ITERATIVE METHODS IN LINEAR ALGEBRA
R. Beauwens and P. de Groen (Editors)
Elsevier Science Publishers B.V. (North-Holland)

UPPER EIGENVALUE BOUNDS AND RELATED MODIFIED INCOMPLETE FACTORIZATION STRATEGIES

Y. NOTAY*

Abstract. Upper eigenvalue bounds are obtained for the preconditioning by modified incomplete factorizations. To each bound is associated a factorization algorithm or "strategy", which consists in computing the approximate factors so that an a priori chosen upper bound is satisfied. The efficiency of the method, which clearly depends then on the behaviour of the associated smallest eigenvalue, is analyzed through complementary results.

Key words. Iterative methods for linear systems, acceleration of convergence, preconditioning

1. Introduction. We consider the iterative solution of large, sparse linear systems

$$Au = b \tag{1.1}$$

whose system matrix is positive definite, or, including the case of consistent singular systems (A singular with $b \in \mathcal{R}(A)$), positive semidefinite. In this context, the use of the preconditioned conjugate gradient method is of particular interest due to its optimal convergence properties. It requires the choice of a positive definite [1] preconditioning matrix B and the convergence rate is then classically estimated by the following formula

$$k_\varepsilon \le \frac{1}{2}\sqrt{\kappa}\, ln\frac{2}{\varepsilon} \tag{1.2}$$

where k_ε denotes the maximal number of iterations necessary to reduce the relative error in the A-(semi)norm by a factor ε, while the (generalized) spectral condition number κ is given by

$$\kappa = \frac{\nu_{\max}(B^{-1}A)}{\nu_{\min}(B^{-1}A)} \tag{1.3}$$

where

$$\begin{aligned} \nu_{\max}(B^{-1}A) &= \max\{\, \nu \mid \nu \in \sigma(B^{-1}A)\,\}\,, \\ \nu_{\min}(B^{-1}A) &= \min\{\, \nu \mid \nu \in \sigma(B^{-1}A),\ \nu > 0\}\,. \end{aligned} \tag{1.4}$$

One may obtain sharper estimates than (1.2) when the eigenvalues of $B^{-1}A$ are not regularly distributed between $\nu_{\max}$ and $\nu_{\min}$. In particular, attractive results are found when some upper eigenvalues are isolated [2, 19, 7], but it is important to note in this respect that these bounds are actually not valid in presence of rounding errors [15], and therefore that, as experimentally observed in [27], one too large eigenvalue is sufficient for slowing down the convergence.

---

* Chargé de Recherche FNRS, Université Libre de Bruxelles, Service de Métrologie Nucléaire (CP 165), 50 av. F.D. Roosevelt, B-1050 Brussels, Belgium. Author's research is supported by the "Fonds National de la Recherche Scientifique" (Belgium).

[1] For ease of presentation only, we do not consider the extension that allows, in the singular case, the preconditioner itself to be singular [23]. Singular preconditionings of singular systems are widely discussed in [20, 24, 25].

The main quality one may ask for a preconditioner is thus of reducing the condition number, although interesting convergence rates may still be obtained when a few lower eigenvalues are isolated [19, 7], bounds whose reliability is proven in [26].

The modified incomplete factorization methods are precisely known to be efficient for reducing the condition number. In these methods, the diagonal factor P of the approximate factorization $B = (P - F^t)P^{-1}(P - F)$ (where F is strictly upper triangular) of A is computed from the relation $Bx = Ax + \Lambda Dx$, where x is a positive vector such that $Ax \geq 0$, $D = diag(A)$ and $\Lambda = (\lambda_i \delta_{ij})$ is a nonnegative matrix of small parameters which we call the perturbation term. Following Gustafsson [17], the version with $\Lambda = 0$ (the "unperturbed strategy") is often satisfying. However, its results are sometimes disappointing, due to an unacceptably large value of $\nu_{\max}(B^{-1}A)$. This explains the development of perturbed strategies, with perturbations λ_i that are sufficiently large to significantly reduce $\nu_{\max}(B^{-1}A)$, and sufficiently small to keep $\nu_{\min}(B^{-1}A)$ close to 1 ($\Lambda = 0$ implies that $\nu_{\min}(B^{-1}A) = 1$). Perturbed factorization algorithms have been proposed in [5, 10, 17]. All these methods are based on the same philosophy, that is the perturbations are chosen in such a way that a given upper eigenvalue bound, obtained from Theorem 3.1 of [12], or from some (earlier) less refined version of this result, is satisfied.

New upper spectral bounds have been obtained in [6] and in [22] for generalized SSOR preconditioners, and it is possible as suggested by Axelsson [4] to design alternate factorization strategies on the basis of these new results. The first objective addressed in the present contribution is the extension of these new algorithms to higher order approximate factorizations. We further discuss the lowest eigenvalue behaviour. For each strategy, we obtain conditions under which perturbations are not required, implying that $\nu_{\min}(B^{-1}A) = 1$. When these conditions are only approximately satisfied, one may expect that the perturbations will be sufficiently small to keep $\nu_{\min}(B^{-1}A)$ close to 1, and, summarizing more involved lower eigenvalue analyses, we finally state how precisely the restriction that "these conditions are approximately satisfied" has to be understood. These conclusions are illustrated through numerical examples.

