

# SOLVING THE PROBLEM OF SIMULTANEOUS DIAGONALIZATION OF COMPLEX SYMMETRIC MATRICES VIA CONGRUENCE\*

MIGUEL D. BUSTAMANTE<sup>†</sup>, PAULINE MELLON<sup>‡</sup>, AND M. VICTORIA VELASCO<sup>‡</sup>

**Abstract.** We provide a solution to the problem of simultaneous *diagonalization via congruence* of a given set of  $m$  complex symmetric  $n \times n$  matrices  $\{A_1, \dots, A_m\}$ , by showing that it can be reduced to a possibly lower-dimensional problem where the question is rephrased in terms of the classical problem of simultaneous *diagonalization via similarity* of a new related set of matrices. We provide a procedure to determine in a finite number of steps whether or not a set of matrices is simultaneously diagonalizable by congruence. This solves a long-standing problem in the complex case.

**Key words.** simultaneous diagonalization by congruence, simultaneous diagonalization by similarity, linear pencil

**AMS subject classifications.** 15A, 65K, 90C, 94A

**DOI.** 10.1137/19M1280430

**1. Introduction.** The aim of this paper is to characterize when a given set of  $n \times n$  complex symmetric matrices,  $A_1, \dots, A_m$ , are simultaneously diagonalizable via congruence (SDC), namely, when there exists a nonsingular  $n \times n$  complex matrix  $P$  such that

$$P^T A_i P \text{ is diagonal for all } 1 \leq i \leq m.$$

Since Weierstrass in 1868 [24] gave sufficient conditions for the simultaneous diagonalization by congruence of two real symmetric matrices, several authors [7, 9, 14] have extended those results and there have been applications of the real results to areas as diverse as quadratic programming [13, 27, 2], variational analysis [8], signal processing [15, 16, 20], and medical imaging analysis [1, 5, 17, 23], among others. In the case of complex matrices, Hong, Horn, and Johnson laid the framework in the 1980s for the particular case of unitary transformations by proving [10, 11] that there is a unitary  $U$  satisfying that  $U^T A_i U$  is diagonal for all  $i$  if, and only if, the set  $\{A_i \overline{A_j} : 1 \leq i, j \leq m\}$  is a commuting family. In a more general case they solved the problem for pairs of complex symmetric (or Hermitian) matrices with the restriction that at least one of them be nonsingular. We provide a solution in the general case of complex symmetric matrices by translating it into a simpler problem, at a possibly reduced dimension, regarding simultaneous diagonalizability by similarity of a new set of related matrices. We do this using the concept of matrix pencils so that the general problem is reduced to (possibly) lower dimensions a priori by calculating the intersection of the kernels of the matrices  $A_1, \dots, A_m$ . Once this is done, reduced  $r \times r$

\*Received by the editors August 12, 2019; accepted for publication (in revised form) by F. Tisseur July 10, 2020; published electronically October 20, 2020.

<https://doi.org/10.1137/19M1280430>

**Funding:** The work of the authors was partially supported by the AEI/FEDER, UE Spanish project MTM2016-76327-C3-2-P. The work of the third author was also partially supported by the Distinguished Visitor Grant of the School of Mathematics and Statistics, University College Dublin.

<sup>†</sup>School of Mathematics and Statistics, University College Dublin, Belfield, Dublin 4, Ireland (miguel.bustamante@ucd.ie, pauline.mellon@ucd.ie).

<sup>‡</sup>Departamento de Análisis Matemático, Facultad de Ciencias, Universidad de Granada, 18071 Granada, Spain (vvelasco@ugr.es).

matrices ( $r \leq n$ )  $\tilde{A}_i$  (Lemma 10 below) can be dealt with in a more standard manner, thanks to the existence of a nonsingular matrix pencil. This allows us to obtain fairly simple necessary and sufficient conditions for SDC in Theorem 14 below.

The authors were initially motivated to tackle SDC for complex symmetric matrices by a problem that arose naturally in the area of evolution algebras.

We recall here that an evolution algebra is defined as a commutative algebra  $A$  over  $\mathbb{C}$  for which there exists a basis  $\tilde{B} = \{\tilde{e}_i : i \in \Lambda\}$  such that  $\tilde{e}_i \tilde{e}_j = 0$ , for every  $i, j \in \Lambda$  with  $i \neq j$ . In other words, the multiplication table of  $A$  relative to  $\tilde{B}$  is diagonal. Such a basis is called natural. Evolution algebras were introduced in [19] and [18] in the study of non-Mendelian genetics and are, in general, not associative. The problem concerning the authors was to determine when a given algebra  $A$  is an evolution algebra. In other words, if  $B$  is a basis of  $A$  and the multiplication table of  $A$  with respect to  $B$  is not diagonal, we established the conditions under which there exists a natural basis  $\tilde{B}$  of  $A$ , giving  $A$  the structure of an evolution algebra. If  $B = \{e_1, \dots, e_n\}$  and

$$(1.1) \quad e_i e_j = \sum_{k=1}^n m_{ijk} e_k, \quad i, j = 1, \dots, n,$$

we define the *structure matrices* of  $A$  with respect to  $B$  as the  $n \times n$  matrices  $M_k(B) = (m_{ijk})_{1 \leq i, j \leq n}$  for  $k = 1, \dots, n$ . Notice that the structure matrices  $M_k(B)$  are symmetric and complex because  $A$  is commutative and its base field is  $\mathbb{C}$ . A main result in [4] proves that  $A$  is an evolution algebra if, and only if, the complex symmetric matrices  $M_1(B), \dots, M_n(B)$  are simultaneously diagonalizable via congruence.

There are other areas where the complex results might be applied. One of the most important applications is in the area of signal processing, in particular in the classical problem of blind source separation [3, 25, 26]. In its simplest form, and appropriate to our notation, this problem amounts to finding a nonsingular complex matrix  $Q$  relating  $n$  sets of measurements (denoted by the complex random vector  $x$  of dimension  $n$ ) and  $n$  statistically independent, *but unknown*, sources (denoted by the complex random vector  $s$  of dimension  $n$ ), via the linear relation  $x = Q^* s$ , where  $Q^*$  denotes the conjugate transpose of  $Q$ . To investigate how this relates to the SDC problem we discuss a method introduced in [25]. Consider a generalized second characteristic function, defined in terms of the  $x$  variables as

$$\psi_x(\tau) := \ln E[\exp(\tau^T \bar{x})], \quad \tau \in \mathbb{C}^n,$$

where bar denotes complex conjugation and  $E$  denotes the expectation. In terms of the  $s$  variables, this reads

$$\psi_x(\tau) = \ln E[\exp((Q\tau)^T \bar{s})] =: \psi_s(\mu), \quad \mu = Q\tau.$$

Let us consider  $m$  so-called processing points  $\tau^{(1)}, \dots, \tau^{(m)}$ . We now define the following complex symmetric matrices  $A_1, \dots, A_m$  by their components:

$$(A_j)_{k\ell} := \left. \frac{\partial^2 \psi_x(\tau)}{\partial \tau_k \partial \tau_\ell} \right|_{\tau=\tau^{(j)}}, \quad k, \ell = 1, \dots, n, \quad j = 1, \dots, m.$$

It is then easy to show

$$A_j = Q^T D_j Q, \quad (D_j)_{k\ell} := \left. \frac{\partial^2 \psi_s(\mu)}{\partial \mu_k \partial \mu_\ell} \right|_{\mu=P\tau^{(j)}}, \quad k, \ell = 1, \dots, n, \quad j = 1, \dots, m,$$

where the matrices  $D_j$  are diagonal, due to the statistical independence of the components of  $s$ . This is, of course, the SDC problem, and its solution provides the complex matrix  $Q$  that allows one to unveil the unknown independent sources starting from an arbitrary set of measurements. In real-life applications, experimental or numerical errors will lead to matrices  $A_1, \dots, A_m$  that are not exactly SDC, so “approximate joint diagonalization” is the correct concept, which consists of the variational problem of finding a complex nonsingular matrix  $P$  such that  $P^T A_j P$  is as diagonal as possible, in some metric (see, for example, [3, 6, 26]).

