

# Structured Hölder condition numbers for multiple eigenvalues

Daniel Kressner\*

María José Peláez<sup>‡</sup>

Julio Moro<sup>‡</sup>

December 7, 2008

## Abstract

The sensitivity of a multiple eigenvalue of a matrix under perturbations can be measured by its Hölder condition number. Various extensions of this concept are considered. A meaningful notion of structured Hölder condition numbers is introduced and it is shown that many existing results on structured condition numbers for simple eigenvalues carry over to multiple eigenvalues. The structures investigated in more detail include real, Toeplitz, Hankel, symmetric, skew-symmetric, Hamiltonian, and skew-Hamiltonian matrices. Furthermore, unstructured and structured Hölder condition numbers for multiple eigenvalues of matrix pencils are introduced. Particular attention is given to symmetric/skew-symmetric, Hermitian and palindromic pencils. It is also shown how matrix polynomial eigenvalue problems can be covered within this framework.

## 1 Introduction

Eigenvalue condition numbers asymptotically measure the sensitivity of an eigenvalue with respect to perturbations. If  $\lambda$  is a simple eigenvalue of a matrix  $A \in \mathbb{C}^{n \times n}$  then it is well known that  $\lambda$  is differentiable with respect to perturbations in  $A$  and that the eigenvalue  $\hat{\lambda}(\epsilon)$  of the perturbed matrix  $A + \epsilon E$  admits the expansion

$$\hat{\lambda} = \lambda + (y^H E x) \epsilon + O(\epsilon^2), \quad \epsilon \rightarrow 0, \quad (1)$$

where  $x$  and  $y$  are, respectively, a right and a left eigenvector of  $A$  (normalized so that  $|y^H x| = 1$ ) corresponding to  $\lambda$ . Then the absolute condition number for  $\lambda$ , defined as

$$\kappa(A, \lambda) = \lim_{\epsilon \rightarrow 0} \sup_{\substack{\|E\| \leq 1 \\ E \in \mathbb{C}^{n \times n}}} \frac{|\hat{\lambda} - \lambda|}{\epsilon}, \quad (2)$$

is given by  $\kappa(A, \lambda) = \|x\|_2 \|y\|_2$  for any unitarily invariant norm  $\|\cdot\|$ . One way to show this is to consider  $E = yx^H / (\|x\|_2 \|y\|_2)$ , which – by inserting in (1) – can be seen to attain the supremum in (2).

---

\*Seminar für angewandte Mathematik, ETH Zürich, Switzerland, [kressner@math.ethz.ch](mailto:kressner@math.ethz.ch). The research of this author was supported by the DFG Emmy Noether fellowship KR 2950/1-1 and the Swedish Foundation for Strategic Research under grant A3 02:128.

<sup>‡</sup>Departamento de Matemáticas, Universidad Carlos III de Madrid, Avda. Universidad, 30, 28911-Leganés (Spain), [mpe1aez,jmoro}@math.uc3m.es](mailto:{mpe1aez,jmoro}@math.uc3m.es). The research of these authors was partially supported by the Spanish Ministerio de Ciencia y Tecnología under grant MTM2006-05361.

Much of the recent research on eigenvalue condition numbers has been devoted to the case when the perturbation  $E$  is known to be in a set  $\mathbb{S} \subset \mathbb{C}^{n \times n}$  of structured matrices. In this case, it is more appropriate to restrict the supremum in (2) to  $E \in \mathbb{S}$ , giving rise to the structured eigenvalue condition number  $\kappa(A, \lambda; \mathbb{S})$ . In [9, 40], computable expressions of  $\kappa(A, \lambda; \mathbb{S})$  for general linear structures  $\mathbb{S}$  have been developed. This has been extended to smooth nonlinear structures in [16]. A simplified expression for zero-structured matrices can be found in [28]. Trivially,  $\kappa(A, \lambda; \mathbb{S}) \leq \kappa(A, \lambda)$ . It naturally gives rise to the question whether  $\kappa(A, \lambda; \mathbb{S})$  can be much smaller than  $\kappa(A, \lambda)$ , or, in other words, whether  $\lambda$  can be much less sensitive to structured perturbations than to unstructured ones. For surprisingly many structures  $\mathbb{S}$  the answer to this question is negative in the sense that  $\kappa(A, \lambda; \mathbb{S})$  is always at most within a small factor of  $\kappa(A, \lambda)$ . This has been shown for  $\mathbb{S} = \mathbb{R}^{n \times n}$  in [3], as well as for real skew-symmetric, skew-Hermitian, Hankel, Toeplitz, Hamiltonian, persymmetric, circulant, orthogonal, unitary, and related structures in [16, 32]. Practically relevant examples for which  $\kappa(A, \lambda; \mathbb{S}) \ll \kappa(A, \lambda)$  is possible include complex skew-symmetric [32], zero-structured [28], and symplectic matrices [16].

If  $\lambda$  is a multiple eigenvalue of algebraic multiplicity  $m$ , there is generally not an expansion of the form (1). Instead,  $\lambda$  bifurcates into  $m$  perturbed eigenvalues  $\hat{\lambda}_k(\epsilon)$ , each admitting a fractional expansion

$$\hat{\lambda}_k = \lambda + \alpha_k^{\gamma_k} \epsilon^{\gamma_k} + o(\epsilon), \quad \epsilon \rightarrow 0, \quad k = 1, \dots, m, \quad (3)$$

with  $\alpha_k > 0$  and  $0 < \gamma_k \leq 1$  [20, 42, 27]. Under generically satisfied conditions on  $E$ , Lidskii's theory [20] states that each Jordan block  $J_{n_j}(\lambda)$  of size  $n_j \times n_j$  gives rise to  $n_j$  perturbed eigenvalues satisfying the expansion (3) with  $\gamma_k = 1/n_j$ . Motivated by these results, the *Hölder condition number* for  $\lambda$  is defined in [27] as the pair

$$\kappa(A, \lambda) = (n_1, \alpha), \quad (4)$$

where  $1/n_1$  is the smallest possible power  $\gamma_k$  of  $\epsilon$  in (3) for any perturbation  $E$ . The scalar  $\alpha^{1/n_1} > 0$  is the largest possible magnitude of the coefficient of  $\epsilon^{1/n_1}$  for all  $E$  with  $\|E\| \leq 1$ . While  $n_1$  happens to be the size of the largest Jordan block belonging to  $\lambda$ , we have

$$\alpha^{1/n_1} = \lim_{\epsilon \rightarrow 0} \sup_{\substack{\|E\| \leq 1 \\ E \in \mathbb{C}^{n \times n}}} \max_{k=1, \dots, m} \frac{|\hat{\lambda}_k - \lambda|}{\epsilon^{1/n_1}}. \quad (5)$$

(see also [5, p. 156] for a similar definition of condition number for multiple eigenvalues, and [4] for its relationship with  $\kappa(A, \lambda)$ ). An explicit formula for  $\alpha$  can be found in [27], see also Section 2. Let us emphasize that for certain nongeneric perturbations  $E$ , the value of  $\gamma_k$  can be larger than  $1/n_1$  for all  $\hat{\lambda}_k$ . This is demonstrated by the following example [43, 27]. The characteristic polynomial of

$$A + \epsilon E = \left[ \begin{array}{ccc|cc} 0 & 1 & 0 & & \\ & 0 & 1 & & \\ & & 0 & \epsilon & \\ \hline & & & 0 & 1 \\ \epsilon & & & & 0 \end{array} \right] \quad (6)$$

is  $\epsilon^2 - \lambda^5$ . Thus,  $\gamma_k = 2/5$  for all  $\hat{\lambda}_k$  in (3) while  $1/n_1 = 1/3$ .

The purpose of this paper is to investigate various extensions of the condition number (4). In particular, we are interested in the case when  $E$  is restricted to a set  $\mathbb{S}$  of structured matrices, leading to the notion of a *structured Hölder condition number*  $\kappa(A, \lambda; \mathbb{S}) = (n_{\mathbb{S}}, \alpha_{\mathbb{S}})$ . We begin by noting that there exist structures  $\mathbb{S}$  for which  $n_{\mathbb{S}}$  can be smaller than  $n_1$ . Consider, for instance, the following example, taken from [29]:  $\mathbb{S}$  is the set of complex skew-symmetric matrices,

$$A = \begin{bmatrix} 0 & 1 & 0 & 1 \\ -1 & 0 & -i & 0 \\ 0 & i & 0 & i \\ -1 & 0 & -i & 0 \end{bmatrix} \in \mathbb{S}, \quad (7)$$

and  $\lambda = 0$  with geometric multiplicity two and largest Jordan block of size 3, i.e.,  $n_1 = 3$ . However, any complex skew-symmetric perturbation  $E \in \mathbb{S}$  gives rise to  $O(\epsilon^{1/2})$  perturbed eigenvalues so, according to Definition 3.1 below,  $n_{\mathbb{S}} = 2 < n_1$ .

However, for most structures under consideration we have  $n_{\mathbb{S}} = n_1$ . In this case, it makes sense to compare  $\alpha_{\mathbb{S}}$  with its unstructured counterpart  $\alpha$ . As will be shown in Section 3, many of the results from [3, 16, 32] on structured condition numbers for simple eigenvalues carry over to multiple eigenvalues. A notable exception to  $n_{\mathbb{S}} = n_1$  are complex skew-symmetric matrices, whose zero eigenvalues may exhibit  $n_{\mathbb{S}} < n_1$ ; this is exemplified by (7). The relation  $n_{\mathbb{S}} = n_1$  holds for nonzero eigenvalues of a complex skew-symmetric matrix, but  $\alpha_{\mathbb{S}}$  can be significantly smaller than  $\alpha$ , a fact that has been already observed in [32, 16]. In this paper, we not only provide additional insight by deriving explicit expressions for  $\alpha_{\mathbb{S}}$ , but also cover the more general class of matrices that are skew-symmetric with respect to an orthosymmetric bilinear form.

Hölder condition numbers for the generalized eigenvalues of a regular matrix pencil  $A - \lambda B$  can be defined similarly employing the perturbation expansions of Langer and Najman [17, 18, 19], see also [6]. Structured Hölder condition numbers for eigenvalues of pencils can be defined analogously, and have received lately some attention: results for simple eigenvalues of linearly structured pencils can be found in [9], and for multiple eigenvalues of definite Hermitian matrix pencils in [35, 36]. The problem of estimating the (multiple) eigenvalue sensitivities for parameter-dependent matrix pencils is closely related, see [1, 37, 45] and references therein. To our knowledge, the results provided in this paper on structured Hölder condition numbers for real, symmetric/skew-symmetric, Hermitian, as well as palindromic matrix pencils are new, even for simple eigenvalues. Furthermore, this framework also allows to cover matrix polynomial eigenvalue problems by imposing block companion structure.

The rest of this paper is organized as follows. In Section 2 we recall definitions and provide some basic results on unstructured and structured Hölder condition numbers for multiple eigenvalues of matrices. Section 3 is devoted to structured Hölder condition numbers for real, Toeplitz, and Hankel matrices, as well as for matrix classes that form Jordan or Lie algebras associated with an orthosymmetric bilinear or sesquilinear form. Section 4 is concerned with Hölder condition numbers for multiple eigenvalues of generalized eigenvalue problems, first for (structured) matrix pencils and then for matrix polynomials via companion form. Finally, some conclusions and open issues not addressed in this paper can be found in Section 5.

## 2 Preliminaries

In the following, we summarize the part of the discussion of Lidskii's results in [27] that leads to the condition number (4). Let  $\lambda$  be an eigenvalue of  $A \in \mathbb{C}^{n \times n}$  and let  $n_1$  be the size of the largest Jordan block corresponding to  $\lambda$ . The Jordan canonical form of  $A$ , can be written as

$$\left[ \begin{array}{c|c} J & 0 \\ \hline 0 & \tilde{J} \end{array} \right] = \left[ \begin{array}{c} Q \\ \hline \tilde{Q} \end{array} \right] A \left[ \begin{array}{c|c} P & \tilde{P} \end{array} \right], \quad (8)$$

where

$$\left[ \begin{array}{c} Q \\ \hline \tilde{Q} \end{array} \right] \left[ \begin{array}{c|c} P & \tilde{P} \end{array} \right] = I \quad (9)$$

and  $J$  consists of all  $n_1 \times n_1$  Jordan blocks corresponding to  $\lambda$ . Specifically, we have

$$J = \text{diag}(\Gamma_1^1, \dots, \Gamma_1^{r_1}), \quad \Gamma_1^1 = \dots = \Gamma_1^{r_1} = \begin{bmatrix} \lambda & 1 & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda \end{bmatrix} \in \mathbb{C}^{n_1 \times n_1}. \quad (10)$$

The block  $\tilde{J}$  contains all Jordan blocks corresponding to  $\lambda$  with dimension smaller than  $n_1$ , as well as all Jordan blocks corresponding to eigenvalues different from  $\lambda$ .

The columns of  $P$  form  $r_1$  linearly independent Jordan chains of length  $n_1$ , each of which starts with an eigenvector of  $A$ . Collecting these starting vectors in an  $n \times r_1$  matrix  $X$  yields

$$X = \left[ \begin{array}{c} P e_1, P e_{n_1+1}, \dots, P e_{(r_1-1)n_1+1} \end{array} \right], \quad (11)$$

where  $e_i$  denotes the  $i$ th unit vector of length  $n$ . Similarly we collect in

$$Y = \left[ \begin{array}{c} Q^H e_{n_1}, Q^H e_{2n_1}, \dots, Q^H e_{r_1 n_1} \end{array} \right] \quad (12)$$

the left eigenvectors chosen from the  $r_1$  independent Jordan chains of length  $n_1$  appearing as rows of  $Q$ . Note that each column of  $Y$  represents a left eigenvector of  $A$  belonging to  $\lambda$ . Notice also that the relation (9) implies  $Y^H X = I$  if  $n_1 = 1$ , and  $Y^H X = 0$  otherwise. With these preparations we can state a highly abridged version of Lidskii's result.

**Theorem 2.1 ([20, 27])** *Let  $E \in \mathbb{C}^{n \times n}$  such that  $Y^H E X$  is invertible, where  $X$  and  $Y$  are defined as above. Then there are  $n_1 r_1$  eigenvalues of the perturbed matrix  $A + \epsilon E$  admitting a perturbation expansion*

$$\hat{\lambda}_k = \lambda + (\xi_k)^{1/n_1} \epsilon^{1/n_1} + o(\epsilon^{1/n_1}), \quad k = 1, \dots, r_1, \quad (13)$$

where  $\xi_1, \dots, \xi_{r_1}$  are the eigenvalues of  $Y^H E X$ .

Since  $X$  and  $Y$  have linearly independent columns, the invertibility of  $Y^H E X$  is generically satisfied for a general perturbation  $E$  in  $\mathbb{C}^{n \times n}$ . Within a set  $\mathbb{S}$  of structured perturbations  $E$ , however, it may happen that  $Y^H E X$  is not generically invertible. Fortunately, the result of Theorem 2.1 remains valid even if  $Y^H E X$  is singular. This follows from a very general theory by Moro, Burke, and Overton [27] on the connection between Newton diagrams and eigenvalue perturbation expansions.

**Remark 2.2** (see [27, Pg. 809]) *Let  $E \in \mathbb{C}^{n \times n}$  such that  $Y^H EX$  is singular. Then each of the  $\beta < r_1$  nonzero eigenvalues  $\xi_1, \dots, \xi_\beta$  of  $Y^H EX$  gives rise to  $n_1$  perturbation expansions of the form (13). The remaining  $r_1 - \beta$  zero eigenvalues correspond to expansions where the exponent of the leading nonzero perturbation term is strictly larger than  $1/n_1$ .*

Theorem 2.1 implies that, for sufficiently small  $\epsilon$ , the worst-case change in  $\lambda$  is caused by an eigenvalue of  $Y^H EX$  that is as large as possible in magnitude. This motivates the following definition.

**Definition 2.3** ([27]) *Let  $\lambda$  be an eigenvalue of  $A \in \mathbb{C}^{n \times n}$ , and let the Jordan canonical form of  $A$  be given by (8). The (absolute) Hölder condition number of  $\lambda$  is given by  $\kappa(A, \lambda) = (n_1, \alpha)$ , where  $n_1$  is the size of the largest Jordan block associated with  $\lambda$  in (8) and*

$$\alpha = \sup_{\substack{\|E\| \leq 1 \\ E \in \mathbb{C}^{n \times n}}} \rho(Y^H EX), \quad (14)$$

where  $\rho(\cdot)$  denotes the spectral radius of a matrix.

