Linear Algebra and Its Applications

# On the Quadratic Convergence of the Complex HZ Method for the Positive Definite Generalized Eigenvalue Problem

--Manuscript Draft--

| | |
|---|---|
| **Manuscript Number:** | |
| **Article Type:** | Regular Issue |
| **Keywords:** | positive definite generalized eigenvalue problem; Jacobi method; quadratic convergence |
| **Corresponding Author:** | Vjeran Hari, Ph.D.<br>University of Zagreb<br>Zagreb, CROATIA |
| **First Author:** | Vjeran Hari, Ph.D. |
| **Order of Authors:** | Vjeran Hari, Ph.D. |

# On the Quadratic Convergence of the Complex HZ Method for the Positive Definite Generalized Eigenvalue Problem[☆]

Vjeran Hari[a,*]

[a] *Department of Mathematics, Faculty of Science, University of Zagreb, Bijenička 30 10000 Zagreb, Croatia*

**Abstract**

The paper proves the quadratic convergence of the complex HZ method for solving the positive definite generalized eigenvalue problem. The proof is made for a general cyclic pivot strategy in the case of simple eigenvalues and for any wavefront pivot strategy in the case of simple or double eigenvalues. The proof is valid for the real HZ method. The preliminary numerical tests confirm the theoretical results.

*Keywords:* positive definite generalized eigenvalue problem, Jacobi method, quadratic convergence
*2010 MSC:* 65F15, 65G05

## 1. Introduction

In this paper we consider the asymptotic convergence of the complex HZ method [7, 17] for the positive definite generalized eigenvalue problem (PGEP)

$$Ax = \lambda Bx\,, \quad x \neq 0 \tag{1.1}$$

with full complex Hermitian matrices $A$, $B$ such that $B$ is positive definite.

On contemporary parallel computing machines the block Jacobi methods seem to be the best choice for solving the problem (1.1) with large matrices $A$ and $B$ (see [20, 24]). This is no surprise since the block Jacobi methods can be nicely adapted to the conditions required by the modern computational environments. In the core of each block method lies a kernel algorithm whose task is to diagonalize the block pivot submatrices $\hat{A}$, $\hat{B}$ at each step. The matrices $\hat{A}$, $\hat{B}$ are of smaller size, usually of order 64, 128 or 256, they are Hermitian and $\hat{B}$ is positive definite. The main requirement for the kernel

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

[*]Corresponding author
*Email address:* hari@math.hr (Vjeran Hari)

algorithm is to solve the PGEP with matrices $\hat{A}$, $\hat{B}$ accurately and efficiently. During computation the block pivot submatrices are most of the time nearly diagonal. The kernel algorithm should perform its task very fast and accurate on such matrices. These two requirements are well met by the element-wise Jacobi methods for the PGEP. So far, only three element-wise methods for the PGEP are known: the complex Falk-Langemeyer (FL) method [16, 4, 25], the complex Cholesky-Jacobi (CJ) method [15, 14] and the complex Hari-Zimmermann or shorter HZ method [7, 17]. As has been explained in [20], the HZ method seems to be more suitable to serve as a kernel algorithm than the FL method. The CJ method is pretty new and has been less researched.

The complex HZ method was derived and analyzed in [7]. Its global convergence has been proved in [17] under a large class of generalized serial strategies from [12]. The numerical tests indicate that the method might have an important property, the high relative accuracy, provided that both matrices $A$ and $B$ are positive definite and the spectral condition numbers of $\Delta_A A \Delta_A$ and $\Delta_B B \Delta_B$ are small for some diagonal matrices $\Delta_A$ and $\Delta_B$.

Here we prove that the convergence of the HZ method is asymptotically quadratic. The proof is made for a general cyclic pivot strategy and, with a better bound, for a wavefront strategy. The latter class of wavefront strategies [23] include the known serial pivot strategies. The results are an immediate generalization of the known results of Wilkinson [27] for the standard Jacobi method. The analysis presented here paves the way for the quadratic convergence proof of the block Jecobi method from [18] for the same problem.

The paper is divided into 8 sections. In Section 2 we describe the method, present its algorithm and define the notion of the quadratic convergence. In Section 3 we prove several auxiliary results that are needed for the quadratic convergence proof, which is given in Section 4. As an application, in Section 4 we briefly prove that the same result holds for the real HZ method from [13]. In Section 5 we present some numerical tests that confirm the theoretical results. The conclusions and proposals for future work are given in Section 6. Sections 7 is acknowledgements and Section 8 is an appendix where we have proved some lengthy and less important results.

## 2. Description of the Method

Let $A$ and $B$ be complex Hermitian matrices of order $n$ such that $B$ is positive definite. At the beginning of the process, the (complex) HZ method uses the congruence transformation:

$$A \mapsto A^{(0)} = DAD, \quad B \mapsto B^{(0)} = DBD, \quad D = \mathrm{diag}(B)^{-\frac{1}{2}}, \tag{2.1}$$

which makes the diagonal elements of $B^{(0)}$ equal to 1. Then $(A^{(0)}, B^{(0)})$ is taken as the starting matrix pair for the iterative process

$$A^{(k+1)} = Z_k^* A^{(k)} Z_k, \qquad B^{(k+1)} = Z_k^* B^{(k)} Z_k, \qquad k \geq 0, \tag{2.2}$$

where $A^{(0)}$ and $B^{(0)}$ are defined by (2.1). In (2.2) each transformation matrix $Z_k$ is an elementary plane matrix. It is a nonsingular matrix which differs from the identity matrix $I_n$ in two diagonal elements $z^{(k)}_{i(k)i(k)}$, $z^{(k)}_{j(k)j(k)}$ and two off-diagonal elements $z^{(k)}_{i(k)j(k)}$, $z^{(k)}_{j(k)i(k)}$, where $1 \le i(k) < j(k) \le n$. The subscripts $i = i(k)$, $j = j(k)$ are called *pivot indices*, $(i, j)$ is *pivot pair* and

$$\hat{Z}_k = \begin{bmatrix} z^{(k)}_{ii} & z^{(k)}_{ij} \\ z^{(k)}_{ji} & z^{(k)}_{jj} \end{bmatrix}, \quad k \ge 0, \tag{2.3}$$

is called *pivot submatrix* of $Z_k$. In MATLAB notation $\hat{Z}_k = Z_k([i\,j],[i\,j])$. If $\hat{Z}_k$ is as in (2.3), we shall briefly denote it by $\hat{Z}_k = (z^{(k)}_{ij})$. The transition $(A^{(k)}, B^{(k)}) \mapsto (A^{(k+1)}, B^{(k+1)})$ is called the $k$th *step* of the method.

The method is designed to preserve the unit diagonal of each iteration matrix $B^{(k)}$. This way each $B^{(k)}$ is almost optimally symmetrically scaled that can be obtained by a diagonal matrix. In other words, for the spectral condition number of $B^{(k)}$, we have (see [26]), $\kappa_2(B^{(k)}) \approx \min_\Delta \kappa_2(\Delta B^{(k)}\Delta)$, $\Delta$ is diagonal. We also have (see [17, Lemma 4.2(i)],[13, Lemma 4.1])

$$\|B^{(k)}\|_2 \le n, \quad \|A^{(k)}\|_2 \le n\mu, \quad k \ge 0; \qquad \mu = \max_{\lambda\in\sigma(A,B)} |\lambda|, \tag{2.4}$$

where $\mu$ is the spectral radius of the initial matrix pair $(A, B)$. Hence both sequences $(B^{(k)},\, k \ge 0)$ and $(A^{(k)},\, k \ge 0)$ are bounded.

### *2.1. The cyclic pivot strategies*

The way of selecting pivot pairs is called *pivot strategy*. A pivot strategy can be identified with the function $\mathsf{I} : \mathcal{N}_0 \to \mathcal{P}_n$, where $\mathcal{N}_0 = \{0, 1, 2, \ldots\}$, $\mathcal{P}_n = \{(r, t);\, 1 \le r < t \le n\}$. We see that $\mathcal{P}_n$ contains $N = n(n-1)/2$ pairs of indices. If $\mathsf{I}$ is a periodic function, then $\mathsf{I}$ is called *periodic pivot strategy*. Let $\mathsf{I}$ be the periodic (pivot) strategy with period $P$. If $P = N$ and $\{\mathsf{I}(k) : k = 0, 1, \ldots, P-1\} = \mathcal{P}_n$, then $\mathsf{I}$ is called *cyclic* strategy. If the HZ method is defined by some cyclic pivot strategy we speak of the cyclic HZ method.

For $t \ge 1$, the transition $(A^{((t-1)N)}, B^{((t-1)N)}) \mapsto (A^{(tN)}, B^{(tN)})$ is called the $t$th *cycle* or *sweep* of the method. The most common cyclic strategies are the row- and column-cyclic ones. In the row-cyclic strategy the pivot pair repeatedly runs through the sequence of $N = n(n-1)/2$ pairs:

$$(1,2), (1,3), \ldots, (1,n),\ (2,3), \ldots, (2,n),\ (3,4),\ \ldots\ , (n-1,n),$$

while in the column-cyclic strategy it runs through the sequence: $(1,2), (1,3)$, $(2,3), (1,4), (2,4), (3,4), \ldots$ ,$(1,n), (2,n),\ \ldots$,$(n-1,n)$. The common name for these two pivot strategies is *serial* strategies.

The set of serial pivot strategies can be enlarged to the set of *generalized serial strategies* which was introduced and discussed in [12] and later used in [13, 14, 17, 18]. This class encompasses the most important cyclic strategies, including the serial, *wavefront* and *weak wavefront* ones from [23] and [19].

The row-cyclic strategy can be upgraded to the strategy proposed by de Rijk [2]. Essentially, under that strategy the transformation linking the pair $(A^{(0)}, B^{(0)})$ with $(A^{(N)}, B^{(N)})$ has the form

$$PZ_{12}\cdots Z_{1n}I_{2,r_2}Z_{23}\cdots Z_{2n}I_{3,r_3}Z_{34}\cdots Z_{3n}\ \cdots\ I_{n-2,r_{n-2}}Z_{n-2,n-1}Z_{n-2,n}Z_{n-1,n},$$

where $P$ is the permutation matrix that orders the diagonal elements of $P^TA^{(0)}P$ in descending order and $I_{i,r_i}$ is the transposition matrix defined by $r_i$ which is determined by: $a_{r_i,r_i} = \max\{a_{tt};\ t \geq i\}$. As we shall see in Section 5, the de Rijk strategy has some advantages over the row-cyclic strategy.

*2.2. The global and quadratic convergence, the stopping criterion*

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

$$S(A,B) = \left[\mathsf{S}^2(A) + \mathsf{S}^2(B)\right]^{1/2}, \tag{2.5}$$

where generally, $\mathsf{S}(X) = \|X - \mathrm{diag}(X)\|_F$. Here $\|X\|_F = \sqrt{\mathrm{trace}(X^*X)}$ is the Frobenius norm of $X$.

The complex HZ method is *convergent on the pair* $(A, B)$ if the sequence of generated pairs satisfies $(A^{(k)}, B^{(k)}) \to (\Lambda, I_n)$ as $k \to \infty$. Here $I_n$ is the identity matrix and $\Lambda$ is a diagonal matrix of the eigenvalues of $(A, B)$. The method is *globally convergent* if it is convergent on every initial pair.

The global convergence of the HZ method under the generalized serial strategies has been proved in [17].

The cyclic method is asymptotically *quadratically convergent* on some set of matrix pairs if

$$S(A^{(N)}, B^{(N)}) \leq c_n S^2(A^{(0)}, B^{(0)})$$

holds for every matrix pair $(A, B)$ from that set. Such a set is characterized by some requirements, one of them is that $S(A^{(0)}, B^{(0)})$ is sufficiently small. Here $c_n$ is a constant which may depend on $n$.

If both matrices $A$ and $B$ are positive definite, one can stop the iteration if the current matrices satisfy the condition

$$|a_{rs}| \leq \mathrm{tol}\sqrt{|a_{rr}a_{ss}|}, \qquad |b_{rs}| \leq \mathrm{tol}, \qquad 1 \leq r < s \leq n. \tag{2.6}$$

This condition is usually checked at the end of each cycle. It warrants the high relative accuracy of the computed eigenvalues provided the initial pair of positive definite matrices is *well-behaved* [13] and if the method is proved to have the high relative accuracy property, as the numerical tests indicate. If $A$ is indefinite then we must rely on the values $S(A^{(tN)}, B^{(tN)})$, $t \geq 0$.

*2.3. The complex HZ algorithm*

By complex HZ algorithm we mean the algorithm that is used in one step of the complex HZ method. It computes the pivot submatrix $\hat{Z}_k$ and applies it to the appropriate rows and columns of the current matrices. The complex

HZ algorithm has been derived in [7, 17] and here we briefly describe it. We consider step $k$, omit the appearance of $k$ and denote the pivot indices by $i$, $j$.

The input to the algorithm is the pair $(\hat{A}, \hat{B})$ of pivot submatrices,

$$\hat{A} = \begin{bmatrix} a_{ii} & a_{ij} \\ \bar{a}_{ij} & a_{jj} \end{bmatrix}, \qquad \hat{B} = \begin{bmatrix} 1 & b_{ij} \\ \bar{b}_{ij} & 1 \end{bmatrix}, \qquad \begin{array}{c} a_{ij} = |a_{ij}|\, e^{\imath\alpha_{ij}}, \\ b_{ij} = |b_{ij}|\, e^{\imath\beta_{ij}}. \end{array} \tag{2.7}$$

The principal part of the output are the pivot submatrix $\hat{Z}$,

$$\hat{Z} = \frac{1}{\tau} \begin{bmatrix} \cos\phi & -e^{\imath\alpha}\sin\phi \\ e^{-\imath\beta}\sin\psi & \cos\psi \end{bmatrix} = \begin{bmatrix} c1 & -s1 \\ s2 & c2 \end{bmatrix}, \quad \tau = \sqrt{1-|b_{ij}|^2} \tag{2.8}$$

and $\hat{A}' = \hat{Z}^T \hat{A} \hat{Z}$.

Let $\sigma_{ij} = 1$ ($\sigma_{ij} = -1$) provided that $a_{ii} - a_{jj} \geq 0$ ($a_{ii} - a_{jj} < 0$),

$$u_{ij} + \imath v_{ij} = e^{-\imath\beta_{ij}} a_{ij}, \qquad u_{ij}, v_{ij} \in \mathbf{R}, \tag{2.9}$$

$$\gamma_{ij} = \arg\left(\frac{a_{ii} - a_{jj}}{2} + \imath v_{ij}\right) + (1 - \sigma_{ij})\frac{\pi}{2}, \qquad -\frac{\pi}{2} \leq \gamma_{ij} \leq \frac{\pi}{2}. \tag{2.10}$$

The elements of $\hat{Z}$ are computed from the expressions:

$$\begin{array}{rcl} 2\cos^2\phi & = & 1 - |b_{ij}|\sin(2\theta) + \sqrt{1-|b_{ij}|^2}\cos(2\theta)\cos(\gamma_{ij}), \\ 2\cos^2\psi & = & 1 + |b_{ij}|\sin(2\theta) + \sqrt{1-|b_{ij}|^2}\cos(2\theta)\cos(\gamma_{ij}), \\ e^{\imath\alpha}\sin\phi & = & \frac{e^{\imath\beta_{ij}}}{2\cos\psi}[\sin(2\theta) + |b_{ij}| + \imath\sqrt{1-|b_{ij}|^2}\cos(2\theta)\sin(\gamma_{ij})], \\ e^{-\imath\beta}\sin\psi & = & \frac{e^{-\imath\beta_{ij}}}{2\cos\phi}[\sin(2\theta) - |b_{ij}| - \imath\sqrt{1-|b_{ij}|^2}\cos(2\theta)\sin(\gamma_{ij})], \end{array} \tag{2.11}$$

where $\phi\,,\, \psi \in [0, \pi/2]$. For the angles $\theta$ and $\gamma_{ij}$, we have

$$\tan(2\theta) = \sigma_{ij}\frac{2u_{ij} - (a_{ii} + a_{jj})|b_{ij}|}{\sqrt{1-|b_{ij}|^2}\sqrt{(a_{ii} - a_{jj})^2 + 4v_{ij}^2}}, \quad -\frac{\pi}{4} \leq \theta \leq \frac{\pi}{4}, \tag{2.12}$$

$$\cos(\gamma_{ij}) = \frac{|a_{ii} - a_{jj}|}{\sqrt{(a_{ii} - a_{jj})^2 + 4v_{ij}^2}}, \qquad \sin(\gamma_{ij}) = \sigma\frac{2v_{ij}}{\sqrt{(a_{ii} - a_{jj})^2 + 4v_{ij}^2}}. \tag{2.13}$$

For the diagonal elements, we have

$$a'_{ii} = [\cos^2\phi a_{ii} + \sin^2\psi a_{jj} + 2\cos\phi\sin\psi\,\Re(e^{-\imath\beta}a_{ij})]/(1-|b_{ij}|^2), \tag{2.14}$$

$$a'_{jj} = [\sin^2\phi a_{ii} + \cos^2\psi a_{jj} - 2\cos\psi\sin\phi\,\Re(e^{-\imath\alpha}a_{ij})]/(1-|b_{ij}|^2). \tag{2.15}$$

One can show that in floating point arithmetic, the diagonal elements of $\hat{B}'$ are computed with tiny relative errors while $b'_{ij}$ is computed as zero. This does not apply to $a'_{ij}$, which can be computed by the formula (we use: $c_\varphi = \cos\varphi$, $s_\varphi = \sin\varphi$, $\varphi \in \{\phi\,,\, \psi\}$):

$$a'_{ij} = [c_\phi c_\psi a_{ij} - \bar{a}_{ij}e^{\imath(\alpha+\beta)}s_\phi s_\psi + (a_{jj}e^{\imath\beta}c_\psi s_\psi - a_{ii}e^{\imath\alpha}c_\phi s_\phi)]/(1-|b_{ij}|^2).$$

If $\tan(2\theta)$ has the form $0/0$ then the pivot submatrices are proportional: $\hat{A} = a_{ii}\hat{B}$ (see [17]). From the relation (2.9) we see that $v_{ij} = 0$, $u_{ij} = a_{ii}|b_{ij}|$ and *we choose* $\theta = 0$, $\gamma_{ij} = 0$. Then $\hat{Z}$ reduces to the form

$$\hat{Z} = \frac{1}{\tau}\begin{bmatrix} \rho & -\xi \\ -\bar{\xi} & \rho \end{bmatrix}, \quad \xi = \frac{b_{ij}}{2\rho}, \quad \rho = \frac{\sqrt{1+|b_{ij}|}+\sqrt{1-|b_{ij}|}}{2}. \tag{2.16}$$

The matrix $\hat{Z}$ from (2.16) is a direct extension of the real one from [13, Section 2.3]. In this case we have $a'_{ii} = a_{ii}$ and $a'_{jj} = a_{jj}$.

