

The hyperbolic Schur decomposition (extended)

Šego, Vedran

2012

MIMS EPrint: **2012.116**

Manchester Institute for Mathematical Sciences
School of Mathematics

The University of Manchester

Reports available from: <http://eprints.maths.manchester.ac.uk/>

And by contacting: The MIMS Secretary
School of Mathematics
The University of Manchester
Manchester, M13 9PL, UK

ISSN 1749-9097

The hyperbolic Schur decomposition (extended)[☆]

Vedran Šego^{a,b}

^a*Faculty of Science, University of Zagreb, Croatia*

^b*School of Mathematics, The University of Manchester, Manchester, M13 9PL, UK*

Abstract

We propose a hyperbolic counterpart of the Schur decomposition, with the emphasis on the preservation of structures related to some given hyperbolic scalar product. We give results regarding the existence of such a decomposition and research the properties of its block triangular factor for various structured matrices.

Keywords: indefinite scalar products, hyperbolic scalar products, Schur decomposition, Jordan decomposition, quasitriangularization, quasidiagonalization, structured matrices

2000 MSC: 15A63, 46C20, 65F25

This is the extended version of the submitted paper. The added parts of the text have the light blue background.

1. Introduction

The Schur decomposition $A = UTU^*$, sometimes also called Schur's unitary triangularization, is a unitary similarity between any given square matrix $A \in \mathbb{C}^{n \times n}$ and some upper triangular matrix $T \in \mathbb{C}^{n \times n}$. Such a decomposition has a structured form for various structured matrices, i.e., T is diagonal if and only if A is normal, real diagonal if and only if A is Hermitian, positive (nonnegative) real diagonal if and only if A is positive (semi)definite and so on.

Furthermore, the Schur decomposition can be computed in a numerically stable way, making it a good choice for calculating the eigenvalues of A (which are the diagonal elements of T) as well as the various matrix functions (for more details, see [11]). Its structure preserving property allows to save time and memory when working with structured matrices. For example, computing the value of some function of a Hermitian matrix is reduced to working with a diagonal matrix, which involves only evaluation of the diagonal elements.

[☆]Version of December, 2012.

Email address: vsego@math.hr (Vedran Šego)

Unitary matrices are very useful when working with the traditional Euclidean scalar product $\langle x, y \rangle = y^* x$, as their columns form an orthonormal basis of \mathbb{C}^n . However, many applications require a nonstandard scalar product which is usually defined by $[x, y]_J = y^* J x$, where J is some nonsingular matrix, and many of these applications consider Hermitian or skew-Hermitian J . The hyperbolic scalar product defined by a signature matrix $J = \text{diag}(j_1, \dots, j_n)$ ($j_k \in \{-1, 1\}$) arises frequently in applications. It is used, for example, in the theory of relativity and in the research of the polarized light. More on the applications of such products can be found in [10, 13, 14, 17].

The Euclidean matrix decompositions have some nice structure preserving properties even in nonstandard scalar products, as shown by Mackey, Mackey and Tisseur [16], but it is often worth looking into versions of such decompositions that respect the structures related with the given scalar product. There is plenty of research on the subject, i.e., hyperbolic SVD [17, 23], $J_1 J_2$ -SVD [9], two-sided hyperbolic SVD [20], hyperbolic CS decomposition [8, 10] and indefinite QR factorization [19].

There are many advantages of using decompositions related to some specific, nonstandard scalar product, as such decompositions preserve structures related to a given scalar product. They can simplify calculation and provide a better insight into the structures of such structured matrices.

In this paper we investigate the existence of a decomposition which would resemble the traditional Schur decomposition, but with respect to the given hyperbolic scalar product. In other words, our similarity matrix should be unitary-like (orthonormal, to be more precise) with respect to that scalar product.

As we shall see, a hyperbolic Schur decomposition can be constructed, but not for all square matrices. Furthermore, we will have to relax conditions on both U and T . The matrix U will be hyperexchange (a column-permutation of the matrix unitary with respect to J). The matrix T will have to be block upper triangular with diagonal blocks of order 1 and 2. Both of these changes are quite usual in hyperbolic scalar products. For example, they appear in the traditional QR vs. the hyperbolic QR factorizations [19].

Some work on the hyperbolic Schur decomposition was done by Ammar, Mehl and Mehrmann [1, Theorem 8], but with somewhat different focus. They have assumed to have a partitioned $J = I_p \oplus (-I_q)$, in the paper denoted as $\Sigma_{p,q}$, for which they have observed a Schur-like similarity through unitary factors (without permuting J), producing more complex triangular factors. Also, their decomposition is applicable only to the set of J -unitary matrices, in the paper denoted as the Lie group $\mathcal{O}_{p,q}$.

In the symplectic scalar product spaces, Schur-like decomposition was researched by Lin, Mehrmann and Xu [15], by Ammar, Mehl and Mehrmann [1], and by Xu [21, 22].

In section 2, we provide a brief overview of the definitions, properties and other results relating to the hyperbolic scalar products that will be used later. In section 3, the definition and the construction of the hyperbolic Schur decomposition are presented. We also provide sufficient requirements for its existence

and examples showing why such a decomposition does not exist for all matrices. In section 4 we observe various properties of the proposed decomposition. We finalize the results by providing the necessary and the sufficient conditions for the existence of the hyperbolic Schur decomposition of J -Hermitian matrices in section 5.

The notation used is fairly standard. The capital letters refer to matrices and their blocks, elements are denoted by the appropriate lowercase letter with two subscript indices, while lowercase letters with a single subscript index represent vectors (including matrix columns). By $J = \text{diag}(\pm 1)$ we denote a diagonal signature matrix defining the hyperbolic scalar product, while P and P_k (for some indices k) denote permutation matrices. We use

$$\mathcal{S}_n := [\delta_{i,n+1-j}] = \begin{bmatrix} & & 1 \\ & \dots & \\ 1 & & \end{bmatrix} = \mathcal{S}_n^{-1}$$

for the standard involutory permutation (see [6, Example 2.1.1.]), \mathcal{J} for a Jordan matrix and $\mathcal{J}_k(\lambda)$ for a single Jordan block of order k associated with the eigenvalue λ . Vector e_k denotes k^{th} column of the identity matrix and \otimes denotes the Kronecker product. Symbol \oplus is used to describe a diagonal concatenation of matrices, i.e., $A \oplus B$ is a block diagonal matrix with the diagonal blocks A and B .

Also a standard notation, but somewhat incorrect in terms of indefinite scalar products, is $|v| := \sqrt{|[v, v]|}$. This is the “norm” of vector v induced by the scalar product $[\cdot, \cdot]$. This is not a norm *per se*, as there may exist nonzero vectors v such that $[v, v] = 0$. However, we use this norm only on the vectors for which it is well defined and has the usual norm properties, which makes the notation correct.

2. The hyperbolic scalar products

As mentioned in the introduction, an *indefinite scalar product* is defined by a nonsingular Hermitian indefinite matrix $J \in \mathbb{C}^{n \times n}$ as $[x, y]_J = y^* J x$. When J is known from the context, we simply write $[x, y]$ instead of $[x, y]_J$.

When J is a signature matrix, i.e., $J = \text{diag}(\pm 1) := \text{diag}(j_{11}, j_{22}, \dots, j_{nn})$, where $j_{kk} \in \{-1, 1\}$ for all k , the scalar product is referred to as *hyperbolic* and takes the form

$$[x, y]_J = y^* J x = \sum_{i=1}^n j_{ii} x_i \bar{y}_i.$$

Throughout this paper we assume that all considered scalar products are hyperbolic, unless stated otherwise.

Indefinite scalar products have another important property which, unfortunately, causes a major problem with the construction of the decomposition. A

vector $v \neq 0$ is said to be *J-degenerate* if $[v, v] = 0$; otherwise, we say that it is *J-nondegenerate*. Degenerate vectors are sometimes also called *J-neutral*.

If $[v, v] < 0$ for some vector v , we say that v is *J-negative*, while we call it *J-positive* if $[v, v] > 0$. When J is known from the context, we simply say that the vector is *degenerate*, *nondegenerate*, *neutral*, *negative* or *positive*.

We extend this notion to matrices as well: a matrix A is *J-degenerate* if $\text{rank } A^*JA < \text{rank } A$. Otherwise, we say that A is *J-nondegenerate*. Again, if J is known from the context, we simply say that A is *degenerate* or *nondegenerate*.

We say that the vector v is *J-normalized*, or just *normalized* when J is known from the context, if $|[v, v]| = 1$. As in the Euclidean scalar product, if a vector v is given, then the vector

$$v' = \frac{1}{|v|}v = \frac{1}{\sqrt{|[v, v]|}}v \quad (1)$$

is a normalization of v . Note that degenerate vectors cannot be normalized. Also, for a given vector $x \in \mathbb{C}^n$, $\text{sign}[\xi x, \xi x]$ is constant for all $\xi \in \mathbb{C} \setminus \{0\}$. This means that the normalization (1) does not change the sign of the scalar product, i.e.,

$$\text{sign}[v, v] = \text{sign}[v', v'] = [v', v'].$$

Like in the Euclidean scalar products, we define the *J-conjugate transpose* (or *J-adjoint*) of A with respect to a hyperbolic J , denoted as $A^{[*]J}$, as $[Ax, y]_J = [x, A^{[*]J}y]_J$ for all vectors $x, y \in \mathbb{C}^n$. It is easy to see that $A^{[*]J} = JA^*J$. Again, if J is known from the context, we simply write $A^{[*]}$.

The usual structured matrices are defined naturally. A matrix H is called *J-Hermitian* (or *J-selfadjoint*) if $H^{[*]} = H$, i.e., if JH is Hermitian. A matrix U is said to be *J-unitary* if $U^{[*]} = U^{-1}$, i.e., if $U^*JU = J$.

Like their traditional counterparts, *J-unitary* matrices are orthonormal with respect to $[\cdot, \cdot]_J$. However, unlike in the Euclidean scalar product, in hyperbolic scalar products we have a wider class of matrices orthonormal with respect to J (which are not necessarily unitary with respect to the same scalar product), called *J-hyperexchange* matrices.

We say that U is *J-hyperexchange* if $U^*JU = P^*JP$ for some permutation P . Although the term “hyperexchange” is quite common, we often refer to such matrices as *J-orthonormal*, to emphasize that their columns are *J-orthonormal* vectors. In other words, if the columns of U are denoted as u_i , then

$$[u_i, u_i] = \pm 1, \quad [u_i, u_j] = 0, \quad \text{for all } i, j.$$

More on the definitions and properties related to hyperbolic (and, more generally, indefinite) scalar products can be found in [6].

Throughout the paper, we often consider the diagonal blocks of a given matrix A . In order to keep the relation with the given hyperbolic scalar product induced by some $J = \text{diag}(\pm 1)$, we introduce the term *the corresponding part of J*. Let $J = \text{diag}(j_1, \dots, j_n)$, $j_k \in \{-1, 1\}$, define a hyperbolic scalar product

and let $A = [A_{ij}] \in \mathbb{C}^{n \times n}$ be a blockmatrix with elements a_{pq} partitioned in the blocks A_{ij} in a way that each diagonal block A_{kk} is of order 1 or 2. Observing the block A_{kk} for a given k , the *corresponding part of J* , here denoted as J' , is defined as $J' = [j_p]$ if $A_{kk} = [a_{pp}]$ is of order 1 and as $J' = \text{diag}(j_p, j_{p+1})$ if

$$A_{kk} = \begin{bmatrix} a_{pp} & a_{p,p+1} \\ a_{p+1,p} & a_{p+1,p+1} \end{bmatrix}$$

is of order 2.

3. Definition and existence of the hyperbolic Schur decomposition

In this section we present the definition and the main results regarding the hyperbolic Schur decomposition of a matrix $A \in \mathbb{C}^{n \times n}$.

Like other hyperbolic generalizations of the Euclidean decompositions, this one also has a hyperexchange matrix instead of a unitary one, as well as the block structured factor instead of a triangular/diagonal factor of the Schur decomposition. Block upper triangular matrices with the diagonal blocks of order 1 and 2 are usually referred to as *quasitriangular* (see [18, Section 3.3]). Similarly, if T is block diagonal with the diagonal blocks of order 1 and 2, we refer to it as *quasidiagonal*.

