On Kogbetliantz's SVD Algorithm in the Presence of Clusters

J.-P. Charlier and P. Van Dooren
Philips Research Laboratory
Brussels, Belgium

Submitted by G. W. Stewart

ABSTRACT

We consider matrices with off-diagonal blocks of small norm and derive tight bounds for the approximation of their singular values by those of their diagonal blocks. These results are used to show that triangular matrices with clusters of singular values must possess a principal submatrix of "nearly" diagonal form. From the latter we then derive results pertaining to the quadratic convergence of Kogbetliantz's algorithm for computing the SVD, in the presence of clusters.

1. INTRODUCTION

Kogbetliantz's algorithm for the singular-value decomposition (SVD) of an arbitrary matrix [5, 6] can be viewed as a natural extension of Jacobi's method for the eigenvalue decomposition of a symmetric, Hermitian, or normal matrix [3, 11]. While the latter has extensively been studied for its convergence properties [3, 12, 14, 15, 16], convergence results are much scarcer for Kogbetliantz's algorithm [3, 9].

As shown in this paper, this is partly due to the fact that not all properties can be extended from the eigenvalue decomposition of a Hermitian matrix (or a normal matrix) to the singular-value decomposition of a general matrix. In Section 2 we show that there are more natural similarities between these two decompositions if one restricts oneself to the SVD of *triangular matrices*. Indeed, the norm of the off-diagonal part (or the "off-norm") of a Hermitian matrix H and of an arbitrary *triangular* matrix A are bounded—up to an appropriate constant—by the "span" of the eigenvalues $[\lambda_{\max}(H) - \lambda_{\min}(H)]$ and of the singular values $[\sigma_{\max}(A) - \sigma_{\min}(A)]$, respectively. This plays a crucial role in convergence properties of Kogbetliantz's algorithm in the presence of clusters.

52 Vanderbilt Ave., New York, NY 10017 0024-3795/87/$3.50

In Section 3 we derive bounds for the approximation of singular values of "nearly" block diagonal matrices by those of the diagonal blocks. These results naturally extend earlier results [2, 13] about the eigenvalues of Hermitian matrices (along the way we derive slightly sharper bounds for the latter case as well). In contrast with Section 2, here the obtained extensions also hold without restricting oneself to the SVD of triangular matrices. The results show that if the norm of the off-diagonal blocks A_{12} and A_{21} of a square matrix A is ε-small and the singular values of A_{11} and A_{22} are separate at least by δ ($\delta > 2\varepsilon$), then the singular values of the diagonal submatrices are ε^2/δ-close to those of A. Here by "η-small" we mean "having norm bounded by η," while "η-close" means "having a distance bounded by η".

In Section 4 we combine the results of the two previous sections to derive bounds for the off-norm of any $k \times k$ principal submatrix approximating a cluster of k singular values of a *triangular* matrix A. These results are similar to the bounds obtained for Hermitian matrices in [17], but again hold for triangular matrices only. In Section 5, they are shown to be crucial for proving the quadratic convergence of Kogbetliantz's algorithm in the presence of repeated singular values. That the obtained theoretical bounds also reflect the true behavior of the method is then finally illustrated by a number of examples.

2. OFF-NORMS IN TERMS OF EIGENVALUES OR SINGULAR VALUES

The eigenvalue decomposition of a Hermitian matrix

$$H = U\Lambda U^*, \tag{1}$$

and the singular-value decomposition of an arbitrary matrix,

$$A = U\Sigma V^*, \tag{2}$$

are closely related, since the Hermitian matrix $H \doteq AA^*$ has a decomposition (1) related with (2) via $\Lambda \doteq \Sigma^2$. This relation leads to a number of similar developments for these two decompositions, such as Jacobi's algorithm and Kogbetliantz's algorithm, or perturbation results of eigenvalues of H and singular values of A [13]. In this section we present a result for which this similarity holds *for triangular matrices only*. It relates the "off-norm" of a matrix with the "span" of its eigenvalues or singular values.

LEMMA 1. *Let H be a Hermitian matrix and have extremal eigenvalues $\lambda_{\min}$ and $\lambda_{\max}$. Then putting $H = D + H_{\mathrm{off}}$, where D is diagonal and H_{off} contains the rest of the matrix H, we have*

$$\|H_{\mathrm{off}}\|_F \leqslant \frac{\sqrt{n}}{2}(\lambda_{\max} - \lambda_{\min}) \tag{3}$$

$$\|H_{\mathrm{off}}\|_2 \leqslant (\lambda_{\max} - \lambda_{\min}). \tag{4}$$

Proof. The shifted matrix

$$\hat{H} = H - \frac{\lambda_{\max} + \lambda_{\min}}{2} I \tag{5}$$

has 2-norm equal to $(\lambda_{\max} - \lambda_{\min})/2$, and likewise

$$\hat{D} = D - \frac{\lambda_{\max} + \lambda_{\min}}{2} I \tag{6}$$

has 2-norm less than $(\lambda_{\max} - \lambda_{\min})/2$, being the diagonal of $\hat{H}$. Since H_{off} is also the off-diagonal part of $\hat{H}$, (3) follows. Moreover, because $H_{\mathrm{off}} = \hat{H} - \hat{D}$, (4) follows from

$$\|H_{\mathrm{off}}\|_2 \leqslant \|\hat{H}\|_2 + \|\hat{D}\|_2. \qquad \blacksquare \tag{7}$$

COROLLARY 1. *The above lemma also holds for a normal matrix N on replacing $(\lambda_{\max} + \lambda_{\min})/2$ by $\lambda_{\mathrm{center}}$, the center of the smallest circle enclosing all the eigenvalues of N, and $(\lambda_{\max} - \lambda_{\min})/2$ by $\lambda_{\mathrm{radius}}$, the radius of that circle.*

Proof. Trivial. Since $N - \lambda_{\mathrm{center}} I$ is also normal, it has 2-norm equal to $\lambda_{\mathrm{radius}}$ and then the above reasoning can be followed. ■

REMARK 1. Lemma 1 provides rather tight bounds, as shown by the following example. Consider the matrix

$$H = \begin{pmatrix} 0 & I_{n/2} \\ I_{n/2} & 0 \end{pmatrix},$$

where n is even. Since its eigenvalues are ± 1, we have

$$\lambda_{\max} - \lambda_{\min} = 2$$

and

$$\|H_{\mathrm{off}}\|_F = \sqrt{n} = \frac{\sqrt{n}}{2}(\lambda_{\max} - \lambda_{\min})$$

which equals the upper bound (3). For the 2-norm we have

$$\|H_{\mathrm{off}}\|_2 = \|H\|_2 = 1 = \tfrac{1}{2}(\lambda_{\max} - \lambda_{\min}),$$

which is only a factor 2 away from the upper bound (4).

