

An extended Hessenberg form for Hamiltonian matrices

Micol Ferranti *Bruno Iannazzo* *Thomas Mach*
Raf Vandebril

Report TW 665, November 2015



KU Leuven
Department of Computer Science
Celestijnenlaan 200A – B-3001 Heverlee (Belgium)

An extended Hessenberg form for Hamiltonian matrices

Micol Ferranti *Bruno Iannazzo* *Thomas Mach*
Raf Vandebril

Report TW 665, November 2015

Department of Computer Science, KU Leuven

Abstract

A unitary symplectic similarity transformation for certain Hamiltonian matrices to extended Hamiltonian Hessenberg form is presented. Whereas the classical Hessenberg form links to Krylov subspaces, the extended Hessenberg form links to extended Krylov subspaces. The presented algorithm generalizes thus the classic reduction to Hamiltonian Hessenberg form and offers more freedom in the choice of extended Hamiltonian forms, to be used within an extended Hamiltonian QR algorithm. Theoretical results identifying the structure of the extended Hamiltonian Hessenberg form and proofs of uniqueness of the reduction process are included. Numerical experiments confirm the validity of the approach.

Keywords : Hamiltonian eigenvalue problems, extended Hessenberg forms, QR algorithm

MSC : Primary : 65F15, Secondary : 15A18, 15A23, 93B60.

An extended Hessenberg form for Hamiltonian matrices*

Micol Ferranti[‡] Bruno Iannazzo[§] Thomas Mach[‡] Raf Vandebril[‡]

1 Introduction

The QR eigenvalue method is one of the most used algorithms to compute eigenvalues of medium sized matrices [8, 9]. It is a two step method: the matrix is first transformed, via unitary similarity transformations, to a convenient condensed form, whose eigenvalues are then, in the second step, computed via a suitable iterative method. This paper deals with the first step of a Hamiltonian QR method, i.e., the reduction to a convenient condensed Hamiltonian form. The second step, the actual QR algorithm, is discussed in the forthcoming paper [6].

The typical generic condensed form is the well known Hessenberg form [17]. However, if the original matrix exhibits particular properties, the preservation of the structure through the whole procedure is advisable, e.g. in model order reduction [7]. For this reason, some special formulations of the QR method have been developed for specific classes of matrices, i.e., symmetric, unitary, etc. On the other hand, it has been shown that the classic Hessenberg form is not the only possible choice as an intermediate condensed form [15]. A much wider family of matrices can be used within a QR-like algorithm. This family generalizes classes such as Hessenberg and Hessenberg-like, and these matrices are called extended Hessenberg matrices.¹

A Hamiltonian matrix is of the form

$$H = \begin{bmatrix} A & G \\ F & -A^H \end{bmatrix} \in \mathbb{C}^{2n \times 2n}, \quad (1)$$

where $F = F^H$ and $G = G^H$. The eigenvalues of a Hamiltonian matrix are symmetric with respect to the imaginary axis. To preserve this symmetry within a QR method, it is desirable to work exclusively with Hamiltonian matrices during the iterative process. An algorithm with this feature is said to be a Hamiltonian QR algorithm.

Unfortunately, the generic classical Hessenberg form does not retain the Hamiltonian structure. Thus, many attempts have been made to design a suitable Hamiltonian condensed form, see, for example, [4, 5, 13]. Until now, a Hamiltonian QR algorithm has only been found for matrices H , partitioned as in (1), whose bottom-left block F has rank one [5]. This algorithm makes use of the so called Hamiltonian Hessenberg form.

In this paper we will show how the arguments presented in [15] can be adapted to the Hamiltonian context, in order to derive new Hamiltonian condensed forms, which extend the classical Hamiltonian Hessenberg form, for matrices whose bottom-left block has rank one.

[‡]Department of Computer Science, KU Leuven, Celestijnenlaan 200A, 3001 Leuven (Heverlee), Belgium; (`{micol.ferranti, thomas.mach, raf.vandebril}@cs.kuleuven.be`).

[§]Dipartimento di Matematica e Informatica, Università di Perugia, Via Vanvitelli 1, 06123 Perugia, Italy; (`bruno.iannazzo@dmf.unipg.it`).

*The research was partially supported by the Research Council KU Leuven, projects CREA-13-012 Can Unconventional Eigenvalue Algorithms Supersede the State of the Art, OT/11/055 Spectral Properties of Perturbed Normal Matrices and their Applications, and CoE EF/05/006 Optimization in Engineering (OPTEC); by the Fund for Scientific Research–Flanders (Belgium) project G034212N Reestablishing Smoothness for Matrix Manifold Optimization via Resolution of Singularities; and by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office, Belgian Network DYSCO (Dynamical Systems, Control, and Optimization). The second author was partly supported by INdAM through GNCS Project 2015.

¹The name extended refers to the link with extended Krylov subspaces, see [12] for details.

The paper is organized as follows. In Section 2 and 3 we present some preliminary definitions and results, which are necessary in the remainder of the text. A brief summary of the techniques introduced in [15] is included. In Section 4 several theoretical results are presented, and the definition of extended Hamiltonian Hessenberg form is given. In Section 5 the transformation to extended Hamiltonian Hessenberg form is described. Also, the essential uniqueness of the transformation is proved. Section 6 is devoted to numerical experiments. Conclusions are presented in Section 7.

The following notation is used throughout this paper: the element of M in position (i, j) is denoted by $m_{i,j}$, its i -th column is indicated by m_i , and we depict with $M(i : j, k : \ell)$ the submatrix of M consisting of rows i up to j , and columns k up to ℓ . M^T refers to the transpose of M , \overline{M} to the conjugate of M , and $M^H = \overline{M}^T$ to its Hermitian transpose. A matrix M is said to be unitary when $MM^H = M^H M = I$, and Hermitian when $M = M^H$. A per-Hermitian matrix $M \in \mathbb{C}^{n \times n}$ is a matrix that is Hermitian with respect to its anti-diagonal, i.e., $m_{n-j+1, n-i+1} = \overline{m}_{ij}$.

2 The extended Hessenberg form

Here, we recall some material regarding the extended Hessenberg form, which is a generalization of the classic Hessenberg form introduced in the paper [15], to which we refer the reader for a full detailed treatise.

The central idea consists of computing a QR factorization $M = QR$, where R is upper triangular and the unitary factor Q is factored in a sequence of $n - 1$ special unitary transformations following a prescribed ordering.

Before introducing the extended Hamiltonian Hessenberg form, we will briefly describe and analyze the fundamental objects, namely unitary core transformations, on which the whole reasoning is based.

Definition 1. *A unitary core transformation is a matrix $G \in \mathbb{C}^{n \times n}$ that is equal to the identity, except for a unitary 2-by-2 block $G_i = G(i : i + 1, i : i + 1)$, which is called active part.*

All the core transformations used in this paper are unitary, thus we will omit the adjective unitary from now on. One special class of core transformations are rotations, whose active part is

$$\begin{bmatrix} c & s \\ -\overline{s} & \overline{c} \end{bmatrix},$$

with $|c|^2 + |s|^2 = 1$. All reasonings presented here are valid for any core transformation, but rotations were used to implement the algorithm. We will use the same notation as in [15] to depict core transformations in a figurative way: the symbol $\overleftrightarrow{\lrcorner}$ refers to a single core transformation. The two arrows point at the rows on which the core transformation acts. We now recall some basic properties and operations on core transformations.

The product of two core transformations acting on the same rows of a matrix is a core transformation. For this reason, if two consequent core transformations are applied to the same rows, we will directly depict them as one, and we will call it a *fusion* of core transformations, $\overleftrightarrow{\lrcorner} \overleftrightarrow{\lrcorner} = \overleftrightarrow{\lrcorner}$.

Another, less trivial operation, is the *turnover*, if U_k and W_k are two core transformations acting on rows k and $k + 1$ of a matrix, and V_{k+1} is a core transformation acting on rows $k + 1$ and $k + 2$, then three core transformations \widetilde{U}_{k+1} , \widetilde{V}_k and \widetilde{W}_{k+1} exist, such that \widetilde{U}_{k+1} and \widetilde{W}_{k+1} act on rows $k + 1$ and $k + 2$, \widetilde{V}_k acts on rows k and $k + 1$, and $U_k V_{k+1} W_k = \widetilde{U}_{k+1} \widetilde{V}_k \widetilde{W}_{k+1}$.

This result can graphically be depicted by

$$\overleftrightarrow{\lrcorner} \overleftrightarrow{\lrcorner} \overleftrightarrow{\lrcorner} = \overleftrightarrow{\lrcorner} \overleftrightarrow{\lrcorner} \overleftrightarrow{\lrcorner},$$

and can be extended to arbitrarily long sequences of core transformations; see Lemma 2.5 in [15]. This allows one to modify the configuration of core transformations, moving the upper-left one to the bottom-right part, or vice versa.

