H}_m = \{h_{i,j}\}_{1 \leqslant i \leqslant j+1; 1 \leqslant j \leqslant m}$
11. Compute $y_m = \mathrm{argmin}_y \|\beta e_1 - \overline{H}_m y\|_2$, and $x_m = x_0 + Z_m y_m$

Observe that the main storage requirements of Algorithm 3.1 are the vectors $z_1, \ldots, z_m \in \mathbb{R}^n$.

Consider now the right preconditioned problem given by

$$BM^{-1}u = f, \tag{13}$$

where $x = M^{-1}u$. Let $u_0 = Mx_0$. The right preconditioned GMRES minimizes the residual norm $\|f - BM^{-1}u\|_2$, among all vectors u from the affine subspace

$$u_0 + \mathscr{K}_m^R = u_0 + \text{span}\{r_0, BM^{-1}r_0, \ldots, (BM^{-1})^{m-1}r_0\}.$$

That is, $u_m = u_0 + V_m y_m$, where here V_m is a matrix whose columns are an orthonormal basis of $\mathscr{K}_m^R$. The Arnoldi relation in this case is

$$BM^{-1}V_m = V_{m+1}\overline{H}_m^R. \tag{14}$$

Thus $x_m = M^{-1}u_m = x_0 + M^{-1}V_m y_m$, so that x_m can be computed directly from y_m, and in fact

$$x_m \in x_0 + M^{-1}\mathscr{K}_m^R. \tag{15}$$

A standard algorithm for right preconditioned GMRES would be similar to Algorithm 3.1, with the appropriate changes, with one set of m vectors in $\mathbb{R}^n$ as main storage requirement.

Remark 3.1. We point out that there is a close relationship between left and right preconditioned GMRES; see, e.g., [23, Section 9.3.4]. In fact, it can be seen that $M^{-1}\mathscr{K}_m^R = \mathscr{K}_m^L$ (cf. (15)), and therefore the columns of both Z_m and $M^{-1}V_m$ are bases of the same space $\mathscr{K}_m^L$. Since the columns of Z_{m+1} are orthogonal, there exists a non-singular upper triangular matrix

$$U_{m+1} = \left[\begin{array}{c|c} U_m & u_{m+1} \\ 0^{\mathrm{T}} & \end{array}\right]$$

such that

$$M^{-1}V_{m+1} = Z_{m+1}U_{m+1} = [Z_m|z_{m+1}]U_{m+1}. \tag{16}$$

Thus from (14), premultiplying by M^{-1}, and using (16) we obtain $M^{-1}BZ_mU_m = Z_{m+1}U_{m+1}\overline{H}_m^R$. Comparing this with the Arnoldi relation for left preconditioning (11), we conclude that

$$\overline{H}_m^L = U_{m+1}\overline{H}_m^R U_m^{-1};$$

cf. [16] where a similar relation is found in a different context.

4. Convergence bounds for minimal residual methods

GMRES is in fact an implementation of the generalized conjugate residual method (GCR) [14] where the same minimization

$$\|r_m\|_2 = \min_{x\in x_0+\mathscr{K}_m} \|f - Bx\|_2 \tag{17}$$

is sought where

$$\mathscr{K}_m = \mathscr{K}_m(B, r_0) = \text{span}\{r_0, Br_0, B^2r_0, \ldots, B^{m-1}r_0\}.$$

The difference is that while in GMRES, as we have seen, the basis used for $\mathscr{K}_m$ has orthogonal vectors, in GCR one constructs a basis of $\mathscr{K}_m$ which is $B^{\mathrm{T}}B$-orthogonal. There are also implementation differences. For example, as we have seen, in GMRES, the minimization problem is transformed into one of reduced size. This is performed with the QR factorization of $\overline{H}_m$, where the orthogonal matrix Q is not explicitly computed.

Thus, convergence analysis of GCR and GMRES is the same assuming exact arithmetic. We only mention GCR here to apply the convergence bounds developed for it to GMRES. There are two classical convergence bounds for these methods given in [15,14, Theorem 3.3]. We present these bounds assuming the linear system (1) with no preconditioning, i.e., $M = I$. The first of these bounds assumes that $(B + B^{\mathrm{T}})/2$, the symmetric part of B, is positive definite, i.e., that B is positive real. In this case, one has that

$$\|r_m\|_2 \leqslant \left(1 - \frac{c^2}{C^2}\right)^{m/2} \|r_0\|_2, \tag{18}$$

where for each real vector x,

$$c = \min_{x\neq 0} \frac{(x, Bx)}{(x, x)} \quad \text{and} \quad C = \max_{x\neq 0} \frac{\|Bx\|_2}{\|x\|_2}. \tag{19}$$

The bound (18) has been mentioned in conjunction with Schwarz preconditioners; see, e.g., [22,30,34], although Cai and Zou [12] present an example where the operator is not positive real (in the l_2 norm), and therefore, this bound is not applicable.

Our aim is to consider a different norm in the minimization (17). Several authors explored the theory of such a different norm, either explicitly or implicitly, and mostly in a formal manner for the classification of Krylov subspace methods; see [2,3, Chapter 12, 13,19,20,31–33], and also [16] for a weighted norm and [1] for a recent use of these ideas in a different context.

