ON THE GLOBAL CONVERGENCE OF THE COMPLEX HZ METHOD*

VJERAN HARI†

Abstract. The paper considers a Jacobi method for solving the generalized eigenvalue problem $Ax = \lambda Bx$, where A and B are complex Hermitian matrices and B is positive definite. The method is a proper generalization of the standard Jacobi method for the Hermitian matrix A to the matrix pair (A, B). The paper derives the method and proves its global convergence under the large class of generalized serial pivot strategies. If both matrices are positive definite, it can be implemented as a one-sided method. It then solves the initial problem as the generalized singular value problem. Its main application is to serve as a kernel algorithm in a block Jacobi method for the same problem with large matrices A and B. The block Jacobi methods are methods of choice on contemporary CPU and GPU computing architectures. The proposed algorithm is very efficient on pairs of almost diagonal matrices, and diagonalization of such matrices is the main task of the kernel algorithm. The numerical tests indicate the high relative accuracy of the method on certain pairs of positive definite matrices.

Key word. Generalized eigenvalue problem, Jacobi method, global convergence

AMS subject classifications. 65F15, 65G05

1. INTRODUCTION.

We consider the positive definite generalized eigenvalue problem (PGEP)

$$Ax = \lambda Bx\,, \quad x \neq 0$$

with full Hermitian matrices A, B of order n, such that B is positive definite.

On contemporary parallel CPU and GPU computing machines block Jacobi methods have proved to be the methods of choice for solving that problem [18, 20]. In the core of those block methods lies the 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, typically of order 32–256, they are Hermitian and if B (or A) is positive definite then $\hat{B}$ ($\hat{A}$) is also such. The main task for a kernel algorithm is to solve PGEP with matrices $\hat{A}$, $\hat{B}$ accurately and efficiently. During the computation the block pivot submatrices will be most of the time nearly diagonal. So, the kernel algorithm has to perform its task quickly and accurately on such matrices. These two requirements are well met by the element-wise Jacobi methods for the PGEP. This raises the question of what is really known about complex Jacobi methods for the PGEP?

To this date, we know of three Jacobi methods for PGEP. These are the complex Falk-Langemeyer method [11], the complex Cholesky-Jacobi method [10, 14] and the complex HZ method [6]. All three methods simultaneously diagonalize the pivot submatrices at each step. Let us briefly highlight the main characteristics of these methods.

The first one is the proper generalization of the real Falk-Langemeyer (FL) method [3, 21, 16] to complex matrices. The method is characterized by the requirement that the transformation matrix has unit diagonal. That ensures simpler transformation formulas and application of BLAS1 `caxpy` and `zaxpy` computational routines. Additional accuracy can be obtained if the floating-point fused multiply and add operation

*Submitted to the editors May 30, 2019.

Funding: This work has been fully supported by Croatian Science Foundation under the project: IP_09_2014_3670

†Faculty of Science, Department of Mathematics, University of Zagreb, Bijenička cesta 30, 10000 Zagreb

is used, computing $\alpha\beta+\gamma$ with a single round. The shortcoming of the FL method lies in the fact that the norms of iteration matrices $A^{(k)}$ and $B^{(k)}$ can increase. So, periodically, one has to check those norms and apply some appropriate congruence transformation to "normalize" them. This slows down the computation, especially on distributed memory parallel machines. Namely, each check for renormalization costs. There is no simple rule when to apply that procedure because timing depends on the characteristics of the matrices. Also, the global and quadratic convergence of the complex method have not been proved. Numerical tests indicate the high relative accuracy of the method on "well-behaved" pairs of positive definite matrices. These are the pairs (A,B) for which the spectral condition numbers of $\kappa_2(D_A A D_A)$ and $\kappa_2(D_B B D_B)$ are small for some diagonal matrices D_A and D_B.

The complex Cholesky-Jacobi (CJ) method was introduced in [10] and its global convergence has been proved in [14]. It is a proper generalization of the real CJ method from [9]. Numerical tests imply the great potential of that method, in the first place for its presumably high relative accuracy on well-behaved pairs of positive definite matrices. It is a pretty new method, so it was less researched.

The third method is one we deal with in this paper. It is a direct generalization of the real one from [9]. Actually, the complex and real methods were derived and analyzed already in [6]. The real method was later used by Novaković et all [18] and was named "Hari-Zimmermann variant of the Falk-Langemeyer method". Later, in [9] we called it simply the HZ method. In [6] the complex HZ method was derived and its asymptotic quadratic convergence was proved under the general cyclic and the serial pivot strategies. In the sequel HZ (FL, CJ) method will mean the complex HZ (FL, CJ) method.

Like the FL method, the HZ method diagonalizes the pivot submatrices at each step. However, instead of simplifying the transformation matrix it simplifies the iteration matrices $B^{(k)}$ by requiring that they have unit diagonal. So, a preliminary step for the HZ method is needed to reduce the diagonal elements of B to ones. This is accomplished by the diagonal congruence transformation

$$A \mapsto A^{(0)} = DAD, \quad B \mapsto B^{(0)} = DBD, \quad D = \operatorname{diag}(B)^{-\frac{1}{2}}. \tag{1.1}$$

Then $(A^{(0)}, B^{(0)})$ is taken as the initial pair for the HZ method. The method preserves the unit diagonal of $B^{(k)}$ for $k \geq 0$ which stabilizes the iterative process. Namely, each $B^{(k)}$ is already almost optimally symmetrically scaled that can be made by a diagonal matrix [22], i.e. $\kappa_2(B^{(k)}) \approx \min_{D_B} \kappa_2(D_B B^{(k)} D_B)$. This also means that the HZ method has no problem with renormalizations. It is a proper generalization of the standard Jacobi method for Hermitian matrices. The principal shortcoming of HZ is that its transformations are slightly more expensive. Compared to the FL method this is no drawback and numerical tests of the real and complex methods on large matrices, using parallel machines [18, 20], have confirmed the advantage of the HZ approach. Here we derive the HZ method and prove its global convergence.

The paper is divided into 5 sections. In Section 2, we briefly describe the method. In Section 3 we derive the HZ algorithm, which determines one step of the method. Here we also define the global and quadratic convergence and provide a numerical example that sheds some light on accuracy and quadratic convergence of the method. In Section 4, we prove the global convergence of the method under the large class of generalized serial strategies from [13]. In Section 5, we point out some open problems and anticipate future work.

2. Description of the Method. Let A and B be complex Hermitian matrices of order n and let B be positive definite. The HZ method is the iterative process of the form

$$A^{(k+1)} = Z_k^* A^{(k)} Z_k, \qquad B^{(k+1)} = Z_k^* B^{(k)} Z_k, \qquad k \geq 0, \tag{2.1}$$

where $A^{(0)}$ and $B^{(0)}$ are defined by relation (1.1). In (2.1) each transformation matrix Z_k is *elementary plane matrix*. It is a nonsingular matrix which differs from the identity matrix I_n in one principal submatrix $\hat{Z}_k$,

$$\hat{Z}_k = Z_k([ij],[ij]) = \begin{bmatrix} z_{ii}^{(k)} & z_{ij}^{(k)} \\ z_{ji}^{(k)} & z_{jj}^{(k)} \end{bmatrix}, \quad k \geq 0, \tag{2.2}$$

where we used MATLAB notation. The subscripts $i = i(k)$, $j = j(k)$ are called *pivot indices*, (i,j) is *pivot pair* and $\hat{Z}_k$ is *pivot submatrix* of Z_k. If $\hat{Z}_k$ is as in (2.2), we shall briefly denote it by $\hat{Z}_k = (z_{ij}^{(k)})$. The transition $(A^{(k)}, B^{(k)}) \mapsto (A^{(k+1)}, B^{(k+1)})$ is called the kth *step* of the method. The way of selecting pivot pairs is a *pivot strategy*. The most common (pivot) strategies are the column- and row-cyclic ones. In the column-cyclic strategy the pivot pair repeatedly runs through the sequence of $N = n(n-1)/2$ pairs:

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

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

$$(A^{((t-1)N)}, B^{((t-1)N)}) \mapsto (A^{(tN)}, B^{(tN)})$$

is called the tth *cycle* or *sweep* of the method. In [13] the set of serial pivot strategies has been enlarged to the set of *generalized serial strategies*. The global convergence of general Jacobi processes under the generalized serial strategies were considered in [13], and the obtained results were used in [9, 14].

The algorithm for computing the elements of $\hat{Z}_k$ has been derived in [6]. It is based on the following theorem, which is a generalization to complex matrices, of the Gose's result [4].

THEOREM 2.1 ([7]). *Let $\hat{B} = (b_{ij})$ and $\hat{B}' = diag(b'_{ii}, b'_{jj})$ be positive definite Hermitian matrices of order two. Then there exist a nonsingular matrix $\hat{F}$ of order two, such that $\hat{B}' = \hat{F}^* \hat{B} \hat{F}$. Each $\hat{F}$ satisfying that property has the form*

$$\hat{F} = \frac{1}{\cos\gamma} \begin{bmatrix} \frac{1}{\sqrt{b_{ii}}} & \\ & \frac{1}{\sqrt{b_{jj}}} \end{bmatrix} \begin{bmatrix} \cos\phi & e^{\imath\alpha}\sin\phi \\ -e^{-\imath\beta}\sin\psi & \cos\psi \end{bmatrix} \begin{bmatrix} e^{\imath\omega_i}\sqrt{b'_{ii}} & \\ & e^{\imath\omega_j}\sqrt{b'_{jj}} \end{bmatrix},$$

where ω_i, ω_j are real, $\phi, \psi, \gamma \in [0\,, \frac{\pi}{2}]$, and

$$\sin\gamma = \frac{|b_{ij}|}{\sqrt{b_{ii}b_{jj}}}, \qquad \cos\gamma = |\cos\phi\cos\psi + e^{\imath(\alpha-\beta)}\sin\phi\sin\psi|$$

holds.

