UNIVERSITÀ DI PISA
DIPARTIMENTO DI INFORMATICA

TECHNICAL REPORT: TR-12-11

A Condensed Representation of Almost Normal Matrices

Roberto Bevilacqua Gianna M. Del Corso

20 September 2012

ADDRESS: Largo B. Pontecorvo 3, 56127 Pisa, Italy. TEL: +39 050 2212700 FAX: +39 050 2212726

A CONDENSED REPRESENTATION OF ALMOST NORMAL MATRICES*

ROBERTO BEVILACQUA † AND GIANNA M. DEL CORSO‡

Abstract. In this paper we study the structure of almost normal matrices, that is the matrices for which there exists a rank-one matrix C such that $A^HA - AA^H = CA - AC$. Necessary and sufficient conditions for a matrix to belong to the class are given and a canonical representation as a block tridiagonal matrix is shown. The approach is constructive and in the paper it is explained how, starting from a 1×1 or 2×2 matrix we can generate almost normal matrices. Moreover, given an $n \times n$ almost normal matrix we can compute the block tridiagonal representation with a finite procedure.

Perturbation of normal matrices, canonical representation 65F99;15A21;15B99

1. Introduction. Normal matrices play an important theoretical role in the field of numerical linear algebra. A square complex matrix is called normal if

$$A^HA - AA^H = 0,$$

where A^H is the conjugate transpose of A. Over the years many equivalent conditions have been found [13, 8], and it has been discovered that the class can be partitioned in accordance with a parameter s, where s is the minimal degree of a polynomial such that $A^H = p_s(A)$, $s \leq n-1$. A milestone in the study of normal(s) matrices are the results related to the Faber-Manteuffel theorem [11, 10, 9] where it is proved that only normal(s) matrices can have short (s-term) recurrence when Arnoldi process is applied. Equivalently, the theorem proves that the Hessenberg matrix we obtain applying Arnoldi process from any possible starting vector to a normal(s) matrix is an $(s+2)$-band Hessenberg matrix. Moreover the structure is preserved under QR steps, making it possible to use implicit methods, and reducing the cost per step of QR iterations. Faber-Manteuffel theorem is not able to discover however the possible hidden structure in the matrix when the degree s is for example $n-1$ as it happens for unitary matrices. It is well known however that unitary Hessenberg matrices have a rank-one structure in the upper triangular part [5]. This structure is reveled by considering that any normal matrix is such that $A^H = p_\ell(A)/q_m(A)$ with p_ℓ and q_m polynomials with degree ℓ and m respectively. For unitary matrices $q_m(A) = A$, with $m = 1$ and $p_\ell(A) = I$, with $\ell = 0$.

Despite it is simple [6] to prove that the Hessenberg form of any normal matrix such that A^H is a rational function in A is such that all the submatrices taken out above the $\max\{|\ell - m|, 1\}$ diagonal[1] have rank at most m, it is uncertain if the contrary is true. In [2] the authors try to extend the Faber-Manteuffel theorem to a wider class of matrices, namely the matrices A such that $A^H = \frac{p_\ell(A)}{q_m(A)} + C_k$, being C_k a rank k matrix. They determined sufficient conditions in order that a conjugate gradient method can be implemented using a short multiple recursion. Necessary conditions, although claimed, are not proved in the general case [1]. Given the hardness of studying this class of matrices, in this paper we consider the simpler case where instead of rational functions we have polynomials and a rank-one correction. In the literature, there are many results concerning the spectrum of a perturbed Hermitian matrix [16] or general normal matrices [18, 15], or examples of implicit QR algorithms for the computation of the eigenvalues of low rank perturbations of symmetric or Hermitian matrices [19, 22]. The notion or quasi-normal, or almost normal matrices have been introduced earlier in the literature [21, 17], in this paper however we give a different definition of the

* This research was partially supported by the Italian Ministry for Education, University and Scientific Research under program PRIN 2008: 20083KLJEZ

†bevilacq@di.unipi.it

‡delcorso@di.unipi.it

[1] When $\ell = 0$ and $m = 1$ as in the case A is unitary, the rank-one structure extends up to the main diagonal.

class and we are mainly interested in the description of a condensed representation of the matrices of the class, which eventually can lead to algorithmic advantages for the computation of eigenvalues or for the solution of linear systems. In particular, we say that an $n \times n$ complex matrix A is *almost normal* if there exists a matrix C with rank at most one such that $A^H - C$ commutes with A. In the nonderogatory case, if A is almost normal then there exists a polynomial $p_s(\cdot)$ such that $A^H = p_s(A) + C$ for a polynomial with degree s, $s \leq n-1$. In the derogatory case, however this is no more true because, matrices can commute without being polynomials in each other. In this paper we do not consider this polynomial formulation and the degree of the polynomial will nor play a role, because we take the union over all the possible s. We investigate the structure of almost normal matrices with a constructive approach. Our approach has the same flavor of that carried out by Ikramov and Elsner in [14] and [7] for normal matrices and by Ghasemi Kamalvand and Ikramov [12] for low rank perturbations of normal matrices. The condensed form introduced in our paper can be obtained by a finite algoritheoremic procedure which is described in the theorems proved in the paper.

Normal matrices are fully characterized by the Schur canonical form that is diagonal. However, the Schur form seems not to characterize almost normal matrices. In fact, if A is a non-derogatory almost normal matrix, that is $A^H = p_s(A) + C$, and we consider the Schur canonical form of A, we have

$$T^H = p_s(T) + \tilde{C}, \quad \text{where } \tilde{C} \text{ is a suitable rank-one matrix.}$$

Since T is upper triangular, we obtain that all the minors taken out from the strictly upper triangular part of T have rank at most one. However this is not a sufficient condition for A to be almost normal.

The paper is organized as follows. In Section 2 we introduce the class of almost normal matrices and we give necessary conditions for a matrix to be in the class. In particular we show that any almost normal matrix can unitarily reduced to a matrix with a tridiagonal block structure, where the diagonal blocks are related to each other and the entries of the off diagonal blocks depend on those of the successive more external block. In Section 3 we show that the necessary conditions turn out to be sufficient as well. A recursive argument allows to border any almost normal matrix with suitable entries to obtain another larger matrix of the class. In Section 4 it is shown that any almost normal matrix is a rank-one perturbation of a normal matrix and Section 5 contains some concluding remarks.

2. Structure of the class: Necessary conditions. Let A be such that $A^H = p_s(A) + C$, where $p_s(\cdot)$ is a polynomial with degree s and C is a rank-one matrix. If A is nonderogatory, this condition is equivalent to $A^H - C$ commuting with A, that is

$$(A^H - C)A = A(A^H - C).$$

In this paper we study the class of of almost normal matrices that we define as follows.

DEFINITION 2.1. *An $n \times n$ complex matrix A is almost normal if there exists a matrix C with rank at most one such that*

$$A^H A - AA^H = CA - AC. \tag{2.1}$$

Note that, since C has rank at most one, the matrix $CA - AC$ has rank two or zero (in the latter case A is normal). Moreover, because matrix $A^H A - AA^H$ appearing on the left-hand side of (2.1) is Hermitian, it can be diagonalized with unitary transformation Q, that is $Q^H(A^H A - AA^H)Q = D$, where D is a real diagonal matrix. To keep the notation simple, we will still call A the matrix obtained applying unitary transformations which preserve the structural properties.

The matrix A^HA-AA^H has zero trace, and rank zero or two, hence there exists a non-negative real number $\alpha \geq 0$ such that

$$A^HA-AA^H = \left[\begin{array}{c|cc} O & \multicolumn{2}{c}{O} \\ \hline O & \alpha & 0 \\ & 0 & -\alpha \end{array}\right]. \tag{2.2}$$

From (2.2), setting

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad A_{22} \in \mathbb{C}^{2\times 2} \tag{2.3}$$

we get the following blocks relations

$$A_{11}^H A_{11} - A_{11}A_{11}^H = A_{12}A_{12}^H - A_{21}^H A_{21}, \tag{2.4}$$

$$A_{11}^H A_{12} - A_{11}A_{21}^H = A_{12}A_{22}^H - A_{21}^H A_{22}, \tag{2.5}$$

$$A_{22}^H A_{22} - A_{22}A_{22}^H = A_{21}A_{21}^H - A_{12}^H A_{12} + \begin{bmatrix} \alpha & \\ & -\alpha \end{bmatrix}. \tag{2.6}$$

REMARK 1. *Note that equations* (2.2), (2.4), (2.5) *and* (2.6) *still hold whenever we apply to A unitary block diagonal transformations of this kind*

$$\begin{bmatrix} P & O \\ O & S \end{bmatrix},$$

where P is an $(n-2)\times(n-2)$ unitary matrix and S is a 2×2 phase matrix.

