Eigenvalue perturbation theory of structured matrices under generic structured rank one perturbations: Symplectic, orthogonal, and unitary matrices*

Christian Mehl ‡ Volker Mehrmann‡ André C. M. Ran § Leiba Rodman ¶

Abstract

We study the perturbation theory of structured matrices under structured rank one perturbations, with emphasis on matrices that are unitary, orthogonal, or symplectic with respect to an indefinite inner product. The rank one perturbations are not necessarily of arbitrary small size (in the sense of norm). In the case of sesquilinear forms, results on selfadjoint matrices can be applied to unitary matrices by using the Cayley transformation, but in the case of real or complex symmetric or skew-symmetric bilinear forms additional considerations are necessary. For complex symplectic matrices, it turns out that generically (with respect to the perturbations) the behavior of the Jordan form of the perturbed matrix follows the pattern established earlier for unstructured matrices and their unstructured perturbations, provided the specific properties of the Jordan form of complex symplectic matrices are accounted for. For instance, the number of Jordan blocks of fixed odd size corresponding to the eigenvalue 1 or -1 have to be even. For complex orthogonal matrices, it is shown that the behavior of the Jordan structures corresponding to the original eigenvalues that are not moved by perturbations follows again the pattern established earlier for unstructured matrices, taking into account the specifics of Jordan forms of complex orthogonal matrices. The proofs are based on general results developed in the paper concerning Jordan forms of structured matrices (which include in particular the classes of orthogonal and symplectic matrices) under structured rank one perturbations. These results are presented and proved in the framework of real as well as of complex matrices.

Key Words: symplectic matrix, orthogonal matrix, unitary matrix, indefinite inner product, Cayley transformation, perturbation analysis, generic perturbation, rank one perturbation.

Mathematics Subject Classification: 15A63, 15A21, 15A57, 47A55, 93B10, 93B35, 93C73.

1 Introduction

In this paper, we study rank one perturbations of matrices that are symplectic, orthogonal, or unitary with respect to an indefinite inner product. This work extends the investigations on matrices with symmetry structures started in [16] and continued in [17] and [18].

‡Institut für Mathematik, Technische Universität Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany. Email: {mehl,mehrmann}@math.tu-berlin.de.

§Afdeling Wiskunde, Faculteit der Exacte Wetenschappen, Vrije Universiteit Amsterdam, De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands, and Unit for BMI, North-West University, Potchefstroom, South Africa. E-mail: ran@few.vu.nl.

¶College of William and Mary, Department of Mathematics, P.O.Box 8795, Williamsburg, VA 23187-8795, USA. E-mail: lxrodm@gmail.com. *The research of this author was supported by* by Plumeri Award for Faculty Excellence at the College of William and Mary.

*This research was supported by *Deutsche Forschungsgemeinschaft*, through the DFG Research Center Matheon *Mathematics for key technologies* in Berlin.

Let $\mathbb{F}$ denote either the field of complex numbers $\mathbb{C}$ or the field of real numbers $\mathbb{R}$ and let I_n denote the $n \times n$ identity matrix. The superscript $(\cdot)^T$ denotes the transpose and $(\cdot)^*$ denotes the conjugate transpose of a matrix or vector. We will sometimes use the superscript $(\cdot)^\star$ to denote either $(\cdot)^T$ or $(\cdot)^*$. If $H \in \mathbb{F}^{n\times n}$ is an invertible matrix inducing an inner product on $\mathbb{F}^n$, then the names of important classes of matrices $A \in \mathbb{F}^{n\times n}$ with symmetry structures with respect to that inner product are listed in the following table.

	$H^* = H$	$H^T = H$	$H^T = -H$
	$\mathbb{F} = \mathbb{C}, \star = *$	$\mathbb{F} \in \{\mathbb{C}, \mathbb{R}\}, \star = T$	$\mathbb{F} \in \{\mathbb{C}, \mathbb{R}\}, \star = T$
$A^\star H = HA$	H-selfadjoint	H-symmetric	H-skew-Hamiltonian
$A^\star H = -HA$	H-skew-adjoint	H-skew-symmetric	H-Hamiltonian
$A^\star HA = H$	H-unitary	H-orthogonal	H-symplectic

Clearly, in the case of a nondegenerate skew-symmetric bilinear form, i.e., if H is invertible, the dimension n of the space $\mathbb{F}^n$ has to be even. A very important special case in applications are the classes obtained with the matrix

$$H := J := \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}.$$

In this case one typically drops the prefix "H-" in the name of the matrix classes. Hamiltonian and symplectic matrices occur in control theory, in particular linear quadratic and H_∞ optimal control, see for example [3, 19, 27] and the references therein, and in mechanics [9].

In recent years, the theory of low rank perturbations of matrices, operators, and matrix polynomials has been developed starting in the 1980's with [11, 13, 26]; and later works in this area include [1, 4, 5, 6, 21, 22, 23, 24]. Structured rank one perturbations of complex H-Hamiltonian and complex H-symmetric matrices have been discussed in [16], while H-selfadjoint matrices have been considered in [17]. The case of H-skew-adjoint matrices can be easily reduced to the case of H-selfadjoint matrices by multiplication with the imaginary unit, since A is H-skew-adjoint if and only if iA is H-selfadjoint. This trick is not possible in the case of bilinear forms, but here structured rank one perturbations for H-skew-Hamiltonian or H-skew-symmetric matrices do not make sense, because those matrices always have even rank and thus a nontrivial H-skew-Hamiltonian or H-skew-symmetric perturbation will necessarily have rank two at least. Therefore, in [12] the class of H-positive-real matrices was considered instead of the class of H-skew-symmetric matrices. This approach allowed the study of H-positive-real rank one perturbations of H-skew-symmetric matrices.

The classes of H-unitary, H-orthogonal, and H-symplectic matrices can be linked to H-selfadjoint, H-skew-symmetric, and H-Hamiltonian matrices via the so-called Cayley transformations that we will review in Section 2. These transformations can be used to carry over all results on H-selfadjoint matrices to H-unitary matrices and all results on H-Hamiltonian matrices to most H-symplectic matrices, excluding only those that have both the eigenvalues 1 and -1. The case of H-orthogonal matrices, however, takes a special role, because H-skew-symmetric matrices do not allow structured rank one perturbations. In contrast, structured rank one perturbations of H-orthogonal matrices are possible as we will show in Section 3, where we will also include two surprising examples illustrating the effect of structured rank one perturbations on H-orthogonal matrices.

Since the approach via the Cayley transformation cannot be used in that case, we will use a different approach to analyze the effects of structured rank one perturbations using canonical forms that we will present in Section 4. Based on these forms, we present three general results on generic structured rank one perturbations in Section 5 that we will apply to H-orthogonal and H-symplectic matrices in Sections 6 and 7. In Section 8, we then investigate the simplicity of new eigenvalues of the perturbed matrices from Sections 6 and 7.

Throughout the paper, we use of the following notation: The spectrum of a matrix $A \in \mathbb{F}^{n\times n}$ is denoted by $\sigma(A)$. The symbols R_n and Σ_n denote the $n \times n$ reverse identity and the $n \times n$ reverse identity with alternating signs, respectively, i.e.,

$$R_n = \begin{bmatrix} 0 & & 1 \\ & \iddots & \\ 1 & & 0 \end{bmatrix}, \quad \Sigma_n = \begin{bmatrix} 0 & & (-1)^0 \\ & \iddots & \\ (-1)^{n-1} & & 0 \end{bmatrix}.$$

For $a_0, \dots, a_{n-1} \in \mathbb{C}$ we denote by

$$\mathrm{Toep}(a_0, \dots, a_{n-1}) := \begin{bmatrix} a_0 & a_1 & \dots & a_{n-1} \\ 0 & a_0 & \ddots & \vdots \\ 0 & 0 & \ddots & a_1 \\ 0 & 0 & 0 & a_0 \end{bmatrix} \tag{1.1}$$

the $n \times n$ upper triangular Toeplitz matrix with $\begin{bmatrix} a_0 & \cdots & a_{n-1} \end{bmatrix}$ as its first row. A special case is $\mathcal{J}_n(\lambda) = \mathrm{Toep}(\lambda, 1, 0, \dots, 0)$ that is the upper triangular $n \times n$ Jordan block associated with the eigenvalue λ. It is well known that a matrix T commutes with $\mathcal{J}_n(\lambda)$ if and only if T is of the form (1.1), see [8]. The unit coordinate column vector with 1 in the ith position and zeros elsewhere will be denoted by $e_i \in \mathbb{R}^n$ (n is understood from context).

2 Cayley transformations

In this section, we review the Cayley transformations. Recall (see, e.g., [10]) that if $\alpha, w \in \mathbb{C}$ satisfy $|\alpha| = 1$ and $w \neq \overline{w}$, then for a matrix $A \in \mathbb{C}^{n \times n}$ with $w \notin \sigma(A)$ its *Cayley transformation* is given by

$$U := \mathcal{C}_{\alpha,w}(A) := \alpha(A - \overline{w}I_n)(A - wI_n)^{-1} \tag{2.1}$$

and satisfies $\alpha \notin \sigma(U)$. Its inverse transformation is given by the formula

$$A := \mathcal{C}^{-1}_{\alpha,w}(U) := (wU - \overline{w}\alpha I_n)(U - \alpha I_n)^{-1}, \tag{2.2}$$

which can be applied to all matrices U that do not have α as an eigenvalue. It is well known that if A and U are related by one of the formulas (2.1) or (2.2), then A is H-selfadjoint if and only if U is H-unitary. Cleary for any H-selfadjoint matrix A the parameter w can be chosen such that w is not in the spectrum of A and similarly for any H-unitary matrix U one can choose α excluding all unimodular eigenvalues of U. Moreover, if $U = \mathcal{C}_{\alpha,w}(A)$ and $U' = \mathcal{C}_{\alpha,w}(A')$ then

$$\begin{aligned} U' - U &= \alpha(A' - wI_n)^{-1}\big((A' - \overline{w}I_n)(A - wI_n) - (A' - wI_n)(A - \overline{w}I_n)\big)(A - wI_n)^{-1} \\ &= \alpha(\overline{w} - w)(A' - wI_n)^{-1}(A' - A)(A - wI_n)^{-1}. \end{aligned}$$

Thus, U' is a rank one perturbation of U if and only if A' is a rank one perturbation of A. Therefore, we obtain the following result:

Meta-Theorem: *For any theorem on structured rank one perturbations for H-selfadjoint matrices, there is a corresponding result for H-unitary matrices.*

We refrain from explicitly listing all those results, but refer the reader to the H-selfadjoint case discussed in [17] instead.

In the case of bilinear forms, the situation is different. Here, only the classical Cayley transformations

$$\mathcal{C}_{+1}(A) := (A + I_n)(A - I_n)^{-1}, \quad \mathcal{C}_{-1}(A) := (A - I_n)(A + I_n)^{-1}$$

that are inverses of each other, map H-Hamiltonian or H-skew-symmetric matrices to H-symplectic or H-orthogonal matrices, respectively. Clearly, $\mathcal{C}_\mu(A)$ is only defined if $\mu \in \{+1, -1\}$ is not an eigenvalue of A. Elementary calculations for $U = \mathcal{C}_\mu(A)$ and $U' = \mathcal{C}_\mu(A')$ yield

$$U' - U = 2\mu(A' - \mu I_n)^{-1}(A - A')(A - \mu I_n)^{-1}, \tag{2.3}$$

so again U' is a rank one perturbation of U if and only if A' is a rank one perturbation of A. Again, we can use this observation to carry over results on structured rank one perturbations from H-Hamiltonian matrices to H-symplectic matrices, but only in the case, where our H-symplectic matrix under consideration does not have both $+1$ and -1 as eigenvalues. In the case of H-orthogonal matrices, the Cayley transformations are of no use for the investigation of structured rank one perturbations, because H-skew-symmetric matrices of rank one do not exist. (At first sight, formula (2.3) may look like being a contradiction then, because rank one perturbations of H-orthogonal matrices do exist as we will show in Section 3. However, a closer look at the results presented in the following sections reveals that if U' and U are H-orthogonal such that $U' - U$ has rank one, then both $+1$ and -1 occur in the union of the spectra of U' and U, so the formula cannot be applied in that case.)

3 H-orthogonal matrices: two surprising examples

In this section, we illustrate the effect of structured rank one perturbations on H-orthogonal matrices, where H is an invertible complex symmetric matrix. We start with a lemma that characterizes structured rank one perturbations of H-orthogonal matrices.

Lemma 3.1 *Let $H \in \mathbb{F}^{n\times n}$ be an invertible symmetric matrix, and suppose that $\widetilde{U}, U \in \mathbb{F}^{n\times n}$ are H-orthogonal. If $\operatorname{rank}(U - \widetilde{U}) = 1$, then there exists a vector $u \in \mathbb{F}^n$ such that $u^T Hu \neq 0$ and*

$$\widetilde{U} = (I - \frac{2}{u^T Hu} uu^T H)U. \tag{3.1}$$

Conversely, for any $u \in \mathbb{F}^n$ with $u^T Hu \neq 0$, such a matrix $\widetilde{U}$ is H-orthogonal.

Proof. Since $\widetilde{U} - U$ has rank one, there exists two nonzero vectors $u, v \in \mathbb{F}^n$ such that $\widetilde{U} = U + uv^T$. Writing out the identity $\widetilde{U}^T H\widetilde{U} - H = 0$ in terms of U and u, v, we obtain

$$U^T Huv^T + vu^T HU + vu^T Huv^T = 0.$$

Multiplying this equation from the right by v, we obtain

$$U^T Huv^T v + vu^T HUv + vu^T Huv^T v = (v^T v) \cdot (U^T Hu) + (u^T HUv + u^T Huv^T v) \cdot v = 0. \tag{3.2}$$

This identity states that the vectors $U^T Hu$ and v are linearly dependent. Since v is nonzero, this implies the existence of a constant c such that $cv = U^T Hu$, in fact,

$$c = -\frac{u^T HUv}{v^T v} - u^T Hu$$

by (3.2). Replacing in the latter expression the formula $u^T HU$ with cv^T, we see that $c = -c - u^T Hu$, i.e., $c = -\frac{u^T Hu}{2}$. In particular, $u^T Hu \neq 0$ (otherwise $U^T Hu$ and thus u would be zero), and also formula (3.1) follows.

Conversely, we have

$$\begin{aligned} & (I - \frac{2}{u^T Hu} uu^T H)^T H(I - \frac{2}{u^T Hu} uu^T H) = \\ = \ & H - \frac{4}{u^T Hu} Huu^T H + \frac{4}{(u^T Hu)^2} H^T uu^T Huu^T H = H, \end{aligned}$$

i.e., $(I - \frac{2}{u^T Hu} uu^T H)$ is H-orthogonal which implies the H-orthogonality of $\widetilde{U}$ in (3.1). $\square$

Remark 3.2 Let $H \in \mathbb{F}^{n\times n}$ be invertible and symmetric and $u \in \mathbb{F}^n$ with $u^T H u \neq 0$. Setting

$$E_H := I - \frac{2}{u^T H u} u u^T H, \tag{3.3}$$

it is an easy computation to see that $E_H^2 = I$, i.e., E_H is its own inverse, and

$$\det E_H = \det(I - \frac{2}{u^T H u} u u^T H) = 1 - \frac{2}{u^T H u}(u^T H u) = -1.$$

The fact that $\det E_H = -1$ has an important consequence. It is well known that the group of real H-orthogonal matrices has four connected components if H is indefinite and has two connected components if H is (positive or negative) definite; see, e.g. [10, Section 6.5], whereas the group of complex H-orthogonal matrices has two connected components characterized by the value of the determinant that can assume the two values 1 and -1. (This fact can be deduced from a topological isomorphism between group of complex I-orthogonal $n \times n$ matrices and the product of the real I-orthogonal $n\times n$ group with $\mathbb{R}^{n(n-1)/2}$, see for example[20, Section 1.4].) Since $\det E_H = -1$, it follows that a rank one perturbation $\widetilde{S} = E_H S$ will result in a change of the sign of the determinant, i.e., the perturbed matrix will be from a different connected component than the original one. In particular this means that *there do not exist structured rank one perturbations of arbitrarily small norm*, because a perturbation of sufficiently small norm would stay in the component of the original matrix.

The following examples illustrate the interesting effect of structured rank one perturbations of H-orthogonal matrices.

Example 3.3 Consider the matrices

$$S = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad H = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Then S is H-orthogonal and $\det S = 1$. If $u \in \mathbb{F}^4$ is such that $u^T H u \neq 0$ and E is as in (3.3), then generically $\widetilde{S} = ES$ has the Jordan canonical form

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}.$$

This is in contrast to the case of unstructured perturbations which generically would have resulted in a Jordan canonical form with one 2×2 block associated with the eigenvalue 1 and two simple eigenvalues distinct from 1, see [16, 21].