To simplify $\hat{F}$, we can require that $\omega_i = \omega_j = 0$, i.e. that the diagonal elements of $\hat{F}$ are real and nonnegative. Furthermore, by replacing α, β by $\alpha+\pi$, $\beta+\pi$, respectively, we can move the $-$ sign from $-e^{-\imath\beta}\sin\psi$ to $e^{\imath\alpha}\sin\phi$.

3. Derivation of the HZ algorithm. As has been described earlier, the initial step (1.1) makes the diagonal elements of $B^{(0)}$ equal to one. The method is designed to retain that property. We shall consider step k of the method. To simplify notation, we omit the superscript k, denote the current matrices by $A = (a_{rs})$, $B = (b_{rs})$ and those obtained after completing step k by $A' = (a'_{rs})$, $B' = (b'_{rs})$. The pivot submatrices are denoted by $\hat{A} = (a_{ij})$, $\hat{B} = (b_{ij})$, where i, j are pivot indices. We assume $b_{ii} = 1$ and $b_{jj} = 1$. The transformation matrix is denoted by Z and its pivot submatrix by $\hat{Z}$.

We shall construct $\hat{Z}$ such that the following conditions hold

$$a'_{ij} = 0, \quad b'_{ij} = 0, \qquad b'_{ii} = 1, \quad b'_{jj} = 1, \qquad z_{ii} \geq 0, \quad z_{jj} \geq 0,$$

Since $b_{ii} = b_{jj} = 1$, Theorem 2.1 shows that $\hat{Z}$ can be sought in the form

$$\text{(3.1)} \qquad \hat{Z} = \frac{1}{\sqrt{1 - |b_{ij}|^2}} \begin{bmatrix} \cos\phi & -e^{\imath\alpha}\sin\phi \\ e^{-\imath\beta}\sin\psi & \cos\psi \end{bmatrix}, \quad \phi, \psi \in [0\,,\, \frac{\pi}{2}].$$

Let us recall the formulas linked to the complex Jacobi rotation which diagonalizes the Hermitian matrix $\hat{H} = (h_{ij})$ of order 2. If we write c_ϑ, s_ϑ for $\cos(\vartheta)$, $\sin(\vartheta)$, respectively, then from the equation

$$\begin{bmatrix} c_\vartheta & e^{\imath\varsigma}s_\vartheta \\ -e^{-\imath\varsigma}s_\vartheta & c_\vartheta \end{bmatrix} \begin{bmatrix} h_{ii} & h_{ij} \\ \bar{h}_{ij} & h_{jj} \end{bmatrix} \begin{bmatrix} c_\vartheta & -e^{\imath\varsigma}s_\vartheta \\ e^{-\imath\varsigma}s_\vartheta & c_\vartheta \end{bmatrix} = \begin{bmatrix} h'_{ii} & 0 \\ 0 & h'_{jj} \end{bmatrix},$$

one obtains

$$\varsigma = \arg(h_{ij}), \qquad \tan(2\vartheta) = \frac{2|h_{ij}|}{h_{ii} - h_{jj}}$$

and

$$h'_{ii} = h_{ii} + |h_{ij}|\tan(\vartheta), \qquad h'_{jj} = h_{jj} - |h_{ij}|\tan(\vartheta).$$

In these formulas the angle ϑ need not be restricted to $[-\pi/4\,,\,\pi/4]$.

To derive $\hat{Z}$, we follow the lines from [6]. The matrix $\hat{Z}$ sought for in the form

$$\text{(3.2)} \qquad \hat{Z} = \hat{R}_1\hat{D}\hat{R}_2\hat{\Phi},$$

where $\hat{R}_1$, $\hat{R}_2$ are complex rotations and $\hat{D}$, $\hat{\Phi}$ are diagonal matrices, $\hat{\Phi}$ being also unitary. Let

$$\begin{aligned} \hat{A}_1 &= \hat{R}_1^*\hat{A}\hat{R}_1, & \hat{B}_1 &= \hat{R}_1^*\hat{B}\hat{R}_1, \\ \hat{A}_2 &= \hat{D}^*\hat{A}_1\hat{D}, & \hat{B}_2 &= \hat{D}^*\hat{B}_1\hat{D}, \\ \hat{A}_3 &= \hat{R}_2^*\hat{A}_2\hat{R}_2, & \hat{B}_3 &= \hat{R}_2^*\hat{B}_2\hat{R}_2, \\ \hat{A}' &= \hat{\Phi}^*\hat{A}_3\hat{\Phi}, & \hat{B}' &= \hat{\Phi}^*\hat{B}_3\hat{\Phi}, \end{aligned}$$

and note that

$$\hat{A}' = \hat{Z}^*\hat{A}\hat{Z}, \qquad \hat{B}' = \hat{Z}^*\hat{B}\hat{Z}.$$

The complex rotation $\hat{R}_1$ has the role of Jacobi rotation which diagonalizes $\hat{B}$. Since the diagonal elements of $\hat{B}$ are equal to 1, the rotation angle can be chosen as $\pm\pi/4$. Choosing it to be $-\pi/4$, we obtain

$$\text{(3.3)} \qquad \hat{R}_1 = \begin{bmatrix} \cos(-\frac{\pi}{4}) & -e^{\imath\beta_{ij}}\sin(-\frac{\pi}{4}) \\ e^{-\imath\beta_{ij}}\sin(-\frac{\pi}{4}) & \cos(-\frac{\pi}{4}) \end{bmatrix} = \frac{\sqrt{2}}{2}\begin{bmatrix} 1 & e^{\imath\beta_{ij}} \\ -e^{-\imath\beta_{ij}} & 1 \end{bmatrix},$$

where

$$\beta_{ij} = \arg(b_{ij}). \tag{3.4}$$

The diagonal elements of $\hat{B}_1$ are no longer equal to 1, so the transformation with $\hat{D}$ is used to make them 1 again. We have

$$\hat{B}_1 = \begin{bmatrix} 1-|b_{ij}| & 0 \\ 0 & 1+|b_{ij}| \end{bmatrix}, \quad \hat{D} = \begin{bmatrix} 1/\sqrt{1-|b_{ij}|} & 0 \\ 0 & 1/\sqrt{1+|b_{ij}|} \end{bmatrix}. \tag{3.5}$$

Now, we have obtained $\hat{B}_2 = I_2$. Since $\hat{R}_2$ and $\hat{\Phi}$ are unitary, we have $\hat{B}' = \hat{B}_3 = I_2$.

To determine $\hat{R}_2$ and $\hat{\Phi}$, we have to compute $\hat{A}_2$. One easily obtains

$$\hat{A}_2 = \begin{bmatrix} \frac{1}{1-|b_{ij}|}\left(\frac{a_{ii}+a_{jj}}{2} - u_{ij}\right) & \frac{e^{\imath\beta_{ij}}}{\sqrt{1-|b_{ij}|^2}}\left(\frac{a_{ii}-a_{jj}}{2} + \imath w_{ij}\right) \\ \frac{e^{-\imath\beta_{ij}}}{\sqrt{1-|b_{ij}|^2}}\left(\frac{a_{ii}-a_{jj}}{2} - \imath w_{ij}\right) & \frac{1}{1+|b_{ij}|}\left(\frac{a_{ii}+a_{jj}}{2} + u_{ij}\right) \end{bmatrix}, \tag{3.6}$$

where

$$u_{ij} + \imath w_{ij} = e^{-\imath\beta_{ij}} a_{ij}, \qquad u_{ij}, v_{ij} \in \mathbf{R}. \tag{3.7}$$

The matrix R_2 is chosen as complex Jacobi rotation which diagonalizes $\hat{A}_2$. We write it in the form

$$\hat{R}_2 = \begin{bmatrix} \cos(\theta + \frac{\pi}{4}) & -e^{\imath\alpha_{ij}} \sin(\theta + \frac{\pi}{4}) \\ e^{-\imath\alpha_{ij}} \sin(\theta + \frac{\pi}{4}) & \cos(\theta + \frac{\pi}{4}) \end{bmatrix}. \tag{3.8}$$

From the relation (3.6) we obtain

$$\tan(2(\theta + \frac{\pi}{4})) = \frac{\frac{2}{\sqrt{1-|b_{ij}|^2}}\left|\frac{a_{ii}-a_{jj}}{2} + \imath w_{ij}\right|}{\frac{1}{1-|b_{ij}|}\left(\frac{a_{ii}+a_{jj}}{2} - u_{ij}\right) - \frac{1}{1+|b_{ij}|}\left(\frac{a_{ii}+a_{jj}}{2} + u_{ij}\right)}$$

$$= \frac{\sqrt{1-|b_{ij}|^2}\ |\ a_{ii} - a_{jj} + 2\imath w_{ij}\ |}{(a_{ii}+a_{jj})|b_{ij}| - 2u_{ij}},\ \ \theta + \frac{\pi}{4} \in [-\pi/4, \pi/4],$$

$$\alpha_{ij} = \beta_{ij} + \arg\left(\frac{a_{ii}-a_{jj}}{2} + \imath w_{ij}\right). \tag{3.9}$$

Note that

$$e^{\imath\alpha_{ij}} \sin(\theta + \frac{\pi}{4}) = e^{\imath(\alpha_{ij} + (1-\sigma_{ij})\frac{\pi}{2})}(\sigma_{ij} \sin(\theta + \frac{\pi}{4})), \quad \sigma_{ij} \in \{-1, 1\}.$$