A similar result for the singular values of an arbitrary triangular matrix is now derived:

LEMMA 2. *Let A be a square upper triangular (complex) matrix with extremal singular values $\sigma_{\min}$ and $\sigma_{\max}$. Then, putting $A = D + A_{\mathrm{off}} = D + A_{\mathrm{up}}$, we have*

$$\|A_{\mathrm{off}}\|_2 \leqslant \|A_{\mathrm{off}}\|_F \leqslant \sqrt{n-1}\,(\sigma_{\max} - \sigma_{\min}). \tag{8}$$

Proof. Denote $\sigma_+ = (\sigma_{\max} + \sigma_{\min})/2$ and $\sigma_- = (\sigma_{\max} - \sigma_{\min})/2$. Using the singular-value decomposition of A, we define

$$A = U\Sigma V^* \quad \text{and} \quad \hat{A} = A - \sigma_+ UV^* = A - \sigma_+ Q. \tag{9}$$

It then easily follows [4] that $\hat{A}$ has 2-norm equal to σ_-. The upper part Q_{up} of Q satisfies

$$\|Q_{\mathrm{up}}\|_F^2 = \sum_{i=1}^{n-1} \left\|(Q_{\mathrm{up}})_i\right\|_2^2, \tag{10}$$

where $(Q_{\mathrm{up}})_i$ is that part of the ith row of Q_{up} to the right of the diagonal. The row $(Q_{\mathrm{up}})_i$ is also the bottom row of a submatrix Q_{12} of Q in the partition

$$Q = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix}, \tag{11}$$

where Q_{11} is $i\times i$. From $\|Q_{12}Q_{12}^*\|_2=\|I-Q_{11}Q_{11}^*\|_2=\|I-Q_{11}^*Q_{11}\|_2=\|Q_{21}^*Q_{21}\|_2$, it follows that

$$\|Q_{12}\|_2=\|Q_{21}\|_2, \tag{12}$$

and, since $\sigma_+Q_{21}=-\hat{A}_{21}$ (A being upper triangular),

$$\left\|(Q_{\text{up}})_i\right\|_2\leqslant\|Q_{12}\|_2=\frac{1}{\sigma_+}\left\|\hat{A}_{21}\right\|_2\leqslant\frac{\sigma_-}{\sigma_+}. \tag{13}$$

Using this in (10), we have

$$\sigma_+^2\|Q_{\text{up}}\|_F^2\leqslant(n-1)\sigma_-^2. \tag{14}$$

By a similar argument to (10), we have for $\hat{A}_{\text{up}}$

$$\left\|\hat{A}_{\text{up}}\right\|_F^2\leqslant(n-1)\sigma_-^2. \tag{15}$$

Finally, from $A_{\text{off}}=A_{\text{up}}=\hat{A}_{\text{up}}+\sigma_+Q_{\text{up}}$ and (15), the result (8) follows by the triangle inequality. ∎

REMARK 2. The above lemma of course does not hold for full matrices, since one can always construct a matrix A with $\|A_{\text{off}}\|=\|A\|$ as well for the 2-norm as for the Frobenius norm. Bounds that always hold for full matrices are $\|A_{\text{off}}\|_F\leqslant\|A\|_F$ and $\|A_{\text{off}}\|_2\leqslant2\|A\|_2$ (the latter follows from $A_{\text{off}}=A-D$ and $\|D\|_2\leqslant\|A\|_2$).

These bounds for the off-norms of arbitrary and triangular matrices are now checked for their tightness on a small example whose singular values are very close to 1, i.e. a matrix A close to a unitary one.

EXAMPLE 1. We generated randomly a 15×15 matrix A very close to a unitary one, such that $\sigma_{\max}-\sigma_{\min}\approx1.25\times10^{-14}$. Then a QR decomposition of A was made and the triangular factor R, which of course has the *same* singular values as A, was considered. Below, we give the off-norms of these two matrices and their bounds:

Actual off-norms	Bounds
$\|A_{\text{off}}\|_2 = 1.215$	$2\sigma_{\max} \approx 2.$
$\|A_{\text{off}}\|_F = 3.765$	$\sqrt{\Sigma_i \sigma_i^2} \approx \sqrt{n} \approx 3.873$
$\|R_{\text{off}}\|_2 = 1.09\times 10^{-14}$	$\sqrt{n-1}\,(\sigma_{\max} - \sigma_{\min}) \approx 4.68\times 10^{-14}$
$\|R_{\text{off}}\|_F = 1.99\times 10^{-14}$	$\sqrt{n-1}\,(\sigma_{\max} - \sigma_{\min}) \approx 4.68\times 10^{-14}$

This example illustrates also what improvements can be obtained by considering triangular factors of matrices when computing its singular values using an iterative algorithm as e.g. that of Kogbetliantz (see Example 3 later on).

3. PERTURBATION THEOREMS FOR EIGENVALUES AND SINGULAR VALUES

Here we obtain perturbation bounds for the eigenvalues and singular values of "nearly" block diagonal (and hence square) matrices. These bounds look very much like those obtained by Stewart in [13] but are slightly sharper. They are derived using a slightly different approach, for which we need the following lemma.

LEMMA 3. *Let A, B, and C be square matrices of the same order, and*

$$AX - XB = C. \tag{16}$$

Assume that A and B are Hermitian, and $\sigma_{\min}(X) > 0$. If the eigenvalues of A and B are ordered in a similar manner, then

$$|\lambda_i(A) - \lambda_i(B)| \leqslant \frac{\|C\|_2}{\sigma_{\min}(X)}. \tag{17}$$

If, moreover, X is Hermitian, then one has the stronger result

$$|\lambda_i(A) - \lambda_i(B)| \leqslant \|A - B\|_2 \leqslant \frac{\|C\|_2}{\sigma_{\min}(X)}. \tag{18}$$

Proof. This is inspired by a result given by Parlett [10, pp. 230–231], and we follow here a related argument. Use the singular value decomposition of X: $X = U\Sigma V^*$, and define $\hat{A} = U^*AU$, $\hat{B} = V^*BV$, $\hat{C} = U^*CV$. Equation (16) is then rewritten as

$$\hat{A}\Sigma - \Sigma\hat{B} = \hat{C} = (\hat{A} - \hat{B})\Sigma + (\hat{B}\Sigma - \Sigma\hat{B}), \tag{19}$$

where the last term is skew-Hermitian. Therefore, for any eigenvector u of $\hat{A} - \hat{B}$, we have

$$u^*(\hat{A} - \hat{B})\Sigma u = \operatorname{Re}(u^*\hat{C}u). \tag{20}$$

Choosing u (normalized) to correspond to the dominant eigenvalue of $\hat{A} - \hat{B}$ (i.e. $|\lambda| = \|\hat{A} - \hat{B}\|_2$), we obtain

$$\|\hat{A} - \hat{B}\|_2 u^*\Sigma u = |\operatorname{Re}(u^*\hat{C}u)| \leqslant |u^*\hat{C}u| \leqslant \|\hat{C}\|_2 = \|C\|_2 \tag{21}$$

and, since $u^*\Sigma u \geqslant \sigma_{\min}(X)$,

$$\|\hat{A} - \hat{B}\|_2 \leqslant \frac{\|C\|_2}{\sigma \min(X)}. \tag{22}$$

Then (17) follows from (see [4, p. 269]):

$$|\lambda_i(A) - \lambda_i(B)| = |\lambda_i(\hat{A}) - \lambda_i(\hat{B})| \leqslant \|\hat{A} - \hat{B}\|_2. \tag{23}$$

