SIAM J. SCI. COMPUT.
Vol. 16, No. 2, pp. 378–399, March 1995

ORDERING STRATEGIES FOR MODIFIED BLOCK INCOMPLETE FACTORIZATIONS*

MAGOLU MONGA-MADE†

Abstract. In this study the aim is to first investigate how the rate of convergence of unperturbed modified block incomplete factorization methods depends on the discrete partial differential equation (PDE) to solve, and next, on this basis, to provide simple practical rules for easily selecting orderings that result in rapid convergence. It emerges from the analysis here that for discrete PDEs, ordering schemes that optimize the rate of convergence strongly depend on the variation of both the PDE coefficients and the mesh sizes. The arguments made, which bring to light the appreciable potentialities of modified methods, also display insight into why block versions are more robust than pointwise ones.

Key words. partial differential equations, discretizations, large sparse linear systems, orderings, preconditionings, conjugate gradient acceleration

AMS subject classifications. 65F10, 65N20, 65F35, 65B99

1. Introduction.

Guided by various motivations, different ordering techniques have been proposed in the literature for the numerical solution of partial differential equations (PDEs). These range from those proposed to reduce the size of fill-in in direct methods (see, e.g., [30]) or to keep control of fill-ins in iterative methods (see, e.g., [6]), to those designed to increase parallelism (see, e.g., [21], [22], [24], [26], [28], [50], and [51]) or to allow some degree of vectorization (see, e.g., [1], [22], [23], [29], [52], and [56]) in iterative methods. The major drawback with most orderings proposed in the context of incomplete factorization methods is that, compared with natural orderings, they seriously degrade the speed of convergence. Several attempts have been made to explain why this is the case (see the above-mentioned papers); most of them concentrated on pointwise unmodified incomplete factorizations and discussed arguments related to the size and the structure of the remainder matrices.

In the case of pointwise modified versions, it is worth mentioning Eijkhout's interesting analysis [27], where it is shown, through existence criteria that generalize earlier results by Notay [47], that certain parallel orderings (including the well-known red-black) give rise to singular preconditioners, unless small (say $\mathcal{O}(h^2)$ for discrete PDEs with mesh size parameter h) diagonal perturbations are added during the factorization process. In the latter case, one in general results in an $\mathcal{O}(h^{-2})$ spectral condition number, contrasting with $\mathcal{O}(h^{-1})$ for natural orderings in favorable circumstances. In [35], Kuo and Chan used a Fourier-like analysis to rigorously establish, for the model Poisson equation with Dirichlet boundary conditions on a square, the $\mathcal{O}(h^{-2})$ behavior of the spectral condition number determined by the red-black $\mathcal{O}(h^2)$-perturbed modified incomplete factorizations. Another complication with the latter methods is that they require some delicate parameter estimation. To circumvent these inconveniences, Beauwens [9] advocated *appropriately* allowing some fill-ins; it turns out from the theoretical and numerical arguments that he provided that, in so doing, a big improvement in the convergence rate can be achieved for some parallel ordering techniques, with a reasonable level of fill-ins.

In this work, we shall deal with unperturbed modified incomplete factorization methods, focussing on blockwise versions. Our purpose is to examine more closely the extent to which the upper spectral bound theories investigated in [41] can help to determine in advance,

*Received by the editors November 25, 1992; accepted for publication (in revised form) January 31, 1994. This research was supported by the "Programme d'Impulsion en Technologie de l'Information" financed by the Belgian State, under contract IT/IF/14.

†Universite Libre de Bruxelles, Service de Metrologie Nucleaire (CP 165), 50 avenue F.D. Roosevelt, B-1050 Brussels, Belgium (mmmago@ulb.ac.be).

for a given discrete PDE, whether a specified ordering scheme would result in an $\mathcal{O}(h^{-1})$ spectral condition number or not. The key idea is that the sufficient conditions discussed in [41] are amenable to interpretation in terms of both the PDE coefficients and the mesh sizes. Our arguments also provide a framework within which the relative superiority of block preconditionings over pointwise ones may be explained.

The rest of our exposition is outlined as follows. Definitions and notation we need are gathered in §2. Section 3 consists of a short review of the unperturbed modified block incomplete factorizations; it includes a general description of the preconditionings and some relevant results on both the existence analysis and the condition number theory. In §4, the main contribution of the work, special attention is paid to elaborate ordering strategies with the condition number analysis as basic support, which allows us to take advantage of the potentialities of unperturbed modified methods. To illustrate in a more quantitative way the practical interest of our findings, some experimental results are reported in §5. Section 6 is devoted to concluding remarks and further perspectives.

2. Terminology and notation.

2.1. General notation. The symbols A^t, A^+, $\mathcal{N}(A)$, $\mathcal{R}(A)$, $\rho(A)$, $\lambda_{\min}(A)$, and $\lambda_{\max}(A)$ stand for, respectively, the transpose, the Moore–Penrose inverse [12], the null space, the range, the spectral radius, the smallest, and the largest *nonzero* eigenvalues of the matrix A. By a $\{1\}$-*inverse* of a matrix A we mean any matrix X such that $AXA = A$.

By $\mathcal{P}_{S,T}$ we denote the projector with range S and null space T (S and T are complementary subspaces in $\mathbb{C}^n$).

We use e to denote the vector whose components are all equal to unity. By a (0, 1)-matrix we mean a matrix whose nonzero entries are equal to 1.

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 $A \geq 0$ ($A > 0$).

A real square matrix A is called an *M-matrix* if there exists a nonnegative number s such that $sI - A \geq 0$ with $\rho(sI - A) \leq s$. A symmetric M-matrix is called a *Stieltjes matrix*. This definition of Stieltjes matrix differs from the one commonly used in the literature (see, e.g., [13] and [59]) in that the matrix is allowed to be singular.

The following lemma gathers some needed properties of the Stieltjes matrix borrowed from [13].

LEMMA 2.1. *Let A be a symmetric matrix such that* $\operatorname{offdiag}_p(A) \leq 0$ (*see* §2.3 *for the notation*). *Then the following conditions are equivalent:*

(1) *A is a Stieltjes matrix;*

(2) *A is nonnegative definite;*

(3) $\exists x > 0$ *such that* $Ax \geq 0$.

Further, if A is irreducible and if any of the above conditions is satisfied, then the following statements are also equivalent:

(a) *A is singular;*

(b) $\exists x > 0$ *such that* $\mathcal{N}(A) = \operatorname{Span}\{x\}$*;*

(c) $\forall x > 0 : Ax \geq 0 \Rightarrow Ax = 0$.

2.2. Hadamard multiplication. The Hadamard product $A * B$ of the matrices $A = (a_{ij})$ and $B = (b_{ij})$ of the same dimensions is the element by element multiplication, i.e., $(A * B)_{ij} = a_{ij}b_{ij}$. The unit matrix with respect to the Hadamard multiplication, denoted ε, is the matrix whose entries are all equal to unity.

2.3. Partitionings. Any partitioning of an n-vector $x = (x_I)$ into block components x_I of dimensions n_I, $I = 1, 2, \ldots, M$ (with $\sum_{I=1}^{M} n_I = n$) is uniquely determined by a partitioning $\pi = (\pi_I)_{1 \le I \le M}$ of the set of the first n integers. *We assume throughout the paper that all n-vectors are partitioned according to a given such partitioning.* The same partitioning π induces also a partitioning of any $n \times n$ matrix A into block components A_{IJ} of dimensions $n_I \times n_J$ and we shall similarly assume that all $n \times n$ matrices are partitioned in this way. Lower case indices refer to scalar entries and capital indices to block entries.

A matrix that is block diagonal (resp., tridiagonal, triangular) relative to the π-partitioning will be referred to as π-diagonal (resp., π-tridiagonal, π-triangular). We use the notation $\mathrm{diag}_\pi(A)$ (resp., $\mathrm{tridiag}_\pi(A)$) to denote the π-diagonal (resp., π-tridiagonal) matrix whose block diagonal (resp., block tridiagonal) part coincides with that of A, and $\mathrm{offdiag}_\pi(A)$ to denote $A - \mathrm{diag}_\pi(A)$. In the case of the usual point partitioning, we set $\pi = p$.

2.4. Standard point LU-factorization. Given a Stieltjes matrix A, by its *standard point LU-factorization* we understand the factorization $A = U^t P^+ U$ such that U is p-upper triangular and $P = \mathrm{diag}_p(U)$.

2.5. Graph notions. All graph concepts considered hereafter refer to *ordered* undirected graphs [30], [32] unless specifically stated otherwise. We recall that the quotient graph of the matrix A with respect to the partitioning $\pi = (\pi_I)_{1 \le I \le M}$, denoted $\mathcal{G}(A)/\pi$, may be identified with the node set $\{1, 2, \dots, M\}$ together with the edge set $\mathcal{E}(\mathcal{G})$ defined by $\{I, J\} \in \mathcal{E}(\mathcal{G})$ if and only if $A_{IJ} \ne 0$ or $A_{JI} \ne 0$. For the sake of easy reference, we recall from [7], [11], and [32] the following additional notions.

An *increasing path* in a graph $\mathcal{G}$ is a path $I_0, I_1, \dots, I_\ell$ such that $I_0 < I_1 \cdots < I_\ell$.

The *maximal increasing length* $\ell(\mathcal{S})$ of a nonempty subset $\mathcal{S}$ of the node set of the graph $\mathcal{G}$ is the length of a longest increasing path in the subgraph of $\mathcal{G}$ induced by $\mathcal{S}$. We further set $\ell(\emptyset) = -1$.

For any node I of a graph $\mathcal{G}$, the *ascent* $As(I)$ of I is defined as

$$As(I) = \{J;\ \text{there exists an increasing path from } J \text{ to } I\}.$$

Observe that $I \in As(I)$ because a path of zero length is an increasing path.