Hence adding $(1-\sigma_{ij})\frac{\pi}{2}$ to α_{ij} implies changing $\theta + \frac{\pi}{4}$ to $\sigma_{ij}(\theta + \frac{\pi}{4})$ in the relation (3.8). For $\sigma_{ij} = -1$ it means that $\tan(\theta + \frac{\pi}{4})$ and $\tan(2(\theta + \frac{\pi}{4}))$ change the sign. The value of σ_{ij} is determined from the requirement

$$-\frac{\pi}{2} \le \alpha_{ij} - \beta_{ij} \le \frac{\pi}{2}, \tag{3.10}$$

which is used in the global convergence proof. From the relation (3.9) one concludes that

$$\sigma_{ij} = \begin{cases} 1, & a_{ii} - a_{jj} \ge 0, \\ -1, & a_{ii} - a_{jj} < 0. \end{cases} \tag{3.11}$$

Since $\tan(2\theta + \pi/2) = -1/\tan(2\theta)$, we obtain

$$\tan(2\theta) = \sigma_{ij} \frac{2u_{ij} - (a_{ii} + a_{jj})|b_{ij}|}{\sqrt{1 - |b_{ij}|^2}\sqrt{(a_{ii} - a_{jj})^2 + 4v_{ij}^2}}, \quad -\frac{\pi}{4} \leq \theta \leq \frac{\pi}{4} \tag{3.12}$$

and

$$\alpha_{ij} = \beta_{ij} + \arg\left(\frac{a_{ii} - a_{jj}}{2} + \imath v_{ij}\right) + (1 - \sigma_{ij})\frac{\pi}{2}. \tag{3.13}$$

This choice of σ_{ij} also ensures that this complex algorithm is a proper generalization of the real HZ algorithm from [9]. Indeed, if all matrices are real, we have $u_{ij} = a_{ij}$, $v_{ij} = 0$ and $\sigma_{ij}\sqrt{(a_{ii} - a_{jj})^2 + 4v_{ij}^2} = a_{ii} - a_{jj}$ and the complex algorithm reduces to the real one.

From Theorem 2.1 (together with the comment regarding the $-$ sign in $(1,2)$-element of $\hat{F}$), and from the fact that $b_{ii} = b_{jj} = 1 = b'_{ii} = b'_{jj}$, we conclude that the general form of $\hat{F}$ that reduces $\hat{B}$ to I_2 reads

$$\hat{F} = \frac{1}{\sqrt{1 - |b_{ij}|^2}} \begin{bmatrix} \cos\phi & -e^{\imath\alpha}\sin\phi \\ e^{-\imath\beta}\sin\psi & \cos\psi \end{bmatrix} \begin{bmatrix} e^{\imath\omega_i} & \\ & e^{\imath\omega_j} \end{bmatrix}, \tag{3.14}$$

where

$$\cos\phi \geq 0, \quad \cos\psi \geq 0, \quad \sin\phi \geq 0, \quad \sin\psi \geq 0. \tag{3.15}$$

Let $\hat{G} = \hat{R}_1\hat{D}\hat{R}_2$. Then $\hat{G}^*\hat{A}\hat{G}$ is diagonal and $\hat{G}^*\hat{B}\hat{G} = I_2$. So, $\hat{G}$ can be represented as $\hat{F}$ from the relations (3.14)–(3.15). If we find that representation of $\hat{G}$, we can set $\hat{\Phi} = \operatorname{diag}(e^{-\imath\omega_i},\ e^{-\imath\omega_j})$ and work with the transformation $\hat{G}\hat{\Phi}$. In other words, $\hat{Z} = \hat{G}\hat{\Phi}$ will be the matrix from the relation (3.1).

From the relations (3.3), (3.5), (3.8), we have

$$\hat{G} = \frac{1}{2} \begin{bmatrix} \frac{1}{\sqrt{1-|b_{ij}|}} & \frac{e^{\imath\beta_{ij}}}{\sqrt{1+|b_{ij}|}} \\ -\frac{e^{-\imath\beta_{ij}}}{\sqrt{1-|b_{ij}|}} & \frac{1}{\sqrt{1+|b_{ij}|}} \end{bmatrix} \begin{bmatrix} c - s & -e^{\imath\alpha_{ij}}(c+s) \\ e^{-\imath\alpha_{ij}}(c+s) & c - s \end{bmatrix}, \tag{3.16}$$

where c and s stand for $\cos\theta$ and $\sin\theta$, respectively. Let $\hat{G} = (g_{ij})$. After a simple calculation, one obtains

$$g_{ii} = \frac{1}{\sqrt{1 - |b_{ij}|^2}} \frac{1}{2} \left[\sqrt{1 + |b_{ij}|}(c - s) + e^{\imath(\beta_{ij} - \alpha_{ij})}\sqrt{1 - |b_{ij}|}(c + s)\right],$$

$$g_{ij} = \frac{1}{\sqrt{1 - |b_{ij}|^2}} \frac{1}{2} \left[-e^{\imath\alpha_{ij}}\sqrt{1 + |b_{ij}|}(c + s) + e^{\imath\beta_{ij}}\sqrt{1 - |b_{ij}|}(c - s)\right],$$

$$g_{ji} = \frac{1}{\sqrt{1 - |b_{ij}|^2}} \frac{1}{2} \left[-e^{-\imath\beta_{ij}}\sqrt{1 + |b_{ij}|}(c - s) + e^{-\imath\alpha_{ij}}\sqrt{1 - |b_{ij}|}(c + s)\right],$$

$$g_{jj} = \frac{1}{\sqrt{1 - |b_{ij}|^2}} \frac{1}{2} \left[e^{\imath(\alpha_{ij} - \beta_{ij})}\sqrt{1 + |b_{ij}|}(c + s) + \sqrt{1 - |b_{ij}|}(c - s)\right].$$

Let us equate $\hat{G} = \hat{F}$, where $\hat{F}$ is from the relation (3.14). Comparing the elements of $\hat{F}$ with the elements g_{ii}, g_{ij}, g_{ji}, g_{jj} of $\hat{G}$ and taking into account the conditions

(3.15), we obtain

$$(3.17)\quad \begin{cases} 2\cos^2\phi &= 1-|b_{ij}|\sin(2\theta)+\sqrt{1-|b_{ij}|^2}\cos(2\theta)\cos(\alpha_{ij}-\beta_{ij}),\\ 2\sin^2\phi &= 1+|b_{ij}|\sin(2\theta)-\sqrt{1-|b_{ij}|^2}\cos(2\theta)\cos(\alpha_{ij}-\beta_{ij}),\\ 2\cos^2\psi &= 1+|b_{ij}|\sin(2\theta)+\sqrt{1-|b_{ij}|^2}\cos(2\theta)\cos(\alpha_{ij}-\beta_{ij}),\\ 2\sin^2\psi &= 1-|b_{ij}|\sin(2\theta)-\sqrt{1-|b_{ij}|^2}\cos(2\theta)\cos(\alpha_{ij}-\beta_{ij}). \end{cases}$$

Since we want positive $\cos\phi$ and $\cos\psi$ in $\hat{Z}$, it suffices to apply the square root to the appropriate equations in (3.17).

It remains to determine $e^{\imath\omega_i}$, $e^{\imath\omega_j}$, $e^{\imath\alpha}$ and $e^{-\imath\beta}$. Obviously, ω_i and ω_j will be the arguments of g_{ii} and g_{jj}. This implies

$$(3.18)\quad \begin{cases} e^{\imath\omega_i} &= [\sqrt{1+|b_{ij}|}(c-s)+e^{\imath(\beta_{ij}-\alpha_{ij})}\sqrt{1-|b_{ij}|}(c+s)]/(2\cos\phi),\\ e^{\imath\omega_j} &= [e^{\imath(\alpha_{ij}-\beta_{ij})}\sqrt{1+|b_{ij}|}(c+s)+\sqrt{1-|b_{ij}|}(c-s)]/(2\cos\psi). \end{cases}$$

Finally, $e^{\imath\alpha}$ and $e^{-\imath\beta}$ will be obtained from the relations

$$\begin{aligned} e^{\imath\alpha}e^{\imath\omega_j} &= [e^{\imath\alpha_{ij}}\sqrt{1+|b_{ij}|}(c+s)-e^{\imath\beta_{ij}}\sqrt{1-|b_{ij}|}(c-s)]/(2\sin\phi),\\ e^{-\imath\beta}e^{\imath\omega_i} &= [-e^{-\imath\beta_{ij}}\sqrt{1+|b_{ij}|}(c-s)+e^{-\imath\alpha_{ij}}\sqrt{1-|b_{ij}|}(c+s)]/(2\sin\psi). \end{aligned}$$

These two relations together with (3.18) imply

$$(3.19)\quad \begin{cases} e^{\imath\alpha} &= \frac{e^{\imath\beta_{ij}}}{2\sin\phi\cos\psi}[\sin(2\theta)+|b_{ij}|+\imath\sqrt{1-|b_{ij}|^2}\cos(2\theta)\sin(\alpha_{ij}-\beta_{ij})],\\ e^{-\imath\beta} &= \frac{e^{-\imath\beta_{ij}}}{2\sin\psi\cos\phi}[\sin(2\theta)-|b_{ij}|-\imath\sqrt{1-|b_{ij}|^2}\cos(2\theta)\sin(\alpha_{ij}-\beta_{ij})]. \end{cases}$$

To obtain the off-diagonal elements $e^{\imath\alpha}\sin\phi$ and $e^{-\imath\beta}\sin\psi$, it remains to remove $\sin\phi$ and $\sin\psi$ from the denominators on the right-hand sides of (3.19).

