Borobia, A., Canogar, R. & De Terán, F. (2021). On the Consistency of the Matrix Equation $X^\top AX = B$ when B is Symmetric. *Mediterranean Journal of Mathematics*, 18(2), 40.

DOI: 10.1007/s00009-020-01656-7

On the consistency of the matrix equation $X^\top AX = B$ when B is symmetric

Alberto Borobia, Roberto Canogar and Fernando De Terán

Abstract. We provide necessary and sufficient conditions for the matrix equation $X^\top AX = B$ to be consistent when B is a symmetric matrix, for all matrices A with a few exceptions. The matrices A, B, and X (unknown) are matrices with complex entries. We first see that we can restrict ourselves to the case where A and B are given in *canonical form for congruence* and, then, we address the equation with A and B in such form. The characterization strongly depends on the canonical form for congruence of A. The problem we solve is equivalent to: given a complex bilinear form (represented by A) find the maximum dimension of a subspace such that the restriction of the bilinear form to this subspace is a symmetric non-degenerate bilinear form.

Mathematics Subject Classification (2010). 15A21, 15A24 .

Keywords. Matrix equation, transpose, congruence, $\top$-Riccati equation, Canonical Form for Congruence, symmetric matrix, bilinear form.

1. Introduction

We are interested in providing necessary and sufficient conditions for the equation

$$X^\top AX = B \tag{1.1}$$

to be consistent. Here, A and B are square complex matrices (that is, matrices with entries in the complex field $\mathbb{C}$) not necessarily of the same size, X is the unknown, and $M^\top$ denotes the transpose of the matrix M.

Equation (1.1) arises in several settings, in particular related to *bilinear forms* and matrix *congruence*. These two concepts are connected, since two square matrices A and B of the same size represent the same bilinear form with respect to different bases if and only if A and B are congruent. By definition, this is equivalent to say that Eq. (1.1) has a nonsingular solution.

.

In order to determine whether two particular square matrices A and B of the same size are congruent, one should ask for invariants or intrinsic properties of A and B that characterize this equivalence relation or, moreover, for a canonical form. Canonical forms for congruence are known since, at least, the 1930s [19, p. 139], but we follow the one in [17] (and we will refer to it as the CFC). Using the CFC, we can get a characterization for the existence of a nonsingular solution of Eq. (1.1), namely, this happens if and only if A and B have the same CFC. Then, this characterization solves, theoretically, the question on whether or not Eq. (1.1) has a nonsingular solution.

But, what happens if we remove the constraint on the nonsingularity of the solution? More precisely, we may allow X to be not only nonsingular, but even rectangular. So, assume that $X \in \mathbb{C}^{n\times m}$ is a solution of Eq. (1.1). In order for (1.1) to be well defined, it must be $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{m\times m}$, namely, A and B must be square, but not necessarily of the same size. In this setting, Eq. (1.1) may have a solution with A and B not being congruent (in particular, this happens if they are not of the same size). So congruence is just the extreme case, $m = n$, of the general equation (1.1) (with X nonsingular).

The other extreme case of Eq. (1.1) is when $m = 1$. In this case, Eq. (1.1) becomes $x^\top Ax = \beta$, with $\beta \in \mathbb{C}$. This gives a quadratic form (note that, if $\beta \neq 0$, the equation has a solution whenever A is not skew-symmetric). So our problem is placed in between these two extreme cases, namely, congruence and quadratic forms.

Nonetheless, congruence plays a key role in our general strategy. More precisely, we will see that Eq. (1.1) is consistent if and only if the equation obtained after replacing A and B by their respective CFCs is consistent as well. Since we are interested in characterizing when Eq. (1.1) is consistent, we will assume most of the time that A and B are already given in CFC.

The CFC is a block-diagonal form containing blocks of three different types (see Theorem 1). A natural approach to solve Eq. (1.1) when both A and B are in CFC is to partition the solution X into blocks, conformally with the partitions of A and B, and then try to solve individually all the equations corresponding to each pair of canonical blocks of A and B in order to get a solution of the whole equation. However, this approach presents a relevant obstacle when applied to Eq. (1.1) (see Section 2), and to analyze the solvability of Eq. (1.1) with A and B in CFC seems to be, in general, a very hard task.

Nonetheless, we have succeeded in obtaining necessary and sufficient conditions for Eq. (1.1) to be consistent when B is symmetric. We want to emphasize that Eq. (1.1) can be consistent with B being symmetric and A being non-symmetric. However, when A is symmetric, if Eq. (1.1) is consistent, then B must be symmetric as well. Then, the case where A is symmetric is a particular case of the one we are interested in. Moreover, as we will see, the case where B is symmetric is much richer than the case where A is symmetric, and the characterization for the consistency is more complex. In particular, the characterization for consistency of Eq. (1.1) when A is symmetric can be stated in a very elementary way without explicitly using the CFC of A (see Lemma 2.2). By contrast, the characterization of the consistency when B is symmetric requires the knowledge of the CFC of A (see Theorem 8). We want to note that the characterization in Theorem 8 is not complete, since it does not cover the case where a particular kind of one type of blocks appear in CFC(A) (namely, blocks of either the form $H_4(1)$ or $H_2(-1)$, see Theorem 1). Nevertheless, this characterization is almost complete, since it covers most instances of CFC(A) or, in other words, is valid for most matrices A.

Our strategy to get the characterization for the consistency of Eq. (1.1) when B is symmetric consists in first obtaining a necessary condition in terms of the canonical blocks of the CFC of A. Then, we show that this condition is sufficient by analyzing Eq. (1.1) for A being a single block of each of the different types in the CFC (excluding the blocks of the form $H_4(1)$ and $H_2(-1)$ mentioned in the previous paragraph). In other words, we show that, when the necessary condition is satisfied, then a block diagonal solution exists. We want to note that, in most of the cases where Eq. (1.1) is consistent, we have provided an explicit solution. Therefore, our proof for the consistency is, in many cases, a constructive proof (up to the matrices that take A and B to their CFC).

Something that is important to emphasize is that, despite (1.1) is a nonlinear (in particular, quadratic) equation, in this work we have used techniques, tools, and developments from linear algebra.

There are several references in the literature that deal with Eq. (1.1). In particular, several papers have been devoted, since the 1960s, to count the number of either general solutions or solutions with some particular property (like having some fixed rank) and with A and B being either arbitrary or having some specific structure (like "alternate") for matrices over finite fields [2–5, 15, 20]. It has also appeared in [14] to determine which $m \times n$ matrices over fields with characteristic 2 have a generalized inverse or pseudoinverse, and also to count the number of such matrices for finite fields with characteristic 2. More recently, this equation has arisen in connection with several applications, like image deblurring problems [13] and dynamics generalized equilibrium (DSEG) problems (see [1] and the references therein). Numerical methods to compute the minimal solution to the more general *nonsymmetric* $\top$*-Riccati equation* (with real matrices) have been proposed in the recent work [1].

Some other related equations to Eq. (1.1) have been of interest in the recent years. Among them, we cite the so-called "generalized Yang-Baxter matrix equation", $AXA = XAX$, which is also a nonlinear equation that is analyzed using linear algebra techniques in several papers, like [12]. The equation $XAX = B$, which resembles very much Eq. (1.1) (it is the same equation without transposing the second appearance of the unknown X) has been analyzed in [18] for A, B being symmetric or skew-symmetric, and using also appropriate canonical forms of the coefficient matrices in order to reduce the equation to a more manageable expression. The previous equations are nonlinear and do not involve transposition. Some other recent references deal with linear equations close to Eq. (1.1), and involving transposition. For instance, the works [6–11], where the solution of Sylvester-like equations $XA+AX^\top = 0$, $AX+X^\top B = 0$, or $AXB+CX^\top D = E$ has been considered. In all these references, both the coefficients and the unknown are complex matrices, like in the present work.

The paper is organized as follows. In Section 2, we present the notation and recall the basic notions and tools that are used throughout the manuscript (like the CFC). We also settle the basic approach that we follow to analyze the consistency of Eq. (1.1), and present some elementary technical results. Section 2.1 is devoted to characterize the consistency of Eq. (1.1) when A is symmetric. The core of the manuscript are Sections 3–7. In these sections, we study the consistency of Eq. (1.1) when B is symmetric. We first present, in Theorem 2, a necessary condition for Eq. (1.1) to be consistent. This condition depends on the size of A, the rank of B, and the number of blocks of certain types appearing in the CFC of A. In Sections 4–6 we analyze the consistency of Eq. (1.1) for A being a single block of the CFC. This allows us to present, in Theorem 8 (which is the main result of this paper), a characterization for the consistency of Eq. (1.1) when the CFC of A does not contain blocks of either the form $H_4(1)$ or $H_2(-1)$, by proving that the condition of Theorem 2 is sufficient. In Section 8 we summarize the contributions of this paper and indicate some lines of further research.

We close the Introduction by recovering the idea that A represents a bilinear form over $\mathbb{C}^n$. In this case, if Eq. (1.1), with $B = I_m$, has a solution X, then the rank of X is necessarily m, so the columns of X form the basis of an m-dimensional linear subspace of $\mathbb{C}^n$. It is noteworthy that the restriction of the bilinear form to this subspace is a symmetric and non-degenerate bilinear form. So, from this point of view, the problem that we solve in this work is the following: given a bilinear form A (with the few exceptions mentioned above), we find the maximum dimension m_A that a subspace of $\mathbb{C}^n$ can have so that the restriction of A to this subspace is symmetric and non-degenerate. Moreover, we provide a basis of such a subspace. Note that this problem is trivial when A is symmetric, since then $m_A = \operatorname{rank} A$ and the basis can be found by means of the Autonne-Takagi factorization [16, Cor. 4.4.4(c)].

