GLOBAL AND QUADRATIC CONVERGENCE OF THE BLOCK JACOBI METHOD FOR HERMITIAN MATRICES UNDER THE DE RIJK PIVOT STRATEGY[*]

VJERAN HARI[†]

Abstract. This paper provides a proof of global and quadratic convergence of the block Jacobi method for Hermitian matrices under the de Rijk pivot strategy. Also global and quadratic convergence of the element-wise Jacobi method under the same pivot strategy is proved. It is shown that sharp quadratic convergence bounds can be deduced from an estimate that is obtained in the global convergence proof. Numerical tests illustrate the behavior of the methods under the de Rijk pivot strategy.

Key words. eigenvalue problem, block Jacobi method, de Rijk pivot strategy, global convergence, quadratic convergence

AMS subject classifications. 65F15

1. Introduction. In this paper we consider the element-wise and block Jacobi method for solving the eigenvalue problem of Hermitian matrices under a special "row-cyclic-like" pivot strategy named after de Rijk. Although this pivot strategy was first mentioned by Hestenes [37], its first proper consideration and practical usage was made by P. P. M. de Rijk. In [7], de Rijk used it for solving the singular value decomposition (SVD) problem with the one-sided Jacobi method. He used the one-sided method for its "properties of parallelism and vectorizability".

In the meantime, the Jacobi method has evolved, and today various versions, upgrades, modifications, and generalizations are used to solve different eigenvalue problems. The standard Jacobi method for real symmetric or complex Hermitian matrices has become reliable [55], efficient [10, 11, 13, 14], highly accurate [6, 8, 44, 47, 49], and a commonly used method. It has an inherent parallelism [15], and it can easily be modified to work with blocks instead of elements. The modifications that work with blocks are referred to as block Jacobi methods. Their code makes use of the BLAS3 routines. In this way the methods better exploit the memory hierarchy system of contemporary computers. The one-sided block method is even more efficient and accurate. The standard and block Jacobi methods are best known for their high relative accuracy when applied to a positive definite symmetric matrix.

The standard (element-wise or simple) Jacobi method for Hermitian or real symmetric matrices has been studied theoretically for more than 65 years. In particular, its global convergence has been proved for many different pivot strategies [16, 17, 18, 23, 31, 36, 38, 41, 42, 43, 48, 54, 56, 58, 60], and what is missing today is a global convergence proof under a general cyclic strategy and under the de Rijk pivot strategy. Its quadratic convergence behavior is well understood and is solved for the cyclic strategies [19, 22, 57, 61], but a proof has been missing for the de Rijk strategy.

As for the block Jacobi method, global convergence has been proved for the classical [63], serial [4, 12], and generalized serial pivot strategy [3, 30]. Quadratic convergence has been proved for the classical pivot strategy [52]. Recently, a new pivot strategy has been introduced by the name *dynamic ordering*. This *dynamic pivot strategy* in a way generalizes and

[†]University of Zagreb, Faculty of Science, Department of Mathematics, Bijenička cesta 30, HR-10000 Zagreb, Croatia (hari@math.hr), ORCID: 0000-0001-7283-0458.

V. HARI

parallelizes the classical pivot strategy [1, 2]. Under this strategy, global [64] and quadratic convergence [51] of the Hermitian Jacobi method has been proved. Again, global and quadratic convergence of the block method has not been considered for the de Rijk strategy.

It is known that the simple serial Jacobi method needs less cycles if the diagonal elements are occasionally ordered into the non-increasing ordering. This reordering of the diagonal elements can be done, say, after each cycle or after each step. However, if the columns of the pivot submatrix of the transformation matrix are occasionally swapped, then we lose a condition that is needed in the global convergence proof. The de Rijk pivot strategy uses the row-cyclic strategy, but before it operates along the rth row, it finds the largest among the diagonal elements h_{rr}, $h_{r+1,r+1}$, …,h_{nn} and brings it to the position (r, r). In this way, during $n - r$ steps, when the annihilation of the off-diagonal elements in the rth row takes place, the diagonal element h_{rr} remains the largest diagonal element in the lower-right principal submatrix of order $n - r + 1$. This makes the angles in the rotation matrices smaller and makes the whole process more stable. This is done for each row, and after a few cycles, all diagonal elements become non-increasingly ordered and remain such until convergence.

What this pivot strategy does within one cycle of the process resembles the pattern of swapping the diagonal elements in the Cholesky factorization method with diagonal pivoting. If the one-sided Jacobi method is used, then one cycle under the de Rijk pivot strategy resembles the QR factorization with column pivoting and the Modified Gram-Schmidt procedure [9, Chapter 6]. Simply said, the de Rijk pivot strategy gradually orders the diagonal elements into the non-increasing ordering. Reducing the number of sweeps means reducing the total number of steps, and this contributes to the accuracy and efficiency of the method. This pivot strategy has proved to be efficient even for the complex HZ methods, which solves the generalized eigenvalue problem [26, Section 3.3]. The de Rijk pivot strategy can be efficiently employed in the core (or kernel) algorithm that is used in the block Jacobi method.

In this paper we first generalize the de Rijk pivot strategy to work with the block Jacobi method. Then we prove global and quadratic convergence of the complex block and simple Jacobi method for Hermitian matrices under that pivot strategy. The proofs also hold for the real (block and simple) Jacobi method. The estimates obtained in the global convergence proof are so sharp that we use them to prove asymptotic quadratic convergence with bounds that are equal to the sharpest bounds for the row-cyclic method.

The paper is organized as follows. In Section 2 we define the block Kogbetliantz method and the block Jacobi method for Hermitian matrices. We also describe the de Rijk pivot strategy for the simple methods and define it for the block methods. This can be done in essentially two ways: one that changes the block-matrix partition during the process and another one that leaves it intact. In Section 3 we gather some known facts that are later used in the global convergence proofs. They are here presented in condensed form, although they come from different research problems. In Section 4 we consider the block Jacobi method for Hermitian matrices under the de Rijk pivot strategy. In particular, we prove an estimate for the off-norm reduction during one cycle of the method. In Section 5 the obtained estimate is used to prove global and quadratic convergence of the block and simple Jacobi method under the de Rijk pivot strategy. In Section 6 we briefly discuss why the de Rijk pivot strategy is not the best option for the Kogbetliantz method. Finally, in Section 7 we describe how we have tested the simple and block Jacobi method under the de Rijk pivot strategy in MATLAB. The numerical tests confirm the presented theory. In Section 8 we give the conclusion of our theoretical and numerical investigation.

2. The block Jacobi methods and the de Rijk pivot strategy. Let $\mathbf{A}$ be a real or complex matrix of order n. Let

(2.1)
$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ A_{21} & A_{22} & \cdots & A_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mm} \end{bmatrix} \begin{matrix} n_1 \\ n_2 \\ \vdots \\ n_m \end{matrix}, \qquad A_{rt} \in \mathbf{C}^{n_r \times n_t}, \quad 1 \le r < t \le m,$$
$$\begin{matrix} n_1 & n_2 & \cdots & n_m \end{matrix}$$

be a *block-matrix partition* of $\mathbf{A}$, where the diagonal blocks A_{rr}, $1 \le r \le m$, are square. The block-matrix partition (2.1) of $\mathbf{A}$ is determined by the partition $\pi_{n,m} = (n_1, \dots, n_m)$ of n, where $n_r \ge 1$, for all $1 \le r \le m$, and $n_1 + \cdots + n_m = n$. We will consider n and m to be constant and mostly use the notation π instead of $\pi_{n,m}$.

Let us consider the block Jacobi process,

(2.2)
$$\mathbf{A}^{(k+1)} = [\mathbf{U}^{(k)}]^* \mathbf{A}^{(k)} \mathbf{V}^{(k)}, \quad k \ge 1, \qquad \mathbf{A}^{(1)} = \mathbf{A},$$

with unitary matrices $\mathbf{U}^{(k)}$, $\mathbf{V}^{(k)}$, $k \ge 1$. In the relation (2.2), X^* denotes the Hermitian transpose of X. The matrices $\mathbf{U}^{(k)} = (U_{rt}^{(k)})$ and $\mathbf{V}^{(k)} = (V_{rt}^{(k)})$ carry the same block-matrix partition as $\mathbf{A}$ and differ from the identity matrix $\mathbf{I}_n$ in one principal submatrix of order $n_{i(k)} + n_{j(k)}$. These principal submatrices are denoted by $\hat{\mathbf{U}}^{(k)}$ and $\hat{\mathbf{V}}^{(k)}$ and are called *pivot submatrices* of $\mathbf{U}^{(k)}$ and $\mathbf{V}^{(k)}$, respectively. They are located at the intersection of the block-rows and block-columns $i(k)$, $j(k)$ of $\mathbf{U}^{(k)}$ and $\mathbf{V}^{(k)}$,

(2.3)
$$\hat{\mathbf{U}}^{(k)} = \begin{bmatrix} U_{i(k)i(k)}^{(k)} & U_{i(k)j(k)}^{(k)} \\ U_{j(k)i(k)}^{(k)} & U_{j(k)j(k)}^{(k)} \end{bmatrix}, \qquad \hat{\mathbf{V}}^{(k)} = \begin{bmatrix} V_{i(k)i(k)}^{(k)} & V_{i(k)j(k)}^{(k)} \\ V_{j(k)i(k)}^{(k)} & V_{j(k)j(k)}^{(k)} \end{bmatrix}, \quad k \ge 1.$$

We will write $\mathbf{U}^{(k)} = \mathcal{E}(i(k), j(k); \hat{\mathbf{U}}^{(k)})$, where $\mathcal{E} = \mathcal{E}_\pi$ is the mapping that constructs the $n \times n$ matrix $\mathbf{U}^{(k)}$ from the input data $i(k)$, $j(k)$ and $\hat{\mathbf{U}}^{(k)}$. Using $\mathcal{E}$ we have that $\mathbf{V}^{(k)} = \mathcal{E}(i(k), j(k); \hat{\mathbf{V}}^{(k)})$. The *pivot indices* i and j are functions of k. When k is known, we omit it and write i, j instead of $i(k)$, $j(k)$, respectively. Say, when k is clear from the context or when we focus on position (i, j) in the matrix, we will write $\mathbf{U}_{ij}^{(k)}$ instead of $\mathbf{U}_{i(k)j(k)}^{(k)}$. We call the matrices $\mathbf{U}^{(k)}$, $\mathbf{V}^{(k)}$ *unitary elementary block matrices*.

The same notation is used for the pivot submatrix $\hat{\mathbf{A}}^{(k)}$ of $\mathbf{A}^{(k)}$. The matrices $A_{ij}^{(k)}$ and $A_{ji}^{(k)}$ will be called *pivot blocks* of $\mathbf{A}^{(k)}$. In step k we construct the pivot submatrices $\hat{\mathbf{U}}^{(k)}$ and $\hat{\mathbf{V}}^{(k)}$ from the pivot submatrix $\hat{\mathbf{A}}^{(k)}$.

In the special case when $\pi = (1, 1, \dots, 1)$, the prefix *block* is omitted, and we speak of the element-wise or simple Jacobi method.

2.1. The block and simple Kogbetliantz methods. A typical task of $\hat{\mathbf{U}}^{(k)}$ and $\hat{\mathbf{V}}^{(k)}$ is to diagonalize $\hat{\mathbf{A}}^{(k)}$:

(2.4)
$$[\hat{\mathbf{U}}^{(k)}]^* \begin{bmatrix} A_{i(k)i(k)}^{(k)} & A_{i(k)j(k)}^{(k)} \\ A_{j(k)i(k)}^{(k)} & A_{j(k)j(k)}^{(k)} \end{bmatrix} \hat{\mathbf{V}}^{(k)} = \begin{bmatrix} A_{i(k)i(k)}^{(k+1)} & 0 \\ 0 & A_{j(k)j(k)}^{(k+1)} \end{bmatrix},$$

where $A_{i(k)i(k)}^{(k+1)}$ and $A_{j(k)j(k)}^{(k+1)}$ are diagonal matrices. This choice of $\mathbf{U}^{(k)}$ and $\mathbf{V}^{(k)}$ is leading us to the block Kogbetliantz method [4]. It is used to compute the singular value decomposition

of the initial matrix $\mathbf{A}$. Hence, it is natural to require that (2.4) is the SVD of $\hat{\mathbf{A}}^{(k)}$. In that case the diagonal matrix on the right-hand side of (2.4) has non-negative diagonal elements ordered non-increasingly.

For the block Kogbetliantz method, it is a good idea to diagonalize the diagonal blocks A_{rr}, $1 \le r \le m$, from (2.1) before the iteration process begins (see [24]). This procedure can be described by the relation

$$(2.5) \qquad \mathbf{A}^{(1)} = [\mathbf{U}^{(0)}]^* \mathbf{A} \mathbf{V}^{(0)},$$

where $\mathbf{U}^{(0)}$ and $\mathbf{V}^{(0)}$ are appropriate block-diagonal unitary matrices. More precisely, the relation (2.5) describes m consecutive SVDs of the diagonal blocks A_{rr}, $1 \le r \le m$. The relations (2.5) and (2.4) imply that all diagonal blocks $A_{rr}^{(k)}$, $1 \le r \le m$, $k \ge 1$, remain to be diagonal and that their non-negative diagonal elements are ordered non-increasingly. In particular, it means that in each pivot submatrix, almost half of the off-diagonal elements are zero. This can be used to simplify and accelerate the diagonalization of the pivot submatrices $\hat{\mathbf{A}}^{(k)}$, $k \ge 1$.

As mentioned above, if $n_1 = \cdots = n_m = 1$, then we speak of the *original, element-wise, or simple Kogbetliantz method*. For the simple Kogbetliantz method, the starting matrix $\mathbf{A}^{(1)}$ is typically obtained by the transformation

$$(2.6) \qquad \mathbf{A}^{(1)} = \mathbf{D}^{(0)} \mathbf{A} \quad \text{or} \quad \mathbf{A}^{(1)} = \mathbf{A} \mathbf{D}^{(0)},$$

where $\mathbf{D}^{(0)}$ is the diagonal matrix of unit complex numbers that makes the diagonal of $\mathbf{A}^{(1)}$ nonnegative. If $\mathbf{A}$ is real, then $\mathbf{D}^{(0)}$ is a diagonal orthogonal matrix. The block and simple Kogbetliantz methods can be combined in such a way that the simple method is used to diagonalize the pivot submatrix of the block method in each step.

As it is well-known, one can apply the QR factorization to $\mathbf{A}$ and start the iterative process with a triangular $\mathbf{A}$. If the original matrix $\mathbf{A}$ was not square, then this would be a standard procedure. A triangular $\mathbf{A}$ is well suited for the simple and block Kogbetliantz method if the pivot strategy is row- or column-cyclic. Then the triangular form is in a way preserved, and it significantly reduces the computational cost. Furthermore, the quadratic asymptotic convergence of the simple Kogbetliantz method can be proved even in the presence of multiple singular values [4, 22]. In addition, the simple Kogbetliantz method computes the singular values to high relative accuracy provided that $\mathbf{A}$ can be symmetrically scaled by a diagonal matrix in such a way that the spectral condition number of the scaled matrix is small [45]. Gobal and quadratic convergence of the block Kogbetliantz method has been proven in [4] for the case of simple singular values.

If $\mathbf{A}$ is real, then $\mathbf{U}^{(k)}$, $\mathbf{V}^{(k)}$, $\mathbf{A}^{(k)}$, $k \ge 0$, are real, and we speak of the *real* block and simple Kogbetliantz method.

2.2. A special block iterative process. Here we introduce a block iterative process that will be used in the convergence analysis.

Consider the linear operator $\mathcal{N}_{\ell r}$, whose domain is the vector space of complex (or real) square matrices of order n. Let $\mathbf{X}$ be a square matrix of order n, and let $\mathbf{X} = (X_{st})$ denote its block-matrix partition defined by π. Then $\mathcal{N}_{\ell r}(\mathbf{X})$ sets the blocks $X_{\ell r}, X_{r\ell}, X_{\ell\ell}, X_{rr}$ to zero and leaves other blocks intact. We see that $\mathcal{N}_{\ell r}$ depends on the partition π of n.

Let $\mathbf{A}^{(1)} = (A_{st}^{(1)})$ be a complex square matrix of order n, partitioned in accordance to π. Consider the sequence $(\mathbf{A}^{(k)}, \ k \ge 1)$ obtained by the rule

$$\mathbf{A}^{(k+1)} = [\mathbf{U}_{i(k)j(k)}^{(k)}]^* \mathcal{N}_{i(k)j(k)}(\mathbf{A}^{(k)}) \mathbf{V}_{i(k)j(k)}^{(k)}$$
$$(2.7) \qquad = \mathcal{N}_{i(k)j(k)} \left([\mathbf{U}_{i(k)j(k)}^{(k)}]^* \mathbf{A}^{(k)} \mathbf{V}_{i(k)j(k)}^{(k)} \right), \qquad k \ge 1,$$

where $\mathbf{U}^{(k)}_{i(k)j(k)} = \mathcal{E}(i(k), j(k); \hat{\mathbf{U}}^{(k)})$, $\mathbf{V}^{(k)}_{i(k)j(k)} = \mathcal{E}(i(k), j(k); \hat{\mathbf{V}}^{(k)})$, and $\hat{\mathbf{U}}^{(k)}$, $\hat{\mathbf{V}}^{(k)}$ are arbitrary unitary matrices of order $n_{i(k)} + n_{j(k)}$.

For this process we have

$$
\tag{2.8} \mathbf{off}^2(\mathbf{A}^{(k+1)}) = \mathbf{off}^2(\mathbf{A}^{(k)}) - \mathbf{off}^2(\hat{\mathbf{A}}^{(k)}), \qquad k \geq 1,
$$

where

$$
\tag{2.9} \mathbf{off}(\mathbf{X}) = \|\Omega(\mathbf{X})\|_F, \qquad \Omega(\mathbf{X}) = \mathbf{X} - \mathrm{diag}(\mathbf{X}).
$$

Here $\mathbf{X}$ is any square matrix of order n, $\mathrm{diag}(\mathbf{X})$ is the *diagonal part* of $\mathbf{X}$, and $\| \cdot \|_F$ is the Frobenius norm. The quantity $\mathbf{off}(\mathbf{X})$ is called the *off-norm* or the *departure from the diagonal form* of $\mathbf{X}$. We use notation $\mathbf{off}^2(\mathbf{X})$ instead of $[\mathbf{off}(\mathbf{X})]^2$.

Another useful measure is $\mathbf{Off}(\mathbf{X})$. It is defined by

$$
\tag{2.10} \mathbf{Off}(\mathbf{X}) = \|\mathbf{\Omega}(\mathbf{X})\|_F, \qquad \mathbf{\Omega}(\mathbf{X}) = \mathbf{X} - \mathrm{diag}(X_{11}, \ldots, X_{mm}),
$$

and it can be called the *block off-norm* or the *departure from the block-diagonal form* of $\mathbf{X}$. For the process (2.7) we have

$$
\tag{2.11}
\begin{aligned}
\mathbf{Off}^2(\mathbf{A}^{(k+1)}) &= \mathbf{Off}^2(\mathbf{A}^{(k)}) - \mathbf{Off}^2(\hat{\mathbf{A}}^{(k)}) \\
&= \mathbf{Off}^2(\mathbf{A}^{(k)}) - \left(\|A^{(k)}_{ij}\|^2_F + \|A^{(k)}_{ji}\|^2_F \right), \qquad k \geq 1.
\end{aligned}
$$

Note that the block matrix $\hat{\mathbf{A}}$ is defined by the partition $(n_{i(k)}, n_{j(k)})$ of $n_{i(k)} + n_{j(k)}$. The iterative process (2.7) is more general then the block Kogbetliantz method because $\hat{\mathbf{U}}^{(k)}$, $\hat{\mathbf{V}}^{(k)}$ are arbitrary unitary matrices for every $k \geq 1$. The relation (2.8) holds also for the simple and block Kogbetliantz method, and (2.11) holds for the block Kogbetliantz method.

2.3. The associated sequences $(\dot{\mathbf{A}}^{(k)},\ k \geq 1)$ and $(\overset{\bullet}{\mathbf{A}}^{(k)},\ k \geq 1)$. Let us return to the block Kogbetliantz method. It generates the sequence of matrices $(\mathbf{A}^{(k)},\ k \geq 1)$ by the rule (2.2), with (2.5) as the initial transformation. Let

$$
\tag{2.12} \dot{\mathbf{A}}^{(k)} = \Omega(\mathbf{A}^{(k)}), \quad \overset{\bullet}{\mathbf{A}}^{(k)} = \mathbf{\Omega}(\mathbf{A}^{(k)}), \qquad k \geq 1.
$$

Since the diagonal blocks of $\mathbf{A}^{(k)}$ are diagonal, we have $\dot{\mathbf{A}}^{(k)} = \overset{\bullet}{\mathbf{A}}^{(k)}$, i.e., $\dot{\mathbf{A}}^{(k)}$ is both the *off-diagonal* and the *block-off-diagonal part* of $\mathbf{A}^{(k)}$.

With the sequence of matrices $(\mathbf{A}^{(k)},\ k \geq 1)$ generated by the block Kogbetliantz method, we associate the sequences $(\dot{\mathbf{A}}^{(k)},\ k \geq 1)$ and $(\overset{\bullet}{\mathbf{A}}^{(k)},\ k \geq 1)$, and we briefly call them the *associated sequences*.

The associated sequences can be obtained from $\dot{\mathbf{A}}^{(1)} = \Omega(\mathbf{A}^{(1)}) = \overset{\bullet}{\mathbf{A}}^{(1)}$ by the rule (2.7) provided that we have at disposal the transformation matrices $\mathbf{U}^{(k)}$, $\mathbf{V}^{(k)}$, $k \geq 1$, which are obtained by applying the block Kogbetliantz method to $\mathbf{A}^{(1)}$. This procedure can be coded efficiently if we keep a vector of length n that holds the diagonal of the matrices $\mathbf{A}^{(k)}$, $k \geq 1$, generated by the block Kogbetliantz method.

In the case of the simple Kogbetliantz method, we have only one associated sequence, that is, $(\dot{\mathbf{A}}^{(k)},\ k \geq 1)$, defined by the relation (2.12).

2.4. Block and simple Jacobi methods for Hermitian matrices. If $\mathbf{H}$ is Hermitian and we are solving the eigenvalue problem for $\mathbf{H}$, then the Hermitian property of the initial matrix has to be preserved during the process. This leads us to the block and simple Jacobi method for Hermitian matrices. Now, we have $\mathbf{V}_{i(k)j(k)}^{(k)} = \mathbf{U}_{i(k)j(k)}^{(k)}$, $k \geq 0$, and the relation (2.2) can be written as

$$\mathbf{H}^{(k+1)} = [\mathbf{U}^{(k)}]^* \mathbf{H}^{(k)} \mathbf{U}^{(k)}, \qquad k \geq 0,$$

where $\mathbf{H}^{(0)} = \mathbf{H}$, $\mathbf{H}^{(1)}$ is the Hermitian matrix whose diagonal blocks are diagonal and $\mathbf{U}^{(0)}$ is a block-diagonal unitary matrix. The relation (2.4) is the eigenvalue decomposition of the pivot submatrix $\hat{\mathbf{H}}^{(k)}$.

The relations (2.7) and (2.8) are both valid, and the matrices $\dot{\mathbf{H}}^{(k)}$, $k \geq 1$, are Hermitian with zero diagonal blocks. Therefore, we have $\dot{\mathbf{H}}^{(k)} = \overset{\bullet}{\mathbf{H}}^{(k)}$, $k \geq 1$.

If $\mathbf{H}$ is real and symmetric, then the generated matrices $\mathbf{H}^{(k)}$, $k \geq 0$, are real and symmetric, while $\mathbf{U}^{(k)}$, $k \geq 0$, are orthogonal.

2.5. Pivot strategies. Here we define what a pivot strategy for the block methods is, and then we consider a special one, the *de Rijk pivot strategy*.

The selection of pivot pairs is defined by a *pivot strategy*. A pivot strategy can be identified with a function $\mathsf{I} : \mathcal{N} \to \mathcal{P}_m$, where

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

Here, $\mathcal{P}_m$ contains pairs of indices which address the blocks in the upper block-triangle of the matrix $\mathbf{A}$. It contains $M = m(m-1)/2$ pairs. For a given m with $3 \leq m \leq n$, the function I can be defined in many ways. If I is periodic, then I is called *periodic pivot strategy*. Let I be a periodic pivot strategy with the basic period P. If $P = M$ and $\{\mathsf{I}(k) : k = 1, \ldots, P\} = \mathcal{P}_m$, then I is called a *cyclic* strategy. Then the first cycle consists of the first M steps of the method. More about pivot strategies can be found in [24, 41, 58] and especially in [30, Section 3].

The most commonly used cyclic strategy is the well-known *row-cyclic* strategy. It is defined by the *row-wise ordering* of the set $\mathcal{P}_m$, i.e., by the sequence of pairs

$$(1,2), \ (1,3), \ \ldots, (1,m), \ (2,3), \ (2,4), \ \ldots, (2,m), \ (3,4), \ \ldots, (m-1,m).$$

Using the row-cyclic process (2.2), for the matrix $\mathbf{A}^{(M+1)}$ obtained at the end of the first cycle, we have

$$\begin{aligned}
\mathbf{A}^{(M+1)} &= [\mathbf{U}^{(1)}\mathbf{U}^{(2)} \cdots \mathbf{U}^{(M)}]^* \mathbf{A}^{(1)} [\mathbf{V}^{(1)}\mathbf{V}^{(2)} \cdots \mathbf{V}^{(M)}] \\
&= [\mathbf{U}^{(1:m-1)}\mathbf{U}^{(m:2m-3)} \cdots \mathbf{U}^{(M)}]^* \mathbf{A}^{(1)} [\mathbf{V}^{(1:m-1)}\mathbf{V}^{(m:2m-3)} \cdots \mathbf{V}^{(M)}] \\
&= [\mathbf{U}_{1,2:m}\mathbf{U}_{2,3:m} \cdots \mathbf{U}_{m-1,m}]^* \mathbf{A}^{(1)} [\mathbf{V}_{1,2:m}\mathbf{V}_{2,3:m} \cdots \mathbf{V}_{m-1,m}],
\end{aligned}$$

where

$$(2.13) \qquad \mathbf{U}^{(r_1:r_2)} = \mathbf{U}^{(r_1)} \cdots \mathbf{U}^{(r_2)}, \qquad \mathbf{V}^{(r_1:r_2)} = \mathbf{V}^{(r_1)} \cdots \mathbf{V}^{(r_2)}, \ \ 1 \leq r_1 \leq r_2 \leq m,$$

$$(2.14) \qquad \mathbf{U}_{r,r+1:m} = \mathbf{U}_{r,r+1} \cdots \mathbf{U}_{rm}, \ \ \mathbf{V}_{r,r+1:m} = \mathbf{V}_{r,r+1} \cdots \mathbf{V}_{rm}, \ 1 \leq r \leq m-1.$$

In the special case when $r_2 = r_1$, we have $\mathbf{U}^{(r_1:r_1)} = \mathbf{U}^{(r_1)}$, $\mathbf{V}^{(r_1:r_1)} = \mathbf{V}^{(r_1)}$. Continuing the process, in the cycle t one simply adds $(t-1)M$ to the superscripts of the terms in the above relations. We will also use the notation $\mathbf{V}_{r,r+1:m}^{[t]}$, $1 \leq r \leq m-1$, where $\mathbf{V}_{r,r+1:m}^{[t]} = \mathbf{V}_{r,r+1}^{[t]} \cdots \mathbf{V}_{rm}^{[t]}$. Here, the matrices $\mathbf{V}_{rs}^{[t]}$ belong to the cycle t of the method. A similar notation is used for the left-hand transformations $\mathbf{U}_{rs}^{[t]}$.

2.5.1. The de Rijk pivot strategy. The de Rijk pivot strategy [7] is the row-cyclic strategy which is "interrupted" at most $m - 1$ times within each cycle. It can be described by considering one full cycle of the process (2.2), say, the first cycle. For the de Rijk (pivot) strategy we use the notation that was introduced for the row-cyclic strategy. If we consider just one cycle of the method, then we can simplify the notation by using just the subscripts or superscripts.

