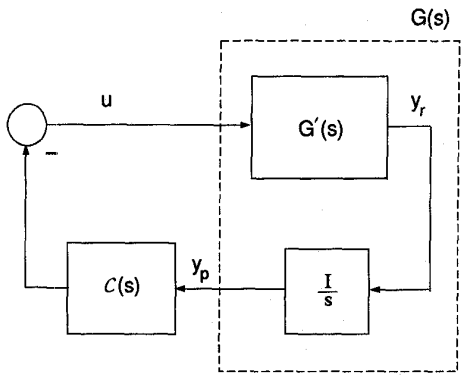

Fig. 2. Flexible structure example.

Example: Consider the rotational motion of a flexible spacecraft with three torque actuators (one for each orthogonal axis of rotation) and 3 collocated attitude sensors. The 3×3 transfer function from the torque input to the attitude (position) output, y_p, is given by

$$G(s) = \frac{G'(s)}{s} \tag{29}$$

where

$$G'(s) = \frac{\alpha_0}{s} + \sum_{i=1}^{p} \frac{\alpha_i s}{s^2 + 2\rho_i \omega_i s + \omega_i^2} \tag{30}$$

where $G'(s)$ is the transfer function from the torque input to the attitude rate $y_r(=\dot{y}_p)$; $\alpha_i = \alpha_i^T \geq 0$ $(i = 0, 1, 2, \cdots, p)$, and α_0 is nonsingular; $\omega_i(>0)$ represents the natural frequency, and $\rho_i \geq 0$ represents the inherent damping ratio for the ith elastic mode $(i = 1, 2, \cdots, p)$. It can be verified that $G'(s)$ is PR and has no zeros at the origin (it may have zeros at other locations on the imaginary axis). Therefore from Theorem 1, it can be stabilized by any MSPR controller that has no poles on the imaginary axis, except possibly at the origin. Let $\mathcal{C}(s)$ denote a 3×3 stable transfer function which has no zeros on the imaginary axis, and suppose $H(s) = [\mathcal{C}(s)/s]$ is MSPR. Then $H(s)$ stabilizes $G'(s)$. Examining the block diagram in Fig. 2, $\mathcal{C}(s)$ stabilizes $G(s)$. In other words, a flexible spacecraft, which has zero-frequency rigid-body modes as well as damped or undamped elastic modes, is stabilized by the controller $\mathcal{C}(s)$ which has the above properties. The stability does not depend on the number of elastic modes or the parameter values, and is therefore robust.

References

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

[2] J. H. Taylor, "Strictly positive real functions and Lefschetz–Kalman–Yakubovich (LKY) lemma," *IEEE Trans. Circuits Syst.*, vol. 21, no. 3, pp. 310–311, Mar. 1974.

[3] P. Ioannou and G. Tao, "Frequency-domain conditions for strictly positive real functions," *IEEE Trans. Automat. Contr.*, vol. 32, pp. 53–54, Jan. 1987.

[4] G. Tao and P. Ioannou, "Strictly positive real matrices and the Lefschetz–Kalman–Yakubovich lemma," *IEEE Trans. Automat. Contr.*, vol. 33, pp. 1183–1185, Dec. 1988.

[5] J. T. Wen, "Time domain and frequency domain conditions for strict positive realness," *IEEE Trans. Automat. Contr.*, vol. 33, pp. 988–992, Oct. 1988.

[6] R. Lozano-Leal and S. M. Joshi, "Strictly positive real functions revisited," *IEEE Trans. Automat. Contr.*, vol. 35, pp. 1243–1245, Nov. 1990.

[7] H. Kaufman, I. Bar-Kana, and K. Sobel, *Direct Adaptive Control Algorithms: Theory and Applications*. New York: Springer Verlag, 1994.

[8] B. D. O. Anderson, "A system theory criterion for positive real matrices," *SIAM J. Contr.*, vol. 5, pp. 171–182, May 1967.

[9] T. Kailath, *Linear Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1980.

[10] R. W. Newcomb, *Linear Multiport Synthesis*. New York: McGraw-Hill, 1966.

