

SRC TR 87-9-r1

**m-Form Numerical Range and the
Computation of the Structured
Singular Value**

by

M.K.H. Fan and A.L. Tits

m-Form Numerical Range and the Computation of the Structured Singular Value

Michael K.H. Fan *

André L. Tits +

Electrical Engineering Department and Systems Research Center

University of Maryland, College Park, MD 20742.

Abstract. The concept of *structured singular value* was recently introduced by Doyle (Proc. IEE, vol. 129, pp. 242-250, 1982) as a tool for the analysis and synthesis of feedback systems with structured uncertainties. In this paper an equivalent expression for the structured singular value is proposed, leading to an alternative algorithm for its computation. The new approach is based on the geometric properties of the generalized numerical range of certain matrices. Similar to previously considered schemes, the algorithm proposed here is proven to give the correct value for block-structures of size up to 3. For larger sizes, insight is gained in the question of the possible 'gap' between the structured singular value and its known upper bound.

Please address all correspondence to A.L. Tits.

Key words. structured singular value, numerical range, field of values.

This research was supported in part by the National Science Foundation under grants No. DMC-84-51515 and CDR-85-00108. During part of the time the research was performed, the first author was a Fellow of the Minta Martin Foundation, College of Engineering, University of Maryland.

* Phone number: 301-454-8832

+ Phone number: 301-454-6861

1 Introduction and Preliminaries

The concept of *structured singular value* was recently introduced by Doyle [1] as a tool for the analysis and synthesis of feedback systems with structured uncertainties. In this paper an equivalent expression for the structured singular value is proposed, leading to an alternative algorithm for its computation. The new approach is based on the properties of the *m-form numerical range* (or *m-form field of values*) (see, e.g., [2–6]) of certain matrices. Similar to previously considered schemes, the algorithm proposed here is proven to give the correct value for block-structures of size up to 3. For larger sizes, insight is gained in the question of the possible ‘gap’ between the structured singular value and its known upper bound.

Throughout the paper, given any complex matrix A , we denote by $\bar{\sigma}(A)$ its largest singular value and by A^H its complex conjugate transpose. If A is Hermitian, $A > 0$ (resp. $A \geq 0$) expresses that A is positive definite (resp. positive semi-definite). Given any complex vector x , x^T indicates its transpose, x^H its complex conjugate transpose and $\|x\|$ its Euclidean norm. The unit sphere in $\mathbb{C}^n$ is denoted by ∂B , i.e., $\partial B = \{x \in \mathbb{C}^n, \|x\| = 1\}$. Given a set S , $\text{co}S$ denotes its convex hull. A *block-structure* of size m is any m -tuple $\mathcal{K} = (k_1, \dots, k_m)$ of positive integers.¹ Given a block-structure $\mathcal{K}$ of size m , we make use of the family of diagonal matrices

$$\mathcal{D} = \{ \text{block diag } (d_1 I_{k_1}, \dots, d_m I_{k_m}) : d_i \in (0, \infty) \}$$

and of the projection matrices

$$P_i = \text{block diag } (O_{k_1}, \dots, O_{k_{i-1}}, I_{k_i}, O_{k_{i+1}}, \dots, O_{k_m}), \quad i = 1, \dots, m,$$

¹This corresponds, in the terminology of [1], to structures with no repeated blocks.

where, for any positive integer k , I_k is the $k \times k$ identity matrix and O_k the $k \times k$ zero matrix.

Definition 1. [7]² The *structured singular value* of a complex $n \times n$ matrix M with respect to the block-structure $\mathcal{K} = (k_1, \dots, k_m)$ of size m , where $n = \sum_{i=1}^m k_i$, is the nonnegative scalar³

$$\mu = \max_{x \in \partial B} \{ \|Mx\| : \|P_i x\| \|Mx\| = \|P_i Mx\|, i = 1, \dots, m \} . \quad (1)$$

□

Notice in particular that, if $\mathcal{K} = (n)$, the structured singular value is equal to the largest singular value $\bar{\sigma}(M)$.

The question of how to compute μ has been addressed by several authors [1,7–10]. Doyle [1] showed that, while in general

$$\mu \leq \inf_{D \in \mathcal{D}} \bar{\sigma}(DM D^{-1}) , \quad (2)$$

if $m \leq 3$,

$$\mu = \inf_{D \in \mathcal{D}} \bar{\sigma}(DM D^{-1}) . \quad (3)$$

He exhibited an open set of counterexamples to (3) for the case $m = 4$ [11] (see also [12]). The minimization problem in (2) has no stationary point other than its global minimizers [8] and algorithms exist to solve it [1,13]. Another approach for computing μ is to solve directly optimization problem

²This definition of the structured singular value, while computationally more tractable, is equivalent to that originally proposed by Doyle [1]: $\mu = 0$ if there is no block diagonal matrix Δ , with i th block of size k_i , such that $\det(I + M\Delta) = 0$; $\mu = (\min\{\bar{\sigma}(\Delta) : \det(I + M\Delta) = 0\})^{-1}$ otherwise, where the ‘min’ is taken over the same family of block diagonal matrices.