We describe this strategy for the simple and block Kogbetliantz method. The definition automatically extends to the simple and block Jacobi method for Hermitian matrices.

2.5.2. The simple and block Kogbetliantz methods under the de Rijk strategy. In [7] this pivot strategy was used in the context of computing the SVD by the one-sided Jacobi method. Here it is defined in the context of the Kogbetliantz method. If $\pi = (1, \ldots, 1)$, then we have $m = n$ and $M = N$, where

$$N = \frac{n(n-1)}{2}.$$

We use the variables m and M in those relations that are valid for both the block and simple Kogbetliantz method. In the relations that are specific for the simple method, we use n and N instead of m and M, respectively.

The first cycle is described as follows:

$$(2.15) \qquad \mathbf{A}^{(M+1)} = \mathbf{A}^{[1]} = [\mathbf{U}^{[1]}]^* \mathbf{A}^{(1)} \mathbf{V}^{[1]},$$

where

$$\mathbf{U}^{[1]} = \mathbf{P}_1 \mathbf{U}^{(1:m-1)} \mathbf{P}_2 \mathbf{U}^{(m:2m-3)} \cdots \mathbf{P}_{m-2} \mathbf{U}^{(m-2:m-1)} \mathbf{P}_{m-1} \mathbf{U}^{(M)},$$
$$\mathbf{V}^{[1]} = \mathbf{P}_1 \mathbf{V}^{(1:m-1)} \mathbf{P}_2 \mathbf{V}^{(m:2m-3)} \cdots \mathbf{P}_{m-2} \mathbf{V}^{(m-2:m-1)} \mathbf{P}_{m-1} \mathbf{V}^{(M)},$$

or, using the notation (2.14),

$$(2.16) \qquad \mathbf{U}^{[1]} = \mathbf{P}_1 \mathbf{U}_{1,2:m} \mathbf{P}_2 \mathbf{U}_{2,3:m} \cdots \mathbf{P}_{m-2} \mathbf{U}_{m-2,m-1:m} \mathbf{P}_{m-1} \mathbf{U}_{m-1,m},$$
$$(2.17) \qquad \mathbf{V}^{[1]} = \mathbf{P}_1 \mathbf{V}_{1,2:m} \mathbf{P}_2 \mathbf{V}_{2,3:m} \cdots \mathbf{P}_{m-2} \mathbf{V}_{m-2,m-1:m} \mathbf{P}_{m-1} \mathbf{V}_{m-1,m}.$$

Note that in each subsequent cycle, the permutation matrices can differ from those in the previous cycle. Next, we consider separately the simple and the block method under that pivot strategy.

2.5.3. The simple Kogbetliantz method under the de Rijk strategy. We consider the case $m = n$ and $M = N$. Let

$$(2.18) \qquad s_\ell = 1 + 2 + \cdots + \ell = \frac{\ell(\ell+1)}{2}, \quad 1 \le \ell \le n-1; \qquad s_0 = 0.$$

Since $s_{n-1} = N = 1 + 2 + \cdots + n - 1$, from (2.18) we obtain $N - s_{n-1} = 0$ and

$$N - s_\ell = N - (1 + 2 + \cdots + \ell) = n - 1 + (n-2) + \cdots + (\ell+1), \quad 1 \le \ell < n-1,$$

and consequently we have

$$N - s_{n-r} + 1 = \begin{cases} 1, & r = 1, \\ 1 + (n-1) + (n-2) + \cdots + (n-(r-1)), & 2 \le r \le n-1. \end{cases}$$

Note that in step $k = N - s_{n-r} + 1$, the transformation $\mathbf{A}^{(k+1)} = [\mathbf{P}_r \mathbf{U}^{(k)}]^* \mathbf{A}^{(k)} [\mathbf{P}_r \mathbf{V}^{(k)}]$ takes place. Let us now define the permutation matrices. We have

$$(2.19) \qquad \mathbf{P}_r = \mathbf{I}_{rr'}, \quad r' \geq r, \quad 1 \leq r \leq n - 1,$$

where $\mathbf{I}_{rr'}$ is the transposition matrix and r' is the smallest index such that

$$(2.20) \qquad a_{r'r'}^{(N-s_{n-r}+1)} = \max_{r \leq \ell \leq n} a_{\ell\ell}^{(N-s_{n-r}+1)}, \quad 1 \leq r \leq n - 1.$$

Thus, before the annihilation of the elements in the rth row and column begins, the rows and columns r and r' are swapped. The relations (2.20) and (2.19) imply that the permutational transformation makes the (r, r)-element larger than or equal to the elements at positions (ℓ, ℓ), $\ell > r$. If $r' = r$, then we have $\mathbf{P}_r = \mathbf{I}_n$, and no swap occurs.

We see that the *de Rijk strategy tries to order the diagonal elements into the non-increasing ordering during the process*.

2.5.4. The block Kogbetliantz method under the de Rijk strategy. In this case we have $m < n$ and $M < N$. We have to extend the de Rijk pivot strategy to be suitable for the block method. We propose two solutions which we call the (pivot) strategies *bdR1* and *bdR2*.

2.5.5. The *bdR1* strategy. In the above relations $\mathbf{P}_1,\dots,\mathbf{P}_{m-1}$ are *block transposition matrices* of order n which are used to swap the block-rows and block-columns r and r'. Let

$$(2.21) \qquad \mathbf{s}_r = n_1 + \cdots + n_r, \quad 1 \leq r \leq m; \qquad \mathbf{s}_0 = 0.$$

If $r' = r$, then $\mathbf{P}_r = \mathbf{I}_n$. Otherwise, we have

$$(2.22) \quad \mathbf{P}_r = \begin{bmatrix} I_{\mathbf{s}_{r-1}} & & & & & \\ & O_{n_r n_{r'}} & & & I_{n_r} & \\ & & I_{\mathbf{s}_{r'-1}-\mathbf{s}_r} & & & \\ & I_{n_{r'}} & & O_{n_{r'} n_r} & & \\ & & & & & I_{n-\mathbf{s}_{r'}} \end{bmatrix}, \qquad 1 \leq r < r' \leq m,$$

where the zero diagonal blocks $O_{n_r n_{r'}}$ and $O_{n_{r'} n_r}$ have the dimensions $n_r \times n_{r'}$ and $n_{r'} \times n_r$, respectively. This means that $\mathbf{P}_r$ does not carry the same block-matrix partition as $\mathbf{A}$. If in the relation (2.22) $r' = r + 1$, then the diagonal block $I_{\mathbf{s}_{r'-1}-\mathbf{s}_r}$ is omitted.

The subscript r' is chosen to be the smallest one that satisfies

$$(2.23) \qquad a_{\mathbf{s}_{r'-1}+1,\mathbf{s}_{r'-1}+1}^{(M-s_{m-r}+1)} = \max_{r \leq \ell \leq m} a_{\mathbf{s}_{\ell-1}+1,\mathbf{s}_{\ell-1}+1}^{(M-s_{m-r}+1)}, \quad 1 \leq r \leq m - 1.$$

Note that for each r, $1 \leq r \leq m - 1$, the position $(\mathbf{s}_{r-1} + 1, \mathbf{s}_{r-1} + 1)$ is exactly where the first diagonal element of the rth diagonal block lies. This first diagonal element is also the largest diagonal element of the rth diagonal block. The superscript $M - s_{m-r} + 1$ indicates the stage of the block process just after the block at position $(r - 1, m)$ has been zeroed. Note that

$$M - s_{m-r} + 1 = \begin{cases} 1, & r = 1, \\ 1 + (m - 1) + (m - 2) + \cdots + (m - (r - 1)), & 2 \leq r \leq m - 1. \end{cases}$$

The matrices $\mathbf{P}_k$, $1 \leq k \leq m - 1$, are such that the whole diagonal blocks of the current matrix are swapped. Note that the similarity transformation with $\mathbf{P}_k$ also changes (permutes) the current partition π that defines the block-matrix partition. Therefore, for computational

reasons it is advantageous when all entries of π (except maybe the last one) are equal. We call this pivot strategy *block de Rijk strategy of the first kind* or shorter the *bdR1* strategy.

REMARK 2.1. Note that $a_{\mathbf{s}_{r'-1}+1,\mathbf{s}_{r'-1}+1}^{(M-s_{m-r}+1)}$ from the relation (2.23) equals the spectral norm of $A_{r'r'}^{(M-s_{m-r}+1)}$. Instead of using the spectral norm, one can use some other matrix norm, say, the Frobenius norm of $A_{r'r'}^{(M-s_{m-r}+1)}$, to select r'. All later convergence results will hold equally well for both ways of selecting r'.

2.5.6. The *bdR2* strategy. Alternatively, if we want to maintain the same block-matrix partition in all $\mathbf{A}^{(k)}$, then we can use the following definition of $\mathbf{P}_r$,

$$(2.24) \quad \mathbf{P}_r = \begin{bmatrix} I_q & \\ & I_{q+1,(q+1)'}I_{q+2,(q+2)'}\cdots I_{q+n_r,(q+n_r)'} \end{bmatrix}, \quad q = \mathbf{s}_{r-1}, \ 1 \le r \le m.$$

Here, $I_{q+k,(q+k)'}$, with $(q+k)' \ge q+k$, $1 \le k \le n_r$, is a transposition matrix which brings the largest diagonal element in the lower-right part of the matrix (of order $n - q - k$) to the $(q+k, q+k)$-position via a similarity transformation. First, $(q+1)'$ is found, then $(q+2)'$, and finally $(q+n_r)'$. Then the rows and columns $q+1, q+2, \ldots, q+n_r$ are interchanged with those of $(q+1)', (q+2)', \ldots, (q+n_r)'$, respectively.

There are several drawbacks of this pivot strategy. After the swapping of the rows and columns, we have to reorder the diagonal elements to ensure that those in each diagonal block are ordered non-increasingly. Furthermore, the diagonal blocks may fail to be diagonal.

However, for this choice of $\mathbf{P}_r$ it seems natural to discard the requirement that the diagonal blocks are diagonal and that the diagonal elements are non-increasingly ordered. This implies that the transformation (2.5) is not needed. Instead, we can assume that the relation (2.6) holds. The condition (2.6) makes the diagonal elements real and nonnegative, and this can be used to simplify the core algorithm which diagonalizes the pivot submatrix.

If we assume that the simple Kogbetliantz method under the de Rijk strategy is used to diagonalize the pivot submatrix $\hat{\mathbf{A}}^{(k)}$, then the algorithm will transform $\hat{\mathbf{A}}^{(k)}$ into a diagonal matrix with nonnegative diagonal elements ordered non-increasingly (this is proved later in the paper). In particular, when annihilations in the block row r are completed (for any $1 \le r \le m-1$), then all diagonal blocks of the current matrix will be diagonal with nonnegative diagonal elements ordered non-increasingly. We will call this pivot strategy *block de Rijk strategy of the second kind* or shorter the *bdR2 strategy*.

In the case $m = n$, both definitions of the de Rijk strategy for the block Kogbetliantz method reduce to the definition of the de Rijk strategy for the simple Kogbetliantz method. The same will hold in the case of the block and simple Jacobi method for Hermitian matrices.

Using the de Rijk strategy, the total number of cycles needed to diagonalize the initial matrix is generally reduced. This is accomplished by a gradual reordering of the diagonal elements. Ultimately, all diagonal elements are ordered non-increasingly.

2.5.7. The case of Hermitian matrices. As we will see, our convergence analysis requires a certain symmetry property of the iterated matrices. This is automatically achieved if the initial matrix is real symmetric or complex Hermitian, and the method is the real or the complex (simple or block) Jacobi method. These Jacobi methods solve the eigenvalue problem for $\mathbf{H}$.

To shorten the presentation, we will consider only the complex block Jacobi method for the Hermitian matrix $\mathbf{H}$. In the case of the block method we will specially consider the case when $\hat{\mathbf{U}}^{(k)}$ is computed by the simple Jacobi method under the de Rijk pivot strategy. This choice ensures that the diagonal elements of each diagonal block are non-increasingly ordered. Alternatively, we can use, say, the row-cyclic strategy, and after computing $\hat{\mathbf{U}}^{(k)}$,

V. HARI

the columns of $\hat{\mathbf{U}}^{(k)}$ can be permuted so that the diagonal elements of $[\hat{\mathbf{U}}^{(k)}]^*\hat{\mathbf{H}}^{(k)}\hat{\mathbf{U}}^{(k)}$ are ordered non-increasingly.

In the case of the simple (i.e., element-wise) Jacobi method under the de Rijk strategy, the initial update, like in the relation (2.5), is omitted. In the case of the block Jacobi method under the *bdR1* strategy, it is replaced by $\mathbf{H}^{(1)} = [\mathbf{U}^{(0)}]^*\mathbf{H}\,\mathbf{U}^{(0)}$, where $\mathbf{U}^{(0)}$ is the block-diagonal unitary matrix which makes the diagonal blocks of $\mathbf{H}^{(1)}$ diagonal and orders the diagonal elements in each diagonal block non-increasingly. In the case of the *bdR2* strategy, the relation (2.5) is omitted.

3. Some auxiliary results. Here we present some known results which we will use as tools in our convergence analysis.

3.1. A minor modification of the results of Bujanović and Drmač. First, we slightly modify [4, Lemma 4.1] such that the results hold for the associated process of the block Kogbetliantz method for triangular matrices. We use the following notation: The tth largest singular value of a matrix $\mathbf{X}$ is denoted by $\sigma_t(\mathbf{X})$. The largest and smallest singular values of $\mathbf{X}$ are $\sigma_{\max}(\mathbf{X})$ and $\sigma_{\min}(\mathbf{X})$, respectively. Recall that $\sigma_{\max}(\mathbf{X}) = \|\mathbf{X}\|_2$, where $\|\mathbf{X}\|_2$ is the spectral norm of $\mathbf{X}$. As mentioned earlier, the Frobenius norm is denoted by $\|\mathbf{X}\|_F$.

The next lemma can be applied to both products $\mathbf{U}^{(1)}\mathbf{U}^{(2)}\cdots\mathbf{U}^{(m-1)}$ and $\mathbf{V}^{(1)}\mathbf{V}^{(2)}\cdots\mathbf{V}^{(m-1)}$. Let

$$(3.1)\quad \mathbf{R}^{(k)} = \mathbf{R}^{(k)}_{1,k+1} = \mathcal{E}(1,k+1;\hat{\mathbf{R}}^{(k)}), \quad \hat{\mathbf{R}}^{(k)} = \begin{bmatrix} C_1^{(k)} & S_1^{(k)} \\ S_2^{(k)} & C_2^{(k)} \end{bmatrix}, \quad 1 \le k \le m-1,$$

where $C_1^{(k)}$, $C_2^{(k)}$, $\hat{\mathbf{R}}^{(k)}$, $\mathbf{R}^{(k)}$ are square matrices of order n_1, n_{k+1}, $n_1 + n_{k+1}$, n, respectively.

LEMMA 3.1. ([4, Lemma 3.1]) *The product*

$$\mathbf{R}^{(1:m-1)} = (R^{(1:m-1)}_{st}) = \mathbf{R}^{(1)}\mathbf{R}^{(2)}\cdots\mathbf{R}^{(m-1)}$$

is an upper block triangular matrix, except for the first block-column. In particular, we have

$$R^{(1:m-1)}_{11} = C_1^{(1)}C_1^{(2)}\cdots C_1^{(m-1)}, \quad R^{(1:m-1)}_{ss} = C_2^{(s-1)}, \quad s = 2,\ldots,m$$
$$R^{(1:m-1)}_{s1} = S_2^{(s-1)}\left(C_1^{(s)}\cdots C_1^{(m-1)}\right), \qquad s = 2,\ldots,m$$
$$R^{(1:m-1)}_{1t} = \left(C_1^{(1)}\cdots C_1^{(t-2)}\right)S_1^{(t-1)}, \qquad t = 2,\ldots,m$$
$$R^{(1:m-1)}_{st} = S_2^{(s-1)}\left(C_1^{(s)}\cdots C_1^{(t-2)}\right)S_1^{(t-1)}, \qquad 2 \le s < t \le m$$
$$R^{(1:m-1)}_{st} = 0, \qquad m \ge s > t \ge 2.$$

If a left circular shift of the block columns is made, then an upper block Hessenberg matrix is obtained. To illustrate the form of $\mathbf{R}^{(1:m-1)}$ we have used MATLAB with symbolic arithmetic.

EXAMPLE 3.2. Consider the following MATLAB script:

```
syms  c1 c2 c3 c4 c5 c6 s1 s2 s3 s4 s5 s6 ...
      C1 C2 C3 C4 C5 C6 S1 S2 S3 S4 S5 S6 R U P
cii = sym([c1 c2 c3 c4 c5 c6]); sij = sym([s1 s2 s3 s4 s5 s6]);
cjj = sym([C1 C2 C3 C4 C5 C6]); sji = sym([S1 S2 S3 S4 S5 S6]);
n=7;%                          The row-cyclic pivot strategy
for i = 1:n-1
 U=sym(eye(n)); k=0;
 for j = i+1:n
  k = k+1; R = elem_plane(n,i,j,cii(k),sij(k),sji(k),cjj(k));
  U = U*R;
 end
 disp(U); pause
end
function R = elem_plane(n,i,j,rii,rij,rji,rjj)
  R = sym(eye(n)); R(i,i)=rii; R(i,j)=rij; R(j,i)=rji; R(j,j)=rjj;
end
```

The script assumes $m = n = 7$, computes the matrix $\mathbf{R}^{(1:m-1)}$, and displays it just before reaching the pause statement for the first time. The script uses notation which is different but still linked to the notation in Lemma 3.1:

$$\texttt{ck} \leftrightarrow C_1^{(k)}, \quad \texttt{sk} \leftrightarrow S_1^{(k)}, \quad \texttt{Ck} \leftrightarrow C_2^{(k)}, \quad \texttt{Sk} \leftrightarrow S_2^{(k)}, \quad k = 1, 2, \ldots, m-1.$$

We copied and edited the matrix obtained after the first pause statement. In particular, we deleted all occurrences of the multiplication operator $\star$ and of the commas that separate matrix elements. To save some horizontal space, we replaced all occurrences of ck, sk, Ck, Sk with $C_1^k, S_1^k, C_2^k, S_2^k$ instead of $C_1^{(k)}, S_1^{(k)}, C_2^{(k)}, S_2^{(k)}$, respectively. Here is the result $\mathbf{R}^{(1:6)}$:

$$
\begin{bmatrix}
C_1^1 C_1^2 C_1^3 C_1^4 C_1^5 C_1^6 & S_1^1 & C_1^1 S_1^2 & C_1^1 C_1^2 S_1^3 & C_1^1 C_1^2 C_1^3 S_1^4 & C_1^1 C_1^2 C_1^3 C_1^4 S_1^5 & C_1^1 C_1^2 C_1^3 C_1^4 C_1^5 S_1^6 \\
S_2^1 C_1^2 C_1^3 C_1^4 C_1^5 C_1^6 & C_2^1 & S_2^1 S_1^2 & S_2^1 C_1^2 S_1^3 & S_2^1 C_1^2 C_1^3 S_1^4 & S_2^1 C_1^2 C_1^3 C_1^4 S_1^5 & S_2^1 C_1^2 C_1^3 C_1^4 C_1^5 S_1^6 \\
S_2^2 C_1^3 C_1^4 C_1^5 C_1^6 & 0 & C_2^2 & S_2^2 S_1^3 & S_2^2 C_1^3 S_1^4 & S_2^2 C_1^3 C_1^4 S_1^5 & S_2^2 C_1^3 C_1^4 C_1^5 S_1^6 \\
S_2^3 C_1^4 C_1^5 C_1^6 & 0 & 0 & C_2^3 & S_2^3 S_1^4 & S_2^3 C_1^4 S_1^5 & S_2^3 C_1^4 C_1^5 S_1^6 \\
S_2^4 C_1^5 C_1^6 & 0 & 0 & 0 & C_2^4 & S_2^4 S_1^5 & S_2^4 C_1^5 S_1^6 \\
S_2^5 C_1^6 & 0 & 0 & 0 & 0 & C_2^5 & S_2^5 S_1^6 \\
S_2^6 & 0 & 0 & 0 & 0 & 0 & C_2^6
\end{bmatrix}.
$$

Since the symbolic arithmetic maintains the order in which operations are applied, we can safely replace scalar matrix elements obtained by the MATLAB code with the appropriate blocks of our choice.

The displayed matrix shows how the blocks $R_{st}^{(1:6)}$ of $\mathbf{R}^{(1:6)}$ from Lemma 3.1 are composed of the blocks from $\hat{\mathbf{R}}^{(k)}$, $1 \le k \le m-1$. Using the script, the matrix $\mathbf{R}^{(1:m-1)}$ can be visualized for any m (cf. [4, Example 3.1]).

Next, we assume that $\hat{\mathbf{R}}^{(k)}$ and consequently $\mathbf{R}^{(k)}$ are unitary matrices. Using the Cosine-Sine decomposition of $\hat{\mathbf{R}}^{(k)}$ from (3.1), we can write

$$
\hat{\mathbf{R}}^{(k)} = \begin{bmatrix} C_1^{(k)} & S_1^{(k)} \\ S_2^{(k)} & C_2^{(k)} \end{bmatrix} = \begin{bmatrix} W_1^{(k)} & \\ & W_2^{(k)} \end{bmatrix} \begin{bmatrix} \mathsf{C}^{(k)} & \mathsf{S}^{(k)} \\ -(\mathsf{S}^{(k)})^* & \tilde{\mathsf{C}}^{(k)} \end{bmatrix} \begin{bmatrix} Z_1^{(k)} & \\ & Z_2^{(k)} \end{bmatrix},
$$

where

$$
\mathsf{C}^{(k)} = \begin{bmatrix} \cos \Phi^{(k)} & \\ & I_{n_i - n_j} \end{bmatrix}, \quad \mathsf{S}^{(k)} = \begin{bmatrix} \sin \Phi^{(k)} \\ 0 \end{bmatrix}, \quad \tilde{\mathsf{C}}^{(k)} = \cos \Phi^{(k)}, \qquad n_i \geq n_j,
$$

$$
\mathsf{C}^{(k)} = \cos \Phi^{(k)}, \quad \mathsf{S}^{(k)} = \begin{bmatrix} 0 & \sin \Phi^{(k)} \end{bmatrix}, \quad \tilde{\mathsf{C}}^{(k)} = \begin{bmatrix} I_{n_j - n_i} & \\ & \cos \Phi^{(k)} \end{bmatrix}, \qquad n_i < n_j,
$$

$$
\cos \Phi^{(k)} = \operatorname{diag} \left(\cos \varphi_1^{(k)}, \ldots, \cos \varphi_{\min\{n_i, n_j\}}^{(k)} \right),
$$
$$
\sin \Phi^{(k)} = \operatorname{diag} \left(\sin \varphi_1^{(k)}, \ldots, \sin \varphi_{\min\{n_i, n_j\}}^{(k)} \right),
$$

where $\cos \Phi^{(k)}$ and $\sin \Phi^{(k)}$ are nonnegative diagonal matrices. The angles $\varphi_1^{(k)}$ are called the *canonical angles from the CS decomposition* of $\hat{\mathbf{R}}^{(k)}$ or simply the *canonical angles* of $\hat{\mathbf{R}}^{(k)}$. For simplicity, sometimes we will call them the canonical angles of $\mathbf{R}^{(k)}$.

LEMMA 3.3. *For an arbitrary unitary matrix* $\mathbf{U}$ *of order* $n_p + n_q$, *partitioned according to the partition* (n_p, n_q), *there exists a permutation matrix* $\mathbf{P}$ *such that in the block-matrix*

$$
\tilde{\mathbf{U}} = \mathbf{U}\mathbf{P} = \begin{bmatrix} \tilde{U}_{11} & \tilde{U}_{12} \\ \tilde{U}_{21} & \tilde{U}_{22} \end{bmatrix} \begin{matrix} n_p \\ n_q \end{matrix}
$$

both diagonal blocks are square and nonsingular with the minimum singular value satisfying

$$
\sigma_{\min}(\tilde{U}_{11}) = \sigma_{\min}(\tilde{U}_{22}) \geq f(n_p, n_q),
$$
$$
(3.2) \qquad f(n_p, n_q) \equiv \max \left\{ \frac{3}{\sqrt{(n_q + 1)(4^{n_p} + 6n_p - 1)}}, \binom{n_p + n_q}{n_p}^{-1/2} \right\}.
$$

Proof. The first bound within the braces is from [12], while the second one is obtained from [28, Lemma 3.3] by setting $\epsilon = 0$. $\square$

Let $\mathbf{X} = (X_{st})$ be a square matrix of order n, partitioned according to π. We will use the following notation:

$$
(3.3) \quad \mathbf{Off_L}(\mathbf{X}) = \left[\sum_{t=1}^{m-1} \sum_{\ell=t+1}^{m} \|X_{\ell t}\|_F^2 \right]^{1/2}, \quad \mathbf{Off_R}(\mathbf{X}) = \left[\sum_{\ell=1}^{m-1} \sum_{t=\ell+1}^{m} \|X_{\ell t}\|_F^2 \right]^{1/2},
$$

$$
(3.4) \qquad \boldsymbol{\gamma}_t(\mathbf{X}) = \left[\sum_{\ell=t+1}^{m} \|X_{\ell t}\|_F^2 \right]^{1/2}, \quad \boldsymbol{\rho}_t(\mathbf{X}) = \left[\sum_{\ell=t+1}^{m} \|X_{t\ell}\|_F^2 \right]^{1/2}, \quad 1 \leq t \leq m - 1.
$$

From (2.10) and (3.3) we obtain

$$
\mathbf{Off}^2(\mathbf{X}) = \mathbf{Off_L^2}(\mathbf{X}) + \mathbf{Off_R^2}(\mathbf{X}).
$$

The following lemma is a slight modification of [4, Lemma 4.1]. Here, we consider the special block Jacobi process from Section 2.2, and we use the notation (2.3) for the pivot submatrices of $\mathbf{U}^{(k)}$ and $\mathbf{V}^{(k)}$. We also use the relation (2.13). Thus, for each $k = 1, 2, \ldots, m - 1$, $\hat{\mathbf{U}}^{(k)}$, $\hat{\mathbf{V}}^{(k)}$ are arbitrary unitary matrices of order $n_1 \times n_{k+1}$.

LEMMA 3.4. *Let $\mathbf{G}$ be a triangular matrix of order n, partitioned according to π and such that all diagonal blocks of $\mathbf{G}$ are zero. Let*

$$\mathbf{G}^{(m)} = (\mathbf{U}^{(1:m-1)})^* \mathbf{G} \mathbf{V}^{(1:m-1)}$$

be the matrix obtained by applying the first $m-1$ steps of the special block Jacobi process under the row-cyclic strategy to $\mathbf{G}$.

If $\mathbf{G}$ is upper-triangular, then there exists a $\boldsymbol{\psi}_1 \in [0, \pi/2]$ such that

$$(3.5) \qquad \boldsymbol{\gamma}_1(\mathbf{G}^{(m)}) = \mathbf{Off}(\mathbf{G}^{(m)}) \sin \boldsymbol{\psi}_1, \quad \mathbf{Off}_{\mathbf{R}}(\mathbf{G}^{(m)}) = \mathbf{Off}(\mathbf{G}^{(m)}) \cos \boldsymbol{\psi}_1,$$

where

$$(3.6) \qquad \sin^2 \boldsymbol{\psi}_1 \leq 1 - \sigma_{\min}^2(V_{11}^{(2)} V_{11}^{(3)} \cdots V_{11}^{(m-1)}) \leq 1 - \prod_{k=2}^{m-1} \sigma_{\min}^2(V_{11}^{(k)}).$$

Here $V_{11}^{(k)}$ is the $n_1 \times n_1$ block of $\hat{\mathbf{V}}^{(k)}$, $1 \leq k \leq m-1$.