General terminology and notation. All vectors belong to C^n; all matrices are $n \times n$ real matrices. The symbols A^t, $\mathcal{R}(A)$, $\sigma(A)$ and $\rho(A)$ denote, respectively, the transpose, the range, the spectrum and the spectral radius of the matrix A. The order relation between real matrices and vectors is the usual componentwise order : if $A = (a_{ij})$ and $B = (b_{ij})$ then $A \leq B(A < B)$ if $a_{ij} \leq b_{ij}(a_{ij} < b_{ij})$ for all i, j; A is called nonnegative (positive) if $A \geq 0(A > 0)$. If $A = (a_{ij})$, we denote by $diag(A)$ the (diagonal) matrix whose entries are $a_{ii}\delta_{ij}$ and we let $offdiag(A) = A - diag(A)$.

2. Modified incomplete factorizations of a Stieltjes matrix.

As usual when dealing with approximate factorization methods, one has to introduce the assumption that the system matrix is a Stieltjes one, i.e. that it is nonnegative definite and has nonpositive offdiagonal entries. More general cases are included through the reduction to the Stieltjes case by means of spectral equivalence [5], or by the technique introduced in [12], or by a combination of both these approaches [26].

An important property [13] is then that the existence of a positive vector x such that Ax is nonnegative is then ascertained, which allows us to compute the upper triangular matrix U with $P = diag(U)$ according to the following algorithm, where the β_{ij} which are nonzero are equal to 1 while the parameters λ_i are nonnegative :

for $i = 1, \ldots, n$ *set* :

$$u_{ij} = a_{ij} - \beta_{ij} \sum_{k<i} \frac{u_{ki}u_{kj}}{p_{kk}}, \quad j = i+1, \ldots, n \tag{2.1}$$

$$p_{ii}^{(0)} = a_{ii} - x_i^{-1} \left(\sum_{k<i} \sum_{j>i} (1-\beta_{ij}) \frac{u_{ki}u_{kj}}{p_{kk}} x_j + \sum_{k<i} \sum_{j<i} (1-\beta_{ji}) \frac{u_{ki}u_{kj}}{p_{kk}} x_j \right) - \sum_{k<i} \frac{u_{ki}^2}{p_{kk}} \tag{2.2}$$

$$p_{ii} = p_{ii}^{(0)} + \lambda_i a_{ii} \tag{2.3}$$

We say that U is the upper triangular factor and $P = diag(U)$ the diagonal factor of the modified incomplete factorization

$$B = U^t P^+ U \tag{2.4}$$

of A associated with $x, \Lambda = (\lambda_i \delta_{ij})$ and $\beta = (\beta_{ij})$, where P^+ is the diagonal matrix satisfying

$$(P^+)_{ii} = \begin{cases} p_{ii}^{-1} & \text{if } p_{ii} \neq 0 \\ 0 & \text{if } p_{ii} = 0 \end{cases} \tag{2.5}$$

Noting $F = offdiag(U)$ and $D = diag(A)$, the basic properties of this algorithm may be summarized as follows (see [25] or [26] for a proof) :

(1) The algorithm cannot fail, U is an M-matrix and B nonnegative definite
(2) U is regular *iff* P is regular *iff* B is regular
(3) $offdiag(U + U^t) \leq offdiag(A)$ and $(P - F)x \geq Ax \geq 0$
(4) $Bx = Ax + \Lambda Dx \geq Ax \geq 0$
(5) if B is regular, then $\Lambda = 0 \Rightarrow \nu_{\min}(B^{-1}A) = 1$

It follows from these properties that B will be an admissible preconditioner if one can ensure its regularity, while on the other hand solving a linear system involving B will be cheap provided that one has chosen β sufficiently sparse.

The essential discussion is thus that of the spectral properties of $B^{-1}A$, which are actually closely related to the values used for the parameters or "perturbations" λ_i. The remaining part of this paper precisely deals with the development and the analysis of what we call the factorization strategies, which are nothing but the methods or the policies according to which these perturbations are chosen.

The simplest choice consists in setting the perturbation term Λ equal to zero. This "unperturbed" strategy, referred to as the Strategy no 1 in [10, 24, 25], has the advantage to guaranty a lower eigenvalue bound equal to 1. Its behaviour depends thus on that of the associated largest eigenvalue. Unfortunately, this behaviour is often not satisfying, see for instance the examples given in [27, 14, 26].

This gives rise to the development of factorization strategies based on the opposite philosophy, that is on the a priori control of the largest eigenvalue $\nu_{\max}(B^{-1}A)$ in place on that of $\nu_{\min}(B^{-1}A)$. Of course, the development of such "perturbed" strategies requires the statement of appropriate upper eigenvalue theorems. These points will be addressed in section 4. Before, we shall introduce some graph notions, which is the price to pay for a purely algebraic presentation of the theory.

3. Graph notation. To each (triangular or symmetric) matrix A we associate an ordered undirected graph with node set $[1,n]$ and such that a pair of node (i,j) belongs to the edge set $iff\ \ i \neq j$ and $a_{ij} \neq 0$ or $a_{ji} \neq 0$. We refer to [18] for general terminology about graphs, but we shall also need the following more specific graph notions.