While the bound (18) and the constants (19) where originally derived using the Euclidean inner product and the associated norm, they are valid for minimal residual methods using any inner product and its induced norm; see, e.g., [13, Section 6.1, 18, 29, 22, Section 4.2, 34]. In other words, as long as $c > 0$ and C is bounded, as defined in (19) with the proper inner product and norm, then, the bound (18) applies to a minimal residual method where the minimization is taken in the same norm. In the next section we provide an appropriate inner product and corresponding energy norm for the left and right preconditioned generalized minimal residual method.

5. Preconditioned GMRES minimizing some energy norm

In this section, we derive GMRES versions minimizing the energy norm of the residual. We discuss first the left preconditioned problem (9) minimizing the A-norm of the residual, where A is a symmetric positive definite matrix. The right preconditioned problem is treated later in the section.

In terms of implementation of left preconditioned GMRES with the A-inner product, it suffices to replace appropriately each inner product in Algorithm 3.1, i.e., in steps 1, 5, and 8. For example, in step 5, we would have

$$h_{i,j} := (z, z_i)_A = z^{\mathrm{T}} A z_i. \tag{20}$$

In this manner, the vectors $z_1, \ldots, z_m$ are A-orthonormal, i.e.,

$$Z_m^{\mathrm{T}} A Z_m = I. \tag{21}$$

Note that this is different than the situation in [1] where an orthogonal basis (with respect to the Euclidean inner product) is kept. We point out that the usual Arnoldi relation (11) still holds here, but the basis matrix Z_m and the upper Hessenberg matrix $\overline{H}_m = \overline{H}_m^L$ here are different than in (11).

If $x - x_0 = Z_m y$, $y \in \mathbb{R}^m$, i.e., writing $x \in x_0 + \mathscr{K}_m(M^{-1}B, M^{-1}r_0)$ using the A-orthonormal basis, because of (21), we have that

$$\|M^{-1}f - M^{-1}Bx\|_A = \|M^{-1}r_0 - M^{-1}BZ_m y\|_A \tag{22}$$
$$= \|Z_{m+1}\beta e_1 - Z_{m+1}\overline{H}_m y\|_A = \|\beta e_1 - \overline{H}_m y\|_2, \tag{23}$$

and this is why we maintain the minimization in step 11 of Algorithm 3.1 in the l_2 norm also here. In summary, by replacing the inner products, we have a GMRES version minimizing the companion norm, i.e., the A-norm, but the smaller minimization problem (12) is still performed in the l_2 norm, in the same usual manner, e.g., using the QR factorization of $\overline{H}_m$. Let us denote by y_m^A the minimizer in (23), $x_m^A = x_0 + Z_m y_m^A$, and $r_m^A = f - Bx_m^A$, so we can distinguish the iterates and residuals of the method which minimizes the energy norm.

We observe that this algorithm, i.e., preconditioned GMRES minimizing the A-norm, can be implemented with only one matrix–vector product with A and one solution of the form $Mz = v$ per iteration, and by storing a set of additional vectors $\tilde{z}_i = Az_i$.

The preceding discussion holds for any symmetric positive definite matrix A. In the particular case where A is the discretization of (3), and B is the discretization of (2), we can use the results of Section 2 to obtain bounds on the operators used here. Specifically, Theorem 2.1 implies that there exist constants C_p and c_p such that for all real vectors x,

$$\frac{(x, M^{-1}Bx)_A}{(x,x)_A} \geqslant c_p \quad \text{and} \quad \|M^{-1}Bx\|_A \leqslant C_p \|x\|_A. \tag{24}$$

These bounds are the counterparts to (7) and (8). The bound (18) is valid for the A-norm, and the minimization in (22) is also in the same A-norm. Thus, the combination of AS with this version of GMRES has the following convergence bound independent of the finite element mesh size and the number of local problems

$$\|M^{-1}r_m^A\|_A \leqslant \left(1 - \frac{c_p^2}{C_p^2}\right)^{m/2} \|M^{-1}r_0\|_A. \tag{25}$$

We remark that this convergence bound points to the interplay between the choice of the energy norm used in the minimal residual method, i.e., the symmetric positive definite matrix A, and the choice of preconditioner M^{-1}. The idea is that the matrix $M^{-1}B$ must be positive real in the A-inner product, i.e., $c_p > 0$. In general, A should be chosen as the elliptic highest order term derivative of B. By selecting a coarse mesh and local problems sufficiently small, local Poincaré inequalities force the positiveness of $M^{-1}B$ in the A-norm. Note that from (24), $c_p/C_p \leqslant 1$. The closer the ratio c_p/C_p is to 1, the smaller is the factor in parenthesis in the convergence bound (25).

