Algebraic Multilevel Incomplete Factorization methods for five-point matrices

M.R. Larin *

Computing Center SD RAS, Novosibirsk, RUSSIA

November 29, 1996

Abstract

In the paper [11] an iterative multilevel incomplete factorization method for five-point difference matrices has been proposed. In the present paper the general formulation of the method is given and the quality of the preconditioning matrix is improved by shifted and scaled Chebyshev matrix polynomials. In conclusion experimental results on standard test problems, which have confirmed the theory, are presented and discussed.

Keywords: Algebraic MultiLevel Iteration method, Iterative Incomplete Factorization method, Preconditioned Conjugate Gradient method, Chebyshev polynomials, generalized eigenvalue problem.

1 Introduction

This work concerns the solution of the linear system of equations

$$Ax = b, \tag{1.1}$$

*This work is a part of NWO Netherlands-Russian programme, dossiernr. 047.003.017, and partly supported by the Russian Foundation for the Basic Research, (96-01-01770).

where A is a sparse symmetric positive definite M-matrix or a Stieltjes matrix of order N. To solve the system (1.1) the preconditioned conjugate gradient (PCG) method is widely used.

Recently a main topic of many papers has been to get an optimal order preconditioner for the solution of the system (1.1), for which the rate of convergence of a preconditioned iterative method does not depend on N, and the total computational complexity is proportional to N. In particular, algebraic multilevel [1, 2, 4, 5, 7] and multigrid [9, 12, 13] methods allow us to construct preconditioners with these properties.

In the paper [11] an iterative multilevel incomplete factorization method for five-point difference matrices has been proposed. Its main difference from the earlier, suggested by Axelsson and Neytcheva [5] is in that instead of using an approximation of the first pivoting block for obtaining a new matrix on the lower level as its Schur's complement we have proposed to use the iterative incomplete factorization method [10] to construct an approximation of the Schur's complement as a new matrix on the lower level, which has the structure similar to that of the original matrix. However, the rate of convergence of the method is not optimal since the condition number of $M^{-1}A$ is the magnitude of $O(\sqrt{N})$.

In the present paper the general formulation of the method is given, and the quality of the preconditioning matrix is improved by shifted and scaled Chebyshev matrix polynomials in order to gain an optimal order of the rate of convergence saving the optimal computational complexity.

The paper is organized as follows. In Section 2 the algorithm of constructing the preconditioning matrix M with the proof of its computational correctness is described. Some conditions of attaining an optimal order of computational complexity and an optimal rate of convergence are derived in Sections 3 and 4, respectively. In the final section of the paper experemental results on standard test problems are presented.

2 Construction of the preconditioning matrix

To construct a multilevel preconditioning matrix M we usually have to define a sequence of matrices $A^{(k)}$, $k = 0, 1, \ldots, L-1, L$ of order n_k, each of which is an approximation of the Schur complement of the previous one, starting with $A^{(0)} = A$.

Let us consider a sequence of nested sets of nodes $\{X_k\}$ corresponding to the sequence of matrices $\{A^{(k)}\}$ such that

$$\frac{n_k}{n_{k+1}} = \rho_k \geq \rho > 1, \tag{2.1}$$

i.e., the number of vertices n_k decreases in a geometric ratio. Note that there are several algorithms of constructing the sequence $\{X_k\}$, see [4, 6, 7, 11, 14].

Now define the sequence of marices $\{A^{(k)}\}$. To do this consider the following block matrix form of $A^{(k)}$, $k \geq 0$,

$$A^{(k)} = \begin{bmatrix} A_{11}^{(k)} & A_{12}^{(k)} \\ A_{21}^{(k)} & A_{22}^{(k)} \end{bmatrix} \begin{matrix} \} \ X_k \backslash X_{k+1} \\ \} \ X_{k+1} \end{matrix} . \tag{2.2}$$

Note that the method of definition of sets X_k is defined a sparsity structure of blocks of the matrices $A^{(k)}$.

Next we consider block LU factorization of the matrix $A^{(k)}$

$$A^{(k)} = \begin{bmatrix} A_{11}^{(k)} & 0 \\ A_{21}^{(k)} & I \end{bmatrix} \begin{bmatrix} I & A_{11}^{(k)^{-1}} A_{12}^{(k)} \\ 0 & S^{(k+1)} \end{bmatrix},$$

where $S^{(k+1)}$ is the Schur complement of the matrix $A^{(k)}$ defined by

$$S^{(k+1)} = A_{22}^{(k)} - A_{21}^{(k)} (A_{11}^{(k)})^{-1} A_{12}^{(k)}. \tag{2.3}$$

We define now the matrix $A^{(k+1)}$ as the following approximation of the matrix $S^{(k+1)}$

$$A^{(k+1)} = A_{22}^{(k)} - \overline{A_{21}^{(k)} (A_{11}^{(k)})^{-1} A_{12}^{(k)}} - \theta Q^{(k+1)}, \tag{2.4}$$

where θ, $0 \leq \theta \leq 1$ is an iterative parameter, $\overline{C}$ is an "approximation" of the matrix C, i.e., the matrix, for which entries of C outside a chosen pattern are deleted, and $Q^{(k+1)}$ is a diagonal matrix defined from the row sum criteria

$$A^{(k+1)} e = S^{(k+1)} e \text{ at } \theta = 1, \tag{2.5}$$

for the positive vector $e = (1, 1, ..., 1)^T$. Note that due to the approximation we can always guarantee that the structure of a new matrix $A^{(k+1)}$ be

similar to that of the original matrix $A^{(k)}$ by deletion and diagonal compensation of undesirable off-diagonal entries, which destruct the chosen structure. Moreover, below we will show that this matrix is also a symmetric positive definite M-matrix. Hence, we can apply the above-defined process to it and repeat this process until the matrix $A^{(L)}$, corresponding to a coarse mesh is obtained.

