

A power method for the structured singular value

Andy Packard^{# b} Michael K.H. Fan[‡] John Doyle[#]

1 Abstract

This paper presents an iterative algorithm to compute lower bounds for the structured singular value. The algorithm resembles a mixture of power methods for eigenvalues and singular values, which is not surprising, since the structured singular value can be viewed as a generalization of both. If the algorithm converges, a lower bound for μ results. We prove that μ is always an equilibrium point of the algorithm, however, since in general there are many equilibrium points, some heuristic ideas to achieve convergence are presented. Extensive numerical experience with the algorithm is discussed.

2 Introduction

The structured singular value, μ , is a useful tool for matrix perturbation problems, [Doy]. Computation of μ is difficult, and usually, upper and lower bounds are all that can be reliably computed. Upper bounds give conservative estimates of the sizes of allowable perturbations, while lower bounds yield "problem perturbations". In [FanT], the calculation of μ is reformulated as a smooth optimization problem. As with all of the known exact expressions for μ , the function to be maximized has local maximums which are not global, so in general the method yields only lower bounds for μ . Similar comments can be made for the ideas in [Doy] and [Hel], as well as the algorithm in this paper. The contribution here is yet another lower bound algorithm to aid in the analysis of robustness of systems with structured uncertainty.

3 Definitions

We consider matrices $M \in \mathbb{C}^{n \times n}$. $\bar{\sigma}(M)$ and $\rho(M)$ respectively denote the *maximum singular value* and the *spectral radius* of M . M^* denotes the complex conjugate transpose of M .

In the definition of $\mu(M)$, there is an underlying structure Δ , (a prescribed set of block diagonal matrices) which depends on the **uncertainty** and **performance objectives** of each problem. Defining the structure involves specifying three things; the type of each block, the total number of blocks, and their dimensions. There are two types of blocks in this formulation—*repeated scalar* and *full* blocks. Two non-negative integers, s and f , represent the number of *repeated scalar* blocks and the number of *full* blocks, respectively. To bookkeep their dimensions, we introduce positive integers $r_1, \dots, r_s; m_1, \dots, m_f$. The i 'th scalar block is $r_i \times r_i$, while the j 'th full block is $m_j \times m_j$. With those integers given, we define Δ as

$$\Delta := \{\text{diag} [\delta_1 I_{r_1}, \dots, \delta_s I_{r_s}, \Delta_1, \dots, \Delta_f] : \delta_i \in \mathbb{C}, \Delta_j \in \mathbb{C}^{m_j \times m_j}\} \quad (3.1)$$

and

$$\mathbf{B}\Delta := \{\Delta \in \Delta : \bar{\sigma}(\Delta) \leq 1\}$$

[#] Dept. of Electrical Eng., 116-81, Caltech, Pasadena, CA 91125, & Honeywell Systems and Research Center, Systems and Control Sciences, MN65-2500, 3660 Technology Drive, Minneapolis, MN 55418.

[‡] Dept. of Electrical Eng. & Systems Research Center, Univ. of Maryland, College Park, MD 20742.

^b Dept. of Electrical and Computer Eng., Univ. of California, Santa Barbara, CA 93106.

For consistency among all the dimensions, we must have $\sum_{i=1}^s r_i + \sum_{j=1}^f m_j = n$. The *full blocks* do not have to be square, but restricting them as such saves a great deal in terms of notation.

Definition 3.1 For $M \in \mathbb{C}^{n \times n}$, $\mu_\Delta(M)$ is defined

$$\mu_\Delta(M) := \frac{1}{\min_{\Delta \in \Delta} \{\bar{\sigma}(\Delta) : \det(I + M\Delta) = 0\}} \quad (3.2)$$

unless no $\Delta \in \Delta$ makes $I + M\Delta$ singular, then $\mu_\Delta(M) = 0$.

We can easily calculate $\mu_\Delta(M)$ when Δ is one of two extreme sets.

- If $\Delta = \{\delta I_n : \delta \in \mathbb{C}\}$, then $\mu_\Delta(M) = \rho(M)$.
- If $\Delta = \mathbb{C}^{n \times n}$, then $\mu_\Delta(M) = \bar{\sigma}(M)$.

Obviously, a general Δ as in (3.1) satisfies $\{\delta I : \delta \in \mathbb{C}\} \subset \Delta \subset \mathbb{C}^{n \times n}$. Hence directly from the "minimization" in Definition 3.1, we can conclude that $\rho(M) \leq \mu_\Delta(M) \leq \bar{\sigma}(M)$. These bounds alone are not sufficient for our purposes, because the gap between ρ and $\bar{\sigma}$ can be arbitrarily large. We refine them by considering transformations on M that **do not affect** $\mu_\Delta(M)$, but **do affect** ρ and $\bar{\sigma}$. To do this, define the following two subsets of $\mathbb{C}^{n \times n}$

$$\mathcal{Q} = \{\Delta \in \Delta : \Delta^* \Delta = I_n\} \quad (3.3)$$

$$\mathcal{D} = \left\{ \text{diag} [D_1, \dots, D_s, d_1 I_{m_1}, \dots, d_f I_{m_f}] : D_i \in \mathbb{C}^{r_i \times r_i}, D_i = D_i^* > 0, d_j \in \mathbb{R}, d_j > 0 \right\} \quad (3.4)$$

Note that for any $\Delta \in \Delta$, $Q \in \mathcal{Q}$, and $D \in \mathcal{D}$, several identities hold, namely: $Q^* \in \mathcal{Q}$, $Q\Delta \in \Delta$, $\Delta Q \in \Delta$, $\bar{\sigma}(Q\Delta) = \bar{\sigma}(\Delta Q) = \bar{\sigma}(\Delta)$, $D\Delta = \Delta D$. Consequently, we have:

Theorem 3.2 For all $Q \in \mathcal{Q}$ and $D \in \mathcal{D}$

$$\mu_\Delta(MQ) = \mu_\Delta(QM) = \mu_\Delta(M) = \mu_\Delta(DMD^{-1})$$

and,

Theorem 3.3

$$\max_{Q \in \mathcal{Q}} \rho(QM) \leq \max_{\Delta \in \mathbf{B}\Delta} \rho(\Delta M) = \mu_\Delta(M) \leq \inf_{D \in \mathcal{D}} \bar{\sigma}(DMD^{-1}) \quad (3.5)$$

An important question is "when are the bounds in (3.5) actually equalities?". A main result from [Doy] is that the left inequality is always an equality. We state this as a theorem.

