

## CONVERGENCE OF THE CYCLIC AND QUASI-CYCLIC BLOCK JACOBI METHODS\*

VJERAN HARI<sup>†</sup> AND ERNA BEGOVIĆ KOVAČIĆ<sup>‡</sup>

**Abstract.** This paper studies the global convergence of the block Jacobi method for symmetric matrices. Given a symmetric matrix  $A$  of order  $n$ , the method generates a sequence of matrices by the rule  $A^{(k+1)} = U_k^T A^{(k)} U_k$ ,  $k \geq 0$ , where  $U_k$  are orthogonal elementary block matrices. A class of generalized serial pivot strategies is introduced, significantly enlarging the known class of weak wavefront strategies, and appropriate global convergence proofs are obtained. The results are phrased in the stronger form:  $S(A') \leq cS(A)$ , where  $A'$  is the matrix obtained from  $A$  after one full cycle,  $c < 1$  is a constant, and  $S(A)$  is the off-norm of  $A$ . Hence, using the theory of block Jacobi operators, one can apply the obtained results to prove convergence of block Jacobi methods for other eigenvalue problems such as the generalized eigenvalue problem. As an example, the results are applied to the block  $J$ -Jacobi method. Finally, all results are extended to the corresponding quasi-cyclic strategies.

**Key words.** eigenvalues, block Jacobi method, pivot strategies, global convergence

**AMS subject classifications.** 65F15

**1. Introduction.** The main incentive for writing this paper was a need to expand the class of “convergent strategies” for the block Jacobi method for symmetric matrices. With a large choice of classes at our disposal we can prove global convergence of other block-wise or element-wise Jacobi-type methods and even apply it to related problems, e.g., to the generalized eigenvalue or singular value problem; see [29]. The techniques we are about to employ use the theory of block Jacobi operators, which was described in [18].

Over the last two decades the Jacobi method has emerged as a method of choice for the eigenvalue computation for dense symmetric matrices. This is mostly due to its inherent parallelism and high relative accuracy on well-behaved matrices. Although the original method is very old [23] and it had been one of the first methods to be implemented on computers, it was forgotten in the 1970s after appearance of the QR and Divide and Conquer method. Already in 1971, Sameh [31] showed how to adapt the serial Jacobi method to parallel processing. Later, in 1992, Demmel and Veselić [4] proved high relative accuracy of the method on well-behaved symmetric positive definite matrices. Following their breakthrough, the method came back to the focus of the current researchers. Drmač and Veselić [7, 8] showed that, even on standard one-processor computers, the method can be modified to become faster than the QR method while still retaining its distinguished property: high relative accuracy. Nowadays, the Jacobi method is well understood. On the one hand, its asymptotic convergence was considered in [15, 30, 35], and its global convergence was studied in [9, 10, 11, 16, 21, 22, 24, 25, 28, 32]. On the other hand, its high relative accuracy was considered in [4, 26, 27, 34], while its efficiency was investigated in [7, 8]. The method has also been implemented as a standard LAPACK routine.

With the development of CPU and GPU parallel computing platforms, it has been found that a sensible way of increasing numerical efficiency of the method involves using a one-sided algorithm, together with BLAS 3 subroutines, which can nicely exploit cache memory hierarchy. The matrix description of such a method is called the block Jacobi method. This block method is always implemented as a one-sided block (Jacobi or  $J$ -Jacobi) algorithm

---

\*Received September 22, 2016. Accepted February 21, 2017. Published online on May 12, 2017. Recommended by M. Hochstenbach. This work has been fully supported by Croatian Science Foundation under the project IP-09-2014-3670.

<sup>†</sup>Department of Mathematics, University of Zagreb, Bijenička 30, 10000 Zagreb, Croatia (hari@math.hr).

<sup>‡</sup>Faculty of Chemical Engineering and Technology, University of Zagreb, Marulićev trg 19, 10000 Zagreb, Croatia (ebegovic@fkit.hr).

because high efficiency and high relative accuracy are warranted then. However, in the global and asymptotic convergence considerations, the results are cast in terms of a two-sided block Jacobi method. The first global convergence results for the block Jacobi methods were given in [3, 5, 18, 19, 20]. These papers considered the most common serial pivot strategies and strategies equivalent to them.

The aim of this paper is to further develop the global convergence theory for the block Jacobi method and to provide a large class of usable pivot strategies for which convergence can be established rigorously. In general, our class consists of more than  $4 \cdot 2! \cdot 3! \cdots m!$  cyclic strategies, where  $m$  is the number of block-columns in the block-matrix partition of a symmetric matrix of order  $n$ . These strategies include the weak wavefront ones from [32] and many others. As a byproduct of this research, we can now prove that every cyclic (element-wise or block) Jacobi method for symmetric matrices of order 4 is globally convergent; see [1, 2]. In addition, we consider a similar class of quasi-cyclic strategies and derive the corresponding convergence results. The block analogue of the strategy that is used in the LAPACK implementation of the Jacobi method lies in that class.

The convergence results are given in the “stronger form”,

$$S(A') \leq cS(A), \quad 0 \leq c < 1.$$

Here,  $A$  is the initial symmetric matrix of order  $n$ ,  $A'$  is obtained from  $A$  after applying one sweep of some cyclic or quasi-cyclic block Jacobi method,  $S(\cdot)$  is the departure from diagonal form, and  $c$  is a constant depending on  $n$  and the block-matrix partition but not on  $A$ . Such a result allows for the use of the theory of block Jacobi operators. Hence, it can be utilized to prove the global convergence of other Jacobi-type methods, designed for different eigenvalue problems. As an application, we will apply it to the block J-Jacobi method from [20]. Some of the results presented here can be found in the unpublished thesis [1].

The paper is divided into six sections of the main text and an appendix. In Section 2 we present the basic concepts linked to a block Jacobi method for symmetric matrices. Special attention is paid to cyclic and quasi-cyclic pivot strategies and to the ways of enlarging significantly the number of “convergent strategies”. The concepts of equivalent, weakly equivalent, and permutation equivalent strategies are used. Another useful tool is introduced, the so-called block Jacobi annihilators and operators for symmetric matrices, and some basic results related to them are proved. In Section 3 we introduce a class of generalized serial strategies and prove the corresponding global convergence results. In Section 4 we briefly introduce a similar class of quasi-cyclic pivot strategies and prove the appropriate convergence results. As an application, in Section 5 we prove the global convergence of the block J-Jacobi method under the strategies from the newly introduced classes. Section 6 announces future work. Finally, to make the paper easier to read, we move all lengthy and technical proofs to Appendix A.

**2. Basic concepts and notation.** We introduce the basic definitions linked to the block Jacobi method for symmetric matrices. Special attention is paid to the cyclic and quasi-cyclic pivot strategies. Later we deal with more advanced concepts like the block Jacobi annihilators and operators.

**2.1. Block Jacobi method.** Let  $A$  be a square matrix of order  $n$  and let  $\pi$  be an integer partition of  $n$ ,

$$(2.1) \quad \pi = (n_1, n_2, \dots, n_m), \quad n_i \geq 1, \quad 1 \leq i \leq m, \quad n_1 + n_2 + \dots + n_m = n.$$

Then  $\pi$  determines the block-matrix partition of  $A$ ,

$$(2.2) \quad A = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1m} \\ A_{21} & A_{22} & & A_{2m} \\ \vdots & & \ddots & \vdots \\ A_{m1} & A_{m2} & \dots & A_{mm} \end{bmatrix} \begin{matrix} n_1 \\ n_2 \\ \vdots \\ n_m \end{matrix},$$

where the diagonal blocks  $A_{11}, \dots, A_{mm}$  are square matrices of order  $n_1, \dots, n_m$ , respectively. Relation (2.2) will be schematically denoted by  $A = (A_{rs})$ .

Since we consider the global convergence of the block Jacobi method for symmetric matrices, we assume that  $A$  is symmetric. A block Jacobi method is determined by the partition  $\pi$ , some pivot strategy, and the algorithm. The partition is chosen in accordance with the capacity of the hierarchical cache memory of the computer. Typically, the code presumes  $n_1 = n_2 = \dots = n_{m-1}, n_m = n - (m-1)n_1 \leq n_1$ . In our analysis we consider it arbitrary but unchanged over the iterations. Actually, there are situations when it makes sense to change  $\pi$  during the process, but these are linked to the asymptotic convergence of the method.

The block Jacobi method uses orthogonal *elementary block matrices* as transformation matrices. An orthogonal elementary block matrix  $\mathbf{U}_{ij}$  has the form (see [18])

$$(2.3) \quad \mathbf{U}_{ij} = \begin{cases} \begin{bmatrix} I & & & \\ & U_{ii} & U_{ij} & \\ & U_{ji} & U_{jj} & \\ & & & I \end{bmatrix} \begin{matrix} n_i \\ n_j \end{matrix} & \text{if } i < j, \text{ or} \\ \begin{bmatrix} I & & \\ & U_{ii} & \\ & & I \end{bmatrix} \begin{matrix} n_i \end{matrix} & \text{if } i = j, \end{cases}$$

where it is presumed that the block-matrix partition is determined by  $\pi$  from the relation (2.1). Since  $i$  and  $j$  address the blocks, they can be called *block pivot indices*, but for brevity we simply call them *pivot indices*. Similarly,  $(i, j)$  is the *pivot pair* and

$$(2.4) \quad \widehat{U}_{ij} = \begin{cases} \begin{bmatrix} U_{ii} & U_{ij} \\ U_{ji} & U_{jj} \end{bmatrix} & \text{if } i < j, \\ U_{ii} & \text{if } i = j, \end{cases} \quad \text{or} \quad \widehat{U}_{ij} = U_{ii} \quad \text{if } i = j,$$

is the *pivot submatrix* of  $\mathbf{U}_{ij}$ . When the indices  $(i, j)$  are clear from the context, we will also write  $\widehat{U}$  instead of  $\widehat{U}_{ij}$ . We can build an orthogonal elementary block matrix using the function  $\mathcal{E}$  which imbeds any orthogonal matrix  $\widetilde{U}$  of order  $n_i + n_j$  (or  $n_i$  if  $i = j$ ) into the identity matrix  $I_n$ , so that  $\mathbf{U}_{ij} = \mathcal{E}(i, j, \widetilde{U})$  implies  $\widehat{U}_{ij} = \widetilde{U}$ . The mapping  $\mathcal{E}$  depends on the partition  $\pi$ .

Each *block Jacobi method* is an iterative processes of the form

$$(2.5) \quad A^{(k+1)} = U_k^T A^{(k)} U_k, \quad k \geq 0, \quad A^{(0)} = A,$$

where  $U_k, k \geq 0$ , are orthogonal elementary block matrices. Let  $A^{(k)} = (A_{rs}^{(k)})$ . We say that  $A^{(k+1)}$  is obtained or generated from  $A^{(k)}$  at *step*  $k$  via the recursion (2.5). Let  $U_k = \mathcal{E}(i(k), j(k), \widehat{U}_k)$ . Then  $i(k), j(k)$  are the pivot indices and  $(i(k), j(k))$  is the pivot pair at step  $k$ . For brevity, we will often omit  $k$  and denote the pivot indices simply by  $i, j$  and the pivot pair by  $(i, j)$ . The way of selecting the pivot pair at each step is called a *pivot strategy*.

At step  $k$  the block Jacobi method diagonalizes the pivot submatrix of  $A^{(k)}$ . Thus, if  $i < j$ , the pivot blocks  $A_{ij}^{(k)}$  and  $A_{ji}^{(k)}$  are annihilated and the affected diagonal blocks  $A_{ii}^{(k)}$  and  $A_{jj}^{(k)}$  are diagonalized. If  $\widehat{A}^{(k)}$  denotes the pivot submatrix of order  $n_i + n_j$  at step  $k$ , it is transformed as follows:

$$(2.6) \quad \begin{bmatrix} \Lambda_{ii}^{(k+1)} & 0 \\ 0 & \Lambda_{jj}^{(k+1)} \end{bmatrix} = \begin{bmatrix} U_{ii}^{(k)} & U_{ij}^{(k)} \\ U_{ji}^{(k)} & U_{jj}^{(k)} \end{bmatrix}^T \begin{bmatrix} A_{ii}^{(k)} & A_{ij}^{(k)} \\ (A_{ij}^{(k)})^T & A_{jj}^{(k)} \end{bmatrix} \begin{bmatrix} U_{ii}^{(k)} & U_{ij}^{(k)} \\ U_{ji}^{(k)} & U_{jj}^{(k)} \end{bmatrix},$$

where  $\Lambda_{ii}^{(k+1)}$  and  $\Lambda_{jj}^{(k+1)}$  are diagonal. If  $i = j$ , then just  $A_{ii}^{(k)}$  is diagonalized. For the diagonalization of the pivot submatrix, one can choose any method for solving the symmetric eigenvalue problem. Typically, one applies a standard (element-wise) Jacobi method for its high relative accuracy [26] and efficiency on nearly diagonal matrices.

As has been explained in [18], it is preferable to preprocess the initial matrix by  $m$  block Jacobi steps with pivot pairs  $(1, 1), \dots, (m, m)$  so that in the starting matrix the diagonal blocks are actually diagonal submatrices. This preprocessing is depicted below for the case  $\pi = (3, 2, 1, 2)$ :

$$A = \begin{array}{|c|c|c|c|} \hline x & x & x & x \\ \hline x & x & x & x \\ \hline x & x & x & x \\ \hline x & x & x & x \\ \hline x & x & x & x \\ \hline x & x & x & x \\ \hline x & x & x & x \\ \hline \end{array} \quad \mapsto \quad A^{(0)} = \begin{array}{|c|c|c|c|} \hline x & 0 & 0 & x \\ \hline 0 & x & 0 & x \\ \hline 0 & 0 & x & x \\ \hline x & x & x & x \\ \hline x & x & x & 0 \\ \hline x & x & x & 0 \\ \hline x & x & x & x \\ \hline x & x & x & x \\ \hline x & x & x & 0 \\ \hline x & x & x & 0 \\ \hline \end{array}.$$

Once the diagonal blocks are diagonalized, all later steps will preserve that property. It means that at each step the pivot indices will satisfy  $i < j$ , which unifies and simplifies the algorithm. In this regard the blocks  $A_{ii}$  and  $A_{jj}$  on the right side of the relation (2.6) can be replaced by  $\Lambda_{ii}$  and  $\Lambda_{jj}$ , respectively. Therefore, in the sequel it is presumed that the diagonal blocks of each  $A^{(k)}$  are diagonal and for the pivot indices  $i < j$  holds.

**2.2. Pivot strategies.** Each pivot strategy can be identified with a function  $I : \mathbb{N}_0 \rightarrow \mathcal{P}_m$ , where  $\mathbb{N}_0 = \{0, 1, 2, 3, \dots\}$  and  $\mathcal{P}_m = \{(r, s) | 1 \leq r < s \leq m\}$ . If  $I$  is a periodic function with period  $T$ , then we say that  $I$  is a *periodic pivot strategy*. In this paper we consider two types of periodic strategies: cyclic and quasi-cyclic ones.

If  $T = M \equiv \frac{m(m-1)}{2}$  and  $\{(i(0), j(0)), (i(1), j(1)), \dots, (i(T-1), j(T-1))\} = \mathcal{P}_m$ , then we say that the pivot strategy is *cyclic*. It immediately follows that, during any  $M$  successive steps of the method, all off-diagonal blocks are annihilated exactly once. Such block Jacobi method is also said to be cyclic and the transition from  $A^{((r-1)M)}$  to  $A^{(rM)}$  is called the  $r$ th *cycle* or *sweep* of the method.

If  $T \geq M$  and  $\{(i(0), j(0)), (i(1), j(1)), \dots, (i(T-1), j(T-1))\} = \mathcal{P}_m$ , then the strategy is called *quasi-cyclic*. Thus, during any  $T$  successive steps of the method, each off-diagonal block is annihilated at least once. The corresponding block Jacobi method is called *quasi-cyclic* and the transition from  $A^{((r-1)T)}$  to  $A^{(rT)}$  is called the  $r$ th *quasi-cycle* or *sweep* of the method.

Let us examine cyclic and quasi-cyclic strategies more closely. For  $\mathcal{S} \subseteq \mathcal{P}_m$  we denote by  $\mathcal{O}(\mathcal{S})$  the set of all finite sequences containing the elements of  $\mathcal{S}$ , assuming that each pair from  $\mathcal{S}$  appears at least once in each sequence. If  $I$  is a cyclic or quasi-cyclic strategy with period  $T$ , then  $\mathcal{O}_I$  stands for the sequence  $I(0), I(1), \dots, I(T-1) \in \mathcal{O}(\mathcal{P}_m)$ , generated by the first  $T$  steps (i.e., by the first *sweep*) of the method. Conversely, if

$\mathcal{O} \in \mathcal{O}(\mathcal{P}_m)$ ,  $\mathcal{O} = (i_0, j_0), (i_1, j_1), \dots, (i_{T-1}, j_{T-1})$ , then the periodic strategy  $I_{\mathcal{O}}$  is defined by  $I_{\mathcal{O}}(k) = (i_{\tau(k)}, j_{\tau(k)})$ , where  $\tau(k)$  is the unique integer satisfying  $0 \leq \tau(k) \leq T-1$  and  $k \equiv \tau(k) \pmod{T}$ ,  $k \geq 0$ .

These two functions,  $\mathcal{O} \mapsto I_{\mathcal{O}}$  and  $I \mapsto \mathcal{O}_I$ , enable us to investigate the cyclic and quasi-cyclic strategies by studying the sequences from  $\mathcal{O}(\mathcal{P}_m)$ . Note that, if  $I$  is cyclic, then  $\mathcal{O}_I$  is simply an ordering of  $\mathcal{P}_m$ . We will also use the term *pivot ordering* in this case, while if  $I$  is quasi-cyclic, we will use the term *pivot sequence*.

An *admissible transposition* of  $\mathcal{O} \in \mathcal{O}(\mathcal{S})$ ,  $\mathcal{S} \subseteq \mathcal{P}_m$ , is any transposition of two adjacent terms in  $\mathcal{O}$ ,

$$(i_r, j_r), (i_{r+1}, j_{r+1}) \rightarrow (i_{r+1}, j_{r+1}), (i_r, j_r),$$

provided that the sets  $\{i_r, j_r\}$  and  $\{i_{r+1}, j_{r+1}\}$  are disjoint. We also say that such pairs  $(i_r, j_r)$  and  $(i_{r+1}, j_{r+1})$  *commute*. The number of pairs in  $\mathcal{O}$  is denoted by  $|\mathcal{O}|$  and it is called the *length* of  $\mathcal{O}$ .

DEFINITION 2.1. Two sequences  $\mathcal{O}, \mathcal{O}' \in \mathcal{O}(\mathcal{S})$ ,  $\mathcal{S} \subseteq \mathcal{P}_m$ , are said to be

- (i) equivalent (we write  $\mathcal{O} \sim \mathcal{O}'$ ) if one can be obtained from the other by a finite set of admissible transpositions;
- (ii) shift-equivalent ( $\mathcal{O} \stackrel{s}{\sim} \mathcal{O}'$ ) if  $\mathcal{O} = [\mathcal{O}_1, \mathcal{O}_2]$  and  $\mathcal{O}' = [\mathcal{O}_2, \mathcal{O}_1]$ , where  $[ \ , \ ]$  stands for concatenation and the length of  $\mathcal{O}_1$  is called *shift length*;
- (iii) weakly equivalent ( $\mathcal{O} \stackrel{w}{\sim} \mathcal{O}'$ ) if there exist  $\mathcal{O}_i \in \mathcal{O}(\mathcal{S})$ ,  $0 \leq i \leq r$ , such that every two adjacent terms in the sequence  $\mathcal{O} = \mathcal{O}_0, \mathcal{O}_1, \dots, \mathcal{O}_r = \mathcal{O}'$  are equivalent or shift-equivalent.

One can verify that  $\sim$ ,  $\stackrel{s}{\sim}$ , and  $\stackrel{w}{\sim}$  are equivalence relations on  $\mathcal{O}(\mathcal{S})$ . If three or more sequences are connected by  $\sim$  or  $\stackrel{s}{\sim}$  one can omit the mid terms because of the transitivity property of equivalence relation. Hence, if  $\mathcal{O} \stackrel{w}{\sim} \mathcal{O}'$ , then there is a sequence  $\mathcal{O} = \mathcal{O}_0, \mathcal{O}_1, \dots, \mathcal{O}_r = \mathcal{O}'$  such that

$$(2.7) \quad \text{either } \mathcal{O}_0 \sim \mathcal{O}_1 \stackrel{s}{\sim} \mathcal{O}_2 \sim \mathcal{O}_3 \stackrel{s}{\sim} \mathcal{O}_4 \dots \quad \text{or } \mathcal{O}_0 \stackrel{s}{\sim} \mathcal{O}_1 \sim \mathcal{O}_2 \stackrel{s}{\sim} \mathcal{O}_3 \sim \mathcal{O}_4 \dots$$

Two pivot strategies  $I_{\mathcal{O}}$  and  $I_{\mathcal{O}'}$  are equivalent (resp. shift-equivalent, weakly equivalent) if the corresponding sequences  $\mathcal{O}$  and  $\mathcal{O}'$  are equivalent (resp. shift-equivalent, weakly equivalent).

The most common cyclic pivot strategies are the row-cyclic one,  $I_{\text{row}} = I_{\mathcal{O}_{\text{row}}}$ , and the column-cyclic one,  $I_{\text{col}} = I_{\mathcal{O}_{\text{col}}}$ , which are defined by the “row-wise” and “column-wise” orderings of  $\mathcal{P}_m$ :

$$\begin{aligned} \mathcal{O}_{\text{row}} &= (1, 2), (1, 3), \dots, (1, m), (2, 3), \dots, (2, m), \dots, (m-1, m) \quad \text{and} \\ \mathcal{O}_{\text{col}} &= (1, 2), (1, 3), (2, 3), \dots, (1, m), (2, m), \dots, \dots, (m-1, m). \end{aligned}$$

The common name for them is *serial strategies*. The cyclic pivot strategies which are equivalent (resp. weakly equivalent) to the serial ones are also called *wavefront* (resp. *weakly wavefront*) strategies; see [32].

DEFINITION 2.2. Let  $\mathcal{O} \in \mathcal{O}(\mathcal{S})$ ,  $\mathcal{S} \subseteq \mathcal{P}_m$ ,  $\mathcal{O} = (i_0, j_0), (i_1, j_1), \dots, (i_r, j_r)$ . Then

$$\mathcal{O}^{\leftarrow} = (i_r, j_r), \dots, (i_1, j_1), (i_0, j_0) \in \mathcal{O}(\mathcal{S})$$

is the reverse (or inverse) sequence to  $\mathcal{O}$ . If  $\mathcal{S} = \mathcal{P}_m$ , we say that the pivot strategy  $I_{\mathcal{O}^{\leftarrow}} = I_{\mathcal{O}^{\leftarrow}}$  is reverse (inverse) to  $I_{\mathcal{O}}$ .

Obviously, we have  $\mathcal{O}^{\leftarrow\leftarrow} = \mathcal{O}$  and hence  $I_{\mathcal{O}^{\leftarrow\leftarrow}} = I_{\mathcal{O}}$  for  $\mathcal{O} \in \mathcal{O}(\mathcal{P}_m)$ .

LEMMA 2.3. Let  $\mathcal{O}, \mathcal{O}' \in \mathcal{O}(\mathcal{S})$ ,  $\mathcal{S} \subseteq \mathcal{P}_m$ . Then  $\mathcal{O}' \stackrel{w}{\sim} \mathcal{O}$  if and only if  $\mathcal{O}'^{\leftarrow} \stackrel{w}{\sim} \mathcal{O}^{\leftarrow}$ .

*Proof.* From formula (2.7), we see that it is sufficient to prove the assertion for the relations  $\sim$  and  $\stackrel{s}{\sim}$ . Let  $\mathcal{O}$  and  $\mathcal{O}^{\leftarrow}$  be as in Definition 2.2. Suppose  $\mathcal{O}'$  is obtained from  $\mathcal{O}$  by applying one admissible transposition. Then for some  $0 \leq t < r$  we have  $\{i_t, j_t\} \cap \{i_{t+1}, j_{t+1}\} = \emptyset$ . If  $t \geq 1$ , then

$$\begin{aligned} \mathcal{O}' &= (i_0, j_0), \dots, (i_{t-1}, j_{t-1}), (i_{t+1}, j_{t+1}), (i_t, j_t), \dots, (i_r, j_r), \\ \mathcal{O}'^{\leftarrow} &= (i_r, j_r), \dots, (i_t, j_t), (i_{t+1}, j_{t+1}), (i_{t-1}, j_{t-1}), \dots, (i_0, j_0), \end{aligned}$$

and obviously  $\mathcal{O}'^{\leftarrow} \sim \mathcal{O}^{\leftarrow}$ . If  $t = 0$ ,  $\mathcal{O}' = (i_1, j_1), (i_0, j_0), \dots, (i_r, j_r)$  and  $\mathcal{O}'^{\leftarrow} = (i_r, j_r), \dots, (i_0, j_0), (i_1, j_1)$ , so we also have  $\mathcal{O}'^{\leftarrow} \sim \mathcal{O}^{\leftarrow}$ . If  $\mathcal{O}'$  is the result of applying more than one admissible transposition to  $\mathcal{O}$ , then the proof proceeds by applying the above argument several times.

Suppose that  $\mathcal{O}'^{\leftarrow} \sim \mathcal{O}^{\leftarrow}$  holds. By the implication proved in the preceding paragraph we have  $\mathcal{O}'^{\leftarrow\leftarrow} \sim \mathcal{O}^{\leftarrow\leftarrow}$ , and this is the same as  $\mathcal{O}' \sim \mathcal{O}$ .

Now, let  $\mathcal{O}' \stackrel{s}{\sim} \mathcal{O}$ . Suppose  $\mathcal{O}' = (i_{t+1}, j_{t+1}), \dots, (i_r, j_r), (i_0, j_0), \dots, (i_t, j_t)$ . Then  $\mathcal{O}'^{\leftarrow} = (i_t, j_t), \dots, (i_0, j_0), (i_r, j_r), \dots, (i_{t+1}, j_{t+1})$ , which is shift-equivalent to  $\mathcal{O}^{\leftarrow} = (i_r, j_r), \dots, (i_1, j_1), (i_0, j_0)$  with shift equal to  $r - t$ . Now we know that  $\mathcal{O}'^{\leftarrow} \stackrel{s}{\sim} \mathcal{O}^{\leftarrow}$  implies  $\mathcal{O}'^{\leftarrow\leftarrow} \stackrel{s}{\sim} \mathcal{O}^{\leftarrow\leftarrow}$ , and this is the same as  $\mathcal{O}' \stackrel{s}{\sim} \mathcal{O}$ .  $\square$

To visually depict an ordering  $\mathcal{O}$  of  $\mathcal{P}_m$  we make use of the symmetric matrix  $M_{\mathcal{O}} = (m_{rt})$  of order  $m$ , defined by the rule

$$m_{i(k)j(k)} = m_{j(k)i(k)} = k, \quad k = 0, 1, \dots, M-1, \quad M = m(m-1)/2.$$

We set  $m_{rr} = -1$ ,  $1 \leq r \leq m$ , but since the pairs  $(r, r)$  do not appear in  $\mathcal{O}$ , we will rather use  $*$  to represent  $-1$ .

