On the Existence of Schur-like Forms for Matrices with Symmetry Structures

Christian Mehl[1]

Abstract

Schur-like forms are developed for matrices that have a symmetry structure with respect to an indefinite inner product induced by a Hermitian and unitary Gram matrix. It is characterized under which conditions these forms can be computed by structure-preserving unitary transformations. The main results combines and generalizes the two well-known results from the literature that on the one hand any normal matrix can be unitarily diagonalized and on the other hand a Hamiltonian matrix can be transformed to Hamiltonian Schur form via a unitary similarity transformation if and only if its purely imaginary eigenvalues satisfy certain conditions that involve the sign characteristic of the matrix under consideration.

1 Introduction

It is well known that Hermitian matrices, skew-Hermitian matrices, and unitary matrices are unitarily diagonalizable. More generally, this is true for normal matrices, i.e., for matrices $A \in \mathbb{C}^{n,n}$ satisfying $AA^H = A^H A$ - a class of matrices that generalizes the classes of Hermitian, skew-Hermitian, and unitary matrices. The fact that unitary matrices can be used as transformation matrices for diagonalizing normal matrices is important because of two fundamental properties: firstly, similarity transformations with unitary matrices preserve the given Hermitian, skew-Hermitian, unitary, or normal structure of the matrix - a fact that it crucial for the development of structure-preserving algorithms for the solution of the given eigenvalue problem - and secondly, unitary matrices are optimally conditioned. Therefore, the normal eigenvalue problem can be considered to be a well-behaved problem

---

✉ Christian Mehl
mehl@math.tu-berlin.de

[1] Institut für Mathematik, Technische Universität Berlin, MA 4-5, 10623 Berlin, Germany

in the sense that a diagonalization can be performed by backward stable and structure-preserving algorithms.

The picture changes drastically, if one considers matrices that carry a symmetry structure with respect to an indefinite inner product, i.e., with respect to a nondegenerate Hermitian form that is not necessarily positive definite or with respect to a nondegenerate skew-Hermitian form. An important example are *Hamiltonian matrices*, i.e., matrices $A \in \mathbb{R}^{2n,2n}$ satisfying $A^\top J + JA = 0$, where J denotes the skew-symmetric matrix

$$J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}. \tag{1}$$

The identity $A^\top J = -JA$ can be interpreted as skew-symmetry of the matrix A with respect to the skew-symmetric inner product induced by J. The corresponding *Hamiltonian eigenvalue problem* arises in many application, e.g., in system theory and the theory of Algebraic Riccati equations, see [1, 5, 11] and the references therein. For practical reasons, one often switches to the complex version of the problem which leads to the consideration of matrices $A \in \mathbb{C}^{2n,2n}$ satisfying $A^H J + JA = 0$. Typically, these matrices are also called *Hamiltonian* in the Numerical Linear Algebra community and we will follow this convention in this paper. It should be noted though that other communities prefer the terminology *J-skew-adjoint* for such matrices (e.g., see [2]) in order to avoid confusion with complex matrices $A \in \mathbb{C}^{2n,2n}$ satisfying $A^\top J + JA = 0$ which are called Hamiltonian as well.

For the solution of the Hamiltonian eigenvalue problem, the so-called *Hamiltonian Schur form* was suggested in [12] as a target form. This is a Hamiltonian matrix of the block form

$$\begin{bmatrix} R & B \\ 0 & -R^H \end{bmatrix}, \tag{2}$$

where $R \in \mathbb{C}^{n,n}$ is upper triangular. It is straightforward to see that this is just a permutation of the classical upper triangular Schur form of a complex matrix and as a consequence the eigenvalues can be read off from the diagonal. It was shown in [12] that this form can always be achieved under a unitary symplectic similarity transformation provided that the given Hamiltonian matrices does not have eigenvalues on the imaginary axis. Recall that a matrix $S \in \mathbb{C}^{2n,2n}$ is called symplectic (following the convention in the Numerical Linear Algebra community) if $S^H J S = J$. It is easily checked that similarity transformations with symplectic matrices preserve the Hamiltonian structure and are therefore the base of structure-preserving algorithms for the solution of the Hamiltonian eigenvalue problem. However, since the condition number of symplectic matrices may be arbitrarily large, it is favorable to further restrict oneself to the class of unitary symplectic similarity transformations in order to guarantee numerical stability.

Surprisingly, there are Hamiltonian matrices that cannot be transformed to Hamiltonian Schur form via a unitary symplectic similarity transformation. As an obvious example, consider the matrix J which is both Hamiltonian and symplectic (and even unitary). Clearly, if $U \in \mathbb{C}^{2n,2n}$ is unitary and symplectic then $U^{-1}JU = U^H JU = J$, so J is invariant under any unitary symplectic similarity transformations and hence cannot be transformed to Hamiltonian Schur form. It is clear from the form in (2) that a necessary condition for the existence of a Hamiltonian Schur form is that the purely imaginary eigenvalues of the given matrix have even multiplicity. Indeed, if λ is an eigenvalue of R then $-\bar{\lambda}$ is an eigenvalue of $-R^H$, so any purely imaginary eigenvalue of R will also be one of $-R^H$. The example of J

however shows that this condition is not sufficient. The long open problem of characterizing all Hamiltonian matrices that can be transformed to Hamiltonian Schur form was finally solved in [6] with the help of a newly developed structured canonical form for Hamiltonian matrices.

At first sight, one may come to the conclusion that the nonexistence of the Hamiltonian Schur form is related to the fact that in contrast to the Euclidean inner product the two fundamental properties of structure-preservation and numerical stability are now partitioned among two sets of transformation matrices instead of only one. Thus, to enable both features, one has to restrict oneself to the set of unitary symplectic matrices which is a much "smaller" subset (in terms of dimension as a manifold) than the sets of symplectic matrices or unitary matrices. This conclusion, however, turns out to be wrong as it is well-known that the Hamiltonian Schur form exists under unitary symplectic similarity transformations if and only if it exists under similarity transformations that are symplectic only. This equivalence can easily be shown by applying a structure-preserving QR decomposition to the transformation matrix, see, e.g., [6] for details. Thus, both in the case of normal matrices with respect to the Euclidean inner product and in the case of Hamiltonian matrices, the actual problem is the computation of a Schur-like form by similarity transformations from the group of matrices that are unitary with respect to the considered inner product and therefore preserve the given symmetry structure of the matrix they are acting on.

We will show in this paper that the diagonal Schur form of normal matrices and the Hamiltonian Schur form of Hamiltonian matrices are two extreme cases of a much more general Schur-like form for matrices carrying a symmetry structure with respect to an indefinite inner product. To treat the problem in full generality, we will consider a generalization of normality of matrices in an indefinite inner product space, the so-called *polynomial H-normality* which will be introduced in Section 2 where we will also review the basic theory of indefinite inner products. In Section 3 we will formulate the main result of this paper and develop a Schur-like form for polynomially H-normal matrices for the case of a Hermitian and unitary Gram matrix H. Then we will discuss how the Schur form of normal matrices and the Hamiltonian Schur form can be deduced as special cases of this result. The proof of the main result will then be given in Section 4 followed by a short summary in Section 5.