Properties of the Mixed μ Problem and Its Bounds

Peter M. Young and John C. Doyle

***Abstract*— Upper and lower bounds for the mixed μ problem have recently been developed, and here we examine the relationship of these bounds to each other and to μ. A number of interesting properties are developed, and the implications of these properties for the robustness analysis of linear systems and the development of practical computation schemes are discussed. In particular we find that current techniques can only guarantee easy computation for large problems when μ equals its upper bound, and computational complexity results prohibit this possibility for general problems. In this context we present some special cases where computation is easy and make some direct comparisons between mixed μ and "Kharitonov-type" analysis methods.**

I. Introduction

The mixed μ problem inherits many of the (appropriately generalized) properties of complex μ, but in some aspects the mixed μ problem can be fundamentally different from the complex μ problem. In this paper we study the relationship between μ and its bounds for the mixed case and compare these findings with the corresponding results for the complex case. A number of interesting properties of μ and its bounds are developed. As well as providing useful insight into the nature of the worst-case solution, these results have important implications for the robustness analysis of linear systems with real parametric uncertainty and the development of practical computation schemes for such problems.

In Section III we examine some of the basic properties of the mixed μ problem, and upper and lower bounds for mixed μ are presented in Section IV [1], [2]. These bounds are used in [3] to form the basis of practical computational software for the mixed μ problem which is currently available as part of the μ-Tools Toolbox [4].

Section V examines the properties of the worst-case solution for general mixed μ problems, and in Section VI we examine some special cases of the mixed μ problem for which computation may be easier. In particular, we briefly consider the rank-one mixed μ problem and its relationship to "Kharitonov-type" analysis results. It is shown that for this problem we have sufficient structure to alleviate the computational complexity of the general problem (which is NP hard) and develop an exact solution. The case of real matrices is also considered in detail. It is shown that for real matrices and nonrepeated parameters, then the mixed and complex μ upper bounds are identical. This result has implications for the use of frequency

Manuscript received January 14, 1994; revised January 10, 1995.

P. M. Young is with the Department of Electrical Engineering, Colorado State University, Fort Collins, CO 80523 USA.

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

Publisher Item Identifier S 0018-9286(96)00387-X.

sweeps versus state-space μ tests and time-invariant versus time-varying perturbations.

II. Notation and Definitions

The notation used here is fairly standard. For any square complex matrix M we denote the transpose by M^T and the complex conjugate transpose by M^*. The largest singular value and the spectral radius are denoted by $\bar{\sigma}(M)$ and $\rho(M)$, respectively. The real spectral radius is defined as $\rho_R(M) = \max\{|\lambda|: \lambda \text{ is a real eigenvalue of } M\}$ with $\rho_R(M) = 0$ if M has no real eigenvalues. For a Hermitian matrix M, then $\bar{\lambda}(M)$ denotes the largest (real) eigenvalue. For any complex vector x, then x^T denotes the transpose, x^* the complex conjugate transpose, and $|x|$ the Euclidean norm. We denote the $k \times k$ identity matrix and zero matrix by I_k and O_k, respectively (occasionally we will drop the subscripts when the size is clear from context).

The definition of the structured singular value, μ, is dependent upon the underlying block structure of the uncertainties which is defined as follows. Given a matrix $M \in \mathbf{C}^{n\times n}$ and three nonnegative integers m_r, m_c, and m_C with $m \doteq m_r + m_c + m_C \leq n$, the block structure $\mathcal{K}(m_r, m_c, m_C)$ is an m-tuple of positive integers

$$\mathcal{K} = (k_1, \cdots, k_{m_r}, k_{m_r+1}, \cdots, k_{m_r+m_c}, k_{m_r+m_c+1}, \cdots, k_m) \quad (1)$$

where we require $\sum_{i=1}^{m} k_i = n$ so these dimensions are compatible with M. This now determines the set of allowable perturbations, namely define

$$X_{\mathcal{K}} = \Big\{\Delta = \text{block diag}\big(\delta_1^r I_{k_1}, \cdots, \delta_{m_r}^r I_{k_{m_r}}, \delta_1^c I_{k_{m_r+1}}, \cdots, \delta_{m_c}^c I_{k_{m_r+m_c}}, \Delta_1^C, \cdots, \Delta_{m_C}^C\big): \delta_i^r \in \mathbf{R}, \delta_i^c \in \mathbf{C}, \Delta_i^C \in \mathbf{C}^{k_{m_r+m_c+i}\times k_{m_r+m_c+i}}\Big\}. \quad (2)$$