2. Basic approach and definitions

Throughout the manuscript, I_n and 0_n denote, respectively, the identity and the null matrix with size $n \times n$. By $0_{m\times n}$ we denote the null matrix of size $m \times n$. By $\mathfrak{i}$ we denote the imaginary unit (namely, $\mathfrak{i}^2 = -1$), and by e_j we denote the jth canonical vector of the appropriate size (namely, the jth column of the identity matrix).

When considering the question on whether Eq. (1.1) is consistent or not, a useful tool is the *canonical form for congruence* (CFC). In order to recall the CFC we first need to introduce the following matrices:

$$J_k(\lambda) := \begin{bmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \lambda & 1 \\ & & & \lambda \end{bmatrix}$$

is a $k \times k$ *Jordan block associated with* $\lambda \in \mathbb{C}$; for each $k \geqslant 1$, let Γ_k be the $k \times k$ matrix

$$\Gamma_k := \begin{bmatrix} 0 & & & & & (-1)^{k+1} \\ & & & & \iddots & (-1)^k \\ & & & -1 & \iddots & \\ & & 1 & 1 & & \\ & -1 & -1 & & & \\ 1 & 1 & & & & 0 \end{bmatrix} \qquad (\Gamma_1 = [1]);$$

and, for each $\lambda \in \mathbb{C}$ and each $k \geqslant 1$, $H_{2k}(\lambda)$ is the $2k \times 2k$ matrix

$$H_{2k}(\lambda) := \begin{bmatrix} 0 & I_k \\ J_k(\lambda) & 0 \end{bmatrix},$$

where $J_k(\lambda)$ is a $k \times k$ Jordan block associated with λ.

Theorem 1. (Canonical form for congruence, CFC) [17, Th. 1.1]. *Each square complex matrix is congruent to a direct sum, uniquely determined up to permutation of addends, of canonical matrices of the following three types*

Type 0	$J_k(0)$
Type I	Γ_k
Type II	$H_{2k}(\mu)$, $0 \neq \mu \neq (-1)^{k+1}$ (μ *is determined up to replacement by* μ^{-1})

The CFC is the basic tool in our strategy to analyze the consistency of Eq. (1.1). More precisely, let C_A and C_B be, respectively, the CFCs of A and B. Then, there are two nonsingular matrices R and S such that

$$A = R^\top C_A R \qquad \text{and} \qquad B = S^\top C_B S.$$

Now, (1.1) is equivalent to

$$X^\top (R^\top C_A R) X = S^\top C_B S \Leftrightarrow (RXS^{-1})^\top C_A (RXS^{-1}) = C_B.$$

With the change of variables $Y = RXS^{-1}$, the previous equation reads

$$Y^\top C_A Y = C_B, \tag{2.1}$$

so (1.1) is consistent if and only if (2.1) is consistent. Note that, in (2.1), the coefficients matrices are given in CFC. As a consequence, when analyzing the consistency of Eq. (1.1), we may restrict ourselves to the case where the coefficient matrices A and B are already given in CFC.

A natural approach to address the solution of Eq. (1.1), when A and B are given in CFC, is to partition the unknown X conformally with the block partition of A and B. This is, for instance, the approach followed in [6] for the equation $XA + AX^\top = 0$ using the CFC, in [8] for the equation $AX + X^\top B = 0$ using the *Kronecker canonical form* of the matrix pencil $A - \lambda B^\top$, or in [18] for the equation $XAX = B$ using canonical forms for the so-called *simultaneous contragredient transformation*. However, when trying this approach with Eq. (1.1), some relevant difficulties arise. Let us illustrate these obstructions assuming that both A and B consist of exactly two diagonal blocks of any of the types described in Theorem 1. More precisely, assume that A and B are of the form

$$A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}, \qquad \text{and} \qquad B = \begin{bmatrix} B_1 & 0 \\ 0 & B_2 \end{bmatrix}.$$

Then, we partition the unknown X conformally with the previous block partitions, as

$$X = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}.$$

Now, multiplying by blocks in Eq. (1.1) with the previous partitions and equating by blocks, we get the system of matrix equations

$$\begin{array}{rcl} X_{11}^\top A_1 X_{11} + X_{21}^\top A_2 X_{21} & = & B_1 \\ X_{11}^\top A_1 X_{12} + X_{21}^\top A_2 X_{22} & = & 0 \\ X_{12}^\top A_1 X_{11} + X_{22}^\top A_2 X_{21} & = & 0 \\ X_{12}^\top A_1 X_{12} + X_{22}^\top A_2 X_{22} & = & B_2 \end{array}.$$

The previous system contains equations which are not of the form (1.1) and, moreover, the blocks X_{ij} are mixed in these equations. This happens even if either A or B consists only of just one canonical block. As a consequence, to address the solution of Eq. (1.1) using this approach does not seem to be appropriate. Nonetheless, there are some particular and very elementary cases where this strategy is useful (see, for instance, Lemmas 2.2 and 2.3). Also, our approach to prove that the necessary condition for Eq. (1.1) to be consistent when B is symmetric is also sufficient (Section 7) uses the fact that, in order to obtain a solution when A is a direct sum of canonical blocks of different types, it is enough to get a solution for the direct sum of canonical blocks of the same type, and then the solution for the direct sum of all blocks is obtained as a block-diagonal matrix by plugging-in every individual solution.

As we have mentioned in the Introduction, our aim is to look for necessary and sufficient conditions for Eq. (1.1) to be consistent when B is symmetric. When A is symmetric, B is necessarily so, and the characterization in this case is elementary (see Section 2.1). However, when B is symmetric, the equation $X^\top AX = B$ can be consistent, with A not being symmetric. Consider, for instance,

$$X^\top \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} X = \begin{bmatrix} 1 \end{bmatrix},$$

which has a solution $X = \left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$. This makes the problem on the consistency of Eq. (1.1), with B symmetric, a more interesting problem. We will see that the characterization for the consistency in this case is far from being so simple as when A is symmetric, and it strongly depends on the CFC of A.

2.1. The case where A is symmetric

If A is symmetric and $X^\top AX = B$ is consistent, then B is also symmetric. In this case, the canonical form for congruence of both A and B consists of Type-0 and Type-I blocks of size 1×1, that is,

$$\mathrm{CFC}(A) = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}, \qquad \mathrm{CFC}(B) = \begin{bmatrix} I_s & 0 \\ 0 & 0 \end{bmatrix}.$$

The necessary and sufficient condition for Eq. (1.1) to be consistent in this case is simply $r \geqslant s$. It is obviously necessary, and to see that it is sufficient just take $X = \left[\begin{smallmatrix} I_s & 0 \\ 0 & 0 \end{smallmatrix}\right]$ as a solution. This is stated in the following result.

Lemma 2.1. *Let $A \in \mathbb{C}^{n\times n}$ be symmetric and $B \in \mathbb{C}^{m\times m}$. Then, the equation $X^\top AX = B$ has a solution if and only if B is also symmetric and* $\operatorname{rank} A \geqslant \operatorname{rank} B$.

Note that Lemma 2.1 extends [16, Th. 4.5.12] to matrices A and B not necessarily being of the same size and X not necessarily invertible.

2.2. Some technical results

In this section, we present three elementary results that will be used later.

As we saw in (2.1), in order to analyze the consistency of Eq. (1.1), we may restrict ourselves to the case where the coefficient matrices A and B are already given in CFC. If B is symmetric, its CFC is of the form $B = I_{m_1} \oplus 0_{m_2}$. The following result allows us to get rid of the null diagonal blocks.

Lemma 2.2. *Let $A = \left[\begin{smallmatrix} \widetilde{A} & 0 \\ 0 & 0_d \end{smallmatrix}\right]$ and $B = \left[\begin{smallmatrix} \widetilde{B} & 0 \\ 0 & 0_k \end{smallmatrix}\right]$. Then, the equation $X^\top AX = B$ is consistent if and only if the equation $X^\top \widetilde{A}X = \widetilde{B}$ is consistent.*

Proof. Let us assume first that $X^\top AX = B$ is consistent, with A and B as in the statement. Let us partition

$$X = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix},$$

where X_{22} has size $d \times k$. Then, $X^\top AX = B$ can be written as

$$\begin{bmatrix} X_{11}^\top & X_{21}^\top \\ X_{12}^\top & X_{22}^\top \end{bmatrix} \begin{bmatrix} \widetilde{A} & 0 \\ 0 & 0_d \end{bmatrix} \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} = \begin{bmatrix} X_{11}^\top \widetilde{A} X_{11} & X_{11}^\top \widetilde{A} X_{12} \\ X_{12}^\top \widetilde{A} X_{11} & X_{12}^\top \widetilde{A} X_{12} \end{bmatrix} = \begin{bmatrix} \widetilde{B} & 0 \\ 0 & 0_k \end{bmatrix},$$

so, in particular, X_{11} is a solution of $X^\top \widetilde{A}X = \widetilde{B}$.

Conversely, assume that $X^\top \widetilde{A}X = \widetilde{B}$ has a solution X_{11}. Then, $X = \left[\begin{smallmatrix} X_{11} & 0 \\ 0 & 0_{k\times d} \end{smallmatrix}\right]$ is a solution of $X^\top AX = B$. □

The next result deals also with Eq. (1.1) with B symmetric, already in CFC, and of full rank.

Lemma 2.3. *If $X^\top AX = I_{m+1}$ is consistent, then $X^\top AX = I_m$ is consistent as well, for all $m \geqslant 1$.*

Proof. If X_0 is a solution of $X^\top AX = I_{m+1}$, then $X_0 \begin{bmatrix} I_m \\ 0_{1\times m} \end{bmatrix}$ is a solution of $X^\top AX = I_m$. □

The last result shows the transitivity of the consistency of Eq. (1.1).