Let $\psi^{(k)} \in [0, \pi/2]$ be the maximal canonical angle of $\hat{\mathbf{V}}^{(k)}$ such that $\cos \psi^{(k)} = \sigma_{\min}(V_{11}^{(k)})$, $1 \leq k \leq m-1$. Then there exist angles $\check{\psi}^{(k)}$, $0 \leq \check{\psi}^{(k)} \leq \psi^{(k)}$, $2 \leq k \leq m-1$, such that

$$(3.7) \qquad\qquad \sin^2 \boldsymbol{\psi}_1 = 1 - \prod_{k=2}^{m-1} \cos^2 \check{\psi}^{(k)}.$$

If $\mathbf{G}$ is lower-triangular, then there exists a $\boldsymbol{\phi}_1 \in [0, \pi/2]$ such that

$$(3.8) \qquad \boldsymbol{\rho}_1(\mathbf{G}^{(m)}) = \mathbf{Off}(\mathbf{G}^{(m)}) \sin \boldsymbol{\phi}_1, \quad \mathbf{Off}_{\mathbf{L}}(\mathbf{G}^{(m)}) = \mathbf{Off}(\mathbf{G}^{(m)}) \cos \boldsymbol{\phi}_1,$$

where

$$(3.9) \qquad \sin^2 \boldsymbol{\phi}_1 \leq 1 - \sigma_{\min}^2(U_{11}^{(2)} U_{11}^{(3)} \cdots U_{11}^{(m-1)}) \leq 1 - \prod_{k=2}^{m-1} \sigma_{\min}^2(U_{11}^{(k)}).$$

If $\phi^{(k)} \in [0, \pi/2]$ is the maximal canonical angle of $\hat{\mathbf{U}}^{(k)}$ such that $\cos \phi^{(k)} = \sigma_{\min}(U_{11}^{(k)})$, then there exist angles $\check{\phi}^{(k)}$, $0 \leq \check{\phi}^{(k)} \leq \phi^{(k)}$, $2 \leq k \leq m-1$, such that

$$(3.10) \qquad\qquad \sin^2 \boldsymbol{\phi}_1 = 1 - \prod_{k=2}^{m-1} \cos^2 \check{\phi}^{(k)}.$$

Proof. The relations (3.5) and (3.6) are obtained following the lines of the proof of [4, Lemma 4.1]. The only difference comes from our assumption that $\mathbf{U}^{(k)}$ and $\mathbf{V}^{(k)}$ are arbitrary unitary elementary block matrices.

The relations (3.8) and (3.9) are obtained from (3.5) and (3.6) by considering the process

$$[\mathbf{G}^{(k+1)}]^* = [\mathbf{V}^{(k)}]^* [\mathbf{G}^{(k)}]^* \mathbf{U}^{(k)}, \qquad k \geq 1; \qquad \mathbf{G}^{(1)} = \mathbf{G},$$

which is obtained by applying the Hermitian conjugate operator to the original process

$$\mathbf{G}^{(k+1)} = [\mathbf{U}^{(k)}]^* \mathbf{G}^{(k)} \mathbf{V}^{(k)}, \qquad k \geq 1; \qquad \mathbf{G}^{(1)} = \mathbf{G}.$$

The relations (3.7), (3.10) are obtained from (3.6), (3.9), respectively. $\square$

3.2. Nearly diagonal Hermitian matrices. In the global convergence analysis of the Jacobi method for Hermitian matrices, we will consider the case when several diagonal elements are affiliated to the same multiple eigenvalue. This can occur in the later stage of the iterative process. To this end we recall some of the known results concerning nearly diagonal Hermitian matrices [21, 62].

Let $\mathbf{H}$ be a Hermitian matrix of order n, and let

$$(3.11) \quad \lambda_1 = \cdots = \lambda_{\varsigma_1} > \lambda_{\varsigma_1+1} = \cdots = \lambda_{\varsigma_2} > \cdots > \lambda_{\varsigma_{\tilde{p}-1}+1} = \cdots = \lambda_{\varsigma_{\tilde{p}}}, \quad \varsigma_{\tilde{p}} = n,$$

be the non-increasing ordering of the eigenvalues of $\mathbf{H}$. In the case $\tilde{p} = 1$, we have $\mathbf{H} = \lambda_1 \mathbf{I}_n$. This trivial case is not interesting, so we assume $\tilde{p} > 1$.

If we set $\varsigma_0 = 0$, the relation (3.11) implies that

$$(3.12) \qquad\qquad\qquad \nu_r = \varsigma_r - \varsigma_{r-1}$$

is the multiplicity of λ_{ς_r}. Now, let $\lambda_{\varsigma_0} = \lambda_0 = \infty$, $\lambda_{\varsigma_{\tilde{p}+1}} = -\infty$, and

$$(3.13) \qquad\qquad 3\delta_r = \min\{\lambda_{\varsigma_{r-1}} - \lambda_{\varsigma_r}, \, \lambda_{\varsigma_r} - \lambda_{\varsigma_{r+1}}\}, \qquad 1 \le r \le \tilde{p}.$$

From (3.13) we conclude that $3\delta_r$ is the gap in the spectrum of $\mathbf{H}$ associated with the eigenvalues λ_{ς_r}. If

$$(3.14) \qquad\qquad\qquad \delta = \min_{1 \le r \le \tilde{p}} \delta_r,$$

then 3δ is the minimum gap in the spectrum of $\mathbf{H}$.

Now, suppose that $\|\Omega(\mathbf{H})\|_2 < \delta$. Let $\mathbf{P}$ be a permutation matrix such that the diagonal elements of $\tilde{\mathbf{H}} = \mathbf{P}^*\mathbf{H}\mathbf{P} = (\tilde{h}_{rt})$ are ordered non-increasingly. Then by the spectral theorem for Hermitian matrices, we have $|\lambda_t - \tilde{h}_{tt}| \le \|\Omega(\tilde{\mathbf{H}})\|_2 = \|\Omega(\mathbf{H})\|_2 < \delta$, $1 \le t \le n$. Let us partition $\tilde{\mathbf{H}}$ into a block-matrix $\tilde{\mathbf{H}} = (\tilde{\mathbf{H}}_{rt})$, where the block $\tilde{\mathbf{H}}_{rt}$ has dimension $\nu_r \times \nu_t$. This block-matrix partition is defined by $\tilde{\pi} = (\nu_1, \nu_2, \ldots, \nu_{\tilde{p}})$. We call $\tilde{\pi}$ the *natural partition* while $\pi = (n_1, n_2, \ldots, n_p)$ is the *basic partition*.

As the following lemma shows, each diagonal block $\tilde{\mathbf{H}}_{rr}$ can be approximated by $\lambda_{\varsigma_r} \mathbf{I}_{\nu_r}$. For computational purposes and for theoretical simplicity, we will use the Frobenius norm instead of the spectral norm.

LEMMA 3.5. *Let $\mathbf{H}$ be a Hermitian matrix of order n, and let*

$$(3.15) \qquad\qquad\qquad \mathbf{off}(\mathbf{H}) \le \delta,$$

where δ is defined by (3.14). Then there is a permutation matrix $\mathbf{P}$ such that for the matrix $\tilde{\mathbf{H}} = \mathbf{P}^*\mathbf{H}\mathbf{P} = (\tilde{\mathbf{H}}_{rt})$, partitioned according to the natural partition $\tilde{\pi}$, we have for all $1 \le r \le \tilde{p}$,

$$(3.16) \qquad \|\tilde{\mathbf{H}}_{rr} - \lambda_{\varsigma_r}\mathbf{I}_{\nu_r}\|_F \le \frac{0.44}{\delta_r} \sum_{\substack{t=1 \\ t \ne r}}^{\tilde{p}} \|\tilde{\mathbf{H}}_{rt}\|_F^2 \le 0.22 \frac{\mathbf{Off}^2(\tilde{\mathbf{H}})}{\delta_r} \le 0.22 \frac{\mathbf{off}^2(\mathbf{H})}{\delta_r}.$$

We also have

$$(3.17) \qquad \sum_{r=1}^{\tilde{p}} \|\tilde{\mathbf{H}}_{rr} - \lambda_{\varsigma_r}\mathbf{I}_{\nu_r}\|_F^2 \le 0.1 \frac{\mathbf{Off}^4(\tilde{\mathbf{H}})}{\delta^2} \le 0.1 \frac{\mathbf{off}^4(\mathbf{H})}{\delta^2}.$$

Proof. The first inequality in (3.16) is proved in [21, Lemma 2.1] for real symmetric matrices, but the proof naturally extends to complex Hermitian matrices. The second inequality in (3.16) holds because $\tilde{\mathbf{H}}$ is Hermitian. The last inequality in (3.16) comes from the fact that $\mathbf{Off}(\tilde{\mathbf{H}}) \le \mathbf{off}(\tilde{\mathbf{H}})$ and $\mathbf{off}(\tilde{\mathbf{H}}) = \mathbf{off}(\mathbf{H})$ (see (2.9) and (2.10)). The assertion (3.17) is implied by (3.16). We have also used: $0.22 \cdot 0.44 = 0.0968 < 0.1$. $\square$

4. The off-norm reduction related to one cycle of the block Jacobi method for Hermitian matrices. Here, we consider the first cycle of the block Jacobi method for Hermitian matrices under the de Rijk pivot strategy. We first consider the core algorithm. Then we derive certain important estimates related to the first $m - 1$ steps of the block method. After that we derive the off-norm reduction bounds for the whole cycle.

Suppose $\mathbf{H}$ is a Hermitian matrix of order n, and we are solving the eigenvalue problem $\mathbf{H}x = \lambda x$, $x \neq 0$. We apply the block Jacobi method under the *bdR1* pivot strategy to $\mathbf{H}$ and consider the first cycle of the iterative process. Later, at the end of Section 4.3, we will consider the pivot strategy *bdR2*.

Using the relations (2.15) and (2.16)–(2.17) from Section 2.5.1, we have

$$\mathbf{H}^{(M+1)} = \mathbf{H}^{[1]} = [\mathbf{U}^{[1]}]^* \mathbf{H}^{(1)} \mathbf{U}^{[1]},$$

(4.1) $$\mathbf{U}^{[1]} = \mathbf{P}_1 \mathbf{U}_{1,2:m} \mathbf{P}_2 \mathbf{U}_{2,3:m} \mathbf{P}_3 \mathbf{U}_{3,4:m} \cdots \mathbf{P}_{m-2} \mathbf{U}_{m-2,m-1:m} \mathbf{P}_{m-1} \mathbf{U}_{m-1,m}.$$

Here,

(4.2) $$\mathbf{H}^{(1)} = [\mathbf{U}^{(0)}]^* \mathbf{H} \, \mathbf{U}^{(0)},$$

where the unitary matrix $\mathbf{U}^{(0)}$ diagonalizes the diagonal blocks of $\mathbf{H}$ and orders the diagonal elements of each diagonal block of $\mathbf{H}^{(1)}$ non-increasingly.

Using the expression for $\mathbf{U}^{[1]}$ from the relation (4.1), we can split the process in such a way that in the first cycle the method generates the sequence of $M + m - 1$ matrices:

(4.3)
$$\mathbf{H}_{\mathbf{P}_1}^{(1)}, \mathbf{H}^{(2)}, \mathbf{H}^{(3)}, \dots, \mathbf{H}^{(m)}, \mathbf{H}_{\mathbf{P}_2}^{(m)}, \mathbf{H}^{(m+1)}, \dots, \mathbf{H}^{(2m-2)}, \mathbf{H}_{\mathbf{P}_3}^{(2m-2)}, \mathbf{H}^{(2m-1)},$$
$$\dots, \mathbf{H}^{(3m-5)}, \dots, \mathbf{H}_{\mathbf{P}_{m-2}}^{(M-2)}, \mathbf{H}^{(M-1)}, \mathbf{H}^{(M)}, \mathbf{H}_{\mathbf{P}_{m-1}}^{(M)}, \mathbf{H}^{(M+1)}.$$

Now, let us pay attention to the superscripts appearing in the sequence of matrices (4.3). Set $\tau(0, m) = 1$, and for $r = 1, 2, \dots, m - 1$, let

(4.4) $$\tau(r, m) = 1 + (m - 1) + (m - 2) + \cdots + (m - r) = \frac{(2m - r - 1)r}{2} + 1.$$

The definition (4.4) also covers the case $r = 0$. We have

$$\tau(1, m) = m, \quad \tau(2, m) = 2m - 2, \ \dots, \tau(m - 2, m) = M, \quad \tau(m - 1, m) = M + 1.$$

For each r, $1 \leq r \leq m - 1$, we have

(4.5) $$\mathbf{H}_{\mathbf{P}_r}^{(\tau(r-1,m))} = \mathbf{P}_r^* \mathbf{H}^{(\tau(r-1,m))} \mathbf{P}_r,$$

(4.6) $$\mathbf{H}^{(\tau(r-1,m)+1)} = [\mathbf{U}^{(\tau(r-1,m))}]^* \mathbf{H}_{\mathbf{P}_r}^{(\tau(r-1,m))} \mathbf{U}^{(\tau(r-1,m))},$$

(4.7) $$\mathbf{H}^{(k+1)} = [\mathbf{U}^{(k)}]^* \mathbf{H}^{(k)} \mathbf{U}^{(k)}, \quad k = \tau(r - 1, m) + 1, \dots, \tau(r, m) - 1.$$

If $\mathbf{P}_r = \mathbf{I}_n$ for some r, then $\mathbf{H}_{\mathbf{P}_r}^{(\tau(r-1,m))} = \mathbf{H}^{(\tau(r-1,m))}$, and the term $\mathbf{H}_{\mathbf{P}_r}^{(\tau(r-1,m))}$ can be removed from the sequence of matrices defined by (4.5)–(4.7). After the similarity transformation with $\mathbf{P}_r$ is completed, the partition π of n has to be updated by interchanging the entries r and r'. From the relations (4.5)–(4.7), we see that π can be updated at most $m - 1$ times in one cycle. As we have already done, *we will use the same notation π for all updated partitions during the whole iterative process.*

To simplify notation, when we write $\lim_{k \to \infty} \mathbf{H}^{(k)}$ or $(\mathbf{H}^{(k)}, k \geq 1)$, we assume that the matrices $\mathbf{H}_{\mathbf{P}_r}^{(\tau(r-1,m))}$ are included in the sequence $(\mathbf{H}^{(k)}, k \geq 1)$.

4.1. The core algorithm. As we have discussed earlier, the *core* or *kernel algorithm* of the block method is used to diagonalize the pivot submatrix. The core algorithm has to be a globally convergent method so that in each step of the block method, the matrix $\hat{\mathbf{U}}^{(k)}$ exists. To prove global convergence of the block method, we may have to modify the core algorithm (see [12, 28]).

4.1.1. The basic and modified core algorithm. By the *basic core algorithm* we mean a method that is used for solving the Hermitian eigenvalue problem of the pivot submatrix $\hat{\mathbf{H}}^{(k)}$, such as the QR method or some convergent simple or block Jacobi method, without using additional permutations, as it is discussed below.

The *modified core algorithm* is designed to achieve two goals:
  (a) The unitary transformation matrices $\mathbf{U}^{(k)}$ are such that

$$(4.8) \qquad\qquad \sigma_{\min}(U^{(k)}_{i(k)i(k)}) \geq C_\pi > 0, \qquad k \geq 1,$$

where

$$(4.9) \qquad\qquad C_\pi = \min_{1 \leq p < q \leq m} f(n_p, n_q)$$

and $f(n_p, n_q)$ is defined in (3.2).
  (b) The unitary transformation matrix $\mathbf{U}^{(k)}$ is such that the diagonal elements of $H^{(k+1)}_{i(k)i(k)}$ and of $H^{(k+1)}_{j(k)j(k)}$ are ordered non-increasingly.

The condition (b) ensures that the diagonal elements of every diagonal block of any iteration matrix $\mathbf{H}^{(k)}$ are ordered non-increasingly. The condition (a) is used in the convergence proof. We call $\mathbf{U}^{(k)}$ satisfying conditions (a) and (b) the *modified transformation matrix*.

A straightforward way to compute $\hat{\mathbf{U}}^{(k)}$ such that it satisfies conditions (a) and (b) is as follows: First, compute any eigenvector matrix $\hat{\mathbf{W}}^{(k)}$ of $\hat{\mathbf{H}}^{(k)}$. For this purpose the basic core algorithm is used. Then find a permutation matrix $\hat{\mathbf{P}}^{(k)}$ such that $\hat{\mathbf{W}}^{(k)}\hat{\mathbf{P}}^{(k)}$ satisfies condition (a). Finally, find the permutation matrix $\hat{\mathbf{Q}}^{(k)} = \text{diag}(Q^{(k)}_{i(k)i(k)}, Q^{(k)}_{j(k)j(k)})$, where $Q^{(k)}_{i(k)i(k)}$ $(Q^{(k)}_{j(k)j(k)})$ is a permutation matrix of order $n_{i(k)}$ $(n_{j(k)})$ that orders the diagonal elements of $H^{(k+1)}_{i(k)i(k)}$ $(H^{(k+1)}_{j(k)j(k)})$ non-increasingly. Then set $\hat{\mathbf{U}}^{(k)} = \hat{\mathbf{W}}^{(k)}(\hat{\mathbf{P}}^{(k)}\hat{\mathbf{Q}}^{(k)})$. It is easy to see that $\mathbf{U}^{(k)} = \mathcal{E}(i(k), j(k); \hat{\mathbf{U}}^{(k)})$ satisfies both conditions (a) and (b).

4.1.2. The relaxed modified core algorithm. In certain situations we can simplify the matrix $\hat{\mathbf{P}}^{(k)}$ or even replace it with the identity matrix. Such a permutation matrix will be denoted by $\hat{\mathbf{P}}^{(k)}_0$. In [12], the matrix $\hat{\mathbf{P}}^{(k)}$ is computed using the QR factorization with column pivoting of the $n_i \times (n_i + n_j)$ block $[W^{(k)}_{11} \ W^{(k)}_{12}]$ of $\hat{\mathbf{W}}^{(k)}$. It is computed as a product of transposition matrices, $\hat{\mathbf{P}}^{(k)} = \hat{\mathbf{I}}^{(k)}_{11'}\hat{\mathbf{I}}^{(k)}_{22'}\cdots\hat{\mathbf{I}}^{(k)}_{n_i n'_i}$, where $\hat{\mathbf{I}}^{(k)}_{11'}, \ldots, \hat{\mathbf{I}}^{(k)}_{n_i n'_i}$ are obtained by an appropriate QR factorization algorithm with column pivoting.

The *relaxed modified core algorithm* is obtained by replacing the permutation matrix $\hat{\mathbf{P}}^{(k)}$ with $\hat{\mathbf{P}}^{(k)}_0$. The matrix $\hat{\mathbf{P}}^{(k)}_0$ is obtained from $\hat{\mathbf{P}}^{(k)}$ by adding a simple rule (see [28, Definition 3.2]):

  • if $\tilde{r}$ is the smallest number from the set $\{1, 2, \ldots, n_i\}$ such that $\tilde{r}' > n_i$, then set $\hat{\mathbf{P}}^{(k)}_0 = \hat{\mathbf{I}}^{(k)}_{\tilde{r}\tilde{r}'}\cdots\hat{\mathbf{I}}^{(k)}_{n_i n'_i}$. If such $\tilde{r}$ does not exist, then set $\hat{\mathbf{P}}^{(k)}_0 = \mathbf{I}_n$.

This actually means that $\hat{\mathbf{P}}^{(k)}_0 = \mathbf{I}_n$, provided that $\hat{\mathbf{P}}^{(k)} = \text{diag}(\hat{\mathsf{P}}^{(k)}_{n_{i(k)}}, \mathbf{I}_{n_{j(k)}})$, where $\hat{\mathsf{P}}^{(k)}_{n_{i(k)}}$ is a permutation matrix of order $n_{i(k)}$.

If the relaxed modified core algorithm is used, then we will have $\hat{\mathbf{U}}^{(k)} = \hat{\mathbf{W}}^{(k)}(\hat{\mathbf{P}}_0^{(k)}\hat{\mathbf{Q}}^{(k)})$ and $\mathbf{U}^{(k)} = \mathcal{E}(i(k), j(k); \hat{\mathbf{U}}^{(k)})$. As earlier, the permutation matrix $\hat{\mathbf{Q}}^{(k)}$ serves to ensure that the condition (b) is satisfied.

In [28, Section 3.2]) it has been proved that condition (a) holds with the same bound C_π for both the modified core algorithm and the relaxed modified core algorithm. We note that the matrices $\hat{\mathbf{Q}}^{(k)}$ do not change the singular values of $U_{i(k)i(k)}^{(k)}$ and $U_{j(k)j(k)}^{(k)}$. Therefore, they do not change the relation (4.8).

Often we will use the term *modified core algorithm with the attribute* R instead of the relaxed modified core algorithm.

REMARK 4.1. As it will become evident later (see the proof of Theorem 5.9), it is advantageous to choose the indices $1'$, $2'$,...,n_i' as small as possible. So, if during the computation there are two or more candidates for r', $1 \le r \le n_i$, then we choose the smallest one.

4.1.3. The well-behaved core algorithm. Using [28, Definition 4.1], we say that a method for solving the Hermitian eigenproblem is *well-behaved* if it is globally convergent and for every $\epsilon > 0$ there is a $\zeta > 0$ such that

$$\mathbf{off}(\mathbf{H}) < \zeta \Rightarrow \|\mathbf{H} - \mathbf{\Lambda}\|_F < \epsilon,$$

where $\mathbf{\Lambda}$ is the limit of the sequence of matrices generated by applying the method to $\mathbf{H}$.

If the core algorithm of the block Jacobi method is well-behaved, then the following series of implications hold:

$$\begin{aligned}(4.10)\qquad \lim_{k\to\infty} \mathbf{off}(\mathbf{H}^{(k)}) = 0 &\implies \lim_{k\to\infty} \mathbf{off}(\hat{\mathbf{H}}^{(k)}) = 0\\ &\implies \lim_{k\to\infty}\left([\hat{\mathbf{U}}^{(k)}]^*\hat{\mathbf{H}}^{(k)}\hat{\mathbf{U}}^{(k)} - \hat{\mathbf{H}}^{(k)}\right) = 0.\end{aligned}$$

The Hermitian block Jacobi method satisfies the relation (2.8). Hence, the sequence $(\mathbf{off}(\mathbf{H}^{(k)}), k \ge 0)$ is non-increasing and therefore convergent. If it converges to zero, then by Lemma 3.5 in the later stage of the block process, the diagonal elements are affiliated to the eigenvalues. If the core algorithm is well-behaved, then it will not change the affiliation of the diagonal elements. Then, in the case of the modified core algorithm with or without the attribute R, we will show that the permutations $\mathbf{P}_0^{(k)}$ or $\mathbf{P}^{(k)}$ and $\mathbf{Q}^{(k)}$ do not ruin the affiliation of the diagonal elements.

An example of the well-behaved eigenvalue method for Hermitian matrices is any convergent cyclic simple Jacobi method. In particular, this is true if the pivot strategy is the generalized cyclic strategy [3, 30]. Also, the classical Jacobi method is well-behaved. In the proof of Theorem 5.4 we will show that the simple Jacobi method under the de Rijk pivot strategy is well-behaved.

4.2. The first $m - 1$ steps of the block method. We have

$$(4.11)\qquad \mathbf{H}^{(m)} = [\mathbf{U}^{(1:m-1)}]^*\mathbf{H}_{\mathbf{P}_1}^{(1)}\mathbf{U}^{(1:m-1)}, \quad \mathbf{H}_{\mathbf{P}_1}^{(1)} = \mathbf{P}_1^*\mathbf{H}^{(1)}\mathbf{P}_1,$$

$$\mathbf{U}^{(1:m-1)} = \mathbf{U}^{(1)}\mathbf{U}^{(2)}\cdots\mathbf{U}^{(m-1)} = \mathbf{U}_{12}\mathbf{U}_{13}\cdots\mathbf{U}_{1m} = \mathbf{U}_{1,2:m}.$$

In the following analysis *we presume that the matrices $\mathbf{U}^{(k)}$, $1 \le k \le m-1$, are known since they have been computed by the block Jacobi method*. The matrices $\mathbf{U}_{1j} = \mathcal{E}(1, j; \hat{\mathbf{U}}_{1j})$ are defined by unitary matrices $\hat{\mathbf{U}}_{1j}$ of order $n_1 + n_j$. Each matrix $\hat{\mathbf{U}}_{1j} = \hat{\mathbf{U}}^{(j-1)}$ is partitioned as in the relation (2.3).

Let

$$
(4.12) \qquad \mathbf{H}_{\mathbf{P}_1}^{(1)} = \mathbf{L}^{(1)} + \mathbf{D}^{(1)} + \mathbf{R}^{(1)}, \qquad \mathbf{R}^{(1)} = [\mathbf{L}^{(1)}]^*,
$$

where $\mathbf{D}^{(1)}$ and $\mathbf{L}^{(1)}$ ($\mathbf{R}^{(1)}$) are the block-diagonal and the strictly lower- (upper-) block-triangular part of $\mathbf{H}_{\mathbf{P}_1}^{(1)}$. Then for $\dot{\mathbf{H}}_{\mathbf{P}_1}^{(1)} = \Omega(\mathbf{H}_{\mathbf{P}_1}^{(1)})$, we have $\dot{\mathbf{H}}_{\mathbf{P}_1}^{(1)} = \overset{\bullet}{\mathbf{H}}_{\mathbf{P}_1}^{(1)}$ and

$$
(4.13) \qquad \overset{\bullet}{\mathbf{H}}_{\mathbf{P}_1}^{(1)} = \mathbf{L}^{(1)} + \mathbf{R}^{(1)}.
$$

Let us now define matrices $\mathbf{L}^{(k)}$, $\mathbf{R}^{(k)}$, $k = 2, \dots, m$. They carry the same block-matrix partition as $\mathbf{H}_{\mathbf{P}_1}^{(1)}$ and are computed by the iterative process (cf. (2.7))

$$
(4.14) \quad \mathbf{L}^{(k+1)} = [\mathbf{U}^{(k)}]^* \mathcal{N}_{1,k+1}(\mathbf{L}^{(k)})\mathbf{U}^{(k)}, \quad \mathbf{R}^{(k+1)} = [\mathbf{U}^{(k)}]^* \mathcal{N}_{1,k+1}(\mathbf{R}^{(k)})\mathbf{U}^{(k)}.
$$

By analogy, we say that in step k of the process (4.14), the pivot submatrix of $\mathbf{L}^{(k)}$ ($\mathbf{R}^{(k)}$) is $\hat{\mathbf{L}}^{(k)}$ ($\hat{\mathbf{R}}^{(k)}$). From the relation (4.14) we see that in step k, $1 \le k \le m - 1$, these two pivot submatrices become the null matrices of order $n_1 + n_{k+1}$.

Let us consider the evolution of the matrices $\mathbf{L}^{(k)}$, $\mathbf{R}^{(k)}$ from their origins $\mathbf{L}^{(1)}$, $\mathbf{R}^{(1)}$ to $\mathbf{L}^{(m)}$, $\mathbf{R}^{(m)}$, respectively. It is easy to see that $\mathbf{L}^{(k)}$ ($\mathbf{R}^{(k)}$) is lower- (upper-) block-triangular, except for the first block-row (block-column).

To clarify, we depict $\mathbf{L}^{(k)}$ and $\mathbf{R}^{(k)}$ for the case $m = 8$, $k = 5$, where 0 and X denote a zero and a possibly non-zero block, respectively:

$$
\mathbf{L}^{(5)} = \left[\begin{array}{c|ccccccc} 0 & X & X & X & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & X & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & X & X & 0 & 0 & 0 & 0 & 0 \\ 0 & X & X & X & 0 & 0 & 0 & 0 \\ X & X & X & X & X & 0 & 0 & 0 \\ X & X & X & X & X & X & 0 & 0 \\ X & X & X & X & X & X & X & 0 \end{array}\right], \quad \mathbf{R}^{(5)} = \left[\begin{array}{c|ccccccc} 0 & 0 & 0 & 0 & 0 & X & X & X \\ \hline X & 0 & X & X & X & X & X & X \\ X & 0 & 0 & X & X & X & X & X \\ X & 0 & 0 & 0 & X & X & X & X \\ 0 & 0 & 0 & 0 & 0 & X & X & X \\ 0 & 0 & 0 & 0 & 0 & 0 & X & X \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & X \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].
$$

From the relation (4.14) we see that in the next (that is, in the 5th) step, we will obtain an additional zero (non-zero) block in the first block-column (block-row) of $\mathbf{L}^{(6)}$ and a non-zero (zero) block in the first block-column (block-row) of $\mathbf{R}^{(6)}$. This proves the structure of the zero and non-zero blocks of $\mathbf{L}^{(k)}$, $\mathbf{R}^{(k)}$ for $1 \le k \le m$.