Note that $X_{\mathcal{K}} \subset \mathbf{C}^{n\times n}$ and this block structure is sufficiently general to allow for repeated real scalars, repeated complex scalars, and full complex blocks. Note also that the full complex blocks need not be square, but we restrict them as such for notational convenience. The purely complex case corresponds to $m_r = 0$, and the purely real case to $m_c = m_C = 0$.

Definition 1 [5]: The structured singular value, $\mu_{\mathcal{K}}(M)$, of a matrix $M \in \mathbf{C}^{n\times n}$ with respect to a block structure $\mathcal{K}(m_r, m_c, m_C)$ is defined as

$$\mu_{\mathcal{K}}(M) = \left(\min_{\Delta\in X_{\mathcal{K}}} \{\bar{\sigma}(\Delta): \det(I - \Delta M) = 0\}\right)^{-1} \quad (3)$$

with $\mu_{\mathcal{K}}(M) = 0$ if no $\Delta \in X_{\mathcal{K}}$ solves $\det(I - \Delta M) = 0$.

In this paper we will be concerned directly with the properties of μ and its bounds, rather than how to use μ as a robustness analysis tool. For the reader unfamiliar with μ based techniques, a fairly comprehensive review is given in [6].

To develop upper and lower bounds for μ, we need to define some sets of block diagonal scaling matrices (which are also dependent on the underlying block structure)

$$\mathcal{Q}_{\mathcal{K}} = \{\Delta \in X_{\mathcal{K}}: \delta_i^r \in [-1 \quad 1], \delta_i^{c*}\delta_i^c = 1, \Delta_i^{C*}\Delta_i^C = I_{k_{m_r+m_c+i}}\} \quad (4)$$

$$\mathcal{U}_{\mathcal{K}} = \{U \in \mathcal{Q}_{\mathcal{K}}: U^*U = I_n\} \quad (5)$$

$$\tilde{\mathcal{D}}_{\mathcal{K}} = \{\text{block diag}(D_1, \cdots, D_{m_r+m_c}, d_1 I_{k_{m_r+m_c+1}}, \cdots, d_{m_C} I_{k_m}): D_i = D_i^* \in \mathbf{C}^{k_i\times k_i}, d_i \in \mathbf{R}\} \quad (6)$$

$$\mathcal{D}_{\mathcal{K}} = \{D \in \tilde{\mathcal{D}}_{\mathcal{K}}: D > 0\} \quad (7)$$

$$\mathcal{G}_{\mathcal{K}} = \{\text{block diag}(G_1, \cdots, G_{m_r}, O_{k_{m_r+1}}, \cdots, O_{k_m}): G_i = G_i^* \in \mathbf{C}^{k_i\times k_i}\}. \quad (8)$$

We introduce one further piece of notation. For any two vectors $x, y \in \mathbf{C}^n$, then partition x and y compatibly with the block structure as

$$x = \left[x_{r_1}^T \cdots x_{r_{m_r}}^T x_{c_1}^T \cdots x_{c_{m_c}}^T x_{C_1}^T \cdots x_{C_{m_C}}^T\right]^T$$
$$y = \left[y_{r_1}^T \cdots y_{r_{m_r}}^T y_{c_1}^T \cdots y_{c_{m_c}}^T y_{C_1}^T \cdots y_{C_{m_C}}^T\right]^T \quad (9)$$

where $x_{r_i}, y_{r_i} \in \mathbf{C}^{k_i}$, $x_{c_i}, y_{c_i} \in \mathbf{C}^{k_{m_r+i}}$, $x_{C_i}, y_{C_i} \in \mathbf{C}^{k_{m_r+m_c+i}}$. These will be referred to as the "block components" of x and y.