To prove the correctness of the definition of the sequence of matrices $\{A^{(k)}\}, k = 0, 1, \ldots, L$ it is sufficient to show that this sequence is one of the Stieltjes matrices.

In the following, we will need the following two results, which were proven earlier. Here we rewrite them in our notation.

Lemma 2.1[8]. *Let $A^{(k)}$ be a Stieltjes matrix, then its Schur's complement $S^{(k+1)}$, defined by* (2.3)*, is also a Stieltjes matrix.*

Lemma 2.2[3]. *Let $S^{(k+1)}$ be a Stieltjes matrix, and let the symmetric matrix $A^{(k+1)}$ be the approximation of $S^{(k+1)}$ defined by* (2.4) *and* (2.5)*. Then $A^{(k+1)}$ is a Stieltjes matrix as well.*

A successive application of Lemmata 2.1 and 2.2 to the matrix $A^{(k)}$ shows that if $A^{(k)}$ is the Stieltjes matrix, then $A^{(k+1)}$ is such as well.

The preconditioning matrix M is recursively defined by the sequence of the preconditioning matrices $M^{(k)}$ as follows

$$
\begin{aligned}
& M^{(L)} = A^{(L)}, \\
& \textbf{For } k = L-1 \quad \textbf{to} \quad 0, \\
& \qquad M^{(k)} = \begin{bmatrix} A_{11}^{(k)} & 0 \\ A_{21}^{(k)} & I \end{bmatrix} \begin{bmatrix} I & A_{11}^{(k)^{-1}} A_{12}^{(k)} \\ 0 & Z^{(k+1)} \end{bmatrix},
\end{aligned}
\tag{2.6}
$$

where $Z^{(k+1)}$ is an approximation of the Schur complement defined by one of the following ways:

$$
\begin{aligned}
& (i) \quad Z^{(k+1)} = S^{(k+1)} \left[I - P_{\nu_{k+1}} (M^{(k+1)^{-1}} S^{(k+1)}) \right]^{-1}, \\
& (ii) \quad Z^{(k+1)} = A^{(k+1)} \left[I - P_{\nu_{k+1}} (M^{(k+1)^{-1}} A^{(k+1)}) \right]^{-1},
\end{aligned}
\tag{2.7}
$$

where $P_{\nu_{k+1}}(t)$ is a polynomial of degree ν_{k+1} as

$$P_{\nu_{k+1}}(t) = 1 - a_1 t - \ldots - a_{\nu_{k+1}} t^{\nu_{k+1}},$$

which is small in the interval $I_k = [\underline{t}_k, \overline{t}_k]$ containing all the eigenvalues of $M^{(k+1)^{-1}} S^{(k+1)} (M^{(k+1)^{-1}} A^{(k+1)})$. The choice of the polynomials will be discussed below in detail. Note that by the following Lemma 2.3 the sequence of matrices $M^{(k)}$ is a sequence of sparse symmetric positive definite matrices, so all the eigenvalues of $M^{(k)^{-1}} A^{(k)}$ are real and positive.

Lemma 2.3. *Let $\{A^{(k)}\}$, $k = 0, 1, \ldots, L$ be a sequence of Stieltjes matrices, then the sequence of preconditioning matrices $\{M^{(k)}\}$, defined by* (2.6) *and* (2.7)*, is a sequence of sparse symmetric positive definite matrices.*

Proof. We will prove it by induction on k. It is evident that by definition, $M^{(L)}$ is a symmetric positive definite M-matrix since $M^{(L)} = A^{(L)}$. Now we assume that for some fixed index k the matrix $M^{(k+1)}$ is a sparse symmetric positive definite one. By definition of $M^{(k)}$ we have

$$M^{(k)} = \begin{bmatrix} I & 0 \\ A_{21}^{(k)} A_{11}^{(k)^{-1}} & I \end{bmatrix} \begin{bmatrix} A_{11}^{(k)} & 0 \\ 0 & Z^{(k+1)} \end{bmatrix} \begin{bmatrix} I & A_{11}^{(k)^{-1}} A_{12}^{(k)} \\ 0 & I \end{bmatrix}. \qquad (2.8)$$

We have that $A_{11}^{(k)}$ is a Stieltjes matrix, since $A^{(k)}$ is a Stieltjes matrix. Moreover, by definition of the matrix $M^{(k+1)}$ and properties of the polynomial $P_{\nu_{k+1}}(t)$ we have that $Z^{(k+1)}$ is a sparse symmetric positive definite matrix. Thus, the matrix $M^{(k)}$ is also sparse symmetric and positive definite by the congruence with the block diagonal matrix in (2.8).

□

Remark 1. If the first degree polynomials $P_{\nu_{k+1}}(t) = 1 - t$ are used, then it leads to the method proposed by Il'in and Larin [11], since $Z^{(k+1)} = M^{(k+1)}$ for both above-described versions.

3 Organization of the computational process

It will be recalled that to solve the system (1.1) the preconditioned conjugate gradient method is used.

Since the single preconditioning matrix M on each iteration is used all the entries of the matrices $A^{(k)}$ and their preconditioners $M^{(k)}$, $k = 0, 1, \ldots, L-1, L$ can be calculated once before the iterative process.