³Formula (1) can be interpreted as a maximization of $\|Mx\|$ over all $x \in \partial B$ such that the relative sizes of the blocks are preserved after multiplication by M .

(1). Although local maxima are generally present, global optimality can be checked whenever (3) holds [7]. For such cases, an algorithm proposed in [7] results in a typical speedup of an order of magnitude over algorithms based on (3). However, the question of obtaining an algorithm to compute μ for any block-size remains open. It is hoped that the geometric framework introduced in this paper for the computation of μ will contribute to the progress towards answering this question.

In the next section, we propose a new algorithm for computing the structured singular value, based on the distance between the origin and members of a certain family of subsets $W(\alpha)$ of $\mathbb{R}^m$. This distance can be efficiently computed when these subsets are convex, which is shown to be the case for block-structures of size no larger than 3. It turns out that, in the non-convex case, an upper bound for μ (often *equal* to μ) can be obtained instead. In the final section, it is shown that this upper bound is the same as the one given in (2), and that the ‘gap’ is related to a specific type of non-convexity of $W(\mu^2)$. An algorithm for plotting the boundary of a set $W(\alpha)$ when $m = 2$ is given in the appendix.

2 Concepts and Algorithm

For $i = 1, \dots, m$, and any real number α , let $A_i(\alpha)$ be defined as

$$A_i(\alpha) = \alpha P_i - M^H P_i M,$$

and consider the nonnegative scalar function $c(\cdot)$ defined as

$$c(\alpha) = \min\{\|v\| : v \in W(\alpha)\}$$

where $W(\alpha)$ is the m -form numerical range associated with matrices $A_1(\alpha), \dots, A_m(\alpha)$, i.e.,

$$W(\alpha) = \{v \in \mathbb{R}^m : v_i = x^H A_i(\alpha) x, i = 1, \dots, m, x \in \partial B\} . \quad (4)$$

Our first theorem leads to an equivalent formula for μ .

Theorem 1. $0 \in W(\mu^2)$ and, for all $\alpha > \mu^2$, $0 \notin W(\alpha)$.

Proof. First, let $x^* \in \partial B$ be a global solution of (1), i.e., $\mu = \|Mx^*\|$ and $\mu\|P_i x^*\| = \|P_i Mx^*\|$ for $i = 1, \dots, m$. Then, for $i = 1, \dots, m$,

$$x^{*H} A_i(\mu^2) x^* = \mu^2 \|P_i x^*\|^2 - \|P_i Mx^*\|^2 = 0 .$$

Thus $0 \in W(\mu^2)$. Second, let α be such that $0 \in W(\alpha)$. There must exist some $x \in \partial B$ such that, for $i = 1, \dots, m$,

$$\alpha \|P_i x\|^2 - \|P_i Mx\|^2 = 0 . \quad (5)$$

For any such x , one then has

$$\alpha = \sum_{i=1}^m \alpha \|P_i x\|^2 = \|Mx\|^2 . \quad (6)$$

Relations (5) and (6) imply that x is a feasible point for (1). Hence, by (1),

$$\mu^2 \geq \|Mx\|^2 = \alpha .$$

□

Corollary 1. $c(\mu^2) = 0$ and, for all $\alpha > \mu^2$, $c(\alpha) > 0$. □

Corollary 2.

$$\mu = \inf\{ \sqrt{\alpha} : c(\beta) > 0 \text{ for all } \beta \geq \alpha \} = \inf\{ \sqrt{\alpha} : 0 \notin W(\beta) \text{ for all } \beta \geq \alpha \} .$$

□

Proposition 1. $c(\cdot)$ is continuous and, for any $s \geq 0$ and any real α ,

$$c(\alpha + s) \leq c(\alpha) + s .$$

Proof. Defining $\varphi : \partial B \times \mathbb{R} \rightarrow \mathbb{R}^m$ by

$$\varphi_i(x, \alpha) = x^H A_i(\alpha) x \quad i = 1, \dots, m ,$$

one has, for any real α ,

$$c(\alpha) = \min\{\|\varphi(x, \alpha)\| : x \in \partial B\} .$$

Continuity of $c(\cdot)$ then follows from continuity of φ and compactness of ∂B .