We have  $\rho(Y^H EX) = \rho(EXY^H) \leq \|EXY^H\|_2 \leq \|XY^H\|_2$  for the matrix 2-norm  $\|\cdot\|_2$ . To show equality, we have to construct a perturbation  $E$  so that  $\rho(Y^H EX) = \|XY^H\|_2$  is attained. The following basic lemma helps identify such perturbations.

**Lemma 2.4** *Let*

$$XY^H = U\Sigma V^H$$

*be a singular value decomposition, such that  $U \in \mathbb{C}^{n \times r_1}$ ,  $V \in \mathbb{C}^{n \times r_1}$  have orthonormal columns and  $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_{r_1})$  with  $\sigma_1 \geq \dots \geq \sigma_{r_1} \geq 0$ . Consider  $E = VDU^H$  with  $D = \text{diag}(1, \delta_2, \dots, \delta_{r_1})$  such that  $\delta_j \leq 1$ . Then  $\rho(Y^H EX) = \|XY^H\|_2$ .*

*Proof.* The result follows from

$$\rho(Y^H EX) = \rho(EXY^H) = \rho(VD\Sigma V^H) = \rho(D\Sigma) = \|D\Sigma\|_2 = \|\Sigma\|_2 = \|XY^H\|_2.$$

□

Note that the definition of  $\alpha$  in (14) depends on the norm  $\|\cdot\|$  used in the constraint  $\|E\| \leq 1$ . For unitarily invariant norms, we have the following result.

**Theorem 2.5** ([27]) *The Hölder condition number  $(n_1, \alpha)$  of an eigenvalue  $\lambda$  satisfies  $\alpha = \|XY^H\|_2$  for any unitarily invariant norm  $\|\cdot\|$  in (14).*

*Proof.* Setting  $D = \text{diag}(1, 0, \dots, 0)$  and  $E = VDU^H$  in Lemma 2.4 gives  $\|E\| = 1$  and thus proves  $\alpha = \|XY^H\|_2$ . □

It is important to note that for a specific norm, other choices of  $D$  than the one used in the above proof are possible. E.g., for  $\|\cdot\| \equiv \|\cdot\|_2$ , any  $E$  in the sense of Lemma 2.4 gives  $\|E\|_2 = 1$ . In particular, setting  $D = \Sigma/\sigma_1$  yields

$$E = \frac{YX^H}{\|XY^H\|_2}, \quad (15)$$

which resembles the classical perturbation matrix for simple eigenvalues [43]. This type of perturbation will be often used when proving that the structured and the unstructured condition numbers coincide for the 2-norm. Another class of perturbations which turns out to be very useful is given by the following lemma.

**Lemma 2.6** *Let  $u_1, v_1 \in \mathbb{C}^{n \times n}$  with  $\|u_1\|_2 = \|v_1\|_2 = 1$  be left/right singular vectors belonging to the largest singular value of  $XY^H$ . Choose  $E \in \mathbb{C}^{n \times n}$  such that  $Eu_1 = \beta v_1$  with  $|\beta| = 1$ . Then  $\rho(EXY^H) \geq \|XY^H\|_2$ .*

*Proof.* Let  $XY^H = U\Sigma V^H$  be a singular value decomposition with  $U = [u_1, \dots, u_n]$ ,  $V = [v_1, \dots, v_n]$ . Then

$$\begin{aligned} \rho(EXY^H) &= \rho(EU\Sigma V^H) = \rho(V^H E U \Sigma) = \rho(V^H [\beta v_1, Ev_2, \dots, Ev_n] \Sigma) \\ &= \rho\left(\begin{bmatrix} \beta \|XY^H\|_2 & \star \\ 0 & \star \end{bmatrix}\right) \geq \|XY^H\|_2. \end{aligned}$$

□

### 3 Structured Hölder condition numbers

Throughout the whole section,  $\lambda$  denotes an eigenvalue of  $A$  with Hölder condition number  $\kappa(A, \lambda) = (n_1, \alpha)$ . The matrices  $X$  and  $Y$  are defined as in (11) and (12), respectively.

Restricting the range of admissible perturbations  $E$  from  $\mathbb{C}^{n \times n}$  to a subset  $\mathbb{S} \subset \mathbb{C}^{n \times n}$  leads to a corresponding structured condition number  $\kappa(A, \lambda; \mathbb{S}) = (n_{\mathbb{S}}, \alpha_{\mathbb{S}})$ .

**Definition 3.1** *Let  $\lambda$  be an eigenvalue of  $A \in \mathbb{C}^{n \times n}$ , and let  $\mathbb{S}$  be a subset of  $\mathbb{C}^{n \times n}$ . The (absolute) structured Hölder condition number of  $\lambda$  is given by  $\kappa(A, \lambda; \mathbb{S}) = (n_{\mathbb{S}}, \alpha_{\mathbb{S}})$ , where  $1/n_{\mathbb{S}}$  is the smallest possible power  $\gamma_k$  of  $\epsilon$  in the eigenvalue expansion (3) among all perturbations  $E \in \mathbb{S}$ , while  $\alpha_{\mathbb{S}} > 0$  is the largest possible magnitude of  $\alpha_k$  in (3) for all  $E \in \mathbb{S}$  with  $\|E\| \leq 1$ .*

As shown by example (7), it may happen that  $n_{\mathbb{S}} < n_1$ , but in this paper we focus on the cases when  $n_{\mathbb{S}} = n_1$ . If so, then by Theorem 2.1 and Remark 2.2 we can write

$$\alpha_{\mathbb{S}} = \sup_{\substack{\|E\| \leq 1 \\ E \in \mathbb{S}}} \rho(Y^H E X). \quad (16)$$

Note that the right-hand side in this expression becomes zero if and only if  $n_{\mathbb{S}} < n_1$ .

It turns out that the presence of the spectral radius in (16) considerably complicates the task of finding explicit formulas or reasonable bounds for  $\alpha_{\mathbb{S}}$ . However, we will see that it is often possible to identify structures with  $\alpha_{\mathbb{S}} \approx \alpha$  by constructing a perturbation  $E \in \mathbb{S}$  for which  $\rho(Y^H E X)$  is close to  $\alpha$ .

#### 3.1 Real matrices

As a first example, we point out that restricting the perturbation to be real can, at best, mildly improve the sensitivity of  $\lambda$ . This has been shown for a simple eigenvalue  $\lambda$  in [3]. The following lemma is a generalization to multiple eigenvalues.

**Lemma 3.2** *Let  $A \in \mathbb{C}^{n \times n}$ . We have  $\kappa(A, \lambda; \mathbb{R}^{n \times n}) = (n_1, \alpha_{\mathbb{R}})$  with*

- (i).  $\alpha/2 \leq \alpha_{\mathbb{R}} \leq \alpha$  in any unitarily invariant norm  $\|\cdot\|$ ;
- (ii). and  $\alpha_{\mathbb{R}} = \alpha$  in the matrix 2-norm,  $\|\cdot\| \equiv \|\cdot\|_2$ , provided that  $A$  is a normal matrix.

*Proof.* Decomposing  $XY^H = M_R + \imath M_I$  with  $M_R, M_I \in \mathbb{R}^{n \times n}$  gives  $\|M_R\|_2 \geq \|XY^H\|_2/2$  or  $\|M_I\|_2 \geq \|XY^H\|_2/2$ . W.l.o.g., we may assume  $\|M_R\|_2 \geq \|XY^H\|_2/2$ . Let us consider the perturbation  $E = v_1 u_1^T \in \mathbb{R}^{n \times n}$ , where  $u_1$  and  $v_1$  are normalized left and right singular vectors belonging to the largest singular value of  $M_R$ . Then  $\|E\| = 1$  and

$$\begin{aligned} \rho(EXY^H) &= \rho(v_1 u_1^T (M_R + \imath M_I)) = |u_1^T (M_R + \imath M_I) v_1| \\ &\geq |u_1^T M_R v_1| = \|M_R\|_2 \geq \|XY^H\|_2/2, \end{aligned}$$

which proves  $\alpha_{\mathbb{R}} \geq \alpha/2$ . To show the second part, note that we can choose  $Y = X$  if  $A$  is normal and thus  $E = I \in \mathbb{R}^{n \times n}$  gives  $\rho(EXX^H) = \alpha = \alpha_{\mathbb{R}}$ .  $\square$

**Remark 3.3** *In the case  $r_1 = 1$  (one single Jordan block of largest size  $n_1$ ), we can use the same arguments as in [3] to improve the lower bound of Lemma 3.2 (i) to  $\alpha/\sqrt{2} \leq \alpha_{\mathbb{R}}$ . It is not clear to us whether this slightly stronger result holds for  $r_1 > 1$ .*

Suppose that  $\mathbb{S}$  is a structure such that for any  $E \in \mathbb{S}$  the real and imaginary parts of  $E$  are both in  $\mathbb{S} \cap \mathbb{R}^{n \times n}$ . For a simple eigenvalue  $\lambda$ , Rump [32] has extended the bounds of Lemma 3.2 to structured condition numbers in the sense that restricting the perturbations from  $\mathbb{S}$  to  $\mathbb{S} \cap \mathbb{R}^{n \times n}$  improves the condition number by at most a factor of  $1/\sqrt{2}$ . By a trivial extension of [32, Lemma 3.1], this result holds for the case  $r_1 = 1$  but it is difficult to show that a similarly general result holds for an eigenvalue having multiple Jordan blocks of largest size. The following lemma is only a first step in this direction.

**Lemma 3.4** *Let  $\mathbb{SR}$  be a subset of  $\mathbb{R}^{n \times n}$  and let  $\mathbb{S} = \mathbb{SR} + \imath \mathbb{SR}$  be the set of all matrices with real and imaginary part in  $\mathbb{SR}$ . If there is a rank one matrix  $E \in \mathbb{S}$  with  $\|E\| = 1$  such that  $\alpha_{\mathbb{S}} = \rho(Y^H E X)$  then*

$$\alpha_{\mathbb{S}}/4 \leq \alpha_{\mathbb{SR}} \leq \alpha_{\mathbb{S}}$$

*in the Frobenius and the matrix 2-norm,  $\|\cdot\| \in \{\|\cdot\|_F, \|\cdot\|_2\}$ .*

*Proof.* We can write  $E = v u^H$  for  $u, v \in \mathbb{C}^{n \times n}$  with  $\|u\|_2 = \|v\|_2 = 1$ . Decomposing  $u = u_R + \imath u_I$  and  $v = v_R + \imath v_I$  with  $u_R, u_I, v_R, v_I \in \mathbb{R}^n$  gives

$$\alpha_{\mathbb{S}} = |u^H X Y^H v| = |(u_R^T X Y^H v_R + u_I^T X Y^H v_I) - \imath (u_I^T X Y^H v_R - u_R^T X Y^H v_I)|.$$

At least one of the two bracketed terms in this sum is not smaller than  $\alpha_{\mathbb{S}}/2$  in magnitude. Suppose  $|u_R^T X Y^H v_R + u_I^T X Y^H v_I| \geq \alpha_{\mathbb{S}}/2$  and set  $U = [u_R, u_I]$ ,  $V = [v_R, v_I]$ . Then  $|\text{trace}(U^T X Y^H V)| \geq \alpha_{\mathbb{S}}/2$  implying  $\rho(U^T X Y^H V) \geq \alpha_{\mathbb{S}}/4$ . Thus, the real perturbation  $E_R = V U^T$  (which is the real part of  $E$ ) yields  $\alpha_{\mathbb{SR}} \geq \rho(E_R X Y^H) \geq \alpha_{\mathbb{S}}/4$  while  $\|E_R\|_2 \leq 1$  and  $\|E_R\|_F \leq 1$ , which completes the proof. The case  $|u_I^T X Y^H v_R - u_R^T X Y^H v_I| \geq \alpha_{\mathbb{S}}/2$  is treated analogously.  $\square$

### 3.2 General linear structures

Let us briefly investigate the rather general case that  $\mathbb{S}$  is a linear matrix space in  $\mathbb{F}^{n \times n}$  with  $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ . Using an approach developed by Higham and Higham [9], we consider a fixed basis  $\{M_1, \dots, M_l\}$  of  $\mathbb{S}$  that is orthonormal with respect to the matrix inner product. Then for each perturbation  $E \in \mathbb{S}$  there is a unique vector  $p = [p_1, \dots, p_l]^T \in \mathbb{F}^l$  so that

$E = p_1 M_1 + \dots + p_l M_l$  and  $\|E\|_F = \|p\|_2$ . If  $n_{\mathbb{S}} = n_1$ , the structured condition number  $\kappa(A, \lambda; \mathbb{S}) = (n_1, \alpha_{\mathbb{S}})$  satisfies

$$\alpha_{\mathbb{S}} = \sup_{\substack{\|p\|_2 \leq 1 \\ p \in \mathbb{F}^l}} \rho(p_1 Y^H M_1 X + \dots + p_l Y^H M_l X) \quad (17)$$

for the Frobenius norm  $\|\cdot\| \equiv \|\cdot\|_F$ . Maximizing a nonsymmetric spectral function is known to be a nontrivial optimization problem, see, e.g., [2]. We therefore see little hope to find an explicit expression for  $\alpha_{\mathbb{S}}$  in general. There are two special cases for which  $\alpha_{\mathbb{S}}$  can be (nearly) determined.

1. The case  $r_1 = 1$  ( $X$  and  $Y$  are vectors) can be treated the same way as the case of simple eigenvalues [9, 40]. Defining the *pattern matrix*

$$\mathcal{M} = [\text{vec}(M_1), \dots, \text{vec}(M_l)], \quad (18)$$

where  $\text{vec}$  stacks the columns of a matrix into a long vector, we can write  $\text{vec}(E) = \mathcal{M}p$ , and therefore

$$\alpha_{\mathbb{S}} = \sup_{\substack{\|p\|_2 \leq 1 \\ p \in \mathbb{F}^l}} |p_1 Y^H M_1 X + \dots + p_l Y^H M_l X| = \|(X^T \otimes Y^H) \mathcal{M}\|_2$$

when  $\mathbb{F} = \mathbb{C}$ , or when  $\mathbb{F} = \mathbb{R}$  and  $X, Y \in \mathbb{R}^n$ . For  $\mathbb{F} = \mathbb{R}$  and  $X, Y \notin \mathbb{R}^n$ , we can show as in [16, Section 2] that  $\|(X^T \otimes Y^H) \mathcal{M}\|_2 / \sqrt{2} \leq \alpha_{\mathbb{S}} \leq \|(X^T \otimes Y^H) \mathcal{M}\|_2$ .

2. If  $\mathbb{F} = \mathbb{C}$  and all matrices  $N_j = Y^H M_j X$  are Hermitian then

$$\begin{aligned} \alpha_{\mathbb{S}} &= \sup_{\substack{\|p\|_2 \leq 1 \\ p \in \mathbb{C}^l}} \|p_1 N_1 + \dots + p_l N_l\|_2 \\ &= \sup_{\substack{\|x\|_2 = 1 \\ x \in \mathbb{C}^n}} \|[x^H N_1 x, \dots, x^H N_l x]\|_2. \end{aligned}$$

It follows that

$$\max_i \|N_i\|_2 \leq \alpha_{\mathbb{S}} \leq \sqrt{l} \max_i \|N_i\|_2.$$

### 3.3 Toeplitz and Hankel matrices

In [32], it is proven that the structured pseudospectrum of a matrix  $A \in \mathbb{S}$  coincides with the unstructured pseudospectrum for the following *complex* structures  $\mathbb{S}$ : symmetric, persymmetric, Toeplitz, symmetric Toeplitz, Hankel, persymmetric Hankel, and circulant. This implies, in particular, that  $\kappa(A, \lambda; \mathbb{S}) = \kappa(A, \lambda)$  for all these structures. Hence, some of the results that follow could be stated without proof. However, the proofs provided here explicitly construct structured perturbations that attain  $\kappa(A, \lambda)$ , which might lead to additional insight.

A *Toeplitz matrix* takes the form

$$T = \begin{bmatrix} t_0 & t_{-1} & \dots & t_{-n+1} \\ t_1 & t_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & t_{-1} \\ t_{n-1} & \dots & t_1 & t_0 \end{bmatrix} \in \mathbb{C}^{n \times n}$$

and  $H \in \mathbb{C}^{n \times n}$  is a *Hankel matrix* if  $F_n H$  is Toeplitz, where  $F_n$  is the  $n \times n$  *flip matrix* with ones on the anti-diagonal and zeros everywhere else.