When X is Hermitian, one has $U = V$, and hence $\|A - B\|_2 = \|\hat{A} - \hat{B}\|_2$, from which (18) follows. ■

We now turn to the perturbation result on eigenvalues of nearly block diagonal matrices, where we make use of the parameter κ as defined by Stewart [13] and satisfying $0 < \kappa < 1$.

THEOREM 1. *Let H be a Hermitian matrix partitioned as*

$$H = \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^* & H_{22} \end{pmatrix}. \tag{24}$$

Define $\varepsilon = \|H_{12}\|_F$ *and* $\delta = \min|\operatorname{eig}(H_{11}) - \operatorname{eig}(H_{22})|$, *and assume* $2\varepsilon < \delta$.

Then there is a unitary matrix X of the form

$$X = \begin{pmatrix} I & -P^* \\ P & I \end{pmatrix} \begin{pmatrix} (I+P^*P)^{-1/2} & 0 \\ 0 & (I+PP^*)^{-1/2} \end{pmatrix} \tag{25}$$

with $\|P\|_F < (1+\kappa)\varepsilon/\delta < 2\varepsilon/\delta$, *and such that*

$$X^*HX = \begin{pmatrix} H'_{11} & 0 \\ 0 & H'_{22} \end{pmatrix} \doteq H'. \tag{26}$$

Moreover, if the eigenvalues of H_{11} *and* H'_{11} *are ordered in a similar manner, they satisfy*

$$|\lambda_i(H_{11}) - \lambda_i(H'_{11})| \leqslant \|H_{11} - H'_{11}\|_2 < \frac{(1+\kappa)\varepsilon^2}{\delta} < \frac{2\varepsilon^2}{\delta}. \tag{27}$$

Proof. The definition of δ and ε allows us to apply Theorem 4.1 of Stewart's paper ([13]; see also Theorem 4.7 therein), from which

$$\|P\|_F < \frac{(1+\kappa)\varepsilon}{\delta} < \frac{2\varepsilon}{\delta} \tag{28}$$

follows. On the other hand, $HX = XH'$ contains the equation

$$H_{11}N - NH'_{11} = -H_{12}PN, \tag{29}$$

where $N \doteq (I + P^*P)^{-1/2}$ is Hermitian. Lemma 3 then gives

$$|\lambda_i(H_{11}) - \lambda_i(H'_{11})| \leqslant \|H_{11} - H'_{11}\|_2 \leqslant \frac{\|H_{12}PN\|_2}{\sigma_{\min}(N)} \leqslant \|H_{12}\|_2 \frac{\|PN\|_2}{\sigma_{\min}(N)}. \tag{30}$$

Using the SVD of P, $P = U\Sigma V^*$, one easily derives $N = V(I+\Sigma^2)^{-1/2}V^*$ and $PN = U\Sigma(I+\Sigma^2)^{-1/2}V^*$. This also yields

$$\sigma_{\min}(N) = [1+\sigma^2_{\max}(P)]^{-1/2} \quad \text{and} \quad \|PN\|_2 = \sigma_{\max}(P)[1+\sigma^2_{\max}(P)]^{-1/2}. \tag{31}$$

This and $\|H_{12}\|_2 \leqslant \|H_{12}\|_F = \varepsilon$ in (30) lead to

$$|\lambda_i(H_{11}) - \lambda_i(H'_{11})| \leqslant \|H_{11} - H'_{11}\|_2 \leqslant \varepsilon\sigma_{\max}(P), \tag{32}$$

and the bound (27) follows from this and (28). ∎

Remark 3. The bound (27) is often only a slight improvement on the one obtained by Stewart:

$$|\lambda_i(H_{11}) - \lambda_i(H'_{11})| \leqslant (1+\kappa)\left(1 + \frac{\varepsilon^2}{\delta^2}\right)\frac{\varepsilon^2}{\delta} + 3\|H_{11}\|_2\frac{\varepsilon^4}{\delta^4}, \tag{33}$$

but the independence of the matrix H_{11} turns out to be crucial for our results in the later sections (especially when $\|H_{11}\|_2\varepsilon^2/\delta^3 \gg 1$).

This theorem is now extended to one about the perturbation of the singular values of a nearly block diagonal matrix. Here κ is again a parameter defined in [13] and satisfies $0 < \kappa < 1$.

Theorem 2. *Let A be a full square (complex) matrix partitioned as*

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}. \tag{34}$$

Define $\varepsilon = \|(A_{12}, A_{21}^*)\|_F$ and $\delta = \min|\text{sing}(A_{11}) - \text{sing}(A_{22})|$, and assume $2\varepsilon < \delta$. Then there are unitary matrices X and Y of the form

$$X = \begin{pmatrix} I & -P^* \\ P & I \end{pmatrix}\begin{pmatrix} (I+P^*P)^{-1/2} & 0 \\ 0 & (I+PP^*)^{-1/2} \end{pmatrix}, \tag{35}$$

$$Y = \begin{pmatrix} I & Q^* \\ -Q & I \end{pmatrix}\begin{pmatrix} (I+Q^*Q)^{-1/2} & 0 \\ 0 & (I+QQ^*)^{-1/2} \end{pmatrix} \tag{36}$$

with $\|(P,Q)\|_F < (1+\kappa)\varepsilon/\delta < 2\varepsilon/\delta$, and such that

$$Y^*AX = \begin{pmatrix} A'_{11} & 0 \\ 0 & A'_{22} \end{pmatrix}. \tag{37}$$

Moreover, if the singular values of A_{11} and A'_{11} are ordered in a similar manner, they satisfy

$$|\sigma_i(A_{11}) - \sigma_i(A'_{11})| \leqslant \|A_{11} - A'_{11}\|_2 < \frac{(1+\kappa)\varepsilon^2}{\delta} < \frac{2\varepsilon^2}{\delta}. \tag{38}$$

