STRUCTURED CONDITION NUMBERS FOR INVARIANT SUBSPACES

RALPH BYERS* AND DANIEL KRESSNER†

Abstract. Invariant subspaces of structured matrices are sometimes better conditioned with respect to structured perturbations than with respect to general perturbations. Sometimes they are not. This paper proposes an appropriate condition number $c_{\mathbb{S}}$, for invariant subspaces subject to structured perturbations. Several examples compare $c_{\mathbb{S}}$ with the unstructured condition number. The examples include block cyclic, Hamiltonian, and orthogonal matrices. This approach extends naturally to structured generalized eigenvalue problems such as palindromic matrix pencils.

Key words. Structured eigenvalue problem, invariant subspace, perturbation theory, condition number, deflating subspace, block cyclic, Hamiltonian, orthogonal, palindromic.

AMS subject classifications. 65F15, 65F35.

1. Introduction. An *invariant subspace* $\mathcal{X} \subseteq \mathbb{C}^n$ of a matrix $A \in \mathbb{C}^{n\times n}$ is a linear subspace that stays invariant under the action of A, i.e., $Ax \in \mathcal{X}$ for all $x \in \mathcal{X}$. The computation of such an invariant subspace to solve a real-world problem is virtually always affected by some error, e.g., due to the limitations of finite-precision arithmetic. Instead of $\mathcal{X}$, it is usually the case that only a (hopefully nearby) invariant subspace $\hat{\mathcal{X}}$ of a slightly perturbed matrix $A+E$ is computed, where E represents measurement, modeling, discretization, or roundoff errors. It is therefore important to analyze the influence of perturbations in the entries of A on the accuracy of the invariant subspace $\mathcal{X}$. Stewart [33, 35] developed such a perturbation analysis, yielding a measure on the worst-case sensitivity of $\mathcal{X}$. This measure, the *condition number* $c(\mathcal{X})$, is most appropriate if the only information available on E is that its norm is below a certain perturbation threshold ϵ. Often, however, more information is available, i.e., it is known that the perturbation E preserves some structure of A. For example, if A is a real matrix then it is reasonable to assume that E is also a real matrix. Also, for many classes of structured eigenvalue problems, such as Hamiltonian eigenvalue problems, it is more natural to study and analyze perturbations that respect the structure.

In this paper, we analyze the influence of *structured perturbations*: $A + E \in \mathbb{S}$, where $\mathbb{S}$ is a linear matrix subspace or a smooth submanifold of $\mathbb{C}^{n\times n}$ or $\mathbb{R}^{n\times n}$. This will lead to the notion of a *structured condition number* $c_{\mathbb{S}}(\mathcal{X})$ for an invariant subspace $\mathcal{X}$. It occasionally happens that $c_{\mathbb{S}}(\mathcal{X}) \ll c(\mathcal{X})$, in which case the standard condition number $c(\mathcal{X})$ becomes an inappropriate measure on the actual worst-case sensitivity of $\mathcal{X}$. An extreme example is provided by

$$A = \left[\begin{array}{cc|cc} 0 & -1-\alpha & 2 & 0 \\ 1+\alpha & 0 & 0 & 2 \\ \hline 0 & 0 & 0 & 1-\alpha \\ 0 & 0 & -1+\alpha & 0 \end{array}\right], \quad \mathcal{X} = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right\}, \qquad (1.1)$$

*byers@math.ku.edu, Department of Mathematics, University of Kansas, 405 Snow Hall, Lawrence, KS 66045. This author was partially supported by National Science Foundation awards 0098150 and 0112375 and by the Deutsche Forschungsgemeinschaft Research Center Mathematics for Key Technologies.

†kressner@math.hr, Department of Mathematics, Bijenička 30, 10000 Zagreb, Croatia. This author has been supported by a DFG Emmy Noether fellowship and in part by the Swedish Foundation for Strategic Research under the Frame Programme Grant A3 02:128.

where $\alpha \geq 0$ is considered to be tiny. While $c(\mathcal{X}) = \frac{1}{2\alpha}$, we will see that the structured condition number is given by $c_{\mathbb{S}}(\mathcal{X}) = 1/2$ if the set $\mathbb{S}$ of perturbed matrices is restricted to matrices of the form $A + E = \begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} \\ \hat{A}_{21} & \hat{A}_{22} \end{bmatrix}$ with $\hat{A}_{ij} = \begin{bmatrix} \beta_{ij} & \gamma_{ij} \\ -\gamma_{ij} & \beta_{ij} \end{bmatrix}$ for some $\beta_{ij}, \gamma_{ij} \in \mathbb{R}$.

Structured condition numbers for eigenvectors have been studied in [14, 17] and for invariant subspaces in [24, 26, 39], mostly for special cases. The (structured) perturbation analysis of quadratic matrix equations is a closely related area, which is comprehensively treated in [23, 40]. In this paper, we aim to provide a more general framework for studying structured condition numbers for invariant subspaces, which applies to all structures that form smooth manifolds.

The rest of this paper is organized as follows. In Section 2, we briefly summarize known first-order perturbation results for invariant subspace along with associated notions, such as Sylvester operators and canonical angles. Two conceptually different approaches to the structured perturbation analysis of invariant subspaces for linear structures are described in Section 3. One approach is based on a Kronecker product formulation and pattern matrices, much in the spirit of [9, 14, 22, 31, 41]. Although such an approach yields a computable formula for the structured condition number $c_{\mathbb{S}}(\mathcal{X})$, it gives little or no firsthand information on the relationship between $c_{\mathbb{S}}(\mathcal{X})$ and $c(\mathcal{X})$. The other approach, possibly offering more insight into this relationship, is based on the observation that for several relevant structures, the Sylvester operator associated with an invariant subspace admits an orthogonal decomposition into two operators, one of them is confined to the structure. This property also allows one to develop global perturbation results and to deal with invariant subspaces that are stable under structured perturbations but unstable under unstructured perturbations. Both approaches extend to structures that form smooth manifolds, as shown in Section 3.4. Illustrating the results, Section 4 explains how structured condition numbers for product, Hamiltonian, and orthogonal eigenvalue problems can be derived in a considerably simple manner. The results extend to deflating subspaces of generalized eigenvalue problems, see Section 5, and apply to structured matrix pencils including polindromic matrix pencils.

2. Preliminaries. Given a k-dimensional invariant subspace $\mathcal{X}$ of a matrix $A \in \mathbb{C}^{n\times n}$, we need some basis for $\mathcal{X}$ to begin with. Let the columns of the matrix $X \in \mathbb{C}^{n\times k}$ form such a basis. It is convenient to assume that this basis is orthonormal, which implies that $X^H X$ equals the $k \times k$ identity matrix I_k. If the columns of $X_\perp \in \mathbb{C}^{n\times k}$ form an orthonormal basis for $\mathcal{X}^\perp$, then the orthogonal complement of $\mathcal{X}$, then A has *block Schur decomposition*:

$$[X, X_\perp]^H A [X, X_\perp] = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}, \tag{2.1}$$

where $A_{11} \in \mathbb{C}^{k\times k}$ and $A_{22} \in \mathbb{C}^{(n-k)\times(n-k)}$.

An entity closely associated with $\mathcal{X}$ is the so called *Sylvester operator*

$$\mathbf{T} : R \mapsto A_{22}R - RA_{11}. \tag{2.2}$$

This operator is invertible if and only if A_{11} and A_{22} have no eigenvalue in common, i.e., $\lambda(A_{11}) \cap \lambda(A_{22}) = \emptyset$, see [36, Thm. V.1.3]. The *separation* of A_{11} and A_{22}, $\mathrm{sep}(A_{11}, A_{22})$, is defined as the smallest singular value of $\mathbf{T}$:

$$\mathrm{sep}(A_{11}, A_{22}) := \min_{R\neq 0} \frac{\|\mathbf{T}(R)\|_F}{\|R\|_F} = \min_{R\neq 0} \frac{\|A_{22}R - RA_{11}\|_F}{\|R\|_F}. \tag{2.3}$$

If $\mathbf{T}$ is invertible, this definition implies $\mathrm{sep}(A_{11}, A_{22}) = 1/\|\mathbf{T}^{-1}\|$, where $\|\cdot\|$ is the norm on the space of linear operators $\mathbb{R}^{k\times(n-k)} \to \mathbb{R}^{k\times(n-k)}$ induced by the Frobenius norm. Note that neither the invertibility of $\mathbf{T}$ nor the value of $\mathrm{sep}(A_{11}, A_{22})$ depend on the choice of orthonormal bases for $\mathcal{X}$ and $\mathcal{X}^\perp$. This justifies the following definition.

DEFINITION 2.1. *An invariant subspace is called* simple *if the associated Sylvester operator is invertible.*

We are now prepared to state a first-order perturbation expansion for simple invariant subspaces, which can be proved by the implicit function theorem [37, 39, 25].

THEOREM 2.2. *Let A have a block Schur decomposition of the form (2.1) and assume the invariant subspace $\mathcal{X}$ spanned by the columns of X to be simple. Let $A+E \in \mathcal{B}(A)$ be a perturbation of A, where $\mathcal{B}(A) \subset \mathbb{C}^{n\times n}$ is a sufficiently small open neighborhood of A. Then there exists a uniquely defined analytic function $f : \mathcal{B}_A \to \mathbb{C}^{n\times k}$ so that $X = f(A)$ and the columns of $\hat{X} = f(A+E)$ form a (not-necessarily orthonormal) basis of an invariant subspace of $A+E$. Moreover, $X^H(\hat{X} - X) = 0$ and we have the expansion*

$$\hat{X} = X - X_\perp \mathbf{T}^{-1}(X_\perp^H E X) + \mathcal{O}(\|E\|_F^2), \tag{2.4}$$

with the Sylvester operator $\mathbf{T} : R \mapsto A_{22}R - RA_{11}$.

2.1. Canonical angles, a perturbation bound and $c(\mathcal{X})$. In order to obtain perturbation bounds and condition numbers for invariant subspaces we require the notions of angles and distances between two subspaces.

DEFINITION 2.3. *Let the columns of X and Y form orthonormal bases for the k-dimensional subspaces $\mathcal{X}$ and $\mathcal{Y}$, respectively, and let $\sigma_1 \le \sigma_2 \le \cdots \le \sigma_k$ denote the singular values of $X^H Y$. Then the* canonical angles *between $\mathcal{X}$ and $\mathcal{Y}$ are defined as $\theta_i(\mathcal{X}, \mathcal{Y}) := \arccos \sigma_i$ for $i = 1, \ldots, k$. Furthermore, we set $\Theta(\mathcal{X}, \mathcal{Y}) := \mathrm{diag}(\theta_1(\mathcal{X}, \mathcal{Y}), \ldots, \theta_k(\mathcal{X}, \mathcal{Y}))$.*

Canonical angles can be used to measure the distance between two subspaces. In particular, it can be shown that any unitarily invariant norm $\|\cdot\|_\gamma$ on $\mathbb{C}^{k\times k}$ defines a unitarily invariant metric d_γ on the space of k-dimensional subspaces via $d_\gamma(\mathcal{X}, \mathcal{Y}) = \|\sin[\Theta(\mathcal{X}, \mathcal{Y})]\|_\gamma$, see [36, p. 93].

In the case that one of the subspaces is spanned by a non-orthonormal basis, as in Theorem 2.2, the following lemma provides a useful tool for computing canonical angles.

LEMMA 2.4 ([36]). *Let $\mathcal{X}$ be spanned by the columns of $[I_k, 0]^H$, and $\mathcal{Y}$ by the columns of $[I_k, R^H]^H$. If $\sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_k$ denote the singular values of R then $\theta_i(\mathcal{X}, \mathcal{Y}) = \arctan \sigma_i$ for $i = 1, \ldots, k$.*

This yields the following perturbation bound for invariant subspaces.

COROLLARY 2.5. *Under the assumptions of Theorem 2.2,*

$$\|\Theta(\mathcal{X}, \hat{\mathcal{X}})\|_F \le \frac{\|E\|_F}{\mathrm{sep}(A_{11}, A_{22})} + \mathcal{O}(\|E\|_F^2), \tag{2.5}$$

where $\hat{\mathcal{X}} = \mathrm{range}(\hat{X})$.