Further,

$$\begin{aligned} c(\alpha + s) &= \min_{x \in \partial B} \sqrt{\sum_{i=1}^m (x^H A_i(\alpha + s) x)^2} \\ &= \min_{x \in \partial B} \sqrt{\sum_{i=1}^m (x^H A_i(\alpha) x + s \|P_i x\|^2)^2} . \end{aligned}$$

Using the triangular inequality in $\mathbb{R}^m$, we obtain, since $s \geq 0$,

$$c(\alpha + s) \leq \min_{x \in \partial B} \left(\sqrt{\sum_{i=1}^m (x^H A_i(\alpha) x)^2} + s \sqrt{\sum_{i=1}^m \|P_i x\|^4} \right) .$$

Replacing the second term in the ‘min’ by its constrained maximum, we obtain

$$c(\alpha + s) \leq \min_{x \in \partial B} \left(\sqrt{\sum_{i=1}^m (x^H A_i(\alpha) x)^2} + s \right) = c(\alpha) + s .$$

□

Based on these facts, provided one has an algorithm to compute $c(\alpha)$, μ can be obtained as follows.

Algorithm 1. Computation of μ

Step 0. Set $\alpha_0 = \bar{\sigma}^2(M)$ and $k = 0$.

Step 1. Set $\alpha_{k+1} = \alpha_k - c(\alpha_k)$.

Step 2. Set $k = k + 1$ and go to Step 1. $\square$

Theorem 2. The sequence $\{\alpha_k\}$ generated by Algorithm 1 is monotone nonincreasing and

$$\lim_{k \rightarrow \infty} \alpha_k = \mu^2.$$

Proof. We first show by induction that $\alpha_k \geq \mu^2$ for all k . Clearly, $\alpha_0 = \bar{\sigma}^2(M) \geq \mu^2$. Assuming the claim is true for k and using Proposition 1 and Theorem 1, we can write

$$c(\alpha_k) = c(\mu^2 + (\alpha_k - \mu^2)) \leq c(\mu^2) + \alpha_k - \mu^2 = \alpha_k - \mu^2.$$

Thus, in view of the construction in Step 1 of Algorithm 1, the claim is true for $k + 1$. Now, since the sequence $\{\alpha_k\}$ is monotone nonincreasing, it follows that it converges to a limit α^* satisfying

$$\alpha^* \geq \mu^2.$$

Since $c(\cdot)$ is continuous, the construction in Step 1 now implies, letting $k \rightarrow \infty$,

$$\alpha^* = \alpha^* - c(\alpha^*)$$

and thus $c(\alpha^*) = 0$. The result now follows from Corollary 1. $\square$

Existing algorithms [1,13,14], proposed in slightly different contexts, can be used to compute the distance between the origin and $\text{co}W(\alpha)$, thus yielding $c(\alpha)$, among other instances, when $W(\alpha)$ is convex.⁴ It turns out that such is the case for any real α whenever $m \leq 3$ [4–6]. Hence Algorithm 1

⁴These algorithms are based on repeated computation of an eigenvector associated

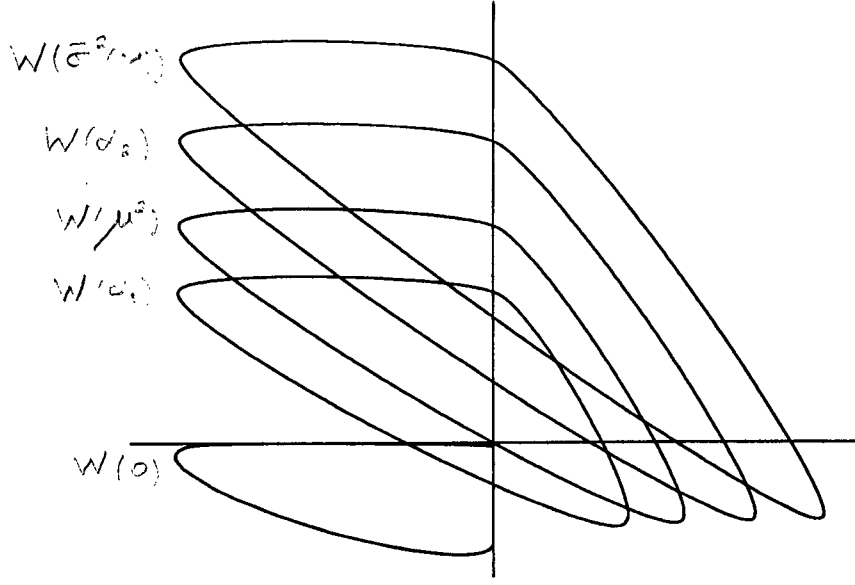


Figure 1: $W(\alpha)$ sets, with $0 < \alpha_1 < \mu^2 < \alpha_2 < \bar{\sigma}^2(M)$

provides an alternative way to compute the structured singular value for block-structures of size no larger than 3. It is not clear whether this algorithm has any computational advantage over previously proposed methods.