To explain this phenomenon, we shall show that in this example there is indeed a Jordan chain of length three, and that one eigenvalue moves from 1 to -1.

Since the geometric multiplicity of each eigenvalue of a rank one perturbed matrix can only reduce by at most one (see, for example, [16]), we know that 1 is still in the spectrum of $\widetilde{S}$. Thus, let x_0 be a vector for which $\widetilde{S}x_0 = x_0$. This is equivalent to $ESx_0 = x_0$, i.e., to $Sx_0 = Ex_0$, where we have used that $E^{-1} = E$. Now

$$Sx_0 = \begin{bmatrix} x_{01} + x_{02} \\ x_{02} \\ x_{03} + x_{04} \\ x_{04} \end{bmatrix} = Ex_0 = \begin{bmatrix} x_{01} \\ x_{02} \\ x_{03} \\ x_{04} \end{bmatrix} - \frac{2u^T H x_0}{u^T H u} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix},$$

where $x_0 = \begin{bmatrix} x_{01} & x_{02} & x_{03} & x_{04} \end{bmatrix}^T$, $u = \begin{bmatrix} u_1 & u_2 & u_3 & u_4 \end{bmatrix}^T$. As u is a generic vector we may assume that u_2 and u_4 are nonzero. Comparing the second coordinates in the equation above, we see that we must have $u^T H x_0 = 0$ and $Sx_0 = x_0$. Now $u^T H x_0 = x_{01}u_4 - x_{03}u_2 = 0$, so we may take without loss of generality $x_0 = \begin{bmatrix} u_2 & 0 & u_4 & 0 \end{bmatrix}^T$. Next, we need to determine a vector $x_1 = \begin{bmatrix} x_{11} & x_{12} & x_{13} & x_{14} \end{bmatrix}^T$ so that $\widetilde{S}x_1 = x_1 + x_0$. Since $u^T H x_0 = 0$ this is equivalent to

$$\begin{aligned} Sx_1 &= E^{-1}(x_1 + x_0) \;=\; Ex_1 + Ex_0 = \\ &= Ex_1 + x_0 - \frac{2u^T H x_0}{u^T H u} u \;=\; Ex_1 + x_0 \;=\; x_1 + x_0 - \frac{2u^T H x_1}{u^T H u} u. \end{aligned}$$

Now

$$Sx_1 = \begin{bmatrix} x_{11} + x_{12} \\ x_{12} \\ x_{13} + x_{14} \\ x_{14} \end{bmatrix} = x_0 + \begin{bmatrix} x_{11} \\ x_{12} \\ x_{13} \\ x_{14} \end{bmatrix} - \frac{2u^T H x_1}{u^T H u} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix},$$

and comparing again the second coordinates we see that $u^T H x_1 = 0$. From the first and the third coordinates we get $x_{12} = u_2$ and $x_{14} = u_4$. Then

$$0 = u^T H x_1 = u_4 x_{11} - u_3 x_{12} - u_2 x_{13} + u_1 x_{14} = u_4 x_{11} - u_3 u_2 - u_2 x_{13} + u_1 u_4,$$

so, without loss of generality we may take $x_1 = \begin{bmatrix} -u_1 & u_2 & -u_3 & u_4 \end{bmatrix}^T$. Continuing, we determine a vector $x_2 = \begin{bmatrix} x_{21} & x_{22} & x_{23} & x_{24} \end{bmatrix}^T$ such that $\widetilde{S}x_2 = x_2 + x_1$. Again, using that $u^T H x_1 = 0$, this is equivalent to

$$Sx_2 = Ex_2 + Ex_1 = x_2 + x_1 - \frac{2u^T H x_2}{u^T H u} u.$$

Expressing this in coordinates it becomes

$$\begin{bmatrix} x_{21} + x_{22} \\ x_{22} \\ x_{23} + x_{24} \\ x_{24} \end{bmatrix} = Ex_2 + Ex_1 = \begin{bmatrix} x_{21} \\ x_{22} \\ x_{23} \\ x_{24} \end{bmatrix} + \begin{bmatrix} -u_1 \\ u_2 \\ -u_3 \\ u_4 \end{bmatrix} - \frac{2u^T H x_2}{u^T H u} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix}.$$

Considering the second and fourth coordinate we must have

$$\frac{2u^T H x_2}{u^T H u} = 1.$$

Using this, the first coordinate gives $x_{22} = -2u_1$, and the third coordinate gives $x_{24} = -2u_3$. Inserting this back into the equation $2u^T H x_2 = u^T H u$, we obtain

$$u^T H x_2 = u_4 x_{21} + 2u_1 u_3 - u_2 x_{23} - 2u_1 u_3 = \frac{1}{2} u^T H u = u_4 u_1 - u_3 u_2.$$

Obviously, we may take $x_{21} = u_1$ and $x_{23} = u_3$, so $x_2 = \begin{bmatrix} u_1 & -2u_1 & u_3 & -2u_3 \end{bmatrix}^T$. In conclusion, we have obtained a Jordan chain x_0, x_1, x_2 of length three for $\widetilde{S}$ corresponding to the eigenvalue 1. For the characteristic polynomial of $\widetilde{S}$ we obtain

$$\begin{aligned} p_{\widetilde{S}}(\lambda) &= \det(\lambda I - \widetilde{S}) = \det(\lambda I - ES) = \det(E)\det(\lambda E - S) = \\ &= -\det(\lambda I - S - \frac{2\lambda}{u^T H u} u u^T H) = \\ &= -\det(\lambda I - S)\det(I - (\lambda I - S)^{-1} \frac{2\lambda}{u^T H u} u u^T H) = \\ &= -\det(\lambda I - S)(1 - \frac{2\lambda}{u^T H u} u^T H (\lambda I - S)^{-1} u), \end{aligned}$$

and a direct computation gives

$$(\lambda I - S)^{-1}u = \begin{bmatrix} \frac{1}{\lambda-1}u_1 - \frac{1}{(\lambda-1)^2}u_2 \\ \frac{1}{\lambda-1}u_2 \\ \frac{1}{\lambda-1}u_3 - \frac{1}{(\lambda-1)^2}u_4 \\ \frac{1}{\lambda-1}u_4 \end{bmatrix}.$$

So,

$$\begin{aligned} u^T H(\lambda I - S)^{-1}u &= \frac{1}{\lambda-1}u_1u_4 - \frac{1}{(\lambda-1)^2}u_2u_4 - \frac{1}{\lambda-1}u_2u_3 + \\ &\quad -\frac{1}{\lambda-1}u_3u_2 + \frac{1}{(\lambda-1)^2}u_4u_2 + \frac{1}{\lambda-1}u_4u_1 \\ &= \frac{2}{\lambda-1}(u_1u_4 - u_2u_3) = \frac{1}{\lambda-1}u^T Hu, \end{aligned}$$

and it follows that

$$p_{\widetilde{S}}(\lambda) = -(\lambda-1)^4(1 - \frac{2\lambda}{\lambda-1}) = (\lambda-1)^3(\lambda+1).$$

The second example is much simpler, but even more surprising.

Example 3.4 Let $\lambda \in \mathbb{F} \setminus \{0\}$ be arbitrary and consider

$$U = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{bmatrix}, \quad H = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \in \mathbb{F}^2.$$

Then U is H-orthogonal and $\det U = 1$. Furthermore, assume that $u^T Hu = 2u_1u_2 \neq 0$. If E is as in (3.3) then

$$E = I_2 - \frac{1}{u_1u_2}\begin{bmatrix} u_1u_2 & u_1^2 \\ u_2^2 & u_1u_2 \end{bmatrix} = \begin{bmatrix} 0 & -\frac{u_1}{u_2} \\ -\frac{u_2}{u_1} & 0 \end{bmatrix}.$$

Thus we obtain that

$$EU = -\begin{bmatrix} 0 & \frac{u_1}{u_2}\lambda^{-1} \\ \frac{u_2}{u_1}\lambda & 0 \end{bmatrix}$$

and this matrix has the eigenvalues $+1$ and -1.

In both examples we have the surprising fact that generically all rank one perturbations of the matrix under consideration will have identical spectrum.

4 Canonical forms

In order to fully explain the phenomena observed in the previous section in the complex case, we will need canonical and simple forms for complex H-orthogonal matrices and H-symplectic matrices, where H is invertible and symmetric, respectively skew-symmetric. We start with canonical forms as they were presented in [15, Theorems 7.5 and 8.5].

Theorem 4.1 (Canonical form for H-orthogonal matrices) *Let $H = H^T$ be an invertible $n \times n$ complex matrix and let $U \in \mathbb{C}^{n\times n}$ be H-orthogonal. Then there exists a nonsingular complex matrix Q such that*

$$Q^{-1}UQ = U_1 \oplus \cdots \oplus U_p, \quad Q^T HQ = H_1 \oplus \cdots \oplus H_p, \tag{4.1}$$

where U_j and H_j have one of the following forms:

1) *blocks associated with eigenvalue* $\lambda_j = \delta = \pm 1$ *of* U *with size* n_j*, where* $n_j \in \mathbb{N}$ *is odd:*

$$U_j = \operatorname{Toep}(\delta, 1, r_2, \ldots, r_{n_j-1}), \quad H_j = \Sigma_{n_j}. \tag{4.2}$$

Moreover, $r_k = 0$ *for odd* k *and the parameters* r_k *for even* k *are real and uniquely determined by the recursive formula*

$$r_2 = \frac{1}{2}\delta, \quad r_k = -\frac{1}{2}\delta \left(\sum_{\nu=1}^{\frac{k}{2}-1} r_{2\cdot\nu} r_{2\cdot(\frac{k}{2}-\nu)} \right), \quad 4 \le k \le n_j;$$

2) *paired blocks associated with eigenvalues* $\lambda_j = \pm 1$*, of size* m_j*, where* $m_j \in \mathbb{N}$ *is even:*

$$U_j = \begin{bmatrix} \mathcal{J}_{m_j}(\lambda_j) & 0 \\ 0 & \left(\mathcal{J}_{m_j}(\lambda_j)\right)^{-T} \end{bmatrix}, \quad H_j = \begin{bmatrix} 0 & I_{m_j} \\ I_{m_j} & 0 \end{bmatrix}, \tag{4.3}$$

3) *blocks associated with a pair of eigenvalues* $(\lambda_j, \lambda_j^{-1}) \in \mathbb{C} \times \mathbb{C}$*, where* $\operatorname{Re}(\lambda_j) > \operatorname{Re}(\lambda_j^{-1})$ *or* $\operatorname{Im}(\lambda_j) > \operatorname{Im}(\lambda_j^{-1})$ *if* $\operatorname{Re}(\lambda_j) = \operatorname{Re}(\lambda_j^{-1})$*, and* $m_j \in \mathbb{N}$*:*

$$U_j = \begin{bmatrix} \mathcal{J}_{m_j}(\lambda_j) & 0 \\ 0 & \left(\mathcal{J}_{m_j}(\lambda_j)\right)^{-T} \end{bmatrix}, \quad H_j = \begin{bmatrix} 0 & I_{m_j} \\ I_{m_j} & 0 \end{bmatrix}.$$

Moreover, the form (4.1) *is unique up to the permutation of blocks. (We highlight that a fixed eigenvalue* μ *may occur multiple times among* $\lambda_1, \ldots, \lambda_p$*. Also, a block associated with* μ *and a fixed size* m *may appear multiple times among the blocks* $U_1, \ldots, U_p$*.)*

Theorem 4.2 (Canonical form for H-symplectic matrices) *Let* $H = -H^T$ *be an invertible* $n \times n$ *complex matrix and let* $S \in \mathbb{C}^{n \times n}$ *be* H*-symplectic. Then* n *is even and there exists a nonsingular complex matrix* Q *such that*

$$Q^{-1} S Q = S_1 \oplus \cdots \oplus S_p, \quad Q^T H Q = H_1 \oplus \cdots \oplus H_p, \tag{4.4}$$

where S_j *and* H_j *have one of the following forms:*

i) even-sized blocks associated with the eigenvalue $\lambda_j = \delta = \pm 1$*, of* S *with size* n_j*, where* $n_j \in \mathbb{N}$ *is even:*

$$S_j = T(\delta, 1, r_2, \ldots, r_{n_j-1}), \quad H_j = \Sigma_{n_j},$$

Moreover, $r_k = 0$ *for odd* k *and the parameters* r_k *for even* k *are real and uniquely determined by the recursive formula*

$$r_2 = \frac{1}{2}\delta, \quad r_k = -\frac{1}{2}\delta \left(\sum_{\nu=1}^{\frac{k}{2}-1} r_{2\cdot\nu} r_{2\cdot(\frac{k}{2}-\nu)} \right), \quad 4 \le k \le n_j;$$

ii) paired blocks associated with the eigenvalues $\lambda_j = \pm 1$*, of* S *with size* m_j*, where* $m_j \in \mathbb{N}$ *is odd:*

$$S_j = \begin{bmatrix} \mathcal{J}_{m_j}(\lambda_j) & 0 \\ 0 & \left(\mathcal{J}_{m_j}(\lambda_j)\right)^{-T} \end{bmatrix}, \quad H_j = \begin{bmatrix} 0 & I_{m_j} \\ -I_{m_j} & 0 \end{bmatrix};$$

iii) blocks associated with a pair of eigenvalues $(\lambda_j, \lambda_j^{-1}) \in \mathbb{C} \times \mathbb{C}$*, satisfying* $\operatorname{Re}(\lambda_j) > \operatorname{Re}(\lambda_j^{-1})$ *or* $\operatorname{Im}(\lambda_j) > \operatorname{Im}(\lambda_j^{-1})$ *if* $\operatorname{Re}(\lambda_j) = \operatorname{Re}(\lambda_j^{-1})$*, where* $m_j \in \mathbb{N}$*:*

$$S_j = \begin{bmatrix} \mathcal{J}_{m_j}(\lambda_j) & 0 \\ 0 & \left(\mathcal{J}_{m_j}(\lambda_j)\right)^{-T} \end{bmatrix}, \quad H_j = \begin{bmatrix} 0 & I_{m_j} \\ -I_{m_j} & 0 \end{bmatrix}.$$

Moreover, the form (4.4) is unique up to the permutation of blocks. (We highlight that a fixed eigenvalue μ may occur multiple times among $\lambda_1, \ldots, \lambda_p$. Also, a block associated with μ and a fixed size m may appear multiple times among the blocks $U_1, \ldots, U_p$.)

Although the canonical forms in Theorem 4.1 and 4.2 display all invariants of the matrix pair under consideration, it will be necessary for the purpose of this paper to further investigate the blocks associated with the eigenvalues ± 1. We start by presenting results on all possible symmetric or skew-symmetric matrices H for which a Jordan block associated with the eigenvalue 1 is H-orthogonal or H-symplectic, respectively.

Proposition 4.3 *Let $n = 2k+1$, where $k \in \mathbb{N}$, and let $\mathcal{J}_n(1)$ be the upper triangular Jordan block of size n with eigenvalue 1. Then the set*

$$\mathcal{V}_n := \{H \in \mathbb{C}^{n \times n} \,|\, \mathcal{J}_n(1)^T H \mathcal{J}_n(1) = H \text{ and } H^T = H\}$$

is a vector space of dimension $k+1$. In particular, any $H = [h_{ij}] \in \mathcal{V}_n$ has the form

$$H = \begin{bmatrix} 0 & 0 & h_{1n} \\ 0 & H_{n-2} & h_n \\ h_{1n} & h_n^T & h_{nn} \end{bmatrix}, \quad h_n := \begin{bmatrix} h_{2n} \\ \vdots \\ h_{n-1,n} \end{bmatrix},$$

where $H_{n-2} \in \mathcal{V}_{n-2}$ and

$$h_{1n} = -h_{2,n-1}; \quad h_{jn} = -h_{j,n-1} - h_{j+1,n-1} \text{ for } j = 2, \ldots, n-2; \quad h_{n-1,n} = -\frac{1}{2} h_{n-1,n-1}, \tag{4.5}$$

and where $h_{n,n} \in \mathbb{C}$ is arbitrary. Moreover, H is uniquely determined by the diagonal elements $h_{k+1,k+1}, \ldots, h_{n,n}$ and for each $m = k+1, \ldots, n$, the entries h_{ij} depending on h_{mm} are only those satisfying $i + j \geq 2m$ and $\min\{i, j\} \leq m$. In particular,

$$h_{jn} = (-1)^{j-1} \frac{1}{2} h_{jj} + \beta_{j+1,j} h_{j+1,j+1} + \cdots + \beta_{n,j} h_{nn} \tag{4.6}$$

for some coefficients β_{ij}, $i = j+1, \ldots, n$ for $j = k+1, \ldots, n-1$.