REMARK 2. *Note that not all the matrices verifying equation* (2.2) *are almost normal, since it might not exist a matrix C such that* (2.1) *holds.*

Next theorem, gives necessary conditions on the structure of the blocks of an almost normal (not normal) matrix.

THEOREM 2.2. *Let A be an $n\times n$, $n\geq 3$ almost normal matrix, not normal, such that equations* (2.1) *and* (2.2) *hold for a given matrix C with rank at most one. Partition A as in* (2.3) *and C accordingly, that is $C = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} \left[\mathbf{y}_1^H, \mathbf{y}_2^H\right]$, for some vectors $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{y}_1$, $\mathbf{y}_2$, with $||\mathbf{x}_1||_2^2 + ||\mathbf{x}_2||_2^2 = 1$. Then we have*

(a) $\mathbf{x}_1 = 0$, $\mathbf{y}_1 = 0$, $\mathbf{x}_2 = \begin{bmatrix} \zeta \\ -\omega \end{bmatrix}$, and $\mathbf{y}_2^H = \kappa[\omega, -\zeta]$, for a nonzero constant κ, and ζ, ω such that $|\zeta|^2 + |\omega|^2 = 1$.

(b) A_{12} and A_{21} have rank at most one and

$$A_{12} = [\omega\mathbf{d}, \zeta\mathbf{d}], \qquad A_{21} = \begin{bmatrix} \zeta\mathbf{f}^H \\ \omega\mathbf{f}^H \end{bmatrix}, \quad \mathbf{d},\mathbf{f} \in \mathbb{C}^{n-2}. \tag{2.7}$$

(c) Denoting $A_{22} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$ we have that $b_{11} = b_{22}$, and moreover $\omega^2 b_{12} - \zeta^2 b_{21} = \alpha/\kappa$.

Proof.

Since A in not normal, C has rank exactly one, so $\mathbf{x}$ and $\mathbf{y}$ are different from zero. For the same reason $\alpha > 0$. Rewriting equation (2.1) and (2.2) in terms of the blocks of A and the generators of C we obtain four matrix equations:

$$\mathbf{x}_1(\mathbf{y}_1^H A_{11} + \mathbf{y}_2^H A_{21}) - (A_{11}\mathbf{x}_1 + A_{12}\mathbf{x}_2)\mathbf{y}_1^H = 0 \tag{2.8}$$
$$\mathbf{x}_1(\mathbf{y}_1^H A_{12} + \mathbf{y}_2^H A_{22}) - (A_{11}\mathbf{x}_1 + A_{12}\mathbf{x}_2)\mathbf{y}_2^H = 0 \tag{2.9}$$
$$\mathbf{x}_2(\mathbf{y}_1^H A_{11} + \mathbf{y}_2^H A_{21}) - (A_{21}\mathbf{x}_1 + A_{22}\mathbf{x}_2)\mathbf{y}_1^H = 0 \tag{2.10}$$
$$\mathbf{x}_2(\mathbf{y}_1^H A_{12} + \mathbf{y}_2^H A_{22}) - (A_{21}\mathbf{x}_1 + A_{22}\mathbf{x}_2)\mathbf{y}_2^H = \begin{bmatrix} \alpha & 0 \\ 0 & -\alpha \end{bmatrix}. \tag{2.11}$$

To prove *(a)* assume by contradiction that $\mathbf{x}_1 \neq 0$. From equation (2.8) and (2.9), applying Lemma 5.1 in the Appendix, we have

$$\mu\mathbf{y}_1^H = (\mathbf{y}_1^H A_{11} + \mathbf{y}_2^H A_{21}) \tag{2.12}$$
$$\mu\mathbf{y}_2^H = (\mathbf{y}_1^H A_{12} + \mathbf{y}_2^H A_{22}), \tag{2.13}$$

where $\mu = (A_{11}\mathbf{x}_1 + A_{12}\mathbf{x}_2)^H\mathbf{x}_1$.

Then from (2.12) and (2.13) it follows that $\mathbf{y}^H$ is a left eigenvector of A corresponding to the eigenvalue μ. From (2.1) and (2.2), we have

$$\left[\begin{array}{c|cc} O & \multicolumn{2}{c}{O} \\ \hline O & \alpha & 0 \\ & 0 & -\alpha \end{array}\right] = \mathbf{x}\mathbf{y}^H A - A\mathbf{x}\mathbf{y}^H = (\mu I - A)\,\mathbf{x}\mathbf{y}^H \tag{2.14}$$

which is a contradiction because $\alpha \neq 0$, and the matrix on the left hand side of equation (2.14) has rank exactly two while the right hand side is a rank-one matrix. Hence $\mathbf{x}_1 = 0$. We can repeat the same reasoning assuming $\mathbf{y}_1 \neq 0$ and using equations (2.8) and (2.10), obtaining that $\mathbf{y}_1$ must be zero.

Equation (2.11) becomes

$$\mathbf{x}_2\,\mathbf{y}_2^H A_{22} - A_{22}\,\mathbf{x}_2\,\mathbf{y}_2^H = \begin{bmatrix} \alpha & 0 \\ 0 & -\alpha \end{bmatrix}. \tag{2.15}$$

Multiplying on the left by $\mathbf{y}_2^H$ and on the right by $\mathbf{x}_2$ we get that the left hand side of (2.15) vanishes, and we obtain the following relation between $\mathbf{y}_2$ and $\mathbf{x}_2$

$$\mathbf{y}_2^H \begin{bmatrix} \alpha & 0 \\ 0 & -\alpha \end{bmatrix} \mathbf{x}_2 = 0.$$

Taking $\mathbf{y}_2^H = (\bar{y}_{n-1}, \bar{y}_n)$ and $\mathbf{x}_2 = (\zeta, -\omega)^T$, with $|\zeta|^2 + |\omega|^2 = 1$, we have

$$\bar{y}_{n-1}\zeta + \bar{y}_n\omega = 0, \tag{2.16}$$

which implies $\mathbf{y}_2^H = \kappa[\omega, -\zeta]$, for some nonzero constant κ. Thus *(a)* is proved.

Rewriting equations (2.9) and (2.10) setting $\mathbf{x}_1 = 0$ and $\mathbf{y}_1 = 0$, since $\mathbf{x}_2 \neq 0$ and $\mathbf{y}_2 \neq 0$, we have

$$A_{12}\mathbf{x}_2 = 0, \quad \mathbf{y}_2^H A_{21} = 0, \tag{2.17}$$

meaning that the matrices A_{12} and A_{21} have rank at most 1. Because of relation (2.17) we have

$$A_{12} = [\omega\mathbf{d}, \zeta\mathbf{d}], \qquad A_{21} = \begin{bmatrix} \zeta\mathbf{f}^H \\ \omega\mathbf{f}^H \end{bmatrix},$$

where $\mathbf{d}$ and $\mathbf{f}$ are $(n-2)$-vectors. Thus *(b)* is proved.

Part *(c)* of the theorem is proved imposing relation (2.15) on the entries of the matrix A_{22}. $\square$

The fact that A_{12} and A_{21} are rank-one matrices with the same proportional parameters is usefull to prove that the structure of blocks A_{12} and A_{21} can be further simplified. This is the first step towards the block tridiagonalization of A.

THEOREM 2.3. *Let A be an almost normal matrix of size n, $n \geq 3$, satisfying equation* (2.2)*, and let A_{12} and A_{21} have the form described by equation* (2.7)*, with $|\omega^2|+|\zeta|^2 = 1$. Then A can be unitarily reduced to a matrix still satisfying equation* (2.2)*, and such that*

$$A_{12} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ |\omega|\,\delta_1 & |\zeta|\,\delta_1 \\ |\omega|\,\delta_2 & |\zeta|\,\delta_2 \end{bmatrix}, \qquad A_{21} = \begin{bmatrix} \mathbf{0}^H & |\zeta|\,\delta_2 & |\zeta|\,\delta_1 \\ \mathbf{0}^H & |\omega|\,\delta_2 & |\omega|\,\delta_1 \end{bmatrix}, \tag{2.18}$$

meaning that ζ and ω can be supposed real non negative. Moreover $|\delta_1| \geq |\delta_2|$.

