The hyperbolic Schur decomposition

Šego, Vedran

2012

MIMS EPrint: **2012.119**

Manchester Institute for Mathematical Sciences
School of Mathematics

The University of Manchester

The hyperbolic Schur decomposition

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

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 decomposition has 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 [10]). Its structure preserving property allows to save on 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.

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 $J = \operatorname{diag}(\pm 1)$ arises frequently in applications.

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

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 [9, 12, 13, 15, 20].

The Euclidean matrix decompositions have some nice structure preserving properties even in nonstandard scalar products, as shown by Mackey, Mackey and Tisseur [14], 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 [15, 21], J_1J_2-SVD [8], two-sided hyperbolic SVD [18], hyperbolic CS decomposition [7, 9] and indefinite QR factorization [17].

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 [17].

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 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.

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

2

some indices k) denote permutation matrices. We use

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

for the standard involutory permutation (see [5, 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. Simbol $\oplus$ 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., $J = \text{diag}(\pm 1) := \text{diag}(j_{11}, j_{22}, \ldots, j_{nn})$, where $j_{kk} \in \{-1, 1\}$ for all k, the scalar product is referred to as *hyperbolic* and takes the form

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

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 rank $A^* J A <$ 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

$$v' = \frac{1}{|v|}v = \frac{1}{\sqrt{|[v,v]|}}v \tag{1}$$

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-hyperexhange 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 [5].

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 [16, Section 3.3]). Similarly, f 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.

It is a well known fact that the first column u_1 of the similarity matrix U in any similarity triangularization $A = UTU^{-1}$ is the eigenvector of A. Also, it is easy to construct a nonsingular matrix with all degenerate columns (see [19, Example 3.1], which cannot be J-normalized. This means that there are some matrices (even diagonalizable ones!) that are not unitarily triangulizable with respect to the scalar product induced by J. However, allowing the blocks of T to be of order 2, we relax that condition, as shown in the Example 3.1.

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.

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 3.1. *Let* $J = \mathrm{diag}(1, 1, -1, -1)$ *and let* $A = S\mathcal{J}S^{-1}$, *where* $\mathcal{J} = \mathrm{diag}(1, 2, 3, 4)$ *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 $A = UTU^{-1}$. Since

$$
S^* J S = \begin{bmatrix}
-1 & -0.994393 & -0.970230 & -0.949852 \\
-0.994393 & -1 & -0.993828 & -0.985075 \\
-0.970230 & -0.993828 & -1 & -0.998151 \\
-0.949852 & -0.985075 & -0.998151 & -1
\end{bmatrix} \tag{2}
$$

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,

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

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

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

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

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

Now, we have:

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

So, v_1 is the eigenvector of

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

i.e.,

$$A_2 v_1 = \lambda' v_1, \quad \lambda' := \left(\frac{(t_{11} - t_{22})^2}{4} + t_{12}t_{21}\right). \tag{4}$$

But, since
$$(A - t_{22}\mathrm{I})(A - t_{11}\mathrm{I}) = (A - t_{11}\mathrm{I})(A - t_{22}\mathrm{I}),$$

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 $v_1 = \alpha_1 s_i + \beta_1 s_j$ and $v_2 = \alpha_2 s_i + \beta_2 s_j$. Obviously, $\alpha_1, \beta_1 \neq 0$ (or v_1 would be an eigenvector of A, hence covered by the case 1). Then, from $J = V^* J V$, we get:

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

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).

Knowing the obstacles, we are now ready to define the hyperbolic Schur decomposition.

Definition 3.2 (Hyperbolic Schur decomposition). *For a given $A \in \mathbb{C}^{n \times n}$ and $J = \operatorname{diag}(\pm 1)$, a hyperbolic Schur decomposition (with respect to J) is any decomposition of one of the following equivalent forms*

$$\begin{aligned} A &= (P_1 U P_2) T (P_1 U P_2)^{-1}, & U^* \tilde{J} U &= \tilde{J}, \quad \tilde{J} = P_1^* J P_1, \\ A &= (VP) T (VP)^{-1}, & V^* J V &= J, & (6) \\ A &= (PW) T (PW)^{-1}, & W^* J' W &= J', \quad J' = P^* J P, & (7) \\ A &= \widehat{V} T \widehat{V}^{-1}, & \widehat{V}^* J \widehat{V} &= P J P^*, & (8) \\ P^* A P &= W T W^{-1}, & W^* J' W &= J', \quad J' = P^* J P, \end{aligned}$$

such that T is quasitriangular.

