



ELSEVIER

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Journal of Computational and Applied Mathematics 187 (2006) 202–226

JOURNAL OF
COMPUTATIONAL AND
APPLIED MATHEMATICS

www.elsevier.com/locate/cam

Restrictive preconditioners for conjugate gradient methods for symmetric positive definite linear systems[☆]

Zhong-Zhi Bai*, Zeng-Qi Wang

State Key Laboratory of Scientific/Engineering Computing, Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and System Sciences, Chinese Academy of Sciences, P.O. Box 2719, Beijing 100080, PR China

Received 15 August 2004; accepted 28 February 2005

Abstract

The restrictively preconditioned conjugate gradient (RPCG) method for solving large sparse system of linear equations of a symmetric positive definite and block two-by-two coefficient matrix is further studied. In fact, this RPCG method is essentially the classical preconditioned conjugate gradient (PCG) method with a specially structured preconditioner. Within this setting, we present algorithmic descriptions of two restrictive preconditioners that, respectively, employ the block Jacobi and the block symmetric Gauss–Seidel matrix splitting matrices as approximations to certain matrices involved in them, and give convergence analyses of the correspondingly induced two PCG methods. Numerical results show that these restrictive preconditioners can lead to practical and effective PCG methods for solving large sparse systems of linear equations of symmetric positive definite and block two-by-two coefficient matrices.

© 2005 Elsevier B.V. All rights reserved.

MSC: 65F10; 65F50; CR: G1.3

Keywords: Linear system; Block two-by-two matrix; Symmetric positive definite matrix; Restrictive preconditioner; Conjugate gradient method

[☆] Subsidized by The Special Funds For Major State Basic Research Projects (no. G1999032803) and The National Natural Science Foundation (no. 10471146), PR China.

* Corresponding author. Tel.: +86 10 6263 0992; fax: +86 10 6254 5820.

E-mail addresses: bzz@lsec.cc.ac.cn (Z.-Z. Bai), zqwang@lsec.cc.ac.cn (Z.-Q. Wang).

1. Introduction

Let \mathbb{R}^n represent the real n -dimensional vector space, and $\mathbb{R}^{n \times n}$ the real $n \times n$ matrix space. In this paper, we will further study several practical restrictive preconditioners and the corresponding preconditioned conjugate gradient methods, called the *restrictively preconditioned conjugate gradient* (RPCG) methods [9], for solving the large sparse systems of linear equations whose coefficient matrices are symmetric positive definite and are of block two-by-two structure.

In general, we first consider the system of linear equations

$$Ax = b, \quad A \in \mathbb{R}^{n \times n} \text{ nonsingular and } x, b \in \mathbb{R}^n, \quad (1.1)$$

where $A \in \mathbb{R}^{n \times n}$ is a large sparse matrix, $x \in \mathbb{R}^n$ the unknown vector, and $b \in \mathbb{R}^n$ a given right-hand-side vector. If $A \in \mathbb{R}^{n \times n}$ is isomorphic to a symmetric positive definite matrix $H \in \mathbb{R}^{n \times n}$, i.e., there exist two nonsingular matrices $P, Q \in \mathbb{R}^{n \times n}$ such that $A = PHQ$, then we can construct a restrictive preconditioner $M = PWQ$ to A , where $W \in \mathbb{R}^{n \times n}$ is a good approximation to $H \in \mathbb{R}^{n \times n}$ and is also a symmetric positive definite matrix (see, for instance [19]). Let $K = Q^{-1}P^T$. Then we can describe the RPCG method [9] for the system of linear equations (1.1) as follows.

Method 1.1 (THE RPCG METHOD FOR $Ax = b$).

Choose $x_0 \in \mathbb{R}^n$, $r_0 = b - Ax_0$.

Solve $Mz_0 = r_0$, and set $p_0 = z_0$.

Solve $Kv_0 = z_0$, and set $q_0 = v_0$.

For $k = 0, 1, 2, \dots$

$$\alpha_k = -\frac{v_k^T r_k}{q_k^T A p_k}$$

$$x_{k+1} = x_k - \alpha_k p_k$$

$$r_{k+1} = r_k + \alpha_k A p_k$$

$$\text{solve } Mz_{k+1} = r_{k+1}$$

$$\text{solve } Kv_{k+1} = z_{k+1}$$

$$\beta_k = \frac{v_{k+1}^T r_{k+1}}{v_k^T r_k}$$

$$p_{k+1} = z_{k+1} + \beta_k p_k$$

$$q_{k+1} = v_{k+1} + \beta_k q_k.$$

Here, $K = Q^{-1}P^T$, and $P, Q \in \mathbb{R}^{n \times n}$ are two nonsingular matrices such that $A = PHQ$ with $H \in \mathbb{R}^{n \times n}$ being a symmetric positive definite matrix, and $M = PWQ$ is the preconditioner of A with W being an approximation to H .

This approach is not only applicable to many standard Krylov subspace methods such as the *conjugate gradient* (CG) [15,12], the conjugate residual [12,13], the CGNR [12,13,18], and the CGNE [12,13,18] methods as well as their preconditioned variants, but also yields many new ones, for solving the system of linear equations (1.1). For details, we refer to [9]. When it is applied to the system of linear

equations (1.1) with the block two-by-two symmetric positive definite coefficient matrix

$$A = \begin{bmatrix} B & E \\ E^T & C \end{bmatrix} \in \mathbb{R}^{n \times n} \quad \text{with } B \in \mathbb{R}^{m \times m}, C \in \mathbb{R}^{\ell \times \ell} \text{ and } m \geq \ell, \quad (1.2)$$

the RPCG method naturally reduces to a special but very effective *preconditioned conjugate gradient* (PCG) method for solving the system of linear equations (1.1)–(1.2). For its detailed description, we refer to [9]. We remark that the case $m \leq \ell$ can be studied in a similar fashion.

At each iteration step of the RPCG method for the system of linear equations (1.1)–(1.2), we need to compute the exact solutions of two sub-systems of linear equations with different right-hand-side vectors but the same coefficient matrix $B \in \mathbb{R}^{m \times m}$ (see [9, Method 3.1]). This is very costly and impractical in actual applications, in particular, when the size of the matrix block B is very large. Therefore, studying variants of this RPCG method that can avoid exact inversion of the matrix block B is of both theoretical and practical importance.

In this paper, we present practical strategies of constructing the transformation matrices P and Q , as well as the approximation matrix W in the RPCG method for the block two-by-two symmetric positive definite system of linear equations (1.1)–(1.2), and consequently, obtain several practical and efficient restrictive preconditioners which lead to iteration methods within the framework of RPCG. These new preconditioners avoid exact inversions of the matrix blocks B and C , and only require inexact solutions of the sub-systems of linear equations with the coefficient matrices $B \in \mathbb{R}^{m \times m}$ and $C \in \mathbb{R}^{\ell \times \ell}$, which are involved in the generalized residual equation at each iteration step of the RPCG method. Therefore, they can result in “inverse-free” variants of the RPCG method for the system of linear equations (1.1)–(1.2) established in [9]. Under reasonable assumptions on the qualities of the approximations to the matrix blocks B and C , we prove the convergence and estimate the convergence rates of these methods. In particular, we present algorithmic descriptions of two special restrictive preconditioners that result from the modified block Jacobi and the modified block symmetric Gauss–Seidel (BSGS) splittings of the matrix W [7,8], and give a convergence theory for the correspondingly induced RPCG methods. Moreover, we also give practical techniques for constructing high-quality approximations to the matrix blocks B and C , as well as the Schur complement of the matrix A . Numerical experiments are implemented for systems of linear equations of form (1.1)–(1.2) that arise from the finite difference discretization of equidistant stepsize, incorporated with the domain decomposition technique, of a second-order self-adjoint elliptic partial differential equations [17], and the results show that our new methods are more robust and effective than the classical PCG methods [15,2,12,18,13].

The organization of the paper is as follows. After establishing the general framework of the practical RPCG methods and demonstrating its convergence theory in Section 2, we describe the algebraic constructions of two special restrictive preconditioners associated with the block Jacobi and the block symmetric Gauss–Seidel splittings, and demonstrate the convergence theorems of the correspondingly induced RPCG methods in Section 3. Practical techniques for constructing high-quality approximations to the matrix blocks B and C are investigated in Section 4, and numerical results are shown in Section 5. Finally, in Section 6, we use a brief conclusion and several remarks to end the paper.

2. General description of practical RPCG methods

For the symmetric positive definite system of linear equations (1.1)–(1.2), in this section we will establish a general paradigm of the restrictive preconditioners and, consequently, obtain a general framework

of practical RPCG methods based on Method 1.1. In addition, we will study their convergence property as well. We remark that the RPCG methods within this framework only involve solutions of sub-systems of linear equations whose coefficient matrices are a given approximation of the matrix block B other than B itself as is involved in the RPCG in [9], or in other words, they allow inexact solutions of the sub-systems of linear equations with the coefficient matrix B and, hence, are inexact variants of those discussed in [9].

