Robustness in the Presence of Mixed Parametric Uncertainty and Unmodeled Dynamics

Michael K. H. Fan, *Member, IEEE,* André L. Tits, *Member, IEEE,* and John C. Doyle

Abstract—**It is shown that, in the case of mixed real parametric and complex uncertainty, the structured singular value can be obtained as the solution of a smooth constrained optimization problem. While this problem may have local maxima, an improved computable upper bound to the structured singular value is derived, leading to a sufficient condition for robust stability and performance.**

Introduction

An inherent tradeoff in modeling is between fidelity and simplicity. It is desirable to have models which closely match reality, yet are still easy to analyze. This tradeoff arises in modeling uncertainty. For example, a single norm bounded perturbation simplifies analysis but may be too conservative. Introducing more structure may improve the model fidelity but typically complicates the analysis. It is often very natural to model uncertainty with real perturbations when, for exmaple, the real coefficients of a differential equation model are uncertain. It is important, however, to remember that such parametric variations are in a model, not in the physical system being modeled. Models with real parametric undertainty are used because, in principle, they allow more accurate representation of some systems.

The structured singular value (SSV or μ) was introduced to study structured uncertainty in linear models [1]–[3]. It is defined in such a way as to give a precise characterization of robust stability and performance, in an H_∞ sense, for a rich variety of uncertainty descriptions (Small μ Theorem [2]). There is a large literature on robustness, particularly for unstructured uncertainty, and connections with H_∞ theory. A brief historical review of the literature most directly influencing the development of μ-based methods will be given after the notation is introduced.

Much of the subsequent research on μ has focused on computational schemes, with reasonable success for problems involving only complex uncertainty. Complex perturbations are typically used to represent uncertainty due to unmodeled dynamics, or to "cover" the variations produced by several real parameters. In the μ framework, complex uncertain blocks also arise for problems of robust performance, and thus practical applications of μ always involve at least one complex block. Although there are important outstanding issues to be resolved in computation of μ for complex perturbations, substantial progress has been made and μ is being applied routinely to large engineering problems. This paper focuses on the computation of μ for mixed real parametric and complex uncertainty, which is fundamentally more difficult than for complex perturbations.

The major issues in computing μ, or its equivalent, are the generality of the problem description, the exactness of analysis, and the ease of computation. Existing methods for real perturbations emphasize just two of these three issues. A general and, in principle, exact method is a brute force global search using a grid of parameter values (e.g., [4], [5]). This inevitably involves an exponential growth in computation as a function of the number of parameters, and taking fewer grid points to avoid this gives up exactness. This "exponential explosion" limits the usefulness of exhaustive global search methods, although simple search in some form will always play an important role in practical control design.

Tentatively, a reasonable requirement on a computational scheme would seem to be that the typical time to compute solutions scales in some polynomial way with the size of the problem. This is compatible with the conventional view in computational linear algebra, where, for example, the QZ iteration for computing singular values is considered an acceptable approach. From this point of view, progress is being made in reducing the computational burden of exact methods [6]–[8], but no general, exact, polynomial-time algorithms are available. Nevertheless, this work suggests some promising research directions, and these will be discussed in the final section.

An approach to obtaining exact results with more modest computation is to restrict the problem description. The best example is Kharitonov's celebrated result for polynomials with coefficients in intervals [9]. While few models with engineering motivation fit the allowable problem description, Kharitonov's theorem inspired a great deal of research. Progress is being made in this direction by allowing more general uncertainty descriptions at the expense of more computation (e.g., [10]–[12]). Unfortunately, the results so far indicate that even a modest departure from the interval polynomial problem leads to exponential explosion.

The approach taken in this paper could be characterized as being very general and computationally attractive, but potentially inexact. Following the methods developed for μ in the case of complex perturbations, the main idea is to get upper and

Manuscript received August 15, 1988; revised August 2, 1989 and April 7, 1990. Paper recommended by Past Associate Editor, J. D. Cobb. The work of M. K. H. Fan and A. L. Tits was supported in part by the National Science Foundation under the Engineering Research Centers Program NSFD-CDR-88-03012 and by the NSF under Grant DMC-84-51515. The work of J. Doyle was supported by NASA, NSF, ONR, and the AFOSR.

M. K. H. Fan was with the University of Maryland, College Park, MD 20742. He is now with the School of Electrical Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0250.

A. L. Tits is with the Department of Electrical Engineering and Systems Research Center, University of Maryland, College Park, MD 20742.

J. C. Doyle is with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125.

IEEE Log Number 9040775.

0018-9286/91/0100-0025$01.00 © 1991 IEEE

lower bounds using local search methods which are computationally inexpensive, but may fail to find global solutions. One then seeks to prove that the local methods yield global solutions, or that the bounds one gets are tight enough to be of value in problems of interest. The strategy taken in this paper has been very successful in the case when all perturbations are complex and appears to have promise for the general case as well, although it is clear that the latter is much more challenging. In the final section, some speculations will be presented on how the results in this paper might be combined with the work in [6]–[8] to obtain an approach which is general, efficient, and exact.

The balance of the paper is organized as follows. In Section I, the SSV framework is introduced and a brief historical review of the literature most influencing the development of μ-based methods is given. In Section II, it is shown that μ can be obtained as the optimal value of a smooth constrained optimization problem. Geometric interpretation of this result is discussed in Section III. The framework of Sections II and III is used in Section IV to derive a new computable upper bound on μ. In Section V, the new bound is shown to be mathematically equivalent to that given in [3] for the case of nonrepeated real scalar blocks. A serious and valid criticism of the literature on μ is the lack of tutorial material, thus limiting the readership primarily to experts. While it is beyond the scope of this paper to correct this deficiency, Section VI contains some simple examples and numerical experiments which should both motivate and illustrate the use of the theory. Also, the next section give some historical background. Finally, Section VII offers some speculation about the future directions for research in this area. All proofs are given in Appendix A. Appendix B contains data used in the experiments of Section VI.

Many of the results of this paper have appeared, without proof, in [13].

I. Preliminaries

A. Framework

Throughout the paper, given any square complex matrix M, we denote by $\bar{\sigma}(M)$ its largest singular value, by $\underline{\sigma}(M)$ its smallest singular value, by $\overline{M}$ its complex conjugate, and by M^H its complex conjugate transpose, and we let $\rho_R(M) = \max\{|\lambda|: \lambda \text{ is a real eigenvalue of } M\}$, with $\rho_R(M) = 0$ if M has no real eigenvalue. If M is Hermitian, we denote by $\bar{\lambda}(M)$ its largest eigenvalue. Given any complex vector x, x^H indicates its complex conjugate transpose and $\|x\|$ its Euclidean norm. The empty set will be denoted by $\varnothing$. Finally, while j (italics) will be used as a running index, $\boldsymbol{j}$ (bold italic) will denote $\sqrt{-1}$.

Given an $n \times n$ complex matrix M and three nonnegative integers m_r, m_c, and m_C, with $m := m_r + m_c + m_C \le n$, *a block structure* $\mathscr{K}$ of dimensions (m_r, m_c, m_C) *associated with* M is an m-tuple of positive integers.

$$\mathscr{K} = (k_1, \cdots, k_{m_r}; k_{m_r+1}, \cdots, k_{m_r+m_c}; k_{m_r+m_c+1}, \cdots, k_m) \tag{1.1}$$

such that $\sum_{q=1}^{m} k_q = n$. Given a block structure $\mathscr{K}$, consider the family of block-diagonal $n \times n$ matrices

$$\mathscr{X}_{\mathscr{K}} = \Big\{\text{block diag}\Big(\delta_1^r I_{k_1}, \cdots, \delta_{m_r}^r I_{k_{m_r}}, \delta_1^c I_{k_{m_r+1}}, \cdots, \delta_{m_c}^c I_{k_{m_r+m_c}}, \Delta_1^C, \cdots, \Delta_{m_C}^C\Big): \delta_q^r \in \mathbb{R}, \delta_q^c \in \mathbb{C}, \Delta_q^C \in \mathbb{C}^{k_{m_r+m_c+q} \times k_{m_r+m_c+q}}\Big\}$$

where for any integer k, I_k denotes the $k \times k$ identity matrix. The "repeated real scalar" blocks $\delta_q^r I_{k_q}$ correspond to parametric uncertainty, one "repeated complex scalar" block $\delta_q^c I_{k_{m_r+q}}$ can be used to represent frequency (see [3] for details; several blocks of the latter type are introduced here mostly for the sake of uniformity) and the "full complex" blocks Δ_q^C correspond to unmodeled dynamics.[1]

Definition 1.1. [1]: The *structured singular value* $\mu_{\mathscr{K}}(M)$ of a complex $n \times n$ matrix M with respect to block-structure $\mathscr{K}$ is 0 if there is no Δ in $\mathscr{X}_{\mathscr{K}}$ such that $\det(I - \Delta M) = 0$, and

$$\Big(\min_{\Delta \in \mathscr{X}_{\mathscr{K}}} \{\bar{\sigma}(\Delta): \det(I - \Delta M) = 0\}\Big)^{-1} \tag{1.2}$$

otherwise. □

Directly from Definition 1.1, it is easily shown that

$$\rho_R(M) \le \mu_{\mathscr{K}}(M) \le \bar{\sigma}(M) \tag{1.3}$$

and that, for any $U, V \in \mathscr{U}_{\mathscr{K}}$, $D \in \mathscr{D}_{\mathscr{K}}$

$$\mu_{\mathscr{K}}(M) = \mu_{\mathscr{K}}(DUMVD^{-1}) \tag{1.4}$$

where

$$\mathscr{U}_{\mathscr{K}} = \{U \in \mathscr{X}_{\mathscr{K}}: UU^H = I\}$$

and

$$\mathscr{D}_{\mathscr{K}} = \Big\{\text{block diag}\big(D_1, \cdots, D_{m_r+m_c}, d_1 I_{k_{m_r+m_c+1}}, \cdots, d_{m_C} I_{k_m}\big): 0 < D_q = D_q^H \in \mathbb{C}^{k_q \times k_q}, d_q > 0\Big\}.$$

Combining these two sharpens (1.3) to[2]

$$\max_{U \in \mathscr{U}_{\mathscr{K}}} \rho_R(MU) \le \mu_{\mathscr{K}}(M) \le \inf_{D \in \mathscr{D}_{\mathscr{K}}} \bar{\sigma}(DMD^{-1}). \tag{1.5}$$

For the purely complex case ($m_r = 0$), the ρ_R inequality is always an equality and the $\bar{\sigma}$ inequality is an equality when $2m_c + m_C < 3$ [1], [14]. The ρ_R expression typically has nonglobal local maxima while every local minimum of the $\bar{\sigma}$ expression is global. Extensive computational experience has suggested that it is easy to obtain $U \in \mathscr{U}_{\mathscr{K}}$ making $\rho_R(MU)$ close to the latter, even when $m_C \gg 1$. These bounds formed the basis for early computational approaches to μ for $m_r = 0$, because local search methods could be used to make the bounds reasonably tight [1], [15], [16].

Unfortunately, when $m_r > 0$ the bounds in (1.5) may be arbitrarily far off, even for problems with engineering motivation. In [3], an improved upper bound was obtained but no practical way to compute it was given. This paper provides an alternative maximization to the ρ_R expression which is equal to $\mu_{\mathscr{K}}(M)$ at its global maximum. The new expression suggests a geometric interpretation based on the concept of "multiform numerical range." Also, an upper bound mathematically equivalent to the one in [3] is obtained, but with much better computational properties.

[1] Note that nonrepeated complex scalars δ^c can be viewed indifferently as repeated complex scalar blocks $\delta^c I_k$, with $k = 1$, or as full complex blocks $\Delta^C \in \mathbb{C}^{1 \times 1}$.

[2] Note that, if $m_r = 0$, ρ_R can be equivalently replaced by the spectral radius ρ.

B. Historical Perspective

In this section, we will briefly review the ideas that most influenced the original development of the μ theory. These remarks are essentially from earlier papers [1], [2], but are repeated here for the convenience of the reader.