A node J of a graph $\mathcal{G}$ is called a *precursor* (*successor*) of another node I of $\mathcal{G}$ if (I, J) belongs to the edge set of $\mathcal{G}$ with $J < I$ ($J > I$); the set of precursors (successors) of I is denoted by $P(I)$ ($S(I)$). We further set $P(\emptyset) = S(\emptyset) = \emptyset$.

An *initial node* (resp., *final node*) in a graph $\mathcal{G}$ is a node I such that $P(I) = \emptyset$ (resp., $S(I) = \emptyset$).

3. Unperturbed modified block incomplete factorizations.

3.1. Description of the preconditioning. Let A be a Stieltjes matrix and x a positive vector such that $Ax \ge 0$. Let β and γ denote some given symmetric (0,1)-matrices different from ε. Let F be the strictly π-upper triangular matrix and P the π-diagonal matrix whose entries are computed from the relation

$$P - F^t - F = A - \beta * (F^t K F) - \Omega, \tag{3.1}$$

where $K = \gamma * P^+$ and Ω is the p-diagonal matrix determined by

$$\Omega x = \left(F^t P^+ F - \beta * (F^t K F)\right) x\ . \tag{3.2}$$

The matrix

$$B = (P - F^t) P^+ (P - F) \tag{3.3}$$

is referred to as the (unperturbed) modified block incomplete LU (or in abbreviated form, MBILU) factorization of A associated with x, β, and γ. The matrix Ω is known as the *modification matrix*; by dropping it one gets the standard *unmodified* block incomplete factorizations.

The perturbed versions of *MBILU* consist in adding some appropriate perturbation term to the p-diagonal of P.

From an implementation viewpoint, the relations (3.1) and (3.2) may be expanded in the algorithmic form

$$\begin{aligned}
&P_{11} = A_{11}, && \\
&F_{1J} = -A_{1J}, && 1 < J \leq M, \\
(3.4)\qquad &P_{II} = A_{II} - \beta_{II} * \left(\sum_{S=1}^{I-1} F_{SI}^t K_{SS} F_{SI} \right) - \Omega_{II}, && 1 < I \leq M, \\
&F_{IJ} = -A_{IJ} + \beta_{IJ} * \left(\sum_{S=1}^{I-1} F_{SI}^t K_{SS} F_{SJ} \right), && 1 < I < J \leq M,
\end{aligned}$$

where $K_{SS} = \gamma_{SS} * P_{SS}^+$ and Ω_{II} denotes the p-diagonal matrix defined by

$$(3.5)\qquad \Omega_{II} x_I = \sum_{S=1}^{I-1} \sum_{J=S+1}^{M} \left(F_{SI}^t P_{SS}^+ F_{SJ} - \beta_{IJ} * \left(F_{SI}^t K_{SS} F_{SJ} \right) \right) x_J .$$

The above quantities may be computed as follows: first compute P_{11} and F_{1J} for $J > 1$; next for $I = 2, 3, \ldots, M$, compute (in order) $K_{I-1,I-1}$, Ω_{II}, P_{II}, and F_{IJ} for $J > I$. In the case of the usual pointwise MILU(0) preconditioner, one has $\pi = p$, $x = e$,

$$\beta_{ij} = \begin{cases} 1 & \text{if } a_{ij} \neq 0, \\ 0 & \text{otherwise,} \end{cases}$$

and

$$\gamma_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

3.2. Practical considerations. We first give a theorem which provides sufficient conditions for the algorithm (3.4)–(3.5) to be carried out trouble-free, as well as some other relevant properties.

THEOREM 3.1. *Let A be an irreducible Stieltjes matrix and x a positive vector such that all π-diagonal entries A_{II} of A are irreducible and $Ax \geq 0$. Let β and γ be symmetric* $(0, 1)$-*matrices and $B = (P - F^t)P^+(P - F)$ stand for the* MBILU *factorization of A associated with x, β, and γ. Set $D = \operatorname{diag}_\pi(A)$ and let F_0 be the strictly π-upper triangular matrix such that $F_0^t + F_0 = -\operatorname{offdiag}_\pi(A)$.*

Define

$$\begin{aligned}
&\mathcal{M} = \{1, 2, \ldots, M\}, \\
&\mathcal{F} = \{I \ : \ I \in \mathcal{M}\ ,\ I \text{ is a final node of } \mathcal{G}(F)/\pi\}, \\
&\mathcal{L} = \mathcal{M} \setminus \mathcal{F}.
\end{aligned}$$

Then:

(1) *P is a Stieltjes matrix and $P - F$ is an M-matrix such that:*
  (1.1) *$\forall I \in \mathcal{L}$: P_{II} is nonsingular and irreducible;*
  (1.2) *$\forall I \in \mathcal{F}$: P_{II} is irreducible;*
  (1.3) *$(P - F)x \geq Ax$;*

(2) *B is nonnegative definite with $Bx = Ax$ and* $\operatorname{offdiag}_p(A - B) \leq 0$*;*

(3) *$F \geq F_0$ and* $\operatorname{offdiag}_p(P) \leq \operatorname{offdiag}_p(D)$*;*

(4) *$P^+F \geq 0$, $PP^+F = F$, and $(I - P^+F)x \geq 0$;*

(5) *$\mathcal{N}(B) = \mathcal{N}(P - F)$;*

(6) *If* $\mathcal{F} = \{M\}$ *then the following assertions are equivalent:*

(6.1) B *is singular;*

(6.2) P_{MM} *is singular;*

(6.3) $(P - F)x = 0$*;*

(6.4) A *is singular;*

(6.5) $\mathcal{N}(B) = \text{Span}\{x\}$.

(7) *If* $\mathcal{F} = \{M\}$ *and any of the conditions* (6.1)–(6.5) *is fulfilled, then* $\mathcal{N}(B) = \mathcal{N}(A)$*;*

(8) *The following assertions are equivalent:*

(8.1) B *is nonsingular;*

(8.2) $\forall I \in \mathcal{F}$: $\exists J \in AS(I)$ *in* $\mathcal{G}(F)/\pi$ *such that* $(Ax)_J \neq 0$.

Proof. Statements (1)–(7) are straightforward generalizations of Theorem 3.2 (1)–(7) of [39]. To prove (8), we first observe that, since $PP^+F = F$ (by (4)), the relation $Bx = Ax$ may also be written

$$(3.6) \qquad (P - F)x = Ax + F^t P^+ (P - F)x \ ,$$

or, equivalently, in a more detailed form,

$$(3.7) \qquad \begin{aligned} ((P - F)x)_I &= (Ax)_I + \sum_{J=1}^{I-1} F^t_{JI} P^+_{JJ} \, ((P - F)x)_J \\ &= (Ax)_I + \sum_{J \in P(I)} F^t_{JI} P^+_{JJ} ((P - F)x)_J \qquad \text{for } 1 \leq I \leq M \ . \end{aligned}$$

Now, (8.1) $\Rightarrow$ (8.2). Indeed, if there exists $S \in \mathcal{F}$ such that for all $J \in AS(S)$, $(Ax)_J = 0$ then (readily by induction through (3.7)) $((P - F)x)_J = 0$ for all $J \in AS(S)$ in $\mathcal{G}(F)/\pi$, whence in particular $P_{SS}x_S = (Px)_S = ((P - F)x)_S = 0$, which shows that P_{SS} is singular. Therefore P, $P - F$, and B are singular.

Finally, assume that (8.2) holds; since for any $I \in \mathcal{F}$, $AS(I) \backslash \{I\} \subset \mathcal{L}$, one has by (1.1) that $P^+_{JJ} = P^{-1}_{JJ} > 0$ for all $J \in AS(I) \backslash \{I\}$, which together with (8.2) and (straightforward induction on) (3.7) implies that for all $I \in \mathcal{F}$, $P_{II}x_I \geq 0$ with $(P_{II}x_I)_i > 0$ for at least one $i \in \pi_I$. Given that P_{II} is irreducible Stieltjes by (1.2), on easily deduces from Lemma 2.1 that P_{II} is nonsingular; combine with (1.1) to conclude that P, $P - F$, and B are nonsingular. □

Under the assumptions of the above theorem, whenever only the last node (M) is a final node in $\mathcal{G}(F)/\pi$, one has that MBILU factorizations (B) of the matrix A are singular if and only if the original matrix itself is so. In this case, to solve the preconditioning system

$$(3.8) \qquad Bz^{(i)} = r^{(i)} \qquad \text{with} \qquad r^{(i)} = b - Au^{(i)},$$

which occurs at each preconditioned conjugate gradient (PCG) iteration, one should resort to any $\{1\}$-inverse $B^{(1)}$ of B, or preferably to its Moore–Penrose inverse, which may be written as ([12, pp. 59–60])

$$(3.9) \qquad B^+ = P_{\mathcal{R}(B),\mathcal{N}(B)} B^{(1)} P_{\mathcal{R}(B),\mathcal{N}(B)},$$

in order to avoid rounding errors spoiling the convergence of the PCG process [33], [39], [40], [47]. Since $\mathcal{N}(B) = \mathcal{N}(A) = \text{Span}\{x\}$ (see Theorem 3.1(6.5), (7)) we also have $\mathcal{R}(B) = \mathcal{R}(A)$, and therefore

$$(3.10) \qquad \forall z \in \mathbb{C}^n \ : \ \mathcal{P}_{\mathcal{R}(A),\mathcal{N}(A)} \, z = \mathcal{P}_{\mathcal{R}(B),\mathcal{N}(B)} \, z = z - \frac{(z, x)}{(x, x)} x \ .$$

The result to come, which is borrowed from [39], provides a practical tool for computing a $\{1\}$-inverse of a singular preconditioning matrix B.