EXAMPLE 2.4. As an illustration, we depict the matrices  $M_{\mathcal{O}_{\text{row}}^{\leftarrow}}$  and  $M_{\mathcal{O}_{\text{col}}^{\leftarrow}}$  for  $m = 5$ .

$$M_{\mathcal{O}_{\text{row}}^{\leftarrow}} = \begin{bmatrix} * & 9 & 8 & 7 & 6 \\ 9 & * & 5 & 4 & 3 \\ 8 & 5 & * & 2 & 1 \\ 7 & 4 & 2 & * & 0 \\ 6 & 3 & 1 & 0 & * \end{bmatrix}, \quad M_{\mathcal{O}_{\text{col}}^{\leftarrow}} = \begin{bmatrix} * & 9 & 8 & 6 & 3 \\ 9 & * & 7 & 5 & 2 \\ 8 & 7 & * & 4 & 1 \\ 6 & 5 & 4 & * & 0 \\ 3 & 2 & 1 & 0 & * \end{bmatrix}.$$

These two matrices give us information on the order in which the off-diagonal blocks in the block-matrix from (2.2) are annihilated during each cycle.

**2.2.1. Permutation equivalent strategies.** Let us introduce yet another equivalence relation on  $\mathcal{O}(\mathcal{P}_m)$  and on the set on cyclic and quasi-cyclic pivot strategies.

Two pivot orderings  $\mathcal{O}, \mathcal{O}' \in \mathcal{O}(\mathcal{P}_m)$  are permutation equivalent if  $M_{\mathcal{O}'} = PM_{\mathcal{O}}P^T$  holds for some permutation matrix  $P$ . In that case we write  $\mathcal{O}' \stackrel{p}{\sim} \mathcal{O}$  and  $I_{\mathcal{O}'} \stackrel{p}{\sim} I_{\mathcal{O}}$ . Recall that each permutation matrix  $P$  of order  $m$  is defined by some permutation  $p$  of the set  $\mathcal{S}_m = \{1, 2, \dots, m\}$  via the relation

$$(2.8) \quad Pe_r = e_{p(r)}, \quad 1 \leq r \leq m.$$

Here  $I_m = [e_1, \dots, e_m]$ . The mapping  $p \mapsto P$  is an isomorphism between the symmetric group on the set  $\mathcal{S}_m$  and the group of permutation matrices of order  $m$ . If  $X = (x_{rt})$  is any square matrix of order  $m$ , then  $PXP^T = (x_{p^{-1}(r), p^{-1}(t)})$ . Hence, if  $\tilde{\mathcal{O}}$  is permutation equivalent to  $\mathcal{O}$  and  $M_{\mathcal{O}} = (m_{rt})$ ,  $M_{\tilde{\mathcal{O}}} = (\tilde{m}_{rt})$ , then  $\tilde{m}_{p(r)p(t)} = m_{rt}$  holds for all  $r, t$ . This relation shows that the  $(r, t)$ -element of  $M_{\mathcal{O}}$  becomes the  $(p(r), p(t))$ -element of  $M_{\tilde{\mathcal{O}}}$ .

Therefore, if  $\tilde{\mathcal{O}} \stackrel{\mathcal{P}}{\sim} \mathcal{O}$  and  $\mathcal{O} = (i_0, j_0), \dots, (i_{M-1}, j_{M-1})$ , then  $\tilde{\mathcal{O}} = (\mathfrak{p}(i_0), \mathfrak{p}(j_0)), \dots, (\mathfrak{p}(i_{M-1}), \mathfrak{p}(j_{M-1}))$ . Here it is presumed that in the case  $\mathfrak{p}(i) > \mathfrak{p}(j)$ , the pair  $(\mathfrak{p}(i), \mathfrak{p}(j))$  in the ordering  $\tilde{\mathcal{O}}$  is replaced by  $(\mathfrak{p}(j), \mathfrak{p}(i))$ .

EXAMPLE 2.5. Let  $m = 4$ ,  $\mathcal{S}_m = \{1, 2, 3, 4\}$ ,  $\mathcal{O} = (1, 2), (2, 3), (2, 4), (3, 4), (1, 3), (1, 4)$  and  $\mathbf{P} = [e_2, e_4, e_3, e_1]$  so that  $\mathfrak{p} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix}$ ,  $\mathfrak{p}^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 3 & 2 \end{pmatrix}$ ,  $\mathbf{P}^T = [e_4, e_1, e_3, e_2]$ . Let  $\tilde{\mathcal{O}}$  be such that  $\mathbf{M}_{\tilde{\mathcal{O}}} = \mathbf{P}\mathbf{M}_{\mathcal{O}}\mathbf{P}^T$ . Since

$$\mathbf{M}_{\tilde{\mathcal{O}}} = \begin{bmatrix} e_4^T \\ e_1^T \\ e_3^T \\ e_2^T \end{bmatrix} \begin{bmatrix} * & 0 & 4 & 5 \\ 0 & * & 1 & 2 \\ 4 & 1 & * & 3 \\ 5 & 2 & 3 & * \end{bmatrix} [e_4, e_1, e_3, e_2] = \begin{bmatrix} * & 5 & 3 & 2 \\ 5 & * & 4 & 0 \\ 3 & 4 & * & 1 \\ 2 & 0 & 1 & * \end{bmatrix},$$

we have  $\tilde{\mathcal{O}} = (2, 4), (3, 4), (1, 4), (1, 3), (2, 3), (1, 2)$ . On the other hand, we have

$$\begin{aligned} \mathcal{O}(\mathfrak{p}) &\equiv (\mathfrak{p}(1), \mathfrak{p}(2)), (\mathfrak{p}(2), \mathfrak{p}(3)), (\mathfrak{p}(2), \mathfrak{p}(4)), (\mathfrak{p}(3), \mathfrak{p}(4)), (\mathfrak{p}(1), \mathfrak{p}(3)), (\mathfrak{p}(1), \mathfrak{p}(4)) \\ &= (2, 4), (3, 4), (1, 4), (1, 3), (2, 3), (1, 2) = \tilde{\mathcal{O}}. \end{aligned}$$

Now it is easy to extend the notion of permutation equivalence from pivot orderings to pivot sequences from  $\mathcal{O}(\mathcal{P}_m)$ .

DEFINITION 2.6. Let  $\mathcal{O} = (i_0, j_0), (i_1, j_1), \dots, (i_{T-1}, j_{T-1}) \in \mathcal{O}(\mathcal{P}_m)$ ,  $T \geq M$ . The sequence  $\mathcal{O}' \in \mathcal{O}(\mathcal{P}_m)$  is permutation equivalent to  $\mathcal{O}$ , and we write  $\mathcal{O}' \stackrel{\mathcal{P}}{\sim} \mathcal{O}$ , if there is a permutation  $\mathfrak{q}$  of the set  $\mathcal{S}_m$  such that  $\mathcal{O}' = (\mathfrak{q}(i_0), \mathfrak{q}(j_0)), (\mathfrak{q}(i_1), \mathfrak{q}(j_1)), \dots, (\mathfrak{q}(i_{T-1}), \mathfrak{q}(j_{T-1}))$ . Then  $\mathcal{O}'$  is denoted by  $\mathcal{O}(\mathfrak{q})$ .

Since  $\mathcal{O} = \mathcal{O}(\mathfrak{e})$ , where  $\mathfrak{e}$  is the identity permutation, we have  $\mathcal{O} \stackrel{\mathcal{P}}{\sim} \mathcal{O}$ . If  $\mathcal{O}' = \mathcal{O}(\mathfrak{q})$  then  $\mathcal{O} = \mathcal{O}'(\mathfrak{q}^{-1})$ . Also if  $\mathcal{O}' = \mathcal{O}(\mathfrak{q})$  and  $\mathcal{O}'' = \mathcal{O}'(\mathfrak{q}')$ , then  $\mathcal{O}'' = \mathcal{O}(\mathfrak{q}' \circ \mathfrak{q})$ , where  $\circ$  denotes the binary operation in the permutation group, which is simply the composition of functions. We conclude that  $\stackrel{\mathcal{P}}{\sim}$  is an equivalence relation on the set  $\mathcal{O}(\mathcal{P}_m)$ . Note that  $\mathcal{O}_r = \mathcal{O}_{r-1}(\mathfrak{q}_r)$ ,  $1 \leq r \leq t$ , implies  $\mathcal{O}_t = \mathcal{O}_0(\mathfrak{q})$  with  $\mathfrak{q} = \mathfrak{q}_t \circ \mathfrak{q}_{t-1} \circ \dots \circ \mathfrak{q}_1$ .

LEMMA 2.7. Let  $\mathcal{O}, \mathcal{O}_1, \mathcal{O}_2, \mathcal{O}_3, \mathcal{O}_4 \in \mathcal{O}(\mathcal{P}_m)$ .

- (i) If  $\mathcal{O} \stackrel{\mathcal{W}}{\sim} \mathcal{O}_1 \stackrel{\mathcal{P}}{\sim} \mathcal{O}_2$ , then there is  $\mathcal{O}' \in \mathcal{O}(\mathcal{P}_m)$  such that  $\mathcal{O} \stackrel{\mathcal{P}}{\sim} \mathcal{O}' \stackrel{\mathcal{W}}{\sim} \mathcal{O}_2$ .
- (ii) If  $\mathcal{O} \stackrel{\mathcal{P}}{\sim} \mathcal{O}_3 \stackrel{\mathcal{W}}{\sim} \mathcal{O}_4$ , then there is  $\tilde{\mathcal{O}} \in \mathcal{O}(\mathcal{P}_m)$  such that  $\mathcal{O} \stackrel{\mathcal{W}}{\sim} \tilde{\mathcal{O}} \stackrel{\mathcal{P}}{\sim} \mathcal{O}_4$ .

*Proof.* (i) It is sufficient to prove this assertion for the two cases: (a)  $\stackrel{\mathcal{W}}{\sim}$  is reduced to  $\sim$  and (b)  $\stackrel{\mathcal{W}}{\sim}$  is reduced to  $\stackrel{\mathcal{S}}{\sim}$ .

(a) It is sufficient to assume that the sequence  $\mathcal{O}_1$  is obtained from  $\mathcal{O}$  by applying one admissible transposition. Let  $\mathcal{O} = (i_0, j_0), \dots, (i_r, j_r), (i_{r+1}, j_{r+1}), \dots, (i_{T-1}, j_{T-1})$  with  $\{i_r, j_r\} \cap \{i_{r+1}, j_{r+1}\} = \emptyset$  so that  $\mathcal{O}_1 = (i_0, j_0), \dots, (i_{r+1}, j_{r+1}), (i_r, j_r), \dots, (i_{T-1}, j_{T-1})$ . If  $\mathcal{O}_2 = \mathcal{O}_1(\mathfrak{q})$ , then

$$\mathcal{O}_2 = (\mathfrak{q}(i_0), \mathfrak{q}(j_0)), \dots, (\mathfrak{q}(i_{r+1}), \mathfrak{q}(j_{r+1})), (\mathfrak{q}(i_r), \mathfrak{q}(j_r)), \dots, (\mathfrak{q}(i_{T-1}), \mathfrak{q}(j_{T-1})).$$

Since  $\mathfrak{q}$  is a bijection from  $\mathcal{S}_m$  onto itself, we have  $\{\mathfrak{q}(i_r), \mathfrak{q}(j_r)\} \cap \{\mathfrak{q}(i_{r+1}), \mathfrak{q}(j_{r+1})\} = \emptyset$ . Therefore, we can set  $\mathcal{O}' = \mathcal{O}(\mathfrak{q})$ , i.e.,

$$\mathcal{O}' = (\mathfrak{q}(i_0), \mathfrak{q}(j_0)), \dots, (\mathfrak{q}(i_r), \mathfrak{q}(j_r)), (\mathfrak{q}(i_{r+1}), \mathfrak{q}(j_{r+1})), \dots, (\mathfrak{q}(i_{T-1}), \mathfrak{q}(j_{T-1})).$$

(b) Let  $\mathcal{O}$  be as in case (a) and assume that the cut has been made after the term  $(i_r, j_r)$ , so that  $\mathcal{O}_1 = (i_{r+1}, j_{r+1}), \dots, (i_{T-1}, j_{T-1}), (i_0, j_0), \dots, (i_r, j_r)$ . Let  $\mathcal{O}_2 = \mathcal{O}_1(\mathfrak{q}) = (\mathfrak{q}(i_{r+1}), \mathfrak{q}(j_{r+1})), \dots, (\mathfrak{q}(i_{T-1}), \mathfrak{q}(j_{T-1})), (\mathfrak{q}(i_0), \mathfrak{q}(j_0)), \dots, (\mathfrak{q}(i_r), \mathfrak{q}(j_r))$ . Obviously,

we can define  $\mathcal{O}'$  to be the same as in case (a). To obtain  $\mathcal{O}_2$  from  $\mathcal{O}'$ , one has to make the cut just after the term  $(q(i_r), q(j_r))$ , i.e., to use the shift  $r + 1$ .

(ii) The proof is quite similar to the proof of (i). First, consider the case when  $\overset{w}{\sim}$  is reduced to  $\sim$ , and  $\sim$  is given by one admissible transposition that interchanges the terms at positions  $r$  and  $r + 1$  as in case (a) above. Let  $\mathcal{O}$  be as in case (i) and denote  $\mathcal{O}_3 = \mathcal{O}(\tilde{q})$ . Then  $\mathcal{O}_4$  is the same as  $\mathcal{O}_2$ , provided that  $q$  is replaced by  $\tilde{q}$  and  $\tilde{\mathcal{O}}$  is like  $\mathcal{O}_1$ . Next, consider the case when  $\overset{w}{\sim}$  is reduced to  $\overset{s}{\sim}$  and  $\mathcal{O}_3 = \mathcal{O}(\tilde{q})$ . Assume that the cut has been made as in case (b) behind the term  $(\tilde{q}(i_r), \tilde{q}(j_r))$ . Then  $\mathcal{O}_4$  is the same as  $\mathcal{O}_2$  (from case (b)) provided that  $q$  is replaced by  $\tilde{q}$  and  $\tilde{\mathcal{O}}$  is like  $\mathcal{O}_1$ .  $\square$

PROPOSITION 2.8. *Let  $\mathcal{O}, \mathcal{O}', \mathcal{O}_1, \mathcal{O}_2, \mathcal{O}_3, \dots, \mathcal{O}_{2t} \in \mathcal{O}(\mathcal{P}_m)$ ,  $t \geq 1$ . If*

$$(2.9) \quad \mathcal{O} \overset{w}{\sim} \mathcal{O}_1 \overset{p}{\sim} \mathcal{O}_2 \overset{w}{\sim} \mathcal{O}_3 \overset{p}{\sim} \mathcal{O}_4 \overset{w}{\sim} \dots \overset{w}{\sim} \mathcal{O}_{2t-1} \overset{p}{\sim} \mathcal{O}_{2t} \overset{w}{\sim} \mathcal{O}',$$

then there exist  $\mathcal{O}'_0, \tilde{\mathcal{O}}_0 \in \mathcal{O}(\mathcal{P}_m)$  such that  $\mathcal{O} \overset{p}{\sim} \mathcal{O}'_0 \overset{w}{\sim} \mathcal{O}'$  and  $\mathcal{O} \overset{w}{\sim} \tilde{\mathcal{O}}_0 \overset{p}{\sim} \mathcal{O}'$ .

If  $\mathcal{O}_{2r} = \mathcal{O}_{2r-1}(q_r)$ , for  $1 \leq r \leq t$ , then  $\mathcal{O}'_0 = \mathcal{O}(q)$  and  $\mathcal{O}' = \tilde{\mathcal{O}}_0(q)$  with  $q = q_t \circ q_{t-1} \circ \dots \circ q_1$ .

*Proof.* Using assertion (i) (resp. assertion (ii)) of the previous lemma, one can gradually move all appearances of  $\overset{p}{\sim}$  to the left (resp. right) end of the chain (2.9). First, the leftmost (resp. rightmost)  $\overset{p}{\sim}$  is moved. The leftmost (resp. rightmost) part of the chain takes the form

$$\mathcal{O} \overset{p}{\sim} \mathcal{O}'_1 \overset{w}{\sim} \mathcal{O}_2 \overset{w}{\sim} \mathcal{O}_3 \overset{p}{\sim} \mathcal{O}_4 \dots \quad (\text{resp. } \dots \mathcal{O}_{2t-3} \overset{p}{\sim} \mathcal{O}_{2t-2} \overset{w}{\sim} \mathcal{O}_{2t-1} \overset{w}{\sim} \tilde{\mathcal{O}}_{2t} \overset{p}{\sim} \mathcal{O}').$$

Note that two consecutive  $\overset{w}{\sim}$  can be replaced by one, so we can remove  $\mathcal{O}_2$  (resp.  $\mathcal{O}_{2t-1}$ ) from the obtained chain. Next,  $\overset{p}{\sim}$  that links  $\mathcal{O}_3$  and  $\mathcal{O}_4$  (resp.  $\mathcal{O}_{2t-3}$  and  $\mathcal{O}_{2t-2}$ ) is moved. The leftmost (resp. rightmost) part of the chain takes the form

$$\mathcal{O} \overset{p}{\sim} \mathcal{O}'_1 \overset{p}{\sim} \mathcal{O}'_3 \overset{w}{\sim} \mathcal{O}_4 \overset{w}{\sim} \mathcal{O}_5 \dots \quad (\text{resp. } \dots \mathcal{O}_{2t-4} \overset{w}{\sim} \mathcal{O}_{2t-3} \overset{w}{\sim} \tilde{\mathcal{O}}_{2t-2} \overset{p}{\sim} \tilde{\mathcal{O}}_{2t} \overset{p}{\sim} \mathcal{O}').$$

Continuing this way one ultimately obtains

$$(2.10) \quad \begin{aligned} & \mathcal{O} \overset{p}{\sim} \mathcal{O}'_1 \overset{p}{\sim} \mathcal{O}'_3 \overset{p}{\sim} \dots \overset{p}{\sim} \mathcal{O}'_{2t-1} \overset{w}{\sim} \mathcal{O}_{2t} \overset{w}{\sim} \mathcal{O}' \\ & (\text{resp. } \mathcal{O} \overset{w}{\sim} \mathcal{O}_1 \overset{w}{\sim} \tilde{\mathcal{O}}_2 \overset{p}{\sim} \dots \overset{p}{\sim} \tilde{\mathcal{O}}_{2t-2} \overset{p}{\sim} \tilde{\mathcal{O}}_{2t} \overset{p}{\sim} \mathcal{O}'). \end{aligned}$$

Here  $\mathcal{O}_{2t}$  (resp.  $\mathcal{O}_1$ ) can be removed. Note that  $\overset{p}{\sim}$  is an equivalence relation. Hence by the transitivity property, the leftmost part of the chain  $\mathcal{O} \overset{p}{\sim} \mathcal{O}'_1 \overset{p}{\sim} \mathcal{O}'_3 \overset{p}{\sim} \dots \overset{p}{\sim} \mathcal{O}'_{2t-1}$  can be replaced by  $\mathcal{O} \overset{p}{\sim} \mathcal{O}'_{2t-1}$  (and similarly for the rightmost part of the chain). To complete the proof of the first assertion one has to rename  $\mathcal{O}'_{2t-1}$  as  $\mathcal{O}'_0$  (resp.  $\tilde{\mathcal{O}}_2$  as  $\tilde{\mathcal{O}}_0$ ).

The proof of the second assertion is the same, but one can use more information. Now, from the proof of the preceding lemma we know that in the final chain (2.10) we have  $\mathcal{O}'_1 = \mathcal{O}(q_1)$  and  $\mathcal{O}'_{2r-1} = \mathcal{O}'_{2r-3}(q_r)$ , for  $2 \leq r \leq t$ , (resp.  $\mathcal{O}' = \tilde{\mathcal{O}}(q_t)$  and  $\tilde{\mathcal{O}}_{2r} = \tilde{\mathcal{O}}_{2r-2}(q_{r-1})$ ,  $2 \leq r \leq t$ ). Hence  $\mathcal{O}'_0 = \mathcal{O}'_{2t-1} = \mathcal{O}(q)$  (resp.  $\mathcal{O}' = \tilde{\mathcal{O}}_2(q) = \tilde{\mathcal{O}}_0(q)$ ) with  $q = q_t \circ q_{t-1} \circ \dots \circ q_1$ .  $\square$

Obviously, the first or the last (or both) appearance of  $\overset{w}{\sim}$  in the chain (2.9) can be omitted provided that  $\mathcal{O} = \mathcal{O}_1$  and  $\mathcal{O}' = \mathcal{O}_{2t}$  (resp.  $\mathcal{O} = \mathcal{O}_1$  and  $\mathcal{O}' = \mathcal{O}_{2t}$ ).

Two sequences  $\mathcal{O}, \mathcal{O}' \in \mathcal{O}(\mathcal{P}_m)$  can be linked via a long chain like (2.9), which may include all equivalence relations  $\sim, \overset{s}{\sim}, \overset{w}{\sim}$ , and  $\overset{p}{\sim}$  introduced so far. Proposition 2.8 shows that each such chain can be reduced to a short chain that uses just one  $\overset{p}{\sim}$  and one  $\overset{w}{\sim}$ . Furthermore, weak equivalence can be written in the most compact form (2.7).



DEFINITION 2.9. Two sequences  $\mathcal{O}, \mathcal{O}' \in \mathcal{O}(\mathcal{P}_m)$  are connected by a chain of equivalence relations if there exist  $\mathcal{O}_1, \dots, \mathcal{O}_r \in \mathcal{O}(\mathcal{P}_m)$ ,  $r \geq 0$ , such that in the sequence  $\mathcal{O}, \mathcal{O}_1, \dots, \mathcal{O}_r, \mathcal{O}'$  each two neighboring terms are linked to  $\sim, \overset{s}{\sim}, \overset{w}{\sim}$ , or  $\overset{p}{\sim}$ . The chain is in canonical form if it looks like (2.7) with one  $\overset{p}{\sim}$  placed in front of it or after it.

We conclude that every chain of equivalence relations can be reduced to the canonical form.

Let us return to the block matrix  $A$  from relation (2.2). To a partition  $\pi = (n_1, \dots, n_m)$  we associate the  $n$ -tuple  $(s_1, \dots, s_m)$ ,

$$s_i = s_{i-1} + n_i, \quad 2 \leq i \leq m, \quad s_1 = n_1.$$

Note that the sequence of the first  $n$  natural numbers,  $1, 2, \dots, n$ , can be written as  $1, \dots, s_1, s_1 + 1, \dots, s_2, \dots, s_{m-1} + 1, \dots, s_m$ , which is the same as

$$s_1 - n_1 + 1, \dots, s_1, s_2 - n_2 + 1, \dots, s_2, \dots, s_m - n_m + 1, \dots, s_m.$$

Here,  $s_r - n_r + 1, \dots, s_r$  are indices of the columns (resp. rows) which define the  $r$ th block-column (resp. block-row) of  $A$ . Let  $\mathcal{O} \in \mathcal{P}_m$  and  $\tilde{\mathcal{O}} = \mathcal{O}(p)$  so that we have  $\tilde{\mathcal{O}} = (p(i_0), p(j_0)), (p(i_1), p(j_1)), \dots, (p(i_{T-1}), p(j_{T-1}))$ . The permutation  $p$  of the set  $\mathcal{S}_n = \{1, 2, \dots, n\}$  associated with  $p$  is defined by:

$$p = \begin{pmatrix} s_1 - n_1 + 1, & \dots & s_1, & s_2 - n_2 + 1, & \dots & s_m - n_m + 1, & \dots & s_m \\ s_{p(1)} - n_{p(1)} + 1, \dots & s_{p(1)}, & s_{p(2)} - n_{p(2)} + 1, \dots & s_{p(m)} - n_{p(m)} + 1, \dots & s_{p(m)} \end{pmatrix}.$$

Using the same rule (2.8), we obtain the permutation matrix  $P$  of order  $n$ , associated with  $p$ . It satisfies  $Pe_t = e_{p(t)}$ ,  $1 \leq t \leq n$ . The matrix  $P$  has the form  $P = [E_{p(1)} \cdots E_{p(m)}]$ , where each  $E_{p(k)}$  is an  $n \times n_{p(k)}$  matrix (i.e., a single block-column) that differs from the zero matrix only in its  $p(k)$ th block-row,

$$E_{p(k)} = \begin{bmatrix} 0 & \cdots & 0 & I_{p(k)} & 0 & \cdots & 0 \end{bmatrix}^T, \quad 1 \leq k \leq m.$$

$n_1 \qquad \qquad \qquad n_{p(k)} \qquad \qquad \qquad n_m$

Let  $A^{(T)}$  be the matrix obtained from  $A$  by applying one sweep of the quasi-cyclic block Jacobi method defined by  $I_{\mathcal{O}}$ . The iterative process has the form (2.5). Hence using  $P$  and  $P^T$ , we can write

$$\begin{aligned} (2.11) \quad PA^{(T)}P^T &= P(U_{T-1}^T U_{T-2}^T \cdots U_0^T P^T) PAP^T (PU_0 \cdots U_{T-2} U_{T-1}) P^T \\ &= (PU_{T-1}^T P^T) \cdots (PU_0^T P^T) PAP^T (PU_0 P^T) \cdots (PU_{T-1} P^T) \\ &= \tilde{U}_{T-1}^T \cdots \tilde{U}_0^T (PAP^T) \tilde{U}_0 \cdots \tilde{U}_{T-1}, \end{aligned}$$

where

$$(2.12) \quad \tilde{U}_k = PU_k P^T, \quad 0 \leq k \leq T-1.$$

Each  $\tilde{U}_k$  is an orthogonal elementary block matrix whose pivot pair is  $(p(i_k), p(j_k))$ . We can interpret the process (2.11) as a quasi-cyclic block Jacobi method defined by  $I_{\tilde{\mathcal{O}}}$ . When it is applied to  $\tilde{A} = PAP^T$ , after one sweep it results in  $\tilde{A}^{(T)} = PA^{(T)}P^T$ . Indeed, at step  $k$  of that process we have

$$(2.13) \quad \tilde{A}^{(k+1)} = \tilde{U}_k^T \tilde{A}^{(k)} \tilde{U}_k, \quad 0 \leq k \leq T-1,$$

where  $\tilde{A}^{(k)} = PA^{(k)}P^T$  for any  $k$ . For the matrix  $\tilde{A}^{(k+1)}$  we know that its  $(p(i_k), p(j_k))$ -pivot submatrix of order  $n_{p(i_k)} + n_{p(j_k)}$  is diagonal. Hence, it is a quasi-cyclic block Jacobi

method. Moreover, if  $U_k$  is in the class  $\text{UBCE}_{\pi}(\varrho)$  from Section 2.3, the same will be true for  $\tilde{U}_k$ .

The process (2.13) is a block Jacobi method for the matrix  $\tilde{A}$  which carries a block-matrix partition defined by  $\pi_p = (n_{p(1)}, \dots, n_{p(m)})$ . Thus, if  $A$  is replaced by  $\tilde{A}$  and the permutation  $\pi$  by  $\pi_p$ , then the block method (2.13) is defined by the pivot sequence  $\tilde{\mathcal{O}} = \mathcal{O}(p) \in \mathcal{O}(\mathcal{P}_m)$ . We can formally write  $(\pi, \mathcal{O}, A) \xrightarrow{\mathcal{P}} (\pi_p, \mathcal{O}(p), PAP^T)$ . This is equivalent to  $(\pi_{p^{-1}}, \mathcal{O}(p^{-1}), P^T AP) \xrightarrow{\mathcal{P}} (\pi, \mathcal{O}, A)$ .

