Operators
and
Matrices
Volume 3, Number 3 (2009), 401–438

MATRICES WITH NORMAL DEFECT ONE

DMITRY S. KALIUZHNYI-VERBOVETSKYI, ILYA M. SPITKOVSKY
AND HUGO J. WOERDEMAN

(*Communicated by L. Rodman*)

Abstract. A $n \times n$ matrix A has normal defect one if it is not normal, however can be embedded as a north-western block into a normal matrix of size $(n+1)\times(n+1)$. The latter is called a minimal normal completion of A. A construction of all matrices with normal defect one is given. Also, a simple procedure is presented which allows one to check whether a given matrix has normal defect one, and if this is the case — to construct all its minimal normal completions. A characterization of the generic case for each n under the assumption $\operatorname{rank}(A^*A - AA^*) = 2$ (which is necessary for A to have normal defect one) is obtained. Both the complex and the real cases are considered. It is pointed out how these results can be used to solve the minimal commuting completion problem in the classes of pairs of $n \times n$ Hermitian (resp., symmetric, or symmetric/antisymmetric) matrices when the completed matrices are sought of size $(n+1)\times(n+1)$. An application to the $2 \times n$ separability problem in quantum computing is described.

1. Introduction

A matrix $N \in \mathbb{C}^{n\times n}$ is called *normal* if $N^*N = NN^*$. For a non-normal $A \in \mathbb{C}^{n\times n}$ it is natural to inquire what is the smallest $p \in \mathbb{N}$ such that A has a *normal completion*

$$\begin{bmatrix} A & * \\ * & * \end{bmatrix} \in \mathbb{C}^{(n+p)\times(n+p)}. \tag{1.1}$$

This smallest p is called the *normal defect of* A, denoted $\operatorname{nd}(A)$, and a normal completion of size $(n+\operatorname{nd}(A))\times(n+\operatorname{nd}(A))$ is called *minimal*.

The normal completion problem as above was introduced in [12], and it was observed there that among completions (1.1) there exist those being scalar multiples of a unitary matrix. The smallest value of p required for such a completion is called the *unitary defect of* A, denoted $\operatorname{ud}(A)$, and the corresponding completions are called *minimal unitary completions of* A. In fact, $\operatorname{ud}(A)$ is simply the number (counting the multiplicities) of the singular values of A different from $\|A\|$, and is therefore strictly less than n. Moreover, it was shown in [12] how a minimal unitary completion of A can be constructed using the singular value decomposition (SVD) of A. Obviously, $\operatorname{nd}(A) \leqslant \operatorname{ud}(A)$.

Mathematics subject classification (2000): 15A57; 47A20; 47B15; 81P68.
Keywords and phrases: normal defect, minimal normal completions, complex symmetric operators and matrices, unitary and symmetric extensions, commuting completions, separability problem.
I. M. Spitkovsky and H. J. Woerdeman are partially supported by NSF grants DMS-0456625 and DMS-0500678, respectively.

Paper OaM-03-24

It is easy to find examples of matrices with $\mathrm{nd}(A) < \mathrm{ud}(A)$. For instance, if A is normal and not a multiple of a unitary matrix then $\mathrm{nd}(A) = 0 < \mathrm{ud}(A)$. However, in all such examples known until recently, the matrix A was *unitarily reducible*, that is, unitarily similar to a block diagonal matrix with more than one block. It is then natural to ask (see [12] and a further discussion in [8]) whether the equality $\mathrm{nd}(A) = \mathrm{ud}(A)$ holds for all unitarily irreducible matrices $A \in \mathbb{C}^{n\times n}$. We will show in Examples 2.23 and 2.24 that this question has a negative answer.

In this paper, we study matrices with normal defect one and their minimal normal completions. We notice some appearances of these matrices in the literature: the relations between the spectra of such matrices and of their minimal normal completions, along with applications to the location problem for the roots of a complex polynomial and its derivative, were considered in [9]; it was shown in [8] that weighted shift matrices with unimodular weights have normal defect one.

All matrices A with normal defect one must satisfy

$$\mathrm{rank}(A^*A - AA^*) = 2 \tag{1.2}$$

(see Corollary 2.6), which throughout the paper will be referred to as the *rank condition*. The manifold of $n\times n$ matrices satisfying (1.2) will be denoted $\mathfrak{M}_n$.

For 2×2 matrices, the unitary defect is at most 1. Therefore, any non-normal matrix A of size 2×2 has normal defect 1, and the (necessary) rank condition is also sufficient for $\mathrm{nd}(A) = 1$. The sufficiency of the rank condition takes also place for 3×3 matrices (Corollary 2.11), while for larger matrices it is not always the case, see Example 2.12.

In Section 2, we obtain several equivalent characterizations of matrices $A \in \mathbb{C}^{n\times n}$ with $\mathrm{nd}(A) = 1$. One of them (Theorem 2.1) serves for construction of all matrices A with $\mathrm{nd}(A) = 1$. Another one (Theorem 2.3) is used to describe a procedure which allows one to check whether $\mathrm{nd}(A) = 1$, and if this is the case — to construct all minimal normal completions of A; see Section 2.2. Finally, the characterization in Theorem 2.4 becomes handy when solving a separability problem originated in quantum computing; see description of Section 5 below. Section 2.3 provides a further analysis which allows us to refine the procedure from Section 2.2 and to describe the generic situation for each n under the assumption that rank condition (1.2) holds.

The minimal normal completion problem in the setting of real matrices is treated in a separate Section 3. The real counterpart of the normal defect of a matrix $A \in \mathbb{R}^{n\times n}$, denoted $\mathrm{rnd}(A)$, is defined. We show that $\mathrm{rnd}(A) = 1$ if and only if $\mathrm{nd}(A) = 1$. (The question on whether $\mathrm{rnd}(A) = \mathrm{nd}(A)$ for an arbitrary $A \in \mathbb{R}^{n\times n}$ remains open.) However, the results in the real case are not immediate consequences of their complex counterparts, and required an additional study. The real counterpart of Theorem 2.1 is obtained for matrices $A \in \mathbb{R}^{n\times n}$ of even size n only, while a construction of all real matrices with $\mathrm{rnd}(A) = 1$ in the case of odd n is left as an open problem. The real analogue of Theorem 2.3, as well as the procedure for verification that $\mathrm{rnd}(A) = 1$ and for construction of all minimal real normal completions, have a slightly different form which splits into two cases. The generic situation in each matrix dimension is also described, however in the real case the analysis happens to be more straightforward than its counterpart in the complex case.

In Section 4, we show how to restate our results from Sections 2 and 3 in terms of commuting completions of a pair of Hermitian (resp., symmetric and antisymmetric) matrices, where the completed matrices are also Hermitian (resp., symmetric and antisymmetric). The results for pairs of Hermitian matrices are used then to solve an analogous problem in the class of pairs of symmetric matrices.

In Section 5, we use the connection between the normal completion problem and the $2\times n$ separability problem, that was established in [13], to obtain Theorem 5.1 which gives easily verifiable necessary and sufficient conditions for a positive semidefinite matrix $M\in\mathbb{C}^{2n\times 2n}$ with a rank one Schur complement to be $2\times n$ separable. Moreover, a new proof is given for the result by Woronowicz [14] (see Theorem 5.2), which establishes, for $n\leqslant 3$, the $2\times n$ separability for a positive semidefinite matrix $M\in\mathbb{C}^{2n\times 2n}$ satisfying the Peres test.

2. The complex case

2.1. Construction of matrices with normal defect one.

THEOREM 2.1. *Let $A\in\mathbb{C}^{n\times n}$ be not normal. The following statements are equivalent:*

(i) $\mathrm{nd}(A)=1$.

(ii) *There exist a contraction matrix $C\in\mathbb{C}^{n\times n}$ with $\mathrm{ud}(C)=1$, a diagonal matrix $D\in\mathbb{C}^{n\times n}$, and a scalar $\mu\in\mathbb{C}$ such that*

$$A=CDC^{*}+\mu I_n. \tag{2.1}$$

(iii) *There exist a unitary matrix $V\in\mathbb{C}^{n\times n}$, a normal matrix $N\in\mathbb{C}^{n\times n}$, and scalars $t\colon 0\leqslant t<1$, $\mu\in\mathbb{C}$ such that*

$$V^{*}AV=MNM+\mu I_n, \tag{2.2}$$

where $M=\mathrm{diag}(1,\dots,1,t)$.

Proof. $(i)\Longleftrightarrow(ii)$ Let $\mathrm{nd}(A)=1$, and let $\begin{bmatrix}A & x\\ y^{*} & z\end{bmatrix}\in\mathbb{C}^{(n+1)\times(n+1)}$ be a minimal normal completion of A. Then there exist a diagonal matrix $\Lambda\in\mathbb{C}^{n\times n}$, a scalar $\mu\in\mathbb{C}$, and a unitary matrix $U=\begin{bmatrix}U_{11} & U_{12}\\ U_{21} & U_{22}\end{bmatrix}\in\mathbb{C}^{(n+1)\times(n+1)}$ such that

$$\begin{bmatrix}A & x\\ y^{*} & z\end{bmatrix}=\begin{bmatrix}U_{11} & U_{12}\\ U_{21} & U_{22}\end{bmatrix}\begin{bmatrix}\Lambda & 0\\ 0 & \mu\end{bmatrix}\begin{bmatrix}U_{11}^{*} & U_{21}^{*}\\ U_{12}^{*} & \overline{U}_{22}\end{bmatrix}. \tag{2.3}$$

The latter equality is equivalent to the following system:

$$\begin{aligned}
A &= U_{11}\Lambda U_{11}^{*}+\mu U_{12}U_{12}^{*}=U_{11}(\Lambda-\mu I_n)U_{11}^{*}+\mu I_n, &(2.4)\\
x &= U_{11}\Lambda U_{21}^{*}+\mu U_{12}\overline{U}_{22}=U_{11}(\Lambda-\mu I_n)U_{21}^{*}, &(2.5)\\
y^{*} &= U_{21}\Lambda U_{11}^{*}+\mu U_{22}U_{12}^{*}=U_{21}(\Lambda-\mu I_n)U_{11}^{*}, &(2.6)\\
z &= U_{21}\Lambda U_{21}^{*}+\mu U_{22}\overline{U}_{22}=U_{21}(\Lambda-\mu I_n)U_{21}^{*}+\mu. &(2.7)
\end{aligned}$$

Setting $C = U_{11}$ and $D = \Lambda - \mu I_n$, we obtain (2.1) from (2.4).

Conversely, if (2.1) holds, we set $U_{11} = C$, $\Lambda = D + \mu I_n$ and obtain (2.4). For $U = \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix}$ a minimal unitary completion of C, we define $x, y \in \mathbb{C}^n$ and $z \in \mathbb{C}$ by (2.5)–(2.7). Then (2.3) holds, i.e., the matrix $\begin{bmatrix} A & x \\ y^* & z \end{bmatrix} \in \mathbb{C}^{(n+1)\times(n+1)}$ is a normal completion of A, and thus $\operatorname{nd} A = 1$.

(ii)$\Longleftrightarrow$*(iii)* If *(ii)* holds, let $C = V \operatorname{diag}(1,\dots,1,t)W^*$ be the SVD of C (here $V, W \in \mathbb{C}^{n\times n}$ are unitary, $0 \leqslant t < 1$, and $M = \operatorname{diag}(1,\dots,1,t) \in \mathbb{C}^{n\times n}$). Then, clearly, $N = W^* D W$ is normal, and (2.2) follows.

Conversely, if (2.2) holds, then $N = W^* D W$ with D diagonal and W unitary. Clearly, for $C = V \operatorname{diag}(1,\dots,1,t)W^*$ we have $\operatorname{ud}(C) = 1$, and (2.1) follows. $\square$

REMARK 2.2. Observe that the matrix A given by (2.2) happens to be normal if and only if the product MNM is normal, that is

$$MNM^2N^*M = MN^*M^2NM. \tag{2.8}$$

Since N itself is normal, (2.8) holds if and only if

$$MNZN^*M = MN^*ZNM, \tag{2.9}$$

where $Z = \operatorname{diag}(0,\dots,0,1)$. Partitioning N as

$$N = \begin{bmatrix} N_0 & g \\ h^* & \alpha \end{bmatrix},$$

where α is scalar, and rewriting (2.9) block-wise, we see that it is equivalent to

$$gg^* = hh^*, \quad t\alpha h = t\overline{\alpha} g.$$

These conditions mean simply that g differs from h by a scalar multiple of absolute value one and, if $t\alpha \neq 0$, this scalar must be $\alpha/\overline{\alpha}$. Consequently, A is not normal if and only if this is not the case.

Observe also that if $t \neq 0$ (so that M is invertible) and N is also invertible, then (2.8) can be written as

$$M^2N^*N^{-1} = N^{-1}N^*M^2. \tag{2.10}$$

But N is normal, so that N^* commutes with N^{-1}. Condition (2.10) therefore means simply that N^*N^{-1} $(= N^{-1}N^*)$ commutes with M^2. In other words, A in this case is normal if and only if $e_n := \operatorname{col}(0,\dots,0,1)$ is an eigenvector of N^*N^{-1}. In turn, this happens if and only if e_n belongs to the sum of eigenspaces of N with the corresponding eigenvalues lying on the same line through the origin.

Representation (2.1) or (2.2) in Theorem 2.1, together with Remark 2.2, allow one to construct all matrices A with $\operatorname{nd}(A) = 1$. However, as we mentioned in Section 1, this does not give an easy way to check whether a given matrix has normal defect one. A procedure for this is our further goal.

2.2. Identification of matrices with $\mathrm{nd}(A) = 1$ and construction of all minimal normal completions of A

In the following two theorems, we establish necessary and sufficient conditions for a matrix A to have normal defect one, and for any matrix A with $\mathrm{nd}(A) = 1$ we describe all its minimal normal completions. Here and throughout the rest of the paper, we set $\mathbb{T} = \{z \in \mathbb{C} \colon |z| = 1\}$.

THEOREM 2.3. *Let $A \in \mathbb{C}^{n\times n}$. Then*

(i) $\mathrm{nd}(A) = 1$ *if and only if* $\mathrm{rank}(A^*A - AA^*) = 2$ *and the equation*

$$PA^*(x_1u_1 + x_2u_2) = PA(\overline{x}_2u_1 + \overline{x}_1u_2) \tag{2.11}$$

has a solution pair $x_1, x_2 \in \mathbb{C}$ satisfying

$$|x_1|^2 - |x_2|^2 = d. \tag{2.12}$$

Here $u_1, u_2 \in \mathbb{C}^n$ are the unit eigenvectors of the matrix $A^*A - AA^*$ corresponding to its nonzero eigenvalues $\lambda_1 = d (> 0)$ and $\lambda_2 = -d$, and

$$P = I_n - u_1u_1^* - u_2u_2^*$$

is the orthogonal projection of $\mathbb{C}^n$ onto $\mathrm{null}(A^*A - AA^*)$.

(ii) *If* $\mathrm{nd}(A) = 1$, *x_1 and x_2 satisfy* (2.11) *and* (2.12), *and $\mu \in \mathbb{T}$ is arbitrary then the matrix*

$$B = \begin{bmatrix} A & \mu(x_1u_1 + x_2u_2) \\ \overline{\mu}(x_2u_1^* + x_1u_2^*) & z \end{bmatrix} \tag{2.13}$$

is a minimal normal completion of A. Here

$$z = a_{11} - \frac{1}{d}\left(x_2(a_{12}\overline{x}_1 - \overline{a}_{21}x_2) + x_1(\overline{a}_{12}x_1 - a_{21}\overline{x}_2)\right) \tag{2.14}$$

and

$$a_{11} = u_1^*Au_1, \quad a_{12} = u_1^*Au_2, \quad a_{21} = u_2^*Au_1. \tag{2.15}$$

All minimal normal completions of A arise in this way.

THEOREM 2.4. *Let $A \in \mathbb{C}^{n\times n}$. Then* $\mathrm{nd}(A) = 1$ *if and only if there exist linearly independent $x, y \in \mathbb{C}^n$ such that*

$$A^*A - AA^* = xx^* - yy^* \tag{2.16}$$

and the vectors x, y, A^*x, Ay are linearly dependent. In this case, there exist $z \in \mathbb{C}$ and $\nu \in \mathbb{T}$ such that the matrix

$$B = \begin{bmatrix} A & \nu x \\ y^* & z \end{bmatrix} \tag{2.17}$$

is normal.

In order to prove Theorems 2.3 and 2.4 we will need several auxiliary statements.

LEMMA 2.5. *Let $A \in \mathbb{C}^{n\times n}$. Then* $\mathrm{nd}(A) = 1$ *if and only if there exist linearly independent vectors $x, y \in \mathbb{C}^n$ satisfying* (2.16) *and a scalar $z \in \mathbb{C}$ such that*

$$(A - zI_n)^* x = (A - zI)y. \tag{2.18}$$

Proof. If $\mathrm{nd}(A) = 1$ then there exists a normal matrix $B = \begin{bmatrix} A & x \\ y^* & z \end{bmatrix} \in \mathbb{C}^{(n+1)\times(n+1)}$. The identity $B^*B = BB^*$ is equivalent to (2.16)&(2.18) (the identity $x^*x = y^*y$ follows from (2.16) since $\mathrm{trace}(A^*A - AA^*) = 0$, and is therefore redundant). Clearly, x and y are linearly independent, otherwise the right-hand side of (2.16) is 0 and A is normal.