An obvious influence was the work in so-called robust multivariable control systems from the late 1970's (see, for example, [17]) which in turn drew heavily on earlier work in stability analysis (e.g., [18]–[22]), particularly, small gain and circle theorems. These theorems established sufficient conditions for stability of nonlinear components connected in feedback. The emphasis in the robustness work was on small gain type conditions involving singular values that were *both* necessary and sufficient for stability of *sets* of linear systems involving a single norm bounded but otherwise unconstrained perturbation. Another emphasis for much of the robustness theory was on using singular value plots to generalize Bode magnitude plots to multivariable systems.

While methods based on singular values were gaining in popularity, it became evident that their assumption of unstructured uncertainty was too crude for many applications. Furthermore, the problem of robust performance was not adequately treated. Freudenberg *et al.* [23] studied these issues using differential sensitivity and suggested that something more than singular values was needed. It was a natural step to introduce structured uncertainty of the type considered in this paper (see [24] for an early treatment). The so-called conservativeness of singular values rested on the fact that the bounds in (1.3) could be arbitrarily far off, and research was begun to provide improved estimates of μ, with an initial focus on the nonrepeated, complex case ($m_r = m_c = 0$).

It was obvious that the sharper bounds in (1.5) could help alleviate the conservativeness somewhat. These bounds were similar to the multiplier methods that were used in nonlinear stability analysis to reduce the conservativeness of small gain type methods [21], but the use of both upper and lower bounds, and the questions of how close the bounds were and how to efficiently compute them were new and open. That the lower bound is equal to μ is relatively straightforward and not surprising. What is remarkable, even in retrospect, is that the upper bound is also an equality for $m_C \leq 3$ and close for $m_C \geq 4$. The equality results were first proven in [1], while the $m_C \geq 4$ case has only experimental evidence. Although by now that evidence is extensive, it remains an important open question to further characterize the exact nature of the upper bound for $m_C \geq 4$.

There was substantial numerical evidence for the upper bound results some time before they were proven. Engineers at Honeywell's Systems and Research Center, particularly J. Wall, began routinely using a simple generalization of Osborne's routine [25] to approximate the upper bound in (1.5) and gradient search to find a local maximum for the lower bound. Osborne's algorithm minimizes the Frobenius norm rather than the maximum singular value, and the scalings produced can be used to approximate the upper bound. The consistent closeness of the bounds, usually within a few percent, suggested that there was a deeper connection between the bounds. Ironically, minimizing the Frobenius norm remains the cheapest method of approximating the upper bound. Safonov [26] suggested a somewhat less general approximation to the upper bound based on Perron eigenvectors which is comparable to Osborne's in speed and accuracy.

While the μ framework arises naturally in studying robust stability with structured uncertainty, it also can be used directly to treat the problem of robust *performance* with structured uncertainty [2]. This is a consequence of the intimate connection between μ and linear fractional transformations (LFT) [3], [14]. In retrospect, it is clear that Redheffer had developed the foundation of this connection in his work on LFT in the late 1950's [27], [28]. In fact, Redheffer had even proven that the upper bound in (1.5) was an equality for the case where $m_r = m_c = 0$ and $m_C = 2$. While Redheffer's results were not well known in the control community until the μ theory was already well developed, the rediscovery of his work has since had an important influence, not only on the further development of μ but in other areas as well (e.g., see [29]).

II. A Smooth Optimization Problem

Definition 1.1 suggests that one consider matrices $\Delta \in \mathscr{X}_{\mathscr{K}}$ such that, for some nonzero x

$$\Delta M x = x. \tag{2.1}$$

Without loss of generality, x has unit length, i.e.,

$$x \in \partial B := \{x \in \mathbb{C}^n : \|x\| = 1\}.$$

In view of the structure of Δ, (2.1) imposes some constraints on "subvectors" of x, corresonding to x being split according to structure $\mathscr{K}$. To help the reader's intuition, we first consider the case of uncertainty consisting of a single, possibly repeated, uncertain real parameter and a single complex block.

A. Key Ideas in a Special Case

As a simple special case, consider the block structure

$$\mathscr{K} = (k_r; ; k_C) \tag{2.2}$$

and the corresponding family of matrices

$$\mathscr{X}_{\mathscr{K}} = \{\text{block diag}\,(\delta^r I_{k_r}, \Delta^c) : \delta^r \in \mathbb{R},\ \Delta^c \in \mathbb{C}^{k_C \times k_C}\}.$$

Let $Q_r \in \mathbb{R}^{k_r \times n}$, $Q_C \in \mathbb{R}^{k_C \times n}$ be projection matrices defined by

$$Q_r = [I_{k_r} \quad O], \qquad Q_C = [O \quad I_{k_C}].$$

The constraints on x implied by (2.1) are

$$\delta^r Q_r M x = Q_r x \tag{2.3}$$

$$\Delta^c Q_C M x = Q_C x. \tag{2.4}$$

In order for (2.4) to be achieved for some Δ^c, $\bar{\sigma}(\Delta^c) \leq b$, for given $b \geq 0$,it is necessary and sufficient that

$$b\|Q_C M x\| \geq \|Q_C x\|.$$

Similarly, for (2.3) to be achieved for some $|\delta^r| \leq b$, it is necessary (but no more sufficient) that

$$b\|Q_r M x\| \geq \|Q_r x\|. \tag{2.5}$$

Equation (2.1) now implies additional constraints, namely (if $x_i \neq 0$, $i = 1, \cdots, k_r$)

$$\frac{(Mx)_i}{x_i} = \frac{(\overline{Mx})_j}{\bar{x}_j}, \qquad i, j = 1, \cdots, k_r \tag{2.6}$$

(including the case $i = j$), as these are equivalent to the existence of $\zeta \in \mathbb{R}$ such that

$$\zeta(Mx)_i = x_i, \qquad i = 1, \cdots, k_r.$$

Thus, $\mu_{\mathscr{X}}(M) = b_*^{-1}$, with b_* the smallest b for which some x exists that satisfies all these constraints. Letting $\theta = b^{-1}$ and removing the assumption that x has nonzero components, one obtains the following results for structure (2.2), a special case of Theorem 2.1 below:

$$\mu_{\mathscr{X}}(M) = \begin{cases} 0 & \text{if } \mathscr{S}_{\mathscr{X}}(M) = \varnothing \\ \max\{\theta : \|Q_r Mx\| \geq \theta \|Q_r x\|, \|Q_C Mx\| \geq \theta \|Q_C x\| \quad \text{for some } x \in \mathscr{S}_{\mathscr{X}}(M)\} & \text{otherwise} \end{cases}$$

with

$$\mathscr{S}_{\mathscr{X}}(M) = \{x \in \partial B : \bar{x}_j(Mx)_i = x_i(\overline{Mx})_j, i, j = 1, \cdots, k_r\} \qquad (m_r = m_C = 1, m_c = 0).$$

Note that the constraint

$$\bar{x}_j(Mx)_i = x_i(\overline{Mx})_j, \qquad i, j = 1, \cdots, k_r$$

can be equivalently expressed as

$$x^H M^H E^{ij} x = x^H E^{ij} Mx \qquad i, j = 1, \cdots, k_r$$

where E^{ij} is any $n \times n$ matrix whose only nonzero entry is in position (i, j), or as

$$x^H M^H G x = x^H G M x \qquad \text{for all } G \in \{\text{block diag}\,(G_{k_r}, O)\}$$

where G_{k_r} ranges over the set of all $k_r \times k_r$ complex matrices or equivalently over the set of all $k_r \times k_r$ Hermitian matrices (as any complex matrix M can be decomposed as $M = M_1 + jM_2$, with M_1, M_2 Hermitian).

B. General Case

It is readily checked that repeated *complex* scalar blocks ($m_c > 0$) imply constraints of the form (to be compared to (2.6) above)

$$\frac{(Mx)_i}{x_i} = \frac{(Mx)_j}{x_j}$$

with i and j ranging over indexes corresponding to the block under consideration. With this in mind, extension to the case of a general block structure presents mostly notational rather than conceptual difficulty. Consider the projection matrices Q_q, $q = 1, \cdots, m$, defined by

$$Q_q = \text{block row}\,(O_{k_q \times k_1}, \cdots, O_{k_q \times k_{q-1}}, I_{k_q}, O_{k_q \times k_{q+1}}, \cdots, O_{k_q \times k_m})$$

where, for any positive integers k, k', O_k is the $k \times k$ zero matrix and $O_{k \times k'}$ is the $k \times k'$ zero matrix. Also, for $q = 1, \cdots, m_r + m_c$, consider the index set J_q defined by

$$J_q = \left\{ \sum_{p=1}^{q-1} k_p + 1, \sum_{p=1}^{q-1} k_p + 2, \cdots, \sum_{p=1}^{q} k_p \right\}.$$

The result obtained in Section II-A is then generalized as follows (see proof in Appendix A).

Theorem 2.1: For any matrix M and associated structure $\mathscr{X}$

$$\mu_{\mathscr{X}}(M) = \begin{cases} 0 & \text{if } \mathscr{S}_{\mathscr{X}}(M) = \varnothing \\ \max\{\theta : \|Q_q Mx\| \geq \theta \|Q_q x\|, q = 1, \cdots, m \quad \text{for some } x \in \mathscr{S}_{\mathscr{X}}(M)\} & \text{otherwise} \end{cases} \tag{2.7}$$

with

$$\mathscr{S}_{\mathscr{X}}(M) = \left\{ x \in \partial B : x_i(\overline{Mx})_j = \bar{x}_j(Mx)_i, (i, j) \in \bigcup_{q=1}^{m_r} J_q \times J_q; \; x_i(Mx)_j = x_j(Mx)_i, (i, j) \in \bigcup_{q=m_r+1}^{m_r+m_c} J_q \times J_q \right\}. \quad \square$$

Formula (2.7) for $\mu_{\mathscr{X}}(M)$ amounts to a constrained maximization over θ and x in $\mathbb{R} \times \mathbb{C}^n$. It has some definite computational advantages over the formula defining $\mu_{\mathscr{X}}(M)$ in Definition 1.1. The number of variables is limited, the objective and constraints are inexpensive to evaluate and, after squaring all the norms, objective and constraints become smooth. However, again, (2.7) may have local maxima which are not global and it is not clear whether the global maximum can be easily obtained.

Remark 2.1: Finding a point $x \in \mathscr{S}_{\mathscr{X}}(M)$ may not always be simple. Yet, in the following two cases, such a point is readily available: i) if $m_C > 0$ then $x \in \mathscr{S}_{\mathscr{X}}(M)$ whenever $\|x\| = 1$ and $x_i = 0$ for all $i \in J_q$, $q = 1, \cdots, m_r + m_c$; ii) if M has a real eigenvalue, then $x \in \mathscr{S}_{\mathscr{X}}(M)$ with x any corresponding unit length eigenvector. $\square$

Remark 2.2: It should be clear that in the purely real case, i.e., when M is a real matrix matrix and $m_c = m_C = 0$, (2.7) still holds if x is restricted to be real. $\square$

Remark 2.3: Some of the constraints defining $\mathscr{S}_{\mathscr{X}}(M)$ may seem to be redundant. In fact, it can be shown that (2.7) is no longer valid if any of these constraints is removed but that for a generic matrix M (in particular, for one with no zero entries), roughly 25% of the constraints are indeed redundant. $\square$

Remark 2.4: Following [3], let us define the spectrum of M with respect to $\mathscr{X}$ as

$$\text{sp}_{\mathscr{X}}(M) = \{\Delta \in \mathscr{X}_{\mathscr{X}} : \det(M - \Delta) = 0\}$$

and let

$$\gamma_{\mathscr{X}}(M) = \begin{cases} 0 & \text{if } \text{sp}_{\mathscr{X}}(M) = \varnothing \\ \sup\limits_{\Delta \in \text{sp}_{\mathscr{X}}(M)} \underline{\sigma}(\Delta) & \text{otherwise.} \end{cases}$$

Then it is simple to check that

$$\gamma_{\mathscr{X}}(M) = \sup_{\substack{\theta > 0 \\ \theta_1, \cdots, \theta_m \geq 0}} \{\theta : \|Q_q Mx\| = \theta_q \|Q_q x\| \quad \theta_q \geq \theta, q = 1, \cdots, m \quad \text{for some } x \in \mathscr{S}_{\mathscr{X}}(M)\} \tag{2.8a}$$

if the feasible set in (2.8a) is nonempty, and

$$\gamma_{\mathscr{X}}(M) = 0 \tag{2.8b}$$

otherwise, Clearly (2.8) is very similar to (2.7). Yet, as pointed out in [3], while $\gamma_{\mathscr{X}}(M) = \mu_{\mathscr{X}}(M)$ if $m_r = 0$ (no parametric uncertainty), equality does not hold in general, but rather

$$\gamma_{\mathscr{X}}(M) \le \mu_{\mathscr{X}}(M). \tag{2.9}$$

□

In the case when $m_c = 0$, $\mathscr{S}_{\mathscr{X}}(M)$ can be expressed in a simple form, to be used in the following sections.