Proof. If $\mathbf{d} = \mathbf{f} = \mathbf{0}$ then A_{12} and A_{21} have the required form. Let $\mathbf{d} \neq \mathbf{0}$, and assume $n > 4$, the case $\mathbf{f} \neq \mathbf{0}$ can be treated analogously. From equation (2.4) we have

$$A_{11}^H A_{11} - A_{11} A_{11}^H = A_{12} A_{12}^H - A_{21}^H A_{21}. \tag{2.19}$$

Substituting (2.7) we have

$$A_{11}^H A_{11} - A_{11} A_{11}^H = (|\omega|^2 + |\zeta|^2)(\mathbf{d}\mathbf{d}^H - \mathbf{f}\mathbf{f}^H) = \mathbf{d}\mathbf{d}^H - \mathbf{f}\mathbf{f}^H.$$

Consider the matrix $B = A_{11}^H A_{11} - A_{11} A_{11}^H$. B is Hermitian and hence diagonalizable with unitary transformation U, moreover its trace is zero, then there exists a real nonnegative value β such that

$$U^H B U = U^H (A_{11}^H A_{11} - A_{11} A_{11}^H) U = \left[\begin{array}{c|cc} O & \multicolumn{2}{c}{O} \\ \hline O & \beta & 0 \\ & 0 & -\beta \end{array}\right]. \tag{2.20}$$

Let us reduce A so that B has already the form (2.20).
Splitting $\mathbf{d} = (\mathbf{d}_1; \mathbf{d}_2)$ and $\mathbf{f} = (\mathbf{f}_1; \mathbf{f}_2)$ according with the partition (2.20) we get

$$\begin{aligned} \mathbf{d}_1 \mathbf{d}_1^H - \mathbf{f}_1 \mathbf{f}_1^H &= 0 \\ \mathbf{d}_1 \mathbf{d}_2^H - \mathbf{f}_1 \mathbf{f}_2^H &= 0 \\ \mathbf{d}_2 \mathbf{d}_2^H - \mathbf{f}_2 \mathbf{f}_2^H &= \begin{bmatrix} \beta & 0 \\ 0 & -\beta \end{bmatrix}. \end{aligned} \tag{2.21}$$

If $\beta = 0$, applying Lemma 5.1, from equations (2.21) we have $\mathbf{d} = \mu \mathbf{f}$, with $|\mu| = 1$. We can construct an Householder-like transformation P such that

$$P\mathbf{d} = \gamma \begin{bmatrix} \mathbf{0} \\ 1 \\ 1 \end{bmatrix},$$

where $|\gamma| = \|\mathbf{d}\|_2/\sqrt{2}$ and $\gamma = |\gamma| \exp^{i\theta}$, with $\theta = -\arg(\bar{d}_{n-3} + \bar{d}_{n-2})$, being $\mathbf{d}_2 = (d_{n-3}, d_{n-2})^T$. Setting $\mu = \exp^{i\varphi}$, we can take the unitary matrix Q as

$$Q = \begin{bmatrix} \exp^{i(\varphi/2-\theta)}\, P & \\ & I_2 \end{bmatrix},$$

which is such that

$$Q^H A Q = \left[\begin{array}{cc|c} \multicolumn{2}{c|}{P^H A_{11} P} & \begin{array}{cc} 0 & 0 \\ \hline \omega\delta & \zeta\delta \\ \omega\delta & \zeta\delta \end{array} \\ \hline O & \begin{array}{cc} \zeta\delta & \zeta\delta \\ \omega\delta & \omega\delta \end{array} & A_{22} \end{array}\right], \qquad (2.22)$$

where $\delta = |\gamma| \exp^{i\varphi/2}$, and hence A_{12} and A_{21} have the desired structure.
If $\beta \neq 0$, assume that $\mathbf{d}_1 \neq \mathbf{0}$, we have from Lemma 5.1 that $\mathbf{d}_2 = \mu \mathbf{f}_2$. This is however a contradiction because otherwise $\mathbf{d}_2 \mathbf{d}_2^H - \mathbf{f}_2 \mathbf{f}_2^H = (|\mu|^2 - 1)\mathbf{f}_2 \mathbf{f}_2^H$ will have rank at most one, which is not possible because $\beta \neq 0$. Hence $\mathbf{d} = (\mathbf{0}; \mathbf{d}_2)$ and $\mathbf{f} = (\mathbf{0}; \mathbf{f}_2)$.
Setting $\mathbf{d} = (\mathbf{0}; \Delta_1; \Delta_2)$, and $\mathbf{f} = (\mathbf{0}; \Phi_1; \Phi_2)$, we need to prove that we can reduce A to a matrix where $\bar{\Phi}_1 = \Delta_2$ and $\bar{\Phi}_2 = \Delta_1$.
Rewriting (2.21) we have

$$\begin{aligned} \Delta_1 \bar{\Delta}_1 - \Phi_1 \bar{\Phi}_1 &= \beta \\ \Delta_2 \bar{\Delta}_2 - \Phi_2 \bar{\Phi}_2 &= -\beta \\ \Delta_1 \bar{\Delta}_2 &= \Phi_1 \bar{\Phi}_2. \end{aligned}$$

Summing together the first two equations we have

$$|\Delta_1|^2 + |\Delta_2|^2 = |\Phi_1|^2 + |\Phi_2|^2$$

that, together with $|\Delta_1||\Delta_2| = |\Phi_1||\Phi_2|$, implies that we can either have $|\Delta_1| = |\Phi_1|$ and $|\Delta_2| = |\Phi_2|$ or $|\Delta_1| = |\Phi_2|$ and $|\Delta_2| = |\Phi_1|$. However the first case is not acceptable when $\beta \neq 0$.

By means of phase transformations we get $\bar{\Phi}_1 = \Delta_2$ and $\bar{\Phi}_2 = \Delta_1$, proving that

$$A_{12} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \omega\Delta_1 & \zeta\Delta_1 \\ \omega\Delta_2 & \zeta\Delta_2 \end{bmatrix}, \quad A_{21} = \begin{bmatrix} \mathbf{0}^H & \zeta\Delta_2 & \zeta\Delta_1 \\ \mathbf{0}^H & \omega\Delta_2 & \omega\Delta_1 \end{bmatrix}, \qquad (2.23)$$

eventually with $\Delta_1 = \Delta_2$ if $\beta = 0$. Let us prove that there exist $\tilde{\omega}$ and $\tilde{\zeta}$ real and nonnegative such that (2.18) holds.

Writing the possibly complex numbers ω and ζ in polar form, we have $\omega = |\omega| \exp^{i\theta_\omega}$, and $\zeta = |\zeta| \exp^{i\theta_\zeta}$. Let $\phi = (\theta_\zeta - \theta_\omega)/2$, and consider the phase matrix

$$S = \begin{bmatrix} I_{n-2} & & \\ & \exp^{-i\phi} & \\ & & \exp^{i\phi} \end{bmatrix},$$

then

$$SAS^{-1} = \begin{bmatrix} A_{11} & \tilde{A}_{12} \\ \tilde{A}_{21} & \tilde{A}_{22} \end{bmatrix},$$

where

$$\tilde{A}_{12} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ |\omega|\,\delta_1 & |\zeta|\,\delta_1 \\ |\omega|\,\delta_2 & |\zeta|\,\delta_2 \end{bmatrix}, \qquad \tilde{A}_{21} = \begin{bmatrix} \mathbf{0}^H & |\zeta|\,\delta_2 & |\zeta|\,\delta_1 \\ \mathbf{0}^H & |\omega|\,\delta_2 & |\omega|\,\delta_1 \end{bmatrix}, \qquad (2.24)$$

where $\delta_1 = \Delta_1 \exp^{i(\theta_\omega+\theta_\zeta)/2}$ and $\delta_2 = \Delta_2 \exp^{i(\theta_\omega+\theta_\zeta)/2}$. The block $\tilde{A}_{22}$ is not modified respect to A_{22} in its diagonal entries and equation (2.20) still holds. Finally, if $|\delta_1| < |\delta_2|$, rows and columns $n-3$ and $n-2$ can be permuted without destroying the structure properties used in this theorem.

If $n = 3$ or $n = 4$ the blocks A_{12} and A_{21} do not have the zeros in the first rows and columns, since they have respectively size 1×2 (when $n = 3$) and 2×2 (if $n = 4$).□

Theorem 2.2 assumes A not normal. However, it is in general possible that the leading principal $(n-2) \times (n-2)$ minor is normal. Next Corollary gives necessary and sufficient condition for A_{11} to be normal.

COROLLARY 2.4. *Under the hypothesis of Theorem 2.3 the matrix A_{11} is normal if and only if $|\delta_1| = |\delta_2|$.*

Proof. If A is reduced in the form described in Theorem 2.3, from (2.4) we have

$$A_{11}^H A_{11} - A_{11} A_{11}^H = -A_{21}^H A_{21} + A_{12} A_{12}^H. \tag{2.25}$$

Consider the right hand side of (2.25), using the equality (2.18), we have

$$A_{11}^H A_{11} - A_{11} A_{11}^H = \begin{bmatrix} O & & \\ & |\delta_1|^2 - |\delta_2|^2 & \\ & & |\delta_2|^2 - |\delta_1|^2 \end{bmatrix}.$$

Then we have that A_{11} is normal iff $|\delta_1| = |\delta_2|$. □

As underlined in remark 2, not all the matrices satisfying equation (2.2) are almost normal. Next theorem gives necessary and sufficient conditions for a matrix A to belong to that class, that is for the existence of a rank-one matrix C such that $A^H A - AA^H = CA - AC$. The theorem will be used in Section 3.1 for proving sufficient conditions.