Moreover, by definition of the preconditioner the solution of the system with the matrix M breaks up into a sequence of problems with the matrix $M^{(k)}$ on each level, which breaks up into forward

$$z_1^{(k)} = A_{11}^{(k)^{-1}} r_1^{(k)},$$
$$z_2^{(k)} = r_2^{(k)} - A_{21}^{(k)} z_1^{(k)},$$

and back substitutions

$$y_2^{(k)} = Z^{(k+1)^{-1}} z_2^{(k)},$$
$$y_1^{(k)} = z_1^{(k)} - A_{11}^{(k)^{-1}} A_{12}^{(k)} z_2^{(k)},$$

where the solution of the system with the matrix $Z^{(k+1)}$ is found by the method proposed in [1]

Solve $M^{(k+1)}x = a_{\nu_{k+1}} w.$

For $r = 1$ **to** $\nu_{k+1} - 1,$

Calculate $h = A^{(k+1)}x($ or $= S^{(k+1)}x)$

Solve $M^{(k+1)}x = h + a_{\nu_{k+1}-r} w.$

Now defining the polynomial degrees ν_k as usual

$$\begin{array}{llllllllll} \nu_0 & = 1, & \nu_1 & = 1, & \ldots, & \nu_{\mu-1} & = 1, & \nu_\mu & = \nu, \\ \nu_{\mu+1} & = 1, & \nu_{\mu+2} & = 1, & \ldots, & \nu_{2\mu} & = 1, & \nu_{\mu+(\mu+1)} & = \nu, \\ \ldots & & \ldots & & \ldots, & \ldots & & \ldots & \\ \nu_{r(\mu+1)} & = 1, & \nu_{r(\mu+1)+1} & = 1, & \ldots, & \nu_L = 1, & & & \end{array} \tag{3.1}$$

where $\mu, 0 \le \mu \le L$ is an integer parameter, $r = [L/\mu]$ and $[q]$ is an integer part of q, $q \in R$, and applying a recursive technique suggested in [5] we obtain the standard condition on the upper bounds of polynomial degrees

$$\nu < \rho^{\mu+1}, \tag{3.2}$$

under which the whole computational cost is proportional to the number of nodes on the fine mesh.

4 Rate of convergence

It is well known that the number of iterations $n(\varepsilon)$, which is necessary to satisfy

$$\frac{\|u^n - \hat{u}\|_A}{\|u^0 - \hat{u}\|_A} \leq \varepsilon, \tag{4.1}$$

where ε is a prescribed small number, for the PCG method is

$$n(\varepsilon) \leq \frac{1}{2} \ln(\frac{2}{\varepsilon})\sqrt{\kappa} + 1,$$

where κ is the condition number of the matrix $M^{-1}A$, which is defined by

$$\kappa = cond(M^{-1/2}AM^{-1/2}) = cond(M^{-1}A) = \frac{\lambda_{max}(M^{-1}A)}{\lambda_{min}(M^{-1}A)}, \tag{4.2}$$

where $\lambda_{max}(M^{-1}A)$ and $\lambda_{min}(M^{-1}A)$ are the maximal and the minimal eigenvalue of the matrix $M^{-1}A$, respectively.

Thus, for definition of the condition of the optimal order of the rate of convergence we have to study the condition number of $M^{-1}A$. The analysis will be made by a recursive relation between the condition numbers for every two adjoining levels. Note that the below analysis is closed to one, which has been done by [4] for another framework.

Let us recall that the polynomials $P_{\nu_{k+1}}(t)$ can have various degrees and values of coefficients with respect to a level, however, they have to satisfy

$$0 \leq P_{\nu_{k+1}}(t) < 1,$$

for all $t \in I_{k+1}$, $P_{\nu_{k+1}}(0) = 1$, where the interval I_{k+1} is defined as

$$(i) \quad I_{k+1} = \left[\inf_x \frac{(S^{(k+1)}x, x)}{(M^{(k+1)}x, x)} \ , \ \sup_x \frac{(S^{(k+1)}x, x)}{(M^{(k+1)}x, x)} \right],$$

$$(ii) \quad I_{k+1} = \left[\inf_x \frac{(A^{(k+1)}x, x)}{(M^{(k+1)}x, x)} \ , \ \sup_x \frac{(A^{(k+1)}x, x)}{(M^{(k+1)}x, x)} \right].$$

Moreover, let us also recall a simple and useful result, which was proven earlier. Here we rewrite it in our notation.

Lemma 4.1[3]. *For $S^{(k+1)}$ and $A^{(k+1)}$, defined by* (2.3), (2.4) *and* (2.5), *there is a constant $\alpha_{k+1} < 1$ such as*

$$0 < \alpha_{k+1}(S^{(k+1)}x, x) \leq (A^{(k+1)}x, x) \leq (S^{(k+1)}x, x)$$

for all $x \in R^{n_{k+1}}$.

To estimate the condition number of $M^{-1}A$ we will need the estimate of the condition numbers of $M^{(k+1)^{-1}}A^{(k+1)}$ and $M^{(k+1)^{-1}}S^{(k+1)}$, which are defined through the estimates for the condition number of $Z^{(k+1)^{-1}}S^{(k+1)}$.