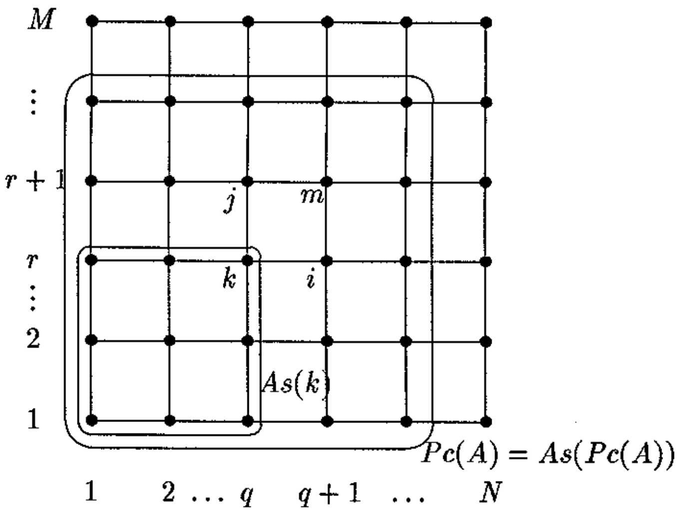
A node k belongs to the set of *precursors* $P(i)$ of a node i if (i,k) belongs to the edge set with $k < i$. The set of common precursors of a pair of nodes i,j is given by $Pc(i,j) = P(i) \cap P(j)$, and we further define the set of common precursors of the matrix A by $Pc(A) = \cup_{i \neq j} Pc(i,j)$

An *increasing path* is a path $i_o, i_1, i_2, \ldots, i_\ell$ such that $i_o < i_1 < i_2 < \ldots < i_\ell$.

For any node i, we define the *ascent* $As(i)$ of i as the set of nodes k such that there exists an increasing path from k to i (note that $i \in As(i)$ because a path of zero length is an increasing path). We further define the ascent of a nonempty subset M of $[1,n]$ by $As(M) = \cup_{i \in M} As(i)$.

The *maximal increasing length* $\ell(A)$ of the matrix A is the maximal length of an increasing path in its graph, and the maximal increasing length $\ell(M)$ of a nonempty subset M of $[1,n]$ is the maximal length of an increasing path whose nodes all belong to M.

These definitions are illustrated on Figure 1 for the graph of the matrix obtained by the 5-point finite difference approximation of a second order elliptic PDE on a rectangular region and when using the lexicographic ordering.



$$\{i,j\} = P(m) \quad and \quad \{k\} = Pc(i,j)$$
$$\ell(A) = \ell(Pc(A)) + 2 = \ell(As(Pc(A))) + 2 = N + M - 2$$
$$\ell(As(k)) = q + r - 2 \ \ (= \mathcal{O}(\ell) \ \ for\ most\ of\ nodes)$$

FIG. 1. *Illustration of graph notions for a lexicographically ordered graph*

4. Perturbed factorization strategies. For each strategy, we first give the upper eigenvalue *theorem* which gives rise to it. In these theorems, we make the following general assumptions which will be referred to as (GA) :

A is a symmetric nonnegative definite matrix, $U = P - F$ *with* $P = diag(U)$ *is an upper triangular M-matrix such that* $offdiag(U + U^t) \leq offdiag(A)$, $B = U^t P^+ U$, x *is a* *positive vector such that* $Px \geq Fx$ *and* $Bx \geq Ax \geq 0$.

One easily verifies that these assumptions are all satisfied within the framework of modified incomplete factorizations of Stieltjes matrices as defined above. Note for completeness that some of them are even sometimes superfluous, see [26] for details.

We then give the factorization *algorithm* which actually reduces to a practical rule for achieving (2.3), a *corollary* which displays the basic properties of the algorithm as they straightforwardly follow from the application of the previously given theorem, and finally a *property* which may be viewed as a first lower eigenvalue analysis result. The property associated with Strategy no 2 is a simple reformulation and extension to the singular case of a result first proved in [8], while the properties associated to both other strategies are original results.

We assume sometimes for convenience the irreducibility of A, more general cases being included through the reduction of A to its canonical block diagonal form.

Strategy no 2. The following upper bound theorem is the most widely used for the analysis of modified incomplete factorizations. Its algebraic formulation is due to Beauwens end Quenon [11]. The present version, somewhat more refined and which includes the singular case, is already proved in [24].

THEOREM 1. *Assume (GA), A irreducible and* $Pc(U) \neq \emptyset$. *If, in addition,*

$$\tau\ (Px)_i \geq (Fx)_i \quad \text{for all } i \in Pc(U) \tag{4.1}$$

for some $0 < \tau < 1$, *then* B *is regular with*

$$\nu_{\max}(B^{-1}A) \leq \frac{1}{1-\tau} \tag{4.2}$$

ALGORITHM 1. *A being a Stieltjes matrix, x a positive vector such that* $Ax \geq 0$ *and the* β_{ij}, $1 \leq i < j \leq n$ *some given (0,1) values, compute the upper triangular matrix* $U = P - F$ *with* $P = diag(U)$ *according to (2.1),(2.2) and*

$$p_{ii} = \begin{cases} \max(p_{ii}^{(0)}, (Fx)_i/(\tau x)_i) & \text{if } \ i \in Pc(U) \\ p_{ii}^{(0)} & \text{if } \ i \notin Pc(U) \end{cases} \tag{4.3}$$

where τ, $0 < \tau < 1$, *is a parameter.*

COROLLARY 1. *Assume A irreducible. Either* $Pc(U) = \emptyset$ *which implies* $B = A$, *or B is regular with* $\nu_{\max}(B^{-1}A) \leq 1/(1-\tau)$.