LEMMA 4.2. *Let the matrices $\mathbf{L}^{(1)}$ and $\mathbf{R}^{(1)}$ be defined by* (4.13) *and $\mathbf{L}^{(k)}$, $\mathbf{R}^{(k)}$, $2 \le k \le m$, by* (4.14). *Then we have*

$$
(4.15) \qquad \overset{\bullet}{\mathbf{H}}^{(k)} = \mathbf{L}^{(k)} + \mathbf{R}^{(k)}, \qquad \mathbf{R}^{(k)} = [\mathbf{L}^{(k)}]^*, \qquad 2 \le k \le m.
$$

Proof. Note that in the relation (4.15) we have $H_{st}^{(k)} = L_{st}^{(k)} + R_{st}^{(k)}$, $s \ne t$, and at least one of the blocks $L_{st}^{(k)}$ or $R_{st}^{(k)}$ is zero. Let us extend the relation (4.15) so that for $k = 1$ instead of $\overset{\bullet}{\mathbf{H}}^{(1)} = \mathbf{L}^{(1)} + \mathbf{R}^{(1)}$ we have the relation (4.13). We prove the so updated relation (4.15) by induction with respect to k.

For $k = 1$ the updated relation (4.15) holds since it reduces to (4.13). Suppose that the updated relation (4.15) holds for k, $1 \le k < m$. We have to prove that it holds for $k + 1$. So, let us consider step k of the block Jacobi method and of the process (4.14), which computes $\mathbf{L}^{(k+1)}$ ($\mathbf{R}^{(k+1)}$) from $\mathbf{L}^{(k)}$ ($\mathbf{R}^{(k)}$).

In step k, the pivot submatrix $\hat{\mathbf{H}}^{(k)}$ becomes a diagonal matrix of order $n_1 + n_{k+1}$. It means that $\dot{\mathbf{H}}^{(k+1)}$ has zero diagonal blocks and additionally the $(1, k + 1)$- and $(k + 1, 1)$-blocks are zero.

From the relation (4.14) we see that the pivot submatrices $\hat{\mathbf{L}}^{(k)}$ and $\hat{\mathbf{R}}^{(k)}$ become the null matrices of order $n_1 + n_{k+1}$. From the zero-nonzero structure of $\mathbf{L}^{(k)}$ and $\mathbf{R}^{(k)}$, we know that all diagonal blocks of $\mathbf{L}^{(k+1)}$ and $\mathbf{R}^{(k+1)}$ are zero.

Using the induction hypothesis, i.e., the relation (4.15), from the relation (4.14) we obtain

$$\dot{\mathbf{H}}^{(k+1)} = [\mathbf{U}^{(k)}]^* \boldsymbol{\mathcal{N}}_{1,k+1}(\dot{\mathbf{H}}^{(k)})\mathbf{U}^{(k)} = [\mathbf{U}^{(k)}]^* \boldsymbol{\mathcal{N}}_{1,k+1}(\mathbf{L}^{(k)} + \mathbf{R}^{(k)})\mathbf{U}^{(k)}$$
$$= [\mathbf{U}^{(k)}]^* \boldsymbol{\mathcal{N}}_{1,k+1}(\mathbf{L}^{(k)})\mathbf{U}^{(k)} + [\mathbf{U}^{(k)}]^* \boldsymbol{\mathcal{N}}_{1,k+1}(\mathbf{R}^{(k)})\mathbf{U}^{(k)} = \mathbf{L}^{(k+1)} + \mathbf{R}^{(k+1)}.$$

This proves the induction step and the lemma. □

Let $\mathbf{E}_t$ be the $n \times (n - t)$ matrix defined by

$$(4.16) \qquad \mathbf{E}_t = \begin{bmatrix} 0 \\ \mathbf{I}_{n-t} \end{bmatrix}, \qquad 1 \le t \le n - 1.$$

If $\mathbf{X}$ is a square matrix of order n, then $\mathbf{E}_t^* \mathbf{X} \mathbf{E}_t$ is the right bottom principal submatrix of $\mathbf{X}$ of order $n - t$. From the relation (4.14) we see that we can apply Lemma 3.4 to the triangular matrices $\mathbf{L}^{(0)}$ and $\mathbf{R}^{(0)}$. We obtain the following lemma:

LEMMA 4.3. *Let $\mathbf{H}$ be a Hermitian matrix of order n partitioned according to π. Let us apply to $\mathbf{H}$ the first $m - 1$ steps of the block Jacobi method under the bdR1 pivot strategy. The method generates matrices $\mathbf{H}^{(1)}, \mathbf{H}_{\mathbf{P}_1}^{(1)}, \mathbf{H}^{(2)}, \ldots, \mathbf{H}^{(m)}$ and $\mathbf{U}^{(0)}, \mathbf{U}^{(1)}, \ldots, \mathbf{U}^{(m-1)}$ which satisfy the relations (4.5)–(4.7).*

Let $\mathbf{L}^{(k)}, \mathbf{R}^{(k)}, 1 \le k \le m$, be defined by (4.12)–(4.14). Then the relation (4.15) holds, and for the matrix $\dot{\mathbf{H}}^{(m)}$ we have

$$(4.17) \qquad \dot{\mathbf{H}}^{(m)} = \dot{\mathbf{H}}^{(m)} = \mathbf{L}^{(m)} + \mathbf{R}^{(m)}, \qquad \mathbf{R}^{(m)} = [\mathbf{L}^{(m)}]^*.$$

There exists a $\varphi_1 \in [0, \pi/2]$ such that

$$(4.18) \qquad \boldsymbol{\gamma}_1(\mathbf{R}^{(m)}) = \mathbf{Off}(\mathbf{R}^{(m)}) \sin \varphi_1, \quad \mathbf{Off}_{\mathbf{R}}(\mathbf{R}^{(m)}) = \mathbf{Off}(\mathbf{R}^{(m)}) \cos \varphi_1,$$

$$(4.19) \qquad \boldsymbol{\rho}_1(\mathbf{L}^{(m)}) = \mathbf{Off}(\mathbf{L}^{(m)}) \sin \varphi_1, \quad \mathbf{Off}_{\mathbf{L}}(\mathbf{L}^{(m)}) = \mathbf{Off}(\mathbf{L}^{(m)}) \cos \varphi_1,$$

where

$$(4.20) \qquad \sin^2 \varphi_1 \le 1 - \sigma_{\min}^2(U_{11}^{(2)} U_{11}^{(3)} \cdots U_{11}^{(m-1)}) \le 1 - \prod_{k=2}^{m-1} \sigma_{\min}^2(U_{11}^{(k)}).$$

If $\varphi^{(k)} \in [0, \pi/2]$ is the maximal canonical angle from the Cosine-Sine decomposition of $U_{11}^{(k)}$ such that $\cos \varphi^{(k)} = \sigma_{\min}(U_{11}^{(k)})$, then there exist angles $\check{\varphi}^{(k)}$, $0 \le \check{\varphi}^{(k)} \le \varphi^{(k)}$, $2 \le k \le m - 1$, such that

$$(4.21) \qquad \sin^2 \varphi_1 = 1 - \prod_{k=2}^{m-1} \cos^2 \check{\varphi}^{(k)}.$$

We also have

$$(4.22) \qquad \boldsymbol{\rho}_1(\mathbf{L}^{(m)}) = \boldsymbol{\rho}_1(\mathbf{H}^{(m)}) = \boldsymbol{\gamma}_1(\mathbf{H}^{(m)}) = \boldsymbol{\gamma}_1(\mathbf{R}^{(m)}),$$

and

$$(4.23) \qquad \mathbf{Off}_{\mathbf{L}}(\mathbf{L}^{(m)}) = \frac{\sqrt{2}}{2} \|\mathbf{E}_{n_1}^* \dot{\mathbf{H}}^{(m)} \mathbf{E}_{n_1}\|_F = \mathbf{Off}_{\mathbf{R}}(\mathbf{R}^{(m)}).$$

Proof. The relation (4.17) has been proved in Lemma 4.2. The relations (4.18)–(4.21) follow directly from Lemma 3.4. One applies Lemma 3.4 to the matrices $\mathbf{L}^{(k)}$ and $\mathbf{R}^{(k)}$. Since $\mathbf{V}^{(k)} = \mathbf{U}^{(k)}$ and $\mathbf{R}^{(k)} = [\mathbf{L}^{(k)}]^*$, for all k, we have $\psi^{(k)} = \phi^{(k)} = \varphi^{(k)}$, $2 \le k \le m - 1$. Finally, since $\dot{\mathbf{H}}^{(m-1)}$ is Hermitian, the relations (4.22) and (4.23) are obtained from (4.17). $\square$

The Hermitian matrix $\dot{\mathbf{H}}^{(m)} = \mathbf{L}^{(m)} + \mathbf{R}^{(m)}$ is made of the zero blocks on the diagonal and at positions $(1, m)$, $(m, 1)$. Other blocks can be nonzero. The structure of $\dot{\mathbf{H}}^{(m)}$ for the case $m = 8$ is shown below:

$$\dot{\mathbf{H}}^{(m)} = \dot{\mathbf{H}}^{(8)} = \begin{bmatrix} 0 & L_{12}^{(8)} & L_{13}^{(8)} & L_{14}^{(8)} & L_{15}^{(8)} & L_{16}^{(8)} & L_{17}^{(8)} & 0 \\ R_{21}^{(8)} & 0 & R_{23}^{(8)} & R_{24}^{(8)} & R_{25}^{(8)} & R_{26}^{(8)} & R_{27}^{(8)} & R_{28}^{(8)} \\ R_{31}^{(8)} & L_{32}^{(8)} & 0 & R_{34}^{(8)} & R_{35}^{(8)} & R_{36}^{(8)} & R_{37}^{(8)} & R_{38}^{(8)} \\ R_{41}^{(8)} & L_{42}^{(8)} & L_{43}^{(8)} & 0 & R_{45}^{(8)} & R_{46}^{(8)} & R_{47}^{(8)} & R_{48}^{(8)} \\ R_{51}^{(8)} & L_{52}^{(8)} & L_{53}^{(8)} & L_{54}^{(8)} & 0 & R_{56}^{(8)} & R_{57}^{(8)} & R_{58}^{(8)} \\ R_{61}^{(8)} & L_{62}^{(8)} & L_{63}^{(8)} & L_{64}^{(8)} & L_{65}^{(8)} & 0 & R_{67}^{(8)} & R_{68}^{(8)} \\ R_{71}^{(8)} & L_{72}^{(8)} & L_{73}^{(8)} & L_{74}^{(8)} & L_{75}^{(8)} & L_{76}^{(8)} & 0 & R_{78}^{(8)} \\ 0 & L_{82}^{(8)} & L_{83}^{(8)} & L_{84}^{(8)} & L_{85}^{(8)} & L_{86}^{(8)} & L_{87}^{(8)} & 0 \end{bmatrix}.$$

4.3. The remaining steps of the first cycle.
Let us consider the transformation $\mathbf{H}^{(m)} \mapsto \mathbf{H}_{\mathbf{P}_2}^{(m)}$, where

$$\mathbf{H}_{\mathbf{P}_2}^{(m)} = \mathbf{P}_2^* \mathbf{H}^{(m)} \mathbf{P}_2 = \begin{bmatrix} I_{n_1} & \\ & P_{n-n_1} \end{bmatrix}^* \mathbf{H}^{(m)} \begin{bmatrix} I_{n_1} & \\ & P_{n-n_1} \end{bmatrix}.$$

The similarity transformation with $\mathbf{P}_2$ interchanges the block-columns (block-rows) 2 and $2'$, where $2' \ge 2$. This transformation does not change the Frobenius norm of the first block-column and block-row. In addition, it does not change $\|\mathbf{E}_{n_1}^* \dot{\mathbf{H}}^{(m)} \mathbf{E}_{n_1}\|_F$. Since $\mathbf{E}_{n_1}^* \dot{\mathbf{H}}_{\mathbf{P}_2}^{(m)} \mathbf{E}_{n_1}$ is Hermitian we have

$$\begin{aligned}
\mathbf{Off_R}(\mathbf{E}_{n_1}^* \mathbf{H}_{\mathbf{P}_2}^{(m)} \mathbf{E}_{n_1}) &= \frac{\sqrt{2}}{2} \|\mathbf{E}_{n_1}^* \dot{\mathbf{H}}_{\mathbf{P}_2}^{(m)} \mathbf{E}_{n_1}\|_F = \mathbf{Off_L}(\mathbf{E}_{n_1}^* \mathbf{H}_{\mathbf{P}_2}^{(m)} \mathbf{E}_{n_1}) \\
(4.24) \qquad\qquad &= \frac{\sqrt{2}}{2} \mathbf{Off}(\mathbf{E}_{n_1}^* \mathbf{H}^{(m)} \mathbf{E}_{n_1}).
\end{aligned}$$

All later transformations in the first cycle do not change the Frobenius norm of the first block-column and block-row. Therefore, we have

$$(4.25) \qquad \boldsymbol{\rho}_1(\mathbf{H}^{(M+1)}) = \boldsymbol{\rho}_1(\mathbf{H}^{(m)}) = \boldsymbol{\gamma}_1(\mathbf{H}^{(m)}) = \gamma_1(\mathbf{H}^{(M+1)}).$$

Let us now consider the remaining $M - m + 1$ steps of the first cycle. In particular, we consider steps that nullify pivot blocks in the rth block-row and block-column.

LEMMA 4.4. *Suppose we are in the stage of the process when all blocks in the block-row and block-column $r - 1$ have all been zeroed. We assume $2 \le r \le m - 1$. Then we have:*

(i) *The similarity transformation with $\mathbf{P}_r$ does not change the Frobenius norm of any of the first $r - 1$ block-rows and block-columns of $\dot{\mathbf{H}}^{(\tau(r-1,m))}$ and of $\mathbf{H}^{(\tau(r-1,m))}$. It also does not change $\mathbf{Off}(\mathbf{E}_{\mathbf{s}_{r-1}}^* \mathbf{H}^{(\tau(r-1,m))} \mathbf{E}_{\mathbf{s}_{r-1}})$.*

(ii) *The similarity transformation with $\mathbf{U}^{(k)}$, $\tau(r - 1, m) \le k \le \tau(r, m) - 1$, does not change the Frobenius norm of any of the first $r - 1$ block-rows and block-columns of $\dot{\mathbf{H}}^{(\tau(r-1,m))}$ and of $\mathbf{H}^{(\tau(r-1,m))}$.*

(iii) *The similarity transformation with* $\mathbf{U}^{(k)}$, $\tau(r-1,m) \leq k \leq \tau(r,m)-1$, *does not increase* $\mathbf{off}(\mathbf{E}_{\mathbf{s}_{r-1}}^* \mathbf{H}^{\tau(r-1,m)} \mathbf{E}_{\mathbf{s}_{r-1}})$. *We have*

$$
(4.26) \quad
\begin{aligned}
\mathbf{Off}^2 &(\mathbf{E}_{\mathbf{s}_{r-1}}^* \mathbf{H}^{(\tau(r,m))} \mathbf{E}_{\mathbf{s}_{r-1}}) \\
&= \mathbf{Off}^2(\mathbf{E}_{\mathbf{s}_{r-1}}^* \mathbf{H}^{(\tau(r-1,m))} \mathbf{E}_{\mathbf{s}_{r-1}}) - \sum_{k=\tau(r-1,m)}^{\tau(r,m)-1} \mathbf{off}^2(\hat{\mathbf{H}}^{(k)}),
\end{aligned}
$$

where $\mathbf{s}_{r-1}$ *and* $\mathbf{E}_{\mathbf{s}_{r-1}}$ *are defined by the relations* (2.21) *and* (4.16), *respectively. In the relation* (4.26) *we assumed the notation* $\hat{\mathbf{H}}^{(\tau(r-1,m))} = \hat{\mathbf{H}}_{\mathbf{P}_r}^{(\tau(r-1,m))}$.

Proof. The assertions (i) and (ii) are implied by the fact that all transformations are unitary and the Frobenius norm is invariant under unitary transformations. The assertion (iii) follows from the relation (2.8) or (2.11), which holds for any block Jacobi method for Hermitian matrices. Here, it holds for the block Jacobi method applied to $\mathbf{E}_{\mathbf{s}_{r-1}}^* \mathbf{H}^{(\tau(r-1,m))} \mathbf{E}_{\mathbf{s}_{r-1}}$ under the *bdR1* pivot strategy. We have also used the second assertion of (i). □

Let

$$
(4.27) \quad \mathbf{H}_r^{(0)} = \mathbf{E}_{\mathbf{s}_{r-1}}^* \mathbf{H}_{\mathbf{P}_r}^{(\tau(r-1,m))} \mathbf{E}_{\mathbf{s}_{r-1}}, \qquad 1 \leq r \leq m-1.
$$

In particular, we have $\mathbf{H}_1^{(0)} = \mathbf{H}_{\mathbf{P}_1}^{(1)}$. The matrix $\mathbf{H}_r^{(0)}$ is the bottom-right principal submatrix of $\mathbf{H}_{\mathbf{P}_r}^{(\tau(r-1,m))}$ of dimension $n - \mathbf{s}_{r-1} = n_r + \cdots + n_m$. Let

$$
\mathbf{H}_r^{(0)} = \mathbf{L}_r^{(0)} + \mathbf{D}_r^{(0)} + \mathbf{R}_r^{(0)}, \quad \mathbf{R}_r^{(0)} = [\mathbf{L}_r^{(0)}]^*, \qquad 1 \leq r \leq m-1,
$$

where $\mathbf{D}_r^{(0)}$ and $\mathbf{L}_r^{(0)}$ ($\mathbf{R}_r^{(0)}$) are the block-diagonal and the strictly lower- (upper-) block-triangular part of $\mathbf{H}_r^{(0)}$. Then we have

$$
(4.28) \quad \dot{\mathbf{H}}_r^{(0)} = \mathbf{L}_r^{(0)} + \mathbf{R}_r^{(0)}, \qquad 1 \leq r \leq m-1.
$$

Next, we define matrices $\mathbf{L}_r^{(k)}$, $\mathbf{R}_r^{(k)}$, $1 \leq k \leq m-r$. They carry the same block-matrix partition as $\mathbf{H}_r^{(0)}$, and they are defined as follows (cf. the relation (4.14)):

$$
(4.29) \quad \mathbf{L}_r^{(k)} = [\mathbf{U}_{r,r+k}]^* \mathcal{N}_{r,r+k}(\mathbf{L}_r^{(k-1)}) \mathbf{U}_{r,r+k}, \quad 1 \leq k \leq m-r,
$$

$$
(4.30) \quad \mathbf{R}_r^{(k)} = [\mathbf{U}_{r,r+k}]^* \mathcal{N}_{r,r+k}(\mathbf{R}_r^{(k-1)}) \mathbf{U}_{r,r+k}, \quad 1 \leq k \leq m-r.
$$

LEMMA 4.5. *For the matrices* $\mathbf{L}_r^{(k)}$, $\mathbf{R}_r^{(k)}$, $1 \leq k \leq m-r$, *defined by* (4.29), (4.30), *and* (4.28), *we have*

$$
\dot{\mathbf{H}}_r^{(k)} = \mathbf{L}_r^{(k)} + \mathbf{R}_r^{(k)}, \qquad \mathbf{R}_r^{(k)} = [\mathbf{L}_r^{(k)}]^*, \qquad 1 \leq k \leq m-r, \quad 1 \leq r \leq m-1,
$$

where

$$
\mathbf{H}_r^{(k)} = \mathbf{E}_{\mathbf{s}_{r-1}}^* \mathbf{H}^{(\tau(r-1,m)+k)} \mathbf{E}_{\mathbf{s}_{r-1}}, \quad 1 \leq k \leq m-r, \quad 1 \leq r \leq m-1.
$$

Proof. The proof uses induction with respect to k and is almost identical to the proof of Lemma 4.2. □

The structure of $\dot{\mathbf{H}}_r^{(k)}$ for the case $m = 8$, $r = 4$, $k = 3$ is shown below. The superscript ℓ stands for $\tau(r-1,m)+k = \tau(3,8)+3 = (2 \cdot 8 - 3 - 1) \cdot 3/2 + 3 = 21$, and we use the notation $\mathbf{H}^{(\ell)} = (H_{st}^{(\ell)})$, $\mathbf{L}_r^{(k)} = (L_{st}^{(\ell)})$, $\mathbf{R}_r^{(k)} = (R_{st}^{(\ell)})$:

$$\dot{\mathbf{H}}^{(\ell)} = \left[\begin{array}{c|c} H_1^{(\ell)} & H_2^{(\ell)} \\ \hline [H_2^{(\ell)}]^* & \dot{\mathbf{H}}_r^{(k)} \end{array} \right] = \left[\begin{array}{ccc|ccccc} 0 & H_{12}^{(\ell)} & H_{13}^{(\ell)} & H_{14}^{(\ell)} & H_{15}^{(\ell)} & H_{16}^{(\ell)} & H_{17}^{(\ell)} & H_{18}^{(\ell)} \\ H_{21}^{(\ell)} & 0 & H_{23}^{(\ell)} & H_{24}^{(\ell)} & H_{25}^{(\ell)} & H_{26}^{(\ell)} & H_{27}^{(\ell)} & H_{28}^{(\ell)} \\ H_{31}^{(\ell)} & H_{32}^{(\ell)} & 0 & H_{34}^{(\ell)} & H_{35}^{(\ell)} & H_{36}^{(\ell)} & H_{37}^{(\ell)} & H_{38}^{(\ell)} \\ \hline H_{41}^{(\ell)} & H_{42}^{(\ell)} & H_{43}^{(\ell)} & 0 & L_{45}^{(\ell)} & L_{46}^{(\ell)} & 0 & R_{48}^{(\ell)} \\ H_{51}^{(\ell)} & H_{52}^{(\ell)} & H_{53}^{(\ell)} & R_{54}^{(\ell)} & 0 & R_{56}^{(\ell)} & R_{57}^{(\ell)} & R_{58}^{(\ell)} \\ H_{61}^{(\ell)} & H_{62}^{(\ell)} & H_{63}^{(\ell)} & R_{64}^{(\ell)} & L_{65}^{(\ell)} & 0 & R_{67}^{(\ell)} & R_{68}^{(\ell)} \\ H_{71}^{(\ell)} & H_{72}^{(\ell)} & H_{73}^{(\ell)} & 0 & L_{75}^{(\ell)} & L_{76}^{(\ell)} & 0 & R_{78}^{(\ell)} \\ H_{81}^{(\ell)} & H_{82}^{(\ell)} & H_{83}^{(\ell)} & L_{84}^{(\ell)} & L_{85}^{(\ell)} & L_{86}^{(\ell)} & L_{87}^{(\ell)} & 0 \end{array} \right].$$

Now, we apply Lemma 3.4 to the matrices $\mathbf{L}_r^{(k)}$ and $\mathbf{R}_r^{(k)}$, $1 \le k \le m - r$. Following the lines in the proof of Lemma 4.3, we obtain that there exist real numbers (angles) $\boldsymbol{\varphi}_r \in [0, \pi/2]$ such that

(4.31) $\boldsymbol{\gamma}_1(\mathbf{R}_r^{(m-r)}) = \mathbf{Off}(\mathbf{R}_r^{(m-r)}) \sin \boldsymbol{\varphi}_r, \quad \mathbf{Off}_{\mathbf{R}}(\mathbf{R}_r^{(m-r)}) = \mathbf{Off}(\mathbf{R}_r^{(m-r)}) \cos \boldsymbol{\varphi}_r,$

(4.32) $\boldsymbol{\rho}_1(\mathbf{L}_r^{(m-r)}) = \mathbf{Off}(\mathbf{L}_r^{(m-r)}) \sin \boldsymbol{\varphi}_r, \quad \mathbf{Off}_{\mathbf{L}}(\mathbf{L}_r^{(m-r)}) = \mathbf{Off}(\mathbf{L}_r^{(m-r)}) \cos \boldsymbol{\varphi}_r,$

where

$$\sin^2 \boldsymbol{\varphi}_r \le 1 - \sigma_{\min}^2(U_{11}^{(\tau(r-1,m)+1)} U_{11}^{(\tau(r-1,m)+2)} \cdots U_{11}^{(\tau(r,m)-1)})$$

(4.33)

$$\le 1 - \prod_{k=\tau(r-1,m)+1}^{\tau(r,m)-1} \sigma_{\min}^2(U_{11}^{(k)}), \qquad 1 \le r \le m - 2.$$

If for each k and r, $\tau(r-1, m) + 1 \le k \le \tau(r-1, m) - 1$, $\varphi_r^{(k)} \in [0, \pi/2]$ is the maximal canonical angle from the Cosine-Sine decomposition of $\hat{\mathbf{U}}^{(k)}$ such that $\cos \varphi_r^{(k)} = \sigma_{\min}(U_{11}^{(k)})$, then there exist angles $\check{\varphi}_r^{(k)}$, $0 \le \check{\varphi}_r^{(k)} \le \varphi^{(k)}$ such that

$$\sin^2 \boldsymbol{\varphi}_r = 1 - \prod_{k=\tau(r-1,m)+1}^{\tau(r,m)-1} \cos^2 \check{\varphi}_r^{(k)}, \qquad 1 \le r \le m - 2.$$

For each r, $1 \le r \le m - 2$, we have

(4.34) $\boldsymbol{\rho}_1(\mathbf{L}_r^{(m-r)}) = \boldsymbol{\rho}_1(\mathbf{H}_r^{(m-r)}) = \boldsymbol{\gamma}_1(\mathbf{H}_r^{(m-r)}) = \boldsymbol{\gamma}_1(\mathbf{R}_r^{(m-r)}),$

$$\mathbf{Off}_{\mathbf{L}}(\mathbf{L}_r^{(m-r)}) = \mathbf{Off}_{\mathbf{R}}(\mathbf{R}_r^{(m-r)}) = \frac{\sqrt{2}}{2} \mathbf{Off}(\mathbf{E}_{\mathbf{s}_r}^* \mathbf{H}^{(\tau(r,m))} \mathbf{E}_{\mathbf{s}_r})$$

(4.35)

$$= \frac{\sqrt{2}}{2} \mathbf{off}(\mathbf{H}_{r+1}^{(0)}).$$

Combining the relations (4.35) and (4.32), we obtain for $1 \le r \le m - 2$,

$$\mathbf{off}^2(\mathbf{H}_{r+1}^{(0)}) = 2\mathbf{Off}_{\mathbf{L}}^2(\mathbf{L}_r^{(m-r)}) = 2\mathbf{Off}^2(\mathbf{L}_r^{(m-r)}) \cos^2 \boldsymbol{\varphi}_r = \mathbf{Off}^2(\mathbf{H}_r^{(m-r)}) \cos^2 \boldsymbol{\varphi}_r$$

(4.36)

$$= \left(\mathbf{off}^2(\mathbf{H}_r^{(0)}) - \sum_{k=1}^{m-r} \mathbf{off}^2(\hat{\mathbf{H}}_r^{(k)}) \right) \cos^2 \boldsymbol{\varphi}_r \le \mathbf{off}^2(\mathbf{H}_r^{(0)}) \cos^2 \boldsymbol{\varphi}_r.$$

Here, we have used the notation $\hat{\mathbf{H}}_r^{(1)} = \hat{\mathbf{H}}_{\mathbf{P}_r}^{(\tau(r-1,m))}$, $\hat{\mathbf{H}}_r^{(k)} = \hat{\mathbf{H}}^{(\tau(r-1,m)+k-1)}$, $2 \le k \le m - r$, $1 \le r \le m - 2$. From the relation (4.36) we obtain

(4.37) $\mathbf{off}^2(\mathbf{H}_{r+1}^{(0)}) \le \mathbf{off}^2(\mathbf{H}) \cos^2 \boldsymbol{\varphi}_1 \cos^2 \boldsymbol{\varphi}_2 \cdots \cos^2 \boldsymbol{\varphi}_r, \quad 1 \le r \le m - 2.$

In the last line we have used $\mathbf{H}_1^{(0)} = \mathbf{H}_{\mathbf{P}_1}^{(1)}$ and $\mathbf{off}(\mathbf{H}_{\mathbf{P}_1}^{(1)})) = \mathbf{off}(\mathbf{H}^{(1)}) = \mathbf{off}(\mathbf{H})$.

