J. Korean Soc. Ind. Appl. Math. Vol.23, No.3, 267–282, 2019 http://doi.org/10.12941/jksiam.2019.23.267

ON REDUCTION OF K-ALMOST NORMAL AND K-ALMOST CONJUGATE NORMAL MATRICES TO A BLOCK TRIDIAGONAL FORM

K. NIAZI ASIL[1], AND M. GHASEMI KAMALVAND[1†]

[1]DEPARTMENT OF MATHEMATICS, LORESTAN UNIVERSITY, KHORRAMABAD, IRAN
Email address: kniaziasil@yahoo.com, ghasemi.m@lu.ac.ir

ABSTRACT. This paper examines how one can build a block tridiagonal structure for k-almost normal matrices and also for k-almost conjugate normal matrices. We shall see that these representations are created by unitary similarity and unitary congruance transformations, respectively. It shall be proven that the orders of diagonal blocks are $1, k+2, 2k+3, \ldots$, in both cases. Then these block tridiagonal structures shall be reviewed for the cases where the mentioned matrices satisfy in a second-degree polynomial. Finally, for these processes, algorithms are presented.

1. INTRODUCTION

We know that the normal matrix in the field of numerical linear algebra has special features which caused researchers to be attracted to this area and similar definitions have been created. Below, three examples of these new matrices which are inspired by normal ones have been introduced. They are conjugate normal matrices, k-almost normal matrices and k-almost conjugate normal matrices, respectively. These new matrices, in turn, have ramarkable properties and are sometimes similar to the properties of the normal matrix. In this paper, we consider an interesting feature of the normal matrices. This property is known as the unitary tridiagonalizability. The details of this property are discussed in the theorem (1.1). The same issue also is expressed about conjugate normal matrices and has been proven, (see [1]). In this paper, the aim is to get the similar structures for k-almost normal matrices and also for k-almost conjugate normal matrices. Our main tool for achieving this important goal are generalized Krylov subspaces and an algorithm similar to generalized Lanczos for these subspaces.
Continuing this section, we mention the definitions, some examples and some properties of k-almost normal and k-almost conjugate normal matrices and then, in the second section we prove that every k-almost normal matrix is unitarily similar to a block tridiagonal matrix and we get some interesting results about the diagonal blocks orders. In secton three we obtain similar results for k-almost conjugate normal matrices but this time through an unitary congruence, not an unitary similarity. In Section four, we present an algorithm containing the details of the

Received by the editors August 21 2019; Accepted September 11 2019; Published online September 25 2019.
2000 *Mathematics Subject Classification.* 15A24 - 15B99.
Key words and phrases. k-almost normal matrix, k-almost conjugate normal matrix, block tridiagonal matrix.
† Corresponding author.

computations of the two preceding sections. Section five contains some examples. Finally, we come to the conclusion of the paper.

Definition 1.1. *A square complex matrix A is said to be conjugate normal when $AA^* = \overline{A^*A}$ where A^* is the conjugate transpose of A.*

This category of matrices in the theory of the unitary congruence plays a similar role as normal matrices have in unitary similarity discussions.

Definition 1.2. *An $n \times n$ complex matrix A is said to be k-almost normal when there exists a matrix $C \in M_n(\mathbb{C})$ of rank k, such that $A(A^* - C) = (A^* - C)A$.*

In [2], the structure of these matrices for $k = 1$ has been investigated. Also, necessary and sufficient conditions for a matrix to belong to this class of matrices are given and a canonical representation as a block tridiagonal matrix is shown. In [3], some statements for k-almost normal matrices are expressed and it is proved that every k-almost normal matrix can be unitarily reduced to a block tridiagonal form where each block diagonal size is maximum of $2k$. This reduction relies on the block Lanczos method, starting from a basis of the column space $A^*A - AA^*$, for k-almost normal A.

Example 1.1. *Let $A = \begin{pmatrix} 1 & 1 \\ 1 & i \end{pmatrix}$. Then, A is 1-almost normal by selecting $C = \begin{pmatrix} 2i & 0 \\ 0 & 0 \end{pmatrix}$.*

Definition 1.3. *An $n \times n$ complex matrix A is called a k-almost conjugate normal matrix for which there exists a matrix $C \in M_n(\mathbb{C})$ of rank k, such that $A(A^* - C) = \overline{(A^* - C)A}$. In other words, $AA^* - AC = A^T\overline{A} - \overline{CA}$.*

Example 1.2. *Let $A = \begin{pmatrix} 0 & e \\ f & 0 \end{pmatrix}$ and $C = \begin{pmatrix} a & \frac{f^2-e^2}{f} \\ 0 & \bar{a} \end{pmatrix}$, in which $e, f \in \mathbb{R}$ where $f \neq 0$ and $a \in \mathbb{C}$. If $a = 0$ than A is 1-almost conjugate normal and if $a \neq 0$ then, A is 2-almost conjugate normal matrix. That is , $AA^* - A^T\overline{A} = AC - \overline{CA}$.*

Example 1.3. *Let $A = diag(a_{11}, ..., a_{nn})$, where each a_{ii} is real or pure imaginary. Then A is k-almost conjugate normal, for $k = 1, ..., n$. For this, let $C = diag(c_{11}, ..., c_{nn})$, in which the belonging of c_{ii} to $\mathbb{R}$ or $i\mathbb{R}$ is same as a_{ii}.*

Example 1.4. *Let $A = \begin{pmatrix} i & -1 \\ 1 & 0 \end{pmatrix}$. Then, by choosing $C = \begin{pmatrix} -2i & 0 \\ 0 & 0 \end{pmatrix}$, A is 1-almost conjugate normal.*

Proposition 1.1. *Below are listed some of the trivial properties of k-almost normal and k-almost conjugate normal matrices.*

(1). If $n \times n$ matrix A is k-almost (conjugate) normal, then $\bar{A}, A^T, A^*$ are also k-almost (conjugate) normal matrices.

(2). Suppose that $A \in M_n$ and $B \in M_m$ are k-almost (conjugate) normal and t-almost (conjugate) normal matrices, respectively. Then, $A \oplus B$ is $(k + t)$-almost (conjugate) normal matrix.

(3). Every $n \times n$ diagonal matrix is k-almost normal, for $k = 1, ..., n$.

(4). matrix $A \in M_n$ is k-almost (conjugate) normal if and only if U^*AU is k-almost (conjugate) normal, for every unitary matrix U.