Lemma 2.4. *If both $XAX^\top = B$ and $YBY^\top = C$ are consistent, then $ZAZ^\top = C$ is also consistent.*

Proof. Let X_0 be a solution of $XAX^\top = B$, and let Y_0 be a solution of $YBY^\top = C$. Since $(X_0Y_0)^\top A(X_0Y_0) = Y_0^\top(X_0^\top AX_0)Y_0 = Y_0^\top BY_0 = C$, then $Z_0 = X_0Y_0$ is a solution of $ZAZ^\top = C$ □

3. A necessary condition

In this Section, we introduce a necessary condition on A for Eq. (1.1) to be consistent when B is symmetric.

Lemma 2.2 guarantees that, when looking for the consistency of (1.1), there is no loss of generality in assuming that A and B are given in CFC, that A has no blocks of type $J_1(0)$, and that $B = I_m$ (the CFC of symmetric invertible $m \times m$ matrices). In particular, in the next theorem we get a necessary condition for Eq. (1.1) to be consistent with such B. In the statement we have included, however, the case where CFC(A) contains blocks of type $J_1(0)$, for the sake of completeness.

Theorem 2. *Let $A \in \mathbb{C}^{n\times n}$ be a matrix whose* CFC *has*

(i) *exactly d Type-0 blocks with size 1;*
(ii) *exactly r Type-0 blocks with odd size greater than 1;*
(iii) *exactly s Type-I blocks with odd size;*
(iv) *exactly t Type-II blocks of the form $H_{4v}(1)$, and*
(v) *an arbitrary number of Type-0, Type-I, and Type-II blocks with other sizes.*

Then, in order for

$$X^\top AX = I_m \tag{3.1}$$

to be consistent, it must be

$$n - d \geqslant 2m - r - s - 2t. \tag{3.2}$$

Proof. As mentioned before, we may assume that A is given in CFC.

Assume first that $d = 0$. In the conditions of the statement, we can write

$$\begin{aligned} A \;=\; & J_{2m_1-1}(0) \oplus \cdots \oplus J_{2m_r-1}(0) \bigoplus J_{2m_{r+1}}(0) \oplus \cdots \oplus J_{2m_{r+h}}(0) \bigoplus \\ & \Gamma_{2\widehat{m}_1-1} \oplus \cdots \oplus \Gamma_{2\widehat{m}_s-1} \bigoplus \Gamma_{2\widehat{m}_{s+1}} \oplus \cdots \oplus \Gamma_{2\widehat{m}_{s+g}} \bigoplus \\ & H_{2\widetilde{m}_1}(\mu_1) \oplus \cdots \oplus H_{2\widetilde{m}_u}(\mu_u) \bigoplus H_{4\widetilde{m}_{u+1}-2}(-1) \oplus \cdots \oplus H_{4\widetilde{m}_{u+v}-2}(-1) \bigoplus \\ & H_{4\widetilde{m}_{u+v+1}}(1) \oplus \cdots \oplus H_{4\widetilde{m}_{u+v+t}}(1), \end{aligned}$$

where $m_i > 1$, for $i = 1, \dots, r$; $\widehat{m}_i \geqslant 1$, for $i = 1, \dots, s+g$; $\check{m}_i \geqslant 1$, for $i = 1, \dots, u+v+t$; and $\mu_j \neq 0, \pm 1$, for $j = 1, \dots, u$. Let us set

$$\begin{array}{l} M := m_1 + \cdots + m_{r+h}, \qquad \widehat{M} := \widehat{m}_1 + \cdots + \widehat{m}_{s+g}, \qquad \text{and} \\ \widecheck{M} := \check{m}_1 + \cdots + \check{m}_u + 2\check{m}_{u+1} + \cdots + 2\check{m}_{u+v} + 2\check{m}_{u+v+1} + \cdots + 2\check{m}_{u+v+t}. \end{array}$$

Then, we have

$$n = 2(M + \widehat{M} + \widecheck{M}) - r - s - 2v, \tag{3.3}$$
$$\operatorname{rank}(A - A^\top) = 2(M + \widehat{M} + \widecheck{M}) - 2r - 2s - 2v - 2t. \tag{3.4}$$

Equation (3.3) is immediate. In order to get (3.4), it suffices to check that the rank of the skew-symmetric part of the canonical blocks involved in this identity is the following

$$\operatorname{rank}(J_n(0) - J_n(0)^\top) = \begin{cases} n-1 & \text{if } n \text{ is odd}, \\ n & \text{if } n \text{ is even}, \end{cases} \tag{3.5}$$
$$\operatorname{rank}(\Gamma_n - \Gamma_n^\top) = \begin{cases} n-1 & \text{if } n \text{ is odd}, \\ n & \text{if } n \text{ is even}, \end{cases} \tag{3.6}$$
$$\operatorname{rank}(H_{2n}(\mu) - H_{2n}(\mu)^\top) = 2n \quad \text{if } \mu \neq 0, \pm 1, \tag{3.7}$$
$$\operatorname{rank}(H_{4n-2}(-1) - H_{4n-2}(-1)^\top) = 4n-2, \tag{3.8}$$
$$\operatorname{rank}(H_{4n}(1) - H_{4n}(1)^\top) = 4n-2. \tag{3.9}$$

The first identity of (3.5) (i.e. for n odd) is a consequence of the fact that any skew-symmetric matrix with odd size is singular, together with the fact that $e_1, \dots, e_{n-1}$ belong to the column space of $J_n(0) - J_n(0)^\top$. To get the second identity in Equation (3.5) (i.e. for n even) we can prove that $\det(J_n(0) - J_n(0)^\top) = \pm 1$ when n is even. This can be done by induction, spanning the determinant across the first row, then across the first column, and then using induction in the remaining minor, namely $\det(J_{n-2}(0) - J_{n-2}(0)^\top)$.

Equation (3.6) is a consequence of the identity

$$\Gamma_n - \Gamma_n^\top = \begin{cases} \begin{bmatrix} & & & 0 \\ & & \iddots & -2 \\ & 0 & \iddots & \\ 0 & 2 & & \end{bmatrix} & \text{if } n \text{ is odd}, \\ \begin{bmatrix} & & -2 \\ & \iddots & \\ -2 & & \\ 2 & & \end{bmatrix} & \text{if } n \text{ is even}. \end{cases}$$

Finally, Equations (3.7)–(3.9) follow from the identity:

$$H_{2n}(\mu) - H_{2n}(\mu)^\top = \left[\begin{array}{cccc|cccc} & & & & 1-\mu & -1 & & \\ & & & & & \ddots & \ddots & \\ & & & & & & 1-\mu & -1 \\ & & & & & & & 1-\mu \\ \hline \mu-1 & & & & & & & \\ 1 & \ddots & & & & & & \\ & \ddots & \mu-1 & & & & & \\ & & 1 & \mu-1 & & & & \end{array}\right].$$

From (3.3)–(3.4) we conclude

$$n - \operatorname{rank}(A - A^\top) = r + s + 2t. \tag{3.10}$$

Now, transposing (3.1) and subtracting, we get

$$X^\top (A - A^\top) X = 0. \tag{3.11}$$

Since $X \in \mathbb{C}^{n\times m}$ is a solution of (3.1), it must be $\operatorname{rank} X = \operatorname{rank} X^\top = m$. Using this fact, together with the well-know inequality (see [16, page 13])

$$\operatorname{rank}(PQ) \geqslant \operatorname{rank} P + \operatorname{rank} Q - k, \qquad \text{for } P \in \mathbb{C}^{p\times k} \quad \text{and} \quad Q \in \mathbb{C}^{k\times q}, \tag{3.12}$$

we obtain

$$\operatorname{rank}(X^\top (A - A^\top)) \geqslant m + \operatorname{rank}(A - A^\top) - n,$$

and, then, using this inequality and (3.10), we get

$$\begin{aligned} \dim(\operatorname{Nul}(X^\top (A - A^\top)) &= n - \operatorname{rank}(X^\top (A - A^\top)) \\ &\leqslant n - (m + \operatorname{rank}(A - A^\top) - n) \\ &= (n - m) + (n - \operatorname{rank}(A - A^\top)) \\ &= n - m + r + s + 2t. \end{aligned}$$

Equation (3.11) implies that the column space of $X \in \mathbb{C}^{n\times m}$ is contained in the null space of $X^\top (A - A^\top)$. Since $\operatorname{rank} X = m$, the column space of X has dimension m, so it must be

$$m \leqslant n - m + r + s + 2t$$

or, in other words,

$$n \geqslant 2m - r - s - 2t,$$

as wanted for the case in which $d = 0$.

Assume now that $d > 0$. Then A can be written as in the statement of Lemma 2.2, with $\widetilde{A} \in \mathbb{C}^{(n-d)\times(n-d)}$ being the matrix obtained from A by deleting the d blocks $J_1(0)$. By Lemma 2.2, $X^\top AX = I_m$ is consistent if and only if $X^\top \widetilde{A} X = I_m$ is consistent. As we have just seen, the condition

$$n - d \geqslant 2m - r - s - 2t$$

is necessary for $X^\top \widetilde{A} X = I_m$ to be consistent, so it is also necessary for $X^\top AX = I_m$ to be consistent. □

We want to emphasize that Theorem 2 provides a necessary condition for Eq. (1.1) to be consistent, when B is symmetric, that covers all possible matrices A. To see this, just note that conditions (i)–(v) in the statement are not restrictive at all, but just a particular description of the CFC of an arbitrary matrix A.