Lemma 4.2. *For* $S^{(k+1)}$ *and* $Z^{(k+1)}$ *defined by* (2.3) *and* (2.7) *the following inequalities for all* $x \in R^{n_{k+1}}$ *are valied:*

$$(i) \quad 0 < \frac{(S^{(k+1)}x, x)}{(Z^{(k+1)}x, x)} \leq 1,$$
$$(ii) \quad 0 < \frac{(S^{(k+1)}x, x)}{(Z^{(k+1)}x, x)} \leq \alpha_{k+1}^{-1}.$$

Proof. From the definition of $Z^{(k+1)}$ and the results of Lemma 4.1 after simple transformations we obtain the desired results for both cases

$$\frac{(S^{(k+1)}x, x)}{(Z^{(k+1)}x, x)} = \frac{(S^{(k+1)}x, x)}{(S^{(k+1)}\left[I - P_{\nu_{k+1}}(M^{(k+1)^{-1}}S^{(k+1)})\right]^{-1}x, x)} \leq$$
$$\leq 1 - P_{\nu_{k+1}}(M^{(k+1)^{-1}}S^{(k+1)}) \leq 1,$$
$$\frac{(S^{(k+1)}x, x)}{(Z^{(k+1)}x, x)} = \frac{(S^{(k+1)}x, x)}{(A^{(k+1)}x, x)} \cdot \frac{(A^{(k+1)}x, x)}{(A^{(k+1)}\left[I - P_{\nu_{k+1}}(M^{(k+1)^{-1}}A^{(k+1)})\right]^{-1}x, x)} \leq$$
$$\leq \alpha_{k+1}^{-1}(1 - P_{\nu_{k+1}}(M^{(k+1)^{-1}}A^{(k+1)})) \leq \alpha_{k+1}^{-1}.$$

□

Lemma 4.3. *For* $A^{(k)}$ *and* $M^{(k)}$ *the following inequalities for all* $x \in R^{n_k}$, $x_2 \in R^{n_{k+1}}$ *hold:*

$$(i) \quad 0 < \inf_{x_2} \frac{(S^{(k+1)}x_2, x_2)}{(Z^{(k+1)}x_2, x_2)} \leq \frac{(A^{(k)}x, x)}{(M^{(k)}x, x)} \leq 1,$$
$$(ii) \quad 0 < \inf_{x_2} \frac{(S^{(k+1)}x_2, x_2)}{(Z^{(k+1)}x_2, x_2)} \leq \frac{(A^{(k)}x, x)}{(M^{(k)}x, x)} \leq \alpha_{k+1}^{-1}.$$

Proof. From factarization forms of $A^{(k)}$ and $M^{(k)}$ one can see that

$$\frac{(A^{(k)}x, x)}{(M^{(k)}x, x)} = \frac{\left(\begin{bmatrix} A_{11}^{(k)} & 0 \\ 0 & S^{(k+1)} \end{bmatrix} x, x\right)}{\left(\begin{bmatrix} A_{11}^{(k)} & 0 \\ 0 & Z^{(k+1)} \end{bmatrix} x, x\right)} = \frac{(A_{11}^{(k)}x_1, x_1) + (S^{(k+1)}x_2, x_2)}{(A_{11}^{(k)}x_1, x_1) + (Z^{(k+1)}x_2, x_2)},$$

from which and by the results of Lemma 4.3 we obtain the lower estimate

$$\frac{(A^{(k)}x,x)}{(M^{(k)}x,x)} \geq \inf_x \frac{(A^{(k)}x,x)}{(M^{(k)}x,x)} = \inf_{x_2} \frac{(S^{(k+1)}x_2,x_2)}{(Z^{(k+1)}x_2,x_2)} > 0$$

for both versions, and the upper estimates

$$\frac{(A^{(k)}x,x)}{(M^{(k)}x,x)} \leq \sup_x \frac{(A^{(k)}x,x)}{(M^{(k)}x,x)} = \sup_{x_2} \frac{(S^{(k+1)}x_2,x_2)}{(Z^{(k+1)}x_2,x_2)} = 1$$

for version (i), and

$$\frac{(A^{(k)}x,x)}{(M^{(k)}x,x)} \leq \sup_x \frac{(A^{(k)}x,x)}{(M^{(k)}x,x)} = \sup_{x_2} \frac{(S^{(k+1)}x_2,x_2)}{(Z^{(k+1)}x_2,x_2)} = \alpha_{k+1}^{-1}$$

for version (ii).

□

Now we make use of the above-obtained results for the definition of the interval I_{k+1} for both versions.

Theorema 4.1. *The interval I_{k+1} is*

$$(i) \quad I_{k+1} = \left[\inf_x \frac{(S^{(k+1)}x,x)}{(Z^{(k+1)}x,x)}, \alpha_{k+1}^{-1}\right],$$
$$(ii) \quad I_{k+1} = \left[\inf_x \frac{(S^{(k+1)}x,x)}{(Z^{(k+1)}x,x)}, \alpha_{k+2}^{-1}\right].$$

Proof. Using the equality

$$\frac{(S^{(k+1)}x,x)}{(M^{(k+1)}x,x)} = \frac{(S^{(k+1)}x,x)}{(A^{(k+1)}x,x)} \cdot \frac{(A^{(k+1)}x,x)}{(M^{(k+1)}x,x)}$$

and by the results of Lemmata 4.1–4.3 we obtain the desired result.