Before providing a formal definition, let us first show obstacles which will explain why these changes are necessary.

The classical proof of the Schur decomposition is constructive and can be found in many books, for example in [12, Theorem 2.3.1] or [7, Theorem 7.1.3]. The basic idea is an iterative reduction. If u_1 is a normalized eigenvector of $A \in \mathbb{C}^{n \times n}$ with the corresponding eigenvalue λ , then we can construct an orthonormal basis $\{u_1, u_2, \dots, u_n\}$ of \mathbb{C}^n and let $U = [u_1 \dots u_n]$. Then $AU = UT$, where

$$T = \left[\begin{array}{c|c} \lambda & * \\ \hline 0 & T' \end{array} \right].$$

We then proceed with the same reduction on T' , until it reaches the order 1, and multiply all the constructed unitary matrices (appropriately embedded in I_n). Since this product is again unitary, we obtain the Schur decomposition of A .

In the hyperbolic scalar product spaces, the situation is more complicated, because some matrices – even some diagonalizable ones! – have only degenerate eigenvectors. This is important because such vectors cannot be normalized and the above process of triangularization fails.

Example 3.1. *The eigenvectors of the diagonalizable matrix $A = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}$ and the nondiagonalizable matrix $B = \begin{bmatrix} 3 & -1 \\ 1 & 1 \end{bmatrix}$ are all degenerate with respect to the*

scalar product defined by $J = \text{diag}(1, -1)$.

An example can be constructed for a more general case. Let $n = 2k$ for some $k \in \mathbb{N}$ and let $J = \mathbf{I}_k \oplus (-\mathbf{I}_k)$. Define $C = S\mathcal{J}S^{-1}$, where \mathcal{J} is any Jordan matrix and $S = [s_{ij}]$ is such that

$$s_{ij} = \begin{cases} 1, & i \leq j, \\ -1, & i > j. \end{cases}$$

It is easy to see that the columns of S , consisting only of the elements -1 and 1 , are all degenerate (although, the matrix S itself is not!).

The previous example may seem a bit discouraging, but it is worth noting that, for any J , if all eigenvectors of a diagonalizable matrix A are degenerate, then not all of them can be mutually orthogonal as well.

Indeed, let $A = S\mathcal{J}S^{-1}$ be a Jordan decomposition of A . Since S and J are nonsingular, S^*JS is nonsingular as well. The (j, i) th element of S^*JS is $s_j^*Js_i = [s_i, s_j]$, where $s_k = Se_k$. If all columns of S are degenerate, then all diagonal elements of S^*JS are zero and, since S^*JS is nonsingular, for each i there is a $j \neq i$ such that $[s_i, s_j] \neq 0$.

How does this help? If all eigenvectors are degenerate, i.e., not normalizable, we can pick two which are not mutually orthogonal. Using their linear combinations (we need two linearly independent ones) as the basis in one step of a Schur decomposition leads to a block triangular matrix

$$T = \left[\begin{array}{c|c} T'_1 & * \\ \hline 0 & T'_2 \end{array} \right],$$

where the block T'_1 is of order 2. Therefore, the triangular part of the hyperbolic Schur decomposition actually has to be quasitriangular, so that the hyperbolic Schur decomposition is applicable to a much bigger set of matrices. This is not unlike the real Euclidean case: some real matrices are not real orthogonally (although they are complex unitary triangularizable), but they are all real orthogonally quasitriangularizable (see [7, Section 7.4.1]).

For better clarity, we say that a block X_{jj} of a matrix X is *nontrivial* if it cannot be split into smaller blocks without losing the block structure of the matrix. If $X = \bigoplus_{i=1}^k X_{ii}$ is block diagonal, we say that X_{jj} is nontrivial if it cannot be split as $X_{jj} = Y_1 \oplus Y_2$ for some square blocks Y_1 and Y_2 . For a block triangular X , we say that X_{jj} is nontrivial if it cannot be split as $X_{jj} = \begin{bmatrix} Y_{11} & Y_{12} \\ & Y_{22} \end{bmatrix}$, where Y_{11} and Y_{22} are square blocks. Since we mostly deal with blocks of order 2, the nontrivial blocks will be those of order 1 and those of order 2 with one of the nondiagonal elements (in case X is block diagonal) or the bottom left element (in case X is a block triangular) being nonzero.

The described notation of nontrivial blocks is just a descriptive way of saying

that we cannot split X into smaller blocks while preserving its block diagonal or block triangular structure. It is used to simplify some of the statements and proofs by reducing the number of observed cases without losing generality.

Example 3.2 (Nontrivial blocks). *Let*

$$X = \begin{bmatrix} 1 & 2 & 3 & 4 \\ & 5 & 6 & 7 \\ & & 8 & 9 \\ & & 10 & 11 \end{bmatrix}.$$

This is a quasitriangular matrix, with the diagonal blocks

$$X_{11} = [1], \quad X_{22} = [5], \quad X_{33} = \begin{bmatrix} 8 & 9 \\ 10 & 11 \end{bmatrix}.$$

Here, X_{11} , X_{22} and X_{33} are considered nontrivial, while the block

$$\begin{bmatrix} 1 & 2 \\ & 5 \end{bmatrix}$$

is not, even though it technically is a (triangular) block of order 2. Similarly, for a quasidiagonal matrix

$$Y = \begin{bmatrix} 1 & 2 & & \\ & 3 & & \\ & & 4 & \\ & & & 5 \end{bmatrix},$$

blocks

$$Y_{11} = \begin{bmatrix} 1 & 2 \\ & 3 \end{bmatrix}, \quad Y_{22} = [4], \quad Y_{33} = [5]$$

are considered nontrivial, while the block

$$\begin{bmatrix} 4 & \\ & 5 \end{bmatrix}$$

is not.

Furthermore, if Y is observed as a quasitriangular matrix, then Y_{11} is also not nontrivial. Instead, all nontrivial blocks are of order 1 (defined by the diagonal elements 1, 3, 4 and 5).

Referring to the blocks as nontrivial allows us to skip (in proofs) various cases which are already covered but this may not be clear if left completely unmentioned. For example, when the proof related to quasidiagonal matrices covers the blocks of order 1, it implicitly also covers the blocks of order 2 of the form $\text{diag}(\xi_1, \xi_2)$. So, when we refer to the block of order 2 as being nontrivial, we avoid the trivial (and already covered) case of a diagonal block of order 2 and we are free to assume that the offdiagonal elements in such blocks are nonzero, without further cluttering the proof with the explanations why we are allowed to do so. The same goes for quasitriangular matrices, except that in this case the nontriviality property means that only the subdiagonal element of a nontrivial block of order 2 is considered to be nonzero.

Unfortunately, allowing the triangular factor to have blocks is not enough. We also need permutations of the decomposing matrix and the scalar product generator J . The following example shows why this is needed, simultaneously illustrating a general approach to proving that some matrices do not have a hyperbolic Schur decomposition.

Example 3.3. Let $J = \text{diag}(1, 1, -1, -1)$ and let $A = SJS^{-1}$, where $J = \text{diag}(1, 2, 3, 4)$ and

$$S = \begin{bmatrix} \frac{1}{5\sqrt{3}} & \frac{1}{17\sqrt{15}} & \frac{1}{195\sqrt{7}} & \frac{1}{257\sqrt{255}} \\ \frac{2}{5\sqrt{3}} & \frac{4}{17\sqrt{15}} & \frac{8}{195\sqrt{7}} & \frac{16}{257\sqrt{255}} \\ \frac{4}{5\sqrt{3}} & \frac{16}{17\sqrt{15}} & \frac{64}{195\sqrt{7}} & \frac{256}{257\sqrt{255}} \\ \frac{8}{5\sqrt{3}} & \frac{64}{17\sqrt{15}} & \frac{512}{195\sqrt{7}} & \frac{4096}{257\sqrt{255}} \end{bmatrix}.$$

We shall now show that no J -unitary V and quasitriangular T exist such that $A = VTV^{-1}$. Since

$$S^*JS = \begin{bmatrix} -1 & -0.994393 & -0.970230 & -0.949852 \\ -0.994393 & -1 & -0.993828 & -0.985075 \\ -0.970230 & -0.993828 & -1 & -0.998151 \\ -0.949852 & -0.985075 & -0.998151 & -1 \end{bmatrix} \quad (2)$$

to 6 significant digits, all eigenvectors of A (i.e., columns of S) are normalized negative vectors. Furthermore, if we denote them by s_i , then $-1 < [s_i, s_j] < 0$ for all $i \neq j$.

Let us now assume that there exist a J -unitary V and a quasitriangular T such that $A = VTV^{-1}$. We distinguish the following possibilities:

1. The first block of T is of order 1. Then v_1 is obviously a J -normalized eigenvector of A . But, this is impossible, since V is J -unitary, meaning that $V^*JV = J$, so $[v_1, v_1] = 1$, while $[s_i, s_i] < 0$ for all i and, since normalization does not change the sign of the vector's scalar product by itself, $[v_1, v_1] < 0$.

2. The first block of T is a nontrivial block of order 2 (i.e., $t_{21} \neq 0$). Then it is easy to see that

$$Av_1 = t_{11}v_1 + t_{21}v_2, \quad Av_2 = t_{12}v_1 + t_{22}v_2.$$

In other words,

$$(A - t_{11}\mathbf{I})v_1 = t_{21}v_2, \quad (A - t_{22}\mathbf{I})v_2 = t_{12}v_1.$$

Multiplying the second equality with t_{21} and substituting $t_{21}v_2$ with the expression from the first equality, we get:

$$t_{12}t_{21}v_1 = (A - t_{22}\mathbf{I})t_{21}v_2 = (A - t_{22}\mathbf{I})(A - t_{11}\mathbf{I})v_1.$$

In other words, v_1 is the eigenvector of $(A - t_{22}\mathbf{I})(A - t_{11}\mathbf{I})$. Using the same argument, we see that v_2 is the eigenvector of $(A - t_{11}\mathbf{I})(A - t_{22}\mathbf{I})$. Furthermore,

$$(A - t_{22}\mathbf{I})(A - t_{11}\mathbf{I})v_1 = A^2v_1 - (t_{11} + t_{22})Av_1 + t_{11}t_{22}v_1 = t_{12}t_{21}v_1.$$

Now, we have:

$$\begin{aligned} 0 &= A^2v_1 - (t_{11} + t_{22})Av_1 + (t_{11}t_{22} - t_{12}t_{21})v_1, \\ &= \left(A - \frac{t_{11} + t_{22}}{2}\mathbf{I} \right)^2 v_1 - \left(\frac{(t_{11} - t_{22})^2}{4} + t_{12}t_{21} \right) v_1. \end{aligned}$$

So, v_1 is the eigenvector of

$$A_2 := \left(A - \frac{t_{11} + t_{22}}{2}\mathbf{I} \right)^2, \quad (3)$$

i.e.,

$$A_2v_1 = \lambda'v_1, \quad \lambda' := \frac{(t_{11} - t_{22})^2}{4} + t_{12}t_{21}. \quad (4)$$

But, since

$$(A - t_{22}\mathbf{I})(A - t_{11}\mathbf{I}) = (A - t_{11}\mathbf{I})(A - t_{22}\mathbf{I}),$$

v_1 and v_2 are both (linearly independent) eigenvectors of A_2 , with the same eigenvalue λ' .

Since A and A_2 are diagonalizable, every eigenvector of A is also an eigenvector of A_2 . Moreover, since the eigenvalues of A are distinct (they are 1, 2, 3 and 4), A_2 has at most one eigenvalue of multiplicity 2 (this is easily seen from its Jordan decomposition). In other words, its eigenspaces have dimensions at most 2, so v_1 and v_2 are linear combinations of s_i and s_j for some $i \neq j$.