Notation By $\mathcal{J}_n(\lambda)$ we denote the $n \times n$ upper triangular Jordan block associated with the eigenvalue λ. The reverse identity of size n is denote by R_n, i.e.,

$$
R_n = \begin{bmatrix} 0 & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & 0 \end{bmatrix}.
$$

If $H \in \mathbb{C}^{n,n}$ is a Hermitian matrix, then its inertia index is denoted by (π, ν, ζ), where π, ν, and ζ are the numbers of positive, negative and zero eigenvalues, respectively, each counted with algebraic multiplicities.

2 Indefinite Inner Products and Polynomially *H*-normal Matrices

Let $H \in \mathbb{C}^{n,n}$ be Hermitian and invertible. Then H defines an indefinite inner product on $\mathbb{C}^n$ via

$$
[x, y] := x^H H y, \quad x, y \in \mathbb{C}^n.
$$

As in the case of positive definite inner product, we will call H the *Gram matrix* of the indefinite inner product. If $A \in \mathbb{C}^{n,n}$ then the matrix $A^{[*]} := H^{-1}A^H H$ is called the *H-adjoint* of A, because it is the unique matrix satisfying the identity $[Ax, y] = [x, A^{[*]}y]$ for all $x, y \in \mathbb{C}^n$. The matrix A is called

- H-selfadjoint, if $A^{[*]} = A$, or, equivalently, $A^H H = HA$,
- H-skew-adjoint, if $A^{[*]} = -A$, or, equivalently, $A^H H + HA = 0$,
- and H-unitary, if $A^{[*]} = A^{-1}$, or, equivalently, $A^H HA = H$.

In the following, we will restrict ourselves to Hermitian inner products, because any skew-Hermitian inner product can easily be transformed to a Hermitian inner product by multiplying the corresponding Gram matrix with the imaginary unit i. In particular, a matrix $A \in \mathbb{C}^{2n,2n}$ is Hamiltonian if and only if A is (iJ)-skew-adjoint, where J is the matrix as in (1).

Canonical forms for H-selfadjoint, H-skew-adjoint, and H-unitary matrices are well known, see, e.g., [2, 8]. More generally, one can define the set of H-normal matrices as the set of all matrices $A \in \mathbb{C}^{n,n}$ satisfying $A^{[*]}A = AA^{[*]}$. Unfortunately, this set turns out to be "too big", because it was shown in [3] that the problem of classifying H-normal matrices is a wild problem and hence canonical forms cannot be obtained. Therefore, it was suggested in [10] to consider the set of polynomially H-normal matrices instead. A matrix $A \in \mathbb{C}^{n,n}$ is called *polynomially H-normal* if there exists a polynomial p in one variable such that $A^{[*]} = p(A)$. The polynomial p is then called the *H-normality polynomial* of A. It is easily checked that any polynomially H-normal matrix is H-normal, but the converse is not true, see [10]. Still, the set of polynomially H-normal matrices is large enough to contain the sets of H-selfadjoint, H-skew-adjoint, and H-unitary matrices. Indeed, if the matrix A is H-selfadjoint or H-skew-adjoint, then it is polynomially H-normal with H-normality polynomial $p(t) = t$ or $p(t) = -t$, respectively, and since $A^{[*]} = A^{-1}$ and the inverse of a matrix is always a polynomial in that matrix, it follows that also any H-unitary matrix is polynomially H-normal.

The major advantage of polynomially H-normal matrices over H-normal matrices is the fact that a complete classification is available under the following equivalence relation.

Remark 1 Let $H \in \mathbb{C}^{n,n}$ be Hermitian and invertible and let $A \in \mathbb{C}^{n,n}$ be polynomially H-normal with H-normality polynomial p. If $T \in \mathbb{C}^{n,n}$ is invertible, then $T^{-1}AT$ is polynomially $T^H HT$-normal with H-normality polynomial p. In particular, the relation

$$(A_1, H_1) \sim (A_2, H_2) : \quad \Leftrightarrow \quad \exists T \in GL_n(\mathbb{C}) : A_2 = T^{-1}A_1T \wedge H_2 = T^H H_1 T \quad (3)$$

is an equivalence relation on the set of pairs (A, H), where H is Hermitian and invertible and A is polynomially H-normal with H-normality polynomial p.

The following canonical form for polynomially H-normal matrices was developed in [8, Theorem 6.1]. Here, $T(\alpha_1, \ldots, \alpha_n)$ denotes an upper triangular $n \times n$ Toeplitz matrix that has $\begin{bmatrix} \alpha_1 & \ldots & \alpha_n \end{bmatrix}$ as its first row.

Theorem 1 (Canonical form for polynomially H-normal matrices) *Let the matrix $A \in \mathbb{C}^{n,n}$ be polynomially H-normal with H-normality polynomial p. Then there exists a nonsingular matrix T such that*

$$T^{-1}AT = A_1 \oplus \cdots \oplus A_q, \quad T^*HT = H_1 \oplus \cdots \oplus H_q, \quad (4)$$

where A_j is H_j-indecomposable and where A_j and H_j have one of the following forms:

i) *blocks associated with eigenvalues $\lambda_j \in \mathbb{C}$ satisfying $p(\lambda_j) = \overline{\lambda_j}$:*

$$A_j = \lambda_j I_{n_j} + e^{i\theta_j} T(0, 1, ir_{j,2}, \ldots, ir_{j,n_j-1}), \quad H_j = \sigma_j R_{n_j}, \tag{5}$$

where $n_j \in \mathbb{N}$, $\sigma_j \in \{1, -1\}$, $\theta_j \in [0, \pi)$, and $r_{j,2}, \ldots, r_{j,n_j-1} \in \mathbb{R}$;

ii) *blocks associated with a pair (λ_j, μ_j) of eigenvalues with $\mu_j = \overline{p(\lambda_j)} \neq \lambda_j$, $\overline{p(\mu_j)} = \lambda_j$, and $Re(\lambda_j) > Re(\mu_j)$ or $Im(\lambda_j) > Im(\mu_j)$ if $Re(\lambda_j) = Re(\mu_j)$:*

$$A_j = \begin{bmatrix} \mathcal{J}_{m_j}(\lambda_j) & 0 \\ 0 & p\left(\mathcal{J}_{m_j}(\lambda_j)\right)^H \end{bmatrix}, \quad H_j = \begin{bmatrix} 0 & I_{m_j} \\ I_{m_j} & 0 \end{bmatrix}, \tag{6}$$

where $m_j \in \mathbb{N}$.

Moreover, the form (4) is unique up to the permutation of blocks, and the parameters θ_j, and $r_{j,2}, \ldots, r_{j,n_j-1}$ in (5) are uniquely determined by λ_j and the coefficients of p, and they can be computed from the identity

$$\overline{\lambda_j} I_{n_j} + e^{-i\theta_j} T(0, 1, -ir_{j,2}, \ldots, -ir_{j,n_j-1}) = p\left(\lambda_j I_{n_j} + e^{i\theta_j} T(0, 1, ir_{j,2}, \ldots, ir_{j,n_j-1})\right).$$

(We highlight that the eigenvalues λ_j in i) are not necessarily pairwise distinct, i.e., the same eigenvalue may occur in different blocks. The same is true for the eigenvalues λ_j, μ_j in ii).)

Besides the eigenvalues and their partial multiplicities, the signs $\sigma_j = \pm 1$ attached to each Jordan block corresponding to an eigenvalue λ_j satisfying $\overline{\lambda_j} = p(\lambda_j)$ are additional invariants under the equivalence relation (3). The list of all signs associated with a fixed eigenvalue λ_j is referred to as the *sign characteristic* of the eigenvalue λ_j extending the terminology in [2] used for H-selfadjoint and H-unitary matrices.