We end this subsection with two remarks. First, the reverse ordering  $\mathcal{O}^{\leftarrow}$  is not the same as  $\mathcal{O}(\tilde{e})$ , where

$$(2.14) \quad \tilde{e} = \begin{pmatrix} 1 & 2 & \dots & m \\ m & m-1 & \dots & 1 \end{pmatrix}.$$

Examples which confirm this claim are shown in Section 3.2.

Second, if in Lemma 2.3 the equivalence relation  $\overset{\sim}{\sim}$  is replaced by  $\overset{\mathcal{P}}{\sim}$ , the assertion will remain to hold. The proof is trivial. Indeed, if in the chain (2.7), which may now include  $\overset{\mathcal{P}}{\sim}$ , only one sequence  $\mathcal{O}_t$  is replaced by  $\mathcal{O}_t^{\leftarrow}$ , then all the sequences have to be replaced by their inverses. Otherwise, the chain can be broken into two chains that are not mutually connected.

**2.3. Global convergence.** A block Jacobi method is *convergent* for  $A$  if the obtained sequence of matrices  $(A^{(k)})$  converges to some diagonal matrix  $\Lambda$ . The method is *globally convergent* if it is convergent for every symmetric matrix  $A$ . This definition assumes that the partition  $\pi$  is arbitrary. In particular, one can take  $\pi = (1, 1, \dots, 1)$ , which means that it is the proper generalization of the standard notion of global convergence. The words “global” and “globally” are often omitted. For example, if one says that the block method converges for some pivot strategy, this means that the method converges for every initial symmetric matrix. For global convergence considerations, it is irrelevant whether the diagonal blocks of the initial matrix are diagonal submatrices. Namely, after some iteration (within the first sweep) this property will be fulfilled, and it will remain to hold until convergence. To measure how much the method has converged, we use the quantity

$$S(A) = \frac{\sqrt{2}}{2} \|A - \text{diag}(A)\|_F = \left[ \sum_{s=1}^{n-1} \sum_{t=s+1}^n |a_{st}|^2 \right]^{\frac{1}{2}},$$

where  $\|X\|_F = \sqrt{\text{trace}(X^T X)}$  stands for the Frobenius norm of  $X$ . In the definition of  $S(A)$  we could have used blocks instead of elements, but since the diagonal blocks are diagonal submatrices, this reduces to the same quantity. Obviously, the convergence of a block Jacobi method applied to  $A$  implies that  $S(A^{(k)}) \rightarrow 0$  as  $k \rightarrow \infty$ . The converse is true provided that the diagonal elements of  $\text{diag}(\Lambda_{ii}^{(k+1)}, \Lambda_{jj}^{(k+1)})$  from (2.6) are always ordered in some prescribed order, typically nonincreasingly.

**THEOREM 2.10.** *Let  $A$  be a symmetric matrix and  $A^{(k)}$ ,  $k \geq 0$ , be the sequence obtained by applying the block Jacobi method to  $A$ . Let the pivot strategy be cyclic or quasi-cyclic, and assume that  $\lim_{k \rightarrow \infty} S(A^{(k)}) = 0$ .*

- (i) *If the algorithm that diagonalizes the pivot submatrix always delivers  $\text{diag}(\Lambda_{ii}^{(k+1)}, \Lambda_{jj}^{(k+1)})$  with nonincreasingly (resp. nondecreasingly) ordered diagonal elements, then  $\Lambda = \lim_{k \rightarrow \infty} A^{(k)}$  and the diagonal elements of  $\Lambda$  are nonincreasingly (resp. nondecreasingly) ordered.*
- (ii) *If the algorithm that diagonalizes the pivot submatrix is any standard (i.e., element-wise) globally convergent Jacobi method, then  $\Lambda = \lim_{k \rightarrow \infty} A^{(k)}$ .*

*Proof.* The proof has been moved to Appendix A. □

Theorem 2.10 implies that the global convergence problem of the block Jacobi method reduces to the convergence of the sequence  $S(A^{(k)})$ ,  $k \geq 0$ , to zero.

By inspecting the proofs of the results related to the global convergence of the standard cyclic Jacobi method [11, 32], one finds that they hold for block methods, too. We summarize those results as follows.

**THEOREM 2.11.** *If a block Jacobi method converges for some cyclic strategy, then it converges for all strategies that are weakly equivalent to it. The block methods defined by equivalent cyclic strategies generate the same matrices after each full cycle, and within the same cycle they produce the same sets of orthogonal elementary matrices.*

Indeed, the proof for the standard Jacobi method essentially uses the fact that commuting pivot pairs results in commuting the Jacobi rotations. Similarly, the proof for the block method uses the fact that commuting pivot pairs  $(i, j)$  and  $(p, q)$  implies commuting orthogonal elementary matrices  $U_{ij}$  and  $U_{pq}$ . For the convergence of the diagonal elements one should presume conditions like those in Theorem 2.10 for the kernel algorithms. The second part of the theorem holds because it presumes that the block Jacobi method uses the same kernel algorithm. Theorem 2.11 also holds for the quasi-cyclic methods provided that care is taken for the blocks that are annihilated more than once within a sweep.

A sufficient condition for global convergence of serial standard Jacobi methods is the existence of a strictly positive uniform lower bound for the cosines of the rotation angles; see [10]. For the serial block Jacobi methods, a sufficient condition for the global convergence is that the transformation matrices  $U_k$  from relation (2.5) have a strictly positive uniform lower bound for the singular values of the diagonal blocks [5]. This condition also appears in the global convergence analysis of more general serial Jacobi-type methods [18]. Unitary elementary block matrices which satisfy such a property are called *UBC* (uniformly bounded cosine) transformation matrices in [5]. In the same paper it was shown that for every unitary matrix of order  $n$  and every partition  $\varsigma = (n_1, n_2)$  of  $n$ , there exists a permutation matrix  $J$  such that for the leading  $n_1 \times n_1$  block of  $\tilde{U} = UJ$  one has

$$\sigma_{\min}(\tilde{U}_{11}) \geq \gamma_\varsigma > \tilde{\gamma}_n > 0, \quad \gamma_\varsigma = \frac{3}{\sqrt{(4^{n_1} + 6n_2 - 1)(n_2 + 1)}}, \quad \tilde{\gamma}_n = \frac{3\sqrt{2}}{\sqrt{4^n + 26}}.$$

The second inequality, which involves  $\tilde{\gamma}_n$ , has been proved in [18]. Hence, every unitary elementary block matrix can be made UBC by an appropriate permutation of its nontrivial columns.

In this paper we will use UBC transformation matrices. Therefore, for each  $0 < \varrho \leq 1$ , we introduce the class  $UBCE_\pi(\varrho)$  of elementary unitary UBC block matrices as follows. The unitary elementary block matrix  $U_{ij}$  from relation (2.3) belongs to the class  $UBCE_\pi(\varrho)$  if

$$(2.15) \quad \sigma_{\min}(U_{ii}) = \sigma_{\min}(U_{jj}) \geq \varrho\gamma_{ij} > \varrho\tilde{\gamma}_{n_i+n_j} \geq \varrho\tilde{\gamma}_n$$

holds, where

$$(2.16) \quad \gamma_{ij} = \frac{3}{\sqrt{(4^{n_i} + 6n_j - 1)(n_j + 1)}}.$$

If  $\pi$  is understood, we will write  $UBCE(\varrho)$ , while if  $\varrho = 1$ , the parameter  $\varrho$  will also be omitted from the notation. In definitions, statements and ordinary text, at every appearance of  $\varrho$ , we will automatically assume that  $0 < \varrho \leq 1$ . We will use the same notation when  $U_{ij}$  is real, i.e., orthogonal.

Note that  $U_k$  from relation (2.5) is an orthogonal elementary matrix defined by the pivot pair  $(i, j)$ , where  $i = i(k)$ ,  $j = j(k)$ . To make  $U_k$  a UBCE transformation, one has to find the permutation  $J_k$  and then compute  $U_k J_k$ . This can be accomplished (see [5]) by performing the QR factorization with column pivoting of  $[U_{ii}^{(k)} \ U_{ij}^{(k)}]$  from relation (2.6). The QR factorization yields  $\widehat{J}_k$ , which then defines  $J_k$  as  $J_k = \mathcal{E}(i, j, \widehat{J}_k)$ . Then  $\widehat{U}_k$  also belongs to the class  $\text{UBCE}_{\varsigma_{ij}}$ , where  $\varsigma_{ij} = (n_i, n_j)$ . If  $(n_i, n_j)$  is understood,  $\varsigma_{ij}$  will be omitted. One easily verifies that  $U_k J_k$  diagonalizes the pivot submatrix  $\widehat{A}^{(k)}$ , and the similarity transformation with  $J_k$  does not change the Frobenius norm of the affected blocks of  $U_k^T A^{(k)} U_k$ . In addition, one can show that once  $S(A^{(k)})$  is sufficiently small and the diagonal elements affiliated with the same eigenvalue occupy successive positions along the diagonal, the permutations  $J_k$  are no longer needed (see [5]), i.e.,  $J_k$  can be taken to be the identity. If  $\pi = (1, 1, \dots, 1)$ , then one can replace  $\gamma_{ij}$  by  $\sqrt{2}/2$ . The uniform bound  $\sqrt{2}/2\varrho$  is the one from the known Forsythe-Henrici condition [10].

REMARK 2.12. The parameter  $\varrho$  has been introduced for several reasons. First, it simplifies the convergence analysis of the more general iterative process described in Section 5. Second, as will be shown in Sections 3 and 4, the convergence proofs for the symmetric block Jacobi method hold for any  $0 < \varrho \leq 1$ . Finally, for the case  $\varrho = 1$ , the determination of the permutation  $J_k$  requires the QR factorization with column pivoting of an  $n_i \times (n_i + n_j)$  matrix. Possibly, for some smaller  $\varrho$ , an appropriate permutation matrix could be obtained for a smaller cost.

**2.4. Block Jacobi annihilators and operators.** Jacobi annihilators and operators have been introduced in [22] as a tool for proving the global and quadratic convergence of the column-cyclic Jacobi method. Later they have been used for proving the global convergence of some norm-reducing Jacobi-type methods for general matrices [12, 14]. In [1, 17, 18] they have been generalized to cope with the block Jacobi methods. Here we define a class of Jacobi annihilators and operators designed precisely for the block Jacobi method for symmetric matrices. They will be referred to as block Jacobi annihilators and operators. This will move us to a more general point of view of the block Jacobi methods, which can be used in the convergence considerations.

First let us introduce some notation. For an arbitrary  $p \times q$  matrix  $X$ , we define the column vector comprising the columns of  $X$ ,

$$\text{col}(X) = [x_{11}, x_{21}, \dots, x_{p1}, \dots, x_{1q}, \dots, x_{pq}]^T.$$

Let  $\pi = (n_1, \dots, n_m)$  be a partition of  $n$ . Let  $\mathbf{S}_n$  denote the real vector space of symmetric matrices of order  $n$ . Let  $A = (A_{rs}) \in \mathbf{S}_n$  be as in relation (2.2). Its block-matrix partition is determined by  $\pi$ . We define the vector-valued function  $\text{vec}_\pi$  as follows (see [17, 18]),

$$(2.17) \quad \text{vec}_\pi(A) = \begin{bmatrix} c_2 \\ c_3 \\ \vdots \\ c_m \end{bmatrix}, \quad \text{where } c_j = \begin{bmatrix} \text{col}(A_{1j}) \\ \text{col}(A_{2j}) \\ \vdots \\ \text{col}(A_{j-1,j}) \end{bmatrix}, \quad 2 \leq j \leq m.$$

Then

$$(2.18) \quad \text{vec}_\pi : \mathbf{S}_n \rightarrow \mathbb{R}^K, \quad K = N - \sum_{i=1}^m \frac{n_i(n_i - 1)}{2}, \quad N = \frac{n(n - 1)}{2},$$

is a linear operator.

Note that  $\text{vec}_\pi(A)$  contains all off-diagonal elements from the upper block triangular part of  $A$ . They are arranged in  $\text{vec}_\pi(A)$  using double column-wise ordering, one with respect to the blocks in  $A$ , the other with respect to the elements within each block. The function  $\text{vec}_\pi$  is a surjection but not an injection. In order to make it a bijection, we restrict it to the vector subspace  $\mathbf{S}_{0,n}$  of  $\mathbf{S}_n$ , consisting of all matrices from  $\mathbf{S}_n$  whose diagonal blocks (with respect to the block-matrix partition defined by  $\pi$ ) are zero. Let  $\text{vec}_{\pi,0} = \text{vec}_\pi|_{\mathbf{S}_{0,n}}$ . Obviously, the function  $\text{vec}_{\pi,0}$  is an invertible linear operator from  $\mathbf{S}_{0,n}$  to  $\mathbb{R}^K$ . In the following text we will often assume that the partition  $\pi$  is known and it will be omitted from the notation. However, it will be denoted whenever an additional partition is also considered.

If  $a \in \mathbb{R}^K$  and  $A = \text{vec}_0^{-1}(a)$ , then  $A$  is obtained from  $a$  using the block-matrix partition defined by  $\pi$  and the double column-wise ordering as described in relation (2.17). The diagonal blocks are set to zero, and the whole matrix is set to be symmetric. Obviously,  $A$  is uniquely determined by  $a$ .

Beside the linear operators  $\text{vec}$  and  $\text{vec}_0$ , we will make use of the linear operator  $\mathcal{N}_{ij} : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$ , which also uses the block-matrix partition defined by  $\pi$  and sets the pivot submatrix of the argument matrix to zero. When applied to  $A \in \mathbf{S}_n$ ,  $\mathcal{N}_{ij}(A)$  sets the blocks  $A_{ij}$ ,  $A_{ji}$ ,  $A_{ii}$ , and  $A_{jj}$  to zero.

DEFINITION 2.13. Let  $\pi = (n_1, \dots, n_m)$  be a partition of  $n$ , let

$$\widehat{U} = \begin{bmatrix} U_{11} & U_{12} \\ U_{12}^T & U_{22} \end{bmatrix} \begin{matrix} n_i \\ n_j \end{matrix},$$

be an orthogonal matrix of order  $n_i + n_j$ , and let  $U = \mathcal{E}(i, j, \widehat{U})$  be the corresponding elementary block matrix. The transformation  $\mathcal{R}_{ij}(\widehat{U})$  determined by

$$\mathcal{R}_{ij}(\widehat{U})(\text{vec}(A)) = \text{vec}(\mathcal{N}_{ij}(U^T A U)), \quad A \in \mathbf{S}_n,$$

is called the  $ij$ -block Jacobi annihilator. For each pair  $1 \leq i < j \leq m$ ,

$$\mathcal{R}_{ij} = \left\{ \mathcal{R}_{ij}(\widehat{U}) \mid \widehat{U} \text{ is an orthogonal matrix of order } n_i + n_j \right\}$$

is the  $ij$ -class of the block Jacobi annihilators. If all  $\widehat{U}$  are restricted to the class  $UBCE_{\zeta_{ij}}(\varrho)$ , then the resulting  $ij$ -class is denoted by  $\mathcal{R}_{ij}^{UBCE_\pi(\varrho)}$ .

Given  $i, j, \widehat{U}$ , the following algorithm computes the vector  $a' = \mathcal{R}_{ij}(\widehat{U})a$  for  $a \in \mathbb{R}^K$ . It is based on the formula  $\mathcal{R}_{ij}(\widehat{U})a = \text{vec}(\mathcal{N}_{ij}(U^T \text{vec}_0^{-1}(a)U))$ , which can be taken as an equivalent definition of  $\mathcal{R}_{ij}(\widehat{U})$ .

ALGORITHM 2.14 (Computing  $\mathcal{R}_{ij}(\widehat{U})a$ ).

---

$a \in \mathbb{R}^K$	% an arbitrary vector
$A = \text{vec}_0^{-1}(a)$	% this invokes the module which computes
<b>for</b> $r = 1, \dots, m$ <b>do</b>	% the symmetric matrix $A = \text{vec}_0^{-1}(a)$
$A'_{ri} = A_{ri}U_{ii} + A_{rj}U_{ji}$	
$A'_{rj} = A_{ri}U_{ij} + A_{rj}U_{jj}$	
<b>end for</b>	% this part of code computes
<b>for</b> $r = 1, \dots, m$ <b>do</b>	% $A' = U^T A U, U = \mathcal{E}(i, j, \widehat{U})$
$A'_{ir} = U_{ii}^T A_{ir} + U_{ji}^T A_{jr}$	% $\widehat{U}$ is partitioned as in relation (2.4)
$A'_{jr} = U_{ij}^T A_{ir} + U_{jj}^T A_{jr}$	
<b>end for</b>	
$A'_{ij} = 0, A'_{ji} = 0, A'_{ii} = 0, A'_{jj} = 0$	% this part updates $A', A' \leftarrow \mathcal{N}_{ij}(A')$
$a' = \text{vec}(A')$	% the module which computes $\text{vec}(A')$

The matrix  $A'$  from Algorithm 2.14 has the same partition as  $A$ . Note that the mapping  $a \mapsto a'$  is a composition of linear transformations. Therefore, given a basis in  $\mathbb{R}^K$ ,  $\mathcal{R}_{ij}(\widehat{U})$  can be represented by some square matrix of order  $K$ . We will choose the canonic basis  $(e)$ , which consists of the columns of  $I_K$ , and denote the obtained matrix by the bold symbol. Hence,

$$(2.19) \quad \mathcal{R}_{ij}(\widehat{U})(a) = \mathbf{R}_{ij}(\widehat{U})a, \quad a \in \mathbb{R}^K.$$

We will call the matrix  $\mathbf{R}_{ij}(\widehat{U})$  by the same name, the  $ij$ -block Jacobi annihilator, and the appropriate class of matrices will be denoted by

$$\mathcal{R}_{ij} = \left\{ \mathbf{R}_{ij}(\widehat{U}) \mid \widehat{U} \text{ is an orthogonal matrix of order } n_i + n_j \right\}.$$

If all  $\widehat{U}$  are restricted to  $\text{UBCE}(\varrho)$ , the obtained class of block Jacobi annihilators is denoted by  $\mathcal{R}_{ij}^{\text{UBCE}(\varrho)}$ . In the sequel, every mention of the block Jacobi annihilator will refer to the matrix  $\mathbf{R}_{ij}(\widehat{U})$  from relation (2.19).

The following theorem reveals the structure of a block Jacobi annihilator. It is a simplification of [17, Theorem 2.1], and its proof can be found in [1]. The theorem utilizes the function  $\tau(i, j) = (j-1)(j-2)/2 + i$ ,  $1 \leq i < j \leq m$ , and the Kronecker product of matrices. The vectors of length  $K$  and the block Jacobi annihilators of order  $K$  carry the block-partition determined by  $(n_1 n_2, n_1 n_3, n_2 n_3, \dots, n_{m-1} n_m)$ . The spectral norm is denoted by  $\|\cdot\|_2$ .

**THEOREM 2.15** ([1, 17]). *Let  $\pi = (n_1, \dots, n_m)$  be the partition of  $n$  and let  $K, N$  be integers defined in relation (2.18). Let  $(i, j) \in \mathcal{P}_m$ ,  $\mathcal{R} \in \mathcal{R}_{ij}$ ,  $\mathbf{R} = \mathcal{R}(\widehat{U})$ , where  $\widehat{U}$  is an orthogonal matrix of order  $n_i + n_j$ . Then  $\mathbf{R}$  differs from the identity matrix  $I_K$  in exactly  $m-1$  principal submatrices, which are given by the following relations:*

$$\begin{aligned} \mathcal{R}_{\tau(i,j), \tau(i,j)} &= 0, \\ \begin{bmatrix} \mathcal{R}_{\tau(r,i), \tau(r,i)} & \mathcal{R}_{\tau(r,i), \tau(r,j)} \\ \mathcal{R}_{\tau(r,j), \tau(r,i)} & \mathcal{R}_{\tau(r,j), \tau(r,j)} \end{bmatrix} &= \begin{bmatrix} U_{ii}^T \otimes I_{n_r} & U_{ji}^T \otimes I_{n_r} \\ U_{ij}^T \otimes I_{n_r} & U_{jj}^T \otimes I_{n_r} \end{bmatrix}, \quad 1 \leq r \leq i-1, \\ \begin{bmatrix} \mathcal{R}_{\tau(i,r), \tau(i,r)} & \mathcal{R}_{\tau(i,r), \tau(r,j)} \\ \mathcal{R}_{\tau(r,j), \tau(i,r)} & \mathcal{R}_{\tau(r,j), \tau(r,j)} \end{bmatrix} &= \begin{bmatrix} I_{n_r} \otimes U_{ii}^T & S(U_{ji}^T \otimes I_{n_r}) \\ \tilde{S}(I_{n_r} \otimes U_{ij}^T) & U_{jj}^T \otimes I_{n_r} \end{bmatrix}, \quad i+1 \leq r \leq j-1, \\ \begin{bmatrix} \mathcal{R}_{\tau(i,r), \tau(i,r)} & \mathcal{R}_{\tau(i,r), \tau(j,r)} \\ \mathcal{R}_{\tau(j,r), \tau(i,r)} & \mathcal{R}_{\tau(j,r), \tau(j,r)} \end{bmatrix} &= \begin{bmatrix} I_{n_r} \otimes U_{ii}^T & I_{n_r} \otimes U_{ji}^T \\ I_{n_r} \otimes U_{ij}^T & I_{n_r} \otimes U_{jj}^T \end{bmatrix}, \quad j+1 \leq r \leq m, \end{aligned}$$

where

$$\begin{aligned} S &= \begin{bmatrix} I_{n_i} \otimes e_1^T \\ \vdots \\ I_{n_i} \otimes e_{n_r}^T \end{bmatrix} = [I_{n_r} \otimes \bar{e}_1 \quad \dots \quad I_{n_r} \otimes \bar{e}_{n_i}], \\ \tilde{S} &= \begin{bmatrix} I_{n_r} \otimes \tilde{e}_1^T \\ \vdots \\ I_{n_r} \otimes \tilde{e}_{n_j}^T \end{bmatrix} = [I_{n_j} \otimes e_1 \quad \dots \quad I_{n_j} \otimes e_{n_r}]. \end{aligned}$$

Here,  $e_i$ ,  $\bar{e}_i$  and  $\tilde{e}_i$  denote the  $i$ th column of  $I_{n_r}$ ,  $I_{n_i}$  and  $I_{n_j}$ , respectively.

The matrix  $\mathbf{R}$  satisfies  $\|\mathbf{R}\|_2 = 1$ , except in the case  $m = 2$ ,  $(i, j) = (1, 2)$ , when  $\mathbf{R} = 0$ .



If the spectral norm is used instead of the spectral radius, then we have the following result.

**PROPOSITION 2.20.** *Let  $\pi = (n_1, \dots, n_m)$  be a partition of  $n$ ,  $\mathcal{O}, \mathcal{O}' \in \mathcal{O}(\mathcal{P}_m)$ , and  $\mathcal{O}' \stackrel{w}{\sim} \mathcal{O}$ . Let  $\mathcal{O}$  and  $\mathcal{O}'$  be linked by the chain  $\mathcal{O} = \mathcal{O}_0, \mathcal{O}_1, \dots, \mathcal{O}_r = \mathcal{O}'$  as in relation (2.7). Suppose that in the chain there are exactly  $d$  pairs of neighboring terms that are shift equivalent. If*

$$\|\mathcal{J}\|_2 \leq \mu_{\pi, \varrho} \quad \text{for all } \mathcal{J} \in \mathcal{J}_{\mathcal{O}}^{\text{UBCE}(\varrho)},$$

then for any  $d + 1$  block Jacobi operators from  $\mathcal{J}_{\mathcal{O}'}^{\text{UBCE}(\varrho)}$  one has

$$\|\mathcal{J}'_1 \mathcal{J}'_2 \cdots \mathcal{J}'_{d+1}\|_2 \leq \mu_{\pi, \varrho}, \quad \mathcal{J}'_1, \dots, \mathcal{J}'_{d+1} \in \mathcal{J}_{\mathcal{O}'}^{\text{UBCE}(\varrho)}.$$

The constant  $\mu_{\pi, \varrho}$  may depend only on  $\pi$  and  $\varrho$ .

*Proof.* The proof is the same as the proof of [18, Lemma 4.8(ii)]. The role of the set  $\Psi_\pi$  from [18, Lemma 4.8(ii)] is played by the set  $\bigcup_{i < j} \mathcal{R}_{ij}^{\text{UBCE}(\varrho)}$ .  $\square$

**PROPOSITION 2.21.** *Let  $\pi = (n_1, \dots, n_m)$  be a partition of  $n$ ,  $\mathcal{O} \in \mathcal{O}(\mathcal{P}_m)$ , and suppose  $\|\mathcal{J}\|_2 \leq \mu_{\pi, \varrho}$  for all  $\mathcal{J} \in \mathcal{J}_{\mathcal{O}}^{\text{UBCE}(\varrho)}$ , where  $\mu_{\varrho}$  depends on  $\pi$  and  $\varrho$ . Then*

$$\|\tilde{\mathcal{J}}\|_2 \leq \mu_{\pi, \varrho} \quad \text{for all } \tilde{\mathcal{J}} \in \mathcal{J}_{\mathcal{O}^{\leftarrow}}^{\text{UBCE}(\varrho)}.$$

The assertion holds provided that in both appearances the spectral norm is replaced by the spectral radius.

*Proof.* Suppose that  $\mathcal{O} = (i_0, j_0), (i_1, j_1), \dots, (i_{T-1}, j_{T-1}) \in \mathcal{O}(\mathcal{P}_m)$ , and let  $\tilde{\mathcal{J}} \in \mathcal{J}_{\mathcal{O}^{\leftarrow}}^{\text{UBCE}(\varrho)}$  be arbitrary. Then

$$\tilde{\mathcal{J}} = \mathcal{R}_{i(0), j(0)}(\hat{U}_0) \mathcal{R}_{i(1), j(1)}(\hat{U}_1) \cdots \mathcal{R}_{i(T-1), j(T-1)}(\hat{U}_{T-1}),$$

for some orthogonal  $\text{UBCE}(\varrho)$  matrices  $\hat{U}_k$ ,  $0 \leq k \leq T - 1$ , of appropriate sizes.

If we show that  $\tilde{\mathcal{J}}^T \in \mathcal{J}_{\mathcal{O}}^{\text{UBCE}(\varrho)}$ , the first claim will follow from  $\|\tilde{\mathcal{J}}\|_2 = \|\tilde{\mathcal{J}}^T\|_2 \leq \mu_{\pi, \varrho}$ , while the part about the spectral radius will be a consequence of  $\text{spr}(\tilde{\mathcal{J}}) = \text{spr}(\tilde{\mathcal{J}}^T) \leq \mu_{\pi, \varrho}$ . Note that

$$\begin{aligned} \tilde{\mathcal{J}}^T &= \left[ \mathcal{R}_{i(0), j(0)}(\hat{U}_0) \mathcal{R}_{i(1), j(1)}(\hat{U}_1) \cdots \mathcal{R}_{i(T-1), j(T-1)}(\hat{U}_{T-1}) \right]^T \\ &= [\mathcal{R}_{i(T-1), j(T-1)}(\hat{U}_{T-1})]^T \cdots [\mathcal{R}_{i(1), j(1)}(\hat{U}_1)]^T \cdot [\mathcal{R}_{i(0), j(0)}(\hat{U}_0)]^T. \end{aligned}$$

By Corollary 2.17 we know that  $[\mathcal{R}_{ij}(\hat{U}_k)]^T \in \mathcal{R}_{ij}^{\text{UBCE}(\varrho)}$ , hence  $\tilde{\mathcal{J}}^T \in \mathcal{J}_{\mathcal{O}}^{\text{UBCE}(\varrho)}$ .  $\square$

**2.4.1. Permutation equivalence and the block Jacobi operators.** Here we derive a similar result for the block Jacobi operators  $\mathcal{J}_{\mathcal{O}}$  and  $\mathcal{J}_{\mathcal{O}(\mathfrak{p})}$ .