III. Basic Properties of Mixed μ

The mixed μ problem inherits many of the basic properties of the complex μ problem (see [5]). In some aspects, however, the mixed μ problem can be fundamentally different from the complex μ problem. In this section we will present some basic properties of the mixed μ problem and contrast them with the corresponding results for the complex μ problem.

From the definition of mixed μ in (3), one may readily derive the following properties (the complex μ versions of these were originally presented in [5]):

a) $\mu_{\mathcal{K}}(\gamma M) = |\gamma|\mu_{\mathcal{K}}(M)$ for all $M \in \mathbf{C}^{n\times n}$ and $\gamma \in \mathbf{R}$.
b) $\mu_{\mathcal{K}}(I_n) = 1$ for any block structure.
c) $\mu_{\mathcal{K}}(\Delta) = \bar{\sigma}(\Delta)$ for all $\Delta \in X_{\mathcal{K}}$.
d) $m_r = 0, m_c = 0, m_C = 1 \implies \mu_{\mathcal{K}}(M) = \bar{\sigma}(M)$.
e) $m_r = 0, m_c = 1, m_C = 0 \implies \mu_{\mathcal{K}}(M) = \rho(M)$.
f) $m_r = 1, m_c = 0, m_C = 0 \implies \mu_{\mathcal{K}}(M) = \rho_R(M)$.
g) For any $M \in \mathbf{C}^{n\times n}$ the following sequence of inequalities holds

$$\rho_R(M) \leq \mu_{\mathcal{K}}(M) \leq \bar{\sigma}(M).$$

h) For all $\Delta \in X_{\mathcal{K}}, Q \in \mathcal{Q}_{\mathcal{K}}$ then $Q\Delta \in X_{\mathcal{K}}, \Delta Q \in X_{\mathcal{K}}$ with $\bar{\sigma}(Q\Delta) \leq \bar{\sigma}(\Delta), \bar{\sigma}(\Delta Q) \leq \bar{\sigma}(\Delta)$.
i) For all $\Delta \in X_{\mathcal{K}}, U \in \mathcal{U}_{\mathcal{K}}$ then $U\Delta \in X_{\mathcal{K}}, \Delta U \in X_{\mathcal{K}}$ with $\bar{\sigma}(U\Delta) = \bar{\sigma}(\Delta), \bar{\sigma}(\Delta U) = \bar{\sigma}(\Delta)$.
j) $D\Delta D^{-1} = \Delta$ for all $\Delta \in X_{\mathcal{K}}$ and $D \in \mathcal{D}_{\mathcal{K}}$.
k) $\mu_{\mathcal{K}}(QM) = \mu_{\mathcal{K}}(MQ) \leq \mu_{\mathcal{K}}(M)$ for all $M \in \mathbf{C}^{n\times n}$ and $Q \in \mathcal{Q}_{\mathcal{K}}$.
l) $\mu_{\mathcal{K}}(UM) = \mu_{\mathcal{K}}(MU) = \mu_{\mathcal{K}}(M)$ for all $M \in \mathbf{C}^{n\times n}$ and $U \in \mathcal{U}_{\mathcal{K}}$.
m) $\mu_{\mathcal{K}}(DMD^{-1}) = \mu_{\mathcal{K}}(M)$ for all $M \in \mathbf{C}^{n\times n}$ and $D \in \mathcal{D}_{\mathcal{K}}$.
n) For any $M \in \mathbf{C}^{n\times n}$ the following sequence of inequalities holds

$$\max_{U\in\mathcal{U}_{\mathcal{K}}} \rho_R(UM) \leq \max_{Q\in\mathcal{Q}_{\mathcal{K}}} \rho_R(QM) \leq \mu_{\mathcal{K}}(M) \leq \inf_{D\in\mathcal{D}_{\mathcal{K}}} \bar{\sigma}(DMD^{-1}).$$

Note that there are important differences between some of these properties and their complex μ versions. The scaling property a) holds for $\gamma \in \mathbf{C}$ for complex μ problems. The function ρ_R may be replaced by ρ in any of the above for complex μ problems.