THEOREM 3.2. *Let F be a strictly π-upper triangular matrix and P a symmetric π-diagonal matrix such that P_{II} is nonsingular for $I \in \mathcal{L} = \{1, 2, \ldots, M-1\}$ while P_{MM} is an irreducible singular Stieltjes matrix with $L_\mu \underset{\sim}{P_\mu^+} U_\mu$ as its standard point LU-factorization.*

Set $B = (P - F^t)P^+(P - F)$ and let $\widetilde{U}_\mu$ be obtained from $\underset{\sim}{U_\mu}$ by exchanging its last diagonal entry (which is zero) for an arbitrary positive number[1]*. Set $\widetilde{P}_\mu = \operatorname{diag}_p(\widetilde{U}_\mu)$. Let $\widetilde{P}$ denote the π-diagonal matrix defined by*

$$\widetilde{P}_{II} = \begin{cases} P_{II} & \textit{if } I \in \mathcal{L}, \\ \widetilde{U}_\mu^t \widetilde{P}_\mu^{-1} \widetilde{U}_\mu & \textit{if } I = M. \end{cases} \tag{3.11}$$

Then $\widetilde{B}^{-1} = (\widetilde{P} - F)^{-1}\widetilde{P}(\widetilde{P} - F^t)^{-1}$ is a $\{1\}$-inverse of B.

We have so far explained how to cope with a singular preconditioning system that stems from MBILU factorizations of an irreducible singular Stieltjes matrix, whenever the set of final nodes of $\mathcal{G}(F)/\pi$ reduces to the last node. To have the latter requirement fulfilled in unfavorable situations, fill-in should appropriately be accepted outside the main block diagonal part. This should also be done in the case where the original matrix is nonsingular but the condition (8.2) of Theorem 3.1 is not satisfied, in order to avoid a singular preconditioning matrix which is out of the question in the involved case.

Another possible remedy consists in using *dynamic* diagonal perturbations (or additive corrections) [37], [39], [40], [42], but this is outside the scope of the present study.

3.3. Condition number analysis. The number of PCG iterations needed to achieve a given accuracy is known to be bounded above by $\mathcal{O}(\sqrt{\kappa})$, where κ denotes the spectral condition number (say the ratio of the largest to the smallest nonzero eigenvalue) of the preconditioned system; see, e.g., [2], [31], and [34] for the nonsingular case and [47] for the singular case. However, it is true that the rate of convergence also depends on the distribution of the eigenvalues [4], [5], [17], [18], [49], [54], [55], [57], [58]; but, as it is difficult to theoretically predict the overall distribution of eigenvalues, except for some simple problems with specified ordering (see, e.g., [14], [15], [16], [25], [35], [59]), much research focuses on finding analytical bounds for the spectral condition number.

As far as (unperturbed) MBILU factorizations are concerned, one easily deduces from Theorem 3.1(2), by arguing as in the proof of Theorem 4.1 of [38], that the smallest nonzero eigenvalue of the preconditioned system is 1, so that the problem reduces to bounding the largest eigenvalue. For this purpose, simple analytical bounds have been developed in [41] and generalized in [40]. We shall not review all the bounds obtained to date, but rather focus on straightforward application of (the proofs of) two of them, say Theorem 4.2 and Theorem 4.5 of [41], that are suitable for our purpose. Earlier results may be found in [3], [10], [16], [36], [45], and [48].

THEOREM 3.3. *In addition to the assumptions of Theorem* 3.1, *set*

$$\mathcal{J} = \{I \ : \ I \in \mathcal{M} \ , \ I \textit{ is an initial node of } \mathcal{G}(F)/\pi\} \ .$$

Assume further that

$$(Ax)_I > 0 \quad \textit{for all } \ I \in \mathcal{J} \tag{3.12}$$

and

$$((F - F^t)x)_I \le (Ax)_I + c_I(Px)_I \quad \textit{for all } \ I \in \mathcal{L}, \tag{3.13}$$

[1] In practice, one should use 1 to avoid amplifying rounding errors.

with $-1 \le c_I \ll 1$. *Then*

$$\lambda_{\max}(B^+A) \le \mathcal{O}(\ell), \tag{3.14}$$

where ℓ *denotes the maximal increasing length of* $\mathcal{G}(F)/\pi$.

A more concrete expression for the upper spectral bound (3.14) may be computed by using elementary algebraic and graphical manipulations specified in [41]; in particular, if $c_I = 0$ for all $I \in \mathcal{L}$, then (see Remark 4.1 of [41])

$$\lambda_{\max}(B^+A) \le \ell + 1\ . \tag{3.15}$$

From a practical viewpoint, it is worth mentioning that if

$$(Fx)_I > 0 \quad \text{for all } \ I \in \mathcal{L}, \tag{3.16}$$

then, through (a straightforward adaptation of the proof of) Lemma A.3 of [41], inequalities (3.13) are satisfied with

$$c_I = \max_{i \in \pi_I} \max \begin{cases} \left(-1,\ \frac{((F-A)x)_i}{((F+A)x)_i}\right) & \text{if } I \in \mathcal{J}, \\ \left(0,\ \frac{((F-F^t)x)_i}{(Fx)_i}\right) & \text{otherwise.} \end{cases} \tag{3.17}$$

THEOREM 3.4. *Adding to the assumptions of Theorem* 3.1 *that* $(Fx)_I > 0$ *for* $I \in \mathcal{L}$ *and*

$$(F^t - F)x \le Ax, \tag{3.18}$$

then

$$\lambda_{\max}(B^+A) \le \ell + 2, \tag{3.19}$$

where ℓ *denotes the maximal increasing length of* $\mathcal{G}(F)/\pi$.

In practical applications involving discretization of PDEs on rectangular meshes with (average) mesh size h such that $M = \mathcal{O}(h^{-1})$, one has $\ell = \mathcal{O}(M) = \mathcal{O}(h^{-1})$, showing that when all "goes right," the spectral condition number of the preconditioned system is bounded above by $\mathcal{O}(h^{-1})$. We know from the analysis of illustrative examples in [40] and [41] that the decisive role is played by conditions (3.13) and (3.18). Therefore, it would be of high interest to (try to) organize the ordering of the unknowns so as to have the required conditions satisfied at most if not all involved nodes. This is the purpose of the next section.

4. Ordering strategies.

By way of illustration, we consider the following self-adjoint second-order elliptic PDE

$$\begin{aligned} -\frac{\partial}{\partial x}\left(p\frac{\partial}{\partial x}u(x,y)\right) - \frac{\partial}{\partial y}\left(q\frac{\partial}{\partial y}u(x,y)\right) + t\,u(x,y) &= f(x,y) \quad \text{in } \Omega, \\ u(x,y) &= g_0(x,y) \quad \text{on } \Gamma_0, \\ \frac{\partial}{\partial n}u(x,y) &= g_1(x,y) \quad \text{on } \Gamma_1, \\ \frac{\partial}{\partial n}u(x,y) + \omega\,u(x,y) &= g_2(x,y) \quad \text{on } \Gamma_2, \end{aligned} \tag{4.1}$$

where Ω denotes the rectangle $0 < x < X,\ 0 < y < Y$; Γ_0, Γ_1, and Γ_2 are parts (possibly empty) of its boundary Γ; the coefficients $p(x,y)$ and $q(x,y)$ are positive, bounded, and piecewise constant; $t(x,y)$ is nonnegative, bounded, and piecewise constant; while $\omega(x,y)$

is positive and piecewise constant on some rectangular mesh covering the closure $\overline{\Omega}$ of Ω. Of course, in the case where $\Gamma_1 = \Gamma$ and $t(x, y) = 0$ on Ω, it is assumed that the compatibility condition

$$\int_{\Omega} f(x, y)\, d\Omega + \int_{\Gamma} g_1(x, y)\, d\Gamma = 0 \tag{4.2}$$

holds. We use the five-point finite difference approximation (point scheme box integration [46], [59], [60]); the mesh points are numbered *lexicographically* (see Fig. 1). The resulting matrix A is a block tridiagonal irreducibly diagonally dominant Stieltjes matrix, which we partition according to the line partitioning $(\pi_l)_{1 \le l \le M}$, keeping together the unknowns in the x-direction.

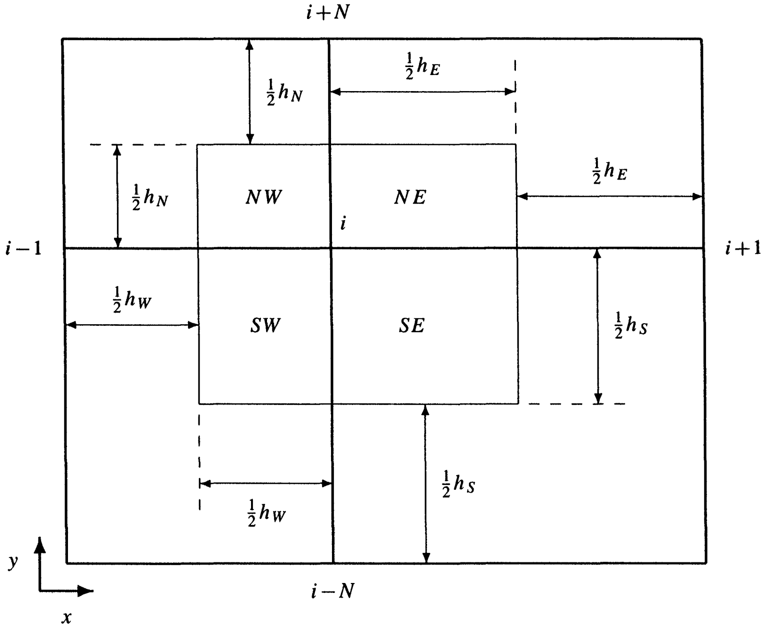

FIG. 1. *Portion of the mesh around the interior grid point i and its four neighbors for the (point scheme) box integration; N denotes the number of unknowns in the x-direction.*