PROPERTY 1. *Assume A irreducible and* $Pc(U) \neq \emptyset$. *If, with* $E = F^t$,

$$(k + \ell_i + 1)^{-1}((F-E)x)_i \leq (Ax)_i \quad \text{for all } i \in As(Pc(U)), \tag{4.4}$$

where k is some nonnegative constant and ℓ_i *the maximal increasing length of* $As(i)$, *and if*

$$\tau \geq 1 - \frac{1}{k + \ell + 2} \tag{4.5}$$

where ℓ *is the maximal increasing length of* $As(Pc(U))$, *then*

$$p_{ii} = p_{ii}^{(0)} \quad \text{for all } i \qquad \text{and} \qquad \sigma(B^{-1}A) \subset [1, k + \ell + 2]$$

Proof. It suffices to prove that, for the Strategy no 1 ($p_{ii} = p_{ii}^{(0)}$ for all i), (4.4) implies (4.1) with $\tau = 1 - (k + \ell + 2)^{-1}$. Now, since $Bx = Ax$ for this strategy, the latter result straightforwardly follows from Lemma 1 of the Appendix applied with $\Delta x = Bx$. □

The last result suggests the following policy for an appropriate use of this strategy: *(1) choose an ordering such that conditions (4.4) are as far as possible satisfied, while keeping an interesting value for the parameter ℓ; (2) use*

$$\tau \approx 1 - \ell^{-1}, \tag{4.6}$$

with ℓ defined as the maximal increasing length of the matrix A, so that the parameter τ does not vary with the level of fill-in allowed.

This strategy has first been proposed in [10]. The idea of controlling the largest eigenvalue through Theorem 1 or some earlier version by the use of appropriately chosen perturbations goes back to the works by Axelsson and Gustafsson [1, 16] and a *dynamic* version, which avoids the a priori evaluation of the perturbations λ_i was already proposed in [5]. Finally, the last refinement consists in eq.(4.6), which provides an easy choice of the parameter τ.

Strategy no 3. The same philosophy gives rise to alternate strategies when used with alternate upper spectral bounds as recently suggested by Axelsson [4] for the preconditioning by (perturbed) generalized SSOR. This third strategy will now be further extended to higher order factorizations on the basis of the following generalization of Lemma 4.5 of [6].

THEOREM 2. *Assume (GA) with A irreducible and let $E = F^t$. If, in addition,*

$$(2 - 1/\lambda)\ Px \geq (A + F + E)x \tag{4.7}$$

for some $\lambda > \frac{1}{2}$, then B is regular with

$$\nu_{\max}(B^{-1}A) \leq \lambda \tag{4.8}$$

Proof. Let $G = A + E + F$. Note that $diag(G) = diag(A)$ while $offdiag(G) \geq 0$. Thus, (4.7) implies that P is regular, which allows us to write

$$B - \lambda^{-1}A = ((1 - 1/\lambda)P - E)P^{-1}((1 - 1/\lambda)P - F) + \lambda^{-1}((2 - 1/\lambda)P - G)$$

Since $H = (2 - 1/\lambda)P - G$ is a Stieltjes matrix (by virtue of $Hx \geq 0$ with $offdiag(H) \leq 0$), this equation proves that $B - \lambda^{-1}A$ is nonnegative definite, whence (4.8). □

ALGORITHM 2. *A being a Stieltjes matrix, x a positive vector such that $Ax \geq 0$ and the β_{ij}, $1 \leq i < j \leq n$ some given (0,1) values, compute the upper triangular matrix $U = P - F$ with $P = diag(U)$ according to (2.1),(2.2) and, noting $E = F^t$,*

$$p_{ii} = \max(p_{ii}^{(0)}, (2 - 1/\lambda)^{-1}((A + F + E)x)_i/x_i) \tag{4.9}$$

where λ, $\lambda > \frac{1}{2}$, is a parameter.