Theorem 3.4 $\max_{Q \in \mathcal{Q}} \rho(QM) = \mu_\Delta(M)$

In [Doy] and [Pac], it is shown that if $2s + f \leq 3$ (recall these are the number of *repeated scalar* and *full* blocks), then the right inequality is also an equality, and for block structures with $2s + f > 3$, there exist matrices for which μ is less than the infimum. Unfortunately, the function $\rho(QM)$ has local maximums which are not global, hence evaluating the expression $\max_{Q \in \mathcal{Q}} \rho(QM)$ using gradient-like techniques is difficult. Moreover, it is expensive, since the cost evaluation and derivative calculation involves eigenvalue/eigenvector computations. Roughly speaking, this paper addresses the lower bound, and develops an iterative algorithm which *quickly* finds points $Q \in \mathcal{Q}$ that are local maximums of the function $r : \mathbf{B}\Delta \rightarrow \mathbb{R}$, defined by $r(\Delta) = \rho(\Delta M)$. Some of the results are generalizations of those found in [FanT] and [DanKL].

We will be interested in local maximums of the function $r(\Delta) = \rho(\Delta M)$, therefore we begin with some facts from perturbation theory, which assist in characterizing local phenomena.

4 Matrix Facts

4.1 Derivatives of eigenvalues

In this section we review the differentiability properties of eigenvalues and eigenvectors of matrices depending analytically on a real variable. All material comes from [Kat].

Suppose $M: \mathbf{R} \rightarrow \mathbf{C}^{n \times n}$ is an analytic function of the real parameter t . If λ_o is an eigenvalue of $M_o := M(0)$ of multiplicity one, then for some open interval containing 0, this eigenvalue is a analytic function of t , as are the eigenvectors associated with it. That is, suppose there are nonzero $x_o, y_o \in \mathbf{C}^n$, satisfying

$$\begin{aligned} y_o^* x_o &= 1 \\ M_o x_o &= \lambda_o x_o \\ y_o^* M &= \lambda_o y_o^* \end{aligned} \quad (4.1)$$

Then there is an $\epsilon > 0$ and analytic functions $x: (-\epsilon, \epsilon) \rightarrow \mathbf{C}^n, y: (-\epsilon, \epsilon) \rightarrow \mathbf{C}^n, \lambda: (-\epsilon, \epsilon) \rightarrow \mathbf{C}$ such that for all $t \in (-\epsilon, \epsilon)$

$$\begin{aligned} y^* x &= 1 \\ M x &= \lambda x \\ y^* M &= \lambda y^* \end{aligned} \quad (4.2)$$

This follows from [Kat]. Hence, we can differentiate and obtain

$$\dot{\lambda}(0) = y_o^* \dot{M}(0) x_o \quad (4.3)$$

4.2 Linear algebra lemmas

The next two lemmas are elementary linear algebra. They will be used in the main theorem of the next section.

Lemma 4.1 *Let $y, x \in \mathbf{C}^n$ with $y \neq 0$ and $x \neq 0$. There exists $d \in \mathbf{R}, d > 0$, such that $y = dx$ if and only if $\operatorname{Re}(y^* W x) \leq 0$ for every $W \in \mathbf{C}^{n \times n}$ satisfying $W + W^* \leq 0$.*

Proof: The “only if” is easy to verify, so we just prove the “if”. As usual, let y_i and x_i denote the i ’th element of y and x , and $W_{i,j}$ denote the i, j element of $W \in \mathbf{C}^{n \times n}$. Begin with any positive integer $i \leq n$. Let W be zero everywhere, except in the i, i element, and set $W_{i,i} = \sigma_i + j\omega_i$ for arbitrary $\sigma_i \leq 0$ and $\omega_i \in \mathbf{R}$. It can be shown that either $x_i = 0, y_i = 0$, or that the phases of x_i and y_i are the same.

Now, let $l \neq k$ be two integers $\leq n$. Let $\omega \in \mathbf{R}$ be arbitrary. Define a matrix W by $W_{l,k} := -e^{-j\omega}, W_{k,l} := e^{j\omega}$ and zero everywhere else. Note that $W + W^* = 0$, so trivially W satisfies the hypothesis. In this case, one can show that $x_i = 0$ if and only if $y_i = 0$, and for nonzero components, the quantity $\left| \frac{y_i}{x_i} \right|$ is independent of i . Define $d > 0$ to be this ratio. For every i , we have $y_i = dx_i$ as desired. $\#$

Lemma 4.2 *Let a and b be two nonzero vectors in \mathbf{C}^n . Then there exists a hermitian, positive definite $D \in \mathbf{C}^{n \times n}$, such that $D b = a$ if and only if $b^* a \in (0, \infty)$.*

5 Eigenvector characterization of local maximums

Consider the function $r: \mathbf{B}\Delta \rightarrow \mathbf{R}$, defined by $r(\Delta) = \rho(\Delta M)$. Recall that $\mu_\Delta(M) = \max_{\Delta \in \mathbf{B}\Delta} r(\Delta)$, and that the global maximum does occur on the “boundary” of $\mathbf{B}\Delta$, \mathcal{Q} . In this section, we characterize the occurrence of a local maximum of r at $\Delta = I$ in terms of the eigenvectors of M . We begin with some notation.