The main open question after Theorem 2 is whether or not condition (3.2) is sufficient. We will show that it is sufficient with very few exceptions (see Theorem 8). The next example shows one of such exceptions.

Example 1. Let

$$A = H_2(-1) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \qquad B = I_1 = 1.$$

Then, we are in the case $n = 2, m = 1, r = s = t = 0$ in the statement of Theorem 2, so (3.2) is satisfied (recall that $\mu = -1$ is allowed in $H_2(\mu)$, since, for $H_{2n}(\mu)$, only $\mu = 0, (-1)^{n+1}$ are not allowed). However, Eq. (1.1) is not consistent, since

$$X^\top AX = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 & x_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0.$$

The other exception for condition (3.2) to be sufficient is the presence of Type-II blocks of the form $H_4(1)$ in the CFC of A (see Theorem 7).

4. The case where CFC(A) is a Type-0 block

We start with the following result, whose proof is straightforward.

Lemma 4.1. *The equation* $X^\top J_{n+1}(0)X = J_n(0)$ *is consistent for* $n \geqslant 1$. *A solution is* $X = \begin{bmatrix} 0_{1\times n} \\ I_n \end{bmatrix}$.

The following result provides a characterization for the consistency of (3.1) when A is a single Type-0 block.

Theorem 3. *Given* $m \geqslant 1$, *the equation*

$$X^\top J_n(0)X = I_m \tag{4.1}$$

is consistent if and only if one of the following situations hold:

1. $n > 1$ *is odd and* $n \geqslant 2m - 1$; *or*
2. n *is even and* $n \geqslant 2m$.

Proof. Clearly $X^\top J_1(0)X = I_m$ is not consistent, since $J_1(0) = [\,0\,]$. The proof for $n > 1$ is divided in two cases:

1. Case $n > 1$ odd. In Theorem 2, the necessary condition (3.2) for $X^\top J_n(0)X = I_m$ to be consistent when n is odd reads $n \geqslant 2m - 1$. Let us see that $n \geqslant 2m - 1$ is also sufficient:

(a) If $n = 2m-1$ with m even, then $X^\top J_{2m-1}(0)X = I_m$ is consistent, and a solution is given by

$$X = \underbrace{\begin{bmatrix} X_{32} \\ 0_{1\times 2} \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} X_{32} \\ 0_{1\times 2} \end{bmatrix}}_{\frac{m-2}{2} \text{ times}} \oplus X_{32}, \quad \text{where } X_{32} = \begin{bmatrix} 1 & 0 \\ 1 & -\mathrm{i} \\ 0 & \mathrm{i} \end{bmatrix}.$$

(b) If $n = 2m-1$ with $m \geqslant 3$ odd, then $X^\top J_{2m-1}(0)X = I_m$ is consistent, and a solution is given by

$$X = \underbrace{\begin{bmatrix} X_{32} \\ 0_{1\times 2} \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} X_{32} \\ 0_{1\times 2} \end{bmatrix}}_{\frac{m-3}{2} \text{ times}} \oplus X_{53},$$

$$\text{where } X_{32} = \begin{bmatrix} 1 & 0 \\ 1 & -\mathrm{i} \\ 0 & \mathrm{i} \end{bmatrix} \quad \text{and} \quad X_{53} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & \frac{\mathrm{i}}{\sqrt{2}} & 0 \\ 0 & -\frac{\mathrm{i}}{\sqrt{2}} & 0 \\ 0 & \frac{\mathrm{i}}{\sqrt{2}} & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

(c) Let $n = 2\widetilde{m} - 1 > 2m - 1$. In 1(a) and 1(b) we have seen that the equation $X^\top J_{2\widetilde{m}-1}(0)X = I_{\widetilde{m}}$ is consistent, and Lemma 2.3 implies that $X^\top I_{\widetilde{m}} X = I_m$ is consistent as well. These results, together with Lemma 2.4, imply that $X^\top J_{2\widetilde{m}-1}(0)X = I_m$ is consistent.

2. Case n even. In Theorem 2, the necessary condition (3.2) for $X^\top J_n(0)X = I_m$ to be consistent when n is even reads $n \geqslant 2m$. Let us see that, if $n \geqslant 2m$, then it is also sufficient:
   (a) If $n = 2$ and $m = 1$, then the equation $X^\top J_2(0)X = I_1$ is consistent and a solution is $X = \left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$.
   (b) Let $n \geqslant 4$, with $n = 2\widetilde{m} \geqslant 2m$. In Lemma 4.1 we saw that $X^\top J_{2\widetilde{m}}(0)X = J_{2\widetilde{m}-1}(0)$ is consistent. In cases 1(a) and 1(b) we have seen that $X^\top J_{2\widetilde{m}-1}(0)X = I_{\widetilde{m}}$ is consistent, and, if $\widetilde{m} > m$, Lemma 2.3 implies that $X^\top I_{\widetilde{m}} X = I_m$ is consistent as well. These results, together again with Lemma 2.4, imply that $X^\top J_{2m}(0)X = I_m$ is consistent.

□

Note that the proof of Theorem 3 is constructive since, in the case where Eq. (4.1) is consistent, we provide an explicit solution of this equation.

5. The case where CFC(A) is a Type-I block

The first result in this section, whose proof is straightforward, will be used later. We denote by

$$R_n := \begin{bmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{bmatrix}_{n\times n}$$

the $n \times n$ antidiagonal matrix whose antidiagonal entries are all equal to 1.

Lemma 5.1. *The equation*

$$X^\top \Gamma_{n+1} X = \Gamma_n$$

is consistent for $n \geqslant 1$. In particular:

(i) *if n is odd, then $X = \begin{bmatrix} 0_{1\times n} \\ R_n \end{bmatrix}$ is a solution;*

(ii) *if n is even, then $X = \begin{bmatrix} 0_{1\times n} \\ \mathfrak{i} R_n \end{bmatrix}$ is a solution.*

The following result, analogous to Theorem 3, characterizes the consistency of (3.1) when A is a single Type-I block.

Theorem 4. *Given $m \geqslant 1$, the equation*

$$X^\top \Gamma_n X = I_m \tag{5.1}$$

is consistent if and only if one of the following situations hold:

1. *n is odd and $n \geqslant 2m - 1$, or*
2. *n is even and $n \geqslant 2m$.*

Proof. 1. Case n odd. In Theorem 2, the necessary condition (3.2) for $X^\top \Gamma_n X = I_m$ to be consistent when n is odd reads $n \geqslant 2m - 1$. Let us see that $n \geqslant 2m - 1$ is also sufficient:

(a) Let $n = 2m-1$. We will see that a solution of $X^\top \Gamma_{2m-1} X = I_m$ is given by the matrix X defined below. To understand the pattern it is important to note that (for both the odd and the even cases) the rows alternate between having a consecutive list of 1's and a consecutive list of $\mathfrak{i}$'s.

$$X^\top = \begin{cases} \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 & 1 & \cdots & 1 & 1 \\ 0 & \mathfrak{i} & \mathfrak{i} & \cdots & \mathfrak{i} & \mathfrak{i} & \cdots & \mathfrak{i} & \mathfrak{i} \\ 0 & 0 & 1 & \cdots & 1 & 1 & \cdots & 1 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \iddots & \iddots & \vdots \\ 0 & 0 & \dots & 0 & 1 & 1 & 0 & \cdots & 0 \end{bmatrix} & \text{if } m \text{ is odd,} \\ \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 & 1 & \cdots & 1 & 1 \\ 0 & \mathfrak{i} & \mathfrak{i} & \cdots & \mathfrak{i} & \mathfrak{i} & \cdots & \mathfrak{i} & \mathfrak{i} \\ 0 & 0 & 1 & \cdots & 1 & 1 & \cdots & 1 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \iddots & \iddots & \vdots \\ 0 & 0 & \dots & 0 & \mathfrak{i} & \mathfrak{i} & 0 & \cdots & 0 \end{bmatrix} & \text{if } m \text{ is even.} \end{cases} \tag{5.2}$$

We start with the case $n = 2m - 1$ when m is odd.

Let $Y_1, \dots, Y_{2m-1}$ be the columns of the matrix $X^\top$ in (5.2). Then

$$\begin{aligned} X^\top \Gamma_{2m-1} &= \begin{bmatrix} Y_1 & Y_2 & \cdots & Y_{2m-1} \end{bmatrix} \Gamma_{2m-1} \\ &= \begin{bmatrix} Y_{2m-1} & Y_{2m-1} - Y_{2m-2} & \cdots & -Y_{m+1} + Y_m & \cdots & Y_3 - Y_2 & -Y_2 + Y_1 \end{bmatrix}. \end{aligned}$$

Finally,

$$(X^\top \Gamma_{2m-1}) X = \begin{bmatrix} B_1 & B_2 & B_3 & B_4 & \dots & B_m \end{bmatrix},$$

where the columns of the product are computed below, in such a way that almost everything cancels out (note that $Y_{m-i} = Y_{m+1+i}$ for $i = 0, \dots, m-1$) after reordering the summands:

$$\begin{aligned} B_1 &= Y_{2m-1} + (Y_{2m-1} - Y_{2m-2}) + (-Y_{2m-2} + Y_{2m-3}) + \cdots + \\ &\quad (Y_3 - Y_2) + (-Y_2 + Y_1) \;=\; 2(Y_{2m-1} - Y_2) + 2(-Y_{2m-2} + Y_3) + \cdots + \\ &\quad 2(-Y_{m+1} + Y_m) + Y_1 \;=\; Y_1 \;=\; e_1; \\ B_2 &= \mathfrak{i}\Big((Y_{2m-1} - Y_{2m-2}) + (-Y_{2m-2} + Y_{2m-3}) + \cdots + (Y_3 - Y_2) + (-Y_2 + Y_1)\Big) \\ &= \mathfrak{i}\Big((Y_{2m-1} - Y_2) + 2(-Y_{2m-2} + Y_3) + \cdots + 2(-Y_{m+1} + Y_m) + (-Y_2 + Y_1)\Big) \\ &= \mathfrak{i}(Y_1 - Y_2) \;=\; e_2; \\ B_3 &= (-Y_{2m-2} + Y_{2m-3}) + (Y_{2m-3} - Y_{2m-4}) + \cdots + (Y_3 - Y_2) + (-Y_2 + Y_1) \\ &= (-Y_{2m-2} + Y_3) + 2(Y_{2m-3} - Y_4) + \cdots + 2(-Y_{m+1} + Y_m) + (Y_3 - Y_2) \\ &= Y_3 - Y_2 \;=\; e_3; \\ &\vdots \\ B_m &= (-Y_{m+1} + Y_m) + (Y_m - Y_{m-1}) \;=\; (Y_m - Y_{m-1}) \;=\; e_m. \end{aligned}$$

The proof for the case $n = 2m-1$ when m is even is analogous.