If $b_{ij} = 0$ and $a_{ij} \neq 0$ then in the above formulas $\arg(b_{ij})$ is replaced by $\arg(a_{ij})$. Hence $\hat{Z}$ is reduced to the complex Jacobi rotation for $\hat{A}$. If in addition $a_{ij} = 0$, then $u_{ij} = v_{ij} = \gamma_{ij} = \theta = 0$, hence $Z$ is reduced to the identity matrix.

In the pseudocode below, $\Re(\omega)$, $\Im(\omega)$ and $\mathrm{conj}(\omega)$ denote the real, imaginary part and complex conjugate of $\omega$. The names of variables in the pseudocode are similar to mathematical notation. Thus, $t2$, $cs2$, $sn2$, $csg$, $sng$ stand for $\tan(2\theta)$, $\cos(2\theta)$, $\sin(2\theta)$, $\cos(\gamma_{ij})$, $\sin(\gamma_{ij})$, respectively.

**%%% The complex HZ algorithm**

`select the pivot pair` $(i,j)$
`if` $a_{ij} \neq 0$ `or` $b_{ij} \neq 0$
  $b = \mathtt{abs}(b_{ij})$; `if` $b = 0$, $eb = a_{ij}/\mathtt{abs}(a_{ij})$; $u = \mathtt{abs}(a_{ij})$; $v = 0$;
  `else`, $eb = b_{ij}/b$; $d = \mathtt{conj}(b_{ij})/b \cdot a_{ij}$; $u = \Re(d)$; $v = \Im(d)$;
  `endif`;
  $e = a_{ii} - a_{jj}$; $\sigma = 1$; `if` $e < 0$, $\sigma = -1$ `endif`;
  $\tau = \sqrt{(1-b)\cdot(1+b)}$; $csg = |e|/\sqrt{e^2+4v^2}$; $sng = \sigma \cdot 2v/\sqrt{e^2+4v^2}$;
  `if` $\mathtt{abs}(2\cdot u - (a_{ii}+a_{jj})\cdot b) = 0$, $sn2 = 0$; $cs2 = 1$;
  `elseif` $\mathtt{abs}(e) + \mathtt{abs}(v) = 0$, $sn2 = 1$; $cs2 = 0$;
  `else`, $t2 = \sigma \cdot (2\cdot u - (a_{ii}+a_{jj})\cdot b)/\sqrt{(e^2+4v^2)\cdot(1-b)\cdot(1+b)}$;
    $cs2 = 1/\sqrt{1+t2^2}$; $sn2 = t2/\sqrt{1+t2^2}$;
  `endif`;
  $c1 = \sqrt{(1+(\tau\cdot cs2\cdot csg - b\cdot sn2))/(2\cdot(1-b)\cdot(1+b))}$;
  $c2 = \sqrt{(1+(\tau\cdot cs2\cdot csg + b\cdot sn2))/(2\cdot(1-b)\cdot(1+b))}$;
  $s1 = eb\cdot(sn2 + b + \imath\,\tau\cdot cs2\cdot sng)/(2\cdot c2\cdot(1-b)\cdot(1+b))$;
  $s2 = \mathtt{conj}(eb)\cdot(sn2 - b - \imath\,\tau\cdot cs2\cdot sng)/(2\cdot c1\cdot(1-b)\cdot(1+b))$;
  $a'_{ii} = c1^2\cdot a_{ii} + |s2|^2\cdot a_{jj} + 2\cdot c1\cdot\Re(s2\cdot a_{ij})$;
  $a'_{jj} = |s1|^2\cdot a_{ii} + c2^2\cdot a_{jj} - 2\cdot c2\cdot\Re(\mathtt{conj}(s1)\cdot a_{ij})$;
  $a'_{ij} = c1\cdot c2\cdot a_{ij} - s1\cdot\mathtt{conj}(s2\cdot a_{ij}) + (c2\cdot a_{jj}\cdot\mathtt{conj}(s2) - c1\cdot a_{ii}\cdot s1)$;
  $a'_{ji} = \mathtt{conj}(a'_{ij})$; $b'_{ij} = 0$; $b'_{ji} = 0$;;
  `for` $k = 1,\ldots,n$, $k \neq i,j$ `do`
    $a'_{ki} = c1\cdot a_{ki} + s2\cdot a_{kj}$; $b'_{ki} = c1\cdot b_{ki} + s2\cdot b_{kj}$;
    $a'_{ik} = \mathtt{conj}(a'_{ki})$; $b'_{ik} = \mathtt{conj}(b'_{ki})$;
    $a'_{kj} = c2\cdot a_{kj} - s1\cdot a_{ki}$; $b'_{kj} = c2\cdot b_{kj} - s1\cdot b_{ki}$;
    $a'_{jk} = \mathtt{conj}(a'_{kj})$; $b'_{jk} = \mathtt{conj}(b'_{kj})$;
  `endfor`
`endif`

Finally, if the eigenvectors are wanted, one can set $F^{(0)} = D$, where $D$ is from the relation (2.1), and in each step make the update: $F^{(k+1)} = F^{(k)} Z_k$. In the case of convergence, after stopping the process, the columns of $F^{(k)}$ will be good approximations of the eigenvectors of the initial pair $(A, B)$.

## 3. Some Auxiliary Results

Here we prove several results that are needed in the quadratic convergence proof of the HZ method. The first subsection sheds some light on the special structure that lies in the nearly diagonal $A^{(0)}$ and $B^{(0)}$, when the pair $(A, B)$ has multiple eigenvalues (see [8, 10]). In the second subsection we prove several lemmas that are needed in the quadratic convergence proof.

### *3.1. Nearly diagonal matrices $A^{(0)}$ and $B^{(0)}$*

Let $A$, $B$ be Hermitian matrices of order $n$ such that $B$ is positive definite. Let the eigenvalues of the pair $(A, B)$ be nonincreasingly ordered,

$$\lambda_1 = \cdots = \lambda_{s_1} > \lambda_{s_1+1} = \cdots = \lambda_{s_2} > \ \cdots \ > \lambda_{s_{p-1}+1} = \cdots = \lambda_{s_p}. \tag{3.1}$$

The case $p = 1$ implies $A = \lambda_1 B$. Then every nonzero vector is an eigenvector belonging to the only eigenvalue $\lambda_1$. So, let $p > 1$.

If we set $s_0 = 0$, we conclude from the relation (3.1) that $n_r = s_r - s_{r-1}$ is the multiplicity of $\lambda_{s_r}$. Let $\lambda_{s_0} = \lambda_0 = \infty$, $\lambda_{s_{p+1}} = -\infty$ and

$$3\delta_t = \min\{\lambda_{s_{t-1}} - \lambda_{s_t}, \lambda_{s_t} - \lambda_{s_{t+1}}\}, \quad 1 \le t \le p. \tag{3.2}$$

We see that $3\delta_t$ is the absolute gap in the spectrum of $(A, B)$ associated with $\lambda_{s_t}$. Let

$$\delta = \min_{1 \le t \le p} \delta_t, \qquad \delta_0 = \frac{\delta}{1 + \mu^2}, \tag{3.3}$$

where $\mu$ is the spectral radius of $(A, B)$. Obviously, $3\delta$ is the minimum absolute gap and for $\delta_0$ we have

$$\delta_0 = \frac{\delta}{1 + \mu^2} \le \frac{\delta}{2\mu} \le \frac{1}{3}. \tag{3.4}$$

Indeed, if $p > 1$ then the worst possible bound for $\delta/(2\mu)$ is obtained when $p = 2$ and $\mu = \lambda_1 = -\lambda_p$. Then $3\delta = 2\mu$. If $B$ has unit diagonal then we have

$$|a_{rr}| = \frac{|e_r^T A e_r|}{e_r^T B e_r} \le \max_{\|x\|_2 = 1} \frac{|x^* A x|}{x^* B x} = \mu\,, \quad 1 \le r \le n. \tag{3.5}$$

Here the last equality sign can serve as definition of $\mu$. The simpler definition, $\mu = \max_{\lambda \in \sigma(A,B)} |\lambda|$, where $\sigma(A, B)$ is the spectrum of the matrix pair $(A, B)$, is given in the relation (2.4).

In the convergence analysis we shall use the following result from [7, Corollary 3.3] or [8, Corollary 3.3].

**Lemma 3.1.** *Let $A$, $B$ be Hermitian matrices of order $n$ such that $B$ is positive definite with unit diagonal. Let the eigenvalues of $(A, B)$ be ordered as in the relation (3.1) and let $\delta$, $\delta_0$ be as in the relation (3.3). If*

$$\sqrt{1+\mu^2}S(A,B) < \delta, \tag{3.6}$$

*then there is a permutation matrix $P$ such that for the matrix $\tilde{A} = P^T AP = (\tilde{a}_{rt})$ we have*

$$2\sum_{l=1}^{n} |\tilde{a}_{ll} - \lambda_l|^2 \leq \frac{S^4(A,B)}{\delta_0^2}.$$

In Lemma 3.1, the condition $\sqrt{1+\mu^2}S(A,B) < \delta$ can be replaced by the simpler but also the stricter one, $S(A,B) < \delta_0$.

Although Lemma 3.1 is sufficient to prove the quadratic convergence of the HZ method in the case of simple eigenvalues, we shall need a more general result for the case of double eigenvalues. To this end, let us assume that (3.6) holds, and let us partition both matrices $\tilde{A} = P^T AP$ and $\tilde{B} = P^T BP$ in accordance with the multiplicities $n_1,\ldots,n_p$:

$$\tilde{A} = \begin{bmatrix} \tilde{A}_{11} & \cdots & \tilde{A}_{1p} \\ \vdots & \ddots & \vdots \\ \tilde{A}_{p1} & \cdots & \tilde{A}_{pp} \end{bmatrix}, \qquad \tilde{B} = \begin{bmatrix} \tilde{B}_{11} & \cdots & \tilde{B}_{1p} \\ \vdots & \ddots & \vdots \\ \tilde{B}_{p1} & \cdots & \tilde{B}_{pp} \end{bmatrix}. \tag{3.7}$$

In the block-matrix partition (3.7), $\tilde{A}_{rr}$ and $\tilde{B}_{rr}$ have dimension $n_r$, $1 \leq r \leq p$. With the partition (3.7) we associate quantity

$$\mathcal{T}(\tilde{A},\tilde{B}) = [\mathsf{T}^2(\tilde{A}) + \mathsf{T}^2(\tilde{B})]^{1/2} \tag{3.8}$$

where for any square matrix $X$ partitioned in accordance with the partition $(n_1,\ldots,n_p)$ of $n$, $\mathsf{T}(X) = \|X - \text{diag}\,(X_{11},\ldots,X_{pp})\|_F$.

**Theorem 3.2.** *[8, Theorem 3.1,Corollary 3.2] Let $A$, $B$ be Hermitian matrices of order $n$ such that $B$ is positive definite with unit diagonal. Let the eigenvalues of $(A, B)$ be ordered as in the relation (3.1) and let $\delta$, $\delta_0$ be as in the relation (3.3). If the condition (3.6) holds then there is a permutation matrix $P$ such that for the matrices $\tilde{A} = P^T AP = (\tilde{A}_{rt})$ and $\tilde{B} = P^T BP = (\tilde{B}_{rt})$, partitioned as in (3.7), we have*

$$\begin{aligned} \|\tilde{A}_{rr} - \lambda_{s_r}\tilde{B}_{rr}\|_F &\leq \frac{1}{\delta_r}\sum_{\substack{t=1\\t\neq r}}^{p} \|\tilde{A}_{rt} - \lambda_{s_r}\tilde{B}_{rt}\|_F^2 \leq \frac{1+\lambda_{s_r}^2}{\delta_r}\sum_{\substack{t=1\\t\neq r}}^{p}(\|\tilde{A}_{rt}\|_F^2 + \|\tilde{B}_{rt}\|_F^2) \\ &\leq \frac{1+\lambda_{s_r}^2}{\delta_r}\frac{\mathcal{T}^2(\tilde{A},\tilde{B})}{2} \leq \frac{1+\lambda_{s_r}^2}{\delta_r}\frac{S^2(A,B)}{2}, \quad 1 \leq r \leq p, \end{aligned} \tag{3.9}$$

$$2\sum_{r=1}^{p} \|\tilde{A}_{rr} - \lambda_{s_r}\tilde{B}_{rr}\|_F^2 \leq \frac{(1+\mu^2)^2\mathcal{T}^4(\tilde{A},\tilde{B})}{\delta^2} = \frac{\mathcal{T}^4(\tilde{A},\tilde{B})}{\delta_0^2} \leq \frac{S^4(A,B)}{\delta_0^2}. \tag{3.10}$$

Let us link the above results to the HZ method. The statements of Lemma 3.1 and Theorem 3.2 can be used in connection with the matrices $(A^{(k)}, B^{(k)})$ whenever $S(A^{(k)}, B^{(k)})$ satisfies the condition (3.6) (recall that each $B^{(k)}$ has unit diagonal). By slightly strengthening the condition (3.6) for the starting pair $(A^{(0)}, B^{(0)})$, we can assure that (3.6) holds during the whole cycle. To make the quadratic convergence proof easier, we shall also need one additional condition on $\mathsf{S}(B^{(0)})$. These two conditions will be sufficient for the quadratic convergence proof. We shall refer to them as *asymptotic assumptions*. They are given below:

$$\mathsf{S}(B^{(0)}) < \frac{1}{2N}, \qquad n \geq 3, \tag{3.11}$$

$$S(A^{(0)}, B^{(0)}) < \frac{\delta}{2\sqrt{1+\mu^2}}, \qquad p \geq 2. \tag{3.12}$$

In the following, we assume that conditions (3.11) and (3.12) hold for $(A^{(0)}, B^{(0)})$. We shall consider the first cycle of the HZ method. To simplify exposition, we shall use the following notation for $0 \leq k \leq N$:

$$\left.\begin{array}{l} a_k = |a^{(k)}_{i(k)j(k)}|, \quad b_k = |b^{(k)}_{i(k)j(k)}|, \quad b_{\max} = \max\limits_{0\leq k\leq N} b_k, \\ \tau_k = \sqrt{1-b_k^2}, \quad x_k = 1/(1-b_k), \quad y_k = 1/(1-b_k^2), \\ e_k = a^{(k)}_{i(k)i(k)} - a^{(k)}_{j(k)j(k)}, \qquad \sigma_k = \left\{ \begin{array}{c} 1 \ \text{ if } \ e_k \geq 0 \\ -1 \ \text{ if } \ e_k < 0 \end{array} \right. \\ u_k + \imath v_k = e^{-\imath \arg(b^{(k)}_{i(k)j(k)})} \cdot a^{(k)}_{i(k)j(k)}, \quad \gamma_k = \gamma_{i(k)j(k)} \\ \epsilon_k = S(A^{(k)}, B^{(k)}), \qquad \mathcal{T}_k = \mathcal{T}(A^{(k)}, B^{(k)}). \end{array}\right\} \tag{3.13}$$

*3.2. The preliminary results*

Here we prove several lemmas. Several lengthier proofs are moved to Section 8 which is an appendix. The first lemma is not directly connected to the method.