Proposition 2.1: Let $\mathscr{G}_{\mathscr{X}}$ be the family of Hermitian matrices

$$\mathscr{G}_{\mathscr{X}} = \left\{\text{block diag}\left(G_1, \cdots, G_{m_r}, O_{k_{m_r+1}}, \cdots, O_{k_m}\right): \quad G_q = G_q^H \in \mathbb{C}^{k_q \times k_q}\right\}$$

and let $\mathscr{E}_{\mathscr{X}}$ be any basis for $\mathscr{G}_{\mathscr{X}}$. Then

$$\begin{aligned}\mathscr{S}_{\mathscr{X}}(M) &\subseteq \left\{x \in \partial B: x^H(M^H G - GM)x = 0 \quad \forall\, G \in \mathscr{G}_{\mathscr{X}}\right\} \\ &= \left\{x \in \partial B: x^H(M^H E - EM)x = 0 \quad \forall\, E \in \mathscr{E}_{\mathscr{X}}\right\}.\end{aligned} \tag{2.10}$$

If $m_c = 0$, then $\mathscr{S}_{\mathscr{X}}(M)$ is equal to the right-hand side of (2.10). □

III. Interpretation in Terms of the Multiform Numerical Range

Formula (2.7) leads to a characterization of the structured singular value in terms of the multiform numerical range of some matrices. The *multiform numerical range* (or *t-form numerical range*) of a t-tuple of $n \times n$ Hermitian matrices $A_1, \cdots, A_t$ is the set

$$W(A_1, \cdots, A_t) = \{f(x): x \in \partial B\}$$

where $f: \mathbb{C}^n \to \mathbb{R}^t$ has components

$$f_q(x) = x^H A_q x, \qquad q = 1, \cdots, t.$$

First, suppoes that $m_r = m_c = 0$. For $\alpha \in \mathbb{R}$, let

$$A_q(\alpha) = \alpha Q_q^T Q_q - M^H Q_q^T Q_q M, \qquad q = 1, \cdots, m_C \tag{3.1}$$

and let $W_{\mathscr{X}}(\alpha)$ be the m_C-form numerical range $W(A_1(\alpha), \cdots, A_{m_C}(\alpha))$. Then[3]

$$\mu_{\mathscr{X}}(M) = \inf_{\alpha \ge 0} \left\{\sqrt{\alpha} : 0 \notin W_{\mathscr{X}}(\alpha) + \mathbb{R}_+^{m_C}\right\} \qquad (m_r = m_c = 0) \tag{3.2}$$

with

$$\mathbb{R}_+^{m_C} = \left\{v \in \mathbb{R}^{m_C}: v_q \ge 0,\ q = 1, \cdots, m_C\right\}.$$

This follows rather directly if one rewrites the constraints in (2.7) as

$$x^H A_q(\theta^2) x \le 0, \qquad q = 1, \cdots, m_C.$$

Suppose now that $m_r \ne 0$. For $q = 1, \cdots, m$ $(m = m_r + m_C)$ and $\alpha \in \mathbb{R}$, let

$$A_q(\alpha) = \alpha Q_q^T Q_q - M^H Q_q^T Q_q M$$

[3]A related result was obtained in [30].

and for $i = 1, \cdots, \sum_{i=1}^{m_r} k_i^2$, let

$$A_{m+q}(\alpha) = j\left(M^H E_q - E_q M\right)$$

where the E_q's are the elements of a basis $\mathscr{E}_{\mathscr{X}}$ of $\mathscr{G}_{\mathscr{X}}$ taken in some arbitrary order, and where the argument α is used for the sake of uniformity of notation. Then $A_q(\alpha) = A_q(\alpha)^H$, $q = 1, \cdots, s$, with

$$s = m + \sum_{i=1}^{m_r} k_i^2$$

and $\mu_{\mathscr{X}}(M) = 0$ if $\mathscr{S}_{\mathscr{X}}(M) = \varnothing$ and

$$\mu_{\mathscr{X}}(M) = \max_{\substack{x \in \partial B \\ \alpha \ge 0}} \left\{\sqrt{\alpha} : x^H A_q(\alpha) x \le 0,\ q = 1, \cdots, m\right\}$$

otherwise. Denoting by $W_{\mathscr{X}}(\alpha)$ the multiform numerical range associated with $A_1(\alpha), \cdots, A_s(\alpha)$, i.e.,

$$W_{\mathscr{X}}(\alpha) = \left\{v \in \mathbb{R}^s: \exists\, x \in \partial B \text{ s.t. } v_q = x^H A_q(\alpha) x, \quad q = 1, \cdots, s\right\}$$

we obtain the following result, to be compared to (3.2). Here, the set $\mathscr{P}_m \subset \mathbb{R}^s$ is defined by

$$\mathscr{P}_m = \left\{v \in \mathbb{R}^s: v_q \ge 0,\ q = 1, \cdots, m;\ v_q = 0,\ q = m+1, \cdots, s\right\}.$$

Theorem 3.1:

$$\mu_{\mathscr{X}}(M) = \inf_{\alpha \ge 0} \left\{\sqrt{\alpha} : 0 \notin W_{\mathscr{X}}(\alpha) + \mathscr{P}_m\right\} \qquad (m_c = 0).$$

□

Let us now define, for any $\alpha \in \mathbb{R}$

$$c_{\mathscr{X}}(\alpha) = \min\left\{N(v): v \in W_{\mathscr{X}}(\alpha) + \mathscr{P}_m\right\}$$

where $N(\cdot)$ is any given norm in $\mathbb{R}^s$ satisfying $N(e_q) \le 1$, with $\{e_q\}$ the canonical basis, and let us consider the following algoritghm.

Algorithm 3.1: (Computation of $\mu_{\mathscr{X}}(M)$ when $m_c = 0$).

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

Step 1: Set $\alpha_{k+1} = \alpha_k - c_{\mathscr{X}}(\alpha_k)$.

Step 2: Set $k = k + 1$ and go to Step 1. □

A key property of $c_{\mathscr{X}}(\cdot)$ (see [30], [31]) still holds here (see proof in Appendix A).

Proposition 3.1: $c_{\mathscr{X}}(\cdot)$ is continuous and, for any $\beta \ge 0$ and $\alpha \in \mathbb{R}$, $c_{\mathscr{X}}(\alpha + \beta) \le c_{\mathscr{X}}(\alpha) + \beta$. □

The following convergence result follows (see [30]).

Theorem 3.2: The sequence $\{\alpha_k\}$ generated by Algorithm 3.1 is monotone nonincreasing and

$$\lim_{k \to \infty} \sqrt{\alpha_k} = \mu_{\mathscr{X}}(M).$$

□

Algorithm 3.1 can be implemented whenever $W_{\mathscr{X}}(\alpha)$ is convex. Since the multiform numerical range of no more than 3 matrices is always convex (provided these matrices have size at least 3×3) [32]–[34], this will always be the case when $s \le 3$, i.e., in the case of 3 or fewer complex blocks ($m_r = 0, m_c = 0$, $m_C \le 3$) or 1 real scalar block and one or no complex block ($m_r = 1$, $k_1 = 1$, $m_c = 0$, $m_C \le 1$).

IV. A Computable Upper Bound

Given any block structure $\mathscr{X}$, the corresponding structured singular value is no greater than the largest singular value, as the

latter corresponds to the least restrictive structure $\mathscr{K} = (;;n)$ [see (1.3)]. The following proposition, a consequence of Theorem 2.1 and Proposition 2.1 (see the proof in Appendix A), provides two intermediate bounds.

Proposition 4.1: For any matrix M and associated block structure $\mathscr{K}$

$$\mu_{\mathscr{K}}(M) \le \eta_{\mathscr{K}}(M) \le \nu_{\mathscr{K}}(M) \le \bar{\sigma}(M) \tag{4.1}$$

where

$$\eta_{\mathscr{K}}(M) = \begin{cases} 0 & \text{if } \mathscr{S}_{\mathscr{K}}(M) = \varnothing \\ \max\limits_{x \in \mathscr{S}_{\mathscr{K}}(M)} \|Mx\| & \text{otherwise} \end{cases}$$

and

$$\nu_{\mathscr{K}}(M) = \sqrt{\max\left\{0, \inf_{G \in \mathscr{G}_{\mathscr{K}}} \bar{\lambda}\left[M^H M + j(GM - M^H G)\right]\right\}}.$$

□

The following theorem is a direct consequence of (1.4) and Proposition 4.1.

Theorem 4.1: For any matrix M and associated block structure $\mathscr{K}$

$$\begin{aligned} \mu_{\mathscr{K}}(M) &\le \inf_{D \in \mathscr{D}_{\mathscr{K}}} \eta_{\mathscr{K}}(DMD^{-1}) \\ &\le \inf_{D \in \mathscr{D}_{\mathscr{K}}} \nu_{\mathscr{K}}(DMD^{-1}) \\ &\le \inf_{D \in \mathscr{D}_{\mathscr{K}}} \bar{\sigma}(DMD^{-1}). \end{aligned} \tag{4.2}$$

□

Theorem 4.1 gives two upper bounds which are less conservative than $\inf_{D \in \mathscr{D}_{\mathscr{K}}} \bar{\sigma}(DMD^{-1})$. However, since $\max_{x \in \mathscr{S}_{\mathscr{K}}(DMD^{-1})} \|DMD^{-1}x\|$ may have local maxima that are not global, attempts to evaluate $\eta_{\mathscr{K}}(DMD^{-1})$ may yield strict lower bounds on this quantity and this may result in underestimation of $\mu_{\mathscr{K}}(M)$. Fortunately, the second upper bound in (4.2) does not suffer from this shortcoming. Indeed, we can write

$$\inf_{D \in \mathscr{D}_{\mathscr{K}}} \nu_{\mathscr{K}}(DMD^{-1}) = \sqrt{\max\left\{0, \inf_{D \in \mathscr{D}_{\mathscr{K}}, G \in \mathscr{G}_{\mathscr{K}}} F(D, G)\right\}} \tag{4.3}$$

where F: $\mathscr{D}_{\mathscr{K}} \times \mathscr{G}_{\mathscr{K}} \to \mathbb{R}$ is defined by

$$F(D, G) = \bar{\lambda}\left[M_D^H M_D + j(GM_D - M_D^H G)\right] \tag{4.4}$$

with $M_D = DMD^{-1}$. Thus, for any $D \in \mathscr{D}_{\mathscr{K}}$, $G \in \mathscr{G}_{\mathscr{K}}$, unless $F(D, G) < 0$ (in which case $\mu_{\mathscr{K}}(M) = 0$), $\sqrt{F(D, G)}$ is an upper bound for $\mu_{\mathscr{K}}(M)$. The same upper bound can be obtained by means of a computationally simpler problem as seen next. For any $\alpha \in \mathbb{R}$ let Φ_α: $\mathscr{D}_{\mathscr{K}} \times \mathscr{G}_{\mathscr{K}} \to \mathbb{R}$ be defined by

$$\Phi_\alpha(D, G) = \bar{\lambda}\left[M^H DM + j(GM - M^H G) - \alpha D\right]. \tag{4.5}$$

Proposition 4.2: i) For any $D \in \mathscr{D}_{\mathscr{K}}$, $G \in \mathscr{G}_{\mathscr{K}}$

$$F(D, G) = \max_{\alpha \in \mathbb{R}} \{\alpha: \Phi_\alpha(D^2, DGD) \ge 0\}$$

which is the only value α for which $\Phi_\alpha(D^2, DGD) = 0$. ii) Moreover

$$\inf_{D \in \mathscr{D}_{\mathscr{K}}, G \in \mathscr{G}_{\mathscr{K}}} F(D, G) = \inf_{D \in \mathscr{D}_{\mathscr{K}}, G \in \mathscr{G}_{\mathscr{K}}} \max_{\alpha \in \mathbb{R}} \{\alpha: \Phi_\alpha(D, G) \ge 0\}.$$

□

Notice that Φ_α is the composition of a convex function ($\bar{\lambda}$) and an affine function, and thus is convex. It can be shown that any local minimizer (D, G) for

$$\inf_{D \in \mathscr{D}_{\mathscr{K}}, G \in \mathscr{G}_{\mathscr{K}}} \max_{\alpha \in \mathbb{R}} \{\alpha: \Phi_\alpha(D, G) \ge 0\}$$

is global and this problem can be solved by means of a simple algorithm [35].