The following values related to the sign characteristic of a fixed eigenvalue will play a crucial role in the characterization when a structured Schur-like form will exist.

Definition 1 Let $H \in \mathbb{C}^{n,n}$ be Hermitian and invertible, let $A \in \mathbb{C}^{n,n}$ be polynomially H-normal, and let $\lambda \in \mathbb{C}$ be an eigenvalue of A that satisfies $\overline{\lambda} = p(\lambda)$. Then the sum of all signs σ_j from the sign characteristic of λ_j attached to blocks of odd size is called the *sign sum* of λ_j and is denoted by $\mathrm{signsum}(\lambda_j)$.

To illustrate Definition 1 consider the matrices

$$A = \mathcal{J}_5(0) \oplus \mathcal{J}_4(0) \oplus \mathcal{J}_3(0) \oplus \mathcal{J}_3(0) \oplus \mathcal{J}_3(0) \oplus \mathcal{J}_2(0) \oplus \mathcal{J}_1(0),$$
$$H = \sigma_1 R_5 \oplus \sigma_2 R_4 \oplus \sigma_3 R_3 \oplus \sigma_4 R_3 \oplus \sigma_5 R_3 \oplus \sigma_6 R_2 \oplus \sigma_7 R_1$$

with $\sigma_i \in \{1, -1\}$ for $i = 1, \ldots, 7$. Then A is polynomially H-normal with H-normality polynomial $p(t) = t$ (in fact, A is H-selfadjoint), and $\overline{0} = p(0)$. The sign sum of the eigenvalue $\lambda = 0$ is then given by

$$\mathrm{signsum}(0) = \sigma_1 + \sigma_3 + \sigma_4 + \sigma_5 + \sigma_7.$$

Note that in accordance with Definition 1 the values σ_2 and σ_6 do not contribute to the sign sum as they are attached to blocks of the even sizes 4 and 2, respectively.

The signsum has an important impact on the inertia index of the given Hermitian matrix defining the indefinite inner product as we will show in the following lemma.

Lemma 1 *Let $H \in \mathbb{C}^{n,n}$ be Hermitian with inertia index $(\pi, \nu, 0)$ and let $A \in \mathbb{C}^{n,n}$ be polynomially H normal. If $\lambda_1, \ldots, \lambda_r \in \mathbb{C}$ are the pairwise distinct eigenvalues of A satisfying $\overline{\lambda_j} = p(\lambda_j)$, then*

$$\pi - \nu = \sum_{j=1}^{r} \text{signsum}(\lambda_j).$$

Proof The proof immediately follows by inspection from the canonical form given in Theorem 1. Indeed, one easily checks that the matrices H_j from blocks of type ii) and blocks of type i) corresponding to an even size n_j contribute equally to the positive and negative eigenvalues of H. On the other hand, the matrix $H_j = \sigma_j R_{n_j}$ of a block as in type i) that corresponds to an odd size $n_j = 2k + 1$ has inertia index $(k + 1, k, 0)$ if $\sigma_j = 1$, and inertia index $(k, k + 1, 0)$ if $\sigma_j = -1$. □

Finally, we recall the concept of neutral subspaces in indefinite inner product spaces. If $H \in \mathbb{C}^{n,n}$ is Hermitian and invertible, then a subspace $\mathcal{V} \subseteq \mathbb{C}^n$ is called *H-neutral* if $[v, w] = 0$ (or, equivalently, $v^H H w = 0$) for all $v, w \in \mathcal{V}$. It is well-known that if $(\pi, \nu, 0)$ is the inertia index of H, then the maximal possible dimension of an H-neutral subspace is equal to $k = \min \frac{n - |\pi - \nu|}{2} = \min\{\pi, \nu\}$. An H-neutral subspace of this maximal dimension k is called a *maximal H-neutral subspace*.

3 Schur-like Forms and Invariant Maximal *H*-neutral Subspaces

In this section, we will develop the main result of this paper: the introduction of a structured Schur-like form for polynomially H-normal matrices combined with a characterization of its existence. As pointed out in the introduction, it was shown in [6] that a Hamiltonian matrix can be transformed to Hamiltonian Schur form via a unitary symplectic similarity transformation if and only if the same can be done via a similarity transformation that is only symplectic. An analysis of the corresponding proof in [6] reveals that the property of the matrix J in (1) to be unitary is a crucial fact in this equivalence. Therefore, we will assume throughout this section that the Gram matrix of the given indefinite inner product is not only Hermitian, but also unitary. Many important examples of Gram matrices such as

$$R_n, \quad \begin{bmatrix} 0 & I_n \\ I_n & 0 \end{bmatrix}, \quad \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix}$$

satisfy this extra condition. (We mention in passing that properties of inner products that are either Hermitian or unitary are discussed in [7].)

Theorem 2 *Let $H \in \mathbb{C}^{n,n}$ be unitary and Hermitian with inertia index $(\pi, \nu, 0)$ and let $A \in \mathbb{C}^{n,n}$ be polynomially H-normal with H-normality polynomial p. Furthermore, let $m := |\pi - \nu|$. Then $n - m$ is even and the following statements are equivalent, where $k := \frac{n-m}{2}$.*

1) *There exists an H-neutral subspace of dimension k that is A-invariant.*
2) *There exists a unitary matrix $U \in \mathbb{C}^{n,n}$ such that*

$$U^{-1}AU = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ 0 & p(B_{11})^H & 0 \\ 0 & B_{32} & B_{33} \end{bmatrix} \quad and \quad U^H H U = \begin{bmatrix} 0 & I_k & 0 \\ I_k & 0 & 0 \\ 0 & 0 & sI_m \end{bmatrix}, \quad (7)$$

where $s = \frac{\pi - \nu}{|\pi - \nu|}$ (if $m \neq 0$ and thus $\pi - \nu \neq 0$, else let $s := 1$), $B_{11} \in \mathbb{C}^{k,k}$ is upper triangular and $B_{33} \in \mathbb{C}^{m,m}$ is diagonal.

3) There exists an invertible matrix $S \in \mathbb{C}^{n,n}$ such that

$$S^{-1} A S = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ 0 & p(B_{11})^H & 0 \\ 0 & B_{32} & B_{33} \end{bmatrix} \quad and \quad S^H H S = \begin{bmatrix} 0 & I_k & 0 \\ I_k & 0 & 0 \\ 0 & 0 & s I_m \end{bmatrix}, \quad (8)$$

where $s = \frac{\pi - \nu}{|\pi - \nu|}$ (if $m \neq 0$ and thus $\pi - \nu \neq 0$, else let $s := 1$), $B_{11} \in \mathbb{C}^{k,k}$ is upper triangular and $B_{33} \in \mathbb{C}^{m,m}$ is diagonal.

4) Let $\lambda_1, \ldots, \lambda_r \in \mathbb{C}$ be the eigenvalues of A satisfying the equation $\bar{\lambda} = p(\lambda)$. Then $s \cdot \mathrm{signsum}(\lambda_i) \geq 0$ for $i = 1, \ldots, r$ and

$$\sum_{i=1}^{r} \mathrm{signsum}(\lambda_i) = sm.$$

Proof The proof of Theorem 2 is rather long and will therefore be presented in a separate section. $\qquad\square$

As mentioned in the introduction, Theorem 2 combines and generalizes two important results from the literature that we will restate below as corollaries. The first results recovers the well-known result on unitary diagonalizability of (I_n-)normal matrices.