Proof that the forms in Definition 3.2 are equivalent is pretty straightforward and can be seen in [19].

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

Proof. Let $A = S\mathcal{J}S^{-1}$, where $\mathcal{J}$ is a diagonal matrix, be a Jordan decomposition of A. Since S is nonsingular, and therefore $S^* J S$ is of full rank, by [17, Theorem 5.3], there exist permutations P_1 and P_2 such that

$$S = P_1 Q R P_2^*, \quad Q^* \tilde{J} Q = \tilde{J}, \quad \tilde{J} = P_1^* J P_1$$

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

$$A = S\mathcal{J}S^{-1} = P_1 Q R P_2^* \mathcal{J} P_2 R^{-1} Q^{-1} P_1^*. \tag{9}$$

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 (9), 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 (7) 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 3.3, 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 3.7.

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 3.4. *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 3.5. *Let $J = \operatorname{diag}(1, -1, 1, -1)$ 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 (8). 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 $\widehat{V}$ 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, $v_1 = xs_1$ 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 $i, j \in \{1, 2\}$. As shown in Example 3.1 (case 2), v_1 and v_2 are both (linearly independent) eigenvectors of A_2, defined in (3), with the same eigenvalue λ', defined in (4).

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 α_{11}, α_{12}, α_{21}, α_{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:

$$\begin{aligned} \pm 1 = [v_1, v_1] &= [\alpha_{11} s_1' + \alpha_{21} s_2', \alpha_{11} s_1' + \alpha_{21} s_2'] \\ &= |\alpha_{11}|^2 [s_1', s_1'] + 2\operatorname{Re}(\alpha_{11}\overline{\alpha_{21}}[s_1', s_2']) + |\alpha_{21}|^2 [s_2', s_2'] = 0. \end{aligned}$$

It is fairly easy to construct a matrix A of order n similar to the one in Example 3.5 for the Schur-like decompositions with bigger diagonal blocks of T. This means some matrices do not have a hyperbolic Schur-like decomposition, regardless of the order of the biggest diagonal block in T, as long as this order is strictly less than n. For a more detailed description, see [19].

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 3.6. *Let* $J = \mathrm{diag}(1, -1, 1, -1)$ *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* (8) *with* $\widehat{V} = [v_1 \ldots v_4]$ *hyperexchange and* T *quasitriangular and let* $s_i = Se_i$. *Using the same argumentation as in Example 3.5, 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 $A_2(\xi) := (A - \xi \mathrm{I})^2$ 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* $\xi = \lambda_2$.

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, very much like the common proof in Schur decomposition (see [11, Theorem 2.3.1], [6, Theorem 7.1.3] or [19]).

Proposition 3.7. *Let* $A \in \mathbb{C}^{n \times n}$ *have a Jordan decomposition* $A = S\mathcal{J}S^{-1}$ *such that $\mathcal{J}$ 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* $J = \mathrm{diag}(\pm 1)$.

Proof. If all Jordan blocks of A are of order 1, the matrix A is diagonalizable and, by Theorem 3.3, has a hyperbolic Schur decomposition. So, we shall only consider a case when A has (exactly one) Jordan block of order 2.

10

$\boxed{\text{Case 1}}$ 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 [5, page 10], v_1 can be expanded to the J-orthonormal basis $\{v_1, v_2, \ldots, v_n\}$. Defining a matrix

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

we see that

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

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 $P^* J P$. 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.

$\boxed{\text{Case 2}}$ 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, 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 $i \in \{1, 3, 4, \ldots, n\}$. Note the absence of the second column, as this one is not an eigenvector, but the generalized eigenvector associated with $\mathcal{J}_2(\lambda_1)$.

$\boxed{\text{Case 2.1}}$ 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',

11

using the fact that $[s_2, s_2] \in \mathbb{R}$, which is valid for all vectors in any indefinite scalar product space:

$$[v_1', v_1'] = [\xi s_1 + s_2, \xi s_1 + s_2] = |\xi|^2 [s_1, s_1] + 2 \operatorname{Re}(\xi [s_1, s_2]) + [s_2, s_2]$$
$$= -2 \operatorname{Re}[s_2, s_2] + [s_2, s_2] = -[s_2, s_2] \neq 0.$$

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 $\{v_1, v_2, \ldots, v_n\}$, define the J-orthonormal matrix V with columns $v_1, \ldots, v_n$ and, by construction, see that

$$A = V \left[\begin{array}{c|c} T_{11} & * \\ \hline 0 & A' \end{array} \right] V^{-1}, \quad V^* J V = P^* J P. \tag{10}$$

Here, T_{11} is of order 2 and the matrix A' is diagonalizable. Hence, by Theorem 3.3, A' has a hyperbolic Schur decomposition, so A has one too.