□

Hence, the polynomials $P_{\nu_k}(t)$ are chosen as

$$P_{\nu_k}(t) = \frac{T_{\nu_k}\left(\frac{\overline{t}_k+\underline{t}_k-2t}{\overline{t}_k-\underline{t}_k}\right)+1}{T_{\nu_k}\left(\frac{\overline{t}_k+\underline{t}_k}{\overline{t}_k-\underline{t}_k}\right)+1}, \tag{4.3}$$

where $T_\nu(t)$ are the Chebyshev polynomials of degree ν,

$$T_0 = 1,\ T_1 = t,\ T_{\nu+1} = 2tT_\nu - T_{\nu-1},$$

and the boundary points of the interval I_k are chosen as

$$\underline{t}_k = \inf_x \frac{(S^{(k+1)}x, x)}{(Z^{(k+1)}x, x)}; \quad \overline{t}_k = \begin{cases} \alpha_k^{-1}, & (i) \\ \alpha_{k+1}^{-1}. & (ii) \end{cases}$$

Note that for the polynomials P_{ν_k} the following equality is valid:

$$\max_{t \in I_k} P_{\nu_k}(t) = P_{\nu_k}(\underline{t}_k),$$

from which we obtain the recursive formula for calculating the lower bounds of intervals

$$\underline{t}_k = 1 - P_{\nu_{k+1}}(\underline{t}_{k+1}). \tag{4.4}$$

Remark 2. For $\nu_k = 1$, due to the definition of $T_1(t) = t$, we have

$$P_{\nu_k}(t) = 1 - \frac{t}{\overline{t}_k} = \begin{cases} 1 - \alpha_k t, & (i) \\ 1 - \alpha_{k+1} t. & (ii) \end{cases}$$

Following (4.2) we now define the condition number of the matrices $M^{(k+1)^{-1}} A^{(k+1)}$ as

$$\kappa_{k+1} = \frac{\overline{t}_{k+1}}{\underline{t}_{k+1}}.$$

Remark 3. In the case $\nu_k = 1$ we have

$$\begin{aligned} &(i) \quad \kappa_k = \frac{1}{1 - P_{\nu_{k+1}}(\underline{t}_{k+1})} = \frac{1}{\alpha_{k+1}\underline{t}_{k+1}} = \alpha_{k+1}^{-1}\kappa_{k+1}, \\ &(ii) \quad \kappa_k = \alpha_{k+1}^{-1} \cdot \frac{1}{1 - P_{\nu_{k+1}}(\underline{t}_{k+1})} = \alpha_{k+1}^{-1} \cdot \frac{1}{\alpha_{k+2}\underline{t}_{k+1}} = \alpha_{k+1}^{-1}\kappa_{k+1}. \end{aligned} \tag{4.5}$$

Remark 4. In the case $\nu_k \neq 1$ due to the definition of κ_k and by (4.3), we have

$$\begin{aligned} &(i) \quad \kappa_k = \frac{1}{1 - P_{\nu_{k+1}}(t_{k+1})} = \frac{T_{\nu_{k+1}}\left(\frac{1+\kappa_{k+1}^{-1}}{1-\kappa_{k+1}^{-1}}\right) + 1}{T_{\nu_{k+1}}\left(\frac{1+\kappa_{k+1}^{-1}}{1-\kappa_{k+1}^{-1}}\right) - 1}, \\ &(ii) \quad \kappa_k = \alpha_{k+1}^{-1} \cdot \frac{1}{1 - P_{\nu_{k+1}}(t_{k+1})} = \alpha_{k+1}^{-1} \cdot \frac{T_{\nu_{k+1}}\left(\frac{1+\kappa_{k+1}^{-1}}{1-\kappa_{k+1}^{-1}}\right) + 1}{T_{\nu_{k+1}}\left(\frac{1+\kappa_{k+1}^{-1}}{1-\kappa_{k+1}^{-1}}\right) - 1}. \end{aligned} \tag{4.6}$$

Using the results of (4.5), (4.6) and due to the choice of polynomial degrees ν_k as in (3.1) we obtain

$$(i) \quad \kappa_k = \left(\prod_{s=k+1}^{k+\mu} \alpha_s^{-1} \right) \cdot \frac{T_{\nu_{k+\mu+1}} \left(\frac{1+\kappa_{k+\mu+1}^{-1}}{1-\kappa_{k+\mu+1}^{-1}} \right) + 1}{T_{\nu_{k+\mu+1}} \left(\frac{1+\kappa_{k+\mu+1}^{-1}}{1-\kappa_{k+\mu+1}^{-1}} \right) - 1},$$

$$(ii) \quad \kappa_k = \left(\prod_{s=k+1}^{k+\mu+1} \alpha_s^{-1} \right) \cdot \frac{T_{\nu_{k+\mu+1}} \left(\frac{1+\kappa_{k+\mu+1}^{-1}}{1-\kappa_{k+\mu+1}^{-1}} \right) + 1}{T_{\nu_{k+\mu+1}} \left(\frac{1+\kappa_{k+\mu+1}^{-1}}{1-\kappa_{k+\mu+1}^{-1}} \right) - 1}.$$

Using the standard technique, which is described in [5], we obtain the final condition on lower bounds of degrees of polynomials

$$\begin{aligned} (i) \quad & \nu > \left(\max_{\xi=1,2,\ldots[L/\mu]} \prod_{s=(\xi-1)(\mu+1)}^{\xi(\mu+1)-1} \alpha_s^{-1} \right)^{\frac{1}{2}}, \\ (ii) \quad & \nu > \left(\max_{\xi=1,2,\ldots[L/\mu]} \prod_{s=(\xi-1)(\mu+1)+1}^{\xi(\mu+1)} \alpha_s^{-1} \right)^{\frac{1}{2}}. \end{aligned} \tag{4.7}$$

On the basis of the above results the following theorem can be formulated.

