Generalized Eigenvalue Problems with Prespecified Eigenvalues

Daniel Kressner* Emre Mengi† Ivica Nakic‡ Ninoslav Truhar§

July 30, 2010

Abstract

We consider the distance from a pencil $A - \lambda B$ (square or non-square) to the nearest pencil $(A + \Delta A_*) - \lambda B$ in 2-norm that has the prespecified eigenvalues $\lambda_1, \ldots, \lambda_\ell$ with algebraic multiplicities summing up to r or greater. A singular value optimization characterization is derived for this problem under mild linear independence and multiplicity assumptions. The corollaries of the singular value optimization characterization are significant. First this provides a singular value formula to determine the nearest pencil whose eigenvalues lie in a compact region in the complex plane. Secondly this partially solves the problem posed in [4] regarding the distance from a non-square $n \times m$ pencil with $n < m$ to the nearest pencil with n eigenvalues. The derived singular value optimization problems are solved by means of BFGS and Lipschitz-based global optimization algorithms.

Key words. Pencils, eigenvalues, optimization of singular values, inverse eigenvalue problems, Lipschitz continuous

AMS subject classifications. 65F15, 65F18, 90C26, 90C56

1 Introduction

Given a pencil $A - \lambda B$ where $A, B \in \mathbb{C}^{n \times m}$ with $n \leq m$. We call a scalar $\tilde{\lambda}$ an *eigenvalue* if there exists a nonzero vector $\tilde{v} \in \mathbb{C}^n$ such that

$$\tilde{v}^*(A - \tilde{\lambda}B) = 0. \tag{1}$$

The vector $\tilde{v}$ is said to be a *(left) eigenvector* associated with $\tilde{\lambda}$ and the pair $(\tilde{\lambda}, \tilde{v})$ is said to be an *eigenpair* of the pencil. Throughout this text we will always assume that B has full row-rank.

*Seminar für angewandte Mathematik, HG G 58.1 Rämistrasse 101, 8092 Zürich, Switzerland (daniel.kressner@sam.math.ethz.ch).

†Department of Mathematics, Koç University, Rumelifeneri Yolu, 34450 Sarıyer-İstanbul, Turkey (emengi@ku.edu.tr).

‡Department of Mathematics, University of Zagreb, Bijenička 30, 10000 Zagreb, Croatia (nakic@math.hr).

§Department of Mathematics, University of Osijek, Trg Ljudevita Gaja 6, HR-31 000 Osijek, Croatia (ntruhar@mathos.hr).

In the square $n \times n$ case the pencil $A - \lambda B$ has exactly n eigenvalues counting their multiplicities. In the non-square case the pencil can have at most n eigenvalues. It is possible that $A - \lambda B$ has full row-rank for all λ meaning that the pencil has no eigenvalues at all. Indeed this is generically the case, since a necessary condition for the satisfaction of (1) is that $m!/\left((m-n)!n!\right)$ polynomials (each corresponding to the determinant of a pencil obtained by picking n columns of $A - \lambda B$ out of m columns) have a common root.

The contribution of this work is two-fold. First given a set of complex scalars $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_\ell$ and an integer r. We derive a singular value optimization characterization for the 2-norm of the smallest perturbation ΔA_* so that the pencil $(A + \Delta A_*) - \lambda B$ has the eigenvalues $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_\ell$ with the sum of their algebraic multiplicities at least equal to r, *i.e.* for the quantity

$$\tau_r^{\tilde{\Lambda}}(A,B) := \inf\{\|\Delta A\|_2 : \tilde{\lambda}_1, \ldots, \tilde{\lambda}_\ell \text{ are eigenvalues of } (A+\Delta A) - \lambda B \text{ with algebraic multiplicities summing up to } r \text{ or greater}\}$$

where $\tilde{\Lambda} = \{\tilde{\lambda}_1, \tilde{\lambda}_2, \ldots, \tilde{\lambda}_\ell\}$. In control theory it is often of interest to design a system whose eigenvalues lie in a certain region in the complex plane. At the very least one desires a control system to be stable. For a stable linear system the distance to a nearest unstable system is introduced in [10] and consequently a singular value optimization characterization is provided. A corollary of the first contribution addresses the converse; for an unstable linear system determine a stable system as close to the original system as possible. Notice that this problem is intrinsically harder than the distance to instability. For the distance to instability it suffices to perturb the system so that *one of the eigenvalues* is in the undesired region. On the other hand to make an unstable system stable one needs to perturb the system so that *all eigenvalues* lie in the region of stability.

The second contribution concerns a question posed in [4]; *for the non-square pencil $A - \lambda B$ what is the smallest perturbation to A and B so that the perturbed pencil has exactly n eigenvalues?* Here we allow perturbations to A only, however we derive a singular value optimization characterization for the more general problem of finding the smallest perturbation to A so that $(A + \Delta A) - \lambda B$ has $r \leq n$ eigenvalues. Specifically given a non-square $n \times m$ pencil $A - \lambda B$ with $m > n$ the quantity for which a singular value characterization is sought is

$$\tau_r(A,B) := \inf\{\|\Delta A\|_2 : (A+\Delta A) - \lambda B \text{ has } r \text{ eigenvalues }\}.$$

In the definition above multiple eigenvalues are counted as many times as their algebraic multiplicities. Suppose $(A - \lambda B)$ is rank deficient at $\tilde{\lambda}_k$ and the algebraic multiplicity of $\tilde{\lambda}_k$ is m_k for $k = 1, \ldots, \ell$. Then the pencil has $\sum_{k=1}^{\ell} m_k$ eigenvalues.

The outline of this paper is as follows. In the next section we remind how one can naturally define the algebraic multiplicity of an eigenvalue for a rectangular pencil. Specifically we make a connection between the eigenvalues of a non-square pencil and the eigenvalues of a relevant matrix. In §3 and §4 we derive rank characterizations for a matrix and a pencil to have the eigenvalues $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_\ell$ with the sum of their algebraic multiplicities equal to r or greater. The rank characterizations depend on the matrix and the pencil linearly, which is crucial for the derivation of the singular value characterizations. In §5 and §6 we deduce the singular value characterizations for $\tau_r^{\tilde{\Lambda}}(A,B)$ and $\tau_r(A,B)$ exploiting the linear rank characterizations. Finally in §7 a numerical technique to solve the singular value optimization problems is briefly outlined and in §8 the numerical technique is applied to examples. We do not claim that the

numerical technique outlined here is as efficient as it could be, neither do we claim that it is reliable. Our mere motivation in describing a numerical technique is to illustrate that the singular value characterizations facilitate the computation of $\tau_r^{\tilde{\Lambda}}(A,B)$ and $\tau_r(A,B)$.

2 Algebraic multiplicity of an eigenvalue for a non-square pencil

As it is well-known for a square pencil the algebraic multiplicity of an eigenvalue can exceed its geometric multiplicity (the dimension of the eigenspace). The definition of the geometric multiplicity of an eigenvalue naturally extends to the non-square case. It can be defined as the dimension of the subspace of left eigenvectors associated with the eigenvalue. However, the definition of the algebraic multiplicity of an eigenvalue for a non-square pencil is more subtle. It appears natural that it must involve the Kronecker canonical form for a rectangular pencil[1]. Kronecker canonical form is rather technical. Below we provide an equivalent simpler definition in terms of characteristic polynomials of square pencils.

Since the matrix B is assumed to have full-rank, the Kronecker canonical form of the pencil $A-\lambda B$ must be of the form [6]

$$T_L(A-\lambda B)T_R = \operatorname{diag}\left(L_{\epsilon_1}(\lambda),\ldots,L_{\epsilon_\mu}(\lambda),J_{\lambda_1}^{(1)}(\lambda),\ldots,J_{\lambda_1}^{(r_1)}(\lambda),\ldots,J_{\lambda_\nu}^{(1)}(\lambda),\ldots,J_{\lambda_\nu}^{(r_\nu)}(\lambda)\right) \tag{2}$$

Above $T_L \in \mathbb{C}^{n\times n}$ and $T_R \in \mathbb{C}^{m\times m}$ are non-singular matrices. The rectangular fat and short matrices $L_{\epsilon_j}(\lambda)$ are the elementary Kronecker blocks that do not have any eigenvalues. The square matrices $J_{\lambda_j}^{(i)}(\lambda)$ for $i=1,\ldots,r_j$ denote the Jordan blocks associated with the eigenvalue λ_j of the pencil $A-\lambda B$. The algebraic multiplicity of the eigenvalue λ_j of the pencil $A-\lambda B$ can be defined as the sum of the sizes of the Jordan blocks $J_{\lambda_j}^{(i)}(\lambda)$ for $i=1,\ldots,r_j$.

Following the conventions for the square case we now provide an equivalent definition for the algebraic multiplicity of an eigenvalue in terms of the multiplicity of the associated root of a characteristic polynomial. Let us denote the characteristic polynomial associated with the Jordan blocks $J_{\lambda_j}^{(i)}(\lambda)$ for $i=1,\ldots,r_j$ by $p_j(\lambda)$ in the Kronecker canonical form (2). If we append rows to the Kronecker form so that it becomes square, the determinant of the resulting square pencil would be either $p(\lambda)=0$ or $p(\lambda)=q(\lambda)\prod_{j=1}^{\nu}p_j(\lambda)$ for some polynomial $q(\lambda)$. This implies that the eigenvalue λ_j is a root of $p(\lambda)$ of multiplicity m_j at least unless $p(\lambda)=0$, where m_j is the sum of the sizes of blocks $J_{\lambda_j}^{(i)}(\lambda)$ for $i=1,\ldots,r_j$. More formally consider the polynomial

$$p_{(A,B,C)}(\lambda) = \det\left(\begin{bmatrix} A-\lambda B \\ C \end{bmatrix}\right)$$

where $C\in\mathbb{C}^{(m-n)\times m}$ and denote the multiplicity of $\tilde{\lambda}$ as a root of $p_{(A,B,C)}(\lambda)$ by $mult\left(p_{(A,B,C)},\tilde{\lambda}\right)$ unless $p_{(A,B,C)}(\lambda)=0$. Clearly the following definition for the algebraic multiplicity is equivalent to the definition in terms of the Jordan blocks in the Kronecker canonical form.