$\boxed{\text{Case 2.2}}$ 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

$$[v_1', v_2'] = [\xi s_1 - s_2, \xi s_1 + s_2] = |\xi|^2 [s_1, s_1] + 2 \operatorname{Im}(\xi [s_1, s_2]) + [s_2, s_2]$$
$$= 2 \operatorname{Im}(\xi [s_1, s_2]).$$

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:

$$[v_1', v_1'] = [\xi s_1 - s_2, \xi s_1 - s_2] = |\xi|^2 [s_1, s_1] - 2 \operatorname{Re}(\xi [s_1, s_2]) + [s_2, s_2]$$
$$= -2 |[s_1, s_2]|^2 \neq 0,$$
$$[v_2', v_2'] = [\xi s_1 + s_2, \xi s_1 + s_2] = |\xi|^2 [s_1, s_1] + 2 \operatorname{Re}(\xi [s_1, s_2]) + [s_2, s_2]$$
$$= 2 |[s_1, s_2]|^2 \neq 0.$$

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 $\{v_1, \ldots, v_n\}$ and a J-orthonormal matrix V. By construction, (10) holds and we can further decompose A', which is again diagonalizable.

$\boxed{\text{Case 2.3}}$ We have now covered all the cases such that $[s_1, s_2] \neq 0$. 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. Using the nonsingularity of S and J, we see that such k always exists.

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 (10), which we omit, as it is unimportant for the proof. $\square$

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 [5, Section 5.6] gives a detailed analysis of J-Hermitian matrices for the case when J (in [5] 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.,

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

However, Proposition 3.7 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 [5, 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 3.8 (J-Hermitian matrix that does not have a hyperbolic Schur decomposition)**.** *Let $J = \operatorname{diag}(1, 1, -1)$ 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 $A = VTV^{-1}$, $V^* J V = P^* J P$, with s_1 being the only eigenvector of A. Note that

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

which means that s_1 is degenerate and J-orthogonal to s_2. Following the previous discussions (see Example 3.5 or the proof of Proposition 3.7), 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.

13

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:

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

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

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

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. In [19] was described how Example 3.8 was constructed.

There also exist J-unitary matrices that do not have a hyperbolic Schur decomposition. One can be constructed in a manner simmilar to the Example 3.8.

Example 3.9 (J-unitary matrix without a hyperbolic Schur decomposition). Let $J = \operatorname{diag}(1, 1, -1)$ 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 3.8.

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 [5].

Theorem 3.10 (*J-nonnegative matrix*). *If $A \in \mathbb{C}^{n \times n}$ is J-nonnegative, then it has a hyperbolic Schur decomposition with respect to any given $J = \mathrm{diag}(\pm 1)$.*

Proof. Since JA is positive semidefinite, we can write $A = JB^*B$ for some B. Then the hyperbolic SVD of B, as described in [21, 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 [21]) is

$$
B = U \begin{bmatrix} \begin{bmatrix} \mathrm{I}_j & \mathrm{I}_j \end{bmatrix} & & \\ & \mathrm{diag}(|\lambda_1|^{1/2}, \ldots, |\lambda_l|^{1/2}) & \\ & & 0 \end{bmatrix} (JV)^*,
$$

where $U^*U = \mathrm{I}$, $(JV)^*J(JV) = V^*JV = P^*JP$, for some permutation P, and $j = \mathrm{rank}\, B - \mathrm{rank}\, BJB^*$. It is easy to see that

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

Note that P^*JP is diagonal and

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

is permutationally similar to

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

which completes the proof of the theorem. $\square$

Remark 3.11. *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 [18, Corollary 5.3]:

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

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. As it turns out, it can be reduced in a way that the 2×2 blocks are of a very specific structure.