Corollary 1 (Schur-form of normal matrices) *Let $A \in \mathbb{C}^{n,n}$ be a normal matrix, i.e., $A^H A = A A^H$. Then A is unitarily diagonalizable.*

Proof By [4] normality with respect to the Euclidean inner product is equivalent to polynomially I_n-normality and hence Theorem 2 can be applied with $H = I_n$. Since I_n has the inertia index $(n, 0, 0)$, we find that $k = 0$. Thus, condition 2) of Theorem 2 states the existence of a unitary matrix $U \in \mathbb{C}^{n,n}$ such that $U^{-1} A U = B_{33}$ is diagonal. $\qquad\square$

Corollary 2 (Hamiltonian Schur-form of Hamiltonian matrices) *Let $A \in \mathbb{C}^{2n,2n}$ be a Hamiltonian matrix, i.e., $A^H J + J A = 0$. Then the following statements are equivalent:*

1) *There exists an n-dimensional subspace of $\mathbb{C}^{2n}$ that is J-neutral and A-invariant.*
2) *There exists a unitary symplectic matrix $Q \in \mathbb{C}^{2n,2n}$ such that $Q^{-1} A Q$ is in Hamiltonian Schur form, i.e.,*

$$Q^{-1} A Q = \begin{bmatrix} B & C \\ 0 & -B^H \end{bmatrix},$$

where $B \in \mathbb{C}^{n,n}$ is upper triangular.
3) *There exists a symplectic matrix $S \in \mathbb{C}^{2n,2n}$ such that $S^{-1} A S$ is in Hamiltonian Schur form, i.e.,*

$$S^{-1} A S = \begin{bmatrix} B & C \\ 0 & -B^H \end{bmatrix},$$

where $B \in \mathbb{C}^{n,n}$ is upper triangular.
4) *For any purely imaginary eigenvalue λ of A, the number of odd partial multiplicities corresponding to λ with sign $+1$ is equal to the number of partial multiplicities corresponding to λ with sign -1.*

Proof First, we recall that A is Hamiltonian if and only if A is H-skew-adjoint with $H = iJ$. In particular, A is polynomially H-normal with H-normality polynomial $p(t) = -t$. We will frequently make use of this fact in the following.

"1) $\Rightarrow$ 2)": It is trivial that the J-neutral subspace in 1) is also iJ-neutral. Moreover, the inertia index of iJ is $(n, n, 0)$ and thus, Theorem 2 implies the existence of a unitary matrix $U \in \mathbb{C}^{2n,2n}$ such that

$$U^{-1}AU = \begin{bmatrix} B & C \\ 0 & -B^H \end{bmatrix} \quad \text{and} \quad U^H(iJ)U = \begin{bmatrix} 0 & I_n \\ I_n & 0 \end{bmatrix}.$$

Setting $Q := U \cdot \operatorname{diag}(I_n, iI_n)$, we obtain that

$$Q^{-1}AQ = \begin{bmatrix} B & iC \\ 0 & -B^H \end{bmatrix} \quad \text{and} \quad Q^H J Q = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix} = J,$$

i.e., Q is unitary and symplectic. This implies 2).

"2) $\Rightarrow$ 3)" is trivial and "3) $\Rightarrow$ 4)" and "4) $\Rightarrow$ 1)" follow immediately from Theorem 2 taking into account that the eigenvalues satisfying $\bar{\lambda} = p(\lambda) = -\lambda$ are exactly the purely imaginary eigenvalues of A, and that signsum$(\lambda) = 0$ is equivalent to the statement that the number of odd partial multiplicities corresponding to λ with sign $+1$ is equal to the number of partial multiplicities corresponding to λ with sign -1. $\qquad\square$

The equivalence of 2), 3) and 4) was proved in [6] while the equivalence of 1) and 4) was proved in [13]. Clearly, the equivalence of 1) and 2) - or 1) and 3) - immediately follows from those two results in the literature and since then this has been implicitly known by many researchers dealing with Hamiltonian matrices. Nevertheless, it seems that a theorem combining all four equivalent conditions into a single result was explicitly formulated only as late as in [9].

Remark 2 The two results in Corollaries 1 and 2 represent the two extreme cases $k = 0$ and $m = 0$ in Theorem 2. Whenever $k, m \neq 0$ as it would be the case for Gram matrices of the form $\operatorname{diag}(I_p, -I_q)$ with $p \neq q$, then the corresponding Schur-like form will have triangular and diagonal parts on the block diagonal as indicated in (7).

We highlight that the transformation in Theorem 2 is a transformation that changes the inner product, but keeps the symmetry structure of the matrix $U^{-1}AU$ linked to the transformed Gram matrix $U^H HU$ in the sense of Remark 1. For the practical use of Theorem 2, we advise to first transform the pair (A, H) to a form (A', H'), where H' already has the form

$$H' = \begin{bmatrix} 0 & I_k & 0 \\ I_k & 0 & 0 \\ 0 & 0 & sI_m \end{bmatrix}$$

with $s \in \{\pm 1\}$. When Theorem 2 is then applied to the pair (A', H') it yields the existence of a unitary matrix U such that $U^{-1}A'U$ is in the Schur-like form as in (7) while $U^H H'U = H'$. The latter conditions just means that the matrix U is not only unitary, but also H'-unitary.

4 Proof of the Main Result

Before we prove Theorem 2, we start with a technical lemma that will be used frequently in the following.

Lemma 2 *Let A_1, $H_1 \in \mathbb{C}^{n_1,n_1}$ and A_2, $H_2 \in \mathbb{C}^{n_2,n_2}$, where H_1 and H_2 are Hermitian and invertible, and let*

$$A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}, \quad H = \begin{bmatrix} H_1 & 0 \\ 0 & H_2 \end{bmatrix}.$$

If A_1 has an invariant H_1-neutral subspace of dimension k_1 and A_2 has an invariant H_2-neutral subspace of dimension k_2, then A has an invariant H-neutral subspace of dimension $k_1 + k_2$.

Proof Let the vectors $v_1, \ldots, v_{k_1} \in \mathbb{C}^{n_1}$ and $w_1, \ldots, w_{k_2} \in \mathbb{C}^{n_2}$ form bases of the A_i-invariant H_i-neutral subspaces for $i = 1, 2$, respectively. Then it is straightforward to verify that the vectors

$$\begin{bmatrix} v_1 \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} v_{k_1} \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ w_1 \end{bmatrix}, \ldots, \begin{bmatrix} 0 \\ w_{k_2} \end{bmatrix} \in \mathbb{C}^{n_1+n_2}$$

form a basis of an A-invariant H-neutral subspace which obviously has dimension $k_1 + k_2$. □

Proof of Theorem 2 "1) $\Rightarrow$ 2)": By switching to an orthonormal basis whose first k columns span the A-invariant H-neutral subspace, we can assume that A and H have the forms