Another important property, of which we will make a massive use in the next sections, is the following. If $M = QR$ is the QR factorization of a matrix M , where Q is a unitary matrix represented by a sequence of core transformations following a certain pattern, then there exists an RQ factorization $M = \tilde{R}\tilde{Q}$, where the transformations in \tilde{Q} follow the same pattern as in Q ; see Lemma 2.6 in [15]. With *pattern* we refer to the order (up to commutativity) in which the core transformations Q_i appear. In practice, this means that we are able to move a bunch of core transformations from one side to the other of an upper triangular matrix at any time, preserving the original configuration.

We are now able to present the extended Hessenberg form and the reduction algorithm. Let $M \in \mathbb{C}^{n \times n}$ be a generic matrix. The goal is to compute unitary matrices V, Q and an upper triangular matrix R such that $V^H MV = QR$ and Q is represented by $n - 1$ core transformations, following a certain pattern fixed in advance. The classical Hessenberg form, which is a special case where the transformations in Q form a descending sequence, is shown in the following example.

Example 2. (*QR factorization of a Hessenberg matrix*)

Let $M \in \mathbb{C}^{n \times n}$ be a Hessenberg matrix. To compute a QR factorization of M we need to annihilate all the nonzero entries of M below the diagonal. To do so, we start applying a first transformation G_1 which eliminates the element $m_{2,1}$. Then we proceed applying a new transformation which eliminates the entry $m_{3,2}$ and so on, until we obtain an upper triangular matrix. The QR factorization of M consists then of $n - 1$ core transformations ordered in a *descending* sequence and an upper triangular matrix. The factorization can be depicted as

$$M = G_1^H \cdots G_{n-1}^H R = \begin{array}{c} \left[\begin{array}{c} \downarrow \\ \downarrow \\ \downarrow \\ \downarrow \end{array} \right] \left[\begin{array}{cccc} \times & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \\ & & & \times & \times \\ & & & & \times \end{array} \right] \end{array}.$$

A descending sequence is not the only possible pattern core transformations can assume. If M is, for example, the inverse of a Hessenberg matrix, its factorization contains an *ascending* sequence of transformations

$$M = G_1^H \cdots G_{n-1}^H R = \begin{array}{c} \left[\begin{array}{c} \downarrow \\ \downarrow \\ \downarrow \\ \downarrow \end{array} \right] \left[\begin{array}{cccc} \times & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \\ & & & \times & \times \\ & & & & \times \end{array} \right] \end{array}.$$

In [15] it was shown that any matrix M can be unitarily transformed into a matrix $\widehat{M} = V^H MV$, factored by $\widehat{M} = QR$ where the core transformations representing the unitary factor Q follow an arbitrarily prescribed pattern. Matrices satisfying such a factorization take the name of *extended Hessenberg* matrices.

To transform M to \widehat{M} we operate on the QR factorization of the matrix M , which typically consists of a pyramid-shaped pattern of core transformations²

$$M = \begin{array}{c} \left[\begin{array}{cccc} \downarrow & & & \\ \downarrow & \downarrow & & \\ \downarrow & \downarrow & \downarrow & \\ \downarrow & \downarrow & \downarrow & \downarrow \end{array} \right] \left[\begin{array}{cccc} \times & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \\ & & & \times & \times \\ & & & & \times \end{array} \right] \end{array}.$$

To get the pyramid, one first needs to apply transformations to the left of M until all the elements below the diagonal are zeroed. This requires in total $\frac{n(n-1)}{2}$ core transformations. We will not go into the details, but one can also obtain an upside down pyramid in the QR factorization.

²If the matrix exhibits particular structure, core transformations could equal the identity and are therefore not shown in the pyramidal structure.

Repeating these procedures $n - 2$ times we eventually obtain a matrix in extended Hessenberg form, where the unitary factor follows a prescribed pattern

$$M_{n-2} = \begin{array}{c} \left[\begin{array}{c} \downarrow \\ \downarrow \\ \downarrow \\ \downarrow \end{array} \right] \begin{bmatrix} \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \\ & & & & \times \end{bmatrix} \end{array}.$$

We will call a pattern, as the one depicted above, a *twisted pattern*, and we will describe it through a *position vector* $\mathbf{p} \in \{\ell, r\}^{n-2}$. For a unitary matrix Q in condensed form $Q = G_1 \cdots G_{n-1}$, the corresponding position vector is given by $\mathbf{p} = (p_1, \dots, p_{n-2})$ where

$$p_i = \begin{cases} \ell, & \text{if } G_i \text{ is on the left of } G_{i+1} \\ r, & \text{if } G_i \text{ is on the right of } G_{i+1} \end{cases}.$$

For example, the pattern chosen for the matrix M_{n-2} above is given by $\mathbf{p} = (\ell, r, r)$. We also define $\text{rev}(\mathbf{p}) = (p_{n-2}, \dots, p_1)$ and $-\mathbf{p} = (-p_1, \dots, -p_{n-2})$, where $-p_i = \ell$ if $p_i = r$ and $-p_i = r$ if $p_i = \ell$.

Note that an extended unitary matrix $Q = G_1 \cdots G_{n-1}$ can always be factored as $Q = Q_d Q_a$, where Q_d and Q_a are respectively a (non necessarily consecutive) descending and an ascending sequence of transformations. The factor Q_d contains all the transformations G_i such that $p_i = \ell$ and the factor Q_a contains all the transformations G_i such that $p_i = r$. The transformation G_{n-1} can be stored in any of the two factors. We call such a factorization a *DA-factorization* of Q and use it later on to prove some theoretical facts.

3 Hamiltonian matrices

A Hamiltonian matrix equals

$$H = \begin{bmatrix} A & G \\ F & -A^H \end{bmatrix} \in \mathbb{C}^{2n \times 2n},$$

where all blocks A , G and F have size $n \times n$ and F , G are Hermitian. Note that this is equivalent to HJ being Hermitian, where

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

In order to preserve the Hamiltonian structure within a Hamiltonian QR algorithm, it is possible to make use of symplectic matrices defined as follows.

Definition 3. A matrix $S \in \mathbb{C}^{2n \times 2n}$ is symplectic if $S^H J S = J$.

Symplectic matrices are particularly interesting in the context of Hamiltonian matrices because of the following result [13].

Lemma 4. Let $H, S \in \mathbb{C}^{2n \times 2n}$ be a Hamiltonian and a symplectic matrix respectively, then $S H S^{-1}$ is Hamiltonian.

In practice, for stability reasons, we will not consider the whole family of symplectic matrices, but restrict ourselves to the family of unitary symplectic matrices, which are described as follows; see [13] for details.

Lemma 5. A unitary matrix $S \in \mathbb{C}^{2n \times 2n}$ is symplectic if and only if

$$S = \begin{bmatrix} U_1 & U_2 \\ -U_2 & U_1 \end{bmatrix},$$

where $U_1, U_2 \in \mathbb{C}^{n \times n}$.

In the classic QR algorithm [17], two matrix forms play a fundamental role: the Schur form, which is the goal of the whole computation, and the Hessenberg form, which is the convenient intermediate form exploited by the iterative method. Both of these forms do, however, not exist for Hamiltonian matrices, and an adaptation of these two concepts to the Hamiltonian setting is required.

Definition 6. Let $H \in \mathbb{C}^{2n \times 2n}$ be a Hamiltonian matrix, the decomposition $H = U^H T U$ is a Hamiltonian Schur form of H if U is unitary and symplectic, and

$$T = \begin{bmatrix} A & G \\ & -A^H \end{bmatrix}$$

is Hamiltonian with the upper-left block $A \in \mathbb{C}^{n \times n}$ upper triangular. A matrix with the same structure as T is a Hamiltonian upper triangular matrix.

Although the Hamiltonian Schur form has similar features to the classic Schur form, in the sense that both reveal the eigenvalues of the matrix along the diagonal, its existence is not for granted in the Hamiltonian setting. Nevertheless, under some assumptions, conditions on the existence are known, as illustrated by the following theorem.

Theorem 7 (Theorem 3.1 in [13]). Let $H \in \mathbb{C}^{2n \times 2n}$ be a Hamiltonian matrix. If H has no eigenvalues on the imaginary axis, then a Hamiltonian Schur form of H exists.

The classic Hamiltonian QR algorithm [5] makes use of an intermediate Hamiltonian condensed form to compute eigenvalues which is named the Hamiltonian Hessenberg form.

Definition 8. A Hamiltonian matrix

$$H = \begin{bmatrix} A & G \\ F & -A^H \end{bmatrix} \in \mathbb{C}^{2n \times 2n}$$

is in Hamiltonian Hessenberg form if $A \in \mathbb{C}^{n \times n}$ is a classical Hessenberg matrix and $F \in \mathbb{C}^{n \times n}$ is zero except for the possibly nonzero entry $f_{n,n}$.

Conditions on the existence of a Hamiltonian Hessenberg form are given in [2] (see also [1, 14]).

In [5], a procedure is described to compute the Hamiltonian Hessenberg form via symplectic unitary similarities. This method is applicable to a subfamily of the Hamiltonian matrices, namely those matrices whose block F has rank at most 1. In this work we extend the Hamiltonian condensed form, thereby still restricting ourselves to a block F of rank 1.

Two other classes of important matrices are the K -Hamiltonian and K -symplectic matrices, defined below.