Figure 1 shows the boundaries of sets $W(\alpha)$ for the matrix⁵

$$M = \begin{bmatrix} 1 - j0.2 & -0.1 + j5 & 0.2 + j3 \\ j1.2 & 0.1 + j0.1 & 0.1 \\ 1 + j2 & 0.3 - j0.2 & 0.1 - j2 \end{bmatrix}$$

with block structure $K = (1, 2)$, for successive values of α equal to 0, 15, $21.1 = \mu^2$, 29, and $36.6 = \bar{\sigma}^2(M)$. The algorithm used for plotting these with the smallest eigenvalue of $\sum_{i=1}^m w_i A_i(\alpha)$, for a certain sequence of vectors $w = [w_1 \dots w_m]^T$. As shown in the Appendix for the case when $m = 2$, each such eigenvector corresponds to a contact point of $\text{co}W(\alpha)$ with the supporting hyperplane orthogonal to w .

⁵ j denotes the square root of -1 .

boundaries is given in the appendix.

3 The Non-Convex Case

For block-structures of size larger than 3, $W(\alpha)$ may not be convex. Algorithms from [1,13,14] then compute, instead of $c(\alpha)$, the value

$$c'(\alpha) = \min\{\|v\| : v \in \text{co}W(\alpha)\}$$

which is a lower bound for $c(\alpha)$. Using this value instead of $c(\alpha)$ in Algorithm 1 would then yield, as the limit of the sequence $\{\sqrt{\alpha_k}\}$, an upper bound μ' for μ , with

$$\mu' = \inf\{\sqrt{\alpha} : c'(\beta) > 0 \text{ for all } \beta \geq \alpha\} . \quad (7)$$

Numerical experiments on randomly selected matrices show that in fact $c'(\alpha)$ is in most cases equal to $c(\alpha)$, even in the non-convex case. This is in agreement with the main result of this section, given in Theorem 3. It states that μ' is always equal to the right hand side of (2), the previously known upper bound on μ .

Before proceeding further, we need to introduce more notation and to recall two results of Doyle [1]. Let r be the multiplicity of $\bar{\sigma}(M)$ as a singular value of M and let $U, V \in \mathbb{C}^{n \times r}$ have as columns r mutually orthogonal left and right unit length singular vectors of M corresponding to $\bar{\sigma}(M)$, respectively. Further, for $i = 1, \dots, m-1$, let $H_i = \bar{\sigma}(M)(U^H P_i U - V^H P_i V) \in \mathbb{C}^{r \times r}$ and define the set $\nabla_2 \subset \mathbb{R}^{m-1}$ by

$$\nabla_2 = \{z \in \mathbb{R}^{m-1} : z_i = e^H H_i e, i = 1, \dots, m-1, e \in \mathbb{C}^r, \|e\| = 1\} .$$

Fact 1. [1] $0 \in \nabla_2$ if, and only if, $\mu = \bar{\sigma}(M)$. $\square$

Fact 2. [1] $0 \in \text{co}\nabla_2$ if, and only if, $\bar{\sigma}(M) = \inf_{D \in \mathcal{D}} \bar{\sigma}(DM D^{-1})$. $\square$

Proposition 2 below shows that ∇_2 (subset of $\mathbb{R}^{m-1}$) is closely related to $W(\bar{\sigma}^2(M))$ (subset of $\mathbb{R}^m$). In proving it, we will make use of the following lemma.

Lemma 1. Let μ' be defined by (7). Then

$$\mu' \leq \inf_{D \in \mathcal{D}} \bar{\sigma}(DM D^{-1}) .$$

Proof. Let α be any real number satisfying

$$\alpha > \left(\inf_{D \in \mathcal{D}} \bar{\sigma}(DM D^{-1}) \right)^2 .$$

There must exist $D = \text{block diag } (d_i I_{k_i}) \in \mathcal{D}$ such that, for any $\beta \geq \alpha$,

$$\beta I - (DM D^{-1})^H (DM D^{-1}) > 0,$$

which implies

$$\beta D^2 - M^H D^2 M > 0,$$

i.e.,

$$\sum_{i=1}^m d_i^2 (\beta P_i - M^H P_i M) > 0$$

so that, for all $x \in \partial B$,

$$\sum_{i=1}^m d_i^2 x^H (\beta P_i - M^H P_i M) x > 0 .$$

Hence, for all $v \in W(\beta)$,

$$\langle v, \lambda \rangle > 0$$

where $\lambda = [d_1^2 \dots d_m^2]^T$. Therefore, for any $\beta \geq \alpha$, $\text{co}W(\beta)$ does not contain the origin, so that $c'(\beta) > 0$. In view of the definition of μ' , we have $\alpha > \mu'^2$, so that $\mu' \leq \inf_{D \in \mathcal{D}} \bar{\sigma}(DM D^{-1})$. $\square$

Proposition 2. The following hold: (i) $0 \in \nabla_2$ if, and only if, $0 \in W(\bar{\sigma}^2(M))$, (ii) $0 \in \text{co}\nabla_2$ if, and only if, $0 \in \text{co}W(\bar{\sigma}^2(M))$.