Note that μ is invariant to matrices in $\mathcal{U}_{\mathcal{K}}$ (which are block unitary) [l)] but not to matrices in $\mathcal{Q}_{\mathcal{K}}$ (which may not be unitary) [k)]. The reason that we introduce the set $\mathcal{Q}_{\mathcal{K}}$ is that it will be seen in Section IV that $\max_{Q\in\mathcal{Q}_{\mathcal{K}}} \rho_R(QM) = \mu_{\mathcal{K}}(M)$, whereas it is easy to construct examples where $\max_{U\in\mathcal{U}_{\mathcal{K}}} \rho_R(UM) \neq \mu_{\mathcal{K}}(M)$. Thus we find that the real parameters may take worst-case values that are not on the boundary of the allowable set. This contrasts with the complex μ case where $\mathcal{Q}_{\mathcal{K}} \equiv \mathcal{U}_{\mathcal{K}}$ so that one may always assume the worst-case perturbation is μ-invariant and on the boundary of the allowable set (i.e., unitary). Note that for mixed problems one may still assume the complex blocks of the worst-case perturbation are unitary.

Mixed μ is invariant to similarity transformations with matrices in $\mathcal{D}_{\mathcal{K}}$ [m)] and this leads to the third inequality in n) which is the usual complex μ upper bound. This upper bound is exact for purely complex block structures with $2m_c + m_C \leq 3$ (see [7]). In general, however, it is not exact for any mixed block structures with $m_r \neq 0$. In fact, a better upper bound for mixed μ problems with $m_r \neq 0$ will be presented in Section IV (see [1]).

IV. Upper and Lower Bounds for Mixed μ

Note that although Definition 1 gives an exact expression for μ, it involves an optimization problem which is not, in general, convex. It is now known that the problem is NP hard [8], so to yield tractable computation schemes, attention has focused on upper and lower bounds for mixed μ problems.

The key to obtaining a lower bound lies in the fact that the μ problem may be reformulated as a real eigenvalue maximization. The following theorem is taken from [2].

Theorem 1 [2]: For any matrix $M \in \mathbf{C}^{n\times n}$ and any compatible block structure $\mathcal{K}$

$$\max_{Q\in\mathcal{Q}_{\mathcal{K}}} \rho_R(QM) = \mu_{\mathcal{K}}(M). \tag{10}$$

The function $\rho_R(QM)$ is nonconcave, and so we can only guarantee to compute a local maximum in (10). Thus we only guarantee to compute a lower bound to μ (which is the global maximum).

Now consider an upper bound for μ. Note that one could, for the purposes of the upper bound, cover the real perturbations with complex ones (and then use the complex μ upper bound) since this would cover the admissible perturbation set $X_{\mathcal{K}}$. This approach, however, does not exploit the phase information that is present in the real perturbations, and hence the bound is frequently poor. The upper bound presented in [1] does exploit this phase information and gives a bound which is never worse than the standard upper bound from complex μ theory [see property n)] and is frequently much better. The following is a slightly modified version of results from [1].

Theorem 2 [1]: For any matrix $M \in \mathbf{C}^{n\times n}$ and any compatible block structure $\mathcal{K}$

$$\mu_{\mathcal{K}}(M) \leq \inf_{D\in\mathcal{D}_{\mathcal{K}}, G\in\mathcal{G}_{\mathcal{K}}} \min_{0\leq\beta\in\mathbf{R}} \{\beta\colon M^*DM + \mathbf{j}(GM - M^*G) - \beta^2 D \leq 0\}. \tag{11}$$

Since the above minimization involves a linear matrix inequality, it is quasi-convex (so that all local minima are global), and hence this bound is computationally tractable. The practical computation of these upper and lower bounds is treated in [3].

V. General Problems

Of course, in general, we cannot guarantee to compute μ exactly, for reasons discussed earlier. We can still gain some insight into the properties of the problem, however, from an examination of the conditions at the minimum of the upper bound.