Definition 9. Let

$$\Phi = \Phi_n \in \mathbb{C}^{n \times n} = \begin{bmatrix} & & 1 \\ & \ddots & \\ 1 & & \end{bmatrix},$$

and

$$K \in \mathbb{C}^{2n \times 2n} = \begin{bmatrix} I & \\ & \Phi \end{bmatrix}.$$

A matrix $H \in \mathbb{C}^{2n \times 2n}$ is K -Hamiltonian if KHK is Hamiltonian. Similarly, we call a matrix $S \in \mathbb{C}^{2n \times 2n}$ K -symplectic if KSK is symplectic.

Note that, if H is a Hamiltonian matrix as in (1), we can construct the corresponding K -Hamiltonian matrix

$$KHK = \begin{bmatrix} A & G\Phi \\ \Phi F & -\Phi A^H \Phi \end{bmatrix},$$

where both blocks ΦF and $G\Phi$ are per-Hermitian.

It is simple to prove that a Hamiltonian matrix H is in Hamiltonian Hessenberg (upper triangular) form if and only if the related K -Hamiltonian matrix KHK is in upper Hessenberg (upper triangular) form. Similarly, K -symplectic matrices play for K -Hamiltonians the same role as that of symplectic matrices in the Hamiltonian framework: they preserve the K -Hamiltonian structure in case of similarity transformations.

K -Hamiltonian matrices are a powerful tool, because they allow one to redefine the classical Hamiltonian QR method: if we switch to the K -Hamiltonian framework, the Hamiltonian QR algorithm is a particular instance of the more general bidirectional implicit QR method described in [16].

In the remainder of the text we will work with both related families, switching from Hamiltonian to K -Hamiltonian at our convenience.

It is also important to consider that the unitary core transformations from Definition 1 cannot be used directly within the (K -)Hamiltonian framework, because they are not (K -)symplectic. For this reason we introduce the following.

Definition 10. A unitary K -symplectic core transformation is a matrix $G \in \mathbb{C}^{2n \times 2n}$ equal to one of the two following forms:

- type I: G is equal to the identity except for two unitary 2-by-2 blocks $G_i = G(i : i+1, i : i+1)$ and $G_{2n-i} = G(2n-i, 2n-i+1, 2n-i : 2n-i+1)$, such that $G_i = \Phi G_{2n-i} \Phi$;
- type II: G is equal to the identity except for the 2-by-2 block

$$G_n = G(n : n+1, n : n+1) = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix},$$

such that all the following equations hold

$$\begin{aligned} \bar{g}_{11}g_{21} - \bar{g}_{21}g_{11} &= 0, & \bar{g}_{12}g_{21} - \bar{g}_{22}g_{11} &= -1, \\ \bar{g}_{11}g_{22} - \bar{g}_{21}g_{12} &= 1, & \bar{g}_{12}g_{22} - \bar{g}_{22}g_{12} &= 0, \\ \bar{g}_{11}g_{11} + \bar{g}_{21}g_{21} &= 1, & \bar{g}_{12}g_{11} + \bar{g}_{22}g_{21} &= 0, \\ \bar{g}_{11}g_{12} + \bar{g}_{21}g_{22} &= 0, & \bar{g}_{12}g_{12} + \bar{g}_{22}g_{22} &= 1. \end{aligned}$$

K -symplectic core rotations of the first type, have two active parts

$$G_i = \begin{bmatrix} c & s \\ -\bar{s} & \bar{c} \end{bmatrix} \text{ and } G_{2n-i} = \begin{bmatrix} \bar{c} & -\bar{s} \\ s & c \end{bmatrix},$$

which will be depicted, similarly to the regular ones, by two arrow-headed brackets $\overleftrightarrow{\lceil}$ on the same vertical line. The condition stated in the second point could seem very strong, but it is fulfilled, for example, by the simple real rotation

$$G_n = \begin{bmatrix} c & s \\ -s & c \end{bmatrix},$$

where both c and s are real, and $c^2 + s^2 = 1$. In this work, we will actually not use core transformations of the second type, but they are compulsory in [6] and included for completeness.

4 The extended Hamiltonian Hessenberg form

In this section we present a series of theoretical results that will lead to the definition of the (K -)Hamiltonian extended Hessenberg form. More precisely, the (K -)Hamiltonian structure will impose some constraints on the pattern of rotations arising in the QR factorization of the extended

Hamiltonian Hessenberg form. We will investigate here the structure of the pattern and we will prove that it is palindromic.

We prove that the pattern of a condensed unitary matrix is reversed through a flip similarity of its Hermitian transpose.

First, observe that if $M = M_1 \oplus M_2 \oplus \cdots \oplus M_t$ is a block diagonal matrix, then $\Phi M \Phi = (\Phi M_t \Phi) \oplus \cdots \oplus (\Phi M_2 \Phi) \oplus (\Phi M_1 \Phi)$. In particular, this implies that for any $1 \leq k < n$ and for any core transformation G_k , there exists another core transformation \tilde{G}_{n-k} , such that $\Phi G_k \Phi = \tilde{G}_{n-k}$.

Theorem 11. *Let $Q \in \mathbb{C}^{n \times n}$ be a condensed unitary matrix whose pattern is determined by the position vector $\mathbf{p} = [p_1, p_2, \dots, p_{n-2}]$ and $n \geq 2$. Then the matrix $\Phi Q^H \Phi$ is a condensed unitary matrix associated to the position vector $\text{rev}(\mathbf{p})$.*

Proof. The proof is based on induction on n . For $n = 2$, the vector \mathbf{p} is void, so there is nothing to prove.

Let $n > 2$ and assume that we have already proved the theorem for matrices of size $n - 1$. We have two cases: $p_1 = r$ or $p_1 = \ell$.

If $p_1 = r$, then there exists a core transformation G_1 and a condensed matrix $Q_1 \in \mathbb{C}^{(n-1) \times (n-1)}$, with pattern $\mathbf{p}^{(1)} = [p_2, \dots, p_{n-2}]$, such that

$$Q = G_1 \begin{bmatrix} 1 & 0 \\ 0 & Q_1 \end{bmatrix}.$$

We get

$$\Phi Q^H \Phi = \Phi \begin{bmatrix} 1 & 0 \\ 0 & Q_1^H \end{bmatrix} \Phi \Phi G_1^H \Phi = \begin{bmatrix} \Phi Q_1^H \Phi & 0 \\ 0 & 1 \end{bmatrix} \tilde{G}_{n-1}.$$

By inductive hypothesis, $\Phi Q_1^H \Phi$ is a condensed unitary matrix with pattern $\text{rev}(\mathbf{p}^{(1)}) = [p_{n-2}, \dots, p_2]$. Thus the product

$$\begin{bmatrix} \Phi Q_1^H \Phi & 0 \\ 0 & 1 \end{bmatrix} \tilde{G}_{n-1}$$

is another condensed unitary matrix. The post-multiplication by \tilde{G}_{n-1} implies that the last letter in the pattern of $\Phi Q^H \Phi$ is $r = p_1$. Therefore the pattern of $\Phi Q^H \Phi$ is $[p_{n-2}, \dots, p_2, p_1]$, which concludes the proof of the first part.

The case $p_1 = \ell$ is specular: since G_1 is a right factor of Q , \tilde{G}_{n-1} becomes a left factor of $\Phi Q^H \Phi$, so the last letter in its pattern is $\ell = p_1$. \square

We indicate by $Q_d(i : j)$ an $n \times n$ unitary matrix obtained as a descending sequence of core transformations $G_i G_{i+1} \cdots G_{j-1}$. A similar notation $Q_a(i : j)$ refers to an ascending sequence.

Lemma 12. *Let $Q_d(i : j) \in \mathbb{C}^{n \times n}$ be a descending sequence of core transformations, then there exists $Q_a(i : j)$ such that $Q_d(i : j)^H = Q_a(i : j)$. Moreover, there exists a descending sequence $Q_d(n - j + 1 : n - i + 1)$, such that $\Phi Q_d(i : j)^H \Phi = Q_d(n - j + 1 : n - i + 1)$.*

Proof. The first part is obtained by writing

$$Q_d(i : j) = G_i G_{i+1} \cdots G_{j-1} \Rightarrow Q_d(i : j)^H = G_{j-1}^H \cdots G_{i+1}^H G_i^H,$$

which shows that $Q_d(i : j)^H$ is an ascending sequence.

For the second part, observe that for any $1 \leq k < n$, and for any core transformation G_k , there exists \tilde{G}_{n-k} such that $\Phi G_k \Phi = \tilde{G}_{n-k}$. This fact yields

$$\begin{aligned} \Phi Q_d(i : j)^H \Phi &= (\Phi G_{j-1}^H \Phi) (\Phi G_{j-2}^H \Phi) \cdots (\Phi G_{i+1}^H \Phi) (\Phi G_i^H \Phi) = \\ &\tilde{G}_{n-j+1} \tilde{G}_{n-j+2} \cdots \tilde{G}_{n-i-1} \tilde{G}_{n-i} = Q_d(n - j + 1 : n - i + 1). \end{aligned}$$

\square

In the same way we can prove the following.