In summary, the solution to the SDC problem provided extends earlier results from the 1980s, solves the initial motivating question for the authors related to evolution algebras, and may have an impact on applications in optimization or signal processing as described above.

In section 2 we provide notation and definitions. In section 3 we solve the SDC problem in the case of complex symmetric matrices and present a finite step procedure to determine whether a given set of matrices is SDC or not. In section 4 we discuss possible avenues of further research.

**2. Notation.** Let  $\mathcal{M}_{n,m}$  denote all  $n \times m$  matrices over  $\mathbb{C}$ . Let  $\mathcal{M}_n := \mathcal{M}_{n,n}$ , let  $\mathcal{MS}_n$  be all symmetric elements in  $\mathcal{M}_n$ , and let  $\mathcal{GL}_n$  be all invertible elements in  $\mathcal{M}_n$ . A diagonal matrix in  $\mathcal{M}_n$  with diagonal entries  $d_1, \dots, d_n$  will be written as  $D = \text{diag}(d_1, \dots, d_n)$ . For  $A \in \mathcal{M}_n$  we denote its  $ij$  component by  $A_{ij}$  or  $(A)_{ij}$  and the zero and identity element in  $\mathcal{M}_n$  are denoted  $0_n$  and  $I_n$ , respectively. We recall that  $A \in \mathcal{M}_n$  is said to be orthogonal if  $A^T = A^{-1}$  (where  $A^T$  denotes the usual transpose of  $A$ ) and is said to be unitary if  $\bar{A}^T = A^{-1}$  (where  $\bar{A}$  denotes the entrywise complex conjugate of  $A$ ); matrices  $A, B \in \mathcal{M}_n$  are said to be congruent if there exists  $P \in \mathcal{GL}_n$  such that  $P^T A P = B$  and are said to be similar if there exists  $P \in \mathcal{GL}_n$  such that  $P^{-1} A P = B$ . Congruent (or similar) matrices have the same rank. In fact,  $A$  and  $B$  are congruent if, and only if, they have the same rank [12, Theorem 4.5.12], and hence  $A$  similar to  $B$  implies  $A$  congruent to  $B$  but the converse does not hold. We introduce the following definitions for a set of matrices in  $\mathcal{M}_n$ .

**DEFINITION 1.** Let  $A_1, \dots, A_m \in \mathcal{M}_n$ . We say  $A_1, \dots, A_m$  are simultaneously diagonalizable via congruence (SDC) if there exists  $P \in \mathcal{GL}_n$  and diagonal matrices  $D_1, \dots, D_m \in \mathcal{M}_n$  such that

$$P^T A_j P = D_j, \quad j = 1, \dots, m.$$

Of course, if  $A_1, \dots, A_m$  are SDC, then they are necessarily symmetric.

**DEFINITION 2.** Let  $L_1, \dots, L_m \in \mathcal{M}_n$ . We say  $L_1, \dots, L_m$  are simultaneously diagonalizable via similarity (SDS) if there exists  $P \in \mathcal{GL}_n$  and diagonal matrices  $D_1, \dots, D_m \in \mathcal{M}_n$  such that

$$P^{-1} L_j P = D_j, \quad j = 1, \dots, m.$$

It is important to remark that even when  $A_1, \dots, A_m$  in Definition 1 or  $L_1, \dots, L_m$  in Definition 2 are real, the resulting matrices  $P$  and  $D_j$  may have to be complex, as illustrated in Example 16 below. The following result is well known (see, for instance, [12, Theorems 1.3.12 and 1.3.21]) and means that SDS is easy to check in practice, in contrast to SDC.

**THEOREM 3.** Let  $L_1, \dots, L_m \in \mathcal{M}_n$ . These matrices are SDS if, and only if, they are all diagonalizable by similarity and they pairwise commute.

**3. Solving the SDC problem.** Let  $S^{2m-1} := \{x \in \mathbb{C}^m : \|x\| = 1\}$ , where  $\|\cdot\|$  denotes the usual Euclidean norm. We use the standard concepts of linear pencil and maximum pencil rank.

DEFINITION 4. Let  $A_1, \dots, A_m \in \mathcal{M}_n$ . Define the associated linear pencil to be the map

$$A : \mathbb{C}^m \longrightarrow \mathcal{M}_n \text{ by } A(\lambda) = \sum_{j=1}^m \lambda_j A_j, \text{ where } \lambda = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix} \in \mathbb{C}^m.$$

Since  $\text{rank} A(\lambda) = \text{rank} A(\frac{\lambda}{\|\lambda\|})$ , for  $\lambda \neq 0$ , it follows that

$$\sup_{\lambda \in \mathbb{C}^m} \text{rank} A(\lambda) = \sup_{\lambda \in S^{2m-1}} \text{rank} A(\lambda).$$

In addition, since  $\{\text{rank} A(\lambda) : \lambda \in S^{2m-1}\} \subseteq \{0, 1, \dots, n\}$ , it follows that the above supremum must be achieved. In other words, there exists some  $\lambda_0 \in S^{2m-1}$  such that

$$\sup_{\lambda \in \mathbb{C}^m} \text{rank} A(\lambda) = \sup_{\lambda \in S^{2m-1}} \text{rank} A(\lambda) = \text{rank} A(\lambda_0).$$

DEFINITION 5. Let  $A_1, \dots, A_m \in \mathcal{M}_n$ . The rank of the associated linear pencil is  $r := \sup_{\lambda \in \mathbb{C}^m} \text{rank} A(\lambda)$ . We refer to  $r$  as the (maximum pencil) rank of  $A_1, \dots, A_m$  and denote it as  $r = \text{rank}(A_1, \dots, A_m)$ . In the case that  $r = n$  we say that the pencil is nonsingular. From above,  $r = \text{rank}(A_1, \dots, A_m) = \text{rank} A(\lambda_0)$  for some  $\lambda_0 \in S^{2m-1}$ .

The following simple lemma is important.