[1] Thanks to Volker Mehrmann for pointing this out during a relevant talk at the workshop IWASEP 8, that was held in Berlin in 2010

Definition 2.1 (Algebraic Multiplicity of an Eigenvalue for Non-square Pencils)**.** *Let $A-\lambda B$ be a pencil of size $n\times m$ with $m > n$. Then we define the algebraic multiplicity of an eigenvalue $\tilde{\lambda}$ of $A-\lambda B$ as*

$$\text{minimum}\left\{mult\left(p_{(A,B,C)},\tilde{\lambda}\right) : C\in\mathbb{C}^{(m-n)\times m} \;\; s.t. \;\; p_{(A,B,C)}(\lambda)\neq 0\right\}.$$

We remark that for invertible matrices $T_L\in\mathbb{C}^{n\times n}$ and $T_R\in\mathbb{C}^{m\times m}$ the transformation

$$A-\lambda B \longrightarrow T_L(A-\lambda B)T_R$$

preserves the eigenvalues of $A-\lambda B$ and their algebraic multiplicities. This is apparent from the Kronecker canonical form as well as the equivalent definition provided above, *i.e.*

$$\begin{aligned}\det\left(\begin{bmatrix} A-\lambda B \\ C\end{bmatrix}\right)=0 \quad &\Longleftrightarrow \quad \det\left(\begin{bmatrix} T_L & 0\\ 0 & I\end{bmatrix}\begin{bmatrix} A-\lambda B \\ C\end{bmatrix}T_R\right)=0\\ &\Longleftrightarrow \quad \det\left(\begin{bmatrix} T_L(A-\lambda B)T_R \\ CT_R\end{bmatrix}\right)=0.\end{aligned}$$

Now since B has full rank, the matrices T_L and T_R (for instance elementary matrices corresponding to row and column operations on the matrix B) can be chosen so that

$$T_L(A-\lambda B)T_R = \begin{bmatrix} A_1 & A_2\end{bmatrix}-\lambda\begin{bmatrix} I_n & 0\end{bmatrix} \tag{3}$$

where $A_1\in\mathbb{C}^{n\times n}, A_2\in\mathbb{C}^{n\times(m-n)}$.

Next we benefit from the controllability canonical form [3, Theorem 2.12, p.70] to conclude with a neat characterization for the eigenvalues of the pencil $A-\lambda B$. Suppose

$$\text{rank}\left(\begin{bmatrix} A_2 & A_1A_2 & A_1^2A_2 & \dots & A_1^{n-1}A_2\end{bmatrix}\right)=r.$$

Then there exists an invertible transformation $T\in\mathbb{C}^{n\times n}$ such that

$$\tilde{A}_1 = TA_1T^{-1}=\begin{bmatrix}\tilde{A}_{11} & \tilde{A}_{12}\\ 0 & \tilde{A}_{22}\end{bmatrix} \quad \text{and} \quad \tilde{A}_2=TA_2=\begin{bmatrix}\tilde{A}_{21}\\ 0\end{bmatrix}. \tag{4}$$

where $\tilde{A}_{11}\in\mathbb{C}^{r\times r}, \tilde{A}_{12}\in\mathbb{C}^{r\times(n-r)}, \tilde{A}_{22}\in\mathbb{C}^{(n-r)\times(n-r)}$ and $\tilde{A}_{21}\in\mathbb{C}^{r\times(m-n)}$. Furthermore the pencil

$$\begin{bmatrix}\tilde{A}_{11} & \tilde{A}_{21}\end{bmatrix}-\lambda\begin{bmatrix} I_r & 0\end{bmatrix}$$

has full rank for all λ. We call the eigenvalues of $\tilde{A}_{22}$ as the *uncontrollable eigenvalues* of the pair (A_1,A_2). This terminology is motivated by the dynamical system $x'(t)=A_1x(t)+A_2u(t)$ where $x(t)$ and $u(t)$ denote the state and input functions, respectively. Transform the state-space of the system by $\tilde{x}(t)=Tx(t)$. Then the eigenvectors associated with the uncontrollable eigenvalues constitute those states that cannot be reached by the transformed system $\tilde{x}'(t)=\tilde{A}_1\tilde{x}(t)+\tilde{A}_2u(t)$ no matter what $u(t)$ is chosen starting from zero initial conditions.

By further applying similarity transformations defined in terms of the state transformation T to (3) we obtain

$$\begin{aligned}TT_L(A-\lambda B)T_R\begin{bmatrix} T^{-1} & 0\\ 0 & I\end{bmatrix} &= T\left(\begin{bmatrix} A_1 & A_2\end{bmatrix}-\lambda\begin{bmatrix} I_n & 0\end{bmatrix}\right)\begin{bmatrix} T^{-1} & 0\\ 0 & I\end{bmatrix}\\ &= \begin{bmatrix}\tilde{A}_{11} & \tilde{A}_{12} & \tilde{A}_{21}\\ 0 & \tilde{A}_{22} & 0\end{bmatrix}-\lambda\begin{bmatrix} I_r & 0 & 0\\ 0 & I_{n-r} & 0\end{bmatrix}\end{aligned}$$

Now since the pencil $\begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{21} \end{bmatrix} - \lambda \begin{bmatrix} I_r & 0 \end{bmatrix}$ has full rank for all λ, it follows that the eigenvalues of $A - \lambda B$ are precisely the eigenvalues of the matrix $\tilde{A}_{22}$. Therefore the pencil $A - \lambda B$ has exactly $n - r$ eigenvalues given by the uncontrollable eigenvalues of the pair (A_1, A_2) as summarized by the following theorem.

Theorem 2.2 (Nonsquare Eigenvalue Problem and Uncontrollable Eigenvalues)**.** *Let $A-\lambda B$ be an $n \times m$ pencil with $n < m$ and such that B has full rank. Suppose also that $A - \lambda B$ has $n - r$ eigenvalues counting their algebraic multiplicities. Then*

(i) *there exist invertible $\hat{T}_L \in \mathbb{C}^{n\times n}$ and $\hat{T}_R \in \mathbb{C}^{m\times m}$ satisfying*

$$\hat{T}_L(A - \lambda B)\hat{T}_R = \begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} & \tilde{A}_{21} \\ 0 & \tilde{A}_{22} & 0 \end{bmatrix} - \lambda \begin{bmatrix} I_r & 0 & 0 \\ 0 & I_{n-r} & 0 \end{bmatrix}$$

where $\tilde{A}_{11} \in \mathbb{C}^{r\times r}, \tilde{A}_{12} \in \mathbb{C}^{r\times(n-r)}, \tilde{A}_{22} \in \mathbb{C}^{(n-r)\times(n-r)}$ and $\tilde{A}_{21} \in \mathbb{C}^{r\times(m-n)}$ such that the pencil

$$\begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{21} \end{bmatrix} - \lambda \begin{bmatrix} I_r & 0 \end{bmatrix}$$

has full rank for all λ,

(ii) *furthermore the eigenvalues of $A-\lambda B$ are same as the eigenvalues of $\tilde{A}_{22}$ with the same algebraic multiplicities.*

3 Rank Characterization for Matrices with Prespecified Eigenvalues

We will first derive a singular value characterization for the distance to the nearest pencil with prespecified eigenvalues whose algebraic multiplicities sum up to r. (The quantity $\tau_r(A, B)$ is the infimum of all these distances over all possible eigenvalues.) But the eigenvalues of the pencil $A - \lambda B$ are deduced to be the same as the eigenvalues of a matrix with the same algebraic multiplicities by Theorem 2.2. Therefore the starting point in our derivation is a rank characterization to check whether a matrix has the prespecified eigenvalues with multiplicities summing up to r.

The first characterization below follows from the Jordan factorization of a matrix.

Theorem 3.1 (Rank Characterization)**.** *Let $A \in \mathbb{C}^{k\times k}$, $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_\ell \in \mathbb{C}$ be distinct complex scalars and $r \in \mathbb{Z}^+$. The scalars $\tilde{\lambda}_j$ for $j = 1, \ldots, \ell$ are the eigenvalues of A with algebraic multiplicities summing up to r or greater if and only if the inequality*

$$\operatorname{rank}\left(\prod_{j=1}^{r}(A - \lambda_j I)\right) \leq k - r$$

holds for some $\lambda_j \in \{\tilde{\lambda}_1, \ldots, \tilde{\lambda}_\ell\}$ for $j = 1, \ldots, r$.

The difficulty with the rank characterization above is that it is in terms of a polynomial of A. When analyzing optimal perturbations, a linearized version of the result above facilitates

the derivation by a great deal. Therefore our next task is to linearize the rank characterization. In essence here we generalize Theorem 2.3 in [12] where a linearized rank characterization was derived for only one eigenvalue to have algebraic multiplicity equal to or greater than r. Note that in what follows throughout this text null(C) always denotes the left null-space of C and whenever we refer to the null-space of a matrix, we always mean the left null-space.

Malyshev in [11] exploited the equivalence[2]

$$\begin{gathered}\lambda \textit{ is an eigenvalue of algebraic multiplicity } \geq 2\\ \Longleftrightarrow\\ \operatorname{rank}(A-\lambda I)^2 \leq n-2\\ \Longleftrightarrow\\ \operatorname{rank}\left(\underbrace{\begin{bmatrix} A-\lambda I & 0\\ \gamma I & A-\lambda I\end{bmatrix}}_{M_2^{\Lambda,\Gamma}}\right) \leq 2n-2.\end{gathered}$$

Now λ is an eigenvalue of A if and only if $\operatorname{rank}(A-\lambda I) \leq n-1$ or equivalently the dimension of $\operatorname{null}(A-\lambda I) \geq 1$. When we multiply $(A-\lambda I)$ with itself, there are potentially vectors (if λ is a defective eigenvalue) that are in $\operatorname{null}(A-\lambda I)^2 \setminus \operatorname{null}(A-\lambda I)$. The second block row of $M_2^{\Lambda,\Gamma}$ takes these vectors contained in $\operatorname{null}(A-\lambda I)^2 \setminus \operatorname{null}(A-\lambda I)$ into account.

More generally we will establish that the null spaces of $\left(\prod_{j=1}^{r}(A-\lambda_j I)\right)$ and the linearized version

$$M_r^{\Lambda,\Gamma}(A) := \begin{bmatrix} A-\lambda_1 I & 0 & & & 0\\ \gamma_{21} I & A-\lambda_2 I & & & 0\\ & & \ddots & & \\ & & & A-\lambda_{r-1} I & 0\\ \gamma_{r1} I & \gamma_{r2} I & & \gamma_{r(r-1)} I & A-\lambda_r I\end{bmatrix} \tag{5}$$

are of same dimension where

$$\begin{aligned}\Lambda &= \begin{bmatrix} \lambda_1 & \lambda_2 & \dots & \lambda_r\end{bmatrix}^T \in \mathbb{C}^r\\ \Gamma &= \begin{bmatrix} \gamma_{21} & \gamma_{31} & \gamma_{32} & \dots & \gamma_{r(r-1)}\end{bmatrix}^T \in \mathbb{C}^{r(r-1)/2}.\end{aligned}$$

The addition of the ith block row of $M_r^{\Lambda,\Gamma}$ contributes to the left null space of $M_r^{\Lambda,\Gamma}(A)$ with the vectors that lies in the set $\operatorname{null}\left(\prod_{j=1}^{i}(A-\lambda_j I)\right) \setminus \operatorname{null}\left(\prod_{j=1}^{i-1}(A-\lambda_j I)\right)$. For instance if λ_i is already repeated m_i times in the set $\{\lambda_1,\dots,\lambda_{i-1}\}$, the dimension of the null space increases by $\dim\left(\operatorname{null}(A-\lambda_i)^{m_i+1}\right) - \dim\left(\operatorname{null}(A-\lambda_i)^{m_i}\right)$ due to the ith block row.