Lemma 13. *Let $Q_a(i : j) \in \mathbb{C}^{n \times n}$ be an ascending sequence of core transformations, then there exists $Q_d(i : j)$ such that $Q_a(i : j)^H = Q_d(i : j)$. Moreover, there exists an ascending sequence $Q_a(n - j + 1 : n - i + 1)$, such that $\Phi Q_a(i : j)^H \Phi = Q_a(n - j + 1 : n - i + 1)$.*

Another interesting result about K -symplectic matrices is the following.

Theorem 14. *Let $T \in \mathbb{C}^{n \times n}$ be a condensed unitary K -symplectic matrix with pattern described by $\mathbf{p} \in \{\ell, r\}^{n-2}$. Then \mathbf{p} is anti-palindromic, in other words, $\text{rev}(\mathbf{p}) = -\mathbf{p}$.*

Proof. We know from Theorem 11 that $\Phi T^H \Phi$ has pattern $\text{rev}(\mathbf{p})$. On the other hand, it is easy to see that the pattern of T^H is $-\mathbf{p}$. Since any unitary K -symplectic matrix can be written (see Lemma 5) as

$$T = \begin{bmatrix} U_1 & U_2 \Phi \\ -\Phi U_2 & \Phi U_1 \Phi \end{bmatrix},$$

we have

$$\Phi T^H \Phi = \begin{bmatrix} U_1^H & U_2^H \Phi \\ -\Phi U_2^H & \Phi U_1^H \Phi \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} T^H \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}.$$

The transformation $T^H \rightarrow \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} T^H \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}$ preserves the pattern. Thus $\text{rev}(\mathbf{p}) = -\mathbf{p}$. \square

Condensed K -symplectic unitary matrices are not used explicitly in the algorithm for the reduction to extended Hessenberg form, but the previous property will be used later, when defining the extended K -Hamiltonian Hessenberg form.

To simplify the presentation, we introduce the matrices

$$P = \begin{bmatrix} I_n & 0 \\ 0 & -I_n \end{bmatrix}, \quad L = P \Phi_{2n} = \begin{bmatrix} 0 & \Phi_n \\ -\Phi_n & 0 \end{bmatrix} = KJK, \quad (2)$$

where J and K are the same as in Section 3. Both K -Hamiltonian and K -symplectic matrices can be characterized in terms of L as follows.

Lemma 15. *The matrix $H \in \mathbb{C}^{2n \times 2n}$ is K -Hamiltonian if and only if LH is Hermitian if and only if $H = LH^H L$. The matrix $T \in \mathbb{C}^{2n \times 2n}$ is K -symplectic if and only if $TLT^H = L$, and is unitary symplectic if and only if $LT = TL$.*

Proof. A matrix H is K -Hamiltonian if and only if KHK is Hamiltonian. That is, $KHKJ$ is Hermitian. Notice that $L^H = -L$, so $KHKJ = -JKH^H K$, which is equivalent to $HL = -LH^H = (HL)^H$ which, in turn, is equivalent to $H = LH^H L$.

A matrix T is K -symplectic if and only if $KT^H K$ is symplectic. This means, by definition, $KT^H K J K T^H K = J$, which is equivalent to $TLT^H = L$. Moreover, If T is unitary, we have $TL = LT$. \square

To proceed further, we need to recall that the patterns arising in the factorization of a compressed unitary matrix are closely linked to rational Krylov matrices. Suppose $M = QR$ and let \mathbf{p} be the position vector of Q . We then construct the associated Krylov matrix $K_{\mathbf{p}}(M, v)$ as follows. The first column is always the vector v , the others are of the form $M^j v$, where j can be positive or negative. More precisely, the $(i - 1)$ -st column is determined by p_i . Suppose that the first $i - 1$ columns of $K_{\mathbf{p}}(M, v)$ have been determined, then the i -th column is $M^j v$, where

$$j = \begin{cases} \max\{k \in \mathbb{Z} : M^k v \text{ is one of the first } i - 1 \text{ columns}\} + 1, & \text{if } p_i = \ell \\ \min\{k \in \mathbb{Z} : M^{-k} v \text{ is one of the first } i - 1 \text{ columns}\} - 1, & \text{if } p_i = r \end{cases}.$$

Since \mathbf{p} has only $n - 2$ elements, the last vector in the Krylov sequence is not defined. In fact it does not matter, p_{n-1} can be chosen as ℓ or r indifferently, as the complete span of the subspace will be identical. The following result was proven in [15].

Lemma 16. *Let $M = QR$ be a nonsingular matrix and R an upper triangular matrix. Suppose the unitary factor Q is in irreducible condensed form, which means none of its core transformations are equal to the identity. If \mathbf{p} is the associated position vector, then $K_{\mathbf{p}}(M, e_1)$ is upper triangular.*

Another result, which is particularly handy in dealing with K -Hamiltonian matrices, is the following.

Lemma 17. *Let $M = QR$ be a nonsingular matrix and R an upper triangular matrix. Suppose the unitary factor Q is in irreducible condensed form and follows the pattern \mathbf{p} . Then the matrix $K_{\text{rev}(\mathbf{p})}(\Phi M^H \Phi, e_1) = \Phi K_{\text{rev}(\mathbf{p})}(M^H, e_n)$ is upper triangular.*

Proof. We observe that if $M = QR$, then $\Phi M^H \Phi = \Phi R^H \Phi \Phi Q^H \Phi$, where $\Phi R^H \Phi$ is upper triangular and $\Phi Q^H \Phi$ is a condensed unitary matrix whose pattern is $\text{rev}(\mathbf{p})$ (see Theorem 11). Moving $\Phi Q^H \Phi$ to the other side of $\Phi R^H \Phi$, we get an upper triangular matrix \tilde{R} and a condensed unitary matrix \tilde{Q} with the same pattern as $\Phi Q^H \Phi$, such that $\Phi M^H \Phi = \tilde{Q} \tilde{R}$. Using Lemma 16, we have that $K_{\text{rev}(\mathbf{p})}(\Phi M^H \Phi, e_1)$ is upper triangular and the proof is completed. \square

Another technical result is the following.

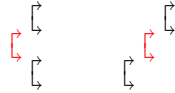
Lemma 18. *Let $M \in \mathbb{C}^{n \times n}$ be a nonsingular matrix and $\mathbf{p} \in \{\ell, r\}^{n-2}$ a position vector. If $M = QR$, where Q is an irreducible condensed unitary matrix whose pattern is determined by \mathbf{p} , then the matrix $K_{\mathbf{p}'}(M, e_1)$ is not upper triangular, for each $\mathbf{p}' \neq \mathbf{p}$.*

Proof. Consider first the case $p_1 \neq p'_1$. If $p_1 = \ell$, then there exists a core transformation G_1 acting just on the rows 1 and 2 and a unitary matrix $Q = \begin{bmatrix} 1 & 0 \\ 0 & Q' \end{bmatrix}$, with Q' irreducible, such that $M = G_1 QR$. We depict by $\tilde{e}_i \in \mathbb{C}^n$ any vector whose first i entries are the only nonzero entries. The second column of $K_{\mathbf{p}}(M, e_1)$ is $M e_1 = G_1 Q R e_1 = G_1 \tilde{e}_1 = \tilde{e}_2$. The second column of $K_{\mathbf{p}'}(M, e_1)$ is $M^{-1} e_1 = R^{-1} Q^H G_1^H e_1 = R^{-1} Q^H \tilde{e}_2$. Given that $Q^H \tilde{e}_2 \neq \tilde{e}_2$, the matrix $K_{\mathbf{p}'}(M, e_1)$ is not upper triangular. The case $p_1 = r$ is similar.

Now assume that $s > 1$, $p_i = p'_i$ for each $i < s$ and $p_s \neq p'_s$. We can distinguish four cases

1. $p_s = \ell$, $p'_s = r$ and $p_{s-1} = p'_{s-1} = r$;
2. $p_s = r$, $p'_s = \ell$ and $p_{s-1} = p'_{s-1} = r$.
3. $p_s = r$, $p'_s = \ell$ and $p_{s-1} = p'_{s-1} = \ell$;
4. $p_s = \ell$, $p'_s = r$ and $p_{s-1} = p'_{s-1} = \ell$;

In case 1, the transformation G_s is the corner of a bend sequence (r, ℓ) in the pattern \mathbf{p} , while it is in the middle of an ascending sequence in the pattern \mathbf{p}' .



This means that the s -th column of $K_{\mathbf{p}}(M, e_1)$ and $K_{\mathbf{p}'}(M, e_1)$ is of the type $\tilde{e}_s = M^{-k} e_1$ with k positive. Thus the $(s+1)$ -st column of $K_{\mathbf{p}'}(M, e_1)$ is $M^{-1} \tilde{e}_s$ by induction.

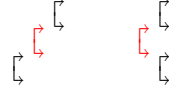
There exist descending sequences Q_{d_0} and Q_{d_2} and an ascending sequence Q_a , such that $M = Q_{d_0} G_s Q_{d_2} Q_a R$. The leftmost and rightmost transformations of Q_{d_2} are of the type G_{s+1} and $G_{t'}$ respectively, with $t' > s$. On the other hand, Q_{d_0} does not act on row s .