Because $A \in \mathbb{R}^{n \times n}$ is symmetric positive definite, its diagonal blocks $B \in \mathbb{R}^{m \times m}$ and $C \in \mathbb{R}^{\ell \times \ell}$ are symmetric positive definite, too. Let $L_B \in \mathbb{R}^{m \times m}$ and $L_C \in \mathbb{R}^{\ell \times \ell}$ be nonsingular matrices such that

$$L_B^{-1} B L_B^{-T} = J_B \quad \text{and} \quad L_C^{-1} C L_C^{-T} = J_C. \quad (2.1)$$

Here we require that

$$J_B \approx I \quad \text{and} \quad J_C \approx I$$

hold, with I the identity matrix whose dimension can be inferred from the context. Take

$$P = \begin{bmatrix} L_B & O \\ E^T L_B^{-T} & L_C \end{bmatrix} \quad \text{and} \quad Q = \begin{bmatrix} L_B^T & L_B^{-1} E \\ O & L_C^T \end{bmatrix} = P^T, \quad (2.2)$$

where O denotes the zero matrix. Then by direct computations we have

$$A = P H Q = P H P^T,$$

where

$$H = \begin{bmatrix} J_B & (I - J_B) \bar{E} \\ \bar{E}^T (I - J_B) & J_C - \bar{E}^T \bar{E} - \bar{E}^T (I - J_B) \bar{E} \end{bmatrix} \quad (2.3)$$

is a symmetric positive definite matrix,

$$\bar{E} = L_B^{-1} E L_C^{-T}$$

and

$$K = Q^{-1} P^T = P^{-T} P^T = I.$$

Now, we can define the restrictive preconditioner as

$$M = P W P^T,$$

where

$$W \approx \bar{W} := \begin{bmatrix} I & (I - J_B) \bar{E} \\ \bar{E}^T (I - J_B) & \bar{S} \end{bmatrix} \approx H \quad \text{and} \quad \bar{S} = I - \bar{E}^T \bar{E}. \quad (2.4)$$

Because

$$P^{-1} = \begin{bmatrix} L_B^{-1} & O \\ -L_C^{-1} E^T L_B^{-T} L_B^{-1} & L_C^{-1} \end{bmatrix},$$

by letting

$$t = (t^{(1)T}, t^{(2)T})^T \equiv P^{-1} r = P^{-1} (r^{(1)T}, r^{(2)T})^T$$

and

$$\tilde{t} = (\tilde{t}^{(1)T}, \tilde{t}^{(2)T})^T \equiv P^T z = P^T (z^{(1)T}, z^{(2)T})^T$$

with

$$r^{(1)}, z^{(1)}, t^{(1)}, \tilde{t}^{(1)} \in \mathbb{R}^m \quad \text{and} \quad r^{(2)}, z^{(2)}, t^{(2)}, \tilde{t}^{(2)} \in \mathbb{R}^\ell,$$

we can solve the generalized residual equation $Mz = r$ involved in Method 1.1 through the formulas:

- $L_B t^{(1)} = r^{(1)}, L_B^T \bar{t}^{(1)} = t^{(1)};$
- $L_C t^{(2)} = r^{(2)} - E^T \bar{t}^{(1)};$
- $W \tilde{t} = t;$
- $L_C^T z^{(2)} = \tilde{t}^{(2)};$
- $L_B \bar{t}^{(1)} = E z^{(2)}, L_B^T z^{(1)} = \tilde{t}^{(1)} - \bar{t}^{(1)}.$

Consequently, we obtain the following special form of Method 1.1 for solving the symmetric positive definite system of linear equations (1.1)–(1.2).

Method 2.1 (THE RPCG METHOD FOR LINEAR SYSTEM (1.1)–(1.2)).

Choose $x_0 \in \mathbb{R}^n, r_0 = b - Ax_0$.

Let $r_0 = (r_0^{(1)T}, r_0^{(2)T})^T, z_0 = (z_0^{(1)T}, z_0^{(2)T})^T$, and

solve $L_B t^{(1)} = r_0^{(1)}$ and $L_B^T \bar{t}^{(1)} = t^{(1)}$

solve $L_C t^{(2)} = r_0^{(2)} - E^T \bar{t}^{(1)}$

solve $W \tilde{t} = t$, with $t = (t^{(1)T}, t^{(2)T})^T$ and $\tilde{t} = (\tilde{t}^{(1)T}, \tilde{t}^{(2)T})^T$

solve $L_C^T z_0^{(2)} = \tilde{t}^{(2)}$

solve $L_B \bar{t}^{(1)} = E z_0^{(2)}$ and $L_B^T z_0^{(1)} = \tilde{t}^{(1)} - \bar{t}^{(1)}.$

Set $p_0 = z_0$.

For $k = 0, 1, 2, \dots$

$$\alpha_k = -\frac{z_k^T r_k}{p_k^T A p_k}$$

$$x_{k+1} = x_k - \alpha_k p_k$$

$$r_{k+1} = r_k + \alpha_k A p_k$$

Let $r_{k+1} = (r_{k+1}^{(1)T}, r_{k+1}^{(2)T})^T, z_{k+1} = (z_{k+1}^{(1)T}, z_{k+1}^{(2)T})^T$, and

solve $L_B t^{(1)} = r_{k+1}^{(1)}$ and $L_B^T \bar{t}^{(1)} = t^{(1)}$

solve $L_C t^{(2)} = r_{k+1}^{(2)} - E^T \bar{t}^{(1)}$

solve $W \tilde{t} = t$, with $t = (t^{(1)T}, t^{(2)T})^T$ and $\tilde{t} = (\tilde{t}^{(1)T}, \tilde{t}^{(2)T})^T$

solve $L_C^T z_{k+1}^{(2)} = \tilde{t}^{(2)}$

solve $L_B \bar{t}^{(1)} = E z_{k+1}^{(2)}$ and $L_B^T z_{k+1}^{(1)} = \tilde{t}^{(1)} - \bar{t}^{(1)}$

$$\beta_k = \frac{z_{k+1}^T r_{k+1}}{z_k^T r_k}$$

$$p_{k+1} = z_{k+1} + \beta_k p_k.$$

In fact, Method 2.1 is essentially the classical PCG method with the special structured preconditioner $M = PWP^T$, or in other words, it is the conjugate gradient method in the M^{-1} -inner product (see [11,14,1] and references therein).

Because the block two-by-two matrix A in (1.2) is symmetric positive definite, from [2, Lemma 9.1] we know that the strengthened Cauchy–Buniakowskii–Schwarz (CBS) inequality holds, i.e., there exists a nonnegative constant $\gamma \in [0, 1)$ such that

$$|y^T Ez| \leq \gamma (y^T By)^{1/2} (z^T Cz)^{1/2}. \quad (2.5)$$

It turns out that in many applications, the constant γ does not depend on the dimension of the matrix. As has been shown in [3,10], this situation occurs, for instance, in the variational and finite element solution of elliptic boundary value problems (see also [4,6,5]).

Based on this fact, by making use of the convergence theorem about Method 1.1 (see [9]) we can obtain the following convergence result for Method 2.1.

Theorem 2.1. *Let $A \in \mathbb{R}^{n \times n}$ be a symmetric positive definite matrix of the block two-by-two structure (1.2), and P and H be the matrices defined by (2.2) and (2.3) such that $A = PHP^T$. Let $M = PWP^T$, where $W \in \mathbb{R}^{n \times n}$ is a symmetric positive definite matrix approximating the matrix $\bar{W} \in \mathbb{R}^{n \times n}$ defined by (2.4). Assume that*

$$y^T J_B y \leq 1 \quad \text{and} \quad z^T J_C z \leq 1 \quad (2.6)$$

hold for all normalized vectors $y \in \mathbb{R}^m$ and $z \in \mathbb{R}^\ell$, where J_B and J_C are the matrices defined in (2.1). Then

- (i) both \bar{S} and \bar{W} defined in (2.4) are symmetric positive definite matrices;
- (ii) if there exist two positive constants $v_B \in (0, 1)$ and $v_C \in (0, 1)$ such that

$$y^T J_B y \geq 1 - v_B \quad \text{and} \quad z^T J_C z \geq 1 - v_C \quad (2.7)$$

hold for all normalized vectors $y \in \mathbb{R}^m$ and $z \in \mathbb{R}^\ell$, the iteration sequence $\{x_k\} \subset \mathbb{R}^n$ generated by Method 2.1 satisfies

$$\|x_k - x^*\|_A \leq 2 \left(\frac{\sqrt{\sigma(v, \gamma) \kappa(W^{-1} \bar{W})} - 1}{\sqrt{\sigma(v, \gamma) \kappa(W^{-1} \bar{W})} + 1} \right)^k \|x_0 - x^*\|_A, \quad (2.8)$$

provided $v \in (0, 1)$ is such that $\gamma < \Gamma(v)$, where

$$\Gamma(v) = \frac{\sqrt{1-v} (\sqrt{4-3v^2} - v)}{2(1+v)}$$

and

$$v = \max\{v_B, v_C\}, \quad \sigma(v, \gamma) = \frac{1 - v - \gamma v \sqrt{1-v} - \gamma^2}{(1-v)^2 - \gamma v \sqrt{1-v} - (1+v)\gamma^2},$$

$\kappa(W^{-1} \bar{W})$ represents the Euclidean condition number of the matrix $(W^{-1} \bar{W})$, and $x^* = A^{-1}b$ is the exact solution of the system of linear equations (1.1)–(1.2).