4.3.1. Revisiting the *bdR2* pivot strategy. Now, let us assume that the permutation matrices $\mathbf{P}_r$ are as in the relation (2.24). It means that we are considering the *bdR2* pivot strategy. When using the *bdR2* pivot strategy, all diagonal blocks of $\mathbf{H}_r^{(m-r)}$, $1 \leq r \leq m-2$, are diagonal. The same is true for the diagonal blocks of each $\mathbf{H}^{\tau(r-1,m)}$, $1 \leq r \leq m$.

This change in the definition of the permutation matrices has almost no impact on the above analysis and the obtained relations. Let us verify this. The relations (4.11), (4.13), and the Lemmas 4.2, 4.3 hold. Also, the relations (4.24), (4.25), and Lemma 4.4 hold.

The relations (4.27)–(4.30) and Lemma 4.5 hold with the *bdR2* pivot strategy. Also, the relations (4.31)–(4.35) hold. The relations (4.36) and (4.37) hold for both block pivot strategies.

One can verify that in the relations (4.26), (4.35), (4.36), (4.37) all appearances of $\mathbf{off}^2(\cdot)$ can be replaced by $\mathbf{Off}^2(\cdot)$ provided that only *bdR1* is used.

5. The main results for the Hermitian block and simple Jacobi methods. Here, we prove convergence of the block and simple Jacobi method for Hermitian matrices under the de Rijk pivot strategy. For the block method it means that we consider both the *bdR1* and *bdR2* strategy. The results automatically hold for the block and simple Jacobi method for real symmetric matrices.

First, we prove convergence to diagonal form of the block method. Then we prove global convergence of the simple Jacobi method. After that we prove global convergence of the block method. The obtained results for the off-norm reduction per cycle are so sharp that we use them to prove asymptotic quadratic convergence of the block and simple Jacobi method.

The problems with the global convergence proof of the block Kogbetliantz method are considered in the next section.

5.1. Convergence to diagonal form of the block method. The first result does not explicitly specify what kind of core algorithm is used.

PROPOSITION 5.1. *Let $\mathbf{H}$ be a Hermitian matrix of order n partitioned according to π. Let us apply the first cycle of the block Jacobi method under the de Rijk pivot strategy to $\mathbf{H}$. The method generates matrices $\mathbf{H}^{(1)}$, $\mathbf{H}_{\mathbf{P}_1}^{(1)}, \ldots, \mathbf{H}^{(M+1)}$ and $\mathbf{U}^{(0)}, \mathbf{U}^{(1)}, \ldots, \mathbf{U}^{(M)}$ which satisfy the relations* (4.5)–(4.7). *We have*

$$\mathbf{off}^2(\mathbf{H}^{(M+1)}) = \mathbf{off}^2(\mathbf{H})(1 - \cos^2 \boldsymbol{\varphi}_1 \cdots \cos^2 \boldsymbol{\varphi}_{m-2})$$

$$(5.1) \qquad \leq \mathbf{off}^2(\mathbf{H}) \left(1 - \prod_{\substack{k=1 \\ i(k) \neq j(k)-1}}^{M-1} \sigma_{\min}^2(U_{i(k)i(k)}^{(k)}) \right).$$

Here, the angles $\boldsymbol{\varphi}_k$ are defined by the relations (4.32), (4.29), (4.28), *and $U_{i(k)i(k)}^{(k)}$ is the upper-left block of $\hat{\mathbf{U}}^{(k)}$ as in the relation* (2.3).

Proof. At the end of the cycle we obtain the matrix $\mathbf{H}^{(M+1)} = (H_{st}^{(M+1)})$. For this matrix we have $H_{m-1,m}^{(M+1)} = 0$, $\dot{\mathbf{H}}^{(M+1)} = \overset{\bullet}{\mathbf{H}}{}^{(M+1)}$, and

$$\mathbf{off}^2(\mathbf{H}^{(M+1)}) = \|(\overset{\bullet}{\mathbf{H}}{}^{(M+1)}\|_F^2 = 2 \sum_{r=1}^{m-2} \boldsymbol{\rho}_r^2(\mathbf{H}^{(M+1)}),$$

where $\boldsymbol{\rho}_r(\mathbf{H}^{(M+1)}) = \boldsymbol{\rho}_1(\mathbf{E}_{\mathbf{s}_r}^* \mathbf{H}^{(M+1)} \mathbf{E}_{\mathbf{s}_r})$ and $\boldsymbol{\rho}_1(\cdot)$ is defined in (3.4). Using the first two assertions of Lemma 4.4, the relations (4.34), (4.32), and Lemma 4.5, we have

$$2\boldsymbol{\rho}_r^2(\mathbf{H}^{(M+1)}) = 2\boldsymbol{\rho}_1^2(\mathbf{H}_r^{(m-r)}) = 2\boldsymbol{\rho}_1^2(\mathbf{L}_r^{(m-r)}) = 2\mathbf{off}^2(\mathbf{L}_r^{(m-r)}) \sin^2 \boldsymbol{\varphi}_r$$

$$= \mathbf{off}^2(\mathbf{H}_r^{(m-r)}) \sin^2 \boldsymbol{\varphi}_r \leq \mathbf{off}^2(\mathbf{H}_r^{(0)}) \sin^2 \boldsymbol{\varphi}_r, \qquad 1 \leq r \leq m-2.$$

Using (4.37) we obtain

$$
\begin{aligned}
(5.2) \quad \mathbf{off}^2(\mathbf{H}^{(M+1)}) &\leq \mathbf{off}^2(\mathbf{H}) \sum_{r=1}^{m-2} \cos^2 \varphi_1 \cdots \cos^2 \varphi_{r-1} sin^2\varphi_r \\
&= \mathbf{off}^2(\mathbf{H})(1 - \cos^2 \varphi_1 \cos^2 \varphi_2 \cdots \cos^2 \varphi_{m-2}).
\end{aligned}
$$

The equality in (5.2) is easy to prove (cf. [36]). Setting $c_r^2 = \cos^2 \varphi_r$, $s_r^2 = \sin^2 \varphi_r$, $1 \leq r \leq m - 2$, we have

$$
\begin{aligned}
\sum_{r=1}^{m-2} & c_1^2 \cdots c_{r-1}^2 s_r^2 \\
&= s_1^2 + c_1^2 s_2^2 + c_1^2 c_2^2 s_3^2 + c_1^2 c_2^2 c_3^2 s_4^2 + \cdots + c_1^2 c_2^2 \cdots c_{m-3}^2 s_{m-2}^2 \\
&= 1 - c_1^2 + c_1^2(1 - c_2^2) + c_1^2 c_2^2(1 - c_3^2) + c_1^2 c_2^2 c_3^2(1 - c_4^2) + \cdots + c_1^2 c_2^2 \cdots c_{m-3}^2(1 - c_{m-2}^2) \\
&= 1 - c_1^2 c_2^2 \cdots c_{m-2}^2.
\end{aligned}
$$

The relation (4.33) implies

$$
(5.3) \quad \cos^2 \varphi_r \geq \prod_{k=\tau(r-1,m)+1}^{\tau(r,m)-1} \sigma_{\min}^2(U_{i(k)i(k)}^{(k)}), \qquad 1 \leq r \leq m - 2.
$$

The proof of the proposition follows from (5.2) and (5.3). □

Note that the condition $i(k) \neq j(k) - 1$ in the relation (5.1) means that pivot positions $(i(k), i(k))$ and $(j(k), j(k))$ are not adjacent. This fact will be used in the proof of the quadratic convergence of the method in the case of double eigenvalues.

The following theorem is a consequence of Proposition 5.1. It does not specify what kind of modified core algorithm is used.

THEOREM 5.2. *If the block Jacobi method uses the modified transformation matrices, then we have*

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

Proof. The method generates the sequences of matrices $(\mathbf{H}^{(k)}, k \geq 1)$ and $(\mathbf{U}^{(k)}, k \geq 0)$ using the relations (4.5), (4.6), (4.7). If we apply Proposition 5.1 to the matrix $\mathbf{H}^{[t]} = \mathbf{H}^{(tM+1)}$, for $t = 0, 1, 2, \ldots$, we obtain

$$
(5.4) \quad \mathbf{off}^2(\mathbf{H}^{((t+1)M+1)}) \leq \mathbf{off}^2(\mathbf{H}^{(tM+1)}) \left(1 - \prod_{\substack{k=tM+1 \\ i(k) \neq j(k)-1}}^{(t+1)M-1} \sigma_{\min}^2(U_{i(k)j(k)}^{(k)}) \right).
$$

Since the modified transformations are used, we have $\sigma_{\min}(U_{i(k)j(k)}^{(k)}) \geq C_\pi > 0$, $k \geq 1$, where C_π is defined by (4.9), (3.2). The relation (5.4) implies

$$
(5.5) \quad \mathbf{off}^2(\mathbf{H}^{[t+1]}) \leq \mathbf{off}^2(\mathbf{H}^{[t]}) \left[1 - (C_\pi^2)^{(m-1)(m-2)/2} \right], \qquad t \geq 0.
$$

In (5.5) we have used the fact that each set $\mathcal{S}_t$,

$$
\mathcal{S}_t = \{k; \ tM + 1 \leq k \leq (t+1)M, \ i(k) \neq j(k) - 1\}, \qquad t \geq 0,
$$

has exactly $(m-2) + (m-3) + \cdots + 1 = (m-1)(m-2)/2$ elements. From (5.5) we obtain

$$(5.6) \qquad \mathbf{off}^2(\mathbf{H}^{[t]}) \leq \mathbf{off}^2(\mathbf{H}^{[0]}) \left[1 - C_\pi^{(m-1)(m-2)}\right]^t, \qquad t \geq 0.$$

The relation (5.6) implies

$$(5.7) \qquad \lim_{t \to \infty} \mathbf{off}(\mathbf{H}^{[t]}) = 0.$$

The proof of the theorem is implied by the relation (5.7) and the inequalities

$$\mathbf{off}(\mathbf{H}^{[t+1]}) = \mathbf{off}(\mathbf{H}^{((t+1)M+1)}) \leq \mathbf{off}(\mathbf{H}^{(k)}) \leq \mathbf{off}(\mathbf{H}^{(tM+1)}) = \mathbf{off}(\mathbf{H}^{[t]}),$$

which hold for every $t \geq 0$ and every k satisfying $tM + 1 \leq k \leq (t+1)M$. We also have used $\mathbf{off}(\mathbf{H}^{(\tau(r-1,m)+1)}) \leq \mathbf{off}(\mathbf{H}_{\mathbf{P}_r}^{(\tau(r-1,m))}) = \mathbf{off}(\mathbf{H}^{(\tau(r-1,m))})$. $\qquad \square$

REMARK 5.3. From the relation (5.4) it is obvious that Theorem 5.2 holds even if we do not require that the unitary transformations $\mathbf{U}_{i(k)j(k)}^{(k)}$, $i(k) + 1 = j(k)$ are modified. Instead, we only need to demand for such a k that after making the pivot submatrix $\hat{\mathbf{H}}^{(k)}$ diagonal, the diagonal elements are ordered non-increasingly.

5.2. Global convergence of the simple Jacobi method. An immediate corollary of Theorem 5.2 is obtained for the case $m = n$. Then we have $\pi = (1, 1, \ldots, 1)$. Let us recall one step of the simple Jacobi method for Hermitian matrices. Since the initial transformation (4.2) is not needed, we set $\mathbf{H}^{(1)} = \mathbf{H}$. In step k, the method computes the unitary plane matrix (complex Jacobi rotation) $\mathbf{U}^{(k)} = \mathcal{E}(i, j; \hat{\mathbf{U}}^{(k)})$,

$$\hat{\mathbf{U}}^{(k)} = \begin{bmatrix} \cos\phi_k & -e^{\imath\alpha_k}\sin\phi_k \\ e^{-\imath\alpha_k}\sin\phi_k & \cos\phi_k \end{bmatrix}.$$

The Jacobi angle ϕ_k and the phase α_k are computed using the formulas

$$(5.8) \qquad \tan(2\phi_k) = \frac{2|h_{ij}^{(k)}|}{h_{ii}^{(k)} - h_{jj}^{(k)}}, \qquad -\frac{\pi}{4} \leq \phi_k \leq \frac{\pi}{4}, \qquad \alpha_k = \arg(h_{ij}^{(k)}),$$

where $\mathbf{H}^{(k)} = (h_{rs}^{(k)})$. The diagonal elements $h_{ii}^{(k)}$, $h_{jj}^{(k)}$ are updated using

$$(5.9) \qquad h_{ii}^{(k+1)} = h_{ii}^{(k)} + |h_{ij}^{(k)}|\tan\phi_k, \qquad h_{jj}^{(k+1)} = h_{jj}^{(k)} - |h_{ij}^{(k)}|\tan\phi_k.$$

The off-norm is updated by

$$\mathbf{off}^2(\mathbf{H}^{(k+1)}) = \mathbf{off}^2(\mathbf{H}^{(k)}) - 2|h_{ij}^{(k)}|^2, \qquad k \geq 1.$$

In the case of a real symmetric $\mathbf{H}$ and the real Jacobi method, one has to set $\alpha_k = 0$ in the relation (5.8) and $2|h_{ij}^{(k)}|$ has to be replaced by $2h_{ij}^{(k)}$. Also, in (5.9), the term $|a_{ij}^{(k)}|\tan\phi_k$ has to be replaced by $h_{ij}^{(k)}\tan\phi_k$.

THEOREM 5.4. *Let* $\mathbf{H}$ *be a Hermitian matrix of order* n. *Let us apply to* $\mathbf{H}$ *the simple Jacobi method under the de Rijk pivot strategy. Then we have*

$$\lim_{k \to \infty} \mathbf{H}^{(k)} = \mathbf{\Lambda},$$

where $\mathbf{\Lambda}$ *is the diagonal matrix of the eigenvalues of* $\mathbf{H}$ *ordered non-increasingly. Furthermore, we have*

$$(5.10) \qquad \mathbf{off}(\mathbf{H}^{((t+1)N+1)}) \leq \mathbf{off}(\mathbf{H}^{(tN+1)}) \left[1 - \prod_{\substack{k=tN+1 \\ i(k) \neq j(k)-1}}^{(t+1)N-1} \cos^2 \phi_k \right]^{1/2}, \qquad t \geq 0,$$

where ϕ_k *are the Jacobi angles. Both assertions also hold for the case of a real symmetric matrix* $\mathbf{H}$ *and the real Jacobi method.*

Proof. Since

$$\sigma_{\min}(U_{11}^{(k)}) = |\cos \phi_k| = \cos \phi_k \geq \cos\left(\frac{\pi}{4}\right) = \frac{\sqrt{2}}{2}, \qquad k \geq 1,$$

all $\mathbf{U}^{(k)}$ are modified transformation matrices. Applying Theorem 5.2, we obtain $\lim_{k \to \infty} \mathbf{off}(\mathbf{H}^{(k)}) = 0$. The estimate (5.10) follows directly from the relation (5.4). Hence, it remains to show that the diagonal elements $h_{11}^{(k)}, h_{22}^{(k)}, \dots, h_{nn}^{(k)}$ converge to the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ such that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$.

Suppose the process has reached the stage in which

$$(5.11) \qquad \mathbf{off}(\mathbf{H}^{(k)}) < \delta, \qquad k \geq k_0,$$

where δ is from (3.14). Let us assume that the eigenvalues are numbered so that (3.11) holds. By the assertion (3.16) of Lemma 3.5, we conclude that each diagonal element of $\mathbf{H}^{(k)}$, $k \geq k_0$, has to be at most 0.22δ away from an eigenvalue of $\mathbf{H}$. Let us define intervals $\mathcal{D}_r$,

$$\mathcal{D}_r = \{z \in \mathbf{R}; |z - \lambda_r| < 0.22\delta\}, \qquad 1 \leq r \leq \tilde{p},$$

where $\tilde{p}$ is from the natural partition of n. These intervals are disjoint, and the minimum distance between two adjacent intervals equals $3\delta - 2 \cdot 0.22\delta = 2.56\delta$. From Lemma 3.5 we conclude that each interval $\mathcal{D}_r$ contains exactly ν_r diagonal elements, where ν_r is defined by (3.12).

Now, let us consider one cycle or sweep of the Jacobi method starting with the matrix $\mathbf{H}^{(tN+1)}$ such that $tN + 1 \geq k_0$. In each Jacobi step within that cycle, the diagonal elements cannot change the intervals in which they are located. This is seen from the relation (5.9). In particular, we see that the affected diagonal elements $h_{ii}^{(k)}$ and $h_{jj}^{(k)}$ can change at most by the amount $|h_{ij}^{(k)} \tan \phi_k| \leq |h_{ij}^{(k)}| \leq \mathbf{off}(\mathbf{H}^{(k)})/\sqrt{2} < \delta/\sqrt{2}$, which is much smaller than 2.56δ.

To simplify notation we can assume $k_0 = 1$. Then we have $t = 0$, and we consider the first cycle. Our aim is to show that

$$(5.12) \qquad h_{\varsigma_{r-1}+1, \varsigma_{r-1}+1}^{(N+1)}, \dots, h_{\varsigma_r, \varsigma_r}^{(N+1)} \in \mathcal{D}_r, \qquad 1 \leq r \leq \tilde{p}.$$

While the Jacobi steps cannot change the affiliation of the diagonal elements, the permutational transformation can change the ordering of the diagonal elements along the diagonal.

Let us consider the first ν_1 permutational transformations. The first permutation brings the largest diagonal element to position $(1, 1)$. All later permutational transformations cannot

change the affiliation of that diagonal element to λ_1 because they do not affect it. Hence, $h_{11}^{(N+1)} \in \mathcal{D}_1$.

The second permutational transformation brings the largest diagonal element of $\mathbf{E}_1^* \mathbf{H}^{(n)} \mathbf{E}_1$ to the second diagonal position in the matrix $\mathbf{H}_{\mathbf{P}_2}^{(n)}$. Now, the same argument as earlier applies to it. Finally, the last of the first ν_1 permutational transformations brings the largest diagonal element of the matrix $\mathbf{E}_{\nu_1-1}^* \mathbf{H}^{(N-s_{n-\nu_1}+1)} \mathbf{E}_{\nu_1-1}$ to the (ν_1, ν_1)-position in $\mathbf{H}_{\mathbf{P}_{\nu_1}}^{(N-s_{n-\nu_1}+1)}$. Note that $N - s_{n-\nu_1} = n - 1 + (n - 2) + \cdots + (n - (\nu_1 - 1))$. Now, $\mathcal{D}_1$ contains ν_1 diagonal elements of $\mathbf{H}_{\mathbf{P}_{\nu_1}}^{(N-s_{n-\nu_1})}$, and they are located in the first ν_1 diagonal positions. The interval $\mathcal{D}_1$ cannot contain more diagonal elements, so it is now full. All later permutational transformations in the first cycle do not affect the first ν_1 diagonal elements. This proves (5.12) for $r = 1$.

The same argument applies to the next ν_2 permutational transformations. For $\ell = 1, \ldots, \nu_2$, the permutation $\mathbf{P}_{\nu_1+\ell}$ brings the largest diagonal element of $\mathbf{E}_{\nu_1+\ell-1}^* \mathbf{H}^{(N-s_{n-\nu_1-\ell+1})} \mathbf{E}_{\nu_1+\ell-1}$, to the diagonal position $\nu_1 + \ell$. Continuing this process we obtain the relation (5.12).

If $\mathbf{H}$ is real symmetric, then the real Jacobi method is applied. It is easy to verify that all statements and arguments in the proof remain to hold. □

From the proof of Theorem 5.4 we see that the simple Jacobi method for the Hermitian (symmetric) matrix under the de Rijk strategy is well-behaved. Indeed, for a sufficiently large k, neither the Jacobi step nor the permutational transformation can change the affiliation of the diagonal elements.

5.3. Quadratic convergence of the simple method. Theorem 5.4 implies that the simple Jacobi method under the de Rijk pivot strategy is *quadratically convergent*. This is implied by the fact that, eventually, the process essentially reduces to the row-cyclic Jacobi method.

In the case of simple eigenvalues, the proof can use the technique from [36], and it takes just a few lines of text. Let $\mathbf{\Lambda}^{(k)}$, $k \geq k_0$, denote the diagonal matrix of the eigenvalues of $\mathbf{H}$ ordered so that the diagonal elements of $\mathbf{H}^{(k)}$ are affiliated to the corresponding diagonal elements of $\mathbf{\Lambda}^{(k)}$ (more details on $\mathbf{\Lambda}^{(k)}$ are given in the next section). Here k_0 is such that (5.11) holds. We consider the stage of the process when $k \geq tN + 1 \geq k_1 \geq k_0$, and k_1 is such that the diagonal elements of $\mathbf{\Lambda}^{(k_1)}$ are ordered non-increasingly.

Using a variant of Bernoulli's inequality we have

$$\prod_k \cos^2 \phi_k = \prod_k (1 - \sin^2 \phi_k) \geq 1 - \sum_k \sin^2 \phi_k.$$

This implies

$$1 - \prod_k \cos^2 \phi_k \leq \sum_k \sin^2 \phi_k.$$

Hence, from the relations (5.10) and (5.8) we obtain

$$\mathbf{off}^2(\mathbf{H}^{([t+1])}) \leq \mathbf{off}^2(\mathbf{H}^{[t]}) \sum_{\substack{k=tN+1 \\ i(k)\neq j(k)-1}}^{(t+1)N-1} \sin^2 \phi_k \leq 0.0763 \frac{\mathbf{off}^4(\mathbf{H}^{[t]})}{\delta^2}$$

or

$$(5.13) \qquad \mathbf{off}(\mathbf{H}^{[t+1]}) \leq 0.2763 \frac{\mathbf{off}^2(\mathbf{H}^{[t]})}{\delta}, \qquad t \geq 0.$$

Here we have used

$$|\sin\phi_k| \le |\tan\phi_k| \le \frac{1}{2}|\tan(2\phi_k)| = \frac{|h_{i(k)j(k)}^{(k)}|}{|h_{i(k)i(k)}^{(k)} - h_{j(k)j(k)}^{(k)}|} \le \frac{|h_{i(k)j(k)}^{(k)}|}{2.56\delta},$$

and

$$\sum_{\substack{k=tN+1 \\ i(k)\ne j(k)-1}}^{(t+1)N-1} |h_{i(k)j(k)}^{(k)}|^2 \le \sum_{k=tN+1}^{(t+1)N} |h_{i(k)j(k)}^{(k)}|^2 = \frac{1}{2}\left(\mathbf{off}^2(\mathbf{H}^{[t]}) - \mathbf{off}^2(\mathbf{H}^{([t+1]})\right)$$

$$\le \frac{1}{2}\mathbf{off}^2(\mathbf{H}^{[t]}).$$

Since the diagonal elements of each $\mathbf{\Lambda}^{(k)}$, $tN \le k \le (t+1)N$, are ordered non-increasingly and the angles ϕ_k, $i(k) = j(k) - 1$, do not appear in the estimate, the presented proof holds in the case of simple and double eigenvalues. In this way we have proved the following theorem:

THEOREM 5.5. *If the eigenvalues of the initial Hermitian matrix* $\mathbf{H}$ *are simple or double, then the Jacobi method under the de Rijk pivot strategy is quadratically convergent, and the relation* (5.13) *holds.*

In the general case when multiple eigenvalues are present, we are confident that the quadratic convergence of the method can be proved by following the lines of the proof of [22, Theorem 2.8(i)]. The proof will use the fact that the permutations $\mathbf{P}^{(k)}$ can swap only the diagonal elements that lie within the same diagonal block in the natural partition of the current matrix. The proof will be much longer and more complicated than the one presented here for the special cases of simple and double eigenvalues.

5.4. Global convergence of the block Jacobi method. In Theorem 5.2 we have shown that the block Jacobi method which uses the modified core algorithm (with or without attribute R) converges to diagonal form under the de Rijk pivot strategy. To simplify terminology we will simply say that the *modified block Jacobi method* converges to diagonal form. The results proved here use ideas and notation from [28, Section 4]. For the sake of completeness of the presentation, we will provide all the necessary details.

Combining the convergence to diagonal form with Lemma 3.5, we conclude that for large enough k, say, for $k \ge k_0$, the diagonal elements of $\mathbf{H}^{(k)}$ are affiliated to the eigenvalues of $\mathbf{H}$. By this we mean that the condition (3.15) holds for each $\mathbf{H}^{(k)}$, $k \ge k_0$. Note that k_0 depends on $\mathbf{H}$, on the block pivot strategy (*bdR1* or *bdR2*), and on the modified core algorithm that is part of the block Jacobi method.

For $k \ge k_0$, let $\mathbf{\Lambda}^{(k)}$ denote a diagonal matrix of the eigenvalues of $\mathbf{H}$ such that

$$(5.14) \qquad\qquad \lim_{k\to\infty}(\mathbf{H}^{(k)} - \mathbf{\Lambda}^{(k)}) = 0.$$

Here, each diagonal entry of $\mathbf{\Lambda}^{(k)}$ is an eigenvalue of $\mathbf{H}$ that is closest to the corresponding diagonal entry of $\mathbf{H}^{(k)}$. The modified block Jacobi method is globally convergent if and only if for every initial Hermitian matrix $\mathbf{H}$, the associated sequence $(\mathbf{\Lambda}^{(k)}, k \ge k_0)$ is convergent. So, we have to find conditions under which the sequence $(\mathbf{\Lambda}^{(k)}, k \ge k_0)$ becomes stationary.

From the relation (5.14) we see that it makes sense to transfer some terminology and notation from $\mathbf{H}^{(k)}$ to $\mathbf{\Lambda}^{(k)}$. So, we call $\hat{\mathbf{\Lambda}}^{(k)}$ a pivot submatrix of $\mathbf{\Lambda}^{(k)}$ if it is extracted from $\mathbf{\Lambda}^{(k)}$ in the same way as $\hat{\mathbf{H}}^{(k)}$ is extracted from $\mathbf{H}^{(k)}$. Then $\hat{\mathbf{\Lambda}}^{(k)} = \mathrm{diag}(\Lambda_{i(k)i(k)}^{(k)}, \Lambda_{j(k)j(k)}^{(k)})$, and the diagonal elements of $H_{i(k)i(k)}^{(k)}$ ($H_{j(k)j(k)}^{(k)}$) are affiliated to the diagonal elements

of $\Lambda_{i(k)i(k)}^{(k)}$ $(\Lambda_{j(k)j(k)}^{(k)})$. Since the diagonal elements of $H_{i(k)i(k)}^{(k)}$ $(H_{j(k)j(k)}^{(k)})$ are ordered non-increasingly, we conclude that the same is true for the diagonal elements of $\Lambda_{i(k)i(k)}^{(k)}$ $(\Lambda_{j(k)j(k)}^{(k)})$.

In Theorem 5.2 we have used the modified core algorithm, so we can write

$$(5.15) \qquad \hat{\mathbf{U}}^{(k)} = \hat{\tilde{\mathbf{U}}}^{(k)}\hat{\tilde{\mathbf{P}}}^{(k)}, \qquad \hat{\tilde{\mathbf{P}}}^{(k)} \in \{\hat{\mathbf{P}}^{(k)}\hat{\mathbf{Q}}^{(k)}, \hat{\mathbf{P}}_0^{(k)}\hat{\mathbf{Q}}^{(k)}\}, \qquad k \geq 1,$$

and consequently we have $\mathbf{U}^{(k)} = \tilde{\mathbf{U}}^{(k)}\tilde{\mathbf{P}}^{(k)}$, $k \geq 1$. Here $\hat{\tilde{\mathbf{U}}}^{(k)}$ and $\hat{\mathbf{U}}^{(k)}$ are computed by the basic and modified core algorithm, respectively.