THEOREM 2.5. *Let A be an $n \times n$ complex matrix, with $n \geq 3$. Partition A as follows*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad A_{11} \in \mathbb{C}^{(n-2)\times(n-2)}.$$

If

(a) Equation (2.2) holds with $\alpha \neq 0$,

(b) A_{12} and A_{21} are structured as in equation (2.7), with ω and ζ nonnegative reals such that $\omega^2 + \zeta^2 = 1$,

(c) $A_{22} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{11} \end{bmatrix}$.

Then A is almost normal (not normal) meaning that there exists a rank one matrix C such that $A^H A - AA^H = CA - AC$, if and only if $\omega^2 b_{12} - \zeta^2 b_{21} \neq 0$.

Proof. If A is almost normal, not normal, then for Theorem 2.2 the rank-one matrix $C = \begin{pmatrix} \mathbf{0} \\ \mathbf{x}_2 \end{pmatrix} \left(\mathbf{0}^H, \mathbf{y}_2^H\right)$ is such that $\mathbf{x}_2 = \begin{bmatrix} \zeta \\ -\omega \end{bmatrix}$, $\mathbf{y}_2^H = \kappa\,(\omega, -\zeta)$. With such a choice

$$CA - AC = \left[\begin{array}{c|cc} O & \multicolumn{2}{c}{O} \\ \hline O & \kappa(\omega^2 b_{12} - \zeta^2 b_{21}) & 0 \\ & 0 & -\kappa(\omega^2 b_{12} - \zeta^2 b_{21}) \end{array}\right] = \left[\begin{array}{c|cc} O & \multicolumn{2}{c}{O} \\ \hline O & \alpha & 0 \\ & 0 & -\alpha \end{array}\right].$$

Since $\alpha \neq 0$ (A is not normal) then $(\omega^2 b_{12} - \zeta^2 b_{21}) \neq 0$. Vice versa, if $(\omega^2 b_{12} - \zeta^2 b_{21}) \neq 0$, ω and ζ being the coefficients in (2.7), a rank-one matrix C such that $A^H A - AA^H = CA - AC$ is given by $C = \begin{pmatrix} \mathbf{0} \\ \mathbf{x}_2 \end{pmatrix} \left(\mathbf{0}^H, \mathbf{y}_2^H\right)$ with $\mathbf{x}_2 = \begin{bmatrix} \zeta \\ -\omega \end{bmatrix}$, $\mathbf{y}_2^H = \kappa\,(\omega, -\zeta)$, and with

$$\kappa = \frac{\alpha}{\omega^2 b_{12} - \zeta^2 b_{21}}.$$

$\square$

We are now ready to prove the main result of this paper. Notice that it is not required A to be almost normal, i.e. the existence of a rank one matrix C such that $A^HA - AA^H = CA - AC$ is not assumed. It states that any almost normal matrix unitary reduced to a form such that (2.2) holds, has a leading principal submatrix with the same structure.

THEOREM 2.6. *Let A be an $n \times n$ complex matrix, with $n \geq 6$. Partition A as follows*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad A_{11} \in \mathbb{C}^{(n-2)\times(n-2)}.$$

If

(a) Equation (2.2) *holds with $\alpha \neq 0$,*

(b) A_{12} and A_{21} are structured as in equation (2.18)*, with ω and ζ real nonnegative such that $\omega^2 + \zeta^2 = 1$ and $|\delta_1| > |\delta_2|$,*

(c) $A_{22} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{11} \end{bmatrix}$,

then $\omega \neq \zeta$ and A can be unitary reduced to a matrix such that the block A_{11} has the form

$$A_{11} = \left[\begin{array}{cc|c} \multicolumn{2}{c|}{\multirow{2}{*}{$\hat{A}_{11}$}} & O \\ \cline{3-3} & & \begin{matrix} \omega_1\delta_1^{(1)} & \zeta_1\delta_1^{(1)} \\ \omega_1\delta_2^{(1)} & \zeta_1\delta_2^{(1)} \end{matrix} \\ \hline O & \begin{matrix} \zeta_1\delta_2^{(1)} & \zeta_1\delta_1^{(1)} \\ \omega_1\delta_2^{(1)} & \omega_1\delta_1^{(1)} \end{matrix} & \hat{A}_{22} \end{array}\right], \qquad (2.26)$$

and satisfies (a), (c). Moreover the parameters of A_{11} are such that $|\delta_1^{(1)}| \geq |\delta_2^{(1)}|$, ω_1 and ζ_1 real non negative, $\omega_1 > \zeta_1$ with $\omega_1^2 + \zeta_1^2 = 1$. In particular equation (2.2) *for A_{11}*

$$A_{11}^HA_{11} - A_{11}A_{11}^H = \left[\begin{array}{c|cc} O & \multicolumn{2}{c}{O} \\ \hline O & \beta & 0 \\ & 0 & -\beta \end{array}\right],$$

holds with $\beta > 0$.

Proof. Since $n > 6$, we can partition A_{11} as follows

$$A_{11} = \begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} \\ \hat{A}_{21} & \hat{A}_{22} \end{bmatrix}, \text{where } \hat{A}_{11} \in \mathbb{C}^{(n-4)\times(n-4)}, \hat{A}_{22} \in \mathbb{C}^{2\times 2}.$$

Rewriting equation (2.5) in block form we have

$$\begin{bmatrix} \hat{A}_{11}^H & \hat{A}_{21}^H \\ \hat{A}_{12}^H & \hat{A}_{22}^H \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \omega\delta_1 & \zeta\delta_1 \\ \omega\delta_2 & \zeta\delta_2 \end{bmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \zeta\bar{\delta}_2 & \omega\bar{\delta}_2 \\ \zeta\bar{\delta}_1 & \omega\bar{\delta}_1 \end{bmatrix} A_{22} - \begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} \\ \hat{A}_{21} & \hat{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \zeta\bar{\delta}_2 & \omega\bar{\delta}_2 \\ \zeta\bar{\delta}_1 & \omega\bar{\delta}_1 \end{bmatrix} - \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \omega\delta_1 & \zeta\delta_1 \\ \omega\delta_2 & \zeta\delta_2 \end{bmatrix} A_{22}^H = 0 \qquad (2.27)$$

Set $\hat{A}_{12} = [\mathbf{g}, \mathbf{h}]$, and $\hat{A}_{21}^H = [\mathbf{r}, \mathbf{s}]$, where $\mathbf{g}, \mathbf{h}, \mathbf{r}$ and $\mathbf{s}$ are $n-4$ complex vectors. Looking at the upper $(n-4)$ equations of (2.27) we get

$$\begin{aligned} \omega(\delta_1\,\mathbf{r} + \delta_2\,\mathbf{s}) - \zeta(\bar{\delta}_2\,\mathbf{g} + \bar{\delta}_1\,\mathbf{h}) &= 0 \\ \zeta(\delta_1\,\mathbf{r} + \delta_2\,\mathbf{s}) - \omega(\bar{\delta}_2\,\mathbf{g} + \bar{\delta}_1\,\mathbf{h}) &= 0. \end{aligned}$$

Writing the four equations of (2.27) involving the entries a_{ij} of $\hat{A}_{22}$ and the entries b_{ij} of A_{22} we get

$$\omega\delta_1\bar{a}_{11}+\omega\delta_2\bar{a}_{21}-\zeta\bar{\delta}_1 a_{12}-\zeta\bar{\delta}_2 a_{11}=\omega\delta_1\bar{b}_{11}-\omega\bar{\delta}_2 b_{21}+\zeta\delta_1\bar{b}_{12}-\zeta\bar{\delta}_2 b_{11} \quad (2.28)$$
$$-\omega\bar{\delta}_1 a_{12}-\omega\bar{\delta}_2 a_{11}+\zeta\delta_1\bar{a}_{11}+\zeta\delta_2\bar{a}_{21}=\omega\delta_1\bar{b}_{21}-\omega\bar{\delta}_2 b_{11}+\zeta\delta_1\bar{b}_{11}-\zeta\bar{\delta}_2 b_{12} \quad (2.29)$$
$$\omega\delta_1\bar{a}_{12}+\omega\delta_2\bar{a}_{22}-\zeta\bar{\delta}_1 a_{22}-\zeta\bar{\delta}_2 a_{21}=\omega\delta_2\bar{b}_{11}-\omega\bar{\delta}_1 b_{21}+\zeta\delta_2\bar{b}_{12}-\zeta\bar{\delta}_1 b_{11} \quad (2.30)$$
$$-\omega\bar{\delta}_1 a_{22}-\omega\bar{\delta}_2 a_{21}+\zeta\delta_1\bar{a}_{12}+\zeta\delta_2\bar{a}_{22}=-\omega\bar{\delta}_1 b_{11}+\omega\delta_2\bar{b}_{21}+\zeta\delta_2\bar{b}_{11}-\zeta\bar{\delta}_1 b_{12} \quad (2.31)$$