Proof The proof proceeds by induction on k. For $k = 0$ the result is obvious. Thus, let $k > 0$ and $H_n \in \mathcal{V}_n$, and partition

$$H = (h_{ij}) = \begin{bmatrix} h_{11} & g_n^T & h_{1n} \\ g_n & H_{n-2} & h_n \\ h_{1n} & h_n^T & h_{nn} \end{bmatrix} \quad \text{and} \quad \mathcal{J}_n(1) = \begin{bmatrix} 1 & e_1^T & 0 \\ 0 & \mathcal{J}_{n-2}(1) & e_{n-2} \\ 0 & 0 & 1 \end{bmatrix},$$

Using $\mathcal{J}_{n-2}(1) - I_{n-2} = \mathcal{J}_{n-2}(0)$, we obtain

$$0 = \mathcal{J}_n(1)^T H_n \mathcal{J}_n(1) - H_n = \begin{bmatrix} 0 & h_{11} e_1^T + g_n^T \mathcal{J}_{n-2}(0) & g_n^T e_{n-2} \\ h_{11} e_1 + \mathcal{J}_{n-2}(0)^T g_n & * & * \\ e_{n-2}^T g_n & * & * \end{bmatrix}. \tag{4.7}$$

This implies $e_{n-2}^T g_n = 0$ (i.e., the last entry of the vector g_n is zero) and $h_{11} e_1 + \mathcal{J}_{n-2}(0)^T g_n = 0$. The latter identity implies that the first $n-1$ entries of g_n are zero (and consequently $g_n = 0$) and that h_{11} is zero. Thus, the identity (4.7) reduces to

$$0 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \mathcal{J}_{n-2}(1)^T H_{n-2} \mathcal{J}_{n-2}(1) - H_{n-2} & \mathcal{J}_{n-2}(1)^T H_{n-2} e_{n-2} + \mathcal{J}_{n-2}(0)^T h_n + h_{1n} e_1 \\ 0 & e_{n2}^T H_{n-2} \mathcal{J}_{n-2}(1) + h_n^T \mathcal{J}_{n-2}(0) + h_{1n} e_1^T & e_{n-2}^T H_{n-2} e_{n-2} + h_n^T e_{n-2} + e_{n-2}^T h_n \end{bmatrix},$$

which in particular implies that $H_{n-2} \in \mathcal{V}_{n-2}$. Furthermore, we obtain the equations

$$0 = e_{n-2}^T H_{n-2} e_{n-2} + h_n^T e_{n-2} + e_{n-2}^T h_n = h_{n-1,n-1} + 2h_{n-1,n}$$

and

$$0 = \mathcal{J}_{n-2}(1)^T H_{n-2} e_{n-2} + \mathcal{J}_{n-2}(0)^T h_n + h_{1n} e_1 = \begin{bmatrix} h_{2,n-1} \\ h_{2,n-1} + h_{3,n-1} \\ \vdots \\ h_{n-2,n-1} + h_{n-1,n-1} \end{bmatrix} + \begin{bmatrix} 0 \\ h_{2,n} \\ \vdots \\ h_{n-2,n} \end{bmatrix} + \begin{bmatrix} h_{1,n} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

which both together imply (4.5). Thus, h_{jn} is uniquely determined for $j = 1, \ldots, n-1$ and h_{nn} is arbitrary. Using the induction hypothesis on H_{n-2} and in particular, that H_{n-2} is uniquely determined by the diagonal elements $h_{k+1,k+1}, \ldots, h_{n-1,n-1}$ and for each $m = k+1, \ldots, n-1$, the entries h_{ij} depending on h_{mm} are only those satisfying $i + j \geq 2m$ and $\min\{i,j\} \leq m$, the claim concerning the entries depending on h_{mm} follows directly from (4.5). Similarly, (4.6) follows by induction using (4.5).
□

Example 4.4 Let $k = 2$, i.e., $n = 2k + 1 = 5$. Then any $H \in \mathcal{V}_5$ has the form

$$H = h_{33} \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 & \frac{3}{2} \\ 0 & 0 & 1 & -\frac{1}{2} & \frac{1}{2} \\ 0 & -1 & -\frac{1}{2} & 0 & 0 \\ 1 & \frac{3}{2} & \frac{1}{2} & 0 & 0 \end{bmatrix} + h_{44} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & -\frac{1}{2} \\ 0 & 0 & -1 & -\frac{1}{2} & 0 \end{bmatrix} + h_{55} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

for some parameters h_{33}, h_{44}, h_{55}.

Proposition 4.5 *Let $n = 2k$, where $k \in \mathbb{N}$, and let $\mathcal{J}_n(1)$ be the upper triangular Jordan block of size n with eigenvalue 1. Then the set*

$$\mathcal{U}_n := \{H \in \mathbb{C}^{n \times n} \,|\, \mathcal{J}_n(1)^T H \mathcal{J}_n(1) = H \text{ and } H^T = -H\}$$

is a vector space of dimension k. In particular, any $H = [h_{ij}] \in \mathcal{U}_n$ has the form

$$H = \begin{bmatrix} 0 & 0 & h_{1n} \\ 0 & H_{n-2} & h_n \\ -h_{1n} & -h_n^T & 0 \end{bmatrix}, \quad h_n := \begin{bmatrix} h_{2n} \\ \vdots \\ h_{n-1,n} \end{bmatrix},$$

where $H_{n-2} \in \mathcal{U}_{n-2}$ and

$$h_{1n} = -h_{2,n-1}; \quad h_{jn} = -h_{j,n-1} - h_{j+1,n-1} \text{ for } j = 2, \ldots, n-3; \quad h_{n-2,n} = -h_{n-2,n-1}, \tag{4.8}$$

and where $h_{n-1,n} \in \mathbb{C}$ is arbitrary. Moreover, H is uniquely determined by $h_{k,k+1}, \ldots, h_{n-1,n}$ and for each $m = k, \ldots, n-1$, the entries h_{ij} depending on $h_{m,m+1}$ are only those satisfying $i + j \geq 2m + 1$ and $\min\{i,j\} \leq m$. In particular,

$$h_{jn} = (-1)^{j-1} h_{j,j+1} + \beta_{j+1,j} h_{j+1,j+2} + \cdots + \beta_{n-1,j} h_{n-1,n} \tag{4.9}$$

for some coefficients β_{ij}, $i = j+1, \ldots, n$ for $j = k, \ldots, n-1$.

Proof. The proof proceeds by induction on k. For $k = 1$ the result is obvious. Thus, let $k > 1$ and $H_n \in \mathcal{U}_n$, and partition

$$H = (h_{ij}) = \begin{bmatrix} 0 & g_n^T & h_{1n} \\ -g_n & H_{n-2} & h_n \\ -h_{1n} & -h_n^T & 0 \end{bmatrix} \quad \text{and} \quad \mathcal{J}_n(1) = \begin{bmatrix} 1 & e_1^T & 0 \\ 0 & \mathcal{J}_{n-2}(1) & e_{n-2} \\ 0 & 0 & 1 \end{bmatrix},$$

As in the proof of Proposition 4.3, we obtain

$$0 = \mathcal{J}_n(1)^T H_n \mathcal{J}_n(1) - H_n = \begin{bmatrix} 0 & g_n^T \mathcal{J}_{n-2}(0) & g_n^T e_{n-2} \\ -\mathcal{J}_{n-2}(0)^T g_n & * & * \\ -e_{n-2}^T g_n & * & * \end{bmatrix}, \tag{4.10}$$

which implies $g_n = 0$. Thus, the identity (4.10) reduces to

$$0 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \mathcal{J}_{n-2}(1)^T H_{n-2} \mathcal{J}_{n-2}(1) - H_{n-2} & \mathcal{J}_{n-2}(1)^T H_{n-2} e_{n-2} + \mathcal{J}_{n-2}(0)^T h_n + h_{1n} e_1 \\ 0 & -e_{n2}^T H_{n-2} \mathcal{J}_{n-2}(1) - h_n^T \mathcal{J}_{n-2}(0) - h_{1n} e_1^T & e_{n-2}^T H_{n-2} e_{n-2} - h_n^T e_{n-2} + e_{n-2}^T h_n \end{bmatrix},$$

which in particular implies that $H_{n-2} \in \mathcal{U}_{n-2}$. The identity in the $(3,3)$-block being trivial, we obtain

$$0 = \mathcal{J}_{n-2}(1)^T H_{n-2} e_{n-2} + \mathcal{J}_{n-2}(0)^T h_n + h_{1n} e_1 = \begin{bmatrix} h_{2,n-1} \\ h_{2,n-1} + h_{3,n-1} \\ \vdots \\ h_{n-3,n-1} + h_{n-2,n-1} \\ h_{n-2,n-1} \end{bmatrix} + \begin{bmatrix} 0 \\ h_{2,n} \\ \vdots \\ h_{n-3,n} \\ h_{n-2,n} \end{bmatrix} + \begin{bmatrix} h_{1,n} \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}$$

which both together imply (4.8). Thus, h_{jn} is uniquely determined for $j = 1, \ldots, n-2$ and $h_{n-1,n}$ is arbitrary (and $h_{nn} = 0$). Using the induction hypothesis on H_{n-2}, the claim concerning the entries depending on $h_{m,m+1}$ follows directly from (4.8). Similarly, (4.9) follows by induction using (4.8). □

Example 4.6 Let $k = 3$, i.e., $n = 2k = 6$. Then each $H \in \mathcal{U}_6$ has the form

$$H = h_{34} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 2 \\ 0 & 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ -1 & -2 & -1 & 0 & 0 & 0 \end{bmatrix} + h_{45} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \end{bmatrix} + h_{56} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 \end{bmatrix}$$

for some parameters h_{34}, h_{45}, h_{56}.

As a consequence of Theorem 4.1 as well as Propositions 4.3 and 4.5, we obtain the following partial simple form, where blocks associated with the same eigenvalue are grouped together and ordered by size, in contrast to the forms of Theorem 4.1 and 4.2.

Corollary 4.7 *Let $H = H^T$ be an $n \times n$ invertible complex matrix. Let $U \in \mathbb{C}^{n \times n}$ be H-orthogonal and let $\lambda \in \mathbb{C}$ be an eigenvalue of U with partial multiplicities $n_1 > \cdots > n_m$ occurring with the multiplicities $\ell_1, \ldots, \ell_m$, respectively, i.e., the algebraic multiplicity of λ is $a = \ell_1 n_1 + \cdots + \ell_m n_m$. Then there exists a nonsingular complex matrix $\widetilde{Q}$ such that*

$$\widetilde{Q}^{-1} U \widetilde{Q} = \widehat{U} \oplus \widetilde{U}, \quad \widetilde{Q}^T H \widetilde{Q} = \widehat{H} \oplus \widetilde{H},$$

where $\sigma(\widehat{U}) = \{\lambda, \frac{1}{\lambda}\}$, $\sigma(\widetilde{U}) \subseteq \mathbb{C} \setminus \{\lambda, \frac{1}{\lambda}\}$, and where $\widehat{U}$ and $\widehat{H}$ have the same size and the following forms:

1) *if $\lambda \notin \{+1, -1\}$ then*

$$\widehat{U} = \begin{bmatrix} U_1 & 0 \\ 0 & U_1^{-T} \end{bmatrix}, \quad \widehat{H} = \begin{bmatrix} 0 & I_a \\ I_a & 0 \end{bmatrix},$$

where

$$U_1 = \left(\bigoplus_{j=1}^{\ell_1} \mathcal{J}_{n_1}(\lambda) \right) \oplus \left(\bigoplus_{j=1}^{\ell_2} \mathcal{J}_{n_2}(\lambda) \right) \oplus \cdots \oplus \left(\bigoplus_{j=1}^{\ell_m} \mathcal{J}_{n_m}(\lambda) \right);$$

2) *if* $\lambda \in \{+1, -1\}$ *then*

$$\widehat{U} = \lambda\left(U^{(1)} \oplus U^{(2)} \oplus \cdots \oplus U^{(m)}\right), \qquad \widehat{H} = H^{(1)} \oplus H^{(2)} \oplus \cdots \oplus H^{(m)}, \tag{4.11}$$

where the matrices $U^{(i)}, H^{(i)}$, $i = 1, \ldots, m$ *have the following forms:*

2a) *if* n_i *is odd, then*

$$U^{(i)} = \bigoplus_{j=1}^{\ell_i} \mathcal{J}_{n_i}(1), \quad H^{(i)} = \bigoplus_{j=1}^{\ell_i} H^{(i,j)}, \tag{4.12}$$

where $H^{(i,j)} \in \mathcal{V}_{n_i}$ *with* $(1, n_i)$*-entry* $h^{(i,j)}_{1,n_i} \neq 0$ *for* $k = 1, \ldots, n_i$ *and* $j = 1, 2, \ldots, \ell_i$*, and where* $\mathcal{V}_{n_i}$ *is as in Proposition* 4.3*;*

2b) *if* n_i *is even, then* ℓ_i *is even and*

$$U^{(i)} = \bigoplus_{s=1}^{\frac{1}{2}\ell_i} \begin{bmatrix} \mathcal{J}_{n_i}(1) & 0 \\ 0 & \mathcal{J}_{n_i}(1) \end{bmatrix}, \quad H^{(i)} = \bigoplus_{s=1}^{\frac{1}{2}\ell_i} \begin{bmatrix} 0 & H^{(i,s)} \\ -H^{(i,s)} & 0 \end{bmatrix}, \tag{4.13}$$

where $H^{(i,s)} \in \mathcal{U}_{n_i}$ *with* $(1, n_i)$*-entry* $h^{(i,s)}_{1,n_i} \neq 0$ *for* $k = 1, \ldots, n_i$ *and* $s = 1, 2, \ldots, \frac{1}{2}\ell_i$*, and where* $\mathcal{U}_{n_i}$ *is as in Proposition* 4.5*. (Note that the matrix* $H^{(i)}$ *in (4.13) is indeed symmetric, because* $H^{(i,s)} \in \mathcal{U}_{n_i}$ *is skew-symmetric.)*

Proof. Part 1) immediately follows from Theorem 4.1 by applying appropriate block permutations.

For Part 2), it is sufficient to consider the case $\lambda = 1$, because the corresponding argument for the case $\lambda = -1$ follows from considering $-U$ instead of U. Next consider a single pair $(\mathcal{J}_{n_i}(1), H^{(i,j)})$ of blocks as in (4.12). Obviously, $H^{(i,j)}$ is symmetric and invertible, and by Proposition 4.3 it follows that $\mathcal{J}_{n_i}(1)$ is $H^{(i,j)}$-orthogonal. Applying Theorem 4.1 to the pair $(\mathcal{J}_{n_i}(1), H^{(i,j)})$, we find that there exists a nonsingular matrix Q_{ij} such that

$$Q_{ij}^{-1}\mathcal{J}_{n_i}(1)Q_{ij} = \mathrm{Toep}(1, 1, r_2, r_3, \ldots, r_{n_i-1}) \quad \text{and} \quad Q_{ij}^T H^{(i,j)} Q_{ij} = \Sigma_{n_i},$$

where $r_2, \ldots, r_{n_i-1}$ are as in (4.2). But this means that in the canonical form of Theorem 4.1, we can replace a pair $(\mathrm{Toep}(1, 1, r_2, \ldots, r_{n_i-1}), \Sigma_{n_i})$ of blocks of the form (4.2) by the equivalent pair $(\mathcal{J}_{n_i}(1), H^{(i,j)})$ with blocks as in (4.12). A similar argument shows that a pair of blocks of the form (4.3) can be replaced by an equivalent pair

$$\left(\begin{bmatrix} \mathcal{J}_{n_i}(1) & 0 \\ 0 & \mathcal{J}_{n_i}(1) \end{bmatrix}, \begin{bmatrix} 0 & H^{(i,s)} \\ -H^{(i,s)} & 0 \end{bmatrix}\right)$$

with blocks as in (4.13). Thus, the result follows from Theorem 4.1 by applying suitable transformations on each block of the form (4.2) or (4.3) and finally applying appropriate block transformations that group those blocks together and order them by size. □

We obtain an analogous corollary for H-symplectic matrices. The proof is analogous to the one of Corollary 4.7 and is therefore omitted.