Finally, there are noteworthy instances where the new upper bound is *equal* to the structured singular value. The next proposition establishes a sufficient condition on Φ_α for this to hold.

Proposition 4.3: Let $m_c = 0$. Let $(D_*, G_*) \in \mathscr{D}_{\mathscr{K}} \times \mathscr{G}_{\mathscr{K}}$ and let

$$\alpha_* = \max\{\alpha: \Phi_\alpha(D_*, G_*) \ge 0\}.$$

If Φ_{α_*} is differentiable at (D_*, G_*) with vanishing derivative, then

$$\begin{aligned} \mu_{\mathscr{K}}(M) &= \min_{D \in \mathscr{D}_{\mathscr{K}}} \nu_{\mathscr{K}}(DMD^{-1}) = \nu_{\mathscr{K}}(D_*^{1/2} D M_*^{-1/2}) \\ &= \sqrt{\max\{0, F(D_*^{1/2}, D_*^{-1/2} G_* D_*^{-1/2})\}} \end{aligned}$$

where $D_*^{1/2}$ is the positive definite square root of D_* and $D^{-1/2}_*$ is its inverse. □

An important case where the conditions of Proposition 4.3 hold is as follows.

Theorem 4.2: Let $m_c = 0$. Suppose that $\inf_{D \in \mathscr{D}_{\mathscr{K}}, G \in \mathscr{G}_{\mathscr{K}}} F(D, G)$ is achieved, say at $(\hat{D}, \hat{G})$, and that the corresponding largest eigenvalue in (4.4) is simple. Then

$$\begin{aligned} \mu_{\mathscr{K}}(M) &= \inf_{D \in \mathscr{D}_{\mathscr{K}}} \nu_{\mathscr{K}}(DMD^{-1}) = \nu_{\mathscr{K}}(\hat{D}M\hat{D}^{-1}) \\ &= \sqrt{\max\{0, F(\hat{D}, \hat{G})\}}. \end{aligned}$$

V. Correspondence with the Linear Fractional Transformation Approach

In [3], it is shown that in the case of nonrepeated real scalar blocks ($k_q = 1$, $q = 1, \cdots, m_r$), given $\alpha > 0$, a sufficient condition to ensure that $\mu_{\mathscr{K}}(M) \le \alpha$ is that

$$\inf_{D \in \mathscr{D}_{\mathscr{K}}} \inf_{C \in \mathscr{C}_{\mathscr{K}}} \bar{\sigma}\left[jC + (I - C^2)^{1/2} D\left(\frac{M}{\alpha}\right) D^{-1}\right] \le 1 \tag{5.1}$$

where

$$\mathscr{C}_{\mathscr{K}} = \{\operatorname{diag}(c_1, c_2, \cdots, c_{m_r}, 0, \cdots, 0): c_i \in (-1, 1)\}.$$

Using the bijection from $(-1, 1)$ to $\mathbb{R}$

$$c \mapsto g = \frac{c}{\sqrt{1 - c^2}}$$

it is easily checked that condition (5.1) is equivalent to

$$\inf_{D \in \mathscr{D}_{\mathscr{K}}} \inf_{G \in \mathscr{G}_{\mathscr{K}}} \bar{\sigma}\left[\left(D\left(\frac{M}{\alpha}\right) D^{-1} + jG\right)(I + G^2)^{-1/2}\right] \le 1. \tag{5.2}$$

The following proposition, which holds whether or not there are repeated scalar blocks, connects (5.2) with (4.3).

Proposition 5.1: Let $\alpha > 0$. Then (5.2) holds if and only if

$$\inf_{D \in \mathscr{D}_{\mathscr{K}}} \inf_{G \in \mathscr{G}_{\mathscr{K}}} F(D, G) \leq \alpha^2. \tag{5.3}$$

Moreover, the infimum is achieved in (5.2) if and only if it is achieved in (5.3). □

Thus, (5.1) implies $\mu_{\mathscr{K}}(M) \leq \alpha$ in the general case, provided $\mathscr{C}_{\mathscr{K}}$ consists now of Hermitian block-diagonal matrices in the place of the scalars, and the infimum of all positive α's satisfying (5.1) is identical to the second upper bound in (4.3). The advantage of (4.3) is that it has much better computational properties. The characterizations in (5.1) and (5.2) may still be useful in the context of μ-synthesis, which uses the upper bounds and H_∞ optimal control to synthesize controllers, This is under investigation.

VI. Examples

The main result of this paper is that obtained in Section IV of a computable upper bound to the structured singular value with respect to a structure involving both complex blocks and real scalars. It was pointed out that the new bound is, in general, sharper than that of formula (1.5) corresponding to the assumption that all uncertainties are allowed to take on complex values. The purpose of the numerical experiments reported here is to illustrate this last statement.

We begin with an example that can be given some engineering motivation. Although a complete tutorial on the use of μ in analyzing control systems is beyond the scope of this paper, this example illustrates some of the key issues. Also, it is an example where the correct answer for real perturbations is known, so we can compare this to what is obtained using the methods of this paper. We begin with the transfer function model

$$g(s) = \sum_{i=1}^{n} \frac{\alpha_i \omega_i s}{s^2 + 2\zeta_i \omega_i s + \omega_i^2 (1 + r_i \delta_i)}$$

with α_i, ω_i, r_i, and ζ_i all positive constants, and the δ_i representing real perturbations which have been normalized such that $-1 < \delta_i < 1$.

Several different physical problems could motivate a model of this type. A mechanical system consisting of an interconnection of n masses and springs would be the simplest example. Uncertainty in the value of the spring constants would lead naturally to perturbations entering in this way. A very similar problem would be the first n modes of a flexible structure with uncertainty in the stiffness of the materials. In either case, the numerator dynamics are consistent with the assumption that the control input is a force and the output is a velocity measurement at the same location as the force input. In the flexible structure literature this is referred to as a collocated sensor and actuator. If this were a model of a flexible structure to be used in control design, we might want to consider uncertainty in the damping as well, and would probably add additional perturbations to cover unmodeled modes.

Another way that an uncertainty description like the one above could naturally arise is when the δ_i do not represent any particular known physical mechanism, but are used to capture the regularities that might be found in input–output data. In any case, it is important to recognize that, strictly speaking, parameters and perturbations are mathematical objects that occur in our models, not in the physical systems being modeled. We use explicit representations of uncertainty because we want models which are useful for control design, but coming up with such models for actual physical systems can often be quite challenging. For this example, we will take $n = 3$, $\omega_1 = 0.5$, $\omega_2 = 1$, $\omega_3 = 2$, and $\zeta_i = 0.01$, $\alpha_i = 0.2$, $r_i = 1$ for all i. This would correspond to a fairly lightly damped system.

Suppose we use a unity feedback system with a disturbance occurring at the same location as the force input. Then denoting the output by y, the input by u, and the disturbance by d, we get

$$y(s) = g(s)(u(s) + d(s))$$
$$u(s) = -y(s).$$

Suppose we are interested in internal stability of this feedback system as the δ_i vary. It is easily verified by examining the Nyquist plot for g that:

i) The system is stable if, and only if $\delta_i > -1$ for all i; if for any i, $\delta_i = -1$ then the system has an open-loop pole-zero cancellation at $s = 0$ and cannot be stabilized.

ii) The magnitude (in the H_∞ sense) of the closed-loop transfer function from d to y is less than 1 for all $s = j\omega$ if $\delta_i > -1$ for all i.

To apply the methods in this paper it is necessary to obtain the interconnection structure of Fig. 1 with

$$H(s) = \begin{bmatrix} H_{11}(s) & H_{12}(s) \\ H_{21}(s) & H_{22}(s) \end{bmatrix} = C(sI - A)^{-1}B$$

where $H_{11}(s)$, $H_{12}(s)$, $H_{21}(s)$, and $H_{22}(s)$ are 3×3, 3×1, 1×3, and 1×1, respectively, and where

$$A = \begin{bmatrix} 0.00 & 1.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ -0.25 & -0.11 & 0.00 & -0.20 & 0.00 & -0.40 \\ 0.00 & 0.00 & 0.00 & 1.00 & 0.00 & 0.00 \\ 0.00 & -0.10 & -1.00 & -0.22 & 0.00 & -0.40 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 1.00 \\ 0.00 & -0.10 & 0.00 & -0.20 & -4.00 & -0.44 \end{bmatrix}$$

$$B = \begin{bmatrix} 0.00 & 0.00 & 0.00 & 0.00 \\ -0.25 & 0.00 & 0.00 & 1.00 \\ 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & -1.00 & 0.00 & 1.00 \\ 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & -4.00 & 1.00 \end{bmatrix}$$

$$C = \begin{bmatrix} 1.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 1.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 1.00 & 0.00 \\ 0.00 & 0.10 & 0.00 & 0.20 & 0.00 & 0.40 \end{bmatrix}.$$

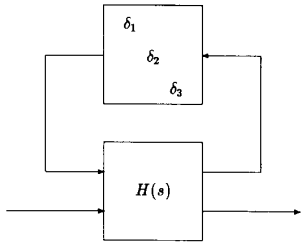


Fig. 1.

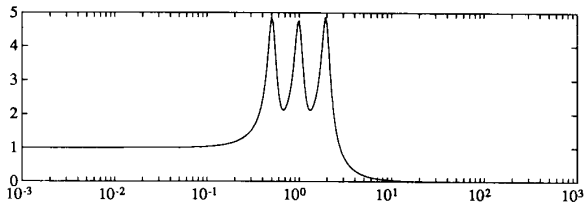


Fig. 2. $\inf_{D\in\mathscr{D}_{\mathscr{K}}} \nu_{\mathscr{K}}(DH_{11}(j\omega)D^{-1})$ with $\mathscr{K} = (;;1,1,1)$.

Computation of A, B, and C is tedious but straightforward, and easily done by computer. For a discussion of how to obtain these interconnection structures for general problems, see, for example [36].

Concerning the robust stability question, the Small μ Theorem [2] asserts that, for given $\delta > 0$, the system in Fig. 1 is stable for all δ_i satisfying $-\delta < \delta_i < \delta$ if and only if

$$\sup_{\omega} \mu_{\mathscr{K}}(H_{11}(j\omega)) \le 1/\delta$$

with the block structure $\mathscr{K} = (1,1,1;;)$. In view of i) in the foregoing discussion, it follows that

$$\sup_{\omega} \mu_{\mathscr{K}}(H_{11}(j\omega)) = 1. \tag{6.1}$$

Computation of $\inf_{D\in\mathscr{D}_{\mathscr{K}}} \nu_{\mathscr{K}}(DH_{11}(j\omega)D^{-1})$ using the algorithm of [35] reveals that

$$\inf_{D\in\mathscr{D}_{\mathscr{K}}} \nu_{\mathscr{K}}(DH_{11}(j\omega)D^{-1}) = \begin{cases} 1, & \text{if } \omega = 0; \\ 0, & \text{if } \omega \ne 0. \end{cases}$$

Theorem 4.1 together with (6.1) thus implies that

$$\mu_{\mathscr{K}}(H_{11}(j\omega)) = \begin{cases} 1, & \text{if } \omega = 0; \\ 0, & \text{if } \omega \ne 0. \end{cases}$$

(Note that the upper bound is thus exact at all frequencies.) This example illustrates that if $m_r \ne 0$, then $\mu_{\mathscr{K}}(M)$ is not necessarily a continuous function of M. As seen in Fig. 2,[4] it also illustrates that the "complex" upper bound,[5] corresponding to the structure $(;;1,1,1)$, can be quite poor an estimate of the "real" structured singular value.