Conversely, if $x, y \in \mathbb{C}^n$ are linearly independent, $z \in \mathbb{C}$, and (2.16)&(2.18) hold then the matrix $B = \begin{bmatrix} A & x \\ y^* & z \end{bmatrix} \in \mathbb{C}^{(n+1)\times(n+1)}$ is normal. Since the right-hand side of (2.16) is not 0, the matrix A is not normal, thus $\mathrm{nd}(A) = 1$. $\square$

COROLLARY 2.6. *If* $\mathrm{nd}(A) = 1$ *then* $\mathrm{rank}(A^*A - AA^*) = 2$.

If the rank condition is satisfied, one can find the unit eigenvectors u_1 and u_2 of the matrix $A^*A - AA^*$ corresponding to its eigenvalues $\lambda_1 = d (> 0)$ and $\lambda_2 = -d$, which are determined uniquely up to a scalar factor. There is more freedom in a choice of other eigenvectors $u_3, \ldots, u_n$, which form an orthonormal basis of $\mathrm{null}(A^*A - AA^*)$. Suppose that such vectors are chosen. Then $U = \begin{bmatrix} u_1 & u_2 & u_3 & \ldots & u_n \end{bmatrix} \in \mathbb{C}^{n\times n}$ is a unitary matrix, and the matrix $\widetilde{A} = U^*AU$ satisfies

$$\widetilde{A}^*\widetilde{A} - \widetilde{A}\widetilde{A}^* = H, \tag{2.19}$$

where

$$H = \mathrm{diag}(d, -d, 0, \ldots, 0) \in \mathbb{C}^{n\times n}. \tag{2.20}$$

LEMMA 2.7. *If $\widetilde{A} \in \mathbb{C}^{n\times n}$ satisfies* (2.19) *then $\widetilde{A}$ has the form*

$$\widetilde{A} = \begin{bmatrix} a_{11} & a_{12} & u \\ a_{21} & a_{22} & v \\ w^* & q^* & S \end{bmatrix},$$

where a_{ij} ($i, j = 1, 2$) are scalars,

$$a_{11} = a_{22}, \tag{2.21}$$

and $u^*, v^*, w^*, q^* \in \mathbb{C}^{n-2}$ satisfy

$$uu^* = qq^*, \quad vv^* = ww^*, \quad uv^* = wq^*, \quad uw^* = vq^*. \tag{2.22}$$

Proof. We have

$$\operatorname{trace}(\widetilde{A}H) = \operatorname{trace}(\widetilde{A}\widetilde{A}^*\widetilde{A} - \widetilde{A}^2\widetilde{A}^*) = 0,$$

which implies (2.21). Then, from the equality

$$(\widetilde{A}^*\widetilde{A} - \widetilde{A}\widetilde{A}^*)_{12} = 0$$

we obtain that $uv^* = wq^*$. Next, from the observation

$$\begin{aligned}\operatorname{trace}(\widetilde{A}^*H\widetilde{A} + \widetilde{A}H\widetilde{A}^*) &= \operatorname{trace}(\widetilde{A}^*(\widetilde{A}^*\widetilde{A} - \widetilde{A}\widetilde{A}^*)\widetilde{A} + \widetilde{A}(\widetilde{A}^*\widetilde{A} - \widetilde{A}\widetilde{A}^*)\widetilde{A}^*) \\ &= \operatorname{trace}(\widetilde{A}^{*2}\widetilde{A}^2 - \widetilde{A}^*\widetilde{A}\widetilde{A}^*\widetilde{A} + \widetilde{A}\widetilde{A}^*\widetilde{A}\widetilde{A}^* - \widetilde{A}^2\widetilde{A}^{*2}) = 0\end{aligned}$$

we obtain

$$uu^* - vv^* + ww^* - qq^* = 0, \tag{2.23}$$

and from the equality

$$\operatorname{trace}((\widetilde{A}^*\widetilde{A} - \widetilde{A}\widetilde{A}^*)_{33}) = 0$$

(where the 33-block is of size $(n-2)\times(n-2)$) we obtain

$$uu^* + vv^* = ww^* + qq^*. \tag{2.24}$$

Combining (2.23) and (2.24), we obtain that $uu^* = qq^*$ and $vv^* = ww^*$. From the observation

$$\begin{aligned}\operatorname{trace}(\widetilde{A}H\widetilde{A}) &= \operatorname{trace}(\widetilde{A}(\widetilde{A}^*\widetilde{A} - \widetilde{A}\widetilde{A}^*)\widetilde{A}) \\ &= \operatorname{trace}(\widetilde{A}\widetilde{A}^*\widetilde{A}^2 - \widetilde{A}^2\widetilde{A}^*\widetilde{A}) = 0\end{aligned}$$

we obtain that $uw^* = vq^*$. □

LEMMA 2.8. *Suppose that*

$$\widetilde{A} = \begin{bmatrix} a_{11} & a_{12} & u \\ a_{21} & a_{11} & v \\ w^* & q^* & S \end{bmatrix} \in \mathbb{C}^{n\times n} \tag{2.25}$$

satisfies (2.19) *with* H *as in* (2.20). *Then the matrix* $\widetilde{B} = \begin{bmatrix} \widetilde{A} & x \\ y^* & z \end{bmatrix} \in \mathbb{C}^{(n+1)\times(n+1)}$ *is normal if and only if*

$$x = \operatorname{col}(x_1, x_2, 0, \ldots, 0) \in \mathbb{C}^n, \quad y = \operatorname{col}(y_1, y_2, 0, \ldots, 0) \in \mathbb{C}^n, \tag{2.26}$$

$$|x_1|^2 - |x_2|^2 = d, \tag{2.27}$$

$$y_1 = e^{i\theta}\overline{x}_2, \quad y_2 = e^{i\theta}\overline{x}_1, \tag{2.28}$$

for some $\theta \in \mathbb{R}$, *and the following identities hold:*

$$(\overline{a}_{11} - \overline{z})x_1 + \overline{a}_{21}x_2 = (a_{11} - z)y_1 + a_{12}y_2, \tag{2.29}$$

$$\overline{a}_{12}x_1 + (\overline{a}_{11} - \overline{z})x_2 = a_{21}y_1 + (a_{11} - z)y_2, \tag{2.30}$$

$$u^*x_1 + v^*x_2 = w^*y_1 + q^*y_2. \tag{2.31}$$

Proof. It follows from Lemma 2.5 applied to the matrix $\widetilde{A}$ as above (see also Lemma 2.7 which justifies that $a_{22} = a_{11}$) that $\mathrm{nd}(\widetilde{A}) = 1$ and $\widetilde{B} = \begin{bmatrix} \widetilde{A} & x \\ y^* & z \end{bmatrix} \in \mathbb{C}^{(n+1)\times(n+1)}$ is a minimal normal completion of $\widetilde{A}$ if and only if x and y are linearly independent and (2.16)&(2.18) hold with A replaced by $\widetilde{A}$. Since in this case $H = xx^* - yy^*$, the vectors x and y have the form (2.26). Indeed, for any vector $h \in \mathbb{C}^n$ which is orthogonal to y and not orthogonal to x, we have

$$0 \neq Hh = xx^*h \in \mathrm{range}(H) \cap \mathrm{span}(x).$$

Similarly, for any vector $g \in \mathbb{C}^n$ which is orthogonal to x and not orthogonal to y, we have

$$0 \neq Hg = -yy^*g \in \mathrm{range}(H) \cap \mathrm{span}(y),$$

thus both x and y are in $\mathrm{range}(H)$. Next, the identity $H = xx^* - yy^*$ holds if and only if

$$|x_1|^2 - |x_2|^2 = d = |y_2|^2 - |y_1|^2, \quad x_1\overline{x}_2 = y_1\overline{y}_2,$$

or equivalently, (2.28) holds with some $\theta \in \mathbb{R}$. Clearly, (2.18) with A replaced by $\widetilde{A}$, is equivalent to (2.29)–(2.31). □

REMARK 2.9. If $\widetilde{A}$ is as in Lemma 2.8 and $\widetilde{B} = \begin{bmatrix} \widetilde{A} & x \\ y^* & z \end{bmatrix} \in \mathbb{C}^{(n+1)\times(n+1)}$ is a minimal normal completion of $\widetilde{A}$ then so is

$$\begin{bmatrix} \widetilde{A} & \mu x \\ \overline{\mu} y^* & z \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ 0 & \overline{\mu} \end{bmatrix} \begin{bmatrix} \widetilde{A} & x \\ y^* & z \end{bmatrix} \begin{bmatrix} I_n & 0 \\ 0 & \mu \end{bmatrix}$$

for any $\mu \in \mathbb{T}$. Therefore, if x, y and z are as in Lemma 2.8 then the matrix

$$\begin{bmatrix} a_{11} & a_{12} & u & e^{-i\theta/2}x_1 \\ a_{21} & a_{11} & v & e^{-i\theta/2}x_2 \\ w^* & q^* & S & 0 \\ e^{-i\theta/2}x_2 & e^{-i\theta/2}x_1 & 0 & z \end{bmatrix}$$

is a minimal normal completion of $\widetilde{A}$. This observation leads to the following statement.

LEMMA 2.10. *Suppose that*

$$\widetilde{A} = \begin{bmatrix} a_{11} & a_{12} & u \\ a_{21} & a_{11} & v \\ w^* & q^* & S \end{bmatrix} \in \mathbb{C}^{n\times n}$$

satisfies (2.19) *with* H *as in* (2.20)*. Then* $\mathrm{nd}(A) = 1$ *if and only if there exist* $x_1, x_2 \in \mathbb{C}$ *satisfying* (2.12) *and such that*

$$u^*x_1 + v^*x_2 = w^*\overline{x}_2 + q^*\overline{x}_1. \tag{2.32}$$

In this case, the matrix

$$\begin{bmatrix} a_{11} & a_{12} & u & x_1 \\ a_{21} & a_{11} & v & x_2 \\ w^* & q^* & S & 0 \\ x_2 & x_1 & 0 & z \end{bmatrix}, \tag{2.33}$$

where z *is given by* (2.14)*, is a minimal normal completion of* $\widetilde{A}$.

Proof. The statement follows from Lemma 2.8, Remark 2.9, and the observation that, if $y_1 = \overline{x}_2$ and $y_2 = \overline{x}_1$, then (2.31) becomes (2.32). Solving (2.29) or (2.30) (which are equivalent in this case) for z gives (2.14). $\square$

Proof of Theorem 2.3. (i) By Corollary 2.6 the rank condition, $\operatorname{rank}(A^*A - AA^*) = 2$, is necessary for A to have normal defect one, thus we can assume that this condition holds. Let u_1 and u_2 be the unit eigenvectors of the matrix $A^*A - AA^*$ corresponding to its eigenvalues $\lambda_1 = d(>0)$ and $\lambda_2 = -d$, and let $u_3, \ldots, u_n$ be an arbitrary orthonormal basis of $\operatorname{null}(A^*A - AA^*)$. Define a unitary matrix $U = \begin{bmatrix} u_1 & \ldots & u_n \end{bmatrix} \in \mathbb{C}^{n\times n}$ and an isometry $U' = \begin{bmatrix} u_3 & \ldots & u_n \end{bmatrix} \in \mathbb{C}^{n\times(n-2)}$. Then

$$U'U'^* = P, \tag{2.34}$$

and the matrix $\widetilde{A} = U^*AU$ has the form (2.25), where the scalars a_{ij} are defined by (2.15),

$$u = u_1^*AU', \quad v = u_2^*AU', \quad w^* = U'^*Au_1, \quad q^* = U'^*Au_2. \tag{2.35}$$

According to Lemma 2.10, $\operatorname{nd}(\widetilde{A}) = 1$ (and hence $\operatorname{nd}(A) = 1$) if and only if (2.32) is satisfied with some $x_1, x_2 \in \mathbb{C}$ subject to (2.12). By (2.35), equation (2.32) can be written as

$$(U'^*A^*u_1)x_1 + (U'^*A^*u_2)x_2 = (U'^*Au_1)\overline{x}_2 + (U'^*Au_2)\overline{x}_1.$$

Multiplying on the left by U' and taking into account that $U' \colon \mathbb{C}^{n-2} \to \mathbb{C}^n$ is an isometry satisfying (2.34), we obtain an equivalent equation

$$\mathbf{u}^*x_1 + \mathbf{v}^*x_2 = \mathbf{w}^*\overline{x}_2 + \mathbf{q}^*\overline{x}_1, \tag{2.36}$$

with the vectors

$$\mathbf{u}^* = PA^*u_1, \quad \mathbf{v}^* = PA^*u_2, \quad \mathbf{w}^* = PAu_1, \quad \mathbf{q}^* = PAu_2. \tag{2.37}$$

Note that these vectors are defined independently of the choice of $u_3, \ldots, u_n$. Since (2.36) is equivalent to (2.11), this proves part (i) of this theorem.

(ii) If $\operatorname{nd}(A) = 1$ and $\widetilde{A}$ is defined as in part (i), then $\operatorname{nd}(\widetilde{A}) = 1$. By Lemma 2.10, for any $x_1, x_2 \in \mathbb{C}$ satisfying (2.11) (or equivalently, (2.32)) and (2.12), the matrix in (2.33) is a minimal normal completion of $\widetilde{A}$. By Remark 2.9, so is

$$\widetilde{B} = \begin{bmatrix} a_{11} & a_{12} & u & \mu x_1 \\ a_{21} & a_{11} & v & \mu x_2 \\ w^* & q^* & S & 0 \\ \overline{\mu} x_2 & \overline{\mu} x_1 & 0 & z \end{bmatrix}, \tag{2.38}$$

with an arbitrary $\mu \in \mathbb{T}$. By Lemma 2.8, all minimal normal completions of $\widetilde{A}$ arise in this way. Since $\widetilde{A} = U^*AU$ and U is unitary, all minimal normal completions of A have the form

$$B = \begin{bmatrix} U & 0 \\ 0 & 1 \end{bmatrix} \widetilde{B} \begin{bmatrix} U^* & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} A & \mu(x_1u_1 + x_2u_2) \\ \overline{\mu}(x_2u_1^* + x_1u_2^*) & z \end{bmatrix},$$

where $\widetilde{B}$ is any of the matrices defined in (2.38). This proves part (*ii*). □

Proof of Theorem 2.4. If $\mathrm{nd}(A) = 1$ then there exists a normal matrix

$$B = \begin{bmatrix} A & x \\ y^* & z \end{bmatrix} \in \mathbb{C}^{(n+1)\times(n+1)}.$$

By Lemma 2.5, (2.16) holds and $A^*x - \overline{z}x = Ay - zy$, i.e., x, y, A^*x, Ay are linearly dependent.

Conversely, suppose that $A^*A - AA^* = xx^* - yy^*$ is satisfied with some linearly independent vectors $x, y \in \mathbb{C}^n$, and the vectors x, y, A^*x, Ay are linearly dependent. Clearly, in this case $\mathrm{rank}(A^*A - AA^*) = 2$. Choose an orthonormal eigenbasis $u_1, \ldots, u_n$ of the matrix $A^*A - AA^*$ and define a unitary matrix $U = \begin{bmatrix} u_1 & \ldots & u_n \end{bmatrix}$ as in the proof of Theorem 2.3. By Lemma 2.7, the matrix $\widetilde{A} = U^*AU$ has the form (2.25). Define new linearly independent vectors $\widetilde{x} = U^*x$, $\widetilde{y} = U^*y$. Then $\widetilde{A}^*\widetilde{A} - \widetilde{A}\widetilde{A}^* = \widetilde{x}\widetilde{x}^* - \widetilde{y}\widetilde{y}^*$. As in the proof of Lemma 2.8, we conclude that (2.26)–(2.28) hold with $\widetilde{x}$ and $\widetilde{y}$ in the place of x and y. Since x, y, A^*x, Ay are linearly dependent, so are $\widetilde{x}, \widetilde{y}, \widetilde{A}^*\widetilde{x}, \widetilde{A}\widetilde{y}$, i.e., the matrix

$$\begin{bmatrix} \widetilde{x} & \widetilde{y} & \widetilde{A}^*\widetilde{x} & \widetilde{A}\widetilde{y} \end{bmatrix} = \begin{bmatrix} \widetilde{x}_1 & \widetilde{y}_1 & \overline{a}_{11}\widetilde{x}_1 + \overline{a}_{21}\widetilde{x}_2 & a_{11}\widetilde{y}_1 + a_{12}\widetilde{y}_2 \\ \widetilde{x}_2 & \widetilde{y}_2 & \overline{a}_{12}\widetilde{x}_1 + \overline{a}_{11}\widetilde{x}_2 & a_{21}\widetilde{y}_1 + a_{11}\widetilde{y}_2 \\ 0 & 0 & u^*\widetilde{x}_1 + v^*\widetilde{x}_2 & w^*\widetilde{y}_1 + q^*\widetilde{y}_2 \end{bmatrix} \in \mathbb{C}^{n\times 4}$$

has rank less than 4. Therefore $u^*\widetilde{x}_1 + v^*\widetilde{x}_2$ and $w^*\widetilde{y}_1 + q^*\widetilde{y}_2$ are linearly dependent. The identities $\widetilde{y}_1 = e^{i\theta}\overline{\widetilde{x}_2}$ and $\widetilde{y}_2 = e^{i\theta}\overline{\widetilde{x}_1}$, together with the first three identities in (2.22), imply that there is $\phi \in \mathbb{R}$ such that

$$u^*\widetilde{x}_1 + v^*\widetilde{x}_2 = (w^*\widetilde{y}_1 + q^*\widetilde{y}_2)e^{i\phi} = (w^*\overline{\widetilde{x}_2} + q^*\overline{\widetilde{x}_1})e^{i(\theta+\phi)}.$$

Putting $x_1^0 = \widetilde{x}_1 e^{-i\frac{\theta+\phi}{2}}$ and $x_2^0 = \widetilde{x}_2 e^{-i\frac{\theta+\phi}{2}}$, we obtain

$$u^*x_1^0 + v^*x_2^0 = w^*\overline{x_2^0} + q^*\overline{x_1^0}.$$

By Lemma 2.10, $\mathrm{nd}(\widetilde{A}) = 1$, and therefore $\mathrm{nd}(A) = 1$.