Let x and y be right and left eigenvectors of M , associated with an eigenvalue λ . Partition x and y compatibly with the block structure

$$x = \begin{bmatrix} x_{r_1} \\ x_{r_2} \\ \vdots \\ x_{r_s} \\ x_{m_1} \\ x_{m_2} \\ \vdots \\ x_{m_f} \end{bmatrix}, \quad y = \begin{bmatrix} y_{r_1} \\ y_{r_2} \\ \vdots \\ y_{r_s} \\ y_{m_1} \\ y_{m_2} \\ \vdots \\ y_{m_f} \end{bmatrix} \quad (5.1)$$

where $x_{r_i}, y_{r_i} \in \mathbf{C}^{r_i}$ and $x_{m_j}, y_{m_j} \in \mathbf{C}^{m_j}$ for each i and j . We call these the “block components” of x and y , and for technical reasons, we make a nondegeneracy assumption - for every $i, y_{r_i}^* x_{r_i} \neq 0$, and for each $j, x_{m_j} \neq 0, y_{m_j} \neq 0$.

Theorem 5.1 *Let $M \in \mathbf{C}^{n \times n}$ be given, and suppose $\lambda_o > 0$ is a distinct eigenvalue of M , with right and left eigenvectors x and y respectively, and assume that these vectors satisfy the nondegeneracy assumption above. Suppose that $\rho(M) = \lambda_o$. If the function $r: \mathbf{B}\Delta \rightarrow \mathbf{R}$ defined by $r(\Delta) = \rho(\Delta M)$ has a local maximum (with respect to the set $\mathbf{B}\Delta$) at $\Delta = I$, then there exists a $D \in \mathcal{D}$ such that $D^{-1}y = Dx$.*

Proof: Let $G \in \Delta$ with $G + G^* \leq 0$. Obviously, G appears as

$$\operatorname{diag}[g_1 I_{r_1}, \dots, g_s I_{r_s}, G_1, \dots, G_f] \quad (5.2)$$

where $\operatorname{Re}(g_i) \leq 0$, and $G_j + G_j^* \leq 0$ for all i and j . Obviously, at $t = 0$, $e^{Gt} = I$, and $e^{Gt} \in \mathbf{B}\Delta$ for all $t \geq 0$. Define a matrix function $W: \mathbf{R} \rightarrow \mathbf{C}^{n \times n}$ by $W(t) := e^{Gt} M$. Note that at $t = 0$, λ_o is a simple eigenvalue of $W(0)$, with x and y the right and left eigenvectors. For some nonempty interval containing 0, this eigenvalue is always simple, and hence there is an analytic function of the real variable t , $\lambda(t)$, defined on that interval, such that $\lambda(t)$ is an eigenvalue of $W(t)$ for all t and $\lambda(0) = \lambda_o$. It is easy to calculate $\dot{\lambda}(0)$, namely

$$\dot{\lambda}(0) = y^* \dot{W}(0) x = \lambda_o y^* G x \quad (5.3)$$

By hypothesis, $\lambda_o > 0, \rho(M) = \lambda_o$ and the function $\rho(\Delta M)$ has a local maximum (with respect to $\mathbf{B}\Delta$) at $\Delta = I$. Therefore

$$\operatorname{Re} \left(\frac{d}{dt} \lambda(t) \Big|_{t=0} \right) \leq 0 \quad (5.4)$$

which says that the magnitude of λ must be nonincreasing at $t = 0$. Using the “block notation” of (5.1) and substituting (5.2) and (5.3) into (5.4) yields

$$\operatorname{Re} \left(\sum_{i=1}^s g_i y_{r_i}^* x_{r_i} + \sum_{j=1}^f y_{m_j}^* G_j x_{m_j} \right) \leq 0. \quad (5.5)$$

This must hold for arbitrary $G \in \Delta$ satisfying $G + G^* \leq 0$. Applying lemmas 4.1 and 4.2 we conclude that for each i , there is a $D_i = D_i^* \in \mathbf{C}^{n \times n}, D_i > 0$ such that $y_{r_i} = D_i x_{r_i}$ and for each j , there is a $d_j \in \mathbf{R}, d_j > 0$ such that $y_{m_j} = d_j x_{m_j}$. Arranging all of these D_i ’s and d_j ’s into one block diagonal \bar{D} , and taking the hermitian square root proves the lemma. $\#$

Remarks: The only restrictive assumption we have made in the above lemma is that the eigenvalue λ_o is distinct. This assures differentiability. Since λ_o is a solution of a $\max_{\Delta \in \mathbf{B}\Delta} \max_i |\lambda_i(M\Delta)|$, it is likely that at the maximum it will be distinct.

6 Decomposition

Consider matrices in $\mathbf{C}^{n \times n}$, and a compatible block structure Δ , with integers $r_1, \dots, r_s, m_1, \dots, m_f$ defining the dimensions of the blocks, as outlined in section 3. Define \mathbf{X} to be the following set of block diagonal, hermitian matrices:

$$\mathbf{X} := \{ \operatorname{diag} [Z_1, \dots, Z_s, z_1, \dots, z_{f-1}] : Z_i = Z_i^* \in \mathbf{C}^{r_i \times r_i}, z_j \in \mathbf{R} \} \quad (6.1)$$

This is a real inner product space (of dimension $\sum_{i=1}^s r_i^2 + f - 1$) with inner product defined by

$$P, T \in \mathbf{X} \quad \langle P, T \rangle := \text{tr}(PT) \quad (6.2)$$

which, in terms of the blocks that make up P and L is just

$$\langle P, T \rangle = \sum_{i=1}^s \text{tr}(P_i T_i) + \sum_{j=1}^{f-1} p_j t_j \quad (6.3)$$

Let M be a complex matrix with SVD

$$M = \beta UV^* + U_2 \Sigma_2 V_2^*. \quad (6.4)$$

In this setting, β is any singular value of M , not necessarily $\bar{\sigma}(M)$, but none of the singular values in Σ_2 should equal β . We use the integer $r > 0$, to denote the multiplicity of β . Hence $U, V \in \mathbb{C}^{n \times r}$, $U^* U = V^* V = I_r$, $U_2, V_2 \in \mathbb{C}^{n \times (n-r)}$, $U_2^* U_2 = V_2^* V_2 = I_{n-r}$.