(5). matrix $A \in M_n$ is k-almost (conjugate) normal if and only if $T = (A^* - C)A - A(A^* - C)$, $(T = (A^T - \overline{C})\overline{A} - A(A^* - C))$, is semidefinite positive for some matrix $C \in M_n$ where $rankC = k$.

Proof. Arguments of (1) and (2) are straightforward. For part (3), let C a diagonal matrix with $rankC = k$, for each $k = 1, ..., n$. For (4), suppose that A is k-almost normal. I.e., there is matrix C where $rankC = k$ such that $A(A^* - C) = (A^* - C)A$. Then, for any unitary matrix we have

$$U^*AU(U^*A^*U - U^*CU) = U^*A(A^* - C)U = U^*(A^* - C)AU = (U^*A^*U - U^*CU)U^*AU.$$

This result, together with that $rankC = rankU^*CU = k$ proves that U^*AU is a k-almost normal matrix. For sufficiency, let $U = I_n$. (The same for the conjugate mode). For (5), it is clear that for k-almost normal matrix A, there exists matrix C where $rankC = k$ and $(A^* - C)A - (A^* - C)A = 0$ and the zero matrix is a semidefinite positive matrix. Conversely, suppose that T is a semidefinite positive matrix. Then, $traceT = 0$. But the zero matrix is the only semidefinite matrix with trace 0. Thus, $T = 0$. I.e., A is a k-almost normal matrix. (The same for the conjugate mode). □

We consider a special theorem that has been proven for normal and conjugate normal matrices and the goal is its reviewing for k-almost normal and k-almost conjugate normal matrices, respectively.

Theorem 1.1. *A normal $n \times n$ matrix A can be brought by a finite orthogonal process to the block tridiagonal form*

$$H = \begin{pmatrix} H_{11} & H_{12} & & \\ H_{21} & H_{22} & H_{23} & \\ & H_{32} & H_{33} & \ddots \\ & & \ddots & \ddots \end{pmatrix}, \tag{1.1}$$

where the diagonal blocks $H_{11}, H_{22}, ...$ are square and their orders are generically given by integers $1, 2, ...$. If A satisfies the equations $g(A, A^*) = 0$, where $g(x, y)$ is a polynomial of degree $m \ll n$, then, starting from $i = m$, the orders of diagonal blocks H_{ii} in matrix H stabilize at the value m.

Proof. See [4]. □

This theorem was a prelude to the next task in [1], to prove a similar statement for conjugate normal matrices. In fact, its authors offered a generalization of the CSYM algorithm to the conjugate normal matrices mode. They performed this by focusing on the generalized Krylov subspaces. CSYM algorithm is an algorithm based on the symmetric matrix reduction to a tridiagonal form, via unitary congruence, (see [5]).

2. A BLOCK TRIDIAGONAL REPRESENTATION FOR k-ALMOST NORMAL MATRICES

In [2], theorem 8 implies that for 1-almost normal matrix A, there exists a normal matrix N and a constant h such that $A = N + hC$, where C is the same rank one matrix such that $AA^* - A^*A = AC - CA$. On the other hand, in [6], within theorem 4, it is seen that with the assumption that N is an n-by-n matrix such that the generalized Lanczos peocess brings N to a block tridiagonal form in which the orders of the diagonal blocks do not exceed the scalar ω_0 then, for every matrix R of rank $k \ll n$, the matrix $A = N + R$ can be brought to a block tridiagonal form where the orders of the diagonal blocks do not exceed $(2k+1)\omega_0$. So, choosing $k = 1$, these two theorems together yield that every 1-almost normal matrix can be unitarily reduced to a block tridiagonal form in which the diagonal blocks sizes are treble of the diagonal blocks sizes of its normal part, i.e., N.
As we know, the Lanczos algorithm is a technique for reduction of a Hermitian matrix to a tridiagonal form and its foundation is orthonormalization of the following power sequence

$$x, Ax, A^2x, A^3x, \dots, \tag{2.1}$$

in which x is an arbitrary non-zero vector in $\mathbb{C}^n$. Suppose that the produced space by this sequence is $\mathbb{C}^n$ and a linear operator $\mathbb{A}$ acting in $\mathbb{C}^n$ is associated with the matrix A. Then its matrix representation under the orthonormal basis obtained from orthogonalization of sequence (2.1), relying upon the Lanczos method, has a tridiagonal structure. For normal non-Hermitian matrix A, the following generalized power sequence is considered,

$$x, Ax, A^*x, A^2x, AA^*x, A^{*2}x, A^3x, \dots. \tag{2.2}$$

Without reduction of generality, again suppose that the space generated by the sequence (2.2) is the entire $\mathbb{C}^n$ and let $\mathbb{A}$ as its associated operator. Then $\mathbb{A}$ has a block tridiagonal structure under the basis obtained during the generalized Lanczos process on sequence (2.2). In this form, the diagonal blocks are square and their orders are the consecutive integers $1, 2, 3, ...$, respectively, (Theorem 1.1). For details, refer to [7] and [4].
In this section, we want to obtain the block tridiagonal structures for k-almost normal matrices. For this, we use an algorithm similar to generalized Lanczos algorithm for the following generalized power sequence.

$$x, Ax, (A^* - C)x, A^2x, A(A^* - C)x, (A^* - C)Ax, (A^* - C)^2x, A^3x, \dots.$$

But, by the definition 1.2 some of the above terms are repetitive and can be removed. Thus, we have the following sequence

$$x, Ax, (A^* - C)x, A^2x, A(A^* - C)x, (A^* - C)^2x, A^3x, \dots. \tag{2.3}$$

The i-th layer is comprised of vectors as follows,

$$A^\alpha (A^* - C)^\beta, \ \alpha + \beta = i.$$

Proposition 2.1. *A matrix $A \in M_n$ of rank k may be written as a sum of k rank 1 matrices. That is, $A = x_1 {y_1}^* + ... + x_k {y_k}^*$, for $x_i, y_i \in \mathbb{C}^n$, $i = 1, ..., k$.*

Proof. See [8], p.62. □

$rank C = k$. Thus, by this proposition, there exist vectors x_i, y_i, $1 \leq i \leq k$, such that $C = \sum_{i=1}^{k} x_i y_i^*$. So, we would be able to rearrange some terms of (2.3).
The second term of the first layer:

$$(A^* - C)x = A^*x - (x_1y_1^*x + x_2y_2^*x + \ldots + x_ky_k^*x)$$
$$= A^*x + \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_k x_k,\ \alpha_j = -y_j^* x,\ 1 \leq j \leq k.$$