(b) Let $n = 2\widetilde{m}-1 > 2m-1$. We have just seen in (a) that $X^\top \Gamma_{2\widetilde{m}-1} X = I_{\widetilde{m}}$ is consistent, and, in Lemma 2.3 we saw that $X^\top I_{\widetilde{m}} X = I_m$ is consistent as well. Using Lemma 2.4, we conclude that $X^\top \Gamma_{2\widetilde{m}-1} X = I_m$ is consistent as well.

2. Case n even. In Theorem 2, the necessary condition (3.2) for the consistency of $X^\top \Gamma_n X = I_m$ when n is even reads $n \geqslant 2m$. Let us see that it is also sufficient. Suppose that $n = 2\widetilde{m} \geqslant 2m$. In Lemma 5.1 we saw that $X^\top \Gamma_{2\widetilde{m}} X = \Gamma_{2\widetilde{m}-1}$ is consistent. In the "n odd case" we have seen that $X^\top \Gamma_{2\widetilde{m}-1} X = I_{\widetilde{m}}$ is consistent and, if $\widetilde{m} > m$, in Lemma 2.3 we saw that $X^\top I_{\widetilde{m}} X = I_m$ is also consistent. Then, Lemma 2.4 implies that $X^\top \Gamma_{2\widetilde{m}} X = I_m$ is consistent.

□

Note that the proof of Theorem 4 is constructive. More precisely, when Eq. (5.1) is consistent, we provide an explicit solution of this equation, like in the proof of Theorem 3 with Eq. (4.1).

6. The case where CFC(A) is a Type-II block

Recall that $H_{2n}(\mu)$ is a Type-II block if and only if $\mu \neq 0, (-1)^{n+1}$. The first result in this section, whose proof is straightforward, is the analogue of Lemmas 4.1 and 5.1 for Type-II blocks.

Lemma 6.1. *For any complex μ and any $n \geqslant 1$, the equation*

$$X^\top H_{2n+2}(\mu) X = H_{2n}(\mu)$$

is consistent. A particular solution is the $(2n+2)\times(2n)$ *matrix*

$$X = \begin{bmatrix} I_n & 0_n \\ 0_{1\times n} & 0_{1\times n} \\ 0_n & I_n \\ 0_{1\times n} & 0_{1\times n} \end{bmatrix}.$$

The following three results are the analogues of Theorems 3 and 4 for a single Type-II block $H_{2k}(\mu)$, depending on whether $\mu \neq 0, \pm 1$; $\mu = -1$; or $\mu = 1$.

Theorem 5. *The equation*

$$X^\top H_{2m}(\mu) X = I_m, \qquad \mu \neq 0, \pm 1 \tag{6.1}$$

is consistent.

Proof. Let us first consider the cases $m = 1, 2$, and then the case $m \geqslant 3$.

1. Case $m = 1$. Let

$$X^\top H_2(\mu) X = \begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} 0 & 1 \\ \mu & 0 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} ab(1+\mu) \end{bmatrix}.$$

   Then a solution is $X = \begin{bmatrix} 1 & \frac{1}{1+\mu} \end{bmatrix}^\top$.
2. Case $m = 2$. Let $X = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & \frac{1}{\mu-1} & 0 \end{bmatrix}^\top$, so

$$X^\top H_4(\mu) X = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & \frac{1}{\mu-1} & 0 \end{bmatrix} \left[\begin{array}{cc|cc} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline \mu & 1 & 0 & 0 \\ 0 & \mu & 0 & 0 \end{array}\right] \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & \frac{1}{\mu-1} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \mu+1 & \frac{\mu}{\mu-1} \\ \frac{\mu}{\mu-1} & \frac{1}{\mu-1} \end{bmatrix},$$

   and let

$$\det(X^\top H_4(\mu) X) = \frac{-1}{(\mu-1)^2} \neq 0.$$

   Then $X^\top H_4(\mu) X$ is symmetric and nonsingular, so its CFC is I_2. Therefore, $X^\top H_4(\mu) X = I_2$ is consistent.
3. Case $m \geqslant 3$. Consider the following matrices Z_a of size $m \times m$ and X_a of size $2m \times m$:

$$Z_a := \begin{bmatrix} 1 & \frac{1}{\mu-1} & 0 & \cdots & 0 & 0 \\ 0 & 1 & \frac{1}{\mu-1} & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & \frac{1}{\mu-1} \\ 0 & 0 & 0 & \cdots & & a \end{bmatrix} \qquad \text{and} \qquad X_a := \begin{bmatrix} I_m \\ Z_a \end{bmatrix}.$$

Then

$$X_a^\top H_{2m}(\mu)X_a = \begin{bmatrix} I_m & Z_a^T \end{bmatrix} \begin{bmatrix} 0 & I_m \\ J_m(\mu) & 0 \end{bmatrix} \begin{bmatrix} I_m \\ Z_a \end{bmatrix} = Z_a^\top J_m(\mu) + Z_a \quad (6.2)$$

$$= \frac{1}{\mu-1} \begin{bmatrix} \mu^2-1 & \mu & 0 & \cdots & 0 & 0 \\ \mu & \mu^2 & \mu & \cdots & 0 & 0 \\ 0 & \mu & \mu^2 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \mu & \mu^2 & \mu \\ 0 & \cdots & 0 & 0 & \mu & a(\mu^2-1)+1 \end{bmatrix}.$$

This matrix is symmetric. So, if it has rank equal to k, then its CFC is $I_k \oplus 0_{m-k}$. Clearly the rank is, at least, $m-1$, since the submatrix of size $(m-1)\times(m-1)$ obtained by deleting the first row and the last column has determinant equal to $(\frac{\mu}{\mu-1})^{m-1} \neq 0$. If we prove that, for any $\mu \neq 0, \pm 1$, there exists some value of a such that $\det(X_a^\top H_{2m}(\mu)X_a) \neq 0$, then $\operatorname{rank}(X_a^\top H_{2m}(\mu)X_a) = m$, which implies that the CFC of $X_a^\top H_{2m}(\mu)X_a$ is I_m, and the proof is finished.

For $k = 1, \ldots, m-2$, consider the matrix $Y_{m-k}(\mu)$ of size $(m-k)\times(m-k)$ obtained by deleting the last k rows and the last k columns of $X_a^\top H_{2m}(\mu)X_a$, multiplied by $\mu-1$, that is,

$$Y_{m-k}(\mu) := \begin{bmatrix} \mu^2-1 & \mu & 0 & \cdots & 0 & 0 \\ \mu & \mu^2 & \mu & \ddots & 0 & 0 \\ 0 & \mu & \mu^2 & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \mu & \mu^2 & \mu \\ 0 & \cdots & 0 & 0 & \mu & \mu^2 \end{bmatrix}.$$

Note that

$$\det(Y_2(\mu)) = \det \begin{bmatrix} \mu^2-1 & \mu \\ \mu & \mu^2 \end{bmatrix} = \mu^4 - 2\mu^2 = 0 \Leftrightarrow \mu \in \{0, \pm\sqrt{2}\} \quad (6.3)$$

and

$$\det(Y_3(\mu)) = \det \begin{bmatrix} \mu^2-1 & \mu & 0 \\ \mu & \mu^2 & \mu \\ 0 & \mu & \mu^2 \end{bmatrix} = \mu^6 - 3\mu^4 + \mu^2 = 0 \Leftrightarrow \mu \in \left\{0, \pm\sqrt{\frac{3\pm 5}{2}}\right\}. \quad (6.4)$$

The only common root of $Y_2(\mu)$ and $Y_3(\mu)$ is $\mu = 0$.

Suppose now that, for some $\mu_0 \neq 0$ and some $k \geqslant 3$, we have

$$\det(Y_k(\mu_0)) = \det(Y_{k+1}(\mu_0)) = 0.$$

Spanning the determinant of $Y_{k+1}(\mu_0)$ across the last row we obtain

$$\det(Y_{k+1}(\mu_0)) = \mu^2 \det(Y_k(\mu_0)) - \mu^2 \det(Y_{k-1}(\mu_0)),$$

and, equivalently,

$$\det(Y_{k-1}(\mu_0)) = \det(Y_k(\mu_0)) - \frac{\det(Y_{k+1}(\mu_0))}{\mu^2}.$$

This implies that

$$\det(Y_2(\mu_0)) = \det(Y_3(\mu_0)) = \cdots = \det(Y_{k+1}(\mu_0)) = 0,$$

which contradicts (6.3) and (6.4), since there is no $\mu_0 \neq 0$ which makes $Y_2(\mu_0)$ and $Y_3(\mu_0)$ to be singular at the same time. Therefore, we conclude that, for any $k \geqslant 2$ and any $\mu_0 \neq 0$, either $\det(Y_k(\mu_0)) \neq 0$ or $\det(Y_{k+1}(\mu_0)) \neq 0$.