LEMMA 6. Let  $A_1, \dots, A_m \in \mathcal{M}_n$  and let  $r = \text{rank}(A_1, \dots, A_m) = \text{rank} A(\lambda_0)$  for some  $\lambda_0 \in S^{2m-1}$ . Then

$$\dim \left( \bigcap_{j=1}^m \ker A_j \right) = n - r \text{ if, and only if, } \bigcap_{j=1}^m \ker A_j = \ker A(\lambda_0).$$

*Proof.* Clearly  $\bigcap_{j=1}^m \ker A_j \subseteq \ker A(\lambda)$  and hence

$$(3.1) \quad \dim \left( \bigcap_{j=1}^m \ker A_j \right) \leq n - \text{rank} A(\lambda) \text{ for all } \lambda \in \mathbb{C}^m.$$

In particular, for maximum pencil rank  $r = \text{rank} A(\lambda_0)$  we have  $\bigcap_{j=1}^m \ker A_j \subseteq \ker A(\lambda_0)$ , and  $\dim(\ker A(\lambda_0)) = n - r$  then gives the result.  $\square$

We will see later that  $\dim(\bigcap_{j=1}^m \ker A_j) = n - r$  is necessary for  $A_1, \dots, A_m$  to be SDC and, in this case, it follows from the above that the subspace  $\ker A(\lambda_0)$  is actually independent of the point  $\lambda_0$  satisfying  $r = \text{rank} A(\lambda_0)$ .

### 3.1. The SDC problem for $n \times n$ matrices with nonsingular pencils.

We now solve the SDC problem for symmetric matrices  $A_1, \dots, A_m \in \mathcal{M}_n$ , in the particular case that  $\text{rank}(A_1, \dots, A_m) = n$ . The proof follows ideas from [12, Theorem 4.5.17], [14, Lemma 1], and [22, p. 230]. In particular, the simple observation that if  $A(\lambda)$  is invertible, then

$$(P^T A(\lambda) P)(P^{-1} A(\lambda)^{-1} A_j P) = P^T A_j P, \text{ for } j = 1, \dots, m, \text{ and any } P \in \mathcal{GL}_n$$

motivates our first main result and proves it in the obvious direction.

THEOREM 7. Let  $A_1, \dots, A_m \in \mathcal{MS}_n$  have maximum pencil rank  $n$ . For any  $\lambda_0 \in \mathbb{C}^m$  with  $\text{rank} A(\lambda_0) = n$  then

$A_1, \dots, A_m$  are SDC if, and only if,  $A(\lambda_0)^{-1}A_1, \dots, A(\lambda_0)^{-1}A_m$  are SDS.

*Proof.* Let  $\lambda_0 \in \mathbb{C}^m$  satisfy  $\text{rank} A(\lambda_0) = n$ . In the forward direction, we assume that  $A_1, \dots, A_m$  are SDC and let  $P \in \mathcal{GL}_n$  satisfy that  $P^T A_j P$  is diagonal for  $j = 1, \dots, m$ . Then  $P^T A(\lambda_0) P$  is diagonal and invertible giving that

$$P^{-1} A(\lambda_0)^{-1} A_j P = (P^T A(\lambda_0) P)^{-1} (P^T A_j P)$$

is diagonal for  $1 \leq j \leq m$  and we are done.

In the opposite direction, assume that  $A(\lambda_0)^{-1}A_1, \dots, A(\lambda_0)^{-1}A_m$  are SDS and let  $P \in \mathcal{GL}_n$  satisfy that  $D^{(j)} := P^{-1} A(\lambda_0)^{-1} A_j P$  is diagonal for  $j = 1, \dots, m$ . We define symmetric matrices  $B_j := P^T A_j P$  and  $B(\lambda_0) := P^T A(\lambda_0) P$  to give

$$(3.2) \quad B_j = B(\lambda_0) D^{(j)}, \quad j = 1, \dots, m.$$

Taking the transpose of this latter equation implies that  $B(\lambda_0)$  commutes with  $D^{(j)}$  for  $j = 1, \dots, m$ . Componentwise, this means that for all  $1 \leq k, l \leq n$ ,

$$(B(\lambda_0))_{k\ell} (D^{(j)})_{\ell\ell} = (D^{(j)})_{kk} (B(\lambda_0))_{k\ell} \text{ for all } j = 1, \dots, m.$$

In particular, for all  $1 \leq k, l \leq n$ ,

$$(3.3) \quad (B(\lambda_0))_{k\ell} = 0 \quad \text{if} \quad (D^{(j)})_{kk} \neq (D^{(j)})_{\ell\ell} \text{ for any } j = 1, \dots, m.$$

Write  $D^{(j)} = \text{diag}(\alpha_1^j, \dots, \alpha_n^j)$ , for  $1 \leq j \leq m$ , and let  $p_j$  satisfy

$$\alpha_1^j = \dots = \alpha_{p_j}^j \neq \alpha_{p_j+1}^j$$

( $p_j$  is the length of the first run of identical diagonals in  $D^{(j)}$ ) and define  $n_1 := \min_{1 \leq j \leq m} p_j$ . Define  $\alpha_1^{(j)} := \alpha_1^j$  and  $\alpha_2^{(j)} := \alpha_{n_1+1}^j$  so that

$$\alpha_1^{(j)} = \alpha_1^j = \dots = \alpha_{n_1}^j \text{ for all } 1 \leq j \leq m,$$

and we may write

$$D^{(j)} = \alpha_1^{(j)} I_{n_1} \oplus \text{diag}(\alpha_2^{(j)}, \alpha_{n_1+2}^j, \dots, \alpha_n^j) \text{ for all } 1 \leq j \leq m,$$

and there is some  $j \in \{1, \dots, m\}$  for which  $\alpha_1^{(j)} \neq \alpha_2^{(j)}$  ( $I_n$  denotes the  $n \times n$  identity matrix). We repeat a similar process twice more on  $\text{diag}(\alpha_2^{(j)}, \alpha_{n_1+2}^j, \dots, \alpha_n^j)$  to find  $n_2, n_3$ , and  $\alpha_3^{(j)} := \alpha_{n_1+n_2+1}^j$  so that, for all  $1 \leq j \leq m$ ,

$$D^{(j)} = \alpha_1^{(j)} I_{n_1} \oplus \alpha_2^{(j)} I_{n_2} \oplus \alpha_3^{(j)} I_{n_3} \oplus \text{diag}(\alpha_{n_1+n_2+n_3+1}^j, \dots, \alpha_n^j),$$

while there is some  $j$  for which  $\alpha_1^{(j)} \neq \alpha_2^{(j)}$  and some  $k$  for which  $\alpha_2^{(k)} \neq \alpha_3^{(k)}$ .

If now  $\alpha_1^{(j)} = \alpha_3^{(j)}$  for all  $1 \leq j \leq m$ , then we may reorder the diagonal entries to amalgamate  $\alpha_1^{(j)} I_{n_1}$  and  $\alpha_3^{(j)} I_{n_3}$ , namely, there is an orthogonal permutation matrix  $R \in \mathcal{GL}_n$  with

$$R^T D^{(j)} R = R^{-1} D^{(j)} R = \alpha_1^{(j)} I_{n_1+n_3} \oplus \alpha_2^{(j)} I_{n_2} \oplus \text{diag}(\alpha_{n_1+n_2+n_3+1}^j, \dots, \alpha_n^j)$$

for all  $1 \leq j \leq m$ .