Observe that $\omega\neq\zeta$. Otherwise, assume by contradiction that $\omega=\zeta$, then the left hand side of equations (2.28) and (2.29) are the same, and subtracting them we get

$$\delta_1(\bar{b}_{21}-\bar{b}_{12})=\bar{\delta}_2(b_{12}-b_{21}). \quad (2.32)$$

From equation (2.2) on the block in position (2,2) we have

$$A_{12}^H A_{12}+A_{22}^H A_{22}-A_{21}A_{21}^H-A_{22}A_{22}^H=\begin{bmatrix}\alpha & \\ & -\alpha\end{bmatrix}.$$

Using the fact $\omega^2-\zeta^2=0$, we have

$$|b_{21}|^2-|b_{12}|^2=\alpha\neq 0.$$

From the absolute value of (2.32) and observing that $|\bar{b}_{21}-\bar{b}_{12}|=|b_{12}-b_{21}|\neq 0$, we get $|\delta_1|=|\delta_2|$ which contradicts the hypothesis because $|\delta_1|>|\delta_2|$.

Hence, because $\omega\neq\zeta$, with two linear combinations we get

$$(\omega^2-\zeta^2)(\bar{\delta}_2\,\mathbf{g}+\bar{\delta}_1\,\mathbf{h})=0 \quad (2.33)$$
$$(\omega^2-\zeta^2)\,(\delta_1\,\mathbf{r}+\delta_2\,\mathbf{s})=0 \quad (2.34)$$

Since $\omega^2-\zeta^2\neq 0$, the matrices $\hat{A}_{12}=[\mathbf{g},\mathbf{h}]$, and $\hat{A}_{21}^H=[\mathbf{r},\mathbf{s}]$ have rank at most one. In particular there exist two $n-4$ vectors $\mathbf{d}^{(1)}$ and $\mathbf{f}^{(1)}$ such that

$$\hat{A}_{12}=\left[\omega_1\,\mathbf{d}^{(1)},\zeta_1\,\mathbf{d}^{(1)}\right],\quad \hat{A}_{21}=\begin{bmatrix}\zeta_1\,\mathbf{f}^{(1)H}\\ \omega_1\,\mathbf{f}^{(1)H}\end{bmatrix},$$

where $\zeta_1=\kappa\bar{\delta}_2$ and $\omega_1=-\kappa\bar{\delta}_1$ and κ is such that $|\zeta_1|^2+|\omega_1|^2=1$. It follows that $|\omega_1|>|\zeta_1|$.

We need now to prove that the diagonal entries of $\hat{A}_{22}$ are the same, that is $a_{11}=a_{22}$.

With some linear combinations of equations (2.28)-(2.31) we obtain

$$(\omega^2-\zeta^2)\bar{\delta}_2(a_{11}-a_{22})=0$$
$$(\omega^2-\zeta^2)\bar{\delta}_1(a_{11}-a_{22})=0.$$

Since we cannot have $\delta_1=\delta_2=0$, we get $a_{11}=a_{22}$ as claimed, proving property *(c)* for A_{11}.

Because of the structure of A_{12} and A_{21} described in *(b)* and using (2.4), we get

$$A_{11}^H A_{11}-A_{11}A_{11}^H=\begin{bmatrix}O & & \\ & |\delta_1|^2-|\delta_2|^2 & \\ & & |\delta_2|^2-|\delta_1|^2\end{bmatrix}. \quad (2.35)$$

proving *(a)* for A_{11}, and the property that $\beta=|\delta_1|^2-|\delta_2|^2$ is positive.

We note that matrix A_{11} satisfies the hypothesis of Theorem 2.3. Hence block $\hat{A}_{12}$ and $\hat{A}_{21}$ have the structure described in (2.26) with $|\delta_1^{(1)}| \geq |\delta_2^{(1)}|$. Note that the phase transformation which makes ω_1 and ζ_1 real nonnegative, modifies the external border, changing the polar part in the representation of δ_1 and δ_2: in detail, setting

$$\delta_1 = |\delta_1| \exp(i\varepsilon_1), \quad \delta_2 = |\delta_2| \exp(i\varepsilon_2), \quad \kappa = |\kappa| \exp(i\varepsilon_3),$$

the new value for δ_1 and δ_2 are respectively

$$|\delta_1| \exp(i(\varepsilon_1 + \varepsilon_2 - \varepsilon_3)), \quad -|\delta_2| \exp(i(\varepsilon_1 + \varepsilon_2 - \varepsilon_3)),$$

finding that $\delta_2 |\delta_1| = -\delta_1 |\delta_2|$.
□

Theorem 2.6 does not consider the cases A is a matrix of size $n < 6$. However, if $n = 3$, $A_{11} = a_{11}$ collapses to a number and the thesis does not make any sense, also because hypothesis (b) is not true. If $n = 4$, we do not have $\delta_1^{(1)}$ and $\delta_2^{(1)}$ because A_{11} is 2×2 and then coincides with $\hat{A}_{22}$. If $n = 5$ the structure of (2.26) simplifies, $\hat{A}_{11}$ is just a number and $|\delta_1^{(1)}| = |\delta_2^{(1)}|$.

Note that if A is almost normal, and hence Theorem 2.6 can be applied, the matrix A_{11} is not, in general almost normal because matrix $\hat{C}$ such that $A_{11}^H A_{11} - A_{11} A_{11}^H = \hat{C} A_{11} - A_{11} \hat{C}$ might not exist. However if $\omega_1^2 a_{12} - \zeta_1^2 a_{21} \neq 0$ then, for Theorem 2.5, A_{11} is almost normal as well.

Summarizing for any $n \times n$ almost normal, not normal matrix A, with $n > 4$, partitioned as

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad \text{where } A_{11} \in \mathbb{C}^{(n-2)\times(n-2)}, \quad A_{22} \in \mathbb{C}^{2\times 2},$$

then either A_{11} is normal or we can unitary reduce A to the following structure

$$A = \left[\begin{array}{cc|c|c} \multicolumn{2}{c|}{\hat{A}_{11}} & \begin{array}{cc} \multicolumn{2}{c}{O} \\ \hline \omega_1 \delta_1^{(1)} & \zeta_1 \delta_1^{(1)} \\ \omega_1 \delta_2^{(1)} & \zeta_1 \delta_2^{(1)} \end{array} & O \\ \hline O & \multicolumn{1}{|c|}{\begin{array}{cc} \zeta_1 \delta_2^{(1)} & \zeta_1 \delta_1^{(1)} \\ \omega_1 \delta_2^{(1)} & \omega_1 \delta_1^{(1)} \end{array}} & \hat{A}_{22} & \begin{array}{cc} \omega\delta_1 & \zeta\delta_1 \\ \omega\delta_2 & \zeta\delta_2 \end{array} \\ \hline \multicolumn{2}{c|}{O} & \begin{array}{cc} \zeta\delta_2 & \zeta\delta_1 \\ \omega\delta_2 & \omega\delta_1 \end{array} & A_{22} \end{array} \right] \tag{2.36}$$

that is the blocks in position $(1,2)$ and $(2,1)$ of A_{11} have the same structure of the more external blocks and the entries of the internal blocks are related to parameters appearing in the more external blocks.

Previous theorem does not account for the case $|\delta_1| = |\delta_2|$, but in Corollary 2.4 we proved that in that case A_{11} is normal. Next theorem gives additional necessary conditions on the parameters and on the diagonal form of A_{11}.

THEOREM 2.7. *Let A be an $n \times n$ complex matrix, with $n > 4$. Partition A as follows*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad A_{11} \in \mathbb{C}^{(n-2)\times(n-2)}.$$

If

(a) Equation (2.2) *holds with* $\alpha \neq 0$,

(b) A_{12} *and* A_{21} *are structured as in equation* (2.18)*, with* ω *and* ζ *real nonnegative such that* $\omega^2 + \zeta^2 = 1$ *and* $|\delta_1| = |\delta_2|$,

(c) $A_{22} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{11} \end{bmatrix}$,

then the matrix A_{11} *is normal and by a unitary transformation we can reduce A to a matrix such that*

(1) $\delta_1 = \delta_2 = \delta$,

(2) and either $\delta = 0$ *(meaning that is A block diagonal), or the matrix* $\omega\delta A_{11}^H - \zeta\bar{\delta}A_{11}$ *has as eigenvector* $\mathbf{e}_{n-3} + \mathbf{e}_{n-2}$ *corresponding to an eigenvalue*

$$\omega(\delta\bar{b}_{11} - \bar{\delta}b_{21}) + \zeta(\delta\bar{b}_{12} - \bar{\delta}b_{11}).$$

Proof. For hypotheses (a), (b) and (c) we can apply Corollary 2.4 concluding that if $|\delta_1| = |\delta_2|$ then A_{11} is normal.