The last statement of the theorem is obtained as follows. Let

$$x^0 = \mathrm{col}(x_1^0, x_2^0, 0, \ldots, 0), \quad y^0 = \mathrm{col}(\overline{x}_2^0, \overline{x}_1^0, 0, \ldots, 0),$$

and

$$z = a_{11} - \frac{1}{d}\left(x_2^0(a_{12}\overline{x}_1^0 - \overline{a}_{21}x_2^0) + x_1^0(\overline{a}_{12}x_1^0 - a_{21}\overline{x}_2^0)\right)$$

(see (2.14)). By Lemma 2.10, the matrix $\widetilde{B} = \begin{bmatrix} \widetilde{A} & x^0 \\ y^{0*} & z \end{bmatrix}$ is a minimal normal completion of $\widetilde{A}$. Then

$$B = \begin{bmatrix} U & 0 \\ 0 & e^{i\frac{\phi-\theta}{2}} \end{bmatrix} \begin{bmatrix} \widetilde{A} & x^0 \\ y^{0*} & z \end{bmatrix} \begin{bmatrix} U^* & 0 \\ 0 & e^{i\frac{\theta-\phi}{2}} \end{bmatrix} = \begin{bmatrix} A & e^{-i\phi}x \\ y^* & z \end{bmatrix}$$

is a minimal normal completion of A, i.e., we obtain (2.17) with $\nu = e^{-i\phi}$. □

Applying Theorem 2.4 to 3×3 matrices, we obtain the following.

COROLLARY 2.11. *A matrix $A \in \mathbb{C}^{3\times 3}$ has normal defect one if and only if*

$$\mathrm{rank}(A^*A - AA^*) = 2.$$

Proof. The necessity of the rank condition has been established in Corollary 2.6. The sufficiency follows from Theorem 2.4, since any four vectors in $\mathbb{C}^3$ are linearly dependent. □

In the following example, we show that for $n > 3$ the rank condition (1.2) is not sufficient for A to have normal defect one.

EXAMPLE 2.12. Let

$$A = \begin{bmatrix} 0 & 0 & 1 & -i \\ 2 & 0 & 0 & 0 \\ 0 & 1 & \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{2}} \\ 0 & -i & \frac{i}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}.$$

Note that A $(= \widetilde{A})$ is already of the form (2.25). We have

$$A^*A - AA^* = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Equation (2.32) in this case takes the form

$$x_1 \begin{bmatrix} 1 \\ i \end{bmatrix} = \overline{x}_1 \begin{bmatrix} 1 \\ -i \end{bmatrix},$$

and it has no solutions with $|x_1|^2 - |x_2|^2 = 2 > 0$. Thus, by Lemma 2.10, $\mathrm{nd}(A) > 1$.

REMARK 2.13. If $\operatorname{rank}(A^*A-AA^*)=2$ and $u_1,u_2\in\mathbb{C}^n$ are the unit eigenvectors of A^*A-AA^* corresponding to the eigenvalues $\lambda_1=d(>0)$ and $\lambda_2=-d$, then the vectors $x=\sqrt{d}u_1$ and $y=\sqrt{d}u_2$ satisfy (2.16). Indeed, u_1 and u_2 are orthogonal, hence linearly independent, and $\operatorname{span}(u_1,u_2)=\operatorname{range}(A^*A-AA^*)$. For arbitrary $a,b\in\mathbb{C}$, we have

$$(A^*A-AA^*)(au_1+bu_2)=d(au_1-bu_2)=d(u_1u_1^*-u_2u_2^*)(au_1+bu_2),$$

therefore $A^*A-AA^*=d(u_1u_1^*-u_2u_2^*)$.

However, as the following example shows, these x and y do not necessarily satisfy the conditions of Theorem 2.4.

EXAMPLE 2.14. Let

$$A=\begin{bmatrix}0&0&\frac{1}{\sqrt{2}}&\frac{i}{\sqrt{2}}\\ 0&0&1&i\\ 1&\frac{1}{\sqrt{2}}&\frac{\sqrt{3}}{2}&-\frac{\sqrt{3}}{2}i\\ i&\frac{i}{\sqrt{2}}&-\frac{\sqrt{3}}{2}i&-\frac{\sqrt{3}}{2}\end{bmatrix}.$$

Then

$$A^*A-AA^*=\begin{bmatrix}1&0&0&0\\ 0&-1&0&0\\ 0&0&0&0\\ 0&0&0&0\end{bmatrix}=e_1e_1^*-e_2e_2^*,$$

where e_1 and e_2 are standard basis vectors, which are the eigenvectors of the matrix A^*A-AA^* corresponding to its eigenvalues $\lambda_1=1$ and $\lambda_2=-1$. However, the vectors $x=e_1$, $y=e_2$, $A^*x=\operatorname{col}(0,0,\frac{1}{\sqrt{2}},-\frac{i}{\sqrt{2}})$, $Ay=\operatorname{col}(0,0,\frac{1}{\sqrt{2}},\frac{i}{\sqrt{2}})$ are linearly independent. On the other hand, we have $\operatorname{nd}(A)=\operatorname{ud}(A)=1$: one of minimal normal completions of A (in fact, its minimal unitary completion) is

$$B=\begin{bmatrix}0&0&\frac{1}{\sqrt{2}}&\frac{i}{\sqrt{2}}&\sqrt{2}\\ 0&0&1&i&-1\\ 1&\frac{1}{\sqrt{2}}&\frac{\sqrt{3}}{2}&-\frac{\sqrt{3}}{2}i&0\\ i&\frac{i}{\sqrt{2}}&-\frac{\sqrt{3}}{2}i&-\frac{\sqrt{3}}{2}&0\\ -1&\sqrt{2}&0&0&0\end{bmatrix}.$$

We are able now to describe a procedure to determine whether $\operatorname{nd}(A)=1$ for a given matrix $A\in\mathbb{C}^{n\times n}$, i.e., whether equation (2.11) in Theorem 2.3 has a solution pair $x_1,x_2\in\mathbb{C}$ satisfying (2.12). Moreover, this procedure allows one to find all such solutions, and then, applying part (*ii*) of Theorem 2.3, all minimal normal completions of an arbitrary matrix A with $\operatorname{nd}(A)=1$.

The procedure

Begin

Step 1. Verify the condition $\operatorname{rank}(A^*A - AA^*) = 2$. If it is satisfied – go to Step 2. Otherwise, stop: $\operatorname{nd}(A) > 1$.

Step 2. Rewrite (2.11) in the form (2.36), where $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$, $\mathbf{q}$ are defined in (2.37) (see Theorem 2.3 for the definition of P, u_1, and u_2). Let

$$\mathbf{u} = \mathbf{u}_R + i\mathbf{u}_I,\ \mathbf{v} = \mathbf{v}_R + i\mathbf{v}_I,\ \mathbf{q} = \mathbf{q}_R + i\mathbf{q}_I,\ \mathbf{w} = \mathbf{w}_R + i\mathbf{w}_I,$$

where $\mathbf{u}_R^T, \mathbf{u}_I^T, \mathbf{v}_R^T, \mathbf{v}_I^T, \mathbf{q}_R^T, \mathbf{q}_I^T, \mathbf{w}_R^T, \mathbf{w}_I^T \in \mathbb{R}^n$, and let

$$x_1 = x_{R1} + ix_{I1}, \quad x_2 = x_{R2} + ix_{I2},$$

where $x_{R1}, x_{I1}, x_{R2}, x_{I2} \in \mathbb{R}$. Then (2.36) becomes

$$\begin{bmatrix} \mathbf{u}_R^T - \mathbf{q}_R^T & \mathbf{u}_I^T + \mathbf{q}_I^T & \mathbf{v}_R^T - \mathbf{w}_R^T & \mathbf{v}_I^T + \mathbf{w}_I^T \\ -\mathbf{u}_I^T + \mathbf{q}_I^T & \mathbf{u}_R^T + \mathbf{q}_R^T & -\mathbf{v}_I^T + \mathbf{w}_I^T & \mathbf{v}_R^T + \mathbf{w}_R^T \end{bmatrix} \begin{bmatrix} x_{R1} \\ x_{I1} \\ x_{R2} \\ x_{I2} \end{bmatrix} = 0. \tag{2.39}$$

Denote

$$\mathbf{Q} = \begin{bmatrix} \mathbf{u}_R^T - \mathbf{q}_R^T & \mathbf{u}_I^T + \mathbf{q}_I^T & \mathbf{v}_R^T - \mathbf{w}_R^T & \mathbf{v}_I^T + \mathbf{w}_I^T \\ -\mathbf{u}_I^T + \mathbf{q}_I^T & \mathbf{u}_R^T + \mathbf{q}_R^T & -\mathbf{v}_I^T + \mathbf{w}_I^T & \mathbf{v}_R^T + \mathbf{w}_R^T \end{bmatrix}.$$

Find $m = \operatorname{rank}(\mathbf{Q})$.

Step 3. Depending on m, consider the following cases.

(1) $m = 0$. In this case, $\mathbf{u} = \mathbf{v} = \mathbf{q} = \mathbf{w} = 0$, and then (2.36) holds with any $x_1, x_2 \in \mathbb{C}$ such that $|x_1|^2 - |x_2|^2 = d$. Therefore, $\operatorname{nd}(A) = 1$. Go to Step 4.

(2) $1 \leqslant m \leqslant 3$. In this case, (2.39) has nontrivial solutions. Let $\mathbf{F} \in \mathbb{R}^{4\times(4-m)}$ be a matrix whose columns are linearly independent solutions of (2.39). Then $x = \operatorname{col}(x_{R1}, x_{I1}, x_{R2}, x_{I2}) \in \mathbb{R}^4$ is a solution of (2.39) if and only if $x = \mathbf{F}h$, with $h \in \mathbb{R}^{4-m}$. Setting $\mathbf{F} = \begin{bmatrix} \mathbf{F}_1 \\ \mathbf{F}_2 \end{bmatrix}$ where $\mathbf{F}_1, \mathbf{F}_2 \in \mathbb{R}^{2\times(4-m)}$, write $|x_1|^2 - |x_2|^2 > 0$ as

$$h^T(\mathbf{F}_1^T\mathbf{F}_1 - \mathbf{F}_2^T\mathbf{F}_2)h > 0.$$

Therefore, $\operatorname{nd}(A) = 1$ if and only if the matrix $\mathbf{K} = \mathbf{F}_1^T\mathbf{F}_1 - \mathbf{F}_2^T\mathbf{F}_2$ has at least one positive eigenvalue. If this is not the case, stop: $\operatorname{nd}(A) > 1$. Otherwise, for any h in the level hyper-surface $h^T\mathbf{K}h = d$, define

$$\begin{bmatrix} x_{R1} \\ x_{I1} \end{bmatrix} = \mathbf{F}_1 h, \quad \begin{bmatrix} x_{R2} \\ x_{I2} \end{bmatrix} = \mathbf{F}_2 h,$$

and thus obtain $x_1 = x_{R1} + ix_{I1}$, $x_2 = x_{R2} + ix_{I2}$ satisfying (2.36) and such that $|x_1|^2 - |x_2|^2 = d$. Go to Step 4.

(3) $m = 4$. In this case, (2.39), and hence (2.36), has no nontrivial solutions, and $\mathrm{nd}(A) > 1$. Stop.

Step 4. For each pair $x_1, x_2 \in \mathbb{C}$ obtained in Step 3, find minimal normal completions of A as described by (2.13)–(2.15).

End

REMARK 2.15. If $m = 1$ then $\mathbf{K}$ always has a positive eigenvalue and $\mathrm{nd}(A) = 1$. Indeed, in this case $\mathbf{F}$ is a full rank matrix of size 4×3. Since $\mathrm{null}(\mathbf{F}_2) \neq \{0\}$ and $\mathrm{null}(\mathbf{F}) = \{0\}$, for a nonzero vector $h \in \mathrm{null}(\mathbf{F}_2)$ we have

$$h^T \mathbf{K} h = h^T \mathbf{F}_1^T \mathbf{F}_1 h > 0.$$

EXAMPLE 2.16. It was shown in [8] that for any weighted shift matrix

$$A = \begin{bmatrix} 0 & a_1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & a_{n-1} \\ 0 & \dots & \dots & \dots & 0 \end{bmatrix}$$

with weights $a_j \in \mathbb{T}$ (which is, clearly, unitarily irreducible as having a single cell in its Jordan form), $\mathrm{nd}(A) = \mathrm{ud}(A) = 1$, and for $n \geqslant 4$ all its minimal normal completions B are also minimal unitary completions and have the form

$$B = \begin{bmatrix} 0 & a_1 & 0 & \dots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & & \ddots & \ddots & 0 & \vdots \\ \vdots & & & \ddots & a_{n-1} & 0 \\ 0 & \dots & \dots & \dots & 0 & \zeta \\ \rho & 0 \dots & \dots & \dots & \dots & 0 \end{bmatrix}$$

with $\zeta, \rho \in \mathbb{T}$. Our procedure gives an alternative proof of this result — we leave the details to the reader as an exercise. We also note that, as was mentioned in [8], for $n = 2$ or 3 there exist non-unitary minimal normal completions of such weighted shift matrices. Our procedure gives the full description of these completions B. Namely, for $n = 2$

$$B = \begin{bmatrix} 0 & a_1 & \mu x_2 \\ 0 & 0 & \mu x_1 \\ \overline{\mu} x_1 & \overline{\mu} x_2 & \overline{a}_1 x_2^2 + a_1 x_1 \overline{x}_2 \end{bmatrix},$$

where $\mu \in \mathbb{T}$ and $x_1, x_2 \in \mathbb{C}$: $|x_1|^2 - |x_2|^2 = 1$ are arbitrary, and for $n = 3$

$$B = \begin{bmatrix} 0 & a_1 & 0 & \mu x_2 \\ 0 & 0 & a_2 & 0 \\ 0 & 0 & 0 & \mu x_1 \\ \overline{\mu} x_1 & 0 & \overline{\mu} x_2 & 0 \end{bmatrix},$$

where $\mu \in \mathbb{T}$ and $x_1, x_2 \in \mathbb{C}$: $|x_1|^2 - |x|_2^2 = 1$, $\overline{a}_1 x_2 = a_2 \overline{x}_2$ are arbitrary.

We will present more applications of this method in Examples 2.23 and 2.24.

2.3. The generic case.

The procedure described in Section 2.2, which is based on part (i) of Theorem 2.3, allows one to check whether a given matrix $A \in \mathbb{C}^{n\times n}$ has normal defect one, and if this is the case — to solve the system of equations (2.11)–(2.12). Part (ii) of Theorem 2.3 describes all minimal normal completions of A. That procedure verifies first the rank condition, and then uses only the two nonzero eigenvalues, $\lambda_1 = d$ and $\lambda_2 = -d$, and the two corresponding unit eigenvectors, u_1 and u_2, of $A^*A - AA^*$. The vector equation (2.39) in that procedure is equivalent to a system of $2n$ real scalar linear equations with 4 unknowns.

In this section, we show how the procedure in Section 2.2 can be refined by using a special choice of the eigenbasis for the matrix $A^*A - AA^*$, i.e., a special construction of orthonormal eigenvectors $u_3, \ldots, u_n$ corresponding to the zero eigenvalue. This additional analysis is rewarded by obtaining a system of $n-2$ (as opposed to $2n$) real linear equations with 4 unknowns. Moreover, it allows us to describe the generic situation under the assumption that the rank condition is satisfied. The refined procedure is based on the following theorem (the proof of which is given later in this section).