Let

$$(5.16) \qquad\qquad \tilde{\mathbf{H}}^{(k+1)} = [\tilde{\mathbf{U}}^{(k)}]^*\mathbf{H}^{(k)}\tilde{\mathbf{U}}^{(k)}, \qquad k \geq 1.$$

Then we have $\mathbf{H}^{(k+1)} = [\tilde{\mathbf{P}}^{(k)}]^*\tilde{\mathbf{H}}^{(k+1)}\tilde{\mathbf{P}}^{(k)}$ and $\mathbf{off}(\tilde{\mathbf{H}}^{(k+1)}) = \mathbf{off}(\mathbf{H}^{(k+1)})$ for $k \geq 1$.

LEMMA 5.6. *Let* $\mathbf{H}$ *be a Hermitian matrix of order* n*, and let the sequences* $(\mathbf{H}^{(k)}, k \geq 1)$ *and* $(\mathbf{U}^{(k)}, k \geq 0)$ *be obtained by applying the modified block Jacobi method to* $\mathbf{H}$ *under the de Rijk pivot strategy. If the basic core algorithm is well-behaved, then the assertions (i)–(iii) hold:*

(i) $\displaystyle\lim_{k\to\infty}\left(\hat{\tilde{\mathbf{U}}}^{(k)}\hat{\mathbf{\Lambda}}^{(k)} - \hat{\mathbf{\Lambda}}^{(k)}\hat{\tilde{\mathbf{U}}}^{(k)}\right) = 0;$

(ii) if k *is sufficiently large and* $\sigma(\Lambda_{i(k)i(k)}^{(k)}) \cap \sigma(\Lambda_{j(k)j(k)}^{(k)}) = \emptyset$*, then*
$\hat{\mathbf{P}}^{(k)} = \mathrm{diag}(\mathbf{P}_{i(k)i(k)}^{(k)}, I_{n_{j(k)}});$

(iii) if the eigenvalues of $\mathbf{H}$ *are simple, then for a sufficiently large* k *we have*

$$\lim_{k\to\infty}\left(|\hat{\tilde{\mathbf{U}}}^{(k)}| - \mathbf{I}_{n_{i(k)}+n_{j(k)}}\right) = 0.$$

Here $\hat{\tilde{\mathbf{U}}}^{(k)}$ *and* $\hat{\mathbf{P}}^{(k)}$ *are from the relation* (5.15).

Proof. (i) This assertion follows directly from the relations (5.14) and (4.10). Since the basic core algorithm is well-behaved, we have replaced $\hat{\mathbf{U}}^{(k)}$ in the relation (4.10) by $\hat{\tilde{\mathbf{U}}}^{(k)}$.

(ii) The assertion (i) can be written in the form

$$(5.17) \qquad\qquad \hat{\tilde{\mathbf{U}}}^{(k)}\hat{\mathbf{\Lambda}}^{(k)} - \hat{\mathbf{\Lambda}}^{(k)}\hat{\tilde{\mathbf{U}}}^{(k)} = \hat{\mathbf{E}}^{(k)}, \qquad \lim_{k\to\infty}\hat{\mathbf{E}}^{(k)} = 0.$$

Let $\hat{\tilde{\mathbf{U}}}^{(k)} = (\hat{\tilde{u}}_{st}^{(k)})$, $\hat{\mathbf{E}}^{(k)} = (\hat{e}_{st}^{(k)})$, $\hat{\mathbf{\Lambda}}^{(k)} = \mathrm{diag}(\hat{\boldsymbol{\lambda}}_1^{(k)}, \ldots, \hat{\boldsymbol{\lambda}}_{n_{i(k)}+n_{j(k)}}^{(k)})$. From (5.17) we conclude that there is an integer k_1, $k_1 \geq k_0$, such that

$$(5.18) \qquad |\hat{\tilde{u}}_{st}^{(k)}| = \frac{|\hat{e}_{st}^{(k)}|}{|\hat{\boldsymbol{\lambda}}_s^{(k)} - \hat{\boldsymbol{\lambda}}_t^{(k)}|} \leq \frac{|\hat{e}_{st}^{(k)}|}{3\delta}, \qquad \hat{\boldsymbol{\lambda}}_s^{(k)} \neq \hat{\boldsymbol{\lambda}}_t^{(k)}, \qquad k \geq k_1.$$

Here, 3δ is the minimum gap in $\sigma(\mathbf{H})$. By the assumption in (ii), we obtain from (5.18)

$$(5.19) \qquad \left[\|\tilde{U}_{i(k)j(k)}^{(k)}\|_F^2 + \|\tilde{U}_{j(k)i(k)}^{(k)}\|_F^2\right]^{1/2} \leq \frac{\|\hat{\mathbf{E}}^{(k)}\|_F}{3\delta}, \qquad k \geq k_1.$$

From (5.19) we see that for a sufficiently large k, we can make the off-diagonal blocks of the unitary matrix $\hat{\tilde{\mathbf{U}}}^{(k)}$ arbitrary close to the appropriate zero matrices. Therefore, the diagonal blocks $\tilde{U}_{i(k)i(k)}^{(k)}$ and $\tilde{U}_{j(k)j(k)}^{(k)}$ approach some unitary matrices of order $n_{i(k)}$ and $n_{j(k)}$, respectively. Hence, when the QR factorization algorithm with column pivoting is applied to

$[\tilde{U}_{i(k)i(k)}^{(k)} \; \tilde{U}_{i(k)j(k)}^{(k)}]$, it delivers the permutation matrix $\hat{\mathbf{P}}^{(k)} = \mathbf{I}_{11'}\mathbf{I}_{22'} \cdots \mathbf{I}_{n_{i(k)}n'_{i(k)}}$ in which all second subscripts $1', 2', \ldots, n'_{i(k)}$ are from the set $\{1, 2, \ldots, n_{i(k)}\}$. This proves (ii).

(iii) If the eigenvalues of $\mathbf{H}$ are simple, then the relation (5.18) implies that the off-diagonal elements of $\hat{\tilde{\mathbf{U}}}^{(k)}$ tend to zero as $k \to \infty$. This proves (iii) and the lemma. □

As a final note, the different block pivot strategies *bdR1* and *bdR2* generate different permutation matrices $\mathbf{P}^{(k)}$, but the proof of Lemma 5.6 stays the same.

From the proof of Lemma 5.6, we see that $\mathbf{P}^{(k)}$ can differ from $\mathbf{I}_n$ for an arbitrary large k. Therefore, we cannot prove convergence of $(\mathbf{\Lambda}^{(k)}, \; k \geq k_0)$ without additional assumptions. However, we can make use of the property R of the modified algorithm.

THEOREM 5.7. *The modified block Jacobi method is globally convergent under the de Rijk pivot strategy provided the basic core algorithm is well-behaved, the modified core algorithm has the property* R*, and one of the following conditions holds:*

(a) the eigenvalues of the initial matrix $\mathbf{H}$ are simple,

(b) for a sufficiently large k, the diagonal elements of $\mathbf{H}^{(k)}$ affiliated with the same eigenvalue lie in one diagonal block of the current block-matrix partition.

Proof. As the block method uses the modified core algorithm, the iteration matrix $\mathbf{H}^{(k)}$ converges to diagonal form. Using the assertions (ii) and (iii) of Lemma 5.6 and the assumption that the modified core algorithm has the property R, we conclude that both conditions (a) and (b) imply that for a sufficiently large k, we have $\hat{\mathbf{P}}_0^{(k)} = \mathbf{I}_{n_{i(k)}+n_{j(k)}}$, and therefore we have $\hat{\mathbf{U}}^{(k)} = \hat{\tilde{\mathbf{U}}}^{(k)}\hat{\mathbf{Q}}^{(k)}$.

In case (a) we have $\hat{\mathbf{Q}}^{(k)} = \mathbf{I}_{n_{i(k)}+n_{j(k)}}$. Indeed, since the basic core algorithm is well-behaved, the diagonal elements of both $\tilde{\mathbf{H}}^{(k+1)}$ and $\mathbf{H}^{(k)}$ are affiliated to the same eigenvalues. We know that the diagonal elements of each diagonal block of $\tilde{\mathbf{H}}^{(k+1)}$ are in non-increasing order. The condition (a) implies that they are in descending order. The same must be true for the diagonal elements of $\mathbf{H}^{(k)}$. Therefore, the permutation matrix $\mathbf{Q}^{(k)}$ has to be the identity.

From the relation (5.18) we conclude that $\hat{\tilde{\mathbf{U}}}^{(k)}$ tends to diagonal form. Hence, we obtain

$$\lim_{k\to\infty}(|\mathbf{U}^{(k)}| - \mathbf{I}_n) = \lim_{k\to\infty}(|\tilde{\mathbf{U}}^{(k)}| - \mathbf{I}_n) = 0.$$

Consequently, for a large enough k, $\|\operatorname{diag}(\mathbf{H}^{(k+1)}) - \operatorname{diag}(\mathbf{H}^{(k)})\|_F$ is arbitrarily small, and the diagonal elements of $\mathbf{H}^{(k)}$ cannot change their affiliation to the eigenvalues of $\mathbf{H}$.

In case (b) the proof is similar although a bit more complicated. Recall that the diagonal elements of $\Lambda_{i(k)i(k)}^{(k)}$ $(\Lambda_{j(k)j(k)}^{(k)})$ are ordered non-increasingly. Let $\tau_{i(k)}$ $(\tau_{j(k)})$ be the partition of $n_{i(k)}$ $(n_{j(k)})$ determined by the diagonal elements of $\Lambda_{i(k)i(k)}^{(k)}$ $(\Lambda_{j(k)j(k)}^{(k)})$ (cf. (3.11), (3.12)), and let $\hat{\tau}^{(k)} = (\tau_{i(k)}, \tau_{j(k)})$ be the partition of $n_{i(k)} + n_{j(k)}$. The partition $\hat{\tau}^{(k)}$ is linked to the sequence $\hat{\boldsymbol{\lambda}}_1^{(k)}, \ldots, \hat{\boldsymbol{\lambda}}_{n_{i(k)}+n_{j(k)}}^{(k)}$ of the diagonal elements of $\hat{\mathbf{\Lambda}}^{(k)}$. The assumption in (b) implies that $\hat{\tau}^{(k)}$ is a subpartition of the partition $(n_{i(k)}, n_{j(k)})$ of $\hat{\mathbf{H}}^{(k)}$. We can call the block-matrix partition defined by $\hat{\tau}^{(k)}$ the *local natural partition* of $\hat{\mathbf{\Lambda}}^{(k)}$ and of $\hat{\mathbf{H}}^{(k)}$.

Let us partition $\hat{\tilde{\mathbf{U}}}^{(k)}$, $\hat{\mathbf{Q}}^{(k)}$, and $\hat{\mathbf{U}}^{(k)}$ using the local natural partition of $\hat{\mathbf{H}}^{(k)}$. The condition (b) and the relation (5.18) imply that $\hat{\tilde{\mathbf{U}}}^{(k)}$ tends to block-diagonal form. Because the basic core algorithm is well-behaved, condition (b) implies that for a large enough k, the matrix $\hat{\mathbf{Q}}^{(k)}$ is block diagonal. Hence, for a large enough k, $\hat{\mathbf{U}}^{(k)}$ is arbitrary close to the block-diagonal form determined by $\hat{\tau}^{(k)}$. Let us prove

$$(5.20) \qquad \lim_{k\to\infty}\left([\hat{\mathbf{U}}^{(k)}]^*\hat{\mathbf{H}}^{(k)}\hat{\mathbf{U}}^{(k)} - \hat{\mathbf{H}}^{(k)}\right) = 0.$$

From the relation (5.14) we have

(5.21) $$\hat{\mathbf{H}}^{(k)} = \hat{\mathbf{\Lambda}}^{(k)} + \hat{\mathbf{\Xi}}^{(k)}, \qquad \lim_{k \to \infty} \hat{\mathbf{\Xi}}^{(k)} = 0.$$

Using (5.21), for a large enough k, we have

$$\|[\hat{\mathbf{U}}^{(k)}]^* \hat{\mathbf{H}}^{(k)} \hat{\mathbf{U}}^{(k)} - \hat{\mathbf{H}}^{(k)}\|_F \leq \|[\hat{\mathbf{U}}^{(k)}]^* \hat{\mathbf{\Lambda}}^{(k)} \hat{\mathbf{U}}^{(k)} - \hat{\mathbf{\Lambda}}^{(k)}\|_F + 2\|\hat{\mathbf{\Xi}}^{(k)}\|_F.$$

The matrices $\hat{\mathbf{\Lambda}}^{(k)}$ and $\hat{\mathbf{U}}^{(k)}$ carry the same local natural partition, $\hat{\mathbf{\Lambda}}^{(k)}$ is diagonal, while $\hat{\mathbf{U}}^{(k)}$ tends to block-diagonal form. This implies

$$\lim_{k \to \infty} \left([\hat{\mathbf{U}}^{(k)}]^* \hat{\mathbf{\Lambda}}^{(k)} \hat{\mathbf{U}}^{(k)} - \hat{\mathbf{\Lambda}}^{(k)} \right) = 0.$$

The relation (5.20) is implied from the last three relations.

Since $[\hat{\mathbf{U}}^{(k)}]^* \hat{\mathbf{H}}^{(k)} \hat{\mathbf{U}}^{(k)}$ is diagonal, (5.20) implies

$$\| \operatorname{diag}(\mathbf{H}^{(k+1)}) - \operatorname{diag}(\mathbf{H}^{(k)})\|_F = \|[\hat{\mathbf{U}}^{(k)}]^* \hat{\mathbf{H}}^{(k)} \hat{\mathbf{U}}^{(k)} - \operatorname{diag}(\hat{\mathbf{H}}^{(k)})\|_F$$
$$\leq \|[\hat{\mathbf{U}}^{(k)}]^* \hat{\mathbf{H}}^{(k)} \hat{\mathbf{U}}^{(k)} - \hat{\mathbf{H}}^{(k)}\|_F + \|\hat{\mathbf{H}}^{(k)} - \operatorname{diag}(\hat{\mathbf{H}}^{(k)})\|_F \; \to \; 0 \qquad \text{as } k \to \infty.$$

As in the case (a) we conclude that for a large enough k, the diagonal elements of $\mathbf{H}^{(k)}$ cannot change their affiliation to the eigenvalues of $\mathbf{H}$. □

Note that condition (a) is contained in (b). We separated it because it is the most common case in practice. In the general case, when the diagonal elements of $\mathbf{H}^{(k)}$ affiliated with the same eigenvalue remain scattered across several blocks, we have to add an additional property to the modified block Jacobi method (see [28, Definition 4.2]).

DEFINITION 5.8. *The block Jacobi method has a property* D *if it is convergent to diagonal form and for a sufficiently large k the diagonal elements of the associated matrices $\mathbf{\Lambda}^{(k)}$ are in non-increasing ordering.*

For the modified block Jacobi method, the first condition of Definition 5.8 is fulfilled. It is an open question how to most efficiently implement the property D in the block Jacobi method. Obviously, for a large enough t it will suffice to apply just one permutational similarity transformation on $\mathbf{H}^{[t]}$ which orders the diagonal elements non-increasingly.

THEOREM 5.9. *The modified block Jacobi method is globally convergent under the de Rijk pivot strategy provided that it has the property* D*, the basic core algorithm is well-behaved, and the modified core algorithm has the property* R.

Proof. Our goal is to prove that $(\mathbf{\Lambda}^{(k)}, \; k \geq k_0)$ is a stationary sequence. Theorem 5.7 implies that it remains to consider the case when the initial matrix $\mathbf{H}$ has at least one multiple eigenvalue which fails to satisfy condition (b).

Since the basic core algorithm is well-behaved, from the relation (5.16) we conclude that the matrix $\mathbf{\Lambda}^{(k)}$, $k > k_0$, is associated with both $\mathbf{H}^{(k)}$ and $\tilde{\mathbf{H}}^{(k+1)}$.

Let $k_0' = \max\{k_0, k'\}$, where k' is such that each $\mathbf{\Lambda}^{(k)}$, $k > k'$, has non-increasing diagonal elements (see Definition 5.8). We consider step k of the block Jacobi method such that k satisfies $k \geq k_1 \geq k_0'$ and where k_1 will be addressed later.

To simplify notation we will omit the superscript k, and in the following analysis we will denote the matrices (see Section 4.1)

$$\hat{\mathbf{H}}^{(k)}, \hat{\mathbf{\Lambda}}^{(k)} = \operatorname{diag}(\hat{\boldsymbol{\lambda}}_1^{(k)}, \dots, \hat{\boldsymbol{\lambda}}_{n_i+n_j}^{(k)}), \; \hat{\mathbf{U}}^{(k)}, \hat{\mathbf{P}}^{(k)}, \hat{\mathbf{P}}_0^{(k)}, \hat{\mathbf{Q}}^{(k)}, \hat{\tilde{\mathbf{U}}}^{(k)}, \hat{\tilde{\mathbf{H}}}^{(k+1)}, \mathbf{H}^{(k+1)}$$

by

$$\hat{\mathbf{H}}, \;\; \hat{\mathbf{\Lambda}} = \operatorname{diag}(\hat{\boldsymbol{\lambda}}_1, \dots, \hat{\boldsymbol{\lambda}}_{n_i+n_j}), \;\;\; \hat{\mathbf{U}}, \;\;\; \hat{\mathbf{P}}, \;\;\; \hat{\mathbf{P}}_0, \;\;\; \hat{\mathbf{Q}}, \;\;\; \hat{\tilde{\mathbf{U}}}, \;\;\; \hat{\tilde{\mathbf{H}}}', \;\;\; \mathbf{H}',$$

respectively. Here, $\hat{\tilde{\mathbf{H}}}' = \hat{\tilde{\mathbf{U}}}^* \hat{\mathbf{H}} \hat{\tilde{\mathbf{U}}}$. We also denote $n_{i(k)}$ $(n_{j(k)})$ by n_i (n_j) and $\hat{\mathbf{\Lambda}} = \mathrm{diag}(\Lambda_{ii}, \Lambda_{jj})$.

Let $\hat{\mathbf{\Lambda}}' = \mathrm{diag}(\hat{\boldsymbol{\lambda}}_1', \ldots, \hat{\boldsymbol{\lambda}}_{n_i+n_j}')$ be associated with $\hat{\mathbf{H}}' = \hat{\mathbf{U}}^* \hat{\mathbf{H}} \hat{\mathbf{U}}$. Then $\hat{\mathbf{\Lambda}}' = [\hat{\mathbf{P}}_0 \hat{\mathbf{Q}}]^* \hat{\mathbf{\Lambda}} \hat{\mathbf{P}}_0 \hat{\mathbf{Q}}$, and our aim is to show $\hat{\mathbf{\Lambda}}' = \hat{\mathbf{\Lambda}}$. Obviously, $\hat{\mathbf{\Lambda}}' = \hat{\mathbf{\Lambda}}$ implies $\mathbf{\Lambda}' = \mathbf{\Lambda}$. Since the modified block Jacobi method has the property D, the diagonal elements of $\mathbf{\Lambda}$ are in non-increasing ordering. Thus, $\hat{\boldsymbol{\lambda}}_1 \geq \cdots \geq \hat{\boldsymbol{\lambda}}_{n_i+n_j}$. Since we consider the case $\hat{\boldsymbol{\lambda}}_{n_i} = \hat{\boldsymbol{\lambda}}_{n_i+1}$, we have four possibilities:

(a) $\hat{\mathbf{\Lambda}} = \hat{\boldsymbol{\lambda}}_1 \mathbf{I}_{n_i+n_j}$,

(b) $\hat{\boldsymbol{\lambda}}_1 = \cdots = \hat{\boldsymbol{\lambda}}_r > \hat{\boldsymbol{\lambda}}_{r+1} \geq \cdots \geq \hat{\boldsymbol{\lambda}}_{n_i+n_j}$, $n_i < r < n_i + n_j$,

(c) $\hat{\boldsymbol{\lambda}}_1 \geq \cdots \geq \hat{\boldsymbol{\lambda}}_r > \hat{\boldsymbol{\lambda}}_{r+1} = \cdots = \hat{\boldsymbol{\lambda}}_{n_i+n_j}$, $1 \leq r < n_i$,

(d) $\hat{\boldsymbol{\lambda}}_1 \geq \cdots \geq \hat{\boldsymbol{\lambda}}_r > \hat{\boldsymbol{\lambda}}_{r+1} = \cdots = \hat{\boldsymbol{\lambda}}_s > \hat{\boldsymbol{\lambda}}_{s+1} \geq \cdots \geq \hat{\boldsymbol{\lambda}}_{n_i+n_j}$,
$$1 \leq r < n_i < s < n_i + n_j.$$

(a) If $\hat{\mathbf{\Lambda}} = \hat{\boldsymbol{\lambda}}_1 \mathbf{I}_{n_i+n_j}$, then all diagonal elements of $\hat{\mathbf{H}}$ and $\hat{\mathbf{H}}'$ are affiliated with the multiple eigenvalue $\hat{\boldsymbol{\lambda}}_1$. Then for any permutation matrix $\hat{\mathbf{P}}_0$ and $\hat{\mathbf{Q}}$, we have $\hat{\mathbf{\Lambda}}' = [\hat{\mathbf{P}}_0 \hat{\mathbf{Q}}]^* \hat{\mathbf{\Lambda}} \hat{\mathbf{P}}_0 \hat{\mathbf{Q}} = \hat{\mathbf{\Lambda}}$.

(b) In this case we can write $\hat{\mathbf{\Lambda}} = \mathrm{diag}(\hat{\mathbf{\Lambda}}_1, \hat{\mathbf{\Lambda}}_2)$, where $\hat{\mathbf{\Lambda}}_1 = \hat{\boldsymbol{\lambda}}_1 \mathbf{I}_r$. Let us partition $\hat{\tilde{\mathbf{U}}}$, $\hat{\mathbf{P}}$, and $\hat{\mathbf{Q}}$ in accordance with the partition $(r, n_i + n_j - r)$ of $n_i + n_j$:

$$(5.22) \quad \hat{\tilde{\mathbf{U}}} = \begin{bmatrix} \hat{\tilde{\mathbf{U}}}_{11} & \hat{\tilde{\mathbf{U}}}_{12} \\ \hat{\tilde{\mathbf{U}}}_{21} & \hat{\tilde{\mathbf{U}}}_{22} \end{bmatrix}, \quad \hat{\mathbf{P}} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{21} & \mathbf{P}_{22} \end{bmatrix}, \quad \hat{\mathbf{Q}} = \begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{bmatrix} = \begin{bmatrix} Q_{ii} & \\ & Q_{jj} \end{bmatrix}.$$

Here, $\hat{\tilde{\mathbf{U}}}_{11}$, $\mathbf{P}_{11}$, and $\mathbf{Q}_{11}$ have dimension r, $r > n_i$, and Q_{ii} has dimension n_i. As in the proof of Lemma 5.6(ii), we conclude that for a sufficiently large k_1, $\|\hat{\tilde{\mathbf{U}}}_{12}\|_F$ and $\|\hat{\tilde{\mathbf{U}}}_{21}\|_F$ can be made so small that the QR algorithm with column pivoting of $[\tilde{U}_{ii} \ \tilde{U}_{ij}]$ yields $\hat{\mathbf{P}} = \mathrm{diag}(\mathbf{P}_{11}, \mathbf{I}_{n_i+n_j-r})$. Since the modified core algorithm has the property R, the matrix $\hat{\mathbf{P}}$ is replaced by $\hat{\mathbf{P}}_0 = \mathrm{diag}(\mathbf{P}_{11}^{(0)}, \mathbf{I}_{n_i+n_j-r})$, where $\mathbf{P}_{11}^{(0)}$ is of order r.

As for $\hat{\mathbf{Q}}$, we have $\hat{\mathbf{Q}} = \mathrm{diag}(Q_{ii}, Q_{jj})$, and let us consider the product $\hat{\mathbf{Q}}^* \hat{\mathbf{\Lambda}} \hat{\mathbf{Q}}$. This product reduces to two products: $Q_{ii}^* \hat{\boldsymbol{\lambda}}_1 \mathbf{I}_{n_i} Q_{ii} = \hat{\boldsymbol{\lambda}}_1 \mathbf{I}_{n_i}$ and $Q_{jj}^* \Lambda_{jj} Q_{jj} = \Lambda_{jj}$. Here we used the fact that Q_{jj} orders the diagonal of Λ_{jj} non-increasingly and that the diagonal elements of Λ_{jj} are already ordered non-increasingly. We obtain

$$\hat{\mathbf{\Lambda}}' = \hat{\mathbf{Q}}^* \hat{\mathbf{P}}_0^* \hat{\mathbf{\Lambda}} \hat{\mathbf{P}}_0 \hat{\mathbf{Q}} = \hat{\mathbf{Q}}^* \begin{bmatrix} [\mathbf{P}_{11}^{(0)}]^* \hat{\boldsymbol{\lambda}}_1 \mathbf{I}_r \mathbf{P}_{11}^{(0)} & \\ & \hat{\mathbf{\Lambda}}_2 \end{bmatrix} \hat{\mathbf{Q}} = \hat{\mathbf{Q}}^* \hat{\mathbf{\Lambda}} \hat{\mathbf{Q}} = \hat{\mathbf{\Lambda}}.$$

(c) In this case we have $\hat{\mathbf{\Lambda}} = \mathrm{diag}(\hat{\mathbf{\Lambda}}_1, \hat{\mathbf{\Lambda}}_2)$, $\hat{\mathbf{\Lambda}}_2 = \hat{\boldsymbol{\lambda}}_{n_i+n_j} \mathbf{I}_{n_i+n_j-r}$, $1 \leq r < n_i$. Let us partition $\hat{\tilde{\mathbf{U}}}$, $\hat{\mathbf{P}}$, and $\hat{\mathbf{Q}}$ in accordance with $\hat{\mathbf{\Lambda}}$. We obtain the relation (5.22) where $\hat{\tilde{\mathbf{U}}}_{11}$, $\mathbf{P}_{11}$, and $\mathbf{Q}_{11}$ have dimension r. For a sufficiently large k_1, $\|\hat{\tilde{\mathbf{U}}}_{12}\|_F$ and $\|\hat{\tilde{\mathbf{U}}}_{21}\|_F$ can be made so small that the Householder QR algorithm with column pivoting of $[\tilde{U}_{ii} \ \tilde{U}_{ij}]$ yields (see Remark 4.1) $\hat{\mathbf{P}} = \mathrm{diag}(\mathbf{P}_{11}, \mathbf{P}_{22})$. Here we used the fact that $\hat{\tilde{\mathbf{U}}}_{11}$ is close to a unitary matrix. Since the modified core algorithm has the property R, the matrix $\hat{\mathbf{P}}$ is replaced by $\hat{\mathbf{P}}_0 = \mathrm{diag}(\mathbf{I}_r, \mathbf{P}_{22})$. Now, as the matrix $\mathbf{P}_{22}$ has the same dimension as $\hat{\mathbf{\Lambda}}_2$, we have $\hat{\mathbf{P}}_0^T \hat{\mathbf{\Lambda}} \hat{\mathbf{P}}_0 = \hat{\mathbf{\Lambda}}$.

We also have $\mathbf{Q}_{12} = 0$, $\mathbf{Q}_{21} = 0$, so that $\hat{\mathbf{Q}} = \mathrm{diag}(\mathbf{Q}_{11}, \mathbf{Q}_{22}) = \mathrm{diag}(Q_{ii}, Q_{jj})$. This is implied by the assumption $\hat{\boldsymbol{\lambda}}_r > \hat{\boldsymbol{\lambda}}_{r+1}$. Now, the permutation matrix Q_{ii} orders the diagonal

elements of Λ_{ii} non-increasingly, although they have already been ordered non-increasingly. Therefore, we have $Q_{ii}^* \Lambda_{ii} Q_{ii} = \Lambda_{ii}$, and the proof of $\hat{\mathbf{\Lambda}}' = \hat{\mathbf{\Lambda}}$ is completed similarly as in the case (b).