The Small μ Theorem also gives a precise characterization of robust performance. Namely, the system in Fig. 1 is stable for all δ_i satisfying $-\delta < \delta_i < \delta$ and the worst-case performance (i.e., the worst-case H_∞ gain from d to y) is strictly less than δ if and only if

$$\sup_{\omega} \mu_{\hat{\mathscr{K}}}(H(j\omega)) \le 1/\delta$$

[4]Figs. 2–17 were plotted using MATLAB.

[5]Equal to the SSV in this three block case.

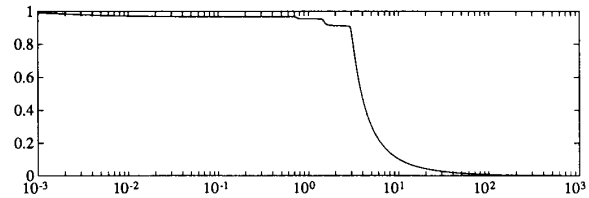


Fig. 3. $\inf_{D\in\mathscr{D}_{\mathscr{K}}} \nu_{\mathscr{K}}(DH(j\omega)D^{-1})$ with $\mathscr{K} = (1,1,1;;1)$.

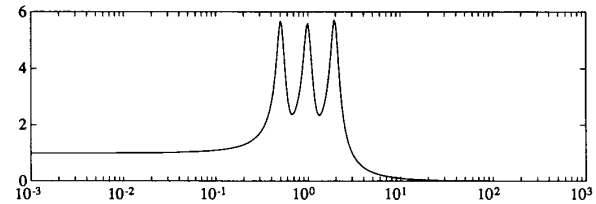


Fig. 4. $\inf_{D\in\mathscr{D}_{\mathscr{K}}} \nu_{\mathscr{K}}(DH(j\omega)D^{-1})$ with $\mathscr{K} = (;;1,1,1,1,1)$.

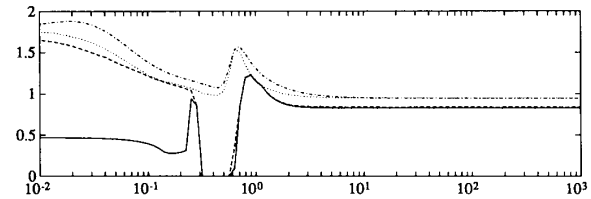


Fig. 5.

with the augmented block structure $\hat{\mathscr{K}} = (1,1,1;;1)$. In view of i) and ii) in the discussion above, it follows that

$$\sup_{\omega} \mu_{\hat{\mathscr{K}}}(H(j\omega)) = 1.$$

The results of computation of $\inf_{D\in\mathscr{D}_{\hat{\mathscr{K}}}} \nu_{\hat{\mathscr{K}}}(DH(j\omega)D^{-1})$ for $\omega \ge 0$ using the algorithm of [35] are plotted in Fig. 3. Again, it is seen that the supremum over frequency of this upper bound is identical to that of the structured singular value.[6] For comparison, a plot of the "complex" upper bound, corresponding to the structure $(;;1,1,1,1)$, is shown in Fig. 4.

The final examples involve square transfer matrices $H(s) = C(sI - A)^{-1}B + E$ of dimension 2×2 and 5×5 and with A, B, C, and E generated randomly. We computed the upper bound to $\mu_{\mathscr{K}}(H(j\omega))$ obtained in Section IV on a grid of values of ω (logarithmically spaced, with 20 points per decade) for the following structures $\mathscr{K}$, all consisting of scalar blocks only: i) all scalars are allowed to take on complex values, ii) some of the scalars are restricted to be real, and iii) all the scalars are restricted to be real.In the 2×2 case (two scalar blocks), the structured singular value in case iii) was also computed exactly (it can be evaluated by finding the roots of a system of two bilinear equations in two variables). A typical sample of the results we obtained is displayed in Figs. 5–17. The four curves in Figs. 5–14 correspond (top to bottom) to structures $(;;1,1)$, $(1;;1)$, and $(1,1;;)$ (for the bottom two curves), respectively.[7] In Figs. 15–17, the three curves correspond to structures $(;;1,1,1,1,1)$, $(1,1;;1,1,1)$, and $(1,1,1,1,1;;)$, respectively. For validation purposes, the matrices corresponding to Figs. 5 and 15 are given in Appendix B.

[6]Computation of $\mu_{\hat{\mathscr{K}}}(H(j\omega))$ for all ω was not attempted as it requires, for each frequency, computation of the global maximum of a constrained optimization problem of the form (2.7).

[7]In Fig. 7, the bottom two curves coincide.

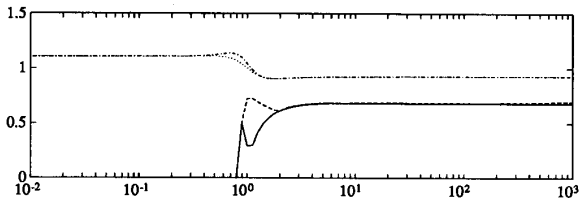


Fig. 6.

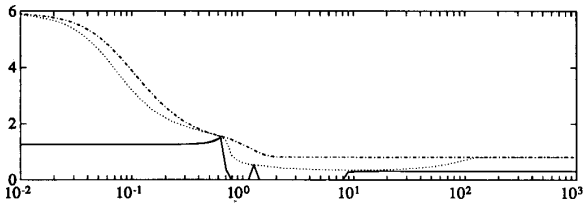


Fig. 7.

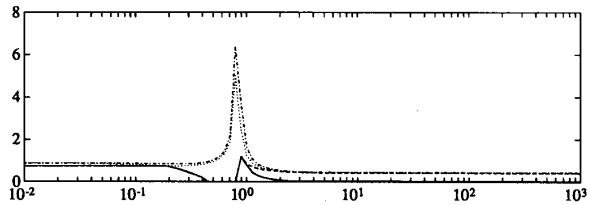


Fig. 8.

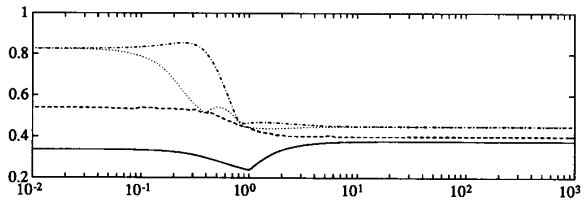


Fig. 9.

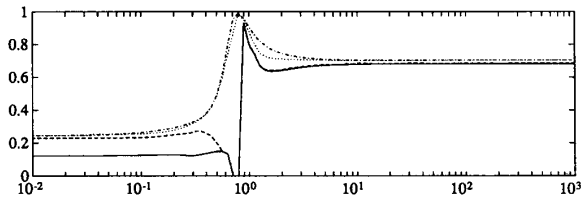


Fig. 10.

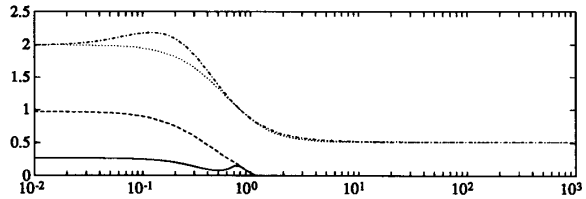


Fig. 11.

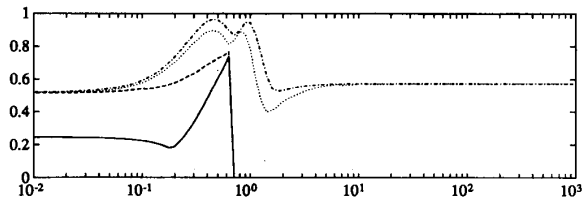


Fig. 12.

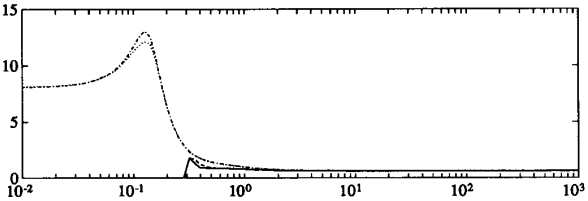


Fig. 13.

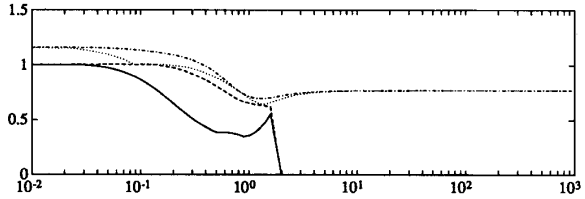


Fig. 14.

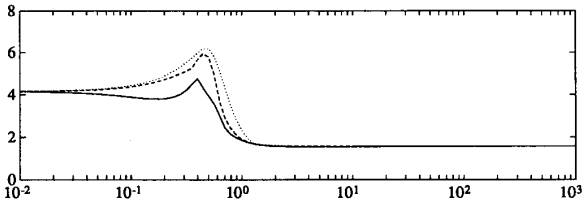


Fig. 15.

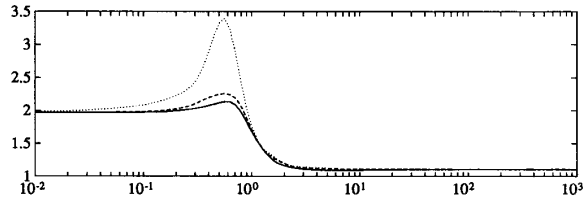


Fig. 16.

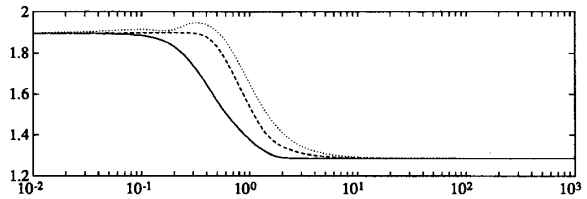


Fig. 17.

Clearly, in most cases, the upper bound derived in Section IV is significantly less conservative than that obtained by assuming possibly complex uncertainty. Note that in many cases (Figs. 5, 7, 8, 10, and 13) the maximum over frequency of the new upper bound is essentially identical to the maximum over frequency of the exact structured singular value.

VII. Future Research

The main result of this paper is the computable upper bound in (4.5) for μ with structures having both real and complex blocks. While this yields a tremendous improvement over the upper bound in (1.5), it still does not give an exact method for

computing μ. A lower bound for μ can be found from (2.7) by local search,[8] but since (2.7) may have local maxima which are not global this may not yield μ. It would be reasonably inexpensive to compute these two bounds and, obviously, μ would be between them. However, it is possible that the bounds could be far apart.

What are the prospects for a general, exact, computationally attractive method for computing μ for real perturbations? One promising possibility is suggested by a research direction initiated by de Gaston [6]–[8]. This work begins with an upper and lower bound which may also be far apart. The bounds are refined, however, by partitioning the domain of the perturbations and computing the bounds for subdomains. While the growth of this tree of subdomains can be exponential, it is conjectured that the number of subdomains can be kept manageable by standard tree-pruning involving comparing the bounds of different domains. Unfortunately, the bounds themselves have exponential explosion. A promising research direction is to combine a subdomain partitioning scheme with the bounds in this paper. This would eliminate the exponential explosion in the computation of the bounds. Then the critical issue would be the growth in the number of subdomains on which the bounds would be computed. If a scheme could be found to keep this growth manageable, it would lead to the desired efficient, general, exact method of computing μ. This idea is currently under investigation.

Appendix A

The following lemma is used in the proof of Theorem 2.1 below.

Lemma A: Let $\theta > 0$, $x \in \mathbb{C}^n$, and let $q \in \{1, \cdots, m_r + m_c\}$. Suppose that

$$\| Q_q M x \| \geq \theta \| Q_q x \|. \tag{A.1}$$

Then i)

$$x_i (\overline{Mx})_j = \bar{x}_j (Mx)_i, \qquad i, j \in J_q \tag{A.2}$$

if and only if for some $\delta \in [-1, 1]$

$$\delta (Mx)_i = \theta x_i, \qquad i \in J_q \tag{A.3}$$

and ii)

$$x_i (Mx)_j = x_j (Mx)_i, \qquad i, j \in J_q \tag{A.4}$$

if and only if (A.3) holds for some $\delta \in \mathbb{C}$, $|\delta| \leq 1$.