THEOREM 2.17. *Let* $A \in \mathbb{C}^{n\times n}$ *satisfy the rank condition* (1.2) *and let* u_1 *and* u_2 *be the unit eigenvectors of the matrix* $A^*A - AA^*$ *corresponding to its eigenvalues* $\lambda_1 = d (> 0)$ *and* $\lambda_2 = -d$. *Then*

(i) *There exist orthonormal vectors* $u_3, \ldots, u_n \in \mathrm{null}(A^*A - AA^*)$ *(and thus the matrix* $W = \begin{bmatrix} u_1 & \ldots & u_n \end{bmatrix} \in \mathbb{C}^{n\times n}$ *is unitary) such that the matrix* $\widetilde{A} = W^*AW$ *has the form*

$$\widetilde{A} = \begin{bmatrix} a_{11} & a_{12} & u \\ a_{21} & a_{11} & v \\ v^T & u^T & S \end{bmatrix}, \tag{2.40}$$

with a_{ij}*'s defined in* (2.15).

(ii) $\mathrm{nd}(A) = 1$ *if and only if the equation*

$$\mathrm{Im}(u^* x_1 + v^* x_2) = 0 \tag{2.41}$$

has a solution pair $x_1, x_2 \in \mathbb{C}$ *satisfying*

$$|x_1|^2 - |x_2|^2 = d. \tag{2.42}$$

(*iii*) *If* $\mathrm{nd}(A) = 1$, x_1 *and* x_2 *satisfy* (2.41) *and* (2.42), *and* $\mu \in \mathbb{T}$ *is arbitrary then the matrix* B *defined in* (2.13) *is a minimal normal completion of* A.

All minimal normal completions of A *arise in this way.*

REMARK 2.18. The matrix W in Theorem 2.17 can be constructed explicitly, as will become clear from the proof of the theorem.

Let $A \in \mathbb{C}^{n\times n}$ satisfy the rank condition. We define the vectors $\mathbf{u}^*, \mathbf{v}^*, \mathbf{w}^*, \mathbf{q}^* \in \mathrm{range}(P) \subset \mathbb{C}^n$ by (2.37) (see also Theorem 2.3 for the definition of P, u_1, and u_2). Since these vectors can be viewed as the images of vectors $u^*, v^*, w^*, q^* \in \mathbb{C}^{n-2}$ under an isometry so that (2.22) holds (see Lemma 2.7 and the proof of Theorem 2.3), we have

$$\mathbf{u}\mathbf{u}^* = \mathbf{q}\mathbf{q}^*, \quad \mathbf{v}\mathbf{v}^* = \mathbf{w}\mathbf{w}^*, \quad \mathbf{u}\mathbf{v}^* = \mathbf{w}\mathbf{q}^*, \quad \mathbf{u}\mathbf{w}^* = \mathbf{v}\mathbf{q}^*. \tag{2.43}$$

The first three equalities mean that the linear operator

$$X\colon\ \mathrm{span}(\mathbf{u}^T, \mathbf{v}^T) \longrightarrow \mathrm{span}(\mathbf{q}^*, \mathbf{w}^*) \tag{2.44}$$

defined via

$$X\colon \mathbf{u}^T \longmapsto \mathbf{q}^*, \quad \mathbf{v}^T \longmapsto \mathbf{w}^* \tag{2.45}$$

is a well defined unitary operator. In order to interpret the last equality in (2.43) we need an intermission for some definitions and results on (complex) symmetric operators and matrices (see, e.g., [7, Section 4.4]).

For a subspace $\mathscr{H}$ in $\mathbb{C}^k$, denote its complex dual by

$$\overline{\mathscr{H}} := \{h \in \mathbb{C}^k : \overline{h} \in \mathscr{H}\}.$$

We will say that a $\mathbb{C}$-linear operator $L\colon \mathscr{H} \to \overline{\mathscr{H}}$ is *symmetric* if $h^T Lg = g^T Lh$ (or, equivalently, $\langle Lg, \overline{h}\rangle = \langle Lh, \overline{g}\rangle$ in the standard inner product in $\mathbb{C}^k$) for all $g, h \in \mathscr{H}$. It is clear that a $\mathbb{C}$-linear operator $L\colon \mathbb{C}^k \to \mathbb{C}^k$ is symmetric if and only if its matrix in a standard basis of $\mathbb{C}^k$ is complex symmetric, i.e, $L = L^T$. In general, a $\mathbb{C}$-linear operator $L\colon \mathscr{H} \to \overline{\mathscr{H}}$ is symmetric if and only if its matrix in any pair of orthonormal bases $\mathscr{B} = \{h_j\}_{j=1}^k \subset \mathscr{H}$ and $\overline{\mathscr{B}} = \{\overline{h}_j\}_{j=1}^k \subset \overline{\mathscr{H}}$ is complex symmetric, i.e., $\langle Lh_j, \overline{h}_i\rangle = \langle Lh_i, \overline{h}_j\rangle$, $i, j = 1, \ldots, k$.

We can restate this also in a coordinate-free form. Let $\mathscr{G}$ and $\mathscr{H}$ be two subspaces in $\mathbb{C}^k$. For a $\mathbb{C}$-linear operator $L\colon \mathscr{G} \to \overline{\mathscr{H}}$ we define its transpose as the unique $\mathbb{C}$-linear operator $L^T\colon \mathscr{H} \to \overline{\mathscr{G}}$ which satisfies $h^T Lg = g^T L^T h$ (or, equivalently, $\langle Lg, \overline{h}\rangle = \langle Lh, \overline{g}\rangle$) for all $g \in \mathscr{G}$, $h \in \mathscr{H}$). Explicitly, $L^T h = \overline{L^* \overline{h}}$ for every $h \in \mathscr{H}$. This, in particular, implies that $(LM)^T = M^T L^T$ for two $\mathbb{C}$-linear operators L, M. If one interprets a vector $h \in \mathscr{H}$ as an operator $h\colon \mathbb{C} \to \mathscr{H}$ then its transpose h^T can be interpreted as the operator $h^T\colon \overline{\mathscr{H}} \to \mathbb{C}$, and then the identity $h^T Lg = g^T L^T h$ can be viewed also as $g^T L^T h = (h^T Lg)^T$. Finally, a $\mathbb{C}$-linear operator $L\colon \mathscr{H} \to \overline{\mathscr{H}}$ is symmetric if and only if $L = L^T$.

We also observe that a matrix of a $\mathbb{C}$-linear operator $L\colon \mathscr{G} \to \overline{\mathscr{H}}$ in a pair of orthonormal bases $\mathscr{B}_1$ and $\overline{\mathscr{B}_2}$ and a matrix of $L^T\colon \mathscr{H} \to \overline{\mathscr{G}}$ in the pair of orthonormal bases $\mathscr{B}_2$ and $\overline{\mathscr{B}_1}$ are transposes of each other.

LEMMA 2.19. *Let* $\mathscr{H}$ *and* $\mathscr{L}$ *be subspaces in* $\mathbb{C}^k$ *such that* $\mathscr{H} \subset \mathscr{L}$, *and let* $Y\colon \mathscr{H} \to \overline{\mathscr{L}}$ *be an isometry such that* $P_{\overline{\mathscr{H}}}Y\colon \mathscr{H} \to \overline{\mathscr{H}}$ *is a symmetric operator. Then there exists a unitary and symmetric operator* $\widetilde{Y}\colon \mathscr{L} \to \overline{\mathscr{L}}$ *such that* $\widetilde{Y}|_{\mathscr{H}} = Y$.

Proof. We have

$$\mathscr{L} = \mathscr{H} \oplus \mathscr{G} \oplus \mathscr{K},$$

where $\mathscr{G} = \overline{\mathrm{range}(P_{\overline{L \ominus H}}Y)}$ and $\mathscr{K} = \mathscr{L} \ominus (\mathscr{H} \oplus \mathscr{G})$. The Takagi decomposition (see, e.g., [7]) of the symmetric operator $P_{\overline{\mathscr{H}}}Y$, in a coordinate-free form, is

$$P_{\overline{\mathscr{H}}}Y = U\Sigma U^T,$$

where $U\colon \mathbb{C}^{\dim(\mathscr{H})} \to \overline{\mathscr{H}}$ is a unitary operator such that $Ue_1, \ldots, Ue_{\dim(\mathscr{H})}$ are the eigenvectors of $P_{\overline{\mathscr{H}}}Y\overline{P_{\overline{\mathscr{H}}}Y}$, and $\Sigma\colon \mathbb{C}^{\dim(\mathscr{H})} \to \mathbb{C}^{\dim(\mathscr{H})}$ is an operator whose matrix in the standard basis of $\mathbb{C}^{\dim(\mathscr{H})}$ is diagonal, with the singular values of $P_{\overline{\mathscr{H}}}Y$ on the diagonal. The operator $P_{\overline{\mathscr{G}}}Y$ can be represented as

$$P_{\overline{\mathscr{G}}}Y = V(I_{\mathbb{C}^{\dim(\mathscr{H})}} - \Sigma^2)^{1/2}U^T,$$

where $V\colon \mathbb{C}^{\dim(\mathscr{H})} \to \overline{\mathscr{G}}$ is a coisometry, with

$$V|_{\mathrm{range}(I_{\mathbb{C}^{\dim(\mathscr{H})}} - \Sigma^2)}\colon \mathrm{range}(I_{\mathbb{C}^{\dim(\mathscr{H})}} - \Sigma^2) \to \overline{\mathscr{G}}$$

unitary. Define the operator $J\colon \mathscr{K} \to \overline{\mathscr{K}}$ for some pair of orthonormal bases $\mathscr{B} = \{\kappa_j\}_{j=1}^{\dim(\mathscr{K})} \subset \mathscr{K}$, $\overline{\mathscr{B}} = \{\overline{\kappa}_j\}_{j=1}^{\dim(\mathscr{K})} \subset \overline{\mathscr{K}}$ as

$$J\left(\sum_{j=1}^{\dim(\mathscr{K})} \alpha_j \kappa_j\right) = \sum_{j=1}^{\dim(\mathscr{K})} \alpha_j \overline{\kappa}_j, \quad \alpha_1, \ldots, \alpha_{\dim(\mathscr{K})} \in \mathbb{C}.$$

Clearly, J is symmetric, and the matrix of J in the pair of bases $\mathscr{B}$ and $\overline{\mathscr{B}}$ is $I_{\dim(\mathscr{K})}$. It is straightforward to verify that the operator

$$\widetilde{Y} := \begin{bmatrix} U\Sigma U^T & U(I_{\mathbb{C}^{\dim(\mathscr{H})}} - \Sigma^2)^{1/2}V^T & 0 \\ V(I_{\mathbb{C}^{\dim(\mathscr{H})}} - \Sigma^2)^{1/2}U^T & -V\Sigma V^T & 0 \\ 0 & 0 & J \end{bmatrix} : \begin{matrix} \mathscr{H} \\ \oplus \\ \mathscr{G} \\ \oplus \\ \mathscr{K} \end{matrix} \to \begin{matrix} \overline{\mathscr{H}} \\ \oplus \\ \overline{\mathscr{G}} \\ \oplus \\ \overline{\mathscr{K}} \end{matrix}$$

has the desired properties. □

COROLLARY 2.20. *The unitary operator* X *defined by* (2.44)–(2.45) *can be extended to a unitary and symmetric operator* $\widetilde{X}\colon \mathrm{range}(\overline{P}) \to \mathrm{range}(P)$.

Proof. Since $\mathbf{u}^T, \mathbf{v}^T \in \operatorname{range}(\overline{P})$ and $\mathbf{q}^*, \mathbf{w}^* \in \operatorname{range}(P)$ (see (2.37)), the operator X can be viewed as an isometry $X\colon \operatorname{span}(u^T, v^T) \to \operatorname{range}(P)$. The last identity in (2.43) means that the operator

$$P_{\operatorname{span}(\mathbf{u}^*,\mathbf{v}^*)}X\colon \operatorname{span}(\mathbf{u}^T, \mathbf{v}^T) \to \operatorname{span}(\mathbf{u}^*, \mathbf{v}^*)$$

is symmetric. Then the statement of this corollary follows from Lemma 2.19, where we set $k = n-2$,

$$\mathscr{H} = \operatorname{span}(\mathbf{u}^T, \mathbf{v}^T) = \operatorname{span}(\overline{P}A^T\overline{u}_1, \overline{P}A^T\overline{u}_2),$$

$Y = X$ and $\mathscr{L} = \operatorname{range}(\overline{P})$. $\square$

REMARK 2.21. It can be shown that the unitary and symmetric operator $\widetilde{X}$ in Corollary 2.20 can be constructed bypassing Lemma 2.19 and using instead the following remarkable theorem from [14]: If $A \in \mathbb{C}^{n\times n}$ and $x, y \in \mathbb{C}^n$ are such that $A^*A - AA^* = xx^* - yy^*$ then there exists an antiunitary involution ι on $\mathbb{C}^n$ such that $\iota x = y$ and $\iota A \iota = A^*$. (A mapping $\iota\colon \mathbb{C}^n \to \mathbb{C}^n$ is called an antiunitary involution if $\iota^2 = I_n$ and $\langle \iota h, \iota g\rangle = \langle g, h\rangle$ for every $h, g \in \mathbb{C}^n$, in the standard inner product in $\mathbb{C}^n$.) Our Lemma 2.19 seems to be of independent interest, and can be applied to other problems as well.

Proof of Theorem 2.17. (i) The operator $\widetilde{X}$ in Corollary 2.20, which is constructed as in Lemma 2.19 for the given matrix A, is symmetric and unitary, and thus has a Takagi factorization (see [7]) $\widetilde{X} = GG^T$ where $G\colon \mathbb{C}^{n-2} \to \operatorname{range}(P)$ is unitary. One can view G as an isometry $G'\colon \mathbb{C}^{n-2} \to \mathbb{C}^n$ with the same range as G:

$$\operatorname{range}(G') = \operatorname{range}(G) = \operatorname{range}(P).$$

Clearly, the columns $u_3, \ldots, u_n$ of the standard matrix of G' are orthonormal, and hence, together with u_1 and u_2, form an orthonormal eigenbasis of $A^*A - AA^*$. We also have $G'G'^* = P$. We then extend $\widetilde{X}$ to the operator $X' = G'G'^T\colon \mathbb{C}^n \to \mathbb{C}^n$. The operator represented by the matrix A in the standard basis of $\mathbb{C}^n$, in the basis $u_1, \ldots, u_n$ has the matrix

$$\widetilde{A} = \begin{bmatrix} u_1^* \\ u_2^* \\ G'^* \end{bmatrix} A \begin{bmatrix} u_1 & u_2 & G' \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & u_1^*AG' \\ a_{21} & a_{11} & u_2^*AG' \\ G'^*Au_1 & G'^*Au_2 & G'^*AG' \end{bmatrix}.$$

We have

$$\begin{aligned} (u_1^*AG')^T &= G'^TA^T\overline{u}_1 = G'^*G'G'^TA^T\overline{u}_1 = G'^*X'A^T\overline{u}_1 = G'^*X'\overline{P}A^T\overline{u}_1 \\ &= G'^*PAu_2 = G'^*Au_2, \end{aligned}$$

and similarly,

$$(u_2^*AG')^T = G'^*Au_1.$$

Setting $u = u_1^*AG'$, $v = u_2^*AG'$, $S = G'^*AG'$, and $W = \begin{bmatrix} u_1 & u_2 & u_3 & \ldots & u_n \end{bmatrix}$, we see that $\widetilde{A} = W^*AW$ has the form (2.40), which proves part (i).

(*ii*) It follows from Lemma 2.10 that $\operatorname{nd}(A) = \operatorname{nd}(\widetilde{A}) = 1$ if and only if there exist $x_1, x_2 \in \mathbb{C}$ satisfying (2.42) such that

$$u^* x_1 + v^* x_2 = v^T \overline{x}_2 + u^T \overline{x}_1.$$

Since the last equation is equivalent to (2.41), this proves part (*ii*).

(*iii*) This part is proved in the same way as part (*ii*) of Theorem 2.3. $\square$

Let $u, v \in \mathbb{C}^{n-2}$ be as in Theorem 2.17, $u = u_R + iu_I$, $v = v_R + iv_I$, where u_R^T, u_I^T, v_R^T, $v_I^T \in \mathbb{R}^{n-2}$. Let $x_1 = x_{R1} + ix_{I1}$, $x_2 = x_{R2} + ix_{I2}$, where x_{R1}, x_{I1}, x_{R2}, $x_{I2} \in \mathbb{R}$. Then (2.41) can be written as

$$Qx = 0, \tag{2.46}$$

where

$$Q = \begin{bmatrix} -u_I^T & u_R^T & -v_I^T & v_R^T \end{bmatrix} \in \mathbb{R}^{(n-2)\times 4}, \quad x = \begin{bmatrix} x_{R1} \\ x_{I1} \\ x_{R2} \\ x_{I2} \end{bmatrix} \in \mathbb{R}^4. \tag{2.47}$$

REMARK 2.22. Observe that replacing $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$, and $\mathbf{q}$ in (2.39) by u, v, $\overline{v}$, and $\overline{u}$ as in Theorem 2.17 we obtain an equivalent condition, i.e., equation (2.46) replaces (2.39) with the matrix Q replacing $\mathbf{Q}$. Thus instead of $2n$ real linear equations with 4 unknowns we obtain $n-2$ real linear equations with 4 unknowns. Let $m = \operatorname{rank}(Q)$. Then, for all possible cases of m, the procedure for checking whether $\operatorname{nd}(A) = 1$, and if this is the case — for constructing all minimal normal completions of A, is exactly the same as described is Section 2.2, with Q in the place of $\mathbf{Q}$.