Let $v_1 = \alpha_1s_i + \beta_1s_j$ and $v_2 = \alpha_2s_i + \beta_2s_j$. Obviously, $\alpha_1, \beta_1 \neq 0$ (or v_1 would be an eigenvector of A , hence covered by the case 1). Then, from $J = V^*JV$, we get:

$$\begin{aligned} 1 = j_{11} &= [v_1, v_1] = |\alpha_1|^2[s_i, s_i] + |\beta_1|^2[s_j, s_j] + 2\operatorname{Re}(\alpha_1\overline{\beta_1}[s_i, s_j]) \\ &= -(|\alpha_1|^2 + |\beta_1|^2 - 2[s_i, s_j]\operatorname{Re}(\alpha_1\overline{\beta_1})). \end{aligned} \quad (5)$$

From (2) we see that for $i \neq k$, $-1 < [s_i, s_k] < 0$. Using (5) and the well known fact that $|\operatorname{Re}(z)| \leq |z|$, i.e., $\operatorname{Re}(z) \geq -|z|$, for all $z \in \mathbb{C}$, we see that

$$\begin{aligned} -1 &= |\alpha_1|^2 + |\beta_1|^2 - 2[s_i, s_j] \operatorname{Re}(\alpha_1 \overline{\beta_1}) \\ &= |\alpha_1|^2 + |\beta_1|^2 + 2|[s_i, s_j]| \operatorname{Re}(\alpha_1 \overline{\beta_1}) \\ &\geq |\alpha_1|^2 + |\beta_1|^2 - 2|[s_i, s_j]| \cdot |\alpha_1| \cdot |\beta_1| \\ &> |\alpha_1|^2 + |\beta_1|^2 - 2|\alpha_1| \cdot |\beta_1| = (|\alpha_1| - |\beta_1|)^2 \geq 0, \end{aligned}$$

which is an obvious contradiction.

Since both possible cases have led to a contradiction, the described decomposition does not exist for the pair (A, J) .

Now, the valid question is what kind of permutations should be used on the similarity matrix (permutations of rows, columns or both). To answer this question, we note that the point of having the Schur decomposition is having a convenient structure preserving similarity triangulation. Notably, in the Euclidean case, A is Hermitian if and only if its triangular factor T from the Schur decomposition is diagonal (i.e., triangular Hermitian). Here, we would like to keep as much of that property as is possible.

So, let us assume that $J = \operatorname{diag}(\pm 1)$ is given and let $A = A^{[*]}$ be a given J -Hermitian matrix. Furthermore, let

$$A = (P_1 U P_2) T (P_1 U P_2)^{-1}, \quad (6)$$

such that U is J_1 -unitary and T is J_2 -Hermitian for some J_1 and J_2 which are permutationally similar to J , i.e., they have the same inertia as J , which is the usual approach when dealing with the indefinite scalar products.

The inertia property comes from another structure related condition. Having a unitary-like matrix U in a space equipped with the scalar product induced by some J , we want $U^* J U = \tilde{J}$. By Sylvester's law of inertia, J and \tilde{J} must have the same inertia. This transition is sometimes used to transfer a problem from a general indefinite scalar product space to a hyperbolic one (see [6, Section 4.5]).

How are J , P_1 , P_2 , U , J_1 and J_2 related to each other? One should note that (6) is a unitary-like similarity (with respect to J) of a given matrix with T , which we want to be quasitriangular, as explained before. Therefore, it is a fitting general form for the hyperbolic Schur decomposition.

Now, to explore the connection between J , P_1 , P_2 , U , J_1 and J_2 , we shall use the J -Hermiticity of A :

$$\begin{aligned} (P_1 U P_2) T P_2^* U^{-1} P_1^* &= A = A^{[*]J} = J((P_1 U P_2) T (P_1 U P_2)^{-1})^* J \\ &= J P_1 U^{-*} P_2 T^* P_2^* U^* P_1^* J = J P_1 U^{-*} P_2 J_2 T^{[*]J_2} J_2 P_2^* U^* P_1^* J. \end{aligned}$$

Since we want the decomposition to retain the Hermiticity property, i.e., $T =$

$T^{[*]J_2}$, we conclude that it is fitting that the similarity matrices are equal, i.e.,

$$P_1UP_2 = JP_1U^{-*}P_2J_2.$$

From here,

$$U^* = P_2J_2P_2^*U^{-1}P_1^*JP_1.$$

Since we want U to be J_1 -unitary,

$$J_1 = U^*J_1U = P_2J_2P_2^*U^{-1}P_1^*JP_1J_1U.$$

We see it is fitting that $J_1 = P_1^*JP_1 = P_2J_2P_2^*$, i.e.,

$$J_1 = P_1^*JP_1, \quad J_2 = P_2^*J_1P_2.$$

So, it makes sense to define the hyperbolic Schur decomposition as follows:

$$A = (P_1UP_2)T(P_1UP_2)^{-1}, \quad U^*\tilde{J}U = \tilde{J}, \quad \tilde{J} = P_1^*JP_1. \quad (7)$$

Note that none of this is a proof of anything (including the property of preserving structures for such decompositions), but merely a construction. The proofs follow soon.

We can reduce the two permutations from (7) to only one, as shown in the following proposition.

Proposition 3.4. *If a matrix A has a decomposition (7) for $J = \text{diag}(\pm 1)$, then it also has the following two equivalent decompositions:*

$$A = (VP)T(VP)^{-1}, \quad V^*JV = J, \quad (8)$$

$$A = (PW)T(PW)^{-1}, \quad W^*J'W = J', \quad J' = P^*JP, \quad (9)$$

where $P = P_1P_2$ is a permutation.

The converse also holds, i.e., if A has a decomposition (8) or (9), then it also has a decomposition (7).

Proof. Let A have a decomposition (7) and let $V = P_1UP_1^*$ and $P = P_1P_2$. Then we get:

$$V^*JV = P_1U^*P_1^*JP_1UP_1^* = P_1U^*\tilde{J}UP_1^* = P_1\tilde{J}P_1^* = J,$$

so A has a decomposition (8).

Now, let A have a decomposition (7) and let $W = P_2^*UP_2$ and $P = P_1P_2$. Then we get:

$$\begin{aligned} W^*J'W &= P_2^*U^*P_2P_2^*P_1^*JP_1P_2P_2^*UP_2 = P_2^*U^*\tilde{J}UP_2 = P_2^*U^*\tilde{J}UP_2 \\ &= P_2^*P_1^*JP_1P_2 = J', \end{aligned}$$

so A has a decomposition (9).

To prove that the decompositions (8) and (9) are equivalent, let us assume that A has decomposition (8) and let us define $W = P^*VP$. Then

$$W^*J'W = (P^*V^*P)J'(P^*VP) = P^*VJVP = P^*JP = J',$$

which shows that A has a decomposition (9).

If A has a decomposition (9), we define $V = PWP^*$ and similarly get that it also has a decomposition (8).

The converse (that A has a decomposition (7) if it has (8) or (9)) trivially follows from (9) by defining $P_1 = P$ and $P_2 = I$. \square

Note that (8) is equivalent to

$$A = \widehat{V}T\widehat{V}^{-1}, \quad \widehat{V}^*J\widehat{V} = P^*JP, \quad (10)$$

i.e., \widehat{V} is hyperexchange (orthonormal), so A is orthonormally similar to a quasitriangular matrix T . Also, (9) is equivalent to

$$P^*AP = WTW^{-1}, \quad W^*J'W = J', \quad J' = P^*JP, \quad (11)$$

i.e., some symmetric permutation of A is J' -unitary similar to a quasitriangular matrix T , where J' is equal to J permuted with the same permutation as A .

We are now ready to define the hyperbolic Schur decomposition.

Definition 3.5 (Hyperbolic Schur decomposition). *For a given $A \in \mathbb{C}^{n \times n}$ and $J = \text{diag}(\pm 1)$, a hyperbolic Schur decomposition (with respect to J) is any decomposition of the form (7) or, equivalently, any of the forms (8)–(11), such that T is quasitriangular.*

Throughout the paper we also consider decompositions that resemble the Schur decomposition, but with some of the nontrivial blocks in T of order strictly larger than 2. We refer to such decompositions as the hyperbolic *Schur-like* decompositions.

Naturally, the most interesting question here concerns the existence of such a decomposition. The following theorem shows that all diagonalizable matrices have it.

Theorem 3.6 (Diagonalizable matrix). *If $A \in \mathbb{C}^{n \times n}$ is diagonalizable, then it has a hyperbolic Schur decomposition with respect to any given $J = \text{diag}(\pm 1)$.*

Proof. Let $A = SJS^{-1}$, where J is a diagonal matrix, be a Jordan decomposition of A . Since S is nonsingular, and therefore S^*JS is of full rank, by [19, Theorem 5.3], there exist matrices $Q \in \mathbb{C}^{n \times n}$ and $R \in \mathbb{C}^{n \times n}$ and permutations P_1 and P_2 such that

$$S = P_1QRP_2^*, \quad Q^*\tilde{J}Q = \tilde{J}, \quad \tilde{J} = P_1^*JP_1$$

and R is quasitriangular. Note that, since S is nonsingular, R is nonsingular as well and is therefore invertible. So,

$$A = SJS^{-1} = P_1QRP_2^*JP_2R^{-1}Q^{-1}P_1^*. \quad (12)$$

Since \mathcal{J} is diagonal, $P_2^* \mathcal{J} P_2$ is diagonal as well, which means it is also (block) upper triangular. We already have that R is quasitriangular and R^{-1} has the same block triangular structure as R . In other words,

$$T := R P_2^* \mathcal{J} P_2 R^{-1}$$

quasitriangular. From this and (12), we see that there exists $P := P_1$ such that

$$A = (PQ)T(PQ)^{-1}, \quad Q^* \tilde{J} Q = \tilde{J}, \quad \tilde{J} = P^* J P,$$

so (9) holds, which means that A has a hyperbolic Schur decomposition. \square

Obviously, there are also some nondiagonalizable matrices that have a hyperbolic Schur decomposition. Trivial examples are all Jordan matrices with at least one diagonal block of order greater than 1, since they are by definition both nondiagonalizable and triangular.

In Theorem 3.6, we assume A to be diagonalizable, which is then conveniently used in the proof. Under the assumption of diagonalizability, we can also mimic the traditional proof from the Euclidean case, using the diagonalizability for the convenient choice of the columns of the similarity matrix. Although technically far more complex than in the Euclidean case, this proof is also pretty straightforward and will be used to prove the existence of the hyperbolic Schur decomposition for some nondiagonalizable matrices in Proposition 3.11.

It is only natural to ask if there exists a nondiagonalizable matrix which does not have a hyperbolic Schur decomposition? As the following example shows, such matrices do exist.

Remark 3.7. *When discussing the counterexamples for the existence of a hyperbolic Schur decomposition, we shall often define our matrices via their Jordan decompositions because the (non)existence of a hyperbolic Schur decomposition heavily depends on the degeneracy and mutual J -orthogonality of the (generalized) eigenvectors, i.e., of (some) columns in the similarity matrix S .*

Example 3.8. *Let $J = \text{diag}(1, -1, 1, -1)$ and $A = S \mathcal{J}_4(\lambda) S^{-1}$ for some $\lambda \in \mathbb{C}$ and*

$$S = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & -1 \end{bmatrix}.$$

Let us assume that A has a hyperbolic Schur decomposition (10). As usual, T_{11} denotes the smallest nontrivial top left diagonal block of $T = [t_{ij}]$ (i.e., such that the elements of T beneath T_{11} are zero). In other words, T_{11} is of order 1 if $t_{21} = 0$ and of order 2 otherwise.

We denote the columns of \hat{V} as v_i and the columns of S as s_i . Note that

$$[v_i, v_j] = \begin{cases} \pm 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

Now, if T_{11} is of order 1, then v_1 is the eigenvector of A , i.e., it is collinear with s_1 . In other words, $v_1 = \xi s_1$ for some $\xi \neq 0$. Then

$$1 = |[v_1, v_1]| = |\xi|^2 |[s_1, s_1]| = 0,$$

which is an obvious contradiction, so T_{11} is of order 2.

Let T_{11} be of an order 2, with the elements denoted as t_{ij} , for $i, j \in \{1, 2\}$. Similar as in Example 3.3 (case 2), we define

$$A_2 := \left(A - \frac{t_{11} + t_{22}}{2} \mathbf{I} \right), \quad \lambda' := \frac{(t_{11} - t_{22})^2}{4} + t_{12}t_{21}.$$

Now, the first two columns of V , v_1 and v_2 , are both (linearly independent) eigenvectors of A_2 with the same eigenvalue λ' .