We continue this process of finding  $\alpha_l^{(j)}$ s for  $\text{diag}(\alpha_{n_1+n_2+n_3+1}^j, \dots, \alpha_n^j)$ . We amalgamate, as described above, where necessary so that for some orthogonal  $U \in \mathcal{M}_n$  and all  $1 \leq j \leq m$  we have

$$(3.4) \quad U^T D^{(j)} U = \alpha_1^{(j)} I_{p_1} \oplus \dots \oplus \alpha_d^{(j)} I_{p_d}$$

subject to the condition that for  $1 \leq a < b \leq d$ , there is some  $j \in \{1, \dots, m\}$  with  $\alpha_a^{(j)} \neq \alpha_b^{(j)}$ . Of course,  $d \leq n$  and  $d$  is as small as possible satisfying the above.

We now write  $U^T B(\lambda_0) U$  as a  $d \times d$  block matrix, whose  $(a, b)$  sub-block, denoted here  $[U^T B(\lambda_0) U]_{ab}$ , is of size  $n_a \times n_b$  for  $1 \leq a, b \leq d$ . Then  $U^T B(\lambda_0) U$  commutes with  $U^T D^{(j)} U$  for all  $1 \leq j \leq m$ , since  $B(\lambda_0)$  commutes with  $D^{(j)}$  for all  $1 \leq j \leq m$ , and  $U$  is orthogonal. This commutativity then yields a block version of (3.3).

Specifically, since

$$[U^T D^{(j)} U]_{aa} = \alpha_a^{(j)} I_{n_a}$$

and we have from (3.4) that if  $a < b$ , then  $\alpha_a^{(j)} \neq \alpha_b^{(j)}$ , we get

$$(3.5) \quad [U^T B(\lambda_0) U]_{ab} = 0 \text{ if } a \neq b.$$

In other words, we have a block diagonal decomposition

$$(3.6) \quad U^T B(\lambda_0) U = C_1 \oplus \dots \oplus C_d,$$

where  $C_a \in \mathcal{MS}_{n_a}$ ,  $a = 1, \dots, d$ . As each  $C_a$  must be symmetric, we can diagonalize it via a unitary transformation, as in [12, Corollary 2.6.6(a)]. In other words, for each  $a = 1, \dots, d$  there exists  $V_a \in \mathcal{GL}_{n_a}$  unitary and  $D_a$  a nonnegative diagonal matrix such that

$$(3.7) \quad V_a^T C_a V_a = D_a, \quad 1 \leq a \leq d.$$

that the diagonal entries of  $D_a$  are the singular values of  $C_a$ .

Defining  $V := V_1 \oplus \dots \oplus V_d$  and  $D := D_1 \oplus \dots \oplus D_d$  (diagonal) then (3.6) and (3.7) give

$$(3.8) \quad V^T (U^T B(\lambda_0) U) V = D.$$

Defining now  $Q = PUV$ , and since  $B(\lambda_0) = P^T A(\lambda_0) P$ , (3.8) implies  $Q^T A(\lambda_0) Q = D$ . In addition,

$$\begin{aligned} Q^T A_j Q &= V^T U^T (P^T A_j P) UV = V^T (U^T B_j U) V \\ &= V^T (U^T B(\lambda_0) D^{(j)} U) V \quad \text{from (3.2)} \\ &= V^T (U^T B(\lambda_0) U) (U^T D^{(j)} U) V, \quad \text{since } U \text{ is orthogonal} \\ &= (V^T (U^T B(\lambda_0) U) V) (U^T D^{(j)} U), \quad \text{as } V \text{ commutes with } U^T D^{(j)} U \text{ by (3.4),} \\ &= D (U^T D^{(j)} U) \quad \text{from (3.8)} \\ &= D (\alpha_1^{(j)} I_{p_1} \oplus \dots \oplus \alpha_d^{(j)} I_{p_d}) \quad \text{from (3.4),} \end{aligned}$$

which is clearly diagonal for all  $j = 1, \dots, m$ .  $\square$

We recall from Theorem 3 that  $A(\lambda_0)^{-1} A_1, \dots, A(\lambda_0)^{-1} A_m$  are SDS if, and only if, they are all diagonalizable by similarity and they pairwise commute. It follows now from Theorem 7 that property SDS of the matrices  $A(\lambda_0)^{-1} A_1, \dots, A(\lambda_0)^{-1} A_m$  is independent of the particular  $\lambda_0$  chosen.

### 3.2. The SDC problem for $n \times n$ matrices with arbitrary pencil rank.

#### 3.2.1. Preliminaries: Diagonal matrices.

LEMMA 8. Let  $D_1, \dots, D_m$  be diagonal matrices in  $\mathcal{M}_n$ ,  $D$  be the associated linear pencil, and  $r$  be its maximum pencil rank. Then the following hold:

(i)  $D_1, \dots, D_m$  have zeros in the same  $(n - r)$  diagonal positions and

$$\dim \left( \bigcap_{j=1}^m \ker D_j \right) = n - r.$$

(ii) There is an orthogonal  $Q \in \mathcal{M}_n$  such that  $Q^T D_j Q = \tilde{D}_j \oplus 0_{n-r}$ , where  $\tilde{D}_j \in \mathcal{M}_r$  is diagonal,  $1 \leq j \leq m$ .

Moreover the pencil  $\tilde{D}$  associated to matrices  $\tilde{D}_1, \dots, \tilde{D}_m \in \mathcal{M}_r$  is nonsingular (and if  $\lambda_0$  satisfies  $r = \text{rank} D(\lambda_0)$ , then  $\tilde{D}(\lambda_0) \in \mathcal{GL}_r$ ).

*Proof.* Since  $D$  has maximum pencil rank  $r$ , we choose  $\lambda_0 \in S^{2m-1}$  with  $r = \text{rank} D(\lambda_0)$ . Writing  $D_j = \text{diag}(d_1^j, \dots, d_n^j) \in \mathcal{M}_n$ ,  $1 \leq j \leq m$ , we define vectors

$$u_i = \begin{pmatrix} d_i^1 \\ \vdots \\ d_i^m \end{pmatrix} \in \mathbb{C}^m$$

for  $1 \leq i \leq n$ . By direct calculation

$$D(\lambda) = \text{diag}(\lambda \cdot u_1, \dots, \lambda \cdot u_n) \text{ for all } \lambda = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix} \in \mathbb{C}^m,$$

where  $\cdot$  represents the dot product on  $\mathbb{C}^m$  given by

$$\begin{pmatrix} z_1 \\ \vdots \\ z_m \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ \vdots \\ w_m \end{pmatrix} = \sum_{i=1}^m z_i w_i.$$