Our analysis bears on the conditioning properties of MBILU factorizations

$$B = (P - F^t)P^+(P - F) \tag{4.3}$$

associated with $x = e$, $\beta = \text{tridiag}_p(\varepsilon)$, and $\gamma = \text{tridiag}_p(\varepsilon)$. In the case under consideration here there holds

$$-\,(F^t + F) = \text{offdiag}_\pi(A)\ , \tag{4.4}$$

which implies that the vector expressions $(F - F^t)e$ and $(F^t - F)e$ are directly related to the difference between neighboring coefficients of the original equation. To see what this means in practice, let us consider a portion of the mesh comprising an interior grid point i and its four neighbors $i - N$, $i - 1$, $i + 1$, and $i + N$ as depicted in Fig. 1. To derive the difference equation at node i, the *box integration* technique proceeds, as indicated by its name, by integrating the leading differential equation (4.1) over the four rectangular boxes labeled NE for northeast, SW for southwest, and so on, and delimited by the perpendicular bisectors of the lines joining point i to its four neighbors, the derivatives being approximated by finite difference quotients (or by using the corresponding boundary condition in the case where the node i is associated with Neumann or mixed boundary condition). This yields (see, e.g., [46, pp. 83–88])

$$(Fe)_i = \begin{cases} -a_{i,i+N} = \chi_F & \text{if the node } i+N \text{ is neither outside } \overline{\Omega} \text{ nor} \\ & \text{associated with Dirichlet boundary condition,} \\ 0 & \text{otherwise,} \end{cases} \tag{4.5}$$

and

$$(F^t e)_i = \begin{cases} -a_{i,i-N} = \chi_{F^t} & \text{if the node } i-N \text{ is neither outside } \overline{\Omega} \text{ nor} \\ & \text{associated with Dirichlet boundary condition,} \\ 0 & \text{otherwise,} \end{cases} \tag{4.6}$$

with

$$\chi_F = \frac{q_{NE}h_E + q_{NW}h_W}{2h_N} \tag{4.7}$$

and

$$\chi_{F^t} = \frac{q_{SE}h_E + q_{SW}h_W}{2h_S}, \tag{4.8}$$

where q_{box} denotes the value of the coefficient $q(x, y)$ in the rectangular subdomain box. If the grid point i is associated with a Neumann or mixed boundary condition, the correct formulas may be obtained by simply dropping illogical contributions in (4.7) and (4.8). It is worthwhile to notice that whatever the case, only the coefficient $q(x, y)$ enters into account.

We shall now examine in turn the analytical bounds derived from both Theorems 3.3 and 3.4, seeking information that can help us to (re)order the unknowns in a beneficial way.

From Theorem 3.3. To get $\mathcal{O}(M)$ upper bound on the spectral condition number with a reasonable leading coefficient, condition (3.13) should be satisfied with the constants c_I far away from 1 [40], [41]. Consider a node $i \in \pi_I$ with $I \in \mathcal{L}$.

Assume first that I is an initial node in $\mathcal{G}(F)/\pi = \mathcal{G}(A)/\pi$. By (3.17), there should hold

$$((F - A)e)_i \leq 0 \tag{4.9}$$

or

$$0 < ((F - A)e)_i \ll ((F + A)e)_i \ . \tag{4.10}$$

In the case where the node i is associated with a Neumann or mixed boundary condition, the latter conditions are in general violated since (for realistic problems) $(Fe)_i = \mathcal{O}(1)$, whereas $(Ae)_i$ is at most of the order of the mesh sizes. If $i - N$ is associated with a Dirichlet boundary condition, given that $(Ae)_i \geq \chi_{F^t}$, it is sufficient in order to meet (4.9) or (4.10) that (respectively)

$$\chi_F - \chi_{F^t} \leq 0 \tag{4.11}$$

or

$$0 < \chi_F - \chi_{F^t} \ll \chi_F + \chi_{F^t} \ , \tag{4.12}$$

which in turn hold whenever (respectively)

$$\frac{q_{NW}}{q_{SW}} \leq \frac{h_N}{h_S} \quad \text{and} \quad \frac{q_{NE}}{q_{SE}} \leq \frac{h_N}{h_S} \tag{4.13}$$

or

$$\frac{q_{NW}}{q_{SW}} > \frac{h_N}{h_S} \quad \text{and} \quad \frac{q_{NE}}{q_{SE}} > \frac{h_N}{h_S}, \tag{4.14}$$

with, however,

$$\mathcal{O}\left(\frac{q_{NW}}{q_{SW}}\right) = \mathcal{O}\left(\frac{h_N}{h_S}\right) = \mathcal{O}\left(\frac{q_{NE}}{q_{SE}}\right) .$$

Assume now that I is not an initial node. By (3.17) again, one should not have $\chi_{F^t} \ll \chi_F$, which leads essentially to (4.13) and (4.14) as sufficient conditions. Properly interpreted, the above discussion may be summarized in the following way.

PROPERTY 1.

(1) *None of the nodes i that belong to block component π_I such that I is an initial node in $G(F)/\pi$ is associated with a Neumann or a mixed boundary condition.*

(2) *The coefficient $q(x, y)$ does not increase (by orders of magnitude) more rapidly than the mesh size in the y-direction, along the increasing (with respect to the selected node ordering) y-direction.*

From Theorem 3.4. The required condition, say (3.18), is equivalent to

$$((F^t - F)e)_i \leq (Ae)_i \qquad \text{for all } i \in \pi_I, \quad I = 1, 2, \dots, M. \tag{4.15}$$

The latter inequality is obviously satisfied when I is an initial node; it does not hold in general if I is a final node and i is associated with a Neumann or mixed boundary condition. In all the remaining cases, the fulfillment of the involved inequality is guaranteed whenever

$$\chi_{F^t} - \chi_F \leq 0 . \tag{4.16}$$

Properly handled, the above considerations amount to the following requirements.

PROPERTY 2.

(1) *None of the nodes i that belong to block component π_I such that I is a final node in $G(F)/\pi$ is associated with Neumann or with mixed boundary condition.*

(2) *The coefficient $q(x, y)$ does not decrease more rapidly than the mesh size in the y-direction, along the increasing y-direction.*

It turns out from numerical experiments (see, e.g., [39] and [40] or the next section) that $\mathcal{O}(M)$ behavior is ensured for the spectral condition number even in the case where inequality (4.15) does not hold at nodes $i \in \pi_I$ for which I is a final node, i.e., nodes i for which $((F^t - F)e)_i = (F^t e)_i > (Ae)_i$, at least whenever there is not any possibility of fill-in outside the π-diagonal part. This bottleneck has to be counted as a shortcoming of the presently available block case analysis, which is less developed than the pointwise one [8]. This leads us to propose the following *empirical* condition.

PROPERTY 3.

(1) *The coefficient $q(x, y)$ does not decrease more rapidly than the mesh size in the y-direction, along the increasing y-direction.*

(2) *There is no possibility of fill-in outside the π-diagonal part.*

Observe that in the case where the mesh size in the y-direction is constant, the second condition in Property 1 and Property 2 reduce, respectively, to

— *The coefficient $q(x, y)$ does not increase (locally) by orders of magnitude along the increasing y-direction.*

— *The coefficient $q(x, y)$ does not decrease along the increasing y-direction.*

It has to be stressed that neither the variation of $q(x, y)$ along the x-direction, nor that of the mesh size in the x-direction, nor that of $p(x, y)$ play any role, which reflects a property of block preconditionings known long from numerical experiments, that their performance is little influenced by strong discontinuities, anisotropy, etc...., compared with pointwise preconditionings ($\pi = p$) where the variation of $p(x, y)$ along the x-direction as well as the variation of the mesh size in the x-direction have to be taken into consideration; in other words, the class of PDE matrices for which unperturbed modified blockwise preconditionings give rise to an $\mathcal{O}(h^{-1})$ spectral condition number is larger (by far) than that for the corresponding pointwise factorizations. *Observe on the other hand that if the line partitioning is organized by grouping together the unknowns in the y-direction, one would deal only with the variation of the coefficient $p(x, y)$ along the x-direction and that of the mesh size in the x-direction.*

Clearly, all the indications provided by the above discussion may be used as guidelines to order the unknowns, passing from a subregion of the involved domain to another only when the variation of the coefficients $p(x, y)$ and/or $q(x, y)$ and that of the mesh size(s) are favorable, in order to ensure nice behavior for the spectral condition number associated with a given unperturbed block method. Another remedy consists in readjusting the mesh sizes so as to satisfy the required condition (see, e.g., (4.13) or (4.14) for Property 1); unfortunately, such a procedure is unrealistic given that it may lead to coarse grids or to excessively fine grids.

All the arguments presented above can readily be extended to deal with other discretization schemes, factorization processes, and boundary value problems, including three-dimensional (3D) cases. For instance, to cope with a 3D problem that involves the elliptic operator

$$Lu = -\frac{\partial}{\partial x}\left(p\frac{\partial}{\partial x}u\right) - \frac{\partial}{\partial y}\left(q\frac{\partial}{\partial y}u\right) - \frac{\partial}{\partial z}\left(r\frac{\partial}{\partial z}u\right) + t\,u, \tag{4.17}$$

where p, q, and r are positive, bounded, and piecewise constant while t is nonnegative, bounded, and piecewise constant, assuming that the seven-point finite difference approximation on a uniform rectangular grid of mesh size h in all the three directions is used and one opts for the *line* partitioning, keeping together the unknowns in the x-direction and ignoring fill-in outside the main block diagonal part, one would deal only with the variation of the coefficients q and r along, respectively, the y- and the z-direction.