Corollary 4.8 *Let* $H = -H^T$ *be an invertible* $n \times n$ *skew-symmetric complex matrix, let* $S \in \mathbb{C}^{n\times n}$ *be* H*-symplectic, and let* $\lambda \in \mathbb{C}$ *be an eigenvalue of* S *with partial multiplicities* $n_1 > \cdots > n_m$ *occurring with the multiplicities* $\ell_1, \ldots, \ell_m$*, respectively, i.e., the algebraic multiplicity of* λ *is given by* $a = \ell_1 n_1 + \cdots + \ell_m n_m$*. Then there exists a nonsingular matrix* $\widetilde{Q}$ *such that*

$$\widetilde{Q}^{-1} S \widetilde{Q} = \widehat{S} \oplus \widetilde{S}, \quad \widetilde{Q}^T H \widetilde{Q} = \widehat{H} \oplus \widetilde{H},$$

where $\sigma(\widehat{S}) = \{\lambda, \frac{1}{\lambda}\}$*,* $\sigma(\widetilde{S}) \subseteq \mathbb{C} \setminus \{\lambda, \frac{1}{\lambda}\}$*, and where* $\widehat{S}$ *and* $\widehat{H}$ *have the same size and the following forms:*

1) *if* $\lambda \notin \{+1,-1\}$ *then*

$$\widehat{S} = \begin{bmatrix} S_1 & 0 \\ 0 & S_1^{-T} \end{bmatrix}, \quad \widehat{H} = \begin{bmatrix} 0 & I_a \\ -I_a & 0 \end{bmatrix},$$

where

$$S_1 = \left(\bigoplus_{j=1}^{\ell_1} \mathcal{J}_{n_1}(\lambda)\right) \oplus \left(\bigoplus_{j=1}^{\ell_2} \mathcal{J}_{n_2}(\lambda)\right) \oplus \cdots \oplus \left(\bigoplus_{j=1}^{\ell_m} \mathcal{J}_{n_m}(\lambda)\right);$$

2) *if* $\lambda \in \{+1,-1\}$ *then*

$$\widehat{S} = \lambda\left(S^{(1)} \oplus S^{(2)} \oplus \cdots \oplus S^{(m)}\right), \qquad \widehat{H} = H^{(1)} \oplus H^{(2)} \oplus \cdots \oplus H^{(m)}, \tag{4.14}$$

where the matrices $S^{(i)}$, $H^{(i)}$, $i = 1, \ldots, m$, *have the following forms:*

2a) *if* n_i *is even, then*

$$S^{(i)} = \bigoplus_{j=1}^{\ell_i} \mathcal{J}_{n_i}(1), \quad H^{(i)} = \bigoplus_{j=1}^{\ell_i} H^{(i,j)},$$

where $H^{(i,j)} \in \mathcal{U}_{n_i}$ *with* $(1, n_i)$*-entry* $h_{1,n_i}^{(i,j)} \neq 0$ *for* $k = 1, \ldots, n_i$ *and* $j = 1, 2, \ldots, \ell_i$, *and where* $\mathcal{U}_{n_i}$ *is as in Proposition* 4.5*;*

2b) *if* n_i *is odd, then* ℓ_i *is even and*

$$S^{(i)} = \bigoplus_{s=1}^{\frac{1}{2}\ell_i} \begin{bmatrix} \mathcal{J}_{n_i}(1) & 0 \\ 0 & \mathcal{J}_{n_i}(1) \end{bmatrix}, \quad H^{(i)} = \bigoplus_{s=1}^{\frac{1}{2}\ell_i} \begin{bmatrix} 0 & H^{(i,s)} \\ -H^{(i,s)} & 0 \end{bmatrix},$$

where $H^{(i,s)} \in \mathcal{V}_{n_i}$ *with* $(1, n_i)$*-entry* $h_{1,n_i}^{(i,s)} \neq 0$ *for* $k = 1, \ldots, n_i$ *and* $s = 1, 2, \ldots, \frac{1}{2}\ell_i$, *and where* $\mathcal{V}_{n_i}$ *is as in Proposition* 4.3.

Remark 4.9 Observe that in case $\lambda = -1$ and $n_i > 1$, the blocks $-U^{(i)}$ and $-S^{(i)}$ in (4.11) and (4.14) are not Jordan matrices; they are direct sums of the negatives of $\mathcal{J}_{n_i}(1)$. With the unitary and Hermitian diagonal matrix $\Psi_n := \Sigma_n R_n$ we have $\mathcal{J}_n(-1) = \Psi_n(-\mathcal{J}_n(1))\Psi_n$. To obtain a form analogous to (4.11) but with $\widehat{U}$ a Jordan matrix (in the case $\lambda = -1$), in Corollary 4.7 we replace $\widehat{U}$ by $U^{(1)} \oplus \cdots \oplus U^{(m)}$; replace $\mathcal{J}_{n_i}(1)$ in (4.12) and (4.13) by $\mathcal{J}_{n_i}(-1)$; and replace the requirements that $H^{(i,j)} \in \mathcal{V}_{n_i}$ in part 2a) and $H^{(i,s)} \in \mathcal{U}_{n_i}$ in part 2b), by the requirements that $H^{(i,j)} \in \Psi_{n_i}\mathcal{V}_{n_i}\Psi_{n_i}$ in part 2a) and $H^{(i,s)} \in \Psi_{n_i}\mathcal{U}_{n_i}\Psi_{n_i}$ in part 2b). The same replacements (using $\widehat{S}$ in place of $\widehat{U}$) in Corollary 4.8 will produce a form analogous to (4.14) but with a Jordan matrix in place of $\widehat{S}$, in the case $\lambda = -1$.

The partial simple forms of Corollary 4.7 and 4.8 are reminiscent of the ones in [12]. They still have some freedom in the choice of the matrices from the vector spaces $\mathcal{V}_{n_i}$ and $\mathcal{U}_{n_i}$. This freedom will become handy in the following sections. We mention in passing that a freedom in parameters has also been observed in the reduction to the canonical forms of Theorem 4.1 and 4.4. There, the freedom had been used to set the parameters r_k to zero for odd k.

5 Results on special rank one perturbations

A closer look at the canonical form from the previous section reveals that we have to deal with three different kinds of blocks: (1) blocks associated with a single eigenvalue; (2) blocks associated with a pair of distinct eigenvalues; (3) paired blocks associated with a single eigenvalue. For the first two kinds, two general results describing the behavior of structured rank one perturbations were presented in [16] and will be slightly modified below. We also add a theorem covering the third case. To keep these results as general as possible, we will use the notation * to denote either the transpose ^{T} or the conjugate transpose *.

A set $\Xi \subseteq \mathbb{F}^m$ is said to be *proper algebraic* if it is equal to the set of common zeros of a system of polynomials with coefficients in the field $\mathbb{F}$ in the variables $(w_1, \ldots, w_m) \in \mathbb{F}^m$ and does not coincide with the whole of $\mathbb{F}^m$. Clearly, any proper algebraic set has Lebesgue measure zero. As in [16, 17, 22], we say that a property or a statement - which is a function of m parameters $w \in \mathbb{F}^m$ - *holds generically* if the set of those w's for which it does not hold is contained in a proper algebraic set. A vector $u \in \mathbb{F}^n$ will be called *generic* if it belongs to the complement of a set which is contained in a proper algebraic set. We have the following two theorems which extend corresponding results of [16].

Theorem 5.1 *Let $A \in \mathbb{F}^{n\times n}$ and let $T, G \in \mathbb{F}^{n\times n}$ be invertible such that*

$$T^{-1}AT = \left(\bigoplus_{j=1}^{\ell_1} \mathcal{J}_{n_1}(\widehat{\lambda})\right) \oplus \left(\bigoplus_{j=1}^{\ell_2} \mathcal{J}_{n_2}(\widehat{\lambda})\right) \oplus \cdots \oplus \left(\bigoplus_{j=1}^{\ell_m} \mathcal{J}_{n_m}(\widehat{\lambda})\right) \oplus \widetilde{A}, \tag{5.1}$$

$$T^\star GT = \left(\bigoplus_{j=1}^{\ell_1} G^{(1,j)}\right) \oplus G^{(2)} \oplus \cdots \oplus G^{(m)} \oplus \widetilde{G}, \tag{5.2}$$

where $\widehat{\lambda} \in \mathbb{F}$ and the decompositions (5.1) *and* (5.2) *have the following properties:*

(1) $n_1 > n_2 > \cdots > n_m$*;*

(2) $G^{(j)} \in \mathbb{F}^{\ell_j n_j \times \ell_j n_j}$*, $j = 2, \ldots, m$ and the matrices*

$$G^{(1,j)} = \begin{bmatrix} 0 & \cdots & 0 & g^{(1,j)}_{1,n_1} \\ \vdots & \iddots & g^{(1,j)}_{2,n_1-1} & g^{(1,j)}_{2,n_1} \\ 0 & \iddots & \iddots & \vdots \\ g^{(1,j)}_{n_1,1} & g^{(1,j)}_{n_1,2} & \cdots & g^{(1,j)}_{n_1,n_1} \end{bmatrix}, \quad j = 1, 2, \ldots, \ell_1;$$

are anti-triangular (necessarily invertible);

(3) $\widetilde{G}, \widetilde{A} \in \mathbb{F}^{(n-a)\times(n-a)}$*, where $a = \sum_{j=1}^{m} \ell_j n_j$ and $\sigma(\widetilde{A}) \subseteq \mathbb{C} \setminus \{\widehat{\lambda}\}$.*

If $B \in \mathbb{F}^{n\times n}$ is a rank one matrix of the form $B = uu^\star G$, then generically (with respect to the components of u if $\star = T$, and with respect to the real and imaginary parts of the components of u if $\star = *$), then for all $\tau \in \mathbb{F} \setminus \{0\}$ the matrix $A + \tau B$ has the Jordan canonical form

$$\left(\bigoplus_{j=1}^{\ell_1-1} \mathcal{J}_{n_1}(\widehat{\lambda})\right) \oplus \left(\bigoplus_{j=1}^{\ell_2} \mathcal{J}_{n_2}(\widehat{\lambda})\right) \oplus \cdots \oplus \left(\bigoplus_{j=1}^{\ell_m} \mathcal{J}_{n_m}(\widehat{\lambda})\right) \oplus \widetilde{\mathcal{J}}, \tag{5.3}$$

where $\widetilde{\mathcal{J}}$ contains all the Jordan blocks of $A + \tau B$ associated with eigenvalues different from $\widehat{\lambda}$.

Theorem 5.2 *Let $A \in \mathbb{F}^{n\times n}$ and let $T, G \in \mathbb{F}^{n\times n}$ be invertible matrices such that*

$$T^{-1}AT = \widehat{A} \oplus \breve{A} \oplus \widetilde{A}, \quad T^\star GT = \begin{bmatrix} 0 & \breve{G} \\ \widehat{G} & 0 \end{bmatrix} \oplus \widetilde{G}, \tag{5.4}$$

where the decomposition (5.4) *has the following properties:*

(a)

$$\widehat{A} = \left(\bigoplus_{j=1}^{\ell_1} \mathcal{J}_{n_1}(\widehat{\lambda}) \right) \oplus \left(\bigoplus_{j=1}^{\ell_2} \mathcal{J}_{n_2}(\widehat{\lambda}) \right) \oplus \cdots \oplus \left(\bigoplus_{j=1}^{\ell_m} \mathcal{J}_{n_m}(\widehat{\lambda}) \right),$$

where $n_1 > n_2 > \cdots > n_m$ *and* $\widehat{\lambda} \in \mathbb{F}$*;*

(b) $a = \sum_{j=1}^{m} \ell_j n_j$ *and* $\widehat{G}, \breve{G}, \breve{A} \in \mathbb{F}^{a \times a}$, $\widetilde{G} \in \mathbb{F}^{(n-2a)\times(n-2a)}$*;*

(c) $\sigma(\breve{A}), \sigma(\widetilde{A}) \subseteq \mathbb{C} \setminus \{\widehat{\lambda}\}$.

If $B \in \mathbb{F}^{n\times n}$ *is a rank one perturbation of the form* $B = uu^{\star}G$, $u \in \mathbb{F}^n$*, then generically (with respect to the components of* u *if* $\star = T$*, and with respect to the real and imaginary parts of the components of* u *if* $\star = *$*), then for all* $\tau \in \mathbb{F}\setminus\{0\}$ *the matrix* $A + \tau B$ *has the Jordan canonical form* (5.3)*.*

Theorem 5.2 was stated and proved in [16] (Theorems 3.1 and 3.2 there) for the special case that $\breve{G} = I_a$. However, Theorem 5.2 can be immediately reduced to that case by applying a transformation $(A, G) \mapsto (Q^{-1}AQ, Q^{\star}GQ)$ with the matrix

$$Q = T\left(\begin{bmatrix} I_a & 0 \\ 0 & \breve{G}^{-1} \end{bmatrix} \oplus I_{n-2a} \right).$$

Both Theorem 5.1 and 5.2 were stated and proved in [16] for the matrix $A + B$ only, but not for the family of matrices $A + \tau B$, where $\tau \in \mathbb{F} \setminus \{0\}$. However, the proof given in [16] can be immediately generalized to the more general case. Observe that the fact that (for generic vectors u and v) the parameter τ has no influence in the Jordan structure of the eigenvalue $\widehat{\lambda}$ under consideration, is in line with the results in [22], where rank one perturbations of the form $A + \tau B$ for unstructured matrices A and B are considered, see also the proof of the following theorem given in the appendix that clearly shows that the presence of the parameter τ is harmless in the derivation of the Jordan structure of $\widehat{\lambda}$.

Theorem 5.3 *Let* $A \in \mathbb{F}^{n\times n}$ *and let* $T, G \in \mathbb{F}^{n\times n}$ *be invertible such that*

$$T^{-1}AT = \left(\bigoplus_{j=1}^{\ell_1} \mathcal{J}_{n_1}(\widehat{\lambda}) \right) \oplus \left(\bigoplus_{j=1}^{\ell_2} \mathcal{J}_{n_2}(\widehat{\lambda}) \right) \oplus \cdots \oplus \left(\bigoplus_{j=1}^{\ell_m} \mathcal{J}_{n_m}(\widehat{\lambda}) \right) \oplus \widetilde{A} \qquad (5.5)$$

$$T^{T}GT = \left(\bigoplus_{j=1}^{\frac{1}{2}\ell_1} \begin{bmatrix} 0 & G^{(1,2j-1)} \\ G^{(1,2j)} & 0 \end{bmatrix} \right) \oplus G^{(2)} \oplus \cdots \oplus G^{(m)} \oplus \widetilde{G}, \qquad (5.6)$$

where ℓ_1 *is even,* $\widehat{\lambda} \in \mathbb{F}$*, and the decompositions* (5.5) *and* (5.6) *have the following properties:*

(1) $n_1 > n_2 > \cdots > n_m$*;*

(2) $G^{(j)} \in \mathbb{F}^{\ell_j n_j \times \ell_j n_j}$, $j = 2, \ldots, m$*, the matrices*

$$G^{(1,j)} = \begin{bmatrix} 0 & \cdots & 0 & g^{(1,j)}_{1,n_1} \\ \vdots & \iddots & g^{(1,j)}_{2,n_1-1} & g^{(1,j)}_{2,n_1} \\ 0 & \iddots & \iddots & \vdots \\ g^{(1,j)}_{n_1,1} & g^{(1,j)}_{n_1,2} & \cdots & g^{(1,j)}_{n_1,n_1} \end{bmatrix}, \quad j = 1, 2, \ldots, \ell_1;$$

are anti-triangular (necessarily invertible), and their entries satisfy the following two conditions:

(2a) $g^{(1,2s)}_{n_1,1} = -g^{(1,2s-1)}_{n_1,1}$ *for* $s = 1, \ldots, \frac{\ell_1}{2}$*;*

(2b) *there exists at least one index* $s \in \{1, \ldots, \frac{\ell_1}{2}\}$ *such that at least one of the three values* $g^{(1,2s)}_{n_1,1} + g^{(1,2s-1)}_{n_1-1,2}$, $g^{(1,2s)}_{n_1-1,2} + g^{(1,2s-1)}_{n_1,1}$, *or* $g^{(1,2s)}_{n_1,2} + g^{(1,2s-1)}_{n_1,2}$ *is nonzero.*

(3) $\widetilde{G}, \widetilde{A} \in \mathbb{F}^{(n-a)\times(n-a)}$, *where* $a = \sum_{j=1}^{m} \ell_j n_j$ *and* $\sigma(\widetilde{A}) \subseteq \mathbb{C} \setminus \{\widehat{\lambda}\}$.

If $B \in \mathbb{F}^{n\times n}$ *is a rank one matrix of the form* $B = uu^T G$, *where* $u \in \mathbb{F}$, *then generically (with respect to the components of* u*) for all* $\tau \in \mathbb{F} \setminus \{0\}$ *the matrix* $A + \tau B$ *has the Jordan canonical form*

$$\mathcal{J}_{n_1+1}(\widehat{\lambda}) \oplus \left(\bigoplus_{j=1}^{\ell_1-2} \mathcal{J}_{n_1}(\widehat{\lambda}) \right) \oplus \left(\bigoplus_{j=1}^{\ell_2} \mathcal{J}_{n_2}(\widehat{\lambda}) \right) \oplus \cdots \oplus \left(\bigoplus_{j=1}^{\ell_m} \mathcal{J}_{n_m}(\widehat{\lambda}) \right) \oplus \widetilde{\mathcal{J}}, \tag{5.7}$$

where $\widetilde{\mathcal{J}}$ *contains all the Jordan blocks of* $A + \tau B$ *associated with eigenvalues different from* $\widehat{\lambda}$.

The rather technical proof of Theorem 5.3 is given in the Appendix.

6 Rank one perturbations of H-orthogonal matrices

We finally have all ingredients to prove our main result concerning structured rank one perturbations of H-orthogonal matrices, where $H = H^T$.