Proof. Without loss of generality, we may assume $X = [I, 0]^T$. Since $X^T(\hat{X} - X) = 0$ the matrix $\hat{X}$ must have the form $[I, R^H]^H$ for some $R \in \mathbb{C}^{(n-k)\times k}$. Together with the perturbation expansion (2.4) this implies

$$\|R\|_F = \|\hat{X} - X\|_F \le \|E\|_F / \mathrm{sep}(A_{11}, A_{22}).$$

Inequality (2.5) is proved by applying Lemma 2.4 combined with the expansion $\arctan z = z + \mathcal{O}(z^3)$. □

The derived bound (2.5) is approximately tight. To see this, let V be a matrix such that $\|V\|_F = 1$ and $\|\mathbf{T}^{-1}(V)\|_F = 1/\operatorname{sep}(A_{11}, A_{22})$. Plugging $E = \epsilon X_\perp V X^H$ with $\epsilon > 0$ into the perturbation expansion (2.4) yields

$$\|\Theta(\mathcal{X}, \hat{\mathcal{X}})\|_F = \|\hat{X} - X\|_F + \mathcal{O}(\|\hat{X} - X\|_F^3) = \epsilon/\operatorname{sep}(A_{11}, A_{22}) + \mathcal{O}(\epsilon^2).$$

Hence, we obtain the following *condition number for a simple invariant subspace* $\mathcal{X}$:

$$\begin{aligned} c(\mathcal{X}) &:= \lim_{\epsilon \to 0} \sup \left\{ \|\Theta(\mathcal{X}, \hat{\mathcal{X}})\|_F/\epsilon : \ E \in \mathbb{C}^{n\times n}, \|E\|_F \leq \epsilon \right\} \\ &= 1/\operatorname{sep}(A_{11}, A_{22}) = \|\mathbf{T}^{-1}\|, \end{aligned} \tag{2.6}$$

see also [33, 36]. The condition number $c(\mathcal{X})$ extends to invariant subspaces $\mathcal{X}$ which are not simple by the convention $c(\mathcal{X}) = \infty$. Unlike eigenvalues, invariant subspaces with infinite condition number are generally discontinuous with respect to changes in the matrix entries, i.e., they are unstable under unstructured perturbations [36].

2.2. On the computation of sep. To obtain a computable formula for the quantity $\operatorname{sep}(A_{11}, A_{22})$, a convenient (but computationally expensive) approach is to express the Sylvester operator $\mathbf{T}$, see (2.2), in terms of Kronecker products:

$$\operatorname{vec}(\mathbf{T}(R)) = K_{\mathbf{T}} \cdot \operatorname{vec}(R), \tag{2.7}$$

where the $k(n-k) \times k(n-k)$ matrix $K_{\mathbf{T}}$ is given by

$$K_{\mathbf{T}} = I_k \otimes A_{22} - A_{11}^T \otimes I_{n-k}. \tag{2.8}$$

Here, '⊗' denotes the Kronecker product of two matrices and the vec operator stacks the columns of a matrix in their natural order into one long vector [12]. Note that A_{11}^T denotes the complex transpose of A_{11}. Combining (2.3) with (2.7) yields the formula

$$\operatorname{sep}(A_{11}, A_{22}) = \sigma_{\min}(K_{\mathbf{T}}) = \sigma_{\min}(I_k \otimes A_{22} - A_{11}^T \otimes I_{n-k}), \tag{2.9}$$

where $\sigma_{\min}$ denotes the smallest singular value of a matrix.

Computing the separation based on a singular value decomposition of $K_{\mathbf{T}}$ is costly in terms of memory and computational time. A cheaper estimate of sep can be obtained by applying a norm estimator [15] to $K_{\mathbf{T}}^{-1}$. This amounts to the solution of a few linear equations $K_{\mathbf{T}} x = c$ and $K_{\mathbf{T}}^H x = d$ for particular chosen right hand sides c and d or, equivalently, the solution of a few Sylvester equations $A_{22}X - XA_{11} = C$ and $A_{22}^H X - XA_{11}^H = D$. This approach becomes particularly attractive when A_{11} and A_{22} are already in Schur form, see [1, 5, 18, 19].

3. The structured condition number $c_{\mathbb{S}}(\mathcal{X})$. The condition number $c(\mathcal{X})$ for a simple invariant subspace $\mathcal{X}$ of A provides a first-order bound on the sensitivity of $\mathcal{X}$. This bound is strict in the sense that for any sufficiently small $\epsilon > 0$ there exists a perturbation E with $\|E\|_F = \epsilon$ such that $\|\Theta(\mathcal{X}, \hat{\mathcal{X}})\|_F \approx c(\mathcal{X})\epsilon$. If, however, it is known that the set of admissible perturbations is restricted to a subset $\mathbb{S} \subseteq \mathbb{C}^{n\times n}$ then $c(\mathcal{X})$ may severely overestimate the actual worst-case sensitivity of $\mathcal{X}$. To avoid this effect, we introduce an appropriate notion of structured condition numbers in the sense of Rice [32] as follows.

DEFINITION 3.1. *Let $\mathbb{S} \subseteq \mathbb{C}^{n\times n}$ and let $\mathcal{X}$ be an invariant subspace of $A \in \mathbb{S}$. Then the* structured condition number for $\mathcal{X}$ *with respect to* $\mathbb{S}$ *is defined as*

$$c_{\mathbb{S}}(\mathcal{X}) := \lim_{\epsilon\to 0} \sup_{\substack{A+E\in\mathbb{S}\\ \|E\|_F\leq\epsilon}} \inf\{\|\Theta(\mathcal{X},\hat{\mathcal{X}})\|_F/\epsilon : \hat{\mathcal{X}} \text{ is an invariant subspace of } A+E\}.$$

Note that the structured condition number $c_{\mathbb{S}}(\mathcal{X})$ may be finite even when $\mathcal{X}$ is not simple. This reflects the fact that (as in (1.1) with "$\alpha = 0$") an invariant subspace may be unstable with respect to unstructured perturbation ($c(\mathcal{X}) = \infty$) but stable with respect to structured perturbations ($c_{\mathbb{S}}(\mathcal{X}) < \infty$). If $\mathbb{S} = \mathbb{C}^{n\times n}$, then $c_{\mathbb{S}}(\mathcal{X}) = c(\mathcal{X})$.

If $\mathcal{X}$ is simple, then Definition 3.1 simplifies to

$$c_{\mathbb{S}}(\mathcal{X}) = \lim_{\epsilon\to 0} \sup\Big\{\|\Theta(\mathcal{X},\hat{\mathcal{X}})\|_F/\epsilon :\ A+E\in\mathbb{S}, \|E\|_F\leq\epsilon\Big\}, \tag{3.1}$$

where $\hat{\mathcal{X}}$ is defined in the sense of Theorem 2.2.

As the supremum in (3.1) is taken over a set which is potentially smaller than for the unstructured condition number in (2.6), it is clear that $c_{\mathbb{S}}(\mathcal{X}) \leq c(\mathcal{X})$. Much of the following discussion will be concerned with the question by how far can $c_{\mathbb{S}}(\mathcal{X})$ be below $c(\mathcal{X})$. As a first step, we provide a useful connection between the structured condition number and $\mathbf{T}^{-1}$.

LEMMA 3.2. *Let $\mathcal{X}$ be a simple invariant subspace of a matrix A corresponding to a block Schur decomposition of the form (2.1). Then the structured condition number for $\mathcal{X}$ with respect to $\mathbb{S} \subseteq \mathbb{C}^{n\times n}$ satisfies*

$$c_{\mathbb{S}}(\mathcal{X}) = \lim_{\epsilon\to 0} \sup\Big\{\|\mathbf{T}^{-1}(X_\perp^H E X)\|_F/\epsilon :\ A+E\in\mathbb{S}, \|E\|_F\leq\epsilon\Big\}, \tag{3.2}$$

where $\mathbf{T}$ is the Sylvester operator $\mathbf{T} : R \mapsto A_{22}R - RA_{11}$.

Proof. This statement can be concluded from Theorem 2.2 along the line of arguments that led to the expression (2.6) for the standard condition number. □

3.1. A Kronecker product approach. In the following, we consider perturbations that are linearly structured, i.e., E is known to belong to some linear matrix subspace $\mathbb{L}$. In this case, Lemma 3.2 implies

$$c_{A+\mathbb{L}}(\mathcal{X}) = \sup\Big\{\|\mathbf{T}^{-1}(X_\perp^H E X)\|_F :\ E\in\mathbb{L}, \|E\|_F = 1\Big\}, \tag{3.3}$$

provided that $\mathcal{X}$ is simple.

The Kronecker product representation of $\mathbf{T}$ described in Section 2.2 can be used to turn (3.3) into a computable formula for $c_{A+\mathbb{L}}(\mathcal{X})$. Very similar approaches have been used to obtain expressions for structured condition numbers in the context of eigenvalues [14, 22, 31, 41] and matrix functions [9]. Given an m-dimensional linear matrix subspace $\mathbb{L} \subseteq \mathbb{K}^{n\times n}$ with $\mathbb{K} \in \{\mathbb{R},\mathbb{C}\}$, one can always find an $n^2 \times m$ pattern matrix $M_{\mathbb{L}}$ such that for every $E \in \mathbb{L}$ there exists a uniquely defined parameter vector $p \in \mathbb{K}^m$ with

$$\operatorname{vec}(E) = M_{\mathbb{L}}p, \quad \|E\|_F = \|p\|_2.$$

This implies

$$\operatorname{vec}(\mathbf{T}^{-1}(X_\perp^H E X)) = K_{\mathbf{T}}^{-1}(X^T\otimes X_\perp^H)\operatorname{vec}(E) = K_{\mathbf{T}}^{-1}(X^T\otimes X_\perp^H)M_{\mathbb{L}}p, \tag{3.4}$$

where $K_{\mathbf{T}}$ is defined as in (2.8). Consequently, we have the formula

$$c_{A+\mathbb{L}}(\mathcal{X}) = \sup_{\|p\|_2=1} \|K_{\mathbf{T}}^{-1}(X^T \otimes X_\perp^H)M_{\mathbb{L}}p\|_2 = \|K_{\mathbf{T}}^{-1}(X^T \otimes X_\perp^H)M_{\mathbb{L}}\|_2, \tag{3.5}$$

provided that either $\mathbb{K} = \mathbb{C}$ or all of $\mathbb{K}$, A and $\mathcal{X}$ are real.

If $\mathbb{K} = \mathbb{R}$ but A or $\mathcal{X}$ is complex then problems occur because the supremum in (3.5) is taken with respect to real vectors p but $K_{\mathbf{T}}^{-1}(X^T \otimes X_\perp^H)M$ could be a complex matrix. Nevertheless, one has the following bounds to address such cases, see also [6].

LEMMA 3.3. *Let $\mathbb{L} \subseteq \mathbb{R}^{n\times n}$ be a linear matrix space with pattern matrix $M_{\mathbb{L}}$ and let $\mathcal{X}$ be a simple invariant subspace of $A \in \mathbb{C}^{n\times n}$. Then*

$$\|K_{\mathbf{T}}^{-1}(X^T \otimes X_\perp^H)M_{\mathbb{L}}\|_2/\sqrt{2} \le c_{A+\mathbb{L}}(\mathcal{X}) \le \|K_{\mathbf{T}}^{-1}(X^T \otimes X_\perp^H)M_{\mathbb{L}}\|_2.$$

Proof. Let $B = K_{\mathbf{T}}^{-1}(X^T \otimes X_\perp^H)M_{\mathbb{L}}$ and decompose $B = B^{(R)} + \imath B^{(I)}$ with real matrices $B^{(R)}$ and $B^{(I)}$. Then

$$\frac{1}{\sqrt{2}}\left\|\begin{bmatrix} B_R & -B_I \\ B_I & B_R \end{bmatrix}\right\|_2 \le \left\|\begin{bmatrix} B_R \\ B_I \end{bmatrix}\right\|_2 \le \left\|\begin{bmatrix} B_R & -B_I \\ B_I & B_R \end{bmatrix}\right\|_2 = \|B\|_2.$$

Using $\left\|\left[\begin{smallmatrix} B_R \\ B_I \end{smallmatrix}\right]\right\|_2 = c_{A+\mathbb{L}}(\mathcal{X})$, this concludes the proof. □

3.2. An orthogonal decomposition approach. Although (3.5) provides an explicit expression for $c_{A+\mathbb{L}}(\mathcal{X})$, it tells little about the relationship to the unstructured condition number $c(\mathcal{X})$. In this section, we provide an alternative approach by decomposing the associated Sylvester operator $\mathbf{T} : R \mapsto A_{22}R - RA_{11}$ with respect to the structure.

For this purpose, assume the invariant subspace $\mathcal{X}$ to be simple, and let the columns of X and $X_\perp$ form orthonormal bases of $\mathcal{X}$ and $\mathcal{X}^\perp$, respectively. We set

$$\mathcal{N} := \{X_\perp^H E X : \; E \in \mathbb{L}\},$$

which can be considered as the structure induced by $\mathbb{L}$ in the $(2,1)$ block in a block Schur decomposition (2.1). Moreover, let $\mathcal{M}$ denote the preimage of $\mathcal{N}$ under $\mathbf{T}$. As we assume $\mathcal{X}$ to be simple, we can simply write $\mathcal{M} := \mathbf{T}^{-1}(\mathcal{N})$. Lemma 3.2 shows that the structured condition number of $\mathcal{X}$ is given by

$$c_{A+\mathbb{L}}(\mathcal{X}) = \|\mathbf{T}_s^{-1}\|,$$

where $\mathbf{T}_s$ is the restriction of $\mathbf{T}$ to $\mathcal{M} \to \mathcal{N}$, i.e., $\mathbf{T}_s := \mathbf{T}\big|_{\mathcal{M}\to\mathcal{N}}$. The operator $\mathbf{T}_s$ can be considered as the part of $\mathbf{T}$ that acts on the linear spaces induced by the structure.

