"Numerische Mathematik" manuscript No.
(will be inserted by the editor)

On the Global Convergence of the Block Jacobi Method for the Positive Definite Generalized Eigenvalue Problem

Vjeran Hari

Abstract The paper proves the global convergence of a general block Jacobi method for the generalized eigenvalue problem $\mathbf{A}x = \lambda\mathbf{B}x$ with symmetric matrices $\mathbf{A}$, $\mathbf{B}$ such that B is positive definite. The proof is made for a large class of generalized serial strategies from [26]. The sequence of matrix pairs generated by the block method converges to $(\boldsymbol{\Lambda}, \mathbf{I})$ where $\boldsymbol{\Lambda}$ is a diagonal matrix of the eigenvalues of the initial matrix pair $(\mathbf{A}, \mathbf{B})$ and $\mathbf{I}$ is the identity matrix. The proof is made for the case when the block method uses element-wise HZ or CJ method as the kernel algorithm. In the case of other kernel algorithms, the paper proves that the iteration matrices tend to diagonal form.

Keywords generalized eigenvalue problem · block Jacobi method · global convergence

Mathematics Subject Classification (2000) 65F15

1 Introduction

We consider the global convergence problem of a block Jacobi method for solving the positive definite eigenvalue problem (PGEP) under the class of generalized serial strategies from [26]. The matrix problem is written in the form

$$\mathbf{A}x = \lambda\mathbf{B}x, \qquad x \neq 0,$$

where $\mathbf{A}$ and $\mathbf{B}$ are symmetric matrices such that $\mathbf{B}$ is positive definite.

With the development of high-performance parallel machines, there is a growing need for the software that solves specific mathematical problems. The BLAS (Basic Linear Algebra Subprograms) are routines for performing basic vector and matrix operations. The Level 3 BLAS perform matrix-matrix operations. Because the BLAS are efficient, portable, and widely available, they are commonly used in the development of high-quality linear algebra software, LAPACK for example. In the last decades many known methods have been modified to become BLAS 3 algorithms, i.e. almost entirely rely on the use of Level 3 BLAS

This work has been fully supported by Croatian Science Foundation under the project IP-09-2014-3670.

Faculty of Science, Department of Mathematics, University of Zagreb, Croatia
Tel.: +385-1-4605748, E-mail: hari@math.hr

routines. This seems to be a sensible way to achieve high efficiency across a wide range of serial and parallel computer architectures.

Jacobi methods for the eigenvalue and singular value problems are known for their accuracy (see [5], [35], [9], [6], [36], [8], [31]), [30], [27]) inherent parallelism ([7], [33], [18], [16]) and efficiency (see [10], [11], [12], [19]). The best way to further enhance these characteristics is to modify them to become BLAS 3 algorithms. Such methods are referred to as block diagonalization or block Jacobi methods [21], [1], [32], [2]. The global convergence of various block Jacobi methods has been considered in [20], [13], [23], [24], [4], [25] while the asymptotic quadratic convergence has been considered in [4], [3].

In this paper we consider the block Jacobi method for PGEP which generates two sequences of matrices by the same rule: $\mathbf{A}^{(k+1)} = \mathbf{Z}^{(k)}\mathbf{A}^{(k)}\mathbf{Z}^{(k)}$, $\mathbf{B}^{(k+1)} = \mathbf{Z}^{(k)}\mathbf{B}^{(k)}\mathbf{Z}^{(k)}$, $k \geq 0$. Here the initial $\mathbf{A}$ is symmetric and $\mathbf{B}$ is symmetric positive definite. The transformation matrices $\mathbf{Z}^{(k)}$ are nonsingular elementary block matrices. Their task is to diagonalize the (block) pivot submatrices of $\mathbf{A}^{(k)}$ and $\mathbf{B}^{(k)}$ in each step. As has been shown in [32] the block Jacobi method for PGEP is very fast and accurate, especially if it is implemented as a one-sided method that solves the generalized singular value problem (GSVD). On contemporary parallel machines it is the method of choice for solving the GSVD problem.

The paper proves the global convergence of the block Jacobi method for the cyclic pivot strategies that are generalizations of the known serial (row- and column-cyclic) pivot strategies. These pivot strategies belong to a wide class of generalized serial strategies whose theory is thoroughly explained in [26]. That class encompasses most common sequential and parallel pivot strategies (see [17], [33], [18], [28], [29], [34]).

The paper is divided into 4 sections. In Section 2, we introduce notation and define the block Jacobi method, pivot strategies, and the global convergence. In Section 3, we prove the global convergence of the block Jacobi method. Finally, in Section 4, we briefly discuss open problems and future work.

2 Block Jacobi Methods for PGEP

Here we define basic notation and definitions linked to the block Jacobi method for PGEP.

Let $\mathbf{A}$ and $\mathbf{B}$ be real symmetric matrices of order n such that $\mathbf{B}$ is positive definite. Let

$$\mathbf{A} = \begin{bmatrix} A_{11} & \cdots & A_{1m} \\ \vdots & \ddots & \vdots \\ A_{m1} & \cdots & A_{mm} \end{bmatrix} \begin{matrix} n_1 \\ \vdots \\ n_m \end{matrix}, \qquad \mathbf{B} = \begin{bmatrix} B_{11} & \cdots & B_{1m} \\ \vdots & \ddots & \vdots \\ B_{m1} & \cdots & B_{mm} \end{bmatrix} \begin{matrix} n_1 \\ \vdots \\ n_m \end{matrix} \tag{2.1}$$
$$\begin{matrix} n_1 & \cdots & n_m \end{matrix} \qquad\qquad \begin{matrix} n_1 & \cdots & n_m \end{matrix}$$

be their *block-matrix partitions* where the diagonal blocks A_{ii}, B_{ii}, $1 \leq i \leq m$ are square. The block-matrix partition (2.1) of $\mathbf{A}$ ($\mathbf{B}$) is determined by the partition $\pi_{n,m} = (n_1, \ldots, n_m)$ of n, where $n_i \geq 1$ for all $1 \leq i \leq m$ and $n_1 + \cdots + n_m = n$. Since n and m will be constant, we shall use notation π instead of $\pi_{n,m}$.

A block Jacobi method for PGEP is an iterative process of the form

$$\mathbf{A}^{(k+1)} = [\mathbf{Z}^{(k)}]^T \mathbf{A}^{(k)} \mathbf{Z}^{(k)}, \qquad \mathbf{B}^{(k+1)} = [\mathbf{Z}^{(k)}]^T \mathbf{B}^{(k)} \mathbf{Z}^{(k)}, \qquad k \geq 0, \tag{2.2}$$

where each $\mathbf{Z}^{(k)}$ is an *elementary block matrix* [25], $[\mathbf{Z}^{(k)}]^T$ is the transpose of $\mathbf{Z}^{(k)}$ and $\mathbf{A}^{(0)}$, $\mathbf{B}^{(0)}$ are the starting matrices. Each matrix $\mathbf{Z}^{(k)}$ is a non-singular $n \times n$ matrix of the form

$$\mathbf{Z}^{(k)} = \mathbf{Z}_{ij}^{(k)} = \begin{bmatrix} I & & & \\ & Z_{ii}^{(k)} & Z_{ij}^{(k)} & \\ & & I & \\ & Z_{ji}^{(k)} & Z_{jj}^{(k)} & \\ & & & I \end{bmatrix} \begin{matrix} \\ \}n_i \\ \\ \}n_j \\ \end{matrix},$$

where $\mathbf{Z}^{(k)}$ carries the same partition as $\mathbf{A}$. All elements of $\mathbf{Z}^{(k)}$, except those of the blocks $Z_{ii}^{(k)}$, $Z_{ij}^{(k)}$, $Z_{ji}^{(k)}$ and $Z_{jj}^{(k)}$, are as in the identity matrix $\mathbf{I}_n$. The "block indices" $i = i(k)$, $j = j(k)$ are *pivot indices*, (i, j) is *pivot pair* and

$$\hat{\mathbf{Z}}^{(k)} = \begin{bmatrix} Z_{ii}^{(k)} & Z_{ij}^{(k)} \\ Z_{ji}^{(k)} & Z_{jj}^{(k)} \end{bmatrix}$$

is *pivot submatrix* or the (i, j) – *restriction* of $\mathbf{Z}^{(k)}$. The submatrix $Z_{ij}^{(k)}$ is referred to as the *pivot block* of $\mathbf{Z}^{(k)}$. The way of selecting the pivot pairs is referred to as *pivot strategy*.