**Lemma 3.3.** *Let $r$ be an integer and $w$ a nonnegative real number such that $r \geq 3$, $2rw < 1$. Then*

*(i)* $(1-w)^{-r} \leq 1 + \frac{12}{7} rw$ *(ii)* $(1-w^2)^{-r} \leq 1 + \frac{72}{67} rw^2$

*(iii)* $(1+w)^{r} \leq 1 + \frac{4}{3} rw$ *(iv)* $(1-w)^{-1/2} \leq 1 + \frac{3}{5} w.$

**Proof.** (i) Let $(1-w)^{-1} = 1+\zeta$. Since $rw < 1/2$, $w < 1/6$, $\zeta = w(1-w)^{-1} < 1/5$, $r\zeta \leq (6/5)rw < 3/5$ and $(1-w)^{-r} = (1+\zeta)^r$, we have

$$\begin{aligned} (1-w)^{-r} &= 1 + r\zeta \left[ 1 + \frac{r-1}{2}\zeta + \frac{(r-1)(r-2)}{2\cdot 3}\zeta^2 + \cdots + \zeta^{r-2} \right] + \zeta^r \\ &\leq 1 + r\zeta \left[ 1 + \left(\frac{r-1}{2}\zeta\right) + + \left(\frac{r-1}{2}\zeta\right)^2 + \cdots \right] = 1 + \frac{r\zeta}{1 - \frac{r-1}{2}\zeta} \\ &\leq 1 + \frac{6}{5}\frac{rw}{1-\frac{3}{5}rw} \leq 1 + \frac{12}{7} rw. \end{aligned}$$

The assertion (iii) is proved in an almost identical way while the proof of (ii) uses the substitution $(1-w^2)^{-1} = 1+\zeta'$ and the inequalities $r\zeta' \le (36/35)rw^2$, $r\zeta' < 6/70$. In the proof of (iv) one can use the expansion of the function $(1-w)^{-1/2}$. Another way is to square the left and right side of that inequality, use $1/(1-w) = 1 + w/(1-w)$ cancel out 1 and if $w > 0$, divide by $w$. □

**Lemma 3.4.** *Let $a_k$, $b_k$, $x_k$, $\epsilon_k$ be defined by the relation (3.13). Then*

$$\epsilon_{k+1}^2 \le x_k \left[\epsilon_k^2 - 2(a_k^2 + b_k^2)\right], \qquad k \ge 0.$$

**Proof.** For a given $k$, let $i = i(k)$, $j = j(k)$,

$$\begin{aligned}
a_{i*} &= [a_{i1}^{(k)}, a_{i2}^{(k)}, \ldots, a_{i,i-1}^{(k)}, a_{i,i+1}^{(k)}, \ldots, a_{i,j-1}^{(k)}, a_{i,j+1}^{(k)}, \ldots, a_{in}^{(k)}],\\
a_{j*} &= [a_{j1}^{(k)}, a_{j2}^{(k)}, \ldots, a_{j,i-1}^{(k)}, a_{j,i+1}^{(k)}, \ldots, a_{j,j-1}^{(k)}, a_{j,j+1}^{(k)}, \ldots, a_{jn}^{(k)}],\\
a_{*i} &= [a_{1i}^{(k)}, a_{2i}^{(k)}, \ldots, a_{i-1,i}^{(k)}, a_{i+1,i}^{(k)}, \ldots, a_{j-1,i}^{(k)}, a_{j+1,i}^{(k)}, \ldots, a_{ni}^{(k)}]^T\\
a_{*j} &= [a_{1j}^{(k)}, a_{2j}^{(k)}, \ldots, a_{i-1,j}^{(k)}, a_{i+1,j}^{(k)}, \ldots, a_{j-1,j}^{(k)}, a_{j+1,j}^{(k)}, \ldots, a_{nj}^{(k)}]^T
\end{aligned}$$

where generally, $c^T$ is the transpose of $c$. Let $a'_{i*}$, $a'_{j*}$, $a'_{*i}$, $a'_{*j}$ be the row-and column-vectors built in the same way, but from the elements of $A^{(k+1)}$. The transformation (2.2) implies

$$\begin{bmatrix} a'_{i*} \\ a'_{j*} \end{bmatrix} = \hat{Z}_k^* \begin{bmatrix} a_{i*} \\ a_{j*} \end{bmatrix}, \qquad \begin{bmatrix} a'_{*i} & a'_{*j} \end{bmatrix} = \begin{bmatrix} a_{*i} & a_{*j} \end{bmatrix} \hat{Z}_k.$$

We obtain

$$\left\| \begin{bmatrix} a'_{i*} \\ a'_{j*} \end{bmatrix} \right\|_F^2 \le \|\hat{Z}_k^*\|_2^2 \left\| \begin{bmatrix} a_{i*} \\ a_{j*} \end{bmatrix} \right\|_F^2, \quad \left\| \begin{bmatrix} a'_{*i} & a'_{*j} \end{bmatrix} \right\|_F^2 \le \|\hat{Z}_k\|_2^2 \left\| \begin{bmatrix} a_{*i} & a_{*j} \end{bmatrix} \right\|_F^2.$$

Since $\|\hat{Z}_k^*\|_2^2 = \|\hat{Z}_k\|_2^2 = 1/(1-b_k) = x_k$ (see [17]), we have

$$\begin{aligned}
\mathsf{S}^2(A^{(k+1)}) &\le \mathsf{S}^2(A^{(k)}) - 2a_k^2 + \Big(\|\hat{Z}_k\|_2^2 - 1)(\|a_{i*}\|_F^2 + \|a_{j*}\|_F^2 \\
&\quad + \|a_{*i}\|_F^2 + \|a_{*j}\|_F^2\Big) \le x_k(\mathsf{S}^2(A^{(k)}) - 2a_k^2). \qquad (3.14)
\end{aligned}$$

The same analysis applies to $B^{(k)}$, so we have

$$\mathsf{S}^2(B^{(k+1)}) \le x_k(\mathsf{S}^2(B^{(k)}) - 2b_k^2). \qquad (3.15)$$

Since $B^{(k)}$ has unit diagonal, we have $\epsilon_l^2 = (\mathsf{S}^2(A^{(l)}) + \mathsf{S}^2(B^{(l)})$ for $l = k, k+1$. By adding the inequalities (3.14) and (3.15) we obtain the assertion of the lemma. □

**Lemma 3.5.** *Let $a_k$, $b_k$, $b_{max}$, $x_k$, $\epsilon_k$ be defined by the relation (3.13). If the assumptions (3.11) and (3.12) hold, then under any pivot strategy, we have*

$$\begin{aligned}
\sum_{k=0}^{N-1}(a_k^2+b_k^2) &\leq 0.876914\,\epsilon_0^2, && (3.16)\\
b_{max} &< 0.8795\,\frac{1}{2N}, && (3.17)\\
\begin{bmatrix} \mathsf{S}^2(A^{(k)}) \\ \mathsf{S}^2(B^{(k)}) \\ \epsilon_k^2 \end{bmatrix} &\leq x_0\ \cdots\ x_{k-1} \begin{bmatrix} \mathsf{S}^2(A^{(0)}) \\ \mathsf{S}^2(B^{(0)}) \\ \epsilon_0^2 \end{bmatrix}, \quad 1\leq k\leq N, && (3.18)\\
x_0\ \cdots\ x_{k-1} &< (1-b_{max})^{-N} < 1.75383, \quad 1\leq k\leq N. && (3.19)
\end{aligned}$$

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

**Lemma 3.6.** *Let the assumptions (3.11) and (3.12) hold and let $\delta$ be as in the relation (3.3). Then for each pair of indices $(r,t)$, $1\leq r<t\leq n$, and for each $k$ such that $0\leq k\leq N$, we have:*

$$\textit{either}\quad |a_{rr}^{(k)}-a_{tt}^{(k)}|>2.56154\,\delta \qquad \textit{or}\quad |a_{rr}^{(k)}-a_{tt}^{(k)}|<0.43846\,\delta. \tag{3.20}$$

*The relation (3.20) holds under any pivot strategy.*

**Proof.** By the assumption (3.12) and the assertions (3.18) and (3.19) of Lemma 3.5 we have

$$(1+\mu^2)\epsilon_k^2 \leq 1.75383(1+\mu^2)\,\epsilon_0^2 < \frac{1.75383\,\delta^2}{4} < 0.43846\,\delta^2,\ \ 0\leq k\leq N. \tag{3.21}$$

From the relation (3.21) we see that for each $0\leq k\leq N$, Lemma 3.1 can be applied to the matrix pair $(A^{(k)},B^{(k)})$. We obtain

$$2\sum_{r=0}^{n}|a_{rr}^{(k)}-\lambda_r^{(k)}|^2 \leq \frac{\epsilon_k^4}{\delta_0^2} < (0.43846\delta)^2, \qquad 0\leq k\leq N, \tag{3.22}$$

where $\lambda_1^{(k)},\ \ldots\ ,\lambda_n^{(k)}$ is an ordering of the eigenvalues depending on $k$. If $\lambda_r^{(k)}\neq\lambda_t^{(k)}$ then by the definition of $\delta$ (see (3.2), (3.3)) and by (3.22), we have

$$\begin{aligned}
|a_{rr}^{(k)}-a_{tt}^{k)}| &\geq |\lambda_r^{(k)}-\lambda_t^{(k)}|-|a_{rr}^{(k)}-\lambda_r^{(k)}|-|\lambda_t^{(k)}-a_{tt}^{(k)}| && (3.23)\\
&\geq 3\delta-\sqrt{2|a_{rr}^{(k)}-\lambda_r^{(k)}|^2+2|a_{tt}^{(k)}-\lambda_t^{(k)}|^2} > 2.56154\delta.
\end{aligned}$$

This proves the first part of the assertion (3.20).

If $\lambda_r^{(k)}=\lambda_t^{(k)}$ then using (3.22), we have

$$\begin{aligned}
|a_{rr}^{(k)}-a_{tt}^{(k)}| &= |a_{rr}^{(k)}-\lambda_r^{(k)}+\lambda_t^{(k)}-a_{tt}^{(k)}| \leq |a_{rr}^{(k)}-\lambda_r^{(k)}|+|\lambda_t^{(k)}-a_{tt}^{(k)}|\\
&\leq \sqrt{2|a_{rr}^{(k)}-\lambda_r^{(k)}|^2+2|a_{tt}^{(k)}-\lambda_t^{(k)}|^2} < 0.43846\,\delta,
\end{aligned}$$

which proves the second part of the assertion (3.20). □

From Lemma 3.6 we see that the set

$$\mathcal{S}' = \left\{ k \in \{0, 1, \ldots, N-1\} \,;\, |a^{(k)}_{i(k)i(k)} - a^{(k)}_{j(k)j(k)}| > 2\delta \right\} \tag{3.24}$$

is well defined for any pivot strategy, provided the assumptions (3.11) and (3.12) hold. For simplicity, we use the notation $\sum'_k$ instead of $\sum_{k,\, k\in\mathcal{S}}$.

**Lemma 3.7.** *Let the assumptions (3.11) and (3.12) hold and let $\phi_k$, $\psi_k$, $\epsilon_k$, $\delta$ and $\mu$ be defined by the relations (2.11), (3.13), (3.3) and (3.5). Then*

$$\sum_{k=0}^{N-1}{}' \sin^2 \omega_k \leq 0.60583283(1+\mu^2)\frac{\epsilon_0^2}{\delta^2}, \tag{3.25}$$

*where $\omega_k \in \{\phi_k, \psi_k\}$, $0 \leq k \leq N-1$. The relation (3.25) holds under any pivot strategy. If all eigenvalues $\lambda_i$ are simple then the sum $\sum'_k$ reduces to the usual sum $\sum_k$ and the constant* 0.60583283 *can be replaced by* 0.4736138.

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

To simplify notation, let us now assume that besides (3.11) and (3.12) we also have

$$a^{(0)}_{11} \geq a^{(0)}_{22} \geq \ \ldots \ \geq \ a^{(0)}_{nn}. \tag{3.26}$$

Then the assertions (3.9) and (3.10) of Theorem 3.2 hold for the partition (3.7) of $A^{(0)}$ and $B^{(0)}$. The question arises whether those estimates hold for the first $N$ steps under any pivot strategy. The following lemma states that this is true provided the condition (3.12) is modified to be more stringent.

**Lemma 3.8.** *Let the assumptions (3.11), (3.12) and (3.26) hold. If*

$$\epsilon_0^2 < \frac{\delta_0}{\mu+1}\delta^2, \tag{3.27}$$

*where $\epsilon_0$, $\mu$ and $\delta_0$, $\delta$ are defined by the relations (3.13), (3.5) and (3.3), respectively, then*

$$\|A^{(k)}_{rr} - \lambda_{s_r} B^{(k)}_{rr}\|_F \ \leq \ \frac{1}{\delta_r}\sum_{\substack{t=1\\ t\neq r}}^{p} \|A^{(k)}_{rt} - \lambda_{s_r} B^{(k)}_{rt}\|_F^2, \qquad 1 \leq r \leq p \tag{3.28}$$

*and*

$$2\sum_{r=1}^{p} \|A^{(k)}_{rr} - \lambda_{s_r} B^{(k)}_{rr}\|_F^2 \leq \frac{(1+\mu^2)^2 \mathcal{T}_k^4}{\delta^2} = \frac{\mathcal{T}_k^4}{\delta_0^2} \leq \frac{\epsilon_k^4}{\delta_0^2} \tag{3.29}$$

*hold for every $0 \leq k \leq N$. In the relations (3.28) and (3.29), $\lambda_{s_r}$, $\delta_r$, $\mathcal{T}_k$, $\epsilon_k$ are defined by the relations (3.1), (3.2), (3.13), and the matrices $A^{(k)}$, $B^{(k)}$ are partitioned in accordance with the relation (3.7). The both assertions hold under any pivot strategy.*

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

## 4. The Quadratic Convergence Proof

Here we prove the quadratic convergence of the cyclic HZ method in the case of simple eigenvalues of the pair $(A, B)$. We consider matrix pairs $(A^{(0)}, B^{(0)})$, $(A^{(1)}, B^{(1)}),\dots,(A^{(N)}, B^{(N)})$ and assume that $S(A^{(0)}, B^{(0)})$ is sufficiently small. Thus, we have $p = n$, $3\delta = \min_{r\neq t} |\lambda_r - \lambda_t|$ and $\mathcal{T}_k = \epsilon_k$, $k \geq 0$.

**Theorem 4.1.** *Let the assumptions (3.11), (3.12) hold for the pair $(A^{(0)}, B^{(0)})$ and let the sequence of pairs $((A^{(k)}, B^{(k)}),\ k \geq 0)$ be generated by a cyclic HZ method defined by relations (2.7)–(2.16). If the eigenvalues of the pair $(A^{(0)}, B^{(0)})$ are simple, then*

$$\epsilon_N \leq \sqrt{N(1+\mu^2)}\,\frac{\epsilon_0^2}{\delta}. \tag{4.1}$$

*Moreover, if the pivot strategy is row-cyclic, then*

$$\epsilon_N \leq \sqrt{1+\mu^2}\,\frac{\epsilon_0^2}{\delta}. \tag{4.2}$$

**Proof.** The proof uses the technique developed by J.H. Wilkinson in [27]. We first prove that (4.1) holds for an arbitrary cyclic strategy. Since the transformation of the elements of $A^{(k)}$ use the same formulas as those of $B^{(k)}$, we shall pay our attention only to the elements of $A^{(k)}$, $0 \leq k \leq N$.