Theorem 3.2 (Null Space of the Linearized Matrix)**.** *Let $A \in \mathbb{C}^{k\times k}$ and $\Lambda = \begin{bmatrix} \lambda_1 & \lambda_2 & \dots & \lambda_r\end{bmatrix}^T \in \mathbb{C}^r$. Then*

$$\dim\left(\operatorname{null}\left(\prod_{j=1}^{r}(A-\lambda_j I)\right)\right) = \dim\left(\operatorname{null}\left(M_r^{\Lambda,\Gamma}(A)\right)\right)$$

[2]Indeed the block upper triangular variant of the linearized matrix $M_2^{\Lambda,\Gamma}$ was used in [11].

for any $\Gamma \in \mathbb{C}^{r(r-1)/2}$ *with nonzero components where* $M_r^{\Lambda,\Gamma}(A)$ *is as defined by equation (5).*

Proof. The proof is by induction on r. As the base case for $r = 1$ we have $M_1^{\Lambda,\Gamma}(A) = (A - \lambda_1 I)$ and the dimensions of the null spaces are equal trivially.

Now as the inductive hypothesis suppose that for $r = i$

$$\dim\left(\text{null}\left(\prod_{j=1}^{i}(A-\lambda_j I)\right)\right) = \dim\left(\text{null}\left(M_i^{\Lambda,\Gamma}(A)\right)\right) = m_s.$$

Suppose that $\dim\left(\text{null}\left(\prod_{j=1}^{i+1}(A-\lambda_j I)\right)\right) = m_s + m_a$. We can split the null space of $\left(\prod_{j=1}^{i+1}(A-\lambda_j I)\right)$ into two subspaces,

(i) $\mathcal{S}_1 = \text{null}\left(\prod_{j=1}^{i}(A-\lambda_j I)\right)$ and

(ii) $\mathcal{S}_2 = \text{null}\left(\prod_{j=1}^{i+1}(A-\lambda_j I)\right) \cap \left(\text{null}\left(\prod_{j=1}^{i}(A-\lambda_j I)\right)\right)^{\perp}$.

These two subspaces are clearly orthogonal complements and together they span null $\left(\prod_{j=1}^{i+1}(A-\lambda_j I)\right)$.

Now let $\{v_\ell : \ell = 1, \ldots, m_s\}$ be a basis for $\mathcal{S}_1$. By the inductive hypothesis there exists a basis $\{\tilde{U}_\ell : \ell = 1, \ldots, m_s\}$ for the left null space of $M_i^{\Lambda,\Gamma}(A)$. Define

$$U_\ell := \begin{bmatrix} \tilde{U}_\ell \\ 0 \end{bmatrix} \in \mathbb{C}^{n(i+1)}.$$

It follows that the set $\{U_\ell : \ell = 1, \ldots, m_s\}$ is also linearly independent and each vector in this set lies in left the null space of $M_{i+1}^{\Lambda,\Gamma}(A)$.

On the other hand suppose that $\{v_\ell : \ell = m_s + 1, \ldots, m_s + m_a\}$ is a basis for $\mathcal{S}_2$. Now we describe a vector U_ℓ that is contained in the left null space of $M_{i+1}^{\Lambda,\Gamma}(A)$ corresponding to each vector v_ℓ for $\ell = m_s + 1, \ldots, m_s + m_a$ such that $\{U_\ell : \ell = 1, \ldots, m_s + m_a\}$ is linearly independent. For each $\ell = m_s + 1, \ldots, m_s + m_a$ define the sequence of vectors $v_\ell^{(1)} := v_\ell$ and $v_\ell^{(t+1)}$ is such that

$$\left(v_\ell^{(t+1)}\right)^* := \left(v_\ell^{(t)}\right)^* (A - \lambda_k I)$$

for $t = 1, \ldots, i$. Note that

$$\left(v_\ell^{(i+1)}\right)^* (A - \lambda_{i+1} I) = \left(v_\ell^* \prod_{j=1}^{i}(A-\lambda_j I)\right)(A - \lambda_{i+1} I) = 0$$

because v_ℓ is contained in a basis for $\mathcal{S}_2$. Furthermore the set $\{v_\ell^{(i+1)} : \ell = m_s+1, \ldots, m_s+m_a\}$

is linearly independent, because

$$\begin{aligned}
0 &= \sum_{\ell=m_s+1}^{m_s+m_a} \alpha_\ell \left(v_\ell^{(i+1)}\right)^* = \sum_{\ell=m_s+1}^{m_s+m_a} \alpha_\ell \left(v_\ell^* \prod_{j=1}^{i}(A-\lambda_j I)\right) \\
\implies & \left(since \sum_{\ell=m_s+1}^{m_s+m_a} \alpha_\ell v_\ell \in \mathcal{S}_2\right) \qquad \sum_{\ell=m_s+1}^{m_s+m_a} \alpha_\ell v_\ell = 0 \\
\implies & (since \ \{v_\ell : \ell = m_s+1,\ldots,m_s+m_a\} \ is \ linearly \ independent) \\
& \alpha_\ell = 0 \quad for \ \ell = m_s+1,\ldots,m_s+m_a.
\end{aligned}$$

It can be shown that there exists a vector U_ℓ of the form

$$U_\ell = \left[\begin{array}{cccc} \left(U_\ell^{(1)}\right)^T & \left(U_\ell^{(2)}\right)^T & \ldots & \left(U_\ell^{(i+1)}\right)^T \end{array} \right]^T \tag{6}$$

where $U_\ell^{(i+1)} = v_\ell^{(i+1)}$ and $U_\ell^{(k)} = \tilde{v}_\ell^{(k:i)}$ for $k \le i$ that lies in the null space of $M_{i+1}^{\Lambda,\Gamma}(A)$. Here and in what follows $\tilde{v}_\ell^{(k_1:k_2)}$ for $k_1 \le k_2$ denotes a vector that belongs to span $\left\{v_\ell^{(k_1)}, v_\ell^{(k_1+1)}, \ldots, v_\ell^{(k_2)}\right\}$. Multiplying the vector $(U_\ell)^*$ with the right-most block column of $M_{i+1}^{\Lambda,\Gamma}(A)$ yields

$$\left(v_\ell^{(i+1)}\right)^* (A-\lambda_{i+1}I) = 0.$$

Before proceeding to the other block columns of $M_{i+1}^{\Lambda,\Gamma}(A)$ we first note that there exists a $\tilde{v}_\ell^{(k_1:k_2-1)}$ satisfying

$$\left(\tilde{v}_\ell^{(k_1:k_2-1)}\right)^* (A-\lambda_{k_1}I) = \left(v_\ell^{(k_2)}\right)^* \tag{7}$$

where $k_1 < k_2$. The proof of this fact is also by induction on k_2. When $k_2 = k_1+1$ by definition

$$\left(v_\ell^{(k_1)}\right)^* (A-\lambda_{k_1}I) = \left(v_\ell^{(k_1+1)}\right)^*.$$

Now suppose that for $k_2 = m \ge k_1+1$ there exists a $\tilde{v}_\ell^{(k_1:(m-1))}$ satisfying

$$\left(\tilde{v}_\ell^{(k_1:(m-1))}\right)^* (A-\lambda_{k_1}I) = \left(v_\ell^{(m)}\right)^*. \tag{8}$$

Then

$$\left(v_\ell^{(m+1)}\right)^* = \left(v_\ell^{(m)}\right)^* (A-\lambda_m I) = \left(v_\ell^{(m)}\right)^* (A-\lambda_{k_1}I) + (\lambda_{k_1}-\lambda_m)\left(v_\ell^{(m)}\right)^*$$

and by (8) we obtain

$$\left(v_\ell^{(m+1)}\right)^* = \left(v_\ell^{(m)}\right)^* (A-\lambda_{k_1}I) + (\lambda_{k_1}-\lambda_m)\left(\tilde{v}_\ell^{(k_1:(m-1))}\right)^* (A-\lambda_{k_1}I) = \left(\tilde{v}_\ell^{(k_1:m)}\right)^* (A-\lambda_{k_1}I)$$

which proves (7).

Next we prove that a vector U_ℓ of the form (6) is orthogonal to the first i block columns of $M_{i+1}^{\Lambda,\Gamma}(A)$ by induction exploiting the property (7). As the base case consider the ith block column of $M_{i+1}^{\Lambda,\Gamma}(A)$. Observe that

$$\begin{aligned}\gamma_{(i+1)i}\left(U_\ell^{(i+1)}\right)^* = \gamma_{(i+1)i}\left(v_\ell^{(i+1)}\right)^* &= \left(\gamma_{(i+1)i}v_\ell^{(i)}\right)^*(A-\lambda_i I)\\ &= \left(-U_\ell^{(i)}\right)^*(A-\lambda_i I).\end{aligned}$$

In other words $U_\ell^{(i)}$ can be chosen so that U_ℓ is orthogonal to the ith column block of $M_{i+1}^{\Lambda,\Gamma}(A)$. Now as the inductive hypothesis suppose that $U_\ell^{(k)}$ for $k = m+1,\ldots,(i+1)$ satisfying the form (6) can be chosen so that U_ℓ is orthogonal to the block columns $k = m+1,\ldots,(i+1)$ of $M_{i+1}^{\Lambda,\Gamma}(A)$. Consider the mth block column of $M_{i+1}^{\Lambda,\Gamma}(A)$. We have

$$\sum_{k=m+1}^{i+1}\gamma_{km}\left(U_\ell^{(k)}\right)^* = \sum_{k=m+1}^{i+1}\gamma_{km}\left(\tilde{v}_\ell^{(k:i+1)}\right)^* = \left(\tilde{v}_\ell^{(m+1:i+1)}\right)^*.$$

By property (7) we have $\left(\tilde{v}_\ell^{(m+1:i+1)}\right)^* = \left(\tilde{v}_\ell^{(m:i)}\right)^*(A-\lambda_m I)$ implying

$$\sum_{k=m+1}^{i+1}\gamma_{km}\left(U_\ell^{(k)}\right)^* = \left(\tilde{v}_\ell^{(m:i)}\right)^*(A-\lambda_m I) = \left(-U_\ell^{(m)}\right)^*(A-\lambda_m I).$$

Therefore $U_\ell^{(m)}$ can be chosen so that U_ℓ is orthogonal to the mth block column of $M_{i+1}^{\Lambda,\Gamma}(A)$. By induction we conclude that there exists a U_ℓ of the form (6) lying in the left null space of $M_{i+1}^{\Lambda,\Gamma}(A)$.

The set of vectors $\{U_\ell : \ell = m_s+1,\ldots,m_s+m_a\}$ is linearly independent because of the linear independence of $\{U_\ell^{(i+1)} : \ell = m_s+1,\ldots,m_s+m_a\} = \{v_\ell^{(i+1)} : \ell = m_s+1,\ldots,m_s+m_a\}$. Furthermore since $U_\ell^{(i+1)} = 0$ for $\ell = 1,\ldots,m_s$ it follows that the set $\{U_\ell : \ell = 1,\ldots,m_s+m_a\}$ is also linearly independent.