(d) In this case we partition $\hat{\mathbf{\Lambda}}$, $\hat{\hat{\mathbf{U}}}$, $\hat{\mathbf{P}}$, and $\hat{\mathbf{Q}}$ in accordance with the partition $(r, s - r, n_i + n_j - s)$ of $n_i + n_j$. We have $\hat{\mathbf{\Lambda}} = \text{diag}(\hat{\mathbf{\Lambda}}_1, \hat{\mathbf{\Lambda}}_2, \hat{\mathbf{\Lambda}}_3)$, with $\hat{\mathbf{\Lambda}}_2 = \hat{\mathbf{\lambda}}_s \mathbf{I}_{s-r}$, and

$$
\hat{\hat{\mathbf{U}}} = \begin{bmatrix} \hat{\hat{\mathbf{U}}}_{11} & \hat{\hat{\mathbf{U}}}_{12} & \hat{\hat{\mathbf{U}}}_{13} \\ \hat{\hat{\mathbf{U}}}_{21} & \hat{\hat{\mathbf{U}}}_{22} & \hat{\hat{\mathbf{U}}}_{23} \\ \hat{\hat{\mathbf{U}}}_{31} & \hat{\hat{\mathbf{U}}}_{32} & \hat{\hat{\mathbf{U}}}_{33} \end{bmatrix}, \quad
\hat{\mathbf{P}} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} & \mathbf{P}_{13} \\ \mathbf{P}_{21} & \mathbf{P}_{22} & \mathbf{P}_{23} \\ \mathbf{P}_{31} & \mathbf{P}_{32} & \mathbf{P}_{33} \end{bmatrix}, \quad
\hat{\mathbf{Q}} = \begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} & \mathbf{Q}_{13} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} & \mathbf{Q}_{23} \\ \mathbf{Q}_{31} & \mathbf{Q}_{32} & \mathbf{Q}_{33} \end{bmatrix}.
$$

As above, we conclude that for a sufficiently large k_1, $\|[\hat{\hat{\mathbf{U}}}_{12}, \hat{\hat{\mathbf{U}}}_{13}]\|_F$ and $\|\hat{\hat{\mathbf{U}}}_{23}\|_F$ can be made so small that the QR factorization with column pivoting of $[\bar{U}_{ii} \;\; \bar{U}_{ij}]$ yields $\hat{\mathbf{P}} = \text{diag}(\mathbf{P}_{11}, \mathbf{P}_{22}, \mathbf{I}_{n_i+n_j-s})$. Here $\mathbf{P}_{11} = \mathbf{I}_{11'} \cdots \mathbf{I}_{r,r'}$ with all second subscripts $1', \ldots, r'$ not larger than $r < n_i$. Because the modified core algorithm has the property R, the matrix $\hat{\mathbf{P}}$ is replaced by $\hat{\mathbf{P}}_0 = \text{diag}(\mathbf{I}_r, \mathbf{P}_{22}, \mathbf{I}_{n_i+n_j-s})$. Since $\mathbf{P}_{22}$ has the same dimension as $\hat{\mathbf{\Lambda}}_2 = \hat{\mathbf{\lambda}}_s \mathbf{I}_{s-r}$, we have $\hat{\mathbf{P}}_0^T \hat{\mathbf{\Lambda}} \hat{\mathbf{P}}_0 = \hat{\mathbf{\Lambda}}$.

As for the matrix $\hat{\mathbf{Q}}$, for a sufficiently large k_1 we must have $\mathbf{Q}_{12} = 0$, $\mathbf{Q}_{13} = 0$, $\mathbf{Q}_{23} = 0$, and therefore $\hat{\mathbf{Q}} = \text{diag}(\mathbf{Q}_{11}, \mathbf{Q}_{22}, \mathbf{Q}_{33})$. This is implied by the assumption $\hat{\mathbf{\lambda}}_r > \hat{\mathbf{\lambda}}_{r+1} = \cdots = \hat{\mathbf{\lambda}}_s > \hat{\mathbf{\lambda}}_{s+1}$. Furthermore, $\mathbf{Q}_{11}$ and $\mathbf{Q}_{33}$ are block-diagonal and such that $\mathbf{Q}_{11}^* \hat{\mathbf{\Lambda}}_1 \mathbf{Q}_{11} = \hat{\mathbf{\Lambda}}_1$ and $\mathbf{Q}_{33}^* \hat{\mathbf{\Lambda}}_3 \mathbf{Q}_{33} = \hat{\mathbf{\Lambda}}_3$. Since $\hat{\mathbf{\Lambda}}_2 = \hat{\mathbf{\lambda}}_s \mathbf{I}_{s-r}$ we have $\mathbf{Q}_{22}^* \hat{\mathbf{\Lambda}}_2 \mathbf{Q}_{22} = \hat{\mathbf{\Lambda}}_2$. Thus, $\hat{\mathbf{Q}}^* \hat{\mathbf{\Lambda}} \hat{\mathbf{Q}} = \hat{\mathbf{\Lambda}}$. Together with $\hat{\mathbf{P}}_0^T \hat{\mathbf{\Lambda}} \hat{\mathbf{P}}_0 = \hat{\mathbf{\Lambda}}$ this proves $\hat{\mathbf{\Lambda}}' = \hat{\mathbf{\Lambda}}$ and completes the proof of the theorem. $\square$

Next, we consider the case when the core algorithm is the simple Jacobi method under the de Rijk pivot strategy.

COROLLARY 5.10. *Let the modified block Jacobi method be defined by the de Rijk pivot strategy. Let the basic core algorithm be the element-wise Jacobi method under the de Rijk pivot strategy, and let the modified core algorithm have the property* R. *Then the iterative process will reach the stage when* $\mathbf{off}(\mathbf{H}^{(k)}) \leq \delta$, $k \geq k_0$.

If during just one full cycle, starting with the step $t_0 M$, $t_0 M \geq k_0$, *the modified core algorithm is replaced by the basic core algorithm, then the block Jacobi method will have the property* D, *and therefore it will be globally convergent.*

Proof. By Theorem 5.2 the block Jacobi method converges to diagonal form. We have already shown that the simple Jacobi method under the de Rijk pivot strategy is well-behaved. Hence, Theorem 5.9 implies that it remains to show that the block method has the property D.

We will use the following fact: if the core algorithm is the simple Jacobi method under the de Rijk pivot strategy, then the diagonal elements of each $\hat{\mathbf{H}}^{(k)\,'} = [\hat{\mathbf{U}}^{(k)}]^* \hat{\mathbf{H}}^{(k)} \hat{\mathbf{U}}^{(k)}$ are ordered non-increasingly.

Let us consider the cycle of the process that begins with $t_0 M$, $t_0 M \geq k_0$ where k_0 is from the relation (3.15). The proof mimics the proof of Theorem 5.4. To simplify notation we can assume that $t_0 = 0$. Then the diagonal elements of the first diagonal block of $\mathbf{H}^{(m)}$ are affiliated with the largest eigenvalue(s).

If the multiplicity ν_1 of λ_{ς_1} from (3.11) is not larger than n_1, then the first ν_1 diagonal elements of $H_{11}^{(m)}$ will be affiliated with λ_{ς_1}. Otherwise, all diagonal elements of $H_{11}^{(m)}$ are affiliated with λ_{ς_1}. If $n_1 + \cdots + n_{\ell-1} < \nu_1 \leq n_1 + \cdots + n_\ell$, then we have to consider first $\tau(\ell, m) = m - 1 + (m - 2) + \cdots + (m - \ell)$ steps of the block method. Since the core algorithm orders the diagonal elements of the current pivot submatrix non-increasingly, we conclude that in $\mathbf{H}^{(\tau(\ell,m))}$ the first ν_1 diagonal elements are affiliated to λ_{ς_1}. Repeating the same argument for the diagonal elements that are affiliated to λ_{ς_2} and later to the other

eigenvalues, we obtain that the associated matrix $\mathbf{\Lambda}^{(M)}$ has non-increasingly ordered diagonal elements. This ordering cannot be disrupted by later transformations as shown in the proof of Theorem 5.9. This shows that the smallest k from Definition 5.8 satisfies $k \leq (t_0 + 1)M$.
□

Corollary 5.10 raises a question how to implement it in a practical computation. In which cycle to replace the modified core algorithm with the basic one? Let t count the cycles. It is sufficient to provide a sequence of positive integers, say, $(3, 5, 7, 9, \ldots)$, and then in cycles 3, 5, 7, 9, ... replace the modified with the basic core algorithm. To ensure global convergence, one has to take care that there is an infinite number of cycles that employ the modified core algorithm.

Anyway, after the cycle t_0 (t_0 is from Corollary 5.10) is completed, the modified core algorithm with the property R will be reduced to the basic one.

5.5. Quadratic convergence of the block method. The following quadratic convergence result can be seen as a corollary of Proposition 5.1 and Theorem 5.9.

THEOREM 5.11. *Consider the modified block Jacobi method under the de Rijk pivot strategy such that it has the property* D*, the basic core algorithm is well-behaved, and the modified core algorithm has the property* R*. Then it is quadratically convergent, provided the diagonal elements that converge to the same eigenvalue are contained within the same diagonal block or within two adjacent diagonal blocks of the basic block-matrix partition. For a sufficiently large t_0 we have*

$$\mathbf{off}(\mathbf{H}^{([t+1])}) \leq 0.27622 \frac{\mathbf{off}^2(\mathbf{H}^{[t]})}{\delta}, \qquad t \geq t_0.$$

Proof. By Theorem 5.9 the block Jacobi method is globally convergent. So, there exists a k_0 such that the diagonal elements of $\mathbf{\Lambda}^{(k)}$, $k \geq k_0$, are ordered non-increasingly and the relation (5.11) holds. Let t_0 be the smallest integer such that $t_0 M \geq k_0$. We consider the cycle t of the method such that $t \geq t_0$. To simplify notation, we can assume $t = 0$. In the first cycle, the process starts with $\mathbf{H}^{(1)} = \mathbf{H}^{[0]}$ and ends with $\mathbf{H}^{(M+1)} = \mathbf{H}^{[1]}$.

Instead of using (5.10) and (5.8), which are linked to the simple method, we will use the relation (5.1) from Proposition 5.1. By a variant of Bernoulli's inequality we obtain

$$\mathbf{off}^2(\mathbf{H}^{[1]}) \leq \mathbf{off}^2(\mathbf{H}^{[0]}) \left(1 - \prod_{\substack{k=1 \\ i(k) \neq j(k)-1}}^{M-1} \cos^2 \varphi^{(k)} \right) \leq \mathbf{off}^2(\mathbf{H}^{[0]}) \sum_{\substack{k=1 \\ i(k) \neq j(k)-1}}^{M-1} \sin^2 \varphi^{(k)},$$

where $\cos \varphi^{(k)} = \sigma_{\min}(U_{i(k)i(k)}^{(k)})$ is the cosine of the maximal angle from the Cosine-Sine decomposition of $\hat{\mathbf{U}}^{(k)}$. In the following we will replace the subscripts $i(k)$, $j(k)$ by i, j, respectively.

From the same decomposition of $\hat{\mathbf{U}}^{(k)}$ we get

$$\sin \varphi^{(k)} = \sigma_{\max}(U_{ji}^{(k)}) = \|U_{ji}^{(k)}\|_2 = \|U_{ij}^{(k)}\|_2.$$

From (2.8) we obtain

$$\sum_{\substack{k=1 \\ i \neq j-1}}^{M-1} \|H_{ij}^{(k)}\|_F^2 \leq \frac{1}{2} \mathbf{off}^2(\mathbf{H}^{[0]}).$$

To connect the last two relations, we use $\hat{\mathbf{H}}^{(k)}\hat{\mathbf{U}}^{(k)} = \hat{\mathbf{U}}^{(k)}\hat{\mathbf{H}}^{(k)}{}'$, which is linked to the transformation of the pivot submatrix. Equating the $(1,2)$-submatrices on the left- and right-hand sides of this equation, we obtain

$$H_{ii}^{(k)}U_{ij}^{(k)} - U_{ij}^{(k)}H_{jj}^{(k)\prime} = -H_{ij}^{(k)}U_{jj}^{(k)}.$$

Since the diagonal blocks $H_{ii}^{(k)}$ and $H_{jj}^{(k)\prime}$ are diagonal, we can assume that $H_{ii}^{(k)} = \mathrm{diag}(\xi_1^{(k)}, \ldots, \xi_{n_i}^{(k)})$ and $H_{jj}^{(k)\prime} = \mathrm{diag}(\xi_1^{(k)\prime}, \ldots, \xi_{n_j}^{(k)\prime})$. For the (r,s)-element of $U_{ij}^{(k)}$, we obtain

$$(\xi_s^{(k)\prime} - \xi_r^{(k)})(U_{ij}^{(k)})_{rs} = (H_{ij}^{(k)}U_{jj}^{(k)})_{rs}.$$

Because the relation (5.11) holds, the diagonal elements of $H_{ii}^{(k)}$ and $H_{jj}^{(k)\prime}$ are at most 0.22δ close to the corresponding diagonal elements of $\hat{\mathbf{\Lambda}}^{(k)}$. The diagonal elements converging to the same eigenvalue are contained within the same diagonal block or within two adjacent diagonal blocks. Since $i + 1 < j$ we conclude that

$$|\xi_r^{(k)} - \xi_s^{(k)\prime}| \geq 3\delta - 2 \cdot 0.22\delta = 2.56\delta.$$

Hence, we obtain

$$|(U_{ij}^{(k)})_{rs}| = \frac{|(H_{ij}^{(k)}U_{jj}^{(k)})_{rs}|}{|\xi_r^{(k)} - \xi_s^{(k)\prime}|} \leq \frac{|(H_{ij}^{(k)}U_{jj}^{(k)})_{rs}|}{2.56\delta}, \qquad 1 \leq r \leq n_i < s \leq n_i + n_j,$$

and consequently,

$$\|U_{ij}^{(k)}\|_2 \leq \|U_{ij}^{(k)}\|_F \leq \frac{\|H_{ij}^{(k)}U_{jj}^{(k)}\|_F}{2.56\delta} \leq \frac{\|H_{ij}^{(k)}\|_F\|U_{jj}^{(k)}\|_2}{2.56\delta} \leq \frac{\|H_{ij}^{(k)}\|_F}{2.56\delta}.$$

Combining all previous relations we arrive at

$$\begin{aligned}
\mathbf{off}^2(\mathbf{H}^{[1]}) &\leq \mathbf{off}^2(\mathbf{H}^{[0]}) \sum_{\substack{k=1 \\ i \neq j-1}}^{M-1} \sin^2\varphi^{(k)} = \mathbf{off}^2(\mathbf{H}^{[0]}) \sum_{\substack{k=1 \\ i \neq j-1}}^{M-1} \|U_{ij}^{(k)}\|_2^2 \\
&\leq \frac{\mathbf{off}^2(\mathbf{H}^{[0]})}{2.56^2\delta^2} \sum_{\substack{k=1 \\ i(k) \neq j(k)-1}}^{M-1} \|H_{ij}^{(k)}\|_F^2 \leq \frac{\mathbf{off}^4(\mathbf{H}^{[0]})}{2 \cdot 2.56^2\delta^2}.
\end{aligned}$$

To complete the proof, one has to take the square root of the left- and right-hand side of the obtained inequality.   □

Once the eigenvalues are computed, we can detect from (3.11) the natural partition. Since the basic partition is also known, we can verify *a posteriori* whether the conditions of the corollary are met.

The quadratic convergence proof of the block Jacobi method in the general case is still an open problem. A first approach to solve it could be to try to follow the lines of the proof in [22] and use some of the above relations.

6. The Kogbetliantz method under the de Rijk pivot strategy. If one wants to compute the singular value decomposition of a square matrix, then one can use the simple [5, 35, 39, 40] or block [4, 53] Kogbetliantz method. However, there are several reasons why not to use the de Rijk pivot strategy with this method. To explain it, we will consider the simple method. All arguments that hold for the simple method will also hold for the block method.

The most precious feature of a method is its ability to compute the output data with an accuracy that is close to the one that the problem allows. Such a method is said to have the high relative accuracy property. The QR factorization is a useful tool which reduces a rectangular or square matrix to triangular form. This procedure not only reduces the matrix to a simpler form but it also maintains the property of the matrix that allows for the computation of the SVD to high relative accuracy.

So, a typical procedure for computing the SVD of a square matrix would first use the QR factorization with column pivoting and then would employ some iterative method to compute the SVD of the obtained triangular matrix. An excellent choice for the second procedure is the provably globally convergent row- (or column-) cyclic Kogbetliantz method [34]. The reason is threefold:

- The algorithm is simplified because during the iteration the pivot submatrix is triangular and half of all of the off-diagonal elements are zero. This can be used to reduce the cost associated with one step of the method. Approximately, the number of flops needed for updating the current matrix is one half of the number of flops when the Kogbetliantz method updates the full square matrix.
- The total number of sweeps (and thus of steps) in practical computation is reduced because the method is quadratically convergent [20, 22]. The quadratic convergence is guaranteed in the case of simple singular values. In the case of multiple singular values, the diagonal elements converging to the same multiple singular value have to occupy adjacent positions on the diagonal. This can be achieved by using a special procedure at the end of each cycle.
- The method computes the singular values to high relative accuracy [45].

Now, let us consider what happens when the Kogbetliantz method is applied to an upper-triangular matrix $\mathbf{A}$ under the de Rijk pivot strategy. One can easily see that each permutation matrix that is part of the pivot strategy ruins the special (sometimes called *essentially triangular* or *hierarchically triangular*) form that is created by the row-cyclic strategy [34]. This means that at the end of the first cycle the matrix will become the full square matrix. When applied to a full square matrix under the de Rijk strategy, the Kogbetliantz method has the following disadvantages:

- The method is not provably globally convergent, i.e., there is no proof of global convergence.
- The method is not efficient. Each step is twice as expensive as the step of the row-cyclic method when applied to the triangular matrix. The additional cost comes from the fact that in the case of multiple singular values, the method is not quadratically convergent [19]. Namely, this means more sweeps until convergence.
- There is no proof that the Kogbetliantz method is a high relative accuracy method when applied to a full square matrix.

In conclusion, although the de Rijk pivot strategy can be defined and used with the simple and block Kogbetliantz method, there are reasons why not to use it.

However, the de Rijk pivot strategy can be used with the simple and block one-sided Jacobi method for computing the SVD of a triangular matrix. The global and quadratic convergence results proved in this paper apply directly to those one-sided Jacobi methods.

7. Several numerical examples. We have used MATLAB to observe the behavior of the function $\mathbf{off}(\mathbf{H}^{(k)})$ for all steps k until convergence. Here, $\mathbf{H}$ is the initial real symmetric matrix. We considered the real simple and block Jacobi methods.

The block method is supposed to operate on large matrices. If the computation is done on a parallel computer, it can use any parallel cyclic pivot strategy, and as the core algorithm it then can use the simple Jacobi method. We will see that the de Rijk pivot strategy is a good choice for the simple Jacobi method when it is used as the core algorithm. For smaller matrices the computation can be done on standard computers like a PC. In this case one can use both the block and simple Jacobi method and can employ the de Rijk strategy in both cases.

As has been shown in [13, 14], before employing the one-sided Jacobi method to solve the SVD of a matrix, it is advantageous to apply the QR factorization with column pivoting of the matrix and, oftentimes additionally, the LQ factorization of the obtained triangular factor R. In this way the obtained factor L is closer to diagonal form, and its diagonal entries are arranged in non-increasing ordering. Furthermore, if the initial matrix allows for accurate singular value computation, then the same is true for the computed triangular factors R and L. Then the block or simple Jacobi method makes an excellent choice for the iterative part of the procedure.

If the eigenvalue problem of a symmetric (or complex Hermitian) positive definite matrix H is considered, then one can first use the Cholesky factorization with diagonal pivoting followed by the LQ factorization of the Cholesky factor. Again the one-sided block Jacobi method can be used to solve the SVD of the computed factor L, and the core algorithm can be the simple Jacobi method under the de Rijk pivot strategy.

These important cases direct us to consider the convergence properties of the block or simple Jacobi method for symmetric or complex Hermitian matrices which are somewhat closer to diagonal form and have non-increasing diagonal elements. Here we considered only real matrices and methods.

7.1. The simple Jacobi method under the de Rijk strategy. As mentioned earlier, the simple Jacobi method is a good choice for the core algorithm of the block method. Here are a few facts that justify it:

- the core algorithm operates on matrices of small size, typically between 32 and 512, and
- the pivot submatrices $\hat{\mathbf{H}}^{(k)}$ in the block Jacobi method are most of the time almost diagonal. For almost diagonal symmetric matrices, the simple Jacobi method is highly accurate [44] and efficient since a quadratic asymptotic convergence takes place (see [22, 61] and Section 5.3 here).

The core algorithm is used to compute both the eigenvalues and eigenvectors. Therefore, we will make several tests to see whether the de Rijk strategy minimizes the total number of cycles required to diagonalize a symmetric matrix. To this end we will compare the efficiency of five pivot strategies:

1. the row-cyclic strategy,
2. the row-cyclic strategy modified as follows: before each cycle the permutational similarity transformation is applied to the matrix in such a way that the diagonal elements are ordered non-decreasingly,
3. the row-cyclic strategy modified as in the latest item, only that the diagonal elements are ordered non-increasingly,
4. the de Rijk pivot strategy,
5. the de Rijk pivot strategy, but before applying the Jacobi method, the diagonal elements are ordered non-increasingly.

The corresponding simple Jacobi methods are coded in MATLAB as functions named

`djacobi, djacobi_a, djacobi_d, djac_Rijk,` and `djac_DRijk.`

The input and output arguments are the same for all functions. The input arguments are a symmetric matrix A and a variable `eivec` that determines whether the eigenvectors are to be computed, e.g., the function `djacobi_Rijk` is invoked by

$$[eivs,V,offs,indicator] = djac\_Rijk(A,eivec).$$

Here, `eivs` is the column vector of eigenvalues, `V` is the eigenvector matrix, and `indicator` is a vector of six components.

The first component indicates if some problem was encountered before or during the iteration. The second and third entry count the total number of steps and number of steps in which the transformation was really computed. Namely, in the later stage of the process, the pivot elements become so small that they can be set to zero [55] without performing the similarity transformation. The fourth and fifth entries are the total number of cycles and the number of actual cycles (the third entry divided by $N = n(n-1)/2$). The last entry is the total number of swaps used by the method. This number includes the swaps that are used in the permutational similarity transformations provided they are a part of the method (this refers to the functions `djacobi_a`, `djacobi_d`, `djac_Rijk`, and `djacobi_DRijk`).

The output variable `offs` is a two-column matrix. The first column stores the ordinal numbers of the steps at which the off-norm is computed while the second column contains the corresponding off-norms. In each cycle the off-norm is computed `ndisp` times. The default of `ndisp` is 20. The output `offs` is used to plot the off-norm curve.

Let us describe the stopping criterion of the process. We have used the one from [55], which works well for the two-sided Jacobi methods. Before computing the elements of the rotation matrix, it is tested in the script whether the pivot element a_{ij} can be replaced with zero, and the process continues with the next step. The pivot element is set to zero if $fl(|a_{ii}|+\mu\,|a_{ij}|) = |a_{ii}|$ and $fl(|a_{jj}|+\mu\,|a_{ij}|) = |a_{jj}|$, where $\mu = 100$. Here $fl(e)$ denotes the computed value of the expression e. This implies $\mu|a_{ij}| \leq \varepsilon|a_{ii}|$ and $\mu|a_{ij}| \leq \varepsilon|a_{jj}|$, and consequently $|a_{ij}|/\sqrt{|a_{ii}a_{jj}|} \leq \varepsilon/\mu$. Here, ε is the unit roundoff. Replacing the pivot element by zero in such a way will be referred to as an empty or zero transformation. The process is terminated when all off-diagonal elements become zero. From [6] we easily deduce that replacing a_{ij} by zero in this way can relatively perturb each eigenvalue of the positive definite matrix A by an amount not larger than ε/μ.

We have used two ways of generating the initial symmetric matrix A of order n. The first one generates a positive definite matrix which, for given n, depends on four parameters `k1,k2,k3,` and `kk`. To compute A we have used the following code:

```
function retval = matgen(n,k1,k2,k3,kk)
X=rand(n); X=X.'*X; D=diag(scalvec(n,k1,k2,k3,kk)); X=D*X*D; retval= 0.5*(X+X.');
end

function retval=scalvec(n,k1,k2,k3,k)
d1=logspace(round(k1),round(k2),k); d2=logspace(round(k2),round(k3),n-k+1);
retval=[d1,d2(2:n-k+1)];
end
```

We have chosen n = 512 and `kk=round(n/2)`, i.e., $kk = 256$. The three parameters `k1`, `k2`, `k3` are selected using three loops which begin with the statements: `for k1=5:-4:-3`, `for k2=3:-4:-5`, `for k3= 2:-5:-8`. In this way most of the generated matrices have descending diagonal elements, and their eigenvalues can be computed to high relative accuracy.

For all figures, the vertical lines of the grid separate the cycles. The ticks on the x-axis coincide with the vertical lines. They are labeled by the appropriate number of steps: $1, N+1,$

$2N + 1, 3N + 1, \dots$, where $N = n(n-1)/2 = 130816$. The y-axis is labeled `off-norm` because the y-values of the graphs are obtained from the second column of the appropriate matrices `offs`. In Figure 7.1 we have selected six typical cases from the obtained 27 graphs.
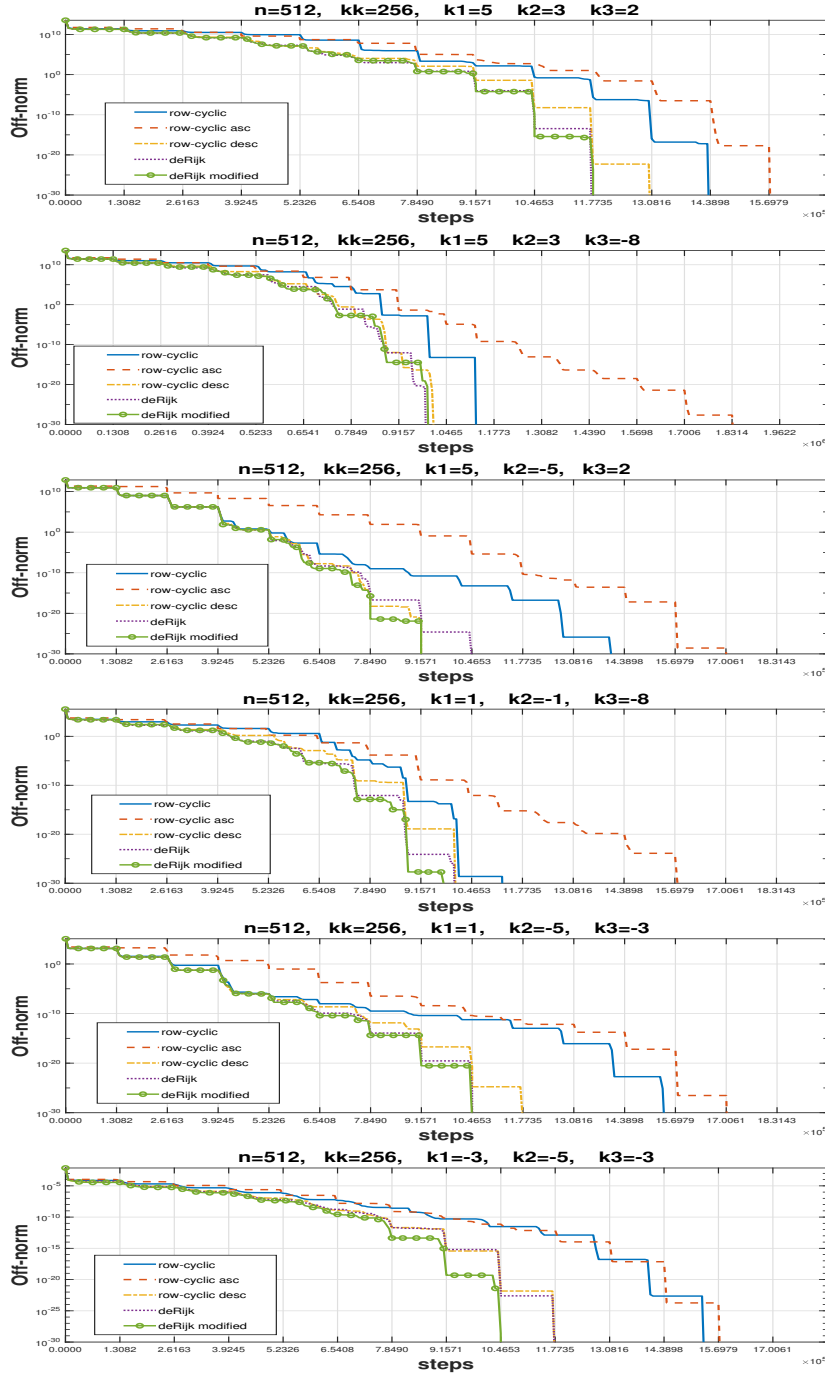

FIG. 7.1. *The graphs of the off-norm function.*

We have labeled the graphs obtained by the functions

> djacobi, djacobi_a, djacobi_d, djac_Rijk,
> djacobi_DRijk

with

> row-cyclic, row-cyclic asc, row-cyclic desc, deRijk,
> deRijk modified,

respectively.

