Using approximate inverses in algebraic multilevel methods

Y. Notay [1]

Service de Métrologie Nucléaire
Université Libre de Bruxelles (C.P. 165)
50, Av. F.D. Roosevelt, B-1050 Brussels, Belgium.
email : ynotay@ulb.ac.be

Report GANMN 97-01

March 1997

[1]Supported by the "Fonds National de la Recherche Scientifique", Chercheur qualifié.

Abstract

This paper deals with the iterative solution of large sparse symmetric positive definite systems. We investigate preconditioning techniques of the two-level type that are based on a block factorization of the system matrix. Whereas the basic scheme assumes an exact inversion of the submatrix related to the first block of unknowns, we analyze the effect of using an approximate inverse instead. We derive condition number estimates that are valid for any type of approximation of the Schur complement and that do *not* assume the use of the hierarchical basis. They show that the two-level methods are stable when using approximate inverses based on modified ILU techniques, or explicit inverses that meet some row-sum criterion. On the other hand, we bring to the light that the use of standard approximate inverses based on convergent splittings can have a dramatic effect on the convergence rate. These conclusions are numerically illustrated on a model example.

keywords: iterative methods for linear systems, acceleration of convergence preconditioning.

AMS CLASSIFICATION : 65F10, 65B99, 65N20.

1 Introduction

To solve large sparse symmetric positive definite systems

$$A\mathbf{u} = \mathbf{b} \tag{1.1}$$

arising from discrete elliptic PDEs, many recent research focus on the design of efficient algebraic multilevel preconditioners. Like standard multigrid [18, 29], these are based on a two-level method recursively used in a V or W cycle algorithm. However, the basic scheme originates here from a block factorization of the system matrix

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} A_{11} & \\ A_{21} & S_A \end{pmatrix} \begin{pmatrix} I & A_{11}^{-1} A_{12} \\ & I \end{pmatrix} , \tag{1.2}$$

where the first block of unknowns corresponds to the fine grid nodes, the second block of unknowns to the coarse grid nodes, and where $S_A = A_{22} - A_{21} A_{11}^{-1} A_{12}$ is the Schur complement of A; since the latter is a dense matrix, some relevant sparse approximation S has to be supplied and the preconditioner writes

$$\widetilde{B} = \begin{pmatrix} A_{11} & \\ A_{21} & S \end{pmatrix} \begin{pmatrix} I & A_{11}^{-1} A_{12} \\ & I \end{pmatrix} . \tag{1.3}$$

Whether considered as a "stand alone" method or as an intermediate step in a full multilevel cycle, the potentialities of this two-level scheme essentially depend on the spectral condition number

$$\kappa(\widetilde{B}^{-1}A) \;=\; \frac{\lambda_{\max}(\widetilde{B}^{-1}A)}{\lambda_{\min}(\widetilde{B}^{-1}A)} \tag{1.4}$$

which should be close to 1 and bounded independently of the grid size [1].

A straightforward computation shows that

$$\widetilde{B}^{-1}A = \begin{pmatrix} I & * \\ & S^{-1} S_A \end{pmatrix} . \tag{1.5}$$

and an essential step in the analysis consists therefore in proving $\mathcal{O}(1)$ upper and lower bounds on the spectrum of $S^{-1} S_A$. Such results exist for the methods that originate from the hierarchical basis multigrid method [5, 7, 8, 9, 10, 11, 25, 26, 27, 30], but also for some other approaches where S is computed according an incomplete Gaussian elimination process [22, 28].

Now, the preconditioner (1.3) involves solving two systems with A_{11} each time it is needed to solve a system with $\widetilde{B}$. Some works pay little attention to that point. The key argument is that A_{11} has condition number $\mathcal{O}(1)$ [1, 5, 10], so that any reasonable iterative scheme converges quickly in a number of steps independent of the grid size. In practice however, the cost of such inner iterations becomes rapidly prohibitive. Most works take this into account and consider preconditioners of the form

$$B = \begin{pmatrix} P & \\ A_{21} & S \end{pmatrix} \begin{pmatrix} I & P^{-1} A_{12} \\ & I \end{pmatrix} , \tag{1.6}$$

[1] throughout this paper, $\lambda_{\max}(C)$ and $\lambda_{\min}(C)$ denote respectively the largest and the smallest eigenvalue of C

where P^{-1} stands for the used approximate inverse of A_{11}, possibly implicitly defined by a very few steps of some iterative procedure. The argument about the conditioning of A_{11} serves then only to show that it is easy to define cost effective approximate inverses such that the spectrum of $P^{-1}A_{11}$ is very well clustered around 1.

Preconditioners of the form (1.6) also arise in another family of algebraic multilevel methods [1, 3, 4, 6, 23], whose primary step consists in approximating A_{11}^{-1} by some sparse matrix B_{11}. The basic scheme assumes then that the Schur complement $A_{22} - A_{21} B_{11} A_{12}$ of the so approximated matrix is computed exactly, but some recent developments [6, 13] introduce versions for which B_{11} is less sparse, leading to a block incomplete factorization (1.6) in which both A_{11} and the Schur complement are approximated.

Now, although numerous works analyze two-level methods of the type (1.6), so far available theoretical results involve the so-called strengthened CBS constant [9, 10, 17] and can therefore be useful only when A is a finite element matrix computed using the *hierarchical* finite element basis functions.

In fact, as is well known [25], the Schur complement S_A is the same for both nodal and hierarchical representations. Hence, as easily seen from (1.5), these analyzes apply to the version (1.3) with exact inversion of A_{11} whatever the used basis.

This is however not true anymore in the general case. In [8, 25, 26], it is shown that the equivalence of the preconditioner (1.6) in the usual nodal basis with its counterpart in the hierarchical basis can be restored by adding to A_{12} and A_{21} a term equal to $(A_{11} - P)$ times some interpolation matrix. As $(A_{11} - P)$ is small when a sufficiently accurate approximate inverse is used, the question of the necessity of this additional term is raised [8]. The purpose of the present paper is to develop a finer analysis of this issue.

In this view, we investigate the eigenvalue distribution associated to preconditioners (1.6) in function of the spectral properties of $P^{-1}A_{11}$ and $S^{-1}S_A$, using only assumptions compatible with A defined from a standard finite difference or finite element scheme.

If A is a (non strictly) diagonally dominant M-matrix, our analysis proves in particular the stability of the two-level method when P is a modified (possibly blockwise) ILU factorization of A_{11} (e.g. [2, 14, 15, 16, 19]), or P^{-1} an explicit approximate inverse of A_{11} satisfying some row-sum criterion [2, Chapter 8].

On the other hand, we show that the use of standard approximate inverses based on convergent splittings can have a much more dramatic effect than expected, leading to a condition number that may grow up to $\mathcal{O}(h^{-4})$ when the mesh size parameter h decreases.