Finally to complete the proof we need to establish that

$$U \notin \operatorname{span}\{U_\ell : \ell = 1,\ldots,m_s+m_a\} \Longrightarrow U \notin \operatorname{null}\left(M_r^{\Lambda,\Gamma}(A)\right) \tag{9}$$

Now partition any vector $U \in \mathbb{C}^{k(i+1)}$ into $i+1$ block components of equal size of the form

$$U_\ell = \begin{bmatrix} \left(U^{(1)}\right)^T & \left(U^{(2)}\right)^T & \ldots & \left(U^{(i+1)}\right)^T \end{bmatrix}^T.$$

Supposing $U \in \operatorname{null}\left(M_r^{\Lambda,\Gamma}(A)\right)$ yields

$$\begin{gathered}\left(U^{(i+1)}\right)^*(A-\lambda_{i+1}) = 0\\ \Longrightarrow\\ U^{(i+1)} \in \operatorname{null}\left(\prod_{j=1}^{i+1}(A-\lambda_j I)\right)\\ \Longrightarrow\\ U \in \operatorname{span}\{U_\ell : \ell = 1,\ldots,m_s+m_a\},\end{gathered}$$

which shows the validity of (9). We conclude that

$$\operatorname{null}\left(M_r^{\Lambda,\Gamma}(A)\right) = \operatorname{span}\left\{U_\ell : \ell = 1, \ldots, m_s + m_a\right\}$$

and consequently $\dim\left(\operatorname{null}\left(M_r^{\Lambda,\Gamma}(A)\right)\right) = m_s + m_a$ as desired. □

Corollary 3.3 (Linearized Rank Characterization)**.** *Let $A \in \mathbb{C}^{k\times k}$ and $\Lambda = \begin{bmatrix} \lambda_1 & \lambda_2 & \ldots & \lambda_r \end{bmatrix}^T \in \mathbb{C}^r$. Then*

$$\operatorname{rank}\left(\prod_{j=1}^{r}(A - \lambda_j I)\right) \leq k - r \quad \Longleftrightarrow \quad \operatorname{rank}\left(M_r^{\Lambda,\Gamma}(A)\right) \leq kr - r$$

for any Γ with nonzero components where $M_r^{\Lambda,\Gamma}(A)$ is as defined by equation (5).

4 Rank Characterization for Pencils with Prespecified Eigenvalues

Now we generalize the rank characterization for matrices in the previous section to pencils by exploiting Theorem 2.2. Given an $n \times m$ pencil $A - \lambda B$ with $n \leq m$. The next result is in terms of a linearized pencil of the form

$$\mathcal{P}_r^{\Lambda,\Upsilon}(A,B) := \begin{bmatrix} A - \lambda_1 B & 0 & & & 0 \\ \upsilon_{21} B & A - \lambda_2 B & & & 0 \\ & & \ddots & & \\ & & & A - \lambda_j B & 0 \\ \upsilon_{r1} B & \upsilon_{r2} B & & \upsilon_{r(r-1)} B & A - \lambda_r B \end{bmatrix} \tag{10}$$

where

$$\begin{aligned} \Lambda &= \begin{bmatrix} \lambda_1 & \lambda_2 & \ldots & \lambda_r \end{bmatrix}^T \in \mathbb{C}^r, \\ \Upsilon &= \begin{bmatrix} \upsilon_{21} & \upsilon_{31} & \upsilon_{32} & \ldots & \upsilon_{r(r-1)} \end{bmatrix}^T \in \mathbb{C}^{r(r-1)/2}. \end{aligned}$$

Theorem 4.1 (Linearized Rank Characterization for Pencils)**.** *Let $A, B \in \mathbb{C}^{n\times m}$ with $n \leq m$, $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_\ell \in \mathbb{C}$ be distinct scalars and $r \in \mathbb{Z}^+$. Suppose also that B has full rank. Then $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_\ell$ are eigenvalues of $A - \lambda B$ with multiplicities summing up to r or greater if and only the inequality*

$$\operatorname{rank}\left(\mathcal{P}_r^{\Lambda,\Upsilon}(A,B)\right) \leq nr - r$$

holds for some $\lambda_j \in \{\tilde{\lambda}_1, \ldots, \tilde{\lambda}_\ell\}$ for $j = 1, \ldots, r$ and for all $\Upsilon \in \mathbb{C}^{r(r-1)/2}$ with non-zero components.

Proof. Suppose that $A - \lambda B$ has k eigenvalues. Then from Theorem 2.2 there exist invertible matrices T_L and T_R such that

$$T_L(A - \lambda B)T_R = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ 0 & A_{22} & 0 \end{bmatrix} - \lambda \begin{bmatrix} I_{n-k} & 0 & 0 \\ 0 & I_k & 0 \end{bmatrix}$$

where $A_{11} \in \mathbb{C}^{(n-k)\times(n-k)}, A_{12} \in \mathbb{C}^{(n-k)\times k}, A_{22} \in \mathbb{C}^{k\times k}$ and $A_{13} \in \mathbb{C}^{(n-k)\times(m-n)}$. (Note that if the pencil is square, then A_{22} is $n \times n$ and A_{11}, A_{12}, A_{13} disappear.) The pencil $[A_{11} \;\; A_{13}] - \lambda[I_{n-k} \;\; 0]$ has full rank for all λ. Additionally the eigenvalues of $A - \lambda B$ are same as the eigenvalues of A_{22} with the same algebraic multiplicities.

Consider the transformation

$$(I_r \otimes T_L)\mathcal{P}_r^{\Lambda,\Upsilon}(A,B)(I_r \otimes T_R) =$$

$$\begin{bmatrix}
A_{11}-\lambda_1 I_{n-k} & A_{12} & A_{13} & 0 & 0 & 0 & & & 0 \\
0 & A_{22}-\lambda_1 I_k & 0 & 0 & 0 & 0 & & & 0 \\
\upsilon_{21} I_{n-k} & 0 & 0 & A_{11}-\lambda_2 I_{n-k} & A_{12} & A_{13} & & & 0 \\
0 & \upsilon_{21} I_k & 0 & 0 & A_{22}-\lambda_2 I_k & 0 & & & 0 \\
 & & & & & & \ddots & & \\
\upsilon_{r1} I_{n-k} & 0 & 0 & & & & A_{11}-\lambda_r I_{n-k} & A_{12} & A_{13} \\
0 & \upsilon_{r1} I_k & 0 & & & & 0 & A_{22}-\lambda_r I_k & 0
\end{bmatrix}.$$

Now the odd row blocks of $(I_r \otimes T_L)\mathcal{P}_r^{\Lambda,\Upsilon}(A,B)(I_r \otimes T_R)$ have full rank due to the fact that the pencil $[A_{11} \;\; A_{13}] - \lambda[I_{n-k} \;\; 0] = [A_{11} - \lambda I_{n-k} \;\; A_{13}]$ has full rank for all λ. We obtain

$$\begin{aligned}
\dim\left(\text{null}\left(\mathcal{P}_r^{\Lambda,\Upsilon}(A,B)\right)\right) &= \dim\left(\text{null}(I_r \otimes T_L)\mathcal{P}_r^{\Lambda,\Upsilon}(A,B)(I_r \otimes T_R)\right) \\
&= \dim\left(\text{null}\left(M_r^{\Lambda,\Upsilon}(A_{22})\right)\right).
\end{aligned}$$

Now by Theorem 3.1 and 3.3 the pencil $A - \lambda B$ or equivalently the matrix A_{22} has eigenvalues $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_\ell$ with algebraic multiplicities summing up to r or greater if and only if

$$\text{rank}\left(M_r^{\Lambda,\Upsilon}(A_{22})\right) \le rk - r \iff \text{rank}\left(\mathcal{P}_r^{\Lambda,\Upsilon}(A,B)\right) \le nr - r,$$

completing the proof. □

5 A singular value characterization for the nearest pencil with prespecified eigenvalues

Given the set of distinct complex scalars $\tilde{\Lambda} = \{\tilde{\lambda}_1, \tilde{\lambda}_2, \ldots, \tilde{\lambda}_\ell\}$. The purpose of this section is to derive a singular value optimization characterization for the distance from a pencil (square or non-square) $A - \lambda B$ to a nearest pencil which has eigenvalues $\tilde{\lambda}_1, \tilde{\lambda}_2, \ldots, \tilde{\lambda}_\ell$ with algebraic multiplicities summing up to r or greater. The technique utilized here is along the lines of the techniques in [12, 13] and the main result of this section generalizes the singular value optimization characterizations given in [12, 13]. Following elementary result [7, Theorem 2.5.3, p.72] will play a central role in our derivation.

Theorem 5.1. *Let $C \in \mathbb{C}^{k\times k}$. Then*

$$\inf\{\|\Delta C\|_2 : \text{rank}(C + \Delta C) \le k - p\} = \sigma_{k-p+1}(C).$$

A consequence of the rank characterization derived in the previous section, in particular Theorem 4.1, is that for all $\Upsilon \in \mathbb{C}^{r(r-1)/2}$ with non-zero components we have

$$\begin{aligned}
\tau_r^{\tilde{\Lambda}}(A,B) := \quad & \inf\{\|\Delta A\|_2 : \tilde{\lambda}_1, \ldots, \tilde{\lambda}_\ell \text{ are eigenvalues of } (A + \Delta A) - \lambda B \text{ with} \\
& \qquad \text{algebraic multiplicities summing up to } r \text{ or greater}\} \\
= \quad & \inf\{\|\Delta A\|_2 : \exists \Lambda \in \mathbb{C}^r \text{ s.t. } \lambda_j \in \tilde{\Lambda} \;\text{ and }\; \text{rank}\left(\mathcal{P}_r^{\Lambda,\Upsilon}(A + \Delta A, B)\right) \le nr - r\} \\
= \quad & \inf_{\Lambda\in\mathbb{C}^r, \lambda_j \in \tilde{\Lambda}} \;\; \inf_{\Delta A}\{\|\Delta A\|_2 : \text{rank}\left(\mathcal{P}_r^{\Lambda,\Upsilon}(A + \Delta A, B)\right) \le nr - r\}
\end{aligned}$$

Let us define

$$\tilde{\tau}_r^{\Lambda}(A,B) := \inf_{\Delta A}\{\|\Delta A\|_2 : \operatorname{rank}\left(\mathcal{P}_r^{\Lambda,\Upsilon}(A+\Delta A,B)\right) \leq nr-r\}$$

where $\Lambda \in \mathbb{C}^r$ so that

$$\tau_r^{\tilde{\Lambda}}(A,B) = \inf_{\Lambda\in\mathbb{C}^r,\lambda_j\in\tilde{\Lambda}} \tilde{\tau}_r^{\Lambda}(A,B).$$

Now Theorem 5.1 implies that

$$\tilde{\tau}_r^{\Lambda}(A,B) \geq \sigma_{nr-r+1}\left(\mathcal{P}_r^{\Lambda,\Upsilon}(A,B)\right)$$

for all $\Upsilon \in \mathbb{C}^{r(r-1)/2}$, or equivalently

$$\tilde{\tau}_r^{\Lambda}(A,B) \geq \kappa_r^{\Lambda}(A,B) := \sup_{\Upsilon\in\mathbb{C}^{r(r-1)/2}} \sigma_{nr-r+1}\left(\mathcal{P}_r^{\Lambda,\Upsilon}(A,B)\right).$$

Above we can only deduce a lower bound for $\tilde{\tau}_r^{\Lambda}(A,B)$ by a straightforward application of Theorem 5.1, because allowable perturbations to $\mathcal{P}_r^{\Lambda,\Upsilon}(A,B)$ are structured and not arbitrary.