We have

$$\begin{aligned}
M^{-1} \tilde{e}_s &= R^{-1} Q_a^H Q_{d_2}^H G_s^H Q_{d_0}^H \tilde{e}_s \\
&= R^{-1} Q_a^H Q_{d_2}^H G_s^H \tilde{e}_s \\
&= R^{-1} Q_a^H Q_{d_2}^H \tilde{e}_{s+1} \quad , \\
&= R^{-1} Q_a^H \tilde{e}_{t'+1} \\
&= \tilde{e}_{t'+1},
\end{aligned}$$

thus $M^{-1}\tilde{e}_s$ is not of the type \tilde{e}_{s+1} and $K_{\mathbf{p}'}(M, e_1)$ is not upper triangular.

In case 2, the transformation G_s is the corner of a bend sequence (r, ℓ) in the pattern \mathbf{p}' , while it is in the middle of an ascending sequence in the pattern \mathbf{p} .



This means that the s -th column of $K_{\mathbf{p}}(M, e_1)$ and $K_{\mathbf{p}'}(M, e_1)$ is of the type $\tilde{e}_s = M^{-k}e_1$ with k positive. Thus the $(s+1)$ -st column of $K_{\mathbf{p}'}(M, e_1)$ is $M\tilde{e}_t$, where $t < s$ is the smallest index of the ascending sequence to which the s -th core transformation belongs.

There exist a descending sequence Q_d and ascending sequences Q_{a_0} , Q_{a_1} and Q_{a_2} such that $M = Q_d Q_{a_1} G_s Q_{a_2} Q_{a_0} R$. The leftmost and the rightmost transformations in Q_{a_1} are of the type $G_{t'}$ and G_{s+1} respectively, with $t' > s$; while the leftmost and the rightmost transformations in Q_{a_2} are of the type G_{s-1} and G_t respectively. Notice that Q_{a_0} does not act on row t .

We have

$$\begin{aligned}
M\tilde{e}_t &= Q_d Q_{a_1} G_s Q_{a_2} Q_{a_0} R \tilde{e}_t \\
&= Q_d Q_{a_1} G_s Q_{a_2} Q_{a_0} \tilde{e}_t \\
&= Q_d Q_{a_1} G_s Q_{a_2} \tilde{e}_t \\
&= Q_d Q_{a_1} G_s \tilde{e}_s \\
&= Q_d Q_{a_1} \tilde{e}_{s+1} \\
&= Q_d \tilde{e}_{t'+1} \\
&= \tilde{e}_{t'+1},
\end{aligned}$$

thus $M\tilde{e}_t$ is not of the type \tilde{e}_{s+1} and $K_{\mathbf{p}'}(M, e_1)$ is not upper triangular.

Cases 3 and 4 are dealt with similarly. \square

Corollary 19 (Uniqueness of the QR factorization). *Let $M \in \mathbb{C}^{n \times n}$ be a nonsingular matrix and $\mathbf{p}, \mathbf{q} \in \{\ell, r\}^{n-2}$ two position vectors. Suppose $M = QR = \tilde{Q}\tilde{R}$, where Q and \tilde{Q} are irreducible condensed unitary matrices with patterns determined by \mathbf{p} and \mathbf{q} respectively, and R, \tilde{R} are upper triangular matrices. Then $\mathbf{p} = \mathbf{q}$.*

Proof. We know that $K_{\mathbf{p}}(M, e_1)$ and $K_{\mathbf{q}}(M, e_1)$ are upper triangular, but Lemma 18 forces $\mathbf{p} = \mathbf{q}$. \square

The previous results apply to any generic nonsingular matrix. If M is, moreover, K -Hamiltonian, something more can be said.

Lemma 20. *Let $H \in \mathbb{C}^{2n \times 2n}$ be a nonsingular K -Hamiltonian matrix and $\mathbf{p} \in \{\ell, r\}^{2n-2}$ a position vector. Suppose $H = QR$, where Q is a condensed unitary matrix with pattern determined by \mathbf{p} . Then the matrix $K_{\text{rev}(\mathbf{p})}(H, e_1)$ is upper triangular.*

Proof. We know from Lemma 17 that $K_{\text{rev}(\mathbf{p})}(\Phi H^H \Phi, e_1)$ is upper triangular. If H is K -Hamiltonian, then

$$\begin{aligned}
K_{\text{rev}(\mathbf{p})}(\Phi H^H \Phi, e_1) &= P K_{\text{rev}(\mathbf{p})}(P \Phi H^H \Phi P, P e_1) \\
&\stackrel{P e_1 = e_1}{=} P K_{\text{rev}(\mathbf{p})}(-L H^H L, e_1) = P K_{\text{rev}(\mathbf{p})}(-H, e_1),
\end{aligned}$$

where P and L are defined in (2). Since the latter matrix is upper triangular, we have that $K_{\text{rev}(\mathbf{p})}(H, e_1)$ is upper triangular. \square

Now we are able to prove the main result of this section.

Theorem 21. Let $H \in \mathbb{C}^{2n \times 2n}$ be a nonsingular K -Hamiltonian matrix and $\mathbf{p} \in \{\ell, r\}^{2n-2}$ a position vector. If $H = QR$, where Q is a condensed unitary matrix following the pattern determined by \mathbf{p} , then \mathbf{p} is palindromic, namely $\mathbf{p} = \text{rev}(\mathbf{p})$.

Proof. The proof is obtained combining Lemma 20 and Lemma 18: from the former we know that both $K_{\mathbf{p}}(H, e_1)$ and $K_{\text{rev}(\mathbf{p})}(H, e_1)$ are upper triangular. From the latter we know that the only upper triangular matrix is $K_{\mathbf{p}}(H, e_1)$, from which we conclude that $\mathbf{p} = \text{rev}(\mathbf{p})$. \square

We have shown that the condensed form of a nonsingular K -Hamiltonian matrix has a palindromic position vector. This means that the central part of the vector (say, (p_n, p_{n+1})) is of the type (ℓ, ℓ) or (r, r) , since $2n - 2$ is even. Thus Q can be factored in just one of the two following ways

1. *Descending type:*

$$Q = \begin{bmatrix} Q_1 & 0 \\ 0 & I \end{bmatrix} G_n \begin{bmatrix} I & 0 \\ 0 & Q_2 \end{bmatrix}; \quad (3)$$

2. *Ascending type:*

$$Q = \begin{bmatrix} I & 0 \\ 0 & \widehat{Q}_2 \end{bmatrix} \widehat{G}_n \begin{bmatrix} \widehat{Q}_1 & 0 \\ 0 & I \end{bmatrix}; \quad (4)$$

where Q_1, Q_2, \widehat{Q}_1 and \widehat{Q}_2 are condensed sequences of core transformations of size n . Notice that, by Theorem 21, if the pattern of Q_1 (\widehat{Q}_1) is \mathbf{p} , then the pattern of Q_2 (\widehat{Q}_2) is $\text{rev}(\mathbf{p})$.

In the first case, there exist an upper triangular matrix \widetilde{R} and a condensed matrix \widetilde{Q}_2 with the same pattern as the one of Q_2 (namely, $\text{rev}(\mathbf{p})$) such that

$$H = \begin{bmatrix} Q_1 & 0 \\ 0 & I \end{bmatrix} G_n \widetilde{R} \begin{bmatrix} I & 0 \\ 0 & \widetilde{Q}_2 \end{bmatrix}. \quad (5)$$

In the second case, the situation is similar: the first and the last matrix of the decomposition of H are swapped.

Extended K -Hamiltonian Hessenberg matrices can be factored as in (3) and (4), the upper triangular matrix R is positioned between two factors as in (5). This factorization is used, as it reveals information, which is not the case if we would have used the classical QR factorization as we will show in Corollary 24. The essential uniqueness of the QR factorization carries over to the factorization above, i.e., if one has two factorizations of an extended K -Hamiltonian Hessenberg matrix of the URV form (5), then these factorizations will be essentially identical. The *essentially* means up to unimodular scaling. This follows easily by rewriting the URV factorization as a QR factorization by using the techniques described in Section 2.

To conclude, we will prove that \widetilde{Q}_2 and $\Phi Q_1^H \Phi$ are essentially identical.

Theorem 22. Let H, Q, R and the palindromic position vector $\mathbf{p} \in \{\ell, r\}^{2n-2}$ be as in Theorem 21. Then H admits the two decompositions

$$H = \begin{bmatrix} Q_1 & 0 \\ 0 & I \end{bmatrix} G_n R \begin{bmatrix} I & 0 \\ 0 & \Phi Q_1^H \Phi \end{bmatrix}; \quad (6)$$

$$H = \begin{bmatrix} I & 0 \\ 0 & Q_2 \end{bmatrix} G_n R \begin{bmatrix} \Phi Q_2^H \Phi & 0 \\ 0 & I \end{bmatrix}; \quad (7)$$

where Q_1 and Q_2 are irreducible condensed matrices of size n , G_n is a core transformation of size $2n$ acting on the rows n and $n + 1$, and R is an upper triangular matrix. Moreover, the matrix $G_n R$ is K -Hamiltonian.