Consider now the right preconditioned system (13) where, as before, $x = M^{-1}u$. Simple calculations give

$$(x,x)_A = (u,u)_{M^{-\mathrm{T}}AM^{-1}}, \tag{26}$$
$$(x, M^{-1}Bx)_A = (u, BM^{-1}u)_{M^{-\mathrm{T}}AM^{-1}}, \tag{27}$$

and

$$(M^{-1}Bx, M^{-1}Bx)_A = (BM^{-1}u, BM^{-1}u)_{M^{-\mathrm{T}}AM^{-1}}. \tag{28}$$

Let $G = M^{-\mathrm{T}}AM^{-1}$. It follows then that we can rewrite the bounds (24) as

$$\frac{(u, BM^{-1}u)_G}{(u,u)_G} \geqslant c_p, \quad \text{and} \quad \|BM^{-1}u\|_G \leqslant C_p\|u\|_G$$

with *the same* constants c_p and C_p. Consequently, M is an optimal right preconditioner for a minimal residual method using the energy norm associated with the symmetric positive definite matrix $G = M^{-\mathrm{T}}AM^{-1}$, i.e., minimizing

$$\|r_0 - BM^{-1}u\|_{M^{-\mathrm{T}}AM^{-1}}. \tag{29}$$

A right preconditioned GMRES such that it minimizes (29) can be implemented from the standard right preconditioned GMRES by using instead the $M^{-\mathrm{T}}AM^{-1}$-inner product. For example, in the construction of the upper Hessenberg matrix one would have

$$h_{i,j} := (w, v_i)_{M^{-\mathrm{T}}AM^{-1}} = (M^{-1}w)^{\mathrm{T}}AM^{-1}v_i. \tag{30}$$

In this manner, the vectors $v_1, \ldots, v_m$ are $M^{-\mathrm{T}}AM^{-1}$-orthonormal, and in a manner similar to the left preconditioning case, we have that

$$\|r_0 - BM^{-1}u\|_{M^{-\mathrm{T}}AM^{-1}} = \|r_0 - BM^{-1}V_m y\|_{M^{-\mathrm{T}}AM^{-1}},$$
$$\|V_{m+1}(\beta e_1 - \overline{H}_m y)\|_{M^{-\mathrm{T}}AM^{-1}} = \|\beta e_1 - \overline{H}_m y\|_2. \tag{31}$$

Therefore, in the implementation of the right preconditioned GMRES which minimizes the $M^{-\mathrm{T}}AM^{-1}$-norm, the smaller least squares problem remains in the l_2 norm. Let us denote by y_m^G the minimizer in (31), $x_m^G = x_0 + Z_m y_m^G$, and $r_m^G = f - Bx_m^G$. Let $Z_m = M^{-1}V_m = [z_1, \ldots, z_m]$, then using identities (26)–(29), we can write

$$\|r_0 - BM^{-1}V_m y\|_{M^{-T}AM^{-1}} = \|M^{-1}r_0 - M^{-1}BM^{-1}V_m y\|_A$$
$$= \|\beta z_1 - M^{-1}BZ_m y\|_A.$$

In other words, for any *fixed* preconditioner M, using right preconditioning and minimizing the $M^{-\mathrm{T}}AM^{-1}$-norm of the residual, produces (in exact arithmetic) the same approximations than if one uses left preconditioning and minimizes the A-norm of the appropriately transformed residual. Furthermore, from (20) and (30) one can see that the upper Hessenberg matrices $\overline{H}_m$ in (23) and (31) are the same matrix, cf. Remark 3.1. We summarize this in the following result.

Proposition 5.1. *For every preconditioner M and every symmetric positive definite matrix A, the minimal residual method for the left preconditioned problem $M^{-1}B\,x = M^{-1}b$ using the A-inner product is completely equivalent (in exact arithmetic) to a minimal residual method for the right preconditioned problem $BM^{-1}u = b$, $M^{-1}u = x$, using the G-inner product, with $G = M^{-T}AM^{-1}$. In particular this holds for $A = I$, i.e., for the Euclidean inner product. Conversely, if we have a right preconditioned problem $BM^{-1}u = b$, $M^{-1}u = x$, with the Euclidean inner product, it is completely equivalent to the left preconditioned problem $M^{-1}Bx = M^{-1}b$ using the A-inner product where $A = M^TM$, so that $M^{-T}AM^{-1} = I$.*

We remark that here we have the same upper Hessemberg matrix $\overline{H}_m$ for both left and right preconditioning, but with different norms, while in Remark 3.1 we have the same norm, but different upper Hessemberg matrices.

From Proposition 5.1 it follows that for the right preconditioned GMRES with $M^{-\mathrm{T}}AM^{-1}$-norm, we have the same convergence bound (25), with the same constants, i.e.,

$$\|r_m^G\|_G \leqslant \left(1 - \frac{c_p^2}{C_p^2}\right)^{m/2} \|r_0\|_G. \tag{32}$$

In terms of implementation, one can then use Algorithm 3.1 with the A-inner product. It goes without saying that while optimality of AS with GMRES is assured, there is the cost of one matrix–vector product with the symmetric positive definite matrix A in each iteration.

6. Bounds for AS/GMRES in Euclidean norm

We begin this section by discussing the constants of equivalency between an energy norm and the Euclidean norm.

Proposition 6.1. *Let H be a symmetric positive definite matrix, and the associated inner product $(x,y)_H = x^{\mathrm{T}}Hy$ and norm $\|x\|_H = (x,x)_H^{1/2}$. Then for any vector x one has*

$$\|x\|_2 \leqslant c_H\|x\|_H, \quad \text{and} \quad \|x\|_H \leqslant C_H\|x\|_2,$$

where $c_H = 1/\sqrt{\lambda_{\min}(H)}$, $C_H = \sqrt{\lambda_{\max}(H)}$, and $\lambda_{\min}(H)$, $\lambda_{\max}(H)$ represent the minimum and maximum eigenvalues of H, respectively.