Suppose that we have $\beta > 0$ as a candidate solution of the upper bound minimization problem (11) (if $\beta = 0$ it trivially equals μ), and if the infimum is achieved, then take D_0, G_0 as the corresponding arguments. If the infimum is not achieved, then D_0, G_0 are defined as follows. First choose sequences D^k, G^k, β_k with $D^k \in \mathcal{D}_{\mathcal{K}}, G^k \in \mathcal{G}_{\mathcal{K}}$ such that $\bar{\lambda}(M^*D^kM + \mathbf{j}(G^kM - M^*G^k) - \beta_k^2 D^k) = 0$ and $\beta_k \downarrow \beta$. Note that we can normalize each element of the sequence such that $\bar{\sigma}(\text{block diag } (D^k, G^k)) = 1$. Then D^k, G^k are bounded, so that by passing to a subsequence we have convergent sequences, and we define D_0, G_0 via the appropriate limits, i.e., $D^k \to D_0$ and $G^k \to G_0$ with $D_0 \in \bar{\mathcal{D}}_{\mathcal{K}}$, $G_0 \in \mathcal{G}_{\mathcal{K}}$ and $D_0 \geq 0$.

Theorem 3: Suppose we have $M \in \mathbf{C}^{n\times n}$ together with D_0, G_0 and $\beta \geq 0$ as above. Then if the maximum eigenvalue of $(M^*D_0M + \mathbf{j}(G_0M - M^*G_0) - \beta^2 D_0)$ is distinct, $\beta = \mu_{\mathcal{K}}(M)$.

Proof: If $\beta = 0$ the result is trivial, so assume $\beta > 0$. Choose x as the unit norm eigenvector, corresponding to the maximum eigenvalue. Then it is easy to check, via a simple perturbation argument, that we must have

$$x^*(M^*DM - \beta^2 D)x \geq 0 \quad \text{for all} \quad D \in \mathcal{D}_{\mathcal{K}}$$
$$x^*(GM - M^*G)x \geq 0 \quad \text{for all} \quad G \in \mathcal{G}_{\mathcal{K}}.$$

But now by continuity, and the definition of $\mathcal{D}_{\mathcal{K}}, \mathcal{G}_{\mathcal{K}}$, this implies that in fact we have

$$x^*(M^*DM - \beta^2 D)x \geq 0 \quad \text{for all} \quad D \in \bar{\mathcal{D}}_{\mathcal{K}}, D \geq 0$$
$$x^*(GM - M^*G)x = 0 \quad \text{for all} \quad G \in \mathcal{G}_{\mathcal{K}}. \tag{12}$$

Now suppose that $(Mx)_i$ and x_i represent one of the block components of (Mx) and x, and that D_i and G_i represent the corresponding blocks for D and G. We will choose every other block of D and G to be identically zero.

Suppose first that we have a full complex block. Choose $D_i = I$ and (12) implies that $|(Mx)_i| \geq \beta|x_i|$. Thus there exists a matrix Δ_i, with $\bar{\sigma}(\Delta_i) \leq 1$, such that $\Delta_i(Mx)_i = \beta x_i$. For a repeated complex scalar block in addition to the above, one may choose D_i as a positive semidefinite matrix with a kernel spanned by $(Mx)_i$. Then we have $D_i(Mx)_i = 0$, so that applying (12) we find that $D_ix_i = 0$. By construction of D_i, this implies there exists a complex scalar δ_i, such that $\delta_i(Mx)_i = \beta x_i$, and from the above we may take $|\delta_i| \leq 1$. Finally for a repeated real scalar block then in addition to the above we may choose $G_i = I$ and (12) implies that $x_i^*(Mx)_i = (Mx)_i^*x_i$. This implies $\delta_i^*|(Mx)_i|^2 = \delta_i|(Mx)_i|^2$ and so we may take $\delta_i \in \mathbf{R}$.

Applying the above relationships to each block component, and stacking them up, we obtain $\Delta \in X_{\mathcal{K}}$, $\bar{\sigma}(\Delta) \leq 1$ such that $\Delta Mx = \beta x$. Thus $\beta \leq \mu_{\mathcal{K}}(M)$ and hence $\beta = \mu_{\mathcal{K}}(M)$. □

This theorem is an extension of a result in [1], where it is additionally assumed that the infimum in the μ upper bound (11) is achieved. In fact, this assumption that the infimum is achieved has strong implications for the worst-case perturbation.