Proof: We prove the first equivalence (the second one follows similarly). First, suppose (A.2) holds. If $(Mx)_i = 0$ for all $i \in J_q$ then, in view of (A.1), $x_i = 0$ for all $i \in J_q$, so that (A.3) holds with any δ. Suppose now that $(Mx)_{i_0} \neq 0$ for some $i_0 \in J_q$ and let

$$\delta = \frac{\theta x_{i_0}}{(Mx)_{i_0}}. \tag{A.5}$$

Let $i \in J_q$. If $(Mx)_i = 0$, then it follows from (A.2) with $(i, j) = (i, i_0)$ that $x_i = 0$, so that (A.3) holds for i. If $(Mx)_i \neq 0$, then using (A.2) with (i, j) successively equal to (i, i_0) and (i_0, i_0), then (A.5) one gets

$$\frac{\theta x_i}{(Mx)_i} = \frac{\theta \bar{x}_{i_0}}{(\overline{Mx})_{i_0}} = \frac{\theta x_{i_0}}{(Mx)_{i_0}} = \delta$$

so that, again, (A.3) holds for i. To prove the converse, suppose now that (A.3) holds. Let $i \in J_q$. If $(Mx)_i = 0$, then (A.3) implies that $x_i = 0$, so that (A.2) holds for i and any $j \in J_q$. Finally, for any (i, j) such that $(Mx)_i \neq 0 \neq (Mx)_j$, (A.3) yields, since δ is real

$$\frac{\theta x_i}{(Mx)_i} = \delta = \frac{\theta x_j}{(Mx)_j} = \frac{\theta \bar{x}_j}{(\overline{Mx})_j}$$

so that (A.2) holds. □

Proof of Theorem 2.1: Let $\bar{\mu}_{\mathscr{X}}(M)$ denote the right-hand side in (2.7). We first show that $\mu_{\mathscr{X}}(M) \geq \bar{\mu}_{\mathscr{X}}(M)$. If $\mathscr{S}_{\mathscr{X}}(M) = \varnothing$, it holds trivially. Otherwise, the feasible set for (2.7) is nonempty. Thus, let (θ, x) be feasible for (2.7). We show that $\mu_{\mathscr{X}}(M) \geq \theta$, which establishes the claim. If $\theta = 0$, this holds trivially. Thus, assume $\theta > 0$. In view of Lemma A, there exists $\delta_q^r \in [-1, 1]$, $q = 1, \cdots, m_r$, $\delta_q^c \in \mathbb{C}$, $|\delta_q^c| \leq 1$, $q = 1, \cdots, m_c$, such that

$$\delta_q^r (Mx)_i = \theta x_i \qquad \forall\, i \in J_q$$

and, for $q = 1, \cdots, m_c$

$$\delta_q^c (Mx)_i = \theta x_i \qquad \forall\, i \in J_{m_r+q}$$

i.e.,

$$\delta_q^r Q_q M x = \theta Q_q x, \qquad q = 1, \cdots, m_r$$

$$\delta_q^c Q_{m_r+q} M x = \theta Q_{m_r+q} x, \qquad q = 1, \cdots, m_c.$$

Feasibility of (θ, x) for (2.7) also implies that there exists $\Delta_q^C \in \mathbb{C}^{k_{m_r+m_c+q} \times k_{m_r+m_c+q}}$, $\bar{\sigma}(\Delta_q^C) \leq 1$, such that

$$\Delta_q^C Q_{m_r+m_c+q} M x = \theta Q_{m_r+m_c+q} x, \qquad q = 1, \cdots, m_C.$$

Thus, the matrix

$$\Delta = \frac{1}{\theta} \text{ block diag} \left(\delta_1^r I_{k_1}, \cdots, \delta_{m_r}^r I_{k_{m_r}}, \delta_1^c I_{k_{m_r+1}}, \cdots, \delta_{m_c}^c I_{k_{m_r+m_c}}, \Delta_1^C, \cdots, \Delta_{m_C}^C \right)$$

is such that $\Delta \in \mathscr{X}_{\mathscr{X}}$, $\bar{\sigma}(\Delta) \leq \theta^{-1}$, and $\Delta M x = x$. The latter implies that $\det(I - \Delta M) = 0$, so that $\mu_{\mathscr{X}}(M) \geq \theta$. Conversely, let us now show that $\mu_{\mathscr{X}}(M) \leq \bar{\mu}_{\mathscr{X}}(M)$. If $\mu_{\mathscr{X}}(M) = 0$, it holds trivially. Thus, suppose $\mu_{\mathscr{X}}(M) > 0$. Let $\delta = \mu_{\mathscr{X}}^{-1}(M)$. By definition of $\mu_{\mathscr{X}}(M)$, there exists $\delta_q^r \in \mathbb{R}$, $q = 1, \cdots, m_r$, $\delta_q^c \in \mathbb{C}$, $q = 1, \cdots, m_c$, $\Delta_q^C \in \mathbb{C}^{k_{m_r+m_c+q} \times k_{m_r+m_c+q}}$, $q = 1, \cdots, m_C$, with

$$|\delta_q^r| \leq 1, \qquad q = 1, \cdots, m_r \tag{A.6a}$$

$$|\delta_q^c| \leq 1, \qquad q = 1, \cdots, m_c \tag{A.6b}$$

$$\bar{\sigma}(\Delta_q^C) \leq 1 \qquad q = 1, \cdots, m_C \tag{A.6c}$$

such that

$$\Delta = \delta \text{ block diag} \left(\delta_1^r I_{k_1}, \cdots, \delta_{m_r}^r I_{k_{m_r}}, \delta_1^c I_{k_{m_r+1}}, \cdots, \delta_{m_c}^c I_{k_{m_r+m_c}}, \Delta_1^C, \cdots, \Delta_{m_C}^C \right) \in \mathscr{X}_{\mathscr{X}}(\delta)$$

satisfies $\det(I - \Delta M) = 0$, i.e., for some $x \in \partial B$

$$\Delta M x = x. \tag{A.7}$$

[8] Provided a feasible point is available.

We show that (δ^{-1}, x) is feasible for (2.7), thus completing the proof. From (A.7) it follows that

$$\delta_q^r Q_q M x = \frac{1}{\delta} Q_q x, \qquad q = 1, \cdots, m_r \tag{A.8a}$$

$$\delta_q^c Q_{m_r+q} M x = \frac{1}{\delta} Q_{m_r+q} x, \qquad q = 1, \cdots, m_c \tag{A.8b}$$

$$\Delta_q^C Q_{m_r+m_c+q} M x = \frac{1}{\delta} Q_{m_r+m_c+q} x, \qquad q = 1, \cdots, m_C. \tag{A.9}$$

In view of Lemma A, it follows from (A.8) that, for $q = 1, \cdots, m_r$

$$x_i(\overline{Mx})_j = \bar{x}_j(Mx)_i \qquad \forall\, i, j \in J_q$$

and, for $q = 1, \cdots, m_c$

$$x_i(Mx)_j = x_j(Mx)_i \qquad \forall\, i, j \in J_{m_r+q}$$

and thus, $x \in \mathscr{S}_{\mathscr{X}}(M)$. Finally (A.6), (A.8), and (A.9) imply that, for $q = 1, \cdots, m$

$$\| Q_q M x \| \geq \frac{1}{\delta} \| Q_q x \|.$$

□

Proof of Proposition 3.1: Defining φ: $\partial B \times \mathbb{R} \to \mathbb{R}^s$ by

$$\varphi_i(x, \alpha) = x^H A_i(\alpha) x \qquad i = 1, \cdots, s$$

one has, for any real α

$$c(\alpha) = \min \{ N(\varphi(x, \alpha) + v) \colon x \in \partial B,\ v \in \mathscr{P}_m \}.$$

Since, when α varies locally around any given $\hat{\alpha}$ and x varies over ∂B, $\varphi(x, \alpha)$ is bounded, it is clear that, for α around $\hat{\alpha}$, v can be restricted to lie in some compact subset $V(\hat{\alpha})$ of $\mathscr{P}_m$. Continuity of $c(\cdot)$ at $\hat{\alpha}$ then follows from continuity of φ and compactness of $\partial B \times V(\hat{\alpha})$. Now, let $\beta \geq 0$ and $\alpha \in \mathbb{R}$. From the definition of $A_i(\alpha)$, $i = 1, \cdots, s$, it follows that, for any $x \in \partial B$

$$\varphi(x, \alpha + \beta) = \varphi(x, \alpha) + \beta Q(x)$$

where $Q(x) \in \mathbb{R}^s$ is given by

$$Q(x) = \begin{bmatrix} \| Q_1 x \|^2 \\ \vdots \\ \| Q_m x \|^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Thus, we have

$$\begin{aligned} c(\alpha + \beta) &= \min_{\substack{x \in \partial B \\ v \in \mathscr{P}_m}} N(\varphi(x, \alpha + \beta) + v) \\ &= \min_{\substack{x \in \partial B \\ v \in \mathscr{P}_m}} N(\varphi(x, \alpha) + v + \beta Q(x)). \end{aligned}$$

Using the triangle inequality we obtain, since $\beta \geq 0$

$$c(\alpha + \beta) \leq \min_{\substack{x \in \partial B \\ v \in \mathscr{P}_m}} \left[N(\varphi(x, \alpha) + v) + \beta N(Q(x)) \right].$$

Since the norm N satisfies $N(e_q) \leq 1$, we have, for $x \in \partial B$

$$N(Q(x)) \leq \sum_{q=1}^{m} N(\| Q_q x \|^2 e_q) \leq \sum_{q=1}^{m} \| Q_q x \|^2 = 1$$

and thus

$$c(\alpha + \beta) \leq \min_{\substack{x \in \partial B \\ v \in \mathscr{P}_m}} \left[N(\varphi(x, \alpha) + v) + \beta \right] = c(\alpha) + \beta.$$

□

Proof of Proposition 4.1: The first inequality holds trivially if $\mathscr{S}_{\mathscr{X}}(M) = \varnothing$. If not, suppose that $(\mu_{\mathscr{X}}(M), x_*)$ solves (2.7). Then, for $q = 1, \cdots, m$, $\| Q_q M x_* \| \geq \mu_{\mathscr{X}}(M) \| Q_q x_* \|$, so that

$$\begin{aligned} \| M x_* \|^2 = \sum_{q=1}^{m} \| Q_q M x_* \|^2 &\geq \mu_{\mathscr{X}}^2(M) \sum_{q=1}^{m} \| Q_q x_* \|^2 \\ &= \mu_{\mathscr{X}}^2(M). \end{aligned}$$

Since $x_* \in \mathscr{S}_{\mathscr{X}}(M)$, this first inequality in (4.1) holds. The second inequality also holds trivially if $\mathscr{S}_{\mathscr{X}}(M) = \varnothing$. If $\mathscr{S}_{\mathscr{X}}(M) \neq \varnothing$, then, in view of Proposition 2.1, for any $x \in \mathscr{S}_{\mathscr{X}}(M)$, $G \in \mathscr{G}_{\mathscr{X}}$

$$\begin{aligned} \| M x \|^2 = x^H M^H M x &= x^H \left[M^H M + j(GM - M^H G) \right] x \\ &\leq \bar{\lambda} \left[M^H M + j(GM - M^H G) \right] \end{aligned}$$

and thus

$$\begin{aligned} \eta_{\mathscr{X}}^2(M) &= \max_{x \in \mathscr{S}_{\mathscr{X}}(M)} \| M x \|^2 \\ &\leq \inf_{G \in \mathscr{G}_{\mathscr{X}}} \bar{\lambda} \left[M^H M + j(GM - M^H G) \right] \\ &\leq \max \left\{ 0, \inf_{G \in \mathscr{G}_{\mathscr{X}}} \bar{\lambda} \left[M^H M + j(GM - M^H G) \right] \right\} \\ &= \nu_{\mathscr{X}}^2(M). \end{aligned}$$

Finally, the last inequality in (4.1) is clear since $\mathscr{G}_{\mathscr{X}}$ contains the zero matrix and since $\bar{\sigma}^2(M) = \bar{\lambda}(M^H M)$. □