Proof. We first prove (i). Because $A \in \mathbb{R}^{n \times n}$ is symmetric positive definite, we know that its Schur complement

$$S = C - E^T B^{-1} E$$

is symmetric positive definite, too. From (2.6) we can obtain the inequalities

$$y^T B y \leq y^T L_B L_B^T y \quad \text{and} \quad z^T C z \leq z^T L_C L_C^T z, \quad \forall y \in \mathbb{R}^m, \quad \forall z \in \mathbb{R}^\ell.$$

Therefore, it holds that

$$y^T (L_B L_B^T)^{-1} y \leq y^T B^{-1} y, \quad \forall y \in \mathbb{R}^m. \quad (2.9)$$

By straightforward computations we have

$$\begin{aligned} \bar{S} &= I - \bar{E}^T \bar{E} \\ &= I - (L_C^{-1} E^T L_B^{-T})(L_B^{-1} E L_C^{-T}) \\ &= L_C^{-1} [L_C L_C^T - E^T (L_B L_B^T)^{-1} E] L_C^{-T}. \end{aligned}$$

It then follows that, for $\forall z \in \mathbb{R}^\ell$,

$$\begin{aligned} z^T \bar{S} z &= z^T L_C^{-1} [L_C L_C^T - E^T (L_B L_B^T)^{-1} E] L_C^{-T} z \\ &\geq z^T L_C^{-1} [C - E^T B^{-1} E] L_C^{-T} z \\ &= z^T L_C^{-1} S L_C^{-T} z, \end{aligned}$$

where the inequality is induced by (2.9). Hence, the matrix \bar{S} is symmetric positive definite.

Evidently, to demonstrate the symmetric positive definiteness of the matrix \bar{W} defined by (2.4), we only need to verify the symmetric positive definiteness of its Schur complement

$$S_{\bar{W}} = \bar{S} - \bar{E}^T (I - J_B)^2 \bar{E}.$$

In fact, because, for all normalized vector $y \in \mathbb{R}^m$, $y^T J_B y \leq 1$ implies $1 \leq y^T J_B^{-1} y$, we obtain for all $z \in \mathbb{R}^\ell$ that

$$\begin{aligned} z^T S_{\bar{W}} z &\geq z^T [\bar{S} - \bar{E}^T (I - J_B) J_B^{-1} (I - J_B) \bar{E}] z \\ &= z^T [I - \bar{E}^T J_B^{-1} \bar{E}] z + z^T \bar{E}^T (I - J_B) \bar{E} z \\ &\geq z^T [I - \bar{E}^T J_B^{-1} \bar{E}] z \\ &\geq z^T L_C^{-1} S L_C^{-T} z. \end{aligned}$$

Hence, $S_{\bar{W}}$ is a symmetric positive definite matrix.

We now turn to demonstrate the validity of (ii). Obviously, it holds that

$$\overline{W} - H = \begin{bmatrix} I - J_B & O \\ O & I - J_C + \overline{E}^T(I - J_B)\overline{E} \end{bmatrix}.$$

By (2.6) we immediately obtain the estimate

$$x^T \overline{W} x \geq x^T H x, \quad \forall x \in \mathbb{R}^n.$$

On the other hand, as (i) implies for $z \in \mathbb{R}^\ell \setminus \{0\}$ that

$$0 < z^T \overline{S} z < z^T z \quad \text{and} \quad z^T \overline{S}^{-1} z > z^T z,$$

we know that

$$\|\overline{D}^{-1}\|_2 = \max \left\{ 1, \max_{z \neq 0} \frac{z^T \overline{S}^{-1} z}{z^T z} \right\} = \|\overline{S}^{-1}\|_2,$$

where we have denoted by

$$\overline{D} = \begin{bmatrix} I & O \\ O & \overline{S} \end{bmatrix}$$

the symmetric positive definite block-diagonal matrix of the matrix \overline{W} . Hence, by (2.4) it holds that

$$\|\overline{D} - \overline{W}\|_2 \leq \|(I - J_B)\overline{E}\|_2 \leq \|\overline{E}\|_2 \|I - J_B\|_2$$

and when

$$\|\overline{S}^{-1}\|_2 \|\overline{E}\|_2 \|I - J_B\|_2 < 1,$$

it holds that

$$\begin{aligned} \|\overline{W}^{-1}\|_2 &= \|[I - (I - \overline{D}^{-1}\overline{W})]^{-1} \overline{D}^{-1}\|_2 \\ &\leq \frac{\|\overline{D}^{-1}\|_2}{1 - \|\overline{D}^{-1}\|_2 \|\overline{D} - \overline{W}\|_2} \\ &= \frac{\|\overline{S}^{-1}\|_2}{1 - \|\overline{S}^{-1}\|_2 \|\overline{D} - \overline{W}\|_2} \\ &\leq \frac{\|\overline{S}^{-1}\|_2}{1 - \|\overline{S}^{-1}\|_2 \|\overline{E}\|_2 \|I - J_B\|_2}. \end{aligned} \tag{2.10}$$

It then follows that, for $\forall x = (y^T, z^T)^T \in \mathbb{R}^n \setminus \{0\}$ with $y \in \mathbb{R}^m$ and $z \in \mathbb{R}^\ell$,

$$\begin{aligned}
 \frac{x^T H x}{x^T \bar{W} x} &\geq 1 - \max_{x \neq 0} \left(1 - \frac{x^T H x}{x^T \bar{W} x} \right) \\
 &= 1 - \max_{x \neq 0} \frac{x^T (\bar{W} - H) x}{x^T \bar{W} x} \\
 &\geq 1 - \|\bar{W}^{-1}(\bar{W} - H)\|_2 \\
 &\geq 1 - \|\bar{W}^{-1}\|_2 \|\bar{W} - H\|_2 \\
 &\geq 1 - \frac{\|\bar{S}^{-1}\|_2}{1 - \|\bar{S}^{-1}\|_2 \|\bar{E}\|_2 \|I - J_B\|_2} \\
 &\quad \times \max\{\|I - J_B\|_2, \|I - J_C\|_2 + \|\bar{E}\|_2^2 \|I - J_B\|_2\} \\
 &\geq 1 - \frac{\|\bar{S}^{-1}\|_2 (1 + \|\bar{E}\|_2^2)}{1 - \|\bar{S}^{-1}\|_2 \|\bar{E}\|_2 \|I - J_B\|_2} \max\{\|I - J_B\|_2, \|I - J_C\|_2\}.
 \end{aligned}$$

Because inequality (2.5), together with (2.9) and (2.7), imply that

$$\begin{aligned}
 z^T \bar{E}^T \bar{E} z &= z^T L_C^{-1} E^T L_B^{-T} L_B^{-1} E L_C^{-T} z \\
 &\leq z^T L_C^{-1} E^T B^{-1} E L_C^{-T} z \\
 &\leq \gamma \sqrt{(z^T L_C^{-1} E^T B^{-1}) B (B^{-1} E L_C^{-T} z)} \sqrt{z^T L_C^{-1} C L_C^{-T} z} \\
 &= \gamma \sqrt{z^T \bar{E}^T J_B^{-1} \bar{E} z} \sqrt{z^T L_C^{-1} C L_C^{-T} z} \\
 &\leq \frac{\gamma}{\sqrt{1 - v_B}} \sqrt{z^T \bar{E}^T \bar{E} z} \sqrt{z^T L_C^{-1} C L_C^{-T} z} \\
 &\leq \frac{\gamma}{\sqrt{1 - v}} \sqrt{z^T \bar{E}^T \bar{E} z} \sqrt{z^T L_C^{-1} C L_C^{-T} z},
 \end{aligned}$$

by making use of (2.6) we can obtain

$$z^T \bar{E}^T \bar{E} z \leq \frac{\gamma^2}{1 - v} z^T L_C^{-1} C L_C^{-T} z \leq \frac{\gamma^2}{1 - v} z^T z$$

and

$$z^T \bar{S} z = z^T (I - \bar{E}^T \bar{E}) z \geq \left(1 - \frac{\gamma^2}{1-v}\right) z^T z,$$

or equivalently,

$$\|\bar{E}\|_2 \leq \frac{\gamma}{\sqrt{1-v}}, \quad \|\bar{S}^{-1}\|_2 \leq \frac{1-v}{1-v-\gamma^2}. \quad (2.11)$$

In addition, noticing that (2.7) implies that

$$\|I - J_B\|_2 \leq v_B \leq v \quad \text{and} \quad \|I - J_C\|_2 \leq v_C \leq v \quad (2.12)$$

and

$$\gamma < \frac{\sqrt{1-v}(\sqrt{v^2+4}-v)}{2},$$

implies that

$$\|\bar{S}^{-1}\|_2 \|\bar{E}\|_2 \|I - J_B\|_2 < 1,$$

we therefore have

$$\begin{aligned} \frac{x^T H x}{x^T \bar{W} x} &\geq 1 - \frac{((1-v)/(1-v-\gamma^2))(1+\gamma^2/(1-v))}{1 - ((1-v)/(1-v-\gamma^2))(\gamma/\sqrt{1-v})v} \max\{v_B, v_C\} \\ &\geq 1 - \frac{(1-v+\gamma^2)v}{1-v-\gamma v \sqrt{1-v}-\gamma^2} \\ &> 0, \end{aligned}$$

when $v \in (0, 1)$ satisfies

$$\gamma < \Gamma(v).$$

Consequently,

$$\kappa(\bar{W}^{-1} H) \leq \frac{1-v-\gamma v \sqrt{1-v}-\gamma^2}{(1-v)^2 - \gamma v \sqrt{1-v} - (1+v)\gamma^2} = \sigma(v, \gamma)$$

and

$$\kappa(M^{-1} A) = \kappa(W^{-1} H) \leq \kappa(W^{-1} \bar{W}) \kappa(\bar{W}^{-1} H) \leq \sigma(v, \gamma) \kappa(W^{-1} \bar{W}).$$

Now, from [9, Theorem 2.1] we know that the RPCG iteration sequence $\{x_k\} \subset \mathbb{R}^n$ satisfies

$$\begin{aligned} \|x_k - x^*\|_A &\leq 2 \left(\frac{\sqrt{\kappa(M^{-1}A)} - 1}{\sqrt{\kappa(M^{-1}A)} + 1} \right)^k \|x_0 - x^*\|_A \\ &\leq 2 \left(\frac{\sqrt{\sigma(v, \gamma)\kappa(W^{-1}\bar{W})} - 1}{\sqrt{\sigma(v, \gamma)\kappa(W^{-1}\bar{W})} + 1} \right)^k \|x_0 - x^*\|_A. \quad \square \end{aligned}$$

3. Several practical restrictive preconditioners

In this section, by particularly choosing the matrix $W \in \mathbb{R}^{n \times n}$ to be the block Jacobi and the *block symmetric Gauss–Seidel* (BSGS) approximations of the matrix $\bar{W} \in \mathbb{R}^{n \times n}$ in (2.4), we can obtain two restrictive preconditioners, called BJ and BSGS, to the block two-by-two symmetric positive definite matrix A in (1.2). Hence, we get two versions of the RPCG methods, called BJ-RPCG and BSGS-RPCG, from Method 2.1 for the system of linear equations (1.1)–(1.2). We will precisely describe them in the following.