COROLLARY 2. *If A is irreducible, then B is regular with $\nu_{\max}(B^{-1}A) \leq \lambda$.*

PROPERTY 2. *If A is irreducible and $E = F^t$, then*

$$\begin{cases} (k + \ell_i + 1)^{-1}((F - E)x)_i \leq (Ax)_i & \textit{for all } i \\ (E - F)x \leq Ax \end{cases} \tag{4.10}$$

where k is some nonnegative constant and ℓ_i the maximal increasing length of $As(i)$. Moreover, if

$$\lambda \geq \frac{k+\ell+2}{2} \tag{4.11}$$

where ℓ is the maximal increasing length of A, then

$$p_{ii} = p_{ii}^{(0)} \quad \text{for all } i \qquad \text{and} \qquad \sigma(B^{-1}A) \subset [1, \tfrac{1}{2}(k+\ell+2)]$$

Proof. Here again, it suffices to prove that, for the Strategy no 1, (4.10) implies (4.7) with $\lambda = (k+\ell+2)/2$. Now, let the diagonal matrix Δ be defined by

$$(\Delta x)_i = (k+\ell_i+2)^{-1}((F-E+A)x)_i \quad \text{for all } i$$

The first inequality (4.10) implies $Ax \geq \Delta x$, showing by Lemma 1 of the Appendix applied with $Bx = Ax$ that

$$(1 - 1/(k+\ell+2))Px \geq (F+A-\Delta)x$$

The required result follows from this inequality together with $\Delta x \leq \frac{1}{2}(F-E+A)x$ (by virtue of the nonnegativity of $(F-E+A)x$). □

This suggests the following policy : *(1) choose an ordering such that conditions (4.10) are as far as possible satisfied, while keeping an interesting value for the parameter ℓ; (2) use*

$$\lambda \approx \frac{\ell}{2}, \tag{4.12}$$

where ℓ is the maximal increasing length of the matrix A.

Strategy no 4. The following theorem generalizes a result first proved in [22].

THEOREM 3. *Assume (GA), A irreducible and let $E = F^t$. If, in addition,*

$$2\,Px \geq (A+F+E)x, \tag{4.13}$$

then B is regular with

$$\nu_{\max}(B^{-1}A) \;\leq\; \rho(M) \;\leq\; \max_i \; y_i^{-1}(My)_i \tag{4.14}$$

for any positive vector y, and where

$$M = (I - P^{-\frac{1}{2}}EP^{-\frac{1}{2}})^{-1} + (I - P^{-\frac{1}{2}}FP^{-\frac{1}{2}})^{-1} \tag{4.15}$$

Proof. Let $G = A + E + F$. Note that $diag(G) = diag(A)$ while $offdiag(G) \geq 0$. Thus, (4.13) implies that P is regular (the diagonal entries of an irreducible symmetric nonnegative definite matrix being positive [26]). Then, since $A = (G-2P)+(P-E)+(P-F)$ with $2P-G$ nonnegative definite, one may write :

$$\nu_{\max}(B^{-1}A) \;=\; \rho((P^{\frac{1}{2}} - EP^{-\frac{1}{2}})^{-1}A(P^{\frac{1}{2}} - P^{-\frac{1}{2}}F)^{-1}) \;\leq\; \rho(M)$$

and (4.14) follows from the nonnegativity of M. □

This theorem does not enable us to satisfy an a priori chosen upper bound. However, the fact that the nonnegative matrix M has a Neumann expansion with about ℓ terms is a good indication that the a posteriori computable upper bound (4.14), when it applies, will indeed be $\mathcal{O}(\ell)$, and suggests therefore a fourth strategy according to which P is computed so as to satisfy (4.13).

ALGORITHM 3. *A being a Stieltjes matrix, x a positive vector such that* $Ax \geq 0$ *and the* β_{ij}, $1 \leq i < j \leq n$ *some given (0,1) values, compute the upper triangular matrix* $U = P - F$ *with* $P = diag(U)$ *according to (2.1),(2.2) and, noting* $E = F^t$,

$$p_{ii} = \max(p_{ii}^{(0)}, \tfrac{1}{2}((A+F+E)x)_i/x_i) \tag{4.16}$$

COROLLARY 3. *If A is irreducible, then B is regular and satisfies (4.14).*

PROPERTY 3. *If A is irreducible and if, with* $E = F^t$,

$$(E-F)x \leq Ax, \tag{4.17}$$

then

$$p_{ii} = p_{ii}^{(0)} \quad \textit{for all } i \qquad \textit{and} \qquad \sigma(B^{-1}A) \subset [1, \ell+2]$$

where ℓ is the maximal increasing length of A.

Proof. One has $2P - A - F - E = (F - E + A) + 2(P - F - A)$, showing that (4.17) implies anyway (4.13), and therefore $p_{ii} = p_{ii}^{(0)}$ for all i when using the Strategy no 4. Further with $y = P^{\frac{1}{2}}x$, one obtains

$$(My)_i = ((I + \sum_{r=1}^{\ell_i^{(a)}} (E')^r)y)_i + ((I + \sum_{r=1}^{\ell_i^{(d)}} (F')^r)y)_i$$

where $E' = P^{-\frac{1}{2}}EP^{-\frac{1}{2}}$, $F' = P^{-\frac{1}{2}}FP^{-\frac{1}{2}}$, and where $\ell_i^{(a)}$ ($\ell_i^{(d)}$) denotes the maximal length of an increasing path ending at i (starting from i) (thus, $\ell_i^{(a)} = \ell(As(i))$. Now, (4.17) implies $E'y \leq P^{-\frac{1}{2}}(F+A)x \leq y$, while $F'y \leq y$. Whence :

$$(My)_i \leq (2 + \ell_i^{(a)} + \ell_i^{(d)})y_i$$

and the upper bound follows from $\ell = \max_{1 \leq i \leq n}(\ell_i^{(a)} + \ell_i^{(d)})$ □

Since there is no parameter to choose, the policy recommended here consists only in choosing an appropriate ordering.