Proof of Proposition 4.2: First, (4.4) can be rewritten as

$$\begin{aligned} F(D, G) &= \max_{\alpha \in \mathbb{R}} \{ \alpha \colon \bar{\lambda} [M_D^H M_D \\ &\qquad + j(GM_D - M_D^H G) - \alpha I] \geq 0 \} \\ &= \max_{\alpha \in \mathbb{R}} \{ \alpha \colon \bar{\lambda} [D^{-1} M^H D^2 M D^{-1} \\ &\qquad + j(GDMD^{-1} - D^{-1} M^H D G) - \alpha I] \geq 0 \}. \end{aligned}$$

Next, since every $D \in \mathscr{D}_{\mathscr{X}}$ is nonsingular, it follows that

$$\begin{aligned} F(D, G) &= \max_{\alpha \in \mathbb{R}} \{ \alpha \colon \bar{\lambda} [M^H D^2 M \\ &\qquad + j(DGDM - M^H D G D) - \alpha D^2] \geq 0 \} \\ &= \max_{\alpha \in \mathbb{R}} \{ \alpha \colon \Phi_\alpha(D^2, DGD) \geq 0 \}. \end{aligned}$$

Also, for given $(D, G) \in \mathscr{D}_{\mathscr{X}} \times \mathscr{G}_{\mathscr{X}}$, since $D > 0$, $\Phi_\alpha(D, G)$ is strictly decreasing as a function of α. Thus i) holds. Claim ii) follows from the fact that the map $\mathscr{D}_{\mathscr{X}} \times \mathscr{G}_{\mathscr{X}} \to \mathscr{D}_{\mathscr{X}} \times \mathscr{G}_{\mathscr{X}}$ defined by $(D, G) \mapsto (D^2, DGD)$ is a bijection. □

Proof of Proposition 4.3: From the fact that the derivative of Φ_{α_*} at (D_*, G_*) exists and vanishes, it follows that, for

any $D \in \mathscr{D}_{\mathscr{X}}$, $G \in \mathscr{G}_{\mathscr{X}}$

$$\frac{d}{dt}\bar{\lambda}\left[M^H(D_* + tD)M + j(G_*M - M^HG_*) - \alpha_*(D_* + tD)\right]|_{t=0} = 0$$

and

$$\frac{d}{dt}\bar{\lambda}\left[M^HD_*M + j((G_* + tG)M - M^H(G_* + tG)) - \alpha_*D_*\right]|_{t=0} = 0.$$

Using a classical result on derivatives (or generalized gradient) of eigenvalues of Hermitian matrices (see, e.g., [37]), we may rewrite these two equalities as

$$x^H{}_*\frac{d}{dt}\left[M^H(D_* + tD)M + j(G_*M - M^HG_*) - \alpha_*(D_* + tD)\right]|_{t=0}x_* = 0 \quad (A.10)$$

and

$$x^H{}_*\frac{d}{dt}\left[M^HD_*M + j((G_* + tG)M - M^H(G_* + tG)) - \alpha_*D_*\right]|_{t=0}x_* = 0 \quad (A.11)$$

where x_* is any unit length eigenvector corresponding to the largest eigenvalue of

$$M^HD_*M + j(G_*M - M^HG_*) - \alpha_*D_*.$$

Equations (A.10) and (A.11) yield, respectively,

$$x^H{}_*(M^HDM - \alpha_*D)x_* = 0 \quad (A.12)$$

and

$$x^H{}_*(GM - M^HG)x_* = 0. \quad (A.13)$$

Since $D \in \mathscr{D}_{\mathscr{X}}$ is arbitrary, (A.12) implies that

$$\|Q_qMx_*\|^2 = \alpha_*\|Q_qx_*\|^2, \qquad q = 1, \cdots, m \quad (A.14)$$

and, since $G \in \mathscr{G}_{\mathscr{X}}$ is also arbitrary, together with (A.13) this implies that $(x_*, \sqrt{\alpha_*})$ is feasible for (2.7) (note that (A.14) implies that α_* is nonnegative). Thus, $\sqrt{\alpha_*} \le \mu_{\mathscr{X}}(M)$. On the other hand, in view of Proposition 4.2ii), the definition of α_* implies that

$$\alpha_* \ge \inf_{D \in \mathscr{D}_{\mathscr{X}},\, G \in \mathscr{G}_{\mathscr{X}}} F(D, G)$$

and it follows from (4.2) and (4.3) that $\sqrt{\alpha_*} \ge \mu_{\mathscr{X}}(M)$. Thus $\sqrt{\alpha_*} = \mu_{\mathscr{X}}(M)$. The claim then follows from Proposition 4.2i). □

Proof of Theorem 4.2: Let $(\hat{D}, \hat{G})$ be a minimizer for $F(D, G)$ and let $\hat{\alpha} = F(\hat{D}, \hat{G})$. Since the largest eigenvalue of $(M_{\hat{D}}^HM_{\hat{D}} + j(\hat{G}M_{\hat{D}} - M_{\hat{D}}^H\hat{G}))$ is simple, we have

$$\operatorname{rank}\left(M_{\hat{D}}^HM_{\hat{D}} + j(\hat{G}M_{\hat{D}} - M_{\hat{D}}^H\hat{G}) - \hat{\alpha}I\right) = n - 1.$$

Given $\hat{D} > 0$, this implies that

$$\operatorname{rank}\left(M^H\hat{D}^2M + j(\hat{D}\hat{G}\hat{D}M - M^H\hat{D}\hat{G}\hat{D}) - \hat{\alpha}\hat{D}^2\right) = n - 1. \quad (A.15)$$

On the other hand, in view of Proposition 4.2i)

$$\Phi_{\hat{\alpha}}(\hat{D}^2, \hat{D}\hat{G}\hat{D}) = 0. \quad (A.16)$$

It follows from (4.5), (A.15), and (A.16) that the largest eigenvalue of the matrix in (A.15) is simple and thus $\Phi_{\hat{\alpha}}$ is differentiable at $(\hat{D}^2, \hat{D}\hat{G}\hat{D})$. Since $(\hat{D}, \hat{G})$ minimizes $F(D, G)$ it follows from Proposition 4.2 i) that the derivative of $\Phi_{\hat{\alpha}}$ vanishes at $(\hat{D}^2, \hat{D}\hat{G}\hat{D})$. The claim then follows from Proposition 4.3. □

Proof of Proposition 5.1: We show that (5.2) holds and the infimum in (5.2) is achieved if and only if (5.3) holds and the infimum in (5.3) is achieved. Extension to the case when the infima are not achieved is left as a simple exercise. Specifically, we show that, given $D \in \mathscr{D}_{\mathscr{X}}$, $G \in \mathscr{G}_{\mathscr{X}}$

$$\bar{\sigma}\left[\left(D\left(\frac{M}{\alpha}\right)D^{-1} + jG\right)(I + G^2)^{-1/2}\right] \le 1$$

if and only if

$$F(D, \alpha G) \le \alpha^2$$

(note that $G \in \mathscr{G}_{\mathscr{X}}$ if and only if $\alpha G \in \mathscr{G}_{\mathscr{X}}$). This equivalence follows from the following sequence of equivalent inequalities, where the notation $M_D = DMD^{-1}$ is used:

$$\bar{\sigma}\left[\left(D\left(\frac{M}{\alpha}\right)D^{-1} + jG\right)(I + G^2)^{-1/2}\right] \le 1$$

$$\left[\left(D\left(\frac{M}{\alpha}\right)D^{-1} + jG\right)(I + G^2)^{-1/2}\right]^H \cdot \left[\left(D\left(\frac{M}{\alpha}\right)D^{-1} + jG\right)(I + G^2)^{-1/2}\right] \le I$$

$$\left[\left(D\left(\frac{M}{\alpha}\right)D^{-1} + jG\right)\right]^H\left[\left(D\left(\frac{M}{\alpha}\right)D^{-1} + jG\right)\right] \le I + G^2$$

$$\left[\left(D\left(\frac{M}{\alpha}\right)D^{-1} + jG\right)\right]^H\left[\left(D\left(\frac{M}{\alpha}\right)D^{-1} + jG\right)\right] - G^2 \le I$$

$$\frac{1}{\alpha^2}(M_D^HM_D + j((\alpha G)M_D - M_D^H(\alpha G))) \le I$$

$$\bar{\lambda}(M_D^HM_D + j((\alpha G)M_D - M_D^H(\alpha G))) \le \alpha^2.$$

□

Appendix B

The A, B, C, E matrices that were used to generated Figs. 5 and 15 are as follows. For Fig. 5

$$A = \begin{bmatrix} -8.0902 \times 10^{-2} & 1.2686 \times 10^{-1} & -3.5762 \times 10^{-1} & -4.8575 \times 10^{-2} & -3.7808 \times 10^{-1} \\ 1.8962 \times 10^{-1} & 4.1289 \times 10^{-1} & 6.1862 \times 10^{-1} & 1.0322 \times 10^{-1} & 7.7984 \times 10^{-1} \\ -4.8243 \times 10^{-1} & 4.1642 \times 10^{-1} & 2.4450 \times 10^{-1} & -5.6185 \times 10^{-1} & -4.8248 \times 10^{-1} \\ -1.8448 \times 10^{-3} & 2.4564 \times 10^{-1} & -6.5441 \times 10^{-2} & 6.7149 \times 10^{-1} & -5.2891 \times 10^{-1} \\ 7.1872 \times 10^{-2} & -3.3682 \times 10^{-1} & 1.2945 \times 10^{-1} & 5.4038 \times 10^{-1} & -4.4566 \times 10^{-2} \end{bmatrix}$$

$$B = \begin{bmatrix} 2.2105 \times 10^{-2} & -1.4448 \times 10^{-1} \\ -4.3877 \times 10^{-2} & 4.2382 \times 10^{-1} \\ -8.3498 \times 10^{-1} & 3.9127 \times 10^{-1} \\ -1.4370 \times 10^{-1} & 2.4036 \times 10^{-1} \\ 1.8585 \times 10^{-1} & -1.7599 \times 10^{-1} \end{bmatrix}$$

$$C = \begin{bmatrix} -5.0249 \times 10^{-2} & 2.3438 \times 10^{-3} & 6.4024 \times 10^{-1} & -1.8836 \times 10^{-1} & -1.3455 \times 10^{-1} \\ -3.6293 \times 10^{-1} & -2.1268 \times 10^{-1} & -2.3698 \times 10^{-1} & -6.4769 \times 10^{-1} & -1.6309 \times 10^{-1} \end{bmatrix}$$

$$E = \begin{bmatrix} 2.3845 \times 10^{-1} & -8.0438 \times 10^{-1} \\ 3.8387 \times 10^{-1} & 6.8429 \times 10^{-1} \end{bmatrix}$$

and for Fig. 15

$$A = \begin{bmatrix} -4.5509 \times 10^{-1} & 5.3934 \times 10^{-1} & -9.0161 \times 10^{-3} & 5.6126 \times 10^{-1} & 2.4023 \times 10^{-1} \\ -4.9641 \times 10^{-2} & -4.0575 \times 10^{-1} & -3.3601 \times 10^{-1} & -5.8569 \times 10^{-1} & -6.1047 \times 10^{-1} \\ -2.8742 \times 10^{-2} & -2.6343 \times 10^{-2} & 2.1769 \times 10^{-2} & 1.7907 \times 10^{-1} & 7.2380 \times 10^{-1} \\ -1.5319 \times 10^{-1} & 2.3885 \times 10^{-1} & -2.7023 \times 10^{-1} & 1.9057 \times 10^{-1} & 2.9318 \times 10^{-1} \\ 5.2573 \times 10^{-1} & 8.1813 \times 10^{-1} & 6.1384 \times 10^{-1} & 2.4381 \times 10^{-2} & 4.9339 \times 10^{-1} \end{bmatrix}$$

$$B = \begin{bmatrix} -2.6256 \times 10^{-1} & 2.5961 \times 10^{-1} & 8.8551 \times 10^{-1} & -6.4732 \times 10^{-2} & 6.3857 \times 10^{-2} \\ 5.1035 \times 10^{-1} & 4.8261 \times 10^{-2} & -1.5665 \times 10^{-1} & -4.1377 \times 10^{-1} & 2.4904 \times 10^{-1} \\ -8.7380 \times 10^{-2} & -1.3578 \times 10^{-2} & -1.8021 \times 10^{-1} & 6.1830 \times 10^{-1} & 7.3116 \times 10^{-1} \\ -3.7540 \times 10^{-1} & -5.2973 \times 10^{-1} & 3.7967 \times 10^{-1} & 3.1948 \times 10^{-1} & -3.4212 \times 10^{-1} \\ 1.5306 \times 10^{-1} & -5.9548 \times 10^{-1} & 5.1851 \times 10^{-1} & 3.0709 \times 10^{-1} & 1.0449 \times 10^{-1} \end{bmatrix}$$