We shall write $\mathbf{Z}_{ij}^{(k)} = \mathscr{E}(i, j; \hat{\mathbf{Z}}^{(k)})$, where $\mathscr{E} = \mathscr{E}_\pi$ is the mapping that constructs the $n \times n$ matrix $\mathbf{Z}_{ij}^{(k)}$ from the input data i, j and $\hat{\mathbf{Z}}^{(k)}$. When emphasis is on pivot indices, we shall write $\mathbf{Z}_{i(k)j(k)}^{(k)}$ instead of $\mathbf{Z}^{(k)}$ or simply $\mathbf{Z}_{ij}$ when k is clear from the context.

The starting matrices $\mathbf{A}^{(0)}$, $\mathbf{B}^{(0)}$ for the process (2.2) are obtained from the initial matrices $\mathbf{A}$, $\mathbf{B}$, via the congruence transformation with the block-diagonal matrix $\mathbf{Z}^{(0)}$,

$$\mathbf{A}^{(0)} = [\mathbf{Z}^{(0)}]^T \mathbf{A} \mathbf{Z}^{(0)}, \qquad \mathbf{B}^{(0)} = [\mathbf{Z}^{(0)}]^T \mathbf{B} \mathbf{Z}^{(0)}, \qquad \mathbf{Z}^{(0)} = \operatorname{diag}(Z_{11}^{(0)}, \ldots, Z_{mm}^{(0)}). \quad (2.3)$$

The purpose of this transformation is to make the diagonal blocks of $\mathbf{A}^{(0)}$ diagonal and the diagonal blocks of $\mathbf{B}^{(0)}$ the identity matrices. Typically, one can choose $Z_{rr}^{(0)}$ to be inverse of the Cholesky factor of B_{rr} postmultiplied by an orthogonal matrix which diagonalizes the updated A_{rr}:

$$Z_{rr}^{(0)} = R_r^{-1} Q_r, \quad R_r^T R_r = B_{rr}, \qquad R_r^{-T} A_{rr} R_r^{-1} = Q_r \Lambda_r Q_r^T, \qquad 1 \le r \le m.$$

For accuracy reasons, a better option is to apply some element-wise Jacobi method for the PGEP (see [27]) to each pair (A_{rr}, B_{rr}), $1 \le r \le n$, collect the transformation matrices in $\mathbf{Z}^{(0)}$ and apply the transformation (2.3) to the off-block-diagonal parts of $\mathbf{A}$ and $\mathbf{B}$.

The process (2.2) is defined, if in each step k, one knows the pivot pair (i, j) and the *algorithm* which computes the pivot submatrix $\hat{\mathbf{Z}}^{(k)}$ from the elements of $\hat{\mathbf{A}}^{(k)}$ and $\hat{\mathbf{B}}^{(k)}$. We call the process *PGEP block Jacobi method* or simply *block Jacobi method*. In each step the block Jacobi method applies the congruence transformation (2.2) which preserves symmetry and positive definiteness of the current matrices.

In step k the *block Jacobi method* diagonalizes $\hat{\mathbf{A}}^{(k)}$ and transforms $\hat{\mathbf{B}}^{(k)}$ *into the identity matrix*. Locally, at the level of pivot submatrices, we have

$$\begin{bmatrix} [Z_{ii}^{(k)}]^T & [Z_{ji}^{(k)}]^T \\ [Z_{ij}^{(k)}]^T & [Z_{jj}^{(k)}]^T \end{bmatrix} \begin{bmatrix} A_{ii}^{(k)} & A_{ij}^{(k)} \\ [A_{ij}^{(k)}]^T & A_{jj}^{(k)} \end{bmatrix} \begin{bmatrix} Z_{ii}^{(k)} & Z_{ij}^{(k)} \\ Z_{ji}^{(k)} & Z_{jj}^{(k)} \end{bmatrix} = \begin{bmatrix} \Lambda_{ii}^{(k)} & \\ & \Lambda_{jj}^{(k)} \end{bmatrix}, \quad (2.4)$$

$$\begin{bmatrix} [Z_{ii}^{(k)}]^T & [Z_{ji}^{(k)}]^T \\ [Z_{ij}^{(k)}]^T & [Z_{jj}^{(k)}]^T \end{bmatrix} \begin{bmatrix} I_{n_i} & B_{ij}^{(k)} \\ [B_{ij}^{(k)}]^T & I_{n_j} \end{bmatrix} \begin{bmatrix} Z_{ii}^{(k)} & Z_{ij}^{(k)} \\ Z_{ji}^{(k)} & Z_{jj}^{(k)} \end{bmatrix} = \begin{bmatrix} I_{n_i} & \\ & I_{n_j} \end{bmatrix}. \quad (2.5)$$

Here each block $A_{ii}^{(k)}$ ($A_{jj}^{(k)}$) is diagonal and of order n_i (n_j). Note that $\Lambda_{ii}^{(k)}$ and $\Lambda_{jj}^{(k)}$ are exactly the diagonal blocks $A_{ii}^{(k+1)}$ and $A_{jj}^{(k+1)}$, respectively. We keep the superscript (k) because the diagonal elements of $\Lambda_{ii}^{(k)}$ and $\Lambda_{jj}^{(k)}$ are the eigenvalues of the pair $(\hat{\mathbf{A}}^{(k)},\hat{\mathbf{B}}^{(k)})$.

If $n_1 = n_2 = \cdots = n_m = 1$, we speak of a *non-block* or *element-wise* Jacobi method.

With each block Jacobi method is associated some *kernel algorithm* whose task is to transform $(\hat{\mathbf{A}}^{(k)},\hat{\mathbf{B}}^{(k)})$ into $(\mathrm{diag}\,(\Lambda_{ii}^{(k)},\Lambda_{jj}^{(k)}),\mathbf{I}_{n_i+n_j})$. For that purpose, one can use an element-wise HZ or CJ method from [27]. As a kernel algorithm, one can also use product of inverse of the Cholesky factor of $\hat{\mathbf{B}}^{(k)}$ with the orthogonal matrix from the spectral decomposition of the updated $\hat{\mathbf{A}}^{(k)}$.

The partition π usually remains constant during the whole process, but it can now and then be adapted to ensure faster asymptotic convergence of the iterative process, especially in the case of multiple eigenvalues of the initial pair $(\mathbf{A},\mathbf{B})$.

Recall that each principal submatrix of a positive definite matrix is positive definite. This implies that all matrices $\hat{\mathbf{B}}^{(k)}$, $k \geq 0$ are positive definite. Furthermore, each block Jacobi method maintains the special form of the block-diagonal parts of $\mathbf{A}^{(0)}$ and $\mathbf{B}^{(0)}$. Namely, we have

$$A_{rr}^{(k)} = \mathrm{diag}\,(A_{rr}^{(k)}), \qquad B_{rr}^{(k)} = I_{n_r}, \qquad 1 \leq r \leq m, \qquad k \geq 0.$$

This property of the block Jacobi method has a double effect on the iterated matrices: it normalizes and balances each $\mathbf{B}^{(k)}$ and it also nicely preconditions $\mathbf{B}^{(k)}$. Namely, $\mathbf{B}^{(k)}$ has now an almost optimal condition that can be obtained by the symmetric diagonal scaling [37]. As it will proved in Lemma 3.3(ii), this property ensures that all matrices $\mathbf{A}^{(k)}$ and $\mathbf{B}^{(k)}$ are bounded by norm.