5. Examples. We report in this section the results of (some of) the experimental studies we have realized on two-dimensional (2D) problems defined on the unit square (the five-point difference approximation of PDE (4.1) with $\Omega = (0, 1) \times (0, 1)$, on a uniform rectangular grid of mesh size h), in order to illustrate in a quantitative way how following the prescribed recommendations does improve the performance of (unperturbed) modified block methods. We first give in Table 1 the list of the seven ordering strategies we shall need to cope with our test problems. The first three are well known, while the remaining ones are proposed here for the purpose of our tests. In Fig. 2, we illustrate the node numbering of some of them (column Cuthill–McKee may be defined similarly to row Cuthill–McKee) on a 5×6 grid, and we give in Figs. 3–5 the pattern of the corresponding matrix, with the exception of the column-wise (resp., reverse column Cuthill–McKee) ordering whose matrix resembles that of the lexicogaphic (resp., reverse row Cuthill–McKee) ordering on a 6×5 grid. Observe that (reverse) row or column Cuthill–McKee orderings substantially reduce the maximal increasing length of $\mathcal{G}(A)$ with respect to the line (row-wise or columnwise) partitioning, which could benefit the spectral condition number. The philosophy of row (column) Cuthill–McKee orderings is simple. The domain is divided into two parts (not necessarily equal) by an invisible horizontal (vertical) line; one then starts from the "middle" of the domain, numbering nodes by rows (columns) alternatively and progressing towards the bottom and the top (the left and the right) boundary of the domain. In this respect, the procedure may be viewed as an extension to the quotient graph $\mathcal{G}(A)/\pi$ (π = *row* or *column*) of Notay's ordering algorithm ([8, p. 680])[2], which in turn generalizes the (reverse) Cuthill–McKee ordering [19], [30][3].

[2] R. Beauwens, Private communication.

[3] Y. Notay, Private communication.

TABLE 1
List of ordering strategies.

Ordering	Abbreviated form
Lexicographic	lexico
Reverse lexicographic	revlexico
Column-wise†	column
Row Cuthill–McKee	rowcm
Reverse row Cuthill–McKee	revrowcm
Column Cuthill–McKee†	colcm
Reverse column Cuthill–McKee†	revcolcm

† Here *lines* are obtained by grouping together the unknowns along the y-direction.

Of course, our list is not exhaustive; any other ordering that meets the required conditions is suitable. Unless specifically stated otherwise, we shall deal with MBILU factorizations associated with $x = e$, $\beta = \text{tridiag}_p(\varepsilon)$, and $\gamma = \text{tridiag}_p(\varepsilon)$. All our examples are particular cases of PDE (4.1) with $t(x, y) = 0$, $\Omega = (0, 1) \times (0, 1)$ and $\Gamma_2 = \emptyset$; for each of which, we specify which ordering strategies are theoretically favorable as well as the property according to which the choice is made. For illustration purposes, the first example is discussed in more detail. Although the right-hand side $f(x, y)$ of PDE (4.1) is irrelevant as far as the choice of ordering is concerned, we give it for later use in numerical experiments.

EXAMPLE 1 (from [58]).

- $\Gamma_0 = \{(x, y);\ 0 \le x \le 1,\ y = 0\}$.
- $\Gamma_1 = \Gamma \backslash \Gamma_0$.
- $p(x, y) = q(x, y) = a(x, y)$ where
$$a(x, y) = \begin{cases} 100 & \text{in } (1/4, 3/4) \times (1/4, 3/4), \\ 1 & \text{elsewhere.} \end{cases}$$
- $$f(x, y) = \begin{cases} 100 & \text{in } (1/4, 3/4) \times (1/4, 3/4), \\ 0 & \text{elsewhere.} \end{cases}$$
- $g_0(x, y) = g_1(x, y) = 0$.

The second condition of Property 1 (Property 2), i.e., here the coefficient $a(x, y)$ does not increase (decrease) along the increasing line numbering direction, is satisfied by both rowcm and colcm orderings. Unfortunately, the first requirement is violated given that (most of) the "middle" nodes, which belong to the first (last) block component or line (see Fig. 2), behave as if they were associated with a Neumann boundary condition, i.e., the row sum of the system matrix A is zero at the involved nodes. The reverse row Cuthill–McKee and the reverse column Cuthill–McKee orderings satisfy all the requirements of Property 3, but preference should be given to the first one because of the presence of Dirichlet boundary conditions at the bottom boundary, which implies a stronger diagonal dominance in $\text{diag}_\pi(A)$ at the first block component. We know from [17] and [20] that the stronger the diagonal dominance in $\text{diag}_\pi(A)$ the faster the (exponential) decay of the elements of P^{-1} from the diagonal.

EXAMPLE 2.

- $\Gamma_0 = \{(x, y);\ 0 \le x \le 1,\ y = 0\}$.
- $\Gamma_1 = \Gamma \backslash \Gamma_0$.
- $p(x, y) = q(x, y) = a(x, y)$ where
$$a(x, y) = \begin{cases} 0.001 & \text{in } (1/4, 3/4) \times (1/4, 3/4), \\ 1 & \text{elsewhere.} \end{cases}$$
- $$f(x, y) = \begin{cases} 10 & \text{in } (1/4, 3/4) \times (1/4, 3/4), \\ 0 & \text{elsewhere.} \end{cases}$$
- $g_0(x, y) = g_1(x, y) = 0$.

26	27	28	29	30
21	22	23	24	25
16	17	18	19	20
11	12	13	14	15
6	7	8	9	10
1	2	3	4	5

lexicographic

5	4	3	2	1
10	9	8	7	6
15	14	13	12	11
20	19	18	17	16
25	24	23	22	21
30	29	28	27	26

reverse lexicographic

6	12	18	24	30
5	11	17	23	29
4	10	16	22	28
3	9	15	21	27
2	8	14	20	26
1	7	13	19	25

column-wise

6	7	8	9	10
16	17	18	19	20
26	27	28	29	30
21	22	23	24	25
11	12	13	14	15
1	2	3	4	5

reverse row CM

21	22	23	24	25
11	12	13	14	15
1	2	3	4	5
6	7	8	9	10
16	17	18	19	20
26	27	28	29	30

row CM

6	18	30	24	12
5	17	29	23	11
4	16	28	22	10
3	15	27	21	9
2	14	26	20	8
1	13	25	19	7

reverse column CM

FIG. 2. *Numbering of nodes on a* 5×6 *grid for different ordering strategies. CM stands for Cuthill–McKee.*

None of the proposed orderings is theoretically indicated (they do not satisfy any of the three properties). Nevertheless, numerical results to come reveal satisfactory behavior with the lexicographic ordering. For this particular example, it is possible to use some sort of an unbalanced reverse row Cuthill–McKee ordering with $\{(x, y);\ 0 \leq x \leq 1,\ y = 3/4\}$ as *dividing line* and last block component (i.e., associated with π_M). Observe that, in so doing, the bottom part meets Property 1, while in the top one the PDE coefficient $a(x, y)$ is constant. The performances are, however, essentially the same as with the lexicographic ordering. Most severe is the fact that in the absence of a Dirichlet boundary condition such a procedure is no more interesting.

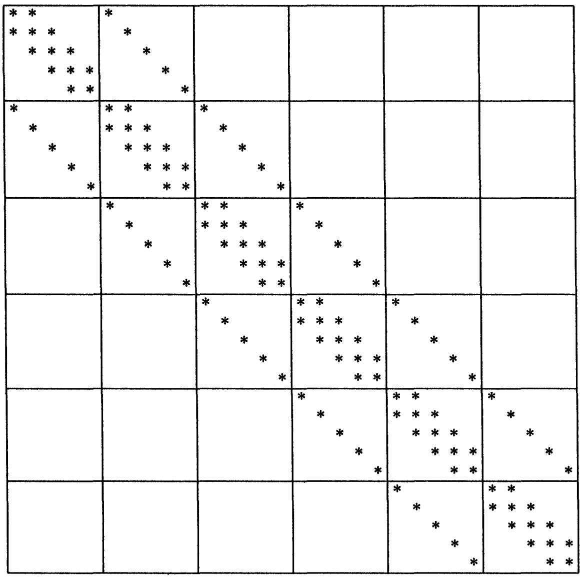

FIG. 3. *Band and block structure of the lexicographic matrix on a* 5 × 6 *grid.*

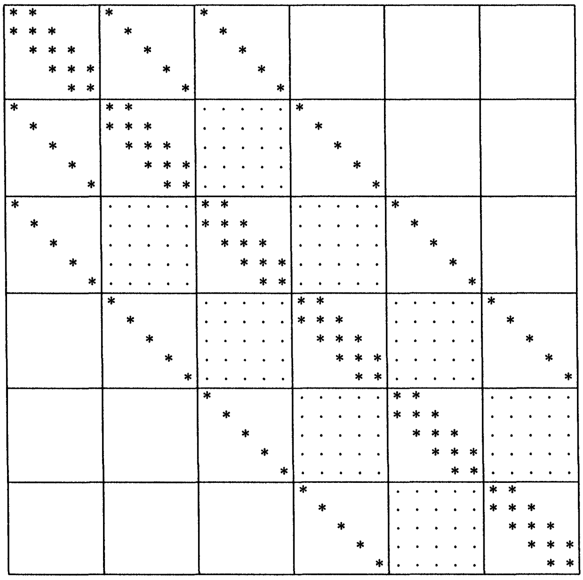

FIG. 4. *Band and block structure of the row Cuthill–McKee matrix on a* 5 × 6 *grid; symbols "." indicate where there is a possibility of fill-in outside the main block diagonal part.*

EXAMPLE 3.