**THEOREM 2.22.** *Let  $\pi = (n_1, \dots, n_m)$  be a partition of  $n$ , and take  $\mathcal{O} \in \mathcal{O}(\mathcal{P}_m)$ . Let  $\mathfrak{p}$  be a permutation of  $\mathcal{S}_m$  and set  $\tilde{\mathcal{O}} = \mathcal{O}(\mathfrak{p})$ .*

(i) *If  $\|\mathcal{J}\|_2 \leq \mu_{\pi, \varrho}$  for all  $\mathcal{J} \in \mathcal{J}_{\mathcal{O}}^{\text{UBCE}_\pi(\varrho)}$ , where  $\mu_{\pi, \varrho}$  only depends on  $\pi$  and  $\varrho$ , then*

$$\|\tilde{\mathcal{J}}\|_2 \leq \mu_{\pi, \varrho} \quad \text{for any } \tilde{\mathcal{J}} \in \mathcal{J}_{\tilde{\mathcal{O}}}^{\text{UBCE}_{\pi \mathfrak{p}}(\varrho)}.$$



(ii) If  $\|\mathcal{J}_1 \mathcal{J}_2 \cdots \mathcal{J}_{d+1}\|_2 \leq \mu_{\pi, \varrho}$ , for all  $\mathcal{J}_1, \dots, \mathcal{J}_{d+1} \in \mathcal{J}_{\mathcal{O}}^{\text{UBCE}_{\pi}(\varrho)}$ , where  $\mu_{\pi, \varrho}$  only depends on  $\pi$  and  $\varrho$ , then

$$\|\tilde{\mathcal{J}}_1 \tilde{\mathcal{J}}_2 \cdots \tilde{\mathcal{J}}_{d+1}\|_2 \leq \mu_{\pi, \varrho} \quad \text{for any } \tilde{\mathcal{J}}_1, \dots, \tilde{\mathcal{J}}_{d+1} \in \mathcal{J}_{\tilde{\mathcal{O}}}^{\text{UBCE}_{\pi_p}(\varrho)}.$$

The assertions also hold provided that every appearance of the spectral norm is replaced by the spectral radius.

*Proof.* (i) Let  $\mathcal{O} = (i_0, j_0), (i_1, j_1), \dots, (i_{T-1}, j_{T-1})$  so that  $\tilde{\mathcal{O}} = (\mathfrak{p}(i_0), \mathfrak{p}(j_0)), (\mathfrak{p}(i_1), \mathfrak{p}(j_1)), \dots, (\mathfrak{p}(i_{T-1}), \mathfrak{p}(j_{T-1}))$ . Let  $\tilde{\mathcal{J}} \in \mathcal{J}_{\tilde{\mathcal{O}}}^{\text{UBCE}_{\pi_p}(\varrho)}$  be arbitrary. Then

$$(2.20) \quad \tilde{\mathcal{J}} = \mathcal{R}_{\mathfrak{p}(i_{T-1}), \mathfrak{p}(j_{T-1})}(\hat{U}_{T-1}) \cdots \mathcal{R}_{\mathfrak{p}(i_1), \mathfrak{p}(j_1)}(\hat{U}_1) \mathcal{R}_{\mathfrak{p}(i_0), \mathfrak{p}(j_0)}(\hat{U}_0),$$

where  $\hat{U}_k$  is an orthogonal  $\text{UBCE}_{\pi_p}(\varrho)$  matrix of order  $n_{\mathfrak{p}(i_k)} + n_{\mathfrak{p}(j_k)}$  for each  $0 \leq k \leq T-1$ .

Let  $a \in \mathbb{R}^K$  be an arbitrary nonzero vector, and consider the computation of  $a' = \tilde{\mathcal{J}}a$ . Using Algorithm 2.14, the vector  $a'$  can be obtained by the following procedure:

- Compute the symmetric matrix  $A^{(0)} \equiv A = \text{vec}_{\pi_p, 0}^{-1}(a)$ .
- Recursively compute:  $A^{(k+1)} = \mathcal{N}_{\mathfrak{p}(i)\mathfrak{p}(j)}(U_k^T A^{(k)} U_k)$ ,  $k = 0, 1, \dots, T-1$ .
- Compute the vector  $a' = \text{vec}_{\pi_p}(A^{(T)})$ .

Here  $U_k = \mathcal{E}(\mathfrak{p}(i_k), \mathfrak{p}(j_k), \hat{U}_k)$ ,  $0 \leq k \leq T-1$ , and the matrices  $A$  and  $A^{(k)}$  carry the matrix block-partition defined by  $\pi_p$ . Let  $P$  be the matrix from relation (2.12), which is defined by  $P e_t = e_{p(t)}$ ,  $1 \leq t \leq n$ , where  $p$  is defined as in Section 2.2.1, and  $\mathfrak{p}$  is from the statement of this lemma. Let  $X$  be a square matrix of order  $n$ , partitioned in accordance with  $\pi_p$ . Then for any  $1 \leq s, t \leq m$ , the transformation  $X \mapsto P^T X P$  changes the partition from  $\pi_p$  to  $\pi$  and moves the block  $X_{\mathfrak{p}(s)\mathfrak{p}(t)}$  to the  $(s, t)$  position. Therefore, we have

$$(2.21) \quad P^T A^{(k+1)} P = P^T \mathcal{N}_{\mathfrak{p}(i)\mathfrak{p}(j)}(U_k^T A^{(k)} U_k) P = \mathcal{N}_{ij}(P^T U_k^T A^{(k)} U_k P) \\ = \mathcal{N}_{ij}\left((P^T U_k P)^T (P^T A^{(k)} P) (P^T U_k P)\right), \quad k = 0, 1, \dots, T-1.$$

If we set  $\bar{A}^{(k)} = P^T A^{(k)} P$ ,  $\bar{A} = P^T A P$ , and  $\bar{U}^{(k)} = P^T U^{(k)} P$ , then the recurrence (2.21) takes the form

$$(2.22) \quad \bar{A}^{(k+1)} = \mathcal{N}_{ij}\left([\bar{U}^{(k)}]^T \bar{A}^{(k)} \bar{U}^{(k)}\right), \quad k = 0, 1, \dots, T-1, \quad \bar{A}^{(0)} = \bar{A}.$$

Obviously, we have  $\text{vec}_{\pi}(\bar{A}) = \text{vec}_{\pi}(P^T A P) = \mathbf{P}a$  for some permutation matrix  $\mathbf{P}$  of order  $K$ . Applying the  $\text{vec}$  function to relation (2.22), one obtains

$$(2.23) \quad \bar{a}^{(k+1)} = \mathcal{R}_{i(k), j(k)}(\hat{U}_k) \bar{a}^{(k)}, \quad k = 0, 1, \dots, T-1, \quad \bar{a}^{(0)} = \mathbf{P}a,$$

where  $\mathcal{R}_{ij}(\hat{U}_k) \in \mathcal{R}_{ij}^{\text{UBCE}_{\pi}(\varrho)}$ . This process is associated with the sequence  $\mathcal{O}$  and results in the final form

$$\mathbf{P}a' = \bar{\mathcal{J}}_{\mathcal{O}} \mathbf{P}a, \quad \bar{\mathcal{J}}_{\mathcal{O}} \in \mathcal{J}_{\mathcal{O}}^{\text{UBCE}_{\pi}(\varrho)}.$$

Because  $a$  is an arbitrary vector, we have  $\tilde{\mathcal{J}} = \mathbf{P}^T \bar{\mathcal{J}}_{\mathcal{O}} \mathbf{P}$ . This implies  $\|\tilde{\mathcal{J}}\|_2 = \|\bar{\mathcal{J}}_{\mathcal{O}}\|_2 \leq \mu_{\pi, \varrho}$ . Since,  $\tilde{\mathcal{J}}$  and  $\bar{\mathcal{J}}_{\mathcal{O}}$  are similar, their spectral radius is the same.

(ii) The proof is similar to the proof of (i). We start our consideration with an arbitrary nonzero vector  $a \in \mathbb{R}^K$  and consider the computation of  $a' = \tilde{\mathcal{J}}_1 \tilde{\mathcal{J}}_2 \cdots \tilde{\mathcal{J}}_{d+1} a$ . Since we have  $d+1$  block Jacobi operators, we will use altogether  $(d+1)T$  block Jacobi annihilators. All we have to change in the proof of (i) is the the range of the index  $k$  in the relations (2.20)–(2.23): instead of  $T-1$ , its largest value will be  $(d+1)T-1$ . At the end we have  $\tilde{\mathcal{J}}_1 \tilde{\mathcal{J}}_2 \cdots \tilde{\mathcal{J}}_{d+1} = \mathbf{P}^T \mathcal{J}_1 \mathcal{J}_2 \cdots \mathcal{J}_{d+1} \mathbf{P}$ , and the conclusion follows.  $\square$

**3. Generalized serial strategies.** The aim of this section is to significantly enlarge the class of the known “convergent” cyclic pivot strategies, namely the serial ones and those that are weakly equivalent to them (the so-called weak wavefront strategies [32]). We study several classes of cyclic pivot strategies, which are generalizations of the serial ones. The first (resp. second) of those classes is defined by the set  $\mathcal{B}_c^{(m)}$  (resp.  $\mathcal{B}_r^{(m)}$ ) of pivot orderings which arise from column-wise (resp. row-wise) orderings of  $\mathcal{P}_m$ . The other two are defined by the first two using reverse orderings. Once the global convergence of the block Jacobi method under these strategies is proved, one can easily expand the obtained set of pivot strategies using the theory of equivalent strategies.

**3.1. The class  $\mathcal{B}_c^{(m)}$ .** We start with the class of cyclic strategies that choose the  $(1, 2)$ -block as the first pivot block, then choose all blocks from the second block-column in some order, etc. At the last stage they choose all blocks from the last block-column in some order. For the precise definition of that class, we denote the set of all permutations of the set  $\{l_1, l_1 + 1, \dots, l_2\}$  by  $\Pi^{(l_1, l_2)}$ . Let

$$(3.1) \quad \mathcal{B}_c^{(m)} = \left\{ \mathcal{O} \in \mathcal{O}(\mathcal{P}_m) \mid \mathcal{O} = (1, 2), (\tau_3(1), 3), (\tau_3(2), 3), \dots, (\tau_m(1), m), \dots, \right. \\ \left. (\tau_m(m-1), m), \quad \tau_j \in \Pi^{(1, j-1)}, 3 \leq j \leq m \right\}.$$

The set  $\mathcal{B}_c^{(m)}$  is a part of the class of *column-wise orderings with permutations* of the set  $\mathcal{P}_m$ , which will be described in Definition 3.4. A typical ordering  $\mathcal{O} \in \mathcal{B}_c^{(6)}$  is represented by  $M_{\mathcal{O}}$  below. The second matrix  $M_{\tilde{\mathcal{O}}}$  is defined by some  $\tilde{\mathcal{O}} \stackrel{w}{\sim} \mathcal{O}$ . Its purpose is to see how far from the “serial structure” this equivalence can push  $\mathcal{O}$ .

$$M_{\mathcal{O}} = \begin{bmatrix} * & 0 & 2 & 4 & 9 & 12 \\ 0 & * & 1 & 5 & 8 & 10 \\ 2 & 1 & * & 3 & 7 & 13 \\ 4 & 5 & 3 & * & 6 & 11 \\ 9 & 8 & 7 & 6 & * & 14 \\ 12 & 10 & 13 & 11 & 14 & * \end{bmatrix}, \quad M_{\tilde{\mathcal{O}}} = \begin{bmatrix} * & 7 & 9 & 0 & 2 & 5 \\ 7 & * & 10 & 13 & 14 & 6 \\ 9 & 10 & * & 11 & 12 & 8 \\ 0 & 13 & 11 & * & 1 & 4 \\ 2 & 14 & 12 & 1 & * & 3 \\ 5 & 6 & 8 & 4 & 3 & * \end{bmatrix}.$$

From  $M_{\mathcal{O}}$  we can see that the permutation  $\tau_j$  from (3.1) is linked to the  $j$ th block-column of the matrix. The next theorem proves the global convergence of the block Jacobi method under the cyclic strategies  $I_{\mathcal{O}}$  defined by the orderings  $\mathcal{O} \in \mathcal{B}_c^{(m)}$ .

**THEOREM 3.1.** *Let  $\pi = (n_1, \dots, n_m)$  be a partition of  $n$ , and let  $\mathcal{O} \in \mathcal{B}_c^{(m)}$ . Let  $A \in \mathbf{S}_n$  be partitioned as in relation (2.2). Suppose that  $A'$  is obtained from  $A$  by applying one sweep of the cyclic block Jacobi method defined by the strategy  $I_{\mathcal{O}}$ . If all transformation matrices are from the class  $UBCE(\varrho)$ , then there are constants  $\eta_{\pi, \varrho}$  (depending only on  $\pi$  and  $\varrho$ ) and  $\tilde{\eta}_{n, \varrho}$  (depending only on  $n$  and  $\varrho$ ) such that*

$$S^2(A') \leq \eta_{\pi, \varrho} S^2(A), \quad 0 \leq \eta_{\pi, \varrho} < \tilde{\eta}_{n, \varrho} < 1.$$

*Proof.* The proof is lengthy and has been moved to Appendix A.  $\square$

We have to explain why we use the two bounds satisfying  $\mu_{\pi, \varrho} < \tilde{\mu}_{n, \varrho} < 1$ . Recall that each block Jacobi method is defined by some partition  $\pi$  of  $n$ . Different partitions define different block Jacobi methods even in the case when the pivot orderings are the same. The second bound  $\tilde{\mu}_{n, \varrho}$  can be used in the global convergence statements for the block Jacobi method when the order  $n$  of the initial matrix is known, while about the pivot ordering it is only known that it belongs to the set  $\bigcup_{3 \leq m \leq n} \mathcal{B}_c^{(m)}$ . It means that, for a given  $m$ , the convergence result holds for the block Jacobi method defined by any  $\pi$  such that  $n_1 + \dots + n_m = n$ .

Combining Theorem 3.1 with Theorem 2.11, we see that we can enlarge the class of “convergent orderings” from  $\mathcal{B}_c^{(m)}$  to the class of all orderings that are weakly equivalent to orderings from  $\mathcal{B}_c^{(m)}$ . Thus, the ordering  $\tilde{\mathcal{O}}$  linked to the above matrix  $M_{\tilde{\mathcal{O}}}$  is also convergent.

The next result is a slight generalization of Theorem 3.1, and it deals with the block Jacobi operators. The role of the block Jacobi operators will be explained in Section 5, especially by inspecting the proof of Theorem 5.1.

**THEOREM 3.2.** *Let  $\pi = (n_1, \dots, n_m)$  be a partition of  $n$ . Take  $\mathcal{O} \in \mathcal{B}_c^{(m)}$ , and let  $\mathcal{J} \in \mathcal{J}_{\mathcal{O}}^{UBCE(\varrho)}$  be a block Jacobi operator. Then there are constants  $\mu_{\pi, \varrho}$  and  $\tilde{\mu}_{n, \varrho}$  depending only on  $\pi$ ,  $\varrho$  and  $n$ ,  $\varrho$ , respectively, such that*

$$\|\mathcal{J}\|_2 \leq \mu_{\pi, \varrho}, \quad 0 \leq \mu_{\pi, \varrho} < \tilde{\mu}_{n, \varrho} < 1.$$

*Proof.* Let  $a \in \mathbb{R}^K$  be an arbitrary nonzero vector and let  $a' = \mathcal{J}a$ . To track how  $a'$  is obtained from  $a$ , we can assume  $\mathcal{J} = \mathcal{R}_{i_{M-1}j_{M-1}} \mathcal{R}_{i_{M-2}j_{M-2}} \dots \mathcal{R}_{i_0j_0}$ , where  $\mathcal{O} = (i_0, j_0), \dots, (i_{M-1}, j_{M-1})$ . If we define

$$(3.2) \quad a^{(k+1)} = \mathcal{R}_{i_k j_k} a^{(k)}, \quad 0 \leq k \leq M-1, \quad a^{(0)} = a,$$

then we obtain  $a' = a^{(M)}$ . Recall that Algorithm 2.14 describes the  $k$ th step of the process (3.2), i.e., how the vector  $a^{(k+1)}$  is obtained from  $a^{(k)}$ . This algorithm computes the matrix  $A^{(k+1)} = \text{vec}_0^{-1}(a^{(k+1)})$  from the matrix  $A^{(k)} = \text{vec}_0^{-1}(a^{(k)})$ . Note that

$$S^2(A) = \|a\|_2^2 \quad \text{and} \quad S^2(A') = \|a'\|_2^2 = \|\mathcal{J}a\|_2^2.$$

If we prove

$$(3.3) \quad S^2(A') \leq \eta_{\pi, \varrho} S^2(A), \quad \eta_{\pi, \varrho} < \tilde{\eta}_{n, \varrho} < 1,$$

and take into account that  $a$  is an arbitrary nonzero vector, then we will straightforwardly obtain

$$\|\mathcal{J}\|_2 = \max_{a \neq 0} \frac{\|\mathcal{J}a\|_2}{\|a\|_2} \leq \mu_{\pi, \varrho} < \tilde{\mu}_{n, \varrho}, \quad \eta_{\pi, \varrho} = \sqrt{\mu_{\pi, \varrho}}, \quad \tilde{\eta}_{n, \varrho} = \sqrt{\tilde{\mu}_{n, \varrho}}.$$

To prove (3.3), we can rely on the proof of Theorem 3.1. Indeed, let us compare the computation of the matrix  $A^{(k+1)}$  from  $A^{(k)}$  using Algorithm 2.14 with the  $k$ th step of the block Jacobi method. If we neglect the diagonal blocks, then both amount to the same procedure except for the fact that the block Jacobi method actually computes the orthogonal elementary matrix which diagonalizes the pivot submatrix, while in process (3.2) that transformation is given via the matrix  $\mathcal{R}_{i_k j_k} = \mathcal{R}_{i_k j_k}(\hat{U}_k)$ . The two procedures will naturally generate different iteration matrices, but all estimates and the whole proof will be the same. The quantity  $\zeta_l$  from relation (A.10) will be different for the two procedures, but all that is needed for the proof is that  $\zeta_l$  is uniformly bounded from below by some positive constant, which is certainly satisfied.  $\square$

In the special case when  $\pi = (1, 1, \dots, 1)$ , all blocks are  $1 \times 1$  matrices, i.e., single elements, so the block method reduces to the standard Jacobi method. In this case we will denote the class  $\mathcal{B}_c^{(n)}$  by  $\mathcal{C}_c^{(n)}$ . Theorem 3.1 then reduces to the following corollary.

**COROLLARY 3.3.** *Let  $A \in \mathbf{S}_n$ ,  $\mathcal{O} \in \mathcal{C}_c^{(n)}$ , and let  $A'$  be obtained from  $A$  by applying one sweep of the cyclic Jacobi method defined by the strategy  $I_{\mathcal{O}}$ , with rotation angles from the interval  $[-\frac{\pi}{4}, \frac{\pi}{4}]$ . Then there is a constant  $\eta_n$  depending only on  $n$  such that*

$$S^2(A') \leq \eta_n S^2(A), \quad 0 \leq \eta_n < 1.$$

*Proof.* The proof has been moved to Appendix A.  $\square$

In this special case, the class  $\mathcal{C}_c^{(n)}$  is a subset of the set of Nazareth's orderings from [28]. However, note that the bounds obtained here are much better than those in [28]. To illustrate that, observe that for  $n = 3$  (resp.  $n = 4$ ) the value of  $\eta_n$  is equal to  $\max\{\frac{3}{4}, \frac{3}{4}\} = \frac{3}{4}$  (resp.  $\max\{\frac{7}{8}, \frac{27}{28}\} = \frac{27}{28}$ ). In [28] the corresponding bounds are  $1 - 1/(3 \cdot 2^{104})$  and  $1 - 1/(6 \cdot 2^{294})$ . This comparison has also been studied in [1, page 57].

**3.2. The classes  $\mathcal{B}_{cp}^{(m)}$ ,  $\mathcal{B}_{rp}^{(m)}$ , and  $\mathcal{B}_{sp}^{(m)}$ .** The same results hold for the class of cyclic pivot strategies which take the pivot blocks from the block-rows. Let

$$\mathcal{B}_r^{(m)} = \left\{ \mathcal{O} \in \mathcal{O}(\mathcal{P}_m) \mid \mathcal{O} = (m-1, m), (m-2, \tilde{\tau}_{m-2}(m-1)), (m-2, \tilde{\tau}_{m-2}(m)), \dots, (1, \tilde{\tau}_1(2)), \dots, (1, \tilde{\tau}_1(m)), \tilde{\tau}_i \in \Pi^{(i+1, m)}, 1 \leq i \leq m-2 \right\}.$$

The set  $\mathcal{B}_r^{(m)}$  is a part of the class of *row-wise orderings with permutations* of the set  $\mathcal{P}_m$  (see Definition 3.4 below). A typical ordering from  $\mathcal{O} \in \mathcal{B}_r^{(6)}$  is represented by  $M_{\mathcal{O}}$  below. The second matrix  $M_{\tilde{\mathcal{O}}}$  is defined by  $\tilde{\mathcal{O}} \stackrel{w}{\sim} \mathcal{O}$ .

$$M_{\mathcal{O}} = \begin{bmatrix} * & 11 & 13 & 12 & 10 & 14 \\ 10 & * & 9 & 7 & 6 & 8 \\ 11 & 9 & * & 5 & 3 & 4 \\ 12 & 6 & 5 & * & 1 & 2 \\ 13 & 7 & 3 & 1 & * & 0 \\ 14 & 8 & 4 & 2 & 0 & * \end{bmatrix}, \quad M_{\tilde{\mathcal{O}}} = \begin{bmatrix} * & 14 & 1 & 0 & 11 & 2 \\ 14 & * & 13 & 10 & 7 & 12 \\ 1 & 13 & * & 9 & 6 & 8 \\ 0 & 10 & 9 & * & 4 & 5 \\ 11 & 7 & 6 & 4 & * & 3 \\ 2 & 12 & 8 & 5 & 3 & * \end{bmatrix}.$$

From the matrix  $M_{\mathcal{O}}$  we can see that the permutation  $\tilde{\tau}_i$  from (3.1) is linked to the  $i$ th block-row of the matrix. It is immediately clear that

$$(3.4) \quad \mathcal{B}_r^{(m)} = \{ \mathcal{O}(\tilde{\epsilon}) \mid \mathcal{O} \in \mathcal{B}_c^{(m)} \},$$

where  $\tilde{\epsilon}$  is defined by relation (2.14). Theorems 3.1 and 3.2 remain to hold for  $\mathcal{O} \in \mathcal{B}_r^{(m)}$ . The proofs are almost identical to those for the case  $\mathcal{O} \in \mathcal{B}_c^{(m)}$ . The version of Theorem 3.2 for  $\mathcal{O} \in \mathcal{B}_r^{(m)}$  follows directly from the original Theorem 3.2 combined with Theorem 2.22(i) (when the permutation  $p$  is specified to be  $\tilde{\epsilon}$ ). Another way how Theorem 3.1 can be established for  $\mathcal{O} \in \mathcal{B}_r^{(m)}$  is to follow the lines of the proof of Theorem 3.5 (below), but then a version of Theorem 3.2 for  $\mathcal{O} \in \mathcal{B}_r^{(m)}$  should be proved first. If  $\pi = (1, 1, \dots, 1)$ , then the class of orderings  $\mathcal{B}_r^{(n)}$  is denoted by  $\mathcal{C}_r^{(n)}$ . Corollary 3.3 holds with the same constant  $\eta_n$  provided that  $\mathcal{C}_c^{(n)}$  is replaced by  $\mathcal{C}_r^{(n)}$ .

We are interested in two more classes of pivot strategies for the block methods. The first (resp. second) one selects the pivot blocks by block-columns (resp. block-rows), but now from the last one to the second one (resp. from the first one to the next-to-last one). They are defined as

$$\overleftarrow{\mathcal{B}}_c^{(m)} = \{ \mathcal{O} \in \mathcal{O}(\mathcal{P}_m) \mid \mathcal{O}^{\leftarrow} \in \mathcal{B}_c^{(m)} \}, \quad \overleftarrow{\mathcal{B}}_r^{(m)} = \{ \mathcal{O} \in \mathcal{O}(\mathcal{P}_m) \mid \mathcal{O}^{\leftarrow} \in \mathcal{B}_r^{(m)} \}.$$

Typical orderings from  $\overleftarrow{\mathcal{B}}_c^{(6)}$  and  $\overleftarrow{\mathcal{B}}_r^{(6)}$  are represented by  $M_{\overleftarrow{\mathcal{O}}_c}$  and  $M_{\overleftarrow{\mathcal{O}}_r}$  below.

$$M_{\overleftarrow{\mathcal{O}}_c} = \begin{bmatrix} * & 14 & 12 & 10 & 5 & 2 \\ 14 & * & 13 & 9 & 6 & 4 \\ 12 & 13 & * & 11 & 7 & 1 \\ 10 & 9 & 11 & * & 8 & 3 \\ 5 & 6 & 7 & 8 & * & 0 \\ 2 & 4 & 1 & 3 & 0 & * \end{bmatrix}, \quad M_{\overleftarrow{\mathcal{O}}_r} = \begin{bmatrix} * & 4 & 3 & 2 & 1 & 0 \\ 4 & * & 5 & 8 & 7 & 6 \\ 3 & 5 & * & 9 & 11 & 10 \\ 2 & 8 & 9 & * & 13 & 12 \\ 1 & 7 & 11 & 13 & * & 14 \\ 0 & 6 & 10 & 12 & 14 & * \end{bmatrix}.$$

As it has already been noticed at the end of Section 2.2, Lemma 2.3 remains to hold if  $\tilde{\sim}$  is replaced with  $\overset{\leftarrow}{\sim}$ . Therefore, relation (3.4) implies  $\overset{\leftarrow}{\mathcal{B}}_r^{(m)} = \{\mathcal{O}(\tilde{e}) \mid \mathcal{O} \in \overset{\leftarrow}{\mathcal{B}}_c^{(m)}\}$ .

**DEFINITION 3.4.** Let  $\mathcal{B}_{cp}^{(m)} = \mathcal{B}_c^{(m)} \cup \overset{\leftarrow}{\mathcal{B}}_c^{(m)}$ ,  $\mathcal{B}_{rp}^{(m)} = \mathcal{B}_r^{(m)} \cup \overset{\leftarrow}{\mathcal{B}}_r^{(m)}$ . The set  $\mathcal{B}_{cp}^{(m)}$  (resp.  $\mathcal{B}_{rp}^{(m)}$ ) is the class of the column-wise (resp. row-wise) orderings with permutations of  $\mathcal{P}_m$ , and  $\{I_{\mathcal{O}} \mid \mathcal{O} \in \mathcal{B}_{cp}^{(m)}\}$  (resp.  $\{I_{\mathcal{O}} \mid \mathcal{O} \in \mathcal{B}_{rp}^{(m)}\}$ ) is the class of the column-cyclic (resp. row-cyclic) strategies with permutations.

The set  $\mathcal{B}_{sp}^{(m)} = \mathcal{B}_{cp}^{(m)} \cup \mathcal{B}_{rp}^{(m)}$  is the class of the serial orderings with permutations of  $\mathcal{P}_m$ , and  $\{I_{\mathcal{O}} \mid \mathcal{O} \in \mathcal{B}_{sp}^{(m)}\}$  is the class of the serial strategies with permutations.