We describe now the generic situation for each n, under the assumption that $\operatorname{rank}(A^*A - AA^*) = 2$. In other words, we obtain certain topological characterization of the set of matrices with normal defect one in each matrix dimension.

The generic case

Let $A \in \mathbb{C}^{n\times n}$ satisfy the rank condition. Consider the following possibilities for the value of n, and describe the situation for each case separately.

$\mathbf{n = 2}$ **or** $\mathbf{n = 3}$ Vectors u, v as in Theorem 2.17 do not arise (when $n = 2$) or are scalars (when $n = 3$). Then $m = \operatorname{rank}(Q) \leqslant 1$, and $\operatorname{nd}(A) = 1$ (for the case where $m = 1$ it follows from Remark 2.15). This gives a new proof of the statement on 2×2 matrices in Section 1 and of Corollary 2.11.

$\mathbf{n = 4}$ **or** $\mathbf{n = 5}$ In these cases, $m \leqslant 2$ (resp., $m \leqslant 3$). Thus, equation (2.46) (see also (2.47)) has nontrivial solutions. Both the situation where the matrix $\mathbf{K}$, constructed from Q instead of $\mathbf{Q}$, has at least one positive eigenvalue (in which case $\operatorname{nd}(A) = 1$) and where $\mathbf{K}$ has no positive eigenvalues (in which case $\operatorname{nd}(A) > 1$) occur on sets with nonempty interior in the relative topology of the manifolds $\mathfrak{M}_4$ and $\mathfrak{M}_5$ (see page 402 for the definition of $\mathfrak{M}_n$).

$\mathbf{n \geqslant 6}$ In this case, generically $m = 4$, thus (2.46) has no nontrivial solutions. Therefore, generically $\operatorname{nd}(A) > 1$. Still, there are matrices A with $\operatorname{nd}(A) = 1$, which can be constructed, e.g., using Theorem 2.1 and Remark 2.2.

2.4. Normal defect and unitary defect

In this section, we give two examples which show that the question in [12] (see also [8]) asking whether $\mathrm{nd}(A) = \mathrm{ud}(A)$ for any unitarily irreducible matrix A has a negative answer. In the first example, A has a single cell in its Jordan form, and in the second example, A has three distinct eigenvalues. We also present all minimal normal completions of A in both examples.

EXAMPLE 2.23. Let

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}.$$

Then

$$A^*A = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}, \quad AA^* = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix},$$

and

$$A^*A - AA^* = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Clearly, the rank condition holds. It follows from Corollary 2.11 that $\mathrm{nd}(A) = 1$. Since the only eigenvalue of A is 1, and $A - I$ is nilpotent of order 3, A has a single cell in its Jordan form, and hence A is unitarily irreducible. The characteristic polynomial of A^*A is

$$p(\lambda) = (2-\lambda)^2(1-\lambda) + 2\lambda - 3.$$

We have $p(0) = 1 > 0$, $p(1) = -1 < 0$, $p(2) = 1 > 0$, and $p(4) = -7 < 0$. Therefore, $p(\lambda)$ has three distinct roots, in intervals $(0,1)$, $(1,2)$, and $(2,4)$, i.e., A has three distinct singular values. Therefore, $\mathrm{ud}(A) = 2$. We also observe that A has the form (2.40). The procedure described in Section 2.2 together with Theorem 2.3 (or its refined version described in Remark 2.22 together with Theorem 2.17) gives that all minimal normal completions of A have the form

$$B = \begin{bmatrix} 1 & 0 & 0 & \mu x_1 \\ 0 & 1 & 1 & \mu x_2 \\ 1 & 0 & 1 & 0 \\ \overline{\mu} x_2 & \overline{\mu} x_1 & 0 & 1 \end{bmatrix},$$

with arbitrary $\mu \in \mathbb{T}$, and $x_1 \in \mathbb{C}$, $x_2 \in \mathbb{R}$ satisfying $|x_1|^2 - x_2^2 = 1$.

EXAMPLE 2.24. Let

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & \frac{3}{2}i \end{bmatrix}.$$

Changing the basis, we obtain $\widetilde{A} = U^*AU$, where

$$U = \begin{bmatrix} 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{i}{\sqrt{2}} & \frac{i}{\sqrt{2}} & 0 \end{bmatrix}$$

is unitary and

$$\widetilde{A} = \begin{bmatrix} \frac{3i}{4} & \frac{i}{4} & \frac{1}{\sqrt{2}} \\ -\frac{7i}{4} & \frac{3i}{4} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix}$$

satisfies

$$\widetilde{A}^*\widetilde{A} - \widetilde{A}\widetilde{A}^* = \begin{bmatrix} 3 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Clearly, $\operatorname{rank}(A^*A - AA^*) = \operatorname{rank}(\widetilde{A}^*\widetilde{A} - \widetilde{A}\widetilde{A}^*) = 2$. By Corollary 2.11, $\operatorname{nd}(A) = 1$. We also observe that $\widetilde{A}$ has the form (2.40). The matrix A is unitarily irreducible. Indeed, matrices

$$\operatorname{Re}(A) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{and} \quad \operatorname{Im}(\mathrm{A}) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{3}{2} \end{bmatrix}$$

do not have common eigenvectors. Next we show that $\operatorname{ud}(A) = 2$. The characteristic polynomial of A^*A,

$$p(\lambda) = (1-\lambda)(2-\lambda)\left(\frac{13}{4} - \lambda\right) + \frac{13}{4}\lambda - \frac{17}{4},$$

has values $p(0) = \frac{9}{4} > 0$, $p(1) = -1 < 0$, $p(2) = \frac{9}{4} > 0$, $p(5) = -9 < 0$. Therefore, $p(\lambda)$ has three distinct roots, in intervals $(0,1)$, $(1,2)$, and $(2,5)$, i.e., A has three distinct singular values. Thus, $\operatorname{ud}(A) = 2$. Note that in this example A has three distinct eigenvalues, $\lambda_1 = i$, $\lambda_2 = \frac{\sqrt{23}+i}{4}$, $\lambda_3 = \frac{-\sqrt{23}+i}{4}$. The procedure described in Section 2.2 together with Theorem 2.3 (or its refined version described in Remark 2.22 together with Theorem 2.17) gives that all minimal normal completions of $\widetilde{A}$ have the form

$$\widetilde{B} = \begin{bmatrix} \frac{3i}{4} & \frac{i}{4} & \frac{1}{\sqrt{2}} & \mu(h_1 + ih_3) \\ -\frac{7i}{4} & \frac{3i}{4} & \frac{1}{\sqrt{2}} & \mu(h_2 - ih_3) \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ \overline{\mu}(h_2 - ih_3) & \overline{\mu}(h_1 + ih_3) & 0 & \frac{h_1h_3 + 5h_2h_3 + i(3 - 2h_1h_2 + 2h_2^2)}{3} \end{bmatrix},$$

with arbitrary $\mu \in \mathbb{T}$, and $h_1, h_2, h_3 \in \mathbb{R}$: $h_1^2 - h_2^2 = 3$. Correspondingly, all minimal normal completions of A have the form

$$B=\begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & \mu\frac{h_1+h_2}{\sqrt{2}} \\ 0 & 1 & \frac{3}{2}i & \mu\frac{2h_3+i(h_2-h_1)}{\sqrt{2}} \\ 0 & \overline{\mu}\frac{h_1+h_2}{\sqrt{2}} & \overline{\mu}\frac{2h_3+i(h_2-h_1)}{\sqrt{2}} & \frac{h_1h_3+5h_2h_3+i(3-2h_1h_2+2h_2^2)}{3} \end{bmatrix}.$$

3. The real case

Let $A\in\mathbb{R}^{n\times n}$. We define the *real normal defect of* A, $\operatorname{rnd}(A)$, as the smallest nonnegative p such that a matrix $\begin{bmatrix} A & * \\ * & * \end{bmatrix}\in\mathbb{R}^{(n+p)\times(n+p)}$ is normal (such a matrix with the minimal possible p is a *minimal real normal completion of* A). It is clear that $\operatorname{rnd}(A)\geqslant\operatorname{nd}(A)$.

We also define the *orthogonal defect of* A as the smallest nonnegative s such that a matrix $\begin{bmatrix} A & * \\ * & * \end{bmatrix}\in\mathbb{R}^{(n+s)\times(n+s)}$ is a multiple of an orthogonal matrix (such a matrix with the minimal possible s is a *minimal orthogonal completion of* A). In fact, the orthogonal defect of A coincides with $\operatorname{ud}(A)$, so that it does not require a separate notation. Indeed, since a minimal orthogonal completion of a real matrix is obtained using the same construction as for a minimal unitary completion (see [12]), the only difference being that the real SVD is involved, the size of this minimal orthogonal completion is the same as for a minimal unitary completion.

Clearly, $\operatorname{rnd}(A)\leqslant\operatorname{ud}(A)$. We will show later (Example 3.5) that there exist orthogonally irreducible matrices A for which the strict inequality takes place.

3.1. Construction of real matrices of even size with real normal defect one

The following theorem is an analogue of Theorem 2.1 for the case of real $n\times n$ matrices with n even.

THEOREM 3.1. *Let $A\in\mathbb{R}^{n\times n}$, where $n=2k$, be not normal. The following statements are equivalent:*

(i) $\operatorname{rnd}(A)=1$.

(ii) *There exist a contraction matrix $C\in\mathbb{R}^{n\times n}$ with $\operatorname{ud}(C)=1$, a block diagonal matrix $D\in\mathbb{R}^{n\times n}$ of the form*

$$D=\operatorname{diag}\left(\begin{bmatrix} \alpha_1 & \beta_1 \\ -\beta_1 & \alpha_1 \end{bmatrix},\dots,\begin{bmatrix} \alpha_\ell & \beta_\ell \\ -\beta_\ell & \alpha_\ell \end{bmatrix},\alpha_{2\ell+1},\dots,\alpha_{2k}\right), \tag{3.1}$$

and a scalar $\mu\in\mathbb{R}$ such that

$$A=CDC^T+\mu I_n. \tag{3.2}$$

(*iii*) *There exist an orthogonal matrix* $V \in \mathbb{R}^{n\times n}$, *a normal matrix* $N \in \mathbb{R}^{n\times n}$, *and scalars* $t, \mu \in \mathbb{R}$, *with* $0 \leqslant t < 1$, *such that*

$$V^T AV = MNM + \mu I_n, \tag{3.3}$$

where $M = \operatorname{diag}(1,\ldots,1,t)$.

Proof. $(i) \Longleftrightarrow (ii)$ Let $\operatorname{rnd}(A) = 1$, and let $\begin{bmatrix} A & x \\ y^T & z \end{bmatrix} \in \mathbb{R}^{(n+1)\times(n+1)}$ be a minimal real normal completion of A. Then (see, e.g., [4, Section IX.13]) there exist a block diagonal matrix $\Lambda \in \mathbb{R}^{n\times n}$ of the form

$$\Lambda = \operatorname{diag}\left(\begin{bmatrix} \mu_1 & \beta_1 \\ -\beta_1 & \mu_1 \end{bmatrix}, \ldots, \begin{bmatrix} \mu_\ell & \beta_\ell \\ -\beta_\ell & \mu_\ell \end{bmatrix}, \mu_{2\ell+1}, \ldots, \mu_{2k} \right),$$

a scalar $\mu \in \mathbb{R}$, and an orthogonal matrix $O = \begin{bmatrix} O_{11} & O_{12} \\ O_{21} & O_{22} \end{bmatrix} \in \mathbb{R}^{(n+1)\times(n+1)}$ such that

$$\begin{bmatrix} A & x \\ y^T & z \end{bmatrix} = \begin{bmatrix} O_{11} & O_{12} \\ O_{21} & O_{22} \end{bmatrix} \begin{bmatrix} \Lambda & 0 \\ 0 & \mu \end{bmatrix} \begin{bmatrix} O_{11}^T & O_{21}^T \\ O_{12}^T & O_{22} \end{bmatrix}. \tag{3.4}$$

Here we used the fact that n is even, and thus the $(n+1)\times(n+1)$ real normal matrix $\begin{bmatrix} A & x \\ y^T & z \end{bmatrix}$ has at least one real eigenvalue. The last equality is equivalent to the following system:

$$A = O_{11}\Lambda O_{11}^T + \mu O_{12}O_{12}^T = O_{11}(\Lambda - \mu I_n)O_{11}^T + \mu I_n, \tag{3.5}$$
$$x = O_{11}\Lambda O_{21}^T + \mu O_{12}O_{22} = O_{11}(\Lambda - \mu I_n)O_{21}^T, \tag{3.6}$$
$$y^T = O_{21}\Lambda O_{11}^T + \mu O_{22}O_{12}^T = O_{21}(\Lambda - \mu I_n)O_{11}^T, \tag{3.7}$$
$$z = O_{21}\Lambda O_{21}^T + \mu O_{22}^2 = O_{21}(\Lambda - \mu I_n)O_{21}^T + \mu. \tag{3.8}$$

Setting $C = O_{11}$ and $D = \Lambda - \mu I_n$, we obtain (3.2) from (3.5).

Conversely, if (3.2) holds, we set $O_{11} = C$, $\Lambda = D + \mu I_n$ and obtain (3.5). For $O = \begin{bmatrix} O_{11} & O_{12} \\ O_{21} & O_{22} \end{bmatrix}$ a minimal orthogonal completion of C, we define $x, y \in \mathbb{R}^n$ and $z \in \mathbb{R}$ by (3.6)–(3.8). Then (3.4) holds, i.e., the matrix $\begin{bmatrix} A & x \\ y^T & z \end{bmatrix} \in \mathbb{R}^{(n+1)\times(n+1)}$ is a real normal completion of A, and thus $\operatorname{rnd}(A) = 1$.

$(ii) \Longleftrightarrow (iii)$ If (ii) holds, let $C = V\operatorname{diag}(1,\ldots,1,t)W^T$ be the SVD of C (here $V, W \in \mathbb{R}^{n\times n}$ are orthogonal, $0 \leqslant t < 1$, and $M = \operatorname{diag}(1,\ldots,1,t) \in \mathbb{R}^{n\times n}$). Clearly, $N = W^T DW$ is normal, and (3.3) follows.

Conversely, if (3.3) holds, then $N = W^T DW$ with D block diagonal of the form (3.1) and W orthogonal. For $C = V\operatorname{diag}(1,\ldots,1,t)W^T$ we have $\operatorname{ud}(C) = 1$, and (3.2) follows. $\square$

REMARK 3.2. Remark 2.2 can be restated in the real case as follows. The matrix A of even size, given by (3.3), is not normal if and only if, in the matrix

$$N = \begin{bmatrix} N_0 & g \\ h^T & \alpha \end{bmatrix}$$

partitioned so that α is scalar, $g \neq h$ and, in the case where $t\alpha = 0$, also $g \neq -h$. Moreover, if both M and N are invertible then A is not normal if and only if the standard basis vector e_n is not an eigenvector of $N^T N^{-1}$. The statement in the last sentence of Remark 2.2 is, in general, not valid in the real case.

Open problem: What is an analogue of Theorem 3.1 for the case of odd n?

In the case of even n, similarly to the complex case, representation (3.2) or (3.3) in Theorem 3.1, along with Remark 3.2, allow one to construct all matrices A with $\operatorname{rnd}(A) = 1$. However, this does not give a way to check whether a given real matrix has real normal defect one. A procedure for this is our further goal.

3.2. Identification of matrices with rnd$(\mathbf{A}) = \mathbf{1}$ and construction of their minimal real normal completions

The following theorem is the real counterpart of Theorem 2.3.

THEOREM 3.3. *Let $A \in \mathbb{R}^{n\times n}$. Then*

(*i*) $\operatorname{rnd}(A) = 1$ *if and only if* $\operatorname{rank}(A^TA - AA^T) = 2$ *and at least one of the equations*

$$(PA^T u_1 - PAu_2)x_1 + (PA^T u_2 - PAu_1)x_2 = 0, \tag{3.9}$$

$$(PA^T u_1 + PAu_2)x_1 + (PA^T u_2 + PAu_1)x_2 = 0, \tag{3.10}$$

has a solution pair $x_1, x_2 \in \mathbb{R}$ satisfying

$$x_1^2 - x_2^2 = d. \tag{3.11}$$

Here $u_1, u_2 \in \mathbb{R}^n$ are the unit eigenvectors of the matrix $A^TA - AA^T$ corresponding to its nonzero eigenvalues $\lambda_1 = d (> 0)$ and $\lambda_2 = -d$, and

$$P = I_n - u_1u_1^T - u_2u_2^T \tag{3.12}$$

is the orthogonal projection of $\mathbb{R}^n$ onto $\operatorname{null}(A^TA - AA^T)$.