The rest of this section is devoted to the derivation of other direction

$$\tilde{\tau}_r^{\Lambda}(A,B) \leq \kappa_r^{\Lambda}(A,B)$$

establishing the singular value optimization characterization

$$\tilde{\tau}_r^{\Lambda}(A,B) := \sup_{\Upsilon\in\mathbb{C}^{r(r-1)/2}} \sigma_{nr-r+1}\left(\mathcal{P}_r^{\Lambda,\Upsilon}(A,B)\right). \tag{11}$$

The upper bound will be deduced by constructing an optimal perturbation ΔA_* such that

(i) $\|\Delta A_*\| = \kappa_r^{\Lambda}(A,B)$, and

(ii) $\operatorname{rank}\left(\mathcal{P}_r^{\Lambda,\Upsilon}(A+\Delta A_*,B)\right) \leq nr-r \quad \exists \Upsilon \in \mathbb{C}^{r(r-1)/2}$.

Suppose $\Upsilon_* \in \mathbb{C}^{r(r-1)/2}$ is such that

$$\kappa_r^{\Lambda}(A,B) = \sigma_{nr-r+1}\left(\mathcal{P}_r^{\Lambda,\Upsilon_*}(A,B)\right).$$

The optimal perturbation ΔA_* will be defined in terms of the block components of the optimal left singular vector U and right singular vector V satisfying

$$\mathcal{P}_r^{\Lambda,\Upsilon_*}(A,B)\, V = \kappa_r^{\Lambda}(A,B)\, U \;\text{ and }\; U^*\, \mathcal{P}_r^{\Lambda,\Upsilon_*}(A,B) = V^*\, \kappa_r^{\Lambda}(A,B). \tag{12}$$

Now partition the left singular and right singular vectors into r blocks of equal size, *i.e.*

$$U = \begin{bmatrix} U_1^T & U_2^T & \dots & U_r^T \end{bmatrix}^T \text{ and } V = \begin{bmatrix} V_1^T & V_2^T & \dots & V_r^T \end{bmatrix}^T$$

where $U_1,\dots,U_r \in \mathbb{C}^n$ and $V_1,\dots,V_r \in \mathbb{C}^m$. In the next two subsections we will show that the perturbation

$$\Delta A_* := -\kappa_r^{\Lambda}(A,B)\mathcal{U}^+\mathcal{V} \tag{13}$$

with

$$\mathcal{U} = \begin{bmatrix} U_1^* \\ U_2^* \\ \vdots \\ U_r^* \end{bmatrix} \in \mathbb{C}^{r\times n} \text{ and } \mathcal{V} = \begin{bmatrix} V_1^* \\ V_2^* \\ \vdots \\ V_r^* \end{bmatrix} \in \mathbb{C}^{r\times m}$$

satisfies properties **(i)** and **(ii)**. Here $\mathcal{U}^+$ denotes the Moore-Penrose pseudoinverse of $\mathcal{U}$. The optimality of ΔA_* will be established under the assumptions that the optimal singular value $\kappa_r^{\Lambda}(A,B)$ is simple and the set $\{U_1,\ldots,U_r\}$ of block components of optimal left singular vector $\mathcal{U}$ is linearly independent. To state these assumptions neatly we introduce the following terminology.

Definition 5.2 (Multiplicity Qualification)**.** *We say that the multiplicity qualification holds at* (Λ,Υ) *for the pencil* $A-\lambda B$ *if the multiplicity of the singular value* $\sigma_{nr-r+1}\left(\mathcal{P}_r^{\Lambda,\Upsilon}(A,B)\right)$ *is one.*

Definition 5.3 (Linear Independence Qualification)**.** *We say that the linear independence qualification holds at* (Λ,Υ) *for the pencil* $A-\lambda B$ *if the set* $\{\bar{U}_1,\bar{U}_2,\ldots,\bar{U}_r\}$ *is linearly independent where* $\bar{U}_1,\ldots,\bar{U}_r\in\mathbb{C}^n$ *is such that*

$$\bar{U} = \begin{bmatrix} \bar{U}_1^T & \bar{U}_2^T & \ldots & \bar{U}_r^T \end{bmatrix}^T$$

with $\bar{U}$ *denoting a left singular vector associated with* $\sigma_{nr-r+1}\left(\mathcal{P}_r^{\Lambda,\Upsilon}(A,B)\right)$.

5.1 The 2-norm of the optimal perturbation

Throughout this section we assume that the multiplicity qualification holds at the optimal (Λ,Υ_*) for the pencil $A-\lambda B$. In general given an analytic matrix-valued function $\mathcal{A}(\gamma)$. If the multiplicity of the jth largest singular value $\sigma_j\left(\mathcal{A}(\tilde{\gamma})\right)$ is one and the singular value is nonzero, then $\sigma_j\left(\mathcal{A}(\gamma)\right)$ is analytic at $\gamma=\tilde{\gamma}$ with the derivative

$$\frac{d\sigma_j\left(\mathcal{A}(\tilde{\gamma})\right)}{d\gamma} = \mathrm{Re}\left(u_j^*\frac{d\mathcal{A}(\tilde{\gamma})}{d\gamma}v_j\right)$$

where u_j and v_j denote a consistent pair of unit left and right singular vectors associated with $\sigma_j\left(\mathcal{A}(\tilde{\gamma})\right)$. (See [11] for details.) Let us define

$$f(\Upsilon) = \sigma_{nr-r+1}\left(\mathcal{P}_r^{\Lambda,\Upsilon}(A,B)\right)$$

and view f as a map from $\mathbb{R}^{r(r-1)}\to\mathbb{R}$ by associating each complex v_{ij} with its real part $\Re v_{ij}$ and imaginary part $\Im v_{ij}$. Now clearly

$$f(\Upsilon_*)=0 \Longleftrightarrow \sup_{\Upsilon\in\mathbb{C}^{r(r-1)/2}} \sigma_{nr-r+1}\left(\mathcal{P}_r^{\Lambda,\Upsilon}(A,B)\right)=0 \Longleftrightarrow \tau_r^{\tilde{\Lambda}}(A,B)=0.$$

Thus the characterization (11) holds trivially when $f(\Upsilon_*)=0$, so without loss of generality we assume $f(\Upsilon_*)\neq 0$. Because of the multiplicity qualification, we conclude that $f(\Upsilon)$ is analytic at Υ_* with the derivatives

$$\frac{\partial f(\Upsilon_*)}{\partial \Re v_{jk}} = \mathrm{Re}\left(U_j^*BV_k\right) \text{ and } \frac{\partial f(\Upsilon_*)}{\partial \Im v_{jk}} = \mathrm{Re}\left(U_j^*(Bi)V_k\right).$$

Furthermore the fact that Υ_* is a global maximizer of f yields that both of the derivatives above are zero. Consequently we obtain the following result.

Theorem 5.4. *Suppose that the multiplicity qualification holds at* (Λ, Υ_*) *for the pencil* $A - \lambda B$ *and* $f(\Upsilon_*) \neq 0$*. Then*

$$U_j^* B V_k = 0$$

for $j = 2, \ldots, r$ *and* $k = 1, \ldots, j-1$.

Now by exploiting Theorem 5.4 we show that $\mathcal{U}\mathcal{U}^* = \mathcal{V}\mathcal{V}^*$. Geometrically this means that the angle between U_i and U_j is same as the angle between V_i and V_j. In other words the set $\{V_1, \ldots, V_r\}$ can be obtained from $\{U_1, \ldots, U_r\}$ by concatenating the vectors in the latter set by zeros and then rotating them.

Theorem 5.5. *Suppose that the assumptions of Theorem 5.4 hold. Then*

$$\mathcal{U}\mathcal{U}^* = \mathcal{V}\mathcal{V}^*. \tag{14}$$

Proof. The property (14) is equivalent to having

$$U_j^* U_k = V_j^* V_k$$

for all $j = 1, \ldots, r$ and $k = 1, \ldots, j$. For such a pair j, k from equation (12) we have

$$U_j^*(A - \lambda_j B) + U_{j+1}^* B \upsilon_{(j+1)j} + U_{j+2}^* B \upsilon_{(j+2)j} + \cdots + U_r^* B \upsilon_{rj} = \kappa_r^{\Lambda_*}(A, B) V_j^* \tag{15}$$

and

$$B\upsilon_{k1} V_1 + B\upsilon_{k2} V_2 + \cdots + B\upsilon_{k(k-1)} V_{k-1} + (A - \lambda_k B) V_k = \kappa_r^{\Lambda_*}(A, B) U_k. \tag{16}$$

By multiplying equation (16) by U_j^* from left we obtain

$$U_j^* B\upsilon_{k1} V_1 + \cdots + U_j^* B\upsilon_{k(k-1)} V_{k-1} + U_j^*(A - \lambda_k B) V_k = U_j^*(A - \lambda_k B) V_k = \kappa_r^{\Lambda_*}(A, B) U_j^* U_k$$

where the first equality follows from Theorem 5.4. Noting that $U_j^*(\lambda_j - \lambda_k) B V_k = 0$ (by Theorem 5.4 if $k < j$) the equation above can be rewritten as

$$U_j^*(A - \lambda_j B) V_k = \kappa_r^{\Lambda_*}(A, B) U_j^* U_k.$$

Now use equation (15) to substitute for $U_j^*(A - \lambda_j B)$ in the last equation, which yields

$$\begin{aligned}(\kappa_r^{\Lambda_*}(A, B) V_j^* - U_{j+1}^* B \upsilon_{(j+1)j} - U_{j+2}^* B \upsilon_{(j+2)j} - \cdots - U_r^* B \upsilon_{rj}) V_k = \\ \kappa_r^{\Lambda_*}(A, B) V_j^* V_k = \quad \kappa_r^{\Lambda_*}(A, B) U_j^* U_k.\end{aligned}$$

Above the first equality is again due to Theorem 5.4. We conclude that $V_j^* V_k = U_j^* U_k$ as desired, since $\kappa_r^{\Lambda_*}(A, B) = f(\Upsilon_*) \neq 0$ by assumption. □

Consider multiplying a vector with $\mathcal{U}^+\mathcal{V}$ from left. The transformation $w^*(\mathcal{U}^+\mathcal{V}) = w^*(\mathcal{U}^+\mathcal{U}\mathcal{U}^+\mathcal{V})$ first orthogonally projects the row vector w^* onto the row-space of $\mathcal{U}$. This is followed by a change of basis from the row-space of $\mathcal{U}$ to the row-space of $\mathcal{V}$. Orthogonal projection cannot increase the length of the vector, and because of property (14) the change of basis does not affect the length. Therefore $\|\mathcal{U}^+\mathcal{V}\| = 1$. A formal proof of the fact that the property (14) implies $\|\mathcal{U}^+\mathcal{V}\| = 1$ was given in [11, Lemma 2] and [12, Theorem 2.5].