Theorem 4.2. *The algebraic multilevel incomplete factorization method for Stieltjes matrices, based on a sequence of matrices* $\{A^{(k)}\}$*, defined by* (2.2), (2.4) *and* (2.5)*, and a sequence of preconditioning matrices* $\{M^{(k)}\}$*, recursively defined by* (2.6), (2.7) *and* (4.3)*, has an optimal order of computational complexity if*

$$\left(\max_{\xi=1,2,\ldots[L/\mu]} \prod_{s=(\xi-1)(\mu+1)}^{\xi(\mu+1)-1} \alpha_s^{-1} \right)^{\frac{1}{2}} < \nu < \rho^{\mu+1}$$

for version (i) or

$$\left(\max_{\xi=1,2,\ldots[L/\mu]} \prod_{s=(\xi-1)(\mu+1)+1}^{\xi(\mu+1)} \alpha_s^{-1} \right)^{\frac{1}{2}} < \nu < \rho^{\mu+1}$$

for version (ii), where α_s *are the upper bounds of the maximal eigenvalues of* $A^{(s)-1}S^{(s)}$*,* ρ *is the lower bound of the progression ratio of the degrees of*

freedom from the level $k+1$ to the level k defined in (2.1), *ν is the degree of the matrix polynomials used in* (2.7) *and $\nu > 1$ at every $(\mu+1)$th step.*

Thus, properly choosing the polynomial degrees we have an optimal rate of convergence, i.e., the condition number of $M^{-1}A$ has the magnitude of $O(1)$, and an optimal order of the total computational complexity, i.e., the number of arithmetic operations is proportional to N.

In conclusion, for completeness of presentation we write down the explicit estimate of the condition number of $M^{-1}A$ using the above results for every two adjoining levels

$$\begin{aligned} &(i) \quad cond(M^{-1}A) \le \left(\prod_{s=1}^{\mu} \alpha_s^{-1}\right) \cdot \frac{T_{\nu_\mu}\left(\frac{1+\kappa_{\mu+1}^{-1}}{1-\kappa_{\mu+1}^{-1}}\right)+1}{T_{\nu_\mu}\left(\frac{1+\kappa_{\mu+1}^{-1}}{1-\kappa_{\mu+1}^{-1}}\right)-1}, \\ &(ii) \quad cond(M^{-1}A) \le \left(\prod_{s=1}^{\mu+1} \alpha_s^{-1}\right) \cdot \frac{T_{\nu_\mu}\left(\frac{1+\kappa_{\mu+1}^{-1}}{1-\kappa_{\mu+1}^{-1}}\right)+1}{T_{\nu_\mu}\left(\frac{1+\kappa_{\mu+1}^{-1}}{1-\kappa_{\mu+1}^{-1}}\right)-1}. \end{aligned} \tag{4.8}$$

As it is readily seen now, under satisfying the conditions (3.1) and (4.7) the rate of convergence is defined by the quality of approximation on the first μ levels and by the value of a limited function depending on the condition numbers of consequent levels. Moreover, it is shown that the quality of preconditioner for version (i) is better, than for version (ii), because the value of constant befor fractions for version (i) is always less than that for version (ii).

Remark 5. Unfortunately, the above analysis does not ensure existence of the parameters μ and ν for which the conditions of Theorem 4.2 on polynomial degrees are satisfied, because the theoretical investigation of α_s is a very difficult problem which depends on the value of the parameter θ and the choice of the approximation pattern in (2.4). Hence, the numerical experiments for checking up the conditions of Theorem 4.2 have been made.

5 Numerical experiments

Let us give the results of the experiments conducted on the iterative solution of system (1.1), which arises from a 5-point finite difference approximation

of the Laplace equation on a rectangular grid. The right-hand side in the system of equations was chosen so that the solution has the form

$$\hat{u}_{ij} = const = 1.$$

Since the preconditioning matrix M defined by (2.6), (2.7) and (4.3) can be calculated one of the two methods, then all the experiments were made twice for every version separately, and the results are presented in Tables 1–2.

For the complete analysis of the quality of the preconditioner the system of equations has been solved by two different iterative methods with the preconditioning matrix M defined by AMLI method for various definitions of the sets X_k on 7×7, 15×15, 31×31 and 63×63 grids.

The following vector was always taken as an initial guess

$$u_{ij}^0 = 2 + 100 \sin^2(\frac{pi\pi}{L+1}) \sin^2(\frac{qj\pi}{M+1}),$$
$$p = q = 1, \quad i = 1, \ldots, n, \quad j = 1, \ldots, n.$$

Iterations were repeated until the condition (4.1) was met under $\varepsilon = 10^{-6}$.

As it is readily seen from the above reasoning for calculation of polynomial coefficients we have to know the spectrum bounds of $M^{(k)^{-1}} S^{(k)} (M^{(k)^{-1}} A^{(k)})$.

To find the lower boundary of these eigenvalues we made use of the recursive formula (4.4) beginning with the last (coarse) level, for which $\underline{t}_L = \overline{t}_L = 1$.

The search for the upper boundary of these eigenvalues is a more complicated problem for whose solution we shall use the Lanczos method [15], which as in the numerical experiments show (see Table 3) has the fast rate of convergence to the maximal eigenvalue of the matrix $M^{(k)^{-1}} S^{(k)} (M^{(k)^{-1}} A^{(k)})$.