The paper is organized as follows: in Section 2, we develop a simple analysis of the behaviour of the condition number when P^{-1} is a standard approximate inverse of A_{11}; our main theoretical results are proved in Section 3, and their application to approximate inverses satisfying a row-sum criterion is discussed in Section 4. Section 5 is devoted to numerical results.

2 Standard approximate inverses

In this section, we derive a *lower* bound on the condition number by evaluating the Rayleigh quotients

$$\overline{r} = \frac{\overline{\mathbf{v}}^t A \overline{\mathbf{v}}}{\overline{\mathbf{v}}^t B \overline{\mathbf{v}}} \qquad , \qquad \underline{r} = \frac{\underline{\mathbf{v}}^t A \underline{\mathbf{v}}}{\underline{\mathbf{v}}^t B \underline{\mathbf{v}}}$$

associated to vectors of the form

$$\overline{\mathbf{v}} = \begin{pmatrix} -P^{-1} A_{12}\,\mathbf{v}_2 \\ \mathbf{v}_2 \end{pmatrix} \qquad , \qquad \underline{\mathbf{v}} = \begin{pmatrix} -A_{11}^{-1} A_{12}\,\mathbf{v}_2 \\ \mathbf{v}_2 \end{pmatrix} .$$

A straightforward computation shows that

$$\begin{aligned}
\overline{\mathbf{v}}^t A\, \overline{\mathbf{v}} &= \mathbf{v}_2^t S_A\, \mathbf{v}_2 + \mathbf{v}_2^t A_{21} \left(A_{11}^{-1} - 2P^{-1} + P^{-1} A_{11} P^{-1} \right) A_{12}\, \mathbf{v}_2 \\
\overline{\mathbf{v}}^t B\, \overline{\mathbf{v}} &= \mathbf{v}_2^t S\, \mathbf{v}_2 \ , \\
\underline{\mathbf{v}}^t A\, \underline{\mathbf{v}} &= \mathbf{v}_2^t S_A\, \mathbf{v}_2 \ , \\
\underline{\mathbf{v}}^t B\, \underline{\mathbf{v}} &= \mathbf{v}_2^t S\, \mathbf{v}_2 + \mathbf{v}_2^t\ A_{21} \left(P^{-1} - 2A_{11}^{-1} + A_{11}^{-1}\, P\, A_{11}^{-1} \right) A_{12}\, \mathbf{v}_2 \ .
\end{aligned}$$

Further, $A_{11}^{-1} - 2P^{-1} + P^{-1} A_{11} P^{-1} = \left(I - P^{-1} A_{11}\right)^2 A_{11}^{-1}$, whereas, letting $\lambda_M = \lambda_{\max}\left(P^{-1} A_{11}\right)$, one has, for all $\mathbf{w}_1$

$$\begin{aligned}
\mathbf{w}_1^t \left(P^{-1} - 2A_{11}^{-1} + A_{11}^{-1}\, P\, A_{11}^{-1} \right) \mathbf{w}_1 &= \mathbf{w}_1^t \ \left(A_{11}^{-1}\, P \right) \left(P^{-1} A_{11} P^{-1} - 2P^{-1} + A_{11}^{-1} \right)\ \mathbf{w}_1 \\
&\geq \lambda_M^{-1}\ \ \mathbf{w}_1^t\ \left(I - P^{-1} A_{11} \right)^2 A_{11}^{-1}\ \mathbf{w}_1 \ .
\end{aligned}$$

Hence, with

$$\begin{aligned}
q_{\mathbf{v}_2} &= \frac{\mathbf{v}_2^t\, S_A\, \mathbf{v}_2}{\mathbf{v}_2^t\, S\, \mathbf{v}_2} \ , \\
g_{\mathbf{v}_2} &= \frac{\mathbf{v}_2^t\, A_{21} \left(I - P^{-1} A_{11} \right)^2 A_{11}^{-1} A_{12}\, \mathbf{v}_2}{\mathbf{v}_2^t\, S_A\, \mathbf{v}_2} \ ,
\end{aligned}$$

one obtains

$$\overline{r} = q_{\mathbf{v}_2} \left(1 + g_{\mathbf{v}_2} \right) \ , \qquad \underline{r} \leq \frac{1}{q_{\mathbf{v}_2}^{-1} + \lambda_M^{-1}\, g_{\mathbf{v}_2}} \ .$$

Since $\lambda_{max}\left(B^{-1} A\right) \geq \overline{r}$, $\underline{r} \geq \lambda_{\min}\left(B^{-1} A\right)$, this implies

$$\kappa \left(B^{-1} A \right) \geq \left(1 + g_{\mathbf{v}_2} \right) \left(1 + \frac{q_{\mathbf{v}_2}}{\lambda_M}\, g_{\mathbf{v}_2} \right) \ , \tag{2.1}$$

showing that the two-level method can be stable with respect the use of the approximate inverse only if $g_{\mathbf{v}_2}$ is bounded above independently of the mesh size.

Now, as is well known, in discrete PDE applications, S_A is spectrally equivalent to the coarse grid discretization matrix. Therefore, $\mathbf{v}_2^t\, S_A\, \mathbf{v}_2 = \mathcal{O}(h^2)$ for smooth vectors, which further implies $\mathbf{v}_2^t\, A_{21}\, A_{11}^{-1}\, A_{12}\, \mathbf{v}_2 \simeq \mathbf{v}_2^t\, A_{22}\, \mathbf{v}_2 = \mathcal{O}(1)$ when one does *not* make use of the hierarchical basis. Hence, $g_{\mathbf{v}_2}$ can be bounded independently of the mesh size only if P^{-1} acts nearly as an exact inverse for these *smooth* vectors, whereas usual convergent splittings are on the contrary known to converge fast for *rough* vectors [18, 29].

To illustrate this, let A be more specifically the matrix associated with the standard five point finite difference approximation of the Laplacian on the unit square with a uniform mesh size h in both directions, where h is such that h^{-1} is an even integer.

Assume further for simplicity that P^{-1} is the approximate inverse resulting from $m \geq 1$ stationary iterations with the Jacobi preconditioning, i.e.

$$\left(I - P^{-1} A_{11} \right) = \left(M^{-1} N \right)^m$$

where $M = diag(A) = 4\,I$ and $N = M - A = -\mathit{offdiag}(A)$.