**THEOREM 3.5.** Let  $\pi = (n_1, \dots, n_m)$  be a partition of  $n$ ,  $\mathcal{O} \in \mathcal{B}_{sp}^{(m)}$ , and let  $\mathcal{J} \in \mathcal{J}_{\mathcal{O}}^{UBCE(\varrho)}$  be a block Jacobi operator. Then there are constants  $\mu_{\pi, \varrho}$  and  $\tilde{\mu}_{n, \varrho}$  depending only on  $\pi$ ,  $\varrho$  and  $n$ ,  $\varrho$ , respectively, such that

$$\|\mathcal{J}\|_2 \leq \mu_{\pi, \varrho}, \quad 0 \leq \mu_{\pi, \varrho} < \tilde{\mu}_{n, \varrho} < 1.$$

*Proof.* If  $\mathcal{O} \in \mathcal{B}_c^{(m)}$  (resp.  $\mathcal{O} \in \mathcal{B}_r^{(m)}$ ), then the theorem reduces to Theorem 3.2 (resp. Theorem 3.2 combined with Theorem 2.22(i)). If  $\mathcal{O} \in \overset{\leftarrow}{\mathcal{B}}_c^{(m)}$  (resp.  $\mathcal{O} \in \overset{\leftarrow}{\mathcal{B}}_r^{(m)}$ ), then additionally Proposition 2.21 is used.  $\square$

The following result is a corollary of Theorem 3.5, but because of its importance, it is stated as a stand-alone result.

**THEOREM 3.6.** Let  $\pi = (n_1, \dots, n_m)$  be a partition of  $n$ ,  $\mathcal{O} \in \mathcal{B}_{sp}^{(m)}$ , and let  $A \in \mathbf{S}_n$  be partitioned as in relation (2.2). Let  $A'$  be obtained from  $A$  by applying one sweep of the cyclic block Jacobi method defined by the strategy  $I_{\mathcal{O}}$ . If all transformation matrices are from the class  $UBCE(\varrho)$ , then there are constants  $\eta_{\pi, \varrho}$  (depending only on  $\pi$ ,  $\varrho$ ) and  $\tilde{\eta}_{n, \varrho}$  (depending only on  $n$ ,  $\varrho$ ) such that

$$S^2(A') \leq \eta_{\pi, \varrho} S^2(A), \quad 0 \leq \eta_{\pi, \varrho} < \tilde{\eta}_{n, \varrho} < 1.$$

*Proof.* Let  $\mathcal{O} = (i_0, j_0), (i_1, j_1), \dots, (i_{M-1}, j_{M-1})$ . The method in the statement of the theorem generates the recurrence relation of the form (2.5). If we observe how the elements in the block upper-triangle are being updated, we arrive at the recursion

$$a^{(k+1)} = \mathcal{R}_{ij} a^{(k)}, \quad k \geq 0, \quad a^{(0)} = a = \text{vec}(A).$$

Here, for each  $k$ ,  $a^{(k)} = \text{vec}(A^{(k)}) \in \mathbb{R}^K$ , and  $\mathcal{R}_{ij} = \mathcal{R}_{ij}(\hat{U}_k)$  is the block Jacobi annihilator associated with step  $k$  of the method. We have  $(i, j) = (i(k), j(k)) = (i_k, j_k)$  for  $0 \leq k \leq M-1$ . After the first sweep is completed, one obtains

$$(3.5) \quad a^{(M)} = \mathcal{J}_{\mathcal{O}} a, \quad \mathcal{J}_{\mathcal{O}} = \mathcal{R}_{i(M-1)j(M-1)} \cdots \mathcal{R}_{i(0)j(0)}.$$

Since all transformation matrices in the block Jacobi method are from the class  $UBCE(\varrho)$ , we have  $\mathcal{R}_{ij}(\hat{U}_k) \in \mathcal{R}_{ij}^{UBCE(\varrho)}$ . Therefore,  $\mathcal{J}_{\mathcal{O}} \in \mathcal{J}_{\mathcal{O}}^{UBCE(\varrho)}$ , and by applying Theorem 3.5 one obtains  $\|\mathcal{J}\|_2 \leq \mu_{\pi, \varrho}$ ,  $\mu_{\pi, \varrho} < \tilde{\mu}_{n, \varrho} < 1$ . Hence, if one takes the Euclidean vector norm of both sides of the left equation in (3.5), it follows that

$$S^2(A^{(M)}) = \|a^{(M)}\|_2^2 \leq \mu_{\pi, \varrho}^2 \|a\|_2^2 = \mu_{\pi, \varrho}^2 S^2(A).$$

It remains to set  $\eta_{\pi, \varrho} = \mu_{\pi, \varrho}^2$  and  $\tilde{\eta}_{n, \varrho} = \tilde{\mu}_{n, \varrho}^2$ .  $\square$

Obviously, since the assertion of Theorem 3.6 holds for any single sweep, we conclude that  $S(A^{(tM)}) \rightarrow 0$  as  $t \rightarrow \infty$ . Since the sequence  $S(A^{(k)})$ ,  $k \geq 0$ , is nonincreasing, one obtains  $S(A^{(k)}) \rightarrow 0$  as  $k \rightarrow \infty$ . Together with Theorem 2.10 this implies global convergence.

In the case  $\pi = (1, 1, \dots, 1)$  we have  $m = n$ . We can write  $\mathcal{C}_{cp}^{(n)}$ ,  $\mathcal{C}_{rp}^{(n)}$ , and  $\mathcal{C}_{sp}^{(n)}$  in places of  $\mathcal{B}_{cp}^{(m)}$ ,  $\mathcal{B}_{rp}^{(m)}$ , and  $\mathcal{B}_{sp}^{(m)}$ , respectively. Once again, Corollary 3.3 holds with the same constant  $\eta_n$  if  $\mathcal{C}_c^{(n)}$  is replaced by  $\mathcal{C}_{sp}^{(n)}$ . Theorem 3.5 also holds with  $\mu_{n,\varrho}$  replacing  $\mu_{\pi,\varrho}$ .

**3.3. Generalized serial strategies.** To further enlarge the class of convergent strategies one can start with elements from  $\mathcal{B}_{sp}^{(m)}$  and use all conceivable chains which comprise the equivalence relations  $\sim$ ,  $\overset{s}{\sim}$ ,  $\overset{w}{\sim}$ , and  $\overset{p}{\sim}$ . Fortunately, using Proposition 2.8 and Definition 2.9, we know that the set obtained this way can actually be obtained by using just one  $\overset{w}{\sim}$  and one  $\overset{p}{\sim}$ .

DEFINITION 3.7. *Let*

$$\begin{aligned}\mathcal{B}_{spg}^{(m)} &= \left\{ \mathcal{O} \in \mathcal{O}(\mathcal{P}_m) \mid \mathcal{O} \overset{p}{\sim} \mathcal{O}' \sim \mathcal{O}'' \text{ or } \mathcal{O} \sim \mathcal{O}' \overset{p}{\sim} \mathcal{O}'', \mathcal{O}'' \in \mathcal{B}_{sp}^{(m)} \right\}, \\ \mathcal{B}_{sg}^{(m)} &= \left\{ \mathcal{O} \in \mathcal{O}(\mathcal{P}_m) \mid \mathcal{O} \overset{p}{\sim} \mathcal{O}' \overset{w}{\sim} \mathcal{O}'' \text{ or } \mathcal{O} \overset{w}{\sim} \mathcal{O}' \overset{p}{\sim} \mathcal{O}'', \mathcal{O}'' \in \mathcal{B}_{sp}^{(m)} \right\},\end{aligned}$$

where the chains are in canonical form and  $\mathcal{O}' \in \mathcal{O}(\mathcal{P}_m)$ . The set  $\mathcal{B}_{sg}^{(m)}$  (resp.  $\{I_{\mathcal{O}} \mid \mathcal{O} \in \mathcal{B}_{sg}^{(m)}\}$ ) is the class of generalized serial pivot orderings of  $\mathcal{P}_m$  (resp. generalized serial pivot strategies). The set  $\mathcal{B}_{spg}^{(m)}$  is a subclass of  $\mathcal{B}_{sg}^{(m)}$  whose elements are linked by chains that do not use shifts.

THEOREM 3.8. *Let  $\pi = (n_1, \dots, n_m)$  be a partition of  $n$ ,  $\mathcal{O} \in \mathcal{B}_{spg}^{(m)}$ , and let  $\mathcal{J} \in \mathcal{J}_{\mathcal{O}}^{\text{UBCE}\pi(\varrho)}$  be the block Jacobi operator. Suppose that either  $\mathcal{O} \overset{p}{\sim} \mathcal{O}' \sim \mathcal{O}'' \in \mathcal{B}_{sp}^{(m)}$  with  $\mathcal{O}' = \mathcal{O}(\mathfrak{q})$ , or  $\mathcal{O} \sim \mathcal{O}' \overset{p}{\sim} \mathcal{O}'' \in \mathcal{B}_{sp}^{(m)}$  with  $\mathcal{O}'' = \mathcal{O}'(\mathfrak{q})$ , for some permutation  $\mathfrak{q}$  of the set  $\mathcal{S}_m$ . Then there exist constants  $\mu_{\pi,\varrho}$  and  $\tilde{\mu}_{n,\varrho}$ , depending only on  $\pi_{\mathfrak{q}}$ ,  $\varrho$  and  $n$ ,  $\varrho$ , respectively, such that*

$$\|\mathcal{J}\|_2 \leq \mu_{\pi_{\mathfrak{q}}}, \quad 0 \leq \mu_{\pi_{\mathfrak{q}},\varrho} < \tilde{\mu}_{n,\varrho} < 1.$$

*Proof.* Let us first consider the case  $\mathcal{O} \overset{p}{\sim} \mathcal{O}' \sim \mathcal{O}'' \in \mathcal{B}_{sp}^{(m)}$  with  $\mathcal{O}' = \mathcal{O}(\mathfrak{q})$ . Let

$$\tilde{\mathcal{B}}_{sp}^{(m)} = \left\{ \mathcal{O}' \in \mathcal{O}(\mathcal{P}_m) \mid \mathcal{O}' \sim \mathcal{O}'', \mathcal{O}'' \in \mathcal{B}_{sp}^{(m)} \right\}.$$

Theorem 3.5 and Lemma 2.19 imply

$$(3.6) \quad \|\mathcal{J}_{\mathcal{O}'}\|_2 \leq \mu_{\pi_{\mathfrak{q}},\varrho}, \quad 0 \leq \mu_{\pi_{\mathfrak{q}},\varrho} < \tilde{\mu}_{n,\varrho} < 1, \quad \mathcal{J}_{\mathcal{O}'} \in \mathcal{J}_{\mathcal{O}'}^{\text{UBCE}\pi_{\mathfrak{q}}(\varrho)}, \quad \mathcal{O}' \in \tilde{\mathcal{B}}_{sp}^{(m)}.$$

Since  $\mathcal{O} = \mathcal{O}'(\mathfrak{q}^{-1})$  and  $\pi_{\mathfrak{q}^{-1}\circ\mathfrak{q}} = \pi$ , formula (3.6) and Theorem 2.22(i) imply

$$(3.7) \quad \|\mathcal{J}_{\mathcal{O}}\|_2 \leq \mu_{\pi_{\mathfrak{q}},\varrho}, \quad 0 \leq \mu_{\pi_{\mathfrak{q}},\varrho} < \tilde{\mu}_{n,\varrho} < 1, \quad \text{for any } \mathcal{J}_{\mathcal{O}} \in \mathcal{J}_{\mathcal{O}}^{\text{UBCE}\pi(\varrho)}.$$

Hence, relation (3.7) holds for  $\mathcal{J}$ , which in turn proves the theorem.

Now suppose that  $\mathcal{O} \sim \mathcal{O}' \overset{p}{\sim} \mathcal{O}'' \in \mathcal{B}_{sp}^{(m)}$  with  $\mathcal{O}'' = \mathcal{O}'(\mathfrak{q})$ . Let

$$\hat{\mathcal{B}}_{sp}^{(m)} = \left\{ \mathcal{O}' \in \mathcal{O}(\mathcal{P}_m) \mid \mathcal{O}' = \mathcal{O}''(\mathfrak{q}^{-1}), \mathcal{O}'' \in \mathcal{B}_{sp}^{(m)} \right\}.$$

By Theorem 3.5 and Theorem 2.22(i) we have

$$\|\mathcal{J}_{\mathcal{O}'}\|_2 \leq \mu_{\pi_{\mathfrak{q}},\varrho}, \quad 0 \leq \mu_{\pi_{\mathfrak{q}},\varrho} < \tilde{\mu}_{n,\varrho} < 1, \quad \mathcal{J}_{\mathcal{O}'} \in \mathcal{J}_{\mathcal{O}'}^{\text{UBCE}\pi(\varrho)}, \quad \mathcal{O}' \in \hat{\mathcal{B}}_{sp}^{(m)}.$$

Here we have used  $\pi_{\mathfrak{q}^{-1}\circ\mathfrak{q}} = \pi$ . The last formula holds for any  $\mathcal{J}_{\mathcal{O}'} \in \mathcal{J}_{\mathcal{O}'}^{\text{UBCE}\pi(\varrho)}$ . Since  $\mathcal{O} \sim \mathcal{O}'$ , Lemma 2.19 implies that the same formula holds for any  $\mathcal{J}_{\mathcal{O}} \in \mathcal{J}_{\mathcal{O}}^{\text{UBCE}\pi(\varrho)}$ . This completes the proof since  $\mathcal{J}$  is just one of those  $\mathcal{J}_{\mathcal{O}}$ .  $\square$

**THEOREM 3.9.** *Let  $\pi = (n_1, \dots, n_m)$  be a partition of  $n$  and  $\mathcal{O} \in \mathcal{B}_{sg}^{(m)}$ . Suppose that the chain connecting  $\mathcal{O}$  to  $\mathcal{O}'' \in \mathcal{B}_{sp}^{(m)}$  contains  $d$  shift equivalences. Moreover, suppose that either  $\mathcal{O} \stackrel{p}{\sim} \mathcal{O}' \stackrel{w}{\sim} \mathcal{O}'' \in \mathcal{B}_{sp}^{(m)}$  with  $\mathcal{O}' = \mathcal{O}(\mathfrak{q})$  holds, or  $\mathcal{O} \stackrel{w}{\sim} \mathcal{O}' \stackrel{p}{\sim} \mathcal{O}'' \in \mathcal{B}_{sp}^{(m)}$  with  $\mathcal{O}'' = \mathcal{O}'(\mathfrak{q})$  holds, for some permutation  $\mathfrak{q}$  of the set  $\mathcal{S}_m$ . Then there exist constants  $\mu_{\pi, \varrho}$  and  $\tilde{\mu}_{n, \varrho}$  depending only on  $\pi$ ,  $\varrho$  and  $n$ ,  $\varrho$ , respectively, such that for any  $d + 1$  block Jacobi operators  $\mathcal{J}_1, \mathcal{J}_2, \dots, \mathcal{J}_{d+1} \in \mathcal{J}_{\mathcal{O}}^{\text{UBCE}_{\pi}(\varrho)}$  one has*

$$\|\mathcal{J}_1 \mathcal{J}_2 \cdots \mathcal{J}_{d+1}\|_2 \leq \mu_{\pi, \varrho}, \quad 0 \leq \mu_{\pi, \varrho} < \tilde{\mu}_{n, \varrho} < 1.$$

*Proof.* Let us first consider the case  $\mathcal{O} \stackrel{p}{\sim} \mathcal{O}' \stackrel{w}{\sim} \mathcal{O}'' \in \mathcal{B}_{sp}^{(m)}$  with  $\mathcal{O}' = \mathcal{O}(\mathfrak{q})$ . Denote

$$\tilde{\mathcal{B}}_{sp}^{(m)} = \{\mathcal{O}' \in \mathcal{O}(\mathcal{P}_m) \mid \mathcal{O}' \stackrel{w}{\sim} \mathcal{O}'', \mathcal{O}'' \in \mathcal{B}_{sp}^{(m)}\}.$$

Theorem 3.5 and Proposition 2.20 imply

$$(3.8) \quad \|\mathcal{J}'_1 \mathcal{J}'_2 \cdots \mathcal{J}'_{d+1}\|_2 \leq \mu_{\pi, \varrho}, \quad \mathcal{J}'_1, \mathcal{J}'_2, \dots, \mathcal{J}'_{d+1} \in \mathcal{J}_{\mathcal{O}'}^{\text{UBCE}_{\pi \mathfrak{q}}(\varrho)}, \mathcal{O}' \in \tilde{\mathcal{B}}_{sp}^{(m)},$$

where  $0 \leq \mu_{\pi, \varrho} < \tilde{\mu}_{n, \varrho} < 1$ . Since  $\mathcal{O} = \mathcal{O}'(\mathfrak{q}^{-1})$  and  $\pi_{\mathfrak{q}^{-1} \circ \mathfrak{q}} = \pi$ , relation (3.8) and Theorem 2.22(ii) give

$$\|\mathcal{J}_1 \mathcal{J}_2 \cdots \mathcal{J}_{d+1}\|_2 \leq \mu_{\pi, \varrho} \quad \text{for any } \mathcal{J}_1, \mathcal{J}_2, \dots, \mathcal{J}_{d+1} \in \mathcal{J}_{\mathcal{O}}^{\text{UBCE}_{\pi}(\varrho)},$$

which proves the theorem.

Now suppose that  $\mathcal{O} \stackrel{w}{\sim} \mathcal{O}' \stackrel{p}{\sim} \mathcal{O}'' \in \mathcal{B}_{sp}^{(m)}$  with  $\mathcal{O}'' = \mathcal{O}'(\mathfrak{q})$ . Let

$$\hat{\mathcal{B}}_{sp}^{(m)} = \{\mathcal{O}' \in \mathcal{O}(\mathcal{P}_m) \mid \mathcal{O}' = \mathcal{O}''(\mathfrak{q}^{-1}), \mathcal{O}'' \in \mathcal{B}_{sp}^{(m)}\}.$$

By Theorem 3.5 and Theorem 2.22(i) we have

$$\|\mathcal{J}_{\mathcal{O}'}\|_2 \leq \mu_{\pi, \varrho}, \quad 0 \leq \mu_{\pi, \varrho} < \tilde{\mu}_{n, \varrho} < 1, \quad \mathcal{J}_{\mathcal{O}'} \in \mathcal{J}_{\mathcal{O}'}^{\text{UBCE}_{\pi}(\varrho)}, \mathcal{O}' \in \hat{\mathcal{B}}_{sp}^{(m)}.$$

Here we used  $\pi_{\mathfrak{q}^{-1} \circ \mathfrak{q}} = \pi$  once again. Since  $\mathcal{O} \stackrel{w}{\sim} \mathcal{O}'$ , Proposition 2.20 completes the proof.  $\square$

We end this section by shifting our attention from block Jacobi operators to cyclic block Jacobi methods, defined by the generalized serial strategies.

**THEOREM 3.10.** *Let  $\pi = (n_1, \dots, n_m)$  be a partition of  $n$ ,  $\mathcal{O} \in \mathcal{B}_{spg}^{(m)}$ , and let  $A \in \mathbf{S}_n$  be partitioned as in relation (2.2). Let  $A'$  be obtained from  $A$  by applying one sweep of the cyclic block Jacobi method defined by the strategy  $I_{\mathcal{O}}$ . If all transformation matrices are from the class  $\text{UBCE}_{\pi \mathfrak{q}}(\varrho)$  for an appropriate permutation  $\mathfrak{q}$  of the set  $\mathcal{S}_m$ , then there are constants  $\eta_{\pi, \varrho}$  and  $\tilde{\eta}_{n, \varrho}$  depending only on  $\pi, \varrho$  and  $n, \varrho$ , respectively, such that*

$$S^2(A') \leq \eta_{\pi, \varrho} S^2(A), \quad 0 \leq \eta_{\pi, \varrho} < \tilde{\eta}_{n, \varrho} < 1.$$

*Proof.* The proof is almost identical to the proof of Theorem 3.6. The difference is that  $\mathcal{O} \in \mathcal{B}_{spg}^{(m)}$  and we use Theorem 3.8 instead of Theorem 3.5.  $\square$

**THEOREM 3.11.** *Let  $\pi = (n_1, \dots, n_m)$  be a partition of  $n$ ,  $\mathcal{O} \in \mathcal{B}_{sg}^{(m)}$ , and let  $A \in \mathbf{S}_n$  be partitioned as in relation (2.2). Suppose that the chain connecting  $\mathcal{O}$  and  $\mathcal{O}'' \in \mathcal{B}_{sp}^{(m)}$  is in the canonical form and contains  $d$  shift equivalences. Let  $A'$  be obtained from  $A$  by applying  $d + 1$  sweeps of the cyclic block Jacobi method defined by the strategy  $I_{\mathcal{O}}$ . If all transformation*

matrices are from the class  $UBCE_{\pi_q}(\varrho)$  for an appropriate permutation  $q$  of the set  $\mathcal{P}_m$ , then there are constants  $\eta_{\pi_q, \varrho}$  and  $\tilde{\eta}_{n, \varrho}$  depending only on  $\pi_q$ ,  $\varrho$  and  $n$ ,  $\varrho$ , respectively, such that

$$S^2(A') \leq \eta_{\pi_q, \varrho} S^2(A), \quad 0 \leq \eta_{\pi_q, \varrho} < \tilde{\eta}_{n, \varrho} < 1.$$

Here  $q$  is an appropriate permutation of the set  $\mathcal{P}_m$ .

*Proof.* The proof follows the lines of the proof of Theorem 3.6. Since we consider  $d + 1$  sweeps, instead of relation (3.5), we will obtain

$$a' = a^{((d+1)M)} = \mathcal{J}_{\mathcal{O}}^{[d+1]} \mathcal{J}_{\mathcal{O}}^{[d]} \cdots \mathcal{J}_{\mathcal{O}}^{[1]} a, \quad \mathcal{O} \in \mathcal{B}_{sg}^{(m)}.$$

Here  $\mathcal{J}_{\mathcal{O}}^{[s]}$  is the block Jacobi operator associated with cycle  $s$  of the block Jacobi method and  $a = \text{vec}(A)$ ,  $a' = \text{vec}(A')$ . From Theorem 3.9 we know that

$$\|\mathcal{J}_{\mathcal{O}}^{[d+1]} \mathcal{J}_{\mathcal{O}}^{[d]} \cdots \mathcal{J}_{\mathcal{O}}^{[1]}\|_2 \leq \mu_{\pi_q, \varrho}, \quad 0 \leq \mu_{\pi_q, \varrho} < \tilde{\mu}_{n, \varrho} < 1,$$

so

$$S^2(A') = \|a'\|_2^2 \leq \mu_{\pi_q, \varrho}^2 \|a\|_2^2 \leq \mu_{\pi_q, \varrho}^2 S^2(A).$$

It remains to set  $\eta_{\pi_q, \varrho} = \mu_{\pi_q, \varrho}^2$  and  $\tilde{\eta}_{n, \varrho} = \tilde{\mu}_{n, \varrho}^2$ .  $\square$

In the case  $\pi = (1, 1, \dots, 1)$ , we have  $m = n$  and we can use notation  $\mathcal{C}_{spg}^{(m)}$  and  $\mathcal{C}_{sg}^{(m)}$  instead of  $\mathcal{B}_{spg}^{(m)}$  and  $\mathcal{B}_{sg}^{(m)}$ , respectively. Corollary 3.3 holds with the same constant  $\eta_n$  if  $\mathcal{C}_c^{(n)}$  is replaced by  $\mathcal{C}_{spg}^{(n)}$ . Theorems 3.9–3.11 also hold with  $\mu_{n, \varrho}$  instead of  $\mu_{\pi_q, \varrho}$ .

It is not easy to count how many pivot orderings are contained in  $\mathcal{B}_{sg}^{(m)}$ . In  $\mathcal{B}_c^{(m)}$  we have  $2! \cdot 3! \cdots (m-1)!$  elements. The sets  $\overleftarrow{\mathcal{B}}_c^{(m)}$ ,  $\mathcal{B}_r^{(m)}$ , and  $\overleftarrow{\mathcal{B}}_r^{(m)}$  have the same number of elements. Hence, for large  $m$  we expect that  $\mathcal{B}_{sp}^{(m)}$  contains  $4 \cdot 2! \cdot 3! \cdots (m-1)!$  elements. For each ordering  $\mathcal{O} \in \mathcal{B}_{sp}^{(m)}$  there are  $m!$  orderings of the form  $\mathcal{O}(p)$ , so for large  $m$  we expect at least  $4 \cdot 2! \cdot 3! \cdots m!$  elements in  $\mathcal{B}_{spg}^{(m)}$  (where we have not taken the equivalences  $\sim$  and  $\tilde{\sim}$  into account). Obviously, for small  $m$  (like  $m = 3, 4, 5$ ) this count is not realistic.

Nevertheless, the results obtained here have been used in [2] to prove that every cyclic Jacobi method for symmetric matrices of order 4 is globally convergent. Note that there are altogether 720 cyclic strategies when  $n = 4$ .

**4. Quasi-cyclic pivot strategies.** Our next step is to enlarge the scope of generalized serial strategies by allowing repetition of some Jacobi steps within one sweep. This leads us to special quasi-cyclic pivot sequences, which are closely related to the orderings from Section 3. This change often leads to faster convergence of the Jacobi method [7, 8, 16]. To keep our consideration within reasonable framework, we can assume that the length of each quasi-cyclic pivot sequence is smaller than  $2M$ , where  $M = \frac{m(m-1)}{2}$ .

Our basic class of quasi-cyclic pivot sequences is derived from the class  $\mathcal{B}_c^{(m)}$ . More precisely,

$$\begin{aligned} \bar{\mathcal{B}}_c^{(m)} = \left\{ \mathcal{O} \in \mathcal{O}(\mathcal{P}_m) \mid \mathcal{O} = (1, 2), (\pi_3(1), 3), (\pi_3(2), 3), \mathcal{O}_3, \dots, (\pi_m(1), m), \right. \\ \left. \dots, (\pi_m(m-1), m), \mathcal{O}_m, \pi_j \in \Pi^{(1, j-1)}, \mathcal{O}_j \in \mathcal{O}(\mathcal{S}_j), \mathcal{S}_j \subseteq \mathcal{P}_j, 3 \leq j \leq m \right\}. \end{aligned}$$

Thus, the quasi-cyclic pivot strategy  $I_{\mathcal{O}}$  defined by some  $\mathcal{O} \in \bar{\mathcal{B}}_c^{(m)}$  selects pivot blocks by block-columns. After the pivot blocks within the  $j$ th block-column have all been annihilated



once, it is allowed to annihilate again any block that lies within the first  $j$  block-columns. Here,  $3 \leq j \leq m$ .

