Bull. Nov. Comp. Center, Num. Anal., 8 (1998), 57–69
© 1998 NCC Publisher

On the theory of algebraic multilevel incomplete factorization methods for the Stieltjes matrices*

Maxim Larin

Recently an algebraic multilevel incomplete factorization method for solving large linear systems with the Stieltjes matrices has been proposed. This method is a combination of two well-known techniques: algebraic multilevel (AMLI) and incomplete factorization. However, the efficiency of this method strongly depends on the choice of the relaxation parameter θ, an optimal value of which depends on the problem to be solved. In the present paper we study this dependence theoretically and propose a new method, that dynamically computes the corresponding problem–dependence optimal value of θ, and use it to construct an approximation of Schur's complement as a new matrix on the lower level in the AMLI framework.

1. Introduction

This work concerns a solution to the linear system of equations

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

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

Recently the main topic of many papers is to get an optimal order preconditioner for the solution to 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 [1–4, 6] methods allow us to construct preconditioners with these properties.

In the paper [11] an algebraic multilevel incomplete factorization method for the Stieltjes matrices has been proposed. Its main difference from the earlier, suggested by Axelsson and Neytcheva [4] is in the following: instead of using an approximation of the first pivoiting block for obtaining a new matrix on the lower level as its Schur's complement we have propose to use

*Supported by the Russian Foundation for Basic Research under Grant 96–01–01770 and by Presidium of Siberian Division of Russian Academy of Sciences.

the iterative incomplete factorization method [9] to construct an approximation of Schur's complement as a new matrix on the lower level, which has the structure similar to that of the original matrix. However, the choice of the optimal value of the relaxation parameter θ *strongly* depends on the problem to be solved. Moreover, there is a *high* sensitivity of the rate of convergence of PCG method with respect to the variation of θ around θ_{opt} [9, 10].

In the present paper, we study this dependence theoretically, basing on the results for iterative incomplete factorization technique [9, 14], and propose a new method that dynamically computes the corresponding problem-dependence value of θ, and is used in order to construct a new matrix on the lower level in the AMLI framework.

The paper is organized as follows. In Section 2 the algorithm of constructing the preconditioning matrix M is described. Some conditions of attaining an optimal order of computational complexity and an optimal rate of convergence are given in Sections 3 and 4, respectively. The main theoretical results are derived in Section 5.

2. Construction of the preconditioning matrix

To construct a multilevel preconditioning matrix M we usually have to define a sequence of matrices $A^{(k)}$ of order n_k, $k = 0, 1, \ldots, L-1, L$, 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 the 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 $\{X_k\}$, for example, see [3, 5, 6, 11, 15].

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

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

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

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

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

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

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

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

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

where θ_{k+1} $(-1 \leq \theta_{k+1} \leq 1)$ is a relaxation parameter, the choice of which will be discussed in Section 5; $\overline{C}$ is an *approximation* of C matrix, i.e., the matrix, for which all the entries of C outside a chosen pattern are neglected, and $Q^{(k+1)}$ is a diagonal matrix defined from the row sum criteria

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

or, the same,

$$Q^{(k+1)}e = A_{21}^{(k)}\left(\left(A_{11}^{(k)}\right)^{-1} - \overline{\left(A_{11}^{(k)}\right)^{-1}}\right)A_{12}^{(k)}e, \tag{2.6}$$

for the positive vector $e = (1, 1, \ldots, 1)^T$. Note that due to the approximation we can always make gain that the structure of a new $A^{(k+1)}$ matrix is similar to that of the original matrix $A^{(k)}$ by deletion and diagonal compensation of undesirable off-diagonal entries, that destruct the chosen structure. Moreover, as it has been shown in [11], 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 $A^{(L)}$ matrix, corresponding to a coarse mesh is obtained.