Proof. The proof is an extension of that of Theorem 1. Let Π be the permutation such that

$$\Pi^*\begin{pmatrix} 0 & A \\ A^* & 0 \end{pmatrix}\Pi = \begin{pmatrix} 0 & A_{11} & 0 & A_{12} \\ A_{11}^* & 0 & A_{21}^* & 0 \\ 0 & A_{21} & 0 & A_{22} \\ A_{12}^* & 0 & A_{22}^* & 0 \end{pmatrix} \doteq \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^* & H_{22} \end{pmatrix} = H. \tag{39}$$

Define a unitary matrix Z by

$$\begin{aligned} Z &\doteq \Pi^*\begin{pmatrix} Y & 0 \\ 0 & X \end{pmatrix}\Pi \\ &= \begin{pmatrix} I & -R^* \\ R & I \end{pmatrix}\begin{pmatrix} (I+R^*R)^{-1/2} & 0 \\ 0 & (I+RR^*)^{-1/2} \end{pmatrix}, \end{aligned} \tag{40}$$

with

$$R \doteq \begin{pmatrix} -Q & 0 \\ 0 & P \end{pmatrix}, \tag{41}$$

as well as

$$H' \doteq Z^*HZ = \begin{pmatrix} 0 & A'_{11} & 0 & 0 \\ A'^*_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & A'_{22} \\ 0 & 0 & A'^*_{22} & 0 \end{pmatrix} \doteq \begin{pmatrix} H'_{11} & 0 \\ 0 & H'_{22} \end{pmatrix}. \tag{42}$$

Since the eigenvalues of H_{11} and H_{22} are plus and minus the singular values of A_{11} and A_{22} respectively, δ is also given by

$$\delta = \min|\operatorname{eig}(H_{11}) - \operatorname{eig}(H_{22})|. \tag{43}$$

As $\|H_{12}\|_F = \|(A_{12}, A_{21}^*)\|_F = \varepsilon$, the application of Theorem 6.3 of [13] gives here

$$\|R\|_F = \|(P,Q)\|_F < \frac{(1+\kappa)\varepsilon}{\delta} < \frac{2\varepsilon}{\delta}. \tag{44}$$

From $HZ = ZH'$, the equation

$$H_{11}N - NH'_{11} = -H_{12}RN, \tag{45}$$

follows, where $N \doteq (I + R^*R)^{-1/2}$. The matrices H_{11} and H'_{11} are Hermitian, and (45) is analogous in form to (29). Reasoning here on H and R as on H and P in the proof of theorem 1, we have

$$|\lambda_i(H_{11}) - \lambda_i(H'_{11})| \leqslant \|H_{11} - H'_{11}\|_2 \leqslant \varepsilon\sigma_{\max}(R), \tag{46}$$

where $\|H_{12}\|_2 \leqslant \|H_{12}\|_F = \varepsilon$ has been taken into account. Finally, the bound (38) follows from (44), (46) and (see [4, p. 286])

$$|\sigma_i(A_{11}) - \sigma_i(A'_{11})| \leqslant \|A_{11} - A'_{11}\|_2 = \|H_{11} - H'_{11}\|_2. \quad \blacksquare \tag{47}$$

EXAMPLE 2. In order to illustrate the bounds of theorems 1 and 2, we consider the following two matrices:

$$H = \begin{pmatrix} 1 & \varepsilon \\ \varepsilon & 1+\delta \end{pmatrix} \quad \text{and} \quad A = \begin{pmatrix} -1 & -\varepsilon \\ \varepsilon & 1+\delta \end{pmatrix},$$

which have off-norms and separations in their spectrum equal to

$$\|H_{12}\|_F = \varepsilon, \qquad \|A_{12}, A_{21}^*\|_F \doteq \varepsilon' = \sqrt{2}\,\varepsilon,$$

$$|\lambda(H_{11}) - \lambda(H_{22})| = \delta, \qquad |\sigma(A_{11}) - \sigma(A_{22})| = \delta,$$

where we assume $0 < 2\varepsilon < \delta \ll 1$. One then easily checks that

$$|\lambda_1(H_{11}) - \lambda_1(H)| = |\sigma_1(A_{11}) - \sigma_1(A)| = 2\frac{\varepsilon^2}{\delta}\left[\frac{1}{1+\sqrt{1+4\varepsilon^2/\delta^2}}\right] < \frac{\varepsilon^2}{\delta}.$$

Notice that for $\varepsilon/\delta \to 0$ (i.e. for $\kappa \to 0$ in [13]) one has

$$|\lambda_1(H_{11}) - \lambda_1(H)| \to \frac{\varepsilon^2}{\delta},$$

which tends to the upper bound (27), while for the matrix A one has

$$|\sigma_1(A_{11}) - \sigma_1(A)| \to \frac{\varepsilon'^2}{2\delta},$$

which is still a factor 2 away from the upper bound (38).

Notice also that in the above theorem we have assumed that the blocks A_{11} and A_{22} are square. Extensions to nonsquare diagonal blocks could be considered, but are not relevant to our later results.

4. TRIANGULAR MATRICES WITH CLUSTERS OF SINGULAR VALUES

In this section we combine the results of the two previous sections to investigate the off-norm of a *triangular* principal submatrix, approximating a cluster of singular values. The following theorem is analogous to Wilkinson's result on Hermitian matrices in [17].

THEOREM 3. *Let A be a square upper triangular (complex) matrix with off-norm $\|A_{\text{off}}\|_F = \varepsilon$. Consider any subset S of k singular values of A. Let η be the width of S:*

$$\max_{\sigma_i, \sigma_j \in S} |\sigma_i - \sigma_j| = \eta, \tag{48}$$

and 2δ its distance from the other singular values:

$$\min_{\sigma_i \in S,\ \sigma_j \notin S} |\sigma_i - \sigma_j| = 2\delta. \tag{49}$$

Then, if $\delta > 2\varepsilon$, the off-norm of the principal submatrix whose diagonal elements in modulus are ε-close to the elements of S, say A_S, is bounded as

$$\|(A_S)_{\text{off}}\|_{2,F} < \sqrt{k-1}\left(\frac{4\varepsilon^2}{\delta} + \eta\right). \tag{50}$$

Proof. Notice that the singular values of A are ε-close to the moduli of the diagonal elements of A, i.e.

$$\left|\sigma_j(A) - |a_{ii}|\right| \leqslant \|A_{\text{off}}\|_2 \leqslant \varepsilon \tag{51}$$

[4, p. 286] for an adequate ordering of the indices j and i. Since $\delta > 2\varepsilon$, the determination of A_S is thus unambiguous. It is easily seen that there exists a symmetric permutation of rows and columns of A such that the permuted matrix can be partitioned as indicated in (34) with $A_{11} \equiv A_S$:

$$\begin{pmatrix} A_S & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \tag{52}$$

and such that the diagonal elements are preserved. Since $\|(A_{12}, A_{21}^*)\|_F \leqslant \varepsilon$, the singular values of A_S and A_{22} are ε-close to those of A [4, p. 286], and thus

$$\min\left|\operatorname{sing}(A_S) - \operatorname{sing}(A_{22})\right| \geqslant 2\delta - 2\varepsilon > \delta. \tag{53}$$

Then, applying Theorem 2, we have that the distance between each singular value of A_S and the corresponding singular value of A is bounded by $2\varepsilon^2/\delta$. By this and (48), it thus follows that

$$\sigma_{\max}(A_S) - \sigma_{\min}(A_S) < \frac{4\varepsilon^2}{\delta} + \eta. \tag{54}$$

Then (50) is deduced from Lemma 2 and the triangularity of A_S. ∎

REMARK 4. Although no constraints are imposed on η, this theorem is especially useful in practice if $\eta \ll \varepsilon^2/\delta$, the bound (50) being then dominated by the term in ε^2/δ. In particular, when S represents a repeated singular value of A, we have $\eta = 0$.