**Theorem 3.5** *Let  $\mathbb{T}$  and  $\mathbb{SY}$  denote the sets of Toeplitz and complex symmetric matrices, respectively. Then the following statements hold in the matrix 2-norm:*

- (i).  $\kappa(A, \lambda; \mathbb{T}) = \kappa(A, \lambda) = (n_1, \|XX^T\|_2)$  for  $A \in \mathbb{T}$ ;
- (ii).  $\kappa(A, \lambda; \mathbb{T} \cap \mathbb{SY}) = \kappa(A, \lambda) = (n_1, \|XX^T\|_2)$  for  $A \in \mathbb{T} \cap \mathbb{SY}$ ;
- (iii).  $\kappa(A, \lambda; \mathbb{T} \cap \mathbb{R}^{n \times n}) = \kappa(A, \lambda) = (n_1, \|XX^T\|_2)$  for  $A \in \mathbb{T} \cap \mathbb{R}^{n \times n}$  and  $\lambda \in \mathbb{R}$ ;
- (iv).  $\kappa(A, \lambda; \mathbb{T} \cap \mathbb{SY} \cap \mathbb{R}^{n \times n}) = \kappa(A, \lambda) = (1, 1)$  for  $A \in \mathbb{T} \cap \mathbb{SY} \cap \mathbb{R}^{n \times n}$ .

*Proof.* A Toeplitz matrix is complex *persymmetric*, meaning that  $F_n T$  is complex symmetric. We can therefore apply Corollary A.2 (i) to conclude  $Y = F_n \bar{X}$ . A Takagi factorization [13, §4.4] of the complex symmetric matrix  $XX^T$  is a special singular value decomposition  $XX^T = U\Sigma U^T$ , where  $U \in \mathbb{C}^{n \times r_1}$  has orthonormal columns and  $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_{r_1})$  with  $\sigma_1 \geq \dots \geq \sigma_{r_1} > 0$ . By [31, Lemma 10.1], there is a Hankel matrix  $H$  with  $\|H\|_2 = 1$  and  $Hu_1 = \bar{u}_1$ , where  $u_1$  denotes the first column of  $U$ . Setting  $E = F_n H \in \mathbb{T}$  gives  $\|E\|_2 = 1$  with  $Eu_1 = F_n \bar{u}_1$ , which completes the proof of (i) by Lemma 2.6.

A symmetric Toeplitz matrix is persymmetric and symmetric; it can thus be block diagonalized by a simple orthogonal transformation:

$$G^T A G = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix},$$

where  $A_{11} \in \mathbb{R}^{\lceil n/2 \rceil \times \lceil n/2 \rceil}$ ,  $A_{22} \in \mathbb{R}^{\lfloor n/2 \rfloor \times \lfloor n/2 \rfloor}$  are complex symmetric and

$$G = \frac{1}{\sqrt{2}} \begin{bmatrix} I & F_{n/2} \\ -F_{n/2} & I \end{bmatrix} \text{ (even } n), \quad G = \frac{1}{\sqrt{2}} \begin{bmatrix} I & 0 & F_{(n-1)/2} \\ 0 & \sqrt{2} & 0 \\ -F_{(n-1)/2} & 0 & I \end{bmatrix} \text{ (odd } n).$$

This folklore result, which can be found for example in [44], shows that  $X = [X_1, X_2]$  with  $X_1 = -F_n X_1$  and  $X_2 = F_n X_2$ . The eigenvectors contained in  $X_1$  and  $X_2$  stem from Jordan blocks in  $A_{11}$  and  $A_{22}$ , respectively. Moreover,  $Y = F_n [\bar{X}_1, \bar{X}_2]$  and

$$\begin{aligned} \alpha_{\mathbb{T} \cap \mathbb{SY}} &= \sup_{\substack{\|E\|_2=1 \\ E \in \mathbb{T} \cap \mathbb{SY}}} \max \left( \rho(EX_1 X_1^T F_n), \rho(EX_2 X_2^T F_n) \right) \\ &= \sup_{\substack{\|E\|_2=1 \\ E \in \mathbb{T} \cap \mathbb{SY}}} \max \left( \rho(EX_1 X_1^T), \rho(EX_2 X_2^T) \right). \end{aligned}$$

From  $X_2^H X_1 = X_2^H F_n F_n X_1 = -X_2^H X_1$  it follows that  $X_2^H X_1 = 0$  and hence

$$\|XX^T\|_2 = \|[X_1, X_2][X_1, X_2]^T\|_2 = \max(\|X_1 X_1^T\|_2, \|X_2 X_2^T\|_2).$$

Let us assume  $\|X_1 X_1^T\|_2 \geq \|X_2 X_2^T\|_2$  (the other case is treated analogously) and let  $X_1 X_1^T = U\Sigma U^T$  be a Takagi factorization. Then  $U = -F_n U$  and by [32, Lemma 2.4] there is a symmetric Toeplitz matrix  $E$  such that  $\|E\|_2 = 1$  and  $Eu_1 = \bar{u}_1$ . The proof of (ii) is completed by applying Lemma 2.6.

Parts (iii) and (iv) are shown by noting that  $\lambda \in \mathbb{R}$  implies  $X \in \mathbb{R}^{n \times r_1}$  and hence the perturbations constructed above can be chosen to be real [32].  $\square$

Theorem 3.5 can be easily extended to Hankel matrices.

**Corollary 3.6** *Let  $\mathbb{HA}$  and  $\mathbb{PS}$  denote the sets of Hankel and persymmetric matrices, respectively. Then the following statements hold in the matrix 2-norm:*

- (i).  $\kappa(A, \lambda; \mathbb{HA}) = \kappa(A, \lambda) = (n_1, \|XX^T\|_2)$  for  $A \in \mathbb{HA}$ ;
- (ii).  $\kappa(A, \lambda; \mathbb{HA} \cap \mathbb{PS}) = \kappa(A, \lambda) = (n_1, \|XX^T\|_2)$  for  $A \in \mathbb{HA} \cap \mathbb{PS}$ ;
- (iii).  $\kappa(A, \lambda; \mathbb{HA} \cap \mathbb{R}^{n \times n}) = \kappa(A, \lambda) = (1, 1)$  for  $A \in \mathbb{HA} \cap \mathbb{R}^{n \times n}$ ;
- (iv).  $\kappa(A, \lambda; \mathbb{HA} \cap \mathbb{PS} \cap \mathbb{R}^{n \times n}) = \kappa(A, \lambda) = (1, 1)$  for  $A \in \mathbb{HA} \cap \mathbb{PS} \cap \mathbb{R}^{n \times n}$ .

*Proof.* A Hankel matrix is complex symmetric, which implies  $Y = \overline{X}$  by Corollary A.2 (i). The rest of the proof is along the lines of the proof of Theorem 3.5 and is therefore omitted.  $\square$

### 3.4 Symmetric, skew-symmetric and Hermitian matrices

The construction used in the proof of Theorem 3.5 only exploits the persymmetry of Toeplitz matrices and it can thus also be used to show  $\kappa(A, \lambda; \mathbb{PS}) = \kappa(A, \lambda)$  for  $A \in \mathbb{PS}$ . More general, we have the following result, which besides per-symmetric ( $M = F_n$ ), also includes symmetric ( $M = I$ ) and pseudo-symmetric ( $M = \text{diag}(I, -I)$ ) matrices.

**Theorem 3.7** *Let  $M \in \mathbb{R}^{n \times n}$  be an orthogonal symmetric matrix and define  $\mathbb{S} = \{A \in \mathbb{C}^{n \times n} : A^T M = MA\}$ . Then the following statements hold for  $A \in \mathbb{S}$  in any unitarily invariant norm:*

- (i).  $\kappa(A, \lambda; \mathbb{S}) = \kappa(A, \lambda) = (n_1, \|XX^T\|_2)$ ;
- (ii).  $\kappa(A, \lambda; \mathbb{S} \cap \mathbb{R}^{n \times n}) = (n_1, \alpha_{\mathbb{S} \cap \mathbb{R}^{n \times n}})$  with  $\|XX^T\|_2/2 \leq \alpha_{\mathbb{S} \cap \mathbb{R}^{n \times n}} \leq \|XX^T\|_2$ .

*Proof.* Corollary A.2 (i) gives  $Y = M\overline{X}$ . Let  $XX^T = U\Sigma U^T$  be a Takagi factorization, then we set  $u_1 = Ue_1$  and  $E = M\overline{u_1}u_1^H$  to obtain  $\|E\| = 1$  and

$$\alpha_{\mathbb{S}} \geq \rho(EXX^T M) = \rho(M\overline{u_1}u_1^H XX^T M) = \rho(u_1^H XX^T \overline{u_1}) = \|XX^T\|_2 = \alpha,$$

This completes the proof of the first part. The proof of the second part is virtually identical with the proof of Lemma 3.2 (i).  $\square$

Using the terminology of [16], Theorem 3.7 is concerned with Jordan algebras associated with the symmetric bilinear form  $\langle x, y \rangle = x^T M y$ . For the corresponding Lie algebras, which are given by  $\mathbb{S} = \{A \in \mathbb{C}^{n \times n} : A^T M = -MA\}$ , it is known that the structured and unstructured condition numbers for simple eigenvalues may widely differ [16, 32]. We have already shown in the introduction a skew-symmetric matrix (7) (corresponding to  $M = I$ ) such that  $n_{\mathbb{S}} < n_1$ . This shows that multiple eigenvalues can have a *qualitatively* better behavior under structured perturbations. The following theorem identifies one such situation, and proves that, in this setting, it can only happen under very specific conditions (namely, for zero eigenvalues with one single Jordan block of largest odd size). Additionally, Theorem 3.8 provides some insight on the expected difference between  $\alpha_{\mathbb{S}}$  and  $\alpha$  whenever  $n_{\mathbb{S}} = n_1$ .

**Theorem 3.8** *Let  $M \in \mathbb{R}^{n \times n}$  be an orthogonal symmetric matrix and define  $\mathbb{S} = \{A \in \mathbb{C}^{n \times n} : A^T M = -MA\}$ . Then the following statements on  $\kappa(A, \lambda; \mathbb{S}) = (n_{\mathbb{S}}, \alpha_{\mathbb{S}})$  hold for  $A \in \mathbb{S}$  in the matrix 2-norm.*

- (i). *If  $\lambda = 0$ ,  $n_1$  is odd, and  $r_1 = 1$ , then  $n_{\mathbb{S}} < n_1$ ;*

(ii). if  $\lambda = 0$ ,  $n_1$  is odd, and  $r_1 > 1$ , then  $n_{\mathbb{S}} = n_1$  and  $\alpha_{\mathbb{S}} = \sqrt{\sigma_1\sigma_2}$ , where  $\sigma_1, \sigma_2$  are the two largest singular values of  $XX^T$  (whereas  $\alpha = \sigma_1$ );

(iii). if  $\lambda = 0$  and  $n_1$  is even, then  $r_1$  is even,  $n_{\mathbb{S}} = n_1$  and  $\alpha_{\mathbb{S}} = \alpha = \left\| X \begin{bmatrix} 0 & -I_{r_1/2} \\ I_{r_1/2} & 0 \end{bmatrix} X^T \right\|_2$ ;

(iv). if  $\lambda \neq 0$  and  $r_1 = 1$ , then  $n_{\mathbb{S}} = n_1$  and  $\alpha_{\mathbb{S}} = \sqrt{\|X\|_2^2\|Y\|_2^2 - |Y^T M X|^2}$ .

(v). if  $\lambda \neq 0$  and  $r_1 > 1$ , then  $n_{\mathbb{S}} = n_1$ .

*Proof.* For  $\lambda = 0$  with  $n_1$  odd, Corollary A.4 (ia) implies  $Y = M\bar{X}$ . Now, if  $n_{\mathbb{S}}$  was equal to  $n_1$  then there would exist some  $E \in \mathbb{S}$  with  $\rho(Y^H E X) = \rho(X^T M E X) > 0$ . This is impossible for  $r_1 = 1$ , since  $X$  is a vector and  $ME$  is skew-symmetric, so  $\rho(X^T M E X) = |X^T M E X| = 0$ . This shows (i).

To show  $n_{\mathbb{S}} = n_1$  for  $\lambda = 0$ , odd  $n_1$  and  $r_1 > 1$ , it is sufficient to construct a perturbation  $E \in \mathbb{S}$  such that  $\rho(X^T M E X) > 0$ . For this purpose, consider a Takagi factorization  $XX^T = U\Sigma U^T$ , where  $U = [u_1, \dots, u_{r_1}]$  has orthonormal columns and  $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_{r_1})$  with  $\sigma_1 \geq \dots \geq \sigma_{r_1} > 0$ . Setting  $E = M[\bar{u}_1, \bar{u}_2] \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} [u_1, u_2]^H$  gives  $E \in \mathbb{S}$ ,  $\|E\|_2 = 1$ , and

$$\alpha_{\mathbb{S}} \geq \rho(X^T M E X) = \rho \left( \begin{bmatrix} 0 & \sigma_2 \\ -\sigma_1 & 0 \end{bmatrix} \right) = \sqrt{\sigma_1\sigma_2} > 0.$$

On the other hand, letting  $\mathbb{SK}$  denote the set of complex skew-symmetric matrices, we have

$$\begin{aligned} \alpha_{\mathbb{S}} &= \sup_{\substack{\|\tilde{E}\|_2 \leq 1 \\ \tilde{E} \in \mathbb{SK}}} \rho(\tilde{E} X X^T) = \sup_{\substack{\|G\|_2 \leq 1 \\ G \in \mathbb{SK}}} \rho(G\Sigma) = \sup_{\substack{\|G\|_2 \leq 1 \\ G \in \mathbb{SK}}} \rho(\Sigma^{1/2} G \Sigma^{1/2}) \\ &\leq \sup_{\substack{\|G\|_2 \leq 1 \\ G \in \mathbb{SK}}} \|\Sigma^{1/2} G \Sigma^{1/2}\|_2 = \sup_{\substack{\|G\|_2 \leq 1 \\ G \in \mathbb{SK}}} \|\tilde{\Sigma} \circ G\|_2, \end{aligned}$$

where  $\tilde{\Sigma} = [\sqrt{\sigma_i\sigma_j}]_{i,j=1}^{r_1}$  and  $\circ$  denotes the Hadamard product. A result by Mathias [24, Corollary 2.6] implies  $\|\tilde{\Sigma} \circ G\|_2 \leq \sqrt{\sigma_1\sigma_2}\|G\|_2$ , which concludes the proof of (ii).

For  $\lambda = 0$  with  $n_1$  odd, Corollary A.4 (ib) implies  $Y = M\bar{X} \begin{bmatrix} 0 & I_{r_1/2} \\ -I_{r_1/2} & 0 \end{bmatrix}$ . To attain  $\rho(Y^H E X) = \alpha = \left\| X \begin{bmatrix} 0 & -I_{r_1/2} \\ I_{r_1/2} & 0 \end{bmatrix} X^T \right\|_2$  we may just use a perturbation of the form (15), which in this case turns out to be  $E = \frac{1}{\alpha} M\bar{X} \begin{bmatrix} 0 & -I_{r_1/2} \\ I_{r_1/2} & 0 \end{bmatrix} X^H \in \mathbb{S}$ . This proves (iii).

If  $\lambda \neq 0$  then  $-\lambda$  is also an eigenvalue with the same Jordan structure as  $\lambda$ . If we let  $\tilde{X}, \tilde{Y}$  denote the matrices of right/left eigenvectors stemming from the  $n_1 \times n_1$  Jordan blocks belonging to  $-\lambda$  then Corollary A.4 (ic) yields  $\tilde{X} = -M\tilde{Y}$  and  $\tilde{Y} = M\tilde{X}$ . This not only implies  $\kappa(A, -\lambda) = \kappa(A, \lambda)$  as well as  $\kappa(A, -\lambda; \mathbb{S}) = \kappa(A, \lambda; \mathbb{S})$ , but also that  $[X, M\tilde{Y}]$  has full column rank. If  $r_1 = 1$  then  $X, Y$  are vectors and we have

$$|Y^H E X| = \rho([M\tilde{Y}, X]^T M E [M\tilde{Y}, X]) = \rho(M E [M\tilde{Y}, X] [M\tilde{Y}, X]^T)$$

for any  $E \in \mathbb{S}$ . Hence, using the arguments from the proof of (ii), we have  $\alpha_{\mathbb{S}} = \sqrt{\sigma_1\sigma_2}$ , where  $\sigma_1$  and  $\sigma_2$  are the two largest singular values of the symmetric matrix  $[M\tilde{Y}, X] [M\tilde{Y}, X]^T$ .