In all examples considered in this paper, we additionally have the property that the operator $\mathbf{T}^\star : Q \mapsto A_{22}^H Q - QA_{11}^H$ satisfies $\mathbf{T}^\star : \mathcal{N} \to \mathcal{M}$. Note that $\mathbf{T}^\star$ is the Sylvester operator dual to $\mathbf{T}$:

$$\langle \mathbf{T}(R), Q\rangle = \langle R, \mathbf{T}^\star(Q)\rangle$$

with the matrix inner product $\langle X, Y\rangle = \mathrm{trace}(Y^H X)$. This implies $\mathbf{T} : \mathcal{M}^\perp \to \mathcal{N}^\perp$, where $^\perp$ denotes the orthogonal complement w.r.t. the matrix inner product. Hence, $\mathbf{T}$ decomposes orthogonally into $\mathbf{T}_s$ and $\mathbf{T}_u := \mathbf{T}\big|_{\mathcal{M}^\perp\to\mathcal{N}^\perp}$, and we have

$$c(\mathcal{X}) = \max\{\|\mathbf{T}_s^{-1}\|, \|\mathbf{T}_u^{-1}\|\}. \tag{3.6}$$

Hence, comparing $c(\mathcal{X})$ with $c_{A+\mathbb{L}}(\mathcal{X})$ amounts to comparing $\|\mathbf{T}_u^{-1}\|$ with $\|\mathbf{T}_s^{-1}\|$.

REMARK 3.4. *The conditions* $\mathbf{T}:\mathcal{M}\to\mathcal{N}$ *and* $\mathbf{T}^\star:\mathcal{N}\to\mathcal{M}$ *imply* $\mathbf{T}^{-1}|_{\mathcal{N}\to\mathcal{M}} = \mathbf{T}_s^{-1}$ *and* $\mathbf{T}^{-\star}|_{\mathcal{M}\to\mathcal{N}} = \mathbf{T}_s^{-\star}$. *Hence* $\|\mathbf{T}_s^{-1}\| = \sqrt{\|(\mathbf{T}^{-\star}\circ\mathbf{T}^{-1})|_{\mathcal{N}\to\mathcal{N}}\|}$, *and the power method can be applied to* $\mathbf{T}^{-\star}\circ\mathbf{T}^{-1}$ *in order to estimate* $\|\mathbf{T}_s^{-1}\|$.

EXAMPLE 3.5. *Consider the embedding of a complex matrix* $B+\imath C$, *with* $B, C\in\mathbb{R}^{n\times n}$, *into a real* $2n\times 2n$ *matrix of the form* $A = \left[\begin{smallmatrix} B & C\\ -C & B\end{smallmatrix}\right]$. *Let the columns of* $Y+\imath Z$ *and* $Y_\perp+\imath Z_\perp$, *where* $Y,Z\in\mathbb{R}^{n\times k}$ *and* $Y_\perp, Z_\perp\in\mathbb{R}^{n\times(n-k)}$, *form orthonormal bases for an invariant subspace of* $B+\imath C$ *and its orthogonal complement, respectively. Then the columns of* $X = \left[\begin{smallmatrix} Y & Z\\ -Z & Y\end{smallmatrix}\right]$ *and* $X_\perp = \left[\begin{smallmatrix} Y_\perp & Z_\perp\\ -Z_\perp & Y_\perp\end{smallmatrix}\right]$ *form orthonormal bases for an invariant subspace* $\mathcal{X}$ *of* A *and* $\mathcal{X}^\perp$, *respectively. This corresponds to the block Schur decomposition*

$$[X, X_\perp]^T A [X, X_\perp] =: \left[\begin{array}{c|c} A_{11} & A_{12}\\ \hline 0 & A_{22}\end{array}\right] = \left[\begin{array}{cc|cc} B_{11} & C_{11} & B_{12} & C_{12}\\ -C_{11} & B_{11} & -C_{12} & B_{12}\\ \hline 0 & 0 & B_{22} & C_{22}\\ 0 & 0 & -C_{22} & B_{22}\end{array}\right],$$

and the associated Sylvester operator is given by $\mathbf{T}: R\mapsto A_{22}R - RA_{11}$.

If we consider perturbations having the same structure as A *then* $\mathbb{L} = \left\{\left[\begin{smallmatrix} F & G\\ -G & F\end{smallmatrix}\right]\right\}$ *and*

$$\mathcal{N} := X_\perp^T\mathbb{L}X = \left\{\begin{bmatrix} F_{21} & G_{21}\\ -G_{21} & F_{21}\end{bmatrix}\right\},\quad \mathcal{N}^\perp = \left\{\begin{bmatrix} F_{21} & G_{21}\\ G_{21} & -F_{21}\end{bmatrix}\right\}.$$

Moreover, we have $\mathbf{T}:\mathcal{N}\to\mathcal{N}$ *and* $\mathbf{T}^\star:\mathcal{N}\to\mathcal{N}$. *The restricted operator* $\mathbf{T}_s := \mathbf{T}\big|_{\mathcal{N}\to\mathcal{N}}$ *becomes singular only if* $B_{11}+\imath C_{11}$ *and* $B_{22}+\imath C_{22}$ *have eigenvalues in common, while* $\mathbf{T}_u := \mathbf{T}\big|_{\mathcal{N}^\perp\to\mathcal{N}^\perp}$ *becomes singular if* $B_{11}+\imath C_{11}$ *and* $B_{22}-\imath C_{22}$ *have eigenvalues in common. Thus, there are situations in which the unstructured condition number* $c(\mathcal{X}) = \max\{\|\mathbf{T}_s^{-1}\|, \|\mathbf{T}_u^{-1}\|\}$ *can be significantly larger than the structured condition number* $c_{\mathbb{S}}(\mathcal{X}) = c_{A+\mathbb{L}}(\mathcal{X}) = \|\mathbf{T}_s^{-1}\|$, *e.g., if* $\imath\gamma$ *is nearly an eigenvalue of* $B_{11}+\imath C_{11}$ *while* $-\imath\gamma$ *is nearly an eigenvalue of* $B_{22}+\imath C_{22}$ *for some* $\gamma\in\mathbb{R}$.

The introductionary example (1.1) is a special case of Example 3.5, where the unstructured condition number tends to infinity as the parameter α tends to zero. The results above imply that the structured condition number is given by

$$c_{\mathbb{S}}(\mathcal{X}) = \inf_{|\beta|^2+|\gamma|^2=1}\left\{\left\|\begin{bmatrix} 0 & 1-\alpha\\ -1+\alpha & 0\end{bmatrix}\begin{bmatrix}\beta & \gamma\\ -\gamma & \beta\end{bmatrix} - \right.\right.$$
$$\left.\left.\begin{bmatrix}\beta & \gamma\\ -\gamma & \beta\end{bmatrix}\begin{bmatrix} 0 & -1-\alpha\\ 1+\alpha & 0\end{bmatrix}\right\|_F\right\}^{-1} = \frac{1}{2}.$$

There is evidence to believe that $c_{\mathbb{S}}(\mathcal{X}) = 1/2$ holds even if $\alpha = 0$. However, all our arguments so far rest on the perturbation expansion in Theorem 2.2, which requires the invariant subspace to be simple; a condition that is not satisfied if $\alpha = 0$. This restriction will be removed in the following section by adapting the global perturbation analysis for invariant subspaces proposed by Stewart [35] and refined by Demmel [10], see also [8].

3.3. Global perturbation bounds. Additionally to the block Schur decomposition (2.1) we now consider the *perturbed* block Schur decomposition

$$[X, X_\perp]^H (A+E)[X, X_\perp] = \begin{bmatrix} A_{11}+E_{11} & A_{12}+E_{12} \\ E_{21} & A_{22}+E_{22} \end{bmatrix} =: \begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} \\ E_{21} & \hat{A}_{22} \end{bmatrix}. \tag{3.7}$$

In order to obtain a formula for $\hat{X}$, a basis for the perturbed invariant subspace $\hat{\mathcal{X}}$ close to $\mathcal{X} = \operatorname{span}(X)$, we look for an invertible matrix of the form $W = \left[\begin{smallmatrix} I & 0 \\ -R & I \end{smallmatrix}\right]$ so that

$$W^{-1} \begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} \\ E_{21} & \hat{A}_{22} \end{bmatrix} W = \begin{bmatrix} \hat{A}_{11} - \hat{A}_{12}R & \hat{A}_{12} \\ E_{21} + R\hat{A}_{11} - \hat{A}_{22}R - R\hat{A}_{12}R & \hat{A}_{22} + R\hat{A}_{12} \end{bmatrix}$$

is in block upper triangular form. This implies that R is a solution of the algebraic Riccati equation

$$\hat{A}_{22}R - R\hat{A}_{11} + R\hat{A}_{12}R = E_{21}. \tag{3.8}$$

To solve this quadratic matrix equation and for deriving the structured condition number with respect to a linear matrix space $\mathbb{L}$ we need to require the following two conditions on $\mathbb{L}$.

A1: Let $\mathcal{N} = \{X_\perp^H F X : F \in \mathbb{L}\}$ and $\hat{\mathbf{T}} : R \mapsto \hat{A}_{22}R - R\hat{A}_{11}$. Then there exists a linear matrix space $\mathcal{M}$, having the same dimension as $\mathcal{N}$, such that $\hat{\mathbf{T}} : \mathcal{M} \to \mathcal{N}$ and $R\hat{A}_{12}R \in \mathcal{N}$ for all $R \in \mathcal{M}$.

A2: The restricted operator $\hat{\mathbf{T}}_s := \hat{\mathbf{T}}\big|_{\mathcal{M}\to\mathcal{N}}$ is invertible.

THEOREM 3.6. *Assume that A1 and A2 hold. If* $(4\|\hat{\mathbf{T}}_s^{-1}\|^2 \, \|\hat{A}_{12}\|_F \, \|E_{21}\|_F) < 1$ *then there exists a solution* $R \in \mathcal{M}$ *of the quadratic matrix equation (3.8) with*

$$\begin{aligned} \|R\|_F &\le \frac{2\|\hat{\mathbf{T}}_s^{-1}\| \, \|E_{21}\|_F}{1 + \sqrt{1 - 4\|\hat{\mathbf{T}}_s^{-1}\|^2 \, \|\hat{A}_{12}\|_F \, \|E_{21}\|_F}} \\ &< 2\|\hat{\mathbf{T}}_s^{-1}\| \, \|E_{21}\|_F. \end{aligned} \tag{3.9}$$

Proof. The result can be proved by constructing an iteration

$$R_0 \leftarrow 0, \quad R_{i+1} \leftarrow \hat{\mathbf{T}}_s^{-1}(E_{21} - R_i\hat{A}_{12}R_i),$$

which is well-defined because $R_i \in \mathcal{M}$ implies $R_i\hat{A}_{12}R_i \in \mathcal{N}$. This approach is very similar to the technique used by Stewart, see [33, 35] or [36, Thm. V.2.11]. In fact, it can be shown in precisely the same way as in [36, Thm. V.2.11] that all iterates R_i satisfy a bound of the form (3.9) and converge to a solution of (3.8). □

Having obtained a solution R of (3.8), a basis for an invariant subspace $\hat{\mathcal{X}}$ of $A+E$ is given by $\hat{X} = X - X_\perp R$. Together with Lemma 2.4, this leads to the following global version of Corollary 2.5.

COROLLARY 3.7. *Under the assumptions of Theorem 3.6 there exists an invariant subspace* $\hat{\mathcal{X}}$ *of* $A+E$ *so that*

$$\|\tan\Theta(\mathcal{X}, \hat{\mathcal{X}})\|_F \le \frac{2\|\hat{\mathbf{T}}_s^{-1}\| \, \|E_{21}\|_F}{1 + \sqrt{1 - 4\|\hat{\mathbf{T}}_s^{-1}\|^2 \, \|\hat{A}_{12}\|_F \, \|E_{21}\|_F}}. \tag{3.10}$$

The quantity $\|\hat{\mathbf{T}}_s^{-1}\|$ in the bound (3.10) can be related to $\|\mathbf{T}_s^{-1}\|$, the norm of the inverse of the unperturbed Sylvester operator, using the following lemma.

LEMMA 3.8. *Assume that A1 holds, and that the Sylvester operator* $\mathbf{T} : R \mapsto A_{22}R - RA_{11}$ *associated with the unperturbed block Schur decomposition (2.1) also satisfies* $\mathbf{T} : \mathcal{M} \to \mathcal{N}$. *If* $\mathbf{T}_s := \mathbf{T}\big|_{\mathcal{M}\to\mathcal{N}}$ *is invertible and* $1/\|\mathbf{T}_s^{-1}\| > \|E_{11}\|_F + \|E_{22}\|_F$, *then* $\hat{\mathbf{T}}_s$ *is also invertible and satisfies*

$$\|\hat{\mathbf{T}}_s^{-1}\| \leq \frac{\|\mathbf{T}_s^{-1}\|}{1 - \|\mathbf{T}_s^{-1}\|(\|E_{11}\|_F + \|E_{22}\|_F)}. \tag{3.11}$$

Proof. Under the given assumptions we have

$$\|I - \mathbf{T}_s^{-1} \circ \hat{\mathbf{T}}_s\| = \sup_{\substack{R\in\mathcal{M} \\ \|R\|_F=1}} \|\mathbf{T}_s^{-1}(E_{22}R - RE_{11})\|_F \leq \|\mathbf{T}_s^{-1}\|(\|E_{11}\|_F + \|E_{22}\|_F) < 1.$$

Thus, the Neumann series

$$\sum_{i=0}^{\infty} (I - \mathbf{T}_s^{-1} \circ \hat{\mathbf{T}}_s)^i \circ \mathbf{T}_s^{-1}$$

converges to $\hat{\mathbf{T}}_s^{-1}$, which proves (3.11). □

Combining Corollary 3.7 with the expansion $\arctan z = z + \mathcal{O}(z^3)$ and Lemma 3.8 yields

$$\|\Theta(\mathcal{X}, \hat{\mathcal{X}})\|_F \leq \|\mathbf{T}_s^{-1}\| \, \|E\|_F + \mathcal{O}(\|E\|_F^2). \tag{3.12}$$

This implies that $c_{A+\mathbb{L}}(\mathcal{X})$, the structured condition number for $\mathcal{X}$, is bounded from above by $\|\mathbf{T}_s^{-1}\|$, even if the operator $\mathbf{T}$ itself is not invertible. To show that the structured condition number and $\|\mathbf{T}_s^{-1}\|$ are actually equal, we require the extra assumption that $\mathbf{T}^\star : \mathcal{N} \to \mathcal{M}$.