Since  $r = \text{rank} D(\lambda_0)$ , we can then assume without loss of generality (up to rearrangement of the basis vectors) that  $\lambda_0 \cdot u_i \neq 0$ ,  $1 \leq i \leq r$ , and  $\lambda_0 \cdot u_j = 0$ ,  $r+1 \leq j \leq n$ . In particular,  $u_i \neq 0$  for  $1 \leq i \leq r$ . Define  $h : \mathbb{C}^m \rightarrow \mathbb{C}$  by  $h(\lambda) = \prod_{i=1}^r \lambda \cdot u_i$ . Since  $h$  is continuous, the set  $A := h^{-1}(\mathbb{C} \setminus \{0\})$  is open in  $\mathbb{C}^m$  and since  $\lambda_0 \in A$ , we have that for some  $s > 0$ ,  $\lambda_0 + v \in A$ , and hence  $h(\lambda_0 + v) \neq 0$  for all  $v \in \mathbb{C}^m$ ,  $\|v\| < s$ . This gives  $\text{rank} D(\lambda_0 + v) \geq r$  and since  $r$  is the maximum rank of  $D(\lambda)$ , it follows that  $\text{rank} D(\lambda_0 + v) = r$  and therefore  $(\lambda_0 + v) \cdot u_j = 0$  for all  $j$  with  $r+1 \leq j \leq n$ . Thus  $v \cdot u_j = 0$  for all  $r+1 \leq j \leq n$  and all  $v \in \mathbb{C}^m$ ,  $\|v\| < s$ . This is impossible unless  $u_j = 0$  for all  $r+1 \leq j \leq n$  (otherwise  $v = \bar{u}_j (\frac{s}{2\|u_j\|})$  will give a contradiction). In other words

$$0 = (u_j)_k = d_j^k = (D_k)_{jj} \text{ for all } r+1 \leq j \leq n \text{ and all } 1 \leq k \leq m,$$

namely,  $D_1, \dots, D_m$  have zeros in the same  $n - r$  diagonal positions. It follows that  $\dim(\bigcap_{j=1}^m \ker D_j) \geq n - r$ . On the other hand, Lemma 6 and (3.1) then imply  $\dim(\bigcap_{j=1}^m \ker D_j) = n - r$  and  $\bigcap_{j=1}^m \ker D_j = \ker D(\lambda_0)$ .

(ii) For  $\lambda_0 \in S^{2m-1}$  with  $r = \text{rank} D(\lambda_0)$ , we see in the proof of (i) that there is an orthogonal (permutation) matrix  $Q \in \mathcal{M}_n$  satisfying  $Q^T D_j Q = \tilde{D}_j \oplus 0_{n-r}$ , where  $\tilde{D}_j \in \mathcal{M}_r$  is diagonal,  $1 \leq j \leq m$ . Let  $\tilde{D}$  be the reduced linear pencil associated to  $\tilde{D}_1, \dots, \tilde{D}_m$ . Then  $Q^T D(\lambda_0) Q = \tilde{D}(\lambda_0) \oplus 0_{n-r}$ , so  $\text{rank} \tilde{D}(\lambda_0) = \text{rank} D(\lambda_0) = r$  and  $\tilde{D}(\lambda_0) \in \mathcal{GL}_r$ .  $\square$

The next theorem enables us, when considering whether or not a set of  $n \times n$  matrices is SDC, to reduce the problem to a set of  $r \times r$  matrices, where  $r$  is the maximum pencil rank.

**THEOREM 9.** *Let  $A_1, \dots, A_m \in \mathcal{MS}_n$  have maximum pencil rank  $r$ . Then*

*$A_1, \dots, A_m$  are SDC if, and only if,  $\dim \left( \bigcap_{j=1}^m \ker A_j \right) = n-r$  and there exists  $P \in \mathcal{GL}_n$*

*with  $P^T A_j P = \tilde{D}_j \oplus 0_{n-r}$ , where  $\tilde{D}_j \in \mathcal{M}_r$  is diagonal,  $1 \leq j \leq m$ .*

*Moreover, if either of the above conditions is satisfied, the pencil  $\tilde{D}$  associated to matrices  $\tilde{D}_1, \dots, \tilde{D}_m \in \mathcal{M}_r$  is nonsingular (and if  $\lambda_0$  satisfies  $r = \text{rank} A(\lambda_0)$ , then  $\tilde{D}(\lambda_0) \in \mathcal{GL}_r$ ).*

*Proof.* Let  $A_1, \dots, A_m \in \mathcal{MS}_n$  have maximum pencil rank  $r$ . Choose any  $\lambda_0 \in S^{2m-1}$  satisfying  $r = \text{rank} A(\lambda_0)$ .

For the forward direction, assume that  $A_1, \dots, A_m$  are SDC. Then there exists  $S \in \mathcal{GL}_n$  and diagonal matrices  $D_1, \dots, D_m$  such that

$$(3.9) \quad S^T A_j S = D_j, \quad j = 1, \dots, m.$$

Let  $D$  be the pencil associated to matrices  $D_1, \dots, D_m$ . Then  $S^T A(\lambda) S = D(\lambda)$  for all  $\lambda \in \mathbb{C}^m$ , so maximum pencil ranks for  $A(\lambda)$  and  $D(\lambda)$  agree and  $r = \text{rank} A(\lambda_0) = \text{rank} D(\lambda_0)$ . From Lemma 8(ii) there then exists an orthogonal  $Q \in \mathcal{M}_n$  such that  $Q^T D_j Q = \tilde{D}_j \oplus 0_{n-r}$ , where  $\tilde{D}_j \in \mathcal{M}_r$  is diagonal,  $1 \leq j \leq m$ , and  $\tilde{D}(\lambda_0) \in \mathcal{GL}_r$ . Then  $P = SQ$  gives

$$P^T A_j P = \tilde{D}_j \oplus 0_{n-r}$$

for  $1 \leq j \leq m$  as desired and for pencil  $\tilde{D}$  associated to matrices  $\tilde{D}_1, \dots, \tilde{D}_m$  clearly  $r = \text{rank} \tilde{D}(\lambda_0)$ .

Since  $S^T A(\lambda_0) S = D(\lambda_0)$  we have  $\ker A(\lambda_0) = S(\ker D(\lambda_0))$ . On the other hand, from Lemma 8(i), Lemma 6, and (3.9) we have  $S(\ker D(\lambda_0)) = S(\bigcap_{j=1}^m \ker D_j) = \bigcap_{j=1}^m \ker A_j$ . In other words  $\bigcap_{j=1}^m \ker A_j = \ker A(\lambda_0)$  has dimension  $n-r$ . The opposite direction is trivial.  $\square$

**3.2.2. The general case of nondiagonal matrices with arbitrary pencil rank.** The following lemma holds regardless of diagonalizability and is key to solving the SDC problem in the general case.

**LEMMA 10.** *Let  $A_1, \dots, A_m \in \mathcal{MS}_n$  have maximum pencil rank  $r$ . Then*

$$\dim \left( \bigcap_{j=1}^m \ker A_j \right) = n-r \quad \text{if, and only if,}$$

*there exists  $Q \in \mathcal{GL}_n$  with*

$$(3.10) \quad Q^T A_j Q = \tilde{A}_j \oplus 0_{n-r}, \quad \text{where } \tilde{A}_j \in \mathcal{MS}_r, \quad 1 \leq j \leq m.$$

Moreover, if either of the above conditions is satisfied, the pencil  $\tilde{A}$  associated to matrices  $\tilde{A}_1, \dots, \tilde{A}_m \in M_r$  is nonsingular (and if  $\lambda_0$  satisfies  $r = \text{rank} A(\lambda_0)$ , then  $\tilde{A}(\lambda_0) \in \mathcal{GL}_r$ ).

*Proof.* Let  $A_1, \dots, A_m \in \mathcal{MS}_n$  have maximum pencil rank  $r$ .