Now we come back to the matrix $X_a^\top H_{2m}(\mu)X_a$ in (6.2). Spanning its determinant across the last row, we get

$$\det(X_a^\top H_{2m}(\mu)X_a) = \frac{1}{(\mu-1)^m}\Big[\big(a(\mu^2-1)+1\big)\det(Y_{m-1}(\mu)) - \mu^2\det(Y_{m-2}(\mu))\Big].$$

Let $\mu_0 \notin \{0, 1, -1\}$. Then:

(a) If $\det(Y_{m-2}(\mu_0)) = s \neq 0$, take $a = \frac{1}{1-\mu^2}$, and so $\det(X_a^\top H_{2m}(\mu_0)X_a) = -\frac{\mu^2}{(\mu-1)^m}s \neq 0$.

(b) If $\det(Y_{m-2}(\mu_0)) = 0$, then $\det(Y_{m-1}(\mu_0)) \neq 0$. Take $a \neq \frac{1}{1-\mu^2}$, so that $\det(X_a^\top H_{2m}(\mu_0)X_a) \neq 0$.

Then, for any $\mu \neq 0, \pm 1$, there exists some a such that

$$\det(X_a^\top H_{2m}(\mu)X_a) \neq 0,$$

and the proof is finished.

□

Theorem 6. *For all m odd, the equation*

$$X^\top H_{2m}(-1)X = I_m \tag{6.5}$$

is consistent if and only if $m \neq 1$.

Proof. When $m = 1$, (6.5) is inconsistent (see Example 1).

Assume that $m \geqslant 3$. Consider the matrix $\begin{bmatrix} I_m & Z \end{bmatrix}$ of size $m \times 2m$ with

$$Z = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 & 1 & 1/2 \\ 0 & 0 & \cdots & 0 & 1 & 0 & 0 \\ 0 & 0 & \cdots & 1 & 0 & 0 & 0 \\ \vdots & \vdots & \iddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & \cdots & 0 & 0 & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Then

$$\begin{bmatrix} I_m & Z \end{bmatrix} \begin{bmatrix} 0 & I_m \\ J_m(-1) & 0 \end{bmatrix} \begin{bmatrix} I_m \\ Z^\top \end{bmatrix} = ZJ_m(-1) + Z^\top = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \iddots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & \cdots & 0 & 0 & 0 \end{bmatrix}$$

is symmetric and nonsingular, and so its CFC is I_m. Therefore (6.5) is consistent.

We can provide a particular solution of (6.5). Consider the $m\times m$ matrix:

$$Y=\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & 1\\ 0 & 1 & 0 & \cdots & 0 & \cdots & 0 & \frac{1}{2} & 0\\ \vdots & \ddots & \ddots & \ddots & \vdots & \iddots & \iddots & \iddots & \vdots\\ 0 & \cdots & 0 & 1 & 0 & \frac{1}{2} & 0 & \cdots & 0\\ 0 & \cdots & 0 & 0 & 1 & 0 & 0 & \cdots & 0\\ 0 & \cdots & 0 & \mathrm{i} & 0 & -\frac{\mathrm{i}}{2} & 0 & \cdots & 0\\ 0 & \iddots & \iddots & \iddots & \vdots & \ddots & \ddots & \ddots & \vdots\\ 0 & \mathrm{i} & 0 & \cdots & 0 & \cdots & 0 & -\frac{\mathrm{i}}{2} & 0\\ \mathrm{i} & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & -\mathrm{i} \end{bmatrix}.$$

Multiplying by blocks, it is straightforward to check that

$$Y\cdot\begin{bmatrix} I_m & Z\end{bmatrix} H_{2m}(-1)\begin{bmatrix} I_m\\ Z^\top\end{bmatrix}\cdot Y^\top = I_m.$$

The matrix $Y\begin{bmatrix} I_m & Z\end{bmatrix}$ is equal to

$$X^\top=\left[\begin{array}{ccccccccc|cccccccccc} 1 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & \frac{1}{2}\\ 0 & 1 & 0 & \cdots & 0 & \cdots & 0 & \frac{1}{2} & 0 & \frac{1}{2} & \ddots & \vdots & \vdots & \vdots & \iddots & 1 & 0 & 0\\ \vdots & \ddots & \ddots & \ddots & \vdots & \iddots & \iddots & \iddots & \vdots & 0 & \ddots & 0 & 0 & 0 & \iddots & \iddots & 0 & 0\\ 0 & \cdots & 0 & 1 & 0 & \frac{1}{2} & 0 & \cdots & 0 & 0 & \ddots & \frac{1}{2} & 0 & 1 & 0 & \cdots & 0 & 0\\ 0 & \cdots & 0 & 0 & 1 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 1 & 0 & 0 & \cdots & 0 & 0\\ 0 & \cdots & 0 & \mathrm{i} & 0 & -\frac{\mathrm{i}}{2} & 0 & \cdots & 0 & 0 & \iddots & -\frac{\mathrm{i}}{2} & 0 & \mathrm{i} & 0 & \cdots & 0 & 0\\ 0 & \iddots & \iddots & \iddots & \vdots & \ddots & \ddots & \ddots & \vdots & 0 & \iddots & 0 & 0 & 0 & \ddots & \ddots & \vdots & \vdots\\ 0 & \mathrm{i} & 0 & \cdots & 0 & \cdots & 0 & -\frac{\mathrm{i}}{2} & 0 & -\frac{\mathrm{i}}{2} & \iddots & \vdots & \vdots & \vdots & \ddots & \mathrm{i} & 0 & 0\\ \mathrm{i} & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 & -\mathrm{i} & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \mathrm{i} & \frac{\mathrm{i}}{2} \end{array}\right] \tag{6.6}$$

and gives a particular solution of (6.5). □

Theorem 7. *For all m even, the equation*

$$X^\top H_{2m}(1)X = I_{m+1} \tag{6.7}$$

is consistent if and only if $m\neq 2$.

Proof. Let us partition $X^\top=\begin{bmatrix} Y & Z\end{bmatrix}$, with $Y,Z\in\mathbb{C}^{(m+1)\times m}$. Then

$$\begin{aligned} X^\top H_{2m}(1)X &= \begin{bmatrix} Y & Z\end{bmatrix}\begin{bmatrix} 0 & I_m\\ J_m(1) & 0\end{bmatrix}\begin{bmatrix} Y^\top\\ Z^\top\end{bmatrix}\\ &= ZJ_m(1)Y^\top + YZ^\top = ZJ_m(0)Y^\top + ZY^\top + YZ^\top. \end{aligned} \tag{6.8}$$

The consistency of (6.7) is equivalent to the matrix (6.8) being symmetric and nonsingular. If this is the case, the CFC of the matrix in (6.8) is I_{m+1}, so the matrix will be congruent to I_{m+1}. As the sum $ZY^\top + YZ^\top$ is symmetric, the matrix in (6.8) is symmetric if and only if $ZJ_m(0)Y^\top$ is symmetric as well. Let us first analyze the case $m=2$, and then the case $m\geqslant 4$:

- $\boxed{m=2}$ We will show that the matrix equation $X^\top H_4(1)X = I_3$ is not consistent. To do this we will show that, if $ZJ_2(0)Y^\top$ is symmetric, then $X^\top H_4(1)X$ must be singular, so it can not be equal to I_3.

  Let Y_1, Y_2 and Z_1, Z_2 denote the columns of Y and Z, respectively. If $ZJ_2(0)Y^\top$ is symmetric, then $ZJ_2(0)Y^\top = YJ_2(0)^\top Z^\top$, which is equivalent to $Z_1Y_2^\top = Y_2Z_1^\top$, and this implies $Z_1=\alpha Y_2$, for some $\alpha\in\mathbb{C}$. If $\operatorname{rank}\begin{bmatrix} Y_1 & Z_1 & Z_2\end{bmatrix} < 3$, then $\operatorname{rank} X < 3$, which, in turn, implies rank $(X^\top H_4(1)X) < 3$, and we are finished. Otherwise, there is some nonzero vector $v \in \mathbb{C}^3$ satisfying

$$\begin{bmatrix} Y_1^\top \\ Z_1^\top \\ Z_2^\top \end{bmatrix} v = \begin{bmatrix} -1 \\ 0 \\ \alpha \end{bmatrix}.$$

Then:

$$\begin{aligned} X^\top H_4(1)Xv &= ZJ_2(0)Y^\top v + (ZY^\top)v + (YZ^\top)v \\ &= Z_1Y_2^\top v + Z_1Y_1^\top v + Z_2Y_2^\top v + Y_1Z_1^\top v + Y_2Z_2^\top v \\ &= 0 - Z_1 + 0 + 0 + \alpha Y_2 = 0. \end{aligned}$$

This means that there is a nonzero vector in the right null space of $X^\top H_4(1)X$, so $X^\top H_4(1)X$ is singular, as claimed.