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \quad \text{and} \quad H = \begin{bmatrix} 0 & H_{12} \\ H_{12}^H & H_{22} \end{bmatrix},$$

where $A_{22}, H_{22} \in \mathbb{C}^{n-k,n-k}$. Since H is unitary, the rows of $H_{12} \in \mathbb{C}^{k,n-k}$ are orthonormal and consequently its singular value decomposition takes the form

$$H_{12} = U_2 \begin{bmatrix} I_k & 0 \end{bmatrix} V_2,$$

where $U_2 \in \mathbb{C}^{k,k}$ and $V_2 \in \mathbb{C}^{n-k,n-k}$ are unitary. Setting $Q := \mathrm{diag}(U_2, V_2)$ we obtain

$$Q^{-1}AQ = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ 0 & B_{22} & B_{23} \\ 0 & B_{32} & B_{33} \end{bmatrix} \quad \text{and} \quad Q^H H Q = \begin{bmatrix} 0 & I_k & 0 \\ I_k & 0 & 0 \\ 0 & 0 & H_{33} \end{bmatrix},$$

where B_{11}, B_{22} are $k \times k$. The zeros in the block positions $(2, 2)$, $(2, 3)$, and $(3, 2)$ are due to the fact that $Q^H H Q$ is still unitary and consequently has orthonormal columns. On the other hand, $Q^H H Q$ is also Hermitian and keeping in mind that its inertia index is $(\pi, \nu, 0)$ and we have $k = \min\{\pi, \nu\}$, the block structure implies that H_{33} is positive or negative definite, depending on $\pi - \nu$ being positive or negative, respectively. (If $\pi = \nu$, then $n = 2k$ and the $(3, 3)$-blocks in $Q^{-1}AQ$ and $Q^H H Q$ are void.) Since H_{33} is also unitary, it follows that $H_{33} = I_m$ or $H_{33} = -I_m$. Applying a unitary transformation of the form $\mathrm{diag}(U_1, U_1, U_3)$ with $U_1 \in \mathbb{C}^{k,k}$, $U_3 \in \mathbb{C}^{m,m}$ if necessary, we can assume without loss of generality that B_{11} and B_{33} are upper triangular. (Note that this transformation will not change $Q^H H Q$.) Finally, we exploit the fact that A is polynomially H-normal with H-normality-polynomial p. This implies

$$\begin{bmatrix} B_{22}^H & B_{12}^H & B_{32}^H \\ 0 & B_{11}^H & 0 \\ B_{23}^H & B_{13}^H & B_{33}^H \end{bmatrix} = Q^H H^{-1} A^H H Q = p(Q^{-1}AQ) = \begin{bmatrix} p(B_{11}) & & * \\ 0 & p\left(\begin{bmatrix} B_{22} & B_{23} \\ B_{32} & B_{33} \end{bmatrix} \right) \end{bmatrix}$$

and hence $B_{22} = p(B_{11})^H$ and $B_{23} = 0$. But then, we find that $B_{33}^H = p(B_{33})$ which implies that B_{33} is normal (i.e., with respect to the Euclidean inner product) and hence diagonal as upper triangular normal matrices are diagonal.

"2) $\Rightarrow$ 3)": This is trivial.

"3) $\Rightarrow$ 4)": The proof proceeds by induction on the number r of eigenvalues satisfying $\bar{\lambda} = p(\lambda)$. If $r = 0$, then the block B_{33} in (8) is void, because it is diagonal and satisfies $B_{33}^H = p(B_{33})$, so all its eigenvalues satisfy $\bar{\lambda} = p(\lambda)$. Thus, we have $m = 0$ and 4) is satisfied by the definition of the empty sum.

Now assume that $r > 0$. Let $\lambda := \lambda_1$, i.e., we have $\bar{\lambda} = p(\lambda)$. Starting from 3) we can assume without loss of generality that A and H are in the form (8), where the eigenvalues on the diagonal of B_{11} and B_{33} are ordered in such a way that all occurrences of λ come first, i.e., A has the form

$$
A = \begin{array}{c} \\ k_1 \\ k_2 \\ k_1 \\ k_2 \\ m_1 \\ m_2 \end{array}
\begin{array}{c} k_1 \quad k_2 \quad\quad k_1 \quad\quad\quad k_2 \quad\quad m_1 \quad m_2 \end{array}
\left[\begin{array}{cccccc}
A_{11} & A_{12} & A_{13} & A_{14} & A_{15} & A_{16} \\
0 & A_{22} & A_{23} & A_{24} & A_{25} & A_{26} \\
0 & 0 & p(A_{11})^H & 0 & 0 & 0 \\
0 & 0 & A_{43} & p(A_{22})^H & 0 & 0 \\
0 & 0 & A_{53} & A_{54} & A_{55} & 0 \\
0 & 0 & A_{63} & A_{64} & 0 & A_{66}
\end{array} \right],
$$

where $k_1, m_1 \geq 0$ and $\sigma(A_{11}), \sigma(p(A_{11})^H), \sigma(A_{55}) \subseteq \{\lambda\}$ and λ is not contained in $\sigma(A_{22}), \sigma(p(A_{22})^H)$, or $\sigma(A_{66})$. If $k_1 > 0$, then letting $X \in \mathbb{C}^{k_1, k_2}$ be the unique solution of the Sylvester equation $X A_{22} - A_{11} X = A_{12}$ (which exists, because the spectra of A_{11} and A_{22} are disjoint) and defining $\widehat{S}$ as the matrix that differs from the block-partitioned identity $\mathcal{I} := I_{k_1} \oplus I_{k_2} \oplus I_{k_1} \oplus I_{k_2} \oplus I_{m_1} \oplus I_{m_2}$ by adding X in the $(1, 2)$-block position, we obtain that a similarity transformation on A with $\widehat{S}$ annihilates the block entry A_{12} of A. The corresponding congruence transformation creates the entry X in the $(3, 2)$-position and the entry X^H in the $(2, 3)$-block-position of H. We can restore H with the congruence transformation with the matrix T which differs from the block-partitioned identity $\mathcal{I}$ by putting $-X^H$ in the $(4, 3)$-block-position. The corresponding similarity applied to $\widehat{S}^{-1} A \widehat{S}$ will change the block-entries $A_{13}, A_{23}, A_{43}, A_{53}, A_{63}$, but not the yet established zero in the $(1, 2)$-block position. For the ease of notation let us rename $T^{-1} \widehat{S}^{-1} A \widehat{S} T$ by A and $T^H \widehat{S}^H H \widehat{S} T$ by H. (Similar renaming steps will occur after each of the following steps without further notice.) We illustrate the A_{12}-elimination-step in the following diagram, where non-zero block-entries, that were effected by the current transformation are marked as bullets. In each substep, the pair (i, j) denotes the block-entry of the transformation matrix that differs from the one in the identity matrix. Observe that the similarity transformation with such a matrix adds the i-th block column to the j-th block column, but the j-th block row to the i-th block row while in the corresponding congruence transformation the i-th block column and i-th block row are added to the j-th block column or j-th block row, respectively:

$$
A : \left[\begin{array}{cccccc}
* & * & * & * & * & * \\
0 & * & * & * & * & * \\
0 & 0 & * & 0 & 0 & 0 \\
0 & 0 & * & * & 0 & 0 \\
0 & 0 & * & * & * & 0 \\
0 & 0 & * & * & 0 & *
\end{array} \right]
\xrightarrow{(1,2)}
\left[\begin{array}{cccccc}
* & 0 & \bullet & \bullet & \bullet & \bullet \\
0 & * & * & * & * & * \\
0 & 0 & * & 0 & 0 & 0 \\
0 & 0 & * & * & 0 & 0 \\
0 & 0 & * & * & * & 0 \\
0 & 0 & * & * & 0 & *
\end{array} \right]
\xrightarrow{(4,3)}
\left[\begin{array}{cccccc}
* & 0 & \bullet & * & * & * \\
0 & * & \bullet & * & * & * \\
0 & 0 & * & 0 & 0 & 0 \\
0 & 0 & \bullet & * & 0 & 0 \\
0 & 0 & \bullet & * & * & 0 \\
0 & 0 & \bullet & * & 0 & *
\end{array} \right],
$$

$$
H: \begin{bmatrix} 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & * & 0 & 0 \\ * & 0 & 0 & 0 & 0 & 0 \\ 0 & * & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & * \end{bmatrix} \xrightarrow{(1,2)} \begin{bmatrix} 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & \bullet & * & 0 & 0 \\ * & \bullet & 0 & 0 & 0 & 0 \\ 0 & * & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & * \end{bmatrix} \xrightarrow{(4,3)} \begin{bmatrix} 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & * & 0 & 0 \\ * & 0 & 0 & 0 & 0 & 0 \\ 0 & * & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & * \end{bmatrix}.
$$