3.1. The BJ preconditioner

If the matrix $W \in \mathbb{R}^{n \times n}$ is taken to be the BJ splitting matrix of the matrix $\bar{W} \in \mathbb{R}^{n \times n}$ in (2.4), i.e.,

$$W = \begin{bmatrix} I & O \\ O & \bar{S} \end{bmatrix}, \quad (3.1)$$

then we obtain the BJ preconditioner $M = PW P^{-1}$ to the original coefficient matrix A in (1.2). From Method 2.1 we can obtain the *block Jacobi-type RPCG* (BJ-RPCG) method for solving the system of linear equations (1.1)–(1.2).

Let

$$\hat{B} = L_B L_B^T \approx B \quad \text{and} \quad \hat{S} = L_C \bar{S} L_C^T \approx S = C - E^T B^{-1} E.$$

Then by making use of the block structures and the internal relationship between the matrices P and W , we can solve the generalized residual equation $Mz = r$, with $z = (z^{(1)T}, z^{(2)T})^T$ and $r = (r^{(1)T}, r^{(2)T})^T$, involved in the BJ-RPCG method, by the following formulas:

- $t^{(1)} = \hat{B}^{-1} r^{(1)}$;
- $z^{(2)} = \hat{S}^{-1} (r^{(2)} - E^T t^{(1)})$;
- $\tilde{t}^{(1)} = \hat{B}^{-1} (E z^{(2)})$;
- $z^{(1)} = t^{(1)} - \tilde{t}^{(1)}$.

Consequently, we obtain the following algorithmic description of the BJ-RPCG method.

Method 3.1 (THE BJ-RPCG METHOD FOR LINEAR SYSTEM (1.1)–(1.2)).

Choose $x_0 \in \mathbb{R}^n$, $r_0 = b - Ax_0$.

Let $r_0 = (r_0^{(1)\top}, r_0^{(2)\top})^\top$, $z_0 = (z_0^{(1)\top}, z_0^{(2)\top})^\top$, and

$$\text{solve } \widehat{B}t^{(1)} = r_0^{(1)}$$

$$\text{solve } \widehat{S}z_0^{(2)} = r_0^{(2)} - E^\top t^{(1)}$$

$$\text{solve } \widehat{B}\widetilde{t}^{(1)} = Ez_0^{(2)}$$

$$\text{compute } z_0^{(1)} = t^{(1)} - \widetilde{t}^{(1)}.$$

Set $p_0 = z_0$.

For $k = 0, 1, 2, \dots$

$$\alpha_k = -\frac{z_k^\top r_k}{p_k^\top A p_k}$$

$$x_{k+1} = x_k - \alpha_k p_k$$

$$r_{k+1} = r_k + \alpha_k A p_k$$

Let $r_{k+1} = (r_{k+1}^{(1)\top}, r_{k+1}^{(2)\top})^\top$, $z_{k+1} = (z_{k+1}^{(1)\top}, z_{k+1}^{(2)\top})^\top$, and

$$\text{solve } \widehat{B}t^{(1)} = r_{k+1}^{(1)}$$

$$\text{solve } \widehat{S}z_{k+1}^{(2)} = r_{k+1}^{(2)} - E^\top t^{(1)}$$

$$\text{solve } \widehat{B}\widetilde{t}^{(1)} = Ez_{k+1}^{(2)}$$

$$\text{compute } z_{k+1}^{(1)} = t^{(1)} - \widetilde{t}^{(1)}$$

$$\beta_k = \frac{z_{k+1}^\top r_{k+1}}{z_k^\top r_k}$$

$$p_{k+1} = z_{k+1} + \beta_k p_k.$$

Evidently, the only difference between Method 3.1 here and Method 3.1 in [9] is that the former involves only an approximation matrix \widehat{B} of B , while the latter involves the matrix block B . Therefore, Method 3.1 here allows approximate inversion of the matrix block B at each iteration. Also, we note that, at each of its iteration steps, the BJ-RPCG method requires solutions of two sub-systems of linear equations with the coefficient matrix $\widehat{B} \in \mathbb{R}^{m \times m}$ and one sub-system of linear equations with the coefficient matrix $\widehat{S} \in \mathbb{R}^{\ell \times \ell}$.

Based on Theorem 2.1, we can demonstrate the following convergence theorem for Method 3.1.

Theorem 3.1. *Let the conditions of Theorem 2.1 be satisfied. Then the iteration sequence $\{x_k\} \subset \mathbb{R}^n$ generated by Method 3.1 satisfies*

$$\|x_k - x^*\|_A \leq 2 \left(\frac{\sqrt{\delta(v, \gamma)} - 1}{\sqrt{\delta(v, \gamma)} + 1} \right)^k \|x_0 - x^*\|_A,$$

provided $v \in (0, 1)$ is such that $\gamma < \Gamma(v)$, where

$$\Gamma(v) = \frac{\sqrt{1-v} \left(\sqrt{4-3v^2} - v \right)}{2(1+v)}$$

and

$$\delta(v, \gamma) = \frac{1 - v + \gamma v \sqrt{1 - v} - \gamma^2}{(1 - v)^2 - \gamma v \sqrt{1 - v} - (1 + v)\gamma^2}.$$

Proof. We first derive bounds for $\kappa(W^{-1}\bar{W})$. From (2.4) we have

$$\bar{W} = W + \begin{bmatrix} O & (I - J_B)\bar{E} \\ \bar{E}^T(I - J_B) & O \end{bmatrix}$$

and hence,

$$W^{-1}\bar{W} = I + W^{-1} \begin{bmatrix} O & (I - J_B)\bar{E} \\ \bar{E}^T(I - J_B) & O \end{bmatrix}.$$

It then follows that, for $\forall x \in \mathbb{R}^n \setminus \{0\}$,

$$\begin{aligned} \frac{x^T \bar{W} x}{x^T W x} &\leq 1 + \|W^{-1}\|_2 \|(I - J_B)\bar{E}\|_2 \\ &\leq 1 + \|\bar{S}^{-1}\|_2 \|\bar{E}\|_2 \|I - J_B\|_2 \\ &\leq 1 + \frac{\gamma v \sqrt{1 - v}}{1 - v - \gamma^2} \end{aligned}$$

and

$$\begin{aligned} \frac{x^T \bar{W} x}{x^T W x} &\geq 1 - \|W^{-1}\|_2 \|(I - J_B)\bar{E}\|_2 \\ &\geq 1 - \|\bar{S}^{-1}\|_2 \|\bar{E}\|_2 \|I - J_B\|_2 \\ &\geq 1 - \frac{\gamma v \sqrt{1 - v}}{1 - v - \gamma^2}, \end{aligned}$$

here we have applied estimates (2.11) and (2.12). Therefore, it holds that

$$\kappa(W^{-1}\bar{W}) \leq \frac{1 - v + \gamma v \sqrt{1 - v} - \gamma^2}{1 - v - \gamma v \sqrt{1 - v} - \gamma^2}.$$

By substituting this estimate into Theorem 2.1 we straightforwardly get the conclusion of Theorem 3.1. \square

3.2. The BSGS preconditioner

If the matrix $W \in \mathbb{R}^{n \times n}$ is taken to be the BSGS splitting matrix of the matrix $\bar{W} \in \mathbb{R}^{n \times n}$ in (2.4), i.e.,

$$W = \begin{bmatrix} I & (I - J_B)\bar{E} \\ O & \bar{S} \end{bmatrix} \begin{bmatrix} I & O \\ O & \bar{S} \end{bmatrix}^{-1} \begin{bmatrix} I & O \\ \bar{E}^T(I - J_B) & \bar{S} \end{bmatrix}, \quad (3.2)$$

then we obtain the BSGS preconditioner $M = PW P^{-1}$ to the original coefficient matrix A in (1.2). From Method 2.1 we can obtain the *block symmetric Gauss–Seidel-type RPCG* (BSGS-RPCG) method for solving the system of linear equations (1.1)–(1.2).