This shows (iv) since

$$\begin{aligned}
\sigma_1\sigma_2 &= \sqrt{\det\left([M\bar{Y}, X][M\bar{Y}, X]^T[M\bar{Y}, \bar{X}][M\bar{Y}, \bar{X}]^T\right)} \\
&= \sqrt{\det\left([M\bar{Y}, \bar{X}]^T[M\bar{Y}, X][M\bar{Y}, X]^T[M\bar{Y}, \bar{X}]\right)} \\
&= \left|\det\left([M\bar{Y}, \bar{X}]^T[M\bar{Y}, X]\right)\right| \\
&= \|X\|_2^2\|Y\|_2^2 - |Y^T M X|^2.
\end{aligned}$$

Unfortunately, the technique of this proof does not extend to the case  $r_1 > 1$ . Still, we can show  $\alpha_{\mathbb{S}} > 0$  but it is not clear how to obtain a good lower or upper bound on  $\alpha_{\mathbb{S}}$ . The full column rank of  $[X, M\bar{Y}]$  implies the existence of an invertible matrix  $L$  such that

$$L^{-1}[X, M\bar{Y}] = \begin{bmatrix} I_{r_1} & \star \\ 0 & I_{r_1} \\ 0 & 0 \end{bmatrix}$$

Setting

$$E = ML^{-T} \begin{bmatrix} 0 & I_{r_1} & 0 \\ -I_{r_1} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} L^{-1} \in \mathbb{S}$$

yields  $\rho(Y^H E X) = \rho(I_r) = 1$  and thus  $\alpha_{\mathbb{S}} > 0$ , completing the proof of (v).  $\square$

**Remark 3.9** *Note that Theorem 3.8 (iv) also improves the results in [16, Theorem 4.3] and [32, Theorem 3.2], which only state bounds, but no explicit formula for the structured condition number of a simple nonzero eigenvalue. Recently, Karow [15] described the limit sets of the structured pseudospectra for complex skew-symmetric matrices, from which Theorem 3.8 (iv) could also be derived.*

Fortunately, the matter of structured condition numbers is much less complicated for Jordan and Lie algebras associated with a sesquilinear form  $\langle x, y \rangle = x^H M y$ .

**Lemma 3.10** *Let  $M \in \mathbb{R}^{n \times n}$  be an orthogonal symmetric or orthogonal skew-symmetric matrix and define  $\mathbb{S} = \{A \in \mathbb{C}^{n \times n} : A^H M = \gamma M A\}$  for a fixed  $\gamma \in \{1, -1\}$ . Then for any  $A \in \mathbb{C}^{n \times n}$ ,  $\kappa(A, \lambda; \mathbb{S}) = \kappa(A, \lambda)$  holds in the matrix 2-norm.*

*Proof.* Let  $XY^H = U\Sigma V^H$  be a singular value decomposition and set  $u_1 = Ue_1, v_1 = Ve_1$ . Then  $\|u_1\|_2 = \|v_1\|_2 = 1$  and by [22, Theorem 8.6] we can find a Hermitian matrix  $H$  such that  $\|H\|_2 = 1$  and  $Hu_1 = \mu M v_1$  for some  $\mu \in \mathbb{C}$  with  $|\mu| = 1$ . Set  $E = \sqrt{\gamma} M H$  if  $M = M^T$ , and  $E = \sqrt{-\gamma} M H$  if  $M = -M^T$ . Then  $E \in \mathbb{S}$  satisfies  $\|E\|_2 = 1$  and  $Eu_1 = \beta v_1$  for some  $|\beta| = 1$ , which implies the result by Lemma 2.6.  $\square$

### 3.5 $J$ -symmetric and $J$ -skew-symmetric matrices

For  $M = J_{2n} = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$ , the structure  $\mathbb{S} = \{A \in \mathbb{C}^{2n \times 2n} : A^H M = \gamma M A\}$  considered in Lemma 3.10 coincides with the set of *complex* skew-Hamiltonian matrices if  $\gamma = 1$ , and with the set of *complex* Hamiltonian matrices if  $\gamma = -1$ . The following two theorems provide results for the closely related structures  $\mathbb{S} = \{A \in \mathbb{C}^{2n \times 2n} : A^T J_{2n} = \gamma J_{2n} A\}$ , including bounds on structured condition numbers for *real* skew-Hamiltonian and Hamiltonian matrices.

**Theorem 3.11** *Let  $M \in \mathbb{R}^{2n \times 2n}$  be an orthogonal skew-symmetric matrix and define  $\mathbb{S} = \{A \in \mathbb{C}^{2n \times 2n} : A^T M = MA\}$ . Then the following statements hold for  $A \in \mathbb{S}$  in the matrix 2-norm:*

$$(i). \quad \kappa(A, \lambda; \mathbb{S}) = \kappa(A, \lambda) = (n_1, \|XJ_{r_1}X^T\|_2);$$

$$(ii). \quad \kappa(A, \lambda; \mathbb{S} \cap \mathbb{R}^{2n \times 2n}) = (n_1, \alpha_{\mathbb{S} \cap \mathbb{R}^{2n \times 2n}}) \text{ with } \|XJ_{r_1}X^T\|_2/4 \leq \alpha_{\mathbb{S} \cap \mathbb{R}^{2n \times 2n}} \leq \|XJ_{r_1}X^T\|_2.$$

*Proof.* Corollary A.6 reveals the relation  $Y = -M\bar{X}J_{r_1}$ . Using a perturbation as in (15), namely  $E = (M^T\bar{X}J_{r_1}X^H)/\|XJ_{r_1}X^T\|_2$ , yields  $E \in \mathbb{S}$  with  $\|E\|_2 = 1$  and

$$\alpha_{\mathbb{S}} \geq \rho(J_{r_1}^T X^T M E X) / \|XJ_{r_1}X^T\|_2 = \|XJ_{r_1}X^T\|_2 = \alpha.$$

To prove the second part, let  $u = u_R + \iota u_I$  and  $v = v_R + \iota v_I$ , with  $u_R, u_I, v_R, v_I \in \mathbb{R}^n$  and  $\|u\|_2 = \|v\|_2 = 1$ , be left/right singular vectors corresponding to the largest singular value of  $K = XJ_{r_1}X^T$ . Then

$$\alpha_{\mathbb{S}} = \|K\|_2 = u^H K v = (u_R^T K v_R + u_I^T K v_I) + \iota(u_R^T K v_I - u_I^T K v_R).$$

At least one of the four terms in this sum is not smaller in magnitude than  $\alpha_{\mathbb{S}}/4$ . Choose this term and let the columns of  $W = [w_1, w_2] \in \mathbb{R}^{2n \times 2}$  contain the two vectors corresponding to it. For example, if  $|u_R^T K v_R| \geq \alpha_{\mathbb{S}}/4$  then  $W = [u_R, v_R]$ . By the skew-symmetry of  $K$ , we may assume that  $u$  and  $v$  satisfy  $v^T u = 0$ , which implies  $\|W\|_2 \leq 1$ . Setting  $E = M^T W J_2 W^T \in \mathbb{S} \cap \mathbb{R}^{2n \times 2n}$  yields  $\|E\|_2 \leq 1$  and

$$\begin{aligned} \alpha_{\mathbb{S} \cap \mathbb{R}^{2n \times 2n}} &\geq \rho(J_{r_1}^T X^T M E X) = \rho(J_{r_1}^T X^T W J_2 W^T X) \\ &= \rho(K W J_2 W^T) = \rho(J_2 W^T K W) \\ &= \rho(\text{diag}(w_2^T K w_1, -w_1^T K w_2)) \geq \alpha_{\mathbb{S}}/4, \end{aligned}$$

where we used the fact that  $w_1^T K w_1 = w_2^T K w_2 = 0$  due to the skew-symmetry of  $K$ .  $\square$

**Theorem 3.12** *Let  $M \in \mathbb{R}^{2n \times 2n}$  be an orthogonal skew-symmetric matrix and define  $\mathbb{S} = \{A \in \mathbb{C}^{2n \times 2n} : A^T M = -MA\}$ . Then the following statements hold for any  $A \in \mathbb{C}^{n \times n}$ :*

$$(i). \quad \kappa(A, \lambda; \mathbb{S}) = (n_1, \alpha_{\mathbb{S}}) \text{ with } \alpha/\sqrt{2} \leq \alpha_{\mathbb{S}} \leq \alpha \text{ in the Frobenius norm and } \alpha_{\mathbb{S}} = \alpha \text{ in the matrix 2-norm;}$$

$$(ii). \quad \kappa(A, \lambda; \mathbb{S} \cap \mathbb{R}^{2n \times 2n}) = (n_1, \alpha_{\mathbb{S} \cap \mathbb{R}^{2n \times 2n}}) \text{ with } \alpha/8 \leq \alpha_{\mathbb{S} \cap \mathbb{R}^{2n \times 2n}} \leq \alpha_{\mathbb{S}} \text{ in the matrix 2-norm.}$$

*Proof.* Let  $u_1, v_1$  with  $\|u_1\|_2 = \|v_1\|_2 = 1$  be the left/right singular vectors belonging to the largest singular value of  $XY^H M$  and define  $\tilde{E} = [v_1, \bar{u}_1] \begin{bmatrix} 0 & \\ 1 & -v_1^T u_1 \end{bmatrix} [v_1, \bar{u}_1]^T$ . Then  $\tilde{E}$  is symmetric and one can show that  $\|\tilde{E}\|_F = \sqrt{2 - |u_1^T v_1|^2} \leq \sqrt{2}$ , see also [23, Theorem 5.6]. Setting  $E = M\tilde{E}/\sqrt{2}$ , we obtain  $E \in \mathbb{S}$  and

$$\alpha_{\mathbb{S}} \geq \rho(Y^H E X) = \rho(\tilde{E} X Y^H M) / \sqrt{2} \geq \|XY^H\|_2 / \sqrt{2},$$

where we applied Lemma 2.6, using the fact that  $\tilde{E}$  maps  $u_1$  to  $v_1$ . In the matrix 2-norm, Theorem 5.7 in [23] implies the existence of a symmetric matrix  $\tilde{E}$  which maps  $u_1$  to  $v_1$  and satisfies  $\|\tilde{E}\|_2 = 1$ . Thus, setting  $E = M\tilde{E}$  shows the second part of (i).

To show (ii), let us decompose  $XY^H M = S + W$ , where  $S = (XY^H + \bar{Y}X^T)/2$  and  $W = (XY^H - \bar{Y}X^T)/2$ . Then  $\alpha = \|XY^H\|_2 \leq \|S\|_2 + \|W\|_2$ . We distinguish two cases, depending on whether the skew-symmetric part  $W$  dominates the symmetric part  $S$ .

1.  $\|S\|_2 \geq \|W\|_2/3$ : Decompose  $S = S_R + \iota S_I$  with real symmetric matrices  $S_R, S_I$ . Then  $\|S_R\|_2 \geq \|S\|_2/2$  or  $\|S_I\|_2 \geq \|S\|_2/2$ . In the first case, let  $u_1$  be a normalized eigenvector belonging to an eigenvalue of  $S_R$  that has magnitude  $\|S_R\|_2$ . Then  $E = Mu_1u_1^T \in \mathbb{S} \cap \mathbb{R}^{2n \times 2n}$  with  $\|E\|_2 = 1$  and

$$\begin{aligned} \alpha_{\mathbb{S} \cap \mathbb{R}^{n \times n}} &\geq \rho(EXY^H) = |u_1^T XY^H Mu_1| = |u_1^T Su_1| \geq |u_1^T S_R u_1| \\ &\geq \frac{\|S\|_2}{2} = \frac{\|S\|_2 + 3\|S\|_2}{8} \geq \frac{\|S\|_2 + \|W\|_2}{8} \geq \frac{\alpha}{8}. \end{aligned}$$

The case  $\|S_I\|_2 \geq \|S\|_2/2$  can be shown analogously.

2.  $\|S\|_2 \leq \|W\|_2/3$ : Decompose  $W = W_R + \iota W_I$  with real skew-symmetric matrices  $W_R, W_I$ . Suppose that  $\|W_R\|_2 \geq \|W\|_2/2$  (once again, the case  $\|W_I\|_2 \geq \|W\|_2/2$  is treated in an analogous manner). Let  $u_1, v_1$  with  $\|u_1\|_2 = \|v_1\|_2 = 1$  be left/right singular vectors belonging to the largest singular value of  $W_R$ . Since  $W_R$  is skew-symmetric, we have  $v_1^T u_1 = 0$ . Setting

$$E = M[u_1, v_1] \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} [u_1, v_1]^T \in \mathbb{S} \cap \mathbb{R}^{2n \times 2n}$$

yields  $\|E\|_2 = 1$  and

$$\alpha_{\mathbb{S} \cap \mathbb{R}^{2n \times 2n}} \geq \rho(EXY^H) = \rho\left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} [u_1, v_1]^T (S + W) [u_1, v_1]\right) = \rho(\Phi),$$

where

$$\Phi = \begin{bmatrix} -\beta & 0 \\ 0 & \beta \end{bmatrix} + \begin{bmatrix} u_1^T S v_1 & v_1^T S v_1 \\ u_1^T S u_1 & u_1^T S v_1 \end{bmatrix}$$

with  $\beta = \|W_R\|_2 + u_1^T W_I v_1$ . We have  $\det(\Phi) = -(\beta + \gamma)(\beta - \gamma)$  with

$$\gamma = \sqrt{(u_1^T S u_1)(v_1^T S v_1) - (u_1^T S v_1)^2}$$

satisfying  $|\gamma| \leq \|S\|_2$ . This shows

$$\begin{aligned} \rho(\Phi) &\geq |\beta| - |\gamma| \geq \|W_R\|_2 - \|S\|_2 \geq \frac{\|W\|_2}{2} - \|S\|_2 \\ &= \frac{\|W\|_2}{2} - \frac{9}{8}\|S\|_2 + \frac{1}{8}\|S\|_2 \geq \frac{\|S\|_2 + \|W\|_2}{8} \geq \frac{\alpha}{8}, \end{aligned}$$

which concludes the proof.

□

Theorem 3.12 (ii) reveals that forcing the perturbations in a real Hamiltonian matrix to respect the structure will generally only have a mild positive effect on the accuracy of multiple eigenvalues. However, it should be emphasized that condition numbers provide little insight on the direction in which perturbed eigenvalues are likely to move, an issue which is crucial to decide whether a purely imaginary eigenvalue of a Hamiltonian matrix stays on the imaginary axis or not under (structured) perturbations, something which is often important in applications. For results in this direction, see [1, 26] and references therein.

## 4 Generalized eigenvalue problems

### 4.1 Matrix pencils

Langer and Najman [17, 18, 19] extended Lidskii's perturbation theory, obtaining eigenvalue perturbation expansions for analytic matrix functions  $L(\lambda)$ . They used the local Smith form of  $L(\lambda)$  much in the same way as the Jordan canonical form was used by Lidskii for matrices. In a recent paper, de Terán, Dopico and Moro [6] have investigated the special case  $L(\lambda) = A - \lambda B$ , relating the results by Langer and Najman to the Kronecker-Weierstraß form, which is a more natural canonical form when  $L(\lambda)$  is a matrix pencil. Let us briefly recall these results, restricting our attention to regular matrix pencils ( $A$  and  $B$  are square,  $\det(L(\lambda)) \neq 0$ ).