Theorem 6.1 *Let* $H \in \mathbb{C}^{n\times n}$ *be symmetric and invertible, let* $U \in \mathbb{C}^{n\times n}$ *be* H*-orthogonal, and let* $\lambda \in \mathbb{C}$ *be an eigenvalue of* U. *Assume that* U *has the Jordan canonical form*

$$\left(\bigoplus_{j=1}^{\ell_1} \mathcal{J}_{n_1}(\lambda) \right) \oplus \left(\bigoplus_{j=1}^{\ell_2} \mathcal{J}_{n_2}(\lambda) \right) \oplus \cdots \oplus \left(\bigoplus_{j=1}^{\ell_m} \mathcal{J}_{n_m}(\lambda) \right) \oplus \mathcal{J},$$

where $n_1 > \cdots > n_m$ *and where* $\mathcal{J}$ *with* $\sigma(\mathcal{J}) \subseteq \mathbb{C} \setminus \{\lambda\}$ *contains all Jordan blocks associated with eigenvalues different from* λ. *Furthermore, let* $u \in \mathbb{C}^n$ *be a vector satisfying* $u^T H u \neq 0$ *and let* $B = -\frac{2}{u^T H u} uu^T HU$.

(1) *If* $\lambda \notin \{-1, 1\}$, *then generically with respect to the components of* u, *the matrix* $U + B$ *has the Jordan canonical form*

$$\left(\bigoplus_{j=1}^{\ell_1-1} \mathcal{J}_{n_1}(\lambda) \right) \oplus \left(\bigoplus_{j=1}^{\ell_2} \mathcal{J}_{n_2}(\lambda) \right) \oplus \cdots \oplus \left(\bigoplus_{j=1}^{\ell_m} \mathcal{J}_{n_m}(\lambda) \right) \oplus \widetilde{\mathcal{J}},$$

where $\widetilde{\mathcal{J}}$ *contains all the Jordan blocks of* $U + B$ *associated with eigenvalues different from* λ.

(2) *If* $\lambda \in \{+1, -1\}$ *and if* n_1 *is odd, then generically with respect to the components of* u, *the matrix* $U + B$ *has the Jordan canonical form*

$$\left(\bigoplus_{j=1}^{\ell_1-1} \mathcal{J}_{n_1}(\lambda) \right) \oplus \left(\bigoplus_{j=1}^{\ell_2} \mathcal{J}_{n_2}(\lambda) \right) \oplus \cdots \oplus \left(\bigoplus_{j=1}^{\ell_m} \mathcal{J}_{n_m}(\lambda) \right) \oplus \widetilde{\mathcal{J}},$$

where $\widetilde{\mathcal{J}}$ *contains all the Jordan blocks of* $U + B$ *associated with eigenvalues different from* λ.

(3) *If $\lambda \in \{+1,-1\}$ and if n_1 is even, then ℓ_1 is even and generically with respect to the components of u, the matrix $U+B$ has the Jordan canonical form*

$$\mathcal{J}_{n_1+1}(\lambda) \oplus \left(\bigoplus_{j=1}^{\ell_1-2} \mathcal{J}_{n_1}(\lambda) \right) \oplus \left(\bigoplus_{j=1}^{\ell_2} \mathcal{J}_{n_2}(\lambda) \right) \oplus \cdots \oplus \left(\bigoplus_{j=1}^{\ell_m} \mathcal{J}_{n_m}(\lambda) \right) \oplus \widetilde{\mathcal{J}},$$

where $\widetilde{\mathcal{J}}$ contains all the Jordan blocks of $U+B$ associated with eigenvalues different from λ.

(4) *If $-1 \notin \sigma(U)$, then $-1 \in \sigma(\widetilde{\mathcal{J}})$. Similarly, if $1 \notin \sigma(U)$, then $1 \in \sigma(\widetilde{\mathcal{J}})$.*

Proof. Concerning the parts (1)–(3), we may assume without loss of generality that U and H are in the canonical form of Corollary 4.7. If $\lambda \notin \{+1,-1\}$ or if $\lambda \in \{+1,-1\}$ and n_1 is odd, then (1) and (2) follow immediately from Theorem 5.1 or Theorem 5.2, respectively, applied to U and $G = HU$, where $\tau = -\frac{2}{u^T H u}$. If $\lambda \in \{+1,-1\}$ and n_1 is even, then we can apply Theorem 5.3 to U and $G = HU$ to obtain (3). Indeed, observe that in the notation of Corollary 4.7 and Theorem 5.3 we have that

$$\begin{aligned}
g_{n_1,1}^{(1,2s-1)} &= \lambda h_{n_1,1}^{(1,s)} &&= (-1)^{n_1-1}\lambda h_{1,n_1}^{(1,s)},\\
g_{n_1,1}^{(1,2s)} &= -\lambda h_{n_1,1}^{(1,s)} &&= (-1)^{n_1}\lambda h_{1,n_1}^{(1,s)},\\
g_{n_1-1,2}^{(1,2s-1)} &= \lambda h_{n_1-1,2}^{(1,s)} &&= (-1)^{n_1-2}\lambda h_{1,n_1}^{(1,s)},\\
g_{n_1-1,2}^{(1,2s)} &= -\lambda h_{n_1-1,2}^{(1,s)} &&= (-1)^{n_1-1}\lambda h_{1,n_1}^{(1,s)}.
\end{aligned}$$

Thus, we find that $g_{n_1,1}^{(1,2s-1)} = -g_{n_1,1}^{(1,2s)}$ and $g_{n_1,1}^{(1,2s-1)} + g_{n_1-1,2}^{(1,2s)} = -2\lambda h_{n_1,1}^{(1,s)} \neq 0$, and so the conditions (2a) and (2b) of Theorem 5.3 are satisfied.

For the proof of (4), assume that $-1 \notin U$. Then it follows from the canonical form of Theorem 4.1 that $\det U = 1$. Indeed, all possible blocks in the canonical form have determinant one except for Jordan blocks with odd size that are associated with the eigenvalue -1. Since $U + B = EU$, where $E = I - \frac{2}{u^T H u} u u^T H$ has determinant -1, it follows for the same reason that the H-orthogonal matrix $U+B$ must have the eigenvalue -1 (with odd algebraic multiplicity).

The corresponding statement concerning the eigenvalue $\lambda = +1$ follows from the above by considering $-U$ instead of U. □

Remark 6.2 We highlight that if $+1$ and/or -1 are eigenvalues of U then the assertions (2) and/or (3) apply to either of those eigenvalues. Thus, the fact stated in Theorem 6.1(4) that generically a new eigenvalue is generated at $+1$ or -1 only occurs in the situation that this eigenvalue was not yet in the spectrum of the original matrix. In particular, if both 1 and -1 are eigenvalues of U, then the largest Jordan blocks associated with both 1 and -1 will disappear, but no "new" eigenvalues at ± 1 will be created.

We illustrate the, now no longer surprising, behavior of the eigenvalues $+1$ and -1 with the help of a few simple examples.

Example 6.3 Let $\lambda \in \mathbb{C} \setminus \{0, 1, -1\}$,

$$U_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad U_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad U_3 = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{bmatrix}, \quad H = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \in \mathbb{F}^2.$$

Then, U_1, U_2, U_3 are H-orthogonal. (Note that U_3 is the matrix from Example 3.4.) Furthermore, assume that $u^T H u = 2u_1u_2 \neq 0$. If E_H is as in (3.3) then

$$E_H = I_2 - \frac{1}{u_1u_2} \begin{bmatrix} u_1u_2 & u_1^2 \\ u_2^2 & u_1u_2 \end{bmatrix} = \begin{bmatrix} 0 & -\frac{u_1}{u_2} \\ -\frac{u_2}{u_1} & 0 \end{bmatrix}.$$

Perturbing U_1, we obtain $U_1 + B = E_H U_1 = E_H$, so the perturbed matrix now has eigenvalues $+1$ and -1 each with algebraic multiplicity one. According to the theorem, one of the Jordan blocks of U_1 at eigenvalue 1 has disappeared and a new eigenvalue at -1 has emerged. On the other hand, we obtain that

$$E_H U_2 = \begin{bmatrix} \frac{u_1}{u_2} & 0 \\ 0 & \frac{u_2}{u_1} \end{bmatrix}, \quad E_H U_3 = \begin{bmatrix} 0 & \frac{u_1}{u_2}\lambda^{-1} \\ \frac{u_2}{u_1}\lambda & 0 \end{bmatrix}.$$

So, $E_H U_2$ generically has a reciprocal pair of non-unimodular eigenvalues. According to Theorem 6.1, both eigenvalues at 1 and -1 have disappeared, because their geometric multiplicities were equal to one. No new eigenvalues at ± 1 have appeared.

Finally, $E_H U_3$ has the eigenvalues $+1$ and -1 according to Theorem 6.1, because neither of those have been eigenvalues of U_3.

Example 6.4 Revisiting Example 3.3, we have seen there that a rank one perturbation of the H-orthogonal matrix having two Jordan blocks of size 2 associated with the eigenvalue 1 resulted in an increase of the size of one Jordan block to size 3 and the emergence of the eigenvalue -1. Both observation are in accordance with Theorem 6.1.

7 Rank one perturbations of symplectic matrices

In this section we consider rank one additive perturbations of complex symplectic matrices. We start with a lemma that is analogous to Lemma 3.1.

Lemma 7.1 *Suppose that $J = -J^T$ is an invertible complex $n \times n$ matrix, and let S be J-symplectic. If $\widetilde{S}$ is a J-symplectic matrix such that* $\operatorname{rank}(S - \widetilde{S}) = 1$, *then there is a vector $u \in \mathbb{C}^n$ such that*

$$\widetilde{S} = (I + uu^T J)S$$

Conversely, for any vector $u \in \mathbb{C}^n$, the matrix $\widetilde{S}$ is J-symplectic.

Proof. Set $\widetilde{S} := S + uv^T$ for some vectors u, v. Then, from $\widetilde{S}^T J \widetilde{S} = J$, using also the fact that $u^T J u = 0$ (because J is skew-symmetric) it follows that

$$S^T J u v^T + v u^T J S = 0.$$

From this, we see that v is a multiple of $S^T J u$, say $v = c S^T J u$, and so

$$\widetilde{S} = S - c u u^T J S.$$

Writing $-c = a^2$ (which is possible for the complex number c), and incorporating a into the vector u, we see that general additive rank one perturbations of the J-symplectic matrix S are of the form

$$\widetilde{S} = (I + uu^T J)S.$$

On the other hand, it is easily seen that for any vector u the matrix $\widetilde{S}$ is J-symplectic. Indeed, for that it suffices to note that $I + uu^T J$ is J-symplectic, which is immediate from

$$(I - Juu^T)J(I + uu^T J) = J - Juu^T J + Juu^T J - Juu^T Juu^T J = J. \quad \square$$

Observe that in contrast to the H-orthogonal case, we have $\det(I + uu^T J) = 1 + u^T J u = 1$. Also, the norm of the additive perturbation $uu^T JS$ can be arbitrarily small.

Theorem 7.2 *Let $J \in \mathbb{C}^{n\times n}$ be skew-symmetric and invertible, let $S \in \mathbb{C}^{n\times n}$ be J-symplectic, and let $\lambda \in \mathbb{C}$ be an eigenvalue of S. Assume that S has the Jordan canonical form*

$$\left(\bigoplus_{j=1}^{\ell_1} \mathcal{J}_{n_1}(\lambda)\right) \oplus \left(\bigoplus_{j=1}^{\ell_2} \mathcal{J}_{n_2}(\lambda)\right) \oplus \cdots \oplus \left(\bigoplus_{j=1}^{\ell_m} \mathcal{J}_{n_m}(\lambda)\right) \oplus \mathcal{J},$$

where $n_1 > \cdots > n_m$ and where $\mathcal{J}$ with $\sigma(\mathcal{J}) \subseteq \mathbb{C} \setminus \{\lambda\}$ contains all Jordan blocks associated with eigenvalues different from λ. Furthermore, let $u \in \mathbb{C}^n$ and $B = uu^T JS$.

(1) *If $\lambda \notin \{-1, 1\}$, then generically with respect to the components of u, the matrix $S + B$ has the Jordan canonical form*

$$\left(\bigoplus_{j=1}^{\ell_1-1} \mathcal{J}_{n_1}(\lambda)\right) \oplus \left(\bigoplus_{j=1}^{\ell_2} \mathcal{J}_{n_2}(\lambda)\right) \oplus \cdots \oplus \left(\bigoplus_{j=1}^{\ell_m} \mathcal{J}_{n_m}(\lambda)\right) \oplus \widetilde{\mathcal{J}},$$

where $\widetilde{\mathcal{J}}$ contains all the Jordan blocks of $S + B$ associated with eigenvalues different from λ.

(2) *If $\lambda \in \{+1, -1\}$ and if n_1 is even, then generically with respect to the components of u, the matrix $S + B$ has the Jordan canonical form*

$$\left(\bigoplus_{j=1}^{\ell_1-1} \mathcal{J}_{n_1}(\lambda)\right) \oplus \left(\bigoplus_{j=1}^{\ell_2} \mathcal{J}_{n_2}(\lambda)\right) \oplus \cdots \oplus \left(\bigoplus_{j=1}^{\ell_m} \mathcal{J}_{n_m}(\lambda)\right) \oplus \widetilde{\mathcal{J}},$$

where $\widetilde{\mathcal{J}}$ contains all the Jordan blocks of $S + B$ associated with eigenvalues different from λ.

(3) *If $\lambda \in \{+1, -1\}$ and if n_1 is odd, then ℓ_1 is even and generically with respect to the components of u, the matrix $S + B$ has the Jordan canonical form*

$$\mathcal{J}_{n_1+1}(\lambda) \oplus \left(\bigoplus_{j=1}^{\ell_1-2} \mathcal{J}_{n_1}(\lambda)\right) \oplus \left(\bigoplus_{j=1}^{\ell_2} \mathcal{J}_{n_2}(\lambda)\right) \oplus \cdots \oplus \left(\bigoplus_{j=1}^{\ell_m} \mathcal{J}_{n_m}(\lambda)\right) \oplus \widetilde{\mathcal{J}},$$

where $\widetilde{\mathcal{J}}$ contains all the Jordan blocks of $S + B$ associated with eigenvalues different from λ.

Proof. The proof of parts (1)–(3) is analogous to the proof of the corresponding parts of Theorem 6.1 by using Corollary 4.8 and Theorems 5.1–5.3. Indeed, the corresponding computation of the entries of $G = JS$ in the notation of Corollary 4.8 and Theorem 5.3 gives

$$\begin{aligned}
g_{n_1,1}^{(1,2s-1)} &= \lambda h_{n_1,1}^{(1,s)} = (-1)^{n_1-1}\lambda, \\
g_{n_1,1}^{(1,2s)} &= -\lambda h_{1,n_1}^{(1,s)} = -\lambda, \\
g_{n_1-1,2}^{(1,2s-1)} &= \lambda h_{n_1-1,2}^{(1,s)} = (-1)^{n_1-2}\lambda, \\
g_{n_1-1,2}^{(1,2s)} &= -\lambda h_{2,n_1-1}^{(1,s)} = \lambda.
\end{aligned}$$

Thus, as n_1 is odd, we find that $g_{n_1,1}^{(1,2s-1)} = -g_{n_1,1}^{(1,2s)}$ and $g_{n_1,1}^{(1,2s-1)} + g_{n_1-1,2}^{(1,2s)} = 2\lambda \neq 0$ as well as $g_{n_1,1}^{(1,2s)} + g_{n_1-1,2}^{(1,2s-1)} = -2\lambda \neq 0$, and so the conditions (2a) and (2b) of Theorem 5.3 are satisfied. □

Remark 7.3 We mention that in contrast to the H-orthogonal case, generically $+1$ and -1 will never occur as new eigenvalues of the perturbed matrix if they have not yet been eigenvalues of the original matrix. This follows from the fact that generically the new eigenvalues are all simple as we will show in the following section. However, the eigenvalues $+1$ and -1 must both occur with even algebraic multiplicity for symplectic matrices, as it can be easily seen from the canonical form of Theorem 4.2.

8 Simplicity of new eigenvalues

In this section, we investigate the multiplicity of the 'new eigenvalues' of a perturbed H-orthogonal or J-symplectic matrix. Our aim is to show that generically all new eigenvalues will be simple. We start with a lemma that generalizes previous results from [16].