(*ii*) *If* $\operatorname{rnd}(A) = 1$ *then at least one of the following cases occurs:*

Case 1. *If x_1 and x_2 satisfy* (3.9) *and* (3.11) *then the matrix*

$$B_1 = \begin{bmatrix} A & x_1u_1 + x_2u_2 \\ x_2u_1^T + x_1u_2^T & z \end{bmatrix} \tag{3.13}$$

is a minimal real normal completion of A. Here

$$z = a_{11} - \frac{1}{d}(x_1 + x_2)(a_{12}x_1 - a_{21}x_2) \tag{3.14}$$

and

$$a_{11} = u_1^T A u_1, \quad a_{12} = u_1^T A u_2, \quad a_{21} = u_2^T A u_1; \tag{3.15}$$

Case 2. *If x_1 and x_2 satisfy* (3.10) *and* (3.11) *then the matrix*

$$B_2 = \begin{bmatrix} A & x_1u_1 + x_2u_2 \\ -x_2u_1^T - x_1u_2^T & z \end{bmatrix} \tag{3.16}$$

is a minimal real normal completion of A. Here

$$z = a_{11} + \frac{1}{d}(x_1 - x_2)(a_{12}x_1 + a_{21}x_2) \tag{3.17}$$

and a_{ij}'s are defined by (3.15).

Any minimal real normal completion of A arises in this way, i.e., either as in Case 1 or as in Case 2 above.

Proof. Since $\operatorname{rnd}(A) = 1$ implies $\operatorname{nd}(A) = 1$, it follows from Corollary 2.6 that the rank condition is necessary for $\operatorname{rnd}(A) = 1$. For real A it takes the form $\operatorname{rank}(A^TA - AA^T) = 2$. Without loss of generality, we can assume that the rank condition is satisfied. Then we find the unit eigenvectors $u_1, u_2 \in \mathbb{R}^n$ of the matrix $A^TA - AA^T$ corresponding to its eigenvalues $\lambda_1 = d(>0)$ and $\lambda_2 = -d$. Let $u_3, \ldots, u_n$ be an orthonormal basis for $\operatorname{null}(A^TA - AA^T)$. Then $U' = \begin{bmatrix} u_3 & \ldots, u_n \end{bmatrix} \in \mathbb{R}^{n\times(n-2)}$ is an isometry, and

$$U'U'^T = P, \tag{3.18}$$

where P is defined in (3.12). The matrix $\widetilde{A} = U^TAU$, where $U = \begin{bmatrix} u_1 & \ldots & u_n \end{bmatrix}$ is orthogonal, has the form

$$\widetilde{A} = \begin{bmatrix} a_{11} & a_{12} & u \\ a_{21} & a_{11} & v \\ w^T & q^T & S \end{bmatrix} \tag{3.19}$$

(the identity $a_{11} = u_1^TAu_1 = u_2^TAu_2$ is established in the same way as in Lemma 2.7). As in Lemma 2.8, we obtain that $\operatorname{rnd}(\widetilde{A}) = 1$ (and therefore, $\operatorname{rnd}(A) = 1$) if and only if there exist $x_1, x_2, y_1, y_2, z \in \mathbb{R}$ such that the matrix

$$\widetilde{B} = \begin{bmatrix} a_{11} & a_{12} & u & x_1 \\ a_{21} & a_{11} & v & x_2 \\ w^T & q^T & S & 0 \\ y_1 & y_2 & 0 & z \end{bmatrix} \in \mathbb{R}^{(n+1)\times(n+1)} \tag{3.20}$$

is normal if and only if there exist $x_1, x_2, y_1, y_2, z \in \mathbb{R}$ such that

$$
\begin{aligned}
(a_{11}-z)x_1 + a_{21}x_2 &= (a_{11}-z)y_1 + a_{12}y_2, &(3.21)\\
a_{12}x_1 + (a_{11}-z)x_2 &= a_{21}y_1 + (a_{11}-z)y_2, &(3.22)\\
u^T x_1 + v^T x_2 &= w^T y_1 + q^T y_2, &(3.23)\\
x_1^2 - y_1^2 = y_2^2 - x_2^2 &= d, &(3.24)\\
x_1 x_2 &= y_1 y_2. &(3.25)
\end{aligned}
$$

It follows from (3.24) and (3.25) that either $y_1 = x_2$, $y_2 = x_1$ or $y_1 = -x_2$, $y_2 = -x_1$. We will consider these two cases separately.

Case 1: $y_1 = x_2$**,** $y_2 = x_1$**.** Identities (3.21)–(3.24) become

$$
\begin{aligned}
(a_{11}-z)x_1 + a_{21}x_2 &= (a_{11}-z)x_2 + a_{12}x_1, &(3.26)\\
a_{12}x_1 + (a_{11}-z)x_2 &= a_{21}x_2 + (a_{11}-z)x_1, &(3.27)\\
(u^T - q^T)x_1 + (v^T - w^T)x_2 &= 0, &(3.28)\\
x_1^2 - x_2^2 &= d. &(3.29)
\end{aligned}
$$

Clearly, (3.26) and (3.27) are equivalent, and it follows from (3.26) and (3.29) that

$$
z = a_{11} - \frac{a_{12}x_1 - a_{21}x_2}{x_1 - x_2} = a_{11} - \frac{1}{d}(x_1 + x_2)(a_{12}x_1 - a_{21}x_2)
$$

(cf. (2.14)). Next, it follows from (3.19) that

$$
u^T = U'^T A^T u_1, \quad v^T = U'^T A^T u_2, \quad w^T = U'^T A u_1, \quad q^T = U'^T A u_2.
$$

Multiplying both parts of these equalities on the left by U' and taking into account (3.18), we obtain vectors

$$
\begin{aligned}
\mathbf{u}^T = U'u^T = PA^T u_1,\ \mathbf{v}^T &= U'v^T = PA^T u_2, &(3.30)\\
\mathbf{w}^T = U'w^T = PAu_1,\ \ \mathbf{q}^T &= U'q^T = PAu_2. &(3.31)
\end{aligned}
$$

Since U' is an isometry, (3.28) is equivalent to

$$
(\mathbf{u}^T - \mathbf{q}^T)x_1 + (\mathbf{v}^T - \mathbf{w}^T)x_2 = 0. \tag{3.32}
$$

Note that the definition of vectors $\mathbf{u}^T, \mathbf{v}^T, \mathbf{w}^T, \mathbf{q}^T \in \operatorname{range}(P) \subset \mathbb{R}^n$ in (3.30)–(3.31) is independent of U', i.e., of the choice of basis vectors $u_3, \ldots, u_n$ in $\operatorname{range}(P) = \operatorname{null}(A^TA - AA^T)$.

Case 2: $y_1 = -x_2$**,** $y_2 = -x_1$**.** Identities (3.21)–(3.24) become

$$
\begin{aligned}
(a_{11}-z)x_1 + a_{21}x_2 &= -(a_{11}-z)x_2 - a_{12}x_1, &(3.33)\\
a_{12}x_1 + (a_{11}-z)x_2 &= -a_{21}x_2 - (a_{11}-z)x_1, &(3.34)\\
(u^T + q^T)x_1 + (v^T + w^T)x_2 &= 0, &(3.35)\\
x_1^2 - x_2^2 &= d. &(3.36)
\end{aligned}
$$

Clearly, (3.33) and (3.34) are equivalent, and it follows from (3.33) and (3.36) that

$$z = a_{11} + \frac{a_{12}x_1 + a_{21}x_2}{x_1 + x_2} = a_{11} + \frac{1}{d}(x_1 - x_2)(a_{12}x_1 + a_{21}x_2).$$

Next, we obtain vectors $\mathbf{u}^T, \mathbf{v}^T, \mathbf{w}^T, \mathbf{q}^T \in \operatorname{range}(P) \subset \mathbb{R}^n$ as in Case 1. Since U' is an isometry, (3.35) is equivalent to

$$(\mathbf{u}^T + \mathbf{q}^T)x_1 + (\mathbf{v}^T + \mathbf{w}^T)x_2 = 0. \tag{3.37}$$

It follows from the analysis of cases above that $\operatorname{rnd}(A) = 1$ if and only if at least one of the equations (3.32) and (3.37) has a solution pair $x_1, x_2 \in \mathbb{R}$ satisfying (3.11), which proves part (*i*).

If x_1, x_2 satisfy (3.32) (resp., (3.37)) and (3.11) then $y_1 = x_2$, $y_2 = x_1$ and z defined by (3.14) (resp., $y_1 = -x_2$, $y_2 = -x_1$ and z defined by (3.17)) determine the minimal real normal completion $\widetilde{B}$ of the matrix $\widetilde{A}$ by (3.20). Then the matrix

$$B = \begin{bmatrix} U & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & u & x_1 \\ a_{21} & a_{11} & v & x_2 \\ w^T & q^T & S & 0 \\ y_1 & y_2 & 0 & z \end{bmatrix} \begin{bmatrix} U^T & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} A & u_1x_1 + u_2x_2 \\ y_1u_1^T + y_2u_2^T & z \end{bmatrix}$$

is a minimal real normal completion of A. Since $B = B_1$ in Case 1 and $B = B_2$ in Case 2, this proves part (*ii*) of the theorem. □

COROLLARY 3.4. *For a matrix* $A \in \mathbb{R}^{n\times n}$, $\operatorname{rnd}(A) = 1$ *if and only if* $\operatorname{nd}(A) = 1$.

Proof. Since we have $\operatorname{nd}(A) \leqslant \operatorname{rnd}(A)$, it suffices to prove that if $\operatorname{nd}(A) = 1$ then $\operatorname{rnd}(A) = 1$. Suppose that $\operatorname{nd}(A) = 1$. Then, as described in Section 2.2, equation (2.39) has a solution $x = \operatorname{col}(x_{R1}, x_{I1}, x_{R2}, x_{I2}) \in \mathbb{R}^4$ with

$$x_{R1}^2 + x_{I1}^2 > x_{R2}^2 + x_{I2}^2 \tag{3.38}$$

(see Theorem 2.3). The matrix $\mathbf{Q}$ in this (real) case has the form

$$\mathbf{Q} = \begin{bmatrix} \mathbf{u}^T - \mathbf{q}^T & 0 & \mathbf{v}^T - \mathbf{w}^T & 0 \\ 0 & \mathbf{u}^T + \mathbf{q}^T & 0 & \mathbf{v}^T + \mathbf{w}^T \end{bmatrix}.$$

Thus, in this case (2.39) is equivalent to the pair of equations

$$\begin{aligned} (\mathbf{u}^T - \mathbf{q}^T)x_{R1} + (\mathbf{v}^T - \mathbf{w}^T)x_{R2} &= 0, \\ (\mathbf{u}^T + \mathbf{q}^T)x_{I1} + (\mathbf{v}^T + \mathbf{w}^T)x_{I2} &= 0. \end{aligned}$$

Since in (3.38) either $x_{R1}^2 > x_{R2}^2$ or $x_{I1}^2 > x_{I2}^2$, at least one of the equations (3.32) or (3.37) (or equivalently, at least one of the equations (3.9) and (3.10)) has a solution pair $x_1, x_2 \in \mathbb{R}$ with $x_1^2 > x_2^2$ (and thus, a solution pair $x_1, x_2 \in \mathbb{R}$ satisfying $x_1^2 - x_2^2 = d$), which by Theorem 3.3 means that $\operatorname{rnd}(A) = 1$. □

Open problem: Is it true that for any $A \in \mathbb{R}^{n\times n}$ one has $\mathrm{rnd}(A) = \mathrm{nd}(A)$?

As in the complex case, we will describe now a procedure (in this setting based on Theorem 3.3) which allows one to determine whether $\mathrm{rnd}(A) = 1$ for a given matrix $A \in \mathbb{R}^{n\times n}$. Moreover, this procedure allows one to find all solutions of (3.9) and of (3.10) subject to (3.11), and then, applying part (*ii*) of Theorem 3.3, all minimal real normal completions of A when $\mathrm{rnd}(A) = 1$.

The procedure

Begin

Step 1. Check the rank condition. If it holds — go to Step 2. Otherwise, stop: $\mathrm{rnd}(A) > 1$.

Step 2. Write (3.9) in the form (3.32), where $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$, $\mathbf{q}$ are defined by (3.30) and (3.31) (see Theorem 3.3 for the definition of P, u_1, and u_2). Find $m_1 = \mathrm{rank}(\mathbf{u}^T - \mathbf{q}^T, \mathbf{v}^T - \mathbf{w}^T)$.

Step 3. Depending on m_1, consider the following cases.

(1a) $m_1 = 0$. In this case, any $x_1, x_2 \in \mathbb{R}$ with $x_1^2 - x_2^2 = d$ solve (3.32).

(1b) $m_1 = 1$, i.e., $\mathbf{u}^T - \mathbf{q}^T = \alpha b$, $\mathbf{v}^T - \mathbf{w}^T = \beta b$, with some nonzero vector $b \in \mathrm{range}(P)$ and $\alpha, \beta \in \mathbb{R}$, and additionally $|\alpha| \geqslant |\beta|$. In this case, (3.32) is equivalent to $\alpha x_1 + \beta x_2 = 0$, and has no solutions satisfying (3.11).

(1c) $m_1 = 1$, i.e., $\mathbf{u}^T - \mathbf{q}^T = \alpha b$, $\mathbf{v}^T - \mathbf{w}^T = \beta b$, with some nonzero vector $b \in \mathrm{range}(P)$ and $\alpha, \beta \in \mathbb{R}$, and additionally $|\alpha| < |\beta|$. In this case, (3.32) is equivalent to $\alpha x_1 + \beta x_2 = 0$, and has the solutions

$$x_1 = \pm\beta\sqrt{\frac{d}{\beta^2 - \alpha^2}}, \quad x_2 = \mp\alpha\sqrt{\frac{d}{\beta^2 - \alpha^2}}$$

satisfying (3.11).

(1d) $m_1 = 2$. In this case, (3.32) has only a zero solution, and thus has no solutions satisfying (3.11).

Step 4. Write (3.10) in the form (3.37), where $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$, $\mathbf{q}$ are defined by (3.30) and (3.31). Find $m_2 = \mathrm{rank}(\mathbf{u}^T + \mathbf{q}^T, \mathbf{v}^T + \mathbf{w}^T)$.

Step 5. Depending on m_2, consider the following cases.

(2a) $m_2 = 0$. In this case, any $x_1, x_2 \in \mathbb{R}$ with $x_1^2 - x_2^2 = d$ solve (3.37).

(2b) $m_2 = 1$, i.e., $\mathbf{u}^T + \mathbf{q}^T = \gamma h$, $\mathbf{v}^T + \mathbf{w}^T = \delta h$, with some nonzero vector $h \in \mathrm{range}(P)$ and $\gamma, \delta \in \mathbb{R}$, and additionally $|\gamma| \geqslant |\delta|$. In this case, (3.37) is equivalent to $\gamma x_1 + \delta x_2 = 0$, and has no solutions satisfying (3.11).

(2c) $m_2 = 1$, i.e., $\mathbf{u}^T + \mathbf{q}^T = \gamma h$, $\mathbf{v}^T + \mathbf{w}^T = \delta h$, with some nonzero vector $h \in \operatorname{range}(P)$ and $\gamma, \delta \in \mathbb{R}$, and additionally $|\gamma| < |\delta|$. In this case, (3.37) is equivalent to $\gamma x_1 + \delta x_2 = 0$, and has the solutions

$$x_1 = \pm\delta\sqrt{\frac{d}{\delta^2-\gamma^2}}, \quad x_2 = \mp\gamma\sqrt{\frac{d}{\delta^2-\gamma^2}}$$

satisfying (3.11).

(2d) $m_2 = 2$. In this case, (3.37) has only a zero solution, and thus has no solutions satisfying (3.11).

Step 6. $\operatorname{rnd}(A) = 1$ if and only if neither of the combinations (1b)&(2b), (1b)&(2d), (1d)&(2b), (1d)&(2d) occur. If it does, stop: $\operatorname{rnd}(A) > 1$. Otherwise, for each pair $x_1, x_2 \in \mathbb{R}$ obtained at Step 3, find a minimal real normal completion of A as described in (3.13)–(3.15); for each pair $x_1, x_2 \in \mathbb{R}$ obtained at Step 4, find a minimal real normal completion of A as described in (3.15)–(3.17).

End

Of course, if one is interested only in checking whether $\operatorname{rnd}(A) = 1$, the procedure can be terminated as soon as any of cases (1a), (1c), (2a), (2c) occurs.