The vector $\mathbf{v}_2^{(\nu\mu)}$ that interpolates on the coarse grid nodes the function sin $\nu\pi xh$ sin $\mu\pi yh$ is shown in [21] to be an eigenvector of S_A for $\nu = 1, \ldots, \frac{h^{-1}}{2} - 1$, $\mu = 1, \ldots, \frac{h^{-1}}{2} - 1$. The corresponding eigenvalue writes

$$\lambda_{\nu\mu} \;=\; 4\left(\frac{1}{4-2(c_\nu + c_\mu)} + \frac{1}{4-2(c_\nu - c_\mu)} + \frac{1}{4+2(c_\nu - c_\mu)} + \frac{1}{4+2(c_\nu + c_\mu)}\right)^{-1}$$

where $c_\nu = \cos\,\nu\pi h$, $c_\mu = \cos\,\mu\pi h$. For $\nu = \mu$, it takes the simpler expression

$$\lambda_{\nu\nu} = \frac{8(1-c_\nu^2)}{2-c_\nu^2} \quad ,$$

and it is also easy to check that

$$\left(M^{-1}N\right)^2 \left(A_{12}\,\mathbf{v}_2^{(\nu\nu)}\right) = \frac{c_\nu^2}{2}\left(A_{12}\,\mathbf{v}_2^{(\nu\nu)}\right) \; .$$

Hence, since $A_{22} = 4I$ and $(M^{-1}N)^t = M^{-1}N$,

$$\begin{aligned} g_{\mathbf{v}_2^{(\nu\nu)}} &= \left(\frac{c_\nu^2}{2}\right)^m \frac{\mathbf{v}_2^{(\nu\nu)\,t}\,(A_{22} - S_A)\;\mathbf{v}_2^{(\nu\nu)}}{\mathbf{v}_2^{(\nu\nu)\,t}\;S_A\,\mathbf{v}_2^{(\nu\nu)}} \\ &= \left(\frac{c_\nu^2}{2}\right)^m \frac{c_\nu^2}{2(1-c_\nu^2)}\;, \end{aligned}$$

that is, for the smoothest node $\nu = 1$

$$g_{\mathbf{v}_2(11)}^{(11)} \;\simeq\; \left(\frac{1}{2}\right)^{m+1}\;\frac{h^{-2}}{\pi^2}$$

showing with (2.1) that $\kappa(B^{-1}A) \geq \mathcal{O}(h^{-4})$ except if one increases m as the mesh is refined.

3 Eigenvalue bounds

In this section, we prove upper and lower bounds on the spectrum of $B^{-1}A$. The only restrictive assumptions are

$$\mathbf{v}_1^t P \mathbf{v}_1 \;\leq\; \mathbf{v}_1^t A_{11}\mathbf{v}_1 \qquad\qquad \forall \mathbf{v}_1 \tag{3.1}$$

and

$$\mathbf{v}_2^t\,A_{21}P^{-1}\,A_{12}\,\mathbf{v}_2 \;\leq\; (1-\xi)\;\mathbf{v}_2^t\,A_{22}\,\mathbf{v}_2 \;+\; \xi\;\mathbf{v}_2^t\,A_{21}\,A_{11}^{-1}\,A_{12}\,\mathbf{v}_2 \qquad \forall \mathbf{v}_2 \tag{3.2}$$

for some $\xi < 1$. This is relatively weak, since ξ may be negative, but nevertheless sufficient to prevent the problems illustrated in the preceding section. Indeed, whenever $\mathbf{v}_2^t\,A_{22}\,\mathbf{v}_2 = (1+\mathcal{O}(h^2))\,\mathbf{v}_2^t\,A_{21}A_{11}^{-1}A_{12}\,\mathbf{v}_2$, (3.1), (3.2) imply

$$\mathbf{v}_2^t\,A_{21}\,A_{11}^{-1}\,A_{12}\,\mathbf{v}_2 \;\leq\; \mathbf{v}_2^t\,A_{21}\,P^{-1}\,A_{12}\mathbf{v}_2 \;\leq\; \left(1 + (1-\xi)\,\mathcal{O}(h^2)\right)\;\mathbf{v}_2^t\,A_{21}\,A_{11}^{-1}\,A_{12}\,\mathbf{v}_2 \quad ,$$

i.e. P^{-1} has to act nearly as an exact inverse for the corresponding modes.

How to compute a lower bound on ξ for some classes of preconditioners that meet (3.1) is addressed in the next section.

Theorem 3.1 *Let*

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \quad , \quad B = \begin{pmatrix} P & \\ A_{21} & S \end{pmatrix} \begin{pmatrix} I & P^{-1}A_{12} \\ & I \end{pmatrix} \tag{3.3}$$

be symmetric nonnegative definite matrices such that A_{11} and P are invertible.

Assume that

$$\beta\, \mathbf{v}_1^t A_{11} \mathbf{v}_1 \;\leq\; \mathbf{v}_1^t P \mathbf{v}_1 \;\leq\; \mathbf{v}_1^t A_{11} \mathbf{v}_1 \qquad \forall \mathbf{v}_1 \tag{3.4}$$

for some β such that $0 < \beta < 1$ and that

$$\eta\, \mathbf{v}_2^t \left(A_{22} - A_{21} A_{11}^{-1} A_{12}\right) \mathbf{v}_2 \;\leq\; \mathbf{v}_2^t S \mathbf{v}_2 \;\leq\; \zeta\, \mathbf{v}_2^t \left(A_{22} - A_{21} A_{11}^{-1} A_{12}\right) \mathbf{v}_2 \quad \forall \mathbf{v}_2 \tag{3.5}$$

for some η, ζ such that $0 \leq \eta \leq 1 \leq \zeta$.

If

$$\mathbf{v}_2^t A_{21} P^{-1} A_{12} \mathbf{v}_2^t \;\leq\; (1-\xi)\, \mathbf{v}_2^t A_{22} \mathbf{v}_2 \;+\; \xi\, \mathbf{v}_2^t A_{21} A_{11}^{-1} A_{12} \mathbf{v}_2 \qquad \forall \mathbf{v}_2 \tag{3.6}$$

for some $\xi < 1$, then

$$\mathbf{v}^t B \mathbf{v} \;\geq\; \gamma\, \mathbf{v}^t A \mathbf{v} \qquad \forall \mathbf{v} \tag{3.7}$$

where γ is the smallest root of

$$\gamma^2 - \gamma(\eta + 1 - \xi + \beta\xi) + \beta\,\eta$$

and

$$\alpha\, \mathbf{v}^t A \mathbf{v} \;\geq \mathbf{v}^t B \mathbf{v} \qquad \forall \mathbf{v} \tag{3.8}$$

where α is the largest root of

$$\alpha^2 - \alpha(\zeta + 1 - \xi + \beta\xi) + \beta\,\zeta \ .$$

Moreover,

$$\gamma \;\geq\; \frac{\eta\,\beta}{\eta + 1 - \xi(1-\beta)} \tag{3.9}$$

and

$$\alpha \;\leq\; \zeta\left(1 + \tfrac{(1-\beta)(1-\xi)}{\zeta-\beta}\right) \tag{3.10}$$

with, in addition,

$$\gamma \;\geq\; \begin{cases} \eta\,\beta & \text{if } \xi \geq \eta \\ \eta\left(\beta - \sqrt{(1-\beta)(1-\xi)}\right) & \text{if } 0 \leq \xi \leq \eta \ . \end{cases} \tag{3.11}$$