We proceed to define the set $\nabla_{M,\beta}$. Partition U and V compatibly with Δ as

$$U = \begin{bmatrix} A_1 \\ \vdots \\ A_s \\ E_1 \\ \vdots \\ E_f \end{bmatrix} \quad V = \begin{bmatrix} B_1 \\ \vdots \\ B_s \\ F_1 \\ \vdots \\ F_f \end{bmatrix} \quad (6.5)$$

where $A_i, B_i \in \mathbb{C}^{r_i \times r}$, $E_i, F_i \in \mathbb{C}^{m_i \times r}$.

For $\eta \in \mathbb{C}^r$, with $\|\eta\| = 1$, define the following components

$$\begin{aligned} P_i^\eta &= A_i \eta \eta^* A_i^* - B_i \eta \eta^* B_i^* \\ p_j^\eta &= \eta^* (E_j^* E_j - F_j^* F_j) \eta. \end{aligned} \quad (6.6)$$

Let $\nabla_{M,\beta} \subset \mathbf{X}$ be the set of all such P^η .

$$\nabla_{M,\beta} := \left\{ \text{diag} [P_1^\eta, \dots, P_s^\eta, p_1^\eta, \dots, p_{f-1}^\eta] : \begin{aligned} &P_i^\eta, p_i^\eta \text{ in (6.6), } \eta \in \mathbb{C}^r, \|\eta\| = 1 \end{aligned} \right\} \quad (6.7)$$

Note that here we use two subscripts on ∇ . The first is the matrix, and the second is the singular value in question.

Remark: It can be shown, [FreLC] and [Doy], using perturbation theory similar to that in section 4.1, that the derivatives of the r singular values coalesced at β , of the matrix function $e^{D^t M} e^{-D^t}$, are the eigenvalues of the matrix $\beta(U^* D U - V^* D V)$. Hence, if $D \in \mathcal{D}$, and all of the eigenvalues of $U^* D U - V^* D V$ are negative, then D is a descent direction for these r singular values. Simple convexity ideas yield that there is such a $D \in \mathcal{D}$ if and only if $0 \notin \text{co}(\nabla_{M,\beta})$. This was the original motivation for defining the set ∇ , [Doy], [PacD]. It is especially important when considering the computation and characterization of the upper bound, $\inf_{D \in \mathcal{D}} \bar{\sigma}(e^{D^t M} e^{-D^t})$.

The main reason we introduce $\nabla_{M,\beta}$ here is that if there is a singular value, β , of M , and $0 \in \nabla_{M,\beta}$, then β is a lower bound for $\mu(M)$.

Theorem 6.1 Let M and a compatible block structure Δ be given. Suppose β is a singular value of M with multiplicity r . Define $\nabla_{M,\beta}$ as in (6.7). Then $0 \in \nabla_{M,\beta}$ if and only if there exists a vector $x \in \mathbb{C}^n$, a matrix $X_\perp \in \mathbb{C}^{n \times n}$, a matrix $Q \in \mathcal{Q}$, such that $\|x\| = 1$, $x^* X_\perp = 0$, $X_\perp x = 0$, and

$$QM = \beta x x^* + X_\perp \quad (6.8)$$

Proof: Let the SVD of M be

$$M = \beta UV^* + U_2 \Sigma_2 V_2^* \quad (6.9)$$

→ If $0 \in \nabla_{M,\beta}$ then there exists a $\eta \in \mathbb{C}^r$, $\|\eta\| = 1$ such that

$$\begin{aligned} A_i \eta \eta^* A_i^* - B_i \eta \eta^* B_i^* &= 0 \quad i \leq s \\ \eta^* (E_j^* E_j - F_j^* F_j) \eta &= 0 \quad j \leq f-1 \end{aligned}$$

These relations, and the partition in (6.5) imply that there is $Q \in \mathcal{Q}$ such that

$$QU\eta = V\eta \quad (6.10)$$

Define $x \in \mathbb{C}^n$ as the above; $x := QU\eta = V\eta$. Since $\|\eta\| = 1$ and U and V are isometries, $\|x\| = 1$. Simple manipulation of (6.9) and (6.10) gives

$$\begin{aligned} (QM)x &= (QM)V\eta = \beta QU\eta = \sigma V\eta = \beta x \\ x^*(QM) &= \eta^* U^* Q^* (QM) = \beta \eta^* V^* = \beta x^* \end{aligned}$$

Defining $X_\perp := M - \beta x x^*$ completes the decomposition

← Suppose Q, x , and X_\perp are given as in the hypothesis, so that $QM = \beta x x^* + X_\perp$. Define $\tilde{M} := QM$. A singular value decomposition of \tilde{M} is

$$\tilde{M} = \beta (QU)V^* + (QU_2)\Sigma_2 V_2^*$$

Hence β is a singular value of \tilde{M} , and $\tilde{M}x = \beta x$ and $\tilde{M}^*x = \beta x$. Therefore there exists a vector $\eta \in \mathbb{C}^r$, $\|\eta\| = 1$ such that $x = QU\eta = V\eta$. This implies that $0 \in \nabla_{M,\beta}$ as desired. \sharp

It is obvious from the decomposition that β is a lower bound for $\mu(M)$ since β is an **eigenvalue** of $\tilde{M} = QM$. The following corollary follows immediately.

Corollary 6.2 Let M and a compatible block structure Δ be given. Suppose $D \in \mathcal{D}$, and that β is a singular value of DMD^{-1} with multiplicity r . Define $\nabla_{DMD^{-1},\beta}$ as above. Then $0 \in \nabla_{DMD^{-1},\beta}$ if and only if there exists a vector $x \in \mathbb{C}^n$, a matrix $X_\perp \in \mathbb{C}^{n \times n}$, and a matrix $Q \in \mathcal{Q}$ such that $\|x\| = 1$, $x^* X_\perp = 0$, $X_\perp x = 0$, and

$$QDMD^{-1} = \beta x x^* + X_\perp \quad (6.11)$$

The main result of this section is that there is always (almost) a decomposition as in (6.11) with β in fact equal to $\mu(M)$ (remember, *any* β satisfying (6.11) is a lower bound for $\mu(M)$).