**THEOREM 4.1.** *Let  $\pi = (n_1, \dots, n_m)$  be a partition of  $n$ ,  $\mathcal{O} \in \bar{\mathcal{B}}_c^{(m)}$ , and let  $A \in \mathbf{S}_n$  be partitioned as in relation (2.2). Let  $A'$  be obtained from  $A$  by applying one sweep of the quasi-cyclic block Jacobi method defined by the strategy  $I_{\mathcal{O}}$ . If all transformation matrices are from the class  $UBCE(\varrho)$ , then there are constants  $\eta_{\pi, \varrho}$  and  $\tilde{\eta}_{n, \varrho}$ , depending only on  $\pi$ ,  $\varrho$  and  $n$ ,  $\varrho$ , respectively, such that*

$$S^2(A') \leq \eta_{\pi, \varrho} S^2(A), \quad 0 \leq \eta_{\pi, \varrho} < \tilde{\eta}_{n, \varrho} < 1.$$

*Proof.* The proof has been moved to Appendix A. □

Although the quasi-cyclic strategy from [7, 8, 16] belongs to the class of block-oriented strategies, its “full block” analogue is  $I_{\mathcal{O}}$  for some special  $\mathcal{O} \in \bar{\mathcal{B}}_c^{(m)}$ . As has been shown in [19, 20] for large matrices that full-block Jacobi-type methods are generally more efficient than the block-oriented ones. This implies that the Jacobi method from LAPACK can be upgraded to a full block version, and Theorem 4.1 ensures its convergence.

Now, it is easy to prove Theorem 3.2 for  $\mathcal{O} \in \bar{\mathcal{B}}_c^{(m)}$ . The proof remains the same except that the word “cyclic” should be replaced by “quasi-cyclic” and the notation  $\mathcal{B}_c^{(m)}$  should be replaced with  $\bar{\mathcal{B}}_c^{(m)}$ . The case  $m = n$  is treated in the same way.

Following ideas from Section 3, we define  $\bar{\mathcal{B}}_r^{(m)} = \{\mathcal{O}(\tilde{\epsilon}) \mid \mathcal{O} \in \bar{\mathcal{B}}_c^{(m)}\}$ ,

$$\overleftarrow{\bar{\mathcal{B}}}_c^{(m)} = \{\mathcal{O} \in \mathcal{O}(\mathcal{P}_m) \mid \mathcal{O}^{\leftarrow} \in \bar{\mathcal{B}}_c^{(m)}\}, \quad \overleftarrow{\bar{\mathcal{B}}}_r^{(m)} = \{\mathcal{O} \in \mathcal{O}(\mathcal{P}_m) \mid \mathcal{O}^{\leftarrow} \in \bar{\mathcal{B}}_r^{(m)}\}$$

and

$$\bar{\mathcal{B}}_{cp}^{(m)} = \bar{\mathcal{B}}_c^{(m)} \cup \overleftarrow{\bar{\mathcal{B}}}_c^{(m)}, \quad \bar{\mathcal{B}}_{rp}^{(m)} = \bar{\mathcal{B}}_r^{(m)} \cup \overleftarrow{\bar{\mathcal{B}}}_r^{(m)}, \quad \bar{\mathcal{B}}_{sp}^{(m)} = \bar{\mathcal{B}}_{cp}^{(m)} \cup \bar{\mathcal{B}}_{rp}^{(m)}.$$

It is easy to verify that both Theorem 3.5 and Theorem 3.6 hold with  $\bar{\mathcal{B}}_{sp}^{(m)}$  in the place of  $\mathcal{B}_{sp}^{(m)}$ . Finally, we can define

$$\begin{aligned} \bar{\mathcal{B}}_{spg}^{(m)} &= \left\{ \mathcal{O} \in \mathcal{O}(\mathcal{P}_m) \mid \mathcal{O} \stackrel{p}{\sim} \mathcal{O}'' \sim \mathcal{O}' \text{ or } \mathcal{O} \sim \mathcal{O}' \stackrel{p}{\sim} \mathcal{O}'', \mathcal{O}'' \in \bar{\mathcal{B}}_{sp}^{(m)} \right\}, \\ \bar{\mathcal{B}}_{sg}^{(m)} &= \left\{ \mathcal{O} \in \mathcal{O}(\mathcal{P}_m) \mid \mathcal{O} \stackrel{p}{\sim} \mathcal{O}' \stackrel{w}{\sim} \mathcal{O}'' \text{ or } \mathcal{O} \stackrel{w}{\sim} \mathcal{O}' \stackrel{p}{\sim} \mathcal{O}'', \mathcal{O}'' \in \bar{\mathcal{B}}_{sp}^{(m)} \right\}. \end{aligned}$$

It is easy to verify that all four theorems, Theorems 3.8–3.11, hold with  $\bar{\mathcal{B}}_{spg}^{(m)}$  and  $\bar{\mathcal{B}}_{sg}^{(m)}$ .

In the case  $m = n$ , one can reestablish the corresponding results for the standard Jacobi method and the associated Jacobi operators.

**5. Convergence of more general block Jacobi-type methods.** The obtained results for the block Jacobi operators and annihilators can be used to prove convergence of more general block Jacobi-type methods. This section is similar to [18, Section 5], and we will refer to some results from there. First, we prove the main result, and then we apply it to the block J-Jacobi method from [20].

Consider the block Jacobi-type process

$$(5.1) \quad A^{(k+1)} = F_k^T A^{(k)} F_k, \quad k \geq 0, \quad A^{(0)} = A,$$

where  $A$  is a symmetric matrix of order  $n$ , partitioned as in relation (2.2), and  $F_k$ ,  $k \geq 0$ , are elementary block matrices. Their pivot submatrices are only required to be nonsingular. Since all  $F_k$  are nonsingular,  $A \neq 0$  implies  $A^{(k)} \neq 0$  for all  $k$ . The process is said to be of Jacobi-type since it is generally not required that the pivot submatrices are diagonalized. We assume that for the process (5.1) the following assumptions hold:

**A1**  $\mathcal{O} \in \mathcal{B}_{sg}^{(m)}$ , i.e., the pivot strategy  $I_{\mathcal{O}}$  of the process is a generalized serial one.

**A2** There is a sequence of orthogonal elementary block matrices  $U_k$ ,  $k \geq 0$ , such that

$$\lim_{k \rightarrow \infty} (F_k - U_k) = 0.$$

**A3** For the diagonal block  $F_{ii}^{(k)}$  of  $F_k$  one has

$$\sigma = \liminf_{k \rightarrow \infty} \sigma^{(k)} > 0, \quad \text{where } \sigma^{(k)} = \sigma_{\min}(F_{ii}^{(k)}), \quad k \geq 0.$$

The first assumption **A1** deserves a comment. By Definition 3.7, the set  $\mathcal{B}_{sg}^{(m)}$  is defined using a single permutation equivalence. In order to make use of Theorems 3.9 and 3.11, we can presume that in Definition 3.7 either  $\mathcal{O}' = \mathcal{O}(q)$  or  $\mathcal{O}'' = \mathcal{O}'(q)$  holds for some permutation  $q$ .

Because of the condition **A2**, in the assumption **A3** one can replace  $F_{ii}^{(k)}$  by  $U_{ii}^{(k)}$ . From the CS decomposition of the orthogonal  $(n_i + n_j) \times (n_i + n_j)$  matrix  $\tilde{U}_{ij}$ , we have  $\sigma_{\min}(U_{ii}^{(k)}) = \sigma_{\min}(U_{jj}^{(k)})$ . Therefore, in the definition of  $\sigma^{(k)}$ , instead of  $F_{ii}^{(k)}$ , one can use  $F_{jj}^{(k)}$ ,  $U_{ii}^{(k)}$ , or  $U_{jj}^{(k)}$ . Recall that  $F_{ii}^{(k)}$ ,  $F_{jj}^{(k)}$ , stands for  $F_{i(k)i(k)}^{(k)}$ ,  $F_{j(k)j(k)}^{(k)}$ , respectively.

Since for each  $U_k$ , there is a permutation matrix  $P_k$  that makes  $U_k P_k$  a UBCE matrix, condition **A2** shows that for large enough  $k$ , each  $F_k P_k$  will be arbitrarily close to some UBCE matrix. However,  $U_k$ , and therefore also  $P_k$ , is generally not known, while  $F_k$  is available. Thus, one can perform the QR factorization with column pivoting of  $[F_{ii}^{(k)} F_{ij}^{(k)}]$  to obtain  $P_k$  and then replace  $F_k$  by  $F_k P_k$ . The corresponding matrix  $\tilde{U}_k = U_k P_k$  may not be from UBCE(1), but it is certainly from UBCE( $\varrho$ ) for some  $0 < \varrho < 1$  when  $k$  is large enough.

**THEOREM 5.1.** *Let  $\pi = (n_1, \dots, n_m)$  be a partition of  $n$ ,  $\mathcal{O} \in \mathcal{B}_{sg}^{(m)}$ , and let  $A \in \mathbf{S}_n$ ,  $A \neq 0$ , be partitioned as in relation (2.2). Let the sequence of matrices  $(A^{(k)}; k \geq 0)$  be generated by the block Jacobi-type process (5.1). If the assumptions **A1–A3** are met, then the following two assertions are equivalent:*

- (i)  $\lim_{k \rightarrow \infty} \frac{S(\hat{A}_{ij}^{(k+1)})}{\|A^{(k)}\|_F} = 0$ , where  $\hat{A}_{ij}^{(k+1)} = \hat{F}_k^T \hat{A}_{ij}^{(k)} \hat{F}_k$ ,
- (ii)  $\lim_{k \rightarrow \infty} \frac{S(A^{(k)})}{\|A^{(k)}\|_F} = 0$ .

*Proof.* The proof has been moved to Appendix A.  $\square$

Thus, condition (i) is sufficient for the convergence of  $S(A^{(k)})$  to zero. In the case of block-wise or element-wise Jacobi methods (i.e., those that diagonalize the pivot submatrix at each step) condition (i) is trivially fulfilled. Note that  $S(\hat{A}_{ij}^{(k+1)})$  and  $S(A^{(k)})$  are being divided by  $\|A^{(k)}\|_F$ , which is appropriate since the theorem deals with nonorthogonal transformations.

In some applications the following corollaries can be used.

**COROLLARY 5.2.** *Theorem 5.1 holds provided that the recurrence relation (5.1) is replaced by*

$$(5.2) \quad A^{(k+1)} = F_k^T A^{(k)} F_k + E^{(k)}, \quad k \geq 0,$$

where  $\lim_{k \rightarrow \infty} S(E^{(k)})/\|A^{(k)}\|_F = 0$  and  $E^{(k)} \neq -F_k^T A^{(k)} F_k$ ,  $k \geq 0$ . The last condition on  $E^{(k)}$  can be replaced by the requirement that  $E^{(k)} = 0$  whenever  $A^{(k)} = 0$  for some  $k$ .

*Proof.* Comparing with the proof of Theorem 5.1, the only difference appears in the definition of each vector  $g^{(k)}$ , now including the vector  $e^{(k)}$ , which in turn results from the matrix  $E^{(k)}$ .  $\square$

**COROLLARY 5.3.** *Let  $A \neq O$  be a matrix of order  $n$ , and let the sequence  $A^{(0)} = A, A^{(1)}, \dots$  be generated by a block Jacobi-type process defined by relation (5.2). Assume that the assumptions **A1–A3** hold. Suppose that the sequence  $(A^{(k)}; k \geq 0)$  is bounded and*

$$(5.3) \quad \lim_{t \rightarrow \infty} S(E^{(k)}) = 0.$$

*Then the following two conditions are equivalent:*

- (iii)  $\lim_{k \rightarrow \infty} S(\widehat{A}_{ij}^{(k+1)}) = 0,$
- (iv)  $\lim_{k \rightarrow \infty} S(A^{(k)}) = 0.$

*Proof.* The implication (iv)  $\Rightarrow$  (iii) is obvious. For the converse implication, we use the expressions  $F_k = U_k + (F_k - U_k), k \geq 0,$  to transform the recursion (5.2) into the form  $A^{(k+1)} = U_k^T A^{(k)} U_k + T^{(k)} + E^{(k)}$  with

$$T^{(k)} = (F_k - U_k)^T A^{(k)} U_k + U_k^T A^{(k)} (F_k - U_k) + (F_k - U_k)^T A^{(k)} (F_k - U_k).$$

By the boundedness of the sequence  $(A^{(k)}; k \geq 0)$  combined with assumption **A2**, we have

$$\|T^{(k)}\| \leq \sup \left\{ \|A^{(k)}\|_2; k \geq 0 \right\} \cdot (2\|F_k - U_k\| + \|F_k - U_k\|^2) \rightarrow 0 \quad \text{as } k \rightarrow \infty.$$

This confirms relations (A.16), (A.17), with  $g^{(k)}$  which additionally includes the vectors associated with the matrices  $T^{(k)}$  and  $E^{(k)}$ . The conditions (iii), (5.3), and the latest relation together imply  $\lim_{k \rightarrow \infty} g^{(k)} = 0$ . Following the same lines of the proof of Theorem 5.1, one obtains  $\lim_{s \rightarrow \infty} g^{[s]} = 0$  and  $\lim_{k \rightarrow \infty} a^{(k)} = 0$  if and only if  $\lim_{s \rightarrow \infty} a^{[s]} = 0$ . Therefore, under the conditions of this corollary, the sequence  $(a^{[s]}; s \geq 1)$  has all the properties of the sequence (A.24) from the proof of Theorem 5.1. The rest of the proof closely follows the proof of Theorem 5.1.  $\square$

Let us note that the results in this section hold if the set of pivot orderings  $\mathcal{B}_{sg}^{(m)}$  in the assumption **A1** is replaced with  $\widetilde{\mathcal{B}}_{sg}^{(m)}$ . Also, as has already been explained in [18], it makes sense to rewrite the assumption **A3** in the equivalent form:

**A3** For the diagonal block  $F_{ii}^{(k)}$  of  $F_k$ , one has

$$\sigma = \liminf_{t \rightarrow \infty} \sigma^{[t]} > 0, \quad \sigma^{[t]} = \min_{(t-1)T \leq k \leq tT-1} \sigma_{\min}(F_{ii}^{(k)}),$$

where the quantities  $\sigma^{[t]}$  are labeled by sweeps (i.e., cycles or quasi-cycles).

**5.1. An application to the block  $J$ -Jacobi methods.** The main purpose of Theorem 5.1 is its use in the global convergence considerations of the block Jacobi methods for the generalized eigenvalue problem, say for the HZ method from [29]. However, further research is needed to achieve this goal. Hence, we will choose yet another block method, which is well-understood, important in practice, and for which the newly obtained results can be applied straightforwardly. It is the *full block  $J$ -Jacobi method* from [20], for the pair  $(A, J)$ , where  $A$  is symmetric positive definite and  $J = \text{diag}(I_\nu, -I_{n-\nu})$ . The main application of this method is to solve the simple eigenvalue problem  $Hx = \lambda x$ , where  $H$  is indefinite symmetric matrix, with high relative accuracy. The partition  $\pi = (n_1, \dots, n_m)$  has to comply with the partition  $(\nu, n - \nu)$ , i.e., the first has to be a subpartition of the latter. After preliminary transformations, the problem  $Hx = \lambda x$  is reduced to the generalized eigenvalue problem  $Ax = \lambda Jx$ . All details can be found in [20, 33]. The method uses  $J$ -orthogonal elementary block matrices  $F_k$ , which leave  $J$  intact under congruence transformations  $F_k^T J F_k = J, k \geq 0$ . The iteration process has the form (5.1) with a positive definite matrix  $A = A^{(0)}$ .

In [20] the global convergence of this method was proved under the weak wavefront strategies, and here we prove it for the much larger class of generalized serial strategies.

**THEOREM 5.4.** *Let  $\pi = (n_1, \dots, n_m)$  be a partition of  $n$  so that  $\pi$  refines  $(\nu, n - \nu)$ . The full block  $J$ -Jacobi method defined by the cyclic pivot strategy  $I_{\mathcal{O}}$ ,  $\mathcal{O} \in \mathcal{B}_{sg}^{(m)}$ , which uses UBCE  $J$ -orthogonal transformation matrices, is globally convergent.*

*Proof.* Since the iterates generated by the full block  $J$ -Jacobi method are bounded [20, (3.18)], we can apply Corollary 5.3 with the matrices  $E^{(k)}$ ,  $k \geq 0$ , set to zero. The method is called the full block  $J$ -Jacobi method because at each step it diagonalizes the pivot submatrix. This implies that condition (iii) of Corollary 5.3 is fulfilled.

All we have to do is follow the same lines of the proof of [20, Proposition 3.3], which in turn reduces to checking the validity of the assumptions **A1**–**A3** from Theorem 5.1.

The first assumption is presumed. The second one follows from [20, Proposition 3.2]. Assumption **A3** holds for two reasons. First, for each hyperbolic elementary block transformation  $F_k$ , one has  $\sigma^{(k)} \geq 1$ , and we only have to check **A3** for the orthogonal elementary block transformations. However, they are exactly the same as those in the block Jacobi method for symmetric matrices from Sections 3 and 4 of this paper. Relations (2.15) and (2.16) hold for them, even with  $\varrho = 1$ . Since condition (iii) of Corollary 5.3 is fulfilled, we have  $S(A^{(k)}) \rightarrow 0$  as  $k \rightarrow \infty$ . The rest of the proof requires an analogue of Theorem 2.10 for the  $J$ -Jacobi method. However, all that is needed in the proof is an estimate similar to (A.1) for the diagonal elements of  $A^{(k)}$ . Such an estimate is established in [6, Lemma 1.1].  $\square$

By using results from Section 4 one can easily show that Theorem 5.4 holds for any quasi-cyclic strategy  $I_{\mathcal{O}}$ ,  $\mathcal{O} \in \bar{\mathcal{B}}_{sg}^{(m)}$ .

**6. Conclusion and future work.** So far, a satisfactory research goal has been the global convergence of the block Jacobi method for symmetric matrices established for the serial pivot strategies or those that are equivalent or weakly equivalent to them, so-called wavefront or weak wavefront strategies. All those strategies were obtained from a single cyclic strategy, say the column-cyclic one. Here we have shown how to further enlarge the class of convergent strategies by using the notion of the reverse strategy and that of permutation equivalent strategies. Hence, with each convergent pivot strategy we have associated the whole large class of convergent strategies obtained from it by using four equivalence relations,  $\sim$ ,  $\overset{s}{\sim}$ ,  $\overset{w}{\sim}$ ,  $\overset{p}{\sim}$ , and by using reverse strategies. The next step was to increase the number of classes of convergent strategies obtained this way. For large  $m$ , we have obtained at least  $2!3! \cdots (m-1)!$  such classes of convergent strategies, and we have named their union the class of generalized serial strategies. Furthermore, convergence results for that class are stated and proved in the stronger form, which enables us to formulate and prove similar results for the block Jacobi operators. This makes the block Jacobi operators a tool for proving the global convergence of the block Jacobi methods for other eigenvalue problems, in particular for the generalized eigenvalue problem. As an immediate result, we have proved the global convergence of the full block  $J$ -Jacobi method under any generalized serial pivot strategy.

Future work will include proving global convergence of the (block-wise and element-wise) HZ method [13, 29] for the generalized eigenvalue and singular value problem under the class of generalized serial strategies. We also intend to prove the global convergence of the block Paardekoooper method for skew-symmetric matrices. An immediate consequence of the results from this paper is the proof that in the case  $n = 4$ , all 720 cyclic strategies for the symmetric Jacobi method are convergent [2]. Future research will also be concentrated on the complex block Jacobi methods, first for a single Hermitian matrix and then for a positive definite pair of Hermitian matrices (the complex block  $J$ -Jacobi and the complex HZ methods).

**Acknowledgments.** The authors are indebted to the referees and the editor for their valuable suggestions that improved the paper.

**Appendix A. Proofs omitted in the main text.**

**A.1. Proof of Theorem 2.10.** Suppose that  $A$  is not a multiple of  $I_n$  and assume that for the eigenvalues of  $A$  we have

$$\lambda_1 = \dots = \lambda_{s_1} > \lambda_{s_1+1} = \dots = \lambda_{s_2} > \dots > \lambda_{s_{\omega-1}+1} = \dots = \lambda_{s_\omega}, \quad s_\omega = n,$$

$s_r = \nu_1 + \dots + \nu_r$ ,  $1 \leq r \leq \omega$ , where  $\nu_1, \dots, \nu_\omega$  are the multiplicities of the eigenvalues. Let

$$3\delta = \min_{1 \leq r \leq \omega-1} (\lambda_{s_r} - \lambda_{s_{r+1}}).$$

Obviously, there is an integer  $k_0$  such that

$$S(A^{(k)}) < \delta, \quad k \geq k_0.$$

From [15, Lemma 2.1] we conclude that for  $k \geq k_0$  all diagonal elements of  $A^{(k)}$  lie in the union of shrinking segments

$$\mathcal{D}_r^{(k)} \equiv \{t \mid |t - \lambda_{s_r}| \leq 0.22S(A^{(k)})\} \subset \{t \mid |t - \lambda_{s_r}| \leq 0.22\delta\} \equiv \mathcal{D}_r, \quad 1 \leq r \leq \omega.$$

Furthermore, by the same lemma, each  $\mathcal{D}_r^{(k)}$  contains at least  $\nu_r$  diagonal elements of  $A^{(k)}$ , and then it straightforwardly follows that  $\mathcal{D}_r^{(k)}$  contains exactly  $\nu_r$  diagonal elements of  $A^{(k)}$ . In particular, this implies that for any two diagonal elements of  $A^{(k)}$  we have

$$(A.1) \quad \text{either } |a_{ll}^{(k)} - a_{mm}^{(k)}| \leq 0.44\delta, \quad \text{or } |a_{ll}^{(k)} - a_{mm}^{(k)}| > 2.56\delta, \quad k \geq k_0.$$

Since the sequence  $S(A^{(k)})$ ,  $k \geq 0$ , converges to zero, the proof will be completed if we show that for  $k \geq k_0$  the diagonal elements cannot change their affiliation to eigenvalues. Afterwards, we will also show how the diagonal elements of  $\Lambda$  are ordered along the diagonal.

To establish assertion (i) it is sufficient to prove the first claim only since the second one can be proved in a similar way. Even if the blocks  $A_{ii}^{(k)}$ ,  $1 \leq i \leq m$ , were not diagonalized at the beginning, we can increase  $k_0$ , if needed, so that the assumption of the first claim reads: each  $A_{ii}^{(k)}$ ,  $1 \leq i \leq m$ ,  $k \geq k_0$ , is diagonal with diagonal elements ordered nonincreasingly. To this end we denote  $A_{ii}^{(k)}$  by  $\Lambda_{ii}^{(k)}$ ,  $1 \leq i \leq m$ . We can also assume that  $k_0 = t_0T$ , where  $T \geq M = m(m-1)/2$  is the period of the strategy. The proof will be completed if we can find  $k'_0 \geq k_0$  such that

$$(A.2) \quad \{a_{11}^{(k)}, a_{22}^{(k)}, \dots, a_{s_r, s_r}^{(k)}\} \subset \bigcup_{p=1}^r \mathcal{D}_p, \quad 1 \leq r \leq \omega, \quad k \geq k'_0.$$

Let us consider step  $k$  of the block method with  $k \geq k_0$ . Let  $i = i(k)$ ,  $j = j(k)$ , and let  $\widehat{A}_{ij}^{(k)} = \text{diag}(\Lambda_{ii}^{(k+1)}, \Lambda_{jj}^{(k+1)})$  be the transformed pivot submatrix  $\widehat{A}_{ij}^{(k)}$ . From relation (2.6) and the perturbation theorem for the symmetric matrices, we conclude that

$$(A.3) \quad \begin{aligned} \|\text{diag}(\Lambda_{ii}^{(k+1)}, \Lambda_{jj}^{(k+1)}) - P_k^T \text{diag}(\Lambda_{ii}^{(k)}, \Lambda_{jj}^{(k)}) P_k\|_2 &\leq \|A_{ij}^{(k)}\|_2 \\ &\leq \frac{\sqrt{2}}{2} S(A^{(k)}) < \frac{\sqrt{2}}{2} \delta \end{aligned}$$

holds for  $k \geq k_0$ , where  $P_k$  are some permutation matrices. From relations (A.3) and (A.1), we obtain the following geometric interpretation of the movement of the diagonal elements during one step of the method. The diagonal elements of  $A^{(k)}$  are points on the real axis situated within small segments around the eigenvalues. After the completion of step  $k$ , they have moved (as points) within the same segment, but (as diagonal elements) their subscripts may have changed.

What happens with a diagonal element  $a_{qq}^{(k_0)}$ , which lies in the segment  $\mathcal{D}_1$ ?

First, suppose that  $\nu_1 \leq n_p$  for all  $1 \leq p \leq m$ . Then each time  $a_{qq}^{(k_0)}$  is affected, it will be a diagonal element of  $\Lambda_{ii}^{(k+1)}$ . Let  $a_{qq}^{(k_0)}$  lie in  $A_{ll}^{(k_0)}$ . Then, for  $k \geq k_0$ ,  $a_{qq}^{(k)}$  is affected when  $i(k) = l$  or  $j(k) = l$ .

If  $i(k) = l$ , then  $a_{qq}^{(k)}$  will remain in the same diagonal block, which is  $\Lambda_{ii}^{(k+1)} = A_{ll}^{(k+1)}$ . If  $\Lambda_{jj}^{(k)}$  contains some diagonal elements from  $\mathcal{D}_1$ , then they will move to  $\Lambda_{ii}^{(k+1)}$ , and thus the number of diagonal elements from  $\mathcal{D}_1$  in  $A_{ll}^{(k+1)}$  will be larger than in  $A_{ll}^{(k)}$ .

If  $j(k) = l$ , then  $a_{qq}^{(k)}$  will move to the new diagonal block  $\Lambda_{ii}^{(k+1)}$ ,  $i < l$ , hence its subscripts will become smaller than or equal to  $s_{l-1}$ . Since the pivot strategy is cyclic or quasi-cyclic, the case  $j(k) = l$  must occur within the current sweep, unless  $l = 1$ . Hence, during the next sweep  $a_{qq}^{(k_0)}$  will move to some diagonal block which lies closer to  $A_{11}^{(k_0)}$  unless  $l = 1$ .

Since  $a_{qq}^{(k_0)}$  is an arbitrary diagonal element of  $\mathcal{D}_1$ , we conclude that within one sweep all diagonal elements from  $\mathcal{D}_1$  not belonging to  $A_{11}^{(k)}$  will decrease their subscripts to such an extent that they become the diagonal elements of other diagonal blocks. This analysis shows that within the first  $m - 1$  sweeps all diagonal elements belonging to  $\mathcal{D}_1$  will be the elements of the first diagonal block.

Now, let  $\nu_1$  be such that  $1 \leq \nu_1 < n$  holds. Then the same analysis shows that during  $m - 1$  sweeps, the diagonal elements affiliated with  $\lambda_1$  will be filling in the first  $\nu_1$  diagonal positions of the matrix. Hence, there is a number  $k_1 \geq (m - 1)T + k_0$  such that the first  $\nu_1$  diagonal elements in  $A^{(k_1)}$  are affiliated with  $\lambda_1$ .

Almost the same analysis shows that within the first  $m - 1$  sweeps the diagonal elements affiliated with  $\lambda_n$  will be filling in the last  $\nu_\omega$  diagonal positions of the matrix. By increasing  $k_1$  if necessary, we can assume that the last  $\nu_\omega$  diagonal elements of  $A^{(k_1)}$  are affiliated with  $\lambda_n$ .