Proof Define the family of second degree polynomials

$$P_a(t) = t^2 - t(a + 1 - \xi + \xi\beta) + a\beta \ ,$$

where $a > 0$. Since $P_a(a) = -a(1-\beta)(1-\xi) \;<\; 0$ and $P_a(\beta) = -\beta(1-\xi)(1-\beta) < 0$, $P_a(t)$ admits two positive roots $t_1 < min(\beta, a) \leq max(\beta, a) < t_2$, which shows that γ, α exist, are positive and satisfy

$$\gamma < \min(\eta, \beta) < 1 \leq \zeta < \alpha \tag{3.12}$$

Further, it is easily checked that

$$\gamma = \eta - \frac{\gamma(1-\beta)(1-\xi)}{\beta-\gamma} \ , \ \alpha = \zeta + \frac{\alpha(1-\beta)(1-\xi)}{\alpha-\beta} \quad . \tag{3.13}$$

We now prove (3.7).

$$B - \gamma A = \begin{pmatrix} P - \gamma\, A_{11} & (1-\gamma)\, A_{12} \\ (1-\gamma)\, A_{21} & S + A_{21}\, P^{-1}\, A_{12} - \gamma\, A_{22} \end{pmatrix}$$

and, since $\gamma < \beta$, (3.4) implies that $P - \gamma\, A_{11}$ is positive definite. Hence $B - \gamma\, A$ is nonnegative definite if so is its Schur complement

$$S_{B-\gamma A} = S - \gamma\, A_{22} + A_{21}\, \left(P^{-1} - (1-\gamma)^2 (P - \gamma\, A_{11})^{-1}\right)\, A_{12} \ .$$

Now, $\mathbf{v}_2^t\, S\, \mathbf{v}_2 \geq \eta\, \mathbf{v}_2^t \left(A_{22} - A_{21}\, A_{11}^{-1}\, A_{12}\right) \mathbf{v}_2$ while, since $\eta - \gamma > 0$, (3.6), (3.13) imply

$$(\eta-\gamma)\, \mathbf{v}_2^t\, A_{22}\, \mathbf{v}_2 = \tfrac{\gamma(1-\beta)(1-\xi)}{\beta-\gamma}\ \mathbf{v}_2^t\, A_{22}\, \mathbf{v}_2 \ \geq \ \tfrac{\gamma(1-\beta)}{\beta-\gamma}\ \mathbf{v}_2^t\, A_{21}\, \left(P^{-1} - \xi\, A_{11}^{-1}\right)\, A_{12}\, \mathbf{v}_2 \ .$$

Hence,

$$\begin{aligned} \mathbf{v}_2^t\, S_{B-\gamma A}\, \mathbf{v}_2 \ \geq \ & \mathbf{v}_2^t\, A_{21} \Big\{ \left(1 + \tfrac{\gamma(1-\beta)}{\beta-\gamma}\right) P^{-1} \\ & \quad - \left(\eta + \tfrac{\xi\,\gamma(1-\beta)}{\beta-\gamma}\right) A_{11}^{-1} - (1-\gamma)^2 (P - \gamma\, A_{11})^{-1} \Big\}\, A_{12}\, \mathbf{v}_2 \quad , \end{aligned}$$

and, using (3.13) to eliminate η,

$$\begin{aligned} \mathbf{v}_2^t\, S_{B-\gamma A}\, \mathbf{v}_2 \ &\geq \ \mathbf{v}_2^t\, A_{21} \left\{ \tfrac{\beta(1-\gamma)}{\beta-\gamma} P^{-1} - \tfrac{\gamma(1-\gamma)}{\beta-\gamma}\, A_{11}^{-1} - (1-\gamma)^2 (P - \gamma\, A_{11})^{-1} \right\} A_{12}\, \mathbf{v}_2 \\ &= \ \tfrac{1-\gamma}{\beta-\gamma}\, \mathbf{v}_2^t\, A_{21}\, P^{-\frac{1}{2}} \left\{ \beta\, I - \gamma X^{-1} - {\scriptstyle (1-\gamma)(\beta-\gamma)} (I - \gamma\, X)^{-1} \right\} P^{-\frac{1}{2}}\, A_{12}\, \mathbf{v}_2 \, , \end{aligned}$$

where $X = P^{-\frac{1}{2}}\, A_{11}\, P^{-\frac{1}{2}}$. Since X is symmetric positive definite,

$$f(X) = \beta\, I - \gamma\, X^{-1} - (1-\gamma)(\beta-\gamma)(I - \gamma\, X)^{-1}$$

will be non negative definite if $f(\lambda) \geq 0$ for all $\lambda \in \sigma(X) \subset [1, \beta^{-1}]$; (3.7) follows then because

$$\begin{aligned} f(\lambda) \ &= \ (1-\gamma\,\lambda)^{-1} \left((1-\gamma\,\lambda)\,(\beta-\gamma\,\lambda^{-1}) - (1-\gamma)(\beta-\gamma)\right) \\ &= \ (1-\gamma\,\lambda)^{-1}\ (1-\lambda^{-1})\,(1-\beta\,\lambda) \ . \end{aligned}$$

The proof of (3.8) is similar.

$$\alpha A - B = \begin{pmatrix} \alpha\, A_{11} - P & (\alpha-1)\, A_{12} \\ (\alpha-1)\, A_{21} & \alpha\, A_{22} - S - A_{21}\, P^{-1}\, A_{12} \end{pmatrix}$$

and, since $\alpha > 1$, (3.4) implies that $\alpha\, A_{11} - P$ is positive definite. Hence, $\alpha\, A - B$ is nonnegative definite if so is its Schur complement

$$S_{\alpha A - B} = \alpha\, A_{22} - S - A_{21} \left(P^{-1} + (\alpha-1)^2 (\alpha\, A_{11} - P)^{-1}\right)\, A_{12} \ .$$