Theorem 6.3 Let $Q_0 \in \mathcal{Q}$ achieve the global optimum for the problem $\max_{Q \in \mathcal{Q}} \rho(QM)$, and suppose that the eigenvalue associated with $\rho(Q_0 M)$ is distinct, real and positive, and call it μ (since it is). If x and y are the right and left eigenvectors of this eigenvalue, and their block components (eq. 5.1) satisfy the nondegeneracy assumption, then there exists a $D \in \mathcal{D}$ such that

$$\begin{aligned} Q_0 D M D^{-1} (Dx) &= \mu Dx \\ x^* D Q_0 D M D^{-1} &= \mu x^* D. \end{aligned} \quad (6.12)$$

Proof: By Theorems 3.3 & 3.4, any global maximizer of $\max_{Q \in \mathcal{Q}} \rho(QM)$, is also a maximizer of $\max_{\Delta \in \mathbf{B}\Delta} \rho(\Delta M)$. Define $\tilde{M} := Q_0 M$, then $\Delta = I$ is a local (in fact global) maximizer for $\max_{\Delta \in \mathbf{B}\Delta} \rho(\Delta \tilde{M})$. Apply Theorem 5.1 to the matrix \tilde{M} to prove the theorem. \sharp

Remarks: Note that this is the decomposition described in Corollary 6.2, since Dx is both a right and left eigenvector of $Q_0 D M D^{-1}$ associated with the eigenvalue μ .

This result was first shown in [FanT], for the case of $s = 0$. It is also similar to the “principle direction alignment” ideas in [DanKL]. Theorem 6.3 is more general though, since it handles repeated scalar blocks, as well as full blocks. Finally, the proof (which implicitly includes Theorem 5.1) is much simpler.

This theorem is **not true** if we consider *local* *maximizations* that are *not global* of the function $\tilde{r}: \mathcal{Q} \rightarrow \mathbb{R}$ defined as $\tilde{r}(Q) := \rho(QM)$.

7 Lower bound power algorithm

How can this decomposition be used? In this section, we propose an iterative algorithm (reminiscent of the power algorithm for spectral radius) to find such decompositions, and therefore get lower bounds for μ .

Rewriting (6.12), and changing notation a bit, we want to find a $Q \in \mathcal{Q}$, $D \in \mathcal{D}$, $\beta > 0$, and $x \in \mathbb{C}^n$ with $\|x\| = 1$ such that

$$\begin{aligned} QDM D^{-1}x &= \beta x \\ D^{-1}M^*DQ^*x &= \beta x \end{aligned}$$

which can in turn be rewritten as

$$\begin{aligned} M(D^{-1}x) &= \beta(D^{-1}Q^*x) \\ M^*(DQ^*x) &= \beta(Dx). \end{aligned}$$

For a given D, Q , and x , define vectors a, b, z , and w by

$$\begin{aligned} b &:= D^{-1}x, & a &:= D^{-1}Q^*x \\ z &:= DQ^*x, & w &:= Dx \end{aligned} \quad (7.1)$$

With this definition, we have $Mb = \sigma a$ and $M^*z = \sigma w$. We can eliminate x from (7.1), and redefining $D \leftarrow D^2$ to get

$$\begin{aligned} b &= Qa = D^{-1}x \\ z &= Da = Q^*w \end{aligned}$$

We would like to write these four new relationships in a manner that does not involve the matrices Q and D . With a few technical conditions, this can be done. In order to simplify the upcoming formulas, we will consider a block structure with $s = 1, f = 1$. By simply duplicating the appropriate formulas for additional blocks, it is straightforward to extend the algorithm to more general structures. Hence the sets \mathcal{D} and \mathcal{Q} look like

$$\mathcal{D} = \{\text{diag}[D_1, d_2 I_{m_1}] : D_1 \in \mathbb{C}^{r_1 \times r_1}, D_1 = D_1^* > 0, d_2 > 0\} \quad (7.2)$$

$$\mathcal{Q} = \{\text{diag}[q_1 I_{r_1}, Q_2] : \bar{q}_1 q_1 = 1, Q_2 \in \mathbb{C}^{m_1 \times m_1}, Q_2^* Q_2 = I_{m_1}\}. \quad (7.3)$$

With respect to this, we will partition the vectors accordingly, so $z = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$, where $z_1 \in \mathbb{C}^{r_1}$ and $z_2 \in \mathbb{C}^{m_1}$, and likewise for the other vectors.

Lemma 7.1 Let r_1 and m_1 be positive integers. Let $z_1, w_1, b_1, a_1 \in \mathbb{C}^{r_1}$ and $z_2, w_2, b_2, a_2 \in \mathbb{C}^{m_1}$ be nonzero vectors with $a_1^* w_1 \neq 0$. Then, there exists a $D \in \mathcal{D}$, and $Q \in \mathcal{Q}$ such that

$$\begin{aligned} b &= Qa, & z &= Q^*w \\ z &= Da, & b &= D^{-1}w \end{aligned}$$

if and only if

$$\begin{aligned} z_1 &= \frac{w_1^* a_1}{|w_1^* a_1|} w_1, & z_2 &= \frac{\|w_2\|}{\|a_2\|} a_2 \\ b_1 &= \frac{a_1^* w_1}{|a_1^* w_1|} a_1, & b_2 &= \frac{\|a_2\|}{\|w_2\|} w_2 \end{aligned}$$

Proof:

→ The relations for z_2 and b_2 follow by direct substitution. For z_1 and b_1 , it is easiest to define an auxiliary variable $x := D^{\frac{1}{2}}b$, and then verify via substitutions.