Proof. Let

$$c_H^2 = \sup_{x\neq 0}\frac{\|x\|_2^2}{\|x\|_H^2} = \left(\inf_{x\neq 0}\frac{\|x\|_H^2}{\|x\|_2^2}\right)^{-1} = \left(\inf_{x\neq 0}\frac{x^{\mathrm{T}}Hx}{x^{\mathrm{T}}x}\right)^{-1} = \frac{1}{\lambda_{\min}(H)},$$

and the first inequality follows. The proof of the second inequality is analogous. □

We relate now the norm of the residual r_m^L of the usual left preconditioned GMRES method minimizing the Euclidean norm, i.e., obtained using Algorithm 3.1, with r_m^A obtained from the left preconditioned GMRES minimizing the energy norm defined by the symmetric positive definite matrix A. We use the constants $c_A = 1/\sqrt{\lambda_{\min}(A)}$, $C_A = \sqrt{\lambda_{\max}(A)}$, and $\kappa(A) = c_A^2C_A^2$, the condition number of A. Then we have that

$$\begin{aligned}\|M^{-1}r_m^L\|_2 &\leqslant \|M^{-1}r_m^A\|_2 \leqslant c_A\|M^{-1}r_m^A\|_A \\ &\leqslant c_A\left(1-\frac{c_p^2}{C_p^2}\right)^{m/2}\|M^{-1}r_0\|_A \\ &\leqslant c_AC_A\left(1-\frac{c_p^2}{C_p^2}\right)^{m/2}\|M^{-1}r_0\|_2 \\ &= \sqrt{\kappa(A)}\left(1-\frac{c_p^2}{C_p^2}\right)^{m/2}\|M^{-1}r_0\|_2,\end{aligned} \tag{33}$$

where the first inequality follows from the fact that r_m is the minimizing residual (in the Euclidean norm), the second from Proposition 6.1, the third from (18), and the constants c_p and C_p come from (24).

In a similar fashion, we obtain bounds for the residual norm of r_m^R obtained using right preconditioned AS/GMRES minimizing the Euclidean norm and relate these to those of r_m^G obtained using right preconditioned AS/GMRES minimizing the G-norm. Using the same arguments, and (32), we have

$$\begin{aligned}\|r_m^R\|_2 &\leqslant \|r_m^G\|_2 \leqslant c_G\|r_m^G\|_G \leqslant c_G\left(1-\frac{c_p^2}{C_p^2}\right)^{m/2}\|r_0\|_G \\ &\leqslant c_GC_G\left(1-\frac{c_p^2}{C_p^2}\right)^{m/2}\|r_0\|_2 \\ &= \sqrt{\kappa(G)}\left(1-\frac{c_p^2}{C_p^2}\right)^{m/2}\|r_0\|_2,\end{aligned} \tag{34}$$

where $c_G = 1/\sqrt{\lambda_{\min}(G)}$, $C_G = \sqrt{\lambda_{\max}(G)}$, and $\kappa(G) = c_G^2C_G^2$ is the condition number of G.

Several observations regarding the bounds (33) and (34) are in order. These bounds show that the (left and right preconditioned) AS/GMRES (using Euclidean norm minimization) is asymptotically optimal, in the sense that other than the factor $\sqrt{\kappa(A)}$ or $\sqrt{\kappa(G)}$ (which do depend on the mesh size) the convergence is independent of the mesh size or the number of subdomains. These fixed factors are eventually overtaken by the other factor being reduced with each iteration. The positive definite matrix A defining the energy norm needs to be sufficiently far from being singular, i.e., $\lambda_{\min(A)}$ sufficiently far from zero, so that the asymptotic behavior takes hold. This is usually the case in practice, and it is illustrated with examples in the next section, where one has that $\lambda_{\min}(A) = \mathcal{O}(1)$, and $\kappa(A) = \mathcal{O}(1/h^2)$. Note that the constants c_p and C_p depend on the existence of the positive definite matrix A (or operator $a(u,v)$) for which (24) hold. In other words, we have derived the asymptotic optimality of AS/GMRES (using Euclidean norm minimization) through the optimality of the method using an energy norm.

We also see from the above bounds that the 2-norm of the usual GMRES residual differs from that of the energy norm GMRES residual by no more than a factor c_A or c_G (which is fixed for all x_0 and all m). Thus, asymptotically, as the residuals go to zero, their norms behave in the same manner. This fact is well illustrated in some examples in the next section.

7. Numerical experiments

We present numerical experiments associated to partial differential equations of the form $-\Delta u + b \cdot \nabla u + k\, u = 1$, with zero Dirichlet boundary conditions on the two-dimensional unit square; these are particular cases of (2). The three cases we investigate are

(A) The Helmholtz equation where we take $b^{\mathrm{T}} = [0,0]$, and two different values $k = -5$ and $k = -120$, the latter being indefinite.

(B) The implicit one-step time discretization of an advection–diffusion equation, where $b^{\mathrm{T}} = [10,20]$, $k = 1$, and upwind discretization is used.