Moreover, as it is readily seen from the definition of the matrices $Z^{(k+1)}$ in (2.7) the main difference of version (i) from version (ii) is that in the process of solution the multiplication of a vector by the matrix $A^{(k+1)}$ is substituted for the same operation with the matrix $S^{(k+1)}$, for solving which we have to solve the system with $A_{11}^{(k)}$ easy and cheap. The latter depends on the choice of the sequence of nested sets $\{X_k\}$, which is defined by dependence on the structure of the initial grid. In the present paper when carrying out the numerical experiments we used the well-known methods of construction of the sequence of nested sets: recursive red-black ordering [4], [14] and recursive block odd-even ordering [11].

Table 1. Recursive red-black ordering

n	L	μ	ν	(i)	(ii)	AxNey	(i)	(ii)	AxNey
7	4	0	1	3	3	3	9	9	9
		0	2	4	4	4	5	9	5
		0	3	3	4	3	3	9	3
		1	2	5	4	5	5	9	5
		1	3	3	4	3	4	9	4
		2	3	4	4	4	11	9	11
15	6	0	1	7	7	7	11	11	11
		0	2	4	7	4	5	8	5
		0	3	3	6	3	3	8	3
		1	2	5	6	5	7	8	7
		1	3	4	6	4	4	8	4
		2	3	7	7	7	12	10	12
31	8	0	1	9	9	9	17	17	17
		0	2	4	6	4	4	7	4
		0	3	3	6	3	3	6	3
		1	2	6	6	6	7	7	7
		1	3	4	5	4	4	6	4
		2	3	7	7	7	12	9	12
63	10	0	1	11	11	11	25	25	25
		0	2	4	5	4	4	6	4
		0	3	3	4	3	3	5	3
		1	2	6	6	6	7	7	7
		1	3	4	4	4	4	5	4
		2	3	7	7	7	12	8	12

Table 1 show the results of the experiments for the preconditioned conjugate gradient method and the preconditioned steepest descent method for recursive red-black ordering. To compare the method presented here with the earlier, suggested by Axelsson and Neytcheva [5], the appropriat experiments were performed. The results of the computations are also shown in Table 1. Note that the preconditioner for the version (i) is identical to the AMLI preconditioner. All the calculations were performed in MATLAB.

On the basis of the experiments the following conclusions can be done.

- The method has an optimal rate of convergence for both versions for all μ and $\nu \neq 1$.
- The method has an optimal order of a computational complexity if $\nu < 2^{\mu+1}$, since $\rho = 2$ as for recursive red-black ordering and for recursive block odd-even reduction.
- The method has a similar behavior with the AMLI method. Moreover, note that the quality of preconditioning matrices for both methods can be improved by using a band approximation of matrices $A_{11}^{(k)^{-1}}$ instead of a standard (diagonal) one.
- The quality of the preconditioner for version (i) is always better than for version (ii) as it was expected (see formula (4.8)).
- Choosing polynomial degrees ν_k as in (3.1) for $\mu = 1$, $\nu = 3$, we have the optimal method with respect to the number of iterations and the total computational costs.

In the Table 2 we shows the the experimental results for both version of the methods, but for another (recursive block odd-even) orderings. As it is readily seen the method has a nearly optimal rate of convergence for both versions for $\mu = 1$ and $\nu = 3$, except a case $n = 63$. For this value of n we loss the optimality. To understanding why it is happening we have to check the condition (4.7) of the Theorem 4.2 on the polynomial degrees.

Table 2. Recursive block odd-even ordering

n	L	μ	ν	(i)	(ii)	(i)	(ii)
7	3	0	1	6	6	11	11
		0	2	4	6	5	11
		0	3	3	6	3	11
		1	2	4	6	5	11
		1	3	3	6	3	11
		2	3	7	6	10	11
15	4	0	1	9	9	34	34
		0	2	5	8	6	12
		0	3	3	8	3	11
		1	2	8	8	11	12
		1	3	5	8	6	11
		2	3	8	10	11	33
31	5	0	1	14	14	98	98
		0	2	6	10	8	19
		0	3	3	8	4	12
		1	2	11	13	19	25
		1	3	7	10	9	17
		2	3	10	14	16	43
63	6	0	1	26	26	488	488
		0	2	10	16	17	37
		0	3	4	10	4	16
		1	2	21	23	65	114
		1	3	10	19	16	56
		2	3	19	20	53	83

Table 3. Examples of the upper bounds of the maximum eigenvalues of matrices $M^{(k)^{-1}} A^{(k)}$, recursive odd-even ordering

n	level number	λ_{max}	nit Lanczos	α_s^{-1}
7	0	1.7312	4	$\alpha_1^{-1} = 1.7312$
	1	1.0000	2	
15	0	5.2194	5	$\alpha_1^{-1} = 1.9201$
	1	2.7183	4	$\alpha_2^{-1} = 2.7183$
	2	1.0000	2	
31	0	17.1119	5	$\alpha_1^{-1} = 1.9830$
	1	8.6291	5	$\alpha_2^{-1} = 3.9731$
	2	2.1719	4	$\alpha_3^{-1} = 2.1719$
	3	1.0000	2	
63	0	91.9444	7	$\alpha_1^{-1} = 1.9969$
	1	46.0430	6	$\alpha_2^{-1} = 4.7754$
	2	9.6417	6	$\alpha_3^{-1} = 9.5462$
	3	1.0100	2	$\alpha_4^{-1} = 1.0100$
	4	1.0000	2	

To check the condition (4.7) we have to know how to compute α_s. Note that to do this there is no need to solve the generalized eigenvalue problem for matrices $S^{(k)}$ and $A^{(k)}$ directly, but we may use the result of Remark 3 and the fact that $\lambda_{min}^{(s)} = 1$ in the case $\mu = 0,\ \nu = 1$ for all the levels. Thus, we obtain the following formula to calculate α_s^{-1}:

$$\alpha_s^{-1} = \frac{\lambda_{max}^{(s)}}{\lambda_{max}^{(s+1)}},$$

where $\lambda_{max}^{(s)}$ is the upper bound of the maximum eigenvalue of $M^{(k)^{-1}} S^{(k)}$ ($M^{(k)^{-1}} A^{(k)}$) defined by the Lanczos method.