In the following, we denote the regular matrix pencil  $A - \lambda B$  by  $(A, B)$ . For a finite eigenvalue  $\lambda$  of  $(A, B)$ , the Kronecker-Weierstraß form implies

$$\left[ \begin{array}{c|c} J & 0 \\ \hline 0 & \tilde{J}_A \end{array} \right] = \left[ \begin{array}{c} Q \\ \hline \tilde{Q} \end{array} \right] A \left[ \begin{array}{c|c} P & \tilde{P} \end{array} \right], \quad \left[ \begin{array}{c|c} I & 0 \\ \hline 0 & \tilde{J}_B \end{array} \right] = \left[ \begin{array}{c} Q \\ \hline \tilde{Q} \end{array} \right] B \left[ \begin{array}{c|c} P & \tilde{P} \end{array} \right], \quad (19)$$

where  $[P, \tilde{P}]$ ,  $\left[ \begin{array}{c} Q \\ \hline \tilde{Q} \end{array} \right]$  are invertible and  $J$  contains all  $r_1$  Jordan blocks of largest size  $n_1$ , see also (10). Similarly for an infinite eigenvalue of  $(A, B)$ , we have

$$\left[ \begin{array}{c|c} I & 0 \\ \hline 0 & \tilde{J}_A \end{array} \right] = \left[ \begin{array}{c} Q \\ \hline \tilde{Q} \end{array} \right] A \left[ \begin{array}{c|c} P & \tilde{P} \end{array} \right], \quad \left[ \begin{array}{c|c} N & 0 \\ \hline 0 & \tilde{J}_B \end{array} \right] = \left[ \begin{array}{c} Q \\ \hline \tilde{Q} \end{array} \right] B \left[ \begin{array}{c|c} P & \tilde{P} \end{array} \right], \quad (20)$$

where  $N$  contains all  $r_1$  nilpotent blocks of largest nilpotency index  $n_1$ . As for the standard eigenvalue problem, we collect the (generalized) right and left eigenvectors contained in  $P$  and  $Q$ :

$$\begin{aligned} X &= \left[ P e_1, P e_{n_1+1}, \dots, P e_{(r_1-1)n_1+1} \right], \\ Y &= \left[ Q^H e_{n_1}, Q^H e_{2n_1}, \dots, Q^H e_{r_1 n_1} \right]. \end{aligned} \quad (21)$$

As in the standard eigenvalue problem, this relationship between  $X, Y$  and  $P, Q$  imposes some normalization on  $X, Y$ . For  $n_1 = 1$ , we have  $Y^H B X = I$  if  $\lambda$  is finite and  $Y^H A X = I$  if  $\lambda$  is infinite. For  $n_1 > 1$ , we have  $Y^H A X = Y^H B X = 0$ .

The following theorem summarizes results from [6].

**Theorem 4.1** *Let  $\lambda$  be a finite eigenvalue of a regular matrix pencil  $(A, B)$  and let  $(E, F) \in \mathbb{C}^{n \times n} \times \mathbb{C}^{n \times n}$  be such that  $Y^H(E - \lambda F)X$  is invertible, where  $X$  and  $Y$  are defined as in (21). Then there are  $n_1 r_1$  eigenvalues  $\hat{\lambda}_k$  of the perturbed pencil  $(A + \epsilon E, B + \epsilon F)$  admitting a perturbation expansion*

$$\hat{\lambda}_k = \lambda + (\xi_k)^{1/n_1} \epsilon^{1/n_1} + o(\epsilon^{1/n_1}), \quad k = 1, \dots, r_1, \quad (22)$$

where  $\xi_1, \dots, \xi_{r_1}$  are the eigenvalues of  $Y^H(E - \lambda F)X$ . For an infinite eigenvalue of  $(A, B)$ , let  $F \in \mathbb{C}^{n \times n}$  be such that  $Y^H F X$  is invertible. Then there are  $n_1 r_1$  eigenvalues  $\hat{\lambda}_k$  of the perturbed pencil  $(A + \epsilon E, B + \epsilon F)$  admitting a perturbation expansion

$$\frac{1}{\hat{\lambda}_k} = (\xi_k)^{1/n_1} \epsilon^{1/n_1} + o(\epsilon^{1/n_1}), \quad k = 1, \dots, r_1, \quad (23)$$

where  $\xi_k$  are the eigenvalues of  $Y^H F X$ .

#### 4.1.1 Hölder condition numbers for multiple eigenvalues of matrix pencils

Throughout the rest of this section,  $\lambda$  denotes a finite or infinite eigenvalue of a regular matrix pencil  $(A, B)$  with the matrices  $X$  and  $Y$  defined as in (21).

Based on Theorem 4.1, we can define a condition number for a multiple eigenvalue of a matrix pencil as follows.

**Definition 4.2** *The (absolute) Hölder condition number for a finite eigenvalue  $\lambda$  is given by  $\kappa(A, B, \lambda) = (n_1, \alpha)$ , where  $n_1$  is the dimension of the largest Jordan block associated with  $\lambda$  and*

$$\alpha = \sup_{\substack{\|E\| \leq w_A, \|F\| \leq w_B \\ E, F \in \mathbb{C}^{n \times n}}} \rho(Y^H(E - \lambda F)X).$$

*The (absolute) Hölder condition number for  $\lambda = \infty$  is given by  $\kappa(A, B, \infty) = (n_1, \alpha)$ , where  $n_1$  is the nilpotency index of  $(A, B)$  and*

$$\alpha = \sup_{\substack{\|F\| \leq w_B \\ F \in \mathbb{C}^{n \times n}}} \rho(Y^H F X).$$

**Remark 4.3** *Note that Definition 4.2 not only depends on the employed matrix norm  $\|\cdot\|$  but also on the choice of nonnegative weights  $w_A$  and  $w_B$ . It is implicitly assumed that  $w_A$  or  $w_B$  are strictly larger than zero, otherwise  $\kappa(A, B, \lambda) = (0, 0)$ . More specifically, we require  $w_A > 0$  for  $\lambda = 0$ ,  $w_B > 0$  for  $\lambda = \infty$ , and  $\max\{w_A, w_B\} > 0$  for any other eigenvalue.*

The role of the weights  $w_A$  and  $w_B$  is to balance the influence of perturbations on  $A$  and  $B$ . For example, if each of the perturbations  $E$  and  $F$  is known to be small compared to the norm of  $A$  and  $B$ , respectively, then it is reasonable to set  $w_A = \|A\|/\sqrt{\|A\|^2 + \|B\|^2}$  and  $w_B = \|B\|/\sqrt{\|A\|^2 + \|B\|^2}$ .

The following lemma represents a direct extension of [27, Theorem 4.2].

**Lemma 4.4** *For any unitarily invariant norm, we have*

$$\kappa(A, B, \lambda) = (n_1, (w_A + w_B|\lambda|)\|XY^H\|_2)$$

*if  $\lambda$  is a finite eigenvalue, and  $\kappa(A, B, \lambda) = (n_1, w_B\|XY^H\|)$  if  $\lambda = \infty$ .*

*Proof.* On the one hand

$$\begin{aligned} \rho(Y^H(E - \lambda F)X) &= \rho((E - \lambda F)XY^H) \leq \|(E - \lambda F)XY^H\|_2 \\ &\leq \|(E - \lambda F)XY^H\| \leq (w_A + w_B|\lambda|)\|XY^H\|_2 \end{aligned}$$

holds for any  $E, F$  satisfying  $\|E\| \leq w_A, \|F\| \leq w_B$ . Hence,  $\alpha \leq (w_A + w_B|\lambda|)\|XY^H\|_2$ . On the other hand, let  $u_1, v_1$  with  $\|u_1\|_2 = \|v_1\|_2 = 1$  be the left/right singular vectors belonging to the largest singular value of  $XY^H$ . Setting  $E = w_A v_1 u_1^H$  and  $F = -\frac{\lambda}{|\lambda|} w_B v_1 u_1^H$  ( $F = 0$  if  $\lambda = 0$ ) yields  $\|E\| \leq w_A, \|F\| \leq w_B$ , and

$$\alpha \geq \rho((w_A + w_B|\lambda|)v_1 u_1^H XY^H) = (w_A + w_B|\lambda|)\rho(u_1^H XY^H v_1) = (w_A + w_B|\lambda|)\|XY^H\|_2.$$

The proof for  $\kappa(A, B, \infty)$  is analogous.  $\square$



Definition 4.2 is based on the distance  $|\hat{\lambda} - \lambda|$  between an eigenvalue  $\lambda$  and a perturbed eigenvalue  $\hat{\lambda}$ . This distance lacks mathematical elegance for generalized eigenvalue problems, since infinite eigenvalues must be treated separately and  $|\hat{\lambda} - \lambda|$  is not invariant under an interchange of  $A$  and  $B$ , i.e.,  $|\hat{\lambda} - \lambda| \neq |1/\hat{\lambda} - 1/\lambda|$ . A more elegant distance concept is offered by the chordal metric

$$\chi(\hat{\lambda}, \lambda) = \frac{|\hat{\lambda} - \lambda|}{\sqrt{|\hat{\lambda}|^2 + 1}\sqrt{|\lambda|^2 + 1}},$$

which naturally includes infinite eigenvalues

$$\chi(\hat{\lambda}, \infty) = \lim_{|\mu| \rightarrow \infty} \chi(\hat{\lambda}, \mu) = \frac{1}{\sqrt{|\hat{\lambda}|^2 + 1}},$$

see [34] for more details. Inserting the perturbation expansions (22) and (23) yields

$$\chi(\hat{\lambda}_k, \lambda) = \frac{|\xi_k \epsilon|^{1/n_1}}{|\lambda|^2 + 1} + o(\epsilon^{1/n_1})$$

and

$$\chi(\hat{\lambda}_k, \infty) = |\xi_k \epsilon|^{1/n_1} + o(\epsilon^{1/n_1}),$$

respectively. This shows that when working in the chordal metric, the  $\alpha$ -part in the Hölder condition number for a finite eigenvalue needs to be divided by  $|\lambda|^2 + 1$  while the Hölder condition number for an infinite eigenvalue remains the same. It is simple to see that this modified condition number has the pleasant property to be continuous at  $|\lambda| = \infty$ .

Whether  $|\hat{\lambda} - \lambda|$  or  $\chi(\hat{\lambda}, \lambda)$  is more appropriate depends on the application. If the ultimate goal of an computation is a finite eigenvalue  $\lambda$ , it can be suspected that  $|\hat{\lambda} - \lambda|$  is practically more relevant. All the following results employ  $|\hat{\lambda} - \lambda|$  but the discussion above reveals that it is rather easy to translate them into the chordal metric setting.

#### 4.1.2 Structured Hölder condition numbers for eigenvalues of matrix pencils

The structured Hölder condition number  $\kappa(A, B, \lambda; \mathbb{S}) = (n_{\mathbb{S}}, \alpha_{\mathbb{S}})$  for some subset  $\mathbb{S} \subset \mathbb{C}^{n \times n} \times \mathbb{C}^{n \times n}$  can be defined in the same way as for the standard eigenvalue problem. In particular, if  $n_{\mathbb{S}} = n_1$  then

$$\alpha_{\mathbb{S}} = \sup_{\substack{\|E\| \leq w_A, \|F\| \leq w_B \\ (E, F) \in \mathbb{S}}} \rho(Y^H(E - \lambda F)X).$$

Some proofs from Section 3 can be rather directly extended to yield results on generalized eigenvalue problems if the structure is separable, i.e.,  $\mathbb{S} = \mathbb{S}_1 \times \mathbb{S}_2$  with  $\mathbb{S}_1, \mathbb{S}_2 \subset \mathbb{C}^{n \times n}$ . The following theorem collects such results.

**Theorem 4.5** *Let  $\kappa(A, B, \lambda) = (n_1, \alpha)$  and  $\kappa(A, B, \lambda; \mathbb{S}_1 \times \mathbb{S}_2) = (n_{\mathbb{S}_1 \times \mathbb{S}_2}, \alpha_{\mathbb{S}_1 \times \mathbb{S}_2})$ .*

- (i). Real matrix pencils: *If  $\mathbb{S}_1 = \mathbb{S}_2 = \mathbb{R}^{n \times n}$ , then  $n_{\mathbb{S}_1 \times \mathbb{S}_2} = n_1$  and  $\alpha/4 \leq \alpha_{\mathbb{S}_1 \times \mathbb{S}_2} \leq \alpha$  hold for  $A, B \in \mathbb{C}^{n \times n}$  in any unitarily invariant norm.*
- (ii). Symmetric matrix pencils: *If  $\mathbb{S}_1 = \mathbb{S}_2 = \{A \in \mathbb{C}^{n \times n} : A^T = A\}$ , then  $n_{\mathbb{S}_1 \times \mathbb{S}_2} = n_1$  and  $\alpha_{\mathbb{S}_1 \times \mathbb{S}_2} = \alpha$  hold for  $A, B \in \mathbb{S}_1$  in any unitarily invariant norm.*

(iii). Real symmetric matrix pencils: If  $\mathbb{S}_1 = \mathbb{S}_2 = \{A \in \mathbb{R}^{n \times n} : A^T = A\}$  then  $n_{\mathbb{S}_1 \times \mathbb{S}_2} = n_1$  and  $\alpha/4 \leq \alpha_{\mathbb{S}_1 \times \mathbb{S}_2} \leq \alpha$  hold for  $A, B \in \mathbb{C}^{n \times n}$  with  $A^T = A, B^T = B$  in any unitarily invariant norm.

(iv). Symmetric/skew-symmetric matrix pencils: If  $\mathbb{S}_1 = \{A \in \mathbb{C}^{n \times n} : A^T = A\}$  and  $\mathbb{S}_2 = \{B \in \mathbb{C}^{n \times n} : B^T = -B\}$  then the following statements hold for  $(A, B) \in \mathbb{S}_1 \times \mathbb{S}_2$  in the matrix 2-norm:

(a) If  $\lambda = \infty, n_1$  is odd, and  $r_1 = 1$ , then  $n_{\mathbb{S}} < n_1$ ;

(b) if  $\lambda = \infty, n_1$  is odd, and  $r_1 > 1$ , then  $n_{\mathbb{S}} = n_1$  and  $\alpha_{\mathbb{S}} = w_B \sqrt{\sigma_1 \sigma_2}$ , where  $\sigma_1, \sigma_2$  are the two largest singular values of  $XX^T$  (whereas  $\alpha = w_B \sigma_1$ );

(c) if  $\lambda = \infty$  and  $n_1$  is even, then  $r_1$  is even,  $n_{\mathbb{S}} = n_1$  and

$$\alpha_{\mathbb{S}} = \alpha = w_B \left\| X \begin{bmatrix} 0 & -I_{r_1/2} \\ I_{r_1/2} & 0 \end{bmatrix} X^T \right\|_2;$$

(d) if  $\lambda = 0$  and  $n_1$  is even, then  $n_{\mathbb{S}} = n_1$  and  $\alpha_{\mathbb{S}} = \alpha = w_A \|XX^T\|_2$ .

(e) if  $\lambda = 0$  and  $n_1$  is odd, then  $r_1$  is even,  $n_{\mathbb{S}} = n_1$  and

$$\alpha_{\mathbb{S}} = \alpha = w_A \left\| X \begin{bmatrix} 0 & -I_{r_1/2} \\ I_{r_1/2} & 0 \end{bmatrix} X^T \right\|_2.$$

(f) if  $\lambda \neq \infty, \lambda \neq 0$  and  $r_1 = 1$ , then  $n_{\mathbb{S}} = n_1$  and  $\alpha_{\mathbb{S}} = w_A \alpha_1 + w_B |\lambda| \alpha_2$ , where  $\alpha_1 = \|X\|_2 \|Y\|_2$  and  $\alpha_2 = \sqrt{\|X\|_2^2 \|Y\|_2^2 - |Y^T X|^2}$ .

(g) if  $\lambda \neq \infty, \lambda \neq 0$  and  $r_1 > 1$ , then  $n_{\mathbb{S}} = n_1$  and  $\alpha_{\mathbb{S}} \geq w_A \|XY^H\|_2$ .

(v). Skew-symmetric matrix pencils: If  $\mathbb{S}_1 = \mathbb{S}_2 = \{A \in \mathbb{C}^{n \times n} : A^T = -A\}$  then  $n_{\mathbb{S}_1 \times \mathbb{S}_2} = n_1, r_1$  is even, and  $\alpha_{\mathbb{S}_1 \times \mathbb{S}_2} = \alpha$  for any  $A, B \in \mathbb{S}_1$  in the matrix 2-norm.

(vi). Hermitian matrix pencils: Let  $\mathbb{S}_j = \{A \in \mathbb{C}^{n \times n} : A^H = \gamma_j A\}$  for  $j \in \{1, 2\}$  and fixed  $\gamma_1, \gamma_2 \in \{1, -1\}$ . Then  $n_{\mathbb{S}_1 \times \mathbb{S}_2} = n_1$  and  $\alpha/\sqrt{2} \leq \alpha_{\mathbb{S}_1 \times \mathbb{S}_2} \leq \alpha$  hold for any  $A, B \in \mathbb{C}^{n \times n}$  in the matrix 2-norm. If, additionally,  $\gamma_1 = \gamma_2$  and  $\lambda \in \mathbb{R}$  then  $\alpha_{\mathbb{S}_1 \times \mathbb{S}_2} = \alpha$ .