5. KOGBETLIANTZ'S SVD ALGORITHM

Kogbetliantz's method [5, 6] for computing the SVD of an arbitrary $m \times n$ matrix A $(m \geqslant n)$, consists of generating a sequence of matrices $A^{(l)}$ as follows

$$\begin{aligned} &U^{(0)} := I_m, \qquad V^{(0)} := I_n, \qquad A^{(0)} := A, \\ &U^{(l)} := U_l U^{(l-1)}, \qquad V^{(l)} := V_l V^{(l-1)}, \qquad A^{(l)} := U_l A^{(l-1)} V_l^* \end{aligned} \tag{55}$$

such that $\|A_{\text{off}}^{(l)}\|_F$ decreases and eventually becomes negligible (of the order of the relative precision of the machine one is working with) [3]. The unitary updating transformations U_l and V_l are chosen to be complex elementary rotations through angles ϕ_l and ψ_l which annihilate the elements of $A^{(l-1)}$ in symmetric positions (i_l, j_l) and (j_l, i_l). For a general matrix, it can be shown [1] (see also references therein) that the computational burden is reduced, on a sequential machine as well as on a parallel one, if a preliminary QR decomposition is performed:

$$A = Q\begin{pmatrix} R \\ 0 \end{pmatrix}. \tag{56}$$

In [1], implementation aspects were investigated, e.g. the possibility of using approximate rotations in the case of triangular matrices, in order to speed up the computation of their singular values. Here, an additional and theoretical advantage of triangular matrices is stressed: the following theorem shows that Kogbetliantz's algorithm converges quadratically in the presence of repeated or very close singular values provided they correspond to adjacent positions on the diagonal. That this is not true for full matrices or when the cluster of singular values is separate is illustrated by Examples 3 and 4. So far, quadratic convergence has been proved only for a full arbitrary matrix with singular values that are sufficiently distant from each other [9].

The theorem given here holds when the off-diagonal elements are annihilated successively by rows (or by columns), in the same manner as in the special cyclic Jacobi method for symmetric or Hermitian matrices [15]. It is to be noted that an equivalent ordering is well suited to parallel implementations [7, 8], where appropriate permutations at each step allow one to perform transformations on neighboring rows and columns only. Convergence of other orderings (as corresponding to the classical Jacobi method) or of variants of the method (as a "threshold" strategy suggested by Wilkinson [17]) is briefly commented on thereafter.

THEOREM 4. *Let A be a square triangular (complex) matrix with a cluster C of k singular values, of width η. Assume that the singular values of*

A are at least 2δ distant from each other, except for those within the cluster:

$$\min_{\sigma_i \text{ or } \sigma_j \notin C} |\sigma_i - \sigma_j| = 2\delta. \tag{57}$$

If the diagonal elements which converge to the singular values of C occupy successive positions on the diagonal, then the convergence of Kogbetliantz's SVD algorithm applied to A is ultimately quadratic. More precisely: if $\|A_{\text{off}}^{(r)}\|_F < \delta/2$ and $\eta = c\|A_{\text{off}}^{(r)}\|_F^2/\delta$, then after one supplementary sweep of $N = n(n-1)/2$ rotations, one has

$$\|A_{\text{off}}^{(r+N)}\|_F < c' \frac{\|A_{\text{off}}^{(r)}\|_F^2}{\delta}, \qquad c' \doteq \sqrt{(k-1)^2(4+c)^2 + 8}\,. \tag{58}$$

Proof. The proof is based on Theorem 3 and the analysis of Paige and Van Dooren [9] for the case when all the singular values are 2δ distant from each other. The argument is directly inspired by that of van Kempen [15], who proved the quadratic convergence of the special Jacobi method for a real symmetric matrix with a multiple eigenvalue.

We first remark that if the diagonal elements corresponding to C are adjacent on the diagonal and if the condition $\|A_{\text{off}}^{(r)}\|_F < \delta/2$ is verified, then they remain adjacent through any subsequent rotation; also, the ordering of the other diagonal elements is maintained [3, Lemma 6]. Now we assume, for convenience and without loss of generality, that after l rotations ($l > r$), the matrix $A^{(l)}$ can be partitioned as

$$A^{(l)} = \begin{pmatrix} A_C^{(l)} & A_{12}^{(l)} \\ A_{21}^{(l)} & A_{22}^{(l)} \end{pmatrix}, \tag{59}$$

where $A_C^{(l)}$ is the $k \times k$ matrix whose diagonal elements converge to the singular values of the cluster. The matrix $A^{(l)}$ is triangular at the completion of a sweep; at any other step l, it is triangular up to a symmetric permutation [8]. Therefore, Theorem 3 is applicable for any l ($> r$), which leads to

$$\|(A_C^{(l)})_{\text{off}}\|_F < \sqrt{k-1}\left(\frac{4\|A_{\text{off}}^{(l)}\|_F^2}{\delta} + \eta\right) \leqslant \sqrt{k-1}\left(\frac{4\|A_{\text{off}}^{(r)}\|_F^2}{\delta} + \eta\right). \tag{60}$$

On the other hand, if the rotation l annihilates $a_{ij}^{(l-1)}$ outside A_C, it can be shown [9, Lemma 1] that the rotation angles ϕ_l and ψ_l satisfy

$$\sqrt{\sin^2\phi_l + \sin^2\psi_l} < \frac{2}{\delta}\left|a_{ij}^{(l-1)}\right|. \tag{61}$$