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

$$M^{(L)} = A^{(L)},$$

For $k = L-1$ **to** 0

$$M^{(k)} = \begin{bmatrix} A_{11}^{(k)} & 0 \\ A_{21}^{(k)} & I \end{bmatrix} \begin{bmatrix} I & (A_{11}^{(k)})^{-1}A_{12}^{(k)} \\ 0 & Z^{(k+1)} \end{bmatrix} \begin{matrix} \}\, X_k \setminus X_{k+1} \\ \}\, X_{k+1} \end{matrix}, \tag{2.7}$$

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

$$\begin{aligned} &\text{(i)} \quad Z^{(k+1)} = S^{(k+1)}\left[I - P_{\nu_{k+1}}\left((M^{(k+1)})^{-1}S^{(k+1)}\right)\right]^{-1}, \\ &\text{(ii)} \quad Z^{(k+1)} = A^{(k+1)}\left[I - P_{\nu_{k+1}}\left((M^{(k+1)})^{-1}A^{(k+1)}\right)\right]^{-1}, \end{aligned} \tag{2.8}$$

where $P_{\nu_{k+1}}(t)$ is a polynomial of ν_{k+1} degree and is defined at the interval $I_k = [\underline{t}_k, \overline{t}_k]$ containing all the eigenvalues of the matrix

$$(M^{(k+1)})^{-1}S^{(k+1)}\Big((M^{(k+1)})^{-1}A^{(k+1)}\Big)$$

as

$$P_{\nu_k}(t) = \frac{T_{\nu_k}\left(\frac{\bar{t}_k+\underline{t}_k-2t}{\bar{t}_k-\underline{t}_k}\right)+1}{T_{\nu_k}\left(\frac{\bar{t}_k+\underline{t}_k}{\bar{t}_k-\underline{t}_k}\right)+1}, \tag{2.9}$$

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

$$T_0(t) = 1, \quad T_1(t) = t, \quad T_{\nu+1}(t) = 2t\,T_\nu(t) - T_{\nu-1}(t).$$

The choice of the degrees of polynomials is derived from the conditions of an optimal order for a whole computational complexity and for a rate of convergence, and will be discussed in the next sections. Moreover, all the eigenvalues of $(M^{(k)})^{-1}A^{(k)}$ are real and positive.

3. Upper bounds for the polynomial degrees

Let us recall that on each level we have to solve a system with $M^{(k)}$ matrix, that is a preconditioner to $A^{(k)}$ matrix. By its structure it breaks up into forward and back substitutions. More precisely, to solve a system with $M^{(k)}$ matrix we have to solve two systems with the diagonal matrix $A_{11}^{(k)}$ and the system with $Z^{(k+1)}$, that was firstly suggested in [1] and is required ν_{k+1} solutions with $M^{(k+1)}$.

Now define the polynomial degrees ν_k as usual

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

where μ $(0 \le \mu \le L)$ is an integer parameter, $r = \lfloor L/\mu \rfloor$, and $\lfloor q \rfloor$ is an integer part of $q \in R$. Applying a recursive technique suggested in [4], 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. Lower bounds for the polynomial degrees

An analysis of the condition number of A matrix to the M preconditioner can be done by comparing the condition numbers on two adjacent levels. Here we only state our main results in this respect, see [11] for the proof.

First of all, for $S^{(k+1)}$ and $A^{(k+1)}$, defined by (2.3) and (2.4)–(2.6), respectively, there are two positive constants $\beta_{k+1} < 1$ and $\alpha_{k+1} \geq 1$ such that

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

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

Now we can easily define the interval I_{k+1} for both versions as follows:

$$I_{k+1} = [\underline{t}_{k+1}, \bar{t}_{k+1}] = \begin{cases} \left[\inf_x \dfrac{(S^{(k+1)}x, x)}{(Z^{(k+1)}x, x)}, \alpha_{k+1}\right] & \text{(i)}, \\ \left[\inf_x \dfrac{(S^{(k+1)}x, x)}{(Z^{(k+1)}x, x)}, \alpha_{k+2}\right] & \text{(ii)}. \end{cases} \tag{4.1}$$

Now due to the definition of polynomials and auxiliary results in [11] we obtain

$$\underline{t}_k = \begin{cases} 1 - P_{\nu_{k+1}}(\underline{t}_{k+1}) & \text{(i)}, \\ \beta_{k+1}\,(1 - P_{\nu_{k+1}}(\underline{t}_{k+1})) & \text{(ii)}. \end{cases} \tag{4.2}$$

Here we have only to emphasize that both α_{k+1} and β_{k+1} depend on the relaxation parameter θ_{k+1}.