We consider the following four mesh and domain decomposition configurations: mesh 64×64 elements decomposed on 4×4 subdomains; mesh 128×128 decomposed on 4×4 or 8×8 subdomains; and mesh 256×256 decomposed on 8×8 subdomains. For each of these cases, we consider three different amounts of overlap $\delta = 0$, $\delta = 1$, or $\delta = 2$. An overlap of $\delta = 0$ indicates one layer of overlapping nodes, i.e., the interface nodes, while $\delta = 1$ or *2* correspond to three or five layers of overlapping nodes, respectively. The coarse space is based on partition of unity with one degree of variables per subdomains [9,25–27]. We consider the additive preconditioner $P^{(1)}$ of (5). We show results using right preconditioning. In the figures we plot the either Euclidean norm or the G-norm the residual using the two strategies: using the right GMRES with G-norm minimization (plotted with $(*)$) and using the standard right GMRES with Euclidean norm minimization (plotted with $(\circ)$). In all cases our tolerance for the relative residual norm is $\varepsilon = 10^{-8}$. Recall that the right GMRES with G-norm minimization is equivalent to the left GMRES with A-norm minimization; see Proposition 5.1.

We present in Figs. 1–3 representative runs for the Helmholtz equation, and in Figs. 4–6 representative runs

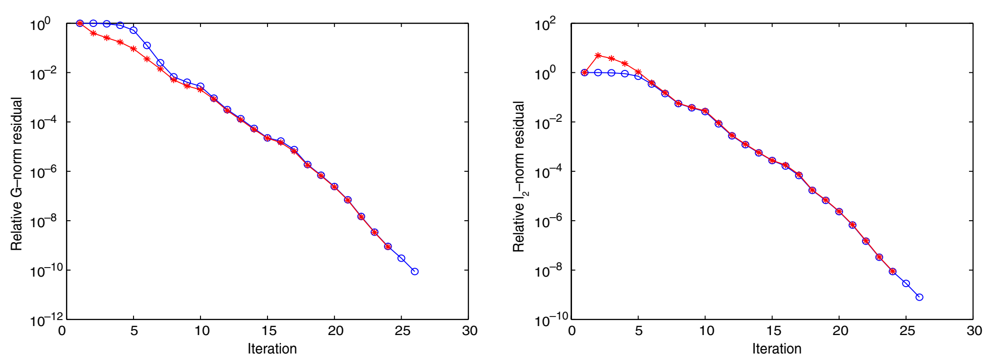

Fig. 1. Problem A. Helmholtz equation with $k = -5$. Relative residual norms for GMRES minimizing the l_2 norm (○), and the G-norm $(*)$. 64×64 grid, 4×4 subdomains, $\delta = 0$. Left: residuals measured in the G-norm. Right: residuals measured in the l_2 norm.

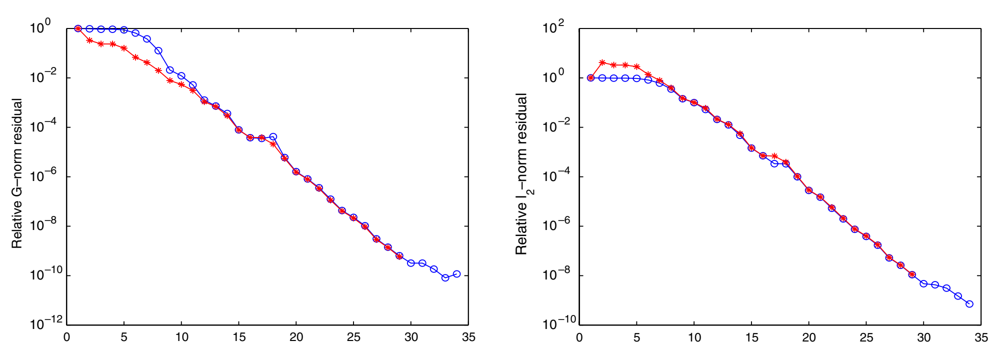

Fig. 2. Problem A. Helmholtz equation with $k = -120$. Relative residual norms for GMRES minimizing the l_2 norm (○), and the G-norm $(*)$. 128×128 grid, 8×8 subdomains, $\delta = 1$. Left: residuals measured in the G-norm. Right: residuals measured in the l_2 norm.

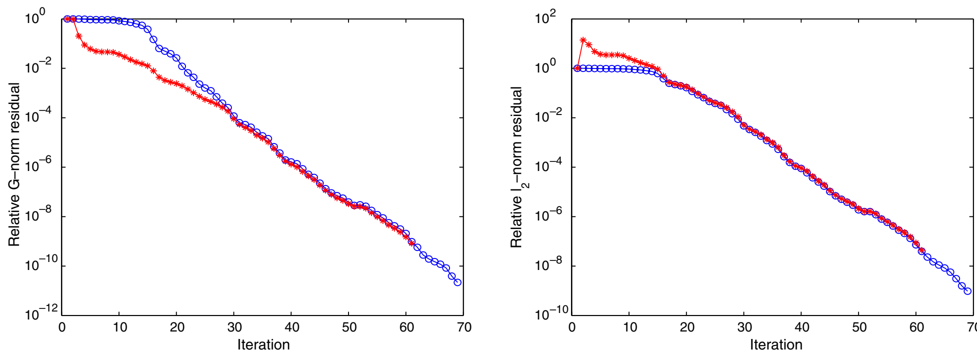

Fig. 3. Problem A. Helmholtz equation with $k = -120$. Relative residual norms for GMRES minimizing the l_2 norm (○), and the G-norm ($\ast$). 256×256 grid, 8×8 subdomains, $\delta = 0$. Left: residuals measured in the G-norm. Right: residuals measured in the l_2 norm.