Since $|a_{ij}^{(l-1)}|^2 = \|A_{\text{off}}^{(l-1)}\|_F^2 - \|A_{\text{off}}^{(l)}\|_F^2$, it thus follows that

$$\sum{}'\left(\sin^2\phi_l + \sin^2\psi_l\right) < \frac{4}{\delta^2}\left\|A_{\text{off}}^{(r)}\right\|_F^2, \tag{62}$$

where Σ' denotes that only the rotations which annihilate elements outside A_C are included. Furthermore, after the annihilation of the first row (starting from the rotation $r+1$ and assuming $A_{21}^{(r)} = 0$ for convenience), one finds, in the same way as in [9],

$$\sum_{i=k+1}^{n}\left|a_{1i}^{(r+n-1)}\right|^2 \leqslant 2\left\|A_{\text{off}}^{(r)}\right\|_F^2 \sum_{l=r+k}^{r+n-1}\left(\sin^2\phi_l + \sin^2\psi_l\right). \tag{63}$$

For the entire first row, one obtains then

$$\begin{aligned}\sum_{i=2}^{n}\left|a_{1i}^{(r+N)}\right|^2 &= \sum_{i=2}^{n}\left|a_{1i}^{(r+n-1)}\right|^2 \\ &\leqslant \left\|\left(A_C^{(r+n-1)}\right)_{\text{off}}\right\|_F^2 + 2\left\|A_{\text{off}}^{(r)}\right\|_F^2 \sum_{l=r+k}^{r+n-1}\left(\sin^2\phi_l + \sin^2\psi_l\right),\end{aligned} \tag{64}$$

since the sum of the squares of the elements in a row is unaltered by the subsequent rotations of the same sweep. Similar inequalities hold for the other rows of A, and from

$$\left\|A_{\text{off}}^{(r+N)}\right\|_F^2 = \sum_{i,j\leqslant k;\, i\neq j}\left|a_{ij}^{(r+N)}\right|^2 + \sum_{i,j>k;\, i\neq j}\left|a_{ij}^{(r+N)}\right|^2,$$

one has then

$$\left\|A_{\text{off}}^{(r+N)}\right\|_F^2 < (k-1)^2\left(\frac{4\left\|A_{\text{off}}^{(r)}\right\|_F^2}{\delta} + \eta\right)^2 + \frac{8}{\delta^2}\left\|A_{\text{off}}^{(r)}\right\|_F^4, \tag{65}$$

by summing the inequalities (64) and taking (60) and (62) into account. Finally, we derive (58) using $\eta = c\|A_{\text{off}}^{(r)}\|_F^2/\delta$. ∎

REMARK 5. The bound (58) is analogous to van Kempen's result [15] for Jacobi's algorithm applied to symmetric matrices, except for the coefficient c'. The following is to be noted about this discrepancy:

(1) The factor k^2 under the square root can be reduced to k by using an appropriately adapted version of Lemma 2 in order to derive a (better) bound for just one row of $(A_C)_{\text{off}}$ at a time. This allows one to write an analogue of (60) for a row, which is then directly used in the summation (65).

(2) van Kempen does not consider clusters, but only multiple eigenvalues $(\eta = c = 0)$.

(3) The bound of van Kempen does not contain k at all. It appears that this is because he simply dismisses (perhaps inaccurately) the contribution of $(A_C)_{\text{off}}$ in his proof.

(4) van Kempen only retains half of the off-diagonal elements of the symmetric matrix in his definition of the off-norm, which then leads to a smaller coefficient c'.

Moreover, when more than one cluster is present the coefficient c' ought to be adapted to (see also van Kempen [15])

$$c' \doteq \sqrt{\sum_{i=1}^{L} (k_i - 1)^2 (4 + c_i)^2 + 8}\,, \tag{66}$$

where we have assumed that there are L clusters with respective size k_i and span $\eta_i = c_i\|A_{\text{off}}\|_F^2/\delta$. These differences in the value of c' do not affect the basic result of the theorem, namely that the convergence is quadratic from one sweep to another. In practice the coefficient c' of (58) or (66) appears to be seriously overestimated (see Example 4).

The following two examples were run on a VAX 780 with relative precision 1.4×10^{-17} in double precision.

EXAMPLE 3. Here we take anew the 15×15 matrix of Example 1, with singular values clustered around 1 ($\eta \approx 1.25 \times 10^{-14}$). In Table 1 we give the off-norms obtained at successive steps of the full Kogbetliantz algorithm

TABLE 1

Step	A	R
0	3.7653D+00	1.9902D − 14
1	1.1735D+00	9.9565D − 15
2	2.2655D − 01	4.1322D − 15
3	1.4825D − 02	6.8737D − 16
4	4.9288D − 04	
5	9.2862D − 05	
6	4.1227D − 06	
7	2.7340D − 07	
8	8.1454D − 08	
9	3.3878D − 08	
10	6.0522D − 09	
11	9.1054D − 10	
12	7.4801D − 12	
13	4.1470D − 16	

applied to A on the one hand, and of the triangular Kogbetliantz algorithm applied to the triangular factor R on the other hand. The iterations were continued until the off-norm was below 6.9×10^{-16} (50 times the relative precision). The numbers in the first line (step 0) are equal to $\|A_{\mathrm{off}}\|_F$ and $\|R_{\mathrm{off}}\|_F$, respectively.

The observed convergence can be explained via two different interpretations:

(1) All singular values belong to *one* cluster of span η. As shown in Example 1, the triangular matrix R has a much smaller off-norm to start with (because of Lemma 2). For this interpretation, Theorem 4 cannot be applied, since δ is undefined.

(2) All singular values do not belong to one cluster. In this case δ is defined but very small ($<\eta$). For R one can rely on Theorem 4 as soon as $\|R_{\mathrm{off}}\|_F$ is smaller than $\delta/2 \approx 10^{-15}$, which occurs only at the end of the convergence. For A the same limit has to be considered before applying the corresponding result [9] for arbitrary matrices with distinct singular values. This is the reason why quadratic convergence is not observed in any of the two cases.

In the proof of Theorem 4 we have used (through [9, 15, 16]) the fact that the diagonal elements approximating singular values of a same cluster occupy adjacent positions. The following example shows that this condition is also a necessary one.