THEOREM 3.9. *Assume that A1 holds with the same matrix space* $\mathcal{M}$ *for all* $\hat{\mathbf{T}}$ *corresponding to a perturbation* $E \in \mathbb{L}$. *Moreover, assume that the Sylvester operator* $\mathbf{T} : R \mapsto A_{22}R - RA_{11}$ *additionally satisfies* $\mathbf{T}^\star : \mathcal{N} \to \mathcal{M}$ *and that* $\mathbf{T}_s := \mathbf{T}\big|_{\mathcal{M}\to\mathcal{N}}$ *is invertible. Then* $c_{A+\mathbb{L}}(\mathcal{X}) = \|\mathbf{T}_s^{-1}\|$.

Proof. By Lemma 3.8, it follows that $\hat{\mathbf{T}}_s$ is invertible for all sufficiently small perturbations E. Thus, the discussion provided above proves $c_{A+\mathbb{L}}(\mathcal{X}) \leq \|\mathbf{T}_s^{-1}\|$. It remains to construct perturbations $E \in \mathbb{L}$ so that

$$\lim_{\|E\|_F\to 0} \|\Theta(\mathcal{X}, \hat{\mathcal{X}})\|_F / \|E\|_F \geq \|\mathbf{T}_s^{-1}\|,$$

where $\hat{\mathcal{X}}$ denotes an invariant subspace of $A + E$ nearest to $\mathcal{X}$. For this purpose, we choose $E_{21} \in \mathcal{N}$ such that $\|E_{21}\|_F = 1$, $\|\mathbf{T}^{-1}(E_{21})\|_F = \|\mathbf{T}_s^{-1}\|$, and consider the perturbation $E = \epsilon X_\perp E_{21} X^H$. Because of (3.12) we may assume that the nearest invariant subspace $\hat{\mathcal{X}}$ of $A+E$ satisfies $\|\Theta(\mathcal{X}, \hat{\mathcal{X}})\|_2 < \pi/2$ for sufficiently small $\epsilon > 0$. In other words, none of the vectors in $\hat{\mathcal{X}}$ is orthogonal to $\mathcal{X}$. This implies the existence of a matrix R such that the columns of $\hat{X} = X - X_\perp R$ form a basis for $\hat{\mathcal{X}}$. Equivalently, R satisfies the matrix equation

$$\mathbf{T}(R) + RA_{12}R = \epsilon E_{21}.$$

If we decompose $R = R_s + R_u$, where $R_s \in \mathcal{M}$ and $R_u \in \mathcal{M}^\perp$, then $\mathbf{T}(R_s) \in \mathcal{N}$ while $\mathbf{T}^\star : \mathcal{N} \to \mathcal{M}$ implies $\mathbf{T}(R_u) \in \mathcal{N}^\perp$. Similarly, $RA_{12}R = Q_s + Q_u$ with $Q_s \in \mathcal{N}$ and $Q_u \in \mathcal{N}^\perp$. Consequently, $\mathbf{T}(R_s) + Q_s = \epsilon E_{21}$ and since $\|Q_s\|_F = \mathcal{O}(\epsilon^2)$ it follows that

$$\lim_{\epsilon \to 0} \|R\|_F/\epsilon \geq \lim_{\epsilon \to 0} \|R_s\|_F/\epsilon = \|\mathbf{T}^{-1}(E_{21})\|_F = \|\mathbf{T}_s^{-1}\|.$$

Combining this inequality with $\|\Theta(\mathcal{X}, \hat{\mathcal{X}})\|_F = \|R\|_F + \mathcal{O}(\epsilon^2)$ yields the desired result. ◻

Let us briefly summarize the discussion on structured condition numbers. If $\mathcal{X}$ is simple then $c_{A+\mathbb{L}}(\mathcal{X})$ is given by $\|\mathbf{T}_s^{-1}\|$. This equality also holds for the case that $\mathcal{X}$ is not simple but stable under structured perturbations, provided that the assumptions of Theorem 3.9 are satisfied. It is easy to see that all these extra assumptions are satisfied by the introductionary example (1.1), showing that $c_{A+\mathbb{L}}(\mathcal{X}) = 1/2$ also holds for $\alpha = 0$.

3.4. Extension to nonlinear structures. So far, we have mainly considered structures $\mathbb{S}$ that form (affine) linear matrix spaces. Nevertheless, the results from the previous subsections can be used to address a smooth manifold $\mathbb{S}$ by observing that the structured condition number with respect to $\mathbb{S}$ equals the one with respect to the tangent space of $\mathbb{S}$ at A. This is a consequence of the following theorem, which is much in the spirit of the corresponding result in [22, Thm. 2.1] for structured eigenvalue condition numbers.

Theorem 3.10. *Let $\mathbb{S}$ be a smooth real or complex manifold and let $\mathcal{X}$ be a simple invariant subspace of $A \in \mathbb{S}$ corresponding to a block Schur decomposition of the form (2.1). Then the structured condition number for $\mathcal{X}$ with respect to $\mathbb{S}$ satisfies*

$$c_{\mathbb{S}}(\mathcal{X}) = \sup\Big\{\|\mathbf{T}^{-1}(X_\perp^H E X)\|_F : \ E \in T_A\mathbb{S}, \|E\|_F = 1\Big\}, \tag{3.13}$$

where $\mathbf{T}$ is the Sylvester operator $\mathbf{T} : R \mapsto A_{22}R - RA_{11}$ and $T_A\mathbb{S}$ is the tangent space of $\mathbb{S}$ at A.

Proof. Let $E \in T_A\mathbb{S}$ with $\|E\|_F = 1$. Then there is a sufficiently smooth curve $G_E : (-\epsilon, \epsilon) \to \mathbb{K}^{n\times n}$ ($\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$) satisfying $G_E(0) = 0$, $G_E'(0) = E$ and $A + G_E(t) \in \mathbb{S}$ for all t. We have $G_E(t) = Et + \mathcal{O}(|t|^2)$ and, by Lemma 3.2,

$$\begin{aligned} c_{A+G_E(\cdot)}(\mathcal{X}) &= \lim_{\epsilon\to 0} \sup\Big\{\|\mathbf{T}^{-1}(X_\perp^H G_E(t)X)\|_F/\epsilon : \ |t| \leq \epsilon\Big\} \\ &= \lim_{\epsilon\to 0} \sup\Big\{\|\mathbf{T}^{-1}(X_\perp^H EtX)\|_F/\epsilon : \ |t| \leq \epsilon\Big\} \\ &= \|\mathbf{T}^{-1}(X_\perp^H EX)\|_F. \end{aligned}$$

The curves $A + G_E(\cdot)$ form a covering of an open neighborhood of $A \in \mathbb{S}$, implying

$$c_{\mathbb{S}}(\mathcal{X}) = \sup\Big\{c_{A+G_E(\cdot)}(\mathcal{X}) : \ E \in T_A\mathbb{S}, \|E\|_F = 1\Big\},$$

which proves (3.13). ◻

Theorem 3.10 admits the derivation of an explicit expression for $c_{\mathbb{S}}(\mathcal{X})$, e.g., by applying the Kronecker product approach from Section 3.1 to $T_A\mathbb{S}$. This requires the computation of a pattern matrix for $T_A\mathbb{S}$; an issue which has been discussed for automorphism groups in [22].

4. Examples. In this section, we illustrate the applicability of the theory developed in the preceding section for product, Hamiltonian and orthogonal eigenvalue problems.

4.1. Block cyclic matrices. Let us consider a matrix product

$$\Pi = A^{(p)} A^{(p-1)} \cdots A^{(1)},$$

where $A^{(1)}, \ldots, A^{(p)} \in \mathbb{C}^{n \times n}$. Computing invariant subspaces of matrix products has applications in several areas, such as model reduction, periodic discrete-time systems and bifurcation analysis, see [42] for a recent survey. In many of these applications, it is reasonable to consider factor-wise perturbations, i.e., the perturbed product $\Pi = (A^{(p)} + E^{(p)})(A^{(p-1)} + E^{(p-1)}) \cdots (A^{(1)} + E^{(1)})$. What seems to be a multilinearly structured eigenvalue problem can be turned into a linearly structured eigenvalue problem associated with the block cyclic matrix

$$A = \begin{bmatrix} 0 & & & A^{(p)} \\ A^{(1)} & \ddots & & \\ & \ddots & \ddots & \\ & & A^{(p-1)} & 0 \end{bmatrix}.$$

To see this, let the columns of the block diagonal matrix $X = X^{(1)} \oplus X^{(2)} \oplus \cdots \oplus X^{(p)}$ with $X^{(1)}, \ldots, X^{(p)} \in \mathbb{C}^{n \times k}$ form a basis for an invariant subspace $\mathcal{X}$ of A. By direct computation, it can be seen that the columns of $X^{(1)}$ form a basis for an invariant subspace of Π. Vice versa, the periodic Schur decomposition [4, 13] shows that any basis $X^{(1)}$ for an invariant subspace of Π can be extended to a basis $X^{(1)} \oplus X^{(2)} \oplus \cdots \oplus X^{(p)}$ for an invariant subspace $\mathcal{X}$ of A.

To perform a structured perturbation analysis for an invariant subspace $\mathcal{X}$ admitting an orthonormal basis $X = X^{(1)} \oplus X^{(2)} \oplus \cdots \oplus X^{(p)}$, we first note that there is an orthonormal basis $X_\perp$ of $\mathcal{X}^\perp$ having the form $X_\perp = X_\perp^{(1)} \oplus X_\perp^{(2)} \oplus \cdots \oplus X_\perp^{(p)}$. This leads to the block Schur decomposition

$$[X, X_\perp]^T A [X, X_\perp] = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix},$$

where $A_{11} \in \operatorname{cyc}(k, k, p)$, $A_{12} \in \operatorname{cyc}(k, n-k, p)$, $A_{22} \in \operatorname{cyc}(n-k, n-k, p)$, and $\operatorname{cyc}(n_1, n_2, p)$ denotes the set of $p \times p$ block cyclic matrices with $n_1 \times n_2$ blocks. The corresponding Sylvester operator is given by $\mathbf{T} : R \mapsto A_{22} R - R A_{11}$.

Factor-wise perturbations in Π correspond to block cyclic perturbations in A, i.e., $\mathbb{S} = \operatorname{cyc}(n, n, p)$. The set $\mathcal{N} = X_\perp^T \mathbb{S} X$ coincides with $\operatorname{cyc}(n-k, k, p)$ and we have $\mathbf{T} : \mathcal{M} \to \mathcal{N}$, where $\mathcal{M}$ equals $\operatorname{diag}(n-k, k, p)$, the set of $p \times p$ block diagonal matrices with $(n-k) \times k$ blocks. Moreover, it can be directly verified that $\mathbf{T}^\star : \mathcal{N} \to \mathcal{M}$. Letting $\mathbf{T}_s = \mathbf{T}\big|_{\mathcal{M} \to \mathcal{N}}$ and $\mathbf{T}_u = \mathbf{T}\big|_{\mathcal{M}^\perp \to \mathcal{N}^\perp}$, we thus have $c_{\mathbb{S}}(\mathcal{X}) = \|\mathbf{T}_s^{-1}\|$ and $c(\mathcal{X}) = \max\{\|\mathbf{T}_s^{-1}\|, \|\mathbf{T}_u^{-1}\|\}$. Note that $\mathcal{M}^\perp$, $\mathcal{N}^\perp$ coincide with the set of all $p \times p$ block matrices with $(n-k) \times k$ blocks that are zero in their block diagonal or block cyclic part, respectively.

Although $\mathbf{T}_s$ is invertible if and only if $\mathbf{T}$ is invertible [25], the following example reveals that there may be significant difference between $\|\mathbf{T}_s^{-1}\|$ and $\|\mathbf{T}^{-1}\|$ (and consequently between the structured and unstructured condition numbers for $\mathcal{X}$).