Let

$$\widehat{B} = L_B L_B^T \approx B \quad \text{and} \quad \widehat{S} = L_C \bar{S} L_C^T \approx S = C - E^T B^{-1} E.$$

Then by making use of the block structures and the internal relationship between the matrices P and W , we can solve the generalized residual equation $Mz = r$, with $z = (z^{(1)T}, z^{(2)T})^T$ and $r = (r^{(1)T}, r^{(2)T})^T$, involved in the BSGS-RPCG method, by the following formulas:

- $t^{(1)} = \widehat{B}^{-1} r^{(1)}$;
- $t^{(2)} = \widehat{S}^{-1} (r^{(2)} - E^T t^{(1)})$;
- $\tilde{t}^{(1)} = \widehat{B}^{-1} (E t^{(2)})$;
- $\hat{t}^{(1)} = \widehat{B}^{-1} (r^{(1)} + B \tilde{t}^{(1)})$;
- $\bar{t}^{(1)} = \hat{t}^{(1)} - \tilde{t}^{(1)}$;
- $t^{(1)} = \widehat{B}^{-1} (B \bar{t}^{(1)})$;
- $\tilde{t}^{(2)} = \widehat{S}^{-1} (E^T (\bar{t}^{(1)} - t^{(1)}))$;
- $z^{(2)} = t^{(2)} - \tilde{t}^{(2)}$;
- $\hat{t}^{(1)} = \widehat{B}^{-1} (E z^{(2)})$;
- $z^{(1)} = \bar{t}^{(1)} - \hat{t}^{(1)}$.

Consequently, we obtain the following algorithmic description of the BSGS-RPCG method.

Method 3.2 (THE BSGS-RPCG METHOD FOR LINEAR SYSTEM (1.1)–(1.2)).

Choose $x_0 \in \mathbb{R}^n$, $r_0 = b - Ax_0$.

Let $r_0 = (r_0^{(1)T}, r_0^{(2)T})^T$, $z_0 = (z_0^{(1)T}, z_0^{(2)T})^T$, and

solve $\widehat{B} t^{(1)} = r_0^{(1)}$
 solve $\widehat{S} t^{(2)} = r_0^{(2)} - E^T t^{(1)}$
 solve $\widehat{B} \tilde{t}^{(1)} = E t^{(2)}$
 solve $\widehat{B} \hat{t}^{(1)} = r_0^{(1)} + B \tilde{t}^{(1)}$
 compute $\bar{t}^{(1)} = \hat{t}^{(1)} - \tilde{t}^{(1)}$
 solve $\widehat{B} t^{(1)} = B \bar{t}^{(1)}$
 solve $\widehat{S} \tilde{t}^{(2)} = E^T (\bar{t}^{(1)} - t^{(1)})$
 compute $z_0^{(2)} = t^{(2)} - \tilde{t}^{(2)}$
 solve $\widehat{B} \hat{t}^{(1)} = E z_0^{(2)}$
 compute $z_0^{(1)} = \bar{t}^{(1)} - \hat{t}^{(1)}$.

Set $p_0 = z_0$.

For $k = 0, 1, 2, \dots$

$$\alpha_k = -\frac{z_k^T r_k}{p_k^T A p_k}$$

$$x_{k+1} = x_k - \alpha_k p_k$$

$$r_{k+1} = r_k + \alpha_k A p_k$$

$$\begin{aligned}
& \text{Let } r_{k+1} = (r_{k+1}^{(1)\text{T}}, r_{k+1}^{(2)\text{T}})^{\text{T}}, z_{k+1} = (z_{k+1}^{(1)\text{T}}, z_{k+1}^{(2)\text{T}})^{\text{T}}, \text{ and} \\
& \text{solve } \widehat{B}t^{(1)} = r_{k+1}^{(1)} \\
& \text{solve } \widehat{S}t^{(2)} = r_{k+1}^{(2)} - E^{\text{T}}t^{(1)} \\
& \text{solve } \widehat{B}\tilde{t}^{(1)} = Et^{(2)} \\
& \text{solve } \widehat{B}\tilde{t}^{(1)} = r_{k+1}^{(1)} + B\tilde{t}^{(1)} \\
& \text{compute } \bar{t}^{(1)} = \tilde{t}^{(1)} - \tilde{t}^{(1)} \\
& \text{solve } \widehat{B}t^{(1)} = B\bar{t}^{(1)} \\
& \text{solve } \widehat{S}\tilde{t}^{(2)} = E^{\text{T}}(\bar{t}^{(1)} - t^{(1)}) \\
& \text{compute } z_{k+1}^{(2)} = t^{(2)} - \tilde{t}^{(2)} \\
& \text{solve } \widehat{B}\tilde{t}^{(1)} = Ez_{k+1}^{(2)} \\
& \text{compute } z_{k+1}^{(1)} = \bar{t}^{(1)} - \tilde{t}^{(1)} \\
& \beta_k = \frac{z_{k+1}^{\text{T}} r_{k+1}}{z_k^{\text{T}} r_k} \\
& p_{k+1} = z_{k+1} + \beta_k p_k.
\end{aligned}$$

Method 3.2 is not discussed in [9], even for its exact variant that is induced by taking $\widehat{B} = B$. Clearly, at each of its iteration steps, the BSGS-RPCG method requires solutions of five sub-systems of linear equations with the coefficient matrix $\widehat{B} \in \mathbb{R}^{m \times m}$ and two sub-systems of linear equations with the coefficient matrix $\widehat{S} \in \mathbb{R}^{\ell \times \ell}$. Therefore, its computing cost is about twice of that of BJ-RPCG method. However, from the convergence theorem established below we can see that the BSGS-RPCG method is faster than the BJ-RPCG method.

Based on Theorem 2.1, we can demonstrate the following convergence theorem for Method 3.2.

Theorem 3.2. *Let the conditions of Theorem 2.1 be satisfied. Then the iteration sequence $\{x_k\} \subset \mathbb{R}^n$ generated by Method 3.2 satisfies*

$$\|x_k - x^*\|_A \leq 2 \left(\frac{\sqrt{\delta(v, \gamma)} - 1}{\sqrt{\delta(v, \gamma)} + 1} \right)^k \|x_0 - x^*\|_A,$$

provided $v \in (0, 1)$ is such that $\gamma < \Gamma(v)$, where

$$\Gamma(v) = \frac{\sqrt{1-v} \left(\sqrt{4-3v^2} - v \right)}{2(1+v)}$$

and

$$\delta(v, \gamma) = \frac{(1-v)v^2\gamma^2 + (1-v-\gamma^2)(1-v-\gamma v\sqrt{1-v}-\gamma^2)}{[(1-v)^2 - \gamma v\sqrt{1-v} - (1+v)\gamma^2](1-v-\gamma^2)}.$$

Proof. We first derive bounds for $\kappa(W^{-1}\bar{W})$. From (2.4) we have

$$W = \bar{W} + \begin{bmatrix} (I - J_B)\bar{E}\bar{S}^{-1}\bar{E}^T(I - J_B) & O \\ O & O \end{bmatrix}.$$

Hence, by making use of (2.6) and Theorem 2.1(i) we can immediately obtain that, for all $x \in \mathbb{R}^n$,

$$x^T W x \geq x^T \bar{W} x,$$

or equivalently,

$$\frac{x^T \bar{W} x}{x^T W x} \leq 1, \quad \forall x \in \mathbb{R}^n \setminus \{0\}.$$

In addition, it holds that, for $\forall x \in \mathbb{R}^n \setminus \{0\}$,

$$\begin{aligned} \frac{x^T W x}{x^T \bar{W} x} &\leq \max_{x \neq 0} \frac{x^T W x}{x^T \bar{W} x} = \|\bar{W}^{-1} W\|_2 \\ &\leq 1 + \|\bar{W}^{-1}\|_2 \|(I - J_B)\bar{E}\bar{S}^{-1}\bar{E}^T(I - J_B)\|_2 \\ &\leq 1 + \|\bar{W}^{-1}\|_2 \|\bar{E}\|_2^2 \|\bar{S}^{-1}\|_2 \|I - J_B\|_2^2. \end{aligned}$$

By applying the estimates (2.10), (2.11) and (2.12) to this inequality, we can further obtain

$$\begin{aligned} \frac{x^T W x}{x^T \bar{W} x} &\leq 1 + \frac{\|\bar{S}^{-1}\|_2^2 \|\bar{E}\|_2^2 \|I - J_B\|_2^2}{1 - \|\bar{S}^{-1}\|_2 \|\bar{E}\|_2 \|I - J_B\|_2} \\ &\leq 1 + \frac{(1 - v)\gamma^2}{(1 - v - \gamma^2)(1 - v - \gamma v \sqrt{1 - v} - \gamma^2)}. \end{aligned}$$

Therefore, it holds that

$$\kappa(W^{-1}\bar{W}) \leq \frac{(1 - v)\gamma^2 + (1 - v - \gamma^2)(1 - v - \gamma v \sqrt{1 - v} - \gamma^2)}{(1 - v - \gamma^2)(1 - v - \gamma v \sqrt{1 - v} - \gamma^2)}.$$

By substituting this estimate into Theorem 2.1 we straightforwardly get the conclusion of Theorem 3.2. \square

Theorems 3.1 and 3.2 show that the effectiveness of these two methods depends on the property of the original problem and the quality of the approximation matrix as well. Therefore, which method is more effective in actual applications needs to be further examined by numerical experiments.

Instead of the BJ and the BSGS splitting matrices defined in (3.1) and (3.2), we can approximate the matrix \bar{W} in (2.4) by the modified BJ and the modified BSGS splitting matrices [7,8] and obtain the approximation matrix W , i.e., we can take W to be of the same form as (3.1) or (3.2) but replace \bar{S} by its approximation (for simplicity of notation, it is still denoted by \bar{S}). This therefore leads to variants of the BJ-PCG method and the BSGS-PCG method, which allow more flexible choices of the approximate Schur complement \hat{S} . Convergence theorems of these two variants can be easily established similarly to Theorems 3.1 and 3.2.