We may use $\mathbf{v}_2^t\,S\ \leq \zeta\,\mathbf{v}_2^t\left(A_{22}-A_{21}\,A_{11}^{-1}A_{12}\right)\mathbf{v}_2$, while, since $(\alpha-\zeta)>0$, (3.6), (3.13) imply

$$(\alpha-\zeta)\mathbf{v}_2^t\,A_{22}\,\mathbf{v}_2=\tfrac{\alpha(1-\beta)(1-\xi)}{\alpha-\beta}\,\mathbf{v}_2^t\,A_{22}\,\mathbf{v}_2\geq\tfrac{\alpha(1-\beta)}{\alpha-\beta}\,\mathbf{v}_2^t\,A_{21}\left(P^{-1}-\zeta A_{11}^{-1}\right)A_{12}\,\mathbf{v}_2\ ,$$

whence

$$\begin{aligned}\mathbf{v}_2^t\,S_{\alpha A-B}\,\mathbf{v}_2\ &\geq\ \mathbf{v}_2^t\,A_{21}\left\{\left(\tfrac{\alpha(1-\beta)}{\alpha-\beta}-1\right)P^{-1}+\left(\zeta-\tfrac{\xi\alpha(1-\beta)}{\alpha-\beta}\right)A_{11}^{-1}\right.\\&\qquad\qquad\left.-(\alpha-1)^2(\alpha A_{11}-P)^{-1}\right\}A_{12}\,\mathbf{v}_2\\&=\ \tfrac{\alpha-1}{\alpha-\beta}\,\mathbf{v}_2^t\,A_{21}\,P^{-\frac{1}{2}}\left\{-\beta I+\alpha X^{-1}-{\scriptstyle(\alpha-1)(\alpha-\beta)}(\alpha X-I)^{-1}\right\}P^{-\frac{1}{2}}A_{12}\mathbf{v}_2\ .\end{aligned}$$

Letting

$$g(X)=-\beta I+\alpha X^{-1}-(\alpha-1)(\alpha-\beta)(\alpha X-I)^{-1}\ ,$$

one has

$$\begin{aligned}g(\lambda)\ &=\ (\alpha\,\lambda-1)^{-1}\left((\alpha\,\lambda-1)\,(-\beta+\alpha\,\lambda^{-1})-(\alpha-1)(\alpha-\beta)\right)\\&=\ \alpha(\alpha\,\lambda-1)^{-1}\ (1-\lambda^{-1})\,(1-\beta\,\lambda)\ .\end{aligned}$$

which is nonnegative for all $\lambda\in[1,\beta^{-1}]$, whence (3.8).

Finally, (3.10) holds by virtue of (3.13), together with $\alpha>\zeta$ and the fact that $\frac{x}{x-\beta}$ is a decreasing function of x for $x>\beta$; (3.9) is deduced from $\gamma\,(\eta+1-\xi(1-\beta))-\beta\,\eta=\gamma^2\geq 0$; (3.11) is proved by checking that

$$\begin{aligned}&P_\eta(\eta\,\beta)\ =\ \eta\,\beta(\xi-\eta)\,(1-\beta)\ ,\\&P_\eta\left(\eta\left(\beta-\sqrt{(1-\beta)(1-\xi)}\right)\right)=\eta(1-\beta)\,(\xi(1-\eta)+(1-\beta)(\eta-\xi))\\&\qquad\qquad+\eta\sqrt{(1-\beta)(1-\xi)}\,(1-\eta\,\beta+(\eta-\xi)(1-\beta))\end{aligned}$$

and are both nonnegative when respectively $\xi\geq\eta$ and $\eta\geq\xi\geq 0$.

∎

4 Approximate inverses satisfying a row-sum criterion

In this section, we examine how to check our main assumptions (3.1), (3.2). Of course (3.1) may hold for any preconditioner with an appropriate scaling. However, since this requires in general an eigenvalue estimation, we find more practical to focus on the classes of approximate inverses for which this condition holds by construction.

About (3.2), note that $A_{22}(1-\xi)-A_{21}\left(P^{-1}-\xi A_{11}^{-1}\right)A_{12}$ is the Schur complement of

$$\begin{pmatrix} P & A_2 \\ A_{21} & A_{22}-\xi\,S_A\end{pmatrix}\ ,$$

i.e. (3.2) holds if and only if the latter matrix is nonnegative definite. Further, to prove this at once for a wide class of preconditioners it suffice to check that

$$\begin{pmatrix} \Delta & A_{12} \\ A_{21} & A_{22}-\xi\,S_A\end{pmatrix}\tag{4.1}$$

is nonnegative definite for some matrix Δ such that

$$\mathbf{v}_1^t\, P\, \mathbf{v}_1 \;\geq\; \mathbf{v}_1^t\, \Delta\, \mathbf{v}_1 \qquad\qquad \forall \mathbf{v}_1\,. \tag{4.2}$$

For instance, if A is an M-matrix, there always exists a positive vector $\mathbf{x} = (\mathbf{x}_1\, \mathbf{x}_2)^t$ such that $A\mathbf{x}$ is nonnegative (e.g. $\mathbf{x} = (1\,1\ldots 1)^t$ when A has nonnegative row-sum) [12]. It is then interesting to consider the diagonal matrix Δ such that $\Delta\mathbf{x}_1 = A_{11}\,\mathbf{x}_1$. Indeed,

$$\widetilde{A} \;=\; \begin{pmatrix} \Delta & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

has then nonpositive offdiagonal entries and nonnegative generalized row-sum since $\widetilde{A}\mathbf{x} = A\mathbf{x}$. It is thus a symmetric M-matrix, therefore nonnegative definite [12], i.e. $\xi = 0$ at the very least. At the end of this section, we show that in fact $\xi \geq \frac{1}{2}$ when A arises from the linear finite element discretization of a 2D second order elliptic PDE.

4.1 MILU factorizations

Assume A_{11} is an M-matrix and let $\mathbf{x}_1$ be a positive vector such that $A_{11}\,\mathbf{x}_1$ is nonnegative.

Then, if P results from a possibly blockwise MILU factorization of A_{11} (e.g. [2, 14, 15, 16, 19]) computed to satisfy the generalized row-sum criterion

$$P\,\mathbf{x}_1 = A_{11}\,\mathbf{x}_1$$

it is well known that $A - P$ is a symmetric M-matrix, whence (3.1). On the other hand, P writes

$$P = (Q - F^t)\, Q^{-1}\, (Q - F)$$

where F is strictly upper triangular and nonnegative and Q a symmetric positive definite M-matrix (which reduces to a diagonal matrix in case of a pointwise factorization). For such matrices, the following theorem shows that $P - \Delta$ is nonnegative definite.

Theorem 4.1 *Let $P = (Q - F^t)\, Q^{-1}\, (Q - F)$ be such that Q is a symmetric positive definite symmetric M-matrix and F is nonnegative and strictly upper triangular.*

Assume that $P\,\mathbf{x}_1$ is nonnegative for some positive vector $\mathbf{x}_1$.

Then,

$$\mathbf{v}_1^t\, P\, \mathbf{v}_1 \;\geq\; \mathbf{v}_1^t\, \Delta\, \mathbf{v}_1 \qquad\qquad \forall \mathbf{v}_1$$

where Δ is the diagonal matrix such that $P\,\mathbf{x}_1 = \Delta\,\mathbf{x}_1$.