The second term of the second layer:

$$A(A^* - C)x = AA^*x - ACx = AA^*x - (Ax_1y_1^*x + Ax_2y_2^*x + \ldots + Ax_ky_k^*x)$$
$$= AA^*x + \alpha_1 Ax_1 + \alpha_2 Ax_2 + \ldots + \alpha_k Ax_k,\ \alpha_j = -y_j^* x,\ 1 \leq j \leq k.$$

The third term of the second layer:

$$(A^* - C)^2 x = A^{*2}x - A^*Cx - CA^*x + C^2x$$
$$= A^{*2}x - A^*(x_1y_1^* + x_2y_2^* + \ldots + x_ky_k^*)x - (x_1y_1^* + x_2y_2^* + \ldots + x_ky_k^*)A^*x$$
$$+C(x_1y_1^* + x_2y_2^* + \ldots + x_ky_k^*)x = A^{*2}x + \alpha_1 A^*x_1 + \alpha_2 A^*x_2 + \ldots + \alpha_k A^*x_k$$
$$+\beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \alpha_1 Cx_1 + \alpha_2 Cx_2 + \ldots + \alpha_k Cx_k$$
$$= A^{*2}x + \alpha_1 A^*x_1 + \alpha_2 A^*x_2 + \ldots + \alpha_k A^*x_k + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k$$
$$+\alpha_1(x_1y_1^* + x_2y_2^* + \ldots + x_ky_k^*)x_1 + \alpha_2(x_1y_1^* + x_2y_2^* + \ldots + x_ky_k^*)x_2 + \ldots$$
$$+\alpha_k(x_1y_1^* + x_2y_2^* + \ldots + x_ky_k^*)x_k = A^{*2}x + \alpha_1 A^*x_1 + \alpha_2 A^*x_2 + \ldots + \alpha_k A^*x_k$$
$$+\beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \gamma_1 x_1 + \gamma_2 x_2 + \ldots + \gamma_k x_k$$
$$= A^{*2}x + \alpha_1 A^*x_1 + \alpha_2 A^*x_2 + \ldots + \alpha_k A^*x_k + \delta_1 x_1 + \delta_2 x_2 + \ldots + \delta_k x_k,$$

where

$$\beta_j = y_j^* A^* x,\ \gamma_j = \alpha_j y_j^* x_j,\ \delta_j = \beta_j + \gamma_j,\ 1 \leq j \leq k.$$

We now turn to the terms of the third layer; the second term of the layer:

$$A^2(A^* - C)x = A^2A^*x - A^2Cx = A^2A^*x + \alpha_1 A^2x_1 + \alpha_2 A^2x_2 + \ldots + \alpha_k A^2x_k.$$

Its third term:

$$A(A^* - C)^2x = AA^{*2}x + \alpha_1 AA^*x_1 + \alpha_2 AA^*x_2 + \ldots + \alpha_k AA^*x_k$$
$$+\delta_1 Ax_1 + \delta_2 Ax_2 + \ldots + \delta_k Ax_k.$$

Its fourth term:

$$(A^* - C)^3x = (A^* - C)(A^{*2} - A^*C - CA^* + C^2)x$$
$$= A^{*3}x - A^{*2}Cx - A^*CA^*x + A^*C^2x - CA^{*2}x - CA^*Cx + C^2A^*x - C^3x$$
$$= A^{*3}x + \alpha_1 A^{2*}x_1 + \alpha_2 A^{2*}x_2 + \ldots + \alpha_k A^{2*}x_k + \beta_1 A^*x_1 + \beta_2 A^*x_2 + \ldots + \beta_k A^*x_k$$
$$+\zeta_1 x_1 + \zeta_2 x_2 + \ldots + \zeta_k x_k,$$

where

$$\zeta_1, \zeta_2, \ldots, \zeta_k \in \mathbf{C},$$

and the same for the other layers. Therefore, the following sequence can be used instead of sequence (2.3).

$$x; Ax, A^*x, x_1, x_2, \ldots, x_k; A^2x, AA^*x, Ax_1, Ax_2, \ldots, Ax_k, A^{*2}x$$
$$, A^*x_1, A^*x_2, \ldots, A^*x_k; A^3x, \ldots. \tag{2.4}$$

In this sequence, the ith layer consists of vectors as follows.

$$A^\alpha A^{*\beta}x,\ A^\gamma A^{*\delta}x_t,\ \alpha+\beta=i,\ \gamma+\delta=i-1,\ t=1,\ldots,k.$$

We define the ith generalized Krylov subspace of sequence (2.4) as follows.

$$K_i = span\{v : v = A^\alpha A^{*\beta}x \text{ or } v = A^\gamma A^{*\delta}x_t,\ \alpha,\beta \le i,\ \gamma,\delta < i,\ t=1,\ldots,k\}.$$

Set $\dim K_i = l_i$ and define $w_i = l_i - l_{i-1}$ as the width of the i-th layer which represents the rate of increase in dimension of the span of the first i layers vectors, in the transition from $i-1$ to i. We set $w_0 = 1$. Now, assume that the sequence (2.4) is orthogonalized up to the ith layer. For this sequence, we have:

(1) For any $q \in K_i$,

$$Aq \in K_{i+1},\ A^*q \in K_{i+1},$$

(2) For every vector $q \in K_i \backslash K_{i-1}$, the following orthogonality is confirmed,

$$Aq \perp K_{i-2},\ A^*q \perp K_{i-2}.$$

The first property can readily be seen. For the second one, if $y \in K_{i-2}$, then

$$(Aq, y) = (q, A^*y)\ ,\ (A^*q, y) = (q, Ay).$$

Ay and A^*y are vectors in the $(i-1)$-th layer, and since the orthogonalization has advanced to the i-th layer, therefore, the desired product is obtained. Suppose that the span of the sequence (2.4) is the entire space $\mathbb{C}^n$ and the orthonormal basis $q_1, \ldots, q_n$ is obtained by the following operations in this sequence;

Let a random nonzero vector q_1. Then assume that the orthonormal and nonzero vectors $q_1, \ldots, q_t$ have already been found that form an orthonormal basis in the subspace K_m. Further, suppose that the last vectors $q_s, \ldots, q_t$ have been formed by using the m-th layer of the sequence (2.4). Now, we construct the vectors of $z = Aq$ for $q = q_s, \ldots, q_t$ and then orthogonalize these vectors with respect to the already accepted vectors q_i belonging to the $(m-1)$-th and m-th layers and also, the already obtained vectors in the current layer, i.e., layer $(m+1)$. After this orthogonalization, in each step, if z is nonzero, this vector is normalized and becomes the new vector q_j. We repeat this orthogonalization process for the vectors $\tilde{z} = A^*q$, where before doing orthogonalization of q, these vecrors have the form $A^*\tilde{q}$, for the vector $\tilde{q} \in \mathbb{C}^n$.