Since $\hat{B}'=I_2$ and $a'_{ij}=0$, it remains to find the expressions for a'_{ii} and a'_{jj}. After that it is easy to apply $\hat{Z}$ ($\hat{Z}^*$) to the appropriate columns (rows) of A and B and thus complete the current iteration step on the pair (A,B). For the diagonal elements we obtain

$$(3.20)\quad \begin{cases} a'_{ii} &= [\cos^2\phi a_{ii}+\sin^2\psi a_{jj}+2\cos\phi\sin\psi\,\Re(e^{-\imath\beta}a_{ij})]/(1-|b_{ij}|^2),\\ a'_{jj} &= [\sin^2\phi a_{ii}+\cos^2\psi a_{jj}-2\cos\psi\sin\phi\,\Re(e^{-\imath\alpha}a_{ij})]/(1-|b_{ij}|^2). \end{cases}$$

It remains to consider the case when $\tan(2\theta)$ has the form 0/0. This happens if and only if

$$a_{ii}=a_{jj},\qquad v_{ij}=0,\qquad e^{-\imath\beta_{ij}}a_{ij}=u_{ij}=a_{ii}|b_{ij}|\,.$$

These 4 conditions are equivalent to

$$(3.21)\qquad a_{ii}=a_{jj},\qquad a_{ij}=a_{ii}b_{ij}.$$

If the conditions in (3.21) hold then we have $\hat{A}=a_{ii}\hat{B}$ and we choose $\theta=0$, $\alpha_{ij}=\beta_{ij}$. In that case we have

$$(3.22)\quad \hat{Z}=\frac{1}{\tau}\begin{bmatrix}\rho & -\xi\\ -\bar{\xi} & \rho\end{bmatrix},\ \xi=\frac{b_{ij}}{2\rho},\ \rho=\frac{\sqrt{1+|b_{ij}|}+\sqrt{1-|b_{ij}|}}{2},\ \tau=\sqrt{1-|b_{ij}|^2}$$

and that matrix $\hat{Z}$ is a direct extension of the real one from one from [9, Section 2.3]. In this case we have $a'_{ii} = a_{ii}$ and $a'_{jj} = a_{jj}$.

Let us make a comment on accuracy issues. In a similar way as in [17, Section 3.2] one can show that setting $\hat{B}' = I_2$ is numerically safe, i. e. 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 have to be the case with a'_{ii}, a'_{jj} and a'_{ij}. Numerical tests show that it is better to compute all those elements. Therefore we provide yet a formula for computing a'_{ij}:

$$(3.23) \qquad a'_{ij} = [\cos\phi\cos\psi a_{ij} + (a_{jj}e^{\imath\beta}\cos\psi\sin\psi - a_{ii}e^{\imath\alpha}\cos\phi\sin\phi) \\ - \bar{a}_{ij}e^{\imath(\alpha+\beta)}\sin\phi\sin\psi\ ]/(1-|b_{ij}|^2).$$

In the later stage of the process, $|a_{ij}|$ will be small and $|a'_{ij}|$ tiny. So, cancelation takes place. Then $\sin\phi$ and $\sin\psi$ will be small, but a_{ii} and a_{jj} can be large. So, we have used the parenthesis to contain maybe those large terms, whose sum will be canceled out with $\cos\phi\cos\psi a_{ij}$. The last term will be tiny since all of its factors will be small.

3.1. The complex HZ algorithm. Here, we organize the obtained formulas in the natural order to obtain the complex HZ algorithm, i. e. the algorithm of one step of the method. Input to the algorithm is the pair of pivot submatrices, i. e. the matrices $\hat{A}$, $\hat{B}$,

$$\hat{A} = \begin{bmatrix} a_{ii} & a_{ij} \\ \bar{a}_{ij} & a_{jj} \end{bmatrix}, \qquad \hat{B} = \begin{bmatrix} 1 & b_{ij} \\ \bar{b}_{ij} & 1 \end{bmatrix},$$

and output consists of the pivot submatrix $\hat{Z}$ of the transformation matrix Z,

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

and of $\hat{A}'$.

In the pseudocode below, $\Re(\omega)$, $\Im(\omega)$, and $\mathrm{conj}(\omega)$ denote the real, imaginary, and complex conjugate of $\omega \in \mathbf{C}$. The names of variables in the pseudocode are linked with names in our mathematical analysis as follows: $t2$, $cs2$, $sn2$, csg, sng stand for $\tan(2\theta)$, $\cos(2\theta)$, $\sin(2\theta)$, $\cos(\alpha_{ij}–\beta_{ij})$, $\sin(\alpha_{ij}–\beta_{ij})$, respectively.

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 which diagonalizes $\hat{A}$.

If in addition $a_{ij} = 0$, then $u = v = sng = t2 = sn2 = 0$, hence Z is the identity matrix.

Finally, if the eigenvectors are wanted, one can set $F^{(0)} = D$, where D is from the relation (1.1), and in each step k, $k \geq 0$, update it: $F^{(k+1)} = F^{(k)}Z_k$. In 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).

Below is a simple pseudocode of the algorithm. It can be "updated" by the formulas (3.20) and (3.23), although the simple one below works quite well.

Algorithm 3.1 The complex HZ algorithm

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

3.2. On the convergence and stopping criterion. To measure advancement of the method we use the quantity $S(A,B)$ defined by

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

where generally, $\|X\|_F = \sqrt{\mathrm{trace}(X^*X)}$ is the Frobenius norm of X. In the following standard convergence definitions A, B are Hermitian and B is positive definite.

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 Λ is a diagonal matrix of eigenvalues and I_n is the identity matrix. The method is *globally convergent* if it is convergent on every initial pair.

The cyclic method is asymptotically *quadratically convergent on the pair* (A, B) if it is convergent on (A, B) and there is a positive integer r_0 such that

$$S(A^{(rN)}, B^{(rN)}) \leq c_n S^2(A^{((r-1)N)}, B^{((r-1)N)}), \qquad r \geq r_0.$$

Here c_n is a constant which may depend on n. The method is *quadratically convergent on some set of matrix pairs* if it is quadratically convergent on every pair from that set.

From [6] we know that such a set consists of the matrix pairs whose eigenvalues are simple.

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

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

This condition is usually checked after completion of each cycle. If the method is high relative accurate on the considered matrix pair then this stopping criterion warrants high relative accuracy of the computed eigenvalues. This claim can be proved using the complex version of [2, Theorem 3.2] (see [11, Theorem 3.2]).

If A is not positive definite, we simply rely upon $S(A, B)$ and Theorem 4.3 for our stopping criterion.

3.3. A few numerical examples. We have used MATLAB to observe behavior of $S(A^{(k)}, B^{(k)})$ for all steps k until convergence, and to inspect accuracy of the computed eigenvalues. The following code was used to compute the initial matrix pair (A, B):

```
n=128; A=hilb(n);A=A-triu(A);A=gallery('minij',n)+eye(n)+1i*(A-A'); A=A+A';
B=rand(n)-1i*0.5*rand(n);  D=diag(logspace(-4,4,n));  B=D*(B'*B)*D; B=B+B';
```

Both matrices are of order 128 and they are positive definite. We have computed the condition numbers of the symmetrically scaled matrices

$$A_S = \mathrm{diag}(A)^{-1/2} A\, \mathrm{diag}(A)^{-1/2}, \qquad B_S = \mathrm{diag}(B)^{-1/2} B\, \mathrm{diag}(B)^{-1/2}.$$

We have obtained: $\kappa_2(A_S) \approx 8.7 \cdot 10^3$, $\kappa_2(B_S) \approx 4.9 \cdot 10^6$. Note that A_S and B_S have unit diagonal.

To gain an insight into the properties of the matrices A and B, we have displayed the following data in Figure 1: the quotient of the diagonal elements of A and B, and the eigenvalues of A, B and of the matrix pair (A, B).

Since the intrinsic MATLAB function `eig` did not compute the eigenvalues of B and of (A, B) with sufficient accuracy, we made the script `ABhermeig(A,B,dg)` which used *variable precision arithmetic* (vpa) with `dg` decimal digits. In `ABhermeig(A,B,dg)` we have used vpa with 32 decimal digits to compute the eigenvalues and eigenvectors of A, B and (A, B). The double precision matrices A and B are first converted to symbolic type, then the output data are computed using vpa, and before exit they are converted to double precision. During computation in vpa, a test is made to ensure that the output data are accurately computed. In particular, the spectral norm of the residual $\|AF - BF\Lambda\|_2 / \|AF\|_2$ is computed in vpa, where F is the matrix of

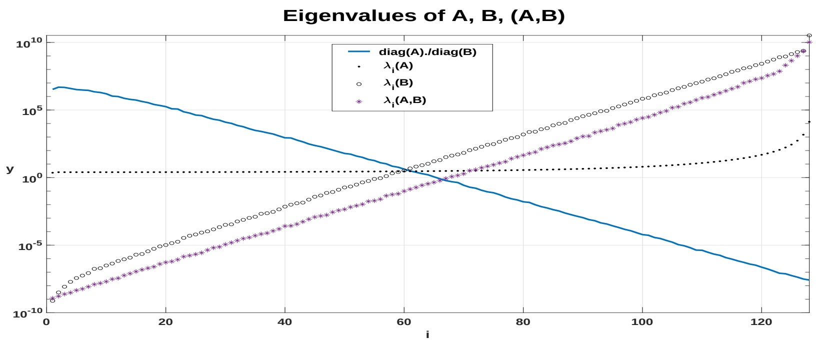

FIG. 1. *The graphs of the eigenvalues of A, B and (A, B)*

eigenvectors and Λ is the diagonal matrix of eigenvalues. In all cases the values of that quantity were smaller than $3 \cdot 10^{-27}$.