Denote by κ_s the condition number of the matrices $M^{(s)^{-1}}A^{(s)}$. Now by the definition of the polynomials $P_{\nu_k}(t)$ and their degrees ν_k as in (3.1) we obtain the following recursive relation:

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

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

Now using the standard technique, that is described in [4], we obtain the final condition on lower bounds of degrees of polynomials

$$\begin{aligned} &\text{(i)} \quad \nu > \left(\max_{\xi=1,\dots,\lfloor L/\mu\rfloor} \prod_{s=(\xi-1)(\mu+1)}^{\xi(\mu+1)-1} \alpha_s\right)^{1/2}, \\ &\text{(ii)} \quad \nu > \left(\max_{\xi=1,\dots,\lfloor L/\mu\rfloor} \prod_{s=(\xi-1)(\mu+1)+1}^{\xi(\mu+1)} \frac{\alpha_s}{\beta_s}\right)^{1/2}. \end{aligned} \tag{4.3}$$

Thus, properly choosing polynomial degrees we have an optimal rate of convergence, i.e., the condition number of $M^{-1}A$ is the magnitude of $O(1)$. Unfortunately, the above analysis does not ensure existence of the parameters μ and ν for which conditions (3.2) and (4.3) on polynomial degrees are satisfied, because the theoretical investigation of α_s and β_s is a very difficult problem, that depends on the value of θ parameter and the choice of the approximation pattern in (2.4). In the next section we will discuss this subject on theoretical grounds of the iterative incomplete factorization method.

5. Theoretical estimates for α_s and β_s

First of all, we have to note that the problem for definition of the spectral bounds of $A^{(k+1)^{-1}}S^{(k+1)}$ is equal to the generalized eigenvalue problem $A^{(k)}v = \lambda \hat{M}^{(k)}v$, where

$$\hat{M}^{(k)} = \begin{bmatrix} A_{11}^{(k)} & 0 \\ A_{21}^{(k)} & I \end{bmatrix} \begin{bmatrix} I & (A_{11}^{(k)})^{-1}A_{12}^{(k)} \\ 0 & A^{(k+1)} \end{bmatrix} \begin{matrix} \} X_k \setminus X_{k+1} \\ \} X_{k+1} \end{matrix} . \tag{5.1}$$

Now to obtain the upper bound on α_{k+1} with respect to a variation of θ_{k+1} we first consider the case when $\theta_{k+1} = 1$. To do this we use the well-known iterative incomplete factorization technique (for detail, see [9]) from which it follows

$$\lambda_{\max}\left((\hat{M}^{(k)})^{-1}A^{(k)}\right) \leq \frac{1}{1-\tau_k}, \tag{5.2}$$

where

$$\tau_k = \max_i \left\{ \left((A_{11}^{(k)})^{-1}A_{12}^{(k)}e\right)_i \right\} < 1. \tag{5.3}$$

From this result we can see that the upper bound on the maximal eigenvalue depends on τ_k for which the inequality

$$\tau_k I e \geq (A_{11}^{(k)})^{-1}A_{12}^{(k)}e \tag{5.4}$$

holds. Here I is an identical matrix and the usual componentwise relation between real matrices (vectors) is used, i.e., $A \geq B$ ($a \geq b$) if $A_{ij} \geq B_{ij}$ ($a_i \geq b_i$) for all i, j.

Unfortunately, it is well-known that the main drawback of the iterative incomplete factorization method is an unpredictable growing of its largest eigenvalue, and hence τ_k, see [8, 10, 17] for example.

Now let us dream a little bit. Analysing (5.4) one can find that we can *control* this growing by including a "fiction" parameter ω into the right-hand side, i.e., if we assume that there is a parameter ω such that $\omega \leq \tau_k^{-1}(1-\gamma)$, where γ is an arbitrary chosen positive value, then the inequality

$$(1-\gamma)Ie \geq \omega (A_{11}^{(k)})^{-1} A_{12}^{(k)} e$$

is valid, and hence, $\lambda_{\max}\Big((\hat{M}^{(k)})^{-1}A^{(k)}\Big) \leq \gamma^{-1}$. Thus, if we find an perturbation ω of the action of $(A_{11}^{(k)})^{-1}$ when the condition $\tau_k \leq 1-\gamma$ is violated, then the desired upper bound γ^{-1} will be satisfied.