In the next step, we eliminate A_{16} by applying a similarity transformation on A that is obtained from $\mathcal{I}$ by changing the $(1, 6)$-block to the solution X of the Sylvester equation $XA_{66} - A_{11}X = A_{16}$. The corresponding congruence transformation on H introduces the matrix X in the $(3, 6)$-block-position and X^H in the $(6, 3)$-block-position. We can restore H as follows: first, we apply a congruence with a matrix that differs from $\mathcal{I}$ by $-sX^H$ in the $(6, 3)$-block-position. This will annihilate the $(3, 6)$- and $(6, 3)$-block entries, but introduce the block $-sXX^H$ in the $(3, 3)$-block entry. The corresponding similarity transformation on A only effects the block-entries A_{23} and A_{26}. Then, we eliminate the $(3, 3)$-block entry of H by a congruence transformation with the matrix that coincides with $\mathcal{I}$ except for having $-\frac{1}{2}X^H X$ as its $(1, 3)$-block. The corresponding similarity transformation on A only changes the block A_{13}. As before, we illustrate this elimination step with the help of a diagram:

$$
A: \begin{bmatrix} * & 0 & * & * & * & * \\ 0 & * & * & * & * & * \\ 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & * & * & 0 & 0 \\ 0 & 0 & * & * & * & 0 \\ 0 & 0 & * & * & 0 & * \end{bmatrix} \xrightarrow{(1,6)} \begin{bmatrix} * & 0 & \bullet & \bullet & * & 0 \\ 0 & * & * & * & * & * \\ 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & * & * & 0 & 0 \\ 0 & 0 & * & * & * & 0 \\ 0 & 0 & * & * & 0 & * \end{bmatrix} \xrightarrow{(6,3)} \begin{bmatrix} * & 0 & * & * & * & 0 \\ 0 & * & \bullet & * & * & * \\ 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & * & * & 0 & 0 \\ 0 & 0 & * & * & * & 0 \\ 0 & 0 & \bullet & * & 0 & * \end{bmatrix} \xrightarrow{(1,3)} \begin{bmatrix} * & 0 & \bullet & * & * & 0 \\ 0 & * & * & * & * & * \\ 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & * & * & 0 & 0 \\ 0 & 0 & * & * & * & 0 \\ 0 & 0 & * & * & 0 & * \end{bmatrix},
$$

$$
H: \begin{bmatrix} 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & * & 0 & 0 \\ * & 0 & 0 & 0 & 0 & 0 \\ 0 & * & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & * \end{bmatrix} \xrightarrow{(1,6)} \begin{bmatrix} 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & * & 0 & 0 \\ * & 0 & 0 & 0 & 0 & \bullet \\ 0 & * & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & \bullet & 0 & 0 & * \end{bmatrix} \xrightarrow{(6,3)} \begin{bmatrix} 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & * & 0 & 0 \\ * & 0 & \bullet & 0 & 0 & 0 \\ 0 & * & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & * \end{bmatrix} \xrightarrow{(1,3)} \begin{bmatrix} 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & * & 0 & 0 \\ * & 0 & 0 & 0 & 0 & 0 \\ 0 & * & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & * \end{bmatrix}.
$$

We continue by eliminating A_{54} with a similarity transformation that differs from $\mathcal{I}$ by the solution of the Sylvester equation $Xp(A_{22})^H - A_{55}X = A_{54}$ in the $(5, 4)$-block-position followed by transformations that restore H. Since the step is analogous to the A_{12}-elimination step, we restrict ourselves to the illustration via a diagram:

$$
A: \begin{bmatrix} * & 0 & * & * & * & 0 \\ 0 & * & * & * & * & * \\ 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & * & * & 0 & 0 \\ 0 & 0 & * & * & * & 0 \\ 0 & 0 & * & * & 0 & * \end{bmatrix} \xrightarrow{(5,4)} \begin{bmatrix} * & 0 & * & \bullet & * & 0 \\ 0 & * & * & \bullet & * & * \\ 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & * & * & 0 & 0 \\ 0 & 0 & \bullet & 0 & * & 0 \\ 0 & 0 & * & * & 0 & * \end{bmatrix} \xrightarrow{(2,5)} \begin{bmatrix} * & 0 & * & * & * & 0 \\ 0 & * & \bullet & * & \bullet & * \\ 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & * & * & 0 & 0 \\ 0 & 0 & * & 0 & * & 0 \\ 0 & 0 & * & * & 0 & * \end{bmatrix},
$$

$$
H: \begin{bmatrix} 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & * & 0 & 0 \\ * & 0 & 0 & 0 & 0 & 0 \\ 0 & * & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & * \end{bmatrix} \xrightarrow{(5,4)} \begin{bmatrix} 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & * & 0 & 0 \\ * & 0 & 0 & 0 & 0 & 0 \\ 0 & * & 0 & 0 & \bullet & 0 \\ 0 & 0 & 0 & \bullet & * & 0 \\ 0 & 0 & 0 & 0 & 0 & * \end{bmatrix} \xrightarrow{(2,5)} \begin{bmatrix} 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & * & 0 & 0 \\ * & 0 & 0 & 0 & 0 & 0 \\ 0 & * & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & * \end{bmatrix}.
$$

Finally, we eliminate A_{14} with a similarity transformation whose relevant entry, the solution of the Sylvester equation $Xp(A_{22})^H - A_{11}X = A_{14}$ is in the $(1, 4)$-block position.

Restoring H imposes another similarity transformation on A with the relevant entry in the $(2, 3)$-block position which effects the entry A_{23} only. This step is depicted as follows:

$$A: \begin{bmatrix} * & 0 & * & * & * & 0 \\ 0 & * & * & * & * & * \\ 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & * & * & 0 & 0 \\ 0 & 0 & * & 0 & * & 0 \\ 0 & 0 & * & * & 0 & * \end{bmatrix} \overset{(1,4)}{\rightsquigarrow} \begin{bmatrix} * & 0 & \bullet & 0 & * & 0 \\ 0 & * & * & * & * & * \\ 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & * & * & 0 & 0 \\ 0 & 0 & * & 0 & * & 0 \\ 0 & 0 & * & * & 0 & * \end{bmatrix} \overset{(2,3)}{\rightsquigarrow} \begin{bmatrix} * & 0 & * & 0 & * & 0 \\ 0 & * & \bullet & * & * & * \\ 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & * & * & 0 & 0 \\ 0 & 0 & * & 0 & * & 0 \\ 0 & 0 & * & * & 0 & * \end{bmatrix},$$

$$H: \begin{bmatrix} 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & * & 0 & 0 \\ * & 0 & 0 & 0 & 0 & 0 \\ 0 & * & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & * \end{bmatrix} \overset{(1,4)}{\rightsquigarrow} \begin{bmatrix} 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & * & 0 & 0 \\ * & 0 & 0 & \bullet & 0 & 0 \\ 0 & * & \bullet & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & * \end{bmatrix} \overset{(2,3)}{\rightsquigarrow} \begin{bmatrix} 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & * & 0 & 0 \\ * & 0 & 0 & 0 & 0 & 0 \\ 0 & * & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & * \end{bmatrix}.$$