Proceeding as in Theorem 2.3 we can unitary reduce A to the form (2.22), obtaining a $\delta = \delta_1 = \delta_2$. In the case $\delta = 0$ the matrix turns out to be block diagonal with A_{11} normal.

Assume $\delta \neq 0$. Using equality (2.5), we have

$$A_{11}^H \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \omega\delta & \zeta\delta \\ \omega\delta & \zeta\delta \end{bmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \zeta\bar{\delta} & \omega\bar{\delta} \\ \zeta\bar{\delta} & \omega\bar{\delta} \end{bmatrix} A_{22} - A_{11} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \zeta\bar{\delta} & \omega\bar{\delta} \\ \zeta\bar{\delta} & \omega\bar{\delta} \end{bmatrix} - \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \omega\delta & \zeta\delta \\ \omega\delta & \zeta\delta \end{bmatrix} A_{22}^H = 0. \tag{2.37}$$

The thesis follows setting $\mathbf{u} = \mathbf{e}_{n-3} + \mathbf{e}_{n-2}$ and rewriting first column of previous equality as

$$\left(\omega\delta A_{11}^H - \zeta\bar{\delta}A_{11}\right)\mathbf{u} = \left((\omega\delta\bar{b}_{11} - \omega\bar{\delta}b_{21}) - (\zeta\bar{\delta}b_{11} + \zeta\delta\bar{b}_{12})\right)\mathbf{u}.$$

Observing that the second column of (2.37) gives the same relation that one obtains using the first column, we can conclude that $\mathbf{e}_{n-3} + \mathbf{e}_{n-2}$ is an eigenvector of the matrix $\left(\omega\delta A_{11}^H - \zeta\bar{\delta}A_{11}\right)$ corresponding to the eigenvalue $(\omega\delta\bar{b}_{11} - \omega\bar{\delta}b_{21}) - (\zeta\bar{\delta}b_{11} + \zeta\delta\bar{b}_{12})$.
□

Summarizing, if the block A_{11} is normal we can have two situations. A is block diagonal, that is

$$A = \left[\begin{array}{c|c} A_{11} & O \\ \hline O & A_{22} \end{array}\right], \text{ and moreover } A_{11} \text{ is normal}$$

otherwise, if A_{12} and A_{21} are not null, A can be unitary reduced to the following structure

$$A = \left[\begin{array}{cc|c} \multicolumn{2}{c|}{A_{11}} & \begin{array}{cc} \multicolumn{2}{c}{O} \\ \hline \omega\delta & \zeta\delta \\ \omega\delta & \zeta\delta \end{array} \\ \hline O & \begin{array}{cc} \zeta\delta & \zeta\delta \\ \omega\delta & \omega\delta \end{array} & A_{22} \end{array}\right]. \tag{2.38}$$

Theorem 2.7 concludes the proof of the structure of almost normal matrices, showing that we can exploit a 2×2 block tridiagonal structure of almost normal matrices until we eventually meet a leading principal block which is normal. In that case, we stop.

3. Sufficient conditions. In this section we show how we can construct an almost normal matrix bordering an $(n-2)\times(n-2)$ almost normal matrix with suitable vectors. This construction can be carried on inductively starting from a normal matrix, or a non normal 1×1 or 2×2 matrices.

THEOREM 3.1. *Let A_{11} be an $(n-2)\times(n-2)$ matrix such that either A_{11} is normal or is not normal but* (2.2) *holds (for A_{11}). Assume A_{11} has the following structure*

$$A_{11} = \left[\begin{array}{cc|cc} \multicolumn{2}{c|}{\multirow{2}{*}{$\hat{A}_{11}$}} & \multicolumn{2}{c}{O} \\ \cline{3-4} & & \omega_1\delta_1^{(1)} & \zeta_1\delta_1^{(1)} \\ & & \omega_1\delta_2^{(1)} & \zeta_1\delta_2^{(1)} \\ \hline O & \begin{array}{cc} \zeta_1\delta_2^{(1)} & \zeta_1\delta_1^{(1)} \\ \omega_1\delta_2^{(1)} & \omega_1\delta_1^{(1)} \end{array} & a_{11} & a_{12} \\ & & a_{21} & a_{11} \end{array}\right],$$

with $|\delta_1^{(1)}| \geq |\delta_2^{(1)}|$, and $\omega^{(1)} > \zeta^{(1)}$. Then we can border A_{11} in infinitely many ways to obtain an $n\times n$ matrix

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad \textit{where } A_{22} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{11} \end{bmatrix} \in \mathbb{C}^{2\times 2},$$

such that (2.2) *holds for A with $\alpha \neq 0$, and A has the structure described in* (2.36). *Moreover we can always choose the new parameters $\delta_1, \delta_2, \omega, \zeta$ and the b_{ij} in such a way that A is almost normal.*

Proof. Assume first that A_{11} is not normal and the eigenvalues of the difference matrix $A_{11}^H A_{11} - A_{11}A_{11}^H$ are $\beta, -\beta$, where $\beta > 0$. We want to prove that if we construct A_{12}, A_{21} and A_{22} as in Theorem 2.6 then the matrix A has the desired properties. Let ω and ζ be non negative real numbers such that $\omega^2+\zeta^2 = 1$, with $\omega \neq \zeta$. Set

$$A_{12} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \omega\delta_1 & \zeta\delta_1 \\ \omega\delta_2 & \zeta\delta_2 \end{bmatrix}, \quad A_{21} = \begin{bmatrix} \mathbf{0}^H & \zeta\delta_2 & \zeta\delta_1 \\ \mathbf{0}^H & \omega\delta_2 & \omega\delta_1 \end{bmatrix},$$

where δ_1 and δ_2 are such that (2.33) and (2.34) hold. This is achieved by setting $\delta_1 = \mu\omega_1$ and $\delta_2 = -\mu\zeta_1$, for any complex constant μ such that $|\mu| = \sqrt{\beta}/(\omega_1^2-\zeta_1^2)$.

Block A_{22} is determined by imposing conditions (2.28)-(2.31) where the unknowns this time are the b_{ij}'s, since

$$\hat{A}_{22} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{11} \end{bmatrix}$$

is given. Combining ζ times equation (2.28) with ω times equation (2.29) we get

$$(\omega^2-\zeta^2)\left(a_{12}\bar{\delta}_1 + a_{11}\bar{\delta}_2\right) = (b_{12}-b_{21})\omega\zeta\bar{\delta}_2 + (\zeta^2\bar{b}_{12} - \omega^2\bar{b}_{21})\delta_1 + b_{11}(\omega^2-\zeta^2)\bar{\delta}_2.$$

Combining ω times equation (2.28) and ζ times equation (2.29) we have

$$(\omega^2-\zeta^2)\left(\bar{a}_{11}\delta_1 + \bar{a}_{21}\delta_2\right) = (\bar{b}_{12}-\bar{b}_{21})\omega\zeta\delta_1 + (\zeta^2 b_{12} - \omega^2 b_{21})\bar{\delta}_2 + \bar{b}_{11}(\omega^2-\zeta^2)\delta_1.$$

We do the following changes of variables $x_1 = b_{11}$, $x_2 = \zeta^2\bar{b}_{12} - \omega^2\bar{b}_{21}$ and $x_3 = b_{12}-b_{21}$. We get the following linear system

$$\begin{bmatrix} (\omega^2-\zeta^2)\bar{\delta}_2 & \delta_1 & \omega\zeta\bar{\delta}_2 \\ (\omega^2-\zeta^2)\bar{\delta}_1 & \delta_2 & \omega\zeta\bar{\delta}_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = (\omega^2-\zeta^2)\begin{bmatrix} a_{12}\bar{\delta}_1 + a_{11}\bar{\delta}_2 \\ a_{21}\bar{\delta}_2 + a_{11}\bar{\delta}_1 \end{bmatrix}.$$

The above linear system is consistent and has infinitely many solutions of the form

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = k \begin{bmatrix} \omega\zeta \\ 0 \\ \zeta^2-\omega^2 \end{bmatrix} + \frac{1}{|\delta_1|^2-|\delta_2|^2} \begin{bmatrix} a_{11}(|\delta_1|^2-|\delta_2|^2)+\delta_1\,\bar{\delta}_2\,a_{21}-\bar{\delta}_1\,\delta_2\,a_{12} \\ (\omega^2-\zeta^2)(\bar{\delta}_1^2\,a_{12}-\bar{\delta}_2^2\,a_{21}) \\ 0 \end{bmatrix}. \tag{3.1}$$