5. Summary of lower eigenvalue analyses. Properties 1, 2 and 3 only establish the existence of a class of model problems for which the control of the largest eigenvalue through the associated theorem is compatible with unperturbed approximate factorizations. We need to go further and analyze the lowest eigenvalue behaviour of Strategies no 2, 3 and 4 when the associated "model problem" conditions (4.4), (4.10) or (4.17) are not satisfied. A first intuitive answer consists in saying that $\nu_{\min}(B^{-1}A)$ should be kept close to 1 when these conditions are approximately satisfied. Of course, we need some further mathematical analysis to know how precisely such a restriction has to be understood. This requires appropriate lower eigenvalue bounds. A technique for obtaining lower eigenvalue bounds for modified incomplete factorizations methods has been described in [9, 21], while another technique has been developed by Axelsson-Barker [5], and further improved by the present author [24]. We have unfortunately no room to explain here how to manage with these results, since their application require some handling of each example. We shall therefore only state the major conclusions one can draw from them. We first write the following conditions :

$$\ell^2 \, \lambda_{\min}(D^{-1}A) \gtrsim 1 \tag{5.1}$$

where $\lambda_{\min}(D^{-1}A)$ denote the first nonzero eigenvalue of $D^{-1}A$ (with $D = diag(A)$),

$$\begin{cases} \frac{1}{\ell}((F-E)x)_i - (Ax)_i \lesssim \ell^{-1} \, a_{ii}x_i & \text{if} \quad i \in I \\ \frac{1}{\ell}((F-E)x)_i - (Ax)_i \lesssim \ell^{-2} \, a_{ii}x_i & \text{if} \quad i \notin I \\ Card(I) \le O(n/\ell) \end{cases} \tag{5.2}$$

and

$$\begin{cases} ((E-F)x)_i - (Ax)_i \lesssim \ell^{-1} \, a_{ii}x_i & \text{if} \quad i \in J \\ ((F-E)x)_i - (Ax)_i \lesssim \ell^{-2} \, a_{ii}x_i & \text{if} \quad i \notin J \\ Card(J) \le O(n/\ell) \end{cases} \tag{5.3}$$

Then, as widely discussed in [26], mainly on the basis of Theorem 4.1 of [24] (which covers both regular and singular cases), one can obtain a lower bound close to 1 provided that :

- for Strategy no 2 : τ is chosen according to (4.6) and conditions (5.1) and (5.2) are satisfied;
- for Strategy no 3 : λ is chosen according to (4.12) and conditions (5.1), (5.2) and (5.3) are satisfied;
- for Strategy no 4 : conditions (5.1) and (5.3) are satisfied.

The condition on the lowest eigenvalue of $D^{-1}A$ is not surprising since the existence of a connection between the lower parts of the spectrum of the preconditioned and unpreconditioned systems is a well known fact [4] for all perturbed modified incomplete factorization methods (including unmodified, relaxed and SSOR methods), while the other conditions may effectively be viewed as weakened forms of the associated "model problem" conditions (4.4), (4.10) or (4.17).

Note finally that, as discussed in [26], condition (5.1) may further be exchanged for a similar condition to be satisfied by the p^{the} non zero eigenvalue of $D^{-1}A$ for p small. Indeed, if other conditions are met, we may then still obtain a lower bound close to 1 for the p^{the} non zero eigenvalue of $B^{-1}A$, and therefore prove an interesting convergence rate by combining this result with a bound analyzing the convergence of the conjugate gradient process in presence of small isolated eigenvalues.

6. Examples. By way of illustration, we consider the five-point finite difference approximation of the boundary value problem

$$-\nabla a(x,y)\nabla u(x,y) = f(x,y) \quad \text{on} \quad \Omega = (0,1)\times(0,1)$$

with (Problem 1)

$$a(x,y) = \begin{cases} d & \text{in } (\frac{1}{2},1)\times(\frac{1}{2},1) \\ 1 & \text{elsewhere,} \end{cases}$$

Dirichlet boundary conditions and a uniform square grid of mesh size h, or (Problem 2)

$$a(x,y) = \begin{cases} 100 & \text{in } (\frac{1}{6},\frac{5}{6}) \times (\frac{1}{6},\frac{5}{6}) \\ 1 & \text{elsewhere,} \end{cases}$$

Dirichlet boundary conditions on the bottom ($y = 0$) boundary and Neumann boundary conditions elsewhere, and a rectangular grid whose mesh $h_x \times h_y$ varies according to

$$h_x = \begin{cases} 4h/3 & \text{if } x \in (\frac{1}{6},\frac{5}{6}) \\ 2h/3 & \text{otherwise,} \end{cases}$$

$$h_y = \begin{cases} 4h/3 & \text{if } y \in (\frac{1}{6},\frac{5}{6}) \\ 2h/3 & \text{otherwise.} \end{cases}$$

The resulting system matrix A is an irreducible Stieltjes matrix which verifies $Ae \geq 0$, where $e = (1 \dots 1)^t$. We use the lexicographic ordering (starting at the left bottom corner). Note that $n = (h^{-1} - 1)^2$ and $\ell = 2h^{-1} - 4$ for Problem 1 while $n = h^{-1}(h^{-1} + 1)$ and $\ell = 2h^{-1} - 1$ for Problem 2 (where ℓ denotes the maximal increasing length of A).