EXAMPLE 4. The following 6×6 symmetric matrix was generated (using MATLAB):

```
A      =

    COLUMNS  1 THRU  3
  0.999999999999992  -0.000000000469621  -0.000000000000014
 -0.000000000469621   0.500000000000004   0.000000000306453
 -0.000000000000014   0.000000000306453   0.999999999999991
 -0.000000000013428   0.000000000000019  -0.000000000000283
  0.000000000000005  -0.000000000017039   0.000000000000027
 -0.000000000014107  -0.000000000000002  -0.000000000000016

    COLUMNS  4 THRU  6
 -0.000000000013428   0.000000000000005  -0.000000000014107
  0.000000000000019  -0.000000000017039  -0.000000000000002
 -0.000000000000283   0.000000000000027  -0.000000000000016
  0.499999999999983   0.000000000000370  -0.000000000000029
  0.000000000000370   0.999999999999991  -0.000000000000023
 -0.000000000000029  -0.000000000000023   0.500000000000016
```

Its eigenvalues are:

$$\begin{array}{ccc} 0.5-3.9510\text{D}-14, & 0.5+4.6213\text{D}-15, & 0.5+3.7428\text{D}-14, \\ 1-4.1619\text{D}-14, & 1-3.6499\text{D}-15, & 1+1.9762\text{D}-14. \end{array}$$

The matrix A has thus two clusters with $\eta_1 \approx 7.7 \times 10^{-14}$ and $\eta_2 \approx 6.1 \times 10^{-14}$. Furthermore, the (Frobenius) off-norm and the distance 2δ are equal to:

$$\text{OFF} = 7.9389 \times 10^{-10}, \qquad 2\delta \approx 0.5$$

Performing one sweep of the symmetric Jacobi algorithm yields

```
A      =

    COLUMNS  1 THRU  3
  1.000000000000015  -0.000000000103790  -0.000000000000009
 -0.000000000103790   0.499999999999998  -0.000000000004981
 -0.000000000000009  -0.000000000004981   0.999999999999960
  0.000000000108905  -0.000000000000016  -0.000000000042436
 -0.000000000000008  -0.000000000009706   0.000000000000000
  0.000000000075669   0.000000000000004   0.000000000011161
```

```
   COLUMNS  4 THRU  6
  0.000000000108905  -0.000000000000008   0.000000000075669
 -0.000000000000016  -0.000000000009706   0.000000000000004
 -0.000000000042436   0.000000000000000   0.000000000011161
  0.499999999999967  -0.000000000001357   0.000000000000000
 -0.000000000001357   1.000000000000000   0.000000000000000
  0.000000000000000   0.000000000000000   0.500000000000037

OFF = 2.4659D-10
```

This clearly lies far above van Kempen's equivalent to the bound (66), which is approximately equal to $(k-1)\sqrt{c_1^2 + c_2^2\|A_{\text{off}}^{(r)}\|_F^2/\delta} = (k-1)\sqrt{\eta_1^2+\eta_2^2} = 2\times 10^{-13}$, since the contribution of the clusters is dominant in (66).

Let us now perform the QR decomposition of the matrix A. The singular values of the factor R are equal to the eigenvalues of A. The triangular matrix R and its off-norm are

```
R      =

   COLUMNS  1 THRU  3
 -0.999999999999992   0.000000000704432   0.000000000000027
  0.0                -0.500000000000004  -0.000000000919360
  0.0                 0.0                -0.999999999999991
  0.0                 0.0                 0.0
  0.0                 0.0                 0.0
  0.0                 0.0                 0.0

   COLUMNS  4 THRU  6
  0.000000000020141  -0.000000000000011   0.000000000021160
 -0.000000000000038   0.000000000051117   0.000000000000005
  0.000000000000426  -0.000000000000055   0.000000000000024
 -0.499999999999953  -0.000000000001110   0.000000000000059
  0.0                -0.999999999999991   0.000000000000036
  0.0                 0.0                 0.500000000000016

OFF = 1.1597D-09
```

and performing one step of the triangular Kogbetliantz algorithm gives

```
R        =

    COLUMNS   1 THRU   3
 -1.000000000000015   0.0                  0.0
  0.000000000311279  -0.499999999999998    0.0
  0.000000000000019   0.000000000007466   -0.999999999999960
 -0.000000000326632   0.000000000000032    0.000000000127393
  0.000000000000017   0.000000000014564    0.000000000000000
  0.000000000226923   0.000000000000008    0.000000000033515

    COLUMNS   4 THRU   6
  0.0                 0.0                  0.0
  0.0                 0.0                  0.0
  0.0                 0.0                  0.0
 -0.499999999999967   0.0                  0.0
  0.000000000002037  -1.000000000000000    0.0
  0.000000000000000   0.000000000000000    0.500000000000037

OFF = 5.2221D-10
```

which thus displays the same behavior as the symmetric Jacobi algorithm. The reason for not obeying the bound (66) (or van Kempen's bound for the symmetric matrix A) is indeed the fact that the diagonal elements corresponding to one cluster are *not* adjacent. We now permute columns and rows of A to yield the reordered matrix A':

```
A'       =

    COLUMNS   1 THRU   3
  0.999999999999992  -0.000000000000014    0.000000000000005
 -0.000000000000014   0.999999999999991    0.000000000000027
  0.000000000000005   0.000000000000027    0.999999999999991
 -0.000000000469621   0.000000000306453   -0.000000000017039
 -0.000000000013428  -0.000000000000283    0.000000000000370
 -0.000000000014107  -0.000000000000016   -0.000000000000023
```

```
    COLUMNS  4 THRU  6
 -0.000000000469621  -0.000000000013428  -0.000000000014107
  0.000000000306453  -0.000000000000283  -0.000000000000016
 -0.000000000017039   0.000000000000370  -0.000000000000023
  0.500000000000004   0.000000000000019  -0.000000000000002
  0.000000000000019   0.499999999999983  -0.000000000000029
 -0.000000000000002  -0.000000000000029   0.500000000000016
OFF = 7.9389D-10
```

and after one sweep of the symmetric Jacobi algorithm we obtain

```
A'     =
    COLUMNS  1 THRU  3
  1.000000000000015  -0.000000000000009  -0.000000000000008
 -0.000000000000009   0.999999999999960   0.000000000000000
 -0.000000000000008   0.000000000000000   1.000000000000000
  0.000000000000000   0.000000000000000   0.000000000000000
  0.000000000000000   0.000000000000000   0.000000000000000
  0.000000000000000   0.000000000000000   0.000000000000000

    COLUMNS  4 THRU  6
  0.000000000000000   0.000000000000000   0.000000000000000
  0.000000000000000   0.000000000000000   0.000000000000000
  0.000000000000000   0.000000000000000   0.000000000000000
  0.499999999999998  -0.000000000000016   0.000000000000004
 -0.000000000000016   0.499999999999967   0.000000000000000
  0.000000000000004   0.000000000000000   0.500000000000037
OFF = 2.9503D-14
```

which is now indeed smaller than the bound 2×10^{-13}. Let us now perform again the QR-decomposition of this reordered matrix:

```
R'     =
    COLUMNS  1 THRU  3
 -0.999999999999992   0.000000000000027  -0.000000000000011
  0.0                -0.999999999999991  -0.000000000000055
  0.0                 0.0                -0.999999999999991
  0.0                 0.0                 0.0
  0.0                 0.0                 0.0
  0.0                 0.0                 0.0
```

```
   COLUMNS   4 THRU  6
  0.000000000704432  0.000000000020141   0.000000000021160
 -0.000000000459680  0.000000000000425   0.000000000000024
  0.000000000025558 -0.000000000000555   0.000000000000035
 -0.500000000000004 -0.000000000000038   0.000000000000005
  0.0               -0.499999999999983   0.000000000000059
  0.0                0.0                 0.500000000000016

OFF = 8.4204D-10
```