For a fixed $k$, $0 \leq k \leq N-1$, the pivot indices $i = i(k)$ and $j = j(k)$ are also fixed. Consider the elements $a_{ij}^{(r)}$, $r = k+1, \dots, N$. Note that $a_{ij}^{(k+1)} = 0$ and after that step $a_{ij}^{(r)}$ changes at most $2(n-2)$ times. By $r_1, \dots, r_s$ $(s \leq 2n-4)$ denote those values of $r$ for which $a_{ij}^{(r)}$ changes in the $r$th step. For simplicity we set $h_t = a_{ij}^{(r_t+1)}$, $0 \leq t \leq s$, where $r_0 = k$ and $h_0 = 0$. Then the relations (2.2), (2.8) and (3.13) imply

$$\left.\begin{aligned} h_1 &= \sqrt{y_{r_1}}\left(0\cdot\cos(\omega'_{r_1}) \pm a^{(r_1)}e^{\imath\nu_1}\sin(\omega_{r_1})\right)\\ h_2 &= \sqrt{y_{r_2}}\left(h_1\cdot\cos(\omega'_{r_2}) \pm a^{(r_2)}e^{\imath\nu_2}\sin(\omega_{r_2})\right)\\ &\vdots\\ h_t &= \sqrt{y_{r_t}}\left(h_{t-1}\cdot\cos(\omega'_{r_t}) \pm a^{(r_t)}e^{\imath\nu_t}\sin(\omega_{r_t})\right) \end{aligned}\right\} \tag{4.3}$$

where $\omega'_{r_t}, \omega_{r_t} \in \{\phi_{r_t}, \psi_{r_t}\}$, $\nu_t \in \{\alpha_{r_t}, -\alpha_{r_t}, \beta_{r_t}, -\beta_{r_t}\}$, while $a^{(r_t)}$ is a certain off-diagonal element of $A^{(r_t)}$. From (4.3) we obtain for $1 \leq t \leq s$,

$$|h_t| \leq \sum_{l=1}^{t} \sqrt{y_{r_l}\cdots y_{r_t}}\,|a^{(r_l)}|\,|\sin(\omega_{r_l})| \leq (1\text{–}b_{\max}^2)^{-t/2}\sum_{l=1}^{t}|a^{(r_l)}|\,|\sin(\omega_{r_l})|. \tag{4.4}$$

Set

$$A^{(k)} = D_A^{(k)} + E^{(k)}, \qquad D_A^{(k)} = \mathrm{diag}(a_{11}^{(k)}, \dots, a_{nn}^{(k)}), \qquad k \geq 0.$$

The matrix $E^{(N)}$ consists exactly of the elements $h_s$. Note that here $s$ is a function of the pivot pair $(i,j)$ (hence of $k$) and the cyclic pivot strategy under consideration. From the relation (4.4) we conclude that

$$\begin{aligned} |E^{(N)}| &\leq (1-b_{\max}^2)^{-(n-2)} \Big( |P^{(1)}|\,|\sin(\omega_1)| + |P^{(2)}|\,|\sin(\omega_2)| + \cdots \\ &\quad + |P^{(N-1)}|\,|\sin(\omega_{N-1})| \Big), \end{aligned} \tag{4.5}$$

where each matrix $P^{(k)}$ contains nonzero elements only at those positions of the $i$'th and $j$'th row and column which have already been pivot positions. The nonzero elements of $P^{(k)}$ are certain elements of $E^{(k)}$ belonging to the $i$'th and $j$'th row and column. Here we use notation $|C| = (|c_{rt}|)$ where $C = (c_{rt})$ is an arbitrary matrix.

By the assertions (3.18) and (3.19) of Lemma 3.5 we obtain

$$\| \, |P^{(k)}| \, \|_F = \|P^{(k)}\|_F \leq \mathsf{S}(A^{(k)}) \leq \sqrt{1.75383}\, \mathsf{S}(A^{(0)}), \quad 1 \leq k \leq N-1. \tag{4.6}$$

Since $n \geq 3$, by Lemma 3.3(ii) and the assertion (3.17) of Lemma 3.5, we obtain

$$(1-b_{\max}^2)^{-(n-2)} \leq \left[(1-b_{\max}^2)^{-2N}\right]^{1/n} \leq \left[1 + \frac{72}{67}\frac{0.8795^2}{3\cdot 2}\right]^{1/3} < 1.0442. \tag{4.7}$$

Finally, using relations (4.5)–(4.7) and the second assertion of Lemma 3.7, we obtain

$$\begin{aligned} \mathsf{S}(A^{(N)}) &= \|E^{(N)}\|_F = \| \, |E^{(N)}| \, \|_F \leq 1.0442\sqrt{1.75383}\, \mathsf{S}(A^{(0)}) \sum_{k=1}^{N-1} |\sin(\omega_k)| \\ &\leq 1.0442\sqrt{1.75383}\, \mathsf{S}(A^{(0)}) \left[ (N-1) \sum_{k=1}^{N-1} \sin^2(\omega_k) \right]^{1/2} \\ &\leq 1.0442\sqrt{1.75383 \cdot 0.4736138}\sqrt{(N-1)(1+\mu^2)}\, \frac{\epsilon_0}{\delta} \mathsf{S}(A^{(0)}) \\ &\leq 0.95168\sqrt{N(1+\mu^2)}\, \frac{\epsilon_0}{\delta} \mathsf{S}(A^{(0)}). \end{aligned} \tag{4.8}$$

The same analysis applies to the matrices $B^{(k)}$, $0 \leq k \leq N$, yielding the same bound connecting $\mathsf{S}(B^{(N)})$ and $\mathsf{S}(B^{(0)})$. Therefore, the first assertion (4.1) follows from (4.8) and the definition of $\epsilon_k$ (see (2.5) and (3.13)).

To prove the assertion (4.2) we apply the analysis that was used in proving the relation (4.3). Recall that by our notation, for $2 \leq t \leq n$, the $(1,t)$-element is annihilated in step $k = t-2$. Thus $a_{1t}^{(t-1)} = 0$ and we consider how that

$(1,t)$-element changes in the next $n-t$ steps. We have

$$\begin{aligned} a_{1t}^{(t)} &= \sqrt{y_{t-1}}\left(0\cdot\cos(\phi_{t-1}) + a_{t+1,t}^{(t-1)} e^{\imath\beta_{t-1}}\sin(\psi_{t-1})\right) \\ a_{1t}^{(t+1)} &= \sqrt{y_t}\ \left(a_{1t}^{(t)}\cdot\cos(\phi_t) + a_{t+2,t}^{(t)} e^{\imath\beta_t}\sin(\psi_t)\right) \\ &\vdots \\ a_{1t}^{(n-1)} &= \sqrt{y_{n-2}}\left(a_{1t}^{(n-2)}\cdot\cos(\phi_{n-2}) + a_{n,t}^{(n-2)} e^{\imath\beta_{n-2}}\sin(\psi_{n-2})\right). \end{aligned}$$

Therefore, for the elements of the first row we have

$$|a_{1t}^{(n-1)}| \le \sum_{r=t}^{n-1} \sqrt{y_{r-1}y_r\cdots y_{n-2}}|\sin(\psi_{r-1})|\,|a_{r+1,t}^{(r-1)}|, \qquad 2\le t\le n-1.$$

Using the Cauchy-Schwarz inequality, we obtain for $2\le t\le n-1$,

$$\begin{aligned} |a_{1t}^{(n-1)}|^2 &\le \sum_{r=t}^{n-1}|a_{r+1,t}^{(r-1)}|^2 \sum_{r=t}^{n-1} y_{r-1}y_r\cdots y_{n-2}\sin^2(\psi_{r-1}) \\ &= \sum_{r=t+1}^{n}|a_{rt}^{(r-2)}|^2 \sum_{r=t-1}^{n-2} y_r y_{r+1}\cdots y_{n-2}\sin^2(\psi_r). \end{aligned} \tag{4.9}$$

Since $a_{1n}^{(n-1)}=0$, the relation (4.9) implies

$$\sum_{t=2}^{n}|a_{1t}^{(n-1)}|^2 \le \left[\sum_{t=2}^{n-1}\sum_{r=t+1}^{n}|a_{rt}^{(r-2)}|^2\right]\sum_{r=1}^{n-2} y_r y_{r+1}\cdots y_{n-2}\sin^2(\psi_r). \tag{4.10}$$

Let us estimate the sum in the brackets. Recall that each $A^{(k)}$ is Hermitian. Hence

$$|a_{rt}^{(r-2)}|^2 = |a_{tr}^{(r-2)}|^2, \quad t<r.$$

Note that $a_{tr}^{(r-2)}$ is the element at $(t,r)$ position just prior to the transformation that annihilates $(1,r)$-element. Therefore its value is the same as of $a_{tr}^{(t-1)}$. Hence, we have

$$\Gamma = \sum_{t=2}^{n-1}\sum_{r=t+1}^{n}|a_{rt}^{(r-2)}|^2 = \sum_{t=2}^{n-1}\sum_{r=t+1}^{n}|a_{tr}^{(t-1)}|^2 = \sum_{r=2}^{n-1}\sum_{t=r+1}^{n}|a_{rt}^{(r-1)}|^2, \tag{4.11}$$

where we switched the indices $t\leftrightarrow r$. To bound $\Gamma$, we use the inequality

$$\sum_{t=r+1}^{n}|a_{1t}^{(r-1)}|^2 + \sum_{t=r+1}^{n}|a_{rt}^{(r-1)}|^2 \le x_{r-2}\sum_{t=r+1}^{n}|a_{1t}^{(r-2)}|^2 + x_{r-2}\sum_{t=r+1}^{n}|a_{rt}^{(0)}|^2, \tag{4.12}$$

which holds for $2\le r\le n-1$. Namely, we have $\|\hat{Z}_{r-2}^*\|_2^2 = 1/(1-b_{r-2}) = x_{r-2}$, $2\le r\le n-1$ (cf. the proof of Lemma 3.4). For $r=2,3,\ldots,n-2$ let us multiply

(4.12) by $x_{n-3}x_{n-4}\cdots x_{r-1}$. We obtain

$$x_{n-3}\cdots x_{r-1}\sum_{t=r+1}^{n}|a_{1t}^{(r-1)}|^2 \;+\; x_{n-3}\cdots x_{r-1}\sum_{t=r+1}^{n}|a_{rt}^{(r-1)}|^2$$
$$\le x_{n-3}\cdots x_{r-2}\sum_{t=r+1}^{n}|a_{1t}^{(r-2)}|^2 + x_{n-3}\cdots x_{r-2}\sum_{t=r+1}^{n}|a_{rt}^{(0)}|^2, \quad 2\le r\le n-2,$$

For $r=n-1$, the last inequality in (4.12) is left unchanged:

$$|a_{1n}^{(n-2)}|^2+|a_{n-1,n}^{(n-2)}|^2 \le x_{n-3}|a_{1n}^{(n-3)}|^2 + x_{n-3}|a_{n-1,n}^{(n-3)}|^2.$$

Let us sum so obtained inequalities. After cancelation, we obtain

$$|a_{n-1,n}^{(n-2)}|^2 + x_{n-3}\sum_{t=n-1}^{n}|a_{n-2,t}^{(n-3)}|^2 + x_{n-3}x_{n-4}\sum_{t=n-2}^{n}|a_{n-3,t}^{(n-4)}|^2+\cdots$$
$$+x_{n-3}\cdots x_1\sum_{t=3}^{n}|a_{2t}^{(1)}|^2 \;+\; z \;\le x_{n-3}\cdots x_0\left(\sum_{t=3}^{n}|a_{1t}^{(0)}|^2+\sum_{t=3}^{n}|a_{2t}^{(0)}|^2\right)$$
$$+\; x_{n-3}\cdots x_1\sum_{t=4}^{n}|a_{3t}^{(0)}|^2+\cdots+x_{n-3}|a_{n-1,n}^{(0)}|^2, \tag{4.13}$$

where $z$,

$$z = x_1\cdots x_{n-3}|a_{13}^{(1)}|^2 + x_2\cdots x_{n-3}|a_{14}^{(2)}|^2 + \;\cdots\; x_{n-3}|a_{1,n-1}^{(n-3)}|^2+|a_{1,n}^{(n-2)}|^2$$

is a nonnegative quantity. Since all $x_k$, $0\le k\le n-3$ are not smaller than 1, we see that $\Gamma$ from the relation (4.11) is not larger than the left side of the inequality (4.13) without $z$. Hence it is not larger than the right side of (4.13). Therefore, we have

$$\Gamma \le x_0\cdots x_{n-3}\frac{\mathsf{S}^2(A^{(0)})}{2}. \tag{4.14}$$

Combining relations (4.10), (4.11) and (4.14), we have

$$\sum_{t=2}^{n}|a_{1t}^{(n-1)}|^2 \le \;x_0\cdots x_{n-3}\cdot y_1\cdots y_{n-2}\,\frac{\mathsf{S}^2(A^{(0)})}{2}\sum_{r=1}^{n-1}\sin^2(\psi_{r-1}). \tag{4.15}$$

During later transformations the elements of the first row can increase by modulus. Since $\|\hat{Z}_k\|_2=\sqrt{x_k}$, it is easy to show that for $2\le i(k)<j(k)\le n$, we have

$$|a_{1,i(k)}^{(k+1)}|^2+|a_{1,j(k)}^{(k+1)}|^2 \le x_k(|a_{1,i(k)}^{(k)}|^2+|a_{1,j(k)}^{(k)}|^2) \le \frac{1}{1-b_{\max}}(|a_{1,i(k)}^{(k)}|^2+|a_{1,j(k)}^{(k)}|^2).$$

Furthermore, we know that after step $k=n-1$ each element of the first row (except of the $(1,2)$-element) will change exactly $n-2$ times. Hence, the latest

relation and relation (4.15) imply

$$\begin{aligned}\sum_{t=2}^{n}|a_{1t}^{(N)}|^2 &\leq \frac{1}{(1-b_{\max})^{n-2}}\sum_{t=2}^{n}|a_{1t}^{(n-1)}|^2\\ &\leq \frac{\mathsf{S}^2(A^{(0)})}{2(1-b_{\max})^{2(n-2)}(1-b_{\max}^2)^{n-2}}\sum_{r=1}^{n-2}\sin^2(\psi_r). \qquad (4.16)\end{aligned}$$

We can make the same analysis for other rows. To this end it is convenient to temporarily denote the angle $\psi_k$ which is used in the annihilation of the $(i,j)$-element by $\psi_{ij}$. Then for the second row we have

$$\begin{aligned}\sum_{t=3}^{n}|a_{2t}^{(N)}|^2 &\leq \frac{\mathsf{S}^2(A^{(n-1)})}{2(1-b_{\max})^{2(n-3)}(1-b_{\max}^2)^{n-3}}\sum_{r=4}^{n}\sin^2(\psi_{2r})\\ &\leq \frac{x_0x_1\cdots x_{n-2}\,\mathsf{S}^2(A^{(0)})}{2(1-b_{\max})^{2(n-3)}(1-b_{\max}^2)^{n-3}}\sum_{r=4}^{n}\sin^2(\psi_{2r}). \qquad (4.17)\end{aligned}$$

Here, we used assertion (3.18) of Lemma 3.5. More generally, for $1\leq r\leq n-2$ we have

$$\begin{aligned}\sum_{t=r+1}^{n}|a_{rt}^{(N)}|^2 &\leq \frac{(1-b_{\max})^{n-(n+n-1+\cdots+n-r+1)}\,\mathsf{S}^2(A^{(0)})}{2(1-b_{\max})^{2(n-(r+1))}(1-b_{\max}^2)^{n-(r+1)}}\sum_{s=r+2}^{n}\sin^2(\psi_{rs})\\ &\leq \frac{1.75383\,\mathsf{S}^2(A^{(0)})}{2(1-b_{\max}^2)^{n-(r+1)}}\sum_{s=r+2}^{n}\sin^2(\psi_{rs}), \qquad (4.18)\end{aligned}$$

where we used assertions (3.18) and (3.19) of Lemma 3.5 and the fact that

$$n-1+\cdots+n-r+1+2(n-(r+1)<n-1+\cdots+n-r+1+n-r+n-r-1\leq N.$$

Summing up the inequalities in (4.18) for $r=1,2,\ldots n-2$ and taking into account $a_{n-1,n}^{(N)}=0$, we obtain

$$\frac{1}{2}\mathsf{S}^2(A^{(N)})\leq\frac{1.75383\,\mathsf{S}^2(A^{(0)})}{2(1-b_{\max}^2)^{n-2}}\sum_{r=1}^{n-2}\sum_{s=r+2}^{n}\sin^2(\psi_{rs}). \qquad (4.19)$$

Using yet Lemma 3.7 (the second assertion), we have

$$\mathsf{S}^2(A^{(N)})\leq\frac{1.75383\,\mathsf{S}^2(A^{(0)})}{(1-b_{\max}^2)^{n-2}}0.4736138(1+\mu^2)\frac{\epsilon_0^2}{\delta^2}. \qquad (4.20)$$

Note that relation (3.17) implies $2(n-2)b_{\max}<2\cdot 0.8795/n<1$. Hence the assertion (ii) of Lemma 3.3 can be applied to obtain

$$\begin{aligned}(1-b_{\max}^2)^{-(n-2)} &\leq 1+\frac{72}{67}(n-2){b_{\max}}^2<1+\frac{72}{67}(n-2)\frac{0.8795^2}{n^2(n-1)^2}\\ &< 1+\frac{72}{67}\frac{0.8795^2}{9\cdot 2}<1.04618032. \qquad (4.21)\end{aligned}$$

Combining the relations (4.21) and (4.20) we obtain

$$\mathsf{S}^2(A^{(N)}) \leq 0.869(1+\mu^2)\frac{\epsilon_0^2}{\delta^2}\mathsf{S}^2(A^{(0)}).$$

The same analysis applies to matrices $B^{(k)}$, $0 \leq k \leq N$ and it yields the same bound connecting $\mathsf{S}(B^{(N)})$ and $\mathsf{S}(B^{(0)})$. Hence, summing up the obtained inequalities, we have

$$\epsilon_N \leq \sqrt{0.869(1+\mu^2)}\frac{\epsilon_0^2}{\delta} \leq 0.932201\sqrt{1+\mu^2}\frac{\epsilon_0^2}{\delta}, \tag{4.22}$$

which proves the theorem. □

The estimates (4.1) and (4.2) are quite analogous to the known ones for the standard Jacobi method for symmetric matrices obtained by Wilkinson [27]. The factor $\sqrt{1+\mu^2}$, which does not appear in the estimates for the standard Jacobi method, originates from the presence of the sum $a_{ii}^{(k)} + a_{jj}^{(k)}$ in the numerator of the ratio defining $\tan(2\theta_k)$. The assumption (3.12) is approximately $\sqrt{1+\mu^2}$ times stronger than the assumption in [27].