*Proof.* If not stated otherwise, it is tacitly assumed that  $\lambda$  is finite (the proofs can be easily modified to cover  $\lambda = \infty$ ).

(i) As in the proof of Lemma 3.2, we can find a real matrix  $\tilde{E}$  with  $\|\tilde{E}\| \leq 1$  such that  $\rho(Y^H \tilde{E} X) \geq \|XY^H\|_2/2$ . We set  $E = w_A \tilde{E}, F = 0$  if  $w_A \geq w_B |\lambda|$ , and  $E = 0, F = w_B \tilde{E}$  otherwise. Then

$$\rho(Y^H (E - \lambda F) X) \geq \frac{w_A + w_B |\lambda|}{2} \rho(Y^H \tilde{E} X) \geq \frac{w_A + w_B |\lambda|}{4} \|XY^H\|_2,$$

which proves (i).

(ii) and (iii) Corollary A.2 (ii) implies  $Y = \bar{X}$  and hence assertions (ii) and (iii) can be shown along the lines of the proof of Theorem 3.7.

(iv) For  $\lambda = \infty$ , the structured canonical form of a symmetric/skew-symmetric pencil imposes the same structure on  $X$  and  $Y$  as for the zero eigenvalue of a skew-symmetric matrix, see Corollary A.4. This implies that  $\alpha/w_B$  coincides with the structured condition number for the zero eigenvalue of  $B$  and hence (a)–(c) follow from Theorem 3.8 (i)–(iii).

For  $\lambda = 0$  and  $n_1$  even, Corollary A.4 (d) yields  $Y = \bar{X}$ , so taking  $E = w_A \bar{X} X^H / \|X X^T\|_2$ ,  $F = 0$  shows (d). If  $\lambda = 0$  and  $n_1$  is odd, then  $r_1$  is even and  $Y = \bar{X} \begin{bmatrix} 0 & I_{r_1/2} \\ -I_{r_1/2} & 0 \end{bmatrix}$ , see Corollary A.4 (c). Let  $u_1, v_1$  be, respectively, left and right singular vectors corresponding to the largest singular value  $\sigma_1$  of the skew-symmetric matrix  $XY^H = X \begin{bmatrix} 0 & -I_{r_1/2} \\ I_{r_1/2} & 0 \end{bmatrix} X^T$ . Then, the pencil  $E - \lambda F$  with  $E = w_A [\bar{u}_1, v_1] \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} [\bar{u}_1, v_1]^T$  and  $F = 0$  is such that  $E \in \mathbb{S}_1, F \in \mathbb{S}_2, \|E\|_2 = w_A$  and

$$\alpha_{\mathbb{S}} \geq \rho \left( EX \begin{bmatrix} 0 & -I_{r_1/2} \\ I_{r_1/2} & 0 \end{bmatrix} X^T \right) = \rho(V^H EU\Sigma) = \rho \left( \begin{bmatrix} w_A \sigma_1 & 0 \\ 0 & * \end{bmatrix} \right) \geq w_A \sigma_1 = \alpha,$$

where  $U\Sigma V^H$  stands for a singular value decomposition of  $X \begin{bmatrix} 0 & -I_{r_1/2} \\ I_{r_1/2} & 0 \end{bmatrix} X^T$ ,  $*$  denotes a nonzero  $(r_1 - 1) \times (r_1 - 1)$  matrix, and we have used that  $v_1^T u_1 = 0$ , since  $X \begin{bmatrix} 0 & -I_{r_1/2} \\ I_{r_1/2} & 0 \end{bmatrix} X^T$  is skew-symmetric. This proves (e).

For finite nonzero  $\lambda$  and  $r_1 = 1$ , we have  $\rho(Y^H(E - \lambda F)X) = |Y^H(E - \lambda F)X|$ , so

$$\alpha_{\mathbb{S}} = \sup_{\substack{\|E\| \leq w_A, \|F\| \leq w_B \\ (E, F) \in \mathbb{S}}} \rho(Y^H(E - \lambda F)X) \leq w_A \sup_{\substack{\|E\| \leq 1 \\ E \in \mathbb{S}_1}} \rho(Y^H E X) + w_B |\lambda| \sup_{\substack{\|F\| \leq 1 \\ F \in \mathbb{S}_2}} \rho(Y^H F X).$$

The supremum over  $E \in \mathbb{S}_1$  is clearly bounded by  $\alpha_1$ , while the supremum over  $F \in \mathbb{S}_2$  is equal to  $\alpha_2$  by Theorem 3.8 (iv). Thus  $\alpha_{\mathbb{S}} \leq w_A \alpha_1 + w_B |\lambda| \alpha_2$ . By [23, Theorem 5.7] there exists a matrix  $\tilde{E} \in \mathbb{S}_1$  with  $\|\tilde{E}\|_2 = \|Y\|_2 / \|X\|_2$  such that  $\tilde{E}X = Y$ . Hence, the symmetric matrix  $E_1 = \frac{\|X\|_2}{\|Y\|_2} \tilde{E}$  has unit 2-norm and attains the upper bound  $\alpha_1$ . Let  $F_2 \in \mathbb{S}_2$  be a matrix with unit 2-norm attaining the maximal value  $\alpha_2$ . Then we may choose  $\gamma_1, \gamma_2 \in \mathbb{C}$ ,  $|\gamma_1| = |\gamma_2| = 1$  in such a way that the pair  $(E, F) = (\gamma_1 w_A E_1, \gamma_2 w_B F_2) \in \mathbb{S}_1 \times \mathbb{S}_2$  satisfies  $\rho(Y^H(E - \lambda F)X) = w_A \alpha_1 + w_B |\lambda| \alpha_2$ . This proves (f).

For finite nonzero  $\lambda$  and  $r_1 > 1$ , recall that, according to the proof of Theorem 3.12, there is a symmetric matrix  $E_1$  with  $\|E_1\|_2 = 1$  and  $\rho(Y^H E_1 X) \geq \|XY^H\|_2$ . Thus, taking  $E = w_A E_1$  and  $F = 0$  leads to (g).

(v) For skew-symmetric/skew-symmetric pencils, Theorem A.5 shows that every eigenvalue has  $r_1$  even. Furthermore, Corollary A.6 (ii) reveals the relationship  $Y = \bar{X} \begin{bmatrix} 0 & I_{r_1/2} \\ -I_{r_1/2} & 0 \end{bmatrix}$ . Hence, if we set  $\tilde{E} = \bar{X} \begin{bmatrix} 0 & I_{r_1/2} \\ -I_{r_1/2} & 0 \end{bmatrix} X^H$ , the perturbation matrices  $E = \frac{w_A}{\|\tilde{E}\|_2} \tilde{E}$ ,  $F = -\frac{w_B}{\|\tilde{E}\|_2} \frac{\bar{\lambda}}{|\lambda|} \tilde{E}$  are such that  $\|E\|_2 = w_A$ ,  $\|F\|_2 = w_B$ , and

$$\rho(Y^H(E - \lambda F)X) = (w_A + |\lambda| w_B) \left\| X^T \begin{bmatrix} 0 & -I_{r_1/2} \\ I_{r_1/2} & 0 \end{bmatrix} X \right\|_2 = \alpha.$$

(vi) As in the proof of Lemma 3.10, we can construct a Hermitian matrix  $\tilde{E}$  such that  $\|\tilde{E}\|_2 = 1$  and  $\rho(\tilde{E}XY^H) = \|XY^H\|_2$ . Let us choose  $\delta \in \{1, -1\}$  such that  $\delta$  matches the

sign of  $\lambda_R$  if  $\gamma_1 = \gamma_2$  and the sign of  $-\lambda_I$  otherwise. Then  $E = w_A\sqrt{\gamma_1}\widetilde{E} \in \mathbb{S}_1$  and  $F = -\delta w_B\sqrt{\gamma_2}\widetilde{E} \in \mathbb{S}_2$  yield

$$\alpha_{\mathbb{S}_1 \times \mathbb{S}_2} \geq \rho((E - \lambda F)XY^H) = |w_A\sqrt{\gamma_1} + \delta w_B\sqrt{\gamma_2}\lambda| \|XY^H\|_2 \geq \frac{w_A + w_B|\lambda|}{\sqrt{2}} \|XY^H\|_2.$$

Note that the last inequality follows from the fact that

$$\begin{aligned} 2|w_A\sqrt{\gamma_1} + \delta w_B\sqrt{\gamma_2}\lambda|^2 - (w_A + w_B|\lambda|)^2 &\geq w_A^2 - 2w_Aw_B|\lambda| + w_B^2|\lambda|^2 \\ &= (w_A - w_B|\lambda|)^2 \geq 0. \end{aligned}$$

If  $\gamma_1 = \gamma_2 = 1$  and  $\lambda \in \mathbb{R}$  then  $|w_A + \delta w_B\lambda| = w_A + w_B|\lambda|$  and the factor  $1/\sqrt{2}$  can be removed.  $\square$

**Remark 4.6** For definite Hermitian matrix pencils, the result of Theorem 4.5 (vi) can be found in [35, 37] for semisimple  $\lambda$  and in [9] for simple  $\lambda$ .

*Palindromic matrix pencils*, which are addressed by the following theorem, provide a practically relevant example for a structure that is not separable, see [12, 21] for more details and applications. With no loss of generality we may assume  $w_A = w_B = 1$  in this case, since  $B = A^T$ .

**Theorem 4.7** Let  $\mathbb{S} = \{(A, A^T) : A \in \mathbb{C}^{n \times n}\}$  denote the set of palindromic matrix pencils, and assume  $w_A = w_B = 1$ . Then the following statements on  $\kappa(A, A^T, \lambda; \mathbb{S}) = (n_{\mathbb{S}}, \alpha_{\mathbb{S}})$  hold for  $A \in \mathbb{C}^{n \times n}$  in the matrix 2-norm.

- (i). If  $\lambda = 1$ ,  $n_1$  is odd, and  $r_1 = 1$ , then  $n_{\mathbb{S}} < n_1$ ;
- (ii). if  $\lambda = 1$ ,  $n_1$  is odd, and  $r_1 > 1$ , then  $n_{\mathbb{S}} = n_1$  and  $\alpha_{\mathbb{S}} = 2\sqrt{\sigma_1\sigma_2}$ , where  $\sigma_1, \sigma_2$  are the two largest singular values of  $XX^T$  (whereas  $\alpha = 2\sigma_1$ );
- (iii). if  $\lambda = 1$  and  $n_1$  is even, then  $r_1$  is even,  $n_{\mathbb{S}} = n_1$  and

$$\alpha_{\mathbb{S}} = \alpha = 2 \left\| X \begin{bmatrix} 0 & -I_{r_1/2} \\ I_{r_1/2} & 0 \end{bmatrix} X^T \right\|_2;$$

- (iv). if  $\lambda = -1$  and  $n_1$  is odd, then  $r_1$  is even,  $n_{\mathbb{S}} = n_1$  and

$$\alpha_{\mathbb{S}} = \alpha = 2 \left\| X \begin{bmatrix} 0 & -I_{r_1/2} \\ I_{r_1/2} & 0 \end{bmatrix} X^T \right\|_2;$$

- (v). if  $\lambda = -1$  and  $n_1$  is even, then  $n_{\mathbb{S}} = n_1$  and  $\alpha_{\mathbb{S}} = \alpha = 2\|XX^T\|_2$ ;
- (vi). if  $\lambda \neq \pm 1$  is finite and  $r_1 = 1$ , then  $n_{\mathbb{S}} = n_1$  and

$$\frac{1}{2}(|1 - \lambda|\alpha_1 + |1 + \lambda|\alpha_2) \leq \alpha_{\mathbb{S}} \leq |1 - \lambda|\alpha_1 + |1 + \lambda|\alpha_2,$$

where  $\alpha_1 = \|X\|_2\|Y\|_2$  and  $\alpha_2 = \sqrt{\|X\|_2^2\|Y\|_2^2 - |Y^T X|^2}$ .

- (vii). if  $\lambda \neq \pm 1$  is finite and  $r_1 > 1$ , then  $n_{\mathbb{S}} = n_1$  and  $\frac{|1-\lambda|}{1+|\lambda|}\alpha \leq \alpha_{\mathbb{S}} \leq \alpha$ ;

(viii). if  $\lambda = \infty$ , then  $n_{\mathbb{S}} = n_1$  and  $\alpha_{\mathbb{S}} = \alpha$ .

*Proof.* If  $\lambda$  is finite and  $n_{\mathbb{S}} = n_1$  then

$$\begin{aligned}\alpha_{\mathbb{S}} &= \sup_{\substack{\|E\|_2 \leq 1 \\ E \in \mathbb{C}^{n \times n}}} \rho(Y^H(E - \lambda E^T)X) \\ &= \frac{1}{2} \sup_{\substack{\|E\|_2 \leq 1 \\ E \in \mathbb{C}^{n \times n}}} \rho\left((1 - \lambda)Y^H(E + E^T)X + (1 + \lambda)Y^H(E - E^T)X\right).\end{aligned}\quad (24)$$

This relation indicates that the analysis of palindromic matrix pencils is closely tied to the analysis of symmetric/skew-symmetric pencils. In fact, it has been shown [14, 30, 33] that the structured canonical form of a palindromic matrix pencil  $(A, A^T)$  can be extracted from the structured canonical form [38] of the symmetric/skew-symmetric pencil  $(A + A^T, A - A^T)$ . In particular, the relation between  $X$  and  $Y$  for an eigenvalue  $\lambda$  of  $(A, A^T)$  coincides with the relation between  $X$  and  $Y$  for the eigenvalue  $-(1 + \lambda)/(1 - \lambda)$  of  $(A + A^T, A - A^T)$ .

Consequently, if  $\lambda = 1$  and  $n_1$  is odd then Corollary A.4 (iia) implies  $Y = \bar{X}$ . If  $\lambda = 1$  and  $n_1$  is even, then  $r_1$  is even and  $Y = \bar{X} \begin{bmatrix} 0 & I_{r_1/2} \\ -I_{r_1/2} & 0 \end{bmatrix}$ . It follows from (24) that

$$\alpha_{\mathbb{S}} = \sup_{\substack{\|E\|_2 \leq 1 \\ E \in \mathbb{C}^{n \times n}}} \rho(Y^H(E - E^T)X) \leq \sup_{\substack{\|E - E^T\|_2 \leq 2 \\ E \in \mathbb{C}^{n \times n}}} \rho(Y^H(E - E^T)X) = 2 \sup_{\substack{\|G\|_2 \leq 1 \\ G \text{ skew-symmetric}}} \rho(Y^HGX).$$

On the other hand,

$$2 \sup_{\substack{\|G\|_2 \leq 1 \\ G \text{ skew-symmetric}}} \rho(Y^HGX) = \sup_{\substack{\|G\|_2 \leq 1 \\ G \text{ skew-symmetric}}} \rho(Y^H(G - G^T)X) \leq \sup_{\substack{\|E\|_2 \leq 1 \\ E \in \mathbb{C}^{n \times n}}} \rho(Y^H(E - E^T)X) \leq \alpha_{\mathbb{S}}.$$

This shows that the structured Hölder condition number for  $\lambda = 1$  of  $(A, A^T)$  essentially coincides with the structured Hölder condition number for the eigenvalue  $\lambda = 0$  of the skew-symmetric matrix  $A - A^T$ . In particular, Theorem 3.8 (i), (ii), and (iii) yield assertions (i), (ii), and (iii) of this theorem.

If  $\lambda = -1$  and  $n_1$  is odd, then Corollary A.4 (iic) implies that  $r_1$  is even and  $Y = \bar{X} \begin{bmatrix} 0 & I_{r_1/2} \\ -I_{r_1/2} & 0 \end{bmatrix}$ . If  $\lambda = -1$  and  $n_1$  is even then  $Y = \bar{X}$ . As above, it follows from (24) that the situation in assertions (iv) and (v) is completely parallel to the one in items (e) and (d), respectively, of Theorem 4.5 (iv). Thus, an analogous choice of symmetric  $E$  proves (iv) and (v).