It is easy to see that if A_2 is nonsingular, then it is similar to $\mathcal{J}_4(\lambda')$, which means it has only one eigenvector. This is a contradiction with the assumption that v_1 and v_2 are linearly independent. Hence, A_2 is singular, i.e., $\lambda' = 0$, and we get $t_{11} + t_{22} = 2\lambda$. A simple calculation yields two eigenvectors of A_2 :

$$s'_1 = [1 \ 1 \ 0 \ 0]^T, \quad s'_2 = [0 \ 0 \ 1 \ 1]^T.$$

Since they span the same eigenspace as v_1 and v_2 , we conclude that

$$v_1 = \alpha_{11}s'_1 + \alpha_{21}s'_2, \quad v_2 = \alpha_{12}s'_1 + \alpha_{22}s'_2.$$

for some $\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}$. Note that

$$[s'_1, s'_1] = 0, \quad [s'_2, s'_2] = 0, \quad [s'_1, s'_2] = 0,$$

which means that s'_1 and s'_2 are both J -degenerate and mutually J -orthogonal. The contradiction is now obvious:

$$\begin{aligned} \pm 1 &= [v_1, v_1] = [\alpha_{11}s'_1 + \alpha_{21}s'_2, \alpha_{11}s'_1 + \alpha_{21}s'_2] \\ &= |\alpha_{11}|^2 [s'_1, s'_1] + 2 \operatorname{Re}(\alpha_{11}\overline{\alpha_{21}} [s'_1, s'_2]) + |\alpha_{21}|^2 [s'_2, s'_2] = 0. \end{aligned}$$

If we have allowed bigger diagonal blocks in T , would all square matrices have a hyperbolic Schur decomposition? As it turns out, the answer is “no”, for any constant upper limit on the size of those blocks. This can be proved by contradiction, as shown in the following remark.

Remark 3.9 (Bigger diagonal blocks in T). *Let us assume that a hyperbolic Schur-like decomposition, such that the biggest diagonal block of the block triangular factor T is of an order at most $k \in \mathbb{N}$, always exists.*

For a given even n , we set $J = (1, -1, 1, -1, \dots, 1, -1)$, i.e., $j_{ii} = (-1)^{i+1}$, of order n and construct a nonsingular matrix S as in the previous example, such that the first $n/2$ columns are J -degenerate and mutually J -orthogonal.

One such matrix is

$$s_{ij} = \begin{cases} 1, & (j \leq n/2 \text{ and } i \leq n - 2j + 2) \text{ or} \\ & (j > n/2 \text{ and } i \leq 2(n - j) + 1), \\ -1, & \text{otherwise.} \end{cases}$$

There is a simpler formula that achieves the same result, but this one was chosen for its likeness with the previous example.

We define $A = S\mathcal{J}_n(\lambda)S^{-1}$ for some $\lambda \in \mathbb{C}$. Repeating the steps from the previous example, we easily see that all eigenvectors of $\prod_{i=1}^j (A - t_{ii}\mathbf{I})$ are J -degenerate and mutually J -orthogonal for all $j \leq k$, showing that A has no hyperbolic Schur-like decomposition with the diagonal blocks of T having the order $k \leq n/2$.

In other words, no constant k would suffice as the upper bound for the size of the diagonal blocks in T .

Finally, let us assume that all matrices of order $n \geq n_0$, for some n_0 , have a hyperbolic Schur-like decomposition with the diagonal blocks of T of order at most $f(n)$ such that $f(n) < n$. In other words, we assume that there is a nontrivial Schur-like decomposition^a for all square matrices of order greater than some fixed n_0 , but with no constant limit on the size of the largest block in T . Instead, we assume that there is a function $f(n)$ which returns the upper bound of such blocks for all matrices of order n .

In this case, we could reapply our decomposition to the blocks of T of order at least n_0 , reducing them further until the biggest block in the triangular factor is smaller than n_0 . This way, we get that n_0 is the constant upper bound for the size of the diagonal blocks in T which is a contradiction with Remark 3.9.

Hence, we conclude that some matrices do not have a hyperbolic Schur-like decomposition even if we allow the diagonal blocks of T to be of an arbitrary size $k < n$. Therefore, it is sufficient for T to have blocks of order 1 and 2. Obviously, for $k = n$ we always have trivial (and useless) hyperbolic Schur-like “decomposition” $A = \mathbf{I}_n A \mathbf{I}_n^{-1}$.

^aBy nontrivial, we mean that the order of the biggest diagonal block in T is strictly smaller than the order of the decomposing matrix.

Since the blocks in T are of an order at most 2, it makes sense to ask if all the matrices with the Jordan blocks of order at most 2 (in their Jordan decomposition) have a hyperbolic Schur decomposition. As the following example shows, this is also, unfortunately, not the case.

Example 3.10. Let $J = \text{diag}(1, -1, 1, -1)$ and let $A = S\mathcal{J}S^{-1}$, where

$$S = \begin{bmatrix} 1 & 1 & 1 & 2 \\ 1 & 2 & 1 & 1 \\ 1 & 2 & -1 & -1 \\ 1 & 1 & -1 & -2 \end{bmatrix}, \quad \mathcal{J} = J_2(\lambda_1) \oplus J_2(\lambda_2), \quad \lambda_1 \neq \lambda_2.$$

Let A have a hyperbolic Schur decomposition (10) with $\widehat{V} = [v_1 \dots v_4]$ hyperexchange and T quasitriangular and let $s_i = Se_i$. Using the same argumentation as in Example 3.8, we see that the top left block of T has to be of order 2. Also, v_1 and v_2 must be the eigenvectors of some $A_2(\xi) := (A - \xi I)^2$ associated with the same eigenvalue (for some ξ). However, v_1 and v_2 must be linearly independent, and the only 2-dimensional eigenspaces of $A_2(\xi)$ are those spanned by:

1. $\{s_1, s_2\}$ for $\xi = \lambda_1$,
2. $\{s_1, s_3\}$ for $\xi = \frac{1}{2}(\lambda_1 + \lambda_2)$, and
3. $\{s_3, s_4\}$ for $\xi = \lambda_2$.

We get

$$S^*JS = \begin{bmatrix} s_j^*Js_i \\ 2 \\ -2 \\ -2 \\ 2 \end{bmatrix},$$

so each of these sets consists of degenerate, mutually J -orthogonal vectors.

It is easy to see that the linear combinations of such vectors are also degenerate (and mutually J -orthogonal), which is a contradiction with the assumption that v_1 and v_2 are columns of a hyperexchange matrix \widehat{V} .

However, if we limit the matrix to have only one Jordan block of order 2 (the rest of the Jordan form being diagonal), it will always have a hyperbolic Schur decomposition, as shown in the following proposition.

Its proof is a constructive one, following the idea of the iterative reduction, as outlined at the beginning of this section.

Proposition 3.11. *Let $A \in \mathbb{C}^{n \times n}$ have a Jordan decomposition $A = SJS^{-1}$ such that J has at most one block of order 2, while all others are of order 1. Then A has a hyperbolic Schur decomposition for any given $J = \text{diag}(\pm 1)$.*

Proof. If all Jordan blocks of A are of order 1, the matrix A is diagonalizable and, by Theorem 3.6, has a hyperbolic Schur decomposition. So, we shall only consider a case when A has (exactly one) Jordan block of order 2.

Case 1 If there is a nondegenerate eigenvector s_1 or A , we can J -normalize it, obtaining the J -normal eigenvector $v_1 = s_1/|s_1|$. As explained in [6, page 10], v_1 can be expanded to the J -orthonormal basis $\{v_1, v_2, \dots, v_n\}$. Defining a matrix

$$V := \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix},$$

we see that

$$A = V \left[\begin{array}{c|c} t_{11} & * \\ \hline 0 & A' \end{array} \right] V^{-1}, \quad V^*JV = P^*JP.$$

We repeat the process on A' until we either get to the block of order 2 or to the A' such that all its eigenvectors are J' -degenerate, where J' is the corresponding (bottom right) part of P^*JP . It is not hard to see that this sequence really gives the hyperbolic Schur-like decomposition such that blocks in T are of order 1, except maybe for the bottom right one which may be of an arbitrary order, which is covered by Case 2.

Case 2 We now focus on A such that all its eigenvectors are J -degenerate, i.e., $[s_i, s_i] = 0$ for all i (except, maybe, the second one, since s_2 is not an eigenvector but (the second) generalized eigenvector associated with the aforementioned block $\mathcal{J}_2(\lambda)$), as this is the only case not resolved by the previously described reductions.

Without the loss of generality, we may assume that

$$\mathcal{J} = \mathcal{J}_2(\lambda_1) \oplus \lambda_2 \oplus \cdots \oplus \lambda_{n-1}.$$

So, by assumption, $[s_i, s_i] = 0$, for all $i \in \{1, 3, 4, \dots, n\}$. Note the absence of the second column, as this one is not an eigenvector, but the generalized eigenvector associated with $\mathcal{J}_2(\lambda_1)$.

Case 2.1 If $[s_1, s_2] \neq 0$ and $[s_2, s_2] \neq 0$, we define

$$v'_1 = \xi s_1 + s_2, \quad v'_2 = s_2.$$

Since s_1 and s_2 are linearly independent, v'_1 and v'_2 are also linearly independent for every $\xi \neq 0$. We shall define the appropriate ξ in a moment. Note that

$$[v'_1, v'_2] = [\xi s_1 + s_2, s_2] = \xi[s_1, s_2] + [s_2, s_2].$$

Since we want to construct a J -orthonormal set, we want $[v'_1, v'_2] = 0$, so we define

$$\xi := -\frac{[s_2, s_2]}{[s_1, s_2]}.$$

Now that v'_1 and v'_2 are J -orthogonal, we need to be able to J -normalize them and, in order to do that, we need them to be nondegenerate. Vector $v'_2 = s_2$ is nondegenerate by assumption. We check the (non)degeneracy of vector v'_1 , using the fact that $[s_2, s_2] \in \mathbb{R}$, which is valid for all vectors in any indefinite scalar product space:

$$\begin{aligned} [v'_1, v'_1] &= [\xi s_1 + s_2, \xi s_1 + s_2] = |\xi|^2[s_1, s_1] + 2\operatorname{Re}(\xi[s_1, s_2]) + [s_2, s_2] \\ &= -2\operatorname{Re}[s_2, s_2] + [s_2, s_2] = -[s_2, s_2] \neq 0. \end{aligned}$$

We define $v_1 = v'_1/|v'_1|$ and $v_2 = v'_2/|v'_2|$, obtaining the J -orthonormal set $\{v_1, v_2\}$. As we did in Case 1, we expand this set to the J -orthonormal basis $\{v_1, v_2, \dots, v_n\}$, define the J -orthonormal matrix V with columns v_1, \dots, v_n and, by construction, see that

$$A = V \left[\begin{array}{c|c} T_{11} & * \\ \hline 0 & A' \end{array} \right] V^{-1}, \quad V^*JV = P^*JP. \quad (13)$$

Here, T_{11} is of order 2 and the matrix A' is diagonalizable. Hence, by Theorem 3.6, A' has a hyperbolic Schur decomposition, so A has one too.

Case 2.2 Let us now assume that $[s_1, s_2] \neq 0$ and $[s_2, s_2] = 0$. We define

$$v'_1 = \xi s_1 - s_2, \quad v'_2 = \xi s_1 + s_2.$$

As before, v'_1 and v'_2 are linearly independent for every $\xi \neq 0$. In order to define the appropriate ξ , note that

$$\begin{aligned} [v'_1, v'_2] &= [\xi s_1 - s_2, \xi s_1 + s_2] = |\xi|^2 [s_1, s_1] + 2 \operatorname{Im}(\xi [s_1, s_2]) + [s_2, s_2] \\ &= 2 \operatorname{Im}(\xi [s_1, s_2]). \end{aligned}$$

Hence, to obtain the J -orthonormality of v'_1 and v'_2 , we define

$$\xi := \overline{[s_1, s_2]},$$

once again getting a J -orthogonal set $\{v'_1, v'_2\}$. As before, we check that these vectors are J -nondegenerate:

$$\begin{aligned} [v'_1, v'_1] &= [\xi s_1 - s_2, \xi s_1 - s_2] = |\xi|^2 [s_1, s_1] - 2 \operatorname{Re}(\xi [s_1, s_2]) + [s_2, s_2] \\ &= -2|[s_1, s_2]|^2 \neq 0, \\ [v'_2, v'_2] &= [\xi s_1 + s_2, \xi s_1 + s_2] = |\xi|^2 [s_1, s_1] + 2 \operatorname{Re}(\xi [s_1, s_2]) + [s_2, s_2] \\ &= 2|[s_1, s_2]|^2 \neq 0. \end{aligned}$$

Next, we J -normalize vectors v'_1 and v'_2 to obtain a J -orthonormal set $\{v_1, v_2\}$, expand it to a J -orthonormal basis $\{v_1, \dots, v_n\}$ and a J -orthonormal matrix V . By construction, (13) holds and we can further decompose A' , which is again diagonalizable.