Now, suppose that $P = \{q_1, \ldots, q_n\}$ provides a basis for space $\mathbb{C}^n$ and define the linear operator $\mathbb{A} : \mathbb{C}^n \to \mathbb{C}^n$ where $\mathbb{A}(x) = Ax$. It is clear that the E basis representation of $\mathbb{A}$ is A for the standard basis $E = \{e_1, \ldots, e_n\}$, (i.e., $[\mathbb{A}]_E = A$). On the other hand, by letting H as its P basis representation, (i.e., $[\mathbb{A}]_P = H$), we have

$$[\mathbb{A}]_E = Q[\mathbb{A}]_P Q^{-1},$$

in which the unitary matrix $Q = \begin{pmatrix} q_1 & q_2 & \ldots & q_n \end{pmatrix}$ is the E-P change of basis matrix. Thus

$$A = QHQ^* \Rightarrow Q^*AQ = H.$$

Hence

$$\begin{pmatrix} q_1^* \\ \vdots \\ q_n^* \end{pmatrix} A \begin{pmatrix} q_1 & \cdots & q_n \end{pmatrix} = H.$$

It shows that for each $1 \leq i, j \leq n$,

$$q_i^* A q_j = h_{ij}. \tag{2.5}$$

In other words, $(Aq_j, q_i) = h_{ij}$. Now, if q_j belongs to the s-th layer of the orthogonalized sequence (2.4), then

$$h_{ij} \neq 0, \ \sum_{l=0}^{s-2} w_l < i \leq \sum_{l=0}^{s+1} w_l.$$

Furthermore, (2.5) yields that $q_j^* A^* q_i = \overline{h}_{ij}$, or equivalently, $(A^* q_i, q_j) = \overline{h}_{ij}$. Similar to what was said, if q_i belongs to the s-th layer of the orthogonalized sequence (2.4), then

$$h_{ij} \neq 0, \ \sum_{l=0}^{s-2} w_l < j \leq \sum_{l=0}^{s+1} w_l.$$

It follows from the above explanations that the number of nonzero entries in the i-th column (row)of H does not exceed $w_{s-1} + w_s + w_{s+1}$, in which s is the index of the layer to which q_i belongs. Thus, we have $AQ = QH$, i.e., A and H are unitarily similar and H has a tridiagonal form (1.1), where H_{ii} as the ith diagonal block has a maximum order of $ik + i + 1$. Thus, the diagonal blocks have maximum orders of

$$1, k + 2, 2k + 3, 3k + 4, \ldots, ik + i + 1, \ldots$$

respectively. In the special case $k = 1$, the maximum orders of diagonal blocks are $1, 3, 5, \ldots,$ respectively. In fact, the following theorem is proved.

Theorem 2.1. *Every k-almost normal matrix in $\mathbb{C}^{n \times n}$ is unitarily similar to a block tridiagonal matrix (1.1), where the diagonal blocks have maximum orders as follows,*

$$1, k + 2, 2k + 3, 3k + 4, \ldots, ik + i + 1, \ldots$$

respectively. In fact, assuming that the basis $q_1, \ldots, q_n$ is obtained for $\mathbb{C}^n$, by orthonormalization of the sequence (2.4) then, $Q^*AQ = H$.

As a special case, assume that the following equality is true for the k-almost normal matrix A,

$$aA^2 + bA^{*2} + cAA^* = dI \tag{2.6}$$

in which b and at least one of the scalars a, c or d is non-zero. Thus, for every arbitrary non-zero vector $x \in \mathbf{C}^n$, $A^{*2}x$ is a linear combination of A^2x, AA^*x, and x. This dependency makes some terms to be repeated, in the sequence (2.4). Thus, the $(k+3)-$th term in the second layer and the last k terms in the third layer can be removed from the sequence and then the widths of layers will be as follows:

$$1, k + 2, 2k + 2, 2k + 2, \ldots.$$

This means stabilizing the width of layers from the second layer on. Thus, in the condensed form of A, i.e., the tridiagonal matrix H, the diagonal blocks are of orders $1, k+2, 2k+2, 2k+$

$2, \ldots$, respectively and H is a band matrix with a constant band width. Consider k-almost normal matrix A for which we know that there exists a matrix $C \in M_n(\mathbb{C})$ of rank k such that $AA^* - A^*A = AC - CA$. Below we see the details of how to build its condensed form, i.e., the matrix H and the orthonormal sequence $q_1, q_2, \ldots, q_n$.

(1) Find vectors $x_1, x_2, \ldots, x_k, y_1, y_2, \ldots, y_k \in \mathbb{C}^n$, such that $C = \sum_{i=1}^{k} x_i y_i^*$.
(2) Find an appropriate initial normal vector

$$q_1 = \frac{1}{\|x\|_2} x.$$

(3) The vector Aq_1 is orthogonalized to q_1; that is,

$$w_2 = Aq_1 - h_{11} q_1,$$

where

$$h_{11} = (Aq_1, q_1). \tag{2.7}$$

Now, we set

$$h_{21} = \| w_2 \|_2, \qquad q_2 = w_2 / h_{21}.$$

(We assume that $h_{21} \neq 0$)
(4) The vector $A^* q_1$ is orthogonalized to q_1 and q_2; thus,

$$w_3 = A^* q_1 - \overline{h_{11}} q_1 - \overline{h_{12}} q_2.$$

Here, the coefficient

$$h_{11} = \overline{(A^* q_1, q_1)} = \overline{(q_1, Aq_1)} = (Aq_1, q_1)$$

is already known (see (2.7)), while

$$h_{12} = \overline{(A^* q_1, q_2)}.$$

After calculating w_3, we set

$$h_{13} = \| w_3 \|_2, \qquad q_3 = w_3 / h_{13}.$$

(5) For $i = 1, 2, 3, \ldots, k$, the vector x_i is orthogonalized to $q_1, q_2, q_3, \ldots, q_r$, that is,

$$w_{3+i} = x_i - (x_i, q_1) q_1 - (x_i, q_2) q_2 - \ldots - (x_i, q_r) q_r,$$

where q_r is the last obtained vector of the orthnormal basis. After finding w_{3+i}, if $\| w_{3+i} \|_2 \neq 0$, we set $q_{r+1} = w_{3+i} / \| w_{3+i} \|_2$.