4. Typical choices of the matrices L_B and L_C

Recalling that we have represented by

$$\widehat{B} = L_B L_B^T \approx B, \quad \widehat{S} = (\text{or } \approx) L_C \bar{S} L_C^T \approx S = C - E^T B^{-1} E.$$

Let

$$\widehat{C} = L_C L_C^T \approx C.$$

In this section, we will discuss possible choices of the matrices L_B and L_C , or in other words, the matrices \widehat{B} , \widehat{C} and \widehat{S} .

Let $M_B \in \mathbb{R}^{n \times n}$ be a symmetric positive definite matrix obtained possibly through *incomplete Cholesky* (IC) factorizations [2,16,18], splitting iterations (e.g., Jacobi, symmetric Gauss–Seidel (SGS), or symmetric successive overrelaxation (SSOR), etc.) [2,12,7,8], or multigrid/multilevel approximations [2], etc., of the matrix block $B \in \mathbb{R}^{m \times m}$. Without loss of generality, we assume that the condition (2.6) is automatically satisfied, i.e.,

$$\frac{y^T B y}{y^T M_B y} \leq 1, \quad \forall y \in \mathbb{R}^m \setminus \{0\}.$$

Because, otherwise, we can turn to consider the shifted matrix

$$M_B := M_B + \beta I, \quad \beta > 0 \text{ a constant,}$$

instead. In addition, we assume that the splitting

$$B = M_B - N_B \quad \text{with } N_B = M_B - B,$$

is convergent, i.e., the spectral radius of the matrix $M_B^{-1} N_B$, denoted by $\rho(M_B^{-1} N_B)$, is less than 1. In fact, such an assumption is not very restrictive and can be satisfied by many splittings such as the (block) symmetric Gauss–Seidel splitting (see, for instance, [17,2,13]). If such a matrix M_B further satisfies the condition (2.7), i.e.,

$$\min_{y \neq 0} \frac{y^T B y}{y^T M_B y} \geq 1 - v$$

for a $v \in (0, 1)$ such that $\gamma < \Gamma(v)$, then we can take $\widehat{B} = M_B$. Otherwise, we take

$$\widehat{B} = B[I - (M_B^{-1} N_B)^{k_B}]^{-1} = M_B \left[\sum_{j=0}^{k_B-1} (M_B^{-1} N_B)^j \right]^{-1},$$

where k_B is a positive integer such that

$$k_B > \frac{\ln(v)}{\ln(\rho(M_B^{-1} N_B))}.$$

Such a matrix \widehat{B} will be symmetric positive definite and satisfy

$$\min_{y \neq 0} \frac{y^T B y}{y^T \widehat{B} y} \geq 1 - v, \quad (4.1)$$

or the condition (2.7).

In fact, \widehat{B} is evidently symmetric since

$$\widehat{B}^T = \left[\sum_{j=0}^{k_B-1} (N_B M_B^{-1})^j \right]^{-1} M_B = M_B \left[\sum_{j=0}^{k_B-1} (M_B^{-1} N_B)^j \right]^{-1} = \widehat{B}.$$

It is positive definite since the smallest eigenvalue of the matrix

$$\begin{aligned} M_B^{-1/2} \widehat{B} M_B^{-1/2} &= M_B^{1/2} \left[\sum_{j=0}^{k_B-1} (M_B^{-1} N_B)^j \right]^{-1} M_B^{-1/2} \\ &= (I - M_B^{-1/2} N_B M_B^{-1/2}) [I - (M_B^{-1/2} N_B M_B^{-1/2})^{k_B}]^{-1} \end{aligned}$$

is bounded from below by the positive constant $(1 - \rho(M_B^{-1} N_B)) / (1 - [\rho(M_B^{-1} N_B)]^{k_B})$. Moreover, \widehat{B} satisfies (4.1) since

$$\begin{aligned} \max_{y \neq 0} \frac{y^T \widehat{B} y}{y^T B y} &\leq \rho(B^{-1} \widehat{B}) = \rho([I - (M_B^{-1} N_B)^{k_B}]^{-1}) \\ &\leq [1 - \rho(M_B^{-1} N_B)^{k_B}]^{-1} := (1 - v)^{-1} \end{aligned}$$

with

$$v = \rho(M_B^{-1} N_B)^{k_B}.$$

The choice of the matrix \widehat{C} can be discussed in an analogous fashion.

According to the matrix $\widehat{S} \in \mathbb{R}^{\ell \times \ell}$, we therefore have the following two typical choices:

- (a) $\widehat{S} = \widehat{C}$;
- (b) $\widehat{S} = \widehat{C} - E^T G E$, and $G \in \mathbb{R}^{m \times m}$ is a sparse approximation to the matrix \widehat{B}^{-1} .

Correspondingly, we obtain two practical versions of the BJ-RPCG method and the BSGS-RPCG method, denoted, respectively, by BJ-RPCG(a), BJ-RPCG(b) and BSGS-RPCG(a), BSGS-RPCG(b).

5. Numerical results

To simplify statements, we first introduce the following necessary notations.

$\Omega = (0, 1) \times (0, 1) \in \mathbb{R}^2$ represents a square domain on the (ξ, η) -plane, and $\partial\Omega$ its boundary. $\text{diag}(\cdot)$ and $\text{tridiag}(\cdot)$ denote a point-diagonal and point-tridiagonal matrix, respectively, and $\text{Diag}(\cdot)$ and $\text{Tridiag}(\cdot)$

denote a block-diagonal and block-tridiagonal matrix, respectively. For example,

$$\text{diag}(t_1, t_2, \dots, t_p) = \begin{bmatrix} t_1 & & & \\ & t_2 & & \\ & & \ddots & \\ & & & t_p \end{bmatrix},$$

$$\text{tridiag}(t^{(-1)}, t_{i,[1:p]}^{(0)}, t^{(1)}) = \begin{bmatrix} t_{i,1}^{(0)} & t^{(1)} & & & \\ t^{(-1)} & t_{i,2}^{(0)} & t^{(1)} & & \\ & \ddots & \ddots & \ddots & \\ & & t^{(-1)} & t_{i,p-1}^{(0)} & t^{(1)} \\ & & & t^{(-1)} & t_{i,p}^{(0)} \end{bmatrix}$$

and

$$\text{Diag}(T_1, T_2, \dots, T_p) = \begin{bmatrix} T_1 & & & \\ & T_2 & & \\ & & \ddots & \\ & & & T_p \end{bmatrix},$$

$$\text{Tridiag}(T^{(-1)}, T_{i,[1:p]}^{(0)}, T^{(1)}) = \begin{bmatrix} T_{i,1}^{(0)} & T^{(1)} & & & \\ T^{(-1)} & T_{i,2}^{(0)} & T^{(1)} & & \\ & \ddots & \ddots & \ddots & \\ & & T^{(-1)} & T_{i,p-1}^{(0)} & T^{(1)} \\ & & & T^{(-1)} & T_{i,p}^{(0)} \end{bmatrix}$$

with $T^{(-1)}, T^{(1)}, T_j$ ($j = 1, 2, \dots, p$) and $T_{i,j}^{(0)}$ ($i, j = 1, 2, \dots, p$) being square submatrices of suitable sizes. In addition, we denote by $n = N \times N$ and $h = 1/(N + 1)$. For a nonlinear function $u : \mathbb{R}^2 \rightarrow \mathbb{R}^1$, we employ $u_{i,j}$ to represent $u(ih, jh)$, i.e., the value of $u(\xi, \eta)$ at the grid point (ih, jh) .

Notation	Description
NIT	Number of iteration steps
CPU	CPU timing
ERR	Absolute error of the approximate solution
$\mathcal{R}(A)$	Restricted matrix of the matrix A
$\mathcal{R}_B(A)$	$\mathcal{R}(A)$ with the same sparse pattern as the matrix B
SGS	Symmetric Gauss–Seidel preconditioner
IC	IC preconditioner with no fill-in [16]
CG	Conjugate gradient method
PCG	Preconditioned CG method

In numerical experiments, the right-hand-side vector $b \in \mathbb{R}^n$ is taken to be $b = (1^2, 2^2, \dots, n^2)^T \in \mathbb{R}^n$, and the exact solution $x^* = A^{-1}b$ of the system of linear equations (1.1)–(1.2) is computed by Matlab with precision 10^{-16} . All iterations are started from an initial vector $x_0 \in \mathbb{R}^n$ having all entries equal to zero, and terminated if either $\varepsilon \leq \varepsilon$ or $IT \geq IT_{\max}$ is satisfied, where $\varepsilon = \|x_k - x^*\|_2 / \|x_0 - x^*\|_2$ denotes the reduction factor with respect to the error of an iteration, “IT” is the number of total iteration steps, and ε and IT_{\max} are, respectively, the prescribed accuracy and the maximum of iteration steps.

For a number set $\mathcal{P} \subset \{(i, j) \mid i, j = 1, 2, \dots, n\}$ and a matrix $A = (a_{ij}) \in \mathbb{R}^{n \times n}$, $\mathcal{R}(A) = (\alpha_{ij}) \in \mathbb{R}^{n \times n}$ is defined by $\alpha_{ij} = a_{ij}$ if $(i, j) \in \mathcal{P}$, and $\alpha_{ij} = 0$ if $(i, j) \notin \mathcal{P}$, $i, j = 1, 2, \dots, n$. If $B = (b_{ij}) \in \mathbb{R}^{n \times n}$, then $\mathcal{R}_B(A) = (\alpha_{ij}) \in \mathbb{R}^{n \times n}$ is defined by $\alpha_{ij} = a_{ij}$ if $b_{ij} \neq 0$, and $\alpha_{ij} = 0$ if $b_{ij} = 0$, $i, j = 1, 2, \dots, n$.