Proof. We will prove the statement only for the descending type. We know that the K -Hamiltonian matrix H can be factored as

$$H = \begin{bmatrix} Q_1 & 0 \\ 0 & I \end{bmatrix} G_n \tilde{R} \begin{bmatrix} I & 0 \\ 0 & \tilde{Q}_2 \end{bmatrix}. \quad (8)$$

Since $H = LH^H L$, where L is defined in (2), we get

$$H = L \begin{bmatrix} I & 0 \\ 0 & \tilde{Q}_2^H \end{bmatrix} \tilde{R}^H G_n^H \begin{bmatrix} Q_1^H & 0 \\ 0 & I \end{bmatrix} L = \begin{bmatrix} \Phi \tilde{Q}_2^H \Phi & 0 \\ 0 & I \end{bmatrix} L \tilde{R}^H L L G_n^H L \begin{bmatrix} I & 0 \\ 0 & \Phi Q_1^H \Phi \end{bmatrix}.$$

Since $L \tilde{R}^H L$ is upper triangular and $L G_n^H L$ a core transformation, there exist a core transformation \hat{G}_n acting on the rows n and $n+1$ and an upper triangular matrix \hat{R} such that

$$H = \begin{bmatrix} \Phi \tilde{Q}_2^H \Phi & 0 \\ 0 & I \end{bmatrix} \hat{G}_n \hat{R} \begin{bmatrix} I & 0 \\ 0 & \Phi Q_1^H \Phi \end{bmatrix}. \quad (9)$$

Applying the essentially uniqueness of the URV factorization to (8) and (9), since the patterns of \tilde{Q}_2 and $\Phi Q_1^H \Phi$ are the same, there exists a diagonal matrix D with $|d_{ii}| = 1$ so that $\tilde{Q}_2 = D \Phi Q_1^H \Phi$. Thus, from (8), we have

$$H = \begin{bmatrix} Q_1 & 0 \\ 0 & I \end{bmatrix} G_n \tilde{R} \begin{bmatrix} I & 0 \\ 0 & \tilde{Q}_2 \end{bmatrix} = \begin{bmatrix} Q_1 & 0 \\ 0 & I \end{bmatrix} G_n R \begin{bmatrix} I & 0 \\ 0 & \Phi Q_1^H \Phi \end{bmatrix},$$

where R is obtained merging D into \tilde{R} .

To conclude the proof, observe that since H is K -Hamiltonian, we have $LH^H L = H$. Since

$$LH^H L = \begin{bmatrix} Q_1 & 0 \\ 0 & I \end{bmatrix} L R^H G_n^H L \begin{bmatrix} I & 0 \\ 0 & \Phi Q_1^H \Phi \end{bmatrix},$$

we have $LR^H G_n L = G_n R$ which implies that $G_n R$ is K -Hamiltonian. \square

The previous theorem shows that there are essentially two possible classes of condensed forms for K -Hamiltonian matrices, yielding the following.

Definition 23 (K -Hamiltonian extended Hessenberg form). *A K -Hamiltonian matrix H is in extended K -Hamiltonian Hessenberg form of descending type with pattern \mathbf{p} if it can be decomposed as in (6), where Q_1 has pattern \mathbf{p} . Similarly, H is in extended K -Hamiltonian Hessenberg form of ascending type with pattern \mathbf{p} if it can be decomposed as in (7), where Q_1 has pattern \mathbf{p} .*

It was already stated before that the URV factorization of the extended K -Hamiltonian Hessenberg matrix reveals more information than its QR factorization would. In fact, the matrix $G_n \tilde{R}$ in the middle is K -Hamiltonian; while a definition of the type $H = QR$, where Q is condensed, would not maintain the structure. The following result follows from Theorem 14 and Theorem 21.

Corollary 24. *Let $H \in \mathbb{C}^{2n \times 2n}$ be a nonsingular K -Hamiltonian matrix and $\mathbf{p} \in \{\ell, r\}^{2n-2}$ a position vector. If $H = QR$, where Q is an irreducible condensed unitary matrix, then Q is not K -symplectic.*

5 The algorithm

In this section we describe the transformation to extended Hamiltonian Hessenberg form in a constructive way. For simplicity, all the operations are executed on K -Hamiltonian matrices, but the results can be translated easily to the Hamiltonian setting by means of permutations, as explained in Definition 9. We will also prove that the transformation is essentially unique.

$$\begin{bmatrix} R\Phi U\Phi & G\Phi \\ |\alpha|^2 e_1 e_n^H & -U^H \Phi R^H \Phi \end{bmatrix} = \left[\begin{array}{c|c} \begin{array}{cccc} \times & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \\ & & & \times \end{array} & \begin{array}{cccc} \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \end{array} \\ \hline & \begin{array}{cccc} \times & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \\ & & & \times \end{array} \end{array} \right].$$

By multiple turnover and fusion operations, we merge the two ascending sequences into the two pyramids, obtaining H of the form

$$\begin{array}{c} \begin{array}{cccc} \times & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \\ & & & \times \end{array} \\ \begin{array}{cccc} \times & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \\ & & & \times \end{array} \\ \begin{array}{cccc} \times & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \\ & & & \times \end{array} \\ \begin{array}{cccc} \times & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \\ & & & \times \end{array} \end{array} \left[\begin{array}{c|c} \begin{array}{cccc} \times & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \\ & & & \times \end{array} & \begin{array}{cccc} \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \end{array} \\ \hline & \begin{array}{cccc} \times & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \\ & & & \times \end{array} \end{array} \right].$$

Finally, we move the remaining descending sequences to the other side of the corresponding upper triangular block. We get

$$H = \begin{bmatrix} \Phi \tilde{U} \Phi & \\ & \Phi \tilde{W}^H \Phi \end{bmatrix} \begin{bmatrix} \tilde{R} & \tilde{G} \Phi \\ |\alpha|^2 e_1 e_n^H & -\Phi \tilde{R}^H \Phi \end{bmatrix} \begin{bmatrix} \tilde{W} \\ \tilde{U}^H \end{bmatrix}$$

$$= \begin{array}{c} \begin{array}{cccc} \times & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \\ & & & \times \end{array} \\ \begin{array}{cccc} \times & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \\ & & & \times \end{array} \\ \begin{array}{cccc} \times & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \\ & & & \times \end{array} \\ \begin{array}{cccc} \times & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \\ & & & \times \end{array} \end{array}$$

This is a specular situation with respect to (11), thus we can proceed in a similar way. Repeatedly applying similarities by K -symplectic transformations, we eliminate the two descending sequences of transformations, and manipulate the two pyramids, in order to get the desired patterns:

$$\tilde{V}^H H \tilde{V} = \begin{bmatrix} I & \\ & \Phi Q^H \Phi \end{bmatrix} \begin{bmatrix} \hat{R} & \hat{G} \Phi \\ |\alpha|^2 e_1 e_n^H & -\Phi \hat{R}^H \Phi \end{bmatrix} \begin{bmatrix} Q \\ I \end{bmatrix}$$

$$= \begin{array}{c} \begin{array}{cccc} \times & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \\ & & & \times \end{array} \\ \begin{array}{cccc} \times & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \\ & & & \times \end{array} \\ \begin{array}{cccc} \times & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \\ & & & \times \end{array} \\ \begin{array}{cccc} \times & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \\ & & & \times \end{array} \end{array},$$

where the K -symplectic unitary matrix \tilde{V} is the product of all the transformations performed, \hat{R} is upper triangular and Q is a unitary matrix which obeys the prescribed pattern \mathbf{p} .

Remark 25. Suppose a matrix H in extended K -Hamiltonian Hessenberg form of descending type is given:

$$H = \begin{bmatrix} Q & \\ & I \end{bmatrix} \begin{bmatrix} \hat{R} & \hat{G}\Phi \\ |\alpha|^2 e_1 e_n^H & -\Phi \hat{R}^H \Phi \end{bmatrix} \begin{bmatrix} I & \\ & \Phi Q^H \Phi \end{bmatrix}.$$

If we apply the K -symplectic transformation

$$\begin{bmatrix} Q^H & \\ & \Phi Q^H \Phi \end{bmatrix} H \begin{bmatrix} Q & \\ & \Phi Q \Phi \end{bmatrix} = \begin{bmatrix} I & \\ & \Phi Q^H \Phi \end{bmatrix} \begin{bmatrix} \hat{R} & \hat{G}\Phi \\ |\alpha|^2 e_1 e_n^H & -\Phi \hat{R}^H \Phi \end{bmatrix} \begin{bmatrix} Q & \\ & I \end{bmatrix},$$

we can get an extended Hessenberg form of ascending type at no additional cost. The reasoning is of course valid also in the inverse case, when switching from ascending to descending type.

5.3 Uniqueness

Once the transformation to extended K -Hamiltonian Hessenberg form is defined, the next natural step is to determine to which extent such a transformation is unique. For the general case, the following result was proved.