Now, that we have at disposal accurate eigenvalues of the pair (A, B), we can compute the relative errors of the eigenvalues computed by other scripts. To this end we have made the script `dsychz_qc(A,B,eivec)` which computes the eigenvalues and eigenvectors using the row-cyclic complex HZ method.

The same script has been used to check the quadratic convergence of the HZ method. The code lines follow the lines of the HZ algorithm as is presented above. The output to `dsychz_qc` are: the eigenvector matrix, the column-vector of eigenvalues, the total number of cycles and steps (`steps`), and matrix `qc`. The matrix `qc` has 5 columns each of length `steps`. The kth row of `qc` is obtained from step k. The columns of `qc` contain the values of $S(A_S^{(k)})$, $S(A^{(k)})$, $S(B^{(k)})$, $S(A^{(k)}, B^{(k)})$, $S(A_S^{(k)}, B^{(k)})$ in their kth component. It has been noticed that the value of $S(B^{(k)})$ is much larger than the values of $S(A_S^{(k)})$ and $S(A^{(k)})$ in the later stage of the process, so the values of $S(A^{(k)}, B^{(k)})$, $S(A_S^{(k)}, B^{(k)})$ are very close to $S(B^{(k)})$. Therefore, they are not depicted in Figure 2. Note that the values of $S(A_S^{(k)})$ and $S(B^{(k)})$ determine when to stop the process.

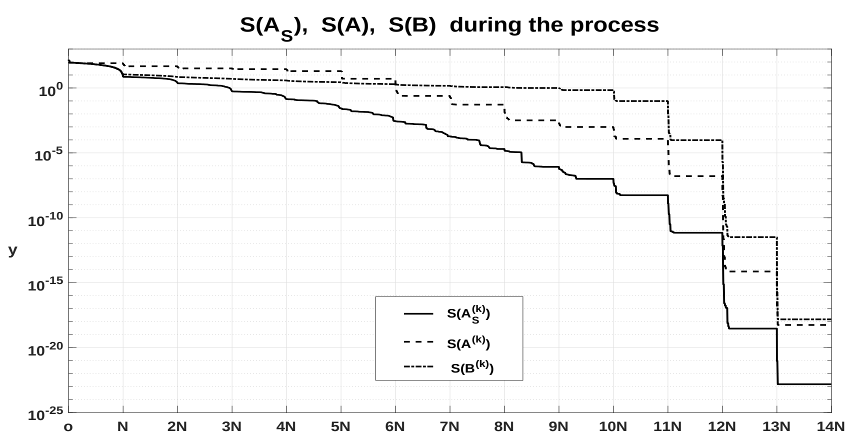

FIG. 2. *The reduction of $S(A_S^{(k)})$, $S(A^{(k)})$ and $S(B^{(k)})$*

We have labeled ticks on x-axis as multiples of N steps, where $N = 128(127)/2 = 8128$. Vertical grids are displayed in accordance with the ticks. We can observe the quadratic convergence behavior of all three functions in the later cycles. Once the quadratic convergence commences, a significant drop of values occurs after each cycle. The delay of the quadratic convergence of $S(A^{(k)}, B^{(k)})$ comes from the fact that $S(A^{(k)})$ and $S(B^{(k)})$ have their own rates of decrease, and when they become aligned $S(A^{(k)}, B^{(k)})$ strongly decreases. We speculate that slower convergence of $S(B^{(k)})$ is a consequence of fact that $\kappa_2(A_S) \ll \kappa_2(B_S)$.

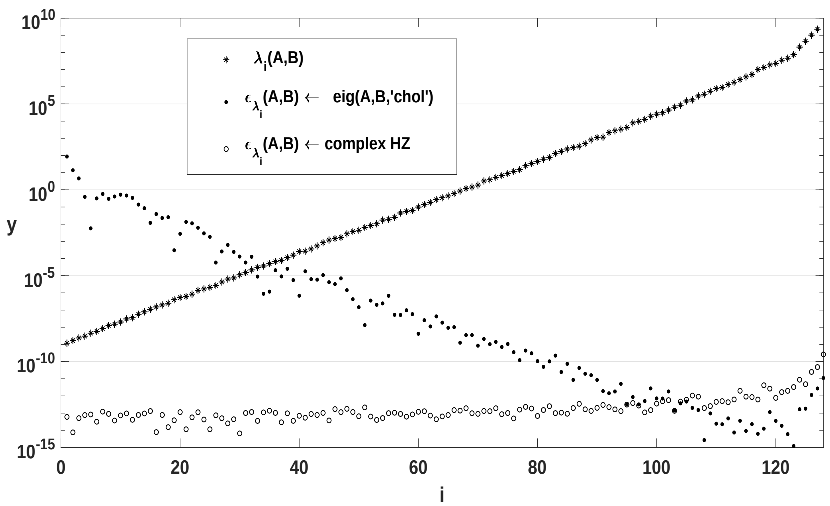

Fig. 3. *The relative errors of the eigenvalues computed by* `eig` *and by the HZ algorithm*

Figure 3 draws the relative errors of the eigenvalues computed by the HZ algorithm and (for comparison reasons) by the MATLAB `eig(A,B,'chol')` function. In the same figure we have added the graph of the eigenvalues of (A, B) to see if there is some correlation between magnitudes of the eigenvalues and the corresponding relative errors. We see that `eig(A,B,'chol')` computed the eigenvalues of the pair (A, B) with large relative errors. The HZ method computed them with high relative accuracy. This is in accordance with the behavior of the real HZ method [9, 17].

Then we have switched A and B in the matrix pair (A, B). The relative errors of the eigenvalues computed by the HZ method are even smaller, which reflects the fact that now B_S has smaller condition number. But in the same time, the relative errors of the eigenvalues computed by `eig(A,B,'chol')` become equally tiny. This seems to be a consequence of the fact that now the diagonal elements of $A^{(0)}$ (computed as `diag(A)./diag(B)`) are increasingly ordered along the diagonal of A. This interesting phenomenon of the QR algorithm was noticed and communicated to the author by Professor Marc Van Barel of Leuven University.

We have made several other numerical experiments and they all indicate that the complex HZ method appears to have high relative accuracy on well-behaved pairs of positive definite matrices. It has been noticed that number of cycles needed to reach the stopping criterion decreases when the algorithm is so modified that it tries to order the diagonal elements in the nonincreasing order (cf. [5, 1]).

We end this section with an example which shows behavior of the method when the matrix A is indefinite and the initial pair (A, B) has both multiple eigenvalues and clusters of eigenvalues. We shall not delve into the construction of the initial pair since it is described in [5], where the quadratic asymptotic convergence of the HZ method has been considered. We display the graphs of the functions $S(A_S^{(k)})$, $S(A^{(k)})$, $S(B^{(k)})$ and $S(A^{(k)}, B^{(k)})$ under the row-cyclic strategy and under the deRijk [1] strategy.

We shall display the most important data linked with (A, B). We have $n = 128$, $\kappa_2(A_S) \approx 5.1 \cdot 10^{11}$, $\kappa_2(B_S) \approx 9.97 \cdot 10^3$, the diagonal elements of $A^{(0)}$ are scattered in the interval $[-538.35, -365.33]$. The pair (A, B) has 10 eigenvalues of multiplicity 10, one cluster of 20 simple eigenvalues around 0 and 8 additional simple eigenvalues. The approximate values of the multiple (simple) eigenvalues are: -732.28, -574.80, -417.32, $-259, 84$, -102.36, 370.08, 527.56, 685.04, 842.58, 1000 (-1000.0, -984.25, -968.50, -952.76, -937.01, -921.26, -905.51, -8.8976). The cluster is made of the eigenvalues whose approximations are: $-4.7 \cdot 10^{-1}$, $-7.96 \cdot 10^{-2}$, $-4.2 \cdot 10^{-3}$, $-4.2 \cdot 10^{-4}$, $-9.7 \cdot 10^{-5}$, $-3.3 \cdot 10^{-5}$, $-1.1 \cdot 10^{-6}$, $-4.9 \cdot 10^{-7}$, $-1.97 \cdot 10^{-8}$, $-8.4 \cdot 10^{-9}$, $2.6 \cdot 10^{-8}$, $1.2 \cdot 10^{-7}$, $8.3 \cdot 10^{-7}$, $2.6 \cdot 10^{-5}$, $3.4 \cdot 10^{-4}$, $2.5 \cdot 10^{-3}$, $1.5 \cdot 10^{-2}$, $8.4 \cdot 10^{-2}$, $3.3 \cdot 10^{-1}$, 7.5. The relative accuracy of the computed eigenvalues has been computed and it is around 10^{-14}, with the exception of the eigenvalues which form the cluster. Their relative accuracy varies from 10^{-13} to $6.5 \cdot 10^{-6}$, the smaller the magnitude of an eigenvalue the lower the relative accuracy. The same can be said for the eigenvalues computed by `eig(A,B,'chol')`. In Figure 4 and Figure 5 are displayed the graphs of the functions. We can see failure of the asymptotic quadratic convergence.