← Let $q_1 = \frac{a_1^* w_1}{|a_1^* w_1|}$, since this is well defined. Likewise, choose $d_2 = \frac{\|w_2\|}{\|a_2\|}$. By assumption, d_2 is well defined, and nonzero.

Obviously, $\|w_2\| = \|z_2\|$, so let Q_2 be any unitary matrix that takes w_2 into z_2 . An easy calculation shows that Q_2 also rotates b_2 into a_2 ,

$$Q_2 b_2 = \frac{1}{d_2} Q_2 w_2 = \frac{1}{d_2} z_2 = a_2.$$

Next, we calculate $a_1^* z_1$. Plugging in gives $a_1^* z_1 = |a_1^* w_1|$. By assumption, this is nonzero, hence Lemma 4.2 yields a hermitian, positive definite D_1 such that $D_1 a_1 = z_1$. As we hope, D_1 takes b_1 into w_1 too,

$$D_1 b_1 = q_1 D_1 a_1 = q_1 z_1 = w_1.$$

Defining D and Q in the obvious manner completes the proof. #

This gives us the main theorem.

Theorem 7.2 Let $M \in \mathbb{C}^{n \times n}$ be given, and let Δ be the two block ($s = 1, f = 1$) structure defined above, with block sizes r_1 and m_1 , where $r_1 + m_1 = n$. Suppose $\beta > 0$ is given. Then there exists $Q \in \mathcal{Q}$, $D \in \mathcal{D}$, $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{C}^n$, $X_\perp \in \mathbb{C}^{n \times (n-1)}$ such that

$$\begin{aligned} \|x\| &= 1, x_1 \neq 0, x_2 \neq 0 \\ x^* X_\perp &= 0, X_\perp x = 0 \\ QDM D^{-1} &= \beta x x^* + X_\perp \end{aligned} \quad (7.4)$$

if and only if there exists nonzero vectors $z_1, w_1, b_1, a_1 \in \mathbb{C}^{r_1}$ and $z_2, w_2, b_2, a_2 \in \mathbb{C}^{m_1}$ with $a_1^* w_1 \neq 0$ and

$$\begin{aligned} \beta a &= Mb \\ z_1 &= \frac{w_1^* a_1}{|w_1^* a_1|} w_1, & z_2 &= \frac{\|w_2\|}{\|a_2\|} a_2 \\ \beta w &= M^* z \\ b_1 &= \frac{a_1^* w_1}{|a_1^* w_1|} a_1, & b_2 &= \frac{\|a_2\|}{\|w_2\|} w_2. \end{aligned} \quad (7.5)$$

Remark: In order to find decompositions using the representation this theorem allows (equation (7.5) - free of Q 's and D 's), we can restrict ourselves to unit vectors a, b, z, w . Why? Suppose we find nonzero vectors satisfying (7.5). Examining these equations, it is clear that scaling z and w by some $\alpha \neq 0$ and scaling b and a by some $\gamma \neq 0$ does not affect any of the equalities in (7.5). Moreover, these equations **always** imply that $\|z\| = \|w\|$, and $\|a\| = \|b\|$, so by proper scaling, all the vectors would be unit norm.

In the above theorem, we have written the conditions (7.5) in a suggestive manner. We will attempt to find solutions to (7.5) in an iterative fashion. In particular, for $i = 1, 2$, let vectors $a_{i,k}, b_{i,k}, z_{i,k}$, and $w_{i,k}$, and positive scalars $\tilde{\beta}_k, \hat{\beta}_k$ evolve as

$$\begin{aligned} \tilde{\beta}_{k+1} a_{k+1} &= M b_k \\ z_{k+1} &= \frac{w_k^* a_{k+1}}{|w_k^* a_{k+1}|} w_k, & z_{k+1} &= \frac{\|w_{k+1}\|}{\|a_{k+1}\|} a_{k+1} \\ \hat{\beta}_{k+1} w_{k+1} &= M^* z_{k+1} \\ b_{k+1} &= \frac{a_{k+1}^* w_{k+1}}{|a_{k+1}^* w_{k+1}|} a_{k+1}, & b_{k+1} &= \frac{\|a_{k+1}\|}{\|w_{k+1}\|} w_{k+1} \end{aligned} \quad (7.6)$$

where $\tilde{\beta}_{k+1}$ and $\hat{\beta}_{k+1}$ are chosen > 0 , so that $\|a_{k+1}\| = \|w_{k+1}\| = 1$.

Note also that if the initial b and w vectors that start the iteration are unit vectors, then at every step, all vectors, a, b, z , and w will be unit length.

Remarks:

7.a Potential problems are:

- $M b_k = 0$ ($M^* z_k = 0$), then a_{k+1} (w_{k+1}) is not well defined.
- $a_{k+1}^* w_{k+1} = 0$, then the vectors z_{k+1} and/or b_{k+1} are not well defined.
- Either $\|w_{k+1}\| = 0$ or $\|a_{k+1}\| = 0$, making b_{k+1} and/or z_{k+1} not well defined.

The heuristic fix when any of these happen is to restart the algorithm at a different initial condition (ie., a new b_1 and w_1).

7.b If the iteration does converge to an equilibrium point, then the β values must be equal, that is $\tilde{\beta} = \hat{\beta}$. This is easy to see: suppose the equations in (7.5) are satisfied (convergence of the algorithm in (7.6)), but the β associated with b and a is $\tilde{\beta}$ and the β associated with z and w is $\hat{\beta}$. The converged equations imply that there exists a $Q \in \mathcal{Q}$ and $D \in \mathcal{D}$ such that $QDM D^{-1}(Db) =$

$\tilde{\beta}(Db)$ and $(QDM D^{-1})^*(Db) = \hat{\beta}(Db)$. Since the β 's are real, they must be equal. Hence, when verifying convergence of the algorithm, it is necessary to begin checking the convergence of the vectors *only after* the β_k and β_k values are nearly equal. This saves some computations early in the iteration.