We give in Table 1 the numerical results obtained for the modified incomplete factorizations associated with $x = e$ and $\beta = 0$. (For the PCG experiment, we test the relative residual error when using the algorithm with zero initial approximation in order to solve (1.1), where $b = Au_s$, u_s being the vector which approximates the function $(1+x)^2(1+y)(2-y)e^{xy}$ on the unit square).

TABLE 1

Numerical results for the test Problem 1 with $d = 10^0, 10^{-3}, 10^3$, and for the test Problem 2.

Strategy	Numerically determined value of $\kappa(B^{-1}A)$					Number of PCG iterations ($h^{-1} = 192$)			
	for $h^{-1} =$					to reach $\Vert r_i \Vert_2 / \Vert r_0 \Vert_2 \leq$			
	12	24	48	96	192	10^{-3}	10^{-5}	10^{-7}	10^{-9}
Problem 1, d=1									
1, 2, 3, 4	3.32	6.85	14.4	30.2	62.7	12	28	44	59
Problem 1, d=0.001									
1, 2	4.49	9.60	20.6	43.7	91.6	14	33	49	66
3 ($\lambda = \ell/2$)	3.68	9.52	30.6	111.	423.	49	83	110	133
4	3.55	9.11	29.8	109.	420.	48	83	110	133
Problem 1, d=1000									
1, 4	3.10	6.75	14.8	31.8	67.5	14	29	44	59
2 ($\tau = 1 - \ell^{-1}$)	3.50	7.21	15.2	32.1	66.9	13	30	45	59
3 ($\lambda = \ell/2$)	3.16	6.83	14.8	31.6	66.4	13	30	45	59
Problem 2									
1	107.	375.	1432.	5625.	21.E3	103	163	222	283
2 ($\tau = 1 - \ell^{-1}$)	91.5	158.	316.	683.	1508.	54	74	94	115
3 ($\lambda = \ell/2$)	236.	850.	3256.	13.E3	53.E3	171	228	281	326
4	249.	951.	3771.	15.E3	62.E3	184	243	301	349

These results show clearly that Strategy no 2 has the widest scope of application. This follows from the fact that condition (5.2) always holds in the second order discrete elliptic PDE context (for the lexicographic ordering used here). Other strategies appear less robust. Strategy no 1 behaves poorly for Problem 2 which satisfies none of the model problem conditions (4.4), (4.10) or (4.17), while the lowest eigenvalue behaviour of Strategies no 3 and 4 become unreliable when condition (5.3) is not satisfied, as occurs in Problem 1 with $d \ll 1$ and in Problem 2.

Note in this respect that one may try to improve the behaviour of Strategies no 1, 3 and 4 by changing the ordering so as to satisfy the model problem condition (4.17), or at least its weakened form (5.3). This seems feasible for the examples considered here whose configurations are particularly simple, but not in general since no ordering procedure is known to automatically achieve this goal (nor even is it known whether such an ordering exists).

Thus, if the bound of Property 2 is the best, indicating a very interesting behaviour of Strategy no 3 (as one could already deduce from the model problem analysis in [3]), one has to remember that this result is obtained under very restrictive conditions, and that Strategy no 3 has actually the narrowest scope of application.

Acknowledgements. We are greatly indebted to Professor Robert Beauwens and Mr Magolu Monga-Made for useful comments.

Appendix.

LEMMA 1. *Let* $U = P - F$ *with* $P = diag(U)$ *be an upper triangular M-matrix. Set* $B = U^tP^+U$ *(where* P^+ *is defined by (2.5)) and* $E = F^t$*, and let* x *be a positive vector and* Δ *a diagonal matrix such that* $\Delta x \leq Bx$*. If, for some subset* M *of [1,n],*

$$(k + \ell_i + 2)^{-1}((F - E + B)x)_i \leq (\Delta x)_i \quad \textit{for all } i \in As(M) \tag{A.1}$$

where k *is nonnegative and* ℓ_i *the maximal increasing length os* $As(i)$*, then*

$$(1 - 1/(k + \ell_i + 2))(Px)_i \geq ((F + B - \Delta)x)_i \tag{A.2}$$

for all $i \in M$*.*

Proof. By induction. Assume (A.2) for all $i \in P(j)$. Note that (A.2) and $\Delta x \leq Bx$ imply $((P - F)x)_i \geq 0$ while $i \in P(j) \Rightarrow (Fx)_i > 0$, so that $((P - F)x)_i \geq 0$ implies $p_{ii} > 0$. Then :

$$\begin{aligned}((P - F - B)_j &= (EP^+(P - F)x)_j \\ &= -\sum_{i \in P(j)} \frac{u_{ij}}{p_{ii}}((P - F)x)_i \\ &\geq -\sum_{i \in P(j)} (k + \ell_i + 2)^{-1}u_{ij}x_i \\ &\geq (k + \ell_j + 1)^{-1}(Ex)_j \\ &\geq \frac{1}{k + \ell_j + 1}((F + B)x)_j - \frac{k + \ell_j + 2}{k + \ell_j + 1}(\Delta x)_j\end{aligned}$$

whence the required result. □

REFERENCES

[1] O. AXELSSON, *A generalized SSOR method*, BIT, 13 (1972), pp. 443–467.