2.1 Pivot Strategies

Let us very briefly introduce the class of generalized serial strategies from [26].

A pivot strategy can be identified with a function $\mathsf{I} : \mathscr{N}_0 \to \mathscr{P}_m$, where

$$\mathscr{N}_0 = \{0,1,2,\ldots\}, \qquad \mathscr{P}_m = \{(r,t);\, 1 \leq r < t \leq m\}.$$

We see that $\mathscr{P}_m$ contains pairs of indices which address the blocks in the upper block-triangle of the matrix $\mathbf{A}$. The set $\mathscr{P}_m$ contains $M = m(m-1)/2$ pairs of indices. Once $n \geq 2$, $m \geq 1$ and $\pi = (n_1,\ldots,n_m)$ are given, the function I can be defined in many ways. If I is a periodic function, then I is called *periodic pivot strategy*. Let I be the periodic (pivot) strategy with period P. If $P \geq M$ ($P = M$) and $\{\mathsf{I}(k) : k = 0,1,\ldots,P-1\} = \mathscr{P}_m$, then I is called *quasi-cyclic* (*cyclic*) strategy.

In this paper, we shall use the class of *generalized serial strategies* which is a subset of the set of cyclic pivot strategies. That class of strategies was introduced in [26] and it includes serial, wavefront and weak-wavefront strategies from [34], then inverses of weak-wavefront strategies and all those cyclic strategies that are permutation equivalent to all of them. Hence they include the modulus strategy [29] and some other cyclic strategies that are used for parallel processing. The theory of generalized serial strategies and of their quasi-cyclic derivatives can be found in [26].

3 The Global Convergence

Here we prove the global convergence of the PGEP block Jacobi method under the class of generalized serial strategies. A similar, but somewhat more complicated proof can be made for the corresponding quasi-cyclic strategies from [26].

To measure advancement of the method, one can use quantity $S(A,B)$,

$$S(A,B) = \left[S^2(A) + S^2(B)\right]^{1/2}, \qquad S(H) = \|H - \mathrm{diag}\,(H)\|_F,$$

where A, B, H are square matrices and $\|X\|_F$ is the Frobenius norm of X. Here $\mathrm{diag}\,(X)$ denotes the diagonal part of X. The measure $S(X)$ is referred to as *departure from diagonal form* or as *off-norm* of X. So, $S(A,B)$ can be called *off-norm of the pair* (A,B).

In the following definition it is assumed that a block Jacobi method is defined by some partition π od n.

Definition 3.1 Let $\mathbf{A}$ and $\mathbf{B}$ be symmetric matrices of order n such that $\mathbf{B}$ is positive definite. Let the sequence of matrix pairs $\left((\mathbf{A}^{(k)},\mathbf{B}^{(k)}),\ k \geq 0\right)$ be generated by applying the block Jacobi method to the pair $(\mathbf{A},\mathbf{B})$. The block Jacobi method is *convergent* on $(\mathbf{A},\mathbf{B})$ if $\mathbf{B}^{(k)} \to \mathbf{I}_n$ and $\mathbf{A}^{(k)} \to \mathbf{\Lambda}$ as $k \to \infty$, where $\mathbf{\Lambda}$ is a diagonal matrix of the eigenvalues of $(\mathbf{A},\mathbf{B})$. If the block Jacobi method is convergent on every pair $(\mathbf{A},\mathbf{B})$ then it is *globally convergent*. The block Jacobi method *converges to diagonal form* if for every pair $(\mathbf{A},\mathbf{B})$, $S(\mathbf{A}^{(k)},\mathbf{B}^{(k)}) \to 0$ as $k \to \infty$.

The global convergence of one-sided block Jacobi methods for solving the GSVD problem is defined using their two-sided counterparts which solve the PGEP with positive definite $\mathbf{A}$ and $\mathbf{B}$. So, the result obtained here has a direct application to one-sided block Jacobi methods.

We shall use notation $\sigma_t(X)$ for the t'th largest singular value of X, and $\sigma_{min}(X)$ for the minimum singular value of X. The convergence proof is based on the following general result from [26].

Theorem 3.2 *Let $\pi = (n_1,\ldots,n_m)$ be a partition of n. Let $\mathbf{H} = (H_{rt}) \neq 0$ be a symmetric matrix of order n partitioned in accordance with π. Let $(\mathbf{H}^{(k)},\ k \geq 0)$ be the sequence generated by applying a block Jacobi-type process to $\mathbf{H}$,*

$$\mathbf{H}^{(k+1)} = \mathbf{F}_k^T \mathbf{H}^{(k)} \mathbf{F}_k, \quad \mathbf{H}^{(0)} = \mathbf{H}, \quad k \geq 0.$$

Here each $\mathbf{F}_k$ is an elementary block matrix defined by the pivot pair $(i(k),j(k))$, where $1 \leq i(k) < j(k) \leq m$. Suppose the following assumptions are satisfied:

(A1) *The pivot strategy is generalized serial*

(A2) *There is a sequence of orthogonal elementary block matrices $(\mathbf{U}^{(k)},\ k \geq 0)$ such that*

$$\lim_{k\to\infty} (\mathbf{F}_k - \mathbf{U}_k) = 0$$

(A3) *For the diagonal block $F^{(k)}_{i(k)i(k)}$ of $\mathbf{F}_k$ one has*

$$\sigma = \liminf_{k\to\infty} \sigma^{(k)} > 0, \qquad \sigma^{(k)} = \sigma_{min}(F^{(k)}_{i(k)i(k)}),\ k \geq 0$$

(A4) *The sequence $(\mathbf{H}^{(k)},\ k \geq 0)$ is bounded.*

Then the following two conditions are equivalent

(i) $\lim_{k\to\infty} S(\hat{\mathbf{F}}_k^T \hat{\mathbf{H}}^{(k)} \hat{\mathbf{F}}_k) = 0$

(ii) $\lim_{k\to\infty} S(\mathbf{H}^{(k)}) = 0.$

Proof The theorem is a special case of [25, Corollary 5.3] with $E^{(k)} = 0$, $k \geq 0$. □

In Theorem 3.2 the pivot submatrix $\hat{\mathbf{H}}^{(k)}$ does not have to be diagonalized in each step, but the condition **(i)** has to hold. Such are the norm-reducing methods of Eberlein [14], [15]. That is the reason why we used the term Jacobi-type methods in Theorem 3.2.

We shall apply Theorem 3.2 to the sequences $(\mathbf{A}^{(k)}, k \geq 0)$ and $(\mathbf{B}^{(k)}, k \geq 0)$ obtained by the block Jacobi method. To this end we shall prove some preparatory results. First, we want to prove that all matrices $\mathbf{A}^{(k)}$, $\mathbf{B}^{(k)}$, generated by the method are bounded. That accounts for the assumption **A4** of Theorem 3.2. Then we want to prove that $B^{(k)}_{i(k)\,j(k)}$ tends to zero as k increases. Once we prove it, the other assertions of Theorem 3.2 will be easy to show.

In the following lemma we use the spectral radius of the matrix pair $(\mathbf{A}, \mathbf{B})$,

$$\mu = \max_{\lambda \in \sigma(\mathbf{A},\mathbf{B})} |\lambda|,$$

where $\sigma(\mathbf{A}, \mathbf{B})$ denotes the spectrum of $(\mathbf{A}, \mathbf{B})$. Here $\mathbf{A}$, $\mathbf{B}$ are arbitrary symmetric matrices such that $\mathbf{B}$ is positive definite.