Then $b_{11} = x_1$ and b_{12} and b_{21} are the solutions of the linear system

$$\begin{bmatrix} \zeta^2 & -\omega^2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} b_{12} \\ b_{21} \end{bmatrix} = \begin{bmatrix} \bar{x}_2 \\ x_3 \end{bmatrix},$$

obtaining

$$b_{12} = \frac{\bar{x}_2}{\zeta^2-\omega^2} - k\,\omega^2, \quad b_{21} = \frac{\bar{x}_2}{\zeta^2-\omega^2} - k\,\zeta^2. \tag{3.2}$$

To complete the proof we need to show that we can always choose k in such a way $\omega^2\,b_{12} - \zeta^2\,b_{21} \neq 0$. In Theorem 2.5 we show that in this case there exists a rank-one matrix C for which A is almost normal. Note that $\omega^2\,b_{12} - \zeta^2\,b_{21} = x_3 - \bar{x}_2$. Hence

$$\omega^2\,b_{12} - \zeta^2\,b_{21} = k\,(\zeta^2-\omega^2) - \frac{1}{|\delta_1|^2-|\delta_2|^2}\,(\omega^2-\zeta^2)(\delta_1^2\,\bar{a}_{12} - \delta_2^2\,\bar{a}_{21}).$$

Since x_2 does not depend on k, $|\delta_1| \neq |\delta_2|$ and $\zeta^2-\omega^2 \neq 0$, we achieve our goal by choosing

$$k \neq \frac{-\delta_1^2\,\bar{a}_{12} + \delta_2^2\,\bar{a}_{21}}{|\delta_1|^2-|\delta_2|^2}.$$

Assume now A_{11} is normal. In accordance with Theorem 2.7, either we can choose $\delta = 0$, or the parameters ω, ζ and δ should be chosen in such a way matrix $\omega\,\delta A_{11}^H - \zeta\,\bar{\delta} A_{11}$ has eigenvector $\mathbf{e}_{n-3} + \mathbf{e}_{n-2}$. If we decide for $\delta = 0$, then we set $A_{12} = O$ and $A_{21} = O$, and A is a block diagonal matrix $A = \begin{bmatrix} A_{11} & O \\ O & A_{22} \end{bmatrix}$, with A_{22} a generic 2×2 matrix with $b_{11} = b_{22}$. With a unitary transformation acting on the last two rows and columns of A we can always obtain a matrix satisfying (2.2). With the particular choice $b_{22} = b_{11}$ and $b_{12} \neq b_{21}$ we get an almost normal matrix and

$$C = \kappa \left[\begin{array}{c|cc} O & \multicolumn{2}{c}{O} \\ \hline O & 1 & -1 \\ & -1 & 1 \end{array}\right],$$

where $\kappa\,(b_{12} - b_{21}) = |b_{21}|^2 - |b_{12}|^2$.

In the case A_{11} is normal and $\mathbf{u} = \mathbf{e}_{n-3} + \mathbf{e}_{n-2}$ is eigenvector of $\omega\,\delta A_{11}^H - \zeta\,\bar{\delta} A_{11}$ we have infinitely many different choices for the entries b_{ij} of A_{22}. Let $A_{11} = QDQ^H$, with D diagonal and Q the unitary matrix of the eigenvectors. Denote by μ_i the eigenvalues of A which are the diagonal entries of D, and let μ be the eigenvalue of A_{11} such that

$$\omega\,\delta\,\bar{\mu} - \zeta\,\bar{\delta}\,\mu = \omega(\delta\,\bar{b}_{11} - \bar{\delta}\,b_{21}) + \zeta(\delta\,\bar{b}_{12} - \bar{\delta}\,b_{11}), \tag{3.3}$$

knowing in advance that we need $b_{11} = b_{22}$. For the necessary conditions proved in Theorem 2.7, we have to find suitable b_{ij}'s such that equality (3.3) holds.

Reasoning similarly to what done for the non normal case, we end up with the following linear equation

$$(\omega^2-\zeta^2)\,\bar{\delta}\,\mu = (\omega^2-\zeta^2)\,\bar{\delta}\,b_{11} + \delta\,(\zeta^2\,\bar{b}_{12} - \omega^2\,\bar{b}_{21}) + \omega\,\zeta\,\bar{\delta}(b_{12} - b_{21}). \tag{3.4}$$

With the same change of variables $x_1 = b_{11}$, $x_2 = \zeta^2\,\bar{b}_{12} - \omega^2\,\bar{b}_{21}$ and $x_3 = b_{12} - b_{21}$, we have that if $\omega \neq z$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = k_1 \begin{bmatrix} \omega\zeta \\ 0 \\ (\zeta^2 - \omega^2) \end{bmatrix} + k_2 \begin{bmatrix} \zeta^2 - \omega^2 \\ \bar{\delta}/\delta \\ 0 \end{bmatrix} + \begin{bmatrix} \mu \\ 0 \\ 0 \end{bmatrix},$$

where k_1 and k_2 are free complex parameters. If $\omega = \zeta$ the above equation (3.4) simplifies and we have

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = k_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + k_2 \begin{bmatrix} 0 \\ \bar{\delta}/\delta \\ -2 \end{bmatrix},$$

where k_1 is a free parameter, $|k_2|$ is free but the angle α_k of k_2 is such that $\alpha_k = \alpha_\delta \pm \pi/2$. where α_δ is the angle of δ.

Thus we have infinitely many solutions in the above systems depending on k_1 and k_2 , and hence choices of the b_{ij}'s. Moreover, similarly to the not normal case, we can always choose k_1 and k_2 in such a way there exists a rank-one matrix C such that $A^H A - AA^H = CA - AC$. For example, possible values guaranteeing $\omega^2\, b_{12} - \zeta^2\, b_{21} \neq 0$ are $k_1 = 0$ and $k_2 \neq 0$ (in the case $\omega = \zeta$, k_2 should still be such that $\alpha_k = \alpha_\delta \pm \pi/2$). □

REMARK 3. *Note that if A_{11} is not normal, for any two different values of k, the corresponding almost normal matrices differ only for the tailing 2×2 principal minor A_{22}. The difference of the two almost normal matrices is a multiple of matrix C itself.*

In the case the principal minor A_{11} is normal, given an almost normal matrix obtained with parameters k_1 and k_2 we can choose a second almost normal matrix B differing from A only for the tailing 2×2 principal block with parameters k_1' and the same k_2 such that $A - B = (k_1 - k_1')C$.

Previous theorem gives the conditions for the recursive construction of an almost normal matrix. We now have to describe the structure of the basic block to start recursion. We have two cases. A_{11} is normal, or it is a 2×2 non normal matrix. If A_{11} is an odd-size normal matrix - eventually of size one - we can border it as described in Theorem 3.1 obtaining an odd-size almost normal matrix. If A_{11} has even size, we end up with an even-size almost normal matrix. Everything works well starting with a with a 2×2 matrix as proved in next theorem.

THEOREM 3.2. *Any 2×2 matrix is almost normal.*

Proof. We can assume A is in the Schur form, that is A is upper triangular,

$$A = \begin{bmatrix} \lambda_1 & \eta \\ 0 & \lambda_2 \end{bmatrix}.$$

We want to show that there exists a rank one matrix C and a polynomial $p(x) = ax + b$ with degree at most one, such that such that $A^H - C = p(A)$. Let ε, $\varepsilon \neq 0$ such that $\lambda_1 + \varepsilon \neq \lambda_2$. Then there exist a and b, unique, such that

$$\bar{\lambda}_1 = a\lambda_1 + b + \varepsilon a$$
$$\bar{\lambda}_2 = a\lambda_2 + b - \frac{1}{\varepsilon}|\eta|^2.$$

In fact

$$\det(\begin{bmatrix} \lambda_1 + \varepsilon & 1 \\ \lambda_2 & 1 \end{bmatrix}) = \lambda_1 + \varepsilon - \lambda_2 \neq 0.$$

Once we found a and b solutions of the linear system, we compute $C = A^H - p(A)$ which is given by

$$C = \begin{bmatrix} \bar{\lambda}_1 - (a\lambda_1 + b) & -a\eta \\ \bar{\eta} & \bar{\lambda}_2 - (a\lambda_2 + b) \end{bmatrix},$$

whose determinant is zero, so that C has rank at most one. □

4. Relationship with normal matrices. We already stressed that when adding a rank-one perturbation to a normal matrix we may end up with a matrix which is *not* almost normal because $A^HA - AA^H$ can have rank up to four. However, next theorem shows that any almost normal matrix can be viewed as a *particular* rank-one correction of a *particular* normal matrix.