We see that the performance of the modified de Rijk pivot strategy was always the best or among the best ones. The number of swaps that occurred in djac_Rijk was the smallest (in the range 173–1340), followed by the number of swaps in djacobi_DRijk (202–1263), djacobi_d (293–1840), and djacobi_a (1219–2286).

The second way generates a positive definite matrix A which, for given n, also depends on the parameters k1, k2, k3, and kk. This time we have chosen $n = 256$, kk=round(n/2). The matrix A is computed using the code which is listed below:

```
[Q,~]=qr(rand(n)); m=round(n/5); nn=10; nn1=20;
d=scalvec(n,k1,k2,k3,kk); for k=m+1:m+nn, d(k)=d(m); end
for k=n-m+1:n-m+nn1, d(k)=d(n-m)+1e-4*sin(pi*k/n); end
A=Q*diag(d)*Q.'; A=0.5*(A+A.');
```

The matrix A has two clusters of eigenvalues; the one of smaller size includes 10 eigenvalues, and the other one has 20 eigenvalues. The second cluster has a width of about 10^{-4}. The width of the first cluster depends on k1, k2, k3, and on the limitations of the computer arithmetic. This width is as small as $\max\{10^{k1}, 10^{k2}, 10^{k3}\}\varepsilon$, where ε is the unit roundoff. The graphs obtained by MATLAB are very similar to the ones presented above. Interestingly, with three of the eight matrices, the performance of djacobi was worse than or equal to the performance of djacobi_a.

7.2. The block Jacobi method under the de Rijk strategy. We have tested four variants of the block methods under the de Rijk pivot strategy. The first two variants are the *bdR1* method with and without the initial sorting of the diagonal elements in the non-increasing ordering. The third and fourth variants are the *bdR2* method with and without the initial sorting of the diagonal elements. In all four block methods, the core algorithm was the same, the simple Jacobi method under the modified de Rijk strategy. This means that the initial sorting of the diagonal elements of the pivot submatrix was used. The input and output parameters are similar as in the scripts related to the simple Jacobi methods, only that the off-norms are computed after certain block-steps. The block iterative process is terminated when in one cycle all transformations within the core algorithm are empty. This automatically implies that **off**$(A) = 0$.

We have used the same algorithms for generating the initial symmetric matrix. Thus, for a given n, the initial matrix is determined by four parameters: k1, k2, k3, and kk. When we used the parameters satisfying k1 $\geq$ k2 $\geq$ k3, kk $= n/2$, the off-norm reduction was almost the same for all four block methods. Therefore, we have chosen $n = 1024$, kk $= 512$, k1 $= 1$, k2 $= 5$, k3 $= -4$. For the block methods, we have chosen the block size to be 4, 8, 16, 32, and 64.

As mentioned earlier, in each figure, the grid's vertical lines mark the beginning of the new cycle. The x-ticks coincide with these vertical lines. In Figure 7.2 all five graphs of the off-norm functions for different choices of the block sizes are displayed.
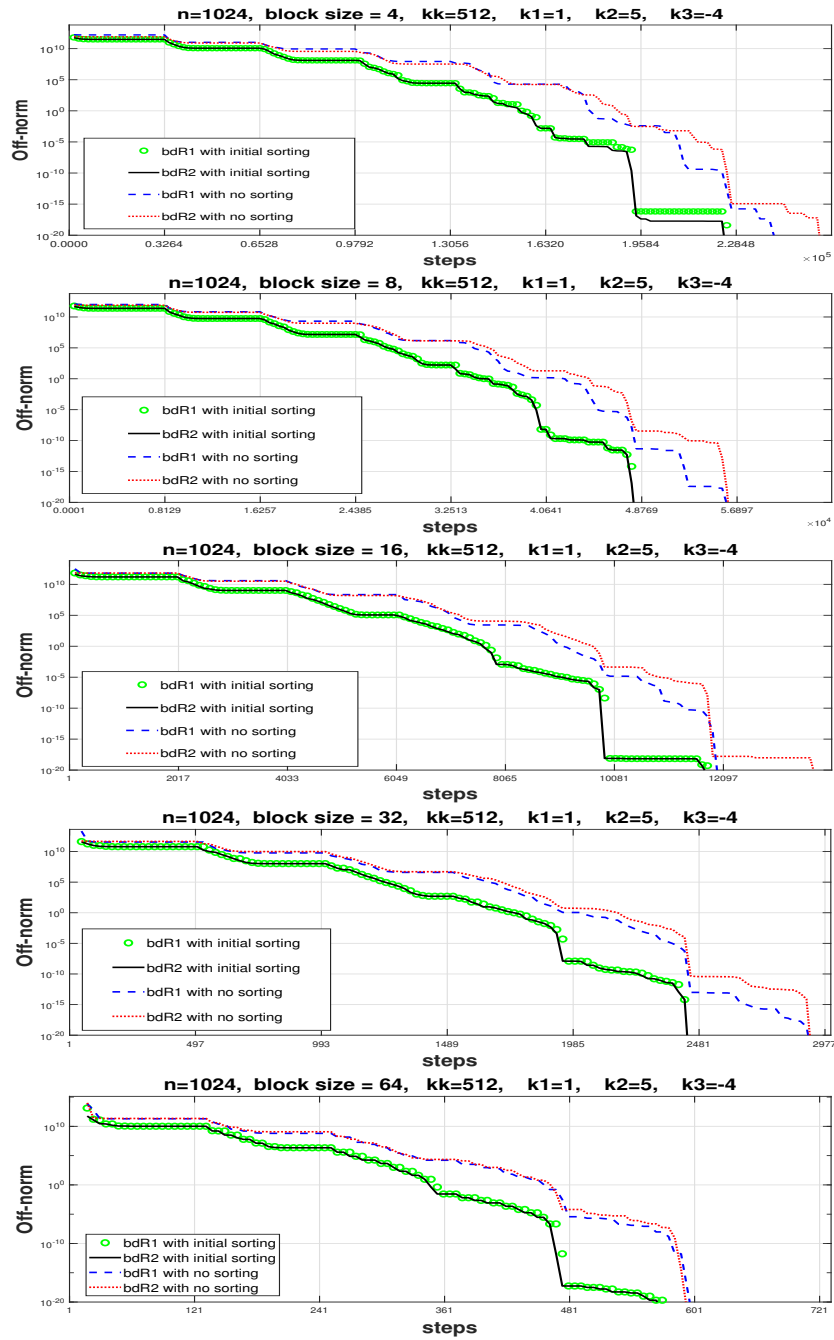
FIG. 7.2. *The graphs of the off-norm function.*

As is expected, the larger the block size, the fewer cycles are needed to complete the process. It can be seen that the initial sorting of the diagonal elements really contributes to the decrease of the total number of cycles. We can also observe the ultimate quadratic reduction of the off-norm per cycle. Next, we considered the off-norm function when the symmetric

matrix has clusters of mutually close eigenvalues. We have chosen $n = 2048$ and the block size 16. The symmetric matrix A is generated in a similar way as in Section 7.1. We have used the function `gensymm_mult` whose listing is given below:

```
function A = gensymm_mult(n,kk,k1,k2,k3,posit,rng_num)
[~,m]=size(posit); d=scalvec(n,k1,k2,k3,kk); rng(rng_num);
for i=1:m
    pi=posit(1,i); ni=posit(2,i); ei=posit(3,i);
    for j=pi:pi+ni, d(j) = d(pi)+ei*(2.0*rand-1.0); end
end
[Q,~]=qr(rand(n)); A=Q*diag(d)*Q.'; A=0.5*(A+A.'); end
```

The code first generates the vector `d` whose entries will be close to the eigenvalues of `A`. This is accomplished by using the `scalvec` function. The two loops are used to replace certain adjacent entries of `d` with repeated entries or clusters of close numbers. The input parameter posit is a three-row matrix. The entries of the first row of `posit` determine locations from which the entries of `d` will be changed. The second row determines how many entries of `d` are changed at each location. The third row determines the width of the cluster at each location. If the width of some cluster is zero, then we have repeated the entries in `d` at that location. The last line of the code makes a similarity transformation of the diagonal matrix `diag(d)` with an orthogonal matrix `Q`. Obviously, since all computations are made in double precision, the exact eigenvalues of `A` and the entries of `d` will not be the same. They can differ by quantities bounded by $\max\{10^{k1}, 10^{k2}, 10^{k3}\}\varepsilon$, where ε is the unit roundoff.

In this way we have generated three matrices of order 2048. The input matrices `posit` are given below:

$$\begin{bmatrix} 146 & 402 & 800 & 1506 \\ 10 & 10 & 10 & 10 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 150 & 400 & 800 & 1500 \\ 10 & 10 & 10 & 10 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 150 & 400 & 800 & 1500 \\ 30 & 40 & 50 & 40 \\ 10^{-5} & 10^{-10} & 0 & 10^{-8} \end{bmatrix}.$$

In Figure 7.3 we display the graphs of the off-norm function linked to the same block methods that we used in past experiments.

From the first graph we see that the asymptotic quadratic convergence occurs when the diagonal elements associated with a tiny cluster of eigenvalues lie in one diagonal block of the block-matrix partition. This is in harmony with Theorem 5.7. It seems that this condition also reduces the total number of cycles. Note that the underlying matrix has clusters of width smaller than 10^{-13}.

The symmetric matrix linked to the second graph has clusters of width not larger than 10^{-11}, and in the later stage of the process the clusters of mutually close diagonal elements are not contained in one diagonal block. This slows the process, but ultimately the quadratic convergence prevails.

From the last graph we see that the process slows down even more when the clusters of eigenvalues are large and their widths vary largely. In the later stage of the process the diagonal elements affiliated with one cluster of eigenvalues spread through more than two diagonal blocks. In this case the process has terminated before the ultimate quadratic convergence has begun.

We end this section by examining the magnitude of $\sigma_{\min}(U^{(k)}_{i(k)i(k)})$ during the process. From Proposition 5.1 and relation (5.1) we conclude that the larger the minimum magnitude of $\sigma_{\min}(U^{(k)}_{i(k)i(k)})$ during one cycle, the better the reduction of the off-norm in that cycle. Our aim is to numerically investigate what block method delivers the largest minimum of $\sigma_{\min}(U^{(k)}_{i(k)i(k)})$ over all iterations.
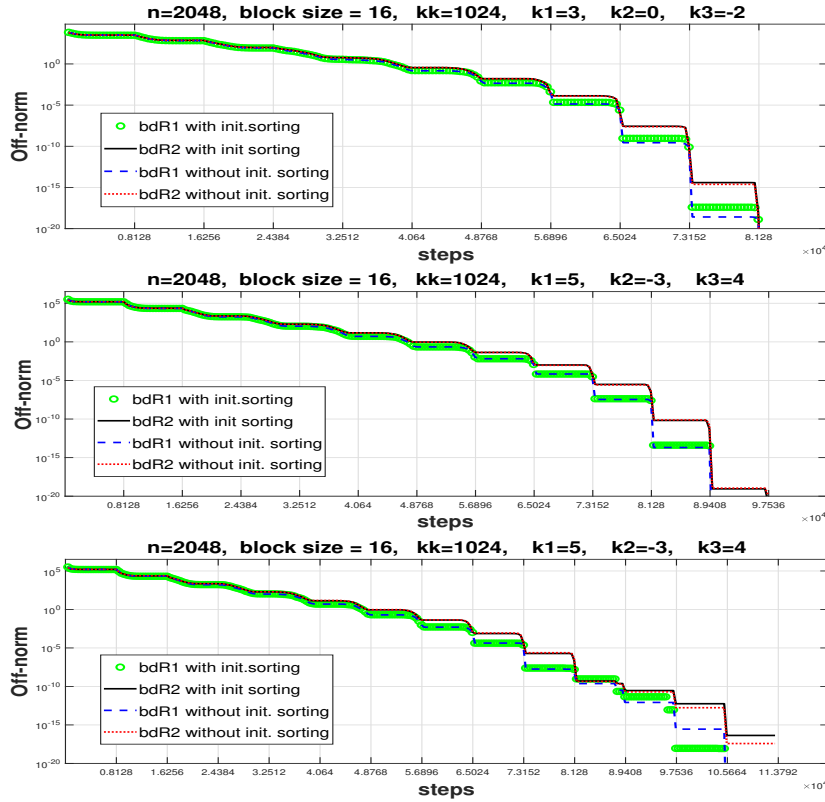
FIG. 7.3. *The graphs of the off-norm function.*

So, we have chosen the matrices associated with Figure 7.2, and we slightly modified our MATLAB functions. The output `offs` now contains $\sigma_{\min}(U_{i(k)i(k)}^{(k)})$ in the second column instead of the off-norm of the current matrix. To this end we also changed the value of `ndisp` from 20 to 100. Namely, only during the first two or three cycles, the value of $\sigma_{\min}(U_{i(k)i(k)}^{(k)})$ was not in close proximity of 1. So, we displayed only the graphs for the first three cycles. In Figure 7.4 are displayed the graphs of $\sigma_{\min}(U_{i(k)i(k)}^{(k)})$ for the matrix determined by the parameters `n = 1024`, `kk = 512`, `k1 = −3`, `k2 = 5`, `k3 = −5`.
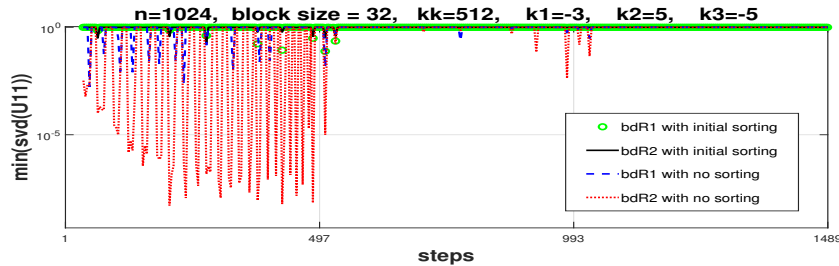

FIG. 7.4. *The graphs of the $\sigma_{\min}(U_{i(k)i(k)}^{(k)})$ function.*

TABLE 7.1
The minimum $\sigma_{\min}(U_{i(k)i(k)}^{(k)})$.

	bdR1-is	bdR2-is	bdR1-ns	bdR2-ns
$\min_k\{\sigma_{\min}(U_{i(k)i(k)}^{(k)})\}$	$2.771 \cdot 10^{-2}$	$5.777 \cdot 10^{-2}$	$6.462 \cdot 10^{-5}$	$4.633 \cdot 10^{-10}$

In Table 7.1 we present the minimum values of $\sigma_{\min}(U_{i(k)i(k)}^{(k)})$ taken over all steps of the process for each method. In the table, the headers *bdR1-is* (*bdR1-ns*) denotes the block Jacobi method under de Rijk strategy with the initial (i.e., with no) sorting. The same notation is used for the block methods with the *bdR2* strategy.

We can see that the initial ordering of the diagonal elements significantly raises $\min_k\{\sigma_{\min}(U_{i(k)i(k)}^{(k)})\}$. This implies that the block transformation matrices have small off-diagonal blocks, which has a beneficial impact on the iterative process. By this we mean the reduction in the total number of cycles and a slightly better accuracy of the output data. In addition it suggests that a modification of the core algorithm might not be necessary.

8. Conclusion and future work. This paper brings two main messages. First, it shows that the de Rijk pivot strategy most likely requires a smaller number of cycles than any cyclic strategy in the simple and block Jacobi method for the Hermitian eigenvalue problem. The price for this is an additional but altogether small number of swaps of columns and rows. This makes that pivot strategy a good candidate to be a part of the core algorithm in the block methods. Second, the problem of the global and quadratic convergence of the Jacobi methods for symmetric (complex Hermitian) matrices is almost completed. It yet remains to prove global convergence of the simple and block Jacobi method under a general cyclic strategy.

As for future work, it is tempting to prove similar results for the Jacobi methods that use other Jacobi-like methods or solve other eigenvalue problems, say, for the J-symmetric [59] and J-Hermitian Jacobi method [32, 33] or for the Cholesky–Jacobi [25, 27, 31] and Hari–Zimmermann [25, 26, 29, 46, 50] method, which are direct extensions of the Jacobi method. They solve the generalized eigenvalue problem.

REFERENCES

[1] M. BEČKA, G. OKŠA, AND M. VAJTERŠIC, *Dynamic ordering for a parallel block-Jacobi SVD algorithm*, Parallel Comput., 28 (2002), pp. 243–262.
[2] ———, *New dynamic orderings for the parallel one-sided block-Jacobi SVD algorithm*, Parallel Process. Lett., 25 (2015), Paper No. 1550003, 19 pages.
[3] E. BEGOVIĆ KOVAČ AND V. HARI, *Convergence of the complex block Jacobi methods under the generalized serial pivot strategies*, Linear Algebra Appl., 699 (2024), pp. 421–458.
[4] Z. BUJANOVIĆ AND Z. DRMAČ, *A contribution to the theory and practice of the block Kogbetliantz method for computing the SVD*, BIT Numer. Math., 52 (2012), pp. 827–849.
[5] J.-P. CHARLIER, M. VANBEGIN, AND P. VAN DOOREN, *On efficient implementations of Kogbetliantz's algorithm for computing the singular value decomposition*, Numer. Math., 52 (1988), pp. 279–300.
[6] J. DEMMEL AND K. VESELIĆ, *Jacobi's method is more accurate than QR*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 1204–1245.
[7] P. P. M. DE RIJK, *A one-sided Jacobi algorithm for computing the singular value decomposition on a vector computer*, SIAM J. Sci. Statist. Comput., 10 (1989), pp. 359–371.
[8] F. M. DOPICO, J. M. MOLERA, AND J. MORO, *An orthogonal high relative accuracy algorithm for the symmetric eigenproblem*, SIAM J. Matrix Anal. Appl., 25 (2003), pp. 301–351.
[9] Z. DRMAČ, *Computing the Singular and the Generalized Singular Values*, PhD. Thesis, University of Hagen, Hagen, 1994.

[10] ———, *Implementation of Jacobi rotations for accurate singular value computation in floating point arithmetic*, SIAM J. Sci. Comput., 18 (1997), pp. 1200–1222.

[11] ———, *A posteriori computation of the singular vectors in a preconditioned Jacobi SVD algorithm*, IMA J. Numer. Anal., 19 (1999), pp. 191–213.

[12] ———, *A global convergence proof for cyclic Jacobi methods with block rotations*, SIAM J. Matrix Anal. Appl., 31 (2009), pp. 1329–1350.

[13] Z. DRMAČ AND K. VESELIĆ, *New fast and accurate Jacobi SVD algorithm. I*, SIAM J. Matrix Anal. Appl., 29 (2007), pp. 1322–1342.

[14] ———, *New fast and accurate Jacobi SVD algorithm. II*, SIAM J. Matrix Anal. Appl., 29 (2007), pp. 1343–1362.

[15] P. J. EBERLEIN AND H. PARK, *Efficient implementation of Jacobi algorithms and Jacobi sets on distributed memory architectures*, J. Parallel Distrib. Comput., 8 (1990), pp. 358–366.

[16] K. V. FERNANDO, *Linear convergence of the row cyclic Jacobi and Kogbetliantz methods*, Numer. Math., 56 (1989), pp. 73–91.

[17] 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.

[18] E. R. HANSEN, *On cyclic Jacobi methods*, J. Soc. Indust. Appl. Math., 11 (1963), pp. 448–459.

[19] V. HARI, *On the quadratic convergence of Jacobi algorithms*, Rad. Mat., 2 (1986), pp. 127–146.

[20] ———, *On the quadratic convergence of the serial singular value decomposition Jacobi methods for triangular matrices*, SIAM J. Sci. Statist. Comput., 10 (1989), pp. 1076–1096.

[21] ———, *On pairs of almost diagonal matrices*, Linear Algebra Appl., 148 (1991), pp. 193–223.

[22] ———, *On sharp quadratic convergence bounds for the serial Jacobi methods*, Numer. Math., 60 (1991), pp. 375–406.

[23] ———, *Convergence of a block-oriented quasi-cyclic Jacobi method*, SIAM J. Matrix Anal. Appl., 29 (2007), pp. 349–369.

[24] ———, *Convergence to diagonal form of block Jacobi-type methods*, Numer. Math., 129 (2015), pp. 449–481.

[25] ———, *Globally convergent Jacobi methods for positive definite matrix pairs*, Numer. Algorithms, 79 (2018), pp. 221–249.

[26] ———, *On the global convergence of the complex HZ method*, SIAM J. Matrix Anal. Appl., 40 (2019), pp. 1291–1310.

[27] ———, *Complex Cholesky-Jacobi algorithm for PGEP*, in International Conference of Numerical Analysis and Applied Mathematics 2018, AIP Conf. Proc. 2116, AIP Publishing, Melville, 2019, Paper No. 450011, 4 pages.

[28] ———, *On the global convergence of the block Jacobi method for the positive definite generalized eigenvalue problem*, Calcolo, 58 (2021), Paper No. 24, 39 pages.

[29] ———, *On the quadratic convergence of the complex HZ method for the positive definite generalized eigenvalue problem*, Linear Algebra Appl., 632 (2022), pp. 153–192.

[30] V. HARI AND E. BEGOVIĆ KOVAČ, *Convergence of the cyclic and quasi-cyclic block Jacobi methods*, Electron. Trans. Numer. Anal., 46 (2017), pp. 107–147.
https://etna.ricam.oeaw.ac.at/vol.46.2017/pp107-147.dir/pp107-147.pdf

[31] ———, *On the convergence of complex Jacobi methods*, Linear Multilinear Algebra, 69 (2021), pp. 489–514.

[32] V. HARI, S. SINGER, AND S. SINGER, *Block-oriented J-Jacobi Methods for Hermitian matrices*, Linear Algebra Appl., 433 (2010), pp. 1491–1512.

[33] ———, *Full block J-Jacobi method for Hermitian matrices*, Linear Algebra Appl., 444 (2014), pp. 1–27.

[34] V. HARI AND K. VESELIĆ, *On Jacobi methods for singular value decompositions*, SIAM J. Sci. Statist. Comput., 8 (1987), pp. 741–754.

[35] V. HARI AND V. ZADELJ-MARTIĆ, *Parallelizing the Kogbetliantz method: a first attempt*, JNAIAM J. Numer. Anal. Ind. Appl. Math., 2 (2007), pp. 49–66.

[36] P. HENRICI AND K. ZIMMERMANN, *An estimate for the norms of certain cyclic Jacobi operators*, Linear Algebra Appl., 1 (1968), pp. 489–501.

[37] M. R. HESTENES, *Inversion of matrices by biorthogonalization and related results*, J. Soc. Indust. Appl. Math., 6 (1958), pp. 51–90.

[38] 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.

[39] E. KOGBETLIANTZ, *Diagonalization of general complex matrices as a new method for solution of linear equations*, in Proceedings of the International Congress of Mathematicians 1954, Vol. II, J. Gerretsen and J. de Groot, eds., North Holland, 1954, pp. 356–357

[40] ———, *Solution of linear equations by biagonalization of coeficient matrix*, Quart. Appl. Math., 13 (1955), pp. 123–132.

[41] F. LUK AND H. PARK, *A proof of convergence for two parallel Jacobi SVD algorithms*, IEEE Trans. Comput., 38 (1989), pp. 806–811.

[42] ———, *On parallel Jacobi orderings*, SIAM J. Sci. Statist. Comput., 10 (1989), pp. 18–26.

[43] W. F. MASCARENHAS, *On the convergence of the Jacobi method for arbitrary orderings*, SIAM J. Matrix Anal. Appl., 16 (1995), pp. 1197–1209.

[44] J. MATEJAŠ, *Accuracy of the Jacobi method on scaled diagonally dominant symmetric matrices*, SIAM J. Matrix Anal. Appl., 31 (2009), pp. 133–153.

[45] J. MATEJAŠ AND V. HARI, *On high relative accuracy of the Kogbetliantz method*, Linear Algebra Appl., 464 (2015), pp. 100–129.

[46] ———, *The high relative accuracy of the HZ method*, Appl. Math. Comput., 433 (2022), Paper No. 127358, 21 pages.

[47] R. MATHIAS, *Accurate eigensystem computations by Jacobi methods*, SIAM J. Matrix Anal. Appl., 16 (1995), pp. 977–1003.

[48] L. NAZARETH, *On the convergence of the cyclic Jacobi method*, Linear Algebra Appl., 12 (1975), pp. 151–164.

[49] V. NOVAKOVIĆ, *Accurate complex Jacobi rotations*, J. Comput. Appl. Math., 450 (2024), Paper No. 116003, 6 pages.

[50] 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.

[51] G. OKŠA, Y. YAMAMOTO, M. BEČKA, AND M. VAJTERŠIC, *Asymptotic quadratic convergence of the parallel block-Jacobi EVD algorithm with dynamic ordering for Hermitian matrices*, BIT Numer. Math., 58 (2018), pp. 1099–1123.

[52] G. OKŠA, Y. YAMAMOTO, AND M. VAJTERŠIC, *Asymptotic quadratic convergence of the serial block-Jacobi EVD algorithm for Hermitian matrices*, Numer. Math., 136 (2017), pp. 1071–1095.

[53] ———, *Convergence to singular triplets in the two-sided block-Jacobi SVD algorithm with dynamic ordering*, SIAM J. Matrix Anal. Appl., 43 (2022), pp. 1238–1262.

[54] 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.

[55] H. RUTISHAUSER, *The Jacobi method for real symmetric matrices*, in Handbook for Automatic Computation, Vol II: Linear Algebra, J. H. Wilkinson and C. Reinsch, eds., Springer, Berlin, 1971, pp. 202–211.

[56] A. H. SAMEH, *On Jacobi and Jacobi-like algorithms for a parallel computer*, Math. Comp., 25 (1971), pp. 579–590.

[57] A. SCHÖNHAGE, *Zur quadratischen Konvergenz des Jacobi-Verfahrens*, Numer. Math., 6 (1964), pp. 410–412.

[58] 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.

[59] K. VESELIĆ, *A Jacobi eigenreduction algorithm for definite matrix pairs*, Numer. Math., 64 (1993), pp. 241–269.

[60] W. VOEVODIN, *Cislennye Metody Linejnoj Algebry*, Nauka, Moscow, 1966.

[61] J. H. WILKINSON, *Note on the quadratic convergence of the cyclic Jacobi process*, Numer. Math., 4 (1962), pp. 296–300.

[62] ———, *Almost diagonal matrices with multiple or close eigenvalues*, Linear Algebra Appl., 1 (1968), pp. 1–12.

[63] Y. YAMAMOTO, Z. LAN, AND S. KUDO, *Convergence analysis of the parallel classical block Jacobi method for the symmetric eigenvalue problem*, JSIAM Lett., 6 (2014), pp. 57–60.

[64] Y. YAMAMOTO, G. OKŠA, AND M. VAJTERŠIC, *On convergence to eigenvalues and eigenvectors in the block-Jacobi EVD algorithm with dynamic ordering*, Linear Algebra Appl., 622 (2021), pp. 19–45.