Lemma 3.3 *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be symmetric matrices of order n such that* $\mathbf{B}$ *is positive definite. Let the sequences of matrices* $(\mathbf{A}^{(k)},\ k \geq 0)$, $(\mathbf{B}^{(k)},\ k \geq 0)$ *be generated by applying the block Jacobi method to the pair* $(\mathbf{A}, \mathbf{B})$ *under an arbitrary pivot strategy. Then the assertions (i)–(iii) hold:*

(i) *The matrices generated by the method are bounded and we have*

$$\|\mathbf{B}^{(k)}\|_2 < n, \qquad \|\mathbf{A}^{(k)}\|_2 \leq \mu \|\mathbf{B}^{(k)}\|_2 < n\mu\,, \qquad k \geq 0 \tag{3.1}$$

(ii) *For the pivot blocks* $B^{(k)}_{i(k)\,j(k)}$ *of* $\mathbf{B}^{(k)}$*, we have*

$$\lim_{k\to\infty} B^{(k)}_{i(k)\,j(k)} = 0$$

(iii) *For the transformation matrices* $\mathbf{Z}^{(k)}$*, we have*

$$\lim_{k\to\infty} \left(\mathbf{Z}^{(k)} - \mathbf{Q}^{(k)}\right) \to 0,$$

where $\mathbf{Q}^{(k)}$ *are orthogonal elementary block matrices.*

Proof (i) The proof of the relation (3.1) is identical to the proof of [27, Lemma 4.1].

(ii) The proof uses the technique from the proof of [27, Proposition 4.1]. So, let $\mathbf{B}^{(k)} = (b^{(k)}_{rs})$, $k \geq 0$ and

$$H(\mathbf{B}^{(k)}) = \frac{\det(\mathbf{B}^{(k)})}{b^{(k)}_{11} b^{(k)}_{22} \cdots b^{(k)}_{nn}} = \det(\mathbf{B}^{(k)}), \quad k \geq 0.$$

By the Hadamard's inequality we have

$$0 < H(\mathbf{B}^{(k)}) \leq 1, \quad k \geq 0.$$

Applying the determinant to the equations (2.2) and (2.5), we have

$$\det(\mathbf{B}^{(k+1)}) = \det^2(\mathbf{Z}^{(k)})\det(\mathbf{B}^{(k)}) \tag{3.2}$$

$$1 = \det^2(\hat{\mathbf{Z}}^{(k)})\det(\hat{\mathbf{B}}^{(k)}). \tag{3.3}$$

Using the singular value decomposition of $B_{ij}^{(k)}$, $B_{ij}^{(k)} = U_{ij}^{(k)}\Sigma_{ij}^{(k)}[V_{ij}^{(k)}]^T$, one obtains

$$\begin{bmatrix} I_{n_i} & B_{ij}^{(k)} \\ [B_{ij}^{(k)}]^T & I_{n_j} \end{bmatrix} = \begin{bmatrix} U_{ij}^{(k)} & \\ & V_{ij}^{(k)} \end{bmatrix} \begin{bmatrix} I_{n_i} & \Sigma_{ij}^{(k)} \\ [\Sigma_{ij}^{(k)}]^T & I_{n_j} \end{bmatrix} \begin{bmatrix} [U_{ij}^{(k)}]^T & \\ & [V_{ij}^{(k)}]^T \end{bmatrix}, \quad k \geq 0, \tag{3.4}$$

where

$$\Sigma_{ij}^{(k)} = \operatorname{diag}\left(\sigma_1(B_{ij}^{(k)}), \ldots, \sigma_{\min\{n_i,n_j\}}(B_{ij}^{(k)})\right), \qquad k \geq 0.$$

We conclude that the middle matrix on the right side of (3.4) is positive definite. Hence, for the spectral norm of $B_{ij}^{(k)}$ we have

$$\|B_{ij}^{(k)}\|_2 = \|\Sigma_{ij}^{(k)}\|_2 = \sigma_1(B_{ij}^{(k)}) < 1, \qquad k \geq 0.$$

From (3.4) we also see that the eigenvalues of $\hat{\mathbf{B}}^{(k)}$ are $1 \pm \sigma_t(B_{ij}^{(k)})$, $1 \leq t \leq \min\{n_i, n_j\}$ and $n_i + n_j - 2\min\{n_i, n_j\}$ eigenvalues are equal to 1. Therefore, we have

$$\det(\hat{\mathbf{B}}^{(k)}) = (1 - \sigma_1^2(B_{ij}^{(k)})) \cdots (1 - \sigma_{\min\{n_i,n_j\}}^2(B_{ij}^{(k)})) < 1, \qquad k \geq 0. \tag{3.5}$$

Since $\det(\mathbf{Z}^{(k)}) = \det(\hat{\mathbf{Z}}^{(k)})$, from the relation (3.3), we obtain

$$|\det(\mathbf{Z}^{(k)})| = \frac{1}{\sqrt{(1 - \sigma_1^2(B_{ij}^{(k)})) \cdots (1 - \sigma_{\min\{n_i,n_j\}}^2(B_{ij}^{(k)}))}} > 1. \tag{3.6}$$

From the relations (3.2) and (3.6) we have

$$H(B^{(k)}) = (1 - \sigma_1^2(B_{ij}^{(k)})) \cdots (1 - \sigma_{\min\{n_i,n_j\}}^2(B_{ij}^{(k)}))H(B^{(k+1)}), \quad k \geq 0. \tag{3.7}$$

Recall that $i = i(k)$, $j = j(k)$, $k \geq 0$. From the relations (3.7) and (3.5) we see that $H(B^{(k)})$ is an increasing sequence of positive real numbers, bounded above by 1. Hence it is convergent with limit ζ, $0 < \zeta \leq 1$. By taking the limit on the both sides of the equation (3.7), after cancelation with ζ, we obtain

$$\lim_{k\to\infty} (1 - \sigma_1^2(B_{i(k)j(k)}^{(k)})) \cdots (1 - \sigma_{\min\{n_{i(k)},n_{j(k)}\}}^2(B_{i(k)j(k)}^{(k)})) = 1\,.$$

In particular, this implies $\|B_{i(k)j(k)}^{(k)}\|_2 = \sigma_1(B_{i(k)j(k)}^{(k)}) \to 0$ as $k \to \infty$, which proves the assertion (ii) of the lemma.

(iii) Let

$$\hat{\mathbf{Z}}^{(k)} = \hat{\mathbf{H}}^{(k)}\hat{\mathbf{Q}}^{(k)}, \qquad [\hat{\mathbf{H}}^{(k)}]^T = \hat{\mathbf{H}}^{(k)}, \qquad [\hat{\mathbf{Q}}^{(k)}]^T\hat{\mathbf{Q}}^{(k)} = \mathbf{I}_{n_i+n_j} \tag{3.8}$$

be the polar decomposition of $\hat{\mathbf{Z}}^{(k)}$ with positive definite $\hat{\mathbf{H}}^{(k)}$ and orthogonal $\hat{\mathbf{Q}}^{(k)}$. From the relation (3.8) and $[\hat{\mathbf{Z}}^{(k)}]^T\hat{\mathbf{B}}^{(k)}\hat{\mathbf{Z}}^{(k)} = \mathbf{I}_{n_i+n_j}$, one obtains $\hat{\mathbf{H}}^{(k)} = [\hat{\mathbf{B}}^{(k)}]^{-1/2}$ and

$$\sigma(\hat{\mathbf{H}}^{(k)}) = \left\{ \frac{1}{\sqrt{1 \pm \sigma_1(B_{ij}^{(k)})}}, \ \cdots\ , \frac{1}{\sqrt{1 \pm \sigma_{\min\{n_i,n_j\}}(B_{ij}^{(k)})}}, 1, \ldots, 1 \right\}.$$