Theorem 5.6. *Suppose that the assumptions of Theorem 5.4 hold. Then*

$$\|\Delta A_*\| = \kappa_r^{\Lambda_*}(A, B).$$

5.2 Satisfaction of the rank condition by the optimally perturbed pencil

Now we suppose that the linear independence qualification holds at (Λ, Υ_*) for the pencil $A - \lambda B$, that is the set $\{U_1, \ldots, U_r\}$ is linearly independent. Next we establish that rank $\left(\mathcal{P}_r^{\Lambda,\Upsilon_*}(A + \Delta A_*, B)\right) \leq nr - r$. We note that the linear independence qualification implies $\mathcal{U}\mathcal{U}^+ = I_r$. In particular $U_j^*\mathcal{U}^+ = e_j^*$ where e_j denotes the jth column of the identity matrix of size r.

Lemma 5.7. *Suppose that the linear independence qualification holds at (Λ, Υ_*) for the pencil $A-\lambda B$. Also suppose that Λ consists of distinct components. Given the scalars $\beta_k, \ldots, \beta_r \in \mathbb{C}$ and a pair of positive integers j, k such that $r \geq k > j$. There exists a $U_{k:r} \in \mathbb{C}^n$ of the form*

$$U_{k:r} = \alpha_k U_k + \alpha_{k+1} U_{k+1} + \cdots + \alpha_r U_r$$

for some $\alpha_k, \ldots, \alpha_r \in \mathbb{C}$ satisfying

$$(U_{k:r})^* (A - \lambda_j B + \Delta A_*) = \beta_k U_k^* B + \beta_{k+1} U_{k+1}^* B + \cdots + \beta_r U_r^* B.$$

Proof. The proof is by induction on k. As the base case suppose $k = r$. Then

$$\begin{aligned} \alpha_r U_r^*(A - \lambda_j B + \Delta A_*) = &\quad \alpha_r U_r^*(A - \lambda_r B + \Delta A_*) + \alpha_r U_r^*(\lambda_r - \lambda_j)B \\ = &\quad \alpha_r U_r^*(\lambda_r - \lambda_j)B. \end{aligned}$$

Hence choose $\alpha_r := \frac{\beta_r}{\lambda_r - \lambda_j}$ to obtain

$$\alpha_r U_r^*(A - \lambda_j B + \Delta A_*) = \beta_r U_r^* B.$$

Suppose, as the inductive hypothesis, that for $k = \ell + 1 > j + 1$ there exists a $U_{\ell+1:r}$ of the form

$$U_{\ell+1:r} = \alpha_{\ell+1} U_{\ell+1} + \alpha_{\ell+2} U_{\ell+2} + \cdots + \alpha_r U_r$$

such that

$$(U_{\ell+1:r})^* (A - \lambda_j B + \Delta A_*) = \tilde{\beta}_{\ell+1} U_{\ell+1}^* B + \tilde{\beta}_{\ell+2} U_{\ell+2}^* B + \cdots + \tilde{\beta}_r U_r^* B.$$

for every $\tilde{\beta}_{\ell+1}, \tilde{\beta}_{\ell+2}, \ldots, \tilde{\beta}_r \in \mathbb{C}$. Now

$$\begin{aligned} U_{\ell:r}^*(A - \lambda_j B + \Delta A_*) = &\quad \alpha_\ell U_\ell^*(A - \lambda_j B + \Delta A_*) + U_{\ell+1:r}^*(A - \lambda_j B + \Delta A_*) \\ = &\quad \alpha_\ell U_\ell^*(A - \lambda_j B + \Delta A_*) + \tilde{\beta}_{\ell+1} U_{\ell+1}^* B + \tilde{\beta}_{\ell+2} U_{\ell+2}^* B + \cdots + \tilde{\beta}_r U_r^* B. \end{aligned}$$

where

$$\begin{aligned} \alpha_\ell U_\ell^*(A - \lambda_j B + \Delta A_*) = &\quad \alpha_\ell U_\ell^*(A - \lambda_\ell B + \Delta A_*) + \alpha_\ell U_\ell^*(\lambda_\ell - \lambda_j)B \\ = &\quad \alpha_\ell \gamma_{(\ell+1)\ell} U_{\ell+1}^* B + \cdots + \alpha_\ell \gamma_{r\ell} U_r^* B + \alpha_\ell U_\ell^*(\lambda_\ell - \lambda_j)B. \end{aligned}$$

Given any $\beta_\ell, \ldots, \beta_r$. We can choose $\alpha_\ell := \frac{\beta_\ell}{\lambda_\ell - \lambda_j}$ and $\alpha_{\ell+1}, \ldots, \alpha_r$ such that $\tilde{\beta}_m = \beta_m - \alpha_\ell \gamma_{m\ell}$ for $m = \ell + 1, \ldots, r$ which yields

$$(U_{\ell:r})^* (A - \lambda_j B + \Delta A_*) = \beta_\ell U_\ell^* B + \beta_{\ell+1} U_{\ell+1}^* B + \cdots + \beta_r U_r^* B$$

as desired.

□

Theorem 5.8. *Suppose that the linear independence qualification holds at* (Λ, Υ_*) *for the pencil* $A - \lambda B$*. Then we have*

$$\operatorname{rank}\left(\mathcal{P}_r^{\Lambda,\Upsilon_*}(A + \Delta A_*, B)\right) \leq nr - r. \tag{17}$$

Proof. First notice that from the right-hand equality in (12) we have

$$\begin{aligned} U_j^*(A - \lambda_j B + \Delta A_*) + U_{j+1}^* \upsilon_{(j+1)j} B + \cdots + U_r^* \upsilon_{rj} B = \\ -V_j^* \kappa_r^{\Lambda_*}(A, B) + U_j^*(A - \lambda_j B) + U_{j+1}^* \upsilon_{(j+1)j} B + \cdots + U_r^* \upsilon_{rj} B = \quad 0 \end{aligned}$$

implying

$$U^* \ \mathcal{P}_r^{\Lambda,\Upsilon_*}(A + \Delta A_*, B) = 0.$$

Let us first assume that the components of Λ are distinct. We next show that, in addition to U, there are vectors lying in the left null space of $\mathcal{P}_r^{\Lambda,\Upsilon_*}(A + \Delta A_*, B)$ of the form

$$U^{(j)} = \left[\begin{array}{ccccccccc} \left(U_1^{(j)}\right)^T & \left(U_2^{(j)}\right)^T & \ldots & \left(U_{j-1}^{(j)}\right)^T & \left(U_j^{(j)}\right)^T & 0 & \ldots & 0 \end{array} \right]^T \in \mathbb{C}^{nr} \tag{18}$$

for $j = 1, \ldots, r-1$ with $U_\ell^{(j)} \in \operatorname{span}\{U_\ell, \ldots, U_r\}$ and $U_j^{(j)} \neq 0$ is . Clearly the set

$$\{U^{(1)}, U^{(2)}, \ldots, U^{(r-1)}, U\}$$

is linearly independent. Therefore the existence of such $U^{(j)}$ that are contained in the left null space of $\mathcal{P}_r^{\Lambda,\Upsilon_*}(A + \Delta A_*, B)$ implies (17).

We prove that a vector $U^{(j)}$ of the form (18) is contained in the left null space of $\mathcal{P}_r^{\Lambda,\Upsilon_*}(A + \Delta A_*, B)$ by induction. Clearly

$$U_j^*(A - \lambda_j B + \Delta A_*) = -U_{j+1}^* \upsilon_{(j+1)j} B - \cdots - U_r^* \upsilon_{rj} B$$

Now by Lemma 5.7 there exists a $U_{j+1:r} \in \operatorname{span}\{U_{j+1}, \ldots, U_r\}$ such that

$$U_{j+1:r}^*(A - \lambda_j B + \Delta A_*) = U_{j+1}^* \upsilon_{(j+1)j} B + \cdots + U_r^* \upsilon_{rj} B.$$

Defining $U_j^{(j)} := U_j + U_{j+1:r} \neq 0$ we have

$$\left(U_j^{(j)}\right)^* (A - \lambda_j B + \Delta A_*) = 0.$$

Suppose that $U_k^{(j)}$ for $k = \ell + 1, \ldots, j$ can be chosen so that $U^{(j)}$ is orthogonal to the column blocks $k = \ell + 1, \ldots, j$ of $\mathcal{P}_r^{\Lambda,\Upsilon_*}(A + \Delta A_*, B)$. Now we construct $U_\ell^{(j)}$ exploiting the structure of the ℓth column block of $\mathcal{P}_r^{\Lambda,\Upsilon_*}(A + \Delta A_*, B)$. First multiply the vector

$$\left[\begin{array}{ccccccccc} 0 & 0 & \ldots & 0 & U_\ell^* & \left(U_{\ell+1}^{(j)}\right)^* & \ldots & \left(U_j^{(j)}\right)^* & 0 & \ldots & 0 \end{array} \right]$$

with the ℓth block column of $\mathcal{P}_r^{\Lambda,\Upsilon_*}(A + \Delta A_*, B)$ resulting

$$\begin{aligned} U_\ell^*(A - \lambda_\ell B + \Delta A_*) + \left(U_{\ell+1}^{(j)}\right)^* \upsilon_{(\ell+1)\ell} B + \cdots + \left(U_j^{(j)}\right)^* \upsilon_{j\ell} B = \\ \left(-U_{\ell+1}^* \upsilon_{(\ell+1)\ell} B - \cdots - U_r^* \upsilon_{r\ell} B\right) + \left(\tilde{\beta}_{\ell+1} U_{\ell+1}^* B + \cdots + \tilde{\beta}_r U_r^* B\right) = \\ \beta_{\ell+1} U_{\ell+1}^* B + \cdots + \beta_r U_r^* B. \end{aligned}$$

for some $\tilde{\beta}_{\ell+1}, \ldots, \tilde{\beta}_r, \beta_{\ell+1}, \ldots, \beta_r \in \mathbb{C}$. Furthermore by Lemma 5.7 there exists a $U_{\ell+1:r} \in \text{span}\{U_{\ell+1}, \ldots, U_r\}$ such that

$$U_{\ell+1:r}^*(A - \lambda_\ell B + \Delta A_*) = -\beta_{\ell+1} U_{\ell+1}^* B - \cdots - \beta_r U_r^* B.$$

Hence defining $U_\ell^{(j)} := U_\ell + U_{\ell+1:r}$ makes $U^{(j)}$ orthogonal to the ℓth block column of $\mathcal{P}_r^{\Lambda,\Upsilon_*}(A + \Delta A_*, B)$. Now the existence of a $U^{(j)}$ of the form (18) lying in the left null space of $\mathcal{P}_r^{\Lambda,\Upsilon_*}(A + \Delta A_*, B)$ follows from induction.