- $\Gamma_0 = \{(x, y);\ 0 \le x \le 1,\ y = 0\} \bigcup \{(x, y);\ 0 \le x \le 1,\ y = 1\}$.
- $\Gamma_1 = \Gamma \backslash \Gamma_0$.
- $p(x, y) = q(x, y) = a(x, y)$ where

$$a(x, y) = \begin{cases} 0.001 & \text{in } (1/4, 3/4) \times (1/4, 3/4), \\ 1 & \text{elsewhere.} \end{cases}$$

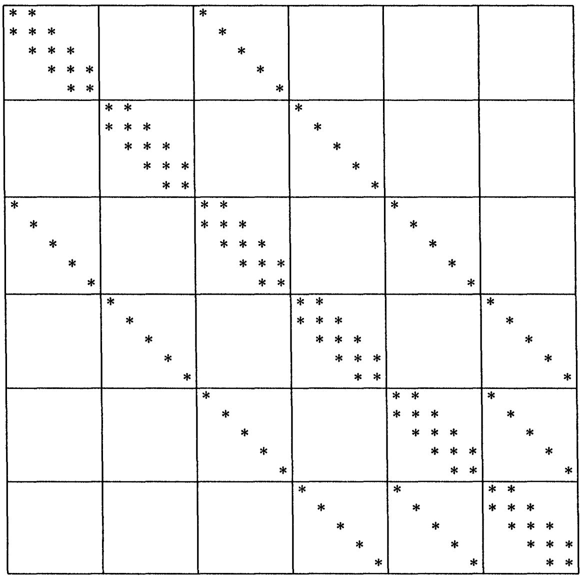

FIG. 5. *Band and block structure of the reverse row Cuthill–McKee matrix on a* 5 × 6 *grid.*

- $$f(x,y)=\begin{cases}0.1 & \text{in } (1/4,3/4)\times(1/4,3/4),\\ 0 & \text{elsewhere.}\end{cases}$$
- $g_0(x,y)=g_1(x,y)=0$.

With the row Cuthill–McKee ordering, Property 2 and condition (8.2) of Theorem 3.1 are satisfied, without the need of any fill-in outside the main block diagonal part. The reverse row Cuthill–McKee ordering should be preferred because it meets Property 1 and no fill-in has to be neglected outside the main block diagonal part to avoid increasing the computational complexity.

EXAMPLE 4.

- $\Gamma_0=\emptyset$.
- $\Gamma_1=\Gamma$.
- $p(x,y)=q(x,y)=a(x,y)$ where

$$a(x,y)=\begin{cases}1 & \text{in } (0,3/4)\times(0,3/4),\\ 0.001 & \text{elsewhere.}\end{cases}$$

- The right-hand side of the (singular) system to solve is chosen such that the function $u_0(x,y)=x(1-x)y(1-y)e^{xy}$ generates a solution on the grid.

The reverse lexicographic, the reverse row Cuthill–McKee, and the reverse column Cuthill–McKee orderings are theoretically favorable by Property 3, but, as mentioned earlier the last two, whose performances are here identical, have the advantage of reducing the maximal increasing length of $\mathcal{G}(F)/\pi$.

EXAMPLE 5.

- $\Gamma_0=\emptyset$.
- $\Gamma_1=\Gamma$.
- $p(x,y)=q(x,y)=a(x,y)$ where

$$a(x,y)=\begin{cases}0.001 & \text{in } (0,1)\times(1/4,3/4),\\ 1 & \text{elsewhere.}\end{cases}$$

- The right-hand side of the (singular) system to solve is chosen such that the function $u_0(x, y) = x(1-x)y(1-y)e^{xy}$ generates a solution on the grid.

As the coefficient $a(x, y)$ is constant along the x-direction whereas it presents both strong decreasing and increasing slopes along the y-direction, the unknowns should be grouped together along the y-direction. Therefore, by Property 3, the columnwise, and preferably the reverse column Cuthill–McKee (to avoid neglecting fill-in outside the main block diagonal part), ordering should be used. This is an example for which there is no natural (row-wise or columnwise) ordering that is favorable to *pointwise* modified incomplete factorizations (MILU(0)). Indeed, whatever the choice, there persists one direction along which the variation of $a(x, y)$ is not beneficial; numerical calculation yields $\kappa(B^+A) \approx 0.8h^{-1.96}$ for both above-mentioned natural orderings.

EXAMPLE 6.

- $\Gamma_0 = \emptyset$.
- $\Gamma_1 = \Gamma$.
- We assume the domain to be subdivided as depicted in Fig. 6 where the coefficients p and q are also specified. This is essentially Stone's problem [53] where, only for ease of implementation, we have discarded the hole inside Ω. We point out that the difficulty does not lie on the presence of the hole but on the (opposite) presentation of anisotropy in Ω_2 and Ω_3.
- The right-hand side of the (singular) system to solve is chosen such that the function $u_0(x, y) = x(1-x)y(1-y)e^{xy}$ generates a solution on the grid.

region	p	q
Ω_1	1	1
Ω_2	1	10^3
Ω_3	10^3	1

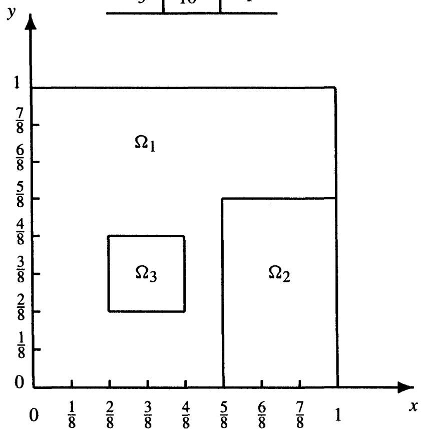

FIG. 6. *Example 6. Specification of the coefficients of the partial differential equation.*

By Property 3, the reverse lexicographic, the reverse row Cuthill–McKee, and the reverse column Cuthill–McKee orderings are suitable for a satisfactory behavior of the spectral condition number. However, the numerical results (to come) indicate that the spectral condition number associated with the reverse lexicographic ordering is, at least for considered values of mesh size, a bit larger than for the lexicographic ordering, and that this misrepresents the rate of convergence measured in number of PCG iterations, the clustering of intermediate eigenvalues being more accentuated for the reverse lexicographic ordering.

We gather in Tables 2–7 the results of our numerical experiments, including the spectral condition number and the number of PCG iterations needed to reduce an initial residual by a prescribed amount. In the case of fill-in outside the main block diagonal part, only the main diagonal of the involved submatrices (see Fig. 4) has been retained. The reason why the results related to the row Cuthill–McKee and the column Cuthill–McKee orderings are sometimes not reported is that, when they give rise to singular preconditioners for nonsingular systems, adding *one diagonal of fill-in* outside the main block diagonal part results in quasi-singular preconditioners. In the singular case, the outcome is a singular preconditioner whose block diagonal factor P includes both singular and quasi-singular nonzero submatrices. Moreover, one loses the relation (4.4) and therefore any control on the key requirements. We believe that the best thing would be to avoid these orderings, e.g., by shifting to a dynamically perturbed method [37], [39], [40], [42]. The computations were carried out in double precision Fortran on a CDC 4680 computer. All the theoretical predictions are corroborated. Very often, an important reduction of the spectral condition number is observed, which reflects on the number of PCG iterations.

TABLE 2

Example 1. *Spectral condition number $\kappa(B^{-1}A)$ associated with MBILU for different ordering strategies. Exponent μ corresponds to the (estimated) asymptotic relationship $\kappa(B^{-1}A) = Ch^{-\mu}$; C denotes a constant and, for $h^{-1} = 192$, the number of PCG iterations to achieve $\|r^{(i)}\|_2 \,/\, \|r^{(0)}\|_2 \le \epsilon$. The symbol "$*$" means theoretical favorable ordering; by "$-$" and "$+$" we understand, respectively, that the preconditioner is singular and that one diagonal fill-in has been accepted outside the main block diagonal part.*

h^{-1} **Ordering**	12	24	48	96	192	μ
			$\kappa(B^{-1}A)$			
lexico	4.29	13.38	51.28	150.4	456.1	1.60
revlexico	3.83	12.81	55.80	176.5	552	1.64
column	4.72	22.79	237.9	2511	10361	2.04
rowcm	-	-	-	-	-	-
colcm	571.3	2718	20901	153500	793210	2.37
colcm$^+$	109	1384	10812	79267	408910	2.37
revrowcm*	2.83	6.40	14.3	29.9	58.3	0.96
revcolcm*	3.54	9.12	22.99	54.41	121.4	1.16
			Number of PCG iterations for $h^{-1} = 192$.			
ϵ	10^{-3}	10^{-5}	10^{-7}	10^{-9}		
lexico	35	50	65	79		
revlexico	35	52	66	83		
column	65	90	114	137		
rowcm	-	-	-	-		
colcm	120	170	>200	>200		
colcm$^+$	126	175	>200	>200		
revrowcm*	24	34	45	56		
revcolcm*	33	46	58	71		

6. Concluding remarks and further perspectives. In agreement with previous studies, our investigations show that the ordering of the unknowns has a strong effect on the rate of convergence of PCG methods. It also emerges from our analysis that, for discrete PDEs, there is no general "good" ordering strategy in the sense of "optimizing" the rate of convergence (the variation of the PDE coefficients as well as that of the mesh size(s) there playing the key role) which means that each problem requires a customized treatment. This also implies that orderings which are elaborated only from implementation considerations (like parallelism or vectorization) can hardly be competitive in a general way on presently available computers. Our arguments, which bring to light the appreciable potentialities of modified incomplete factorizations, also display why the behavior of block methods is less dependent on the above-mentioned variations than that of pointwise methods.