Here $\sigma(\hat{\mathbf{H}}^{(k)})$ is the spectrum of $\hat{\mathbf{H}}^{(k)}$ and it contains exactly $n_i+n_j-2\min\{n_i,n_j\}$ ones. Let $\mathbf{Z}^{(k)}=\mathscr{E}(i(k),j(k);\hat{\mathbf{Z}}^{(k)})$ and $\mathbf{Q}^{(k)}=\mathscr{E}(i(k),j(k);\hat{\mathbf{Q}}^{(k)})$. Then we have

$$\begin{aligned}\|\mathbf{Z}^{(k)}-\mathbf{Q}^{(k)}\|_2 &= \|\hat{\mathbf{Z}}^{(k)}-\hat{\mathbf{Q}}^{(k)}\|_2=\|(\hat{\mathbf{H}}^{(k)}-\mathbf{I}_{n_i+n_j})\,\hat{\mathbf{Q}}^{(k)}\|_2=\|\hat{\mathbf{H}}^{(k)}-\mathbf{I}_{n_i+n_j}\|_2\\ &= \frac{1}{\sqrt{1-\sigma_1(B^{(k)}_{i(k)j(k)})}}-1=\frac{\sigma_1(B^{(k)}_{i(k)j(k)})}{1-\sigma_1(B^{(k)}_{i(k)j(k)})+\sqrt{1-\sigma_1(B^{(k)}_{i(k)j(k)})}}.\end{aligned}$$

By (ii) we have $B^{(k)}_{i(k)j(k)}\to 0$ as $k\to\infty$. Hence $\sigma_1(B^{(k)}_{i(k)j(k)})\to 0$ and $\|\mathbf{Z}^{(k)}-\mathbf{Q}^{(k)}\|_2\to 0$ as $k\to\infty$. This proves the assertion (iii) of the lemma. □

In order to satisfy the condition (**A3**) of Theorem 3.2 the block Jacobi method can be modified in a similar way as it was suggested in [10] and [25] for other block Jacobi methods.

Suppose $\hat{\mathbf{Z}}^{(k)}$ is the pivot submatrix computed by some kernel algorithm in step k. Then equations (2.4) and (2.5) hold. If $\hat{\mathbf{Z}}^{(k)}$ is replaced by $\hat{\mathbf{Z}}^{(k)}\hat{\mathbf{P}}^{(k)}$ where $\hat{\mathbf{P}}^{(k)}$ is some permutation matrix then the right-hand side of the equation (2.5) does not change, while the one of the equation (2.4) has to be replaced by $[\hat{\mathbf{P}}^{(k)}]^T\mathrm{diag}\,(\Lambda^{(k)}_{ii},\Lambda^{(k)}_{jj})\hat{\mathbf{P}}^{(k)}$, which is again a diagonal matrix. So, we would like to replace each $\hat{\mathbf{Z}}^{(k)}$ by $\hat{\mathbf{Z}}^{(k)}\hat{\mathbf{P}}^{(k)}$ in such a way that the condition (**A3**) is satisfied.

We shall use notation

$$\hat{\mathbf{Z}}^{(k)}=\begin{bmatrix}Z^{(k)}_{ii} & Z^{(k)}_{ij}\\ Z^{(k)}_{ji} & Z^{(k)}_{jj}\end{bmatrix},\qquad \tilde{\mathbf{Z}}^{(k)}=\hat{\mathbf{Z}}^{(k)}\hat{\mathbf{P}}^{(k)}=\begin{bmatrix}\tilde{Z}^{(k)}_{ii} & \tilde{Z}^{(k)}_{ij}\\ \tilde{Z}^{(k)}_{ji} & \tilde{Z}^{(k)}_{jj}\end{bmatrix},\qquad k\ge 0.$$

The goal is to find, under acceptable computational cost, permutation matrices $\hat{\mathbf{P}}^{(k)}$ such that

$$\sigma_{min}(\tilde{Z}^{(k)}_{i(k)i(k)})\ge f(n_{i(k)},n_{j(k)})\ge f_1(\pi)\ge f_2(n)>0,\qquad k\ge k_0.$$

Here k_0 is a positive integer and f_1, f_2 are functions that do not depend on k.

By Lemma 3.3(iii) we have

$$\hat{\mathbf{Z}}^{(k)}=\hat{\mathbf{Q}}^{(k)}+\hat{\mathbf{E}}^{(k)},\qquad \lim_{k\to\infty}\hat{\mathbf{E}}^{(k)}=0. \tag{3.9}$$

Since each $\hat{\mathbf{Q}}^{(k)}$ is orthogonal, we have $\lim_{k\to\infty}\|\hat{\mathbf{Z}}^{(k)}\|_2^2=1$.

For $k\ge 0$, let $\mathbf{R}^{(k)}$ be any matrix with orthonormal rows and $n_{i(k)}+n_{j(k)}$ columns. Then $\|\mathbf{R}^{(k)}\hat{\mathbf{Q}}^{(k)}\|_F^2=\|\mathbf{R}^{(k)}\|_F^2$ equals number of rows of $\mathbf{R}^{(k)}$, which is not larger than $n_{i(k)}+n_{j(k)}$. Furthermore, for any orthogonal matrices $\hat{\mathbf{V}}^{(k)}_1$ and $\hat{\mathbf{V}}^{(k)}_2$ of order $n_{i(k)}+n_{j(k)}$, we have

$$\|\mathbf{R}^{(k)}(\hat{\mathbf{V}}^{(k)}_1\hat{\mathbf{Z}}^{(k)}\hat{\mathbf{V}}^{(k)}_2)\|_F=\|\mathbf{R}^{(k)}\hat{\mathbf{V}}^{(k)}_1\hat{\mathbf{Z}}^{(k)}\|_F,\quad \|\mathbf{R}^{(k)}(\hat{\mathbf{V}}^{(k)}_1\hat{\mathbf{Q}}^{(k)}\hat{\mathbf{V}}^{(k)}_2)\|_F=\|\mathbf{R}^{(k)}\|_F,\quad k\ge 0.$$

Hence

$$\lim_{k\to\infty}(\|\mathbf{R}^{(k)}(\hat{\mathbf{V}}^{(k)}_1\hat{\mathbf{Z}}^{(k)}\hat{\mathbf{V}}^{(k)}_2)\|_F^2-\|\mathbf{R}^{(k)}\|_F^2)=0. \tag{3.10}$$

Now, let us consider special matrices $\mathbf{R}^{(k)}$. For each $k\ge 0$, let $[e^{(k)}_1,\dots,e^{(k)}_{n_{i(k)}+n_{j(k)}}]$ be the column partition of $\mathbf{I}_{n_{i(k)}+n_{j(k)}}$ and $\mathbf{G}^{(k)}_r=\left[e^{(k)}_r,e^{(k)}_{r+1},\dots,e^{(k)}_{n_{i(k)}}\right]$, $1\le r\le n_{i(k)}$.