The key observation is now that the block zero pattern obtained for A is invariant under multiplication and thus, all powers of A as well as $p(A)$ will have the same block zero pattern. But then, the fact that A is polynomially H-normal with H-normality polynomial p implies

$$\begin{bmatrix} p(A_{11}) & A_{43}^H & A_{13}^H & A_{23}^H & A_{53}^H & A_{63}^H \\ 0 & p(A_{22}) & 0 & A_{24}^H & 0 & A_{64}^H \\ 0 & 0 & A_{11}^H & 0 & 0 & 0 \\ 0 & 0 & 0 & A_{22}^H & 0 & 0 \\ 0 & 0 & A_{15}^H & A_{25}^H & A_{55}^H & 0 \\ 0 & 0 & 0 & A_{26}^H & 0 & A_{66}^H \end{bmatrix} = H^{-1}A^H H = p(A) = \begin{bmatrix} * & 0 & * & 0 & * & 0 \\ 0 & * & * & * & * & * \\ 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & * & * & 0 & 0 \\ 0 & 0 & * & 0 & * & 0 \\ 0 & 0 & * & * & 0 & * \end{bmatrix},$$

and we obtain that $A_{23} = 0$, $A_{25} = 0$, $A_{43} = 0$, and $A_{63} = 0$. But then, after applying a block permutation, we may assume that A and H have the forms $A = A_1 \oplus A_2$ and $H = H_1 \oplus H_2$ with

$$A_1 = \begin{bmatrix} A_{11} & A_{13} & A_{15} \\ 0 & p(A_{11})^H & 0 \\ 0 & A_{53} & A_{55} \end{bmatrix}, \quad H_1 = \begin{bmatrix} 0 & I_{k_1} & 0 \\ I_{k_1} & 0 & 0 \\ 0 & 0 & s I_{m_1} \end{bmatrix}$$

and

$$A_2 = \begin{bmatrix} A_{22} & A_{24} & A_{26} \\ 0 & p(A_{22})^H & 0 \\ 0 & A_{64} & A_{66} \end{bmatrix}, \quad H_2 = \begin{bmatrix} 0 & I_{k_2} & 0 \\ I_{k_2} & 0 & 0 \\ 0 & 0 & s I_{m_2} \end{bmatrix}.$$

Note that A_1 has only one eigenvalue which is $\lambda_1 = \lambda$. It immediately follows from Lemma 1 that $m_1 s = \mathrm{signsum}(\lambda_1)$ and $s \cdot \mathrm{signsum}\,\lambda_1 = m_1 \geq 0$. On the other hand, the matrix A_2 now has precisely $r - 1$ eigenvalues $\lambda_2, \ldots, \lambda_r$ satisfying $\overline{\lambda_i} = p(\lambda_i)$ and we can apply the induction hypothesis to obtain that $s \cdot \mathrm{signsum}(\lambda_i) \geq 0$ for $i = 2, \ldots, r$ and

$$sm = sm_1 + sm_2 = \mathrm{signsum}(\lambda_1) + \sum_{j=2}^{r} \mathrm{signsum}(\lambda_j)$$

which proves 4).

"4) $\Rightarrow$ 1)": We may assume without loss of generality that A and H are in the forms of Theorem 1. Thus, by Lemma 2, we may consider some of the corresponding diagonal blocks of A and H separately in order to construct an A-invariant H-neutral subspace. We will do this by individually investigating each block in the canonical form of Theorem 1 that

is of type ii) or of type i) having even dimension. Concerning the blocks of type i) having odd dimension, we will have to consider all of them together to obtain the desired dimension of the A-invariant H-neutral subspace. We proceed by first considering the following three special cases, before discussing the general case.

Special Case 1: A is a block of type ii).   In this case, we have $m = 0$, $k = \frac{n}{2}$ and

$$A = \begin{bmatrix} \mathcal{J}_k(\lambda) & 0 \\ 0 & p(\mathcal{J}_k(\lambda))^H \end{bmatrix}, \quad H = \begin{bmatrix} 0 & I_k \\ I_k & 0 \end{bmatrix},$$

where $p(\lambda) \neq \bar{\lambda}$. Obviously, the first k standard basis vectors of $\mathbb{C}^n$ span a k-dimensional A-invariant subspace that is H-neutral.

Special Case 2: A is a block of type i) of even dimension.   In this case, we again have $m = 0$, $k = \frac{n}{2}$ and

$$A = \lambda I_n + e^{i\theta} T(0, 1, ir_2, \ldots, ir_{n-1}), \quad H = \sigma R_n,$$

where $p(\lambda) = \bar{\lambda}$, $\sigma \in \{1, -1\}$, and where $\theta, r_2, \ldots, r_{n-1}$ are as specified in Theorem 1. Again, it is obvious that the first k standard basis vectors of $\mathbb{C}^n$ span a k-dimensional A-invariant subspace that is H-neutral.

Special Case 3: A only consists of blocks of type i) with odd dimension.   In this case, we have

$$A = (\lambda I_{n_1} + e^{i\theta_1} T(0, 1, ir_{1,2}, \ldots, ir_{1,n_1-1})) \oplus \cdots \oplus (\lambda I_{n_\ell} + e^{i\theta_\ell} T(0, 1, ir_{\ell,2}, \ldots, ir_{\ell,n_\ell-1}))$$

$$\text{and} \quad H = \sigma_1 R_{n_1} \oplus \cdots \oplus \sigma_\ell R_{n_\ell},$$

where $n_j = 2k_j + 1$ for some nonnegative integers $k_1, \ldots, k_\ell$. By 4), we have $|\text{signsum}(\lambda)| = m$. Without loss of generality we may assume that $s = 1$ considering $-H$ instead of H otherwise, i.e., we have that $\text{signsum}(\lambda) = m$. It follows that $\ell \geq m$. More precisely, there exists a nonnegative integer α such that $\ell = m + 2\alpha$ and such that $m + \alpha$ signs among $\sigma_1, \ldots, \sigma_\ell$ are positive and α are negative. Without loss of generality, we may assume that among the diagonal blocks the first 2α blocks have alternating signs starting with $\sigma_1 = 1$ and thus the last m blocks all have positive sign. Let us first consider a group of two blocks with signs $+1, -1$, say

$$A_j = \begin{bmatrix} T_{2j-1} & 0 \\ 0 & T_{2j} \end{bmatrix}, \quad H_j = \begin{bmatrix} R_{n_{2j-1}} & 0 \\ 0 & -R_{n_{2j}} \end{bmatrix},$$

where $j \in \{1, \ldots, \alpha\}$ and where $T_i \in \mathbb{C}^{n_i, n_i}$ is an upper triangular matrix having λI_{n_j} as its diagonal. Then it is easy to check that the vectors

$$e_1, \ldots e_{k_{2j-1}}, e_{n_{2j-1}+1}, \ldots, e_{n_{2j-1}+k_{2j}}, e_{k_{2j-1}+1} + e_{n_{2j-1}+k_{2j}+1} \tag{9}$$