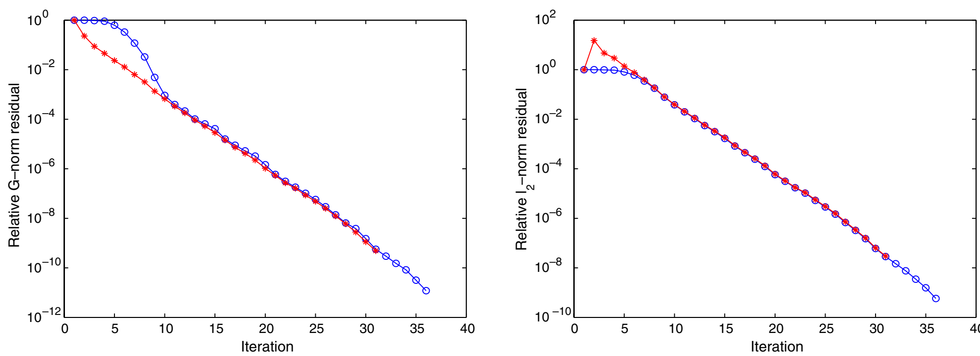

Fig. 4. Problem B. Advection–diffusion equation. Relative residual norms for GMRES minimizing the l_2 norm (○), and the G-norm ($\ast$). 64×64 grid, 4×4 subdomains, $\delta = 0$. Left: residuals measured in the G-norm. Right: residuals measured in the l_2 norm.

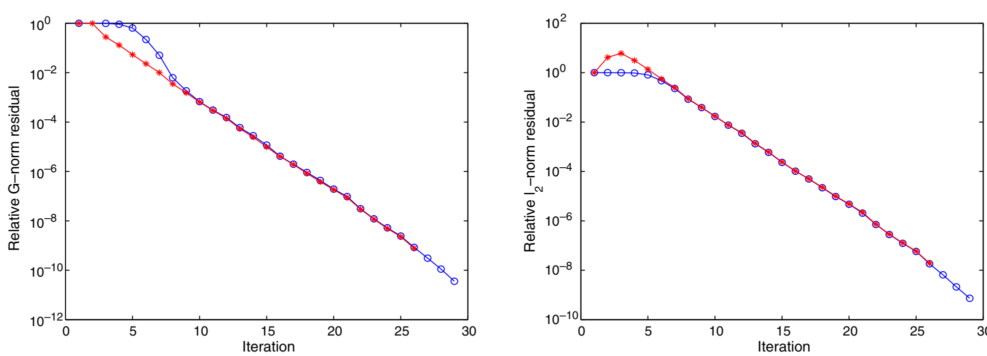

Fig. 5. Problem B. Advection–diffusion equation. Relative residual norms for GMRES minimizing the l_2 norm (○), and the G-norm ($\ast$). 128×128 grid, 4×4 subdomains, $\delta = 2$. Left: residuals measured in the G-norm. Right: residuals measured in the l_2 norm.

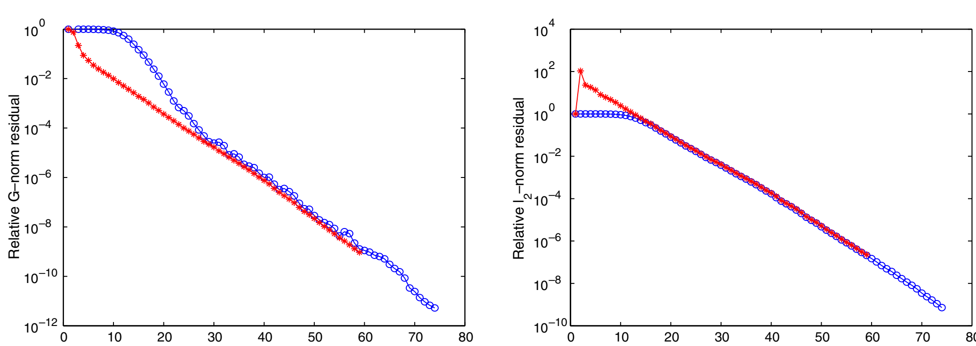

Fig. 6. Problem B. Advection–diffusion equation. Relative residual norms for GMRES minimizing the l_2 norm (○), and the G-norm ($\ast$). 256×256 grid, 8×8 subdomains, $\delta = 0$. Left: residuals measured in the G-norm. Right: residuals measured in the l_2 norm.

for the advection–diffusion equation. In each figure, we show the same problem solved with the standard AS/GMRES minimizing the Euclidean norm, and with the method minimizing the G-norm. We present the same results in two different graphs, one, on the left, measuring the two residuals r_m^R and r_m^G in the G-norm, and the second, on the right, measuring them in the Euclidean norm. It can be appreciated from these figures that, as expected, $\|r_m^G\|_G \leqslant \|r_m^R\|_G$ (left plots), and that $\|r_m^R\|_2 \leqslant \|r_m^G\|_2$ (right plots). It can also been clearly seen how asymptotically the two sequences of residual norms are very close to each other, and that the asymptotic regime begins well before the method reaches the desired tolerance.