Proof First, since Q is an M-matrix, Q^{-1} is nonnegative [12]. From $(I - Q^{-1}F)\,\mathbf{x}_1 = Q^{-1}\,(\Delta\mathbf{x}_1 + F^t\,(I - Q^{-1}F)\,\mathbf{x}_1)$, it is then easily seen by induction that $(I - Q^{-1}F)\,\mathbf{x}_1$ is nonnegative. Let T be the diagonal matrix such that $T\mathbf{x}_1 = Q^{-1}F\mathbf{x}_1$. One has

$$P - \Delta = \left(T - F^t\, Q^{-1}\right)\, Q\, \left(T - Q^{-1}\, F\right) + \left\{Q - T\, Q\, T - (I - T)\, F - F^t(I - T) - \Delta\right\}\,.$$

The first term of the r.h.s. is nonnegative definite and has zero generalized row-sum by construction. On the other hand, the term under brackets, which has therefore also zero generalized row-sum, has nonpositive offdiagonal entries since T is diagonal with diagonal entries less or equal to 1 . Therefore, the term under brackets is a symmetric M-matrix, hence nonnegative definite.

■

4.2 "Compensated" explicit inverses

Assume A_{11} is an M-matrix, and let $\mathbf{x}_1$ be a positive vector such that $A_{11}\,\mathbf{x}_1$ is *positive.*

Since A_{11}^{-1} is nonnegative [12], it is not restrictive to assume from an explicit approximate inverses P^{-1} that it has nonnegative offdiagonal entries not larger than the corresponding ones in A_{11}^{-1}. Assume further that the neglected entries have been "compensated" on the diagonal in such a way that $P^{-1}(A_{11}\,\mathbf{x}_1) = A_{11}^{-1}(A_{11}\,\mathbf{x}_1) = \mathbf{x}_1$; (3.1) follows then from [2, Theorem 8.4]. On the other hand, since $(\Delta - \Delta\, P^{-1}\,\Delta)\,\mathbf{x}_1 = 0$, it is easily seen that $\Delta - \Delta\, P^{-1}\,\Delta$ is a symmetric M-matrix, which implies (4.2) since both Δ and P are positive definite.

4.3 Polynomial acceleration

Here we consider approximate inverses of A_{11} implicitly defined by a very few steps of some preconditioned iterative methods.

This means that

$$P^{-1} = \mathcal{P}_m\left(\tilde{P}^{-1}\,A_{11}\right)\,A_{11}^{-1} \quad , \tag{4.3}$$

where $\tilde{P}$ stands for the used preconditioner and where $\mathcal{P}_m$ is a polynomial such that $\mathcal{P}_m(0) = 0$.

We assume that $\tilde{P}$ satisfies (3.1) an (4.2), for instance that $\tilde{P}$ is defined as in subsection 4.2 or 4.3 above. Then, (3.1) and (4.2) will a fortiori hold for P defined by (4.3) if

$$1 \le \mathcal{P}_m(t) \le t \qquad\qquad \forall t \in \left[1, b^{-1}\right] \ ,$$

where $b = \lambda_{\max}^{-1}\,(P^{-1}\,A_{11})$.

As is well known [24], the best polynomial acceleration is obtained with shifted Chebyshev polynomials of the first kind, i.e.,

$$\mathcal{P}_m(t) = c\left(1 - \frac{T_m\left(\frac{b^{-1}+1-2t}{b^{-1}-1}\right)}{T_m\left(\frac{b^{-1}+1}{b^{-1}-1}\right)}\right)$$

where c is some scaling factor and where $T_m(x)$, $m = 1, 2, \ldots$ obeys the recurrence relation

$$T_0 = 1, \quad T_1 = x, \quad T_{m+1}(x) = 2x\,T_m(x) - T_{m-1}(x)\,, \quad m = 1,\ 2, \ldots$$

Since $|T_m(x)| \le 1$ for $|x| \le 1$, we set

$$c = \left(1 - T_m^{-1}\left(\frac{b^{-1}+1}{b^{-1}-1}\right)\right)^{-1} \quad ,$$

so that

$$\mathcal{P}_m(t) = \frac{T_m\left(\frac{b^{-1}+1}{b^{-1}-1}\right) - T_m\left(\frac{b^{-1}+1-2t}{b^{-1}-1}\right)}{T_m\left(\frac{b^{-1}+1}{b^{-1}-1}\right) - 1}$$

satisfies $\mathcal{P}(t) \ge 1$ for all $t \in [1, b^{-1}]$.

We have no general proof that $t \ge \mathcal{P}_m(t)\ \forall t \in [1, b^{-1}]$. Nevertheless, for $m = 2$, one finds

$$\mathcal{P}_2(t) = (b+1)\,t - b\,t^2 \quad ,$$

whence $t - P_2(t) = b\,t\,(t-1) \geq 0$ for $t \geq 1$. For $m=3$, an explicit check is harder, but one may use the following reasoning: since $T_3(-1) = -1$ and $T_3\left(\frac{b^{-1}+1}{b^{-1}-1}\right) > \left(\frac{b^{-1}+1}{b^{-1}-1}\right)$, one has $\mathcal{P}_3(b^{-1}) < b^{-1}$; therefore $t - \mathcal{P}_3(t)$ is a third degree polynomial that is zero for $t=0$ and $t=1$, positive for $t = b^{-1}$, and negative for $t \to \infty$ (the leading coefficient is < 0); hence, $t - P_3(t)$ cannot have additional root in $]1, b^{-1}[$, i.e. is uniformly positive in that interval.

On the other hand, one will seldomly consider in practice polynomials with degree larger than 3. Indeed, letting

$$\beta^{-1} = \max_{t \in]1, b^{-1}[} \mathcal{P}(t) \quad ,$$

one readily obtains, since $|T_k(x)| \leq 1$ for $|x| \leq 1$,

$$1 - \beta = \frac{2}{T_m\left(\frac{1+b}{1-b}\right) + 1} \quad ,$$

and the resulting value is generally sufficiently small for $m = 2$ or 3. For instance, with $b = \frac{1}{2}$, (which is already pessimistic since A_{11} is well conditioned), one gets $1 - \beta = \frac{1}{9}$ for $m = 2$ and $1 - \beta = \frac{1}{50}$ for $m = 3$.

4.4 Linear finite elements

Here we prove a better lower bound on ξ in the case A arises from the linear finite element discretization of a two dimensional second order elliptic PDE

$$-\nabla\left(\bar{a}(x,y)\,\overline{\nabla}\,u\right) = f \tag{4.4}$$

on a arbitrary "coarse" mesh to which on has added one level of uniform refinement.