On the other hand, we can rewrite (2.4) and (2.6) as follows:

$$A^{(k+1)} = A_{22}^{(k)} - A_{21}^{(k)}\overline{(A_{11}^{(k)})^{-1}} A_{12}^{(k)} - \hat{Q}^{(k+1)},$$
$$\hat{Q}^{(k+1)}e = A_{21}^{(k)}\left(\theta_{k+1}(A_{11}^{(k)})^{-1} - \overline{\theta_{k+1}(A_{11}^{(k)})^{-1}}\right) A_{12}^{(k)} e,$$

i.e., we can consider our method as the iterative incomplete factorization method (when the relaxation parameter is equal to one) with a special choice of approximation for the inverse first pivoting block.

Moreover, let us also recall one simple and useful result that was proven earlier.

Lemma [14]. *Let $F \geq 0$ be a strictly upper triangular matrix and P, Q be nonnegative diagonal matrices. Let B be a matrix such that*

$$Bx \geq 0 \quad \textit{and} \quad \operatorname{offdiag}(B) = \operatorname{offdiag}((P-E)Q(P-F)),$$

where x is some positive vector, offdiag(B) *denotes the off-diagonal part of B and $E = F^T$. If $Px \geq Fx$, then B is nonnegative definite matrix.*

Now we combine these ideas and results with a help of a technique, suggested in [14], to prove the next theorem.

Theorem 1. *Let $A^{(k)}$ be a diagonally dominant Stieltjes matrix defined by* (2.2), *$A^{(k+1)}$ be defined by* (2.4) *and* (2.6), *the preconditioning matrix $\hat{M}^{(k)}$ be defined by* (5.1). *The parameter τ_k is defined by* (5.3) *and, moreover, the additional condition $(1+\theta_{k+1})\tau_k < 2$ holds. Then*

$$\lambda_{\max}\Big((M^{(k)})^{-1}A^{(k)}\Big) \leq \gamma^{-1} = \frac{2}{2-(1+\theta_{k+1})\tau_k}.$$

Proof. Consider the positive block-diagonal matrix W with diagonal blocks such that

$$W_1 = \begin{cases} I & \tau_k \leq 1-\gamma, \\ \tau_k^{-1}(1-\gamma)I & \text{otherwise} \end{cases} \quad \text{and} \quad W_2 = I,$$

and let T be the positive block-diagonal matrix with diagonal blocks such that

$$T_1 = \begin{cases} (1-\gamma)I & \tau_k \le 1-\gamma, \\ \tau_k I & \text{otherwise} \end{cases} \quad \text{and} \quad T_2 = (1-\gamma)I.$$

Hence, by the definitions we have that $W \le I$, $T \le I$ and

$$TW = WT = (1-\gamma)I.$$

Then we define a matrix B_1 as follows

$$\begin{aligned} B_1 &= \hat{D}_1 + \begin{bmatrix} T_1 & 0 \\ A_{21}^{(k)}(A_{11}^{(k)})^{-1} & T_2 \end{bmatrix} \begin{bmatrix} W_1 & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A_{11}^{(k)} & 0 \\ 0 & A^{(k+1)} \end{bmatrix} \begin{bmatrix} T_1 & (A_{11}^{(k)})^{-1}A_{12}^{(k)} \\ 0 & T_2 \end{bmatrix} \\ &= \hat{D}_1 + \begin{bmatrix} T_1 W_1 A_{11}^{(k)} T_1 & W_1 T_1 A_{21}^{(k)} \\ A_{21}^{(k)} W_1 T_1 & (1-\gamma)A^{(k+1)}T_2 + A_{21}^{(k)} W_1 (A_{11}^{(k)})^{-1} A_{12}^{(k)} \end{bmatrix}, \end{aligned}$$

where $\hat{D}_1$ is a block–diagonal matrix with diagonal blocks such that $B_1 e = 0$. Hence, by application of the lemma it is easy to show that B is a nonnegetive definite matrix.