Theorem 26 (Theorem 5.1 in [15]). *Let $M \in \mathbb{C}^{n \times n}$ be a nonsingular matrix and $\mathbf{p} \in \{\ell, r\}^{n-2}$ a position vector. Let V_1 and V_2 be two unitary matrices sharing the first column (up to a unimodular factor) such that*

$$Q_1 R_1 = M_1 = V_1^H M V_1, \quad Q_2 R_2 = M_2 = V_2^H M V_2,$$

where the unitary factors Q_1 and Q_2 in the QR decomposition of M_1 and M_2 obey the pattern specified by \mathbf{p} and, in addition, all transformations differ from the identity. Then M_1 and M_2 are essentially identical.

This is a generalization of the classic implicit Q -theorem, that ensures essential uniqueness in the general nonsingular case.

In the K -Hamiltonian case, a few more considerations are needed. First, notice that also the following result holds.

Theorem 27. *Let $M \in \mathbb{C}^{n \times n}$ be a nonsingular matrix and $\mathbf{p} \in \{\ell, r\}^{n-2}$ a position vector. Let V_1 and V_2 be two unitary matrices sharing the last column (up to a unimodular factor) such that*

$$Q_1 R_1 = M_1 = V_1^H M V_1, \quad Q_2 R_2 = M_2 = V_2^H M V_2,$$

where the unitary factors Q_1 and Q_2 in the QR decomposition of M_1 and M_2 obey the pattern specified by \mathbf{p} and, in addition, all transformations differ from the identity. Then M_1 and M_2 are essentially identical.

Proof. A simple computation shows that

$$(V_1 \Phi)(\Phi K_{\text{rev}(\mathbf{p})}(M_1^H, e_n)) = K_{\text{rev}(\mathbf{p})}(V_1 M_1^H V_1^H, V_1 e_n) = K_{\text{rev}(\mathbf{p})}(M^H, V_1 e_n),$$

and analogously $V_2 K_{\text{rev}(\mathbf{p})}(M_2^H, e_n) = K_{\text{rev}(\mathbf{p})}(M^H, V_2 e_n)$.

Invoking Lemma 17, we know that $\Phi K_{\text{rev}(\mathbf{p})}(M_1^H, e_n)$ and $\Phi K_{\text{rev}(\mathbf{p})}(M_2^H, e_n)$ are upper triangular. Since $V_1 e_n = \beta V_2 e_n$, we have two QR factorizations of the same matrix

$$K_{\text{rev}(\mathbf{p})}(M^H, V_1 e_n) = (V_1 \Phi)(\Phi K_{\text{rev}(\mathbf{p})}(M_1^H, e_n)) = (V_2 \Phi)(\Phi K_{\text{rev}(\mathbf{p})}(M_2^H, e_n)),$$

where the unitary factors share the same first column. Because of Theorem 26, V_1 and V_2 are essentially identical. \square

Now suppose we have computed a transformation to extended K -Hamiltonian Hessenberg form. We will consider the descending form case, the other case is analogous.

$$\begin{aligned}\tilde{V}^H H \tilde{V} &= \begin{bmatrix} V^H & \\ & \Phi V^H \Phi \end{bmatrix} \begin{bmatrix} A & G\Phi \\ \Phi F & -\Phi A^H \Phi \end{bmatrix} \begin{bmatrix} V & \\ & \Phi V \Phi \end{bmatrix} \\ &= \begin{bmatrix} Q & \\ & I \end{bmatrix} \begin{bmatrix} R & \tilde{G}\Phi \\ \Phi \tilde{F} & -\Phi R^H \Phi \end{bmatrix} \begin{bmatrix} I & \\ & \Phi Q^H \Phi \end{bmatrix}.\end{aligned}\quad (12)$$

The following three equations hold:

$$V^H A V = Q R \quad (13)$$

$$-\Phi(V^H A^H V)\Phi = (-\Phi R^H \Phi)(\Phi Q^H \Phi)$$

$$\Phi V^H F V = \Phi \tilde{F} \quad (14)$$

We know that $\text{rank}(F) = 1$, thus $F = v v^H$ for a certain vector v . The only nonzero entry in $\Phi \tilde{F}$ is $\tilde{f}_{1,n}$, thus (14) implies that $V^H F V = |\alpha|^2 e_n e_n^T$, where $v = \alpha V e_n$. This means that the last column of V is determined by the vector v . Now we are ready to present the main result of this section.

Theorem 28 (Essential uniqueness). *Let*

$$H = \begin{bmatrix} A & G\Phi \\ \Phi F & -\Phi A^H \Phi \end{bmatrix}$$

be a K -Hamiltonian matrix, where A is nonsingular and $\text{rank}(F) = 1$. Suppose (12) holds for both the K -symplectic unitary matrices

$$\tilde{V}_1 = \begin{bmatrix} V_1 & \\ & \Phi V_1 \Phi \end{bmatrix}, \quad \tilde{V}_2 = \begin{bmatrix} V_2 & \\ & \Phi V_2 \Phi \end{bmatrix}.$$

Suppose that the corresponding unitary matrices V_1 and V_2 follow the same pattern specified by $\mathbf{p} \in \{\ell, r\}^{n-2}$ and all their transformations differ from the identity, then $H_1 = \tilde{V}_1^H H \tilde{V}_1$ and $H_2 = \tilde{V}_2^H H \tilde{V}_2$ are essentially equal.

Proof. Because of (14), we know that $V_1 e_n = \beta V_2 e_n$. Furthermore, from (13) we obtain $V_1^H A V_1 = Q_1 R_1$ and $V_2^H A V_2 = Q_2 R_2$. By Theorem 27 we deduce that \tilde{V}_1 and \tilde{V}_2 are essentially equal. \square

6 Numerical experiments

In this section we present the results of two types of numerical experiments, the first one estimates the accuracy of the algorithm, the second algorithm illustrates the Ritz-values behavior of the reduction. The algorithm was implemented in Matlab. Random test blocks A and F were generated as follows:

```
A=randn(n)+1i*randn(n);
f=randn(n,1)+1i*randn(n,1);
F=f*f';
```

while the block G was generated in two different ways:

- case 1:

```
G=randn(n)+1i*randn(n);
G=G*G';
```

- case 2:

```
G=randn(n)+1i*randn(n);
G=(0.5)*(G+G');
```

The size n varies between 50 and 500. We tested the accuracy for the following four cases:

1. **Hessenberg case:** descending type, the transformations in Q follow the pattern $\mathbf{p} = (\ell, \dots, \ell)$;
2. **Hessenberg-like case:** ascending type, the transformations in Q follow the pattern $\mathbf{p} = (r, \dots, r)$;
3. **CMV case:** ascending type, the transformations in Q follow the pattern $\mathbf{p} = (\dots, \ell, r, \ell, r)$;
4. **Random case:** both type and position vector are chosen randomly, $\mathbf{p} = \text{randi}(2, n, 1) - 1$.

The errors in Figures 1 and 2 were measured as

$$\Delta = \frac{\|H - \tilde{V}H_1\tilde{V}^H\|_2}{\|H\|_2},$$

where

$$H = \begin{bmatrix} A & G\Phi \\ \Phi F & -\Phi A^H \Phi \end{bmatrix}$$

is the original K -Hamiltonian matrix, H_1 is its extended Hamiltonian Hessenberg form and \tilde{V} is the K -symplectic unitary matrix that gives the transformation. The computations were repeated five times for each size, and the average values are depicted.

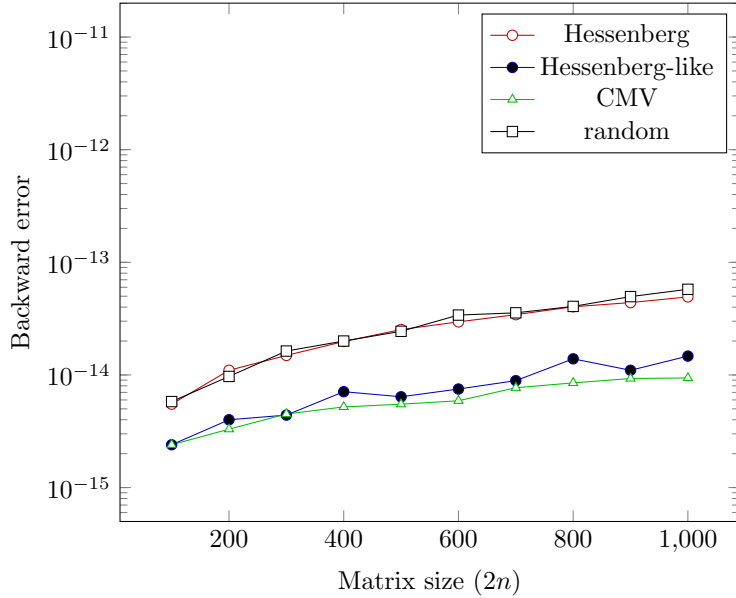


Figure 1: Backward error in case 1.