TABLE 3

Example 2. Spectral condition number $\kappa(B^{-1}A)$ associated with MBILU for different ordering strategies. Exponent μ corresponds to the (estimated) asymptotic relationship $\kappa(B^{-1}A) = Ch^{-\mu}$; C denotes a constant and, for $h^{-1} = 192$, the number of PCG iterations to achieve $\|r^{(i)}\|_2 \,/\, \|r^{(0)}\|_2 \le \epsilon$. By "−" and "+" we understand, respectively, that the preconditioner is singular and that one diagonal fill-in has been accepted outside the main block diagonal part.

h^{-1} **Ordering**	12	24	48	96	192	μ
			$\kappa(B^{-1}A)$			
lexico	2.50	5.26	11.55	26.27	59.33	1.18
revlexico	3.02	8.38	34.06	228.3	1220	2.42
column	4.02	10.89	30.18	206.0	1237	2.59
rowcm	-	-	-	-	-	-
colcm	27.37	115.0	502.9	2202	9535	2.11
colcm$^+$	14.17	60.10	261.5	1140	4921	2.11
revrowcm	3.20	8.60	34.39	228.9	1219	2.41
revcolcm	5.76	21.94	165.1	4547	63028	3.79
		Number of PCG iterations for $h^{-1} = 192$.				
ϵ	10^{-3}	10^{-5}	10^{-7}	10^{-9}		
lexico	24	34	45	55		
revlexico	39	54	66	84		
column	55	76	94	114		
rowcm	-	-	-	-		
colcm	46	67	88	108		
colcm$^+$	52	72	92	111		
revrowcm	37	52	67	83		
revcolcm	81	110	140	169		

TABLE 4

Example 3. Spectral condition number $\kappa(B^{-1}A)$ associated with MBILU for different ordering strategies. Exponent μ corresponds to the (estimated) asymptotic relationship $\kappa(B^{-1}A) = Ch^{-\mu}$; C denotes a constant and, for $h^{-1} = 192$, the number of PCG iterations to achieve $\|r^{(i)}\|_2 \,/\, \|r^{(0)}\|_2 \le \epsilon$. The symbol "$*$" means theoretical favorable ordering; by "+" we understand that one diagonal fill-in has been accepted outside the main block diagonal part.

h^{-1} **Ordering**	12	24	48	96	192	μ
			$\kappa(B^{-1}A)$			
lexico	1.30	2.17	4.14	7.39	15.39	1.06
revlexico	1.30	2.17	4.14	7.39	15.39	1.06
column	1.91	4.64	17.41	185.7	1234	2.73
rowcm*	5.18	10.78	22.27	45.65	92.93	1.03
rowcm^{*+}	2.65	5.51	11.54	23.55	47.99	1.03
colcm	2.26	4.90	14.35	58.99	322.9	2.45
colcm$^+$	1.75	3.91	11.07	38.07	191.4	2.33
revrowcm*	1.34	2.32	4.40	7.68	16.27	1.08
revcolcm	1.91	4.72	20.57	546.0	17405	4.99
		Number of PCG iterations for $h^{-1} = 192$.				
ϵ	10^{-3}	10^{-5}	10^{-7}	10^{-9}		
lexico	17	23	30	37		
revlexico	17	23	30	37		
column	45	61	79	97		
rowcm*	22	33	43	54		
rowcm^{*+}	23	32	42	52		
colcm	34	50	66	81		
colcm$^+$	37	52	67	81		
revrowcm*	17	23	30	37		
revcolcm	61	81	106	128		

TABLE 5

Example 4. Spectral condition number $\kappa(B^+A)$ associated with MBILU for different ordering strategies. Exponent μ corresponds to the (estimated) asymptotic relationship $\kappa(B^+A) = Ch^{-\mu}$; C denotes a constant and, for $h^{-1} = 192$, the number of PCG iterations to achieve $\|r^{(i)}\|_2 \,/\, \|r^{(0)}\|_2 \le \epsilon$. The symbol "$*$" means theoretical favorable ordering.

h^{-1} **Ordering**	12	24	48	96	192	μ
			$\kappa(B^+A)$			
lexico	18.98	103.6	1271	17587	107670	2.61
column	18.98	103.6	1271	17587	107670	2.61
*revlexico**	6.55	17.38	42.28	97.26	230.3	1.24
*revrowcm**	4.56	10.63	24.35	57.14	125.8	1.14
*revcolcm**	4.56	10.63	24.35	57.14	125.8	1.14
	Number of PCG iterations for $h^{-1} = 192$.					
ϵ	10^{-3}	10^{-5}	10^{-7}	10^{-9}		
lexico	60	94	126	156		
column	60	94	126	156		
*revlexico**	31	48	64	80		
*revrowcm**	28	42	57	71		
*revcolcm**	28	42	57	71		

TABLE 6

Example 5. Spectral condition number $\kappa(B^+A)$ associated with MBILU for different ordering strategies. Exponent μ corresponds to the (estimated) asymptotic relationship $\kappa(B^+A) = Ch^{-\mu}$; C denotes a constant and, for $h^{-1} = 192$, the number of PCG iterations to achieve $\|r^{(i)}\|_2 \,/\, \|r^{(0)}\|_2 \le \epsilon$. The symbol "$*$" means theoretical favorable ordering.

h^{-1} **Ordering**	12	24	48	96	192	μ
			$\kappa(B^+A)$			
lexico	5.32	17.65	90.56	408.9	1451	1.83
*column**	18.16	36.04	71.78	143.3	286.2	1.00
revrowcm	21.98	119.8	1247	11561	49756	2.11
*revcolcm**	9.09	18.03	35.90	71.63	143.1	1.00
	Number of PCG iterations for $h^{-1} = 192$.					
ϵ	10^{-3}	10^{-5}	10^{-7}	10^{-9}		
lexico	48	72	99	122		
*column**	33	50	66	82		
revrowcm	84	131	177	225		
*revcolcm**	29	44	60	74		

In practice, depending on the problem at hand, applying our recommendations could require several changes of the direction along which the unknowns are grouped together, which increases the computational complexity. So, compromise should sometimes be made in order to balance things, taking advantage as much as possible of the potential efficiency of modified methods by avoiding as far as possible (the strongest) unfavorable variations of the coefficients and/or mesh sizes, remaining at the same time relatively close to the *simplicity* of natural (row-wise or columnwise) orderings. One way to achieve such a compromise would be to first fix, for each as vast as possible region of the involved domain, the direction along which the unknowns will be gathered together in such a way that (most of) the strongest variations of the coefficients and/or mesh sizes are *absorbed* by the *lines* that determine the partitioning (we recall that variations within *diagonal blocks* or *lines* do not affect the order of magnitude of the preconditioning), and next to move by *lines* so as to neutralize (most of) the remaining unfavorable variations, which is the highest priority. This idea is based on the belief that partially reordering is better than not reordering at all. It would be advisable in the case of a 3D problem to move by planes, treating each of them as in the 2D case.

TABLE 7

Example 6. Spectral condition number $\kappa(B^+A)$ associated with MBILU for different ordering strategies. Exponent μ corresponds to the (estimated) asymptotic relationship $\kappa(B^+A) = Ch^{-\mu}$; C denotes a constant and, for $h^{-1} = 192$, the number of PCG iterations to achieve $\|r^{(i)}\|_2 \,/\, \|r^{(0)}\|_2 \leq \epsilon$. The symbol "$*$" means theoretical favorable ordering.

h^{-1} **Ordering**	24	48	96	192	μ
		$\kappa(B^+A)$			
lexico	37.92	94.00	231.0	599.3	1.38
column	16.76	43.39	104.8	237.9	1.18
*revlexico**	36.04	91.60	234.3	620.9	1.41
*revrowcm**	10.14	22.79	51.15	114.5	1.16
*revcolcm**	19.20	51.17	144	435	1.59
	Number of PCG iterations for $h^{-1} = 192$.				
ϵ	10^{-3}	10^{-5}	10^{-7}	10^{-9}	
lexico	24	46	64	84	
column	20	33	46	58	
*revlexico**	19	39	52	69	
*revrowcm**	14	28	39	51	
*revcolcm**	26	42	58	76	

We have illustrated our technique on finite difference grids for which *natural* partitionings are available. This raises the question of how to handle finite element unstructured grids applied on irregular domains. Another point that deserves attention is the use of reverse row or column Cuthill–McKee orderings combined with *perturbed* modified block incomplete factorization methods [37], [42] (to avoid loss of efficiency) in a parallel environment; the associated matrices (see Fig. 4) may also be viewed as partitioned in some sort of two-line partitioning, and most of the computations relative to the preconditioning may be performed by keeping two (arrays of) processors busy during both the factorization process and the solution of the preconditioning system. Moreover, the idea may be generalized to obtain an *essentially n-line* partitioning [43], [44].

Acknowledgments. The author thanks the anonymous referees whose careful reading resulted in an improvement in the presentation of this paper.

REFERENCES

[1] C.C. ASHCRAFT AND R.G. GRIMES, *On vectorizing incomplete factorization and SSOR preconditioners*, SIAM J. Sci. Statist. Comput., 9 (1988), pp. 122–151.

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

[3] O. AXELSSON AND V. EIJKHOUT, *Vectorizable preconditioners for elliptic difference equations in three space dimensions*, J. Comput. Appl. Math., 27 (1989), pp. 299–321.

[4] O. AXELSSON AND G. LINDSKOG, *On the eigenvalue distribution of a class of preconditioning methods*, Numer. Math., 48 (1986), pp. 479–498.

[5] ———, *On the rate of convergence of the preconditioned conjugate gradient method*, Numer. Math., 48 (1986), pp. 499–523.