Theorem 3.12 (The complete hyperbolic Schur decomposition). *Let $A \in \mathbb{C}^{n \times n}$ have a hyperbolic Schur decomposition for a given $J = \mathrm{diag}(\pm 1)$. Then it has a hyperbolic Schur decomposition* (6) *such that all the nontrivial diagonal 2×2 blocks of T have only degenerate eigenvectors with respect to the corresponding part of $\tilde{J}$.*

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., $J' = \mathrm{diag}(1, -1)$ or $J' = \mathrm{diag}(-1, 1)$). 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 [5, p. 10]). So, the only nonreducible 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 $J' = \pm \mathrm{diag}(1, -1)$. We see that T' must be in one of the following two forms:

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

$$(T')^* J' T' = S^{-*} \mathrm{diag}(\overline{\lambda}, 0) S^* J' S \, \mathrm{diag}(\lambda, 0) S^{-1}$$

$$= S^{-*} \mathrm{diag}(\overline{\lambda}, 0) \begin{bmatrix} 0 & \alpha \\ \overline{\alpha} & 0 \end{bmatrix} \mathrm{diag}(\lambda, 0) S^{-1} = 0, \text{ or}$$

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

$$(T')^* J' T' = S^{-*} \mathcal{J}_2(0)^T S^* J' S \mathcal{J}_2(0) S^{-1}$$

$$= S^{-*} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & \alpha \\ \overline{\alpha} & \beta \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} S^{-1} = 0.$$

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 nonreducible blocks (those that cannot be further reduced by the hyperbolic Schur decomposition) are sometimes referred to as *atomic* (see [4, page

466]) and Theorem 3.12 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 3.12 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 3.2. We also use V, P and $\tilde{J}$ from (6) and (8).

One of the main properties of the traditional Schur decomposition is that it keeps the structure of the matrix unchanged. Not surprisingly, similar properties hold in the hyperbolic case as well.

Proposition 4.1 (J-conjugate transpose). *Let A have a hyperbolic Schur decomposition (6) with respect to $J = \text{diag}(\pm 1)$. Then*

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

Proof. The proof is straightforward. From (6), it follows that

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

As we have seen in Example 3.5, 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

$$A = (A^*)^* = (A^{[*]_J})^{[*]_J}. \qquad \square$$

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 $J = \mathrm{diag}(\pm 1)$, then A is J-Hermitian if and only if T is $\tilde{J}$-Hermitian and quasidiagonal, where $\tilde{J} = P^* J P$.*

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 $\overline{\lambda}$, as shown in [5, Proposition 4.2.3.]. Also, by [5, Corollary 4.2.5], nonreal eigenvalues of J-Hermitian matrices have J-neutral root subspaces, which implies that all their adjoined 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 $J = \mathrm{diag}(\pm 1)$, then A is J-normal if and only if T is $\tilde{J}$-normal quasitriangular, where $\tilde{J} = P^* J P$.*

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 $J = \mathrm{diag}(1, -1, 1, -1)$ 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 comes from? In the Euclidean case, for an upper triangular normal A, we have:

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

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, (11) 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 $J = \operatorname{diag}(\pm 1)$. 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. $\qquad\square$

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 with respect to $J = \operatorname{diag}(\pm 1)$, then A is J-unitary if and only if 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 $\tilde{J}$-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. $\quad\square$

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 [14, 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 $J = \mathrm{diag}(\pm 1)$ and A of the same order be given such that A has a complete hyperbolic Schur decomposition* (6)*. 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 $\tilde{J}$-unitary or $\tilde{J}$-Hermitian, respectively. Since we have assumed a complete hyperbolic Schur decomposition, by Theorem 3.12 the eigenvector of every nondiagonalizable atomic block $\widehat{T}$ is $\widehat{J}$-degenerate, also implying that $\widehat{J} = \pm \mathrm{diag}(1, -1)$. This means that

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

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 $\tilde{J}$-unitary and since it is quasidiagonal, its block $\widehat{T}$ is $\widehat{J}$-unitary. So,

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

Premultiplying by S^* and postmultiplying S, we get

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

From here, we see that $2\lambda\,\mathrm{Re}(\alpha) = 0$. Since $\widehat{T}$ is $\widehat{J}$-unitary and, hence, nonsingular, $\lambda \neq 0$. Obviously, $\mathrm{Re}(\alpha) = 0$, 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 $\widehat{T}$ is $\widehat{J}$-Hermitian. From $\widehat{T}^{[*]} = \widehat{T}$ follows that $\widehat{J}\widehat{T}^* = \widehat{T}\widehat{J}$, so

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

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

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

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

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

Here, we see that $\lambda \in \mathbb{R}$ (which also follows straight from [14, Theorem 7.6]) and $\overline{\alpha} = \alpha$, i.e., α is real, so $[s_1, s_2] \in \mathbb{R}$. $\qquad\square$