- $\boxed{m \geqslant 4}$ Let $X \in \mathbb{C}^{2m\times(m+1)}$ be given by

$$X^\top = \left[\begin{array}{c|c} Y & Z \end{array}\right] = \left[\begin{array}{ccccc|ccccc} 0 & 0 & 0 & \cdots & 0 & 1 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \iddots & 1 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \iddots & 0 & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 1 & \iddots & 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & 1 \\ 1 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & 0 \end{array}\right], \qquad (6.9)$$

with $Y, Z \in \mathbb{C}^{(m+1)\times m}$. We analyze independently the three addends in the right-hand side of (6.8):

- $YZ^\top = Y \begin{bmatrix} I_m \\ 0_{1\times m} \end{bmatrix}^\top = \begin{bmatrix} Y & 0_{m\times 1} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \iddots & 1 & 0 \\ 0 & 0 & 0 & \iddots & 0 & 0 \\ 0 & 0 & 1 & \iddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}.$
- $ZY^\top = (YZ^\top)^\top = \begin{bmatrix} 0 & \cdots & 0 & 0 & 0 & 1 \\ 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & \cdots & 0 & 1 & 0 & 0 \\ \vdots & \iddots & \iddots & \iddots & \vdots & \vdots \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}.$
- $ZJ_m(0)Y^\top = \begin{bmatrix} I_m \\ 0_{1\times m} \end{bmatrix} J_m(0) \begin{bmatrix} 0 & \cdots & 0 & 0 & 0 & 1 \\ 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & \cdots & 0 & 1 & 0 & 0 \\ \vdots & \iddots & \iddots & \iddots & \vdots & \vdots \\ 0 & 1 & 0 & \cdots & 0 & 0 \end{bmatrix}$

$$= \begin{bmatrix} 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & \cdots & 0 & 1 & 0 & 0 \\ \vdots & \iddots & \iddots & \iddots & \vdots & \vdots \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

Adding all together we obtain:

$$ZJ_m(0)Y^\top + ZY^\top + YZ^\top = \begin{bmatrix} 0 & \cdots & 0 & 0 & 0 & 0 & 1 \\ 0 & \cdots & 0 & 0 & 1 & 1 & 0 \\ 0 & \cdots & 0 & 1 & 2 & 0 & 0 \\ \vdots & \iddots & \iddots & \iddots & \iddots & \vdots & \vdots \\ 0 & 1 & 2 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix},$$

which is a symmetric and nonsingular matrix, so we are done.

□

The next remark, which summarizes all results of this section, characterizes the consistency of (3.1) when A is a single Type-II block.

Remark 6.2. Given $m \geqslant 1$, regarding the consistency of

$$X^\top H_{2k}(\mu)X = I_m, \qquad \text{with } \mu \neq 0, (-1)^{k+1}, \tag{6.10}$$

we have the following:

(i) If $\mu \neq 0, \pm 1$, then (6.10) is consistent if and only if $k \geqslant m$.

(ii) If $\mu = -1$, then k is odd, and
  - (a) if $k = 1$, then (6.10) is not consistent;
  - (b) if $k \geqslant 3$ is odd, then (6.10) is consistent if and only if $k \geqslant m$.

(iii) If $\mu = 1$, then k is even, and
  - (a) if $k = 2$, then (6.10) is only consistent for $m = 1, 2$;
  - (b) if $k \geqslant 4$ is even, then (6.10) is consistent if and only if $k \geqslant m - 1$.

Let us check the results item by item. For item (i), the necessity is given in (3.2) of Theorem 2, and the sufficiency follows from Theorem 5, Lemma 2.3, and Lemma 2.4. Item (ii)(a) is part of Theorem 6. For item (ii)(b), the necessity is given in (3.2) of Theorem 2, and the sufficiency follows from Theorem 6, Lemma 2.3, and Lemma 2.4. Regarding (iii)(a): for $m = 1$, the equation $X^\top H_4(1)X = I_1$ has solution $X = \begin{bmatrix} 1 & 1/2 & 0 & 1 \end{bmatrix}^\top$; for $m = 2$, the equation $X^\top H_4(1)X = I_2$ has solution $X = \begin{bmatrix} 1 & 1/2 & 0 & 1 \\ i & -i/2 & 0 & i \end{bmatrix}^\top$; for $m = 3$, the inconsistency of $X^\top H_4(1)X = I_3$ is part of Theorem 7; and, for $m > 3$, the inconsistency is a consequence of the inconsistency for $m = 3$ together with Lemma 2.3. As for item (iii)(b), the necessity is given in (3.2) of Theorem 2, and the sufficiency follows from Theorem 7, Lemma 2.3, and Lemma 2.4.

7. A necessary and sufficient condition

In this section, we prove that (3.2) is a sufficient condition for Eq. (3.1) to be consistent, provided that the CFC of A does not contain blocks of either the form $H_4(1)$ or the form $H_2(-1)$. This is stated in the following result.

Theorem 8. *Let B be a complex symmetric matrix with* $\operatorname{rank} B = m$*, and let* $A \in \mathbb{C}^{n\times n}$ *be a matrix whose* CFC *has*

(i) *exactly d Type-0 blocks of the form $J_1(0)$,*

(ii) *exactly r Type-0 blocks with odd size greater than 1,*
(iii) *exactly s Type-I blocks with odd size,*
(iv) *exactly t Type-II blocks of the form $H_{4k}(1)$, with $k \geqslant 2$,*
(v) *no Type-II blocks of either the form $H_2(-1)$ or $H_4(1)$, and*
(vi) *an arbitrary number of Type-0, Type-I, and Type-II blocks with other sizes.*

Then $X^\top AX = B$ is consistent if and only if

$$n - d \geqslant 2m - r - s - 2t. \tag{7.1}$$

Proof. First, we may assume both A and B are in CFC. In particular, A has d blocks of type $J_1(0)$ and B is of the form $I_m \oplus J_1(0) \oplus \cdots \oplus J_1(0)$. By Lemma 2.2, we can get rid of the $J_1(0)$ blocks in both A and B, so we can assume that A has size $(n-d) \times (n-d)$ and $B = I_m$.

Now, in the conditions of the statement, we can write

$$\begin{array}{rcl} A & = & J_{2m_1-1}(0) \oplus \cdots \oplus J_{2m_r-1}(0) \bigoplus J_{2m_{r+1}}(0) \oplus \cdots \oplus J_{2m_{r+h}}(0) \bigoplus \\ & & \Gamma_{2\widehat{m}_1-1} \oplus \cdots \oplus \Gamma_{2\widehat{m}_s-1} \bigoplus \Gamma_{2\widehat{m}_{s+1}} \oplus \cdots \oplus \Gamma_{2\widehat{m}_{s+g}} \bigoplus \\ & & H_{2\breve{m}_1}(\mu_1) \oplus \cdots \oplus H_{2\breve{m}_u}(\mu_u) \bigoplus H_{4\breve{m}_{u+1}+2}(-1) \oplus \cdots \oplus H_{4\breve{m}_{u+v}+2}(-1) \bigoplus \\ & & H_{4\breve{m}_{u+v+1}}(1) \oplus \cdots \oplus H_{4\breve{m}_{u+v+t}}(1), \end{array}$$

where $m_i > 1$ for $i = 1, \ldots, r$; $\breve{m}_{u+v+k} > 1$ for $k = 1 \ldots, t$; and $\mu_j \neq 0, \pm 1$, for $j = 1, \ldots, u$.

Note that condition (7.1) is equivalent to

$$\begin{array}{rcl} n - d & = & (2m_1 - 1) + \cdots + (2m_r - 1) + 2m_{r+1} + \cdots + 2m_{r+h} + \\ & & (2\widehat{m}_1 - 1) + \cdots + (2\widehat{m}_s - 1) + 2\widehat{m}_{s+1} + \cdots + 2\widehat{m}_{s+g} + \\ & & 2\breve{m}_1 + \cdots + 2\breve{m}_u + (4\breve{m}_{u+1} + 2) + \cdots + (4\breve{m}_{u+v} + 2) + \\ & & 4\breve{m}_{u+v+1} + \cdots + 4\breve{m}_{u+v+t} \\ & \geqslant & 2m - r - s - 2t. \end{array}$$

If we set

$$\begin{array}{rcl} \widetilde{m} : & = & m_1 + \cdots + m_r + m_{r+1} + \cdots + m_{r+h} + \\ & & \widehat{m}_1 + \cdots + \widehat{m}_s + \widehat{m}_{s+1} + \cdots + \widehat{m}_{s+g} + \\ & & \breve{m}_1 + \cdots + \breve{m}_u + (2\breve{m}_{u+1} + 1) + \cdots + (2\breve{m}_{u+v} + 1) + \\ & & (2\breve{m}_{u+v+1} + 1) + \cdots + (2\breve{m}_{u+v+t} + 1), \end{array}$$

then condition (7.1) becomes $\widetilde{m} \geqslant m$. In turn, the statement of the theorem becomes:

$$X^\top AX = I_m \text{ is consistent if and only if } \widetilde{m} \geqslant m.$$

Theorem 2 proves that (7.1) is necessary, so $\widetilde{m} \geqslant m$ is necessary. Let us prove that $\widetilde{m} \geqslant m$ is sufficient as well. Suppose first that $\widetilde{m} = m$. According to Theorem 3, there exist $X_1, \ldots, X_r$ such that

$$X_i^\top J_{2m_i-1}(0) X_i = I_{m_i} \quad \text{for } i = 1, \ldots, r,$$

and there also exist $X_{r+1}, \ldots, X_{r+h}$ such that

$$X_j^\top J_{2m_j}(0) X_j = I_{m_j} \quad \text{for } j = r+1, \ldots, r+h.$$

According to Theorem 4, there exist $Y_1, \ldots, Y_s$ such that

$$Y_i^\top \Gamma_{2\widehat{m}_i-1} Y_i = I_{\widehat{m}_i} \quad \text{for } i = 1, \ldots, s,$$

and there also exist $Y_{s+1}, \dots, Y_{s+g}$ such that

$$Y_j^\top \Gamma_{2\widehat{m}_j} Y_j = I_{\widehat{m}_j} \quad \text{for } j = s+1, \dots, s+g.$$