Now, we have found the vectors $q_1, q_2, \ldots, q_t$, and an orthonormal basis in the generalized subspace Krylov

$$K_1 = \operatorname{span}\{q_1, q_2, \ldots, q_t\},$$

where $t \leq k + 3$.
(6) In this step, the vectors of the latest completed layer are left-multiplied by A and then, these new vectors are orthogonalized to the vectors of the two layers before and also, to the already obtained vectors in the current layer. If this obtained vector has a nonzero norm, set its normalization as the the newest acquired member of the orthogonal basis.

(7) Here, the vectors of the latest completed layer, except for the vectors, which before doing orthogonalization, have the form A^*q, for a vector $q \in \mathbf{C}^n$, are left-multiplied by A^* and then, similar to what was performed in the previous stage, these new vectors are orthogonalized to the vectors of the two layers before and also, to the already obtained vectors in the current layer. If the norm of the resulting vector is nonzero, we shall replace its normalization as the next vector of the orthonormal basis.
The coefficients h_{ij} which were generated during these orthogonalizations, are used as the entries of matrix H.

The subsequent steps are similar.

3. A BLOCK TRIDIAGONAL REPRESENTATION FOR k-ALMOST CONJUGATE NORMAL MATRICES

Inspired by the above arguments, in this section the aim is to have a block tridiagonal representation for k-almost conjugate normal matrices.

Let $A \in M_n$ as a k-almost conjugate normal matrix. I.e., there exists the matrix $C \in M_n(\mathbb{C})$ of rank k such that $A(A^* - C) = \overline{(A^* - C)A}$. Define double orders matrices $\mathcal{A}$ and $\mathcal{C}$ as following,

$$\mathcal{A} = \begin{pmatrix} 0 & \overline{A} \\ A & 0 \end{pmatrix}, \, \mathcal{C} = \begin{pmatrix} 0 & C \\ \overline{C} & 0 \end{pmatrix}.$$

we have

$$\mathcal{A}(\mathcal{A}^* - \mathcal{C}) = (\overline{A}A^T - \overline{AC}) \oplus (AA^* - AC),$$

and

$$(\mathcal{A}^* - \mathcal{C})\mathcal{A} = (A^*A - CA) \oplus (A^T\overline{A} - \overline{CA}).$$

Now, if $\mathcal{A}(\mathcal{A}^* - \mathcal{C}) = (\mathcal{A}^* - \mathcal{C})\mathcal{A}$, then $\overline{A}A^T - \overline{AC} = A^*A - CA$ and vice versa. Thus, we have the following assertion.

Proposition 3.1. *Consider $\mathcal{A}$ and A as above. Then, A is the k-almost conjugate normal matrix if and only if $\mathcal{A}$ is a $2k$-almost normal matrix.*

According to the proposition 2.1, for matrices C and $\mathcal{C}$ in the above discussions, there exist vectors $x_i, y_i \in \mathbb{C}^n$, $i = 1, ..., k$, such that $C = \sum_{i=1}^k x_i y_i^*$ and by it

$$\mathcal{C} = \sum_{i=1}^k \begin{pmatrix} 0 \\ \overline{x}_i \end{pmatrix} \begin{pmatrix} \overline{y}_i \\ 0 \end{pmatrix}^* + \sum_{i=1}^k \begin{pmatrix} x_i \\ 0 \end{pmatrix} \begin{pmatrix} o \\ y_i \end{pmatrix}^*.$$

Now, define $\mathcal{X}_i \in \mathbb{C}^{2n}$ as following,

$$\mathcal{X}_i = \begin{cases} \begin{pmatrix} o \\ \overline{x}_i \end{pmatrix}, & \text{i=1,...,k;} \\ \begin{pmatrix} x_{i-k} \\ 0 \end{pmatrix}, & \text{i=k+1,...,2k.} \end{cases}$$

Choose an arbitrary nonzero initial vector $x \in \mathbb{C}^n$ and define $\mathcal{X} = \begin{pmatrix} x \\ \overline{x} \end{pmatrix}$. According to the previous section, the following sequence can be written for the k-almost normal matrix $\mathcal{A}$,

$$\mathcal{X}, \mathcal{A}\mathcal{X}, \mathcal{A}^*\mathcal{X}, \mathcal{X}_1, ..., \mathcal{X}_{2k}, \mathcal{A}^2\mathcal{X}, \mathcal{A}\mathcal{A}^*\mathcal{X}, \mathcal{A}^{*2}\mathcal{X}, \mathcal{A}\mathcal{X}_1, ..., \mathcal{A}\mathcal{X}_{2k}, \mathcal{A}^*\mathcal{X}_1, ..., \mathcal{A}^*\mathcal{X}_{2k}, ..., \quad (3.1)$$

in which the i-th layer contains vectors as below,

$$\mathcal{A}^\alpha \mathcal{A}^{*\beta}\mathcal{X}, \ \mathcal{A}^\gamma \mathcal{A}^{*\delta}\mathcal{X}_t, \ \alpha + \beta = i, \ \gamma + \delta = i - 1, \ t = 1, ..., 2k$$

and the i-th generalized Krylov subspace is defined as

$$\mathcal{K}_i = span\{V | \ V = \mathcal{A}^\alpha \mathcal{A}^{*\beta}\mathcal{X} \ or \ V = \mathcal{A}^\gamma \mathcal{A}^{*\delta}\mathcal{X}_t, \ \alpha, \beta \le i, \ \gamma, \delta < i, \ t = 1, ..., 2k\}.$$

Let $dim\mathcal{K}_i = \mathcal{L}_i$ and define $\mathcal{W}_i = \mathcal{L}_i - \mathcal{L}_{i-1}$ as the width of the i-th layer. Now, the upper halves of the vectors in the sequence (3.1) form the following sequence,

$$x, \bar{A}\bar{x}, A^*\bar{x}, x_1, ..., x_k, \bar{A}Ax, \bar{A}A^Tx, \bar{A}\bar{x}_1, ..., \bar{A}\bar{x}_k, A^*A^Tx,$$

$$A^*\bar{x}_1, ..., A^*\bar{x}_k, \bar{A}A\bar{A}\bar{x}, \bar{A}AA^*\bar{x}, \bar{A}Ax_1, ..., \bar{A}Ax_k, \bar{A}A^TA^*\bar{x},$$
$$\bar{A}A^Tx_1, ..., \bar{A}A^Tx_k, A^*A^TA^*\bar{x}, A^*A^Tx_1, ..., A^*A^Tx_k, ... \quad (3.2)$$