In the forward direction, assume that  $\mathcal{V} := \bigcap_{j=1}^m \ker A_j$  has  $\dim(\mathcal{V}) = n-r$ . Choose a basis  $v_{r+1}, \dots, v_n$  of  $\mathcal{V}$  and extend by vectors  $v_1, \dots, v_r$  to get a basis  $v_1, \dots, v_n$  of  $\mathbb{C}^n$ . Let  $Q \in \mathcal{GL}_n$  be the matrix whose  $i$ th column is given by the vector  $v_i$ . For  $r+1 \leq i \leq n$ , we have  $v_i \in \ker A_j$  and hence  $Q^T A_j Q(e_i) = Q^T A_j(v_i) = 0$  for all  $1 \leq j \leq m$  (and  $e_i$  is the column vector with 1 in the  $i$ th position and all other entries 0). In other words, columns  $r+1$  to  $n$  of  $Q^T A_j Q$  are identically zero and, since  $Q^T A_j Q$  is symmetric, it follows that

$$(3.11) \quad Q^T A_j Q = \tilde{A}_j \oplus 0_{n-r}, \text{ where } \tilde{A}_j \in \mathcal{MS}_r, \quad 1 \leq j \leq m,$$

as desired. In the opposite direction, assume that (3.10) holds for some  $Q \in \mathcal{GL}_n$ . Then  $Q(0_r \oplus \mathbb{C}^{n-r}) \subseteq \bigcap_{j=1}^m \ker A_j$  so  $\dim(\bigcap_{j=1}^m \ker A_j) \geq n-r$ . Equality now follows from (3.1).

Finally, if the conditions in the statement hold, the reduced pencil  $\tilde{A}$  associated with  $\tilde{A}_1, \dots, \tilde{A}_m \in M_r$  has maximum pencil rank  $r$ , since for any  $\lambda_0 \in \mathbb{C}^m$  with  $r = \text{rank} A(\lambda_0)$  then (3.11) implies that  $\text{rank} \tilde{A}(\lambda_0) = r$  and  $\tilde{A}(\lambda_0) \in \mathcal{GL}_r$ .  $\square$

*Remark 11.* We note that by choosing the basis vectors  $v_1, \dots, v_n$  in the above proof to be orthogonal, with respect to the complex inner product  $\langle z, w \rangle := z \cdot \bar{w}$  on  $\mathbb{C}^n$ ,  $Q$  can be chosen to be unitary.

Lemma 10 therefore allows us to find matrices  $\tilde{A}_1, \dots, \tilde{A}_m$  satisfying (3.10) using only the kernels of the  $A_j$ . This enables us, subject to the condition that  $\dim(\bigcap_{j=1}^m \ker A_j) = n-r$ , to reduce the dimension of the problem by proving that

$$A_1, \dots, A_m \text{ are SDC in } M_n \text{ if, and only if, } \tilde{A}_1, \dots, \tilde{A}_m \text{ are SDC in } M_r.$$

Theorem 7 then motivates the following definition.

**DEFINITION 12** (reduced maximal-rank matrices). *Let  $A_1, \dots, A_m \in \mathcal{MS}_n$  have maximum pencil rank  $r$  and satisfy  $\dim(\bigcap_{j=1}^m \ker A_j) = n-r$ . Let  $\tilde{A}_1, \dots, \tilde{A}_m$  be as in (3.10) and fix  $\lambda_0 \in S^{2m-1}$  with  $r = \text{rank} A(\lambda_0)$ . Reduced pencil  $\tilde{A}$  then has  $\tilde{A}(\lambda_0) \in \mathcal{GL}_r$ .*

*We define the  $r \times r$  matrices*

$$(3.12) \quad L_j (= L_j(\lambda_0)) := \tilde{A}(\lambda_0)^{-1} \tilde{A}_j, \quad 1 \leq j \leq m.$$

*Remark 13.*  $L_1, \dots, L_m$  are not symmetric in general and  $\sum_{j=1}^m (\lambda_0)_j L_j = I_n$ . In addition, Theorem 7 states that  $L_1, \dots, L_m$  are SDS if, and only if,  $\tilde{A}_1, \dots, \tilde{A}_m$  are SDC and consequently the condition is independent of the particular  $\lambda_0$  chosen in the definition. For this reason we write  $L_j$  instead of  $L_j(\lambda_0)$ .

The following is our main theorem.

**THEOREM 14.** *Let  $A_1, \dots, A_m \in \mathcal{MS}_n$  have maximum pencil rank  $r$ . Then*

$$A_1, \dots, A_m \text{ are SDC if, and only if, } \dim \left( \bigcap_{j=1}^m \ker A_j \right) = n-r \text{ and } L_1, \dots, L_m \text{ are SDS,}$$

*where  $L_1, \dots, L_m$  are as in Definition 12 above.*

*Proof.* Let  $A_1, \dots, A_m \in \mathcal{MS}_n$  have maximum pencil rank  $r$  and choose  $\lambda_0 \in S^{2m-1}$  satisfying  $r = \text{rank } A(\lambda_0)$ .

In the forward direction, assume now that  $A_1, \dots, A_m$  are SDC.

From Theorem 9  $\dim(\bigcap_{j=1}^m \ker A_j) = n - r$  and there exists  $P \in \mathcal{GL}_n$  such that

$$(3.13) \quad P^T A_j P = \tilde{D}_j \oplus 0_{n-r}, \text{ where } \tilde{D}_j \in \mathcal{M}_r \text{ is diagonal, } 1 \leq j \leq m.$$

In addition, if  $\tilde{D}$  is the pencil associated to  $r \times r$  matrices  $\tilde{D}_1, \dots, \tilde{D}_m$ , then  $\tilde{D}(\lambda_0) \in \mathcal{GL}_r$ . Lemma 10 then gives  $Q \in \mathcal{GL}_n$  with

$$(3.14) \quad Q^T A_j Q = \tilde{A}_j \oplus 0_{n-r}, \text{ where } \tilde{A}_j \in \mathcal{M}_r, \quad 1 \leq j \leq m,$$

and  $\tilde{A}(\lambda_0) \in \mathcal{GL}_r$  for reduced pencil  $\tilde{A}(\lambda_0) = \sum_{j=1}^m (\lambda_0)_j \tilde{A}_j$ .