Proof. (i) Follows directly from Theorem 1, Fact 1 and the inequality $\mu \leq \bar{\sigma}(M)$. (ii) First suppose that $0 \in \text{co}W(\bar{\sigma}^2(M))$. Then, from the definition of μ' , it follows that

$$\bar{\sigma}^2(M) \leq \mu'.$$

In view of Lemma 1 and Fact 2, we conclude that $0 \in \text{co}\nabla_2$. To prove the converse, suppose now that $0 \in \text{co}\nabla_2$. Since the columns of V have unit length and are mutually orthogonal, ∇_2 can be written as

$$\nabla_2 = \{z \in \mathbb{R}^{m-1} : z_i = e^H H_i e, i = 1, \dots, m-1, e \in \mathbb{C}^r, \|Ve\| = 1\}.$$

Next, for any $e \in \mathbb{C}^r$, let $x = Ve$. Then

$$Mx = \bar{\sigma}(M)Ue$$

so that, for $i = 1, \dots, m-1$,

$$\begin{aligned} e^H H_i e &= \bar{\sigma}(M)x^H \left(\frac{1}{\bar{\sigma}^2(M)} M^H P_i M - P_i \right) x \\ &= \frac{-1}{\bar{\sigma}(M)} x^H (\bar{\sigma}^2(M) P_i - M^H P_i M) x. \end{aligned}$$

Thus,

$$\begin{aligned} \nabla_2 &= \{z \in \mathbb{R}^{m-1} : z_i = \frac{-1}{\bar{\sigma}(M)} x^H (\bar{\sigma}^2(M) P_i - M^H P_i M) x, i = 1, \dots, m-1, \\ &\quad x \in \mathbb{C}^n, \|x\| = 1, x \in \text{range of } V\}. \end{aligned}$$

Clearly, in view of the definition of V , $x \in \text{range of } V$ if, and only if,

$$\sum_{i=1}^m x^H (\bar{\sigma}^2(M) P_i - M^H P_i M) x = x^H (\bar{\sigma}^2(M) I - M^H M) x = 0$$

so that

$$\begin{aligned} \nabla_2 &= \{z \in \mathbb{R}^{m-1} : z_i = \frac{-1}{\bar{\sigma}(M)} x^H (\bar{\sigma}^2(M) P_i - M^H P_i M) x, i = 1, \dots, m-1, \\ &\quad x \in \mathbb{C}^n, \|x\| = 1, \sum_{i=1}^m x^H (\bar{\sigma}^2(M) P_i - M^H P_i M) x = 0\} . \end{aligned}$$

Since $0 \in \text{co}\nabla_2$, it follows that $0 \in \text{co}S$, where $S \subset \mathbb{R}^m$ is given by

$$\begin{aligned} S &= \{v \in \mathbb{R}^m : v_i = x^H (\bar{\sigma}^2(M) P_i - M^H P_i M) x, i = 1, \dots, m, \\ &\quad x \in \mathbb{C}^n, \|x\| = 1, \sum_{i=1}^m v_i = 0\} . \end{aligned}$$

Indeed, if $0 \in \text{co}\nabla_2$, then there exists a finite integer q , vectors $x^j \in \mathbb{C}^n$, $\|x^j\| = 1$, $j = 1, \dots, q$, and scalars $\nu_j \geq 0$, $j = 1, \dots, q$, with $\sum_{j=1}^q \nu_j = 1$, such that, for $i = 1, \dots, m-1$,

$$\sum_{j=1}^q \nu_j x^{jH} (\bar{\sigma}^2(M) P_i - M^H P_i M) x^j = 0$$

with, for $j = 1, \dots, q$,

$$\sum_{i=1}^m x^{jH} (\bar{\sigma}^2(M) P_i - M^H P_i M) x^j = 0$$

and this implies that

$$\sum_{j=1}^q \nu_j x^{jH} (\bar{\sigma}^2(M) P_m - M^H P_m M) x^j = 0.$$

Finally it is clear that

$$S = W(\bar{\sigma}^2(M)) \cap \{v \in \mathbb{R}^m : \sum_{i=1}^m v_i = 0\} ,$$

so that $0 \in \text{co}W(\bar{\sigma}^2(M))$. $\square$

Corollary 3. $0 \in \text{co}W(\bar{\sigma}^2(M))$ if, and only if, $\bar{\sigma}(M) = \inf_{D \in \mathcal{D}} \bar{\sigma}(DMD^{-1})$.