Example 4.1 ([25]). *Let* $p = 2$, $A_{11} = \left[\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right]$ *and* $A_{22} = \left[\begin{smallmatrix} 0 & C \\ D & 0 \end{smallmatrix}\right]$, *where*

$$C = \begin{bmatrix} 10^5 & 10^5 \\ 0 & 10^{-5} \end{bmatrix}, \quad D = \begin{bmatrix} 10^{-5} & 0 \\ 0 & 10^5 \end{bmatrix}.$$

Then the structured condition number is given by

$$c_{\mathbb{S}}(\mathcal{X}) = \left\| \begin{bmatrix} C & -I_2 \\ 0 & D \end{bmatrix}^{-1} \right\|_2 = \sqrt{2} \times 10^5,$$

while the unstructured condition number is much higher,

$$c(\mathcal{X}) = \max \left\{ c_{\mathbb{S}}(\mathcal{X}), \left\| \begin{bmatrix} D & -I_2 \\ 0 & C \end{bmatrix}^{-1} \right\|_2 \right\} = 10^{10}.$$

Other and more detailed approaches to the perturbation analysis for invariant subspaces of (generalized) matrix products, yielding similar results, can be found in [3, 27].

4.2. Hamiltonian matrices. A Hamiltonian matrix is a $2n \times 2n$ matrix A of the form

$$A = \begin{bmatrix} -B & G \\ Q & B^T \end{bmatrix}, \quad G = G^T, \quad Q = Q^T,$$

where $B, G, Q \in \mathbb{R}^{n \times n}$. Hamiltonian matrices arise from, e.g., linear-quadratic optimal control problems and certain quadratic eigenvalue problems, see [2, 29] and the references therein. A particular property of A is that its eigenvalues are symmetric with respect to the imaginary axis. Hence, if A has no purely imaginary eigenvalues, there are n eigenvalues having negative real part. The invariant subspace $\mathcal{X}$ belonging to these n eigenvalues is called the stable invariant subspace. For all $x \in \mathcal{X}$ we have $Jx \perp \mathcal{X}$ with $J = \left[\begin{smallmatrix} 0 & I_n \\ -I_n & 0 \end{smallmatrix}\right]$, a property which makes $\mathcal{X}$ an isotropic vector space [30]. If the columns of $X \in \mathbb{R}^{2n \times n}$ form an orthonormal basis for $\mathcal{X}$, the isotropy of $\mathcal{X}$ implies that $[X, JX]$ is an orthogonal matrix and we have the structured block Schur decomposition

$$[X, X_\perp]^T A [X, X_\perp] = \begin{bmatrix} -\tilde{B} & \tilde{G} \\ 0 & \tilde{B}^T \end{bmatrix}, \quad \tilde{G} = \tilde{G}^T.$$

The corresponding Sylvester operator is given by $\mathbf{T} : R \mapsto \tilde{B}^T R + R\tilde{B}$.

If we restrict the set $\mathbb{S}$ of admissible perturbations to be Hamiltonian then $\mathcal{N} = X_\perp^T \mathbb{S} X$ equals $\mathrm{symm}(n)$, the set of $n \times n$ symmetric matrices, while $\mathcal{N}^\perp = \mathrm{skew}(n)$, the set of $n \times n$ skew-symmetric matrices. It can be directly seen that $\mathbf{T} : \mathcal{N} \to \mathcal{N}$ and, moreover, $\mathbf{T}^\star = \mathbf{T}$. Thus, by letting $\mathbf{T}_s = \mathbf{T}\big|_{\mathcal{N} \to \mathcal{N}}$ and $\mathbf{T}_u = \mathbf{T}\big|_{\mathcal{N}^\perp \to \mathcal{N}^\perp}$, we have $c_{\mathbb{S}}(\mathcal{X}) = \|\mathbf{T}_s^{-1}\|$ and $c(\mathcal{X}) = \max\{\|\mathbf{T}_s^{-1}\|, \|\mathbf{T}_u^{-1}\|\}$. It is known that the expression $\|\tilde{B}^T R + R\tilde{B}\|_F / \|R\|_F$, $R \neq 0$, is always minimized by a symmetric matrix R [7, Thm. 8], which implies $\|\mathbf{T}_u^{-1}\| \leq \|\mathbf{T}_s^{-1}\|$. Hence, the structured and unstructured condition numbers for the stable invariant subspace of a Hamiltonian matrix are always the same.

A more general perturbation analysis for (block) Hamiltonian Schur forms, based on the technique of splitting operators and Lyapunov majorants, can be found in [24].

4.3. Orthogonal matrices. As an orthogonal matrix $A \in \mathbb{R}^{n\times n}$ is normal, the block Schur decomposition associated with a simple invariant subspace $\mathcal{X}$ is block diagonal:

$$[X, X_\perp]^T A [X, X_\perp] = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}.$$

Here, we will assume for convenience that X and $X_\perp$ are real. Both diagonal blocks, $A_{11} \in \mathbb{R}^{k\times k}$ and $A_{22} \in \mathbb{R}^{(n-k)\times(n-k)}$, are again orthogonal matrices.

The set of orthogonal matrices $\mathbb{S} = \{A : \ A^T A = I\}$ forms a smooth real manifold and the tangent space of $\mathbb{S}$ at A is given by $T_A\mathbb{S} = \{AW : \ W \in \operatorname{skew}(n)\}$. According to Theorem 3.10, this implies that the structured condition number is given by

$$\begin{aligned} c_{\mathbb{S}}(\mathcal{X}) &= \sup\Big\{ \|\mathbf{T}^{-1}(X_\perp^T AWX)\|_F : \ W \in \operatorname{skew}(n), \|AW\|_F = 1 \Big\} \\ &= \sup\Big\{ \|\mathbf{T}^{-1}(A_{22}X_\perp^T WX)\|_F : \ W \in \operatorname{skew}(n), \|W\|_F = 1 \Big\} \\ &= \sup\Big\{ \|\mathbf{T}^{-1}(A_{22}W_{21})\|_F : \ W_{21} \in \mathbb{R}^{(n-k)\times k}, \|W_{21}\|_F = 1 \Big\} \\ &= \sup\Big\{ \|\mathbf{T}^{-1}(\tilde{W}_{21})\|_F : \ \tilde{W}_{21} \in \mathbb{R}^{(n-k)\times k}, \|\tilde{W}_{21}\|_F = 1 \Big\} = c(\mathcal{X}), \end{aligned}$$

where $\mathbf{T} : R \mapsto A_{22}R - RA_{11}$. Here we used the fact that the "off-diagonal" block $W_{21} = X_\perp^T WX$ of a skew-symmetric matrix W has no particular structure. Hence, there is no difference between structured and unstructured condition numbers for invariant subspaces of orthogonal matrices.

5. Extension to matrix pencils. In this section, we extend the results of Section 3 to deflating subspaces of matrix pencils. The exposition is briefer than for the standard eigenvalue problem as many of the results can be derived by similar techniques.

Throughout this section it is assumed that our matrix pencil $A - \lambda B$ of interest, with $n \times n$ matrices A and B, is regular, i.e., $\det(A - \lambda B) \not\equiv 0$. The roots $\lambda \in \mathbb{C}$ (if any) of $\det(A - \lambda B) = 0$ are the finite eigenvalues of the pencil. In addition, if B is not invertible, then the pencil has infinite eigenvalues. A k-dimensional subspace $\mathcal{X}$ is called a *(right) deflating subspace* of $A - \lambda B$ if $A\mathcal{X}$ and $B\mathcal{X}$ are both contained in a subspace $\mathcal{Y}$ of dimension k. The regularity of $A - \lambda B$ implies that such a subspace $\mathcal{Y}$ is uniquely defined; we call $\mathcal{Y}$ a *left deflating subspace* and $(\mathcal{X}, \mathcal{Y})$ a *pair of deflating subspaces*, see [36] for a more detailed introduction.

Let $(\mathcal{X}, \mathcal{Y})$ be such a pair of deflating subspaces and let the columns of $X, X_\perp, Y, Y_\perp$ form orthonormal bases for $\mathcal{X}, \mathcal{X}^\perp, \mathcal{Y}, \mathcal{Y}^\perp$, respectively. Then $A - \lambda B$ admits the following *generalized block Schur decomposition*:

$$[Y, Y_\perp]^H (A - \lambda B) [X, X_\perp] = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} - \lambda \begin{bmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{bmatrix}. \tag{5.1}$$

The eigenvalues of $A - \lambda B$ are the union of the eigenvalues of the $k \times k$ pencil $A_{11} - \lambda B_{11}$ and the $(n-k) \times (n-k)$ pencil $A_{22} - \lambda B_{22}$.

An entity closely associated with (5.1) is the *generalized Sylvester operator*

$$\mathbf{T} : (R_r, R_l) \mapsto (A_{22}R_r - R_l A_{11}, B_{22}R_r - R_l B_{11}), \tag{5.2}$$

where R_r and R_l are $(n-k) \times k$ matrices. It can be shown [34] that $\mathbf{T}$ is invertible if and only if the matrix pencils $A_{11} - \lambda B_{11}$ and $A_{22} - \lambda B_{22}$ have no eigenvalues in common. Clearly, this property is independent of the choice of orthonormal bases for $\mathcal{X}$ and $\mathcal{Y}$, justifying the following definition.

Definition 5.1. *Deflating subspaces are called* simple *if the associated generalized Sylvester operator is invertible, i.e., if the matrix pencils $A_{11} - \lambda B_{11}$ and $A_{22} - \lambda B_{22}$ in (5.1) have no eigenvalues in common.*

Provided that $\mathbf{T}$ is invertible, the *separation of two matrix pencils $A_{11} - \lambda B_{11}$* and $A_{22} - \lambda B_{22}$ can be defined via the norm of the inverse of $\mathbf{T}$:

$$\begin{aligned}\operatorname{dif}[(A_{11}, B_{11}), (A_{22}, B_{22})] &:= 1/\sup\Big\{\|\mathbf{T}^{-1}(E_{21}, F_{21})\|_F :\ \|(E_{21}, F_{21})\|_F = 1\Big\}\\ &= 1/\|\mathbf{T}^{-1}\|,\end{aligned}$$

where we let $\|(E_{21}, F_{21})\|_F = \sqrt{\|E_{21}\|_F^2 + \|F_{21}\|_F^2}$. Not surprisingly, it turns out that $\mathbf{T}^{-1}$ governs the sensitivity of $(\mathcal{X}, \mathcal{Y})$ with respect to perturbations in A and B.

Theorem 5.2 ([38, 25]). *Let the matrix pencil $A - \lambda B$ have a generalized block Schur decomposition of the form (5.1) and assume the pair of deflating subspaces $(\mathcal{X}, \mathcal{Y}) = (\operatorname{span}(X), \operatorname{span}(Y))$ to be simple. Let $(A + E, B + F) \in \mathcal{B}(A, B)$ be a perturbation of (A, B), where $\mathcal{B}(A, B) \subset \mathbb{C}^{n\times n} \times \mathbb{C}^{n\times n}$ is a sufficiently small open neighborhood of (A, B). Then there exists an analytic function $f : \mathcal{B}(A, B) \to \mathbb{C}^{n\times k} \times \mathbb{C}^{n\times k}$ so that $(X, Y) = f(A, B)$, and the columns of $(\hat{X}, \hat{Y}) = f(A + E, B + F)$ span a pair of deflating subspaces for the perturbed matrix pencil $(A + E) - \lambda(B + F)$. Moreover, $X^H(\hat{X} - X) = Y^H(\hat{Y} - Y) = 0$, and we have the expansion*

$$(\hat{X}, \hat{Y}) = (X, Y) - (X_\perp R_r, Y_\perp R_l) + \mathcal{O}(\|[E, F]\|^2) \tag{5.3}$$

where $(R_r, R_l) = \mathbf{T}^{-1}(Y_\perp^H E X, Y_\perp^H F X)$ and $\mathbf{T}$ is the generalized Sylvester operator defined in (5.2).

By using similar techniques as in Section 3, it can be concluded from (5.3) that the condition number for $(\mathcal{X}, \mathcal{Y})$, defined as

$$c(\mathcal{X}, \mathcal{Y}) := \lim_{\epsilon\to 0} \sup\Big\{\|(\Theta(\mathcal{X}, \hat{\mathcal{X}}), \Theta(\mathcal{Y}, \hat{\mathcal{Y}}))\|_F/\epsilon :\ E, F \in \mathbb{C}^{n\times n}, \|(E, F)\|_F \le \epsilon\Big\},$$

happens to coincide with $\|\mathbf{T}^{-1}\| = 1/\operatorname{dif}[(A_{11}, B_{11}), (A_{22}, B_{22})]$; a result which goes back to Stewart [34, 35]. If $\operatorname{dif}[(A_{11}, B_{11}), (A_{22}, B_{22})] = 0$, then T is not invertible and, by convention, $c(\mathcal{X}, \mathcal{Y}) = \infty$. Algorithms that estimate dif efficiently by solving only a few generalized Sylvester equations can be found in [20, 21].

It may happen that $\mathcal{X}$ and $\mathcal{Y}$ are not equally sensitive to perturbations. In this case, $c(\mathcal{X}, \mathcal{Y})$ overestimates the sensitivity of one of the deflating subspaces; an aspect emphasized by Sun [38, 39], who has also pointed out that separating the influence of the operator $\mathbf{T}^{-1}$ on $\mathcal{X}$ and $\mathcal{Y}$ resolves this difficulty. However, for the purpose of simplifying the presentation we will only consider joint (structured) condition numbers for $(\mathcal{X}, \mathcal{Y})$.