Finally suppose that the components of Λ are not distinct. Then consider any sequence $\{\hat{\Lambda}\}$ in $\mathbb{C}^r$ with distinct components whose limit is Λ. Then the previous argument applies to deduce that

$$\text{rank}\left(\mathcal{P}_r^{\hat{\Lambda},\Upsilon_*(\hat{\Lambda})}(A + \Delta A_*(\hat{\Lambda}), B)\right) \leq nr - r \Longrightarrow \sigma_{nr-r+1}\left(\mathcal{P}_r^{\hat{\Lambda},\Upsilon_*(\hat{\Lambda})}(A + \Delta A_*(\hat{\Lambda}), B)\right) = 0$$

for all $\hat{\Lambda}$ in the sequence where $\Upsilon_*(\hat{\Lambda})$ is such that

$$\kappa_r^{\hat{\Lambda}}(A, B) = \sigma_{nr-r+1}\left(\mathcal{P}_r^{\hat{\Lambda},\Upsilon_*(\hat{\Lambda})}(A, B)\right) \tag{19}$$

and $\Delta A_*(\hat{\Lambda})$ is the optimal perturbation defined in terms of the singular vectors associated with $\sigma_{nr-r+1}\left(\mathcal{P}_r^{\hat{\Lambda},\Upsilon_*(\hat{\Lambda})}(A, B)\right)$ as in (13). Generically $\Upsilon_*(\hat{\Lambda})$ satisfying (19) is unique for any pair (A, B). We conclude from the continuity of singular values (since without loss of generality the sequence $\{\hat{\Lambda}\}$ can be chosen so that $\gamma_*(\hat{\Lambda})$ is unique for all $\hat{\Lambda}$ and therefore $\gamma_*(\hat{\Lambda})$ and $\Delta A_*(\hat{\Lambda})$ vary with respect to $\hat{\Lambda}$ continuously [2, proposition 4.4]) that

$$\sigma_{nr-r+1}\left(\mathcal{P}_r^{\Lambda,\Upsilon_*}(A + \Delta A_*, B)\right) = 0$$

implying

$$\text{rank}\left(\mathcal{P}_r^{\Lambda,\Upsilon_*}(A + \Delta A_*, B)\right) \leq nr - r.$$

□

We conclude this section with a summary of the singular value optimization characterization derived.

Theorem 5.9 (Characterization for a Nearest Pencil with Prespecified Eigenvalues)**.** *Let $A - \lambda B$ be an $n \times m$ pencil with $n \leq m$, $r \in \mathbb{Z}^+$ and $\tilde{\Lambda}$ be a set consisting of at most r complex scalars.*

(i) *Then*

$$\tau_r^{\tilde{\Lambda}}(A, B) = \inf_{\Lambda \in \mathbb{C}^r, \lambda_j \in \tilde{\Lambda}} \sup_{\Upsilon \in \mathbb{C}^{r(r-1)/2}} \sigma_{nr-r+1}\left(\mathcal{P}_r^{\Lambda,\Upsilon}(A, B)\right)$$

provided that the optimization problem on the right is attained at a (Λ_*, Υ_*) where the multiplicity and linear independence qualifications hold.

(ii) *Furthermore a minimal perturbation ΔA_* such that the pencil $(A + \Delta A_*) - \lambda B$ has the eigenvalues $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_\ell$ with algebraic multiplicities summing up to r or greater is given by (13) but with Λ replaced by Λ_*.*

One potential application in control theory is to ensure that the eigenvalues lie in a particular region in the complex plane. Thus let Ω be a compact region in the complex plane and define the distance

$$\tau_r^{\Omega}(A,B) := \quad \inf\{\|\Delta A\|_2 : r \text{ of the eigenvalues of } (A+\Delta A)-\lambda B \text{ belong to } \Omega\}.$$

The previous derivation applies analogously if the finite set of desired eigenvalues $\tilde{\Lambda}$ is replaced by the desired compact region Ω yielding

$$\tau_r^{\Omega}(A,B) = \inf_{\Lambda\in\mathbb{C}^r,\lambda_j\in\Omega} \quad \sup_{\Upsilon\in\mathbb{C}^{r(r-1)/2}} \sigma_{nr-r+1}\left(\mathcal{P}_r^{\Lambda,\Upsilon}(A,B)\right).$$

An optimal perturbation ΔA_* such that the pencil $(A+\Delta A_*)-\lambda B$ has r of its eigenvalues in Ω is once again given by (13).

6 A singular value characterization for the nearest nonsquare pencil with r eigenvalues

The distance $\tau_r(A,B)$ is the infimum of all distances $\tau_r^{\tilde{\Lambda}_r}(A,B)$ over all $\tilde{\Lambda}_r$ consisting of at most r complex scalars . Hence it immediately follows that

$$\begin{aligned}\tau_r(A,B) = &\quad \inf_{\tilde{\Lambda}_r} \tau_r^{\tilde{\Lambda}_r}(A,B)\\ = &\quad \inf_{\tilde{\Lambda}_r} \quad \inf_{\Lambda\in\mathbb{C}^r,\lambda_j\in\tilde{\Lambda}_r} \tilde{\tau}_r^{\Lambda}(A,B)\\ = &\quad \inf_{\Lambda\in\mathbb{C}^r} \tilde{\tau}_r^{\Lambda}(A,B).\end{aligned}$$

and we obtain the following corollary of Theorem 5.9

Corollary 6.1 (Characterization for a Nearest Pencil with r Eigenvalues)**.** *Let $A-\lambda B$ be an $n\times m$ pencil with $n<m$ and $r\in\mathbb{Z}^+$.*

(i) *Then*

$$\tau_r(A,B) = \inf_{\Lambda\in\mathbb{C}^r} \quad \sup_{\Upsilon\in\mathbb{C}^{r(r-1)/2}} \sigma_{nr-r+1}\left(\mathcal{P}_r^{\Lambda,\Upsilon}(A,B)\right)$$

provided that the optimization problem on the right is attained at a (Λ_*,Υ_*) where the multiplicity and linear independence qualifications hold.

(ii) *Furthermore a minimal perturbation ΔA_* such that the pencil $(A+\Delta A_*)-\lambda B$ has r eigenvalues is given by (13) but with Λ replaced by Λ_*.*

Specifically the distance from a nonsquare $n\times m$ pencil $A-\lambda B$ to the nearest pencil with n eigenvalues can be posed as

$$\tau_n(A,B) = \inf_{\Lambda\in\mathbb{C}^n} \quad \sup_{\Upsilon\in\mathbb{C}^{n(n-1)/2}} \sigma_{n^2-n+1}\left(\mathcal{P}_n^{\Lambda,\Upsilon}(A,B)\right)$$

with minimal perturbation ΔA_* of the form (13).

7 Computational issues

A numerical technique that can be used to compute $\tau_r^{\Omega}(A,B)$ and $\tau_r(A,B)$ exploiting the singular value characterizations were described in [12, 13]. In this section for completeness we briefly remind this numerical technique. We can express the distances of interest as

$$\tau_r^{\Omega}(A,B) = \inf_{\Lambda\in\mathbb{C}^r,\ \lambda_j\in\Omega} g(\Lambda) \qquad \text{and} \qquad \tau_r(A,B) = \inf_{\Lambda\in\mathbb{C}^r} g(\Lambda)$$

where $g:\mathbb{C}^r\to\mathbb{R}$ is defined by

$$g(\Lambda) := \sup_{\Upsilon\in\mathbb{C}^{n(n-1)/2}} \sigma_{nr-r+1}\left(\mathcal{P}_r^{\Lambda,\Upsilon}(A,B)\right).$$

The inner maximization problems are solved by BFGS. The rth smallest singular value function above is not differentiable when its multiplicity is greater than one. In practice this is not a major issue for BFGS, as the multiplicity of the rth smallest singular value is generically one for all Υ for any given Λ. If at a local maximizer Υ_* the multiplicity and linear independence qualifications hold, then Υ_* is indeed a global maximizer meaning $g(\Lambda)$ is retrieved. On the other hand if BFGS converges to a point where one or both of the qualifications are violated, it needs to be restarted with a different initial guess. In practice we almost always observe convergence to a global maximizer immediately without any need for a restart of BFGS.

The function $g(\Lambda)$ is non-convex in general, but it is Lipschitz continuous, *i.e.*

$$|g(\Lambda+\Delta\Lambda) - g(\Lambda)| \le \|\Delta\Lambda\|_2.$$

There are various Lipschitz-based global optimization algorithms in the literature stemming mainly from ideas due to Piyavskii-Shubert [14, 15]. The Piyavskii-Shubert algorithm is based on constructing a piecewise linear approximation lying beneath the Lipschitz function. We used a sophisticated variant of the Piyavskii-Shubert algorithm called DIRECT [8] in practice. DIRECT attempts to estimate the Lipschitz constant locally, which can possibly speed-up the convergence.

The main computational cost involved in the numerical optimization of singular values is the retrieval of the rth smallest singular value of $\mathcal{P}_r^{\Lambda,\Upsilon}(A,B)$ at various Λ and Υ. For this purpose we used the direct solvers in LAPACK as we only experimented with small pencils. For medium to large scale pencils one should rather use iterative algorithms such as the Lanczos method.

8 Numerical Experiments

The software is originally implemented in Fortran. It calls routines from LAPACK for singular value computations, the limited memory BFGS routine written by J. Nocedal [9] for inner maximization problems and an implementation of the DIRECT algorithm by Gablonsky [5] for outer Lipschitz-based minimization. A mex interface file is written so that the Fortran code can be reached from Matlab.

It should be noted that the current implementation is not very reliable mainly due to the numerical solution of the outer Lipschitz minimization problem. This implementation is intended only for small pencils (*e.g.* $n, m < 100$) for now.

8.1 Nearest Pencils with Multiple Eigenvalues

As a corollary of Theorem 5.9 it follows that for a square pencil $A - \lambda B$ the nearest pencil with $\tilde{\lambda}$ as a multiple eigenvalue is given by

$$\tau_2^{\tilde{\lambda}}(A, B) = \sup_{\upsilon} \left(\begin{bmatrix} A - \tilde{\lambda}B & 0 \\ \upsilon B & A - \tilde{\lambda}B \end{bmatrix} \right).$$

Therefore the distance from $A - \lambda B$ to the nearest square pencil with a multiple eigenvalue is characterized by

$$\inf_{\lambda \in \mathbb{C}} \sup_{\upsilon} \sigma_{2n-1} \left(\begin{bmatrix} A - \lambda B & 0 \\ \upsilon B & A - \lambda B \end{bmatrix} \right).$$

Specifically we consider the pencil

$$\underbrace{\begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}}_{\tilde{A}} - \lambda \underbrace{\begin{bmatrix} -1 & 2 & 3 \\ 2 & -1 & 2 \\ 4 & 2 & -1 \end{bmatrix}}_{\tilde{B}}. \tag{20}$$

By solving the singular value optimization problem we obtain the distance from this pencil to the nearest pencil with a multiple eigenvalue as 0.59299. Indeed by equation (13) a nearest pencil turns out to be

$$\begin{bmatrix} 1.91465 & -0.57896 & -1.21173 \\ -1.32160 & 1.93256 & -0.57897 \\ -0.72082 & -1.32160 & 1.91466 \end{bmatrix} - \lambda \begin{bmatrix} -1 & 2 & 3 \\ 2 & -1 & 2 \\ 4 & 2 & -1 \end{bmatrix}$$

with the multiple eigenvalue $\lambda_* = -0.85488$

The ϵ-pseudospectrum of $A - \lambda B$ (subject to perturbations to A only) is the set $\Lambda_\epsilon(A, B)$ consisting of eigenvalues of all pencils $(A + \Delta A) - \lambda B$ such that $\|\Delta A\|_2 \leq \epsilon$. This set can equivalently be defined in terms of singular values as

$$\Lambda_\epsilon(A, B) = \{\lambda \in \mathbb{C} : \sigma_{\min}(A - \lambda B) \leq \epsilon\}.$$

It is well-known that the smallest ϵ such that two components of $\Lambda_\epsilon(A, B)$ coalesce is the distance to the nearest pencil with a multiple eigenvalue. (See [1] for the standard eigenvalue problem when $B = I$, which easily extends for arbitrary invertible B.)