Let $\varepsilon_k>0$ be defined by

$$\varepsilon_k^2=\max_{1\le r\le n_{i(k)}}\left|\|[\mathbf{G}^{(k)}_r]^T\hat{\mathbf{Z}}^{(k)}\|_F^2-(n_{i(k)}-r+1)\right|,\qquad k\ge 0.$$

From the relation (3.10), we have

$$\lim_{k\to\infty} \varepsilon_k = 0. \tag{3.11}$$

The definition of ε_k implies

$$(n_{i(k)} - r + 1) - \varepsilon_k^2 \le \|[\mathbf{G}_r^{(k)}]^T \hat{\mathbf{Z}}^{(k)}\|_F^2 \le (n_{i(k)} - r + 1) + \varepsilon_k^2, \quad 1 \le r \le n_{i(k)}, \quad k \ge 0. \tag{3.12}$$

Let $\eta > 0$ be the smallest positive solution of the equation

$$(2-\eta^2)(3-\eta^2)\cdots(n-\eta^2) = \frac{n!}{1+\eta}.$$

Then for any ζ, $0 < \zeta < \eta$, we have $(1-\zeta^2)(2-\zeta^2)\cdots(n_t-\zeta^2) > n!\cdot(1-\zeta)$ and that implies

$$(1-\zeta^2)(2-\zeta^2)\cdots(n_t-\zeta^2) > n_t!\cdot(1-\zeta), \quad 0 < \zeta < \eta, \ 1 \le n_t \le n, \ 1 \le t \le m. \tag{3.13}$$

Here n_t is from the partition $\pi = (n_1, \ldots, n_m)$. From the relations (3.11), (3.12) and (3.13) we conclude that there exists $k_0 \ge 0$ such that

$$1-\varepsilon_k^2 \le \|\hat{\mathbf{Z}}^{(k)}\|_2^2 \le 1-\varepsilon_k^2, \ (1-\varepsilon_k^2)(2-\varepsilon_k^2)\cdots(n_{i(k)}-\varepsilon_k^2) > n_{i(k)}!\cdot(1-\varepsilon_k), \ k \ge k_0. \tag{3.14}$$

In the following lemma, we derive estimates related to step k of the block Jacobi method when $k \ge k_0$. We shall omit k and denote the pivot indices by i, j, and $\hat{\mathbf{Z}}^{(k)}$, $\hat{\mathbf{Q}}^{(k)}$, ε_k by $\hat{\mathbf{Z}}$, $\hat{\mathbf{Q}}$, ε, respectively.

Lemma 3.4 *Let*

$$\hat{\mathbf{Z}} = \begin{bmatrix} Z_{ii} & Z_{ij} \\ Z_{ji} & Z_{jj} \end{bmatrix}, \quad 1-\varepsilon^2 \le \|\hat{\mathbf{Z}}\|_2^2 \le 1+\varepsilon^2, \ (1-\varepsilon^2)(2-\varepsilon^2)\cdots(n_i-\varepsilon^2) > n_i!\cdot(1-\varepsilon).$$

Let $\hat{\mathbf{P}}$ be the permutation matrix from the QR factorization with column pivoting of $[Z_{ii}\ Z_{ij}]$ and let

$$\begin{bmatrix} \tilde{Z}_{ii} & \tilde{Z}_{ij} \end{bmatrix} = \begin{bmatrix} Z_{ii} & Z_{ij} \end{bmatrix} \hat{\mathbf{P}} = W \begin{bmatrix} R_{ii} & R_{ij} \end{bmatrix}, \qquad W^T W = I_{n_i}, \qquad R_{ii} \text{ is upper-triangular}.$$

Then

$$\sigma_{min}^2(\tilde{Z}_{ii}) = \sigma_{n_i}^2(\tilde{Z}_{ii}) = \sigma_{n_i}^2(R_{ii}) \ge \frac{1-\varepsilon}{(1+\varepsilon^2)^{n_i-1}} \Big/ \begin{pmatrix} n_i+n_j \\ n_i \end{pmatrix}. \tag{3.15}$$

Proof The first two equalities in the relation (3.15) are trivial. To prove the inequality, we apply the QR factorization algorithm with Householder reflectors to $[Z_{ii}, Z_{ij}]$. The algorithm consists of n_i steps. The whole process is described by the matrix equation

$$\begin{bmatrix} I_{n_i-1} & \\ & -1 \end{bmatrix} \cdots \begin{bmatrix} 1 & \\ & H_{n_i-1} \end{bmatrix} H_{n_i} \begin{bmatrix} Z_{ii} & Z_{ij} \end{bmatrix} I_{11'} I_{22'} \cdots I_{n_i n_i'} = \begin{bmatrix} R_{ii} & R_{ij} \end{bmatrix}. \tag{3.16}$$

Here $I_{rr'}$ is the transposition matrix with $r' \in \{r, r+1\ldots., n_i+n_j\}$ and R_{ii} is upper-triangular.

In the first step, the algorithm finds the column of largest Euclidean norm of $[Z_{ii}\ Z_{ij}]$. It is subscripted $1'$. Then the algorithm switches the columns 1 and $1'$ and applies the Householder reflector H_{n_i} from the left side. From the relation (3.12) we obtain by assuming $r = 1$,

$$\mathbf{G}_1^T \hat{\mathbf{Z}}^{(k)} = [Z_{ii}\ Z_{ij}], \qquad \|[Z_{ii}\ Z_{ij}]\|_F^2 \ge n_i - \varepsilon^2.$$

Since $[Z_{ii}\ Z_{ij}]$ has n_i+n_j columns, we obtain

$$|(R_{ii})_{11}|^2 = \|[Z_{ii}\ Z_{ij}]e_{1'}\|_2^2 \geq \frac{n_i-\varepsilon^2}{n_j+n_i}.$$

In the second step the algorithm works with the matrix $\mathbf{G}_2^T H_{n_i}\left[Z_{ii}\ Z_{ij}\right] I_{11'}$. It finds the column of largest norm and switches it with the second column. From the relations (3.10) and (3.12) we obtain, by assuming $\hat{\mathbf{V}}_1 = H_{n_i}$, $\hat{\mathbf{V}}_2 = I_{11'}$ and $r=2$,

$$\|\mathbf{G}_2^T H_{n_i}\left[Z_{ii}\ Z_{ij}\right] I_{11'}\|_F^2 \geq (n_i-1)-\varepsilon^2.$$

Since $\mathbf{G}_2^T H_{n_i}[Z_{ii}\,, Z_{ij}]I_{11'}$ has n_j+n_i-1 nontrivial columns, we obtain

$$|(R_{ii})_{22}|^2 \geq \frac{n_i-1-\varepsilon^2}{n_j+n_i-1}.$$

Continuing this procedure, we obtain

$$|(R_{ii})_{tt}|^2 \geq \frac{n_i-(t-1)-\varepsilon^2}{n_j+n_i-t+1}, \quad 1\leq t\leq n_i. \tag{3.17}$$

Applying the determinant to R_{ii} and to the singular value decomposition of R_{ii}, we have

$$\begin{aligned}\sigma_1^2(R_{ii})\cdots\sigma_{n_i}^2(R_{ii}) &= \det{}^2(R_{ii}) = |(R_{ii})_{11}|^2\cdots|(R_{ii})_{n_in_i}|^2 \geq \frac{(n_i-\varepsilon^2)\cdots(2-\varepsilon^2)(1-\varepsilon^2)}{(n_j+n_i)\cdots(n_j+2)(n_j+1)}\\ &= \frac{n_i!(1-\varepsilon)}{(n_i+n_j)!/n_j!} = (1-\varepsilon)/\tbinom{n_i+n_j}{n_i}.\end{aligned}$$

Here we used (3.17) and the assumptions of the lemma (which are based on the relation (3.14)). Since

$$\sigma_1^2(R_{ii})\cdots\sigma_{n_i-1}^2(R_{ii}) \leq \sigma_1^{2(n_i-1)}(R_{ii}) \leq (1+\varepsilon^2)^{n_i-1},$$

we obtain

$$\sigma_{n_i}^2(R_{ii}) \geq \frac{1-\varepsilon}{(1+\varepsilon^2)^{n_i-1}}/\binom{n_i+n_j}{n_i},$$

which completes the proof of the lemma. □

With $n\geq 2$, $m\geq 2$ and the partition $\pi=(n_1,\ldots,n_m)$, we associate numbers ν_π and $\tilde{\nu}_n$ for which we have

$$\nu_\pi \equiv \left[\max_{r<t}\binom{n_r+n_t}{n_r}\right]^{-1} \geq \left[\binom{n}{[n/2]}\right]^{-1} \equiv \tilde{\nu}_n.$$