Case 2.3 We have now covered all the cases such that $[s_1, s_2] \neq 0$, so we now assume that $[s_1, s_2] = 0$. Let k be such that $[s_1, s_k] \neq 0$. Obviously, $k \neq 2$, so both vectors s_1 and s_k are J -degenerate eigenvectors of A . Note that such k must exist because, otherwise, k -th row and column of S^*JS would be zero, which is contradictory to the assumption that S and J are nonsingular.

We handle this case exactly the same way we did Case 2.2. The only difference is in the exact formula for the block T_{11} in (13), which we omit, as it is unimportant for the proof. \square

Although it may not be very obvious, the condition of having at most one Jordan block of order 2 in A is very important. Without this condition, in Case 2.3 of the proof, we would not be able to make a conclusion about s_1 and s_k being the eigenvectors. Instead, s_k might be a generalized eigenvector associated with a Jordan block of order 2 belonging to the some eigenvector other than s_1 . If the described construction was followed in that case, (13) would not hold, because T_{11} would not be of order 2. Instead, it would be associated to s_1, s_{k-1} and s_k , hence being of dimension 3×2 . We have shown (see Example 3.10) that this is not just a matter of construction, but a matrix A with 2 Jordan blocks of order

2 exists, such that A does not have a hyperbolic Schur decomposition.

We have researched the existence of the hyperbolic Schur decomposition of a given matrix in relation to its Jordan structure. But, the Schur decomposition is particularly interesting as a decomposition that preserves structures (with respect to the corresponding scalar product), so it is only natural to research the existence of the hyperbolic Schur decomposition for J -Hermitian and J -unitary matrices.

A discussion in [6, Section 5.6] gives a detailed analysis of J -Hermitian matrices for the case when J (in [6] denoted as H) has exactly one negative eigenvalue. In the case of the hyperbolic scalar products, the discussion is basically about the Minkowski spaces, i.e.,

$$J = \pm \text{diag}(1, -1, \dots, -1) \quad \text{or} \quad J = \pm \text{diag}(1, \dots, 1, -1).$$

However, Proposition 3.11 cannot be used to conclude that all J -Hermitian matrices in the Minkowski space have a hyperbolic Schur decomposition. The problem arises from the case (iv) in the aforementioned discussion in [6, Section 5.6], which states that a J -Hermitian matrix can have a Jordan block of order 3. Using this case, we construct the following very important example which shows that there really exist such matrices that have no hyperbolic Schur decomposition.

Example 3.12 (J -Hermitian matrix that does not have a hyperbolic Schur decomposition). *Let $J = \text{diag}(1, 1, -1)$ and let*

$$A = \begin{bmatrix} 12 & 11 & -16 \\ 11 & 8 & -13 \\ 16 & 13 & -20 \end{bmatrix} = S\mathcal{J}_3(0)S^{-1}, \quad S = \begin{bmatrix} 3 & 5 & 3 \\ 4 & 5 & 3 \\ 5 & 7 & 4 \end{bmatrix}.$$

*Let us assume that A has a hyperbolic Schur decomposition $A = VTV^{-1}$, $V^*JV = P^*JP$, with s_1 being the only eigenvector of A . Note that*

$$S^*JS = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 2 \end{bmatrix},$$

which means that s_1 is degenerate and J -orthogonal to s_2 . Following the previous discussions (see Example 3.8 or the proof of Proposition 3.11), we see that the top left block of T must be of order 2. But, the associated (first two) columns of V , denoted v_1 and v_2 , must be J -normal, mutually J -orthogonal linear combinations of s_1 and s_2 . This is impossible, since s_1 is degenerate and s_1 and s_2 are J -orthogonal.

To show this, assume that such vectors exist, i.e., we have α_{ij} such that

$$v_1 = \alpha_{11}s_1 + \alpha_{12}s_2, \quad v_2 = \alpha_{21}s_1 + \alpha_{22}s_2.$$

We check the desired properties of v_1 and v_2 . First, J -normality:

$$\begin{aligned} 1 &= [v_1, v_1] = [\alpha_{11}s_1 + \alpha_{12}s_2, \alpha_{11}s_1 + \alpha_{12}s_2] \\ &= |\alpha_{11}|^2[s_1, s_1] + 2\operatorname{Re}(\alpha_{11}\overline{\alpha_{12}}[s_1, s_2]) + |\alpha_{12}|^2[s_2, s_2] = |\alpha_{12}|^2, \\ 1 &= [v_2, v_2] = [\alpha_{21}s_1 + \alpha_{22}s_2, \alpha_{21}s_1 + \alpha_{22}s_2] \\ &= |\alpha_{21}|^2[s_1, s_1] + 2\operatorname{Re}(\alpha_{21}\overline{\alpha_{22}}[s_1, s_2]) + |\alpha_{22}|^2[s_2, s_2] = |\alpha_{22}|^2. \end{aligned}$$

So, $|\alpha_{12}| = |\alpha_{22}| = 1$. Using this result, we analyse the J -orthogonality of v_1 and v_2 :

$$\begin{aligned} 0 &= |[v_1, v_2]| = |[\alpha_{11}s_1 + \alpha_{12}s_2, \alpha_{21}s_1 + \alpha_{22}s_2]| \\ &= |\alpha_{11}\overline{\alpha_{21}}[s_1, s_1] + \alpha_{11}\overline{\alpha_{22}}[s_1, s_2] + \alpha_{12}\overline{\alpha_{21}}[s_2, s_1] + \alpha_{12}\overline{\alpha_{22}}[s_2, s_2]| \\ &= |\alpha_{12}\overline{\alpha_{22}}[s_2, s_2]| = 1, \end{aligned}$$

which is an obvious contradiction, hence no hyperbolic Schur decomposition exists for A .

The previous example is very significant, as it proves that not even all J -Hermitian matrices have a hyperbolic Schur decomposition. However, it may be interesting to show how this example was constructed.

Choosing the Jordan block is simple: the only possible order is 3 (because, if it was smaller, A would have a hyperbolic Schur decomposition, by Proposition 3.11) and the eigenvalue λ is arbitrary (as long as it is real), since A is J -Hermitian if and only if $A - \xi I$ is J -Hermitian for an arbitrary choice of $\xi \in \mathbb{R}$. We chose $\lambda = 0$ to make the example as simple as possible.

The construction of S is more interesting. We start by determining the structure of S^*JS . From the proof of Proposition 3.11 we see that s_1 is degenerate and orthogonal to s_2 , i.e., $[s_1, s_1] = [s_1, s_2] = 0$. So, we can define

$$S^*JS = \begin{bmatrix} 0 & 0 & x \\ 0 & y & a \\ \bar{x} & \bar{a} & b \end{bmatrix},$$

where $x, y \neq 0$ (due to the nonsingularity of S and, so, S^*JS) and $y, b \in \mathbb{R}$ (because S^*JS is Hermitian). Using $A^{[*]} = A$, for $A = S\mathcal{J}_3(0)S^{-1}$, we see that

$$JS^{-*}\mathcal{J}_3(0)^*S^*J = S\mathcal{J}_3(0)S^{-1}.$$

Premultiplying by S^*J and postmultiplying by S , we get

$$\mathcal{J}_3(0)^*S^*JS = S^*JS\mathcal{J}_3(0).$$

It is now obvious that $x = y \in \mathbb{R}$ and $a \in \mathbb{R}$. Having these conditions, it is easy to construct an appropriate matrix S . The first column was chosen as Pitagora's triple (to keep everything in the integer arithmetics), the second column was

set to get $y = 1$ (once again, to keep it as simple as possible), while the third one was chosen to respect the obtained condition $x = y$.

Using the similar approach, we can construct a J -unitary matrix that does not have a hyperbolic Schur decomposition. Unlike in the previous example, we do not have as much freedom in choosing the matrix' eigenvalue. By [16, Theorem 7.2], we see that $|\lambda| = 1$ (because our matrix will have only one Jordan block, so $\lambda = 1/\bar{\lambda}$). For the sake of simplicity, we take our pick at $\lambda = 1$ and construct the appropriate S using the defining relation $U^*JU = J$ for $U = S\mathcal{J}_3(1)S^{-1}$.

Example 3.13 (J -unitary matrix without a hyperbolic Schur decomposition). Let $J = \text{diag}(1, 1, -1)$ and let

$$U = \begin{bmatrix} -1 & -32 & 31 \\ 8 & -8 & 8 \\ 1 & -32 & 33 \end{bmatrix} = S\mathcal{J}_3(1)S^{-1}, \quad S = \frac{1}{8} \begin{bmatrix} 3 & 10 & 23 \\ 4 & 10 & 23 \\ 5 & 14 & 33 \end{bmatrix}.$$

It is easy to see that U is J -unitary. Also, it does not have a hyperbolic Schur decomposition, which can be shown using exactly the same arguments as in Example 3.12.

Even though the above examples show that some J -Hermitian and J -unitary matrices do not have a hyperbolic Schur decomposition, they also show how rare such matrices are: in both cases we had rather strict conditions that had to be met in order to construct them.

Interestingly, a special class of J -Hermitian matrices, referred to as J -nonnegative matrices, always has a hyperbolic Schur decomposition.

We say that a matrix A is J -nonnegative if JA is positive semidefinite and A is J -positive if JA is positive definite. These are, in a way, hyperbolic counterparts of positive definite and semidefinite matrices and find their applications in the research of J -nonnegative spaces and the semidefinite J -polar decomposition. For details, see [3] and [6].

Theorem 3.14 (J -nonnegative matrix). Let $J = \text{diag}(\pm 1)$. If $A \in \mathbb{C}^{n \times n}$ is J -nonnegative, then it has a hyperbolic Schur decomposition with respect to J .

Proof. Since JA is positive semidefinite, we can write $A = JB^*B$ for some B . Then the hyperbolic SVD of B , as described in [23, Section 3], where J is denoted as Φ , A^* as A^H and JV as P (we shall use P as a permutation matrix, which is not explicitly used in [23]) is

$$B = U \begin{bmatrix} \begin{bmatrix} I_j & \\ & I_j \end{bmatrix} & & \\ & \text{diag}(|\lambda_1|^{1/2}, \dots, |\lambda_l|^{1/2}) & \\ & & 0 \end{bmatrix} (JV)^*,$$

where $U^*U = I$, $(JV)^*J(JV) = V^*JV = P^*JP$, for some permutation P , and

$j = \text{rank } B - \text{rank } BJB^*$. It is easy to see that

$$A = JB^*B = V \begin{bmatrix} \begin{bmatrix} \mathbf{I}_j & \mathbf{I}_j \\ \mathbf{I}_j & \mathbf{I}_j \end{bmatrix} & & \\ & \text{diag}(|\lambda_1|, \dots, |\lambda_l|) & \\ & & 0 \end{bmatrix} P^*JPV^{-1}.$$

Note that P^*JP is diagonal and

$$\begin{bmatrix} \mathbf{I}_j & \mathbf{I}_j \\ \mathbf{I}_j & \mathbf{I}_j \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \otimes \mathbf{I}_j$$

is permutationally similar to

$$\bigoplus_{i=1}^j \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \mathbf{I}_j \otimes \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},$$

which completes the proof of the theorem. \square

Remark 3.15. *If A is J -positive, then JA is nonsingular, so $j = 0$ in the above proof, which means that J -positive matrices have a hyperbolic Schur decomposition with T diagonal with positive entries.*

In fact, more is known for this case, as shown in [20, Corollary 5.3]:

$$A = VTV^{-1}, \quad V^{-1} = V^{[*]}, \quad T = J\Sigma, \quad \Sigma = \text{diag}(|\lambda_1|, \dots, |\lambda_n|),$$

i.e., V can be chosen to be J -unitary if we set the signs in T according to those in J .