Lemma 8.1 *Let $A \in \mathbb{C}^{n\times n}$ have the pairwise distinct eigenvalues $\lambda_1,\dots,\lambda_m \in \mathbb{C}$ with algebraic multiplicities $a_1,\dots,a_m$ and let $X \in \mathbb{C}^{n\times n}$. Suppose that the matrix $B(u) = A + uu^T X$ generically (with respect to the entries of u) has the eigenvalues $\lambda_1,\dots,\lambda_m$ with algebraic multiplicities $\widetilde{a}_1,\dots,\widetilde{a}_m$, where $\widetilde{a}_j \le a_j$ for $j = 1,\dots,m$. Furthermore, let $\varepsilon > 0$ be such that the discs*

$$D_i := \left\{\mu \in \mathbb{C} \,\middle|\, |\lambda_i - \mu| < \varepsilon^{2/n}\right\}, \quad i = 1,\dots,m$$

are pairwise disjoint. If for each $j = 1,\dots,m$ there exists a vector $u_j \in \mathbb{C}^n$ with $\|u_j\| < \varepsilon$ such that the matrix $A + u_j u_j^T X$ has exactly $a_j - \widetilde{a}_j$ simple eigenvalues in D_j different from λ_j, then generically (with respect to the entries of u) the eigenvalues of $B(u)$ that are different from the eigenvalues of A are simple.

Proof. For the proof of Lemma 8.1 we follow the lines of the proof of [16, Lemma 2.5]. First we note that by the choice of ε, any matrix $B(u)$ with $\|u\| < K \cdot \min\{1, \varepsilon\}$ has exactly a_i eigenvalues in the disc D_i, where the positive constant K depends only on $\|A\|, \|X\|$ and n. This follows from well-known results on matching distance of eigenvalues of nearby matrices, see for example [25, Section IV.1] and references there. (Concrete formulas are available for K but we do not need them.) We set $\varepsilon' = K \cdot \min\{1, \varepsilon\}$.

Let Ω denote the generic set of vectors u for which $B(u)$ has the eigenvalues $\lambda_1,\dots,\lambda_m$ with algebraic multiplicities $\widetilde{a}_1,\dots,\widetilde{a}_m$. Next, let us fix λ_j and let $\chi(\lambda_j, u)$ be the characteristic polynomial in the independent variable t for the restriction of $B(u)$ to its spectral invariant subspaces corresponding to the eigenvalues of $B(u)$ within D_j. Then the coefficients of $\chi(\lambda_j, u)$ are analytic functions of the components of u (cf. [16, Lemma 2.5]).

Let $q(\lambda_j, u)$ be the number of distinct eigenvalues of $B(u)$ in the disc D_j. Denote by $S(p_1, p_2)$ the Sylvester resultant matrix of the two polynomials $p_1(t)$, $p_2(t)$ and recall that $S(p_1, p_2)$ is a square matrix of size degree$\,(p_1)$ + degree$\,(p_2)$ and that the rank deficiency of $S(p_1, p_2)$ coincides with the degree of the greatest common divisor of the polynomials $p_1(t)$ and $p_2(t)$. We have

$$q(\lambda_j, u) = \operatorname{rank} S\left(\chi(\lambda_j, u), \frac{\partial \chi(\lambda_j, u)}{\partial t}\right) - a_j + 1.$$

The entries of $S(\chi(\lambda_j, u), \frac{\partial \chi(\lambda_j,u)}{\partial x})$ are scalar multiples (which are independent of u) of the coefficients of $\chi(\lambda_j, u)$, and therefore the set $Q(\lambda_j)$ of all vectors $u \in \mathbb{C}^n$, $\|u\| < \varepsilon'$, for which $q(\lambda_j, u, v)$ is maximal is the complement of the set of common zeros of finitely many analytic functions of the components of u. In particular, $Q(\lambda_j)$ is open and dense in

$$\{u \in \mathbb{C}^n \mid \|u\| < \varepsilon'\}.$$

By hypothesis, there exists a vector $u_j \in \mathbb{C}^n$ such that $B(u)$ has exactly $a_j - \widetilde{a}_j$ simple eigenvalues in D_j different from λ_j. If by chance the vector u_j is not in Ω, then we slightly perturb u_j to obtain a new vector $u_j' \in \Omega$ such that $B(u_j)$ has the eigenvalues $\lambda_1,\dots,\lambda_m$ with algebraic multiplicities $\widetilde{a}_1,\dots,\widetilde{a}_m$ and $a_j - \widetilde{a}_j$ simple eigenvalues in D_j different from λ_j. Such choice of u_j' is possible because Ω is generic, the property of eigenvalues being simple persists under small perturbations of $B(u_j)$, and the total number of eigenvalues of $B(u)$ within D_j, counted with multiplicities, is equal to a_j, for every $u \in \mathbb{C}^n$, $\|u\| < \varepsilon'$. Since Ω is open, clearly there exists $\delta > 0$ such that for all $u \in \mathbb{C}^n$ with $\|u - u_j\| < \delta$ the matrix $B(u_j)$ has the eigenvalues $\lambda_1,\dots,\lambda_m$ with algebraic multiplicities $\widetilde{a}_1,\dots,\widetilde{a}_m$ and $a_j - \widetilde{a}_j$

simple eigenvalues in D_j different from λ_j. Since the set of all such vectors u is open in $\mathbb{C}^n$, it follows from the properties of the set $Q(\lambda_j)$ established above that in fact we have

$$q(\lambda_j, u, v) = a_j - \widetilde{a}_j, \quad \text{for all } \ u \in \mathbb{C}^n, \quad \|u - u_j\| < \delta.$$

So for the open set

$$\Omega_j := Q(\lambda_j) \cap \Omega$$

which is dense in $\{u \in \mathbb{C}^n \mid \|u\| < \varepsilon'\}$, we have that all eigenvalues of $B(u)$ within D_j different from λ_j are simple. Now let

$$\Omega' = \bigcap_{j=1}^{m} \Omega_j \subseteq \Omega.$$

Note that Ω' is nonempty as the intersection of finitely many open dense (in $\{u \in \mathbb{C}^n \mid \|u\| < \varepsilon'\}$) sets.

Finally, let $\chi(u)$ denote the characteristic polynomial (in the independent variable t) of $B(u)$. Then the number of distinct roots of $\chi(u)$ is given by

$$\operatorname{rank} S \left(\chi(u), \frac{\partial \chi(u)}{\partial t} \right) - n + 1$$

and therefore, the set of all vectors $u \in \Omega$ on which the number of distinct roots of $\chi(u)$ is maximal, is a generic set. Since Ω' constructed above is nonempty, this maximal number is equal to $\sum_{j=1}^{m}(a_j - \widetilde{a}_j)$, i.e., generically all eigenvalues of $B(u)$ that are different from $\lambda_1, \ldots, \lambda_m$ are simple. □

Theorem 8.2 *Let $J \in \mathbb{C}^{2n\times 2n}$ be skew-symmetric and invertible, let $S \in \mathbb{C}^{2n\times 2n}$ be J-symplectic, and let $B = uu^T JS$, where $u \in \mathbb{C}^{2n}$. Then generically (with respect to the entries of u) the eigenvalues of $S + B$ that are not eigenvalues of S are all simple.*

Proof. In view of Theorem 4.2, we may assume without loss of generality that S and J have the forms

$$S = S_1 \oplus S_2, \quad J = J_1 \oplus J_2,$$

where S_i and J_i have the same size for $i = 1, 2$, and where $\sigma(S_1) = \{1\}$ and $\sigma(S_2) \subseteq \mathbb{C} \setminus \{1\}$. Then we consider rank one perturbations of the form $S + u_i u_i^T JS$, $i = 1, 2$ with

$$u_1 = \begin{bmatrix} v_1 \\ 0 \end{bmatrix}, \quad u_2 = \begin{bmatrix} 0 \\ v_2 \end{bmatrix},$$

where the size of the vector v_i is corresponding to the size of S_i. Note that both perturbations are not generic, because the perturbation with u_i just perturbs the block S_i of S. Nevertheless, it follows from Theorem 7.2 that the behavior of the algebraic multiplicities of the eigenvalues of S_i under a generic perturbation of the form $S_i + v_i v_i^T J_i S_i$ is identical to the behavior of the corresponding eigenvalues of S. Thus, in view of Lemma 8.1 it suffices to construct a rank one perturbation of the form $S_i + v_i v_i^T J_i S_i$ that results in a perturbation that respects the generic behavior of algebraic multiplicities of eigenvalues of S_i and that only has simple eigenvalues in the spectrum different from the spectrum of S_i. However, in both cases, we can apply the Cayley transformation from Section 2, because S_1 does not have the eigenvalue -1 and S_2 does not have the eigenvalue $+1$. The existence of the desired perturbations then follows easily from the results on J-Hamiltonian matrices in [16], in particular Theorem 4.2 there. □

Unfortunately, an analogous approach will not work for H-orthogonal matrices, because rank one perturbations of H-orthogonal matrices of sufficiently small norm may not exists due to the scaling factor $\frac{2}{u^* H u}$ in the formula (3.3). However, numerical tests suggest that indeed new eigenvalues of perturbed H-orthogonal matrices will be generically simple. The following example is in line with that observation.

Example 8.3 Let $k \in \mathbb{N}$, $n = 2k+1$, and $U = \mathcal{J}_n(1)$. Then we will show below that there exists a nonsingular matrix $H = (h_{ij}) \in \mathcal{V}_n$ with $e_n^T H e_n = h_{nn} \neq 0$ such that the H-orthogonal matrix $\widetilde{U} = (I_n - \frac{2}{h_{nn}} e_n e_n^T H)U$ has only simple eigenvalues.

By Proposition 4.3, H is uniquely determined by its entries $h_{k+1,k+1}, \ldots, h_{nn}$. If $h_{nn} = 1$, then $\lambda I_n - \widetilde{U}$ takes the form

$$\lambda I_n - \widetilde{U} = \begin{bmatrix} \lambda - 1 & -1 & 0 & \ldots & 0 \\ 0 & \lambda - 1 & -1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \lambda - 1 & -1 \\ 2h_{1n} & 2(h_{1n} + h_{2n}) & \ldots & 2(h_{n-2,n} + h_{n-1,n}) & 2h_{n-1,n} + 2 + \lambda - 1 \end{bmatrix},$$

and hence its characteristic polynomial $p(\lambda)$ has the form

$$\begin{aligned} p(\lambda) &= \det(\lambda I_n - \widetilde{U}) = (\lambda - 1)^n + \sum_{j=0}^{n-1} (2h_{j,n} + 2h_{j+1,n})(\lambda - 1)^j \quad (\text{with } h_{0n} := 0 \text{ and } h_{nn} = 1) \\ &=: \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1 \lambda + a_0, \end{aligned}$$

where

$$a_j = 2h_{j,n} + \alpha_{j+1,j} h_{j+1,n} + \cdots + \alpha_{n-1,j} h_{n-1,n} + \alpha_{nj}$$

for some coefficients $\alpha_{i,j-1}$, $i = j, \ldots, n-1$ for $j = 0, \ldots, n-1$. Thus a_j depends on h_{in} for $i = j, \ldots, n-1$, but not on h_{in} for $i < j$. From (4.6), we then obtain that

$$a_j = (-1)^{j-1} h_{jj} + \gamma_{j+1,j} h_{j+1,j+1} + \cdots + \gamma_{n-1,j} h_{n-1,n-1} + \gamma_{n,j} \tag{8.1}$$

for some coefficients $\gamma_{j+1,j}, \ldots, \gamma_{n,j}$ for $j = k+1, \ldots, n-1$.

Since $\widetilde{U}$ is H-orthogonal, it follows from Theorem 4.1 that $p(\lambda)$ has the form

$$p(\lambda) = (\lambda - 1)^{\mu_+} (\lambda + 1)^{\mu_-} \prod_{i=1}^{\ell} (\lambda - \lambda_i)^{\mu_i} (\lambda - \frac{1}{\lambda_i})^{\mu_i},$$

for some nonzero values $\lambda_1, \ldots, \lambda_\ell$ and some multiplicities $\mu_+, \mu_-, \mu_1, \ldots, \mu_\ell$. Moreover, since $\det U = 1$ and thus $\det \widetilde{U} = -1$, it follows that μ_- is odd (and hence necessarily μ_+ is even, possibly zero). But then, [14, Corollary 5.9] it follows that $p(\lambda) = \lambda^n p(\frac{1}{\lambda})$. This implies that

$$a_j = a_{n-j}, \quad j = 1, \ldots, \frac{n+1}{2} \quad \text{and} \quad a_0 = 1.$$

Thus, $p(\lambda)$ is already uniquely determined by $a_{n-1}, \ldots, a_{k+1}$ and from (8.1) we see that there is a unique choice of the parameters $h_{k+1,k+1}, \ldots, h_{n-1,n-1}$ such that $a_{n-1} = \cdots = a_{k+1} = 0$ so that the characteristic polynomial $p(\lambda)$ of $\widetilde{U}$ will be $\lambda^n + 1$ for $H \in \mathcal{V}_n$ given by this particular choice of $h_{k+1,k+1}, \ldots, h_{n-1,n-1}$ and $h_{nn} = 1$. In particular, all eigenvalues of $\widetilde{U}$ are simple. If by chance $h_{k+1,k+1} = 0$ and thus H is singular, then choose $h_{k+1,k+1} := \varepsilon > 0$. Since the entries of $p(\lambda)$ depend continuously on $h_{k+1,k+1}$ it will be guaranteed that the eigenvalues of $\widetilde{U}$ are still simple if ε is sufficiently small.

Note that if $\widehat{H}$ symmetric and invertible is given such that $\widehat{U}$ is $\widehat{H}$-orthogonal, where $\widehat{U}$ is similar to $\mathcal{J}_n(1)$, then by Corollary 4.7 the pair $(\widehat{U}, \widehat{H})$ is equivalent to the pair (U, H) with H constructed as above. Thus, we have shown that for any symmetric and invertible matrix $\widehat{H}$ and any $\widehat{H}$-orthogonal matrix $\widehat{U}$ similar to $\mathcal{J}_n(1)$, there exists an H-orthogonal rank one perturbation such that all eigenvalues of the perturbed matrix are simple.

9 Conclusions

We have presented several general results on Jordan forms of real and complex matrices under generic rank one perturbations, within the framework of certain structures imposed on the matrices and their perturbations. These results served as a basis for a study of the perturbation analysis of complex unitary (with respect to a nondegenerate sesquilinear form), orthogonal (with respect to a nondegenerate bilinear form), and symplectic (with respect to a nondegenerate skew-symmetric form) matrices under rank one perturbations that preserve the indicated structure. The forms in question are represented by an invertible hermitian or symmetric matrix H, or skew-symmetric (as the case may be) matrix J. The complex unitary case is disposed of quickly by virtue of the Cayley transform that reduces the unitary case to Hamiltonian matrices whose perturbation analysis was developed earlier. The orthogonal and symplectic cases present additional difficulties because generally speaking they cannot be reduced by the Cayley transform. The main findings of the paper are the following. For a given complex J-symplectic matrix S, a rank one additive perturbation that results again in an H-symplectic matrix, generically (with respect to the vector parameter representing the perturbation) destroys the biggest Jordan block for every eigenvalue of S, except for the case of the eigenvalue ± 1 and the biggest Jordan block corresponding to this eigenvalue is of odd size n_1. In the exceptional case, generically the two biggest blocks are destroyed and one block of size $n_1 + 1$ is created (corresponding to the same eigenvalue ± 1). Moreover, generically the "new" eigenvalues (i.e., those that are not eigenvalues of S) of the perturbed matrix are all simple. For complex H-orthogonal matrices, we have an analogous result, but now the exceptional case applies to the eigenvalues ± 1 when the size of the biggest Jordan block is even. However, we do not claim here the generic simplicity of new eigenvalues as for H-symplectic matrices, the reason being that there exist J-symplectic matrices arbitrarily close in norm to the given J-symplectic matrix S that differ from S by a rank one matrix, but this is not the case for H-orthogonal matrices. An additional phenomenon is observed for H-orthogonal (but not J-symplectic) matrices U, namely, if ± 1 is not an eigenvalue of U, then ± 1 is a new eigenvalue.

References

[1] M.A. Beitia, I. de Hoyos, and I. Zaballa. The change of the Jordan structure under one row perturbations. *Linear Algebra Appl.*, 401:119–134, 2005.

[2] P. Brunovsky. A classification of linear controllable systems. *Kybernetika (Prague)*, 6:173–188, 1970.

[3] B.M. Chen. *Robust and H_∞ Control.* Springer Verlag, London, 2000.

[4] F. De Terán and F. Dopico. Low rank perturbation of Kronecker structures without full rank. *SIAM J. Matrix Anal. Appl.*, 29:496–529, 2007.

[5] F. De Terán, F. Dopico, and J. Moro. Low rank perturbation of Weierstrass structure. *SIAM J. Matrix Anal. Appl.*, 30:538–547, 2008.

[6] F. De Terán and F. Dopico. Low rank perturbation of regular matrix polynomials. *Linear Algebra Appl.*, 430:579–586, 2009.

[7] P.A. Fuhrmann. *Linear Systems and Operators in Hilbert Space.* McGraw Hill, New York, 1981.

[8] F.R. Gantmacher. *Theory of Matrices*, volume 1. Chelsea, New York, 1959.