Here $[n/2]$ is the largest integer smaller than or equal to $n/2$. From Lemma 3.4 we obtain

$$\sigma_{min}^2(\tilde{Z}_{i(k)i(k)}^{(k)})\frac{(1+\varepsilon_k^2)^{n_{i(k)}-1}}{1-\varepsilon_k} \geq \left[\binom{n_{i(k)}+n_{j(k)}}{n_{i(k)}}\right]^{-1} \geq \left[\max_{r<t}\binom{n_r+n_t}{n_r}\right]^{-1} = \nu_\pi, \quad k\geq k_0.$$

By letting $k\to\infty$, we have

$$\liminf_{k\to\infty}\sigma_{min}(\tilde{Z}_{i(k)i(k)}^{(k)}) \geq \nu_\pi \geq \tilde{\nu}_n. \tag{3.18}$$

Remark 3.5 In real computation, we shall have $n_1 = n_2 = \cdots = n_{m-1} = \lfloor n/m \rfloor$, $n_m \le n_1$. Hence the lower bound will be

$$\nu_\pi = \left[\binom{2n_1}{n_1} \right]^{-1} \gg \tilde{\nu}_n. \qquad \square$$

Remark 3.6 To decrease the computational cost, we can modify the algorithm which determines the permutation matrix $\hat{\mathbf{P}}$. It can work with the block row of $\hat{\mathbf{Z}}$ which has fewer rows. If in the proof of Lemma 3.4 we have $n_j < n_i$, the algorithm can work with the block-row $[Z_{ji}, Z_{jj}]$. Then it reduces Z_{jj} to a lower-triangular matrix with positive diagonal elements. Instead of the relation (3.16) we shall have

$$\begin{bmatrix} -1 & \\ & I_{n_i-1} \end{bmatrix} \cdots \begin{bmatrix} H_{n_i-1} & \\ & 1 \end{bmatrix} H_{n_i} \begin{bmatrix} Z_{ji} & Z_{jj} \end{bmatrix} I_{n_i+n_j,1'} \cdots I_{n_i+1,n_j'} = \begin{bmatrix} L_{ji} & L_{jj} \end{bmatrix},$$

where L_{jj} is lower-triangular with nondecreasing diagonal elements. In the first step the algorithm produces zeros in the last column of $[Z_{ji}, Z_{jj}]$ and then it proceeds in an obvious way. We shall generally have $r' \in [1, n_i + n_j - r + 1]$, $r = 1, 2, \ldots, n_j$. Although the lower bound for $\sigma_{min}(L_{jj})$ will be the same as earlier, the gain is in smaller computational cost.

As a final note, instead of reducing Z_{jj} to the lower-triangular matrix, the algorithm can reduce it to the upper-triangular matrix. The obtained bound is the same. $\square$

Now we can prove the main result of the paper.

Theorem 3.7 *The block Jacobi method converges to diagonal form under the class of generalized serial pivot strategies provided the condition*

$$\liminf_{k\to\infty} \sigma_{min}(Z^{(k)}_{i(k)i(k)}) > 0 \tag{3.19}$$

holds. The block method is globally convergent if the kernel algorithm is any convergent HZ or CJ method.

Proof Let $\mathbf{A}$ and $\mathbf{B}$ be symmetric matrices such that $\mathbf{B}$ is positive definite. Let us apply the block Jacobi method to $(\mathbf{A}, \mathbf{B})$ under some generalized serial strategy, thus obtaining the sequence of matrix pairs $\left((\mathbf{A}^{(k)}, \mathbf{B}^{(k)}), k \ge 0 \right)$. To prove

$$\lim_{k\to\infty} S(\mathbf{A}^{(k)}, \mathbf{B}^{(k)}) = 0 \tag{3.20}$$

we apply Theorem 3.2 to the sequences $(\mathbf{A}^{(k)}, k \ge 0)$, $(\mathbf{B}^{(k)}, k \ge 0)$ and show

$$\lim_{k\to\infty} S(\mathbf{A}^{(k)}) = 0, \qquad \lim_{k\to\infty} S(\mathbf{B}^{(k)}) = 0. \tag{3.21}$$

The first condition (**A1**) of Theorem 3.2 is just selection of the pivot strategy. Hence it is satisfied for both sequences $(\mathbf{A}^{(k)}, k \ge 0)$ and $(\mathbf{B}^{(k)}, k \ge 0)$.

From Lemma 3.3(iii) we see that condition (**A2**) of Theorem 3.2 is satisfied for both sequences.

The condition (**A3**) of Theorem 3.2 is satisfied for the sequences $(\mathbf{A}^{(k)}, k \ge 0)$ and $(\mathbf{B}^{(k)}, k \ge 0)$ because it is just the assumption (3.19) of the theorem.

From Lemma 3.3(i) we see that both sequences $(\mathbf{A}^{(k)}, k \ge 0)$ and $(\mathbf{B}^{(k)}, k \ge 0)$ are bounded, so the condition (**A4**) of Theorem 3.2 holds for these sequences of matrices.

Finally, from relations (2.4) and (2.5) we see that

$$S\left([\hat{\mathbf{Z}}^{(k)}]^T\hat{\mathbf{A}}^{(k)}\hat{\mathbf{Z}}^{(k)}\right)=0,\qquad [\hat{\mathbf{Z}}^{(k)}]^T\hat{\mathbf{B}}^{(k)}\hat{\mathbf{Z}}^{(k)}=\mathbf{I}_{n_{i(k)}n_{j(k)}},\qquad k\geq 0.$$

Hence the condition (**i**) of Theorem 3.2 is satisfied for both sequences.

We conclude that the condition (**ii**) of Theorem 3.2 holds for both sequences, which means that the relation (3.21) and consequently the relation (3.20) holds.

To prove the second assertion of the theorem, we have to clarify how the block method operates on matrices $\mathbf{A}^{(k)}=(A_{rt}^{(k)})$, $\mathbf{B}^{(k)}=(B_{rt}^{(k)})$ in step k. First, the kernel algorithm computes matrices $\Lambda_{i(k)i(k)}^{(k)}$, $\Lambda_{j(k)j(k)}^{(k)}$ and $\hat{\mathbf{Z}}^{(k)}$. Then the block method updates the off-block-diagonal blocks of $\mathbf{A}^{(k)}$: $A_{ri(k)}^{(k)}$, $A_{rj(k)}^{(k)}$, $A_{i(k)r}^{(k)}$, $A_{j(k)r}^{(k)}$, $r\notin\{i(k),j(k)\}$, and similar for $\mathbf{B}^{(k)}$.

In [27] the global convergence of the HZ and CJ methods has been proved under the class of generalized serial strategies. As part of the proof, it has been proved that for small enough $S(\hat{\mathbf{A}}^{(k)},\hat{\mathbf{B}}^{(k)})$ the diagonal elements of $\hat{\mathbf{A}}^{(k)}$ cannot change their affiliation with the eigenvalues of the initial pair. In another words, in the final stage of the block process, the kernel algorithm cannot change affiliation of the diagonal elements.

This fact together with $\lim_{k\to\infty}S(\mathbf{A}^{(k)},\mathbf{B}^{(k)})=0$ means that the block method is globally convergent. □

We have shown that any block Jacobi method can be modified to satisfy the condition (3.19) (see the derivation of the relation (3.18)). Hence it can be made convergent to diagonal form or globally convergent (depending on the kernel algorithm), under an arbitrary generalized serial strategy.