Definition 5.3. *Let $\mathbb{S} \subseteq \mathbb{C}^{n\times n} \times \mathbb{C}^{n\times n}$ and let $(\mathcal{X}, \mathcal{Y})$ be a pair of deflating subspaces of a matrix pencil $A - \lambda B$ with $(A, B) \in \mathbb{S}$. Then the* structured condition number for $(\mathcal{X}, \mathcal{Y})$ *with respect to $\mathbb{S}$ is defined as*

$$c_{\mathbb{S}}(\mathcal{X}, \mathcal{Y}) := \lim_{\epsilon\to 0} \sup_{\substack{(A+E,B+F)\in\mathbb{S}\\ \|(E,F)\|_F\le\epsilon}} \inf\left\{ \|(\Theta(\mathcal{X}, \hat{\mathcal{X}}), \Theta(\mathcal{Y}, \hat{\mathcal{Y}}))\|_F/\epsilon \ :\ \begin{array}{r}(\hat{\mathcal{X}}, \hat{\mathcal{Y}}) \text{ is a deflating}\\ \text{subspace pair for}\\ (A+E) - \lambda(B+F)\end{array}\right\}.$$

If $\mathbb{S} = \mathbb{C}^{n\times n} \times \mathbb{C}^{n\times n}$, then $c_{\mathbb{S}}(\mathcal{X}, \mathcal{Y}) = c(\mathcal{X}, \mathcal{Y})$.

A straightforward generalization of Lemma 3.2 relates $c_{\mathbb{S}}(\mathcal{X},\mathcal{Y})$ to the norm of $\mathbf{T}^{-1}$ restricted to a certain subset.

LEMMA 5.4. *Let $(\mathcal{X},\mathcal{Y})$ be a pair of simple deflating subspaces of a matrix pencil $A-\lambda B$ corresponding to a generalized block Schur decomposition of the form (2.1). Then the structured condition number for $\mathcal{X}$ with respect to $\mathbb{S}\subseteq\mathbb{C}^{n\times n}$ satisfies*

$$c_{\mathbb{S}}(\mathcal{X})=\lim_{\epsilon\to 0}\sup\Big\{\|\mathbf{T}^{-1}(Y_\perp^H EX, Y_\perp^H FX)\|_F/\epsilon:\ (A+E,B+F)\in\mathbb{S}, \|(E,F)\|_F\le\epsilon\Big\},$$

where $\mathbf{T}$ is the generalized Sylvester operator defined in (5.2).

5.1. A Kronecker product approach. Using Kronecker products, the generalized Sylvester operator $\mathbf{T}$ can be represented as

$$\operatorname{vec}(\mathbf{T}(R_r,R_l))=K_{\mathbf{T}}\begin{bmatrix}\operatorname{vec}(R_r)\\ \operatorname{vec}(R_l)\end{bmatrix},$$

with the $2k(n-k)\times 2k(n-k)$ matrix

$$K_{\mathbf{T}}=\begin{bmatrix} I_k\otimes A_{22} & -A_{11}^T\otimes I_{n-k}\\ I_k\otimes B_{22} & -B_{11}^T\otimes I_{n-k}\end{bmatrix}.$$

This implies $c(\mathcal{X},\mathcal{Y})=\|\mathbf{T}^{-1}\|=\|K_{\mathbf{T}}^{-1}\|_2$.

In the following, we will assume that the structure $\mathbb{S}$ under consideration takes the form $\mathbb{S}=(A,B)+\mathbb{L}$. Here, $\mathbb{L}$ denotes a linear matrix pencil subspace, i.e., $(E_1,F_1)\in\mathbb{L}$ and $(E_2,F_2)\in\mathbb{L}$ imply $(\alpha E_1+\beta E_2,\alpha F_1+\beta F_2)\in\mathbb{L}$ for all $\alpha,\beta\in\mathbb{K}$, where $\mathbb{K}=\mathbb{R}$ if $\mathbb{L}$ is real or $\mathbb{K}=\mathbb{C}$ if $\mathbb{L}$ is complex. Let m be the dimension of $\mathbb{L}$. Then one can always find a $2n^2\times m$ pattern matrix $M_{\mathbb{L}}$ such that for every $(E,F)\in\mathbb{L}$ there exists a uniquely defined parameter vector $p\in\mathbb{K}^m$ with

$$\begin{bmatrix}\operatorname{vec}(E)\\ \operatorname{vec}(F)\end{bmatrix}=M_{\mathbb{L}}p,\quad \|(E,F)\|_F=\|p\|_2.$$

This yields for $(R_r,R_l)=\mathbf{T}^{-1}(Y_\perp^H EX, Y_\perp^H FX)$ with $(E,F)\in\mathbb{L}$,

$$\begin{bmatrix}\operatorname{vec}(R_r)\\ \operatorname{vec}(R_l)\end{bmatrix}=K_{\mathbf{T}}^{-1}\begin{bmatrix} X\otimes Y_\perp^H & 0\\ 0 & X\otimes Y_\perp^H\end{bmatrix}M_{\mathbb{L}}p.$$

Hence, Lemma 5.4 implies

$$\begin{aligned} c_{(A,B)+\mathbb{L}}(\mathcal{X},\mathcal{Y}) &= \sup_{\substack{p\in\mathbb{K}^m\\ \|p\|_2=1}}\left\|K_{\mathbf{T}}^{-1}\begin{bmatrix} X\otimes Y_\perp^H & 0\\ 0 & X\otimes Y_\perp^H\end{bmatrix}M_{\mathbb{L}}p\right\|_2\\ &= \left\|K_{\mathbf{T}}^{-1}\begin{bmatrix} X\otimes Y_\perp^H & 0\\ 0 & X\otimes Y_\perp^H\end{bmatrix}M_{\mathbb{L}}\right\|_2. \end{aligned}\tag{5.4}$$

Note that the latter equality only holds provided that either $\mathbb{K}=\mathbb{C}$, or all of $\mathbb{K}$, A, B, $\mathcal{X}$ and $\mathcal{Y}$ are real. Otherwise, inequalities analogous to Lemma 3.3 can be derived.

5.2. An orthogonal decomposition approach. In this section, we extend the orthogonal decomposition approach of Section 3.2 to matrix pencils in order to gain more insight into the relationship between the structured and unstructured condition numbers for a pair of deflating subspaces.

For this purpose, assume the pair of deflating subspaces $(\mathcal{X}, \mathcal{Y})$ to be simple, and let the columns of $X, X_\perp, Y, Y_\perp$ form orthonormal bases for $\mathcal{X}, \mathcal{X}^\perp, \mathcal{Y}, \mathcal{Y}^\perp$, respectively. Let

$$\mathcal{N} := \{(Y_\perp^H E X, Y_\perp^H F X) : \ (E, F) \in \mathbb{L}\},$$

and let $\mathcal{M}$ denote the preimage of $\mathcal{N}$ under $\mathbf{T}$, i.e., $\mathcal{M} := \mathbf{T}^{-1}(\mathcal{N})$. Then Lemma 5.4 implies that the structured condition number for $(\mathcal{X}, \mathcal{Y})$ is given by

$$c_{(A,B)+\mathbb{L}}(\mathcal{X}, \mathcal{Y}) = \|\mathbf{T}_s^{-1}\|,$$

where $\mathbf{T}_s$ is the restriction of $\mathbf{T}$ to $\mathcal{M} \to \mathcal{N}$, i.e., $\mathbf{T}_s := \mathbf{T}\big|_{\mathcal{M}\to\mathcal{N}}$.

Let us assume that we additionally have the property that the linear matrix operator

$$\mathbf{T}^\star : (Q_r, Q_l) \mapsto (A_{22}^H Q_r + B_{22}^H Q_l, -Q_r A_{11}^H - Q_l B_{11}^H)$$

satisfies $\mathbf{T}^\star : \mathcal{N} \to \mathcal{M}$. This is equivalent to the condition $\mathbf{T} : \mathcal{M}^\perp \to \mathcal{N}^\perp$, where $^\perp$ denotes the orthogonal complement w.r.t. the inner product $\langle (X_r, X_l), (Y_r, Y_l) \rangle = \operatorname{trace}(Y_r^H X_r + Y_l^H X_l)$. Note that $\mathbf{T}^\star$ can be considered as the linear operator dual to $\mathbf{T}$:

$$\langle \mathbf{T}(R_r, R_l), (Q_r, Q_l) \rangle = \langle (R_r, R_l), \mathbf{T}^\star(Q_r, Q_l) \rangle.$$

The same conclusions as for the matrix case in Section 3.2 can be drawn: $\mathbf{T}$ decomposes orthogonally into $\mathbf{T}_s$ and $\mathbf{T}_u := \mathbf{T}\big|_{\mathcal{M}^\perp\to\mathcal{N}^\perp}$, and we have

$$c(\mathcal{X}, \mathcal{Y}) = \max\{\|\mathbf{T}_s^{-1}\|, \|\mathbf{T}_u^{-1}\|\}.$$

5.3. Global perturbation bounds. To derive global perturbation bounds we consider, additionally to (5.1), the *perturbed* generalized block Schur decomposition

$$[Y, Y_\perp]^H \Big((A+E) - \lambda (B+F) \Big) [X, X_\perp] = \begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} \\ E_{21} & \hat{A}_{22} \end{bmatrix} - \lambda \begin{bmatrix} \hat{B}_{11} & \hat{B}_{12} \\ F_{21} & \hat{B}_{22} \end{bmatrix}. \tag{5.5}$$

The following approach follows the work by Stewart [34, 35], which has been refined by Demmel and Kågström in [11]. In order to obtain bases $(\hat{X}, \hat{Y})$ for a nearby pair of perturbed deflating subspaces $(\hat{\mathcal{X}}, \hat{\mathcal{Y}})$ we look for $(n-k) \times k$ matrices R_r and R_l such that the matrix pencil

$$\begin{bmatrix} I_k & 0 \\ R_l & I_{n-k} \end{bmatrix} \left(\begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} \\ E_{21} & \hat{A}_{22} \end{bmatrix} - \lambda \begin{bmatrix} \hat{B}_{11} & \hat{B}_{12} \\ F_{21} & \hat{B}_{22} \end{bmatrix} \right) \begin{bmatrix} I_k & 0 \\ -R_r & I_{n-k} \end{bmatrix}$$

is in block upper triangular form. This is equivalent to the condition that the pair (R_r, R_l) satisfies the following system of quadratic matrix equations:

$$\begin{aligned} \hat{A}_{22} R_r - R_l \hat{A}_{11} + R_l \hat{A}_{12} R_r &= E_{21}, \\ \hat{B}_{22} R_r - R_l \hat{B}_{11} + R_l \hat{B}_{12} R_r &= F_{21}. \end{aligned} \tag{5.6}$$

The following assumptions on the linear structure $\mathbb{L} \subseteq \mathbb{C}^{n\times n} \times \mathbb{C}^{n\times n}$ are related to the solvability of (5.6), along the lines of the assumptions A1 and A2 for the matrix case:

A3: Let $\mathcal{N} = \{(Y_\perp^H G X, Y_\perp^H H X) : (G,H) \in \mathbb{L}\}$ and

$$\hat{\mathbf{T}} : (R_r, R_l) \mapsto (\hat{A}_{22} R_r - R_l \hat{A}_{11}, \hat{B}_{22} R_r - R_l \hat{B}_{11})$$

Then there exists a linear matrix space $\mathcal{M}$, having the same dimension as $\mathcal{N}$, such that $\hat{\mathbf{T}} : \mathcal{M} \to \mathcal{N}$ and $(R_l \hat{A}_{12} R_r, R_l \hat{B}_{12} R_r) \in \mathcal{N}$ for all $(R_r, R_l) \in \mathcal{M}$.

A4: The restricted operator $\hat{\mathbf{T}}_s := \hat{\mathbf{T}}\big|_{\mathcal{M}\to\mathcal{N}}$ is invertible.

Theorem 5.5. *Assume that A3 and A4 hold. If*

$$\kappa := 4\|\hat{\mathbf{T}}_s^{-1}\|^2 \, \|(\hat{A}_{12}, \hat{B}_{12})\|_F \, \|(E_{21}, F_{21})\|_F < 1$$

then there exists a solution $(R_r, R_l) \in \mathcal{M}$ *of (5.6) with*

$$\|(R_r, R_l)\|_F \le \frac{2\|\hat{\mathbf{T}}_s^{-1}\| \, \|(E_{21}, F_{21})\|_F}{1+\sqrt{1-\kappa}} < 2\|\hat{\mathbf{T}}_s^{-1}\| \, \|(E_{21}, F_{21})\|_F.$$

Proof. It follows from A3 that the iteration

$$(R_0, L_0) \leftarrow (0,0), \quad (R_{i+1}, L_{i+1}) \leftarrow \hat{\mathbf{T}}_s^{-1}(E_{21} - L_i \hat{A}_{12} R_i, F_{21} - L_i \hat{B}_{12} R_i)$$

is well-defined and $(R_i, L_i) \in \mathcal{M}$ for all i. Its convergence and the bound (5.7) can be proved along the lines of the proof of [36, Thm. V.2.11]. □

Any solution (R_r, R_l) of (5.6) yields a pair of deflating subspaces $(\hat{\mathcal{X}}, \hat{\mathcal{Y}})$ of the perturbed pencil $(A+E) - \lambda(B+F)$ with the bases $\hat{X} = X - X_\perp R_r$ and $\hat{Y} = Y - Y_\perp R_l$. Considering the solution constructed in Theorem 5.5, we obtain

$$\|(\tan\Theta(\mathcal{X}, \hat{\mathcal{X}}), \tan\Theta(\mathcal{Y}, \hat{\mathcal{Y}}))\|_F \le \frac{2\|\hat{\mathbf{T}}_s^{-1}\| \, \|(E_{21}, F_{21})\|_F}{1+\sqrt{1-\kappa}}. \tag{5.7}$$

The proof of Lemma 3.8 can be easily adapted to relate $\|\hat{\mathbf{T}}_s^{-1}\|$ to $\|\mathbf{T}_s^{-1}\|$.