EXAMPLE 3.5. In Example 2.23,

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

is a matrix with real entries, and

$$A^TA - AA^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

so that the rank condition is satisfied. By Corollaries 2.11 and 3.4, $\operatorname{rnd}(A) = 1$. We have $u_1 = e_1$, $u_2 = e_2$, and

$$P = I - u_1u_1^T - u_2u_2^T = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Then, in the procedure above, $\mathbf{u}^T = \mathbf{q}^T = 0$, $\mathbf{v}^T = \mathbf{w}^T = e_3$. Since $m_1 = \operatorname{rank}(\mathbf{u}^T - \mathbf{q}^T, \mathbf{v}^T - \mathbf{w}^T) = 0$, as in Case (1a), any $x_1, x_2 \in \mathbb{R}$ with $x_1^2 - x_2^2 = 1$ solve (3.32). We have $y_1 = x_2$ and $y_2 = x_1$. According to (3.14), $z = 1$. Therefore, for any $x_1 \in \mathbb{R}$ such that $|x_1| \geqslant 1$,

$$\begin{bmatrix} 1 & 0 & 0 & x_1 \\ 0 & 1 & 1 & \pm\sqrt{x_1^2-1} \\ 1 & 0 & 1 & 0 \\ \pm\sqrt{x_1^2-1} & x_1 & 0 & 1 \end{bmatrix}$$

is a minimal real normal completion of A. We also have

$$m_2 = \operatorname{rank}(\mathbf{u}^T + \mathbf{q}^T, \mathbf{v}^T + \mathbf{w}^T) = \operatorname{rank}(0, 2e_3) = 1,$$

and as in Case (2c), $h = e_3$, $\gamma = 0$, $\delta = 2$, so that $x_1 = \pm 1 = -y_2, x_2 = 0 = -y_1$. According to (3.17), $z = 1$. Thus,

$$\begin{bmatrix} 1 & 0 & 0 & \pm 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & \mp 1 & 0 & 1 \end{bmatrix}$$

is a minimal real normal completion of A. Therefore, the set of minimal real normal completions of A arises from Cases (1a) and (2c). Note that the minimal real normal completions of A in this example are special cases of minimal normal completions of A as in (2.13), where $x_1 \in \mathbb{R}$: $|x_1| \geqslant 1$, $x_2 = \pm\sqrt{x_1^2 - 1}$, and $\mu = 1$, or where $x_1 = i$, $x_2 = 0$, and $\mu = \pm i$. We know from Example 2.23 that $\operatorname{ud}(A) = 2$ and that A is unitarily (and therefore orthogonally) irreducible. This example shows that $\operatorname{rnd}(A) < \operatorname{ud}(A)$ is possible for an orthogonally irreducible matrix A.

3.3. The generic case

We will describe now the generic situation in each matrix dimension n. As in the complex case, a finer analysis is needed for this. However, in the real case our analysis is more straightforward and does not use a "heavy machinery" of symmetric extensions.

For a real matrix A satisfying the rank condition, it follows from Lemma 2.7 that the following identities hold:

$$\mathbf{u}\mathbf{u}^T = \mathbf{q}\mathbf{q}^T, \quad \mathbf{v}\mathbf{v}^T = \mathbf{w}\mathbf{w}^T, \quad \mathbf{u}\mathbf{v}^T = \mathbf{w}\mathbf{q}^T, \quad \mathbf{u}\mathbf{w}^T = \mathbf{v}\mathbf{q}^T. \tag{3.39}$$

Consequently,

$$(\mathbf{u}+\mathbf{q})(\mathbf{u}-\mathbf{q})^T = 0,\ (\mathbf{v}+\mathbf{w})(\mathbf{v}-\mathbf{w})^T = 0,\ (\mathbf{v}+\mathbf{w})(\mathbf{u}-\mathbf{q})^T = 0,\ (\mathbf{u}+\mathbf{q})(\mathbf{v}-\mathbf{w})^T = 0,$$

i.e., each of the vectors $(\mathbf{u}+\mathbf{q})^T$ and $(\mathbf{v}+\mathbf{w})^T$ is orthogonal to each of the vectors $(\mathbf{u}-\mathbf{q})^T$ and $(\mathbf{v}-\mathbf{w})^T$. Note that the vectors $\mathbf{u}^T$, $\mathbf{v}^T$, $\mathbf{w}^T$, and $\mathbf{q}^T$ belong to $\operatorname{range}(P)$ whose dimension is $n-2$.

Restricting our attention to real matrices in $\mathfrak{M}_n$, we now consider different values of n separately.

$\mathbf{n = 2}$ In this case, vectors $\mathbf{u}^T, \mathbf{v}^T, \mathbf{w}^T$, and $\mathbf{q}^T$ do not arise, thus necessarily $\operatorname{rnd}(A) = 1$. This follows also from the fact that $\operatorname{nd}(A) = 1$ by Corollary 3.4.

$\mathbf{n = 3}$ In this case, vectors $\mathbf{u}^T, \mathbf{v}^T, \mathbf{w}^T$, and $\mathbf{q}^T$ are collinear, and in view of (3.39) either $\mathbf{u}^T = \mathbf{q}^T$ and $\mathbf{v}^T = \mathbf{w}^T$, or $\mathbf{u}^T = -\mathbf{q}^T$ and $\mathbf{v}^T = -\mathbf{w}^T$. Thus either Case (1a) or Case (2a) in the Procedure occurs. Therefore, necessarily $\operatorname{rnd}(A) = 1$ (this follows also from Corollaries 2.11 and 3.4).

n = 4 Generically, $\mathbf{u}^T \neq \pm\mathbf{q}^T$, and $\mathbf{v}^T \neq \pm\mathbf{w}^T$. Since $\dim(\mathrm{range}(P)) = 2$, the vectors $(\mathbf{u}+\mathbf{q})^T$ and $(\mathbf{v}+\mathbf{w})^T$ are collinear and orthogonal to $(\mathbf{u}-\mathbf{q})^T$ and $(\mathbf{v}-\mathbf{w})^T$, which are also collinear. Then both the combination of Case (1b) and Case (2b), and the combination of Case (1c) and Case (2c) (and thus, both $\mathrm{rnd}(A) = 1$ and $\mathrm{rnd}(A) > 1$) occur on the sets whose interior is nonempty in the relative topology of the manifold $\mathfrak{M}_4$. Indeed, the first combination occurs when we fix $\mathbf{u}^T, \mathbf{q}^T$ and make $\mathbf{v}\mathbf{v}^T = \mathbf{w}\mathbf{w}^T$ small enough, and the second combination occurs when we fix $\mathbf{v}^T, \mathbf{w}^T$ and make $\mathbf{u}\mathbf{u}^T = \mathbf{q}\mathbf{q}^T$ small enough.

n = 5 Generically, $\mathbf{u}^T \neq \pm\mathbf{q}^T$, and $\mathbf{v}^T \neq \pm\mathbf{w}^T$. Since $\dim(\mathrm{range}(P)) = 3$, at least one of the pairs of vectors (generically, only one such pair), $(\mathbf{u}+\mathbf{q})^T$ and $(\mathbf{v}+\mathbf{w})^T$ or $(\mathbf{u}-\mathbf{q})^T$ and $(\mathbf{v}-\mathbf{w})^T$, is collinear. As in the case $n = 4$, for a collinear pair, both cases (b) and (c) occur on the sets whose interior is nonempty in the relative topology of $\mathfrak{M}_5$. Thus, combinations of Case (1b) and Case (2d), Case (1d) and Case (2b) (where $\mathrm{rnd}(A) > 1$), and combinations of Case (1c) and Case (2d), Case (1d) and Case (2c) (where $\mathrm{rnd}(A) = 1$) occur on the sets whose interior is nonempty in the relative topology of $\mathfrak{M}_5$.

n ⩾ 6 Since $\dim(\mathrm{range}(P)) \geqslant 4$, the pairs $(\mathbf{u}+\mathbf{q})^T$, $(\mathbf{v}+\mathbf{w})^T$ and $(\mathbf{u}-\mathbf{q})^T$, $(\mathbf{v}-\mathbf{w})^T$ are generically linearly independent. Therefore, the combination of Case (1d) and Case (2d) (corresponding to $\mathrm{rnd}(A) > 1$) occurs generically.

Thus, we see that the generic situation in the real case is similar to the one in the complex case.

4. Commuting completion problems

The problem of finding commuting completions of a N-tuple of $n \times n$ matrices was raised in [3], where a special emphasis was placed on symmetric completions of N-tuples of symmetric matrices. In [8], an inverse completion $(A_{\mathrm{ext}}, B_{\mathrm{ext}})$ of a pair (A,B) was constructed. Namely, A_{ext}, B_{ext} by definition satisfy $A_{\mathrm{ext}}B_{\mathrm{ext}} = \alpha I$ with a non-zero scalar α, and therefore commute. Our results from Sections 2 and 3 can be used to tackle commuting completion problems in the classes of Hermitian (resp., symmetric, or symmetric/antisymmetric) pairs of matrices.

4.1. The commuting Hermitian completion problem.

Let (A_1, A_2) be a pair of Hermitian matrices of size $n \times n$. We define the *commuting Hermitian defect of* A_1 *and* A_2, denoted $\mathrm{chd}(A_1, A_2)$, as the smallest p such that there exist commuting Hermitian matrices $B_1 = \begin{bmatrix} A_1 & * \\ * & * \end{bmatrix}$ and $B_2 = \begin{bmatrix} A_2 & * \\ * & * \end{bmatrix}$ of size $(n+p) \times (n+p)$. We call such a pair (B_1, B_2) of size $(n+\mathrm{chd}(A_1,A_2)) \times (n+\mathrm{chd}(A_1,A_2))$ a *minimal commuting Hermitian completion of* (A_1, A_2).

Since (B_1, B_2) is a commuting Hermitian completion of a pair (A_1, A_2) of Hermitian matrices if and only if $B = B_1 + iB_2$ is a normal completion of $A = A_1 + iA_2$, and therefore $\mathrm{chd}(A_1, A_2) = \mathrm{nd}(A_1 + iA_2)$, the results from Section 2.2 allow one to check whether $\operatorname{chd}(A_1,A_2)=1$, and when this is the case — to construct all minimal commuting Hermitian completions of (A_1,A_2). For example, Theorem 2.3 yields the following.

THEOREM 4.1. *Let $A_1,A_2\in\mathbb{C}^{n\times n}$ be Hermitian.*

(i) $\operatorname{chd}(A_1,A_2)=1$ *if and only if* $\operatorname{rank}(A_1A_2-A_2A_1)=2$ *and the equation*

$$PA_1(t_1u_1-\bar{t}_1u_2)=iPA_2(t_2u_1+\bar{t}_2u_2) \tag{4.1}$$

has a solution pair $t_1,t_2\in\mathbb{C}$ satisfying

$$\operatorname{Re}(\bar{t}_1t_2)=d. \tag{4.2}$$

Here $u_1,u_2\in\mathbb{C}^n$ are the unit eigenvectors of the matrix $2i(A_1A_2-A_2A_1)$ corresponding to its nonzero eigenvalues $\lambda_1=d(>0)$ and $\lambda_2=-d$, and $P=I_n-u_1u_1^*-u_2u_2^*$ is the orthogonal projection of $\mathbb{C}^n$ onto $\operatorname{null}(A_1A_2-A_2A_1)$.

(ii) *If* $\operatorname{chd}(A_1,A_2)=1$, *$t_1$ and t_2 satisfy* (4.1) *and* (4.2)*, and $\mu\in\mathbb{T}$ is arbitrary then the pair (B_1,B_2) of matrices*

$$B_1=\begin{bmatrix} A_1 & \frac{\mu}{2}(t_2u_1+\bar{t}_2u_2)\\ \frac{\bar{\mu}}{2}(\bar{t}_2u_1^*+t_2u_2^*) & z_1\end{bmatrix}, \tag{4.3}$$

$$B_2=\begin{bmatrix} A_2 & \frac{\mu}{2i}(t_1u_1-\bar{t}_1u_2)\\ -\frac{\bar{\mu}}{2i}(\bar{t}_1u_1^*-t_1u_2^*) & z_2\end{bmatrix} \tag{4.4}$$

is a minimal commuting Hermitian completion of (A_1,A_2). Here

$$z_1=u_1^*A_1u_1-\frac{1}{d}\left(\operatorname{Im}(t_2^2u_2^*A_2u_1)+\operatorname{Re}(t_1t_2u_2^*A_1u_1)\right) \tag{4.5}$$

and

$$z_2=u_1^*A_2u_1-\frac{1}{d}\left(\operatorname{Im}(t_1^2u_2^*A_1u_1)-\operatorname{Re}(t_1t_2u_2^*A_2u_1)\right). \tag{4.6}$$

All minimal commuting Hermitian completions of (A_1,A_2) arise in this way.

Proof. Letting $A=A_1+iA_2$, we observe that $A^*A-AA^*=2i(A_1A_2-A_2A_1)$. It is straightforward to verify that, under the change of variables $t_1=x_1-\bar{x}_2$, $t_2=x_1+\bar{x}_2$, condition (2.11) in Theorem 2.3 is equivalent to condition (4.1), (2.12) is equivalent to (4.2), B_1 and B_2 defined by (4.3) and (4.4) are Hermitian and such that $B=B_1+iB_2$ is as in (2.13), z_1 and z_2 defined by (4.5) and (4.6) are real and such that $z=z_1+iz_2$ is as in (2.14). Thus, this theorem is equivalent to Theorem 2.3. $\square$

4.2. The commuting completion problem in the class of pairs of symmetric and antisymmetric matrices.

Let $A_1 = A_1^T \in \mathbb{R}^{n\times n}$ and $A_2 = -A_2^T \in \mathbb{R}^{n\times n}$. It is natural to ask what is the smallest p such that there exist commuting matrices $B_1 = B_1^T = \begin{bmatrix} A_1 & * \\ * & * \end{bmatrix} \in \mathbb{R}^{(n+p)\times(n+p)}$ and $B_2 = -B_2^T = \begin{bmatrix} A_2 & * \\ * & * \end{bmatrix} \in \mathbb{R}^{(n+p)\times(n+p)}$. Such a pair (B_1, B_2) is a *minimal commuting completion of* (A_1, A_2) *in the class of pairs of symmetric and antisymmetric matrices*.

Since (B_1, B_2) is a commuting completion of (A_1, A_2) in this class if and only if $B = B_1 + B_2$ is a real normal completion of $A = A_1 + A_2$, our results from Section 3 can be restated in terms of pairs of matrices in this class. We omit the details, since the reasoning is similar to the one in Section 4.1.

4.3. The commuting symmetric completion problem.

In this section, we consider the commuting completion problem in the class of pairs of symmetric matrices. This is a special case of the problem raised in Degani et al. [3] (see the first paragraph of Section 4) for $N = 2$. The authors of [3] presented an approach to n-dimensional cubature formulae where the cubature nodes are obtained by means of commuting completions of certain matrix tuples. While their commuting completion problem is stated in a certain subclass of tuples of symmetric matrices, some observations were also made for the problem in the whole class. In particular, the question on the minimal possible size of completed matrices was accentuated as important.

Let $A_1 = A_1^T \in \mathbb{R}^{n\times n}$ and $A_2 = A_2^T \in \mathbb{R}^{n\times n}$. We define the *commuting symmetric defect of* A_1 *and* A_2, denoted $\operatorname{csd}(A_1, A_2)$, as the smallest p such that there exist commuting symmetric matrices $B_1 = \begin{bmatrix} A_1 & * \\ * & * \end{bmatrix}, B_2 = \begin{bmatrix} A_2 & * \\ * & * \end{bmatrix} \in \mathbb{R}^{(n+p)\times(n+p)}$. Such a pair (B_1, B_2) of size $(n + \operatorname{csd}(A_1, A_2)) \times (n + \operatorname{csd}(A_1, A_2))$ is a *minimal commuting symmetric completion of the pair* (A_1, A_2).

We note that (B_1, B_2) is a commuting symmetric completion of a pair (A_1, A_2) of real symmetric matrices if and only if $B = B_1 + iB_2$ is a normal, and simultaneously complex symmetric, completion of $A = A_1 + iA_2$. We also observe that a priori

$$\operatorname{csd}(A_1, A_2) \geqslant \operatorname{chd}(A_1, A_2). \tag{4.7}$$

Open problem. Is it true that for any pair (A_1, A_2) of real symmetric matrices one has $\operatorname{csd}(A_1, A_2) = \operatorname{chd}(A_1, A_2)$?

This problem is equivalent to the question whether a minimal normal completion of a complex symmetric matrix can be chosen to be complex symmetric. It is somewhat similar to the open problem stated in Section 3.2, which actually asks whether a minimal normal completion of a real matrix can be chosen to be real. The following theorem shows that, for a pair (A_1, A_2) of real symmetric matrices,

$$\operatorname{csd}(A_1, A_2) = 1 \Longleftrightarrow \operatorname{chd}(A_1, A_2) = 1,$$

which motivates the open problem stated above. Moreover, this theorem shows that if $\operatorname{csd}(A_1,A_2)=1$ then the set of all minimal commuting symmetric completions (B_1,B_2) of (A_1,A_2) can be obtained by putting in Theorem 4.1 $u_2=\overline{u}_1$ and $\mu=1$.