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 $J = \operatorname{diag}(\pm 1)$ and let $A \in \mathbb{C}^{n \times n}$ be both J-Hermitian and J-unitary. Then A has a complete hyperbolic Schur decomposition* (6) *such that each diagonal atomic block $\widehat{T}$ of order 2 is of the form*

$$\widehat{T} = \begin{bmatrix} \alpha & \beta \\ -\overline{\beta} & -\alpha \end{bmatrix}, \quad \widehat{J} = \pm\operatorname{diag}(1, -1),$$

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

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

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

which means that A is involutory and, hence, diagonalizable (see [2, Fact 5.12.13]) which, by Theorem 3.3 means that it has a hyperbolic Schur decomposition. By Theorem 3.12, we can choose such a decomposition in a way that each diagonal block of order 2 in T, denoted $\widehat{T}$, has only degenerate eigenvectors with respect to $\widehat{J}$ (the corresponding part of $\tilde{J}$), which is possible if and only if that part is either $\widehat{J} = \operatorname{diag}(1, -1)$ or $\widehat{J} = \operatorname{diag}(-1, 1)$.

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

$$\mathrm{I}_2 = \widehat{T}^{[*]\widehat{J}}\widehat{T} = \widehat{T}^2. \tag{12}$$

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

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

Since $\widehat{T} = \widehat{T}^{[*]\widetilde{J}}$, 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 (12), 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. □

5. Conclusion

In this paper we have introduced the hyperbolic Schur decomposition of a square matrix with respect to the scalar product induced by $J = \mathrm{diag}(\pm 1)$. We have shown that all diagonalizable matrices have such a decomposition and have given examples of matrices that do not have it. Furthermore, we have shown how such examples can be constructed even if larger blocks in T are allowed.

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 $\widetilde{J}$, which is a common property of hyperbolic decompositions. We have analyzed the properties of such matrices, as well as the properties of atomic (nonreducible) blocks on the diagonal of the quasitriangular factor T.

In Example 3.8 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 3.9 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).

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 3.8.

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 or 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 [17, Theorem 5.3]. But, as shown in the examples in this paper, not all such matrices have a hyperbolic Schur decomposition (see Example 3.5 for $\lambda \neq 0$, Example 3.6 for $\lambda_1, \lambda_2 \neq 0$ and Example 3.9). 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 [17, 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

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. *Indefinite Linear Algebra and Applications.* Birkhäuser, Basel, Switzerland, 2005.

[6] G. Golub and C. F. Van Loan. *Matrix computations. 3rd ed.* Johns Hopkins University Press, Baltimore, MD, USA, 1996.

[7] 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.

[8] 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.

[9] N. J. Higham. *J*-orthogonal matrices: Properties and generation. *SIAM Rev.*, 45(3):504–519, 2003.

[10] N. J. Higham. *Functions of matrices. Theory and computation.* Society for Industrial and Applied Mathematics, 2008.

[11] R. A. Horn and C. R. Johnson. *Matrix Analysis.* Cambridge University Press, Cambridge, UK, 1985.

[12] A. Kılıçman and Z. A. Zhour. The representation and approximation for the weighted Minkowski inverse in Minkowski space. *Math. Comput. Modelling*, 47(3-4):363–371, 2008.

[13] B. C. Levy. A note on the hyperbolic singular value decomposition. *Linear Algebra Appl.*, 277(1–3):135–142, 1998.

[14] D. Mackey, N. Mackey, and F. c. Tisseur. Structured factorizations in scalar product spaces. *SIAM J. Matrix Anal. Appl.*, 27(3):821–850, 2006.

[15] 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.

[16] G. Sewell. *Computational Methods of Linear Algebra, Second Edition.* Wiley, 2005.

[17] S. Singer. Indefinite QR factorization. *BIT*, 46(1):141–161, 2006.

[18] V. Šego. Two-sided hyperbolic SVD. *Linear Algebra Appl.*, 433(7):1265–1275, 2010.

[19] V. Šego. The hyperbolic schur decomposition (extended). MIMS EPrint 2012.116, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, Dec 2012.

[20] H. Xu. An SVD-like matrix decomposition and its applications. *Linear Algebra Appl.*, 368:1–24, 2003.

[21] H. Zha. A note on the existence of the hyperbolic singular value decomposition. *Linear Algebra Appl.*, 240:199–205, 1996.