Next we define a matrix B_2 as follows:

$$B_2 = \hat{D}_2 + \begin{bmatrix} (1-\gamma)A_{11}^{(k)}(I - T_1) & 0 \\ 0 & (1-\gamma)A^{(k+1)}(I - T_2) \end{bmatrix},$$

where $\hat{D}_2$ is a block-diagonal matrix with diagonal blocks such that $B_2 e = 0$. Moreover, we have

$$\text{offdiag}(A_{11}^{(k)}(I - T_1)) \le 0 \quad \text{and} \quad \text{offdiag}(A^{(k+1)}(I - T_2)) \le 0.$$

Hence, using the following well-known fact from the theory of the Stieltjes matrices: "Let B be a matrix such that $Bx \ge 0$ and offdiag$(B) \le 0$, where x is some positive vector, then B is nonnegative definite", it is easy to show that B_2 is nonnegative definite matrix.

Now we construct the auxiliary matrix R defined by

$$R = \hat{M}^{(k)} - \gamma A^{(k)} - B_1 - B_2,$$

and we want show its positive definiteness.

By definitions we have

$$\begin{aligned} R = & \begin{bmatrix} A_{11}^{(k)} & A_{12}^{(k)} \\ A_{21}^{(k)} & A^{(k+1)} + A_{21}^{(k)}(A_{11}^{(k)})^{-1}A_{12}^{(k)} \end{bmatrix} - \gamma \begin{bmatrix} A_{11}^{(k)} & A_{12}^{(k)} \\ A_{21}^{(k)} & A_{22}^{(k)} \end{bmatrix} - \\ & \hat{D}_1 - \begin{bmatrix} T_1 W_1 A_{11}^{(k)} T_1 & W_1 T_1 A_{21}^{(k)} \\ A_{21}^{(k)} W_1 T_1 & (1-\gamma)A^{(k+1)}T_2 + A_{21}^{(k)} W_1 (A_{11}^{(k)})^{-1} A_{12}^{(k)} \end{bmatrix} - \end{aligned}$$

$$
\begin{aligned}
&\hat{D}_2 - \begin{bmatrix} (1-\gamma)A_{11}^{(k)}(I-T_1) & 0 \\ 0 & (1-\gamma)A^{(k+1)}(I-T_2) \end{bmatrix} \\
&= \begin{bmatrix} (1-\gamma)A_{11}^{(k)} & (1-\gamma)A_{12}^{(k)} \\ (1-\gamma)A_{21}^{(k)} & A^{(k+1)} + A_{21}^{(k)}(I-W_1)(A_{11}^{(k)})^{-1}A_{12}^{(k)} - \gamma A_{22}^{(k)} \end{bmatrix} - \\
&\hat{D}_1 - \begin{bmatrix} (1-\gamma)A_{11}^{(k)}T_1 & (1-\gamma)A_{21}^{(k)} \\ (1-\gamma)A_{21}^{(k)} & (1-\gamma)A^{(k+1)}T_2 \end{bmatrix} - \\
&\hat{D}_2 - \begin{bmatrix} (1-\gamma)A_{11}^{(k)}(I-T_1) & 0 \\ 0 & (1-\gamma)A^{(k+1)}(I-T_2) \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 \\ 0 & \gamma A^{(k+1)} + A_{21}^{(k)}(I-W_1)(A_{11}^{(k)})^{-1}A_{12}^{(k)} - \gamma A_{22}^{(k)} \end{bmatrix} - \hat{D}_1 - \hat{D}_2.
\end{aligned}
$$

Now using (2.4) rewrite the second nonzero diagonal block in the following form:

$$
\begin{aligned}
&\gamma A^{(k+1)} + A_{21}^{(k)}(I-W_1)(A_{11}^{(k)})^{-1}A_{12}^{(k)} - \gamma A_{22}^{(k)} \\
&\quad = \gamma\left[A^{(k+1)} - A_{22}^{(k)}\right] + A_{21}^{(k)}(I-W_1)\overline{(A_{11}^{(k)})^{-1}}A_{12}^{(k)} + \\
&\quad\quad A_{21}^{(k)}(I-W_1)\left[(A_{11}^{(k)})^{-1} - \overline{(A_{11}^{(k)})^{-1}}\right]A_{12}^{(k)} \\
&\quad = \gamma\left[-A_{21}^{(k)}\overline{(A_{11}^{(k)})^{-1}}A_{12}^{(k)} - \theta_{k+1}Q^{(k+1)}\right] + A_{21}^{(k)}(I-W_1)\overline{(A_{11}^{(k)})^{-1}}A_{12}^{(k)} + \\
&\quad\quad A_{21}^{(k)}(I-W_1)\left[(A_{11}^{(k)})^{-1} - \overline{(A_{11}^{(k)})^{-1}}\right]A_{12}^{(k)}.
\end{aligned}
$$