Before exploring the properties of the hyperbolic Schur decomposition, we investigate further the quasitriangular factor T . To do this, we define the following very important variant of the hyperbolic Schur decomposition.

Definition 3.16 (The complete hyperbolic Schur decomposition). *Let $J = \text{diag}(\pm 1)$ and $A \in \mathbb{C}^{n \times n}$ be of same order. We say that $A = VTV^{-1}$ is a complete hyperbolic Schur decomposition of the matrix A with respect to J if V is J -orthonormal, T is quasitriangular, and all diagonal blocks T_{kk} of order 2 are irreducible¹.*

Definition 3.16 allows us to work with the fully reduced (in terms of the J -orthonormal similarity) quasitriangular matrices. This makes proofs simpler (by reducing the number of observed cases) and gives us the following properties of the diagonal blocks in a irreducible matrix T . Note that any matrix having a hyperbolic Schur decomposition trivially also has a complete hyperbolic Schur decomposition.

¹A diagonal block T_{kk} is *irreducible* if it cannot be further reduced by the hyperbolic Schur decomposition, i.e., there are no \tilde{J}_k -orthogonal \tilde{V} and triangular (not just quasitriangular!) \tilde{T} such that $T_{kk} = \tilde{V}\tilde{T}\tilde{V}^{-1}$, where \tilde{J}_k is the part of $\tilde{J} := V^*JV$ corresponding to T_{kk} .

Theorem 3.17 (The complete hyperbolic Schur decomposition). *Let $A = VTV^{-1}$ be a complete hyperbolic Schur decomposition of A with respect to some $J = \text{diag}(\pm 1)$. Then all the nontrivial diagonal 2×2 blocks of T have only degenerate eigenvectors with respect to the corresponding part of $\tilde{J} := V^*JV$.*

Furthermore, all such blocks are either degenerate (with respect to the corresponding part of \tilde{J}) or nonsingular.

Proof. It is sufficient to note that, once we have a hyperbolic Schur decomposition, we can further decompose nontrivial 2×2 diagonal blocks of the quasitriangular factor T , either using the traditional Schur decomposition (if the corresponding part of \tilde{J} , denoted J' , is definite, i.e., $J' = \pm I_2$) or the hyperbolic Schur decomposition (if J' is hyperbolic, i.e., $J' = \text{diag}(1, -1)$ or $J' = \text{diag}(-1, 1)$). In the latter case, we can triangularize the observed 2×2 block if and only if it has at least one nondegenerate eigenvector, by J' -normalizing that eigenvector and expanding it to the J' -orthonormal basis, which can always be done for a J' -orthonormal set (see [6, p. 10]). So, the only irreducible 2×2 blocks of T are those with degenerate eigenvectors with respect to the hyperbolic scalar product induced by the corresponding part of \tilde{J} .

For the second part of the theorem, regarding the nontrivial 2×2 diagonal blocks in the factor T , note that all nontrivial singular blocks of order 2 must have rank 1 (otherwise, they are either 0 (which is a trivial block) or nonsingular). Let us observe one of such blocks, denoting it as T' and the corresponding part of \tilde{J} as $J' = \pm \text{diag}(1, -1)$. We see that T' must be in one of the following two forms:

$$1. T' = S \text{diag}(\lambda, 0)S^{-1}, \lambda \neq 0, S^* J' S = \begin{bmatrix} 0 & \alpha \\ \bar{\alpha} & 0 \end{bmatrix}, \alpha \in \mathbb{C} \setminus \{0\}, \text{ which gives}$$

$$\begin{aligned} (T')^* J' T' &= S^{-*} \text{diag}(\bar{\lambda}, 0) S^* J' S \text{diag}(\lambda, 0) S^{-1} \\ &= S^{-*} \text{diag}(\bar{\lambda}, 0) \begin{bmatrix} 0 & \alpha \\ \bar{\alpha} & 0 \end{bmatrix} \text{diag}(\lambda, 0) S^{-1} = 0, \text{ or} \end{aligned}$$

$$2. T' = S \mathcal{J}_2(0) S^{-1}, S^* J' S = \begin{bmatrix} 0 & \alpha \\ \bar{\alpha} & \beta \end{bmatrix}, \alpha \in \mathbb{C} \setminus \{0\}, \beta \in \mathbb{C}, \text{ which gives}$$

$$\begin{aligned} (T')^* J' T' &= S^{-*} \mathcal{J}_2(0)^T S^* J' S \mathcal{J}_2(0) S^{-1} \\ &= S^{-*} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & \alpha \\ \bar{\alpha} & \beta \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} S^{-1} = 0. \end{aligned}$$

So, in both cases $(T')^* J' T' = 0$, which means that T' is J' -degenerate. Since T' was an arbitrary nontrivial diagonal singular block in the complete Schur decomposition, this means that all such blocks are degenerate. \square

The irreducible blocks are sometimes referred to as *atomic* (see [4, page 466]) and Theorem 3.17 states that all nontrivial blocks in the complete hyperbolic

Schur decomposition are atomic and have a specific eigenstructure. Note that each atomic block can have one or two eigenvectors, as Theorem 3.17 makes no statement on the Jordan structure of such blocks.

4. Properties

In this section, we assume that $J = \text{diag}(\pm 1)$ and A are given such that A has a hyperbolic Schur decomposition, as described in Definition 3.5. We also use V , P and \tilde{J} from (8) and (10).

One of the main properties of the traditional Schur decomposition is that it keeps the structure of the matrix unchanged, i.e., if the decomposed matrix A is normal, Hermitian or unitary, then the triangular block T of a Schur decomposition of A will also be normal, Hermitian or unitary, respectively. Not surprisingly, similar properties hold in the hyperbolic case as well.

Proposition 4.1 (*J-conjugate transpose*). *Let A have a hyperbolic Schur decomposition (8) with respect to $J = \text{diag}(\pm 1)$. Then*

$$A^{[*]J} = (VP)T^{[*]\tilde{J}}(VP)^{-1}, \quad \tilde{J} = P^*JP.$$

Proof. The proof is straightforward. From (8), it follows that

$$\begin{aligned} A^{[*]J} &= J((VP)T(VP)^{-1})^*J = J(VP)^{-*}T^*(VP)^*J \\ &= JV^{-*}PT^*P^*V^*J = (VP)P^*(V^*JV)^{-1}PT^*P^*V^*JVP(VP)^{-1} \\ &= (VP)\tilde{J}T^*\tilde{J}(VP)^{-1} = (VP)T^{[*]\tilde{J}}(VP)^{-1}. \quad \square \end{aligned}$$

As we have seen in Example 3.8, some matrices do not have a hyperbolic Schur decomposition. However, if a matrix A has it, then its conjugate transposes (both, the Euclidean and the hyperbolic one) have it as well.

Proposition 4.2 (*Existence of the hyperbolic Schur decomposition for conjugate transposes*). *A matrix A has a hyperbolic Schur decomposition with respect to $J = \text{diag}(\pm 1)$ if and only if A^* and $A^{[*]J}$ have it as well.*

Proof. The proof is a direct consequence of Proposition 4.1. Note that if some matrix X is lower triangular, then $\mathcal{S}_n^{-1}X\mathcal{S}_n$ is upper triangular. Now, we have:

$$\begin{aligned} A^{[*]J} &= (VP)T^{[*]\tilde{J}}(VP)^{-1} = (VP)\mathcal{S}_n\mathcal{S}_n^{-1}T^{[*]\tilde{J}}\mathcal{S}_n\mathcal{S}_n^{-1}(VP)^{-1} \\ &= (VP\mathcal{S}_n)(\mathcal{S}_n^{-1}T^{[*]\tilde{J}}\mathcal{S}_n)(VP\mathcal{S}_n)^{-1}, \end{aligned}$$

which is one possible hyperbolic Schur decomposition of $A^{[*]J}$. The similar proofs can be applied to A^* .

For the other implication, we need only apply what was already proven in the first part and use the fact that

$$A = (A^*)^* = (A^{[*]J})^{[*]J}. \quad \square$$

We now consider J -Hermitian matrices. As seen in the previous section, the hyperbolic Schur decomposition was defined in the way that it keeps J -Hermitianity. This property can be shown directly as well.

Proposition 4.3 (J -Hermitian matrices). *If a matrix A has a hyperbolic Schur decomposition with respect to $J = \text{diag}(\pm 1)$, then A is J -Hermitian if and only if T is \tilde{J} -Hermitian and quasidiagonal, where $\tilde{J} = P^*JP$.*

Proof. From Proposition 4.1, we see that for $A = (VP)T(VP)^{-1}$,

$$A^{[*]J} = (VP)T^{[*]\tilde{J}}(VP)^{-1},$$

so $T = T^{[*]\tilde{J}}$ if and only if $A^{[*]J} = A$.

If T is quasitriangular and \tilde{J} -Hermitian, then it is quasidiagonal. \square

At this point, it is worth noting that the spectrum of a J -Hermitian matrix A is always symmetric with respect to the real axis. Moreover, the Jordan structure associated with λ is the same as that associated with $\bar{\lambda}$, as shown in [6, Proposition 4.2.3.]. Also, by [6, Corollary 4.2.5], nonreal eigenvalues of J -Hermitian matrices have J -neutral root subspaces, which implies that all their adjointed eigenvectors are J -degenerate. This means that each such eigenvalue will participate in some singular atomic block of order 2 in the quasidiagonal matrix T of the matrix' hyperbolic Schur decomposition. Of course, for $J = \pm I$, all eigenvalues are real and such blocks do not exist.

In the following proposition, we consider J -normal matrices. Recall that A is J -normal if $AA^{[*]J} = A^{[*]J}A$.

Proposition 4.4 (J -normal matrices). *If a matrix A has a hyperbolic Schur decomposition with respect to $J = \text{diag}(\pm 1)$, then A is J -normal if and only if T is \tilde{J} -normal quasitriangular, where $\tilde{J} = P^*JP$.*

Proof. This follows directly from Proposition 4.1:

$$AA^{[*]J} = (VP)TT^{[*]\tilde{J}}(VP)^{-1}, \quad A^{[*]J}A = (VP)T^{[*]\tilde{J}}T(VP)^{-1}. \quad \square$$

Unlike the Euclidean case, where a normal triangular matrix is also diagonal, in the hyperbolic case, we have no guarantees that a block triangular J -normal matrix is also block diagonal, as shown in the following example.

Example 4.5 (A block triangular, J -normal matrix which is not block diagonal). *Let $J = \text{diag}(1, -1, 1, -1)$ and let A be the following block triangular matrix:*

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

for some $\xi \in \mathbb{C}$. A simple multiplication shows that $AA^{[*]} = A^{[*]}A$, for any $\xi \in \mathbb{C}$, and A is obviously not block diagonal (with the diagonal blocks of order 2).

Where does this difference between the Euclidean and the hyperbolic case come from? In the Euclidean case, for an upper triangular normal A , we have:

$$\sum_{k \leq n} |a_{1k}|^2 = (AA^*)_{11} = (A^*A)_{11} = |a_{11}|^2, \quad (14)$$

from which we conclude that $\sum_{1 < k \leq n} |a_{1k}|^2 = 0$, so $|a_{1k}| = 0$ for all $k > 1$. Then we do the same for $(AA^*)_{22}$, $(AA^*)_{33}$, etc. But, in the hyperbolic case, for a block upper triangular J -normal A , (14) takes the following form:

$$j_{11} \sum_{k \leq n} j_{kk} |a_{1k}|^2 = (AA^{[*]})_{11} = (A^{[*]}A)_{11} = |a_{11}|^2.$$

Since J contains both positive and negative numbers on its diagonal, the sum on the left hand side of the previous equation may also contain both positive and negative elements, so we can make no direct conclusion about the elements a_{1k} for any k .