For the particular pencil $\tilde{A} - \lambda \tilde{B}$ in (20) the ϵ-pseudospectra are plotted for various ϵ in Figure 1. Two components of the ϵ-pseudospectrum coalesce for $\epsilon = 0.59299$, that is the distance to the nearest pencil with a multiple eigenvalue. Furthermore the point of coalescence is the multiple eigenvalue of the nearest pencil marked by an asterisk in the figure.

8.2 Nearest Rectangular Pencils with at least Two Eigenvalues

Now let us consider the 3×4 pencil

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.1 & 2 & 1 \\ 0 & 0 & 0.3 & 2 \end{bmatrix}}_{\hat{A}} - \lambda \underbrace{\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\hat{B}}.$$

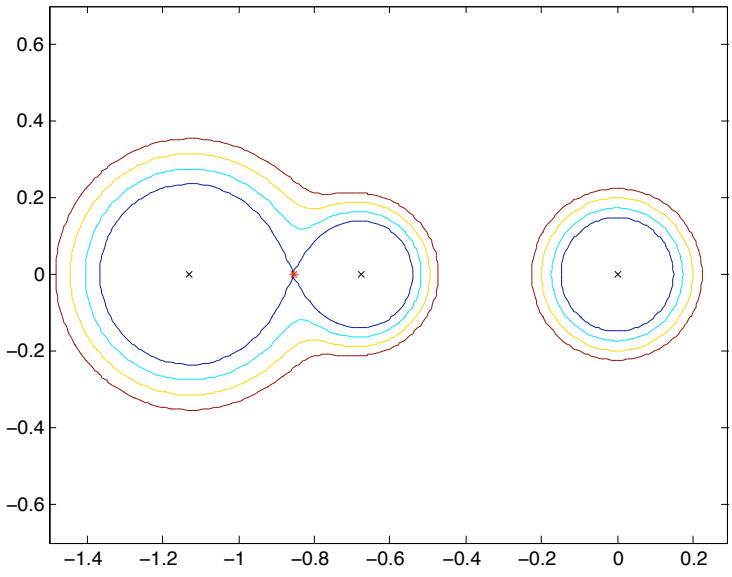

Figure 1: The ϵ-pseudospectra are plotted for various ϵ for the pencil $\tilde{A} - \lambda\tilde{B}$ in (20). The black crosses represent the eigenvalues of $\tilde{A} - \lambda\tilde{B}$. Two components of the ϵ-pseudospectrum coalesce for $\epsilon = 0.59299$, which is the distance to the nearest pencil with a multiple eigenvalue. The asterisk marks the multiple eigenvalue of the nearest pencil.

If the entries $\hat{a}_{ij}$ of $\hat{A}$ are set equal to zero for $j > i$, it is clear that the pencil becomes similar to an elementary Kronecker block of size three. Therefore the pencil has no eigenvalues. However, if the entry $\hat{a}_{22}$ is set equal to zero, then the Kronecker canonical form of the resulting pencil is comprised of an elementary Kronecker block of size one and a block of size two corresponding to finite eigenvalues. Therefore a perturbation with 2-norm equal to 0.1 applied to $\hat{A} - \lambda\hat{B}$ is sufficient to have two eigenvalues.

By corollary 6 the distance to the nearest 3×4 pencil with at least two eigenvalues has the characterization

$$\tau_2(\hat{A}, \hat{B}) = \inf_{\lambda_1, \lambda_2 \in \mathbb{C}} \underbrace{\sup_{\upsilon} \sigma_{2n-1} \left(\begin{bmatrix} \hat{A} - \lambda_1 \hat{B} & 0 \\ \upsilon \hat{B} & \hat{A} - \lambda_2 \hat{B} \end{bmatrix} \right)}_{g(\lambda_1, \lambda_2)}. \tag{21}$$

It turns out that $\tau_2(\hat{A}, \hat{B}) = 0.03927$ based on the numerical solution of the singular value optimization problem above. By equation (13) a nearest pencil is given by

$$\begin{bmatrix} 0.99847 & 0 & 0 & 0.00007 \\ -0.03697 & 0.08698 & 2.00172 & 1.00095 \\ -0.01283 & 0.03689 & 0.30078 & 2.00376 \end{bmatrix} - \lambda \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

with eigenvalues at $\lambda_1 = 2.55144$ and $\lambda_2 = 1.45405$. The level sets of the function $g(\lambda_1, \lambda_2)$ (defined in (21)) over $\mathbb{R}^2$ are illustrated in Figure 2. The optimal eigenvalues $(2.55144, 1.45405)$ are marked by an asterisk in the figure.

9 Concluding Remarks

In this work a singular value characterization is derived for the 2-norm of a smallest perturbation to a square or a non-square pencil $A - \lambda B$ so that the perturbed pencil has the desired set of eigenvalues. The immediate corollaries of this main result are

(i) a singular value characterization for the 2-norm of the smallest perturbation so that the perturbed pencil has a certain number of its eigenvalues in a desired region in the complex plane, and

(ii) a singular value characterization for the 2-norm of the smallest perturbation to a rectangular pencil so that it has a prespecified number of eigenvalues.

We allow perturbations to A only and assume B has full-rank. Derivation of singular value characterizations for the distance problems considered in this paper when A and B are perturbed simultaneously remains open. The development of efficient and reliable computational techniques for the solution of the derived singular value optimization problems is still in progress. As of now the optimization problems can be solved numerically only for small pencils with small number of desired eigenvalues. The main task that needs to be addressed from a computational point of view is a reliable and efficient implementation of the DIRECT algorithm for Lipschitz-based optimization.

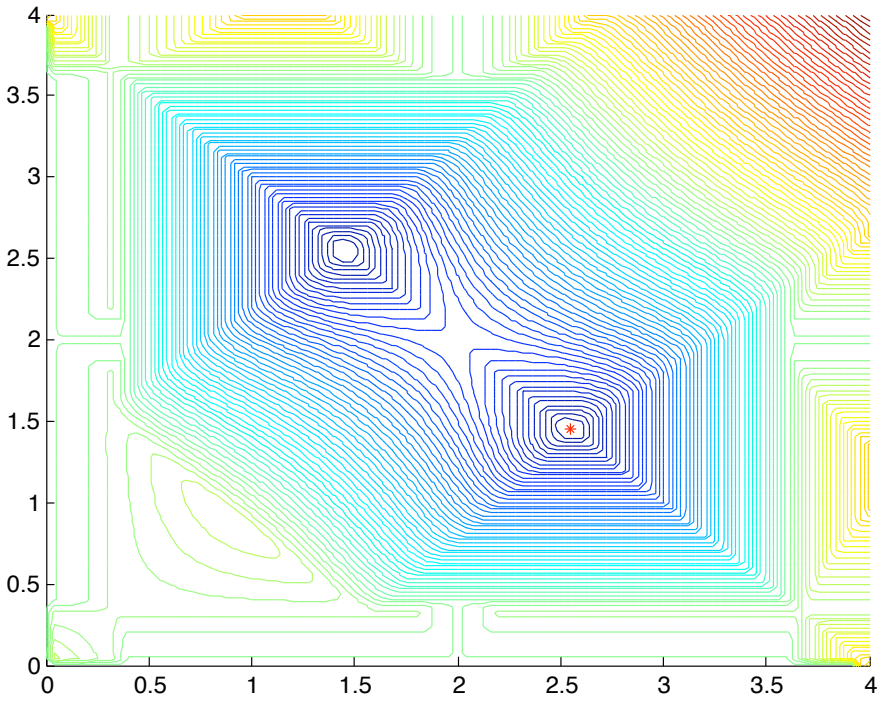

Figure 2: The level-sets of the function $g(\lambda_1, \lambda_2)$ defined in equation (21) over $\mathbb{R}^2$ are plotted. Asterisk marks the numerically computed global minimizer of g, which are the eigenvalues of the nearest pencil with two eigenvalues.

References

[1] R. Alam and S. Bora. On sensitivity of eigenvalues and eigendecompositions of matrices. *Linear Algebra Appl.*, 396:273–301, 2005.

[2] J.F. Bonnans and A. Shapiro. *Perturbation Analysis of Optimization Problems*. Springer-Verlag, New York, 2000.

[3] G. E. Dullerud and F. Paganini. *A Course in Robust Control Theory*. Springer-Verlag, New York, 2000.

[4] G.H. Golub G. Boutry, M. Elad and P. Milanfar. The generalized eigenvalue problem for nonsquare pencils using a minimal perturbation approach. *SIAM J. Matrix Anal. Appl.*, 27:582–600, 2005.

[5] J.M. Gablonsky. *Modifications of the DIRECT algorithm*. PhD thesis, North Carolina State University, Raleigh, North Carolina, 2001.

[6] F. R. Gantmacher. *The Theory of Matrices*, volume II. Chelsea, 1959.

[7] G. Golub and C. Van Loan. *Matrix Computations, 3rd Ed.* The Johns Univ. Press, Baltimore, Maryland, 1996.

[8] D. R. Jones, C. D. Perttunen, and B. E. Stuckman. Lipschitzian optimization without the Lipschitz constant. *J. Optim. Theory Appl.*, 79(1):157–181, 1993.

[9] D. C. Liu and J. Nocedal. On the limited memory BFGS method for large scale optimization. *Math. Programming*, 45(3, (Ser. B)):503–528, 1989.

[10] C.F. Van Loan. How near is a stable matrix to an unstable matrix. *Contemporary Math.*, 47:465–477, 1985.

[11] A.N. Malyshev. A formula for the 2-norm distance from a matrix to the set of matrices with multiple eigenvalues. *Numer.Math.*, 83:443–454, 1999.

[12] E. Mengi. Locating a nearest matrix with an eigenvalue of prespecified algebraic multiplicity. *Numer.Math, to appear*, 2009.

[13] E. Mengi. Nearest linear systems with highly deficient reachable subspaces. *SIAM J. Matrix Anal. Appl., submitted*, 2010.

[14] S.A Piyavskii. An algorithm for finding the absolute extremum of a function. *USSR Comput.Math. and Math. Phys.*, 12:57–67, 1972.

[15] B. Shubert. A sequential method seeking the global maximum of a function. *SIAM J. Numer. Anal.*, 9:379–388, 1972.