The rest of the proof considers the matrix  $A^{(k_1)}$ . In  $A^{(k_1)}$  the first  $\nu_1$  and the last  $\nu_\omega$  diagonal positions are occupied by the diagonal elements from  $\mathcal{D}_1$  and  $\mathcal{D}_\omega$ , respectively. The situation is described by the following block-matrix partition

$$A^{(k_1)} = \begin{bmatrix} A_{s_\varrho}^{(k_1)} & B^{(k_1)} & C^{(k_1)} \\ B^{(k_1)T} & A_{n-s_\varrho-\tilde{s}_\rho}^{(k_1)} & G^{(k_1)} \\ C^{(k_1)T} & G^{(k_1)T} & A_{\tilde{s}_\rho}^{(k_1)} \end{bmatrix}, \quad \begin{aligned} s_\varrho &= n_1 + \cdots + n_\varrho, \\ \tilde{s}_\rho &= n_m + \cdots + n_{m-\rho+1}, \end{aligned}$$

where  $s_\varrho \leq \nu_1 < s_{\varrho+1}$  and  $\tilde{s}_\rho \leq \nu_\omega < \tilde{s}_{\rho+1}$ . In this situation, if the pivot blocks are within  $A_{s_\varrho}^{(k)}$ ,  $B^{(k)}$ ,  $C^{(k)}$ ,  $G^{(k)}$ ,  $A_{\tilde{s}_\rho}^{(k)}$ ,  $k \geq k_1$ , the corresponding steps will make no subscript change in the diagonal elements, or the change will only mean repositions within the same diagonal block. Therefore, our analysis will only consider the central block  $A_{n-s_\varrho-\tilde{s}_\rho}^{(k_1)}$ . The diagonal elements of  $A_{n-s_\varrho-\tilde{s}_\rho}^{(k_1)}$  from  $\mathcal{D}_1$  (resp.  $\mathcal{D}_\omega$ ), if there are any, have already settled within the first (resp. last) positions of  $A_{\varrho+1,\varrho+1}^{(k_1)}$  (resp.  $A_{m-\rho,m-\rho}^{(k_1)}$ ). They will not leave these positions during the next steps.

Then after the following  $m - \varrho - \rho - 1$  or more sweeps, the diagonal elements from  $\mathcal{D}_2$  and  $\mathcal{D}_{\omega-1}$  will settle in. Continuing this consideration we finally obtain the matrix  $A^{(k')}$ ,  $k' > k_0$ , for which relation (A.2) holds. We note that  $k'$  depends on the pivot strategy. For the serial ones, say for the row-cyclic one, the above analysis shows that we can take  $k' = k_0 + M$ .

In order to prove (ii), we consider the diagonalization of the pivot submatrix, which is generally described by relation (2.6). Since we use a globally convergent element-wise Jacobi method, we know that the off-norm sequence of this submatrix tends to zero. In [25] it was proved that the diagonal elements always converge. Thus, the limit  $\text{diag}(\Lambda_{ii}^{(k+1)}, \Lambda_{jj}^{(k+1)})$  exists. Finally, it is known that the diagonal elements are updated by  $\pm \tan \phi_k a_{lm}^{(k)}$ , where  $(l, m)$  is the pivot pair. Therefore, the change is smaller than  $1 \cdot |a_{lm}^{(k)}| \leq \|A_{ij}^{(k)}\|_2 < \sqrt{2}/2 \delta$ . Hence, the diagonal elements cannot change their affiliation to the eigenvalues. This means that the permutation  $P_k$  from relation (A.3) can be taken to be identity.  $\square$

**A.2. Proof of Corollary 2.17.** Set  $\mathcal{R} = \mathcal{R}(\hat{U})$ , where  $\hat{U}$  is as in (2.4). It is sufficient to verify that the transpose of each of the three types of submatrices appearing in Theorem 2.15 is of the same type and possesses the same properties. For the first and the third type, the proof is straightforward:

$$\begin{aligned} \begin{bmatrix} U_{ii}^T \otimes I_{n_r} & U_{ji}^T \otimes I_{n_r} \\ U_{ij}^T \otimes I_{n_r} & U_{jj}^T \otimes I_{n_r} \end{bmatrix}^T &= \begin{bmatrix} U_{ii} \otimes I_{n_r} & U_{ij} \otimes I_{n_r} \\ U_{ji} \otimes I_{n_r} & U_{jj} \otimes I_{n_r} \end{bmatrix} = \begin{bmatrix} V_{ii}^T \otimes I_{n_r} & V_{ji}^T \otimes I_{n_r} \\ V_{ij}^T \otimes I_{n_r} & V_{jj}^T \otimes I_{n_r} \end{bmatrix}, \\ \begin{bmatrix} I_{n_r} \otimes U_{ii}^T & I_{n_r} \otimes U_{ji}^T \\ I_{n_r} \otimes U_{ij}^T & I_{n_r} \otimes U_{jj}^T \end{bmatrix}^T &= \begin{bmatrix} I_{n_r} \otimes U_{ii} & I_{n_r} \otimes U_{ij} \\ I_{n_r} \otimes U_{ji} & I_{n_r} \otimes U_{jj} \end{bmatrix} = \begin{bmatrix} I_{n_r} \otimes V_{ii}^T & I_{n_r} \otimes V_{ji}^T \\ I_{n_r} \otimes V_{ij}^T & I_{n_r} \otimes V_{jj}^T \end{bmatrix}, \end{aligned}$$

with

$$V = \begin{bmatrix} V_{ii} & V_{ij} \\ V_{ji} & V_{jj} \end{bmatrix} = U^T.$$

Note that  $V$  has the same essential properties as  $U$ : dimension and orthogonality (belonging to UBCE( $\varrho$ )). For the second type of submatrices we have

$$\begin{aligned} \begin{bmatrix} I_{n_r} \otimes U_{ii}^T & S(U_{ji}^T \otimes I_{n_r}) \\ \tilde{S}(I_{n_r} \otimes U_{ij}^T) & U_{jj}^T \otimes I_{n_r} \end{bmatrix}^T &= \begin{bmatrix} I_{n_r} \otimes U_{ii} & (I_{n_r} \otimes U_{ij}) \tilde{S}^T \\ (U_{ji} \otimes I_{n_r}) S^T & U_{jj} \otimes I_{n_r} \end{bmatrix} \\ &= \begin{bmatrix} I_{n_r} \otimes V_{ii}^T & S(V_{ji}^T \otimes I_{n_r}) \\ \tilde{S}(I_{n_r} \otimes V_{ij}^T) & V_{jj}^T \otimes I_{n_r} \end{bmatrix}. \end{aligned}$$

To prove the second equality we need some extra work. It is obvious that this equality holds for the corresponding diagonal blocks. To prove that the corresponding (1, 2) blocks are equal, let  $U_{ij} = (u_{st})$ , and note that it is an  $n_i \times n_j$  matrix. If  $e_k^T$  denotes the  $k$ th row of  $I_{n_r}$ , then  $U_{ij}(I_{n_j} \otimes e_k^T)$  is an  $n_i \times n_j n_r$  matrix, and we have

$$\begin{aligned} U_{ij}(I_{n_j} \otimes e_k^T) &= \begin{bmatrix} e_k^T & & \\ & \ddots & \\ & & e_k^T \end{bmatrix} \begin{bmatrix} u_{11} I_{n_r} & \cdots & u_{1n_j} I_{n_r} \\ & \ddots & \\ u_{n_i 1} I_{n_r} & \cdots & u_{n_i n_j} I_{n_r} \end{bmatrix} \\ &= (I_{n_i} \otimes e_k^T) (U_{ij} \otimes I_{n_r}) = (I_{n_i} \otimes e_k^T) (V_{ji}^T \otimes I_{n_r}), \quad 1 \leq k \leq n_r. \end{aligned}$$

Hence,

$$(I_{n_r} \otimes U_{ij}) \tilde{S}^T = \begin{bmatrix} U_{ij}(I_{n_j} \otimes e_1^T) \\ \vdots \\ U_{ij}(I_{n_j} \otimes e_{n_r}^T) \end{bmatrix} = \begin{bmatrix} (I_{n_i} \otimes e_1^T) (V_{ji}^T \otimes I_{n_r}) \\ \vdots \\ (I_{n_i} \otimes e_{n_r}^T) (V_{ji}^T \otimes I_{n_r}) \end{bmatrix} = S(V_{ji}^T \otimes I_{n_r}).$$

Next, let us prove that the corresponding  $(2, 1)$  blocks are equal. Note that  $U_{ji} = V_{ij}^T$  is an  $n_j \times n_i$  matrix. If  $e_k$  denotes the  $k$ th column of  $I_{n_r}$ , then for each  $1 \leq k \leq n_r$  we have

$$(U_{ji} \otimes I_{n_r})(I_{n_i} \otimes e_k) = U_{ji} I_{n_i} \otimes I_{n_r} e_k = U_{ji} \otimes e_k = (I_{n_j} \otimes e_k)(U_{ji} \otimes I_1) = (I_{n_j} \otimes e_k) U_{ji},$$

and all  $(U_{ji} \otimes I_{n_r})(I_{n_i} \otimes e_k)$  are  $n_j n_r \times n_i$  matrices. Now we have

$$\begin{aligned} (U_{ji} \otimes I_{n_r}) S^T &= (U_{ji} \otimes I_{n_r}) [I_{n_i} \otimes e_1 \cdots I_{n_i} \otimes e_{n_r}] \\ &= [(U_{ji} \otimes I_{n_r})(I_{n_i} \otimes e_1) \cdots (U_{ji} \otimes I_{n_r})(I_{n_i} \otimes e_{n_r})] \\ &= [(I_{n_j} \otimes e_1) U_{ji} \cdots (I_{n_j} \otimes e_{n_r}) U_{ji}] \\ &= [I_{n_j} \otimes e_1 \cdots I_{n_j} \otimes e_{n_r}] \begin{bmatrix} U_{ji} & & \\ & \ddots & \\ & & U_{ji} \end{bmatrix} \\ &= \tilde{S}(I_{n_r} \otimes U_{ji}) = \tilde{S}(I_{n_r} \otimes V_{ij}^T). \quad \square \end{aligned}$$

**A.3. Proof of Theorem 3.1.** First we list some inequalities that will be used. Let  $\sigma_{\min}(X)$  and  $\sigma_{\max}(X)$  ( $= \|X\|_2$ ) denote the smallest and largest singular value of  $X$ . Recall that  $\|X\|_F$  denotes the Frobenius norm. Let  $\|X\|$  denote any matrix norm. We have

$$\begin{aligned} \text{(A.4)} \quad & \| \|X\| - \|Y\| \| \leq \|X + Y\| \leq \|X\| + \|Y\|, \\ & \|X_1 + \cdots + X_t\| \leq \|X_1\| + \cdots + \|X_t\|, \quad t \geq 1, \\ & \max\{\sigma_{\min}(F)\|G\|_F, \sigma_{\min}(G)\|F\|_F\} \leq \|FG\|_F \leq \min\{\sigma_{\max}(F)\|G\|_F, \sigma_{\max}(G)\|F\|_F\}, \\ & \sigma_{\min}(X_1 \cdots X_t) \geq \sigma_{\min}(X_1) \cdots \sigma_{\min}(X_t), \quad t \geq 1, \\ \text{(A.5)} \quad & \sigma_{\max}(X_1 \cdots X_t) \leq \sigma_{\max}(X_1) \cdots \sigma_{\max}(X_t), \quad t \geq 1. \end{aligned}$$

Besides, if  $X = (X_{rs})$  is a block matrix as in relation (2.2), then both for the operator matrix norm and also for the Frobenius norm we have

$$\text{(A.6)} \quad \|X_{rs}\| \leq \|X\|, \quad 1 \leq r, s \leq m.$$

To prove Theorem 3.2, we start with the partition  $\pi = (n_1, \dots, n_m)$  and denote by  $\pi_l = (n_1, \dots, n_l)$  the partition of  $s_l = n_1 + \cdots + n_l$ . Obviously, for  $l = m$  we have  $\pi_m = \pi$  and  $s_m = n$ . The set associated with  $\pi_l$  is  $\mathcal{B}_c^{(l)}$  from (3.1) where  $m$  is replaced by  $l$ .

We will prove the following statement. Let  $l \in \{2, \dots, m\}$ , and let  $A$  be any symmetric matrix of order  $s_l$ , carrying the block-matrix partition defined by  $\pi_l$ . Apply to  $A$  the cyclic block Jacobi method defined by the pivot strategy  $I_{\mathcal{O}}$ ,  $\mathcal{O} \in \mathcal{B}_c^{(l)}$ , thus obtaining the symmetric matrices  $A^{(0)} = A, A^{(1)}, \dots$  defined by the recursion (2.5). If the transformation matrices are from the class  $\text{UBCE}_{\pi_l}(\varrho)$ , then

$$\text{(A.7)} \quad S^2(A^{(L)}) \leq \eta_{\pi_l, \varrho} S^2(A), \quad 0 \leq \eta_{\pi_l, \varrho} < \tilde{\eta}_{s_l, \varrho} < 1, \quad L = \frac{l(l-1)}{2},$$

where the constants  $\eta_{\pi_l, \varrho}$  and  $\tilde{\eta}_{s_l, \varrho}$  in (A.7) depend only on  $\pi_l, \varrho$  and  $s_l, \varrho$ , respectively.

Obviously, for  $l = m$  we obtain the assertion of Theorem 3.2. The proof of (A.7) uses induction on  $l, 2 \leq l \leq m$ .

For  $l = 2$ ,  $A$  is of order  $s_2 = n_1 + n_2$ . Its only pivot block is  $A_{12}$ , so that  $\hat{A} = A$ . One step of the block Jacobi method is needed to diagonalize  $A$ . We have  $S^2(A^{(1)}) = 0$ , so  $\eta_{\pi_2, \varrho} = \tilde{\eta}_{s_2, \varrho} = 0$ .



Assume that assertion (A.7) holds for  $l - 1$ ,  $l \in \{3, \dots, m\}$ , and for the partition  $\pi_{l-1}$  with constants  $0 \leq \eta_{\pi_{l-1}, \varrho} \leq \tilde{\eta}_{s_{l-1}, \varrho} < 1$ . In the induction step, we will prove that (A.7) holds for  $l$ .

Let  $A = (A_{rs})$  be a symmetric matrix of order  $s_l$ , partitioned according to  $\pi_l$ . For an arbitrary ordering  $\mathcal{O}$  from  $\mathcal{B}_c^{(l)}$  apply the cyclic block Jacobi method defined by  $I_{\mathcal{O}}$ . Let the transformation matrices be from the class  $\text{UBCE}_{\pi_l}(\varrho)$ , and let the obtained sequence of matrices be denoted by  $A^{(0)} = A, A^{(1)}, \dots$

Let  $\tilde{L} = (l - 1)(l - 2)/2$ . Let  $\tilde{A} (A_{l-1})$  be the leading submatrix of  $A^{(\tilde{L})}$  ( $A$ ) of order  $s_{l-1}$ . In other words,  $\tilde{A}$  is obtained from  $A_{l-1}$  after completing one full sweep of  $\tilde{L}$  Jacobi steps. During these steps, the last,  $l$ th, block-column of  $A$  has been affected only by the left transformations. Therefore, we have

$$\sum_{i=1}^{l-1} \|\tilde{A}_{il}\|_F^2 = \sum_{i=1}^{l-1} \|A_{il}\|_F^2,$$

where  $A^{(\tilde{L})} = (\tilde{A}_{rs})$ . Let  $0 \leq \epsilon \leq 1$  be such that

$$(A.8) \quad (1 - \epsilon^2)S^2(A) = \sum_{i=1}^{l-1} \|A_{il}\|_F^2.$$

Thus,  $S(A_{l-1}) = \epsilon S(A)$ . The submatrix  $A_{l-1}$  is of order  $s_{l-1}$  and carries the block-matrix partition defined by  $\pi_{l-1}$ . Therefore, the induction hypothesis can be applied. It follows that

$$(A.9) \quad \begin{aligned} S^2(A^{(\tilde{L})}) &= S^2(\tilde{A}) + \sum_{i=1}^{l-1} \|\tilde{A}_{il}\|_F^2 \leq \eta_{\pi_{l-1}, \varrho} S^2(A_{l-1}) + \sum_{i=1}^{l-1} \|\tilde{A}_{il}\|_F^2 \\ &= \eta_{\pi_{l-1}, \varrho} \epsilon^2 S^2(A) + (1 - \epsilon^2) S^2(A) \\ &= (1 - \epsilon^2(1 - \eta_{\pi_{l-1}, \varrho})) S^2(A). \end{aligned}$$

Even though we have  $S^2(A^{(L)}) \leq S^2(A^{(\tilde{L})})$ , we cannot set  $\eta_{\pi_l, \varrho} = 1 - \epsilon^2(1 - \eta_{\pi_{l-1}, \varrho})$  because  $\epsilon$  can be arbitrarily small or zero. We still need to estimate the contribution to the off-norm reduction coming from the last  $l - 1$  steps.

According to relation (3.1), the blocks in the  $l$ th block-column are annihilated in the order:  $(\tau_l(1), l), \dots, (\tau_l(l - 1), l)$ . Let us consider how the block  $\tau_l(i)$  changes until it is annihilated in the  $i$ th step. To simplify notation in this analysis, we write  $\tau$  instead of  $\tau_l$  (the permutation of the set  $\{1, \dots, l - 1\}$ ) until relation (A.12). We have

$$\begin{aligned} \tilde{A}_{\tau(i)l}^{(1)} &= \tilde{A}_{\tau(i)l} U_{ll}^{(\tilde{L})} + \tilde{A}_{\tau(i)\tau(1)} U_{\tau(1)l}^{(\tilde{L})}, \\ \tilde{A}_{\tau(i)l}^{(2)} &= \tilde{A}_{\tau(i)l}^{(1)} U_{ll}^{(\tilde{L}+1)} + \tilde{A}_{\tau(i)\tau(2)} U_{\tau(2)l}^{(\tilde{L}+1)}, \\ &\vdots \\ \tilde{A}_{\tau(i)l}^{(i-1)} &= \tilde{A}_{\tau(i)l}^{(i-2)} U_{ll}^{(\tilde{L}+i-2)} + \tilde{A}_{\tau(i)\tau(i-1)} U_{\tau(i-1)l}^{(\tilde{L}+i-2)}, \\ \tilde{A}_{\tau(i)l}^{(i)} &= 0. \end{aligned}$$

The contribution to the off-norm reduction comes from  $\|\tilde{A}_{\tau(i)l}^{(i-1)}\|_F$ , so we have to estimate it from below. To express  $\tilde{A}_{\tau(i)l}^{(i-1)}$  in terms of the blocks from  $\tilde{A}$ , we multiply the

equation for  $\tilde{A}_{\tau(i)l}^{(i-2)}$  from the right by  $U_{ll}^{(\tilde{L}+i-2)}$ , then multiply the equation for  $\tilde{A}_{\tau(i)l}^{(i-3)}$  by  $U_{ll}^{(\tilde{L}+i-3)}U_{ll}^{(\tilde{L}+i-2)}$ , etc. Finally, we multiply the first equation by  $U_{ll}^{(\tilde{L}+1)} \dots U_{ll}^{(\tilde{L}+i-2)}$  from the right. Then we take the sum of the obtained equations. It follows that

$$\tilde{A}_{\tau(i)l}^{(i-1)} = \tilde{A}_{\tau(i)l} U_{ll}^{(\tilde{L})} U_{ll}^{(\tilde{L}+1)} \dots U_{ll}^{(\tilde{L}+i-2)} + \sum_{k=1}^{i-1} \tilde{A}_{\tau(i)\tau(k)} U_{\tau(k)l}^{(\tilde{L}+k-1)} U_{ll}^{(\tilde{L}+k)} \dots U_{ll}^{(\tilde{L}+i-2)}.$$

Denote

$$(A.10) \quad \zeta_l = \min_{0 \leq k \leq l-3} \{ \sigma_{\min}(U_{ll}^{(\tilde{L}+k)}) \}.$$

Using the inequalities (A.4)–(A.5), for  $1 \leq i \leq l-1$ , we obtain

$$\begin{aligned} \|\tilde{A}_{\tau(i)l}^{(i-1)}\|_F &\geq \left\| \tilde{A}_{\tau(i)l} U_{ll}^{(\tilde{L})} \dots U_{ll}^{(\tilde{L}+i-2)} \right\|_F - \left\| \sum_{k=1}^{i-1} \tilde{A}_{\tau(i)\tau(k)} U_{\tau(k)l}^{(\tilde{L}+k-1)} \dots U_{ll}^{(\tilde{L}+i-2)} \right\|_F \\ &\geq \sigma_{\min}(U_{ll}^{(\tilde{L})} \dots U_{ll}^{(\tilde{L}+i-2)}) \|\tilde{A}_{\tau(i)l}\|_F - \sum_{k=1}^{i-1} \sigma_{\max}(U_{\tau(k)l}^{(\tilde{L}+k-1)} \dots U_{ll}^{(\tilde{L}+i-2)}) \|\tilde{A}_{\tau(i)\tau(k)}\|_F \\ &\geq \sigma_{\min}(U_{ll}^{(\tilde{L})}) \sigma_{\min}(U_{ll}^{(\tilde{L}+1)}) \dots \sigma_{\min}(U_{ll}^{(\tilde{L}+i-2)}) \|\tilde{A}_{\tau(i)l}\|_F - \\ &\quad - \sum_{k=1}^{i-1} \sigma_{\max}(U_{\tau(k),l}^{(\tilde{L}+k-1)}) \sigma_{\max}(U_{ll}^{(\tilde{L}+k)}) \dots \sigma_{\max}(U_{ll}^{(\tilde{L}+i-2)}) \|\tilde{A}_{\tau(i)\tau(k)}\|_F \\ &\geq \zeta_l^{i-1} \|\tilde{A}_{\tau(i),l}\|_F - \sum_{k=1}^{i-1} \|\tilde{A}_{\tau(i)\tau(k)}\|_F. \end{aligned}$$

Here, we have used (A.6) for the transformation matrices, which are orthogonal. Squaring the obtained inequality and then using  $(a-b)^2 \geq \frac{1}{2}a^2 - b^2$ ,  $a, b \in \mathbb{R}$ , and the Cauchy–Schwarz inequality, we obtain

$$(A.11) \quad \|\tilde{A}_{\tau(i)l}^{(i-1)}\|_F^2 \geq \frac{1}{2} \zeta_l^{2(i-1)} \|\tilde{A}_{\tau(i)l}\|_F^2 - (i-1) \sum_{k=1}^{i-1} \|\tilde{A}_{\tau(i)\tau(k)}\|_F^2.$$

Now we have the lower bound for the reduction of  $S^2(A^{(\tilde{L})})$  coming from just one annihilated block in the last block-column. The lower bound coming from all blocks in the last block-column is obtained by summing up these. Using the relations (A.8), (A.9), and (A.11), we have

$$\begin{aligned} (A.12) \quad \sum_{i=1}^{l-1} \|\tilde{A}_{\tau(i)l}^{(i-1)}\|_F^2 &\geq \frac{1}{2} \sum_{i=1}^{l-1} \zeta_l^{2(i-1)} \|\tilde{A}_{\tau(i)l}\|_F^2 - (l-2) \sum_{i=1}^{l-1} \sum_{k=1}^{i-1} \|\tilde{A}_{\tau(i)\tau(k)}\|_F^2 \\ &\geq \frac{1}{2} \zeta_l^{2(l-2)} \sum_{i=1}^{l-1} \|\tilde{A}_{\tau(i)l}\|_F^2 - (l-2) S^2(\tilde{A}) \\ &\geq \frac{1}{2} \zeta_l^{2(l-2)} (1 - \epsilon^2) S^2(A) - (l-2) \eta_{\pi_{l-1}, \varrho} S^2(A_{l-1}) \\ &= \frac{1}{2} \zeta_l^{2(l-2)} (1 - \epsilon^2) S^2(A) - (l-2) \eta_{\pi_{l-1}, \varrho} \epsilon^2 S^2(A) \\ &= f(\epsilon) S^2(A), \end{aligned}$$

where the function  $f : [0, 1] \rightarrow \mathbb{R}$  is defined by

$$f(\epsilon) = \frac{1}{2}\zeta_l^{2(l-2)} - \left( \frac{1}{2}\zeta_l^{2(l-2)} + l\eta_{\pi_{l-1},\varrho} - 2\eta_{\pi_{l-1},\varrho} \right) \epsilon^2.$$

The first derivative of  $f$  is not positive since  $l \geq 2$ . Recall that the transformation matrices are  $\text{UBCE}_{\pi_l}(\varrho)$  and satisfy the relations (2.15) and (2.16). Therefore, we have

$$(A.13) \quad 1 \geq \zeta_l = \min_{0 \leq k \leq l-3} \left\{ \sigma_{\min}(U_{ll}^{(\tilde{L}+k)}) \right\} \geq \varrho \min_{1 \leq i < j \leq l} \gamma_{ij} > \frac{3\sqrt{2}\varrho}{\sqrt{4^{sl} + 26}} > 0.$$

This implies that the first derivative is negative, i.e.,  $f$  is decreasing. Its maximum is at  $\epsilon = 0$ , and its zero is

$$\epsilon_l = \sqrt{\frac{\zeta_l^{2(l-1)}}{\zeta_l^{2(l-1)} + 2l\eta_{\pi_{l-1},\varrho} - 4\eta_{\pi_{l-1},\varrho}}}.$$

In relation (A.12) the left-hand side is nonnegative. Therefore, it is better to replace  $f$  by a nonnegative function  $f_+$ , such that

$$f_+(\epsilon) = \begin{cases} f(\epsilon), & \epsilon \in [0, \epsilon_l], \\ 0, & \epsilon \in [\epsilon_l, 1]. \end{cases}$$

Then  $f'_+(\epsilon) \leq 0$  for  $\epsilon \neq \epsilon_l$  and  $f_+(\epsilon) \geq f(\epsilon)$  for  $0 \leq \epsilon \leq 1$ . From the relations (A.9) and (A.12), we have

$$\begin{aligned} S^2(A^{(L)}) &= S^2(A^{(\tilde{L})}) - \sum_{i=1}^{l-1} \|\tilde{A}_{\tau(i),l}^{(i)}\|_F^2 \leq S^2(A^{(\tilde{L})}) - f_+(\epsilon)S^2(A) \\ &\leq (1 - \epsilon^2(1 - \eta_{\pi_{l-1},\varrho}) - f_+(\epsilon))S^2(A) = g(\epsilon)S(A), \end{aligned}$$

where

$$g(\epsilon) = \begin{cases} 1 - \frac{1}{2}\zeta_l^{2(l-1)} + \epsilon^2(l\eta_{\pi_{l-1},\varrho} - \eta_{\pi_{l-1},\varrho} + \frac{1}{2}\zeta_l^{2(l-1)} - 1), & \epsilon \in [0, \epsilon_l], \\ 1 - \epsilon^2(1 - \eta_{\pi_{l-1},\varrho}), & \epsilon \in [\epsilon_l, 1]. \end{cases}$$

The function  $g$  is differentiable on  $\langle 0, 1 \rangle \setminus \{\epsilon_l\}$  and one has

$$g'(\epsilon) = \begin{cases} \epsilon(2l\eta_{\pi_{l-1},\varrho} - 2\eta_{\pi_{l-1},\varrho} + \zeta_l^{2(l-1)} - 2), & \epsilon \in \langle 0, \epsilon_l \rangle, \\ 2\epsilon(\eta_{\pi_{l-1},\varrho} - 1), & \epsilon \in \langle \epsilon_l, 1 \rangle. \end{cases}$$

For  $\epsilon \in \langle \epsilon_l, 1 \rangle$  we have  $g'(\epsilon) < 0$  since  $\eta_{\pi_{l-1},\varrho} - 1 < 0$ . For  $\epsilon \in \langle 0, \epsilon_l \rangle$ ,  $g'(\epsilon)$  is either positive or negative on the whole interval  $\langle 0, 1 \rangle$ , depending on  $l$ . We conclude that  $g$  is either decreasing on whole segment  $[0, 1]$ , or increasing on  $[0, \epsilon_l]$  and decreasing on  $\langle \epsilon_l, 1 \rangle$ . Thus,  $g$  attains its maximum either at  $\epsilon = 0$  or at  $\epsilon = \epsilon_l$ . Therefore, we have

$$\eta_{\pi_l,\varrho} = \max \{g(0), g(\epsilon_l)\} = \max \left\{ 1 - \frac{1}{2}\zeta_l^{2(l-1)}, 1 - \frac{(1 - \eta_{\pi_{l-1},\varrho})\zeta_l^{2(l-1)}}{\zeta_l^{2(l-1)} + 2(l-2)\eta_{\pi_{l-1},\varrho}} \right\}.$$

From relation (A.13) we see that  $\zeta_l$  is bounded from below by a positive constant that depends on  $\pi_l$  and  $\varrho$ . Therefore, the constant  $\eta_{\pi_l,\varrho}$  depends on  $\pi_l$  and  $\varrho$  and  $0 \leq \eta_{\pi_l,\varrho} < 1$ .