[2] ———, *A class of iterative methods for finite element equations*, Comp. Meth. Appl. Mech. Eng., 9 (1976), pp. 123–137.

[3] ——, *On iterative solution of elliptic difference equations on a mesh connected array of processors*, J. High Speed Comput., 1 (1989), pp. 165–184.

[4] ——, *On the eigenvalue distribution of relaxed incomplete factorization methods and the rate of convergence of conjugate gradient methods*, Technical Report, Departement of Mathematics, Catholic University, Nijmegen, The Netherlands, 1989.

[5] O. AXELSSON AND V. BARKER, *Finite Element Solution of Boundary Value Problems. Theory and Computation*, Academic Press, New York, 1984.

[6] O. AXELSSON AND V. EIJKHOUT, *Robust vectorizable preconditioners for three-dimensional elliptic difference equations with anisotropy*, in Algorithms and Applications on Vector and Parallel Computers, H. te Riele, T. J. Dekker, and H. van der Vorst, eds., Elsevier (North Holland), Amsterdam, 1987, pp. 279–306.

[7] O. AXELSSON AND G. LINDSKOG, *On the rate of convergence of the preconditioned conjugate gradient method*, Numer. Math., 48 (1986), pp. 499–523.

[8] R. BEAUWENS, *Upper eigenvalue bounds for pencils of matrices*, Lin. Alg. Appl., 62 (1984), pp. 87–104.

[9] ——, *Lower eigenvalue bounds for pencils of matrices*, Lin. Alg. Appl., 85 (1987), pp. 101–119.

[10] ——, *Modified incomplete factorization strategies*, in Preconditioned Conjugate Gradient Methods, O. Axelsson and L. Kolotilina, eds., Lectures Notes in Mathematics No. 1457, Springer-Verlag, 1990, pp. 1–16.

[11] R. BEAUWENS AND L. QUENON, *Existence criteria for partial matrix factorizations in iterative methods*, Siam J. Numer. Anal., 13 (1976), pp. 615–643.

[12] R. BEAUWENS AND R. WILMET, *Conditioning analysis of positive definite matrices by approximate factorizations*, J. Comput. Appl. Math., 26 (1989), pp. 257–269.

[13] A. BERMAN AND R. PLEMMONS, *Nonnegative Matrices in the Mathematical Sciences*, Academic Press, New York, 1979.

[14] V. EIJKHOUT, *Beware of unperturbed modified incomplete factorizations*, these proceedings, pp. 583–591.

[15] A. GREENBAUM, *Behavior of slightly perturbed lanczos and conjugate-gradient recurrences*, Lin. Alg. Appl., 113 (1989), pp. 7–63.

[16] I. GUSTAFSSON, *On first order factorization methods for the solution of problems with mixed boundary conditions and problems with discontinuous material coefficients*, Research Report 77.13R, Dept. of Computer Sciences, Chalmers Univ. of Technology and Univ. of Goeteborg, Goeteborg, Sweden, 1977.

[17] ——, *Modified Incomplete Cholesky (MIC) methods*, in Preconditioning Methods. Theory and Applications, D. Evans, ed., Gordon and Breach, New York-London-Paris, 1983, pp. 265–293.

[18] F. HARARY, *Graph Theory*, Addison-Wesley, Reading, 1969.

[19] A. JENNINGS, *Influence of the eigenvalue spectrum on the convergence rate of the conjugate gradient method*, J. Inst. Maths. Applics., 20 (1977), pp. 61–72.

[20] M. MAGOLU, *Sparse block approximate factorizations for singular problems*, these proceedings, pp. 519–529.

[21] ——, *Lower eigenvalue bounds for singular pencils of matrices*, J. Comput. Appl. Math., to appear.

[22] M. MAGOLU AND Y. NOTAY, *On the conditioning analysis of block approximate factorization methods*, Lin. Alg. Appl., to appear.

[23] Y. NOTAY, *Polynomial acceleration of iterative schemes associated with subproper splittings*, J. Comput. Appl. Math., 24 (1988), pp. 153–167.

[24] ——, *Incomplete factorization of singular linear systems*, BIT, 29 (1989), pp. 682–702.

[25] ——, *Solving positive (semi)definite linear systems by preconditioned iterative methods*, in Preconditioned Conjugate Gradient Methods, O. Axelsson and L. Kolotilina, eds., Lectures Notes in Mathematics No. 1457, Springer-Verlag, 1990, pp. 105–125.

[26] ——, *Résolution iterative de systèmes linéaires par factorizations approchées*, PhD thesis, Service de Métrologie Nucléaire, Université Libre de Bruxelles, Brussels, Belgium, 1991.

[27] H. VAN DER VORST, *The convergence behaviour of preconditioned CG and CG-S*, in Preconditioned Conjugate Gradient Methods, O. Axelsson and L. Kolotilina, eds., Lectures Notes in Mathematics No. 1457, Springer-Verlag, 1990, pp. 126–136.