Lemma 5.6. *Assume that A3 holds, and that the unperturbed generalized Sylvester operator* $\mathbf{T}$ *defined in (5.2) also satisfies* $\mathbf{T} : \mathcal{M} \to \mathcal{N}$. *If* $\mathbf{T}_s := \mathbf{T}\big|_{\mathcal{M}\to\mathcal{N}}$ *is invertible and* $1/\|\mathbf{T}_s^{-1}\| > \|(E_{11}, F_{11})\|_F + \|(E_{22}, F_{22})\|_F$, *then* $\hat{\mathbf{T}}_s$ *is also invertible and satisfies*

$$\|\hat{\mathbf{T}}_s^{-1}\| \le \frac{\|\mathbf{T}_s^{-1}\|}{1 - \|\mathbf{T}_s^{-1}\|(\|(E_{11}, F_{11})\|_F + \|(E_{22}, F_{22})\|_F)}. \tag{5.8}$$

Combining (5.7) and (5.8) implies $c_{(A,B)+\mathbb{L}}(\mathcal{X}, \mathcal{Y}) \le \|\mathbf{T}_s^{-1}\|$. Assuming $\mathbf{T}^\star : \mathcal{N} \to \mathcal{M}$, it can be shown that $c_{(A,B)+\mathbb{L}}(\mathcal{X}, \mathcal{Y})$ and $\|\mathbf{T}_s^{-1}\|$ are equal.

Theorem 5.7. *Assume that A3 holds with the same matrix space* $\mathcal{M}$ *for all* $\hat{\mathbf{T}}$ *corresponding to perturbations* $(E, F) \in \mathbb{L}$. *Moreover, assume that the generalized Sylvester operator* $\mathbf{T}$ *defined in (5.2) additionally satisfies* $\mathbf{T}^\star : \mathcal{N} \to \mathcal{M}$ *and that* $\mathbf{T}_s := \mathbf{T}\big|_{\mathcal{M}\to\mathcal{N}}$ *is invertible. Then* $c_{(A,B)+\mathbb{L}}(\mathcal{X}, \mathcal{Y}) = \|\mathbf{T}_s^{-1}\|$.

Proof. To adapt the proof of Theorem 3.9 to matrix pencils, we consider perturbations of the form $(E, F) = (\epsilon Y_\perp E_{21} X^H, \epsilon Y_\perp F_{21} X^H)$, where $(E_{21}, F_{21}) \in \mathcal{N}$ is chosen such that $\|(E_{21}, F_{21})\|_F = 1$ and $\|\mathbf{T}^{-1}(E_{21}, F_{21})\|_F = \|\mathbf{T}_s^{-1}\|$. The bound (5.7) implies for sufficiently small $\epsilon > 0$ that the nearest deflating subspace $(\hat{\mathcal{X}}, \hat{\mathcal{Y}})$ of $(A+E) - \lambda(B+F)$ satisfies

$$\max\{\|\Theta(\mathcal{X}, \hat{\mathcal{X}})\|_2, \|\Theta(\mathcal{Y}, \hat{\mathcal{Y}})\|_2\} < \pi/2.$$

This yields the existence of a matrix pair (R, L) such that the columns of $(\hat{X}, \hat{Y}) = (X - X_\perp R, Y - Y_\perp L)$ form bases for $(\hat{\mathcal{X}}, \hat{\mathcal{Y}})$. Equivalently, (R, L) satisfies

$$\mathbf{T}(R, L) + \mathbf{Q}(R, L) = (\epsilon E_{21}, \epsilon F_{21}),$$

where $\mathbf{Q}(R, L) = (L\hat{A}_{12}R, L\hat{B}_{12}R)$. Let us decompose $(R, L) = (R_s, L_s) + (R_u, L_u)$, where $(R_s, L_s) \in \mathcal{M}$ and $(R_u, L_u) \in \mathcal{M}^\perp$. Then $\mathbf{T}(R_s, L_s) \in \mathcal{N}$ and $\mathbf{T}(R_u, L_u) \in \mathcal{N}^\perp$. This implies, as in the proof of Theorem 3.9,

$$\lim_{\epsilon \to 0} \|(R, L)\|_F/\epsilon \geq \lim_{\epsilon \to 0} \|(R_s, L_s)\|_F/\epsilon = \|\mathbf{T}^{-1}(E_{21}, F_{21})\|_F = \|\mathbf{T}_s^{-1}\|,$$

and consequently $c_{(A,B)+\mathbb{L}}(\mathcal{X}, \mathcal{Y}) \geq \|\mathbf{T}_s^{-1}\|$, which concludes the proof. □

5.4. Nonlinear structures. The following theorem shows that the results of Sections 5.1 and 5.2 can also be used to address matrix pencil structures that form smooth manifolds.

THEOREM 5.8. *Let $\mathbb{S} \subseteq \mathbb{K}^{n \times n} \times \mathbb{K}^{n \times n}$ with $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ be a smooth manifold. Let $(\mathcal{X}, \mathcal{Y})$ be a pair of simple deflating subspaces of $A - \lambda B$ with $(A, B) \in \mathbb{S}$, corresponding to a generalized block Schur decomposition of the form (5.1). Then the structured condition number for $(\mathcal{X}, \mathcal{Y})$ with respect to $\mathbb{S}$ satisfies*

$$c_{\mathbb{S}}(\mathcal{X}, \mathcal{Y}) = \sup\left\{\|\mathbf{T}^{-1}(Y_\perp^H E X, Y_\perp^H F X)\|_F : \ (E, F) \in T_{(A,B)}\mathbb{S}, \|(E, F)\|_F = 1\right\},$$

where $\mathbf{T}$ is the generalized Sylvester operator defined in (5.2) and $T_{(A,B)}\mathbb{S}$ is the tangent space of $\mathbb{S}$ at (A, B).

Proof. The result follows from a rather straightforward extension of the proof of Theorem 3.10. □

5.5. Example: Palindromic matrix pencils. To illustrate the obtained results for structured matrix pencils, let us consider a matrix pencil of the form $A + \lambda A^T$ with $A \in \mathbb{C}^{2n \times 2n}$. A matrix pencil that takes this form is called *palindromic*; it arises, e.g., from structure-preserving linearizations of palindromic matrix polynomials [16, 28]. The following result provides a structured Schur form.

LEMMA 5.9 ([16]). *Let $A \in \mathbb{C}^{2n \times 2n}$, then there exists a unitary matrix $U \in \mathbb{C}^{2n \times 2n}$ such that*

$$U^T A U = \begin{bmatrix} 0 & \cdots & 0 & t_{1,2n} \\ \vdots & \iddots & t_{2,2n-1} & t_{2,2n} \\ 0 & \iddots & \iddots & \vdots \\ t_{2n,1} & t_{2n,2} & \cdots & t_{2n,2n} \end{bmatrix} =: T,$$

i.e., T is anti-triangular.

It should be emphasized that U^T in Lemma 5.9 denotes the complex transpose of U, i.e., $U^T A U$ is *not* similar to A. Nevertheless, $T + \lambda T^T$ is equivalent to $A + \lambda A^T$, implying that the eigenvalues of $A + \lambda A^T$ are given by

$$-t_{1,2n}/t_{2n,1}, \ldots, -t_{n,n+1}/t_{n+1,n}, -t_{n+1,n}/t_{n,n+1}, \ldots, -t_{2n,1}/t_{1,2n}.$$

It follows immediately that the eigenvalues have the following pairing: λ is an eigenvalue of $A + \lambda A^T$ if and only if $1/\lambda$ is an eigenvalue. Zero eigenvalues are included in these pairings as $\lambda = 0$ and $1/\lambda = \infty$.

In the following, we consider the (right) deflating subspace $\mathcal{X}$ belonging to the eigenvalues $-t_{n+1,n}/t_{n,n+1},\dots,-t_{2n,1}/t_{1,2n}$. Let the columns of X and $X_\perp$ form orthonormal bases for $\mathcal{X}$ and $\mathcal{X}^\perp$, respectively. Then Lemma 5.9 implies a structured generalized block Schur decomposition of the form

$$[X_\perp,\, X]^T(A+\lambda A^T)[X,X_\perp] = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} + \lambda \begin{bmatrix} A_{22}^T & A_{12}^T \\ 0 & A_{11}^T \end{bmatrix} \tag{5.9}$$

with $A_{11},A_{22}\in\mathbb{C}^{n\times n}$. Note that this also shows that $\overline{X_\perp}$, obtained from $X_\perp$ by conjugating its entries, spans a left deflating subspace $\mathcal{Y}$ belonging to the eigenvalues $-t_{n+1,n}/t_{n,n+1},\dots,-t_{2n,1}/t_{1,2n}$. We require the following preliminary result for obtaining the structured condition number of $(\mathcal{X},\mathcal{Y})$ with respect to palindromic perturbations.

Lemma 5.10. *Let $C,D\in\mathbb{C}^{n\times n}$, then the matrix equation*

$$CR+\alpha R^TD^T = F, \tag{5.10}$$

where $\alpha\in\{1,-1\}$, has a unique solution R for any $F\in\mathbb{C}^{n\times n}$ if and only if the following two conditions hold for the eigenvalues of $C-\lambda D$:

1. *if $\lambda\neq\alpha$ is an eigenvalue then $1/\lambda$ is* not *an eigenvalue;*
2. *if $\lambda=\alpha$ is an eigenvalue, it has algebraic multiplicity one.*

Proof. The proof can be found in Appendix A. □

The generalized Sylvester operator associated with (5.9) takes the form

$$\mathbf{T}:(R_r,R_l)\mapsto(A_{22}R_r+R_lA_{11},\,A_{11}^TR_r+R_lA_{22}^T).$$

Considering the linear space

$$\mathcal{N} := \left\{(X^TEX,-X^TE^TX):\ E\in\mathbb{C}^{2n\times 2n}\right\} = \left\{(E_{21},-E_{21}^T):\ E_{21}\in\mathbb{C}^{n\times n}\right\},$$

we have $\mathbf{T}:\mathcal{N}\to\mathcal{N}$ and $\mathbf{T}:\mathcal{N}^\perp\to\mathcal{N}^\perp$, where $\mathcal{N}^\perp = \left\{(E_{21},E_{21}^T):\ E_{21}\in\mathbb{C}^{n\times n}\right\}$. Moreover, $(R_lA_{12}R_r,-R_lA_{12}^TR_r)\in\mathcal{N}$ for all $(R_l,R_r)\in\mathcal{N}$. The restricted Sylvester operators $\mathbf{T}_s = \mathbf{T}\big|_{\mathcal{N}\to\mathcal{N}}$ and $\mathbf{T}_u = \mathbf{T}\big|_{\mathcal{N}^\perp\to\mathcal{N}^\perp}$ can be identified with the matrix operators

$$\mathbf{S}_s: R\mapsto A_{22}R-R^TA_{11},\quad \mathbf{S}_u: R\mapsto A_{22}R+R^TA_{11},$$

in the sense that

$$\mathbf{T}_s(R,-R^T) = (\mathbf{S}_s(R),-\mathbf{S}_s(R)^T),\quad \mathbf{T}_u(R,R^T) = (\mathbf{S}_u(R),\mathbf{S}_u(R)^T).$$

In particular, $\mathbf{T}_s$ is invertible if and only if $\mathbf{S}_s$ is invertible, which in turn is equivalent to require $A_{22}-\lambda A_{11}^T$ to satisfy the conditions of Lemma 5.10 for $\alpha=-1$. In this case, all assumptions of Theorem 5.7 are satisfied and the structured condition number for the deflating subspace pair $(\mathcal{X},\mathcal{Y}) = (\operatorname{span}(X),\operatorname{span}(\overline{X_\perp}))$ with respect to $\mathbb{S} = \{(E,-E^T):\ E\in\mathbb{C}^{2n\times 2n}\}$ is given by

$$c_{\mathbb{S}}(\mathcal{X},\mathcal{Y}) = \|\mathbf{T}_s^{-1}\| = \sqrt{2}\,\|\mathbf{S}_s^{-1}\| = \frac{\sqrt{2}}{\inf\{\|A_{22}R-R^TA_{11}\|_F:\ R\in\mathbb{C}^{n\times n},\|R\|_F=1\}}.$$

On the other hand, the unstructured condition number satisfies

$$c(\mathcal{X},\mathcal{Y}) = \sqrt{2}\,\max\{\|\mathbf{S}_s^{-1}\|,\|\mathbf{S}_u^{-1}\|\}.$$

This shows that the unstructured condition number can be much larger than the structured condition number, e.g., if $A_{22} - \lambda A_{11}^T$ has a simple eigenvalue close to -1. If one of the eigenvalues of $A_{22} - \lambda A_{11}^T$ happens to be exactly -1 then $(\mathcal{X}, \mathcal{Y})$ is not stable under unstructured perturbations, but Lemma 5.10 implies that it can still be stable under structured perturbations. In these cases, the use of a computational method that yields structured backward errors is likely to be significantly more accurate than other methods.