By inspecting the quadratic convergence proof of the row-cyclic HZ method, especially the derivation of the relation (4.19), one can see that the angles $\psi_{r,r+1}$, $r = 1, \ldots, n-1$, do not appear in the bound. This fact can be used to prove the quadratic convergence if the eigenvalues have multiplicity at most 2. However, the asymptotic assumption (3.12) has to be replaced by a more stringent one, so that during iteration the diagonal elements converging to the same double eigenvalue remain in adjacent positions on the diagonal.

**Corollary 4.2.** *Let the assumptions (3.11), (3.12), (3.27) hold. Let the eigenvalues of of the pair $(A^{(0)}, B^{(0)})$ be at most double. If the condition*

$$|a_{rr}^{(0)} - a_{ss}^{(0)}| < \delta \quad \Rightarrow \quad s \in \{r-1, r, r+1\} \tag{4.23}$$

*holds, then the row-cyclic HZ method is quadratically convergent and the relation*

$$\epsilon_N \leq 1.055\sqrt{1+\mu^2}\,\frac{\epsilon_0^2}{\delta}. \tag{4.24}$$

*holds. The estimate (4.24), as well as (4.22) in the case of simple eigenvalues, remain to hold if the pivot strategy is any wavefront strategy or the de Rijk strategy.*

**Proof.** The condition (4.23) together with relations (3.12), (3.27), (3.11) and Lemma 3.8 ensure that the diagonal elements affiliated with double eigenvalues take adjacent positions on the diagonal of $A^{(0)}$ and remain such during the whole cycle. This then ensures that the relation (4.19) holds. Now, using Lemma 3.7 (this time with the constant 0.6058327 instead of 0.4736138), we obtain

$$\epsilon_N \leq \sqrt{1.75383 \cdot 0.6058327 \cdot 1.04618032\,(1+\mu^2)}\frac{\epsilon_0^2}{\delta} \leq 1.0543223\sqrt{1+\mu^2}\frac{\epsilon_0^2}{\delta},$$

which proves the bound in (4.24).

For the second assertion, it suffices to show that matrices $A^{(N)}$ and $B^{(N)}$ obtained from $A^{(0)}$ and $B^{(0)}$ using the row-cyclic pivot strategy are the same as those obtained by any wavefront strategy. For the column-cyclic strategy the proof can be found in [7, 3.7 Lemma]. It is based on the proof from [6]. For a general wavefront strategy, the proof is combinatorial and almost identical to that from [23]. The only difference comes from the fact that the transformation matrices are not orthogonal, which is irrelevant for the proof.

Finally, for the de Rijk strategy, one can check whether the main relations in the proof, like relations (4.15), (4.16), (4.17) and (4.18) remain to hold. They do because the transposition $I_{s,r_s}$ does not change the sum of squares of elements in row $l$, $l < i$. □

If $\delta$ is tiny due to a pair of very close eigenvalues then the estimates (4.1) and (4.2) imply that $\epsilon_N$ is not "essentially smaller" than $\epsilon_0$. The following result implies that in such a situation, certain off-diagonal elements of $A^{(N)}$ and $B^{(N)}$ are still essentially smaller than $\epsilon_0$. To simplify notation, we introduce vector $\tilde{\boldsymbol{\delta}}$,

$$\tilde{\boldsymbol{\delta}} = [\tilde{\delta}_1, \ldots, \tilde{\delta}_n] = [\delta_1, \ldots, \delta_1, \ \ldots \ , \delta_p, \ldots, \delta_p]. \tag{4.25}$$

For each $r$, $1 \le r \le p$, $\delta_r$ is defined by the relation (3.2) and it appears in $\tilde{\boldsymbol{\delta}}$ exactly $n_r$ times.

**Corollary 4.3.** *Let the assumptions (3.11), (3.12), (3.26) and (3.27) hold and let the eigenvalues of the pair $(A^{(0)}, B^{(0)})$ be at most double. Let the sequence of pairs $((A^{(k)}, B^{(k)}), k \ge 0)$ be generated by the HZ method under some wavefront strategy. Then we have*

$$\sum_{t=r+1}^{n} \left( |a_{rt}^{(N)}|^2 + |b_{rt}^{(N)}|^2 \right) \le 0.5558 \frac{1+\mu^2}{\tilde{\delta}_r^2} \epsilon_0^4, \qquad 1 \le r \le n-1,$$

*where $\tilde{\delta}_r$ are defined by the relation (4.25). If all eigenvalues are simple than the constant* 0.5558 *can be replaced by* 0.4345.

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

If any of the conditions in Theorem 4.1 fails to hold, one can easily find a matrix pair for which the quadratic convergence fails. The following example sheds light to the failure of the quadratic convergence of the HZ method under the row-cyclic strategy provided that some eigenvalue has multiplicity larger than 2. The analysis for the de Rijk strategy is the same.

**Example 4.4.** *Let $n = 5$ and let the eigenvalues of the initial matrix pair $(A, B)$ satisfy $\lambda_1 > \lambda_2 = \lambda_3 = \lambda_4 > \lambda_5$. Using the notation from (3.1), we have $s_1 = 1$, $s_2 = 4$ and $s_3 = 5$. Suppose the row-cyclic HZ method is applied to $(A, B)$ and we stop the process in step $k$ when the conditions (3.11), (3.12) and (3.27) are met. Then we apply the transformation $(A^{(k)}, B^{(k)}) \mapsto (P^T A^{(k)} P, P^T B^{(k)} P)$ where the permutation matrix $P$ is chosen to order the*

*diagonal elements of $P^TA^{(k)}P$ nonincreasingly. We reset the step counter, so that $(A^{(0)},B^{(0)}) = (P^TA^{(k)}P, P^TB^{(k)}P)$. Then we consider one cycle of the row-cyclic HZ method applied to $(A^{(0)},B^{(0)})$.*

*After the first 5 steps we obtain $(A^{(5)},B^{(5)})$. We consider the subsequent step ($k=6$) and apply the qualitative analysis which uses the asymptotic notation with big O symbol. To simplify notation, we set $A^{(5)} = (a_{rt})$, $B^{(5)} = (b_{rt})$, $A^{(6)} = (a'_{rt})$, $B^{(6)} = (b'_{rt})$ and $\varepsilon = S(A^{(5)},B^{(5)})$. We have*

$$A^{(5)} = \left[\begin{array}{c|ccc|c} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ \hline \bar{a}_{12} & a_{22} & 0 & a_{24} & a_{25} \\ \bar{a}_{13} & 0 & a_{33} & a_{34} & a_{35} \\ \bar{a}_{14} & \bar{a}_{24} & \bar{a}_{34} & a_{44} & a_{45} \\ \hline \bar{a}_{15} & \bar{a}_{25} & \bar{a}_{35} & \bar{a}_{45} & \bar{a}_{55} \end{array}\right], \quad B^{(5)} = \left[\begin{array}{c|ccc|c} 1 & b_{12} & b_{13} & b_{14} & b_{15} \\ \hline \bar{b}_{12} & 1 & 0 & b_{24} & b_{25} \\ \bar{b}_{13} & 0 & 1 & b_{34} & b_{35} \\ \bar{b}_{14} & \bar{b}_{24} & \bar{b}_{34} & 1 & b_{45} \\ \hline b_{15} & b_{25} & b_{35} & b_{45} & 1 \end{array}\right].$$

*From the proof of Theorem 4.1 we know that $\|[a_{12}\ a_{13}\ a_{14}\ a_{15}]\|_2 = O(\varepsilon^2)$, $\|[b_{12}\ b_{13}\ b_{14}\ b_{15}]\|_2 = O(\varepsilon^2)$. Other off-diagonal elements of both matrices are generally equal to $O(\varepsilon)$. In particular, we assume $|a_{24}| + |b_{24}| = O(\varepsilon)$ and $|a_{34}| + |b_{34}| = O(\varepsilon)$.*

*To simplify exposition, let $\lambda_{s_1} = 3$, $\lambda_{s_2} = 2$, $\lambda_{s_3} = 1$, so that $\delta_1 = \delta_2 = \delta_3 = \delta = 1$. By Theorem 3.2 we have*

$$\| \left[\begin{array}{ccc} a_{22} & 0 & a_{24} \\ 0 & a_{33} & a_{34} \\ \bar{a}_{24} & \bar{a}_{34} & a_{44} \end{array}\right] - 2 \left[\begin{array}{ccc} 1 & 0 & b_{24} \\ 0 & 1 & b_{34} \\ \bar{b}_{24} & \bar{b}_{34} & 1 \end{array}\right] \| = 5(\|[a_{12}\ a_{13}\ a_{14}\ a_{15}]\|^2$$
$$+ \|[b_{12}\ b_{13}\ b_{14}\ b_{15}]\|^2 + \|[a_{25}\ a_{35}\ a_{45}]\|^2 + \|[b_{25}\ b_{35}\ b_{45}]\|^2) = O(\varepsilon^2). \quad (4.26)$$

*In step 6 the pivot elements are $a_{24}$ and $b_{24}$. From relations (2.9)–(2.13) and (4.26) we obtain*

$$\begin{aligned} u_{24} + \imath v_{24} &= \frac{a_{24}\bar{b}_{24}}{|b_{24}|} = \frac{(2b_{24} + \alpha_{24})\bar{b}_{24}}{|b_{24}|} = 2|b_{24}| + \alpha'_{24}, \\ \tau_{24} &= \sqrt{1 - |b_{24}|^2} = 1 - O(\varepsilon^2), \\ |\tan(2\theta_{24})| &= \frac{|2u_{24} - (a_{22} + a_{44})|b_{24}|\,|}{\tau_{24}\sqrt{(a_{22} - a_{44})^2 + 4v_{24}^2}} = \frac{|4|b_{24}| - 4|b_{24}| + \beta_{24}|}{\tau_{24}\sqrt{(\beta'_{24})^2 + 4(\alpha''_{24})^2}} \\ &= \frac{|\beta_{24}|}{\tau_{24}\sqrt{(\beta'_{24})^2 + 4(\alpha''_{24})^2}} = \frac{O(\varepsilon^2)}{O(\varepsilon^2)}, \end{aligned}$$

*where $\beta_{24} = O(\varepsilon^2)$, $\beta'_{24} = O(\varepsilon)$ and $\alpha_{24}$, $\alpha'_{24}$, $\alpha''_{24} = O(\varepsilon^2)$. Thus, $\theta_{24}$ can be any value in the segment $[-\pi/4, \pi/4]$. In a similar way we obtain $\cos(\gamma_{24}) = O(\varepsilon^2)/O(\varepsilon^2)$, $\sin(\gamma_{24}) = O(\varepsilon^2)/O(\varepsilon^2)$, so $\gamma_{24}$ can be any value in $[-\pi/2, \pi/2]$. Hence $\phi_{24}$ and $\psi_{24}$ can be any value in $[-\pi/2, \pi/2]$.*

*The same conclusion can be made for $\phi_{34}$ and $\psi_{34}$ in step 8. Now, for $|a'_{23}| + |b'_{23}|$ we have*

$$|a'_{23}| + |b'_{23}| = |\sin\psi_{24}| \cdot (|\bar{a}_{34}| + |\bar{b}_{34}|) = |\sin\psi_{24}| \cdot (|a_{34}| + |b_{34}|) = O(\varepsilon).$$

*By a similar analysis one can see that in the steps* 7, 9 *and* 10, *the contributions to all matrix elements are either* 0 *or equal to* $O(\varepsilon^2)$. *In step* 8 *the element* $a_{24}^{(8)}$ ($b_{24}^{(8)}$) *is obtained from an expression that includes* $a_{23}^{(7)}$ ($b_{23}^{(7)}$), *respectively. Hence they can become equal to* $O(\varepsilon)$. *In any case, we conclude that* $|a_{23}^{(N)}| + |a_{24}^{(N)}| = O(\varepsilon)$ *or* $|b_{23}^{(N)}| + |b_{24}^{(N)}| = O(\varepsilon)$, *which shows the failure of the quadratic convergence.* □

*4.1. The bounds for the real HZ method*

All obtained results can be almost directly applied to the real HZ method. So, the next section can be seen as an application of the previous theory.

If $A$ and $B$ are real then the complex HZ algorithm reduces to the real one. Let us present the formulas of the real HZ algorithm, which can be found in [7, 13]. We have

$$\hat{Z} = \frac{1}{\sqrt{1-(b_{ij})^2}} \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\psi & \cos\psi \end{bmatrix} = \begin{bmatrix} c1 & -s1 \\ s2 & c2 \end{bmatrix}, \tag{4.27}$$

where

$$\left.\begin{array}{rcl} \cos\phi &=& \cos\theta - \xi(\sin\theta + \eta\cos\theta) = \rho\cos\theta - \xi\sin\theta \\ \sin\phi &=& \sin\theta + \xi(\cos\theta - \eta\sin\theta) = \rho\sin\theta + \xi\cos\theta \\ \cos\psi &=& \cos\theta + \xi(\sin\theta - \eta\cos\theta) = \rho\cos\theta + \xi\sin\theta \\ \sin\psi &=& \sin\theta - \xi(\cos\theta + \eta\sin\theta) = \rho\sin\theta - \xi\cos\theta. \end{array}\right\} \tag{4.28}$$

Here

$$\xi = \frac{b_{ij}}{\sqrt{1+b_{ij}}+\sqrt{1-b_{ij}}}, \quad \eta = \frac{b_{ij}}{(1+\sqrt{1+b_{ij}})(1+\sqrt{1-b_{ij}})}, \tag{4.29}$$

$$\rho = 1 - \xi\eta = \xi + \sqrt{1-b_{ij}} = \frac{1}{2}(\sqrt{1+b_{ij}}+\sqrt{1-b_{ij}}). \tag{4.30}$$

One can show that $\rho^2 + \xi^2 = 1$ and (since $|b_{ij}| < 1$) $|\xi| < \rho$. The values of $\sin\theta$ and $\cos\theta$ are computed from $\tan\theta$ while $\tan\theta$ is computed from $\tan(2\theta)$,

$$\tan(2\theta) = \frac{2a_{ij} - (a_{ii}+a_{jj})\,b_{ij}}{\sqrt{1-(b_{ij})^2}\,(a_{ii}-a_{jj})}, \qquad -\frac{\pi}{4} \le \theta \le \frac{\pi}{4}\ . \tag{4.31}$$

If $a_{ii} = a_{jj}$ and $2a_{ij} = (a_{ii}+a_{jj})b_{ij}$ then $\hat{A}$ and $\hat{B}$ are proportional and we choose $\theta = 0$. Then $\hat{Z}$ reduces to the relation (2.16) and it is easy to show that $a'_{ii} = a_{ii}$, $a'_{jj} = a_{jj}$.