Thus for  $R = Q^{-1}P$  (3.13) and (3.14) give

$$(3.15) \quad R^T (\tilde{A}_j \oplus 0_{n-r}) R = \tilde{D}_j \oplus 0_{n-r}, \quad 1 \leq j \leq m.$$

Writing  $R$  as a block matrix,

$$R = \begin{pmatrix} S & T \\ U & V \end{pmatrix},$$

for  $S \in \mathcal{M}_r, V \in \mathcal{M}_{n-r}, U \in \mathcal{M}_{n-r,r}, T \in \mathcal{M}_{r,n-r}$ , it follows from (3.15) and matrix multiplication that

$$(3.16) \quad S^T \tilde{A}_j S = \tilde{D}_j \quad \text{for } 1 \leq j \leq m.$$

Then for reduced  $r \times r$  matrix pencils (and  $\lambda = (\lambda_1, \dots, \lambda_m) \in \mathbb{C}^m$ )

$$\tilde{A}(\lambda) = \sum_{j=1}^m \lambda_j \tilde{A}_j \quad \text{and} \quad \tilde{D}(\lambda) = \sum_{j=1}^m \lambda_j \tilde{D}_j$$

(3.16) gives

$$(3.17) \quad S^T \tilde{A}(\lambda) S = \tilde{D}(\lambda),$$

and, in particular,

$$(3.18) \quad S^T \tilde{A}(\lambda_0) S = \tilde{D}(\lambda_0).$$

Since  $\tilde{A}(\lambda_0)$  and  $\tilde{D}(\lambda_0)$  are invertible, it follows that  $S$  is invertible and combining (3.16) and (3.18) gives

$$\tilde{D}(\lambda_0)^{-1} \tilde{D}_j = S^{-1} \tilde{A}(\lambda_0)^{-1} (S^T)^{-1} S^T \tilde{A}_j S = S^{-1} \tilde{A}(\lambda_0)^{-1} \tilde{A}_j S, \quad 1 \leq j \leq m.$$

In particular,  $S^{-1} \tilde{A}(\lambda_0)^{-1} \tilde{A}_j S$  are diagonal for all  $j = 1, \dots, m$ . In other words the  $r \times r$  matrices

$$L_j = \tilde{A}(\lambda_0)^{-1} \tilde{A}_j \quad \text{for } 1 \leq j \leq m$$

are SDS and we are done.

For the opposite direction, let us assume that  $\dim(\bigcap_{j=1}^m \ker A_j) = n - r$  and that  $L_1, \dots, L_m$  are SDS. Then from Lemma 10 there exists  $Q \in \mathcal{GL}_n$  such that

$$Q^T A_j Q = \tilde{A}_j \oplus 0_{n-r}, \quad j = 1, \dots, m,$$

with  $\tilde{A}_j \in \mathcal{MS}_r$  and  $\tilde{A}(\lambda_0) \in \mathcal{GL}_r \cap \mathcal{MS}_r$ . Construct  $L_j = \tilde{A}(\lambda_0)^{-1} \tilde{A}_j$ ,  $j = 1, \dots, m$  as in Definition 12 above. By hypothesis, these matrices are SDS so from Theorem 7 it follows that  $\tilde{A}_1, \dots, \tilde{A}_m$  are SDC, namely, there exists  $P \in \mathcal{GL}_r$  such that  $P^T \tilde{A}_j P = D_j$  for  $D_j$  diagonal in  $\mathcal{M}_r$ , for all  $j = 1, \dots, m$ . Define  $R := P \oplus I_{n-r} \in \mathcal{GL}_n$ . Then

$$R^T (Q^T A_j Q) R = (P^T \tilde{A}_j P) \oplus 0_{n-r} = D_j \oplus 0_{n-r}, \quad j = 1, \dots, m.$$

Thus, for  $S = QR \in \mathcal{GL}_n$  we have

$$S^T A_j S = D_j \oplus 0_{n-r},$$

diagonal for all  $j = 1, \dots, m$ . Thus,  $A_1, \dots, A_m$  are SDC.  $\square$

**3.3. A procedure to solve the SDC problem.** The above results allow us now to determine in a finite number of steps whether or not a set of matrices are SDC. Given  $A_1, \dots, A_m \in \mathcal{MS}_n$ , let  $r := \text{rank}(A_1, \dots, A_m)$ . From (3.1), we have

$$\dim \left( \bigcap_{j=1}^m \ker A_j \right) \leq n - r$$

and Theorems 14 and 3 now give us the following procedure:

- (1) If  $\dim(\bigcap_{j=1}^m \ker A_j) < n - r$ , then  $A_1, \dots, A_m$  are not SDC.
- (2) If  $\dim(\bigcap_{j=1}^m \ker A_j) = n - r$ , then we calculate  $L_1, \dots, L_m$  from (3.12) for some  $\lambda_0 \in \mathbb{C}^m$  with  $r = \text{rank} A(\lambda_0)$ . If  $L_1, \dots, L_m$  do not pairwise commute, then  $A_1, \dots, A_m$  are not SDC.
- (3) If  $\dim(\bigcap_{j=1}^m \ker A_j) = n - r$ ,  $L_1, \dots, L_m$  do pairwise commute, and if each  $L_1, \dots, L_m$  is diagonalizable by similarity, then  $A_1, \dots, A_m$  are SDC. Otherwise (that is, if any one  $L_j$  is not diagonalizable by similarity)  $A_1, \dots, A_m$  are not SDC.

*Remark 15.* Regarding (1) above, we note that estimation may be sufficient to determine if  $\dim(\bigcap_{j=1}^m \ker A_j) < n - r$ , as in Example 17 below. Regarding (2), since  $\sum_{j=1}^m (\lambda_0)_j L_j = I_n$ , it suffices to check if  $m - 1$  of  $L_1, \dots, L_m$  pairwise commute. Regarding (3), we recall that  $L_i$  are not symmetric in general.

*Example 16* ( $n = 2, m = 2$ ). Let

$$A_1 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

We apply the above procedure.

- (1)  $\ker A_1 = \ker A_2 = \{0\}$  so  $\dim(\bigcap_{j=1}^2 \ker A_j) = 0$ . As  $A_1$  is nonsingular we take  $\lambda_0 = (1, 0)$  so  $A(\lambda_0) = A_1$  and  $\text{rank} A_1 = 2$ . Therefore  $r = 2$  and  $\dim(\bigcap_{j=1}^m \ker A_j) = n - r$  holds and we continue to the next step.
- (2) We compute

$$L_1 = A_1^{-1} A_1 = I_2 \text{ and } L_2 = A_1^{-1} A_2 = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}.$$

These matrices (trivially) commute so we continue to the next step.

(3)  $L_1$  is diagonal.  $L_2$  is diagonalizable by similarity as it has two different eigenvalues:  $d_{\pm} = (1 \pm i\sqrt{3})/2$ . Therefore the matrices  $A_1, A_2$  are SDC. Explicitly,  $P^{-1}L_2P = \text{diag}(d_+, d_-)$  with

$$P = \begin{pmatrix} ad_- & bd_+ \\ -a & -b \end{pmatrix}$$

for any  $a, b \in \mathbb{C}$  with  $ab \neq 0$ . Note that  $P$  cannot be made real by any choice of the constants  $a, b$ . We have, finally,

$$P^T A_1 P = i\sqrt{3} \text{diag}(a^2, -b^2), \quad P^T A_2 P = i\sqrt{3} \text{diag}(a^2 d_+, -b^2 d_-).$$

*Example 17* ( $n = 3, m = 2$ ). Let

$$A_1 = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

We apply the above procedure.

- (1) We calculate  $\ker A_1 = \text{span}\{(0, 0, 1)^T\}$  and  $\ker A_2 = \text{span}\{(0, 1, 0)^T\}$ . Thus  $\ker A_1 \cap \ker A_2 = \{0\}$  so  $\dim(\bigcap_{j=1}^2 \ker A_j) = 0$ .