Although the results show an acceptable accuracy in general, a significant difference exists between case 1 and 2. In the first case, Hessenberg and random patterns produce bigger errors, when compared to Hessenberg-like and CMV. On the other hand, in the second case, the Hessenberg pattern clearly outperforms the others, and even random patterns perform better than both Hessenberg-like and CMV. We deduce that not all possible patterns are equally good for a certain matrix. The strength of our algorithm lies in the freedom of choosing, for each matrix, the condensed form that suits the problem the most. In this way, we are able to achieve the best accuracy

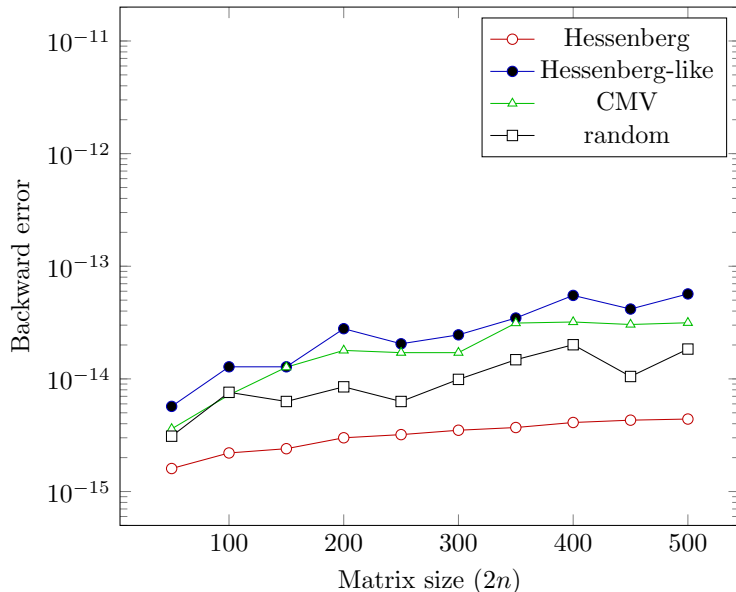


Figure 2: Backward error in case 2.

possible when using the extended Hamiltonian Hessenberg form to compute the eigenvalues of a Hamiltonian matrix [6]. This is the subject of future research.

In a second experiment, we computed the Ritz values of random real extended K -Hamiltonian Hessenberg matrices. First we generated a real K -Hamiltonian matrix H_1 with equispaced eigenvalues in $[-1, 1]$, a rank 1 bottom-left block ΦF_1 and size $2n = 100$. We computed its extended K -Hamiltonian Hessenberg form \tilde{H} in three cases: Hessenberg, Hessenberg-like, and CMV, defined as before. For each $m \in \{1, \dots, 50\}$, we determined the Ritz values as the eigenvalues of the submatrices $\tilde{H}_m = \tilde{H}(n - m + 1 : n + m, n - m + 1 : n + m)$, making use of the Matlab function `eig`.

The first three plots in Figure 3 show the results of these computations. For each value of m on the x -axis, we plot the computed Ritz values on the y -axis. Red crosses stand for Ritz values approximating eigenvalues with an absolute error smaller than $10^{-7.5}$; yellow crosses indicate good approximations with absolute errors between $10^{-7.5}$ and 10^{-5} ; the green crosses are not so good approximations with errors between 10^{-5} and $10^{-2.5}$; blue crosses are bad approximations, with errors bigger than $10^{-2.5}$. These figures do not exactly mimic the behavior one would expect after reading [3, 10, 11]. In the K -Hamiltonian setting it appears that the outer eigenvalues are not found first: the two outer red peaks are not on the outside of the interval for the first three figures. The last figure, however, does behave as expected, with peaks on the outer parts of the interval. This last figure stems from a K -Hamiltonian matrix, where the bottom-left block was set to zero, i.e., we have two completely decoupled eigenvalue problems.

7 Conclusions

In this article we have presented an algorithm to transform a Hamiltonian matrix with lower-left rank one block to a condensed format, generalizing thereby the classical Hamiltonian Hessenberg form. The uniqueness of this decomposition was proved. The numerical experiments revealed new directions for future research: selecting the optimal pattern and predicting the convergence behavior of the Ritz values are still open questions.

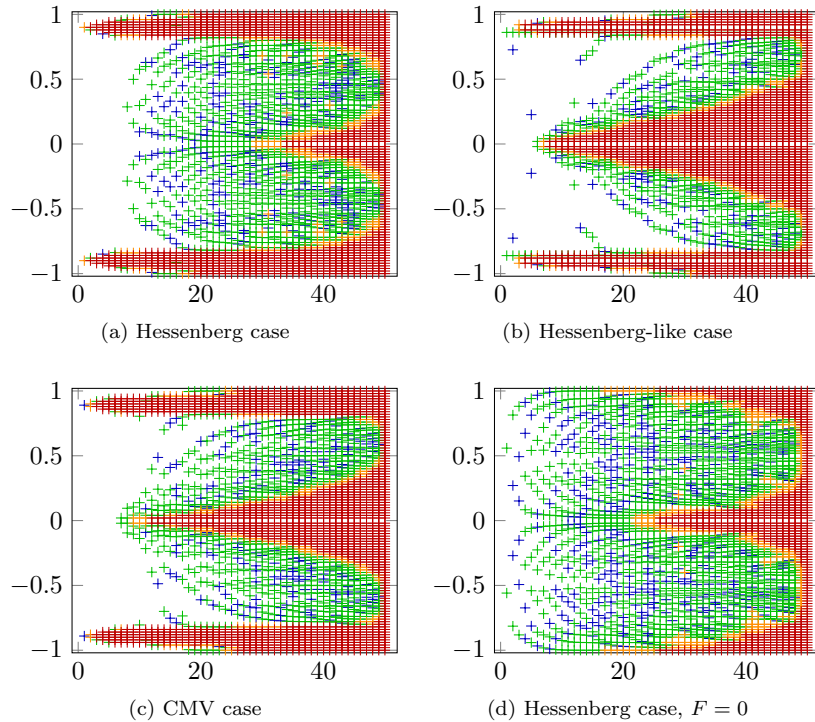


Figure 3: Ritz values of K -Hamiltonian extended Hessenberg matrices in different test cases.

References

- [1] AMMAR G., MEHRMANN V., *A geometric perspective on condensed forms for Hamiltonian matrices*, in Computation and control, II (Bozeman, MT, 1990), vol. 11 of Progr. Systems Control Theory, Birkhäuser Boston, Boston, MA, pp. 1–11 (1991).
- [2] ———, *On Hamiltonian and symplectic Hessenberg forms*, Linear Algebra Appl., 149, pp. 55–72 (1991).
- [3] BECKERMANN B., GÜTTEL S., VANDEBRIL R., *On the convergence of rational Ritz-values*, SIAM J. Matrix Anal. A., 31, pp. 1740–1774 (2010).
- [4] BINI D. A., IANNAZZO B., MEINI B., *Numerical solution of algebraic Riccati equations*, Fundamentals of Algorithms, SIAM (2012).
- [5] BYERS R., *A Hamiltonian QR-algorithm*, SIAM J. Sci. Stat. Comp., 7, pp. 212–229 (1986).
- [6] FERRANTI M., IANNAZZO B., MACH T., VANDEBRIL R., *An extended Hamiltonian QR algorithm*. In preparation (2015).
- [7] FERRANTI M., MACH T., VANDEBRIL R., *Extended Hamiltonian Hessenberg matrices arise in projection based model order reduction*, in Proceedings in Applied Mathematics and Mechanics, 15, pp. 583–584 (2015).
- [8] FRANCIS J. G. F., *The QR Transformation a unitary analogue to the LR transformation—Part 1*, Comput. J., 4, pp. 265–271 (1961).
- [9] ———, *The QR Transformation—Part 2*, Comput. J., 4, pp. 332–345 (1962).

- [10] KUIJLAARS A. B. J., *Which eigenvalues are found by the Lanczos method?*, SIAM J. Matrix Anal. A., 22, pp. 306–321 (2000).
- [11] ———, *Convergence analysis of Krylov subspace iterations with methods from potential theory*, SIAM Rev., 48, pp. 3–40 (2006).
- [12] MACH T., PRANIĆ M. S., VANDEBRIL R., *Computing approximate extended Krylov subspaces without explicit inversion*, Electron. T. Numer. Ana., 40, pp. 414–435 (2013).
- [13] PAIGE C., VAN LOAN C., *A Schur decomposition for Hamiltonian matrices*, Linear Algebra Appl., 41, pp. 11–32 (1981).
- [14] RAINES A. C. III, WATKINS D. S., *A class of Hamiltonian-symplectic methods for solving the algebraic Riccati equation*, Linear Algebra Appl., 205/206, pp. 1045–1060 (1994).
- [15] VANDEBRIL R., *Chasing bulges or rotations? A metamorphosis of the QR-algorithm*, SIAM J. Matrix Anal. A., 32, pp. 217–247 (2011).
- [16] WATKINS D. S., *Bulge exchanges in algorithms of QR-type*, SIAM J. Matrix Anal. A., 19, pp. 1074–1096 (1998).
- [17] ———, *Francis’s algorithm*, Am. Math. Mon., 118, pp. 387–403 (2011).