Then after one sweep of the triangular Kogbetliantz algorithm, we have

```
R'     =

   COLUMNS   1 THRU  3
 -1.000000000000015  0.0                 0.0
  0.000000000000019 -0.999999999999960   0.0
  0.000000000000017  0.000000000000000  -1.000000000000000
  0.000000000000000  0.000000000000000   0.000000000000000
  0.000000000000000  0.000000000000000   0.000000000000000
  0.000000000000000  0.000000000000000   0.000000000000000

   COLUMNS   4 THRU  6
  0.0               0.0                  0.0
  0.0               0.0                  0.0
  0.0               0.0                  0.0
 -0.499999999999998  0.000000000000000   0.000000000000000
  0.000000000000032 -0.499999999999967   0.000000000000000
  0.000000000000008  0.000000000000000   0.500000000000037

OFF = 4.1715D-14
```

which is also smaller than the bound 2×10^{-13}. We draw attention here to the fact that, although the off-norms of A' and R' become of the order of η after one sweep, much smaller values (namely of the order of $\text{OFF}^2/\delta \approx 10^{-18}$) are observed outside the diagonal blocks corresponding to the clusters. This is in fact explained by the bound (65), where the last term is precisely that contribution to the off-norm.

REMARK 6. From what precedes, one concludes that, if the cyclic by rows (or columns) ordering of annihilations is used in the presence of multiple or clustered singular values, Kogbetliantz's algorithm converges quadratically only for triangular matrices and if the cluster(s) are grouped. However, quadratic convergence is also ensured in the two following situations:

(1) If a different ordering is considered, where the off-diagonal element of largest magnitude is annihilated at each step. This corresponds to the classical Jacobi method, for which van Kempen has proven the ultimate quadratic convergence in the case of multiple eigenvalues, either grouped or separate [14]. Notice that this ordering is not suited for a parallel implementation of the method. Also the structure of the matrix (e.g. triangular) is not maintained.

(2) If a "threshold" strategy is applied. This was proposed by Wilkinson [17] for Hermitian matrices and can be directly extended to triangular matrices. At any stage, the rotation is taken as the identity matrix if the element to be annihilated is smaller than a given threshold relative to the current off-norm. Therefore, large angles due to very close diagonal elements are not propagated, which avoids losing quadratic convergence. Here the structure of the matrix is retained during the transformations. On a parallel machine, though, such a strategy might be difficult to apply, unless a local threshold could be defined.

6. CONCLUSION

In this paper we have analyzed the convergence of Kogbetliantz's algorithm for computing the SVD of a matrix. A result was obtained for *triangular matrices* which strongly resembles the convergence properties of Jacobi's algorithm for Hermitian or normal matrices: ultimate quadratic convergence is guaranteed even in the presence of multiple or clustered singular values *provided the diagonal elements corresponding to the singular values of the same cluster occupy adjacent positions.*

Counterexamples were also given to illustrate the lack of quadratic convergence when:

(1) the full Kogbetliantz algorithm is used instead of the triangular one (Example 3);

(2) clusters are not grouped in the triangular Kogbetliantz algorithm, as in Jacobi's algorithm for Hermitian matrices (Example 4).

It is also indicated how this constraint can be removed, e.g. by threshold strategies. Along the way, a number of additional results were obtained:

(1) Sharp bounds for the off-norm of matrices were derived in terms of the span of the spectrum of their eigenvalues or singular values (Section 2).

(2) Perturbation bounds for eigenvalues and singular values of (block) diagonally dominant matrices were sharpened or derived (Section 3).

Although these results are introduced here for proving the quadratic convergence of Kogbetliantz's triangular SVD algorithm, we think that they have their own potential use in other problems.

We want to thank M. Vanbegin of PRLB for a fruitful collaboration related to this work, especially through [1]. During the preparation of this paper, our attention was drawn to (far less complete) results of B. Zhaojun of Fudan University (Shangai, China) along the lines of our Section 5. These results are not published yet.

REFERENCES

1 J.-P. Charlier, M. Vanbegin, and P. Van Dooren, On efficient implementation of Kogbetliantz's algorithm for computing the singular value decomposition, submitted for publication.

2 C. Davis and W. Kahan, The rotation of eigenvectors by a perturbation III, *SIAM J. Numer. Anal.* 7:1–46 (1970).

3 G. Forsythe and P. Henrici, The cyclic Jacobi method for computing the principal values of a complex matrix, *Trans. Amer. Math. Soc.* 94:1–23 (1960).

4 G. Golub and C. Van Loan, *Matrix Computations*, North Oxford Academic, Oxford, 1983.

5 E. Kogbetliantz, Diagonalization of general complex matrices as a new method for solution of linear equations, *Proc. Internat. Congr. Math. Amsterdam* 2:356–357 (1954).

6 E. Kogbetliantz, Solution of linear equations by diagonalization of coefficient matrices, *Quart. Appl. Math.* 13:123–132 (1955).

7 F. Luk, A Triangular Processor Array for Computing the Singular Value Decomposition, Tech. Report 84-625, Computer Science Dept., Cornell Univ., 1984.

8 C. Paige, Computing the generalized singular value decomposition, *SIAM J. Sci. Statist. Comput.*, to appear.

9 C. Paige and P. Van Dooren, A note on the convergence of Kogbetliantz's iterative algorithm for obtaining the singular value decomposition, *Linear Algebra Appl.* 77:301–313 (1986).

10 B. N. Parlett, *The Symmetric Eigenvalue Problem*, Prentice-Hall, Englewood Cliffs, N.J., 1980.

11 A. Ruhe, On the quadratic convergence of the Jacobi method for normal matrices, *BIT* 7:305–313 (1968).

12 A. Schönhage, Zur Konvergenz des Jacobi-Verfahrens, *Numer. Math.* 3:374–380 (1961).
13 G. W. Stewart, Error and perturbation bounds for subspaces associated with certain eigenvalues problems, *SIAM Rev.* 15:727–764 (1973).
14 H. P. M. van Kempen, On the convergence of the classical Jacobi method for real symmetric matrices with non-distinct eigenvalues, *Numer. Math.* 9:11–18 (1966).
15 H. P. M. van Kempen, On the quadratic convergence ot the special cyclic Jacobi method, *Numer. Math.* 9:19–22 (1966).
16 J. H. Wilkinson, Note on the quadratic convergence of the cyclic Jacobi process, *Numer. Math.* 4:296–300 (1962).
17 J. H. Wilkinson, Almost diagonal matrices with multiple or close eigenvalues, *Linear Algebra Appl.* 1:1–12 (1968).

Received June 1986; revised 20 November 1986