For the analysis of the condition (4.7) now using the results from Table 3 one can easier see that the square root of the maximal product of maximum eigenvalues is less than the value of ν for $n = 7, 15, 31$:

$$n = \ 7,\ (\max\{1.7312\})^{1/2} = (1.7312)^{1/2} = 1.3158 < 3,$$

$$n = 15,\ (\max\{2.7183, 1.9201\})^{1/2} = (2.7183)^{1/2} = 1.6487 < 3,$$

$$n = 31,\ (\max\{2.1719, 3.9731 \cdot 1.9830\})^{1/2} = (7.8767)^{1/2} = 2.8066 < 3,$$

and greater than it for $n = 63$:

$$(\max\{1.9969 \cdot 4.7754, 9.5462 \cdot 1.0100\})^{1/2} = (9.5462)^{1/2} = 3.0897 > 3.$$

Note that, however, till now we did not use the iterative parameter θ to accelerate the rate of convergence ($\theta = 1$ in all above mentioned experiments). For example, in the Table 4 the upper bounds of the maximum eigenvalues of matrices $M^{(k)^{-1}} A^{(k)}$ for $\theta = 0.99$ are shown, from which we have for $n = 63$

$$(\max\{1.8861 \cdot 3.4922, 1.6523 \cdot 1.0100\})^{1/2} = (6.5864)^{1/2} = 2.5665 < 3,$$

i.e. the first condition of Theorem 4.2 are satisfied.

Moreover, as it is easy to see from (3.2), in this case the method has an optimal order of computational complexity, since $3 < 2^2$.

Table 4. Examples of the upper bounds of the maximum eigenvalues of matrices $M^{(k)^{-1}}A^{(k)}$, recursive odd-even ordering, $\theta = 0.99$

n	level number	λ_{max}	nit Lanczos	α_s^{-1}
63	0	10.9941	7	
				$\alpha_1^{-1} = 1.8861$
	1	5.8278	6	
				$\alpha_2^{-1} = 3.4922$
	2	1.6688	7	
				$\alpha_3^{-1} = 1.6523$
	3	1.0100	2	
				$\alpha_4^{-1} = 1.0100$
	4	1.0000	2	

Thus, we have proven both theoretical and experimentally that under a destruction of the condition on the polynomial degrees the method loses its optimality. On the other hand, we show that the quality of preconditioners depends on the choice of the sets of nested meshes directly.

In conclusion we can indicate some ways of future investigations:

- Study the influence of the optimal value of the iterative parameter θ on the behavior of the rate of convergence.
- Study the efficacy of this method for anizotropic problems.

References

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

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

[3] O.Axelsson, *The method of diagonal compensation of reduced matrix entries and multilevel iteration*, J. of Comp. and Appl. Math., 38 (1991), pp. 31–43.

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

[5] O.Axelsson and M.Neytcheva, *Algebraic multilevel iteration method for Stieltjes matrices*, Numer. Lin. Alg. Appl., 1 (1994), pp. 213–236.

[6] O.Axelsson, K.Georgiev, M.Mellaard, M.Neytcheva and A.Padiy, *Scalable and Optimal Iterative Solvers for Linear and Nonlinear Problems*, Report 9613, 1996, Catholic University, Nijmegen, The Netherlands.

[7] O.Axelsson and M.Larin, *An algebraic multilevel iteration method for finite element matrices*, (to appear).

[8] F.R. Gantmakher, *The theory of matrices*, Vol 1 & 2, New York, Chelsea, 1959.

[9] Yu.R. Hakopian and Yu.A. Kuznetsov, *Algebraic multigrid / substructuring preconditioners on triangular grids*, Sov. J. Numer. Anal. Math. Modelling, 6 (1989), pp. 453–483.

[10] V.P. Il'in, *Iterative Incomplete Factorization Methods*, World Scientific Publishing Co., Singapore, 1992.

[11] V.P. Il'in and M.R. Larin, *Iterative Multilevel Incomplete Factorization method for 5-point difference matrices*, submitted to NCC Publisher, 1996 (in Russian).

[12] Yu.A. Kuznetsov, *Algebraic Multigrid Domain Decomposition Methods*, Sov. J. Numer. Anal. Math. Modelling, 4 (1989), pp. 351–380.

[13] Yu.Iliash, Yu.Kuznetsov and Yu.Vassilevski, *On the application of fictious domain and strengthened AMG method for locally fitted 3D cartesian meshes to the potential flow problem on a massively parallel computer*, Report 9543, 1995, Catholic University, Nijmegen, Netherlands.

[14] Y. Notay and Z. Ould Amar, *Incomplete factorization preconditioning may lead to multigrid like speed of convergence*, Proceedings of the International Conference AMCA-95, Novosibirsk, Russia, 20-24 June, 1995, edited by A.S. Alekseev and N.S. Bakhvalov, NCC Publisher, 1995, pp. 435–446.

[15] B. Parlett, *The Symmetric Eigenvalue Problem*, Prentice-Hall, Inc., 1980.