We consider the second-order self-adjoint elliptic boundary value problem

$$\begin{cases} -\mu \Delta u + \theta(\xi, \eta)u = f(\xi, \eta), & \text{in } \Omega, \\ u = 0, & \text{on } \partial\Omega, \end{cases} \quad (5.1)$$

where μ is a positive constant, and $\theta(\xi, \eta)$ and $f(\xi, \eta)$ are bounded nonlinear functions. By choosing different μ and $\theta(\xi, \eta)$ we can get a number of testing problems.

Assume that $\Omega_1 = (0, 1) \times (0, \gamma_1)$ and $\Omega_2 = (0, 1) \times (\gamma_2, 1)$ are two subdomains of the domain Ω , with γ_1 and γ_2 being positive constants satisfying $0 < \gamma_2 < \gamma_1 < 1$, and $\Omega_o = \Omega_1 \cap \Omega_2 = (0, 1) \times (\gamma_2, \gamma_1)$ the overlapping between Ω_1 and Ω_2 . If the partial differential equation (5.1) is discretized by centered differences with uniform stepsize h ($0 < h < \gamma_1 - \gamma_2$), and the nodes are labeled by the ordering of the subdomains according to

$$\mathcal{P}_1 = \{(i, j) \mid (ih, jh) \in \Omega_1 \setminus \Omega_o, \ 1 \leq i \leq N, \ 1 \leq j \leq p\},$$

$$\mathcal{P}_2 = \{(i, j) \mid (ih, jh) \in \Omega_2 \setminus \Omega_o, \ 1 \leq i \leq N, \ 1 \leq j \leq q\},$$

$$\mathcal{P}_3 = \{(i, j) \mid (ih, jh) \in \Omega_o, \ 1 \leq i \leq N, \ 1 \leq j \leq s\}$$

and the nodes in each subdomain are labeled by the natural ordering, i.e., from left to right, and bottom to top, then we obtain a system of linear equations of form (1.1)–(1.2), in which $h = 1/(N + 1)$,

$$B = \begin{bmatrix} A_{1,1} & O \\ O & A_{2,2} \end{bmatrix}, \quad E = \begin{bmatrix} A_{1,3} \\ A_{2,3} \end{bmatrix}, \quad C = A_{3,3},$$

where $m = (p + q)N$, $\ell = sN$; $p = \gamma_2/h$, $q = (1 - \gamma_1)/h - 1$ and $s = (\gamma_1 - \gamma_2)/h$ are positive integers satisfying $p + q + s = N$;

$$A_{1,1} = \text{Tridiag}(-\mu I, T_{[1:p]}, -\mu I) \in \mathbb{R}^{pN \times pN},$$

$$A_{2,2} = \text{Tridiag}(-\mu I, T_{[p+1:p+q]}, -\mu I) \in \mathbb{R}^{qN \times qN},$$

$$A_{3,3} = \text{Tridiag}(-\mu I, T_{[p+q+1:N]}, -\mu I) \in \mathbb{R}^{sN \times sN},$$

$$A_{1,3} = \begin{pmatrix} O & O \\ -\mu I_{N \times N} & O \end{pmatrix} \in \mathbb{R}^{pN \times sN}, \quad A_{2,3} = \begin{pmatrix} O & -\mu I_{N \times N} \\ O & O \end{pmatrix} \in \mathbb{R}^{qN \times sN}$$

and for $i = 1, 2, \dots, N$,

$$T_i = \text{tridiag}(\mu^{(1)}, \mu_{[1:N],i}^{(0)}, \mu^{(1)}) \in \mathbb{R}^{N \times N}$$

with

$$\mu^{(1)} = -\mu, \quad \mu_{i,j}^{(0)} = 4\mu + h^2 \theta_{i,j}, \quad 1 \leq i, j \leq N.$$

For more details, we refer the readers to [17].

We remark that the block two-by-two matrix $A \in \mathbb{R}^{n \times n}$ described above differs from the usual block tridiagonal matrix $\tilde{A} \in \mathbb{R}^{n \times n}$ only by a permutation. The latter is the matrix obtained from the centered finite difference discretization with uniform stepsize h of the second-order self-adjoint elliptic boundary value problem (5.1), for which the nodes in the domain Ω are labeled by the natural ordering, i.e., from left to right, and bottom to top. More specifically, let $x = (y^T, z^T, w^T)^T \in \mathbb{R}^n$, with $y \in \mathbb{R}^{pN}$, $z \in \mathbb{R}^{qN}$ and $w \in \mathbb{R}^{sN}$, and let

$$\Pi = \begin{pmatrix} I_{pN \times pN} & O & O \\ O & O & I_{qN \times qN} \\ O & I_{sN \times sN} & O \end{pmatrix} \in \mathbb{R}^{n \times n}.$$

Then Π is a permutation matrix and

$$\tilde{A} = \Pi A \Pi.$$

It follows that the block two-by-two linear system (1.1)–(1.2) is equivalent to the linear system

$$\tilde{A}\tilde{x} = \tilde{b} \quad \text{with } \tilde{x} = \Pi x \text{ and } \tilde{b} = \Pi b \quad (5.2)$$

through the permutation Π . Because a permutation does not introduce any more arithmetic operations, we can claim that the classical CG and PCG methods for linear system (1.1)–(1.2) possess the same theoretical properties and numerical behaviors as those for the linear system (5.2).

In our computing, we take $\gamma_1 = \frac{1}{2} + \frac{1}{8}$, $\gamma_2 = \frac{1}{2}$, $\varepsilon = 10^{-8}$, and $\text{IT}_{\max} = n$. The RPCG with an IC (SGS) preconditioner means that we take $G = \hat{B}^{-1}$ with \hat{B} the IC factorization (SGS splitting matrix) of the matrix B , and \hat{S} the IC factorization (SGS splitting matrix) of the matrix $C - \mathcal{R}_C(E^T G E)$. And the PCG with an IC (SGS) preconditioner means that we take the IC factorization (SGS splitting matrix) of the coefficient matrix A as its preconditioner and apply it to precondition the conjugate gradient method for the system of linear equations (1.1)–(1.2).

Example 5.1. Problem (5.1) with $\mu = 1.0$ and

$$\theta(\xi, \eta) = \frac{2 \tan^2(\pi(\xi - \frac{1}{2}))}{\xi(1 - \eta)} + \frac{2 \tan^2(\pi(\eta - \frac{1}{2}))}{(1 - \xi)\eta}.$$

Example 5.2. Problem (5.1) with $\mu = 1.0$ and

$$\theta(\xi, \eta) = \frac{2 \cos^2(\xi) \sin^2(\eta) \tan^2(\pi(\xi - \frac{1}{2}))}{\xi(1 - \eta)} + \frac{2 \sin^2(\xi) \cos^2(\eta) \tan^2(\pi(\eta - \frac{1}{2}))}{(1 - \xi)\eta}.$$

Table 1
NITs, CPUs and ERRs for Example 5.1

h^{-1} n			96	104	112	120	128	136	144
			9025	10,609	12,321	14,161	16,129	18,225	20,449
BJ-RPCG(a)	IC	NIT	13	13	13	13	13	13	13
		CPU	0.70	0.88	1.00	1.21	1.38	1.63	1.92
		ERR	3.10e−9	2.81e−9	3.16e−9	2.66e−9	3.15e−9	2.52e−9	3.12e−9
	SGS	NIT	12	12	12	12	12	12	12
		CPU	0.69	0.86	0.98	1.16	1.31	1.58	1.93
		ERR	3.96e−9	3.14e−9	3.80e−9	2.99e−9	3.62e−9	2.85e−9	3.45e−9
BJ-RPCG(b)	IC	NIT	12	12	12	12	12	12	12
		CPU	0.82	0.97	1.18	1.37	1.61	1.87	2.25
		ERR	6.53e−9	5.02e−9	6.55e−9	4.82e−9	6.51e−9	4.62e−9	6.44e−9
	SGS	NIT	13	13	13	13	13	13	13
		CPU	0.67	0.80	0.94	1.10	1.25	1.45	1.92
		ERR	9.45e−9	8.45e−9	9.35e−9	8.14e−9	9.13e−9	7.80e−9	8.88e−9
BSGS-RPCG(a)	IC	NIT	12	11	12	11	11	11	11
		CPU	1.47	1.56	1.92	2.11	2.42	2.88	3.43
		ERR	2.43e−9	9.42e−9	2.32e−9	8.91e−9	9.83e−9	8.43e−9	9.38e−9
	SGS	NIT	10	10	10	10	10	10	10
		CPU	1.30	1.63	1.86	2.22	2.51	2.97	3.68
		ERR	8.07e−9	6.89e−9	7.76e−9	6.59e−9	7.44e−9	6.31e−9	7.15e−9
BSGS-RPCG(b)	IC	NIT	10	10	10	10	10	10	10
		CPU	1.43	1.69	2.04	2.35	2.78	3.13	3.84
		ERR	9.22e−9	4.49e−9	9.14e−9	4.29e−9	9.01e−9	4.09e−9	8.87e−9
	SGS	NIT	12	12	12	12	12	12	12
		CPU	1.36	1.62	1.94	2.25	2.58	2.98	3.77
		ERR	9.14e−9	7.48e−9	8.86e−9	7.14e−9	8.52e−9	6.78e−9	8.18e−9
CG		NIT	230	244	266	280	300	314	335
		CPU	1.71	2.21	2.96	3.73	4.44	5.42	6.92
		ERR	9.45e−9	9.82e−9	9.96e−9	9.43e−9	9.79e−9	9.87e−9	9.48e−9
PCG	IC	NIT	38	38	38	38	38	37	38
		CPU	1.01	1.21	1.43	1.67	1.92	2.30	2.63
		ERR	8.02e−9	6.87e−9	7.81e−9	6.55e−9	7.47e−9	9.90e−9	7.12e−9
	SGS	NIT	16	14	16	14	16	14	16
		CPU	0.76	0.84	1.06	1.14	1.92	1.56	1.90
		ERR	3.58e−9	7.75e−9	3.75e−9	7.39e−9	3.82e−9	7.02e−9	3.84e−9

From Tables 1 and 2 we see that for different stepsizes h , the numbers of iteration steps of each of the tested methods, roughly speaking, remain almost the same, however, the CPU times have apparent increases. All methods can obtain high-accuracy approximations to the exact solution of the system of linear equations (1.1)–(1.2). We also see that in the sense of iteration step and CPU time, all RPCG-like methods are much more effective than the CG method, and both IC and SGS preconditioners can considerably improve the numerical behaviors of all tested methods.