Define the layers in this sequence as the projections of the corresponding layers in the sequence (3.1) onto $\mathbb{C}^n$. Thus, the vectors in the first layer are $\overline{A}\overline{x}, A^*\overline{x}, x_1, ..., x_k$ and the vectors $\overline{A}Ax, \overline{A}A^Tx, \overline{A}\overline{x}_1, ..., \overline{A}\overline{x}_k, A^*A^Tx, A^*\overline{x}_1, .., A^*\overline{x}_k$ form the second layer, etc. Define K_i as the span of the vectors in the first i layers of the sequence (3.2) and $\overline{K}_i = \{x \in \mathbb{C}^n | \overline{x} \in K_i\}$ and assume that orthogonalization is done up to the i-th layer. Then, the following two traits can be seen,

(1) For any $q \in K_i$,

$$Aq \in \overline{K}_{i+1}, \ A^Tq \in \overline{K}_{i+1},$$

or equivalently

$$\bar{A}\bar{q} \in K_{i+1}, \ A^*\bar{q} \in K_{i+1}.$$

(2) For every vector $q \in K_i \backslash K_{i-1}$, we have

$$Aq \perp \overline{K}_{i-2}, \ A^Tq \perp \overline{K}_{i-2}.$$

The first property can be verified straightforwardly. However, in each layer, there may be some terms that seemingly do not follow this principle. But in practice, we see that these terms can be written as a linear combination of the other terms in that layer. For the second one, let $p \in \overline{K}_{i-2}$. Then,

$$(Aq, p) = (q, A^*p) = 0, \ and \ (A^Tq, p) = (q, \bar{A}p) = 0,$$

since q is orthogonal to K_{i-1} and $\bar{A}p, A^*p \in K_{i-1}$. As defined in the sequence (3.1), for the sequence (3.2), here also define $dimK_i = l_i$ and $w_i = l_i - l_{i-1}$. Looking at the above explanations result would be that $w_i \le \mathcal{W}_i, i = 1, 2, ...$ In fact, considering that $rank\mathcal{C} = 2k$, then by the previous section, it is clear that $dim\mathcal{K}_i = 2ik + i + 1$. But, we can see that in projecting sequence(3.1) onto $\mathbb{C}^n$, when we consider the upper halves of the vectors in the sequence (3.1), the upper halves of the vectors $\mathcal{X}_1, ..., \mathcal{X}_k$, which all belong to the first layer, are projected onto zero. Thus $w_i \leq ik + i + 1$ and therefore, $w_i \leq \mathcal{W}_i$. Suppose that the span of the sequence (3.2) is the entire space $\mathbb{C}^n$. During the operations similar to those mentioned in the previous chapter, the orthonormal basis $q_1, ..., q_n$ is constructed from the sequence (3.2). With the difference that, here after constructing the vectors of $z = Aq$ for $q = q_s, ..., q_t$, we orthogonalize these vectors to already accepted vectors $\overline{q}_i$ which q_i belong to the $(m-1)$-th, m-th or $(m+1)$-th layers. Then, we repeat this orthogonalization process for vectors $\tilde{z} = A^T q$. Finally, the same bounds are obtained for the number of nonzero entries in the i-th column and i-th row of H and thereby, $AQ = \overline{Q}H$, i.e., A and H are unitarily congruent. Thus, we have the following theorem.

Theorem 3.1. *An k-almost conjugate normal matrix $A \in M_n(\mathbb{C})$ can be brought by an unitary congruence to a block tridiagonal form 1.1 where the diagonal blocks are square and their maximum orders are given by the integers $1, k+2, 2k+3, ..., ik+i+1, ...,$ respectively. In fact, if the orthogonalization of the sequence (3.2) creates an orthonormal basis $q_1, ..., q_n$ for $\mathbb{C}^n$, then by defining $Q = \begin{pmatrix} q_1 & q_2 & \dots & q_n \end{pmatrix}$, it holds that $Q^T AQ = H$.*

Now, we discuss about one particular mode. Suppose that the k-almost conjugate normal matrix A satisfies the equation

$$a\bar{A}A + b\bar{A}A^T + cA^*A^T = dI,$$

in which b and at least one of the scalars a, c, d is no-zero. But this leads to a formula as (2.6) for the k-almost normal matrix $\mathcal{A}$ and thereby, the widths of the layers of the sequence (3.1) will be as follows:

$$1, k+2, 2k+2, 2k+2, \dots .$$

As previously mentioned, this means that in the condensed form of $\mathcal{A}$ the diagonal blocks are of orders $1, k+2, 2k+2, 2k+2, \dots,$ respectively. So similarly, the widths of the layers of sequence (3.2) are given utmost by integers $1, k+2, 2k+2, 2k+2, \dots$. Therefore, the orders of the diagonal blocks of H remain constant at utmost $2k+2$.

Consider k-almost conjugate normal matrix A for which there exists a matrix $C \in M_n(\mathbb{C})$ of rank k such that $AA^* - A^T\overline{A} = AC - \overline{CA}$. The details of the calculations to find the matrix H and orthonormal basis $q_1, ..., q_n$ are as follows.

(1) Find vectors $x_i, y_i \in \mathbb{C}^n$, $i = 1, ..., k$, such that $C = \sum_{i=1}^{k} x_i y_i{}^*$.
(2) Choose a random nonzero starting unit vector q_1.
(3) At this step, the vector Aq_1 is orthogonalized to $\overline{q}_1$; that is,

$$w_2 = Aq_1 - h_{11}\overline{q}_1, \text{ in which } h_{11} = (Aq_1, \overline{q}_1).$$

Then, let

$$h_{21} = \| w_2 \|_2, \; q_2 = \overline{w}_2 / h_{21}$$

(4) In the fifth stage, the vector $A^T q_1$ is orthogonalized to $\overline{q_1}$ and $\overline{q_2}$; Therefore,

$$w_3 = A^T q_1 - h_{11}\overline{q}_1 - h_{12}\overline{q}_2$$

in which $h_{11} = (A^T q_1, \overline{q}_1) = (q_1, A^*\overline{q}_1) = (Aq_1, \overline{q}_1)$ and $h_{12} = (A^T q_1, \overline{q}_2)$. Set $h_{13} = \| w_3 \|_2$ and define $q_3 = \overline{w}_3 / h_{13}$.
(5) For $i = 1, ..., k$, the vector x_i is orthogonalized to the vectors $\overline{q}_1, ..., \overline{q}_r$; Thus,

$$w_{3+i} = x_i - \sum_{t=1}^{r}(x_i, \overline{q}_t)\overline{q}_t,$$

in which q_r is the latest earned vector of the orthonormal basis. After calculating w_{3+i}, if this value is nonzero, set $\overline{q_{r+1}} = w_{3+i}/\| w_{3+i} \|_2$. The outcome of this process so far is the vectors $q_1, ..., q_t$, $t \leq k+3$.