[9] S.K. Godunov and M. Sadkane. Spectral analysis and symplectic matrices with application to the theory of parametric resonance. *SIAM J. Math. Anal. Appl.* 28:1045–1069, 2006.

[10] I. Gohberg, P. Lancaster and L. Rodman. *Indefinite Linear Algebra and Applications.* Birkhäuser, Basel, 2005.

[11] L. Hörmander and A. Melin. A remark on perturbations of compact operators. *Math. Scand.*, 75:255–62, 1994.

[12] D. Janse van Rensburg. *Structured Matrices in Indefinite Inner Product Spaces: Simple Forms, Invariant Subspaces, and Rank-one Perturbations.* Ph.D. thesis, North-West University, Potchefstroom, South Africa, 2012.

[13] M. Krupnik. Changing the spectrum of an operator by perturbation. *Linear Algebra Appl.*, 167:113–118, 1992.

[14] D.S. Mackey, N. Mackey, C. Mehl and V. Mehrmann. Smith forms of palindromic matrix polynomials. *Electron. J. Linear Algebra*, 22:53-91, 2011.

[15] C. Mehl. On classification of normal matrices in indefinite inner product spaces. *Electron. J. Linear Algebra*, 15:50–83, 2006.

[16] C. Mehl, V. Mehrmann, A.C.M. Ran and L. Rodman. Eigenvalue perturbation theory of classes of structured matrices under generic structured rank one perturbations. *Linear Algebra Appl.*, 435:687–716, 2011.

[17] C. Mehl, V. Mehrmann, A.C.M. Ran and L. Rodman. Perturbation theory of selfadjoint matrices and sign characteristics under generic structured rank one perturbations. *Linear Algebra Appl.*, 436:4027–4042, 2012.

[18] C. Mehl, V. Mehrmann, A.C.M. Ran and L. Rodman. Jordan forms of real and complex matrices under rank one perturbations. *Oper. Matrices*, 7: 381–398, 2013.

[19] V. Mehrmann. *The Autonomous Linear Quadratic Optimal Control Problem: Theory and Numerical Solution.* Number 163 in Lecture Notes in Control and Information Sciences. Springer-Verlag, Heidelberg, 1991.

[20] M. Mimura and H. Toda. *Topology of Lie Groups, I and II.* Amer. Math. Soc., Providence, Rhode Island, 1991.

[21] J. Moro and F. Dopico. Low rank perturbation of Jordan structure. *SIAM J. Matrix Anal. Appl.*, 25:495–506, 2003.

[22] A.C.M. Ran and M. Wojtylak. Eigenvalues of rank one perturbations of unstructured matrices. *Linear Algebra Appl.*, 437:589–600, 2012.

[23] S.V. Savchenko. Typical changes in spectral properties under perturbations by a rank-one operator. *Mat. Zametki*, 74:590–602, 2003. (Russian). Translation in Mathematical Notes. 74:557–568, 2003.

[24] S. Savchenko. On the change in the spectral properties of a matrix under a perturbation of a sufficiently low rank. *Funktsional. Anal. i Prilozhen*, 38:85–88, 2004. (Russian). Translation in Funct. Anal. Appl. 38:69–71, 2004.

[25] G.W. Stuart and J.-G. Sun. *Matrix Perturbation Theory.* Academic Press, Boston etc., 1990.

[26] R.C. Thompson. Invariant factors under rank one perturbations. *Canadian J. Math*, 32:240–245, 1980.

[27] K. Zhou, J.C. Doyle, K. Glover. *Robust and Optimal Control.* Prentice Hall, Upper Saddle River, NJ, 1995.

10 Appendix: Proof of Theorem 5.3

In this section we prove Theorem 5.3. The proof follows the same lines as the proof of Theorem 4.2 in [16], but is more general and extends the result that was obtained there. Before we prove Theorem 5.3, we quote two results from [16]. The first one follows from the Brunovsky canonical form, see [2], and also [7, 3], of general multi-input control systems $\dot{x} = Ax + Bu$ under transformations

$$(A, B) \quad \mapsto \quad (C^{-1}(A + BR)C, C^{-1}BD),$$

with invertible matrices C, D and arbitrary matrix R of suitable sizes.

Theorem 10.1 *Let $A \in \mathbb{C}^{n\times n}$ be a matrix in Jordan canonical form*

$$A = \mathcal{J}_{n_1}(\lambda_1) \oplus \cdots \oplus \mathcal{J}_{n_g}(\lambda_g) \oplus \mathcal{J}_{n_{g+1}}(\lambda_{g+1}) \oplus \cdots \oplus \mathcal{J}_{n_\nu}(\lambda_\nu), \tag{10.1}$$

where $\lambda_1 = \cdots = \lambda_g =: \widehat{\lambda} \in \mathbb{C}$, $\lambda_{g+1}, \ldots, \lambda_\nu \in \mathbb{C} \setminus \{\widehat{\lambda}\}$, $n_1 \geq \cdots \geq n_g$. Moreover, let $B = uv^T$, where

$$u = \begin{bmatrix} u_1 \\ \vdots \\ u_\nu \end{bmatrix}, \quad v = \begin{bmatrix} v_1 \\ \vdots \\ v_\nu \end{bmatrix}, \quad u_i, v_i \in \mathbb{C}^{n_i},\ i = 1, \ldots, \nu.$$

Assume that the first component of each vector v_i, $i = 1, \ldots, \nu$ is nonzero. Then the matrix $\mathrm{Toep}\,(v_1) \oplus \cdots \oplus \mathrm{Toep}\,(v_\nu)$ *is invertible, and if we denote its inverse by S, then $S^{-1}AS = A$ and*

$$S^{-1}BS = \left[we_{1,n_1}^T, \ldots, we_{1,n_\nu}^T\right], \tag{10.2}$$

where $w = S^{-1}u$. Moreover, the matrix $S^{-1}(A+B)S$ has at least $g-1$ Jordan chains associated with $\widehat{\lambda}$ of lengths at least $n_2, \ldots, n_g$ given by

$$\begin{array}{lll} e_1 - e_{n_1+1}, & \ldots, & e_{n_2} - e_{n_1+n_2}; \\ e_1 - e_{n_1+n_2+1}, & \ldots, & e_{n_3} - e_{n_1+n_2+n_3}; \\ \vdots & \ddots & \vdots \\ e_1 - e_{n_1+\cdots+n_{g-1}+1}, & \ldots, & e_{n_g} - e_{n_1+\cdots+n_{g-1}+n_g}. \end{array} \tag{10.3}$$

Theorem 10.2 (partial Brunovsky form) *Let*

$$A = \left(\mathcal{J}_{n_1}(\widehat{\lambda})^{\oplus \ell_1}\right) \oplus \cdots \oplus \left(\mathcal{J}_{n_m}(\widehat{\lambda})^{\oplus \ell_m}\right) \oplus \widetilde{A} \in \mathbb{C}^{n\times n},$$

where $n_1 > \cdots > n_m$ and $\sigma(\widetilde{A}) \subseteq \mathbb{C} \setminus \{\widehat{\lambda}\}$. Moreover, let $a = \ell_1 n_1 + \cdots + \ell_m n_m$ denote the algebraic multiplicity of $\widehat{\lambda}$ and let $B = uv^T$, where $u, v \in \mathbb{C}^n$ and

$$v = \begin{bmatrix} v^{(1)} \\ \vdots \\ v^{(m)} \\ \widetilde{v} \end{bmatrix}, \quad v^{(i)} = \begin{bmatrix} v^{(i,1)} \\ \vdots \\ v^{(i,\ell_i)} \end{bmatrix}, \quad v^{(i,j)} \in \mathbb{C}^{n_i}, \quad j = 1, \ldots, \ell_i, \quad i = 1, \ldots, m.$$

Assume that the first component of each vector $v^{(i,j)}$, $j = 1, \ldots, \ell_i$, $i = 1, \ldots, m$ is nonzero. Then the following statements hold:

(1) *The matrix* $S := \left(\bigoplus_{j=1}^{\ell_1} \mathrm{Toep}(v^{(1,j)}) \oplus \cdots \oplus \bigoplus_{j=1}^{\ell_m} \mathrm{Toep}(v^{(m,j)})\right)^{-1} \oplus I_{n-a}$ *exists and satisfies*

$$S^{-1}AS = A, \quad S^{-1}BS = w\left[\underbrace{e_{1,n_1}^T, \ldots, e_{1,n_1}^T}_{\ell_1 \text{ times}}, \ldots, \underbrace{e_{1,n_m}^T, \ldots, e_{1,n_m}^T}_{\ell_m \text{ times}}, z^T\right]$$

where $w = S^{-1}u$ and for some appropriate vector $z \in \mathbb{C}^{n-a}$.

(2) *The matrix $S^{-1}(A+B)S$ has at least $\ell_1+\cdots+\ell_m-1$ Jordan chains associated with $\widehat{\lambda}$ given as follows:*

   a) *ℓ_1-1 Jordan chains of length at least n_1:*

$$\begin{array}{lll} e_1-e_{n_1+1}, & \ldots, & e_{n_1}-e_{2n_1}; \\ \vdots & \ddots & \vdots \\ e_1-e_{(\ell_1-1)n_1+1}, & \ldots, & e_{n_1}-e_{\ell_1 n_1}; \end{array}$$

   b) *ℓ_i Jordan chains of length at least n_i for $i=2,\ldots,m$:*

$$\begin{array}{lll} e_1-e_{\ell_1 n_1+\cdots+\ell_{i-1}n_{i-1}+1}, & \ldots, & e_{n_i}-e_{\ell_1 n_1+\cdots+\ell_{i-1}n_{i-1}+n_i}; \\ e_1-e_{\ell_1 n_1+\cdots+\ell_{i-1}n_{i-1}+n_i+1}, & \ldots, & e_{n_i}-e_{\ell_1 n_1+\cdots+\ell_{i-1}n_{i-1}+2n_i}; \\ \vdots & \ddots & \vdots \\ e_1-e_{\ell_1 n_1+\cdots+\ell_{i-1}n_{i-1}+(\ell_i-1)n_i+1}, & \ldots, & e_{n_i}-e_{\ell_1 n_1+\cdots+\ell_{i-1}n_{i-1}+\ell_i n_i}; \end{array}$$

(3) *Partition $w=S^{-1}u$ as*

$$w=\begin{bmatrix} w^{(1)} \\ \vdots \\ w^{(m)} \\ \widetilde{w} \end{bmatrix}, \quad w^{(i)}=\begin{bmatrix} w^{(i,1)} \\ \vdots \\ w^{(i,\ell_i)} \end{bmatrix}, \quad w^{(i,j)}=\begin{bmatrix} w_1^{(i,j)} \\ \vdots \\ w_{n_i}^{(i,j)} \end{bmatrix}\in\mathbb{C}^{n_i},$$

and let $\lambda_1,\ldots,\lambda_q$ be the pairwise distinct eigenvalues of A different from $\widehat{\lambda}$ having the algebraic multiplicities $r_1,\ldots,r_q$, respectively. Set $\mu_i=\lambda_i-\widehat{\lambda}$, $i=1,2,\ldots,q$.
Then the characteristic polynomial $p_{\widehat{\lambda}}$ of $A+B-\widehat{\lambda}I$ is given by

$$p_{\widehat{\lambda}}(\lambda)=(-\lambda)^a q(\lambda)+\left(\prod_{i=1}^{q}(\mu_i-\lambda)^{r_i}\right)\cdot\left((-\lambda)^a+(-1)^{a-1}\sum_{i=1}^{m}\sum_{j=1}^{\ell_i}\sum_{k=1}^{n_i}w_k^{(i,j)}\lambda^{a-k}\right),$$

where $q(\lambda)$ is some polynomial;

(4) *Write $p_{\widehat{\lambda}}(\lambda)=c_n\lambda^n+\cdots+c_{a-n_1+1}\lambda^{a-n_1+1}+c_{a-n_1}\lambda^{a-n_1}$. Then*

$$c_{a-n_1}=(-1)^{a-1}\left(\prod_{i=1}^{q}\mu_i^{r_i}\right)\left(\sum_{j=1}^{\ell_1}w_{n_1}^{(1,j)}\right);$$

and in the case $n_1>1$ we have in addition that

$$c_{a-n_1+1}=(-1)^a\left(\sum_{\nu=1}^{q}r_\nu\mu_\nu^{r_\nu-1}\prod_{\substack{i=1\\ i\neq\nu}}^{q}\mu_i^{r_i}\right)\left(\sum_{j=1}^{\ell_1}w_{n_1}^{(1,j)}\right)+(-1)^{a-1}\left(\prod_{i=1}^{q}\mu_i^{r_i}\right)\left(\sum_{j=1}^{\ell_1}w_{n_1-1}^{(1,j)}\right),$$

if $n_1-1>n_2$ or, if $n_1-1=n_2$, then

$$\begin{aligned} c_{a-n_1+1} &= (-1)^a\left(\sum_{\nu=1}^{q}r_\nu\mu_\nu^{r_\nu-1}\prod_{\substack{i=1\\ i\neq\nu}}^{q}\mu_i^{r_i}\right)\left(\sum_{j=1}^{\ell_1}w_{n_1}^{(1,j)}\right) \\ &+ (-1)^{a-1}\left(\prod_{i=1}^{q}\mu_i^{r_i}\right)\left(\sum_{j=1}^{\ell_1}w_{n_1-1}^{(1,j)}+\sum_{j=1}^{\ell_2}w_{n_2}^{(2,j)}\right). \end{aligned}$$

The following notation of linear combinations of Jordan chains will be necessary.

Definition 10.3 *Let $A \in \mathbb{C}^{n\times n}$ and let $X = (x_1, \dots, x_p)$ and $Y = (y_1, \dots, y_q)$ be two Jordan chains of A associated with the same eigenvalue $\widehat{\lambda}$ of (possibly different) lengths p and q. Then the sum $X+Y$ of X and Y is defined to be the chain $Z = (z_1, \dots, z_{\max(p,q)})$, where*

$$z_j = \begin{cases} x_j & \text{if } p \geq q \\ y_j & \text{if } p < q \end{cases}, \quad j = 1, \dots, |p-q|$$

and

$$z_j = \begin{cases} x_j + y_{j-p+q} & \text{if } p \geq q \\ y_j + x_{j-q+p} & \text{if } p < q \end{cases}, \quad j = |p-q|+1, \dots, \max(p,q).$$

To illustrate this construction, consider e.g. $X = (x_1, x_2, x_3, x_4)$ and $Y = (y_1, y_2)$, then $X + Y = (x_1, x_2, x_3 + y_1, x_4 + y_2)$.

It is straightforward to check that the sum $Z = X + Y$ of two Jordan chains associated with an eigenvalue $\widehat{\lambda}$ is again a Jordan chain associated with $\widehat{\lambda}$ of the given matrix A, but it should be noted that this sum is not commutative.