Let A_c be the finite element matrix associated to the discretization of (4.4) on this coarse mesh. As is well known [25], $\mathbf{v}_2^t\,A_c\,\mathbf{v}_2 \geq \mathbf{v}_2^t\,S_A\,\mathbf{v}_2$ so that (4.1) will be a fortiori nonnegative definite for some $\xi > 0$ if

$$\mathbf{v}^t \begin{pmatrix} \Delta & A_{12} \\ A_{21} & A_{22} - \xi A_c \end{pmatrix} \mathbf{v} \geq 0 \qquad\qquad \forall \mathbf{v} \tag{4.5}$$

Now, the latter matrix corresponds to the assembly of elementary matrices associated to each coarse triangle. Therefore, (4.5) may be checked considering only these elementary matrices. If we assume further that $\bar{a}(x,y)$ is piecewise constant on the coarse triangulation, one readily obtains that each of these elementary matrices writes (the vertex opposite to mid edge node i is node $i+3, i = 1, \ldots, 3$)

$$\left(\begin{array}{ccc|ccc} 2c_1 & 0 & 0 & 0 & -c_1 & -c_1 \\ 0 & 2c_2 & 0 & -c_2 & 0 & -c_2 \\ 0 & 0 & 2c_3 & -c_3 & -c_3 & 0 \\ \hline 0 & -c_2 & -c_3 & (c_2+c_3)(1-\xi) & \xi c_3 & \xi c_2 \\ -c_1 & 0 & -c_3 & \xi c_3 & (c_1+c_3)(1-\xi) & \xi c_1 \\ -c_1 & -c_2 & 0 & \xi c_2 & \xi c_1 & (c_1+c_2)(1-\xi) \end{array}\right)$$

$$= c_1\,T_1 + c_2 T_2 + c_3\,T_3 \ ,$$

where c_1, c_2, c_3 are positive provided there are no interior angle larger than $\frac{\pi}{2}$. On the other hand, deleting the zero lines and columns in T_1, T_2, T_3, one gets the same 3×3 matrix

$$\tilde{T} = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 1-\xi & \xi \\ -1 & \xi & 1-\xi \end{pmatrix},$$

which is readily found nonnegative definite for $\xi = \frac{1}{2}$.

5 Numerical experiment

We report here the results of a numerical experiment that illustrates our theoretical investigations. In this experiment, A is the matrix resulting from the five point finite difference discretization of the Laplacian on the unit square with uniform mesh size h in both directions, where h is such that h^{-1} is an even integer.

We consider two-level preconditioners of the form (1.6), where the second block is formed with the nodes present in the grid of mesh size $2h$, and where S is taken equal to the matrix associated with the discretization of the Laplacian on this coarse grid.

For P, we consider preconditioners based on a simple pointwise incomplete factorization of A_{11} without fill-in and with a natural ordering of the concerned nodes. Two choices are included: a MILU factorization for which P and A_{11} have same row-sum, and an ILU factorization for which the error matrix $A_{11} - P$ has zero diagonal.

The results are given in Table 1. Recalling that $\kappa(S^{-1} S_A) = 2$ [20], they illustrate the stability of the two-grid method when using MILU preconditioning for A_{11}. On the other hand, the results obtained with ILU preconditioning are quite disastrous, although P is an excellent preconditioner of A_{11} alone.

For the sake of completeness, we have also performed some tests with scaled ILU preconditioners, that is $P = c\,P_{\text{ILU}}$ where: (1) $c = \lambda_{\max}\left(P_{\text{ILU}}^{-1} A_{11}\right)$, as one would have use

MILU preconditioning of A_{11}

	$P^{-1}A_{11}$			$B^{-1}A$		
h^{-1}	λ_{min}	$\lambda_{\max}$	κ	λ_{min}	$\lambda_{\max}$	κ
16	1.00	1.20	1.20	0.51	1.25	2.45
32	1.00	1.21	1.21	0.50	1.27	2.54
64	1.00	1.21	1.21	0.50	1.29	2.58
128	1.00	1.21	1.21	0.50	1.29	2.58

ILU preconditioning of A_{11}

	$P^{-1}A_{11}$			$B^{-1}A$		
h^{-1}	λ_{min}	$\lambda_{\max}$	κ	λ_{min}	$\lambda_{\max}$	κ
16	0.88	1.09	1.24	0.510	1.42	2.78
32	0.88	1.09	1.24	0.380	2.28	6.00
64	0.87	1.09	1.25	0.176	4.97	28.30
128	0.87	1.09	1.25	0.058	15.00	258.00

Table 1: results for MILU and ILU preconditioning of A_{11}

$\kappa\,(B^{-1}A)$

h^{-1}	$c=1.09$	$c=0.87$	$c=0.84$	$c=0.90$
16	4.07	2.56	2.69	2.095
32	15.8	2.63	2.90	2.55
64	120.	2.65	3.66	3.02
128	1475	2.66	9.03	5.20

Table 2: results for scaled ILU preconditioning of A_{11}

on the basis of some theoretical results for hierarchical finite elements [8, 25, 26, 27], which require $\mathbf{v}_1^t\,P\,\mathbf{v}_1 \ge \mathbf{v}_1^t\,A_{11}\,\mathbf{v}_1 \quad \forall \mathbf{v}$; (2) $c=\lambda_{\min}\left(P_{\mathrm{ILU}}^{-1}A_{11}\right)$, to satisfy on the contrary $\mathbf{v}_1^t\,P\,\mathbf{v}_1 \le \mathbf{v}_1^t\,A_{11}\,\mathbf{v}_1 \quad \forall \mathbf{v}$; (3) $c=\lambda_{\min}\left(P_{\mathrm{ILU}}^{-1}A_{11}\right)+0.03$ and $c=\lambda_{\min}\left(P_{\mathrm{ILU}}^{-1}A_{11}\right)-0.03$, which may be seen as attempts to realize the above choice, but with a rough eigenvalue estimation. The results are given in Table 2, and clearly show that $c=\lambda_{\min}\left(P_{\mathrm{ILU}}^{-1}A_{11}\right)$ is the only viable choice.

Finally, returning to the case where P is a MILU factorization of A_{11}, it is interesting to compare the theoretical bounds of Theorem 3.1 to the extremal eigenvalues of $B^{-1}A$ given in Table 1. Since, in this application, $\eta=1, \zeta=2$ [20] and $\xi=\frac{1}{2}$ (see subsection 4.4), (3.7) and (3.8) give respectively

$$\begin{aligned}\lambda_{\max}\,(B^{-1}A) &\le 1.52\ ,\\ \lambda_{\min}\,(B^{-1}A) &\ge 0.47\ ,\end{aligned}$$

which is relatively accurate. One may then use the theoretical bounds to estimate the gain obtained by using a more accurate approximate inverse of A_{11}. Assume for instance that, according to the developments in subsection 4.3, one uses

$$P^{-1} \;=\; (1+b)\,\tilde{P}^{-1} \;-\; b\,\tilde{P}^{-1}\,A_{11}\,\tilde{P}^{-1}$$

where $\tilde{P}$ stands for the simple MILU factorization of A_{11} considered so far, and where b^{-1} is a known upper bound on $\lambda_{\max}\left(\tilde{P}^{-1}\,A_{11}\right)$.