According to Theorem 5, there exist $Z_1, \dots, Z_u$ such that

$$Z_i^\top H_{2\widetilde{m}_i}(\mu_i) Z_i = I_{\widetilde{m}_i} \quad \text{for } i = 1, \dots, u;$$

according to Theorem 6 there exist $Z_{u+1}, \dots, Z_{u+v}$ such that

$$Z_j^\top H_{4\widetilde{m}_j+2}(-1) Z_j = I_{2\widetilde{m}_j+1} \quad \text{for } j = u+1, \dots, u+v;$$

and according to Theorem 7 there exist $Z_{u+v+1}, \dots, Z_{u+v+t}$ such that

$$Z_j^\top H_{4\widetilde{m}_k}(-1) Z_j = I_{2\widetilde{m}_k+1} \quad \text{for } k = u+v+1, \dots, u+v+t.$$

Then $X^\top A X = I_m$ is consistent, with a solution given by

$$X = (X_1 \oplus \cdots \oplus X_{r+h}) \bigoplus (Y_1 \oplus \cdots \oplus Y_{s+g}) \bigoplus (Z_1 \oplus \cdots \oplus Z_{u+v+t}). \quad (7.2)$$

Suppose now that $\widetilde{m} > m$. As we have just proved, $X^\top A X = I_{\widetilde{m}}$ is consistent (with a solution X given in (7.2)). Then, the consistency of $X^\top A X = I_m$ follows from Lemma 2.3 and Lemma 2.4. So $\widetilde{m} \geqslant m$ is sufficient. □

Remark 7.1. In the proof of Theorem 8, we show that Eq. (1.1), with B symmetric and A in the conditions of the statement, is consistent if and only if, after taking A and B to their CFC, the equation is consistent block-wise, where the (three) blocks of A are the direct sum of all Type-0, Type-I, and Type-II blocks in the CFC, respectively (and those of B are identities of the appropriate size).

After Theorem 8, it is natural to ask whether or not allowing the blocks of the form $H_2(-1)$ and $H_4(1)$ to appear in the CFC of A one could get a characterization like the one in this theorem for the consistency of Eq. (1.1) with B symmetric. We have seen in Example 1 and in Theorem 7 that $X^\top H_2(-1) X = I_1$ and $X^\top H_4(1) X = I_3$, respectively, are not consistent, so their inclusion in the CFC of A has an unknown effect on the increase of the value of m (the rank/size of B) in the bound (7.1). Following Remark 7.1, we should look for a sufficient condition like (3.2) that allows one to construct a block-wise solution, looking only at the direct sum of blocks of the same type. However, it may happen that the combined effect of blocks of either the form $H_2(-1)$ or $H_4(1)$, together with some other blocks in the CFC of A, make Eq. (1.1) to be consistent, even if the equation is not consistent block-wise. This is indeed what happens, as the following examples show.

Example 2. Let

$$A = H_2(-1) \oplus J_2(0) = \left[\begin{array}{cc|cc} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right], \qquad B = I_2 = \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & 1 \end{array}\right].$$

With these A and B, Eq. (1.1) is consistent. A particular solution is

$$X = \begin{bmatrix} \mathrm{i} & 0 \\ 0 & 1 \\ 1 & \mathrm{i} \\ 1 & -\mathrm{i} \end{bmatrix}.$$

However, this equation is not consistent block-wise, since $X^\top H_2(-1)X = I_1$ is not consistent (see Example 1).

Example 3. Let

$$A = H_4(1) \oplus J_2(0) = \left[\begin{array}{cccc|cc} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right], \qquad B = I_4 = \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right].$$

With these A and B, Eq. (1.1) is consistent, and a particular solution is

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

Again, the equation is not consistent block-wise, since $X^\top H_4(1)X = I_3$ is not consistent (see Theorem 7).

8. Conclusions and open problems

We have provided necessary and sufficient conditions for the matrix equation (1.1) to be consistent when B is symmetric. We have also extended the well-known characterization of congruence for two symmetric square matrices of the same size, to matrices not necessarily having the same size. The characterization for the case when B is symmetric depends on the CFC of A and B (in particular, $\mathrm{CFC}(B) = I_m$). However, this characterization does not include the case where $\mathrm{CFC}(A)$ contains Type-II blocks of either the form $H_4(1)$ or $H_2(-1)$.

As a continuation of this work, the following lines of research arise:

- To obtain necessary and sufficient conditions, like in Theorem 8, for Eq. (1.1) to be consistent in the case where B is skew-symmetric.
- To get a characterization for Eq. (1.1) to be consistent, with B symmetric and A arbitrary (namely, including the case where its CFC has blocks of the form $H_4(1)$ and $H_2(-1)$).
- To obtain necessary and sufficient conditions for Eq. (1.1) to be consistent when $\mathrm{CFC}(A)$ contains only one block, and B is arbitrary.

- (Quite hard) To get necessary and sufficient conditions for Eq. (1.1) to be consistent, with A and B arbitrary.

Another natural project that can be addressed consist in analyzing the consistency of Eq. (1.1) over the real field (instead of the complex one). That would require to use an appropriate canonical form for congruence over $\mathbb{R}$. Finally, it also remains as an open problem to address the consistency of the related equation $X^*AX = B$, where $(\cdot)^*$ denotes complex conjugation, over the complex field.

Acknowledgments: This work has been partially supported by the *Ministerio de Economía y Competitividad* of Spain through grant MTM2015-65798-P (F. De Terán), and by the *Ministerio de Ciencia, Innovación y Universidades* of Spain through grant MTM2017–90682–REDT.

References

[1] P. Benner, D. Palitta. On the solution of the non-symmetric T-Riccati equation. Preprint. Available at: `https://arxiv.org/pdf/2003.03693.pdf`

[2] P. G. Buckhiester. Rank r solutions to the matrix equation $XAX^t = C$, A alternate, over GF(2^y). *Trans. Amer. Math. Soc.*, 189 (1974) 201–209.

[3] P. G. Buckhiester. Rank r solutions to the matrix equation $XAX^t = C$, A non-alternate, C alternate, over GF(2^y). *Canad. J. Math.*, 26 (1974) 78–90.

[4] P. G. Buckhiester. The number of solutions to the matrix equation $XAX' = C$, A and C nonalternate and of full rank, over GF(2^y). *Math. Nachr.*, 63 (1974) 37–41.

[5] L. Carlitz. Representations by skew forms in a finite field. *Arch. Math.*, V (1954) 19–31.

[6] F. De Terán, F. M. Dopico. The solution of the equation $XA + AX^T = 0$ and its application to the theory of orbits. *Linear Algebra Appl.*, 434 (2011) 44–67.

[7] F. De Terán, F. M. Dopico. Consistency and efficient solution for the Sylvester equation for $*$-congruence: $AX + X^*B = C$. *Electron. J. Linear Algebra*, 22 (2011) 849–863.

[8] F. De Terán, F. M. Dopico, N. Guillery, D. Montealegre, N. Z. Reyes. The solution of the equation $AX + BX^\star = 0$. *Linear Algebra Appl.*, 438 (2010) 2817–2860.

[9] F. De Terán, B. Iannazzo. Uniqueness of solution of a generalized $*$-Sylvester matrix equation. *Linear Algebra Appl.*, 493 (2016) 323–335.

[10] F. De Terán, B. Iannazo, F. Poloni, L. Robol. Solvability and uniqueness criteria for generalized Sylvester-type equations. *Linear Algebra Appl.*, 542 (2018) 501–521.

[11] F. De Terán, B. Iannazo, F. Poloni, L. Robol. Nonsingular systems of generalized Sylvester equations: an algorithmic approach. *Numer. Lin. Algebra Appl.*, 26 (2019) e2261 (29 pages).

[12] J. Ding, N. H. Rhee. Spectral solutions of the Yang-Baxter matrix equation. *J. Math. Anal. Appl.*, 402 (2013) 567–573.

[13] C. Estatico, F. Di Benedetto. Shift-invariant approximations of structured shift-variant blurring matrices. *Numer. Algor.*, 62 (2013) 615–635.

[14] J. D. Fulton. Generalized inverses of matrices over fields of characteristic two. *Linear Algebra Appl.*, 28 (1979) 69–76.

[15] J. H. Hodges. A skew matrix equation over a finite field. *Math. Nachr.*, 17 (1966) 49–55.

[16] R. A. Horn, C. R. Johnson. *Matrix Analysis.* Cambridge University Press, 2nd Ed. New York, 2013.

[17] R. A. Horn, V. V. Sergeichuk. Canonical forms for complex matrix congruence and $*$-congruence. *Linear Algebra Appl.*, 416 (2006) 1010–1032.

[18] D. Kressner, X. Liu. Structured canonical forms for products of (skew-) symmetric matrices and the matrix equation $XAX = B$. *Electron. J. Linear Algebra*, 26 (2013) 215–230.

[19] H. W. Turnbull, A. C. Aitken. *An Introduction to the Theory of Canonical Matrices.* Dover, New York, 1961. (First published by: Blackie & Son Limited, 1932).

[20] H. Wei, Y. Zhang. The number of solutions to the alternate matrix equation over a finite field and a q-identity. *J. Stat. Plan. Infer.*, 94 (2001) 349–358.

Alberto Borobia
UNED
Facultad de Ciencias
C/ Senda del Rey 9
28040 Madrid
e-mail: `aborobia@mat.uned.es`

Roberto Canogar
UNED
Facultad de Ciencias
C/ Senda del Rey 9
28040 Madrid
e-mail: `rcanogar@mat.uned.es`

Fernando De Terán
Universidad Carlos III de Madrid
Departamento de Matemáticas
Avda. Universidad s/n
28911 Leganés
e-mail: `fteran@math.uc3m.es`