- 7.c
- If there were only the first block, which is a repeated scalar block, the iteration would be a power iteration for the largest (in magnitude) eigenvalue of the matrix M . Since μ for 1 repeated scalar block is the spectral radius, the algorithm we have proposed reduces to a valid algorithm in the special case of 1 repeated scalar block.
 - If there were only the second block, which is a full block, the iteration becomes a eigenvalue power algorithm for M^*M , hence it will give the largest singular value of M . Again, with respect to this specific block structure, this is what we want.

Hence, the iteration we have proposed is a mix of two separate, well understood iterations. Both of these converge to the largest eigenvalue/singular value. Therefore, we are led to guess (incorrectly) that this algorithm will converge to the largest β for which a decomposition described in Corollary 6.2 exists, which by Theorem 6.3 is equal to $\mu_\Delta(M)$.

Extensive computational experience has led to the following conclusions:

1. The difficulties described in 7.b above do not *seem* to occur in practice. Matrices whose eigenvector block components (eq. 5.1) do not satisfy the “nonzero” block conditions described at the beginning of section 5 will cause this problem. There is a systematic “reduction” that is possible with these matrices, which produces a smaller matrix \tilde{M} along with a smaller block structure, $\tilde{\Delta}$, such that $\mu_\Delta(M) = \mu_{\tilde{\Delta}}(\tilde{M})$. More research is underway to understand these reductions. In any event, while it is easy to construct matrices where these problems happen, running the algorithm on frequency responses of actual closed loop systems has not been a problem.
2. Limit cycles can occur, and seem to occur more often when there are large repeated scalar blocks. Unlike an equilibrium point, the presence of a stable limit cycle does not immediately give rise to a lower bound for μ .
3. If $s = 0$ (and often when $s > 0$), the algorithm usually performs well, converging quickly, and providing a lower bound which is better (ie. bigger) than $\rho(M)$. Unfortunately, this is not always true. We have found an example that has a stable equilibrium point with the corresponding $\beta < \rho(M)$. With lack of any further insight, we do not bother to reproduce this here. The block structure was five 1×1 blocks. More numerical tests are outlined in the next section.
4. There are many other iterative algorithms besides (7.6) that have decompositions (Theorem 6.3) as equilibrium points. For instance, simply rearranging the order of our iteration in (7.6) will yield a different algorithm, yet the decompositions are still the equilibrium points and vice versa. The most desirable algorithm would have as its only stable equilibrium points, decompositions with large (relative to $\mu(M)$) converged β values. Other iterations schemes may be better suited toward this goal.
5. In general, there are several stable equilibrium points, with **different values of β** . This means that the bound resulting from this algorithm might be simply a lower bound. This is in contrast with the conventional power algorithms for ρ and $\bar{\sigma}$, where only the largest ones are stable.

8 Implementation

This section addresses the implementation of the algorithm, beginning with a heuristic selection of the starting vectors, b_1 and w_1 .

To motivate what follows, suppose $\mu(M) = \inf_{D \in \mathcal{D}} \bar{\sigma}(DMD^{-1})$, and that the infimum is achieved by D_o . Then, from Theorem 6.1, we must have $0 \in \nabla_{D_o M D_o^{-1}, \bar{\sigma}}$. Therefore, if

$$\tilde{M} := D_o M D_o^{-1} = \mu UV^* + U_2 \Sigma_2 V_2^* \quad (8.1)$$

is a singular value decomposition, there is a $\eta \in \mathbb{C}^r$ and $Q \in \mathcal{Q}$ such that

$$\begin{aligned} \tilde{M} V \eta &= \mu Q V \eta \\ \tilde{M}^* (Q V \eta) &= \mu V \eta \end{aligned} \quad (8.2)$$

Hence, with respect to \tilde{M} , (which has $\mu(\tilde{M}) = \mu(M)$), the vectors $b = w := V \eta$ are the correct b and w vectors for the decomposition. We therefore propose the following overall algorithm.

1. Cheaply find a D_{so} that nearly minimizes $\bar{\sigma}(DMD^{-1})$. In [Osb], an algorithm for computing

$$D_F := \arg \min_{D \in \mathcal{D}} \text{trace}(DMD^{-2}M^*D) \quad (8.3)$$

Although there is little theoretical evidence to suggest that D_F is a good choice of D_{so} , our experience has shown that it is.

2. Absorb this into M , ie., define $\tilde{M} := D_{so} M D_{so}^{-1}$
3. choose $b_1 = w_1$ to be a right singular vector associated with $\bar{\sigma}(\tilde{M})$ – note that $\bar{\sigma}(\tilde{M})$ is a upper bound for $\mu_\Delta(M)$
4. perform the iteration on \tilde{M} with these starting vectors

9 Numerical examples

Several numerical tests have been run with this algorithm, in addition to successful use in practical situations (Caltech design studies, Honeywell Systems and Research Center). In this section, two types of random tests are described, along with the numerical results that have been obtained.

9.1 Test # 1

Let n be a positive integer, and let $\Delta \subset \mathbb{C}^{n \times n}$ be any block structure. Let r be a positive integer with $r \leq n$. Using the definition of the set ∇ , we can randomly generate matrices with the property that

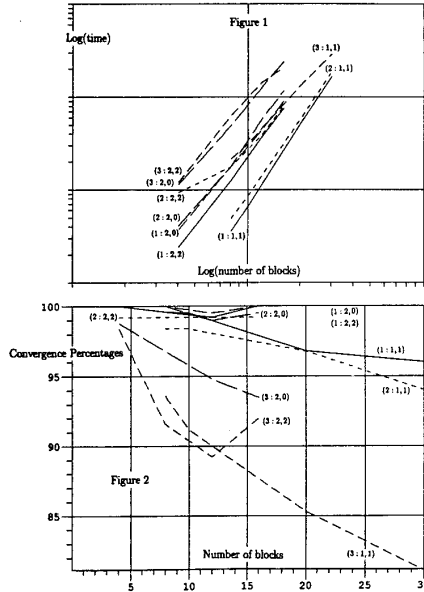
$$\mu_\Delta(M) = \min_{D \in \mathcal{D}} \bar{\sigma}(DMD^{-1}) = 1$$

and at the minimum, there are r singular values coalesced at 1. In [FanT], such a procedure is given for block structures with $s = 0$ - no repeated scalar blocks - the extension of this procedure to block structures having repeated scalar blocks is simple. After running the algorithm on these matrices, several performance measures can be compiled, such as