Remark 7.1. Depending on the problem, and especially if a low tolerance desired, it may turn out to be less expensive to reach the desired tolerance in the energy norm than in the l_2 norm. In addition, the energy norm may be more meaningful. We call the reader's attention to Fig. 6 where $\|M^{-1}r_m^A\|_A$ falls below 10^{-4} after 22 iterations, while it takes 40 iterations for $\|M^{-1}r_m^L\|_2$ to fall below the same tolerance. Thus, in this case the additional cost of one matrix–vector product with the SPD matrix A per step is more than offset by the savings in number of iterations.

We report in Tables 1 and 2 results on runs with the usual AS/GMRES (l_2 norm minimization) with the two cases of the Helmholtz problem considered here, and in Table 3 with the advection–diffusion problem already mentioned. We show the number of iterations to reach a relative residual norm below 10^{-8} for all the meshes described, and three different levels of overlap. In the tables, n stands for the number of points in one side of the mesh, and $n$sub, in parenthesis, the numbers of subdomains in each side of the square considered.

As it can be appreciated in these tables, while the number of iterations is not constant across each row, i.e., for each preconditioner considered, they do not grow unbounded; indeed they only about double when the value of h is reduced by a factor of four, i.e., when the cell size is reduced by a factor of 16.

Table 1
Problem A. Helmholtz equation with k = −5

n ($n$sub)	64 (2)	128 (4)	128 (8)	256 (8)
$\delta = 0$	23	34	30	49
$\delta = 1$	16	23	20	33
$\delta = 2$	13	18	16	27

Number of iterations for AS/GMRES convergence.

Table 2
Problem A. Helmholtz equation with $k = -120$

n ($n$sub)	64 (2)	128 (4)	128 (8)	256 (8)
$\delta = 0$	29	41	50	68
$\delta = 1$	21	28	33	49
$\delta = 2$	18	23	26	39

Number of iterations for AS/GMRES convergence.

Table 3
Problem B. Advection–diffusion equation

n ($n$sub)	64 (2)	128 (4)	128 (8)	256 (8)
$\delta = 0$	35	51	52	73
$\delta = 1$	24	35	38	52
$\delta = 2$	20	28	32	42

Number of iterations for AS/GMRES convergence.

8. Conclusion

We make the case, both theoretically and experimentally, that the two-level additive Schwarz preconditioning is asymptotically optimal when combined with a minimal residual iterative method such as GMRES. The key here is that the methods are optimal when the minimal residual iterative method uses the same energy norm as that used to derive the optimal Schwarz bounds, and the asymptotic optimality of the usual method (minimizing the Euclidean norm) is obtained as a consequence. We also developed an equivalence between left and right preconditioned methods.

Acknowledgements

We would like to thank Valeria Simoncini for her interest in our work and for her useful comments. We also thank the referees for their comments, which helped improved and focus our presentation.

References

[1] Mario Arioli, Daniel Loghin, Andrew J. Wathen, Stopping criteria in finite element problems, Numer. Math. 99 (2005) 381–410.
[2] Steven F. Ashby, Thomas A. Manteuffel, Paul E. Saylor, A taxonomy for conjugate gradient methods, SIAM J. Numer. Anal. 27 (1990) 1542–1568.
[3] Owe Axelsson, Iterative Solution Methods, Cambridge University Press, Cambridge and New York, 1994.
[4] Satish Balay, Kris Buschelman, Victor Eijkhout, William D. Gropp, Dinesh Kaushik, Matthew G. Knepley, Lois Curfman McInnes, Barry F. Smith, Hong Zhang. PETSc Users Manual, Technical Report ANL-95/11 – Revision 2.1.5, Argonne National Laboratory, 2004. Available from: <http://www.mcs.anl.gov/petsc>.
[5] Richard Barrett, Michael W. Berry, Tony F. Chan, James Demmel, June Donato, Jack Dongarra, Victor Eijkhout, Roldán Pozo, Charles Romine, Henk van der Vorst, Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods, SIAM, Philadelphia, 1993.
[6] Michele Benzi, Andreas Frommer, Reinhard Nabben, Daniel B. Szyld, Algebraic theory of multiplicative Schwarz methods, Numer. Math. 89 (2001) 605–639.
[7] Dietrich Braess, Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics, second ed., Cambridge University Press, Cambridge and New York, 2001.
[8] Susanne C. Brenner, L. Ridgway Scott, The mathematical theory of finite element methods, Springer Series Texts in Applied Mathematics, vol. 15, Springer, New York, 1994.
[9] Xiao-Chuan Cai, Maksymilian Dryja, Marcus Sarkis, Restricted additive Schwarz preconditioner with harmonic overlap for symmet-

ric positive definite linear systems, SIAM J. Numer. Anal. 41 (2003) 1209–1231.