[6] E.F. D'AZEVEDO, P.A. FORSYTH, AND WEI-PAI-TANG, *Two variants of minimum discarded fill ordering*, in Iterative Methods in Linear Algebra, R. Beauwens and P. de Groen, eds., North-Holland, Amsterdam, 1992, pp. 603–612.

[7] R. BEAUWENS, *Upper eigenvalue bounds for pencils of matrices*, Linear Algebra Appl., 62 (1984), pp. 87–104.

[8] ———, *Approximate factorizations with S/P consistently ordered M-factors*, BIT, 29 (1989), pp. 658–681.

[9] ———, *Approximate factorizations with modified S/P consistently ordered M-factors*, 1991, Numer. Linear Algebra Appl., 1 (1994), pp. 3–17.

[10] R. BEAUWENS AND M. BEN BOUZID, *Existence and conditioning properties of sparse approximate block factorizations*, SIAM J. Numer. Anal., 25 (1988), pp. 941–956.

[11] R. Beauwens and R. Wilmet, *Conditioning analysis of positive definite matrices by approximate factorizations*, J. Comput. Appl. Math., 26 (1989), pp. 257–269.

[12] A. Ben Israel and T.N.E. Greville, *Generalized Inverses: Theory and Applications*, John Wiley, New York, 1974.

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

[14] T.F. Chan, *Fourier analysis of relaxed incomplete factorization preconditioners*, SIAM J. Sci. Statist. Comput., 12 (1991), pp. 668–680.

[15] T.F. Chan and H.C. Elman, *Fourier analysis of iterative methods for elliptic problems*, SIAM Rev., 31 (1989), pp. 20–49.

[16] T.F. Chan and G.A. Meurant, *Fourier analysis of block preconditioners*, CAM Report 90-04, University of California, Los Angeles, February 1990.

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

[18] P. Concus, G.H. Golub, and D.P. O'Leary, *A generalized conjugate gradient method for the numerical solution of elliptic partial differential equations*, in Sparse Matrix Computations, J. Bunch and D. Rose, eds., Academic Press, New York, 1976, pp. 309–332.

[19] E. Cuthill and J. McKee, *Reducing the bandwidth of sparse symmetric matrices*, in Proceedings of the 24th National Conference of the Association for Computing Machinery, Brandon Press, New Jersey, pp. 157–172.

[20] S. Demko, W.F. Moss, and P.W. Smith, *Decay rates for inverses of band matrices*, Math. Comp., 43 (1984), pp. 491–499.

[21] S. Doi, *On parallelism and convergence of incomplete LU factorizations*, Appl. Numer. Math., 7 (1991), pp. 417–436.

[22] ———, *A Gustafsson-type modification for parallel ordered incomplete LU factorizations*, in Advances in Numerical Methods for Large Sparse Sets of Linear Systems, 7, T. Nodera, ed., Keio University, Japan, 1991.

[23] S. Doi and A. Lichnewsky, *Some parallel and vector implementations of preconditioned iterative methods on Cray*-2, Internat. J. High Speed Comput., 2 (1990), pp. 143–179.

[24] ———, *A graph-theory approach for analysing the effects of ordering on ILU preconditioning*, Res. Report No. 1452, INRIA, France, June 1991.

[25] J.M. Donato and T.F. Chan, *Fourier analysis of incomplete factorization preconditioners for three-dimensional anisotropic problems*, SIAM J. Sci. Statist. Comput., 13 (1992), pp. 319–338.

[26] I.S. Duff and G.A. Meurant, *The effect of ordering on preconditioned conjugate gradients*, BIT, 29 (1989), pp. 635–657.

[27] V. Eijkhout, *Beware of unperturbed modified incomplete factorizations*, in Iterative Methods in Linear Algebra, R. Beauwens and P. de Groen, eds., North-Holland, Amsterdam, 1992, pp. 583–591.

[28] H. Elman and E. Agrón, *Ordering techniques for the preconditioned conjugate gradient method on parallel computers*, Comput. Phys. Comm., 53 (1989), pp. 253–269.

[29] S. Fujino and S. Doi, *Optimizing multicolor ICCG methods on some vectorcomputers*, in Iterative Methods in Linear Algebra, R. Beauwens and P. de Groen, eds., North-Holland, Amsterdam, 1992, pp. 349–358.

[30] A. George and J.L. Liu, *Computer Solution of Large Sparse Positive Definite Systems*, Prentice-Hall, Englewood Cliffs, NJ, 1981.

[31] G.H. Golub and C.F. Van Loan, *Matrix Computations*, 2nd ed., The John Hopkins University Press, Baltimore, MD, 1989.

[32] F. Harary, *Graph Theory*, Addison-Wesley, Reading, MA, 1969.

[33] E.F. Kaasschieter, *Preconditioned conjugate gradients for solving singular systems*, J. Comp. Appl. Math., 24 (1988), pp. 265–275.

[34] S. Kaniel, *Estimates for some computational techniques in linear algebra*, Math. Comp., 20 (1966), pp. 369–378.

[35] C.-C.J. Kuo and T.F. Chan, *Two-color Fourier analysis of iterative algorithms for elliptic problems with red/black ordering*, SIAM J. Sci. Statist. Comput., 11 (1990), pp. 767–793.

[36] M.M. Magolu, *Conditioning analysis of sparse block approximate factorizations*, Appl. Numer. Math., 8 (1991), pp. 25-42.

[37] ———, *Modified block-approximate factorization strategies*, Numer. Math., 61 (1992), pp. 91–110.

[38] ———, *Lower eigenvalue bounds for singular pencils of matrices*, J. Comput. Appl. Math., 39 (1992), pp. 329–351.

[39] ———, *Sparse block approximate factorizations for singular problems*, in Iterative Methods in Linear Algebra, R. Beauwens and P. de Groen, eds., North-Holland, Amsterdam, 1992, pp. 519–529.

[40] ———, *Sparse approximate block factorizations for solving symmetric positive (semi)definite linear systems*, Ph.D. Thesis, Service de Métrologie Nucléaire, Université Libre de Bruxelles, Brussels, Belgium, 1992.

[41] M.M. MAGOLU, *Analytical bounds for block approximate factorization methods*, Linear Algebra Appl., 179 (1993), pp. 33–57.

[42] ———, *Empirically modified block incomplete factorizations*, in Incomplete Decomposition (ILU)-Algorithms, Theory and Applications, W. Hackbush and G. Wittum, eds., Notes on Numerical Fluid Mechanics, Vol. 41, Vieweg, Braunschweig, 1993, pp. 78–87.

[43] ———, *Implementation of parallel block preconditionings on a transputer-based multiprocessor*, Report No. IT/IF/14-10, Université Libre de Bruxelles, July 1993.

[44] ———, *Ordering strategies for parallelizable block preconditionings*, in preparation.

[45] M.M. MAGOLU AND Y. NOTAY, *On the conditioning analysis of block approximate factorization methods*, Linear Algebra Appl., 154-156 (1991), pp. 583–599.

[46] S. NAKAMURA, *Computational Methods in Engineering and Science*, John Wiley, New York, 1977.

[47] Y. NOTAY, *Solving positive (semi)definite linear systems by preconditioned iterative methods*, in Preconditioned Conjugate Gradient Methods, O. Axelsson and L. Kolotilina, eds., Lecture Notes in Mathematics 1457, Springer-Verlag, Berlin, 1990, pp. 105–125.

[48] ———, *Conditioning analysis of modified block incomplete factorizations*, Linear Algebra Appl., 154-156 (1991), pp. 711–722.

[49] ———, *On the convergence rate of the conjugate gradients in the presence of rounding errors*, Numer. Math., 65 (1993), pp. 301–317.

[50] D.P. O'LEARY, *Ordering schemes for parallel processing of certain mesh problems*, SIAM J. Sci. Statist. Comput., 5 (1984), pp. 620–632.

[51] J.M. ORTEGA, *Orderings for conjugate gradient preconditionings*, SIAM J. Optim., 1 (1991), pp. 565–582.

[52] E.L. POOLE AND J.M. ORTEGA, *Multicolor ICCG methods for vector computers*, SIAM J. Numer. Anal., 24 (1987), pp. 1394–1418.

[53] H.L. STONE, *Iterative solution of implicit approximation of multidimensional partial differential equations*, SIAM J. Numer. Anal., 5 (1968), pp. 530–558.

[54] A. VAN DER SLUIS, *The convergence behaviour of conjugate gradients and ritz values in various circumstances*, in Iterative Methods in Linear Algebra, R. Beauwens and P. de Groen, eds., North-Holland, Amsterdam, 1992, pp 49–66.

[55] A. VAN DER SLUIS AND H.A. VAN DER VORST, *The rate of convergence of conjugate gradients*, Numer. Math., 48 (1986), pp. 543–560.

[56] H.A. VAN DER VORST, *High performance preconditioning*, SIAM J. Sci. Statist. Comput., 10 (1989), pp. 1174–1185.

[57] ———, *The convergence behavior of some iterative methods*, in Proceedings of the Fifth International Symposium on Numerical Methods in Engineering 1, R. Gruber, J. Periaux, and R. Shaw, eds., Springer-Verlag, Berlin, 1989, pp. 61–72.

[58] ———, *The convergence behaviour of preconditioned CG and CG-S*, in Preconditioned Conjugate Gradient Methods, O. Axelsson and L. Kolotilina, eds., Lecture Notes in Mathematics 1457, Springer-Verlag, Berlin, 1990, pp. 126–136.

[59] R.S. VARGA, *Matrix Iterative Analysis*, Prentice Hall, Englewood Cliffs, NJ, 1962.

[60] E.L. WACHSPRESS, *Iterative Solution of Elliptic Systems and Application to the Neutron Diffusion Equations of Reactor Physics*, Prentice Hall, Englewood Cliffs, NJ, 1966.