form a basis of an A_j-invariant H_j-neutral subspace. Indeed, partitioning

$$T_{2j-1} = \begin{matrix} k_{2j-1} \\ 1 \\ k_{2j-1} \end{matrix} \begin{bmatrix} \overset{k_{2j-1}}{T_{11}} & \overset{1}{T_{12}} & \overset{k_{2j-1}}{T_{13}} \\ 0 & \lambda & T_{23} \\ 0 & 0 & T_{33} \end{bmatrix} \quad \text{and} \quad T_{2j} = \begin{matrix} k_{2j} \\ 1 \\ k_{2j} \end{matrix} \begin{bmatrix} \overset{k_{2j}}{\tilde{T}_{11}} & \overset{1}{\tilde{T}_{12}} & \overset{k_{2j}}{\tilde{T}_{13}} \\ 0 & \lambda & \tilde{T}_{23} \\ 0 & 0 & \tilde{T}_{33} \end{bmatrix}$$

a simultaneous block permutation on A_j and H_j results in

$$
\begin{bmatrix}
T_{11} & 0 & T_{12} & 0 & T_{13} & 0 \\
0 & \widetilde{T}_{11} & 0 & \widetilde{T}_{12} & 0 & \widetilde{T}_{13} \\
0 & 0 & \lambda & 0 & T_{23} & 0 \\
0 & 0 & 0 & \lambda & 0 & \widetilde{T}_{23} \\
0 & 0 & 0 & 0 & T_{33} & 0 \\
0 & 0 & 0 & 0 & 0 & \widetilde{T}_{33}
\end{bmatrix}
\quad \text{and} \quad
\begin{bmatrix}
0 & 0 & 0 & 0 & R_{k_{2j-1}} & 0 \\
0 & 0 & 0 & 0 & 0 & -R_{k_{2j-1}} \\
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & -1 & 0 & 0 \\
R_{k_{2j-1}} & 0 & 0 & 0 & 0 & 0 \\
0 & -R_{k_{2j-1}} & 0 & 0 & 0 & 0
\end{bmatrix}.
$$

Then transforming the middle 2×2 blocks with the transformation given by the matrix

$$
\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}
$$

produces matrices of the form

$$
\begin{bmatrix} \widehat{T}_1 & \widehat{T}_2 \\ 0 & \widehat{T}_3 \end{bmatrix}
\quad \text{and} \quad
\begin{bmatrix} 0 & \widehat{H} \\ \widehat{H}^H & 0 \end{bmatrix},
\tag{10}
$$

where all blocks have size $\frac{n_{2j-1}+n_{2j}}{2}$. Observe that the first $\frac{n_{2j-1}+n_{2j}}{2}$ columns of the transformation matrix that transform (A_j, H_j) to the pair of the matrices in (10) coincide with the vectors in (9), possibly up to scalar multiples.

Next, we consider a block for $j \in \{2\alpha + 1, \ldots, \ell\}$, i.e.,

$$
A_j = \lambda I_{n_j} + e^{i\theta_j} T(0, 1, ir_{j,2}, \ldots, ir_{1,n_j-1}), \quad H_j = R_{n_j}.
$$

Here, the first k_j standard basis vectors span an A_j-invariant H-neutral subspace.

In view of Lemma 2, we obtain the existence of an A-invariant H-neutral subspace of dimension

$$
\sum_{j=1}^{\alpha} (k_{2j-1} + k_{2j} + 1) + \sum_{j=2\alpha+1}^{\ell} k_j = \sum_{j=1}^{\alpha} \frac{n_{2j-1} + n_{2j}}{2} + \sum_{j=2\alpha+1}^{\ell} \frac{n_j - 1}{2} = \frac{n-m}{2},
$$

where we used that $\ell - 2\alpha = m$.

The general case: Now let A be general having r eigenvalues $\lambda_1, \ldots, \lambda_r \in \mathbb{C}$ satisfying $\overline{\lambda_i} = p(\lambda_i)$, $i = 1, \ldots, r$. Putting together the special cases above in view of our observation, we obtain that there exists an A-invariant H-neutral subspace of dimension

$$
\frac{1}{2}\left(n - \sum_{j=1}^{r} |\text{signsum}(\lambda_i)| \right),
$$

and by 4) this dimension is equal to $\frac{n-m}{2} = k$ as desired. This finishes the proof. $\qquad \square$

5 Conclusions

We have developed a Schur-like form for polynomially H-normal matrices, where H is Hermitian and unitary and characterized under which conditions these forms can be obtained via structure preserving (and unitary) similarity transformations. In particular, the result can be applied to all matrices that are selfadjoint, skew-adjoint, or unitary with respect to an indefinite inner product that has a unitary Hermitian Gram matrix. As two extreme special cases, the unitary diagonalizability of normal matrices and equivalent conditions for the existence of the Hamiltonian Schur form of Hamiltonian matrices have been recovered. While structure-preserving and numerically backward stable algorithms for the numerical

computation of these forms are well known in the mentioned two special cases, it remains to develop such algorithms for the general case.

References

1. Benner, P., Kressner, D., Mehrmann, V.: Skew-Hamiltonian and Hamiltonian eigenvalue problems: theory, algorithms and applications. In: Drmač, Z., Marušić, M., Tutek, Z. (eds.) Proceedings of the Conference on Applied Mathematics and Scientific Computing, pp. 3–39. Springer, Dordrecht (2005)
2. Gohberg, I., Lancaster, P., Rodman, L.: Indefinite Linear Algebra and Applications. Birkhäuser, Basel (2005)
3. Gohberg, I., Reichstein, B.: On classification of normal matrices in an indefinite scalar product. Integral Equ. Oper. Theory **13**, 364–394 (1990)
4. Grone, R., Johnson, C., Sa, E., Wolkowicz, H.: Normal matrices. Linear Algebra Appl. **87**, 213–225 (1987)
5. Lancaster, P., Rodman, L.: Algebraic Riccati Equations. Clarendon Press, Oxford (1995)
6. Lin, W.-W., Mehrmann, V., Xu, H.: Canonical forms for Hamiltonian and symplectic matrices and pencils. Linear Algebra Appl. **302–303**, 469–533 (1999)
7. Mackey, D.S., Mackey, N., Tisseur, F.: On the definition of two natural classes of scalar product MIMS EPrint 2007.64, The University of Manchester, UK (2007)
8. Mehl, C.: On classification of normal matrices in indefinite inner product spaces. Electron. J. Linear Algebra **15**, 50–83 (2006)
9. Mehl, C.: Finite-dimensional inner product spaces and applications in numerical analysis. In: Alpay, D. (ed.) Operator Theory, pp. 431–449. Springer Basel, Basel (2015)
10. Mehl, C., Rodman, L.: Classes of normal matrices in indefinite inner products. Linear Algebra Appl. **336**, 71–98 (2001)
11. Mehrmann, V.: The Autonomous Linear Quadratic Control Problem: Theory and Numerical Solution. Lecture Notes in Control and Information Sciences, vol. 163. Springer, Berlin (1991)
12. Paige, C., Van Loan, C.: A Schur decomposition for Hamiltonian matrices. Linear Algebra Appl. **41**, 11–32 (1981)
13. Ran, A.C.M., Rodman, L.: Stability of invariant maximal semidefinite subspaces I. Linear Algebra Appl. **62**, 51–86 (1984)