THEOREM 4.2. *Let $A_1,A_2\in\mathbb{R}^{n\times n}$ be symmetric.*

(*i*) $\operatorname{csd}(A_1,A_2)=1$ *if and only if* $\operatorname{rank}(A_1A_2-A_2A_1)=2$ *and the equation*

$$PA_1\operatorname{Im}(t_1u_1)=PA_2\operatorname{Re}(t_2u_1) \tag{4.8}$$

has a solution pair $t_1,t_2\in\mathbb{C}$ satisfying

$$\operatorname{Re}(\overline{t}_1t_2)=d. \tag{4.9}$$

Here $u_1\in\mathbb{C}^n$ is the unit eigenvector of the matrix $2i(A_1A_2-A_2A_1)$ corresponding to its eigenvalue $\lambda_1=d(>0)$, and $P=I_n-u_1u_1^*-\overline{u}_1u_1^T$.

(*ii*) *If* $\operatorname{csd}(A_1,A_2)=1$, t_1 *and* t_2 *satisfy* (4.8) *and* (4.9) *then the pair* (B_1,B_2) *of matrices*

$$B_1=\begin{bmatrix}A_1 & \operatorname{Re}(t_2u_1)\\ \operatorname{Re}(t_2u_1)^T & z_1\end{bmatrix}, \tag{4.10}$$

$$B_2=\begin{bmatrix}A_2 & \operatorname{Im}(t_1u_1)\\ \operatorname{Im}(t_1u_1)^T & z_2\end{bmatrix} \tag{4.11}$$

is a minimal commuting symmetric completion of (A_1,A_2). Here

$$z_1=u_1^*A_1u_1-\frac{1}{d}\left(\operatorname{Im}(t_2^2u_1^TA_2u_1)+\operatorname{Re}(t_1t_2u_1^TA_1u_1)\right) \tag{4.12}$$

and

$$z_2=u_1^*A_2u_1-\frac{1}{d}\left(\operatorname{Im}(t_1^2u_1^TA_1u_1)-\operatorname{Re}(t_1t_2u_1^TA_2u_1)\right). \tag{4.13}$$

All minimal commuting symmetric completions of (A_1,A_2) arise in this way.

Proof. (*i*) By (4.7), if $\operatorname{csd}(A_1,A_2)=1$ then $\operatorname{chd}(A_1,A_2)=1$. Therefore, by Theorem 4.1, $\operatorname{rank}(A_1A_2-A_2A_1)=2$ and equation (4.1) has a solution pair $t_1,t_2\in\mathbb{C}$ satisfying (4.2). If u_1 is the unit eigenvector of the Hermitian matrix $2i(A_1A_2-A_2A_1)$ corresponding to its eigenvalue $\lambda_1=d(>0)$ then $\overline{u}_1$ is the unit eigenvector corresponding to the eigenvalue $\lambda_2=-d$. Thus, we can choose in Theorem 4.1 $u_2=\overline{u}_1$. Then $P=I_n-u_1u_1^*-\overline{u}_1u_1^T$ is a real $n\times n$ matrix, and equation (4.1) becomes (4.8).

Conversely, if $\operatorname{rank}(A_1A_2-A_2A_1)=2$ and equation (4.8) (which is equivalent to (4.1) in our case) has a solution pair $t_1,t_2\in\mathbb{C}$ satisfying (4.9) (= (4.2)) then by Theorem 4.1, $\operatorname{chd}(A_1,A_2)=1$. For any such t_1,t_2 the corresponding minimal commuting Hermitian completions (B_1,B_2) of (A_1,A_2) have the form (4.3)–(4.4). We observe that since $u_2=\overline{u}_1$, the matrices B_1 and B_2 are real symmetric if and only if $\mu=1$ or $\mu=-1$. Consequently, $\operatorname{csd}(A_1,A_2)=1$.

(*ii*) If $t_1, t_2 \in \mathbb{C}$ satisfy (4.8)–(4.9) then so do $t_1' = -t_1$ and $t_2' = -t_2$. Therefore, we do not miss any minimal commuting symmetric completions of (A_1, A_2) if in (4.3)–(4.4) we choose t_1, t_2 as above and fix $\mu = 1$. Finally, since (B_1, B_2) constructed in Theorem 4.1 with $\mu = 1$ has in our case the form (4.10)–(4.11), and (4.5)–(4.6) become (4.12)–(4.13), this completes the proof. $\square$

REMARK 4.3. The procedure for checking whether $\operatorname{csd}(A_1, A_2) = 1$, and if this is the case — for finding all minimal commuting symmetric completions of a pair of symmetric matrices (A_1, A_2), can be obtained as the specialization of the procedure mentioned in Section 4.1 (which, in turn, is based on the procedure from Section 2.2) by setting $u_2 = \overline{u}_1$ and $\mu = 1$, see Theorem 4.2 and its preceding paragraph.

5. The separability problem

In the 1980s the use of quantum systems as computing devices started to being explored. The idea gained momentum when Peter Shor [11] presented a quantum algorithm for factoring a large composite integer N that was polynomial in the number of digits in N. An excellent overview article on the subject of quantum computing is [1].

The separability problem occurs when a quantum system is divided into parts. For convenience we consider a bipartite system. The state of the system is described by a density matrix M, a positive semidefinite matrix with trace 1. A state is called *separable* when it can be written as a convex combination of so-called pure separable states, i.e., $\rho = \sum_{i=1}^{k} p_i \, \psi_i \psi_i^* \otimes \phi_i \phi_i^*$ where ψ_i and ϕ_i are (nonzero) state vectors in the spaces corresponding to two parts of the system, and $p_i > 0$. When $\psi_i \in \mathbb{C}^m$ and $\phi_i \in \mathbb{C}^n$, the matrix ρ is called $m \times n$ separable. The number k is referred to as the number of states in the representation.

The problem whether a given state is separable or entangled (= not separable) may be stated as a semi-algebraic one, and is therefore decidable by the Tarski-Seidenberg decision procedure [2]. As it turns out though, the separability problem scales very poorly with the number of variables and these techniques are in general not practical. In fact, the separability problem in its full generality has been shown to be NP-complete [5].

As a consequence of the results of Section 2 we can state a new result for the $2 \times n$ case. Thus we are concerned with matrices

$$M = \begin{bmatrix} A & B^* \\ B & C \end{bmatrix} \geqslant 0. \tag{5.1}$$

Notice that if $M = \sum_{i=1}^{k} P_i \otimes Q_i$ with $P_i \in \mathbb{C}^{2\times 2}$ and $Q_i \in \mathbb{C}^{n\times n}$ positive semidefinite, then $\widetilde{M} = \sum_{i=1}^{k} P_i^T \otimes Q_i$ is positive semidefinite as well. One easily sees that

$$\widetilde{M} = \begin{bmatrix} A & B \\ B^* & C \end{bmatrix} \geqslant 0. \tag{5.2}$$

Thus for (5.1) to have a chance to be $2 \times n$ separable we need (5.2) to hold. This is referred to as the "Peres test"; see [10]. As was observed by several authors, the

$2 \times n$ separability problem for (5.1) can easily be reduced to the case when $A = I$; see, for instance, Proposition 3.1 in [13]. Using Theorem 3.2 in [13], which connects the separability problem to the normal completion problem, we can now state a method for checking separability of (5.1) in the case when $\operatorname{rank}(M) = \operatorname{rank}(\widetilde{M}) = \operatorname{rank}(A) + 1$.

THEOREM 5.1. *Let $B,\ C \in \mathbb{C}^{n \times n}$ be such that*

$$M = \begin{bmatrix} I_n & B^* \\ B & C \end{bmatrix} \geqslant 0, \quad \widetilde{M} = \begin{bmatrix} I_n & B \\ B^* & C \end{bmatrix} \geqslant 0,$$

and suppose that $\operatorname{rank}(M) = \operatorname{rank}(\widetilde{M}) = n + 1$. *Write*

$$C - BB^* = xx^*, \quad C - B^*B = yy^*$$

for some vectors $x, y \in \mathbb{C}^n$. Then M is $2 \times n$ separable if and only if x, y, B^*x, By are linearly dependent. In this case, the minimal number of states in a separable representation of M is $n+1$.

Proof. First notice that $B^*B - BB^* = xx^* - yy^*$.

Suppose that x, y, B^*x, By are linearly dependent. Then by Theorem 2.4 there exists a normal matrix

$$N = \begin{bmatrix} B & \nu x \\ y^* & z \end{bmatrix},$$

where $|\nu| = 1$. But as $(\nu x)(\nu x)^* = C - BB^*$ it follows from Theorem 3.2 in [13] that M is $2 \times n$ separable, and that the minimal number of states in a separable representation of M is $n+1$.

Conversely, suppose that M is $2 \times n$ separable. By Theorem 3.2 in [13] there exists a normal matrix

$$N = \begin{bmatrix} B & S \\ T & P \end{bmatrix}$$

so that $BB^* + SS^* \leqslant C$. But then $SS^* \leqslant xx^*$ and thus $S = xv^*$ with $\|v\| \leqslant 1$. Also $B^*B + T^*T = BB^* + SS^* \leqslant C$, and thus $T^*T \leqslant yy^*$ yielding $T = yw^*$ with $\|w\| \leqslant 1$. In addition, $BT^* + SP^* = B^*S + T^*P$. In particular, $\operatorname{range}(BT^* + SP^*) = \operatorname{range}(B^*S + T^*P)$. Note that $\operatorname{range}(BT^* + SP^*) \subseteq \operatorname{span}(By, x)$ and $\operatorname{range}(B^*S + T^*P) \subseteq \operatorname{span}(B^*x, y)$. But then it follows easily that x, y, B^*x, By are linearly dependent. Indeed, if $BT^* + SP^* = B^*S + T^*P \neq 0$, then $\operatorname{span}(By, x)$ and $\operatorname{span}(B^*x, y)$ must have a nontrivial intersection, and if $BT^* + SP^* = B^*S + T^*P = 0$, then $\operatorname{span}(By, x)$ and $\operatorname{span}(B^*x, y)$ are both at most one dimensional. $\square$

We can now provide a new proof of the following result by Woronowicz [14].

THEOREM 5.2. *Let A, B, C be $n \times n$ matrices with $n \leqslant 3$, so that M and $\widetilde{M}$ are as in* (5.1)–(5.2). *Then M is $2 \times n$ separable.*

We will use a result by Hildebrand which we quote without proof.

LEMMA 5.3. [6, Lemma 2.6] *Let K be a convex cone in a real vector space $\mathscr{H}$ of finite dimension N, and let $\mathscr{L} \subseteq \mathscr{H}$ be a subspace of dimension n. Let $K' = K \cap \mathscr{L}$ and x the generator of an extreme ray in K'. Then the minimal face in K containing x has dimension at most $N - n + 1$.*

Notice that if we consider the cone PSD_n of $n \times n$ complex positive semidefinite matrices, then the minimal face containing $M \geqslant 0$, is the cone $F = \{GCG^* : C \in PSD_k\}$, where $M = GG^*$ with $\mathrm{null}(G) = \{0\}$, and $k = \mathrm{rank}(M)$. In particular, the real dimension of this minimal face is $(\mathrm{rank}(M))^2$.

Proof of Theorem 5.2. Since the case $n < 3$ can be embedded into the case $n = 3$, we will focus on the latter. As the $2 \times n$ separable matrices form a convex cone, it suffices to prove the result for pairs $(M, \widetilde{M})$ that generate extreme rays in the cone of pairs of matrices as in (5.1)–(5.2). If we apply Lemma 5.3 with the choices of $K = PSD_6 \times PSD_6$ and $\mathscr{L}$ the subspace

$$\left\{ \left(\begin{bmatrix} A & B^* \\ B & C \end{bmatrix}, \begin{bmatrix} A & B \\ B^* & C \end{bmatrix} \right) \right\}$$

in the (real) vector space of pairs of Hermitian matrices of size 6×6, then K' is the cone of pairs of matrices as in (5.1)–(5.2). By Lemma 5.3 the minimal faces in K containing extreme rays of K' cannot have dimension greater than $72 - 36 + 1 = 37$. However, the minimal face in K containing $(M, \widetilde{M})$ (which generates an extreme ray in K') has dimension $(\mathrm{rank}(M))^2 + (\mathrm{rank}(\widetilde{M}))^2$, and hence the vector $(\mathrm{rank}(M), \mathrm{rank}(\widetilde{M})) \in \mathbb{R}^2$ lies in the closed disk of radius $\sqrt{37}$ centered at the origin. This now gives that either $\min\{\mathrm{rank}(M), \mathrm{rank}(\widetilde{M})\} \leqslant 3$ or $\max\{\mathrm{rank}(M), \mathrm{rank}(\widetilde{M})\} = 4$. Next, as in Proposition 3.1 in [13] we can assume that $A = I$. If now $\min\{\mathrm{rank}(M), \mathrm{rank}(\widetilde{M})\} \leqslant 3$ we have that $C = BB^* = B^*B$, and thus B is normal, which yields by Theorem 3.2 in [13] that M is $2 \times n$ separable. On the other hand, if $\max\{\mathrm{rank}(M), \mathrm{rank}(\widetilde{M})\} = 4$ we can conclude by Theorem 5.1 that M is $2 \times n$ separable (as 4 vectors in $\mathbb{C}^n$ are always linearly dependent when $n \leqslant 3$). $\square$

It should be noted that the original statement of Woronowicz is formulated in the dual form: if $\Phi : \mathbb{C}^{2\times 2} \to \mathbb{C}^{n\times n}$ is a positive linear map (thus $\Phi(PSD_2) \subseteq PSD_n$) and $n \leqslant 3$, then Φ must be decomposable. That is, Φ must be of the form $\Phi(M) = \sum_{i=1}^k R_i M R_i^* + \sum_{i=1}^l S_i M^T S_i^*$.

Acknowledgement. The authors wish to thank Dr. Roland Hildebrand for pointing out how Theorem 5.1 in conjunction with his results, leads to a proof of the result by Woronowicz.

REFERENCES

[1] C. H. BENNETT AND D. P. DIVINCENZO, *Quantum information and computation*, *Nature*, **404** (2000), 247–255.

[2] J. BOCHNAK, M. COSTE AND M.-F. ROY, *Real algebraic geometry*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol. 36, Springer-Verlag, Berlin, 1998, Translated from the 1987 French original, Revised by the authors.

[3] I. DEGANI, J. SCHIFF AND D. J. TANNOR, *Commuting extensions and cubature formulae*, Numer. Math., **101**, 3 (2005), 479–500.
[4] F. R. GANTMACHER, *The theory of matrices*. Vols. 1, 2. Translated by K. A. Hirsch. Chelsea Publishing Co., New York, 1959 Vol. 1, x+374 pp. Vol. 2, ix+276 pp.
[5] L. GURVITS, *Classical complexity and quantum entanglement*, J. Computer and System Sciences, **69** (2004), 448–484.
[6] R. HILDEBRAND, *Semidefinite descriptions of low-dimensional separable matrix cones*, Linear Algebra Appl. **429** (2008), 901–932.
[7] R. A. HORN AND C. R. JOHNSON, *Matrix analysis*. Corrected reprint of the 1985 original. Cambridge University Press, Cambridge, 1990.
[8] D. KIMSEY AND H. J. WOERDEMAN, *Minimal normal and commuting completions*, Int. J. Inf. Syst. Sci., **4**, 1 (2008), 50–59.
[9] S. M. MALAMUD, *Inverse spectral problem for normal matrices and the Gauss–Lucas theorem*, Trans. Amer. Math. Soc., **357**, 10 (2005), 4043–4064.
[10] A. PERES, *Separability criterion for density matrices*, Phys. Rev. Lett., **77**, 8 (1996) 1413–1415.
[11] P. W. SHOR, *Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer*, SIAM J. Comput., **26**, 5 (1997), 1484–1509.
[12] H. J. WOERDEMAN, *Hermitian and normal completions*, Linear and Multilinear Algebra, **42**, 3 (1997), 239–280.
[13] H. J. WOERDEMAN, *The separability problem and normal completions*, Linear Algebra Appl., **376** (2004), 85–95.
[14] S. L. WORONOWICZ, *Positive maps of low dimensional matrix algebras*, Rep. Math. Phys., **10**, 2 (1976), 165–183.

Dmitry S. Kaliuzhnyi-Verbovetskyi
Department of Mathematics
Drexel University
3141 Chestnut Str.
Philadelphia, PA, 19104
U.S.A.
e-mail: dmitryk@math.drexel.edu

Ilya M. Spitkovsky
Department of Mathematics
College of William and Mary
Williamsburg, VA, 23187-8795
U.S.A.
e-mail: ilya@math.wm.edu

Hugo J. Woerdeman
Department of Mathematics
Drexel University
3141 Chestnut Str.
Philadelphia, PA, 19104
U.S.A.
e-mail: hugo@math.drexel.edu