Proof of Theorem 5.3. Let $\tau \in \mathbb{F} \setminus \{0\}$ be arbitrary. We may assume without loss of generality that A and G are already in the forms (5.5) and (5.6). Furthermore, we may assume $\widehat{\lambda} = 0$, otherwise consider the matrix $A - \widehat{\lambda} I$ instead of A. Then the algebraic and geometric multiplicity a and γ of the eigenvalue zero of A are given by

$$a = \sum_{s=1}^{m} \ell_s n_s, \quad \gamma = \sum_{s=1}^{m} \ell_s,$$

respectively. Let us partition u conformably with the forms (5.5) and (5.6), i.e., we let

$$u = \begin{bmatrix} u^{(1)} \\ \vdots \\ u^{(m)} \\ \widetilde{u} \end{bmatrix}, \quad u^{(i)} = \begin{bmatrix} u^{(i,1)} \\ \vdots \\ u^{(i,\ell_i)} \end{bmatrix}, \quad u^{(i,j)} = \begin{bmatrix} u_1^{(i,j)} \\ \vdots \\ u_{n_i}^{(i,j)} \end{bmatrix} \in \mathbb{C}^{n_i},$$

for $j = 1, \dots, \ell_i$; $i = 1, \dots, m$. Thus, $\widetilde{u} \in \mathbb{C}^{n-a}$. Then the vector $v^T = u^T G$ has the following structure:

$$v = (u^T G)^T = G^T u = \begin{bmatrix} v^{(1)} \\ \vdots \\ v^{(m)} \\ \widetilde{v} \end{bmatrix}, \quad v^{(i)} = \begin{bmatrix} v^{(i,1)} \\ \vdots \\ v^{(i,\ell_i)} \end{bmatrix}, \quad v^{(i,j)} = \begin{bmatrix} v_1^{(i,j)} \\ \vdots \\ v_{n_i}^{(i,j)} \end{bmatrix} \in \mathbb{C}^{n_i},$$

for $j = 1, \dots, \ell_i$ and $i = 1, \dots, m$, where $v^{(1,2s-1)} = (G^{(1,2s)})^T u^{(1,2s)}$ and $v^{(1,2s)} = (G^{(1,2s-1)})^T u^{(1,2s-1)}$, that is

$$v^{(1,2s-1)} = \begin{bmatrix} g_{n_1,1}^{(1,2s)} u_{n_1}^{(1,2s)} \\ g_{n_1-1,2}^{(1,2s)} u_{n_1-1}^{(1,2s)} + g_{n_1,2}^{(1,2s)} u_{n_1}^{(1,2s)} \\ * \\ \vdots \\ * \end{bmatrix}, \quad v^{(1,2s)} = \begin{bmatrix} g_{n_1,1}^{(1,2s-1)} u_{n_1}^{(1,2s-1)} \\ g_{n_1-1,2}^{(1,2s-1)} u_{n_1-1}^{(1,2s-1)} + g_{n_1,2}^{(1,2s-1)} u_{n_1}^{(1,2s-1)} \\ * \\ \vdots \\ * \end{bmatrix}$$

for $s = 1, \ldots, \ell_1/2$. Generically, the hypothesis of Theorem 10.2 is satisfied, i.e., the first entries of the vectors $v^{(i,j)}$ are nonzero. Thus, generically the matrix S as in Theorem 10.2 exists so that $S^{-1}(A + \tau B)S$ is in partial Brunovsky form. In fact, S^{-1} takes the form

$$S^{-1} = \left(\bigoplus_{j=1}^{\ell_1} \mathrm{Toep}\,(v^{(1,j)}) \right) \oplus \cdots \oplus \left(\bigoplus_{j=1}^{\ell_m} \mathrm{Toep}\,(v^{(m,j)}) \right) \oplus I_{n-a},$$

and it follows that

$$S^{-1}\tau BS = w(\underbrace{e_{1,n_1}^T, \ldots, e_{1,n_1}^T}_{\ell_1 \text{ times}}, \ldots, \underbrace{e_{1,n_m}^T, \ldots, e_{1,n_m}^T}_{\ell_m \text{ times}}, z^T) \tag{10.4}$$

for some $z \in \mathbb{C}^{n-a}$, where $w = S^{-1}u$. Thus,

$$w = S^{-1}u = \begin{bmatrix} w^{(1)} \\ \vdots \\ w^{(m)} \\ \widetilde{w} \end{bmatrix}, \quad w^{(i)} = \begin{bmatrix} w^{(i,1)} \\ \vdots \\ w^{(i,\ell_i)} \end{bmatrix}, \quad w^{(i,s)} = \begin{bmatrix} w_1^{(i,j)} \\ \vdots \\ w_{n_i}^{(i,j)} \end{bmatrix} \in \mathbb{C}^{n_i}, \tag{10.5}$$

for $j = 1, \ldots, \ell_i$ and $i = 1, \ldots, m$, where

$$w_{n_1}^{(1,2s-1)} = \tau g_{n_1,1}^{(1,2s)} u_{n_1}^{(1,2s)} u_{n_1}^{(1,2s-1)}, \qquad w_{n_1}^{(1,2s)} = \tau g_{n_1,1}^{(1,2s-1)} u_{n_1}^{(1,2s-1)} u_{n_1}^{(1,2s)} = -w_{n_1}^{(1,2s-1)} \tag{10.6}$$

using the hypothesis (2a), and, provided that $n_1 > 1$,

$$\begin{aligned} w_{n_1-1}^{(1,2s-1)} &= \tau g_{n_1,1}^{(1,2s)} u_{n_1}^{(1,2s)} u_{n_1-1}^{(1,2s-1)} + \tau g_{n_1-1,2}^{(1,2s)} u_{n_1-1}^{(1,2s)} u_{n_1}^{(1,2s-1)} + \tau g_{n_1,2}^{(1,2s)} u_{n_1}^{(1,2s)} u_{n_1}^{(1,2s-1)} \\ w_{n_1-1}^{(1,2s)} &= \tau g_{n_1,1}^{(1,2s-1)} u_{n_1}^{(1,2s-1)} u_{n_1-1}^{(1,2s)} + \tau g_{n_1-1,2}^{(1,2s-1)} u_{n_1-1}^{(1,2s-1)} u_{n_1}^{(1,2s)} + \tau g_{n_1,2}^{(1,2s-1)} u_{n_1}^{(1,2s-1)} u_{n_1}^{(1,2s)} \end{aligned}$$

for $s = 1, \ldots, \ell_1/2$. This implies that

$$\begin{aligned} & w_{n_1-1}^{(1,2s-1)} + w_{n_1-1}^{(1,2s)} \\ = \; & \tau(g_{n_1,1}^{(1,2s)} + g_{n_1-1,2}^{(1,2s-1)}) u_{n_1}^{(1,2s)} u_{n_1-1}^{(1,2s-1)} + \tau(g_{n_1-1,2}^{(1,2s)} + g_{n_1,1}^{(1,2s-1)}) u_{n_1-1}^{(1,2s)} u_{n_1}^{(1,2s-1)} \\ & + \tau(g_{n_1,2}^{(1,2s)} + g_{n_1,2}^{(1,2s-1)}) u_{n_1}^{(1,2s-1)} u_{n_1}^{(1,2s)} \end{aligned}$$

which, by the hypothesis (2b), is generically nonzero.

We will now show in two steps that generically $A + \tau B$ has the Jordan canonical form (5.7). By Theorem 10.2 we know that generically $A + \tau B$ has $\ell_1 - 1$ Jordan chains of length n_1 and ℓ_j Jordan chains of length n_j, $j = 2, \ldots, m$ associated with the eigenvalue zero. (These chains are linearly independent but need not form a basis of the corresponding root subspace of $A + \tau B$ yet, as it may be possible to extend some of the chains.) In the first step, we will show that generically there exists a Jordan chain of length $n_1 + 1$. In the second step, we will show that the algebraic multiplicity of the eigenvalue zero of $A + \tau B$ generically is $\widetilde{a} = (\sum_{s=1}^m \ell_s n_s) - n_1 + 1 = a - n_1 + 1$. Both steps together obviously imply that (5.7) represents the only possible Jordan canonical form for $A + \tau B$.

Step 1: Existence of a Jordan chain of length $n_1 + 1$.

Consider the following Jordan chains associated with the eigenvalue zero of $S^{-1}(A + \tau B)S$ and denoted by $C_{1,s}$ and $C_{i,j}$, respectively:

$$\begin{array}{lll} \text{length } n_1: & C_{1,s}: & e_{2(s-1)n_1+1} - e_{(2s-1)n_1+1}, \ldots, e_{(2s-1)n_1} - e_{2sn_1}, \quad s = 1, \ldots, \frac{\ell_1}{2} \\ \text{length } n_i: & C_{i,j}: & -e_1 + e_{\Sigma_{k=1}^{i-1}\ell_k n_k + (j-1)n_i + 1}, \ldots, -e_{n_i} + e_{\Sigma_{k=1}^{i-1}\ell_k n_k + jn_i}, \quad j = 1, \ldots, \ell_i, \end{array}$$

where $i = 2, \dots, m$. Observe that $C_{i,j}$, $i \neq 1$, are just the Jordan chains from Theorem 10.2 multiplied by -1 while the chains $C_{1,s}$ are linear combinations of the Jordan chains from Theorem 10.2. Namely, in the notation of (10.3), and numbering the chains in (10.3) first, second, etc., from the top to the bottom, we see that the chains $C_{1,1}, \dots, C_{1,\ell_1/2}$ are the negative of the second chain plus the first chain, the negative of the fourth chain plus the third chain, ..., the negative of the $(\ell_1 - 1)$-th chain plus the (ℓ_1)-th chain, respectively. Now consider the Jordan chain

$$C := \left(\sum_{s=1}^{\ell_1/2} \alpha_{1,s} C_{1,s} \right) + \sum_{i=2}^{m} \sum_{j=1}^{\ell_i} \alpha_{i,j} C_{i,j}$$

of length n_1 (see Definition 10.3), and let y denote the n_1-th (and thus last) vector of this chain. We next show that the Jordan chain C can be extended by a certain vector to a Jordan chain of length n_1+1 associated with the eigenvalue zero, for some particular choice of the parameters $\alpha_{i,s}$ (depending on u) such that generically at least one of $\alpha_{1,1}, \dots, \alpha_{1,\ell_1/2}$ is nonzero. To see this, we have to show that y is in the range of $S^{-1}(A + \tau B)S$. First, partition

$$y = \begin{bmatrix} y^{(1)} \\ \vdots \\ y^{(m)} \\ \widetilde{y} \end{bmatrix}, \quad y^{(i)} = \begin{bmatrix} y^{(i,1)} \\ \vdots \\ y^{(i,\ell_i)} \end{bmatrix}, \quad y^{(i,j)} = \begin{bmatrix} y_1^{(i,j)} \\ \vdots \\ y_{n_i}^{(i,j)} \end{bmatrix} \in \mathbb{C}^{n_i},$$

for $j = 1, \dots, \ell_i$; $i = 1, \dots, m$. Then by the definition of y, we have $\widetilde{y} = 0 \in \mathbb{C}^{n-a}$,

$$\begin{aligned} y_{n_1}^{(1,2s-1)} &= \alpha_{1,s}, \quad y_{n_1}^{(1,2s)} = -\alpha_{1,s}, \quad s = 1, \dots, \ell_1/2, \\ y_{n_i}^{(i,j)} &= \alpha_{i,j}, \quad j = 1, \dots, \ell_i; \quad i = 2, \dots, m. \end{aligned}$$

We have to solve the linear system

$$S^{-1}(A + \tau B)Sx = y. \tag{10.7}$$

Partitioning

$$x = \begin{bmatrix} x^{(1)} \\ \vdots \\ x^{(m)} \\ \widetilde{x} \end{bmatrix}, \quad x^{(i)} = \begin{bmatrix} x^{(i,1)} \\ \vdots \\ x^{(i,\ell_i)} \end{bmatrix}, \quad x^{(i,j)} = \begin{bmatrix} x_1^{(i,j)} \\ \vdots \\ x_{n_i}^{(i,j)} \end{bmatrix} \in \mathbb{C}^{n_i},$$

and making the ansatz $\widetilde{x} = 0$, then equation (10.7) becomes (here we use (10.4, (10.5))):

$$w_k^{(i,j)} \left(\sum_{\nu=1}^{m} \sum_{\mu=1}^{\ell_\nu} x_1^{(\nu,\mu)} \right) + x_{k+1}^{(i,j)} = y_k^{(i,j)}, \qquad k=1,\dots,n_i-1;\ j=1,\dots,\ell_i;\ i=1,\dots,m, \tag{10.8}$$

$$w_{n_i}^{(i,j)} \left(\sum_{\nu=1}^{m} \sum_{\mu=1}^{\ell_\nu} x_1^{(\nu,\mu)} \right) = \alpha_{i,j}, \qquad j=1,\dots,\ell_i;\ i=2,\dots,m, \tag{10.9}$$

$$w_{n_1}^{(1,2s-1)} \left(\sum_{\nu=1}^{m} \sum_{\mu=1}^{\ell_\nu} x_1^{(\nu,\mu)} \right) = \alpha_{1,s}, \qquad s=1,\dots,\ell_1/2, \tag{10.10}$$

$$w_{n_1}^{(1,2s)} \left(\sum_{\nu=1}^{m} \sum_{\mu=1}^{\ell_\nu} x_1^{(\nu,\mu)} \right) = -\alpha_{1,s}, \qquad s=1,\dots,\ell_1/2. \tag{10.11}$$

Set $x_1^{(1,1)} = 1$ and $x_1^{(\nu,\mu)} = 0$, for $\mu = 1, \ldots, \ell_\nu$; $\nu = 1, \ldots, m$; $(\nu,\mu) \neq (1,1)$, as well as $\alpha_{i,j} = w_{n_i}^{(i,j)}$ for $j = 1, \ldots, \ell_i$; $i = 2, \ldots, m$ and $\alpha_{1,s} = w_{n_1}^{(1,2s-1)}$ for $s = 1, \ldots, \ell_1/2$. Then (10.9) and (10.10) are satisfied and so is (10.11), because $w_{n_1}^{(1,2s)} = -w_{n_1}^{(1,2s-1)}$ by (10.6). Finally, the equation (10.8) can be solved by choosing $x_{k+1}^{(i,j)} = y_k^{(i,j)} - w_k^{(i,j)}$ for $k = 1, \ldots, n_i - 1$; $j = 1, \ldots, \ell_i$; $i = 1, \ldots, m$.

Step 2: We show that the algebraic multiplicity of the eigenvalue zero of $A + \tau B$ generically is $\widetilde{a} = (\sum_{s=1}^m \ell_s n_s) - n_1 + 1 = a - n_1 + 1$.

Let $\mu_1, \ldots, \mu_q$ denote the pairwise distinct nonzero eigenvalues of A and let $r_1, \ldots, r_q$ be their algebraic multiplicities. Denote by $p_0(\lambda)$ the characteristic polynomial of $A+\tau B$. By Theorem 10.2, the lowest possible power of λ associated with a nonzero coefficient in $p_0(\lambda)$ is $a-n_1$ and the corresponding coefficient c_{a-n_1} is

$$c_{a-n_1} = (-1)^{a-1} \left(\prod_{i=1}^q \mu_i^{r_i} \right) \left(\sum_{j=1}^{\ell_1} w_{n_1}^{(1,j)} \right) = 0,$$

because of (10.6). If $n_1 = 1$ then $\widetilde{a} = a$ and there is nothing to show as the algebraic multiplicity of the eigenvalue zero cannot increase when a generic perturbation is applied. Otherwise, we distinguish the cases $n_2 < n_1 - 1$ and $n_2 = n_1 - 1$. If $n_2 < n_1 - 1$, then by Theorem 10.2 the coefficient c_{a-n_1+1} of λ^{a-n_1+1} in $p_0(\lambda)$ is

$$\begin{aligned}
c_{a-n_1+1} &= (-1)^a \left(\sum_{\nu=1}^q r_\nu \mu_\nu^{r_\nu - 1} \prod_{\substack{i=1\\ i\neq\nu}}^q \mu_i^{r_i} \right) \left(\sum_{j=1}^{\ell_1} w_{n_1}^{(1,j)} \right) + (-1)^{a-1} \left(\prod_{i=1}^q \mu_i^{r_i} \right) \left(\sum_{j=1}^{\ell_1} w_{n_1-1}^{(1,j)} \right) \\
&= (-1)^{a-1} \left(\prod_{i=1}^q \mu_i^{r_i} \right) \left(\sum_{j=1}^{\ell_1} w_{n_1-1}^{(1,j)} \right) \\
&= (-1)^{a-1} \left(\prod_{i=1}^q \mu_i^{r_i} \right) \left(\sum_{s=1}^{\ell_1/2} \Big(\tau(g_{n_1,1}^{(1,2s)} + g_{n_1-1,2}^{(1,2s-1)}) u_{n_1}^{(1,2s)} u_{n_1-1}^{(1,2s-1)} \right. \\
&\qquad \left. + \tau(g_{n_1-1,2}^{(1,2s)} + g_{n_1,1}^{(1,2s-1)}) u_{n_1-1}^{(1,2s)} u_{n_1}^{(1,2s-1)} + \tau(g_{n_1,2}^{(1,2s)} + g_{n_1,2}^{(1,2s-1)}) u_{n_1}^{(1,2s-1)} u_{n_1}^{(1,2s)} \Big) \right)
\end{aligned}$$

by (10.7), where we have used (10.6) in the second equation to show that $\sum_{j=1}^{\ell_1} w_{n_1}^{(1,j)} = 0$. By the hypothesis (2b), it follows that c_{a-n_1+1} generically is nonzero. If, on the other hand, $n_2 = n_1 - 1$, then again by Theorem 10.2 (and using (10.6)) the coefficient c_{a-n_1+1} of λ^{a-n_1+1} in $p_0(\lambda)$ is

$$c_{a-n_1+1} = (-1)^{a-1} \left(\prod_{i=1}^q \mu_i^{r_i} \right) \left(\sum_{s=1}^{\ell_1} w_{n_1-1}^{(1,s)} + \sum_{j=1}^{\ell_2} w_{n_2}^{(2,j)} \right),$$

so in comparison to the case $n_2 < n_1 - 1$, there is an extra term in c_{a-n_1+1} depending on $w_{n_2}^{(2,j)}$, $j = 1, \ldots, \ell_2$. However, each entry $w_{n_2}^{(2,j)}$ only depends on the entries of the vectors $u^{(2,s)}$, $s = 1, \ldots, \ell_2$, so still c_{a-n_1+1} is nonzero generically. In all cases, we have shown that zero is a root of $p_0(\lambda)$ with multiplicity $a - n_1 + 1$. Thus, the algebraic multiplicity of the eigenvalue zero of $A+\tau B$ is $a - n_1 + 1$. Together with Step 1, we obtain that (5.7) generically is the only possible Jordan canonical form of $A + \tau B$. $\square$