EXAMPLE 5.11. *For $n = 1$, we obtain*

$$\|\mathbf{S}_s^{-1}\| = \frac{1}{|A_{22} - A_{11}|}, \quad \|\mathbf{S}_u^{-1}\| = \frac{1}{|A_{22} + A_{11}|}.$$

Hence, if $A_{22}/A_{11} \approx -1$ then $c(\mathcal{X}, \mathcal{Y}) \gg c_{\mathbb{S}}(\mathcal{X}, \mathcal{Y})$.

6. Conclusions.

We have derived directly computable expressions for structured condition numbers of invariant and deflating subspaces for smooth manifolds of structured matrices and matrix pencils. An orthogonal decomposition of the associated Sylvester operators yields global perturbation bounds that remain valid even in cases where the subspace is unstable under unstructured perturbations. It also provides additional insight into the difference between structured and unstructured condition numbers. We have identified structures for which this difference can be significant (block cyclic, palindromic) or negligible (Hamiltonian, orthogonal). Developing efficient structured condition estimators going beyond the simple method mentioned in Remark 3.4 remains an important future task.

The examples suggest some relation between structures that admit the proposed orthogonal decomposition approach and those that admit structured Schur decompositions. However, addressing this question thoroughly requires further investigation.

Acknowledgments. The authors sincerely thank Ji-guang Sun for helpful remarks and discussions on an earlier version of this paper. Portions of this work were completed while the second author was enjoying the hospitality and support of the Department of Mathematics, University of Kansas.

REFERENCES

[1] R. H. Bartels and G. W. Stewart. Algorithm 432: The solution of the matrix equation $AX - BX = C$. *Communications of the ACM*, 8:820–826, 1972.

[2] P. Benner, D. Kressner, and V. Mehrmann. Skew-Hamiltonian and Hamiltonian eigenvalue problems: Theory, algorithms and applications. In Z. Drmač, M. Marušić, and Z. Tutek, editors, *Proceedings of the Conference on Applied Mathematics and Scientific Computing, Brijuni (Croatia), June 23-27, 2003*, pages 3–39. Springer-Verlag, 2005.

[3] P. Benner, V. Mehrmann, and H. Xu. Perturbation analysis for the eigenvalue problem of a formal product of matrices. *BIT*, 42(1):1–43, 2002.

[4] A. Bojanczyk, G. H. Golub, and P. Van Dooren. The periodic Schur decomposition; algorithm and applications. In *Proc. SPIE Conference*, volume 1770, pages 31–42, 1992.

[5] R. Byers. A LINPACK-style condition estimator for the equation $AX - XB^T = C$. *IEEE Trans. Automat. Control*, 29(10):926–928, 1984.

[6] R. Byers and D. Kressner. On the condition of a complex eigenvalue under real perturbations. *BIT*, 44(2):209–215, 2004.

[7] R. Byers and S. Nash. On the singular "vectors" of the Lyapunov operator. *SIAM J. Algebraic Discrete Methods*, 8(1):59–66, 1987.

[8] F. Chatelin. *Eigenvalues of matrices.* John Wiley & Sons Ltd., Chichester, 1993.

[9] P. I. Davies. Structured conditioning of matrix functions. *Electron. J. Linear Algebra*, 11:132–161, 2004.

[10] J. W. Demmel. Computing stable eigendecompositions of matrices. *Linear Algebra Appl.*, 79:163–193, 1986.

[11] J. W. Demmel and B. Kågström. Computing stable eigendecompositions of matrix pencils. *Linear Algebra Appl.*, 88/89:139–186, 1987.

[12] G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, MD, third edition, 1996.

[13] J. J. Hench and A. J. Laub. Numerical solution of the discrete-time periodic Riccati equation. *IEEE Trans. Automat. Control*, 39(6):1197–1210, 1994.

[14] 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.

[15] N. J. Higham. *Accuracy and Stability of Numerical Algorithms*. SIAM, Philadelphia, PA, second edition, 2002.

[16] A. Hilliges, C. Mehl, and V. Mehrmann. On the solution of palindromic eigenvalue problems. In *Proceedings of ECCOMAS, Jyväskylä, Finland*, 2004.

[17] M. E. Hochstenbach and B. Plestenjak. Backward error, condition numbers, and pseudospectra for the multiparameter eigenvalue problem. *Linear Algebra Appl.*, 375:63–81, 2003.

[18] I. Jonsson and B. Kågström. Recursive blocked algorithm for solving triangular systems. I. one-sided and coupled Sylvester-type matrix equations. *ACM Trans. Math. Software*, 28(4):392–415, 2002.

[19] B. Kågström and P. Poromaa. Distributed and shared memory block algorithms for the triangular Sylvester equation with sep^{-1} estimators. *SIAM J. Matrix Anal. Appl.*, 13(1):90–101, 1992.

[20] B. Kågström and P. Poromaa. Computing eigenspaces with specified eigenvalues of a regular matrix pair (A, B) and condition estimation: theory, algorithms and software. *Numer. Algorithms*, 12(3-4):369–407, 1996.

[21] B. Kågström and P. Poromaa. LAPACK-style algorithms and software for solving the generalized Sylvester equation and estimating the separation between regular matrix pairs. *ACM Trans. Math. Software*, 22(1):78–103, 1996.

[22] M. Karow, D. Kressner, and F. Tisseur. Structured eigenvalue condition numbers. Numerical Analysis Report No. 467, Manchester Centre for Computational Mathematics, Manchester, England, 2005.

[23] M. Konstantinov, D.-W. Gu, V. Mehrmann, and P. Petkov. *Perturbation theory for matrix equations*, volume 9 of *Studies in Computational Mathematics*. North-Holland Publishing Co., Amsterdam, 2003.

[24] M. Konstantinov, V. Mehrmann, and P. Petkov. Perturbation analysis of Hamiltonian Schur and block-Schur forms. *SIAM J. Matrix Anal. Appl.*, 23(2):387–424, 2001.

[25] D. Kressner. *Numerical Methods and Software for General and Structured Eigenvalue Problems*. PhD thesis, TU Berlin, Institut für Mathematik, Berlin, Germany, 2004.

[26] D. Kressner. Perturbation bounds for isotropic invariant subspaces of skew-Hamiltonian matrices. *SIAM J. Matrix Anal. Appl.*, 26(4):947–961, 2005.

[27] W.-W. Lin and J.-G. Sun. Perturbation analysis for the eigenproblem of periodic matrix pairs. *Linear Algebra Appl.*, 337:157–187, 2001.

[28] D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann. Palindromic polynomial eigenvalue problems: Good vibrations from good linearizations, 2005. Submitted.

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

[30] V. Mehrmann and D. S. Watkins. Structure-preserving methods for computing eigenpairs of large sparse skew-Hamiltonian/Hamiltonian pencils. *SIAM J. Sci. Comput.*, 22(6):1905–1925, 2000.

[31] S. Noschese and L. Pasquini. Eigenvalue condition numbers: zero-structured versus traditional. Preprint 28, Mathematics Department, University of Rome La Sapienza, Italy, 2004.

[32] J. R. Rice. A theory of condition. *SIAM J. Numer. Anal.*, 3:287–310, 1966.

[33] G. W. Stewart. Error bounds for approximate invariant subspaces of closed linear operators. *SIAM J. Numer. Anal.*, 8:796–808, 1971.

[34] G. W. Stewart. On the sensitivity of the eigenvalue problem $Ax = \lambda Bx$. *SIAM J. Numer. Anal.*, 9:669–686, 1972.

[35] G. W. Stewart. Error and perturbation bounds for subspaces associated with certain eigenvalue problems. *SIAM Rev.*, 15:727–764, 1973.

[36] G. W. Stewart and J.-G. Sun. *Matrix Perturbation Theory*. Academic Press, New York, 1990.

[37] J.-G. Sun. Perturbation expansions for invariant subspaces. *Linear Algebra Appl.*, 153:85–97, 1991.

[38] J.-G. Sun. Perturbation analysis of singular subspaces and deflating subspaces. *Numer. Math.*, 73(2):235–263, 1996.

[39] 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.

[40] J.-G. Sun. Perturbation analysis of algebraic Riccati equations. Technical report UMINF 02-03, Department of Computing Science, University of Umeå, Umeå, Sweden, 2002.

[41] F. Tisseur. A chart of backward errors for singly and doubly structured eigenvalue problems. *SIAM J. Matrix Anal. Appl.*, 24(3):877–897, 2003.

[42] D. S. Watkins. Product eigenvalue problems. *SIAM Rev.*, 47:3–40, 2005.

Appendix A.

Proof of Lemma 5.10. We only have to show the case $\alpha = 1$, as $\alpha = -1$ follows from $\alpha = 1$ after replacing D^T by $-D^T$. First, we prove by induction that (5.10) has a solution for any F if the two conditions hold. For $n = 1$, the first condition implies $C \neq -D$ and thus $R = F/(C + D)$. For $n > 1$, using the generalized Schur decomposition of $C - \lambda D$, we may assume without loss of generality that C and D have upper triangular form. Partition the matrices

$$C = \begin{bmatrix} C_{11} & C_{12} \\ 0 & C_{22} \end{bmatrix}, \; D = \begin{bmatrix} D_{11} & D_{12} \\ 0 & D_{22} \end{bmatrix}, \; F = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix}, \; R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}$$

conformally with no void blocks, then (5.10) can be written as

$$F_{11} = C_{11}R_{11} + C_{12}R_{21} + R_{11}^T D_{11} + R_{21}^T D_{12}, \tag{A.1}$$
$$F_{21} = C_{22}R_{21} + R_{12}^T D_{11} + R_{22}^T D_{12}, \tag{A.2}$$
$$F_{12} = C_{11}R_{12} + C_{12}R_{22} + R_{21}^T D_{22}, \tag{A.3}$$
$$F_{22} = C_{22}R_{22} + R_{22}^T D_{22}. \tag{A.4}$$

By the induction assumption, the matrix equation (A.4) is solvable. Thus, R_{22} can be regarded as known, which turns (A.2)–(A.3), after transposing (A.3), into a generalized Sylvester equation associated with the matrix pencils $C_{22}+\lambda D_{11}^T$ and $D_{22}^T+\lambda C_{11}$. Under the given conditions these two pencils have no eigenvalue in common. Hence, (A.2)–(A.3) is solvable and R_{12} as well as R_{21} can be regarded as known. This turns (A.1) into a matrix equation of the form (5.10) of smaller dimension, which is – by the induction assumption – solvable. The uniqueness of the constructed solution R follows from the fact that (5.10) can be regarded as a square linear system of equations in the entries of R.

For the other direction, consider the linear matrix operator $\mathbf{S} : R \mapsto CR + R^T D^T$. We will make use of the fact that the matrix equation (5.10) is uniquely solvable if and only if $\mathrm{kernel}(\mathbf{S}) = \{0\}$. Suppose that $\lambda = -1$ is an eigenvalue of $C - \lambda D$ and let x be an associated eigenvector. Then the nonzero matrix $R_0 = xx^T D^T$ satisfies

$$\mathbf{S}(R_0) = Cxx^T D^T + Dxx^T D^T = Cxx^T D^T - Cxx^T D^T = 0,$$

i.e., $R_0 \in \mathrm{kernel}(\mathbf{S})$. Now, suppose that $\lambda \neq -1$ and $1/\lambda$ are eigenvalues of $C - \lambda D$ and let x, y be corresponding eigenvectors such that x, y are linearly independent (for $\lambda = 1$ this is only possible if λ has geometric multiplicity at least 2). If $\lambda \neq 0$, the nonzero matrix $R_1 = xy^T D^T - yx^T C^T$ satisfies

$$\begin{aligned} \mathbf{S}(R_1) &= Cxy^T D^T - Cyx^T C^T + Dyx^T D^T - Cxy^T D^T \\ &= -\frac{1}{\lambda} Dyx^T C^T + \frac{1}{\lambda} Dyx^T C^T = 0. \end{aligned}$$

Analogously for $\lambda = 0$, the matrix $R_2 = xy^T C^T - yx^T D^T$ is nonzero and satisfies $\mathbf{S}(R_2) = 0$. It remains to show that $\mathrm{kernel}(\mathbf{S}) \neq \{0\}$ holds if $\lambda = 1$ is an eigenvalue of $C - \lambda D$ with algebraic multiplicity at least 2 but with geometric multiplicity 1. This is, however, an immediate consequence of the fact that $\mathbf{S}$ cannot be nonsingular at isolated points. Hence, if one of the two conditions of Lemma 5.10 is violated then (5.10) is not uniquely solvable, which concludes the proof. $\square$