The diagonal elements of $\hat{B}'$ are ones, while the diagonal elements of $\hat{A}'$ can be computed using the formulas [13],

$$a'_{ii} = a_{ii} + \frac{1}{1-b_{ij}^2}\left[(b_{ij}^2 - \sin^2\phi)\,a_{ii} + 2\cos\phi\sin\psi\,a_{ij} + \sin^2\psi\,a_{jj}\right], \tag{4.32}$$

$$a'_{jj} = a_{jj} - \frac{1}{1-b_{ij}^2}\left[(\sin^2\psi - b_{ij}^2)\,a_{jj} + 2\cos\psi\sin\phi\,a_{ij} - \sin^2\phi\,a_{ii}\right]. \tag{4.33}$$

To prove the quadratic convergence of the real HZ method, we shall use the same asymptotic assumptions (3.11), (3.12) and all of the notation from (3.13) that applies to the real algorithm. Note that all matrices $A^{(k)}$ are symmetric and all $B^{(k)}$ are symmetric positive definite with unit diagonal.

**Theorem 4.5.** *Let the assumptions (3.11), (3.12) hold for real symmetric $A^{(0)}$ and $B^{(0)}$, where $B^{(0)}$ is positive definite with unit diagonal. Let the sequence of pairs $((A^{(k)}, B^{(k)}), k \geq 0)$ be generated by the real cyclic HZ method defined by the relations (4.27)–(4.33).*

*If the eigenvalues of the pair $(A^{(0)}, B^{(0)})$ are simple, then the relation (4.1) holds for any cyclic pivot strategy. If the pivot strategy is row-cyclic or any wavefront strategy, then the relation (4.2) holds.*

*If in addition the conditions (3.27) and (4.23) hold then the relation (4.2) holds for any wavefront strategy even if the eigenvalues are double.*

*Corollary 4.3 holds with constants* 0.4079 *and* 0.2863 *instead of* 0.5558 *and* 4345.

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

## 5. The numerical experiments in MATLAB

The goal of this section is twofold. First, to inspect how the method behaves asymptotically when the eigenvalues of the PGEP are simple, double and multiple. Second, to see whether some special pivot strategies, such as the de Rijk strategy from [2] can reduce total number of cycles. We have made two MATLAB functions, `dsychz_qc(A,B,eivec)` and `dsychz_qcsortd(A,B,eivec)`. In the first (second) function the method uses the row-cyclic (de Rijk) strategy. The de Rijk strategy is essentially the row-cyclic strategy that tries to order the diagonal elements of $A^{(k)}$ in descending order during the process, with modest cost. The function `dsychz_qc(A,B,eivec)` was coded exactly following the lines of the HZ algorithm which is displayed at the end of Section 2. Except for the pivot strategy, all other parameters of the two functions are identical (same input, same output, same statements).

The input to `dsychz_qc(A,B,eivec)` are the initial matrices $A$, $B$ and `eivec` parameter which determines whether the matrix of eigenvectors has to be computed.

The output to `dsychz_qc` are: the eigenvector matrix, the column vector of eigenvalues, the total number of cycles (`cycles`) and steps (`steps`), the variable `info` and `steps`$\times 5$ matrix `qc`. The $k$'th row of `qc` is row-vector with 5 components: $\mathsf{S}(A_S^{(k)})$, $\mathsf{S}(A^{(k)})$, $\mathsf{S}(B^{(k)})$, $S(A^{(k)}, B^{(k)})$ and $S(A_S^{(k)}, B^{(k)})$.

We shall construct 3 matrix pairs. The first has simple, the second has double and the third has multiple eigenvalues. For each pair $(A, B)$ we display 2 figures. The first (second) is related to the method that uses the row-cyclic (de Rijk) strategy. In each figure we display the graphs of 4 functions. The first function is $k \mapsto S(A^{(k)}, B^{(k)})$, $k \geq 0$, then $k \mapsto \mathsf{S}(A^{(k)})$, $k \mapsto \mathsf{S}(B^{(k)})$ and $k \mapsto \mathsf{S}(A_S^{(k)})$, where $A_S^{(k)} = D_k^{-1/2} A^{(k)} D_k^{-1/2}$, $D_k = \mathrm{diag}(|A^{(k)}|)$, $k \geq 0$. These

functions are obtained from the output matrix `qc`. We note that the function $\mathsf{S}(A_S^{(k)})$ is important when both matrices $A$ and $B$ are positive definite. Then the spectral condition numbers of $A_S^{(0)}$ and $B^{(0)}$ determine whether the pair $(A, B)$ is well-behaved [13, 3, 26]. Also, in the stopping criterion (2.6) the quantity $S(A_S^{(k)}, B^{(k)})$ is used. We have displayed the graphs of those 4 functions using the logarithmic scale ($y$-axis only). This is accomplished by the MATLAB `semilogy` function.

Once the quadratic convergence assumptions are met, we expect a significant drop of the function values at the end of every cycle. Therefore we have labeled $x$-axis ticks by 0, ... `cycles`. The tick labeled $t$ corresponds to the step $t \cdot N$.

The matrix pairs have been generated using the following code:

```
function [A,B] = genmatAB(da,db,cndF);
 n=length(da); X= ones(n)-2*rand(n)+1i*rand(n); [U,~] = qr(X);
 D=diag(linspace(1e1^(-cndF/2),1e1^(cndF/2),n)); F=D*U;
 A = F'*diag(da)*F; A=0.5*(A+A'); B = F'*diag(db)*F; B=0.5*(B+B');
end
```

This way we have control over the condition of the transformation matrix $F$: $\kappa_2(F) = 10^{\texttt{cndF}}$. We know that the eigenvalues of the pair $(A, B)$ are (up to the influence of rounding errors) entries of the vector `da`. This is a consequence of the choice of the vector `db`. It contains $n$ units: `db`$= [1, 1, \ldots, 1]$.

Our choice is `n`$= 128$, `cndF`$= 2$. We shall not delve in the construction of the vector `da` which depends on several parameters. Since the figures are displayed for only 3 matrix pairs, we only describe how these 3 vectors `da` are constructed. Then we shall present and comment the graphs of the functions.

*5.1. Simple eigenvalues*

The vector `da` is computed using the code: `da = linspace(1.0,1000.0,n);` Hence both matrices $A$ and $B$ are positive definite and the eigenvalues of the pair $(A, B)$ are very close to the entries of the vector `da`. The characteristic data are: $\delta \approx 2.622$, $\mu = 10^3$, $\kappa_2(A) \approx 10^7$, $\kappa_2(B) \approx 10^4$, $\kappa_2(A_S) \approx 9.89 \cdot 10^6$, $\kappa_2(B_S) \approx 9.93 \cdot 10^3$. The diagonal elements of $A^{(0)}$ make a slowly increasing sequence of numbers: 743.54, 768.07, ..., 760.01. The asymptotic conditions (3.11) and (3.12) take the form:

$$\mathsf{S}(B^{(0)}) < \frac{1}{128 \cdot 127} \approx 6.15 \cdot 10^{-5}, \;\; \epsilon_0 < \frac{\delta}{2\sqrt{1+\mu^2}} \approx \frac{2.622}{2\sqrt{1000001}} \approx 1.311 \cdot 10^{-3}.$$

We expect that after $S(A^{(k)}, B^{(k)})$ reaches the value $10^{-5}$, the quadratic convergence will commence. In Figure 1 (Figure 2) are displayed the graphs of the functions obtained by applying the complex HZ method with the row-cyclic (de Rijk) pivot strategy.

We can see from the graphs that the quadratic convergence commenced after $S(A^{(k)}, B^{(k)})$ reached the value $10^{-5}$. We see that the total number of cycles equals 14 (9) for the row-cyclic (de Rijk) strategy. The behavior of the

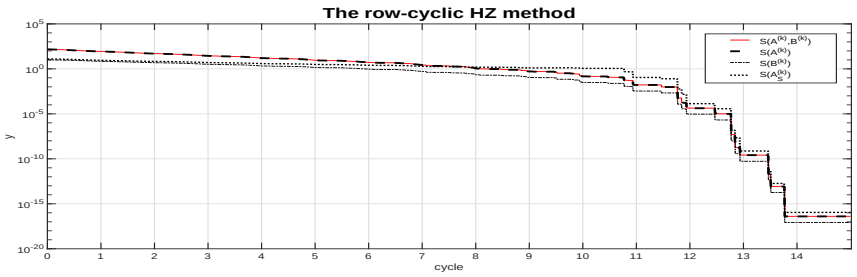


Figure 1: The graphs of the functions $S(A^{(k)}, B^{(k)})$, $\mathsf{S}(A^{(k)})$, $\mathsf{S}(B^{(k)})$, $\mathsf{S}(A_S^{(k)})$

]

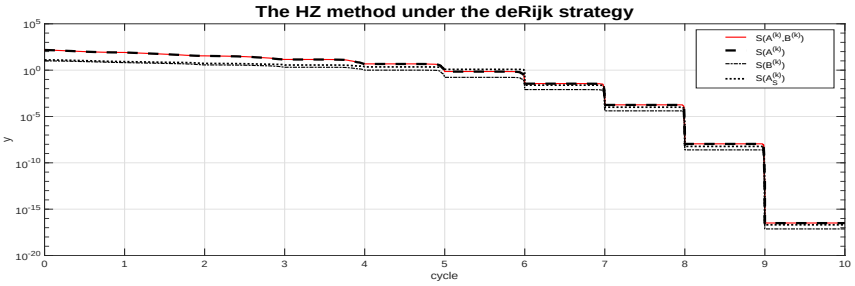


Figure 2: The graphs of the functions $S(A^{(k)}, B^{(k)})$, $\mathsf{S}(A^{(k)})$, $\mathsf{S}(B^{(k)})$, $\mathsf{S}(A_S^{(k)})$

functions in Figure 2 is just perfect. This probably comes from the fact that the descending ordering of the diagonal elements during the process plays an important role in better performance of the method. We note that the diagonal elements of $A^{(0)}$ are increasingly ordered, so the row-cyclic method had to cope with that. On the other hand, under the de Rijk strategy the matrices $A^{(0)}$ and $B^{(0)}$ are symmetrically permuted by permutation $P$ so that $P^T A^{(0)} P$ has nonincreasing ordering of the diagonal entries.

*5.2. Double eigenvalues*

In this case the vector `da` has been generated by the statements:

```
da = linspace(1.0,1000,n);  for i = 1:2:n-1, da(i) = da(i+1); end
```

Again, we have $\mu = 10^3$ and since all components of `da` are double, we have $\delta \approx 5.2$. The asymptotic assumptions (3.11) and (3.12) have the form: $\mathsf{S}(B^{(0)}) < 6.2 \cdot 10^{-5}$ and $\epsilon_0 < 2.6 \cdot 10^{-3}$. The additional condition (3.27) takes the form

$$\epsilon_0 < \left[\frac{\delta_0}{\mu+1}\right]^{1/2} \delta = \left[\frac{\delta}{(\mu+1)(\mu^2+1)}\right]^{1/2} \delta \approx 3.75 \cdot 10^{-4}$$

Thus $\epsilon_0 \approx 10^{-5}$ is the expected threshold. We note that yet the condition (3.26) has to be satisfied. The behavior of the considered functions is displayed in figures 5.2 and 5.2.

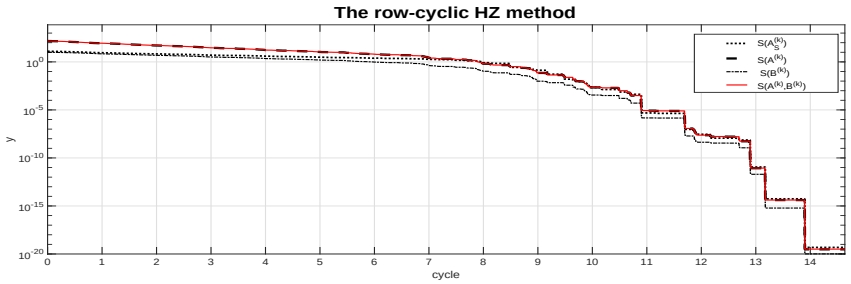


Figure 3: The graphs of the functions $S(A^{(k)}, B^{(k)})$, $\mathsf{S}(A^{(k)})$, $\mathsf{S}(B^{(k)})$, $\mathsf{S}(A_S^{(k)})$

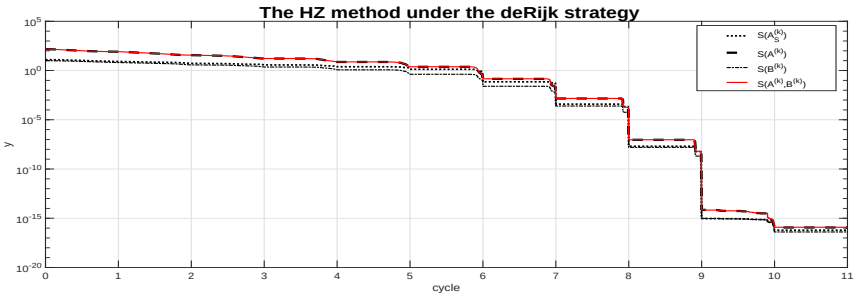


Figure 4: The graphs of the functions $S(A^{(k)}, B^{(k)})$, $\mathsf{S}(A^{(k)})$, $\mathsf{S}(B^{(k)})$, $\mathsf{S}(A_S^{(k)})$

In Figure 3 we see the failure of the quadratic convergence. This comes from the fact that the starting matrix $A^{(0)}$ has the diagonal elements in increasing order and after 11 cycles when $S(A^{(k)}, B^{(k)})$ is reduced to $10^{-5}$, the diagonal elements failed to satisfy the condition (4.23).

On the other hand, the de Rijk strategy will ensure that the condition (4.23) holds after just a few cycles. The quadratic convergence is clearly observed after cycle 6. The total number of cycles is also reduced from 14 to 9. It is interesting to see the slowdown of the quadratic convergence when $S(A^{(k)}, B^{(k)})$ approaches the convergence criterion. This is due to the rounding error influence. During the generation of $A$ and $B$, the double eigenvalues have been perturbed by tiny perturbations and for small $S(A^{(k)}, B^{(k)})$ the relevant $\delta$ is not 5.2 but a quantity close to the unit round-off.

### *5.3. Multiple eigenvalues*

In this case `da` has been generated by the code:

```
da=linspace(-1000, 1000,n);
for i=1:10:n-10, j=i; for k=j+1:j+9, da(k)=da(i); end, end
```

We have 12 multiple eigenvalues of the pair $(A, B)$, each of multiplicity 10. Their approximate values are $-1000$, $-842.52$, $-685.04$, ..., 574.80, 732.28. We also have 8 simple and equidistant eigenvalues in the interval $[889.76, 1000]$. These

simple eigenvalues determine the value of $\delta$. The diagonal elements of $A^{(0)}$ are scattered between 360.58 and 527.92. They are not in a monotone order.

We have $\mu = 10^3$ and $\delta \approx 1.575$. The conditions (3.11) and (3.12) have the form: $\mathsf{S}(B^{(0)}) < 6.2 \cdot 10^{-5}$ and $\epsilon_0 < 7.875 \cdot 10^{-4}$. The condition (3.27) takes the form $\epsilon_0 < 3.967 \cdot 10^{-5}$. If the quadratic convergence occurred, the threshold would be $\epsilon_0 \approx 10^{-5}$. From figures 5.3 and 5.3 we see the failure of the quadratic convergence.

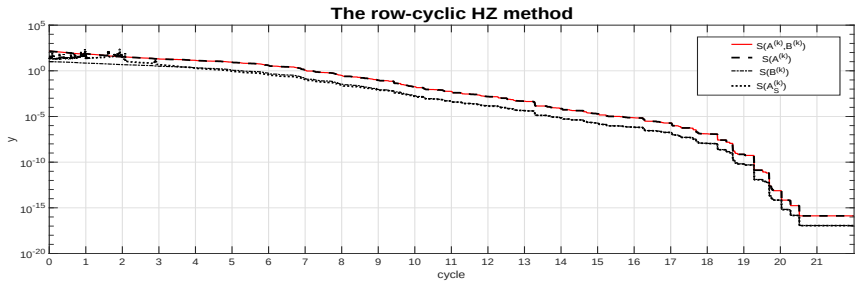


Figure 5: The graphs of the functions $S(A^{(k)}, B^{(k)})$, $\mathsf{S}(A^{(k)})$, $\mathsf{S}(B^{(k)})$, $\mathsf{S}(A_S^{(k)})$

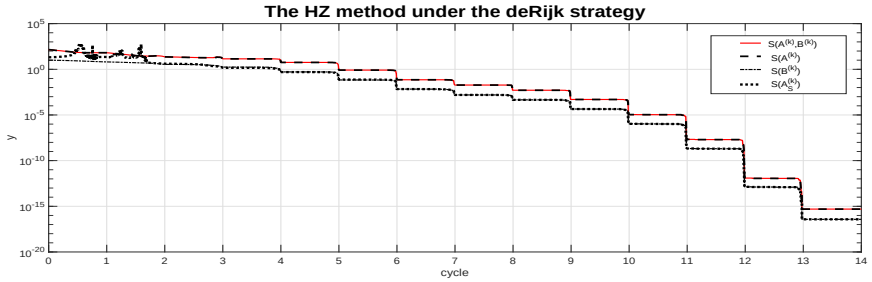


Figure 6: The graphs of the functions $S(A^{(k)}, B^{(k)})$, $\mathsf{S}(A^{(k)})$, $\mathsf{S}(B^{(k)})$, $\mathsf{S}(A_S^{(k)})$

Nevertheless, we see that the row-cyclic method required full 21 cycles to reach the stopping criterion. On the other hand, the de Rijk strategy reduced the number of cycles to 13.