It remains to show that there is a bound  $\tilde{\eta}_{s_l, \varrho}$  for  $S^2(A^{(L)})/S^2(A)$  depending only on  $s_l$  and  $\varrho$  such that  $\eta_{\pi_l, \varrho} \leq \tilde{\eta}_{s_l, \varrho} < 1$ . It will be derived from  $\eta_{\pi_l, \varrho}$ .

If  $\eta_{\pi_l, \varrho} = g(0)$ , then relation (A.13) implies

$$\eta_{\pi_l, \varrho} < 1 - \frac{1}{2} \left( \frac{3\sqrt{2}\varrho}{\sqrt{4^{s_l} + 26}} \right)^{2(l-1)} < 1 - \frac{1}{2} \left( \frac{3\sqrt{2}\varrho}{\sqrt{4^{s_l} + 26}} \right)^{2(s_l-1)} \equiv \eta'_{s_l, \varrho}.$$

If  $\eta_{\pi_l, \varrho} = g(\epsilon_l)$ , then

$$\eta_{\pi_l, \varrho} = \frac{2(l-2) + \zeta_l^{2(l-1)}}{\zeta_l^{2(l-1)} + 2(l-2)\eta_{\pi_{l-1}, \varrho}} \eta_{\pi_{l-1}, \varrho},$$

which implies  $\eta_{\pi_{l-1}, \varrho} \leq \eta_{\pi_l, \varrho}$ . Therefore

$$\eta_{\pi_l, \varrho} = 1 - \frac{(1 - \eta_{\pi_{l-1}, \varrho}) \zeta_l^{2(l-1)}}{\zeta_l^{2(l-1)} + 2(l-2)\eta_{\pi_{l-1}, \varrho}} \leq 1 - \frac{(1 - \eta_{\pi_{l-1}, \varrho}) \zeta_l^{2(l-1)}}{\zeta_l^{2(l-1)} + 2(l-2)} = \frac{2(l-2) + \eta_{\pi_{l-1}, \varrho} \zeta_l^{2(l-1)}}{2(l-2) + \zeta_l^{2(l-1)}},$$

which is equivalent to  $\eta_{\pi_l, \varrho} \leq (l-2)/(\zeta_l^{2(l-1)} + l-2)$ . By (A.13) we have

$$\eta_{\pi_l, \varrho} \leq \frac{l-2}{\zeta_l^{2(l-1)} + l-2} < \frac{l-2}{\left(\frac{3\sqrt{2}\varrho}{\sqrt{4^{s_l} + 26}}\right)^{2(l-1)} + l-2} < \frac{s_l-2}{\left(\frac{3\sqrt{2}\varrho}{\sqrt{4^{s_l} + 26}}\right)^{2(l-1)} + s_l-2} \equiv \eta''_{s_l, \varrho}.$$

Finally, set

$$\tilde{\eta}_{n, \varrho} = \max \{ \eta'_{s_l, \varrho}, \eta''_{s_l, \varrho} \}.$$

This completes the induction step and the proof of assertion (A.7).  $\square$

**A.4. Proof of Corollary 3.3.** We follow the proof of Theorem 3.1 and use the notation from there. Now the blocks become the elements and we can use sharper estimates for the rotation angles. In particular, we can replace the lower bound  $\varrho\gamma_{ij}$  from relation (2.16) by  $\sqrt{2}/2$ . We have  $m = n$ ,  $2 \leq l \leq n$ , and

$$\zeta_l = \min_{0 \leq k \leq l-1} \{ \sigma_{\min}(U_l^{(\tilde{L}+k)}) \} = \min_{0 \leq k \leq l-1} \cos \phi_{\tilde{L}+k} \geq \frac{\sqrt{2}}{2}.$$

Hence, applying that lower bound for  $\zeta_l$ , using the notation  $\eta_l$  for  $\eta_{\pi_l}$ , we obtain

$$\eta_l = \max \{ g(0), g(\epsilon_l) \},$$

where

$$g(0) = 1 - 2^{-l}, \quad g(\epsilon_l) = 1 - \frac{(1 - \eta_{l-1})2^{-l}}{2^{-l} + (l-2)\eta_{l-1}}.$$

This yields the constant  $\eta_n$  by replacing  $l$  by  $n$ .

If the whole analysis is performed on the elements, a somewhat larger constant  $\eta_n$  can be obtained. In [1] it has been shown that

$$\eta_n = \max \left\{ 1 - 2^{1-n}, 1 - \frac{2^{2-n}(1 - \eta_{n-1})}{2^{2-n} + (n-2)\eta_{n-1}} \right\}. \quad \square$$

**A.5. Proof of Theorem 4.1.** The proof is similar to the proof of Theorem 3.1, so let us follow the same lines. The notation  $\pi$ ,  $\pi_l$ , and  $s_l$  has the same meaning as before. The proof uses induction on  $l$ ,  $2 \leq l \leq m$ . Relation (A.7) remains the same, except for  $L$ , which has to take the actual number of steps into account. The same is true for  $\tilde{L}$ . Since  $\mathcal{O} \in \tilde{\mathcal{B}}_c^{(m)}$ , we have

$$L = \frac{l(l-1)}{2} + |\mathcal{O}_3| + \dots + |\mathcal{O}_l|, \quad \tilde{L} = \frac{(l-1)(l-2)}{2} + |\mathcal{O}_3| + \dots + |\mathcal{O}_{l-1}|.$$

As earlier, the matrix  $\tilde{A}$  ( $A_{l-1}$ ) is the leading submatrix of  $A^{(\tilde{L})}$  ( $A$ ) of order  $s_{l-1}$ . Until the end, the proof uses the same lines as the proof of Theorem 3.1. We only note that the final estimate is first obtained for the matrix  $A^{(L-|\mathcal{O}_l|)}$ , but since  $S(A^{(L)}) \leq S(A^{(L-|\mathcal{O}_l|)})$ , it automatically holds for  $S(A^{(L)})$ .  $\square$

**A.6. Proof of Theorem 5.1.** The proof is similar to the proof of [18, Theorem 5.1]. The only difference is that here the considered matrices are real and therefore the block Jacobi annihilators and operators are of order  $K$  (where  $K$  is from relation (2.18)) and not  $2K$  as they are in [18]. Also, the iterative process (5.1) uses the congruence transformation, while in [18] it uses the equivalence transformation. Finally, here we show how the parameter  $\varrho$  is used to avoid the assumption that the matrices  $U_k$  from A2 have to be UBCE. Hence, for the completeness of the paper, we will present a somewhat shorter version of the proof, often referring to the proof of [18, Theorem 5.1]. The complete proof can be found in the thesis [1].

Using the relation  $F_k = U_k + (F_k - U_k)$ ,  $k \geq 0$ , and assumption A2 it is easy to transform the process (5.1) into the form

$$(A.14) \quad A^{(k+1)} = U_k^T A^{(k)} U_k + E^{(k)}, \quad k \geq 0,$$

where the ‘‘perturbation’’ matrices  $E^{(k)}$  satisfy

$$(A.15) \quad \lim_{k \rightarrow \infty} \frac{E^{(k)}}{\|A^{(k)}\|_F} = 0.$$

The matrices  $A$ ,  $A^{(k)}$ ,  $U_k$ ,  $E^{(k)}$ ,  $k \geq 0$ , carry matrix block-partition defined by  $\pi$ . Applying the function  $\text{vec}_\pi$  to both sides of equation (A.14) and using (A.15) together with condition (i), one obtains (cf. [18, Lemma 5.2])

$$(A.16) \quad a^{(k+1)} = \mathcal{R}^{(k)} a^{(k)} + g^{(k)}, \quad k \geq 0,$$

and

$$(A.17) \quad \lim_{k \rightarrow \infty} \frac{g^{(k)}}{\|A^{(k)}\|_F} = 0.$$

Here,  $\mathcal{R}^{(k)}$  is the block Jacobi annihilator determined by the pivot submatrix  $\hat{U}_k$  of  $U_k$  and the pivot pair  $(i(k), j(k))$ , while  $g^{(k)} = \text{vec}(H^{(k)}) + \text{vec}(E^{(k)})$ . The matrix  $H^{(k)}$  of order  $n$  carries the same partition as  $A^{(k)}$ , which differs from the zero-matrix only in the pivot submatrix of order  $n_i + n_j$  where it equals  $\hat{U}_k^T \hat{A}_{ij}^{(k)} \hat{U}_k$ .

Recall that one cycle consists of  $M$  steps. After the first cycle has been completed, relation (A.16) implies that we can write  $a^{[1]} = \mathcal{J}^{[1]} a^{(0)} + g^{[1]}$ , where  $a^{[1]} = a^{(M)}$ ,  $\mathcal{J}^{[1]} = \mathcal{R}^{(M-1)} \dots \mathcal{R}^{(0)}$ , and

$$g^{[1]} = g^{(M-1)} + \sum_{k=0}^{M-2} \mathcal{R}^{(M-1)} \dots \mathcal{R}^{(k+1)} g^{(k)}.$$

By Theorem 2.15 we have  $\|\mathcal{R}^{(k)}\|_2 \leq 1$  for all  $k$ . Hence,  $\|g^{[1]}\| \leq \|g^{(0)}\| + \dots + \|g^{(M-1)}\|$ . Similarly, after  $s$  cycles we have

$$(A.18) \quad a^{[s]} = \mathcal{J}^{[s]} a^{[s-1]} + g^{[s]}, \quad s \geq 1,$$

with  $a^{[s]} = a^{(sM)}$ ,  $\mathcal{J}^{[s]} = \mathcal{R}^{(sM-1)} \dots \mathcal{R}^{((s-1)M)}$  and

$$(A.19) \quad \|g^{[s]}\| \leq \|g^{((s-1)M)}\| + \dots + \|g^{(sM-1)}\|.$$

We will also write  $A^{[s]} = A^{(sM)}$ , so that  $a^{[s]} = \text{vec}(A^{[s]})$ ,  $s \geq 0$ .

Using assumption **A2** and condition (i), it is easy to prove that (see [18, Lemma 5.3])

$$(A.20) \quad \lim_{k \rightarrow \infty} \frac{\|A^{(k+i)}\|_F}{\|A^{(k)}\|_F} = 1 \quad \text{for each } i \geq 0,$$

$$(A.21) \quad \lim_{s \rightarrow \infty} \frac{g^{[s]}}{\|A^{[s]}\|_F} = 0,$$

$$(A.22) \quad \lim_{k \rightarrow \infty} \frac{a^{(k)}}{\|A^{(k)}\|_F} = 0 \quad \text{iff} \quad \lim_{s \rightarrow \infty} \frac{a^{[s]}}{\|A^{[s]}\|_F} = 0.$$

Relation (A.20) is implied by relations (A.14) and (A.15), while relation (A.21) follows directly from (A.19), (A.20), and (A.17). Relation (A.22) is implied by (A.17) and (A.20). From (A.22) it follows that, to prove  $\|a^{(k)}\|/\|A^{(k)}\|_F \rightarrow 0$  as  $k \rightarrow \infty$ , it is sufficient to show that  $\lim_{s \rightarrow \infty} b^{[s]} = 0$  for

$$(A.23) \quad b^{[s]} = \frac{a^{[s]}}{\|A^{[s]}\|_F}, \quad s \geq 0.$$

We transform the iterative process (A.18) into

$$(A.24) \quad b^{[s]} = \mathcal{J}^{[s]} b^{[s-1]} + c^{[s]}, \quad s \geq 1,$$

where

$$(A.25) \quad c^{[s]} = \left( \frac{\|A^{[s-1]}\|_F}{\|A^{[s]}\|_F} - 1 \right) \mathcal{J}^{[s]} b^{[s-1]} + \frac{g^{[s]}}{\|A^{[s]}\|_F}, \quad s \geq 0.$$

By taking the norm of both sides of equation (A.25), we obtain

$$\|c^{[s]}\| \leq \left| \frac{\|A^{[s-1]}\|_F}{\|A^{[s]}\|_F} - 1 \right| \|\mathcal{J}^{[s]}\|_2 \|b^{[s-1]}\| + \frac{\|g^{[s]}\|}{\|A^{[s]}\|_F}, \quad s \geq 0.$$

Relation (A.23) and Theorem 2.15 imply  $\|b^{[s-1]}\| \leq 1$  and  $\|\mathcal{J}^{[s]}\|_2 \leq 1$  for all  $s \geq 0$ . Hence, from (A.20) and (A.21) we have

$$(A.26) \quad \lim_{s \rightarrow \infty} c^{[s]} = 0.$$

The proofs of the preceding relations also hold for any quasi-cyclic Jacobi-type process satisfying assumption **A2** and condition (i) of the theorem. To prove

$$\lim_{s \rightarrow \infty} b^{[s]} = 0,$$

we will additionally use assumptions **A3** and **A1**.

Since **A3** holds, it implies  $\sigma > 0$ . From the definition of  $\sigma$ , we know that there exists  $s_0 \geq 1$  such that

$$\|F_k - U_k\|_2 \leq \frac{1}{4}\sigma \quad \text{and} \quad \sigma_{\min}(F_{ii}^{(k)}) \geq \frac{3}{4}\sigma, \quad s \geq s_0M.$$

By the perturbation theorem for the singular values, we have

$$\begin{aligned} \sigma_{\min}(U_{ii}^{(k)}) &= \sigma_{\min}(F_{ii}^{(k)} - (F_{ii}^{(k)} - U_{ii}^{(k)})) \geq \frac{3}{4}\sigma - \|F_{ii}^{(k)} - U_{ii}^{(k)}\|_2 \\ &\geq \frac{3}{4}\sigma - \|F_k - U_k\|_2 \geq \frac{3}{4}\sigma - \frac{1}{4}\sigma = \frac{1}{2}\sigma, \quad k \geq s_0M. \end{aligned}$$

Set  $\varrho = \frac{1}{2}\sigma$ . We have proved that  $\widehat{U}_k$  belongs to the class  $\text{UBCE}(\varrho)$  provided that  $k \geq s_0M$ . Then, by Definition 2.18, the block Jacobi operators  $\mathcal{J}^{[s]}$ ,  $s \geq s_0$ , from (A.24) are in  $\mathcal{J}_{\mathcal{O}}^{\text{UBCE}(\varrho)}$ .

Next, we have to use assumption **A1**. Since  $\mathcal{O} \in \mathcal{B}_{s_g}^{(m)}$ , we can presume that the chain connecting  $\mathcal{O}$  to  $\mathcal{O}'' \in \mathcal{B}_{s_p}^{(m)}$  (see Definition 3.7) is in canonical form and contains  $d$  shift equivalences. Without loss of generality, we may assume (as in Theorem 3.9) that  $\mathcal{O} \stackrel{\mathcal{P}}{\sim} \mathcal{O}' \stackrel{\mathcal{W}}{\sim} \mathcal{O}'' \in \mathcal{B}_{s_p}^{(m)}$  with  $\mathcal{O}' = \mathcal{O}(\mathfrak{q})$ , or  $\mathcal{O} \stackrel{\mathcal{W}}{\sim} \mathcal{O}' \stackrel{\mathcal{P}}{\sim} \mathcal{O}'' \in \mathcal{B}_{s_p}^{(m)}$  with  $\mathcal{O}'' = \mathcal{O}'(\mathfrak{q})$ , for some permutation  $\mathfrak{q}$  of the set  $\mathcal{S}_m$ .

Applying Theorem 3.9, one concludes that there are constants  $\mu_{\pi_{\mathfrak{q}}, \varrho}$  and  $\tilde{\mu}_{n, \varrho}$  depending only on  $\pi_{\mathfrak{q}}$ ,  $\varrho$  and  $n$ ,  $\varrho$ , respectively, such that

$$(A.27) \quad \|\mathcal{J}^{[s+d]} \dots \mathcal{J}^{[s+1]} \mathcal{J}^{[s]}\|_2 \leq \mu_{\pi_{\mathfrak{q}}, \varrho}, \quad 0 \leq \mu_{\pi_{\mathfrak{q}}, \varrho} < \tilde{\mu}_{n, \varrho} < 1, \quad s \geq s_0.$$

By unfolding the recursion (A.24)  $d$  times, similarly as in the proof of [18, Theorem 5.1], one obtains

$$b^{[s+d]} = \mathcal{J}^{[s+d]} \mathcal{J}^{[s+d-1]} b^{[s-1]} \dots \mathcal{J}^{[s]} b^{[s-1]} + h^{[s]}, \quad s \geq s_0,$$

where  $h^{[s]} \rightarrow 0$  as  $s \rightarrow \infty$ . Here, we have used (A.26) and Theorem 2.15. Taking the Euclidean norm, it follows that

$$\|b^{[s+d]}\| \leq \|\mathcal{J}^{[s+d]} \mathcal{J}^{[s+d-1]} b^{[s-1]} \dots \mathcal{J}^{[s]}\|_2 \|b^{[s-1]}\| + \|h^{[s]}\|, \quad s \geq s_0.$$

This inequality together with (A.27) implies

$$(A.28) \quad \beta_{s+d} \leq \mu_{\pi_{\mathfrak{q}}, \varrho} \beta_{s-1} + \varepsilon_s, \quad s \geq s_0, \quad \text{with} \quad \lim_{s \rightarrow \infty} \varepsilon_s = 0,$$

where  $0 \leq \mu_{\pi_{\mathfrak{q}}, \varrho} < 1$  and  $\beta_s = \|b^{[s]}\|$ ,  $\varepsilon_s = \|h^{[s]}\|$ ,  $s \geq s_0$ . Set  $\alpha_t = \beta_{s_0-1+t(d+1)}$  and  $\eta_t = \varepsilon_{s_0+t(d+1)}$ ,  $t \geq 0$ . Then from relation (A.28) it follows that

$$\alpha_{t+1} \leq \mu_{\pi_{\mathfrak{q}}, \varrho} \alpha_t + \eta_t, \quad t \geq 0, \quad \text{with} \quad \lim_{t \rightarrow \infty} \eta_t = 0.$$

This enables us to apply [12, Lemma 1] to obtain  $\lim_{t \rightarrow \infty} \alpha_t = 0$ , i.e.,

$$\lim_{t \rightarrow \infty} \beta_{s_0-1+t(d+1)} = 0.$$

Relations (A.24) and (A.26) imply

$$\beta_{s_0-1+t(d+1)+r} \leq \beta_{s_0-1+t(d+1)} + \vartheta_{r,t}, \quad 0 \leq r \leq d, \quad t \geq 0,$$

with  $\lim_{t \rightarrow \infty} \vartheta_{r,t} = 0$  for any  $0 \leq r \leq d$ . This proves  $\lim_{s \rightarrow \infty} b^{[s]} = 0$  and, because of (A.22), it also proves  $\lim_{k \rightarrow \infty} a^{(k)} / \|A^{(k)}\|_F = 0$ .

Therefore, it suffices to show that  $\lim_{k \rightarrow \infty} \frac{\sum_{l=1}^m S^2(A_{ll}^{(k)})}{\|A^{(k)}\|_F^2} = 0$ . Let  $\varepsilon > 0$ . Then by condition (i) of the theorem and by relation (A.20), there is an integer  $k_\varepsilon$  such that (cf. [18, Theorem 5.1])

$$(A.29) \quad \frac{S(\widehat{A}_{ij}^{(k+1)})}{\|A^{(k)}\|_F} \leq \varepsilon, \quad \frac{\|A^{(p)}\|_F}{\|A^{(k)}\|_F} \leq 1 + \varepsilon, \quad k - M \leq p < k, \quad k \geq k_\varepsilon.$$

Here  $M$  is the number of steps within one cycle. For given  $k \geq k_\varepsilon + M$  and  $l \in \{1, \dots, m\}$ , let  $q < k$  denote the last step when  $A_{ll}^{(q)}$  was a part of some pivot submatrix. Obviously,  $k_\varepsilon \leq q < k$ . Relation (A.29) implies

$$\frac{S(A_{ll}^{(k)})}{\|A^{(k)}\|_F} = \frac{S(A_{ll}^{(q+1)})}{\|A^{(k)}\|_F} = \frac{S(A_{ll}^{(q+1)})}{\|A^{(q)}\|_F} \frac{\|A^{(q)}\|_F}{\|A^{(k)}\|_F} \leq \varepsilon(1 + \varepsilon),$$

for any  $1 \leq l \leq m$  and any  $k \geq k_\varepsilon + M$ . This proves the theorem.  $\square$

#### REFERENCES

- [1] E. BEGOVIĆ, *Convergence of Block Jacobi Methods*, Ph.D Thesis, Department of Mathematics, University of Zagreb, Zagreb, 2014.
- [2] E. BEGOVIĆ KOVAČ AND V. HARI, *Jacobi method for symmetric matrices of order 4 converges for every cyclic pivot strategy*, Preprint on arXiv, 2017. <https://arxiv.org/abs/1701.02387>
- [3] Z. BUJANOVIĆ AND Z. DRMAČ, *A contribution to the theory and practise of the block Kogbetliantz method for computing the SVD*, BIT, 52 (2012), pp. 827–849.
- [4] J. DEMMEL AND K. VESELIĆ, *Jacobi's method is more accurate than QR*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 1204–1245.
- [5] Z. DRMAČ, *A global convergence proof of cyclic Jacobi methods with block rotations*, SIAM J. Matrix Anal. Appl., 31 (2009), pp. 1329–1350.
- [6] Z. DRMAČ AND V. HARI, *On quadratic convergence bounds for the J-symmetric Jacobi method*, Numer. Math., 64 (1993), pp. 147–180.
- [7] Z. DRMAČ AND K. VESELIĆ, *New fast and accurate Jacobi SVD algorithm. I*, SIAM J. Matrix Anal. Appl., 29 (2007), pp. 1322–1342.
- [8] ———, *New fast and accurate Jacobi SVD algorithm. II*, SIAM J. Matrix Anal. Appl., 29 (2007), pp. 1343–1362.
- [9] K. V. FERNANDO, *Linear convergence of the row cyclic Jacobi and Kogbetliantz methods*, Numer. Math., 56 (1989), pp. 73–91.
- [10] G. E. FORSYTHE AND P. HENRICI, *The cyclic Jacobi method for computing the principal values of a complex matrix*, Trans. Amer. Math. Soc., 94 (1960), pp. 1–23.
- [11] E. R. HANSEN, *On cyclic Jacobi methods*, J. Soc. Indust. Appl. Math., 11 (1963), pp. 448–459.
- [12] V. HARI, *On the global convergence of the Eberlein method for real matrices*, Numer. Math., 39 (1982), pp. 361–369.
- [13] ———, *On Cyclic Jacobi Methods for the Positive Definite Generalized Eigenvalue Problem*, Ph.D. Thesis, University of Hagen, Hagen, 1984.
- [14] ———, *On the convergence of cyclic Jacobi-like processes*, Linear Algebra Appl., 81 (1986), pp. 105–127.
- [15] ———, *On sharp quadratic convergence bounds for the serial Jacobi methods*, Numer. Math., 60 (1991), pp. 375–406.
- [16] ———, *Convergence of a block-oriented quasi-cyclic Jacobi method*, SIAM J. Matrix Anal. Appl., 29 (2007), pp. 349–369.
- [17] ———, *On block Jacobi annihilators*, in Proceedings of the Conference Algorithmy 2009, A. Handlovicova, K. Mikula, P. Frolkovic, and D. Sevcovic, eds., Publishing House of the Slovak University of Technology in Bratislava, Bratislava, 2009, pp. 429–439.
- [18] ———, *Convergence to diagonal form of block Jacobi-type methods*, Numer. Math., 129 (2015), pp. 449–481.
- [19] V. HARI, S. SINGER, AND S. SINGER, *Block-oriented J-Jacobi methods for Hermitian matrices*, Linear Algebra Appl., 433 (2010), pp. 1491–1512.



- [20] ———, *Full block  $J$ -Jacobi method for Hermitian matrices*, *Linear Algebra Appl.*, 444 (2014), pp. 1–27.
- [21] V. HARI AND K. VESELIĆ, *On Jacobi methods for singular value decompositions*, *SIAM J. Sci. Statist. Comput.*, 8 (1987), pp. 741–754.
- [22] P. HENRICI AND K. ZIMMERMANN, *An estimate for the norms of certain cyclic Jacobi operators*, *Linear Algebra and Appl.*, 1 (1968), pp. 489–501.
- [23] C. G. J. JACOBI, *Über ein leichtes Verfahren die in der Theorie der Säcularstörungen vorkommenden Gleichungen numerisch aufzulösen*, *J. Reine Angew. Math.*, 30 (1846), pp. 51–94.
- [24] F. T. LUK AND H. PARK, *On parallel Jacobi orderings*, *SIAM J. Sci. Statist. Comput.*, 10 (1989), pp. 18–26.
- [25] W. F. MASCARENHAS, *On the convergence of the Jacobi method for arbitrary orderings*, *SIAM J. Matrix Anal. Appl.*, 16 (1995), pp. 1197–1209.
- [26] J. MATEJAŠ, *Accuracy of the Jacobi method on scaled diagonally dominant symmetric matrices*, *SIAM J. Matrix Anal. Appl.*, 31 (2009), pp. 133–153.
- [27] R. MATHIAS, *Accurate eigensystem computations by Jacobi methods*, *SIAM J. Matrix Anal. Appl.*, 16 (1995), pp. 977–1003.
- [28] L. NAZARETH, *On the convergence of the cyclic Jacobi method*, *Linear Algebra and Appl.*, 12 (1975), pp. 151–164.
- [29] V. NOVAKOVIĆ, S. SINGER AND S. SINGER, *Blocking and parallelization of the Hari–Zimmermann variant of the Falk–Langemeyer algorithm for the generalized SVD*, *Parallel Comput.*, 49 (2015), pp. 136–152.
- [30] N. H. RHEE AND V. HARI, *On the global and cubic convergence of a quasi-cyclic Jacobi method*, *Numer. Math.*, 66 (1993), pp. 97–122.
- [31] A. H. SAMEH, *On Jacobi and Jacobi-like algorithms for parallel computer*, *Math. Comp.*, 25 (1971), pp. 579–590.
- [32] G. SHROFF AND R. SCHREIBER, *On the convergence of the cyclic Jacobi method for parallel block orderings*, *SIAM J. Matrix Anal. Appl.*, 10 (1989), pp. 326–346.
- [33] I. SLAPNIČAR, *Highly accurate symmetric eigenvalue decomposition and hyperbolic SVD*, *Linear Algebra Appl.*, 358 (2003), pp. 387–424.
- [34] K. VESELIĆ AND V. HARI, *A note on a one-sided Jacobi algorithm*, *Numer. Math.*, 56 (1989), pp. 627–633.
- [35] J. H. WILKINSON, *Note on the quadratic convergence of the cyclic Jacobi process*, *Numer. Math.*, 4 (1962), pp. 296–300.