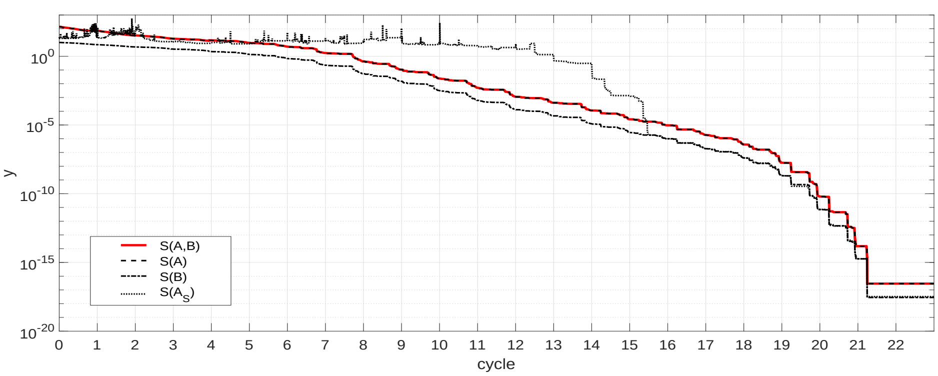

FIG. 4. *The reduction of* $S(A_S^{(k)})$, $S(A^{(k)})$, $S(B^{(k)})$, $S(A^{(k)}, B^{(k)})$ *under the row-cyclic strategy*

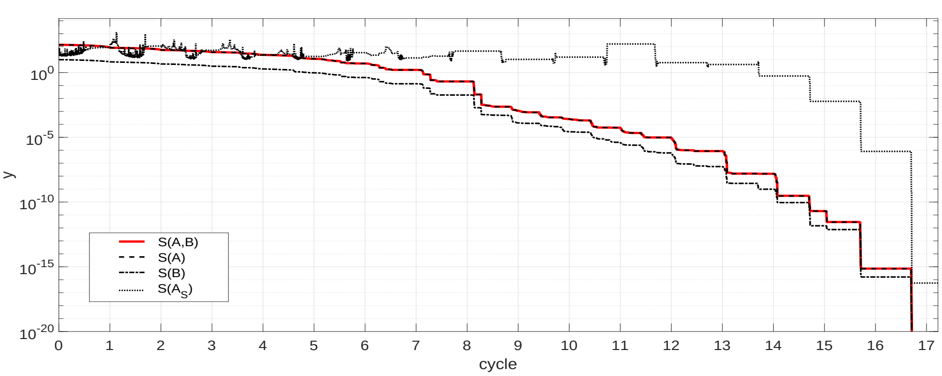

FIG. 5. *The reduction of* $S(A_S^{(k)})$, $S(A^{(k)})$, $S(B^{(k)})$, $S(A^{(k)}, B^{(k)})$ *under the deRijk strategy*

4. The Global Convergence. Here we prove the global convergence of the complex HZ method under the large class of *generalized serial strategies*. This class of cyclic strategies was introduced in [13] and it includes serial, wavefront, weak-wavefront, inverse of weak-wavefront strategies and those cyclic strategies that are permutational equivalent to all of them. Hence they also include the modulus strategy [15, 19] and some other cyclic strategies that are used for parallel processing.

The convergence proof is similar to that of the complex CJ method [14, 10], although it is more complicated. It is based on the following general theorem from [14].

THEOREM 4.1. *Let $H \neq 0$ be a Hermitian matrix and let $(H^{(k)},\ k \geq 0)$ be the sequence generated by applying a Jacobi-type process to H,*

$$H^{(k+1)} = F_k^* H^{(k)} F_k, \quad H^{(0)} = H, \quad k \geq 0.$$

Here each F_k is an elementary plane matrix which acts in the $(i(k), j(k))$ plane, $1 \leq i(k) < j(k) \leq n$. Suppose the following assumptions are satisfied:

(A1) *the pivot strategy is generalized serial*

(A2) *there is a sequence $(U_k,\ k \geq 0)$ of unitary elementary plane matrices such that $\lim_{k\to\infty} (F_k - U_k) = 0$*

(A3) *the diagonal elements of F_k satisfy the condition $\liminf_{k\to\infty} |f_{i(k)i(k)}^{(k)}| > 0$*

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

Then the following two conditions are equivalent

(i) $\lim_{k\to\infty} |h_{i(k)j(k)}^{(k+1)}| = 0$

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

We shall apply Theorem 4.1 to the sequences $(A^{(k)},\ k \geq 0)$ and $(B^{(k)},\ k \geq 0)$ obtained by the HZ method. To this end we shall prove some preparatory results. First, we want to prove that all matrices $A^{(k)}$, $B^{(k)}$, generated by the method are bounded. That accounts for the assumption **A4** of Theorem 4.1. Then we want to prove that $b_{i(k)j(k)}^{(k)}$ tends to zero as k increases. Once we prove it, the other assumptions of Theorem 4.1 will be easy to show.

In the following lemma we use the spectral radius of the matrix pair (A, B),

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

where $\sigma(A, B)$ denotes the spectrum of (A, B).

LEMMA 4.2. *Let A and B be Hermitian matrices of order n such that B is positive definite. Let the sequences of matrices $(A^{(k)},\ k \geq 0)$, $(B^{(k)},\ k \geq 0)$ be generated by applying the complex HZ method to the pair (A, B) under an arbitrary pivot strategy. Then the* assertions *(i)–(iv) hold.*

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

$$\|B^{(k)}\|_2 < n, \qquad \|A^{(k)}\|_2 \leq \mu \|B^{(k)}\|_2 < n\mu \tag{4.1}$$

(ii) *For the pivot element $b_{i(k)j(k)}^{(k)}$ of $B^{(k)}$ we have $\lim_{k\to\infty} b_{i(k)j(k)}^{(k)} = 0$*

(iii) *For the transformation matrices Z_k, we have*

$$\lim_{k\to\infty}(Z_k - U_k) \to 0,$$

where U_k are unitary plane matrices

(iv) *For the diagonal elements of $\hat{U}_k$, we have*

$$|u^{(k)}_{i(k)i(k)}| = |u^{(k)}_{j(k)j(k)}| \geq \frac{\sqrt{2}}{2}, \qquad k \geq 0.$$

Proof. (i) The proof of the relation (4.1) is identical to the proof of [9, Lemma 4.1]. One only has to replace the adjective "symmetric" by "Hermitian".

(ii) The proof follows the lines in the proof of [9, Proposition 4.1]. Let $B^{(k)} = (b^{(k)}_{rs})$ and

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

By the Hadamard's inequality we have

$$0 < H(B^{(k)}) \leq 1, \quad k \geq 0. \tag{4.2}$$

By the relations (2.1) and (3.2) we have

$$H(B^{(k+1)}) = |\det(Z_k)|^2 \det(B^{(k)}) = \frac{1}{1 - |b^{(k)}_{i(k)j(k)}|^2} H(B^{(k)}), \quad k \geq 0.$$

Hence

$$H(B^{(k)}) = \left(1 - |b^{(k)}_{i(k)j(k)}|^2\right) H(B^{(k+1)}), \quad k \geq 0. \tag{4.3}$$

From the relations (4.3) and (4.2) we see that $H(B^{(k)})$ is a nondecreasing sequence of positive real numbers, bounded above by 1. Hence it is convergent with limit ζ, $0 < \zeta \leq 1$. By taking the limit on the both sides of the equation (4.3), after cancelation with ζ, we obtain

$$1 = \lim_{k\to\infty}\left(1 - |b^{(k)}_{i(k)j(k)}|^2\right) = 1 - \lim_{k\to\infty} |b^{(k)}_{i(k)j(k)}|^2,$$

which proves (ii).

(iii) Recall that each Z_k is product $Z_k = R^{(k)}_1 D_k R^{(k)}_2 \Phi^{(k)}$ where $R^{(k)}_1$ and $R^{(k)}_2$ are complex rotations from the relation (3.2) related to step k. Let $U_k = R^{(k)}_1 R^{(k)}_2 \Phi^{(k)}$. Since $\Phi^{(k)}$ is unitary, we have

$$\begin{aligned}
\|Z_k - U_k\|_2 &= \|R^{(k)}_1 (D_k - I_n) R^{(k)}_2 \Phi^{(k)}\|_2 = \|D_k - I_n\|_2 \\
&= \|\mathrm{diag}\left(1/\sqrt{1 - |b^{(k)}_{ij}|} - 1\,,\, 1/\sqrt{1 + |b^{(k)}_{ij}|} - 1\right)\|_2 \\
&= |b^{(k)}_{ij}|/(1 - |b^{(k)}_{ij}| + \sqrt{1 - |b^{(k)}_{ij}|}).
\end{aligned}$$

Hence $\|Z_k - U_k\|_2 \to 0$ as $k \to \infty$. Here we have used the *assertion* (ii).