Proof. Follows directly from Fact 2 and Proposition 2. $\square$

Proposition 3. For any $D \in \mathcal{D}$ and any real α , (i) $0 \in W(\alpha)$ if, and only if, $0 \in W_D(\alpha)$ and (ii) $0 \in \text{co}W(\alpha)$ if, and only if, $0 \in \text{co}W_D(\alpha)$, where $W_D(\alpha)$ is defined by substituting $DM D^{-1}$ for M in the definition (4) of $W(\alpha)$.

Proof. (i) It suffices to show that $0 \in W(\alpha)$ implies $0 \in W_D(\alpha)$. Suppose $0 \in W(\alpha)$, i.e., there exists $x \in \partial B$ and, for $i = 1, \dots, m$, $x^H(\alpha P_i - M^H P_i M)x = 0$. For any $D = \text{block diag } (d_i I_{k_i}) \in \mathcal{D}$, let $y = (Dx)/\|Dx\| \in \partial B$. Then, for $i = 1, \dots, m$,

$$y^H(\alpha P_i - (DM D^{-1})^H P_i (DM D^{-1}))y = \frac{d_i^2}{\|Dx\|^2} x^H(\alpha P_i - M^H P_i M)x = 0$$

which implies $0 \in W_D(\alpha)$. (ii) It also suffices to show that $0 \in \text{co}W(\alpha)$ implies $0 \in \text{co}W_D(\alpha)$. Suppose $0 \in \text{co}W(\alpha)$, then there exists a finite integer q , vectors $x^j \in \mathbb{C}^n$, $\|x^j\| = 1$, $j = 1, \dots, q$, and scalars $\nu_j \geq 0$, $j = 1, \dots, q$, with $\sum_{j=1}^q \nu_j = 1$, such that, for $i = 1, \dots, m$,

$$\sum_{j=1}^q \nu_j x^{jH}(\alpha P_i - M^H P_i M)x^j = 0.$$

For any $D = \text{block diag } (d_i I_{k_i}) \in \mathcal{D}$ and $j = 1, \dots, q$, let $y^j = (Dx^j)/\|Dx^j\| \in \partial B$ and $\beta_j = (\nu_j \|Dx^j\|^2)/(\sum_{l=1}^q \nu_l \|Dx^l\|^2)$. Then, for $i = 1, \dots, m$,

$$\begin{aligned} \sum_{j=1}^q \beta_j y^{jH}(\alpha P_i - (DM D^{-1})^H P_i (DM D^{-1}))y^j &= \\ \frac{d_i^2}{\sum_{l=1}^q \nu_l \|Dx^l\|^2} \sum_{j=1}^q \nu_j x^{jH}(\alpha P_i - M^H P_i M)x^j &= 0. \end{aligned}$$

Since $\beta_j \geq 0$, $j = 1, \dots, q$ and $\sum_{j=1}^q \beta_j = 1$, it follows that $0 \in \text{co}W_D(\alpha)$. $\square$

Theorem 3 below shows that μ' is nothing but the previously known upper bound on μ (see (2)). In proving it (as well as in proving Proposition 5), we will make use of the following proposition, of independent interest,

which gives a sufficient condition verifiable *a priori*, under which the infimum in (2) is achieved. In this proposition, we denote by M_{ij} the ij th block of M for the given structure K , i.e., the $k_i \times k_j$ matrix with (p, q) entry given by the $(\sum_{l=1}^{i-1} k_l + p, \sum_{l=1}^{j-1} k_l + q)$ entry in M .

Proposition 4. Suppose that for any nonempty proper subset I of $\{1, \dots, m\}$ there exist $i \in I, j \notin I$ such that M_{ij} is not identically zero. Then the infimum in (2) is achieved.