Similarly to J -normal and J -Hermitian, we can also analyze J -unitary matrices. Somewhat surprisingly, this property is much closer to the Euclidean case than the previous result regarding J -normal matrices.

Let us first review the structure of quasitriangular hyperexchange matrices, as this will give us more insight into the structure of the J -unitary matrices (which are a special case of the hyperexchange matrices).

Proposition 4.6 (Block triangular hyperexchange matrices). *Let T be a quasitriangular hyperexchange matrix with respect to some given $J = \text{diag}(\pm 1)$. Then T is also quasidiagonal.*

Proof. Since T is a hyperexchange matrix, there exists a permutation P such that $T^*JT = PJP^*$. This means that $T^{-1} = PJP^*T^*J$. Because T is block upper triangular and J and PJP^* are diagonal,

1. T^{-1} is block upper triangular, and
2. T^* and PJP^*T^*J are block lower triangular.

Hence, T^{-1} is both block upper and block lower triangular, i.e., T^{-1} and therefore T are quasidiagonal. \square

We are now ready to analyze the hyperbolic Schur decomposition of a J -unitary matrix.

Proposition 4.7 (J -unitary matrices). *If a matrix A has a hyperbolic Schur decomposition $A = VTV^{-1}$ with respect to $J = \text{diag}(\pm 1)$, then A is J -unitary if and only if the quasitriangular factor T is \tilde{J} -unitary and quasidiagonal, where $\tilde{J} = P^*JP$.*

Proof. By Proposition 4.1, using the same arguments as in the proof of Proposition 4.4, T is J -unitary. The quasidiagonality of T follows directly from Proposition 4.6 and the fact that every \tilde{J} -unitary matrix is also \tilde{J} -hyperexchange. \square

The previous proposition can also be proven in a more straightforward manner, by analyzing the top right block of dimensions $1 \times (n - 1)$ or $2 \times (n - 2)$ in T^*JT and then repeating the process iteratively on the bottom right block of order $n - 1$ or $n - 2$.

Eigenvalues of J -unitary and J -Hermitian matrices have well researched properties, nicely presented in [16, Section 7] with J -unitary matrices being referred to as members of the automorphism group \mathbb{G} and J -Hermitian matrices being referred to as members of the Jordan algebra \mathbb{J} . These results apply for various scalar products, but when it comes to hyperbolic products and a hyperbolic Schur decomposition, more can be said about the atomic blocks in T .

Proposition 4.8 (Nondiagonalizable atomic blocks). *Let $J = \text{diag}(\pm 1)$ and A of the same order be given such that A has a complete hyperbolic Schur decomposition (8). Furthermore, let \widehat{T} be a nondiagonalizable atomic block on the diagonal of T , let \widehat{J} be the corresponding part of J and let s_1 and s_2 denote the columns of the similarity matrix S in the Jordan decomposition $\widehat{T} = S\mathcal{J}_2(\lambda)S^{-1}$. Then the following is true:*

1. *If A is J -unitary, then $|\lambda| = 1$ and $[s_1, s_2]_{\widehat{J}} \in i\mathbb{R} \setminus \{0\}$, and*
2. *If A is J -Hermitian, then $\lambda \in \mathbb{R}$ and $[s_1, s_2]_{\widehat{J}} \in \mathbb{R} \setminus \{0\}$.*

Proof. By Propositions 4.7 and 4.3, if A is J -unitary or J -Hermitian, T is quasidiagonal \widetilde{J} -unitary or \widetilde{J} -Hermitian, respectively. Since we have assumed a complete hyperbolic Schur decomposition, by Theorem 3.17 the eigenvector of every nondiagonalizable atomic block \widehat{T} is \widetilde{J} -degenerate, also implying that $\widehat{J} = \pm \text{diag}(1, -1)$. This means that

$$\widehat{T} = S\mathcal{J}_2(\lambda)S^{-1}, \quad S^*\widehat{J}S = \begin{bmatrix} 0 & \alpha \\ \bar{\alpha} & \beta \end{bmatrix}.$$

Since $S^*\widehat{J}S$ is Hermitian, $\beta \in \mathbb{R}$. Note that $\alpha = [s_2, s_1]_{\widehat{J}} \neq 0$ because $S^*\widehat{J}S$ is nonsingular.

Let us now assume that A is J -unitary. As stated before, T is \widetilde{J} -unitary and since it is quasidiagonal, its block \widehat{T} is \widetilde{J} -unitary. So,

$$\widehat{J} = \widehat{T}^*\widehat{J}\widehat{T} = S^{-*}\mathcal{J}_2(\lambda)^*S^*\widehat{J}S\mathcal{J}_2(\lambda)S^{-1} = S^{-*}\begin{bmatrix} 0 & \alpha \\ \bar{\alpha} & 2\lambda \text{Re}(\alpha) + \beta \end{bmatrix}S^{-1}.$$

Premultiplying by S^* and postmultiplying S , we get

$$\begin{bmatrix} 0 & \alpha \\ \bar{\alpha} & \beta \end{bmatrix} = S^*\widehat{J}S = \begin{bmatrix} 0 & \alpha \\ \bar{\alpha} & 2\lambda \text{Re}(\alpha) + \beta \end{bmatrix}.$$

From here, we see that $2\lambda \text{Re}(\alpha) = 0$. Since \widehat{T} is \widehat{J} -unitary and, hence, nonsingular, $\lambda \neq 0$. Obviously, $\text{Re}(\alpha) = 0$, i.e., α is imaginary, so $[s_1, s_2] \in i\mathbb{R}$.

The J -Hermitian case is similar. For a J -Hermitian A , as before, we conclude that T is \tilde{J} -Hermitian, and its block \hat{T} is \hat{J} -Hermitian. From $\hat{T}^{[*]} = \hat{T}$ follows that $\hat{J}\hat{T}^* = \hat{T}\hat{J}$, so

$$\hat{J}S^{-*}\mathcal{J}_2(\lambda)^*S^* = S\mathcal{J}_2(\lambda)S^{-1}\hat{J}.$$

Premultiplying by $S^*\hat{J}$ and postmultiplying $S^{-*}S$, we get

$$\mathcal{J}_2(\lambda)^* = S^*\hat{J}S\mathcal{J}_2(\lambda)S^{-1}\hat{J}S^{-*} = (S^*\hat{J}S)\mathcal{J}_2(\lambda)(S^*\hat{J}S)^{-1}.$$

Expanding all these matrices and multiplying those on the right hand side, we see that

$$\begin{bmatrix} \bar{\lambda} & \\ 1 & \bar{\lambda} \end{bmatrix} = \begin{bmatrix} \lambda & \\ \bar{\alpha}/\alpha & \lambda \end{bmatrix}.$$

Here, we see that $\lambda \in \mathbb{R}$ (which also follows straight from [16, Theorem 7.6]) and $\bar{\alpha} = \alpha$, i.e., α is real, so $[s_1, s_2] \in \mathbb{R}$. \square

It is worth noting that J -Hermitian J -unitary matrices always have a hyperbolic Schur decomposition with a specific structure of the atomic blocks of order 2, as shown in the following proposition.

Proposition 4.9 (*J -Hermitian J -unitary matrices*). *Let $J = \text{diag}(\pm 1)$ and let $A \in \mathbb{C}^{n \times n}$ be both J -Hermitian and J -unitary. Then A has a complete hyperbolic Schur decomposition (8) such that each diagonal atomic block \hat{T} of order 2 is of the form*

$$\hat{T} = \begin{bmatrix} \alpha & \beta \\ -\bar{\beta} & -\alpha \end{bmatrix}, \quad \hat{J} = \pm \text{diag}(1, -1),$$

for some $\alpha \in \mathbb{R}$, $\beta \in \mathbb{C}$ such that $\alpha^2 = |\beta|^2 + 1$. As before, \hat{J} denotes the part of \tilde{J} corresponding to \hat{T} .

Proof. From the J -unitarity and the J -Hermitianity of A , we get:

$$I = A^{[*]}A = A^2,$$

which means that A is involutory and, hence, diagonalizable (see [2, Fact 5.12.13]) which, by Theorem 3.6 means that it has a hyperbolic Schur decomposition. By Theorem 3.17, we can choose such a decomposition in a way that each diagonal block of order 2 in T , denoted \hat{T} , has only degenerate eigenvectors with respect to \hat{J} (the corresponding part of \tilde{J}), which is possible if and only if that part is either $\hat{J} = \text{diag}(1, -1)$ or $\hat{J} = \text{diag}(-1, 1)$.

By Proposition 4.3, \hat{T} is \hat{J} -Hermitian and, by Proposition 4.7, it is also \hat{J} -unitary. So,

$$I_2 = \hat{T}^{[*]\hat{J}}\hat{T} = \hat{T}^2. \quad (15)$$

Let us denote the elements of \hat{T} as t_{ij} ($i, j \in \{1, 2\}$):

$$\hat{T} = \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix}, \quad \hat{T}^{[*]\hat{J}} = \hat{J}\hat{T}^*\hat{J} = \begin{bmatrix} \bar{t}_{11} & -\bar{t}_{21} \\ -\bar{t}_{12} & \bar{t}_{22} \end{bmatrix}.$$

Since $\widehat{T} = \widehat{T}^{[*]\hat{J}}$, we see that $t_{11}, t_{22} \in \mathbb{R}$ and $t_{21} = -\overline{t_{12}}$. Note that we have chosen \widehat{T} to be nontrivial (because it comes from the complete hyperbolic Schur decomposition), i.e., $t_{12} \neq 0$ or $t_{21} \neq 0$, which means that both of the nondiagonal elements of \widehat{T} are nonzero. Furthermore, from (15), we get

$$\begin{bmatrix} 1 & \\ & 1 \end{bmatrix} = \begin{bmatrix} t_{11}^2 - |t_{12}|^2 & t_{12}(t_{11} + t_{22}) \\ -\overline{t_{12}}(t_{11} + t_{22}) & t_{22}^2 - |t_{12}|^2 \end{bmatrix}.$$

Since $t_{12} \neq 0$, we get $t_{11} = -t_{22}$. Defining $\alpha := t_{11}$ and $\beta := t_{12}$ completes this proof. \square

5. Existence of the hyperbolic Schur decomposition for J -Hermitian matrices

As we have seen in Example 3.12, some J -Hermitian matrices do not have a hyperbolic Schur decomposition. Also, Theorem 3.11 provides a sufficient, but not necessary conditions for the existence of such a decomposition in a general case. In this section we give a necessary and sufficient conditions for the existence of the hyperbolic Schur decomposition of J -Hermitian matrices.

To achieve our goal, we briefly leave the hyperbolic scalar product spaces and move to more general, indefinite scalar product spaces. A detailed analysis of such spaces is given in [6], from where we use Theorem 5.1.1. which states that for every nonsingular indefinite J and for every J -Hermitian A (i.e., $JA = A^*J$) there exists a nonsingular matrix X such that

$$A = X\mathcal{J}X^{-1}, \quad X^*JX = \mathcal{S}, \quad \mathcal{J} = \bigoplus_k \mathcal{J}_k, \quad \mathcal{S} = \bigoplus_k \varepsilon_k \mathcal{S}_k, \quad (16)$$

where \mathcal{J} is a Jordan normal form of A . Blocks \mathcal{J}_k for $k = 1, \dots, \alpha$ are associated with the real eigenvalues $\lambda_1, \dots, \lambda_\alpha$ and blocks \mathcal{J}_k for $k = \alpha + 1, \dots, \beta$ are associated with conjugate pairs of nonreal eigenvalues $\lambda_{\alpha+1}, \dots, \lambda_\beta$ in the upper half-plane, i.e., $\mathcal{J}_k = \mathcal{J}(\lambda_k) \oplus \mathcal{J}(\overline{\lambda}_k)$ for $k = \alpha + 1, \dots, \beta$. Matrices \mathcal{S}_k denote standard involutory permutations of order same as \mathcal{J}_k , $\varepsilon_k \in \{-1, 1\}$ for $k \leq \alpha$ and $\varepsilon_k = 1$ for $k > \alpha$. In [6], J , X and \mathcal{S} are denoted as H , T^{-1} and J , respectively.

The decomposition (16) is referred to as the *canonical form of the pair* (A, J) and has many applications. See [6, Chapter 5] for the applications in the indefinite scalar product spaces and [15] for its symplectic version and the application in the symplectic scalar product spaces.