(*iv*) Note that the diagonal elements of $|\hat{U}_k|$ are equal since $\hat{U}_k$ is unitary of order 2. Nevertheless, we shall find expressions for both $|u^{(k)}_{i(k)i(k)}|$ and $|u^{(k)}_{j(k)j(k)}|$. Since $\hat{\Phi}^{(k)}$ is diagonal and unitary we have

$$|\hat{U}_k| = |\hat{R}^{(k)}_1 \hat{R}^{(k)}_2 \hat{\Phi}^{(k)}| = |\hat{R}^{(k)}_1 \hat{R}^{(k)}_2|, \qquad k \geq 0.$$

From the relations (3.3), (3.4) and (3.8) one easily obtains expressions for the diagonal elements of $|\hat{R}^{(k)}_1 \hat{R}^{(k)}_2|$. They are also the diagonal elements of $|\hat{U}_k|$. We have

$$4|u^{(k)}_{i(k)i(k)}|^2 = |c_k - s_k + e^{\imath\gamma_k}(c_k + s_k)|^2 = 2 + 2\cos(2\theta_k)\cos\gamma_k,$$
$$4|u^{(k)}_{j(k)j(k)}|^2 = |c_k - s_k + e^{-\imath\gamma_k}(c_k + s_k)|^2 = 2 + 2\cos(2\theta_k)\cos\gamma_k,$$

where $c_k = \cos\theta_k$, $s_k = \sin\theta_k$, $\gamma_k = \beta^{(k)}_{i(k)j(k)} - \alpha^{(k)}_{i(k)j(k)}$. This proves the *assertion* (*iv*). Indeed, our choice of σ_{ij} in (3.11) ensures $-\pi/2 \leq \gamma_k \leq \pi/2$ and we also have $-\pi/4 \leq \theta_k \leq \pi/4$. ∎

In the convergence proof we shall need to estimate how close are the diagonal elements of $A^{(k)}$ to the corresponding eigenvalues of the pair $(A^{(k)}, B^{(k)})$. To this end let the eigenvalues of the initial pair (A, B) be nonincreasingly ordered:

$$\lambda_1 = \cdots = \lambda_{s_1} > \lambda_{s_1+1} = \cdots = \lambda_{s_2} > \cdots > \lambda_{s_{p-1}+1} = \cdots = \lambda_{s_p}. \tag{4.4}$$

The case $p = 1$ implies $A = \lambda_1 B$. Then every nonzero vector is an eigenvector belonging to the only eigenvalue λ_1. So, let $p > 1$.

If we set $s_0 = 0$ we conclude from the relation (4.4) that $n_r = s_r - s_{r-1}$ is the multiplicity of λ_{s_r}. Let $\lambda_{s_0} = \lambda_0 = \infty$, $\lambda_{s_{p+1}} = -\infty$ and

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

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

$$\delta = \min_{1\leq t\leq p} \delta_t, \qquad \delta_0 = \frac{\delta}{1+\mu^2}, \tag{4.5}$$

where μ is the spectral radius of (A, B). Obviously, 3δ is the minimum absolute gap and for δ_0 we have

$$\delta_0 = \frac{\delta}{1+\mu^2} \leq \frac{\delta}{2\mu} \leq \frac{1}{3}. \tag{4.6}$$

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 $3\delta = 2\mu$. Note also that

$$|a_{rr}| = \frac{|e_r^T A e_r|}{e_r^T B e_r} \leq \max_{\|x\|_2=1} \frac{|x^* A x|}{x^* B x} = \mu, \quad 1 \leq r \leq n. \tag{4.7}$$

In the convergence theorem we shall need the following result from [8, Corollary 3.3] or from [9, Lemma 4.3] .

LEMMA 4.3. *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 (4.4) and let δ, δ_0 be as in the relation (4.5). If*

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

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

$$2\sum_{l=1}^{n} |\tilde{a}_{ll} - \lambda_l|^2 \leq \frac{S^4(A,B)}{\delta_0^2}. \tag{4.8}$$

In Lemma 4.3, the condition $\sqrt{1+\mu^2}S(A,B) < \delta$ can be replaced by the simpler and stricter one, $S(A,B) < \delta_0$. Similar estimates that include relative distances between $\tilde{a}_{ll}$ and λ_l can be found in [12].

Theorem 4.4. *The complex HZ method is globally convergent under the class of generalized serial pivot strategies.*

Proof. Let us apply Theorem 4.1 to $(B^{(k)},\ k \geq 0)$ and $(A^{(k)},\ k \geq 0)$. In both cases the assumptions ***(A1)***, ***(A2)***, ***(A4)*** and the condition ***(i)*** hold. Indeed, ***(A1)*** is just selection of the pivot strategy while ***(A2)*** and ***(A4)*** are the *assertions* (*iii*) and (*i*) of Lemma 4.2, respectively. The condition ***(i)*** holds because the HZ method diagonalizes the pivot submatrices, that is $a_{i(k)j(k)}^{(k+1)} = 0$ and $b_{i(k)j(k)}^{(k+1)} = 0$ holds for all $k \geq 0$.

It remains to prove the assumption ***(A3)***, that is $\liminf_{k\to\infty} |z_{i(k)i(k)}^{(k)}| > 0$. By the *assertion* (iv) of Lemma 4.2, we have

$$|z_{i(k)i(k)}^{(k)}| \geq |u_{i(k)i(k)}^{(k)}| - |z_{i(k)i(k)}^{(k)} - u_{i(k)i(k)}^{(k)}| \geq \frac{\sqrt{2}}{2} - \|Z_k - U_k\|_2$$

and by the *assertion* (iii) of the same lemma, $\|Z_k - U_k\|_2 \to 0$ as $k \to \infty$. Hence

$$\liminf_{k\to\infty} |z_{i(k)i(k)}^{(k)}| \geq \sqrt{2}/2.$$

From Theorem 4.1 we conclude that $S(A^{(k)}) \to 0$ and $S(B^{(k)}) \to 0$ as $k \to \infty$. Since each $B^{(k)}$ has unit diagonal, it is shown that $B^{(k)} \to I_n$ as $k \to \infty$.

If $\sigma(A,B)$ is singleton, i.e. if $p = 1$ holds in the relation (4.4), the proof is completed. Namely, if $A = \lambda_1 B$, we shall have $A^{(k)} = \lambda_1 B^{(k)}$, $k \geq 0$. In that case the HZ algorithm chooses $\theta_k = 0$, $k \geq 0$, and $\hat{Z}_k$ is computed by the relation (3.22). Since $B^{(k)} \to I_n$, we shall have $A^{(k)} \to \lambda_1 I_n$ as $k \to \infty$.

It remains to prove that the diagonal elements of $A^{(k)}$ converge in the case $p > 1$. This comes down to showing that for large enough k the diagonal elements of $A^{(k)}$ cannot change their eigenvalue affiliations.

Suppose k_0 is so large that we have

$$S(A^{(k)}, B^{(k)}) < \delta_0^2, \qquad k \geq k_0. \tag{4.9}$$

Let us consider step k of the process when $k \geq k_0$. Set $A = (a_{rt}) = A^{(k)}$, $A' = (a'_{rt}) = A^{(k+1)}$, $B = (b_{rt}) = B^{(k)}$, $B' = (b'_{rt}) = B^{(k+1)}$. From the relation (4.6) we see that the assumption (4.9) implies

$$S(A,B) < \delta_0^2 \leq \frac{1}{3}\delta_0 \tag{4.10}$$

and therefore we have

$$|b_{ij}| < \frac{\sqrt{2}}{6}\delta_0 \leq \frac{\sqrt{2}}{18}, \qquad \tau_{ij} = \sqrt{1 - |b_{ij}|^2} > \frac{\sqrt{322}}{18}. \tag{4.11}$$

Using (4.10), the upper bound appearing in (4.8) can be further bounded as follows:

$$\frac{S^4(A,B)}{\delta_0^2} < \frac{\delta_0^2}{81}.$$

Hence, from Lemma 4.3 we can conclude that all diagonal elements of A are contained in the union of disks

$$\mathcal{D}_t = \left\{ x \,:\, |x-\lambda_t| \leq \frac{\sqrt{2}}{18}\delta_0 \right\}, \qquad 1 \leq t \leq n.$$

Since $(\sqrt{2}/18)\delta_0 < 0.0786\delta_0 < 0.0786\delta$, these disks are disjoint. Hence, Lemma 4.3 implies that each disk $\mathcal{D}_t$ contains exactly n_t diagonal elements of A.

The same conclusion holds for the diagonal elements of A'. The proof will be completed if we show that no diagonal element of A can jump from one disk to another.

Suppose a_{ii} is affiliated with λ_r and a_{jj} with λ_t. Then by Lemma 4.3 and the relation (4.10) we have

$$|a_{ii}-\lambda_r|^2 + |a_{jj}-\lambda_t|^2 \leq \frac{S^4(A,B)}{2\delta_0^2} \leq \frac{1}{18}S^2(A,B) < \frac{1}{162}\delta_0^2, \tag{4.12}$$

$$\max\{|a_{ii}-\lambda_r|\,,\,|a_{jj}-\lambda_t|\} \leq \frac{\sqrt{2}}{6}S(A,B) < \frac{\sqrt{2}}{18}\delta_0 < \frac{\sqrt{2}}{18}\delta. \tag{4.13}$$

We consider two cases: **(a)** $\lambda_r \neq \lambda_t$ and **(b)** $\lambda_r = \lambda_t$.