- percentage that converge to an equilibrium point
- percentage that converge to correct value of μ (since we know that $\mu_\Delta(M) = 1$, this is possible to verify)
- time to converge (or conclude that the iteration is not going to converge)

The “size” of the problem (number of blocks, size of blocks, number of maximum singular values coalesced at the infimum - r , etc.) will affect these, and our computational experiments include many combinations. The results are shown below. Figure 1 is a log-log plot of the time versus number of blocks, for different types of blocks, and differ-

ent values of r . The notation is best explained by example: $(3 : 2, 0)$ means that $r = 3$, and each block was a 2×2 repeated scalar block, and if the blocks were 2×2 full blocks, then the notation would be $(r : 2, 2)$ where r corresponds to the number of coalesced singular values at the minimum. Note that as r increases, the power algorithm's performance generally deteriorates. The timings indicate that the overall algorithm has roughly polynomial behavior in b , where b is the number of blocks. Figure 2 shows the percentage of problems for which the power algorithm actually converged, and hence gave a lower bound for $\mu_\Delta(M)$. Although there are several equilibrium points to which the power algorithm could converge, in the cases where convergence was achieved, the average value of the corresponding lower bound was 0.997.



Unfortunately there is no (known) way to generate arbitrary matrices M with $1 = \mu_\Delta(M) < \inf_{D \in D} \bar{\sigma}(DMD^{-1})$. Hence, this numerical experiment *only* tests the power algorithm on the special class of matrices having the property that μ and the upper bound are the same.

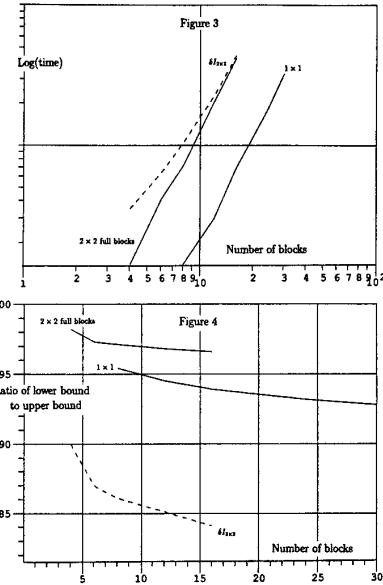
9.2 Test # 2

The most simple test, at least from an implementation standpoint, is to generate random matrices for several different block structures, and let the algorithm run. There are two main drawbacks to this approach. First, since $\mu_\Delta(M)$ is not explicitly known, the only conclusions that can be drawn from these experiments are

- the relationship between problem size and time to converge to an equilibrium point
- an idea of the general "goodness" of the upper bound obtained from the Frobenious norm minimization (eq. 8.3) in relation to the lower bound obtained from convergence of the power algorithm.

Second, and maybe more important, extrapolating the trends that emerge for random matrices in to rules of thumb for matrices actually encountered in practice might be misleading.

In any event, the results from this test are shown in Figures 3 and 4. Figure 3 has the timings for various blocks structures - *in all of these tests, every example converged to an equilibrium point*. Figure 4 shows that ratio of the resulting lower bound to the Frobenious upper bound. This ratio could be increased closer to 1.0 by actually minimizing $\bar{\sigma}(DMD^{-1})$ directly.



9.3 Conclusions

We have found that the algorithm works well in practice (Caltech experiments and Honeywell S&RC). However, it does fail to converge sometime, and often for large block structures, the value that it gives is undoubtedly a μ - value that is lower than μ . In general, we have found that it works better in practice than it does on the matrices generated in Test #1. On the other hand, simply generating random matrices tends to give an overly optimistic opinion of the performance of the power algorithm.

10 References

- DanKL: R.W. Daniel, B. Kouvaritakis, and H. Latchman, "Principal direction alignment: a geometric framework for the complete solution to the μ problem," *IEE Proceedings*, vol. 133, Part D, No. 2, pp. 45-56, March 1986.
- Doy: J.C. Doyle, "Analysis of feedback systems with structured uncertainties", *IEE Proceedings*, vol. 129, Part D, No. 6, pp. 242-250, Nov. 1982.
- FanT: M.K.H. Fan and A.L. Tits, "Characterization and Efficient Computation of the Structured Singular Value", *IEEE Trans. Auto. Control*, vol. AC-31, no. 8, pp. 734-743, Aug. 1986.
- FreLC: J.S. Freudenberg, D.P. Looze and J.B. Cruz,, "Robustness analysis using singular value sensitivities", *Int. J. Control*, vol. 35, no. 1, pp. 95-116, 1982.
- Hel: W. Helton, "A numerical method for computing the structured singular value," *Systems and Control Letters*, pp. 21-26, 1988.
- Kat: T. Kato, "A Short Introduction to Perturbation Theory for Linear Operators," 1982, Springer Verlag, New York.
- Osb: E.E. Osborne, "On Preconditioning of Matrices," *J. Assoc. Comp. Mach.*, 7, pp. 338-345, 1960.
- PacD: A. Packard and J. Doyle, "Structured singular value with repeated scalar blocks," *1988 American Control Conference*, Atlanta, June 1988.
- Pac: A. Packard, PhD thesis, University of California, Berkeley, Dept. of Mechanical Eng., 1988.

11 Acknowledgements

The authors gratefully acknowledge financial support from AFOSR, NASA, ONR and Honeywell SRC IRAD funding.