For finite  $\lambda$  with  $r_1 = 1$  ( $X$  and  $Y$  are vectors), relation (24) implies

$$\alpha_{\mathbb{S}} \leq |1 - \lambda| \sup_{\substack{\|E_1\|_2 \leq 1 \\ E_1 \text{ is symmetric}}} |Y^H E_1 X| + |1 + \lambda| \sup_{\substack{\|E_2\|_2 \leq 1 \\ E_2 \text{ is skew-symmetric}}} |Y^H E_2 X|.$$

As shown in the proof of Theorem 4.5 (iv) (f), the supremum over symmetric  $E_1$  is equal to  $\alpha_1$  and, according to Theorem 3.8 (iv), the one over skew-symmetric  $E_2$  is equal to  $\alpha_2$ . This shows  $\alpha_{\mathbb{S}} \leq |1 - \lambda| \alpha_1 + |1 + \lambda| \alpha_2$ . Now, let  $E_1$  be a symmetric matrix with  $\|E_1\|_2 \leq 1$  and  $|Y^H E_1 X| = \alpha_1$ , and let  $E_2$  be a skew-symmetric matrix with  $\|E_2\|_2 \leq 1$  and  $|Y^H E_2 X| = \alpha_2$ . Then, the matrix  $E = \gamma_1 E_1 + \gamma_2 E_2$  with suitable scalars  $\gamma_1, \gamma_2$  satisfying  $|\gamma_1| = |\gamma_2| = 1$  gives  $|Y^H(E - \lambda E)X| = |1 - \lambda| \alpha_1 + |1 + \lambda| \alpha_2$  with  $\|E\|_2 \leq 2$ , which yields

$$\alpha_{\mathbb{S}} \geq \frac{1}{2} (|1 - \lambda| \alpha_1 + |1 + \lambda| \alpha_2)$$

and concludes the proof of (vi).

For Assertion (vii) we recall that there is a symmetric matrix  $E$  such that  $\|E\|_2 = 1$  and  $\rho(Y^H E X) \geq \|XY^H\|_2$ , see proof of Theorem 3.12. Thus

$$\alpha_{\mathbb{S}} \geq \rho(Y^H(E - \lambda E^T)X) = |1 - \lambda| \rho(Y^H E X) \geq |1 - \lambda| \|XY^H\|_2.$$

This proves the assertion since  $\alpha = (1 + |\lambda|) \|XY^H\|_2$ .

Finally, (viii) is verified by observing that imposing palindromic structure does not change the definition of  $\alpha$  for an infinite eigenvalue.  $\square$

Summarizing the statements of Theorem 4.7, one may conclude that the structured and unstructured (Hölder) condition numbers of a palindromic matrix pencil may differ significantly only for eigenvalues close to 1.

## 4.2 Matrix polynomials

Some seemingly more general variants of the generalized eigenvalue problem, such as polynomial and product eigenvalue problems, can be addressed with the concepts introduced above. We illustrate this point for a matrix polynomial

$$P(\lambda) = \lambda^m A_m + \lambda^{m-1} A_{m-1} + \dots + \lambda A_1 + A_0, \quad A_i \in \mathbb{C}^{n \times n}.$$

Nonzero vectors  $x, y \in \mathbb{C}^{n \times n}$  are called, respectively, right and left eigenvectors belonging to an eigenvalue  $\lambda$  if  $P(\lambda)x = 0$  and  $y^H P(\lambda)x = 0$ , respectively. In the following, we assume that  $P$  is regular, i.e.,  $\det P(\cdot) \not\equiv 0$ . The  $mn \times mn$  matrix pencil

$$A - \lambda B = \begin{bmatrix} -A_{m-1} & -A_{m-2} & \cdots & -A_1 & -A_0 \\ I & 0 & \cdots & 0 & 0 \\ 0 & I & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I & 0 \end{bmatrix} - \lambda \begin{bmatrix} A_m & 0 & \cdots & 0 & 0 \\ 0 & I & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & I \end{bmatrix} \quad (25)$$

is called the *companion form* of  $P$  and represents one of its most common linearizations. It is well known [7] that the eigenvalues of  $(A, B)$  coincide with those of  $P$ .

Because of this one-to-one relation between the eigenvalues, condition numbers for the eigenvalues of  $P$  can be derived from *structured* eigenvalue condition numbers for  $(A, B)$  if the structure only admits perturbations in the blocks  $A_0, \dots, A_m$  of  $A - \lambda B$ . A consequence of this restriction on the perturbations, the resulting eigenvalue condition numbers for the polynomial do not depend on the particular type of linearization chosen. The described approach has the advantage that we can make use of Theorem 4.5 and do not require more general concepts for matrix polynomials.

Following this approach, let us consider the perturbed matrix polynomial  $P + \epsilon \Delta P$  with

$$\Delta P(\lambda) = \lambda^m E_m + \lambda^{m-1} E_{m-1} + \dots + \lambda E_1 + E_0.$$

Equivalently, we can consider the correspondingly perturbed linearization  $(A + \epsilon E, B + \epsilon F)$ , where

$$E = -V \begin{bmatrix} E_{m-1} & E_{m-2} & \cdots & E_1 & E_0 \end{bmatrix}, \quad F = V E_m V^T, \quad (26)$$

and  $V = [I_n, 0, \dots, 0]^T$ . To measure the perturbations, we allow  $n$  nonnegative weights  $w_1, \dots, w_n$ , each corresponding to a coefficient of the matrix polynomial. As in Remark 4.3, to avoid degenerate situations we require  $w_0 > 0$  for  $\lambda = 0$ ,  $w_m > 0$  for  $\lambda = \infty$ , and  $\max\{w_0, \dots, w_m\} > 0$  for any other eigenvalue.

**Definition 4.8** Let  $\lambda$  be a finite eigenvalue of a matrix polynomial  $P$  with the companion form  $(A, B)$  as in (25). Moreover, let  $X$  and  $Y$  be the corresponding eigenvector matrices (21) of  $(A, B)$ . Consider perturbations  $(E, F)$  of the form (26), which preserve the companion form. Then the (absolute) Hölder condition number for  $\lambda$  is given by  $\kappa(P, \lambda) = (n_1, \alpha)$ , where  $n_1$  is the size of the largest Jordan block of  $(A, B)$  associated with  $\lambda$  and

$$\alpha = \sup_{\substack{\|E_i\| \leq w_i \\ E_i \in \mathbb{C}^{n \times n}}} \rho(Y^H(E - \lambda F)X).$$

The (absolute) Hölder condition number for  $\infty$  is given by  $\kappa(P, \infty) = (n_1, \alpha)$ , where  $n_1$  is the nilpotency index of  $(A, B)$  and

$$\alpha = \sup_{\substack{\|E_i\| \leq w_i \\ E_i \in \mathbb{C}^{n \times n}}} \rho(Y^H F X).$$

The results in [11, Lemma 7.2] and [10, Lemma 3.7] show that  $x_1$  and  $y_1$  are right and left eigenvectors belonging to a finite eigenvalue  $\lambda$  of  $P$  if and only if

$$x = \begin{bmatrix} \lambda^{m-1} x_1 \\ \vdots \\ \lambda x_1 \\ x_1 \end{bmatrix}, \quad y = \begin{bmatrix} y_1 \\ (\lambda A_m + A_{m-1})^H y_1 \\ \vdots \\ (\lambda^{m-1} A_m + \lambda^{m-2} A_{m-1} + \cdots + A_1)^H y_1 \end{bmatrix} \quad (27)$$

are right and left eigenvector of  $(A, B)$ , respectively. For  $\lambda = \infty$ , the eigenvectors of  $(A, B)$  are given by  $x = [x_1^H, 0, \dots, 0]^H$  and  $y = [y_1^H, 0, \dots, 0]^H$ . This shows that the matrices  $X$  and  $Y$  defined in (21), containing right and left eigenvectors belonging to a finite (multiple) eigenvalue  $\lambda$  of  $(A, B)$ , take the form

$$X = \begin{bmatrix} \lambda^{m-1} X_1 \\ \vdots \\ \lambda X_1 \\ X_1 \end{bmatrix}, \quad Y = \begin{bmatrix} Y_1 \\ (\lambda A_m + A_{m-1})^H Y_1 \\ \vdots \\ (\lambda^{m-1} A_m + \lambda^{m-2} A_{m-1} + \cdots + A_1)^H Y_1 \end{bmatrix}, \quad (28)$$

where  $X_1$  and  $Y_1$  are matrices of right and left eigenvectors of  $P$ . For an infinite eigenvalue only the first blocks of  $X$  and  $Y$  are nonzero and equal to  $X_1$  and  $Y_1$ , respectively.

The following lemma provides an explicit formula for the Hölder condition number and also shows  $\alpha > 0$  (under the mentioned conditions on the weights), which – strictly speaking – is needed to justify Definition 4.8.

**Lemma 4.9** For a finite eigenvalue  $\lambda$ , we have

$$\kappa(P, \lambda) = (n_1, (w_m |\lambda|^m + w_{m-1} |\lambda|^{m-1} + \cdots + w_0) \|X_1 Y_1^H\|_2)$$

in any unitarily invariant norm  $\|\cdot\|$ , where  $X_1$  and  $Y_1$  are the eigenvector matrices of  $P$  related to the eigenvector matrices  $X$  and  $Y$  of  $(A, B)$  as shown in (28). For an infinite eigenvalue,  $\kappa(P, \lambda) = (n_1, w_m \|X_1 Y_1^H\|_2)$ .

*Proof.* The structure of the matrices  $E$ ,  $F$ ,  $X$  and  $Y$  shown in (26) and (28) implies

$$Y^H(E - \lambda F)X = -Y_1^H(\lambda^m E_m + \lambda^{m-1} E_{m-1} + \cdots + E_0)X_1$$

As in the proof of Lemma 4.4, this shows

$$\rho(Y^H(E - \lambda F)X) \leq (w_m|\lambda|^m + w_{m-1}|\lambda|^{m-1} + \dots + w_0)\|X_1Y_1^H\|_2 \quad (29)$$

Let  $u, v$  with  $\|u\|_2 = \|v\|_2 = 1$  be the left/right singular vectors belonging to the largest singular value of  $X_1Y_1^H$ . Then equality in (29) is attained for the perturbation coefficients

$$E_0 = w_0vu^H, \quad E_1 = w_1 \frac{\bar{\lambda}}{|\lambda|}vu^H, \quad \dots, \quad E_m = w_m \frac{\bar{\lambda}^m}{|\lambda|^m}vu^H,$$

with  $E_1 = \dots = E_m = 0$  for  $\lambda = 0$ . This proves the result for finite  $\lambda$ . For an infinite eigenvalue, the result follows analogously after observing  $Y^HFX = Y_1^HE_mX_1$ .  $\square$

It should be emphasized that  $X_1$  and  $Y_1$  cannot be chosen arbitrarily in Lemma 4.9; the result depends on the normalization of the matrices  $X$  and  $Y$  imposed by (21). To illustrate the effect of this normalization, let  $\lambda$  be a *semisimple* finite eigenvalue of  $P$  and suppose that  $\tilde{X}_1$  and  $\tilde{Y}_1$  contain *arbitrary* bases of right and left eigenvectors belonging to  $\lambda$ . If we let  $\tilde{X}$  and  $\tilde{Y}$  denote the corresponding bases for eigenvectors of  $(A, B)$  then (28) implies

$$\tilde{Y}^HB\tilde{X} = \tilde{Y}_1^HP'(\lambda)\tilde{X}_1.$$

Since  $\lambda$  is semi-simple and finite, the matrix  $\tilde{Y}_1^HP'(\lambda)\tilde{X}_1$  is invertible and

$$X_1 = \tilde{X}_1(\tilde{Y}_1^HP'(\lambda)\tilde{X}_1)^{-1}, \quad Y_1 = \tilde{Y}_1$$

satisfy  $Y_1^HP'(\lambda)X_1 = I$ , which amounts to the condition imposed by (21) for a semi-simple eigenvalue. By Lemma 4.9,

$$\kappa(P, \lambda) = \left(1, (w_m|\lambda|^m + w_{m-1}|\lambda|^{m-1} + \dots + w_0)\left\|\tilde{X}_1(\tilde{Y}_1^HP'(\lambda)\tilde{X}_1)^{-1}\tilde{Y}_1^H\right\|_2\right).$$

For  $r_1 = 1$ , this formula coincides with a result by Tisseur [39, Theorem 5] on the condition number for a simple eigenvalue of a matrix polynomial.

Finally, let us emphasize again that the companion form linearization serves a purely theoretical purpose here. If one admits general, unstructured perturbations to the linearization then the corresponding condition numbers do depend on the linearization, see the discussion in [10, 11]. In particular, [11] shows how to minimize the unstructured condition number for a simple eigenvalue of the linearization. This is useful when applying an unstructured method, as the QZ algorithm, to compute the eigenvalue via the linearization. The extension of these results to multiple eigenvalues would require comparing the result of Lemma 4.9 with the unstructured Hölder condition numbers of a linearization. Also, it could be of interest to study the effect on the Hölder condition numbers if further structure is imposed on the coefficients of the matrix polynomial and this structure is preserved by the linearization [21]. Some results in this direction concerning structured pseudospectra can be found in [8, 41].

## 5 Conclusions

A definition of structured Hölder condition number for multiple eigenvalues, both of matrices and of regular matrix pencils, has been introduced with the purpose of comparing structured and unstructured condition numbers for several classes of structured matrices and pencils.



Moreover, eigenvalues of matrix polynomials can be treated within this framework via linearization through companion form. Like previous Hölder condition numbers in the literature, the structured condition number  $\kappa(A, \lambda; \mathbb{S}) = (n_{\mathbb{S}}, \alpha_{\mathbb{S}})$  has two entries, the first one related to the leading exponent, the second one to the leading coefficient in the asymptotic expansions of perturbed eigenvalues. Although the present paper focuses on the case when the first entry  $n_{\mathbb{S}}$  coincides with the one in the unstructured condition number, some examples are given when this does not happen (see, e.g., (6) and (7)).

According to the results in this paper, the behavior of multiple eigenvalues under structured perturbations does not differ much from the one for simple eigenvalues described in [3, 16, 32], in the sense that the influence of structure on the condition number is usually mild, except in a few, quite specific situations. All these situations seem to be related to a combination of symmetry with skew-symmetry, either for matrices which are skew-symmetric with respect to a symmetric bilinear form (Theorem 3.8, items (ii) and (iv)) or for symmetric/skew-symmetric pencils (Theorem 4.5 (iv)). Palindromic pencils (Theorem 4.7) represent another instance of the interplay between symmetry and skew-symmetry, see equation (24). Understanding why this happens is one of the open questions raised by such results. Also, there are a few cases where all we can say is that  $n_{\mathbb{S}} = n_1$ , with no further information to compare  $\alpha_{\mathbb{S}}$  and  $\alpha$ . Such cases remain as objects of future study.

Another open problem is a more complete picture of what happens in those cases where  $n_{\mathbb{S}} < n_1$ , i.e., whenever structured perturbations induce a behavior qualitatively different from the one induced by unstructured ones.

**Acknowledgments:** The second author wishes to thank the hospitality of both the Institut für Mathematik at TU Berlin and the Department of Computing Science at Umeå University. Part of the research leading to this paper was conducted while she was visiting both institutions. Thanks go also to Volker Mehrmann for many helpful discussions.

## References

- [1] S. Bora and V. Mehrmann. Linear perturbation theory for structured matrix pencils arising in control theory. *SIAM J. Matrix Anal. Appl.*, 28(1):148–169, 2006.
- [2] J. V. Burke, A. S. Lewis, and M. L. Overton. A robust gradient sampling algorithm for nonsmooth, nonconvex optimization. *SIAM J. Optim.*, 15(3):751–779, 2005.
- [3] R. Byers and D. Kressner. On the condition of a complex eigenvalue under real perturbations. *BIT*, 44(2):209–215, 2004.
- [4] F. Chaitin-Chatelin, A. Harrabi, and A. Ilahi. About Hölder condition numbers and the stratification diagram for defective eigenvalues. *Math. Comput. Simul.*, 54(4-5):397–402, 2000.
- [5] F. Chatelin. *Eigenvalues of Matrices*. Wiley, New York, 1993.
- [6] F. de Terán, F. Dopico, and J. Moro. First order spectral perturbation theory of matrix pencils via the Kronecker form, 2006. In preparation.
- [7] I. Gohberg, P. Lancaster, and L. Rodman. *Matrix polynomials*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1982. Computer Science and Applied Mathematics.