$$C = \begin{bmatrix} -1.3036 \times 10^{-1} & -6.6165 \times 10^{-2} & 4.3862 \times 10^{-1} & 6.6575 \times 10^{-2} & 9.5535 \times 10^{-2} \\ -6.8446 \times 10^{-1} & -3.0868 \times 10^{-1} & 1.3999 \times 10^{-1} & -4.5997 \times 10^{-1} & 1.1329 \times 10^{-1} \\ 4.3439 \times 10^{-1} & 3.1450 \times 10^{-2} & -2.2962 \times 10^{-1} & -5.4526 \times 10^{-1} & 2.3548 \times 10^{-1} \\ 3.9745 \times 10^{-1} & 4.0280 \times 10^{-1} & -3.0315 \times 10^{-1} & 7.0834 \times 10^{-1} & 5.4871 \times 10^{-1} \\ -1.0055 \times 10^{-1} & 4.0313 \times 10^{-1} & -3.5256 \times 10^{-1} & 3.9585 \times 10^{-2} & 4.4990 \times 10^{-2} \end{bmatrix}$$

$$E = \begin{bmatrix} 3.4802 \times 10^{-1} & 7.9979 \times 10^{-1} & 4.5178 \times 10^{-1} & -1.5375 \times 10^{-1} & -2.4380 \times 10^{-1} \\ -2.8164 \times 10^{-1} & -1.2299 \times 10^{-1} & 6.6314 \times 10^{-1} & -2.6890 \times 10^{-1} & -1.1213 \times 10^{-1} \\ 2.7349 \times 10^{-1} & -6.7474 \times 10^{-1} & -4.5236 \times 10^{-1} & 3.7879 \times 10^{-1} & -5.2324 \times 10^{-1} \\ -6.7545 \times 10^{-1} & -3.5034 \times 10^{-1} & 3.1562 \times 10^{-1} & -7.0583 \times 10^{-1} & 3.4941 \times 10^{-1} \\ 2.4558 \times 10^{-1} & -1.9770 \times 10^{-1} & -1.8149 \times 10^{-1} & 6.9694 \times 10^{-1} & -2.2861 \times 10^{-1} \end{bmatrix}.$$

Acknowledgment

The authors wish to thank Dr. E. Zafiriou and Dr. A. K. Packard for numerous discussions, L. Lee and P. Young for pointing out inaccuracies in a preliminary version, and anonymous reviewers for providing valuable comments.

References

[1] J. C. Doyle, "Analysis of feedback systems with structured uncertainties," *Proc. IEE*, pt. D, vol. 129, pp. 242–250, 1982.

[2] J. C. Doyle, J. E. Wall, and G. Stein, "Performance and robustness analysis for structured uncertainty," in *Proc. 21st IEEE Conf. Decision Contr.*, Orlando, FL, Dec. 1982, pp. 629–636.

[3] J. C. Doyle, "Structured uncertainty in control system design," in *Proc. 24th IEEE Conf. Decision Contr.*, Fort Lauderdale, FL, Dec. 1985, pp. 260–265.

[4] O. Yaniv and I. M. Horowitz, "A quantative design method for MIMO linear feedbach systems having uncertain plants," *Int. J. Contr.*, vol. 43, pp. 401–421, 1986.

[5] J. Ackermann, *Sampled-Data Control Systems*. New York: Springer-Verlag, 1985.

[6] R. R. E. de Gaston and M. G. Safonov, "Exact calculation of the multiloop stability margin," *IEEE Trans. Automat. Contr.*, vol. 33, pp. 156–171, 1988.

[7] A. Sideris and R. R. E. de Gaston, "Multivariable stability margin calculation with uncertain correlated parameters," in *Proc. 1987 Amer. Contr. Conf.*, Minneapolis, MN, June 1987.

[8] A. Sideris and R. S. Sánchez Peña, "Fast computation of the multivariable stability margin for real interrelated uncertain pa-

rameters," in *Proc. 1988 Amer. Contr. Conf.*, Atlanta, GA, pp. 1483-1488, June 1988.

[9] V. L. Kharitonov, "Asymptotic stability of an equilibrium position of a family of systems of linear differential equations," *Differentsial'nye Uraveniya*, vol. 14, pp. 2086-2088, 1978; (in English) *Differential Equations*, vol. 14, pp. 1483-1485, 1979.

[10] A. C. Bartlett, C. V. Hollot, and L. Huang, "Root locations of an entire polytope of polynomials: It suffices to check the edges," *Math. Contr. Signals, Syst.*, vol. 1, pp. 61-71, 1988.

[11] M. Fu and B. R. Barmish, "Stability of convex and linear combinations of polynomials and matrices arising in robustness problems," in *Proc. 1987 Conf. Inform. Sci. Syst.*, Baltimore, MD, 1987.

[12] I. R. Petersen, "A new extension to Kharitonov's theorem," in *Proc. 26th IEEE Conf. Decision Contr.*, Los Angeles, CA, Dec. 1987, pp. 2070-2075.

[13] M. K. H. Fan, A. L. Tits, and J. C. Doyle, "Robustness in the presence of joint parametric uncertainty and unmodeled dynamics," in *Proc. 1988 Amer. Contr. Conf.*, Atlanta, GA, June 1988, pp. 1195-1200.

[14] A. K. Packard, "What's new with μ: Structured uncertainty in multivariable control," Ph.D. dissertation, Univ. California, Berkeley, 1988.

[15] M. K. H. Fan, "An algorithm to compute the structured singular value," Syst. Res. Center, Univ. Maryland, College Park, Tech. Res. Rep. SRC TR-86-8, 1986.

[16] M. K. H. Fan and A. L. Tits, "Characterization and efficient computation of the structured singular value," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 734-743, 1986.

[17] Special issue on linear multivariable control systems, *IEEE Trans. Automat. Contr.*, vol. AC-26, Feb. 1981.

[18] G. Zames, "On the input-output stability of nonlinear time-varying feedback systems; Part I," *IEEE Trans. Automat. Contr.*, vol. AC-11, pp. 228-238, 1966.

[19] —, "On the input-output stability of nonlinear time-varying feedback systems, Part II," *IEEE Trans. Automat. Contr.*, vol. AC-11, pp. 465-477, 1966.

[20] C. A. Desoer and M. Vidyasagar, *Feedback Systems: Input-Output Properties*. New York: Academic, 1975.

[21] J. C. Willems, *The Analysis of Feedback Systems*. Cambridge, MA: M.I.T. Press, 1971.

[22] M. G. Safonov, *Stability and Robustness of Multivariable Feedback Systems*. Cambridge, MA: M.I.T. Press, 1980.

[23] J. S. Freudenberg, D. P. Looze, J. B. Cruz, "Robustness analysis using singular value sensitivities," *Int. J. Contr.*, vol. 35, pp. 95-116, 1982.

[24] M. G. Safonov, "Tight bounds on the response of multivariable systems with component uncertainty," presented at the 16th Allerton Conf., 1978.

[25] E. E. Osborne, "On pre-conditioning of matrices," *J. Ass. Comput. Mach.*, vol. 7, pp. 338-345, 1960.

[26] M. G. Safnonov, "Stability margins of diagonally perturbed multivariable feedback systems," in *Proc. IEE*, vol. 129, pt. D, pp. 251-256, 1982.

[27] R. M. Redheffer, "Inequalities for a matrix Riccati equation," *J. Math. Mechanics*, vol. 8, pp. 349-367, 1959.

[28] —, "On a certain linear fractional transformation," *J. Math. Phys.*, vol. 39, pp. 269-286, 1960.

[29] J. Doyle, K. Glover, P. Khargonekar, and B. Francis, "State-space solutions to standard $\mathscr{H}_2$ and $\mathscr{H}_\infty$ control problems," *IEEE Trans. Automat. Contr.*, vol. 34, Aug. 1989.

[30] M. K. H. Fan and A. L. Tits, "m-form numerical range and the computation of the structured singular value," *IEEE Trans. Automat. Contr.*, vol. 33, pp. 284-289, 1988.

[31] N.-K. Tsing, "On the norms used in computing the structured singular value," *IEEE Trans. Automat. Contr.*, vol. 35, pp. 201-203, 1990.

[32] S. Friedland and R. Loewy, "Subspaces of symmetric matrices containing matrices with a multiple first eigenvalue," *Pacific J. Math.*, vol. 62, pp. 389-399, 1976.

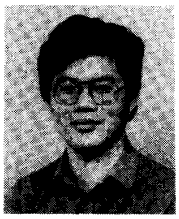
[33] Y.-H. Au-Yeung and Y.-T. Poon, "A remark on the convexity and positive definiteness concerning hermitian matrices," *Southeast Asian Bull. Math.*, vol. 3, pp. 85-92, 1979.

[34] M. K. H. Fan and A. L. Tits, "On the generalized numerical range," *Linear Multilinear Algebra*, vol. 21, pp. 313-320, 1987.

[35] M. K. H. Fan, "User's guide for MUSOL2: A package for computing the structured singular value or its upper bound," Syst. Res. Center, Univ. Maryland, College Park Tech. Res. Rep. *SRC TR* 87-204, 1988.

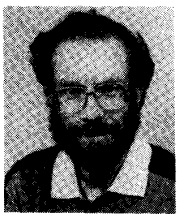
[36] B. G. Morton, "New applications of mu to real-parameter variation problem," in *Proc. 24th Conf. Decision Contr.*, Fort Lauderdale, FL, Dec. 1985, pp. 233-238.

[37] E. Polak and Y. Wardi, "Nondifferentiable optimization algorithm for the design of control systems subject to singular value inequalities," *Automatica, J. IFAC*, vol. 18, pp. 267-283, 1982.



Michael K. H. Fan (M'87) was born in Taiwan, Republic of China, on August 15, 1958. He received the B.S. degree in control engineering from the Chiao-Tong University, Republic of China, and the M.S. and Ph.D. degrees in electrical engineering from the University of Maryland, College Park, in 1980, 1984, and 1986, respectively.

From 1986 to 1989, he was associated with the Systems Research Center, University of Maryland, College Park. Since September 1989, he has been with the School of Electrical Engineering, Georgia Institute of Technology, Atlanta, where he is currently an Assistant Professor. His research interests include robust control, system theory, and optimization-based engineering system design.



André L. Tits (S'71-M'76-S'80-M'80) was born in Verviers, Belgium, on April 13, 1951. He received the Ingénieur Civil degree from the University of Liège, Belgium, and the M.S. and Ph.D. degrees from the University of California, Berkeley, all in electrical engineering, in 1974, 1979, and 1980, respectively.

He has held visiting positions at the University of California, Berkeley, at the Lund Institute of Technology, and at INRIA. Since 1981, he has been with the Department of Electrical Engineering, University of Maryland, College Park, where he is currently a Professor. His research interests include robust control system design as well as numerical methods in optimization and their application to the design of engineering systems.

Dr. Tits was the recipient of a 1985 National Science Foundation Presidential Young Investigator Award.

John C. Doyle received the B.S. and M.S. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge, in 1977, and the Ph.D. degree in mathematics from the University of California, Berkeley, in 1984.

He has been a consultant to Honeywell Systems and Research Center since 1976 and an Associate Professor of Electrical Engineering at the California Institute of Technology, Pasadena, since 1987. His theoretical research interests include modeling and control of uncertain and nonlinear sysems, matrix perturbation problems, and operator theoretic methods. Application interests include flexible structures, chemical process control, flight control, and control of unsteady fluid flow and combustion. Additional interests include gender issues in science and society and impact of control on system design.

Dr. Doyle is the recipient of the Hickernell Award, the Eckman Award, the IEEE Control Systems Society Centennial Outstanding Young Engineer Award, and the Bernard Friedman Award. He is a National Science Foundation Presidential Young Investigator, and ONR Young Investigator, and has coauthored two TRANSACTIONS Best Paper Award winners.