The set of signs $\{\varepsilon_1, \dots, \varepsilon_\alpha\}$ is called the *sign characteristic of the pair* (A, J) . In [15], this notion is adapted to the symplectic scalar product spaces and the sign characterist includes the signs corresponding to all of the Jordan blocks. For more details, see [15, Section 3.2.].

Apart from the applications to the nonstandard scalar product spaces, the sign characteristic plays a significant role in the research of the self-adjoint matrix polynomials [5, Chapter 12].

Throughout this paper, the Jordan normal form of a matrix plays a crucial role in the existence of the hyperbolic Schur decomposition. To better describe it, we associate with each Jordan block $\mathcal{J}_d(\lambda)$ the term *partial multiplicity*, which is the size d of that Jordan block. Obviously, each eigenvalue has as many partial multiplicities as it has associated Jordan blocks and the largest partial multiplicity of each eigenvalue λ of a matrix A corresponds to the degree of a factor $x - \lambda$ in the minimal polynomial μ_A of A .

We are now ready to present and prove the necessary and sufficient conditions for the existence of the hyperbolic Schur decomposition of a J -Hermitian matrix.

Theorem 5.1 (Hyperbolic Schur decomposition of a J -Hermitian matrix). *Let $J = \text{diag}(\pm 1)$ and let $A \in \mathbb{C}^{n \times n}$ be a J -Hermitian matrix of the same order. Then A has a hyperbolic Schur decomposition with respect to J if and only if its real eigenvalues have partial multiplicities at most 2 and its non-real eigenvalues have multiplicities at most 1.*

Proof. Let us first show that if such a decomposition of A exists, then the real eigenvalues of A have partial multiplicities at most 2 and its non-real eigenvalues have multiplicities at most 1.

Let $A = UTU^{-1}$, where U is J -orthonormal, and let $\tilde{J} := U^*JU$. By Proposition 4.3, T is quasidiagonal and \tilde{J} -Hermitian. This means that we can write

$$T = \bigoplus_k T_k, \quad \tilde{J} = \bigoplus_k \tilde{J}_k,$$

where T_k are nontrivial (i.e., either of order 1 or nondiagonal of order 2) and \tilde{J}_k have the same order as T_k . Furthermore, each T_k is \tilde{J}_k -Hermitian. Trivially, this means that all blocks T_k of order 1 are real.

Notice that if T_k of order 2 has non-real eigenvalues $\lambda_1^{(k)}$ and $\lambda_2^{(k)}$, then they form a complex conjugate pair, i.e., $\lambda_1^{(k)} = \overline{\lambda_2^{(k)}}$, due to [16, Theorem 7.6], so their partial multiplicity must be 1. If they are real, their partial multiplicity is 1 or 2, depending on diagonalizability of T_k . Since A is similar to T_k , it has the same Jordan structure, i.e., the same eigenvalues with the same partial multiplicities.

Let us now show that any J -Hermitian matrix A with the described eigenvalues' partial multiplicities has a hyperbolic Schur decomposition with respect to J . We observe the canonical form of (A, J) from (16). By the assumption, all \mathcal{J}_k are of an order at most 2.

Let k be such that \mathcal{J}_k is of order 2. Then $\mathcal{S}_k = \begin{bmatrix} 1 & \\ & 1 \end{bmatrix}$, which is congruent with $\text{diag}(1, -1)$. By the Sylvester's law of inertia, there exists Y_k such that $Y_k^* \mathcal{S}_k Y_k = \text{diag}(1, -1)$. For all k such that \mathcal{J}_k is of order 1 we define $Y_k = 1$, so $Y_k^* \mathcal{S}_k Y_k = I_1$ for such k . Furthermore, we define

$$Y := \bigoplus_k Y_k, \quad U := XY, \quad T := Y^{-1} \mathcal{J} Y. \quad (17)$$

Note that T has the same block structure as \mathcal{J} , so it is quasidiagonal, and

$$Y^*SY = \left(\bigoplus_k Y_k^* \right) \left(\bigoplus_k \varepsilon_k \mathcal{S}_k^* \right) \left(\bigoplus_k Y_k \right) = \bigoplus_k (\varepsilon_k Y_k^* \mathcal{S}_k^* Y_k) =: J', \quad (18)$$

where $J' = \text{diag}(j_1, \dots, j_n)$ for some $j_1, \dots, j_n \in \{-1, 1\}$. Using (16), (17) and (18), we see that

$$\begin{aligned} A &= X\mathcal{J}X^{-1} = UY^{-1}\mathcal{J}YU^{-1} = UTU^{-1}, \\ U^*JU &= (XY)^*J(XY) = Y^*X^*JXY = Y^*SY = J', \end{aligned}$$

which proves that A is J -orthonormally similar to some quasidiagonal T , hence A has a hyperbolic Schur decomposition with respect to J . \square

6. Conclusion

In this paper we have introduced the hyperbolic Schur decomposition of a square matrix with respect to the scalar product induced by $J = \text{diag}(\pm 1)$. We have shown that all diagonalizable matrices have such a decomposition, which means that the set of matrices that don't have a hyperbolic Schur decomposition is a subset of the set of nondiagonalizable matrices, which is a set of measure zero (this follows from [12, Section 2.4.7]). We have also given examples of matrices that do not have such a decomposition.

By its design, the hyperbolic Schur decomposition preserves structures of the structured matrices, albeit with respect to a somewhat changed (symmetrically permuted) J , denoted by \tilde{J} , which is a common property of hyperbolic decompositions. We have analyzed the properties of such matrices, as well as the properties of atomic (irreducible) blocks on the diagonal of the quasitriangular factor T .

In Example 3.12 we have shown that there exist J -Hermitian (and, therefore, J -normal) matrices that do not have a hyperbolic Schur decomposition, even in spaces as simple as the Minkowski space. In Example 3.13 we have provided an example of a J -unitary matrix for which there is no hyperbolic Schur decomposition, also showing that not all hyperexchange matrices have such a decomposition (since J -unitary matrices are a special case).

In section 5, we have given sufficient and necessary conditions under which a J -Hermitian matrix has a hyperbolic Schur decomposition.

Even though each of the important sets of structured matrices (J -normal, J -Hermitian, J -unitary) contains matrices that do not have a hyperbolic Schur decomposition, such examples are very rare, which can be seen from their construction, described in detail, following Example 3.12.

We have shown that those structured matrices that do have a hyperbolic Schur decomposition also have desirable properties. The only exemption from this rule are general J -normal matrices (which are not “nicer”, i.e., neither J -Hermitian nor hyperexchange) for which block triangularity does not imply

block diagonality. It remains to be researched if the block triangular ones can always be block diagonalized via J -orthonormal similarities (i.e., by a hyperbolic Schur decomposition), as always happens in the Euclidean case. Interestingly enough, all important subclasses of J -normal matrices maintain the block diagonality of the factor T .

Another subject that remains to be researched is the algorithm to calculate such a decomposition. The first thing that might come to mind here is the QR algorithm used for the Euclidean Schur decomposition. However, this would not work, at least not as directly as one might hope. Firstly, every nonsingular matrix has a hyperbolic QR factorization, as shown by [19, Theorem 5.3]. But, as shown in the examples in this paper, not all such matrices have a hyperbolic Schur decomposition (see Example 3.8 for $\lambda \neq 0$, Example 3.10 for $\lambda_1, \lambda_2 \neq 0$ and Example 3.13). This means that any algorithm employing the hyperbolic QR factorization to calculate the hyperbolic Schur decomposition will diverge for such matrices. Secondly, many singular matrices have a hyperbolic Schur decomposition, while it is unclear if they also have a hyperbolic QR factorization, since [19, Theorem 5.3] covers only the matrices A such that A^*JA is of full rank. This means that it may be possible for some matrices to have a hyperbolic Schur decomposition which is uncomputable via the hyperbolic QR factorization.

Acknowledgments

Most of the work on this paper was done at the School of Mathematics, University of Manchester, where I was invited as a research visitor by Françoise Tisseur, whom I thank dearly for the opportunity and a great working experience as well as many suggestions that helped me with this work. The paper was also proofread by Nataša Strabić whom I thank for all the suggestions that have considerably improved the paper. I would also like to thank the anonymous referee at the LAA, who made a special impact on this paper and without whose suggestions section 5 would not exist.

References

- [1] G. Ammar, C. Mehl, and V. Mehrmann. Schur-like forms for matrix Lie groups, Lie algebras and Jordan algebras. *Linear Algebra Appl.*, 287(1-3):11–39, 1999.
- [2] D. S. Bernstein. *Matrix Mathematics: Theory, Facts, and Formulas with Application to Linear Systems Theory*. Princeton University Press, Princeton, NJ, USA, 2005.
- [3] Y. Bolshakov, C. V. M. van der Mee, A. Ran, B. Reichstein, and L. Rodman. Extension of isometries in finite-dimensional indefinite scalar product spaces and polar decompositions. *SIAM J. Matrix Anal. Appl.*, 18(3):752–774, July 1997.

- [4] P. Davies and N. Higham. A Schur-Parlett algorithm for computing matrix functions. *SIAM J. Matrix Anal. Appl.*, 25(2):464–485, 2006.
- [5] I. Gohberg, P. Lancaster, and L. Rodman. *Matrix polynomials*. Classics in applied mathematics. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1982.
- [6] I. Gohberg, P. Lancaster, and L. Rodman. *Indefinite Linear Algebra and Applications*. Birkhäuser, Basel, Switzerland, 2005.
- [7] G. Golub and C. F. Van Loan. *Matrix computations*. 3rd ed. Johns Hopkins University Press, Baltimore, MD, USA, 1996.
- [8] E. Grimme, D. Sorensen, and P. Van Dooren. Model reduction of state space systems via an implicitly restarted Lanczos method. 12(1-2):1–31, 1996.
- [9] S. Hassi. *A Singular Value Decomposition of Matrices in a Space with an Indefinite Scalar Product*. Series A, I mathematica, dissertationes no. 79, Annales Academiæ Scientiarum Fennicæ, Helsinki, 1990.
- [10] N. J. Higham. J -orthogonal matrices: Properties and generation. *SIAM Rev.*, 45(3):504–519, 2003.
- [11] N. J. Higham. *Functions of matrices. Theory and computation*. Society for Industrial and Applied Mathematics, 2008.
- [12] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, UK, second edition, 2013.
- [13] A. Kılıçman and Z. A. Zhou. The representation and approximation for the weighted Minkowski inverse in Minkowski space. *Math. Comput. Modelling*, 47(3-4):363–371, 2008.
- [14] B. C. Levy. A note on the hyperbolic singular value decomposition. *Linear Algebra Appl.*, 277(1–3):135–142, 1998.
- [15] W.-W. Lin, V. Mehrmann, and H. Xu. Canonical forms for Hamiltonian and symplectic matrices and pencils. *Linear Algebra Appl.*, 302-303:469–533, 1999.
- [16] D. S. Mackey, N. Mackey, and F. Tisseur. Structured factorizations in scalar product spaces. *SIAM J. Matrix Anal. Appl.*, 27(3):821–850, 2006.
- [17] R. Onn, A. O. Steinhardt, and A. Bojanczyk. The hyperbolic singular value decomposition and applications. Applied mathematics and computing, Trans. 8th Army Conf., Ithaca/NY (USA) 1990, ARO Rep. 91–1, 93–108, 1991.
- [18] G. Sewell. *Computational Methods of Linear Algebra, Second Edition*. Wiley, 2005.

- [19] S. Singer. Indefinite QR factorization. *BIT*, 46(1):141–161, 2006.
- [20] V. Šego. Two-sided hyperbolic SVD. *Linear Algebra Appl.*, 433(7):1265–1275, 2010.
- [21] H. Xu. An SVD-like matrix decomposition and its applications. *Linear Algebra Appl.*, 368:1–24, 2003.
- [22] H. Xu. A numerical method for computing an SVD-like decomposition. *SIAM J. Matrix Anal. Appl.*, 26(4):1058–1082, 2005.
- [23] H. Zha. A note on the existence of the hyperbolic singular value decomposition. *Linear Algebra Appl.*, 240:199–205, 1996.