THEOREM 4.1. *Let A be an almost normal matrix. Then 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 $A^HA - AA^H = CA - AC$.*

Proof. Let A be an almost normal matrix already in the condensed form described in (2.36). In this proof, for the sake of simplicity, we assume A_{11} is not normal; the case A_{11} normal is similar. In particular, as remarked in 3, we can choose infinitely many $A^{(k)}$ that differ from A only for the lower right 2×2 block, taking different values of k in the solution of system (3.1). Among the infinite possible $A^{(k)}$ we prove that we can always find a particular $\tilde{k}$ such that the matrix $A^{(\tilde{k})}$ is normal.

In fact, an almost normal matrix, which is also normal should satisfy equations (2.4), (2.5) and (2.6) with $\alpha = 0$. Let $A_{22}^{(k)} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{11} \end{bmatrix}$ the lower 2×2 diagonal block. From (2.6) setting $\alpha = 0$ we have

$$|b_{21}|^2 - |b_{12}|^2 = (|\delta_1|^2 + |\delta_2|^2)(\zeta^2 - \omega^2).$$

Substituting in previous equality relations (3.2), where the dependence of b_{12} and b_{21} from k is made explicit, we get

$$|k|^2 - k\frac{x_2}{(\zeta^2 - \omega^2)} - \bar{k}\frac{\bar{x}_2}{(\zeta^2 - \omega^2)} - (|\delta_1|^2 + |\delta_2|^2) = 0, \tag{4.1}$$

where x_2 is independent of k. Setting $c = x_2/(\zeta^2 - \omega^2)$, a solution can be obtained expressing k and c in polar form, i.e. $k = |k|\exp^{i\theta_k}$, $c = |c|\exp^{i\phi}$. We get

$$|k|^2 - 2|c||k|\cos(\theta_k + \phi) - (|\delta_1|^2 + |\delta_2|^2) = 0.$$

For any value of θ_k, the previous equation in $|k|$ has always two real roots, one of which is positive. For all these solutions we obtain a normal matrix. In particular, by choosing $\theta_{\tilde{k}} = \pi/2 - \phi$, and $|\tilde{k}| = \sqrt{|\delta_1|^2 + |\delta_2|^2}$, we have that the corresponding $A^{(\tilde{k})}$ is normal and is a rank-one perturbation of A. With the choice of $h = k - \tilde{k}$, we get the thesis.
□

5. Conclusions. In this paper we studied almost normal matrix, i.e. those matrices for which there exists a rank-one perturbation C such that $A^HA - AA^H = CA - AC$ and showed how to transform them to a condensed representation. This condensed form allows to represent almost normal matrices by means of $O(n)$ parameters compared with the $O(n^2)$ representation of the general Hessenberg form.

The approach used is essentially theoretical and has some link with previous condensed representations for normal matrices [7] and low-rank perturbations of normal matrices [12].

These results allow us to explicitly construct matrices in the class, with a recursive argument, bordering smaller matrices with suitable entries.

We are also able to recognize if a matrix A for which $A^H A - AA^H$ is a rank-two matrix is almost normal, looking at the structure of the two outer rows and columns, and applying Theorem 2.2 and Theorem 2.5. This can be verified with a finite procedure.

In [7] and in [3, 12, 20] the reduction to other condensed forms is obtained using generalized Krylov subspace methods. On the contrary, our reduction is obtained applying unitary transformations which diagonalize rank-two matrices (which is a finite procedure) and phase transformations which guarantee that some of the parameters are real.

Another interesting problem is that of designing QR-like algorithms which can take advantage from the block tridiagonal structure as it happens when other rank-structured matrices are considered [4].

As underlined in the introduction, a rank one perturbation of a normal matrix is not, in general, in our class. We intend to study the structure of these rank-one perturbation of normal matrices to see if we can discover structures and have some interesting computational advantages.

Appendix. LEMMA 5.1. *Let* $\mathbf{x}, \mathbf{y}, \mathbf{u}$ *and* $\mathbf{v}$ *be* $n-$*vectors such that*

$$\mathbf{x}\mathbf{y}^H = \mathbf{u}\mathbf{v}^H.$$

If $\mathbf{x} \neq 0$*, then there exists a constant* μ*, such that* $\mathbf{y}^H = \mu \mathbf{v}^H$.

Proof. By multiplying both sides of the equality by $\mathbf{x}^H$ on the left. We get

$$(\mathbf{x}^H \mathbf{x})\mathbf{y}^H = (\mathbf{x}^H \mathbf{u})\mathbf{v}^H,$$

hence setting $\mu = (\mathbf{x}^H \mathbf{u})/\|\mathbf{x}\|^2$ we get $\mathbf{y}^H = \mu \mathbf{v}^H$. $\square$

REFERENCES

[1] T. Barth. *Implementation of the Conjugate Gradient Algorithms using short Multiple Recursions*. Ph.D. Thesis, Center for Computational Mathematics. University of Colorado at Denver, 1996.

[2] T. Barth and T. Manteuffel. Multiple recursion conjugate gradient algorithms part i: Sufficient conditions. *SIAM J. Matrix Anal. Appl.*, 21(3):768–796, February 2000.

[3] R. Bevilacqua, E. Bozzo, and G. M. Del Corso. Transformations to rank structures by unitary similarity. *Linear Algebra and Its Applications*, 402(1-3):126–134, 2005.

[4] R. Bevilacqua, E. Bozzo, and G. M. Del Corso. Qd-type methods for quasiseparable matrices. *SIAM Journal on Matrix Analysis and Applications*, 32(3):723–747, 2011.

[5] D. A. Bini, Y. Eidelman, L. Gemignani, and I. Gohberg. Fast qr eigenvalue algorithms for hessenberg matrices which are rank-one perturbations of unitary matrices. *SIAM J. Matrix Anal. Appl.*, 29(2):566–585, March 2007.

[6] G. M. Del Corso and R. Bevilacqua. On a class of matrices deviating from normality. Structured Numerical Linear and Multilinear Algebra Problems: Analysis, Algorithms, and Applications, Leuven, Belgium, September 2012.

[7] L. Elsner and Kh. D. Ikramov. On a condensed form for normal matrices under finite sequences of elementary unitary similarities. *Lin. Alg. its Appl.*, 254:79–98, 1997.

[8] L. Elsner and Kh. D. Ikramov. Normal matrices: an update. *Lin. Alg. and its Appl.*, 285:291–303, 1998.

[9] V. Faber, J. Liesen, and P. Tichý. The Faber-Manteuffel theorem for linear operators. *SIAM J. Numer. Anal.*, 46(3):1323–1337, 2008.

[10] V. Faber, J. Liesen, and P. Tichý. On orthogonal reduction to hessenberg form with small bandwidth. *Numerical Algorithms*, 51(2):133–142, 2009.

[11] V. Faber and T. Manteuffel. Necessary and sufficient conditions for the existence of a conjugate gradient method. *SIAM J. Numer. Anal.*, 21(2):352–362, 1984.

[12] M. Ghasemi Kamalvand and Kh. D. Ikramov. Low-rank perturbations of normal and conjugate-normal matrices and their condensed forms under unitary similarity and congruences. *Moscow University Computational Mathematics and Cybernetics*, 33(3):109–116, 2009.

[13] R. Grone, C. R. Johnosn, E. M. Sa, and H. Wolkowicz. Normal matrices. *Lin. Alg. and its Appl.*, 87:213–225, 1987.

[14] Kh. D. Ikramov and L. Elsner. On normal matrices with normal principal submatrices. *Journal of Mathematical Science*, 89(6):1631–1651, 1998.

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

[16] W. Kahan. Spectra of nearly hermitain matrices. *Proc. Amer. Math. Soc.*, 48:11–17, 1975.

[17] A. Ruhe. On the closeness of eigenvalues and singular values for almost normal matrices. *Lin. Alg. and its Appl.*, 11(1):87–93, 1975.

[18] J. Sun. On the variation of the specrtum of a normal matrix. *Lin. Alg. and its Appl.*, 246:215–223, 1996.

[19] R. Vandebril and G. M. Del Corso. An implicit multishift *qr*-algorithm for hermitian plus low rank matrices. *SIAM J. Sci. Comp.*, 16(4):2190–2212, 2010.

[20] R. Vandebril and G. M. Del Corso. A unification of unitary similarity transforms to compressed representations. *Numerische Mathematik*, 119(4):641–665, 2011.

[21] N. A. Wiegmann. Quasi-normal matrices and products. *J. Austral. Math. Soc.*, 11:329–339, 1970.

[22] L. Gemignani Y. Eidelman and I. C. Gohberg. Efficient eigenvalue computation for quasiseparable Hermitian matrices under low rank perturbation. *Numerical Algorithms*, 47(3):253–273, 2008.