Table 2
NITs, CPUs and ERRs for Example 5.2

h^{-1} n			104	112	120	128	136	144
			10,609	12,321	14,161	16,129	18,225	20,449
BJ-RPCG(a)	IC	NIT	12	13	12	13	12	13
		CPU	0.82	1.00	1.13	1.48	1.52	1.89
		ERR	9.25e−9	4.00e−9	8.75e−9	3.87e−9	8.31e−9	3.74e−9
	SGS	NIT	12	12	12	12	12	12
		CPU	0.86	0.98	1.18	1.34	1.57	1.93
		ERR	2.40e−9	2.81e−9	2.26e−9	2.67e−9	2.12e−9	2.54e−9
BJ-RPCG(b)	IC	NIT	12	12	12	12	12	12
		CPU	0.97	1.17	1.37	1.64	1.87	2.26
		ERR	3.40e−9	7.82e−9	3.21e−9	7.56e−9	3.05e−9	7.30e−9
	SGS	NIT	13	14	13	13	13	13
		CPU	0.80	1.00	1.09	1.24	1.43	1.79
		ERR	8.74e−9	2.87e−9	8.23e−9	9.97e−9	7.75e−9	9.77e−9
BSGS-RPCG(a)	IC	NIT	11	12	11	12	11	12
		CPU	1.55	1.93	2.17	2.62	2.86	3.66
		ERR	6.61e−9	2.35e−9	6.18e−9	2.29e−9	5.82e−9	2.23e−9
	SGS	NIT	10	10	10	10	10	10
		CPU	1.60	1.87	2.22	2.55	2.98	3.68
		ERR	5.59e−9	6.41e−9	5.30e−9	6.04e−9	5.00e−9	5.67e−9
BSGS-RPCG(b)	IC	NIT	10	10	10	10	10	10
		CPU	1.68	2.03	2.35	2.79	3.12	3.85
		ERR	3.14e−9	9.47e−9	2.95e−9	9.24e−9	2.79e−9	9.01e−9
	SGS	NIT	12	13	12	13	12	13
		CPU	1.63	2.08	2.23	2.80	2.97	4.07
		ERR	7.38e−9	3.07e−9	6.85e−9	3.04e−9	6.38e−9	2.98e−9
CG		NIT	185	199	210	229	239	254
		CPU	1.85	2.21	2.71	3.51	4.36	5.19
		ERR	9.18e−9	9.86e−9	9.86e−9	8.47e−9	9.64e−9	9.98e−9
PCG	IC	NIT	38	38	38	38	37	38
		CPU	1.21	1.43	1.66	1.93	2.29	2.64
		ERR	7.25e−9	8.07e−9	6.73e−9	7.58e−9	9.34e−9	7.17e−9
	SGS	NIT	15	17	15	17	15	17
		CPU	0.87	1.10	1.18	1.46	1.63	1.96
		ERR	5.38e−9	5.11e−9	5.17e−9	5.09e−9	4.95e−9	5.04e−9

According to the preconditioning effect, we have the following observations from both [Tables 1](#) and [2](#):

- (i) In the sense of iteration step, all RPCG-like methods are much faster than the PCG-like methods.
- (ii) However, in the sense of CPU time, all BJ-RPCG-like methods are comparatively effective with the PCG methods. The IC-type preconditioners are comparatively effective with the SGS-type preconditioners for RPCG(a), less effective than the SGS-type preconditioners for both RPCG(b) and PCG, and are equally effective to the SGS-type preconditioners for the BSGS-RPCG-like methods.

In addition, we observe that the BJ-RPCG-like methods outperform the BSGS-RPCG-like methods for all cases.

Here, we should point out again that the main difference between RPCG and PCG is on the preconditioner. The former uses a structured preconditioner, while the latter uses a standard one.

6. Conclusion and remarks

For the large sparse symmetric positive definite systems of linear equations of block two-by-two structures, we have presented a class of restrictive preconditioners by technically and sufficiently using the block structures of the coefficient matrices and, consequently, established a class of practical RPCG methods. Unlike the RPCG methods studied in [9], the new RPCG methods here do not need exact solutions of the sub-systems of linear equations with the coefficient matrices B and C , which are the diagonal blocks of the original coefficient matrix A . Theoretically, the new RPCGs have similar convergence properties to PCGs, however, numerically, they may show better computing behaviors, in particular, when good restrictive preconditioners are easily available. Therefore, our new restrictive preconditioners are practical and effective alternatives of the classical incomplete Cholesky factorization and splitting iteration preconditioners, and the correspondingly induced RPCG methods are competitive iteration solvers to the classical PCG methods for solving the large sparse symmetric positive definite systems of linear equations of block two-by-two structures.

Acknowledgements

The authors thank Ms. Gui-Qing Li for her kind help on coding the test methods, and the referees for their valuable comments and suggestions. Part of this work was done during the first author's visit to Shanghai Key Laboratory of Contemporary Applied Mathematics of Fudan University, so he thanks professors Yang-Feng Su and Zong-Min Wu very much for their kind invitations.

References

- [1] S.F. Ashby, M.J. Holst, T.A. Manteuffel, P.E. Saylor, The role of the inner product in stopping criteria for conjugate gradient iterations, *BIT* 41 (2001) 26–52.
- [2] O. Axelsson, *Iterative Solution Methods*, Cambridge University Press, Cambridge, 1994.
- [3] O. Axelsson, I. Gustafsson, Preconditioning and two-level multigrid methods of arbitrary degree of approximation, *Math. Comput.* 40 (1983) 219–242.
- [4] O. Axelsson, P.S. Vassilevski, Algebraic multilevel preconditioning methods. I, *Numer. Math.* 56 (1989) 157–177.
- [5] O. Axelsson, P.S. Vassilevski, A survey of multilevel preconditioned iterative methods, *BIT* 29 (1989) 769–793.
- [6] O. Axelsson, P.S. Vassilevski, Algebraic multilevel preconditioning methods. II, *SIAM J. Numer. Anal.* 27 (1990) 1569–1590.
- [7] Z.-Z. Bai, A class of modified block SSOR preconditioners for symmetric positive definite systems of linear equations, *Adv. Comput. Math.* 10 (1999) 169–186.
- [8] Z.-Z. Bai, Modified block SSOR preconditioners for symmetric positive definite linear systems, *Ann. Oper. Res.* 103 (2001) 263–282.
- [9] Z.-Z. Bai, G.-Q. Li, Restrictively preconditioned conjugate gradient methods for systems of linear equations, *IMA J. Numer. Anal.* 23 (2003) 561–580.

- [10] V.L. Eijkhout, P.S. Vassilevski, The role of the strengthened Cauchy–Buniakowskii–Schwarz inequality in multilevel methods, *SIAM Rev.* 33 (1991) 405–419.
- [11] V. Faber, T.A. Manteuffel, Necessary and sufficient conditions for the existence of a conjugate gradient method, *SIAM J. Numer. Anal.* 21 (1984) 352–362.
- [12] G.H. Golub, C.F. Van Loan, *Matrix Computations*, third ed., The Johns Hopkins University Press, Baltimore, London, 1996.
- [13] A. Greenbaum, *Iterative Methods for Solving Linear Systems*, SIAM, Philadelphia, PA, 1997.
- [14] M.H. Gutknecht, Changing the norm in conjugate gradient type algorithms, *SIAM J. Numer. Anal.* 30 (1993) 40–56.
- [15] M.R. Hestenes, E.L. Stiefel, Methods of conjugate gradients for solving linear systems, *J. Res. Natl. Bur. Standards Sec. B* 49 (1952) 409–436.
- [16] J.A. Meijerink, H.A. van der Vorst, An iterative solution method for linear systems of which the coefficient matrix is a symmetric M -matrix, *Math. Comput.* 31 (1977) 148–162.
- [17] J.M. Ortega, *Introduction to Parallel and Vector Solution of Linear Systems*, Plenum Press, New York, London, 1988.
- [18] Y. Saad, *Iterative Methods for Sparse Linear Systems*, PWS Publishing Company, Boston, 1996.
- [19] J.-C. Sun, On positive symmetrizable matrices and pre-symmetry iteration algorithms, *Math. Numer. Sin.* 22 (2000) 379–384.