(6) Left-multiply the vectors of the latest completed layer by A and orthogonalize the new resulting vectors to the vectors of the two layers before and also to the already obtained vectors in the current layer. If this obtained vector has a nonzero norm, set its normalization as the the newest acquired member of the orthogonal basis.

(7) The vectors of the latest completed layer, except for the vectors, which before doing orthogonalization, have the form $A^T q$, for a vector $q \in \mathbf{C}^n$, are left-multiplied by A^T and then we do similarly to what was performed in the previous stage.

The coefficients h_{ij} which were generated during these orthogonalizations are used as the entries of the matrix H.

The next steps are the same as above.

4. ALGORITHM AND NUMERICAL COMPUTATIONS

We were eager to give a numerical example of a small-scale 2-almost normal matrix to test the theorem and algorithm. But this example had to be provided for a matrix of at least 12 in size to be a good benchmark. Since it was not possible to do this manually, we had to create the following pseudocode-algorithm and then we coded it with MATLAB software in order to provide examples with it. In the following examples, we have tested the zero elements of the H matrix outside the algorithm and satisfied that they are zero. To do this, we use the following formula:

$$h_{ij} = q_i{}^{\star} A q_j, \ \forall i, j.$$

The elements located at the end of the H-matrix diagonal are zero bacause the algorithm is designed in such a way that creats the arrays of the H-matrix for the previous layers and if we go one layer further these arrays will also be made.

Algorithm 1. Following, an algorithm for reducing an admissible k-almost (conjugate) normal matrix A to its condensed form H is presented.

Input: Matrix A and vectors $x, x_1, x_2, x_3, \ldots, x_k$.

$q_1 = \frac{x}{\|x\|_2}$

$z = Aq_1$

$h_{1,1} = q_1^* z, \ (h_{1,1} = q_1^T z)$

$z = z - h_{1,1}q_1, \ (z = z - h_{1,1}\overline{q}_1)$

$h_{2,1} = \|z\|_2$

$q_2 = z/h_{2,1}, \ (q_2 = \bar{z}/h_{2,1})$

$z = A^* q_1, \ (z = A^T q_1)$

$z = z - \sum_{i=1}^{2} q_i^* z q_i, \ (z = z - \sum_{i=1}^{2} q_i^T z \overline{q}_i)$

$h_{1,2} = \overline{q_2^* A^* q_1}, (h_{1,2} = \overline{q_2^T A^T q_1})$

$h_{1,3} = \|z\|_2$

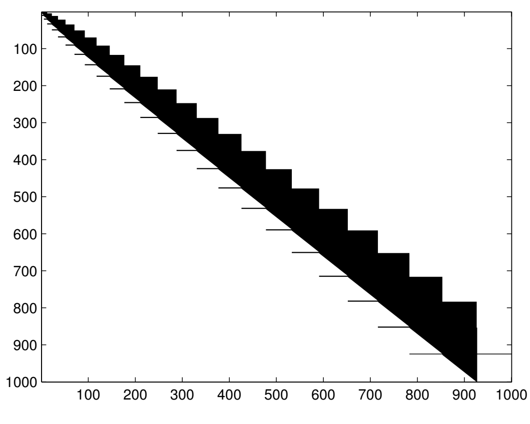
$q_3 = z/\|z\|_2, \ (q_3 = \bar{z}/\|z\|_2)$

$s = 3$
$for\ i = 1\ to\ k$
  $z = x_i - \sum_{j=1}^{s} q_j^* x_i q_j,\ (z = x_i - \sum_{j=1}^{s} q_j^T x_i \overline{q}_j)$
  $if \|z\|_2 > 0$
    $s = s + 1$
    $q_s = z/\|z\|_2,\ (q_s = \bar{z}/\|z\|_2)$
  end
$end\ for$
$c_1 = 1$
$c_2 = s$
$s = s + 1$
$w = 3$
$while\ w \leq n$
  $j = c_{w-2} + 1$
  $p = j$
  $if\ w < 4$
    $d = 1$
  $else$
    $d = c_{w-3} + 1$
  end
  $for\ j = c_{w-2} + 1 : c_{w-1}$
    $z = Aq_j$
    $z = z - \sum_{l=d}^{s-1} q_l^* z q_l,\ (z = z - \sum_{l=d}^{s-1} q_l^T z \overline{q}_l)$
    $if\ \|z\|_2 > 0$
      $for\ l = d\ to\ s - 1$
        $h_{lp} = q_l^* A q_j,\ (h_{lp} = q_l^T A \overline{q}_j)$
      $end\ for$
      $q_s = z/\|z\|_2,\ (q_s = \bar{z}/\|z\|_2)$
      $h_{sj} = \|z\|_2$
      $p = p + 1$
    $end\ if$
    $if\ s < n$
      $s = s + 1$
    $else$
      $break$
    end
  $end\ for$
  $a = s$
  $if\ w < 4$
    $f = 3$
  $else$
    $f = b$
  $end\ if$
  $for j = f\ to\ c_{w-1}$
    $z = A^* q_j,\ (z = A^T q_j)$

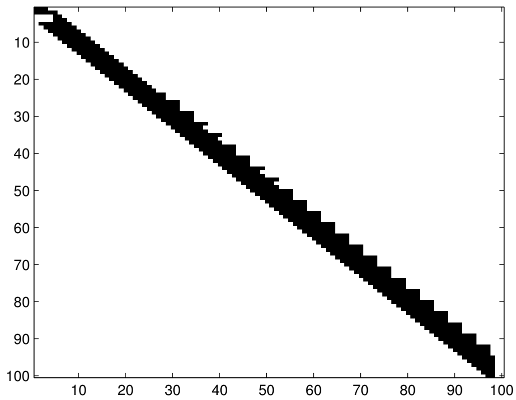
$z = z - \sum_{l=d}^{s-1} q_l^* z q_l, \ (z = z - \sum_{l=d}^{s-1} q_l^T z \overline{q}_l)$
$if \ \|z\|_2 > 0$
$\quad q_s = z/\|z\|_2, \ (q_s = \bar{z}/\|z\|_2)$
$\quad for \ l = d \ to \ s-1$
$\quad\quad h_{jl} = \overline{q_l^* A^* q_j}, \ (h_{jl} = \overline{q_l^T A^T q_j})$
$\quad end \ for$
$\quad h_{js} = \|z\|_2$
$\quad if \ s < n$
$\quad\quad s = s+1$
$\quad else$
$\quad\quad break$
$\quad end$
$end \ for$
$c_w = s-1$
$b = a$
$if \ s < n+1$
$\quad w = w+1$
$else$
$\quad break$
end
$end \ while$

Output: Matrices H and $Q = (q_1, q_2, q_3, \ldots, q_n)$.