Since

$$\det A(\lambda) = \det \begin{pmatrix} \lambda_1 & \lambda_1 & \lambda_2 \\ \lambda_1 & 0 & 0 \\ \lambda_2 & 0 & 0 \end{pmatrix} = 0 \quad \text{for all } \lambda \in \mathbb{C}^2,$$

we have  $r \leq 2$  and then  $n - r \geq 1$ . Therefore  $\dim(\bigcap_{j=1}^2 \ker A_j) < n - r$  and hence  $A_1, A_2$  are not SDC.

**4. Discussion.** In this paper we solved the long-standing problem of simultaneous diagonalization via congruence in the complex symmetric case, providing also an explicit set of steps to solve this problem. The complex case has applications in signal processing, in particular to the problem of blind source separation. This latter problem is based on the exact SDC problem but, due to experimental and numerical errors in obtaining the target matrices  $A_1, \dots, A_m$ , it relies on the so-called approximate joint diagonalization, which is an optimization problem. Our results could shed light on these approximate problems, as these problems usually consider an ad hoc cost function [3, 26], which does not take into account the kernels of the target matrices.

Some optimization-related applications consider the special case where, in the context of our Definition 1, the symmetric matrices  $A_1, \dots, A_m$  are real, and the corresponding transformation matrix  $P$  and resulting diagonal matrices  $D_1, \dots, D_m$  are required to be real. In the context of Theorem 14 above, such a case would impose extra conditions of realness on the eigenvectors and eigenvalues of the reduced matrices  $L_1, \dots, L_m$ .

In many applications in genetics the matrices  $L_1, \dots, L_m$  turn out to commute, but may not necessarily be diagonalizable. Thus, the SDC problem could be relaxed to a weaker problem, namely, that of simultaneous block diagonalization [21].

Further research on building an algorithm to solve the SDC problem will focus on developing an efficient method for finding  $\lambda_0$  such that the pencil  $A(\lambda_0)$  has maximum rank.

**Acknowledgment.** The authors thank the referees for several helpful comments and suggestions and for bringing key references to our attention.

## REFERENCES

- [1] B. AFSARI, *Sensitivity Analysis for the Problem of Matrix Joint Diagonalisation*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 1148–1171.
- [2] R.I. BECKER, *Necessary and sufficient conditions for the simultaneous diagonalability of two quadratic forms*, Linear Algebra Appl., 30 (1980), pp. 129–139.
- [3] A. BELOUCHRANI, K. ABED-MERAÏM, J.-F. CARDOSO, AND E. MOULINES, *A blind source separation technique using second-order statistics*, IEEE Trans. Signal Process., 45 (1997), pp. 434–444.
- [4] M. D. BUSTAMANTE, P. MELLON, AND M.V. VELASCO, *Determining when an algebra is an evolution algebra*, Mathematics, 8 (2020), 1349.
- [5] J.-F. CARDOSO AND A. SOULOUMIAC, *Blind beamforming for non-Gaussian signals*, IEE Proc. F Radar and Signal Process., 140 (1993), pp. 362–370.
- [6] D.T. PHAM AND M. CONGEDO, *Least square joint diagonalisation of matrices under an intrinsic scale constraint*, in Proceedings of the International Conference on Independent Component Analysis and Signal Separation, 2009, T. Adali, C. Jutten, J.M.T. Romano, and A.K. Barros, eds., Lecture Notes in Comput. Sci. 5441, Springer, Berlin.
- [7] J.B. HIRIART-URRUTY, *Potpourri of conjectures and open questions in nonlinear analysis and optimisation*, SIAM Rev., 49 (2007), pp. 255–273.
- [8] J.B. HIRIART-URRUTY AND J. MALICK, *A fresh variational-analysis look at the positive semi-definite matrices world*, J. Optim. Theory Appl., 153 (2012), pp. 551–577.
- [9] J.B. HIRIART-URRUTY AND M. TORKI, *Permanently going back and forth between the “quadraticworld” and the “convexityworld” in optimization*, Appl. Math. Optim., 45 (2002), pp. 169–184.
- [10] Y.P. HONG, R.A. HORN, AND C.R. JOHNSON, *On the reduction of pairs of Hermitian or symmetric matrices to diagonal form by congruence*, Linear Algebra Appl., 73 (1986), pp. 213–226.
- [11] Y.P. HONG AND R.A. HORN, *On simultaneous reduction of families of matrices to triangular or diagonal form by unitary congruences*, Linear Multilinear Algebra, 17 (1985), pp. 271–288.
- [12] R.A. HORN AND C.R. JOHNSON, *Matrix Analysis*, 2nd edition, Cambridge University Press, Cambridge, 2013.
- [13] Y. HSIA, G.X. LIN, AND R.L. SHEU, *A revisit to quadratic programming with one inequality quadratic constraint via matrix pencil*, Pac. J. Optim., 10 (2014), pp. 461–481.
- [14] R. JIANG AND D. LI, *Simultaneous diagonalisation of matrices and its applications in quadratically constrained quadratic programming*, SIAM J. Optim., 26 (2016), pp. 1649–1669.
- [15] L. DE LATHAUWER, *A link between the canonical decomposition in multilinear algebra and simultaneous matrix diagonalisation*, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 642–666.
- [16] D.T. PHAM, *Joint approximate diagonalisation of positive definite matrices*, SIAM. J. Matrix Anal. Appl., 22 (2001), pp. 1136–1152.
- [17] M. SORESENSEN AND P. COMON, *A pair sweeping method for some simultaneous matrix diagonalisation*, in Equipe SIGNAL - Pôle SIS - Février, 2012.
- [18] J.P. TIAN AND P. VOJTECHOVSKY, *Mathematical concepts of evolution algebras in non-mendelian genetics*, Quasigroups Related Systems, 24 (2006), pp. 111–122.
- [19] J.P. TIAN, *Evolution algebras and their applications*, Lecture Notes in Math. 1921, Springer, Berlin, 2008.
- [20] P. TICHAVSKY AND A. YEREDOR, *Fast approximate joint diagonalisation incorporating weight matrices*, IEEE Trans. Signal Process., 57 (2009), pp. 878–891.
- [21] F. UHLIG, *Simultaneous block diagonalization of two real symmetric matrices*, Linear Algebra Appl., 7 (1973), pp. 281–289.
- [22] F. UHLIG, *A recurring theorem about pairs of quadratic forms and extensions: A survey*, Linear Algebra Appl., 25 (1979), pp. 219–237.
- [23] L. WANG, L. ALBERA, A. KACHENOURA, H.Z. SHU, AND L. SENHADJI, *Nonnegative joint diagonalisation by congruence based on LU matrix factorization*, IEEE Signal Process. Lett., 20 (2013), pp. 807–810.

- [24] K. WEIERSTRASS, *Zur Theorie der quadratischen und bilinearen Formen*, Monatsber. Akad. Wiss. 1868, pp. 310–338.
- [25] A. YEREDOR, *Blind source separation via the second characteristic function*, Signal Process., 80 (2000), pp. 897–902.
- [26] A. YEREDOR, *Non-orthogonal joint diagonalization in the least-squares sense with application in blind source separation*, IEEE Trans. Signal Process., 50 (2002), pp. 1545–1553.
- [27] A.Y. YIK-HOI, *A necessary and sufficient condition for simultaneously diagonalisation of two Hermitian matrices and its applications*, Glasg. Math. J., 11 (1970), pp. 81–83.