Proof. By contraposition. Suppose the infimum in (2) is not achieved and let $D^* = \text{block diag } (d_i^* I_{k_i})$, where some of the d_i^* may be infinite, be such that, for some sequence $\{D_k\}$ converging to D^*

$$\lim_{k \rightarrow \infty} \bar{\sigma}(D_k M D_k^{-1}) = \inf_{D \in \mathcal{D}} \bar{\sigma}(D M D^{-1}).$$

Without loss of generality, assume that there exist integers $i, j \in \{1, \dots, m\}$ such that $d_i^* \neq 0$ and $d_j^* = 0$, so that the set

$$I = \{i \in \{1, \dots, m\} : d_i^* \neq 0\}$$

is a nonempty proper subset of $\{1, \dots, m\}$. Since the ij th block of $D M D^{-1}$ is $(d_i/d_j)M_{ij}$, in order that $\bar{\sigma}(D M D^{-1})$ be finite when $D \rightarrow D^*$, it is necessary that, for all $i \in I, j \notin I$,

$$M_{ij} = 0.$$

□

Theorem 3.

$$\mu' = \inf_{D \in \mathcal{D}} \bar{\sigma}(D M D^{-1}).$$

Proof. In view of Lemma 1, it suffices to show that $\mu' \geq \inf_{D \in \mathcal{D}} \bar{\sigma}(D M D^{-1})$. First assume that $\inf_{D \in \mathcal{D}} \bar{\sigma}(D M D^{-1})$ is achieved at D^* and let $M^* =$

D^*MD^{*-1} . From Fact 2 and Proposition 2, it follows that

$$0 \in \text{co}W_{D^*}(\bar{\sigma}^2(M))$$

and, in view of Proposition 3, $0 \in \text{co}W(\bar{\sigma}^2(M^*))$. The claim then follows directly from the definition of μ' . If the infimum in (2) is not achieved, by using Proposition 4, we can always find a matrix E arbitrarily small such that $\inf_{D \in \mathcal{D}} \bar{\sigma}(D(M+E)D^{-1})$ is achieved. The result then follows by continuity. $\square$

Finally, the next proposition shows that the absence of gap between the two quantities in (2) is equivalent to a certain separation condition for $W(\mu^2)$.

Proposition 5.

$$\mu = \inf_{D \in \mathcal{D}} \bar{\sigma}(DM D^{-1}) \quad (8)$$

if, and only if there exists a vector $\lambda \in \mathbb{R}^m$, with $\lambda_i \geq 0$ for $i = 1, \dots, m$, such that $W(\mu^2)$ is contained in the closed half space

$$H(\lambda) = \{v \in \mathbb{R}^m : \langle v, \lambda \rangle \geq 0\}.$$

Furthermore, the infimum is achieved if, and only if all the λ_i 's can be chosen strictly positive. In this case,

$$D^* = \text{block diag}(\sqrt{\lambda_i} I_{k_i})$$

is a minimizer.

Proof. We first prove the second statement. Let $\lambda_i > 0$, $i = 1, \dots, m$, and let $D^* = \text{block diag}(\sqrt{\lambda_i} I_{k_i})$. Using the definition of $W(\mu^2)$, it is easily checked that $W(\mu^2) \subset H(\lambda)$ if, and only if,

$$\sum_{i=1}^m \lambda_i x^H (\mu^2 P_i - M^H P_i M) x \geq 0 \quad \text{for all } x \in \partial B$$

or, equivalently,

$$\mu^2 D^{*2} - M^H D^{*2} M \geq 0.$$

Since D^* is invertible, this can occur if, and only if,

$$\mu^2 I - (D^* M D^{*-1})^H (D^* M D^{*-1}) \geq 0$$

i.e., if, and only if,

$$\bar{\sigma}(D^* M D^{*-1}) \leq \mu.$$

Clearly, in view of (2), this happens if, and only if,

$$\bar{\sigma}(D^* M D^{*-1}) = \mu = \min_{D \in \mathcal{D}} \bar{\sigma}(D M D^{-1}).$$

If the infimum in (8) is not achieved, in view of Proposition 4, we can always find a matrix E arbitrarily small such that $\inf_{D \in \mathcal{D}} \bar{\sigma}(D(M+E)D^{-1})$ is achieved. The result then follows by continuity. $\square$

Referring back to Figure 1, one can check that there does indeed exist a half space $H(\lambda)$ as specified in the proposition.

4 Discussion

By considering the m -form numerical range $W(\alpha)$ of certain matrices, geometric insight has been gained for the computation of the structured singular value μ . In particular, the gap between μ and its upper bound has been related to a certain type of nonconvexity of $W(\alpha)$.

It is easily checked that Theorem 2 still holds if, in Algorithm 1, one replaces $c(\alpha)$ by a function $\tilde{c}(\cdot)$ such that, for some constant $\gamma > 0$, $\gamma c(\alpha) \leq \tilde{c}(\alpha) \leq c(\alpha)$ for all $\alpha \in \mathbb{R}$. Thus the question of obtaining an algorithm to compute μ in the general case would be all but resolved if one comes up with

a way to infallibly check whether or not $W(\alpha)$ contains the origin. Attempts in this direction are in progress [12].

Acknowledgement

The authors wish to thank J.C. Wang for pointing out several errors in an early version of the paper.