If we require that the kernel algorithm orders the diagonal elements of diag $(\Lambda_{ii}^{(k)},\Lambda_{jj}^{(k)})$ from the relation (2.4) in some fixed monotone ordering, then one can use ideas from the proof of [26, Theorem 2.10] to try to prove the global convergence. However, this is of less importance. In practice it is irrelevant in which order the eigenpairs of $(\mathbf{A},\mathbf{B})$ are collected after the process has been terminated.

4 Conclusions and Future Work

In this paper, we have proved the global convergence of the block Jacobi method for the PGEP under the class of generalized serial pivot strategies from [26]. This is the first global convergence proof of some block Jacobi method for the generalized eigenvalue problem. Because of its efficiency [32] this block method is the method of choice for solving the PGEP with large matrices **A** and **B**.

Future work, which is underway, will prove the global convergence of the complex PGEP block Jacobi method. For this research, one has to generalize many results from [26] to hold for the complex block Jacobi method. Since both real and complex block Jacobi methods compute accurately the eigenvalues of the corresponding PGEP, it is an open problem to prove the high relative accuracy of those methods in the case of positive definite matrices **A** and **B**. Finally, an open problem is to prove the asymptotic quadratic convergence of those block methods.

References

1. Bečka, M., Okša, G., Vajtersic, M.: Parallel Block-Jacobi SVD Methods. In: Berry M. et al. (eds) High-Performance Scientific Computing. Springer, London 2012: 185-197.

2. Bečka, M., Okša, G., Vajtersic, M.: New Dynamic Orderings for the Parallel One-Sided Block-Jacobi SVD Algorithm. Parallel Processing Letters 25(2) (2015).
3. Bečka, M., Okša, G., Vajtersic, M.: Asymptotic quadratic convergence of the serial block-Jacobi EVD algorithm for Hermitian matrices. Numer. Math. 136(4), 1071-1095 (2017).
4. Bujanović, Z., Drmač, Z: A contribution to the theory and practice of the block Kogbetliantz method for computing the SVD. BIT 52(4), 827-849 (2012).
5. Demmel, W.J., Veselić, K.: Jacobi's method is more accurate than QR. SIAM J. Mat. Anal. Appl. 13, 1204–1245 (1992).
6. Demmel, W.J., Gu, M., Eisenstat, S., Slapničar, I., Veselić, K., Drmač, Z.: Computing the singular value decomposition with high relative accuracy. Linear Algebra Appl. 299, 21–80 (1999).
7. deRijk, P.P.M.: A one-sided Jacobi algorithm for computing the singular-value decomposition on a vector computer. SIAM J. Sci. Stat. Comp. 10(2), 359-371 (1989).
8. Dopico, F.M., Molera, J.M., Moro J.: An orthogonal high relative accuracy algorithm for the symmetric eigenproblem. SIAM J. Mat. Anal. Appl. 25(2), 301-351 (2003).
9. Drmač, Z.: Implementation of Jacobi rotations for accurate singular value computation in floating-point arithmetic. SIAM J. Sci. Stat. Comp. 18(4), 1200-1222 (1997)
10. Drmač, Z.: A posteriori computation of the singular vectors in a preconditioned Jacobi SVD algorithm. IMA J. Numer. Anal. 19, 191-213 (1999).
11. Drmač, Z., Veselić, K.: New fast and accurate Jacobi SVD algorithm I. SIAM J. Mat. Anal. Appl. 29(4), 1322-1342 (2007).
12. Drmač, Z., Veselić, K.: New fast and accurate Jacobi SVD algorithm II. SIAM J. Mat. Anal. Appl. 29(4), 1343-1362 (2007).
13. Drmač, Z.: A global convergence proof for cyclic Jacobi methods with block rotations. SIAM journal on matrix analysis and applications. 31(3), 1329-1350 (2009).
14. Eberlein, P.J.: A Jacobi method for automatic computation of eigenvalues and eigenvectors of an arbitrary matrix. Siam J. 10, 74-88 (1962).
15. Eberlein, P.J., Boothroyd, J.: Solution to the eigenproblem by a norm-reducing Jacobi-type method. Numer. Math. 11, 1-12 (1968).
16. Eberlein, P.J., Park, H.: Efficient implementation of Jacobi algorithms and Jacobi sets on distributed memory architectures. Journal of Parallel and Distributed Computing, Special issue on Algorithms for Hypercube Computers 8, 358-366 (1990).
17. Hansen, E.R.: On Cyclic Jacobi Methods. SIAM J. Appl. Math. 11, 449-459 (1963).
18. Hari, V., Veselić, K.: On Jacobi methods for singular value decompositions. SIAM J. Sci. Stat. Comput. 8(5), 741-754 (1987).
19. Hari, V.: Accelerating the SVD Block-Jacobi Method. Computing 75, 27-53 (2005).
20. Hari, V.: Convergence of a block-oriented quasi-cyclic Jacobi method. SIAM J. Mat. Anal. Appl. 29(2), 349-369 (2007).
21. Hari, V., Zadelj-Martić, V.: Parallelizing the Kogbetliantz Method: A First Attempt. JNAIAM 2(1-2), 49-66 (2007).
22. Hari, V.: On Block Jacobi Annihilators. Proceedings of Algoritmy 2009, 1-13 (2009).
23. Hari, V., Singer, S., Singer, S.: Block-oriented *J*-Jacobi Methods for Hermitian Matrices. Lin. Alg. and Its Appl. 433 (8-10), 1491-1512 (2010).
24. Hari, V., Singer, S., Singer, S.: Full Block *J*-Jacobi Method for Hermitian Matrices. Lin. Alg. and Its Appl. 444, 1-27 (2014).
25. Hari, V.: Convergence to Diagonal Form of Block Jacobi-type Methods. Numer. Math. 129(3), 449-481 (2015).
26. Hari, V., Begović Kovač, E.: Convergence of the Cyclic and Quasi-cyclic Block Jacobi Methods, Electron. Trans. Numer. Anal. 46, 107-147 (2017).
27. Hari, V.: Globally convergent Jacobi methods for positive definite matrix pairs. Numer. Algor. 79(1), 221249 (2018). DOI: 10.1007/s11075-017-0435-5
28. Luk F., Park, H.: On the equivalence and convergence of parallel Jacobi SVD algorithms. IEEE Tran. Comp. 38(6), 806-811 (1989).
29. Luk F., Park, H.: On parallel Jacobi orderings. SIAM J. Sci. Stat. Comput. 10, 18–26 (1989).
30. Matejaš, J., Hari, V.: On High Relative Accuracy of the Kogbetliantz Method. Lin. Alg. and Its Appl. 464, 100-129 (2015).
31. Matejaš, J.: Accuracy of the Jacobi Method on Scaled Diagonally Dominant Symmetric Matrices. SIAM J. Matrix Anal. Appl. 31(1), 133-153 (2009).
32. Novaković, J., Singer, S., Singer, S.: Blocking and Parallelization of the HariZimmermann Variant of the FalkLangemeyer Algorithm for the Generalized SVD. Parallel Comput., 49, 136-152 (2015).
33. Sameh, A.H.: On Jacobi and Jacobi-like algorithms for parallel computer. Math. Comp. 25, 579-590 (1971).

34. Shroff, G., Schreiber, R.: On the convergence of the cyclic Jacobi method for parallel block orderings. SIAM J. Matrix Anal. Appl. 10, 326–346 (1989).
35. Slapničar, I.: Accurate Symmetric Eigenreduction by a Jacobi Method. Ph.D. thesis, University of Hagen, 1992.
36. Slapničar, I.: Highly Accurate Symmetric Eigenvalue Decomposition and Hyperbolic SVD. Linear Algebra Appl. 358, 387-424 (2002).
37. van der Sluis, A.: Condition numbers and equilibration of matrices. Numer. Math. 14(1), 14-23 (1969).