## 6. Conclusions and Future Work

In this paper we have proved the quadratic asymptotic convergence of the general cyclic complex HZ method in the case of simple eigenvalues. For the wavefront pivot strategies the multiplicities of the eigenvalues can be at most double. The same conclusion holds if a wavefront strategy is replaced by the de Rijk strategy. The same results hold for the real method.

Future work can be concentrated on modifying the method to become quadratically convergent in the case of multiple eigenvalues. In that case a prominent

role will be played by the quantity $\mathcal{T}(\tilde{A}, \tilde{B})$ from the relation (3.8) and by Theorem 3.2. The first steps in that direction are already made in [7]. The most difficult obstacle appears to be detecting the multiplicities of the eigenvalues. In the presence of tiny clusters of eigenvalues, the problem can be reduced to the case of multiple eigenvalues by the technique presented in [9].

Since the de Rijk pivot strategy has proven to reduce the total number of cycles, its importance has increased. So, the global convergence problem of the HZ method under that pivot strategy should be solved by inspecting the proofs and techniques from [5, 1, 17]. Another important problem that remains open is the asymptotic quadratic convergence of the general cyclic block Jacobi method from [18, 11]. Finally, an open problem is to prove the global and asymptotic quadratic convergence of the real and complex, element-wise and block HZ methods under the special parallel strategies from [21, 22].

## 7. ACKNOWLEDGMENTS

The author is indebted to A. Bašić Šiško for carefully reading the paper.

## References

[1] Bujanović, Z., Drmač, Z.: A contribution to the theory and practice of the block Kogbetliantz method for computing the SVD. BIT Numer Math 52 (4), 827–849 (2012)

[2] de Rijk, P. P. M.: A one–sided Jacobi algorithm for computing the singular value decomposition on a vector computer. SIAM J. Sci. Stat. Comp. 10, 359-371 (1989)

[3] Drmač, Z.: A tangent algorithm for computing the generalized singular value decomposition. SIAM J. Numer. Anal. 35 (5), 1804–1832 (1998)

[4] Falk, S., Langemeyer, P.: Das Jacobische Rotations-Verfahren für reel symmetrische Matrizen-Paare I, II. Elektronische Datenverarbeitung 30–43 (1960)

[5] Fernando, K. V.: Linear convergence of the row cyclic Jacobi and Kogbetliantz methods. Numer Math 56, 73–94 (1989)

[6] Hansen, E. R.: On cyclic Jacobi methods. SIAM J. Appl. Math. 11 (2), 448–459 (1963)

[7] Hari, V.: On cyclic Jacobi methods for the positive definite generalized eigenvalue problem. Ph.D. thesis, University of Hagen (1984)

[8] Hari, V.: On pairs of almost diagonal matrices. Linear Algebra and Its Appl. 148, 193–223 (1991)

[9] Hari, V.: On sharp quadratic convergence bounds for the serial Jacobi methods. Numer Math 60, 375–406 (1991)

[10] Hari, V., Drmač, Z.: On scaled almost diagonal Hermitian matrix pairs. SIAM J. Matrix Anal. Appl. 18 (4), 1000–1012 (1997)

[11] Hari, V.: Convergence to diagonal form of block Jacobi-type methods. Numer Math 129 (3), 449–481 (2015)

[12] Hari, V., Begović Kovač, E.: Convergence of the cyclic and quasi-cyclic block Jacobi methods. Electronic transactions on numerical analysis 46 107–147 (2017)

[13] Hari, V.: Globally convergent Jacobi methods for positive definite matrix pairs. Numer Algor 79 (1), 221–249 (2018); https://doi.org/10.1007/s11075-017-0435-5

[14] Hari, V., Begović Kovač, E.: On the convergence of complex Jacobi methods. To appear in Linear and Multilin. Algebra. https://doi.org/10.1080/03081087.2019.1604622

[15] Hari V.: Complex Cholesky–Jacobi algorithm for PGEP, AIP conference proceedings 2116, 450011 (2019); https://doi.org/10.1063/1.5114478

[16] Hari, V.: On the complex Falk-Langemeyer method. Numer Algor (2019). https://doi.org/10.1007/s11075-019-00689-8

[17] Hari, V.: On the global convergence of the complex HZ method. Submitted for publication.

[18] Hari, V.: On the global convergence of the block Jacobi method for the positive definite generalized eigenvalue problem. Submitted for publication.

[19] Luk, F. T. , Park, H.: On parallel Jacobi orderings. SIAM J. Sci. and Stat. Comput. 10 (1) 18–26 (1989)

[20] Novaković, V., Singer, S., Singer, S.: Blocking and parallelization of the Hari–Zimmermann variant of the Falk–Langemeyer algorithm for the generalized SVD. Parallel Comput. 49, 136–152 (2015)

[21] Oksa, G., Yamamoto, Y., Vajtersic. M.: Asymptotic quadratic convergence of the serial block–Jacobi EVD algorithm for Hermitian matrices. Numer Math 136 (4), 1071–1095 (2017)

[22] Oksa, G., Yamamoto, Y., Bečka, M., Vajtersic, M.: Asymptotic quadratic convergence of the two–sided serial and parallel block–Jacobi SVD algorithm. SIAM J. Matrix Anal. Appl. 40, 631–671 (2019); https://doi.org/10.1137/18M1222727

[23] Shroff, G., Schreiber, R.: On the convergence of the cyclyc Jacobi method for parallel block orderings. SIAM J. Matrix Anal. Appl. 10 (3), 326–346 (1989)

[24] Singer, S., Di Napoli, E., Novaković, V., Čaklović, G.: The LAPW method with eigendecomposition based on the HariZimmermann generalized hyperbolic SVD. arXiv:1907.08560 [math.NA]. Submitted for for publication.

[25] Slapničar, I., Hari, V.: On the quadratic convergence of the Falk–Langemeyer method for definite matrix pairs. SIAM J. Matrix Anal. Appl. 12 (1), 84–114 (1991)

[26] van der Sluis, A.: Condition numbers and equilibration of matrices. Numer Math 14 (1), 14–23 (1969)

[27] Wilkinson, J. H.: Note on the quadratic convergence of the cyclic Jacobi Process. Numer Math 4, 296–300 (1962)

## 8. Appendix

### *8.1. Proof of Lemma 3.5*

Using Lemma 3.4 we have

$$\begin{aligned}\epsilon_{k+1}^2 &\le x_k\left[\epsilon_k^2 - 2(a_k^2+b_k^2)\right] \le x_k\left\{x_{k-1}\left[\epsilon_{k-1}^2 - 2(a_{k-1}^2+b_{k-1}^2)\right] - 2(a_k^2+b_k^2)\right\}\\ &\le \ \ldots \ \le x_k x_{k-1}\cdots x_0\epsilon_0^2 - 2\sum_{r=0}^{k} x_k\cdots x_r(a_r^2+b_r^2)\\ &\le x_k(1-b_{\max})^{-k}\epsilon_0^2 - 2x_k\sum_{r=0}^{k}(a_r^2+b_r^2), \qquad 0\le k\le N-1. \quad (8.1)\end{aligned}$$

Thus

$$\epsilon_N/x_{N-1} \le (1-b_{\max})^{-(N-1)}\epsilon_0^2 - 2\sum_{r=0}^{N-1}(a_k^2+b_k^2),$$

and since $\epsilon_N \ge 0$, $x_{N-1} = 1 - b_{N-1} > 0$, we have

$$\sum_{r=0}^{N-1}(a_k^2+b_k^2) \le \frac{1}{2}(1-b_{\max})^{-(N-1)}\epsilon_0^2 \le \frac{1}{2}(1-b_{\max})^{-N}\epsilon_0^2. \quad (8.2)$$

To prove the bound for $b_{\max}$ we set $\beta_k = \mathsf{S}(B^{(k)})/\sqrt{2}$, $k \ge 0$. Then we have

$$b_k \le \beta_k, \qquad k \ge 0. \quad (8.3)$$

From the relation (3.15) we obtain

$$\beta_{k+1} \le \sqrt{x_k}\beta_k \le \frac{\beta_k}{\sqrt{1-\beta_k}}, \qquad k \ge 0. \quad (8.4)$$

By induction with respect to $k$, we obtain

$$\beta_k \le \frac{\beta_0}{\sqrt{1-k\beta_0}} \ \text{ provided that } k\beta_0 < 1.$$

The assumption (3.11) yields $k\beta_0 \le N\beta_0 < \sqrt{2}/4$. Hence

$$\beta_k \le \frac{\sqrt{2}}{4N}/\sqrt{1-k\frac{\sqrt{2}}{4N}} \le \left[\frac{4+\sqrt{2}}{7}\right]^{1/2}\frac{1}{2N}, \quad 0 \le k \ge N. \tag{8.5}$$

From relations (8.3) and (8.5) we obtain

$$b_{\max} \le \max_{0\le k\le N}\beta_k < 0.8794652241\frac{1}{2N},$$

which proves the assertion (3.17).

To prove the assertion (3.16), note that (3.17) implies $2Nb_{\max} < 1$. Since $N \ge 3$, Lemma 3.3(i) implies

$$(1-b_{\max})^{-N} \le 1+\frac{12}{7}Nb_{\max} < 1+\frac{6}{7}\left[\frac{4+\sqrt{2}}{7}\right]^{1/2} < 1.753827335. \tag{8.6}$$

Now, the assertion (3.16) follows from the inequalities (8.2) and (8.6). Since $x_0 \cdots x_{k-1} \le (1-b_{\max})^N$, the relation (8.6) implies the assertion (3.19).

Finally, the assertion (3.18) is implied by the relations (3.14), (3.15), (8.1) and (8.6). Namely, for the first two components, $\mathsf{S}^2(A^{(k)})$ and $\mathsf{S}^2(B^{(k)})$, the proof is quite similar to the proof for $\epsilon^2(A^{(k)})$. □

*8.2. Proof of Lemma 3.7*

Using relation (2.12) and the Cauchy-Schwarz inequality, we obtain for $k \ge 0$,

$$\tan^2(2\theta_k) \le \frac{4+(a_{ii}^{(k)}+a_{jj}^{(k)})^2}{\tau_k^2(e_k^2+4v_k^2)}(u_k^2+b_k^2) \le \frac{4+(a_{ii}^{(k)}+a_{jj}^{(k)})^2}{\tau_k^2 e_k^2}(a_k^2+b_k^2). \tag{8.7}$$

To bound $(a_{ii}^{(k)}+a_{jj}^{(k)})^2$, we use inequality $3\delta \le 2\mu$ (which was already used in relation (3.4)) and relation (3.22). We have

$$\begin{aligned} |a_{ii}^{(k)}+a_{jj}^{(k)}| &= |\lambda_i^{(k)}+\lambda_j^{(k)}+a_{ii}^{(k)}-\lambda_i^{(k)}+a_{jj}^{(k)}-\lambda_j^{(k)}| \\ &\le |\lambda_i^{(k)}+\lambda_j^{(k)}| + \sqrt{2|a_{ii}^{(k)}-\lambda_i^{(k)}|^2+2|a_{jj}^{(k)}-\lambda_j^{(k)}|^2} \\ &< 2\mu+0.43846\delta \le \mu(2+\frac{2}{3}0.43846) < 2.2923067\mu, \quad k\in\mathcal{S}'. \end{aligned}$$

The relations (3.13), (3.20) and (3.17) imply

$$e_k^2 \ge 2.56154^2\delta^2, \qquad k\in\mathcal{S}', \tag{8.8}$$

$$\tau_k^2 \ge 1-{b_{\max}}^2 \ge 1-\left(\frac{0.8795}{2N}\right)^2 \ge 1-\left(\frac{0.43975}{3}\right)^2 > 0.9785133. \tag{8.9}$$

Inserting the obtained lower bounds for $e_k^2$ and $\tau_k^2$ into (8.7), we have

$$\sin^2(2\theta_k) \le \tan^2(2\theta_k) \le 0.8184204 \frac{1+\mu^2}{\delta^2}(a_k^2 + b_k^2), \quad k \in \mathcal{S}'. \tag{8.10}$$

To bound $\max\{\sin^2\phi_k, \sin^2\psi_k\}$ we use relation (2.11) and nonnegativity of $\cos(2\theta_k)$ and $\cos\gamma_k$. We have

$$\begin{aligned} 2\max\{\sin^2\phi_k, \sin^2\psi_k\} &\le 1 - \tau_k\cos(2\theta_k)\cos\gamma_k + b_k|\sin(2\theta_k)| \\ &= \frac{\sin^2(2\theta_k) + \cos^2(2\theta_k)(\sin^2\gamma_k + b_k^2\cos^2\gamma_k)}{1+\tau_k\cos(2\theta_k)\cos\gamma_k} + b_k|\sin(2\theta_k)|, \ k \ge 0. \end{aligned} \tag{8.11}$$

For $k \in \mathcal{S}'$ the relations (8.10), (3.18), (3.19), (3.21), (2.13), (3.16) and (3.13) imply

$$\left.\begin{aligned} \sin^2(2\theta_k) &\le 0.8184204\frac{1+\mu^2}{\delta^2}\frac{\epsilon_k^2}{2} < 0.179422, \\ \cos(2\theta_k) &> \sqrt{1-0.179422} > 0.9058576, \\ |\sin\gamma_k| &\le 2\frac{|v_k|}{|e_k|} \le \frac{2}{2.56154}\frac{a_k}{\delta} \le 0.780781\frac{a_k}{\delta}, \\ \cos\gamma_k &\ge \sqrt{1 - 0.780781^2\cdot\frac{0.876914}{4}} > 0.93078. \end{aligned}\right\} \tag{8.12}$$

Since $3\delta \le 2\mu$, we have

$$\frac{\sqrt{1+\mu^2}}{\delta} > \frac{\mu}{\delta} \ge \frac{3}{2}. \tag{8.13}$$

Using the inequalities (8.12), (8.13), (8.9) and (8.10) we obtain for $k \in \mathcal{S}'$

$$\begin{aligned} \frac{1}{1+\tau_k\cos(2\theta_k)\cos\gamma_k} &< 0.545242402, \\ \sin^2\gamma_k + b_k^2\cos^2\gamma_k &\le 0.609619\frac{1+\mu^2}{\delta^2}(a_k^2+b_k^2), \\ b_k|\sin(2\theta_k)| &\le \sqrt{\frac{4}{9}\frac{1+\mu^2}{\delta^2}(a_k^2+b_k^2)}\cdot\sqrt{0.8184204\frac{1+\mu^2}{\delta^2}(a_k^2+b_k^2)} \\ &\le 0.60311061\frac{1+\mu^2}{\delta^2}(a_k^2+b_k^2). \end{aligned}$$

Inserting these bounds into (8.11) we obtain

$$\begin{aligned} 2\max\{\sin^2\phi_k, \sin^2\psi_k\} &\le \frac{1+\mu^2}{\delta^2}(a_k^2+b_k^2)\,[0.545242402(0.8184204 + 0.609619) \\ &\quad +0.60311061] \le 1.3817383\,\frac{1+\mu^2}{\delta^2}(a_k^2+b_k^2), \qquad k \in \mathcal{S}'. \end{aligned} \tag{8.14}$$

If $p = n$, i.e. if all eigenvalues are simple, then the assumption $p \ge 2$, $n \ge 3$ implies $3\delta \le \mu$. Hence $\sqrt{1+\mu^2}/\delta > \mu/\delta \ge 3$ and therefore

$$\frac{1+\mu^2}{\delta^2} > 3\frac{\sqrt{1+\mu^2}}{\delta} > 9. \tag{8.15}$$

From the relations (8.12) and (8.15) we obtain

$$2\max\{\sin^2\phi_k, \sin^2\psi_k\} \le \frac{1+\mu^2}{\delta^2}(a_k^2+b_k^2)\,[0.545242402(0.8184204+0.609619) \\ +0.301555305] \le 1.080183\,\frac{1+\mu^2}{\delta^2}(a_k^2+b_k^2), \qquad k\in\mathcal{S}'. \quad (8.16)$$

The relations (8.14) and (8.16) imply

$$\max\{\sin^2\phi_k, \sin^2\psi_k\} \le \left\{ \begin{array}{l} 0.69086915\frac{1+\mu^2}{\delta^2}(a_k^2+b_k^2),\ p<n \\ 0.5400915\ \frac{1+\mu^2}{\delta^2}\,(a_k^2+b_k^2),\ p=n \end{array} \right\},\ k\in\mathcal{S}'. \quad (8.17)$$

In the case $b_k = 0$, $a_k \neq 0$ we have $\phi_k = \psi_k = \theta_k$, where

$$\tan(2\theta_k) = 2\frac{|a_k|}{|e_k|} \le \frac{2}{2.56154}\frac{|a_k|}{\delta}, \qquad k\in\mathcal{S}',$$

hence

$$\max\{\sin^2\phi_k, \sin^2\psi_k\} \le \left(\frac{1}{2.56154}\frac{a_k}{\delta}\right)^2, \qquad k\in\mathcal{S}'.$$

Thus, the inequality (8.17) also holds. Now, Lemma 3.7 follows from the inequalities (3.16) and (8.17). □

*8.3. Proof of Lemma 3.8*

Note that the inequality (3.21) holds for any pivot strategy. It implies $\sqrt{1+\mu^2}\epsilon_k \le \sqrt{0.43846}\,\delta$, $0 \le k \le N$, which means that the assumption (3.6) of Lemma 3.1 and Theorem 3.2 holds for all $0 \le k \le N$. By Theorem 3.2, for each $k$, $0 \le k \le N$, there is a permutation matrix $P_k$ such that the inequalities (3.28) and (3.29) hold for the pair $(P_k^* A^{(k)} P_k, P_k^* B^{(k)} P_k)$. From the relation (3.26) we see that $P_0 = I_n$. It remains to prove $P_k = I_n$ for $1 \le k \le N$. We prove it by induction.