- [8] S. Graillat. A note on structured pseudospectra. *J. Comput. Appl. Math.*, 191(1):68–76, 2006.
- [9] D. J. Higham and N. J. Higham. Structured backward error and condition of generalized eigenvalue problems. *SIAM J. Matrix Anal. Appl.*, 20(2):493–512, 1999.
- [10] N. J. Higham, R.-C. Li, and F. Tisseur. Backward error of polynomial eigenproblems solved by linearization. *SIAM J. Matrix Anal. Appl.*, 29(4):1218–1241, 2007.
- [11] N. J. Higham, D. S. Mackey, and F. Tisseur. The conditioning of linearizations of matrix polynomials. *SIAM J. Matrix Anal. Appl.*, 28(4):1005–1028, 2006.
- [12] A. Hilliges, C. Mehl, and V. Mehrmann. On the solution of palindromic eigenvalue problems. In *Proceedings of ECCOMAS, Jyväskylä, Finland, 2004*.
- [13] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1985.
- [14] R. A. Horn and V. V. Sergeichuk. Canonical forms for complex matrix congruence and \*congruence. *Linear Algebra Appl.*, 416(2-3):1010–1032, 2006.
- [15] M. Karow. Structured pseudospectra and the condition of a nonderogatory eigenvalue. Technical report no. 407, DFG Research Center, MATHEON *Mathematics for key technologies* in Berlin, TU Berlin, Str. des 17. Juni 136, D-10623 Berlin, Germany, 2007.
- [16] M. Karow, D. Kressner, and F. Tisseur. Structured eigenvalue condition numbers. *SIAM J. Matrix Anal. Appl.*, 28(4):1052–1068, 2006.
- [17] H. Langer and B. Najman. Remarks on the perturbation of analytic matrix functions. II. *Integral Equations Operator Theory*, 12(3):392–407, 1989.
- [18] H. Langer and B. Najman. Remarks on the perturbation of analytic matrix functions. III. *Integral Equations Operator Theory*, 15(5):796–806, 1992.
- [19] H. Langer and B. Najman. Leading coefficients of the eigenvalues of perturbed analytic matrix functions. *Integral Equations Operator Theory*, 16(4):600–604, 1993.
- [20] V. B. Lidskiĭ. On the theory of perturbations of nonselfadjoint operators. *Ž. Vyčisl. Mat. i Mat. Fiz.*, 6(1):52–60, 1966.
- [21] D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann. Structured polynomial eigenvalue problems: good vibrations from good linearizations. *SIAM J. Matrix Anal. Appl.*, 28(4):1029–1051, 2006.
- [22] D. S. Mackey, N. Mackey, and F. Tisseur.  $\mathbb{G}$ -reflectors: analogues of Householder transformations in scalar product spaces. *Linear Algebra Appl.*, 385:187–213, 2004.
- [23] D. S. Mackey, N. Mackey, and F. Tisseur. Structured mapping problems for matrices associated with scalar products part I: Lie and Jordan algebras. *SIAM J. Matrix Anal. Appl.*, 29(4):1389–1410, 2008.
- [24] R. Mathias. The singular values of the Hadamard product of a positive semidefinite and a skew-symmetric matrix. *Linear and Multilinear Algebra*, 31(1-4):57–70, 1992.

- [25] C. Mehl. On classification of normal matrices in indefinite inner product spaces. *Electron. J. Linear Algebra*, 15:50–83, 2006.
- [26] V. Mehrmann and H. Xu. Perturbation of purely imaginary eigenvalues of Hamiltonian matrices under structured perturbations, 2007. Preprint.
- [27] J. Moro, J. V. Burke, and M. L. Overton. On the Lidskii-Vishik-Lyusternik perturbation theory for eigenvalues of matrices with arbitrary Jordan structure. *SIAM J. Matrix Anal. Appl.*, 18(4):793–817, 1997.
- [28] S. Noschese and L. Pasquini. Eigenvalue condition numbers: zero-structured versus traditional. *J. Comput. Appl. Math.*, 185(1):174–189, 2006.
- [29] M. Peláez and J. Moro. Structured condition numbers of multiple eigenvalues. *Proceedings in Applied Mathematics and Mechanics*, 6(1):67–70, 2006.
- [30] L. Rodman. Bounded and stably bounded palindromic difference equations of first order. *Electron. J. Linear Algebra*, 15:22–49, 2006.
- [31] S. M. Rump. Structured perturbations. I. Normwise distances. *SIAM J. Matrix Anal. Appl.*, 25(1):1–30, 2003.
- [32] S. M. Rump. Eigenvalues, pseudospectrum and structured perturbations. *Linear Algebra Appl.*, 413(2-3):567–593, 2006.
- [33] C. Schröder. A canonical form for palindromic pencils and palindromic factorizations. Technical report, MATHEON, DFG Research Center "Mathematics for key technologies" in Berlin, TU Berlin, 2006.
- [34] G. W. Stewart and J.-G. Sun. *Matrix Perturbation Theory*. Academic Press, New York, 1990.
- [35] J.-G. Sun. A note on local behavior of multiple eigenvalues. *SIAM J. Matrix Anal. Appl.*, 10(4):533–541, 1989.
- [36] J.-G. Sun. Multiple eigenvalue sensitivity analysis. *Linear Algebra Appl.*, 137/138:183–211, 1990.
- [37] J.-G. Sun. Stability and accuracy: Perturbation analysis of algebraic eigenproblems. Technical report UMINF 98-07, Department of Computing Science, University of Umeå, Umeå, Sweden, 1998. Revised 2002.
- [38] R. C. Thompson. Pencils of complex and real symmetric and skew matrices. *Linear Algebra Appl.*, 147:323–371, 1991.
- [39] F. Tisseur. Backward error and condition of polynomial eigenvalue problems. *Linear Algebra Appl.*, 309(1-3):339–361, 2000.
- [40] F. Tisseur. A chart of backward errors for singly and doubly structured eigenvalue problems. *SIAM J. Matrix Anal. Appl.*, 24(3):877–897, 2003.
- [41] F. Tisseur and N. J. Higham. Structured pseudospectra for polynomial eigenvalue problems, with applications. *SIAM J. Matrix Anal. Appl.*, 23(1):187–208, 2001.

- [42] M. I. Višik and L. A. Ljusternik. Solution of some perturbation problems in the case of matrices and self-adjoint or non-selfadjoint differential equations. I. *Russian Math. Surveys*, 15(3):1–73, 1960.
- [43] J. H. Wilkinson. *The Algebraic Eigenvalue Problem*. Clarendon Press, Oxford, 1965.
- [44] D. Xie, X. Hu, and L. Zhang. The solvability conditions for inverse eigenproblem of symmetric and anti-persymmetric matrices and its approximation. *Numer. Linear Algebra Appl.*, 10(3):223–234, 2003.
- [45] H. Xie and H. Dai. On the sensitivity of multiple eigenvalues of nonsymmetric matrix pencils. *Linear Algebra Appl.*, 374:143–158, 2003.

## A Structured canonical forms

This section collects known results on canonical forms for structured matrices and matrix pencils used in this paper. The forms are constructed as direct sums of the following  $m \times m$  matrices:

$$\begin{array}{cc}
 \textit{Jordan block} & \textit{Flip matrix} \\
 J_m(\lambda) = \begin{bmatrix} \lambda & 1 & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda \end{bmatrix}, & F_m = \begin{bmatrix} & & & & 1 \\ & & & \ddots & \\ & & \ddots & & \\ & \ddots & & & \\ 1 & & & & \end{bmatrix}, \\
 \\
 \textit{Signature matrix (m even)} & \textit{Signature matrix (m odd)} \\
 \Sigma_m = \begin{bmatrix} I_{m/2} & \\ & -I_{m/2} \end{bmatrix}, & \Sigma_m = \begin{bmatrix} 0 & I_{(m-1)/2} & \\ 0 & & -I_{(m-1)/2} \\ 0 & 0 & 0 \end{bmatrix}.
 \end{array}$$

Proofs of the following theorems can be found in Thompson’s overview paper [38], and in the more recent and more general work [25].

**Theorem A.1 (Complex symmetric matrix pencils)** *Consider the regular matrix pencil  $A - \lambda B$  with  $A$  and  $B$  complex symmetric matrices. There exists an invertible matrix  $S$  such that*

$$S^T(A - \lambda B)S = (A_1 - \lambda B_1) \oplus \cdots \oplus (A_p - \lambda B_p),$$

where the diagonal blocks take the form

- (i).  $A_j - \lambda B_j = F_{n_j} - \lambda F_{n_j} J_{n_j}(0)$  for an infinite eigenvalue;
- (ii).  $A_j - \lambda B_j = F_{n_j} J_{n_j}(\lambda_j) - \lambda F_{n_j}$  for a finite eigenvalue  $\lambda_j$ .

### Corollary A.2

- (i). Let  $\lambda_j$  be an eigenvalue of the a matrix  $A$  satisfying  $A^T M = M A$  for some real orthogonal symmetric matrix  $M$ . Then the matrices  $X$  and  $Y$  defined in (11)–(12) satisfy the relation  $Y = M \bar{X}$ .

(ii). Let  $\lambda_j$  be a (finite or infinite) eigenvalue of the regular matrix pencil  $A - \lambda B$  with  $A$  and  $B$  complex symmetric. Then the matrices  $X$  and  $Y$  defined in (21) satisfy the relation  $Y = \bar{X}$ .

*Proof.* The first part is proven by applying Theorem A.1 to the pencil  $MA - \lambda M$ . This yields a matrix  $S$  with  $S^{-1} = (F_{n_1} \oplus \cdots \oplus F_{n_p})S^T M$  such that  $S^{-1}AS$  is in Jordan canonical form. The result follows by inspection of (11)–(12). The second part follows directly from combining Theorem A.1 with (21).  $\square$

**Theorem A.3 (Complex skew-symmetric/symmetric matrix pencils)** Consider the regular matrix pencil  $A - \lambda B$  with a complex skew-symmetric  $A$  and a complex symmetric  $B$ . There exists an invertible matrix  $S$  such that

$$S^T(A - \lambda B)S = (A_1 - \lambda B_1) \oplus \cdots \oplus (A_p - \lambda B_p),$$

where the diagonal blocks take the form

- (i).  $A_j - \lambda B_j = F_{n_j} \Sigma_{n_j} - \lambda F_{n_j} J_{n_j}(0)$  for an infinite eigenvalue with  $n_j$  even;
- (ii).  $A_j - \lambda B_j = \begin{bmatrix} 0 & F_{n_j} - \lambda F_{n_j} J_{n_j}(0) \\ -F_{n_j} - \lambda F_{n_j} J_{n_j}(0) & 0 \end{bmatrix}$  for an infinite eigenvalue with  $n_j$  odd;
- (iii).  $A_j - \lambda B_j = F_{n_j} \Sigma_{n_j} - \lambda F_{n_j}$  for a zero eigenvalue with  $n_j$  odd;
- (iv).  $A_j - \lambda B_j = \begin{bmatrix} 0 & F_{n_j} J_{n_j}(\lambda_j) - \lambda F_{n_j} \\ -F_{n_j} J_{n_j}(\lambda_j) - \lambda F_{n_j} & 0 \end{bmatrix}$  for a finite nonzero eigenvalue pair  $\pm \lambda_j$  or a zero eigenvalue  $\lambda_j$  with  $n_j$  even.

#### Corollary A.4

- (i). Let  $\lambda_j$  be an eigenvalue of a matrix  $A$  satisfying  $A^T M = -MA$  for some real orthogonal symmetric matrix  $M$ . Then the following holds for the matrices  $X$  and  $Y$  defined in (11)–(12):
  - (a) if  $\lambda_j = 0$  and  $n_j$  is odd then  $Y = M\bar{X}$ ;
  - (b) if  $\lambda_j = 0$  and  $n_j$  is even then  $Y = M\bar{X} \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$ ;
  - (c) if  $\lambda_j \neq 0$  then the analogously defined matrices  $\tilde{X}$  and  $\tilde{Y}$  for  $-\lambda_j$  satisfy  $\tilde{X} = -Y, \tilde{Y} = X$ .
- (ii). Let  $\lambda_j$  be an eigenvalue of the regular matrix pencil  $A - \lambda B$  with complex symmetric  $A$  and complex skew-symmetric  $B$ . Then the following holds for the matrices  $X$  and  $Y$  defined in (21).
  - (a) if  $\lambda_j = \infty$  and  $n_j$  is odd then  $Y = \bar{X}$ ;
  - (b) if  $\lambda_j = \infty$  and  $n_j$  is even then  $Y = \bar{X} \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$ ;
  - (c) if  $\lambda_j = 0$  and  $n_j$  is odd then  $Y = \bar{X} \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$ ;
  - (d) if  $\lambda_j = 0$  and  $n_j$  is even then  $Y = \bar{X}$ ;
  - (e) if  $\lambda_j \neq 0$  then the analogously defined matrices  $\tilde{X}$  and  $\tilde{Y}$  for  $-\lambda_j$  satisfy  $\tilde{X} = -Y, \tilde{Y} = X$ .

*Proof.* The first part is proven by applying Theorem A.3 to  $A - \lambda M$ . If  $\lambda_j = 0$  and  $n_j$  is odd then Theorem A.3 (iii) yields the relation  $Q = M\bar{P}(F_{n_j} \oplus \cdots \oplus F_{n_j})$  for the matrices  $P$  and  $Q$  defined in (8). An inspection of (11)–(12) verifies (a). If  $\lambda_j = 0$  and  $n_j$  is even then Theorem A.3 (iv) yields

$$Q = M\bar{P} \left( \left[ \begin{array}{cc} 0 & F_{n_j} \\ -F_{n_j} & 0 \end{array} \right] \oplus \cdots \oplus \left[ \begin{array}{cc} 0 & F_{n_j} \\ -F_{n_j} & 0 \end{array} \right] \right).$$

and thus  $Y = M\bar{X} \left( \left[ \begin{array}{cc} 0 & 1 \\ -1 & 0 \end{array} \right] \otimes \cdots \otimes \left[ \begin{array}{cc} 0 & 1 \\ -1 & 0 \end{array} \right] \right)$ . A perfect shuffle of the columns of  $X$  and  $Y$  yields (b). A similar argument leads to (c).

The second part is proven by applying Theorem A.3 to  $B - \lambda A$ .  $\square$

**Theorem A.5 (Complex skew-symmetric matrix pencils)** *Consider the regular matrix pencil  $A - \lambda B$  with  $A$  and  $B$  complex skew-symmetric matrices. There exists an invertible matrix  $S$  such that*

$$S^T(A - \lambda B)S = (A_1 - \lambda B_1) \oplus \cdots \oplus (A_p - \lambda B_p),$$

where the diagonal blocks take the form

$$(i). \quad A_j - \lambda B_j = \left[ \begin{array}{cc} 0 & F_{n_j} - \lambda F_{n_j} J_{n_j}(0) \\ -F_{n_j} + \lambda F_{n_j} J_{n_j}(0) & 0 \end{array} \right] \text{ for an infinite eigenvalue;}$$

$$(ii). \quad A_j - \lambda B_j = \left[ \begin{array}{cc} 0 & F_{n_j} J_{n_j}(\lambda_j) - \lambda F_{n_j} \\ -F_{n_j} J_{n_j}(\lambda_j) + \lambda F_{n_j} & 0 \end{array} \right] \text{ for a finite eigenvalue.}$$

**Corollary A.6**

(i). *Let  $\lambda_j$  be an eigenvalue of a matrix  $A$  satisfying  $A^T M = M A$  for some real orthogonal skew-symmetric matrix  $M$ . Then the matrices  $X$  and  $Y$  defined in (11)–(12) satisfy  $Y = -M\bar{X} \left[ \begin{array}{cc} 0 & I \\ -I & 0 \end{array} \right]$ .*

(ii). *Let  $\lambda_j$  be a (finite or infinite) eigenvalue of the regular matrix pencil  $A - \lambda B$  with  $A$  and  $B$  complex symmetric. Then the matrices  $X$  and  $Y$  defined in (21) satisfy the relation  $Y = -\bar{X} \left[ \begin{array}{cc} 0 & I \\ -I & 0 \end{array} \right]$*

*Proof.* Using Theorem A.5, the proof is analogous to the proof of Corollary A.4 (ib).  $\square$