Futher introduce once more the diagonal matrix $(\tilde{D}_3)_{22}$ such that

$$
\left[(\tilde{D}_3)_{22} - \gamma A_{21}^{(k)}\overline{(A_{11}^{(k)})^{-1}}A_{12}^{(k)} + A_{21}^{(k)}(I-W_1)\overline{(A_{11}^{(k)})^{-1}}A_{12}^{(k)}\right]e = 0.
$$

Note that by the definition of W_1 we have

$$
\begin{aligned}
&\operatorname{offdiag}\left(A_{21}^{(k)}(I-W_1)\overline{(A_{11}^{(k)})^{-1}}A_{12}^{(k)} - \gamma A_{21}^{(k)}\overline{(A_{11}^{(k)})^{-1}}A_{12}^{(k)}\right) \\
&\qquad = -\operatorname{offdiag}\left(\gamma A_{21}^{(k)}\overline{(A_{11}^{(k)})^{-1}}A_{12}^{(k)}\right) \le 0
\end{aligned}
$$

when $W_1 = I$, and

$$\begin{aligned}
&\operatorname{offdiag}\left(A_{21}^{(k)}(I-W_1)\overline{(A_{11}^{(k)})^{-1}}A_{12}^{(k)} - \gamma A_{21}^{(k)}\overline{(A_{11}^{(k)})^{-1}}A_{12}^{(k)}\right)\\
&\quad = \left(1-\tau_k^{-1}(1-\gamma)-\gamma\right)\operatorname{offdiag}\left(A_{21}^{(k)}\overline{(A_{11}^{(k)})^{-1}}A_{12}^{(k)}\right)\\
&\quad = (1-\gamma)(1-\tau_k^{-1})\operatorname{offdiag}\left(A_{21}^{(k)}\overline{(A_{11}^{(k)})^{-1}}A_{12}^{(k)}\right) \le 0
\end{aligned}$$

when $W_1 = \tau_k^{-1}(1-\gamma)I$, and hence, the matrix

$$(\tilde{D}_3)_{22} - \gamma A_{21}^{(k)}\overline{(A_{11}^{(k)})^{-1}}A_{12}^{(k)} + A_{21}^{(k)}(I-W_1)\overline{(A_{11}^{(k)})^{-1}}A_{12}^{(k)}$$

is nonnegative definite. Thus, we obtain

$$R = C + D + \begin{bmatrix} 0 & 0 \\ 0 & A_{21}^{(k)}(I-W_1)\left[(A_{11}^{(k)})^{-1} - \overline{(A_{11}^{(k)})^{-1}}\right]A_{12}^{(k)} \end{bmatrix},$$

where C is a nonnegative block-diagonal matrix:

$$C_{11} = 0, \quad C_{22} = (\tilde{D}_3)_{22} - \gamma A_{21}^{(k)}\overline{(A_{11}^{(k)})^{-1}}A_{12}^{(k)} + A_{21}^{(k)}(I-W_1)\overline{(A_{11}^{(k)})^{-1}}A_{12}^{(k)}$$

and D is a diagonal matrix:

$$D = -\hat{D}_1 - \hat{D}_2 - \begin{bmatrix} 0 & 0 \\ 0 & (\tilde{D}_3)_{22} + \theta_{k+1}Q^{(k+1)} \end{bmatrix}.$$

On the other hand, from the definition of the preconditioning matrix $\hat{M}^{(k)}$ we have

$$\hat{M}^{(k)}e = A^{(k)}e + \begin{bmatrix} 0 & 0 \\ 0 & (1-\theta_{k+1})A_{21}^{(k)}\left[(A_{11}^{(k)})^{-1} - \overline{(A_{11}^{(k)})^{-1}}\right]A_{12}^{(k)} \end{bmatrix} e,$$

and hence

$$Re = (1-\gamma)A^{(k)}e + \begin{bmatrix} 0 & 0 \\ 0 & (1-\theta_{k+1})A_{21}^{(k)}\left[(A_{11}^{(k)})^{-1} - \overline{(A_{11}^{(k)})^{-1}}\right]A_{12}^{(k)} \end{bmatrix} e,$$