Appendix

We describe an algorithm to plot the boundary of $W(\alpha)$ when this set is in $\mathbb{R}^2$ ($m = 2$). Such an algorithm was used to generate the plots of Figure 1. Below, we denote by $\text{int}W(\alpha)$ the interior of $W(\alpha)$ and by $\text{bd}W(\alpha)$ its boundary. Suppose $W(\alpha)$ is strictly convex, i.e., for any $u, v \in W(\alpha)$,

$$\lambda u + (1 - \lambda)v \in \text{int}W(\alpha) \quad \text{for all } \lambda \in (0, 1).$$

Then clearly there is a one to one correspondence between the points of the boundary of $W(\alpha)$ and the support hyperplanes to $W(\alpha)$, namely, for any $u \in \text{bd}W(\alpha)$ there exists a unit vector $h = [\cos \theta \ \sin \theta]^T$ such that u achieves the minimum in

$$\min\{ \langle v, h \rangle : v \in W(\alpha) \}.$$

In view of the definition of $W(\alpha)$, $u_i = x^H A_i(\alpha)x$, for $i = 1, 2$, where x achieves the minimum in

$$\min_{x \in \partial B} \{ x^H (\cos \theta A_1(\alpha) + \sin \theta A_2(\alpha)) x \}$$

i.e., x is a unit length eigenvector corresponding to the smallest eigenvalue of $\cos \theta A_1(\alpha) + \sin \theta A_2(\alpha)$. This leads to the following algorithm.

Algorithm A.

Step 0. Set $\theta = 0$ and $N =$ a large integer.

Step 1. Let x be any unit length eigenvector corresponding to the smallest eigenvalue of $\cos \theta A_1(\alpha) + \sin \theta A_2(\alpha)$. Set

$$y_2 = \begin{bmatrix} x^H A_1(\alpha) x \\ x^H A_2(\alpha) x \end{bmatrix}$$

If $\theta \neq 0$, draw the line segment $\overline{y_1 y_2}$. If $\theta \geq 2\pi$, stop.

Step 2. Set $y_1 = y_2$, $\theta = \theta + 2\pi/N$ and go to Step 1. $\square$

References

- [1] J.C. Doyle, "Analysis of Feedback Systems with Structured Uncertainties," *Proceedings of IEE* 129 (1982), 242–250.
- [2] L. Brickman, "On the Field of Values of a Matrix," *Proc. Amer. Math. Soc.* 12 (1961), 61–66.
- [3] M. Marcus & B.-Y. Wang, "Some Variations on the Numerical Range," *Linear and Multilinear Algebra* 9 (1980), 111–120.
- [4] S. Friedland & R. Loewy, "Subspaces of Symmetric Matrices Containing Matrices with a Multiple First Eigenvalue," *Pacific J. Math.* 62 (1976), 389–399.
- [5] Y.-H. Au-Yeung & Y.-T. Poon, "A Remark on the Convexity and Positive Definiteness Concerning Hermitian matrices," *Southeast Asian Bull. Math.* 3 (1979), 85–92.
- [6] M.K.H. Fan & A.L. Tits, "On the Generalized Numerical Range," *Linear and Multilinear Algebra* (1987), to appear.
- [7] M.K.H. Fan & A.L. Tits, "Characterization and Efficient Computation of the Structured Singular Value," *IEEE Trans. Automat. Control* AC-31 (1986), 734–743.
- [8] M.G. Safonov & J.C. Doyle, "Minimizing Conservativeness of Robustness Singular Values," in *Multivariable Control*, S.G. Tzafestas, ed., D. Reidel Publishing Company, 1984, 197–207.
- [9] B. Kouvaritakis & H. Latchman, "Necessary and Sufficient Stability Criterion for Systems with Structured Uncertainties: the Major Principal Direction Alignment Principle," *Internat. J. Control* 42 (1985), 575–598.

- [10] M.K.H. Fan, "Characterization and Computation of the Structured Singular Value," University of Maryland, PhD Thesis, College Park, MD 20742, 1986.
- [11] J.C. Doyle, private communication, 1984.
- [12] J.C. Wang, M.K.H. Fan & A.L. Tits, "Structured Singular Value and Geometry of the m -Form Numerical Range," presented at the International Symposium on Mathematic Theory of Networks and Systems (MTNS), Pheonix, Arizona, June 1987.
- [13] M.K.H. Fan, "An Algorithm to Compute the Structured Singular Value," Systems Research Center, University of Maryland, College Park, Technical Report TR-86-8, 1986.
- [14] John E. Hauser, "Proximity Algorithms: Theory and Implementation," Electronics Research Laboratory, University of California, Memo No. UCB/ERL M86/53, Berkeley, California, May 1986.