Then,

$$\beta \ge \frac{4\,b}{(b+1)^2}\ ,$$

so that, with $b^{-1}=1.21$, (3.7) , (3.8) give now

$$\begin{aligned}\lambda_{\max}\,(B^{-1}A) &\le 1.075\ ,\\ \lambda_{\min}\,(B^{-1}A) &\ge 0.498\ ,\end{aligned}$$

which is not much different from the extremal eigenvalues of the two-grid method with exact inversion of A_{11}. This even holds with an inaccurate eigenvalue estimation of $\tilde{P}^{-1}\,A_{11}$. For instance, with $b^{-1}=\frac{3}{2}$, one gets $\beta=\frac{24}{25}$ and therefore

$$\begin{aligned}\lambda_{\max}\,(B^{-1}A) &\le 1.179\ ,\\ \lambda_{\min}\,(B^{-1}A) &\ge 0.497\ .\end{aligned}$$

References

[1] O. Axelsson, *On multigrid methods of the two-level type*, in Multigrid Methods, W. Hackbusch and U. Trottenberg, eds., Lectures Notes in Mathematics No. 960, Berlin Heidelberg New York, 1981, Springer-Verlag, pp. 352–367.

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

[3] O. Axelsson and V. Eijkhout, *Analysis of recursive 5-point/9-point factorization method*, in Preconditioned Conjugate Gradient Methods, O. Axelsson and L. Kolotilina, eds., Lectures Notes in Mathematics No. 1457, Springer-Verlag, 1990, pp. 154–173.

[4] O. Axelsson and V. Eijkhout, *The nested recursive two level factorization for nine-point difference matrices*, SIAM J. Sci. Stat. Comput., 12 (1991), pp. 1373–1400.

[5] O. Axelsson and I. Gustafsson, *Preconditioning and two-level multigrid methods of arbitrary degree of approximation*, Math. Comp., 40 (1983), pp. 214–242.

[6] O. Axelsson and M. Neytcheva, *Algebraic multilevel iterations for Stieltjes matrices*, Num. Lin. Alg. Appl., 1 (1994), pp. 213–236.

[7] O. Axelsson and P. S. Vassilevski, *Algebraic multilevel preconditioning methods, I.*, Numer. Math., 56 (1989), pp. 157–177.

[8] O. Axelsson and P. S. Vassilevski, *Algebraic multilevel preconditioning methods, II.*, SIAM J. Numer. Anal., 27 (1990), pp. 1569–1590.

[9] R. E. Bank, *Hierarchical bases and the finite element method*, Acta Numerica, 5 (1996), pp. 1–43.

[10] R. E. Bank and T. F. Dupont, *Analysis of a two-level scheme for solving finite element equations*, Tech. Rep. CNA-159, Center for Numerical Analysis, The University of Texas at Austin, Texas, USA, 1980.

[11] R. E. Bank, T. F. Dupont, and H. Yserentant, *The hierarchical basis multigrid method*, Numer. Math., 52 (1988), pp. 427–458.

[12] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Academic Press, New York, 1979.

[13] E. Botta and A. van der Ploeg, *Preconditioning techniques for matrices with arbitrary sparsity patterns.* Preprint, Department of Mathematics, University of Groningen, The Netherlands, 1995.

[14] T. C. Chan and H. A. van der Vorst, *Approximate and incomplete factorizations.* Preprint 871, Department of Mathematics, University of Utrecht,The Netherlands, 1994.

[15] P. Concus, G. H. Golub, and G. Meurant, *Block preconditioning for the conjugate gradient method*, SIAM J. Sci. Statist. Comput., 6 (1985), pp. 220–252.

[16] T. Dupont, R. P. Kendall, and H. H. Rachford, *An approximate factorization procedure for solving self-adjoint elliptic difference equations*, SIAM J. Numer. Anal., 5 (1968), pp. 559–573.

[17] V. Eijkhout and P. Vassilevski, *The role of the strengthened c.b.s. inequality in multilevel methods*, SIAM Review, 33 (1991), pp. 405–419.

[18] W. Hackbusch, *Multi-grid Methods and Applications*, Springer, Berlin, 1985.

[19] M. M. Magolu, *Ordering strategies for modified block incomplete factorization*, SIAM J. Sci. Comput., 16 (1995), pp. 378–399.

[20] J. Maitre and F. Musy, *The contraction number of a class of two-level methods; an exact evaluation for some finite element subspaces and model problems*, in Multigrid Methods, W. Hackbusch and U. Trottenberg, eds., Lectures Notes in Mathematics No. 960, Berlin Heidelberg New York, 1981, Springer-Verlag, pp. 535–544.

[21] A. Reusken, *Fourier analysis of a robust multigrid method for convection-diffusion equations*, Numer. Math., 71 (1995), pp. 365–398.

[22] A. Reusken, *A multigrid method based on incomplete Gaussian elimination*, J. Num. Lin. Alg. with Appl., 3 (1996), pp. 369–390.

[23] A. van der Ploeg, E. Botta, and F. Wubs, *Nested grids ILU-decomposition (NGILU)*, J. Comput. Appl. Math., 66 (1996), pp. 515–526.

[24] R. S. Varga, *Matrix iterative analysis*, Prentice Hall, Englewood Cliffs, 1962.

[25] P. Vassilevski, *Nearly optimal iterative methods for solving finite element elliptic equations based on the multilevel splitting of the matrix*, Tech. Rep. # 1989-09, Institute for Scientific Computation, University of Wyoming, Laramie, USA, 1989.

[26] P. Vassilevski, *Hybrid V-cycle algebraic multilevel preconditioners*, Math. Comp., 58 (1992), pp. 489–512.

[27] P. Vassilevski, *On two ways of stabilizing the hierarchical basis multilevel method*, SIAM Review, 39 (1997), pp. 1–34.

[28] C. Wagner, W. Kinzelbach, and G. Wittum, *Schur-complement multigrid*, Numer. Math., 75 (1997), pp. 523–545.

[29] P. Wesseling, *An Introduction to Multigrid Methods*, J. Wiley and Sons, Chichester, 1992.

[30] H. Yserentant, *On the multi-level splitting of finite element spaces*, Numer. Math., 49 (1986), pp. 379–412.