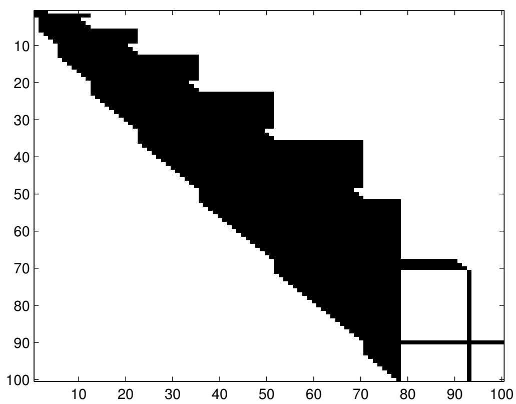
Example 4.1. *For the* 2*- almost (conjugate) normal matrix* A*, mentioned algorithm produces a basis* $q_1, q_2, q_3, \ldots, q_n$*, under which the structure of the block tridiagonal representation matrix* H *is generally as following in which the black parts show the no necessarily zero entries and white parts indicate to zero entries (for example, choose* $n = 1000$*):*



Example 4.2. *Suppose that $A = H + xy^*$ in which H is an $n \times n$ Hermitian matrix and $x, y \in \mathbb{C}^n$. Let $C = yx^* - xy^*$. Then, it is seen that A is 2-almost normal. As an special case, we consider the matrix A so that also satisfies in the equation (2.6). In this case, the diagonal blocks sizes of A do not exceed 6. by letting $n = 100$, the above algorithm for A, creates a condensed form as following.*



Example 4.3. *Suppose that $B = T + xy^T$ in which T is a symmetric matrix in $M_n(\mathbb{C})$ and $x, y \in \mathbb{C}^n$. By choosing $C = yx^T - xy^T$, B is 2-almost conjugate normal and as a special case, by letting $n = 100$, the above algorithm for A, creates the following condensed form.*



P.s: The above graphs are plotted by MATlAB and for random entries.

5. CONCLUSION

We say that an $n \times n$ complex matrix A is k-almost normal if there exists a matrix C with rank k such that $A^* - C$ commutes with A. There are many issues that are reliant on this commutativity condition. For example, there are many results concerning the spectrum of a perturbed Hermitian matrix [9] or general normal matrices [10, 11] or examples of QR algorithms for the computation of the eigenvalues of low rank perturbations of symmetric or Hermitian matrices [12, 13]. As another example, in the operator theory conditions upon this commutativity condition are widely used in the study of structural properties of hypernormal operators [14]. But, what makes it so interesting for us is that we are working on solving methods for linear systems whose coefficient matrices are k-almost normal or k-almost conjugate normal matrices. In this paper, similar to what are done about normal and conjugate normal matrices in [4] and [1], we obtain the block tridiagonal structures for k-almost normal and k-almost conjugate normal matrices via an unitary similarity and an unitary congruence, respectively.

REFERENCES

[1] M.GHASEMI KAMALVAND, KH. D. IKRAMOV, *A method of congruent type for linear systems with conjugate-normal coefficient matrices.* Computational mathematic and physics, Vol. 49, No. 2, 203-216, 2009.

[2] R. BEVILACQUA, G. M. DEL CORSO, *A condensed representation of almost normal matrices.* Linear algebra and its applications, 438, 4408-4425, 2013.

[3] R. BEVILACQUA, G. M. DEL CORSO, L. GEMIGNANI, *Block tridiagonal reduction of perturbed normal and rank structured matrices.* Linear algebra and its applications,1-13, 2013.

[4] L. ELSNER, KH. D. IKRAMOV, *On a condensed form for normal matrices under finite sequences of elementary unitary similarities.* Linear algebra and its applications, 254, 79-98, 1997.

[5] A. BUNSE-GERSTNER, R. STOVER, *On a conjugate-type mathod for solving complex symmetric linear systems.* Linear algebra Appl.287, 105-123, 1999.

[6] M.GHASEMI KAMALVAND, KH. D. IKRAMOV, *Low-rank perturbations of normal and conjugate-normal matrices and their condensed forms under unitary similarities and congruences.* Computational mathematic and physics, Vol. 33, No. 3, 109-116, 2009.

[7] M. DANA, A. G. ZYKOV, KH.D. IKRAMOV, *A minimal residual method for a special class of the linear systems with normal coefficient matrices.* - Comput. Math. Math. Phys., 45, 1854–1863, 2005.

[8] R.A. HORN AND C.R. JOHNSON, *Matrix Analysis.* Cambridge University Press, Cambridge, 1985.

[9] W. KAHAN, *Spectra of nearly Hermitian matrices.* Proc. Amer. Math. Soc. 48, 11-17, 1975.

[10] J. SUN, *On the variation of the spectrum of a normal matrix.* Linear algebra Appl, 246, 215-223, 1996.

[11] I. IPSEN, *Departure from normality and eigenvalue perturbation bounds.* Technical Report TR03-28, NC State University, 2003.

[12] R. VANDEBRIL, G.M. DEL CORSO, *An implicit multishift qr-algorithm for Hermitian plus low rank matrices.* SIAM J. Sci. Comput. 16, 2190-2212, 2010.

[13] L.G.Y. EIDELMAN, I.C. GOHBERG, *Efficient eigenvalue computation for quasiseparable Hermitian matrices under low rank perturbation.* Numer. orithms 47, 253-273, 2008.

[14] M. PUTINAR, *Linear analysis of quadrature domains.* III. J. Math. Anal. Appl, 239(1):101-117, 1999.