(a) Using the relations (4.5), (4.12) and the Cauchy-Schwarz inequality, we have

$$|a_{ii}-a_{jj}| \geq |\lambda_r-\lambda_t| - |a_{ii}-\lambda_r| - |a_{jj}-\lambda_t| > 3\delta - \sqrt{2}\frac{1}{\sqrt{2}\cdot 9}\delta_0 = \frac{26}{9}\delta. \tag{4.14}$$

Let us bound $|a'_{ii}-a_{ii}|$. To this end we denote $\gamma_{ij} = \alpha_{ij} - \beta_{ij}$. From the relations (3.20) and (4.7) we obtain

$$\begin{aligned} \tau_{ij}^2|a'_{ii}-a_{ii}| &= |(|b_{ij}|^2 - \sin^2\phi)a_{ii} + \sin^2\psi a_{jj} + 2\cos\phi\sin\psi u_{ij}| \\ &\leq \mu(\sin^2\phi + \sin^2\psi) + 2\cos\phi\sin\psi\,|a_{ij}| + \mu\,|b_{ij}|^2. \end{aligned} \tag{4.15}$$

From the relations (3.10) and (3.12) we have $\cos\gamma_{ij} \geq 0$ and $\cos(2\theta) \geq 0$, respectively. Hence, from the relation (3.17), we have

$$\begin{aligned} \sin^2\phi + \sin^2\psi &= 1 - \tau_{ij}\cos(2\theta)\cos\gamma_{ij} \leq 1 - (1-|b_{ij}|^2)(1-\sin^2(2\theta))(1-\sin^2\gamma_{ij}) \\ &= \sin^2(2\theta) + \sin^2\gamma_{ij} - \sin^2(2\theta)\sin^2\gamma_{ij} + |b_{ij}|^2\cos^2(2\theta)\cos^2\gamma_{ij} \\ &\leq \tan^2(2\theta) + \tan^2\gamma_{ij} + |b_{ij}|^2, \end{aligned}$$

$$\begin{aligned} 4\cos^2\phi\sin^2\psi &= (1-|b_{ij}|\sin(2\theta))^2 - (1-|b_{ij}|^2)(1-\sin^2(2\theta))(1-\sin^2\gamma_{ij}) \\ &\leq |b_{ij}|^2 + \sin^2(2\theta) + \sin^2\gamma_{ij} + 2|b_{ij}||\sin(2\theta)| \\ &\leq 2\,(\tan^2(2\theta) + \tan^2\gamma_{ij} + |b_{ij}|^2). \end{aligned}$$

We have thus obtained

$$\sin^2\phi + \sin^2\psi + |b_{ij}|^2 \leq \tan^2(2\theta) + \tan^2\gamma_{ij} + 2\,|b_{ij}|^2 \tag{4.16}$$

$$2\cos\phi\sin\psi \leq \sqrt{2}\sqrt{\tan^2(2\theta) + \tan^2\gamma_{ij} + |b_{ij}|^2}. \tag{4.17}$$

Using relations (3.12), (4.11), (4.14) and (4.10), one obtains

$$\tan^2(2\theta) \le \frac{(2|a_{ij}|+2\mu|b_{ij}|)^2}{\tau_{ij}^2(a_{ii}-a_{jj})^2} \le \frac{2(1+\mu^2)S^2(A,B)}{(322/18^2)\cdot(26/9)^2\,\delta^2} \le \frac{2\cdot 18^2\cdot 9^2}{322\cdot 26^2}\frac{S(A,B)}{1+\mu^2} \le 0.2412\frac{S(A,B)}{1+\mu^2}. \tag{4.18}$$

Using (3.9), (4.14), (4.6) and (4.10), we have

$$\tan^2\gamma_{ij} + 2|b_{ij}|^2 \le \frac{4|a_{ij}|^2}{(a_{ii}-a_{jj})^2} + S^2(B) \le \frac{2S^2(A)}{(26/9)^2\,\delta^2} + (\frac{2\mu}{3\delta})^2 S^2(B) \le \frac{4}{9}\frac{(1+\mu^2)S^2(A,B)}{\delta^2} \le \frac{4}{9}\frac{S(A,B)}{1+\mu^2}. \tag{4.19}$$

Combining relations (4.16), (4.18), (4.19) and (4.10), we have

$$\mu(\sin^2\phi+\sin^2\psi+|b_{ij}|^2) \le \mu(0.2412+\frac{4}{9})\frac{S(A,B)}{1+\mu^2} \le 0.686\frac{\mu}{1+\mu^2}S(A,B) \le 0.343\,S(A,B) < 0.1144\,\delta_0. \tag{4.20}$$

In a similar way, from the relations (4.17) and (4.20), we obtain

$$2\cos\phi\sin\psi\,|a_{ij}| \le \sqrt{2}\sqrt{0.343\,S(A,B)} < 0.8283\,\delta_0. \tag{4.21}$$

Combining relation (4.15) with (4.20), (4.21) (4.11), we have

$$|a'_{ii}-a_{ii}| = \frac{1}{1-|b_{ij}|^2}(0.1144+0.8283)\,\delta_0 < 0.9486\,\delta_0 < 0.9486\,\delta. \tag{4.22}$$

Finally, from the relations (4.22) and (4.13) we obtain

$$|a'_{ii}-\lambda_r| \le |a'_{ii}-a_{ii}| + |a_{ii}-\lambda_r| < (0.9486+\frac{\sqrt{2}}{18})\,\delta < 1.03\,\delta\,.$$

We conclude that a_{ii} cannot move from $\mathcal{D}_r$ to any other disk. So, a'_{ii} must remain in $\mathcal{D}_r$.

Quite similar estimates can be made for $|a'_{jj}-\lambda_t|$. But that is not needed. We know that except for a_{ii} and a_{jj} no other diagonal element of A is affected by the transformation. Since a'_{ii} remained in $\mathcal{D}_r$, jump of a_{jj} to any other disk but $\mathcal{D}_t$ would violate the rule on the number of the diagonal elements in the disks.

(b) In this case a_{ii} and a_{jj} both lie in $\mathcal{D}_r$. After the transformation they both have to remain in $\mathcal{D}_r$, because otherwise $\mathcal{D}_r$ and some other disk(s) would violate the rule on the number of the diagonal elements in the disks. Thus, we must have $a'_{ii}, a'_{jj} \in \mathcal{D}_r$, which completes the proof of the theorem. ∎

5. Conclusions and Future Work. The complex HZ method has proved to be a reliable diagonalization method for PGEP. In this paper we have derived its algorithm and have proved the global convergence under the class of generalized serial strategies. The numerical tests indicate that it might be high relative accurate on the set of well-behaved pairs of positive definite matrices.

Future work can be concentrated on proving the asymptotic quadratic convergence of the method and on proving the high relative accuracy of the method for certain classes of matrix pairs. The first problem has already been solved [6, 5] for the case of simple and double eigenvalues, but in the case of multiple eigenvalues the method will need some kind of modification.

Concerning the numerical code, there are many details that can be improved (cf. [20]). In particular, how to reduce the total number of cycles (compare Fig. 4 and Fig. 5), what are the best formulas for updating the diagonal elements of A, what are the most efficient pivot strategies, what is the best stopping criterion, how to implement one-sided version of the method, etc.

Acknowledgments. The author is thankful to the editor J. Barlow and to the anonymous referees for excellent comments which improved the paper.

REFERENCES

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

[2] Z. Drmač, *A tangent algorithm for computing the generalized singular value decomposition*, SIAM J. Num. Anal., 35 (1998), pp. 1804–1832.

[3] S. Falk and P. Langemeyer, *Das Jacobische Rotationsverfahren für reellsymmetrische Matrizenpaare I, II*, ZAMM, 0 (1960), pp. 30–43.

[4] G. Gose, *Das Jacobi–Verfahren für* $Ax = \lambda Bx$, ZAMM, 59 (1979), pp. 93–101.

[5] V. Hari, *On the quadratic convergence of the complex HZ method for the positive definite generalized eigenvalue problem.* Submitted for publ.

[6] V. Hari, *On Cyclic Jacobi Methods for the Positive Definite Generalized Eigenvalue Problem*, PhD thesis, Fernuniversität Hagen, 1984.

[7] V. Hari, *On Jacobi-type congruence transformations for a positive definite matrix*, in Numerical methods and approximation theory II, D. Herceg, ed., Novi Sad, 9 1985, Institute of Mathematics, University of Novi Sad, pp. 27–34.

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

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

[10] V. Hari, *Complex Cholesky–Jacobi algorithm for PGEP*, in International conference of numerical Aaalysis and applied mathematics (ICNAAM 2018), vol. 2116, AIP conference proceedings, 2019, https://doi.org/10.1063/1.5114478.

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

[12] V. Hari and Z. Drmač, *On scaled almost diagonal Hermitian matrix pairs*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 1000–1012.

[13] V. Hari and E. B. Kovač, *Convergence of the cyclic and quasi-cyclic block Jacobi methods*, Electronic transactions on numerical analysis (ETNA), 46 (2017), pp. 107–147.

[14] V. Hari and E. B. Kovač, *On the convergence of complex Jacobi methods*, Linear and Multilin. Algebra, (2019), https://doi.org/10.1080/03081087.2019.1604622.

[15] F. T. Luk and H. Park, *On parallel Jacobi orderings*, SIAM J. Sci. Stat. Comp., 10 (1987), pp. 18–26.

[16] J. Matejaš, *Accuracy of one step of the Falk-Langemeyer method*, Numer. Algor., 68 (2015), pp. 645–670.

[17] J. Matejaš and V. Hari, *Accuracy of the HZ method for PGEP.* Proposed for publ.

[18] V. Novaković, S. Singer, and S. Singer, *Hari-Zimmermann variant of the Falk-Langemeyer algorithm for the generalized SVD*, Parallel Computing, 49 (2015), pp. 136–152.

[19] G. Shroff and R. Schreiber, *On the convergence of the cyclic Jacobi method for parallel block orderings*, SIAM J. Matrix Anal. Appl., 10 (1989), pp. 326–346.

[20] S. Singer, E. Di Napoli, V. Novaković, and G. Čaklović, *The FLAPW methods with eigendecomposition based on the Hari–Zimmermann generalized hyperbolic SVD.* arXiv:1907.08560 [math.NA]. Submitted for publ.

[21] I. Slapničar and V. Hari, *On the quadratic convergence of the Falk–Langemeyer method*, SIAM J. Matrix Anal. Appl., 12 (1991), pp. 84–114.

[22] A. van der Sluis, *Condition numbers and equilibration of matrices*, Numer. Math., 14 (1969), pp. 14–23.