Suppose $P_k = I_n$ for some $k$, $0 \le k \le N-1$. We shall prove $P_{k+1} = I_n$. From the relation (3.22) we obtain

$$|a_{rr}^{(k)} - \lambda_r| < \frac{\sqrt{2}}{2}\cdot 0.43846\delta < 0.31004\delta, \qquad 1 \le r \le n, \qquad (8.18)$$

where $\lambda_1, \ldots, \lambda_n$ is an ordering of the eigenvalues. In the $k$th step only $a_{ii}^{(k)}$ and $a_{jj}^{(k)}$ change. If $\lambda_i \neq \lambda_j$ there are only two possibilities: either

$$|a_{ii}^{(k+1)} - \lambda_i| < 0.31004\delta \qquad \text{and} \qquad |a_{jj}^{(k+1)} - \lambda_j| < 0.31004\delta$$

or

$$|a_{ii}^{(k+1)} - \lambda_j| < 0.31004\delta \qquad \text{and} \qquad |a_{jj}^{(k+1)} - \lambda_i| < 0.31004\delta\,. \qquad (8.19)$$

Any other possibility would lead to a contradiction with the assumption on the multiplicities of the eigenvalues. Therefore, it is sufficient to prove that the inequalities in (8.19) do not hold.

Since the relations (8.18) and (8.19) imply

$$\begin{aligned} |a_{ii}^{(k+1)} - a_{ii}^{(k)}| &\geq |\lambda_i - \lambda_j| - |a_{ii}^{(k+1)} - \lambda_j| - |\lambda_i - a_{ii}^{(k)}| \\ &> 3\delta - 0.31004\delta - 0.31004\delta = 2.37992\delta\,, \end{aligned}$$

it is sufficient to prove that the conditions of Lemma 3.8 imply

$$|a_{ii}^{(k+1)} - a_{ii}^{(k)}| \leq 2.37992\,\delta\,. \tag{8.20}$$

Using the relation (2.14) we obtain

$$\begin{aligned} a_{ii}^{(k+1)} - a_{ii}^{(k)} &= [(b_k^2 - \sin^2\phi_k)a_{ii}^{(k)} + \sin^2\psi_k a_{jj}^{(k)} \\ &\quad - 2\cos\phi_k \sin\psi_k\, \Re(e^{-\imath\beta_k} a_{ij}^{(k)}))]/(1-b_k^2). \end{aligned}$$

Hence

$$\begin{aligned} (1-b_k^2)|a_{ii}^{(k+1)} - a_{ii}^{(k)}| &\leq \max\{|a_{ii}^{(k)}|, |a_{jj}^{(k)}|\}(|(b_k^2 - \sin^2\phi_k| + \sin^2\psi_k) \\ &\quad + \sin^2\psi_k + a_k^2. \end{aligned} \tag{8.21}$$

Note that relation (8.9) implies

$$1/\tau_k^2 = 1/(1-b_k^2) \leq 1/(1-b_{\max}^2) < 1.021959.$$

Since

$$|b_k^2 - \sin^2\phi_k| + \sin^2\psi_k \leq \begin{cases} \sin^2\phi_k + \sin^2\psi_k, & \sin^2\phi_k \geq b_k^2 \\ b_k^2 + \sin^2\psi_k, & \sin^2\phi_k < b_k^2 \end{cases},$$

from the relations (8.13) and (8.17) we obtain

$$|b_k^2 - \sin^2\phi_k| + \sin^2\psi_k \leq 2\cdot 0.690869\frac{1+\mu^2}{\delta^2}(a_k^2+b_k^2) = 1.381738\frac{1+\mu^2}{\delta^2}(a_k^2+b_k^2).$$

Since all matrix pairs $(A^{(k)}, B^{(k)})$ have the same spectral radius $\mu$, relation (3.5) implies

$$\max\{|a_{ii}^{(k)}|, |a_{jj}^{(k)}|\} \leq \mu\,.$$

Finally, from (8.13) and (8.17), we obtain

$$\begin{aligned} \sin^2\psi_k + a_k^2 &\leq 0.690869\frac{1+\mu^2}{\delta^2}(a_k^2+b_k^2) + \frac{4}{9}\frac{1+\mu^2}{\delta^2}(a_k^2+b_k^2) \\ &\leq 1.1353135\frac{1+\mu^2}{\delta^2}(a_k^2+b_k^2). \end{aligned}$$

Inserting the obtained inequalities into (8.21) we have

$$\begin{aligned} |a_{ii}^{(k+1)} - a_{ii}^{(k)}| &\leq 1.021958 \frac{1+\mu^2}{\delta^2}(a_k^2 + b_k^2)\,[1.381738\mu + 1.1353135] \\ &\leq 1.41208(\mu+1)\frac{1+\mu^2}{\delta^2}\frac{\epsilon_k^2}{2} \leq 0.70604\frac{1+\mu^2}{\delta^2}(\mu+1)\epsilon_k^2. \end{aligned}$$

The assertions (3.18), (3.19) of Lemma 3.5 and the assumption (3.27) of this lemma imply

$$|a_{ii}^{(k+1)} - a_{ii}^{(k)}| \leq 0.70604 \cdot 1.75383 \frac{1+\mu^2}{\delta^2}(\mu+1)\epsilon_0^2 < 1.2383\,\delta,$$

which proves (8.20) and the lemma. □

*8.4. Proof of Corollary 4.3*

Since any wavefront pivot strategy delivers the same matrices $A^{(N)}$ and $B^{(N)}$, we can assume that the pivot strategy is row-cyclic.

By Lemma 3.8 we conclude that during the whole cycle, the affiliation of the diagonal elements to the eigenvalues does not change. Since the eigenvalues of the pair $(A^{(0)}, B^{(0)})$ and the diagonal elements of $A^{(0)}$ are nonincreasingly ordered, we conclude that for $1 \leq r \leq n$ and $0 \leq k \leq N$, each $a_{rr}^{(k)}$ is affiliated with $\lambda_r$, where $\lambda_r$ is from the nonincreasing ordering of the eigenvalues (3.1).

A finer estimate than (8.8) can be obtained for the quantity $e_k$. Like in the relation (3.23), we have

$$\begin{aligned} |a_{ii}^{(k)} - a_{jj}^{(k)}| &\geq |\lambda_i - \lambda_j| - \sqrt{2|a_{ii}^{(k)} - \lambda_i|^2 + 2|a_{jj}^{(k)} - \lambda_j|^2} \\ &> 3\max\{\tilde{\delta}_i, \tilde{\delta}_j\} - 0.43846\delta \geq 2.56154\max\{\tilde{\delta}_i, \tilde{\delta}_j\}. \end{aligned}$$

Note that $i = i(k)$, $j = j(k)$. Using that in the proof of Lemma 3.7, the relation (8.17) can be replaced by

$$\sin^2(\omega_k) \leq \left\{ \begin{array}{l} 0.690869 \dfrac{1+\mu^2}{\max\{\tilde{\delta}_i^2, \tilde{\delta}_j^2\}}(a_k^2 + b_k^2),\ p < n \\ 0.5400915 \dfrac{1+\mu^2}{\max\{\tilde{\delta}_i^2, \tilde{\delta}_j^2\}}(a_k^2 + b_k^2),\ p = n \end{array} \right\},\ k \in \mathcal{S}', \qquad (8.22)$$

where $\omega_k \in \{\phi_k, \psi_k\}$. By the relations (4.18), (4.21), (8.22), (3.16) and using $\psi_{ij} = \psi_{i(k)j(k)}$ instead of $\psi_k$, we have

$$\begin{aligned} \sum_{t=r+1}^{n} |a_{rt}^{(N)}|^2 &\leq \frac{1}{2} \cdot 1.75383 \cdot 1.04618032 \frac{\mathsf{S}^2(A^{(0)})}{2} \sum_{s=r+2}^{n} \sin^2(\psi_{rs}) \\ &\leq 0.917411215 \cdot 0.690869 \cdot 0.876915 \frac{1+\mu^2}{\delta_r^2}\mathsf{S}^2(A^{(0)})\epsilon_0^2 \\ &\leq 0.5558 \frac{1+\mu^2}{\tilde{\delta}_r^2}\mathsf{S}^2(A^{(0)})\epsilon_0^2. \qquad (8.23) \end{aligned}$$

Here we have used the fact that $\{(r,r+2),\ldots,(r,n)\}\subseteq\mathcal{S}'$, where $\mathcal{S}'$ is from the relation (3.24). We have also used the assertion (3.16) of Lemma 3.5 and the obvious inequality $\tilde{\delta}_r\leq\max\{\tilde{\delta}_r,\tilde{\delta}_s\}$, $s>r+1$.

The same estimate can be made for the corresponding elements of $B^{(N)}$, thus obtaining

$$\sum_{t=r+1}^{n}|b_{rt}^{(N)}|^2\leq 0.5558\,\frac{1+\mu^2}{\tilde{\delta}_r^2}\mathsf{S}^2(B^{(0)})\epsilon_0^2.$$

Summing the last two inequalities one obtains

$$\sum_{t=r+1}^{n}\left(|a_{rt}^{(N)}|^2+b_{rt}^{(N)}|^2\right)\leq 0.5558\frac{1+\mu^2}{\tilde{\delta}_r^2}\epsilon_0^4,\qquad 1\leq r\leq n-1.$$

In the case of simple eigenvalues, the final bound is multiplied by the factor $0.5400915/0.690869$. □

*8.5. Proof of Theorem 4.5*

Let us check whether the results from Section 3 and Section 4 hold for the real method.

- Lemma 3.1, Lemma 3.3 and Theorem 3.2 are not linked to the HZ algorithm. Lemma 3.1 and Theorem 3.2 obviously hold for real matrices.
- Lemma 3.4 holds because for both the real and complex HZ method we have $\|\hat{Z}_k\|_2=1/\sqrt{1-b_k}=\sqrt{x_k}$, $k\geq 0$ (see [13]).
- Lemma 3.5 holds because its proof uses Lemma 3.4, Lemma 3.3 and the inequalities $b_k\leq\mathsf{S}(B^{(k)})/\sqrt{2}$, $k\geq 0$.
- Lemma 3.6 holds because its proof uses only Lemma 3.5 and Lemma 3.1.
- Let us check whether Lemma 3.7 holds for the real method. From relation (4.31) we see that relation (8.7) holds. Then relations (8.8)– (8.10) also hold because they rely on some previously obtained relations. Next, from relation (4.28), we have

$$\begin{aligned}\max\{\sin^2\phi_k,\sin^2\phi_k\} &\leq \rho_k^2\sin^2\theta_k+\xi_k^2\cos^2\theta_k+\rho_k|\xi_k\,\sin(2\theta_k)|\\ &\leq \max\{\xi_k^2,\sin^2\theta_k\}+|\xi_k\,\sin(2\theta_k)|,\end{aligned}$$

where $\rho_k$ and $\xi_k$ denote $\rho$ and $\xi$ at step $k$, respectively. Using relations (8.9) and (8.13), we obtain

$$\xi_k^2=\frac{b_k^2}{2+2\tau_k}\leq\frac{4}{9}\frac{1+\mu^2}{\delta^2}\frac{b_k^2}{2+2\sqrt{1-0.43975^2/9}}\leq 0.1117145\frac{1+\mu^2}{\delta^2}\,b_k^2,$$

which holds for any $k \geq 0$. On the other hand, from (8.10) we have

$$\begin{aligned} \sin^2\theta_k &\leq 0.204605\frac{1+\mu^2}{\delta^2}(a_k^2+b_k^2), \quad k\in\mathcal{S}', \\ |\xi_k \sin(2\theta_k)| &\leq \sqrt{0.1117145\cdot 0.81842}\frac{1+\mu^2}{\delta^2}b_k\sqrt{a_k^2+b_k^2} \\ &\leq 0.302373\frac{1+\mu^2}{\delta^2}(a_k^2+b_k^2), \quad k\in\mathcal{S}'. \end{aligned}$$

Hence

$$\max\{\sin^2\phi_k, \sin^2\phi_k\} \leq 0.506978\frac{1+\mu^2}{\delta^2}(a_k^2+b_k^2), \quad k\in\mathcal{S}'. \quad (8.24)$$

If $p = n$, using (8.15), we obtain 4 times smaller constant in the bound for $\xi_k^2$. Hence, instead of the constants 0.690869 and 0.5400915, in the relation (8.17) we have 0.506978 and 0.355792, respectively. Therefore, the assertions of Lemma 3.7 can be written in the form

$$\sum_{k=0}^{N-1}{}' \sin^2\omega_k \leq \left\{ \begin{array}{l} 0.445\frac{1+\mu^2}{\delta^2}\epsilon_0^2, \ p<n \\ 0.312\frac{1+\mu^2}{\delta^2}\epsilon_0^2, \ p=n \end{array} \right\}, \ k\in\mathcal{S}'. \quad (8.25)$$

- Let us check whether Lemma 3.8 holds for the real HZ method. By inspecting its proof, we see that the critical checkpoint is relation (8.21). Using relation (4.32) it is easy to see that (8.21) holds for the real method. The rest of the proof is trivial for the real method. Only, the constant 0.690869 can be replaced by 0.506978. So, Lemma 3.8 holds for the real method.

- Let us now check whether Theorem 4.1 holds for the real method. We first consider the general cyclic pivot strategy. The proof is almost the same, only the final constant 0.932 in relation (4.8) can be replaced by 0.7725. Namely, we have $1.0442\sqrt{1.75383\cdot 0.312} < 0.7725$.

  Similar, for the row-cyclic strategy, the constant 0.932201 that appears in relation (4.22) can be reduced to 0.7567. This comes from the fact that the constant 0.4736138 from relation (4.20) can be replaced by 0.312.

- Furthermore, for the real method the constant 1.055 appearing in Corollary 4.2 can be reduced to 0.904. Namely, in the proof the constant 0.6058327 can be replaced by 0.445.

- Finally, the constants 0.5558 and 4345 in the statement of Corollary 4.3 can be reduced to 0.4079 and 0.2863. This comes from replacing the constants 0.690869, 0.5400915 in the proof of the corollary by 0.506978, 0.355792, respectively. □

Dear Editor,

Enclosed please find my manuscript ``On the Quadratic Convergence of the Complex HZ Method for the
Positive Definite Generalized Eigenvalue Problem''. Please consider it for publication in journal Linear
Algebra and Its Applications.

With best regards

Vjeran Hari

Professor at Department of Mathematics
Faculty of Science
University of Zagreb
Bijenicka cesta 30
10000 Zagreb
Croatia
tel.: □1-4605748
email: hari@math.hr