from what follows that

$$\begin{aligned}
D_{11}e &= (1-\gamma)(A^{(k)}e)_1 \ge 0,\\
D_{22}e &= (1-\gamma)(A^{(k)}e)_2 + A_{21}^{(k)}(W_1 - \theta_{k+1}I)\left[(A_{11}^{(k)})^{-1} - \overline{(A_{11}^{(k)})^{-1}}\right]A_{12}^{(k)}e\\
&\ge A_{21}^{(k)}(W_1 - \theta_{k+1}I)\left[(A_{11}^{(k)})^{-1} - \overline{(A_{11}^{(k)})^{-1}}\right]A_{12}^{(k)}e.
\end{aligned}$$

Here we use the fact that $A^{(k)}$ is a strictly diagonally dominant matrix.

Next due to the following well-known fact from the theory of H-matrices: "Let $\mathcal{M}(B)$ be defined as

$$(\mathcal{M}(B))_{ij} = \begin{cases} |b_{ii}| & \text{for } i = j, \\ -|b_{ij}| & \text{for } i \neq j, \end{cases}$$

and be a nonnegative definite matrix, then B is nonnegative definite," the matrix

$$D_{22} + A_{21}^{(k)}(I - W_1)\left[(A_{11}^{(k)})^{-1} - \overline{(A_{11}^{(k)})^{-1}}\right] A_{12}^{(k)}$$

will be nonnegative definite if $W_1 - \theta_{k+1} I \geq I - W_1$. The latter is valid by the assumption of the theorem. Therefore the matrix

$$D + \begin{bmatrix} 0 & 0 \\ 0 & A_{21}^{(k)}(I - W_1)\left[(A_{11}^{(k)})^{-1} - \overline{(A_{11}^{(k)})^{-1}}\right] A_{12}^{(k)} \end{bmatrix}$$

is nonnegative definite. □

Corollary. *To achieve the desired upper bound* γ^{-1} *we have to set*

$$\theta_{k+1} = \frac{2}{\tau_k}(1 - \gamma) - 1.$$

From the results of Theorem 1 we directly obtain the upper bound on α_{k+1} with respect of θ_{k+1}:

$$\alpha_{k+1} \leq \frac{2}{2 - (1 + \theta_{k+1})\tau_k}. \tag{5.5}$$

Proceed to the lower bound on β_{k+1} with respect of a variation of θ_{k+1}. To do this we use the following estimate that has been obtained in [7, 13]:

$$\lambda_{\min}\left((\hat{M}^{(k)})^{-1} A^{(k)}\right) \geq \frac{1}{1 + \xi}, \tag{5.6}$$

where

$$\xi \leq \frac{(1 - \theta_{k+1})(Q_{k+1} e_2, e_2)}{(A_{11}^{(k)} e_1, e_1) + (A_{22}^{(k)} e_2, e_2)} \cdot \lambda_{\min}^{-1}\left(\begin{bmatrix} A_{11}^{(k)} & 0 \\ 0 & A_{22}^{(k)} \end{bmatrix}^{-1} \begin{bmatrix} A_{11}^{(k)} & A_{12}^{(k)} \\ A_{21}^{(k)} & A_{22}^{(k)} \end{bmatrix}\right), \tag{5.7}$$

where $e = (e_1, e_2)^T$. It is known that in practice $\lambda_{\min}([\cdot]^{-1} A)$ is of order $O(n_k^{-1})$, and the order of $(A_{11}^{(k)} e_1, e_1) + (A^{(k+1)} e_2, e_2)$ is $O(n_k)$. Hence, the order of $\lambda_{\min}((\hat{M}^{(k)})^{-1} A^{(k)})$ is proportional to the sum of all perturbations.

Now we collect the above results into

Theorem 2. *Let $A^{(k+1)}$ be a Stieltjes matrix defined by* (2.4) *and* (2.6), *and* $S^{(k+1)}$ *be defined by* (2.3). *Then*

$$\frac{1}{1+C(1-\theta_{k+1})(Q_{k+1}e,e)} \leq \frac{(S^{(k+1)}x,x)}{(A^{(k+1)}x,x)} \leq \frac{2}{2-(1+\theta_{k+1})\tau_k}$$

for all $x \in R^{n_{k+1}}$. *Here* C *is a positive constant, independent of* θ_{k+1} *and* Q_{k+1}, *but it depends on the approximation pattern in* (2.4), *and* τ_k *is defined by* (5.3).