[10] Xiao-Chuan Cai, Olof B. Widlund, Domain decomposition algorithms for indefinite elliptic problems, SIAM J. Sci. Stat. Comput. 13 (1992) 243–258.

[11] Xiao-Chuan Cai, B. Olof, Widlund. Multiplicative Schwarz algorithms for some non-symmetric and indefinite problems, SIAM J. Numer. Anal. 30 (1993) 936–952.

[12] X.-C. Cai, J. Zou, Some observations on the l^2 convergence of the additive Schwarz preconditioned GMRES method, Numer. Linear Algebra Appl. 9 (2002) 379–397.

[13] Michael Eiermann, Oliver G. Ernst, Geometric aspects in the theory of Krylov subspace methods, Acta Numer. 10 (2001) 251–312.

[14] Stanley C. Eisenstat, Howard C. Elman, Martin H. Schultz, Variational iterative methods for non-symmetric systems of linear equations, SIAM J. Numer. Anal. 20 (1983) 345–357.

[15] Howard C. Elman, Iterative Methods for Large, Sparse, Nonsymmetric Systems of Linear Equations, Ph.D. Thesis, Department of Computer Science, Yale Univeristy, Research Report #229, April 1982.

[16] Azeddine Essai, Weighted FOM and GMRES for solving nonsymmetric linear systems, Numer. Algorithms 18 (1998) 277–292.

[17] Andreas Frommer, Daniel B. Szyld, Weighted max norms, splittings, and overlapping additive Schwarz iterations, Numer. Math. 83 (1999) 259–278.

[18] Anne Greenbaum, Iterative methods for solving linear systems, Frontiers in Applied Mathematics, vol. 17, SIAM, Philadelphia, 1997.

[19] Martin H. Gutknecht, Changing the norm in conjugate gradient type algorithms, SIAM J. Numer. Anal. 30 (1993) 40–56.

[20] Martin H. Gutknecht, Miroslav Rozložník, A framework for generalized conjugate gradient methods–with special emphasis on contributions by Rüdiger Weiss, Appl. Numer. Math. 41 (2002) 7–22.

[21] Wolfgang Hackbusch, Iterative solution of large sparse systems of equations, Springer Series in Applied Mathematical Sciences, vol. 95, Springer, New York, Berlin, Heidelberg, 1994.

[22] Alfio Quarteroni, Alberto Valli, Domain Decomposition Methods for Partial Differential Equations, Oxford Science Publications, Clarendon Press, Oxford, 1999.

[23] Yousef Saad, Iterative methods for sparse linear systems, The PWS Publishing Company, Boston, 1996, second ed., SIAM, Philadelphia, 2003.

[24] Yousef Saad, Martin H. Schultz, GMRES: a generalized minimal residual algorithm for solving non-symmetric linear systems, SIAM J. Sci. Stat. Comput. 7 (1986) 856–869.

[25] Marcus Sarkis, Partition of unity coarse space and Schwarz methods with harmonic overlap, in: Luca F. Pavarino, Andrea Toselli (Eds.), Recent Developments in Domain Decomposition Methods, Lecture Notes in Computational Science and Engineering, 23, Springer, Heidelberg, New York, 2002, pp. 75–92.

[26] Marcus Sarkis, A coarse space for elasticity: partition of unity rigid body motions coarse space, in: Zlatko Drmač, Vjeran Hari, Luka Sopta, Zvonimir Tutek, Krešimir Veselić (Eds.), Applied Mathematics and Scientific Computing, Kluwer Academic/Plenum, New York, Boston, Dordrecht, 2003, pp. 253–265.

[27] Marcus Sarkis, Partition of unity coarse spaces: enhanced versions, discontinuous coefficients and applications to elasticity, in: Ismael Herrera et al. (Eds.), Domain Decomposition Methods in Science and Engineering, Publised by the National Autonomous University of Mexico (UNAM), ISBN 970-32-0859-2, Proceedings of the 14th International Conference on Domain Decomposition Methods in Cocoyoc, México, 2003, pp. 149–158.

[28] Barry F. Smith, Peter E. Bjørstad, William. D. Gropp, Domain Decomposition: Parallel Multilevel Methods for Elliptic Partial Differential Equations, Cambridge University Press, Cambridge, 1996.

[29] Gerhard Starke, Field-of-values analysis of preconditioned iterative methods for nonsymmetric elliptic problems, Numer. Math. 78 (1997) 103–117.

[30] Andrea Toselli, Olof B. Widlund, Domain decomposition: Algorithms and theory, Springer Series in Computational Mathematics, 34, Springer, Berlin, Heidelberg, New York, 2005.

[31] Rüdiger Weiss, Error-minimizing Krylov subspace methods, SIAM J. Sci. Comput. 15 (1994) 511–527.

[32] Rüdiger Weiss, Minimization properties and short recurrences for Krylov subspace method, Electron. Trans. Numer. Anal. 4 (1994) 57–75.

[33] Rüdiger Weiss, Parameter-free iterative linear solvers, Mathematical Research Series, vol. 97, Akademie, Berlin, 1996.

[34] Jinchao Xu, Xiao-Chuan Cai, A preconditioned GMRES method for non-symmetric or indefinite problems, Math. Comput. 59 (1992) 311–319.