Finally, using derived results we rewrite conditions (3.2) and (4.3):

$$\text{(i)} \quad \left(\max_{\xi=1,\dots,\lfloor L/\mu \rfloor} \prod_{s=(\xi-1)(\mu+1)}^{\xi(\mu+1)-1} \frac{2}{2-(1+\theta_{k+1})\tau_k} \right)^{1/2} < \nu < \rho^{\mu+1},$$

$$\text{(ii)} \quad \left(\max_{\xi=1,\dots,\lfloor L/\mu \rfloor} \prod_{s=(\xi-1)(\mu+1)+1}^{\xi(\mu+1)} 2\frac{1+C(1-\theta_{k+1})(Q_{k+1}e,e)}{2-(1+\theta_{k+1})\tau_k} \right)^{1/2} < \nu < \rho^{\mu+1},$$

from which we directly derive the condition on the value of relaxation parameter θ_{k+1}:

$$\begin{aligned} &\text{(i)} \quad \theta_{k+1} < \frac{2}{\tau_k}\left(1-\rho^{-2(1+\frac{1}{\mu})}\right) - 1, \\ &\text{(ii)} \quad \frac{\frac{2}{\tau_k}\left[1-(1+C(Q_{k+1}e,e))\,\rho^{-2(1+\frac{1}{\mu})}\right]-1}{1-\frac{2}{\tau_k}C(Q_{k+1}e,e)\rho^{-2(1+\frac{1}{\mu})}} < \theta_{k+1} \leq 1, \end{aligned} \tag{5.8}$$

under which conditions (3.2) and (4.3) on polynomial degrees are satisfied.

Thus, the preliminary results we have reported, demonstrate the global relation between the parameters μ and ν from one hand and the relaxation parameter θ_s on the other hand. Futher we plan to consider this theoretical result in application to the concrete definitions both of the sparsity structure in (2.4) and of the block splitting, i.e., construction of $\{X_k\}$. Another way to improve the quality of the preconditioner which is in progress now [12], is in that we want to use a "local" (choice) relaxation matrix Θ_{k+1} instead of a "global" (uniform) relaxation matrix $\theta_{k+1}I$.

References

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

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

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

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

[5] 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.

[6] O. Axelsson and M. Larin, *An algebraic multilevel iteration method for finite element matrices*, Journal of Applied and Computational Mathematics, **2**, 1998 (to appear).

[7] R. Beauwens, *Modified Incomplete factorization strategies*, Lecture Notes in Mathematics, Springer Verlag, New York, **1457**, 1990, 1–16.

[8] T.F. Chan, *Fourier analysis of relaxed incomplete factorization preconditioners*, SIAM J. Sci. Stat. Comput., **12**, 1991, 668–680.

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

[10] V.P. Il'in and M. Larin, *An iterative multilevel incomplete factorization method for solving five-point system of equations*, Numerical Mathematics, **5**, NCC Publisher, Novosibirsk, 1996, 69–88 (in Russian).

[11] M. Larin, *Algebraic multilevel incomplete factorization method for Stieltjes matrices*, Journal of Computational Mathematics and Mathematical Physics, **38**, 1998 (to appear) (in Russian).

[12] M. Larin, *Block DRIC method* (in progress).

[13] Y. Notay, *Solving positive (semi)definite linear systems by preconditioned iterative methods*, Lecture Notes in Mathematics, Springer Verlag, New York, **1457**, 1990, 105–125.

[14] Y. Notay, *DRIC: a Dynamic Version of the RIC Method*, Numer. Lin. Alg. Appl., **1**, 1994, 511–532.

[15] 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, ed. A.S. Alekseev and N.S. Bakhvalov, NCC Publisher, 1995, 435–446.

[16] B. Parlett, *The Symmetric Eigenvalue Problem*, Prentice–Hall, Inc., 1980.

[17] H.A. van der Vorst, *ICCG and related methods for 3D problems on vector computers*, Comp. Phys. Commun., **53**, 1989, 223–235.