Theorem 4: Suppose we have $M \in \mathbf{C}^{n\times n}$ and a compatible block structure $\mathcal{K}$. Then if the infimization in the μ upper bound (11) is achieved and equals $\mu_{\mathcal{K}}(M)$, we have

$$\max_{U\in\mathcal{U}_{\mathcal{K}}} \rho_R(UM) = \mu_{\mathcal{K}}(M). \tag{13}$$

Proof: Recall that $\max_{U\in\mathcal{U}_{\mathcal{K}}} \rho_R(UM) \leq \mu_{\mathcal{K}}(M)$. Hence the result is trivial for $\mu_{\mathcal{K}}(M) = 0$, so assume $\mu_{\mathcal{K}}(M) > 0$. Then by a simple scaling argument we may, without loss of generality, assume $\mu_{\mathcal{K}}(M) = 1$. Suppose we have the perturbation $Q \in \mathcal{Q}_{\mathcal{K}}$ achieving (10), or in other words $Q \in \mathcal{Q}_{\mathcal{K}}$ and $x \in \mathbf{C}^n$ such that $QMx = x$ with $x \neq 0$. This implies that the block components of the vectors x and Mx satisfy

$$\begin{aligned} q_i^r(Mx)_{r_i} &= x_{r_i} \quad \text{for} \quad i = 1, \cdots, m_r \\ q_i^c(Mx)_{c_i} &= x_{c_i} \quad \text{for} \quad i = 1, \cdots, m_c \\ Q_i^C(Mx)_{C_i} &= x_{C_i} \quad \text{for} \quad i = 1, \cdots, m_C. \end{aligned} \tag{14}$$

Now by assumption we have $D \in \mathcal{D}_{\mathcal{K}}, G \in \mathcal{G}_{\mathcal{K}}$ such that $(M^*DM + \mathbf{j}(GM - M^*G) - D) \leq 0$, so that in particular

$$x^*(M^*DM + \mathbf{j}(GM - M^*G) - D)x \leq 0.$$

Expanding this expression out and substituting for (14), one can derive that

$$\sum_{i=1}^{m_r} \left|D_i^{\frac{1}{2}}(Mx)_{r_i}\right|^2 \leq \sum_{i=1}^{m_r} |q_i^r|^2 \left|D_i^{\frac{1}{2}}(Mx)_{r_i}\right|^2 \tag{15}$$

(note that all the complex blocks q_i^c, Q_i^C are unitary). Since we have $|q_i^r| \leq 1$ for all $i = 1, \cdots, m_r$, this implies that

$$|q_i^r| < 1 \Longrightarrow \left|D_i^{\frac{1}{2}}(Mx)_{r_i}\right| = 0.$$

But in this case, since $D_i > 0$, this implies $(Mx)_{r_i} = 0$, and hence by (14), $x_{r_i} = 0$. Thus we may take $q_i^r = 1$ for all such blocks and still satisfy $QMx = x$ but now with $Q \in \mathcal{U}_{\mathcal{K}}$. □

Note that this theorem says that whenever the upper bound is achieved and equals μ, then the worst-case perturbation may be taken to be on a vertex. Of course for the complex blocks we can always restrict our attention to the boundary of the uncertainty set, but for the real uncertainties this is not the case. It is interesting to consider whether the worst-case perturbations are typically on a vertex or require internal reals. In fact, it can be shown that, at least for rank-one problems, neither case is generic [9], [10].

VI. Special Cases

Note that although the general mixed μ problem is NP hard [8], this does not mean that every μ problem is hard to compute. In particular, if we consider restricted special cases of the general problem, by imposing additional structure on M, then it is possible to find classes of mixed μ problems which we can compute exactly.

A. Some Simple Special Cases

First we briefly consider some elementary special cases for which computation of μ is easy. These results are simple extensions of results for the complex μ case, and we include them here for completeness. The proofs are left to the reader (or see [9]). First note that for Hermitian matrices we can trivially obtain the following result.

Lemma 1: For any Hermitian matrix $M \in \mathbf{C}^{n\times n}$ and any compatible block structure $\mathcal{K}$, then $\mu_{\mathcal{K}}(M) = \bar{\sigma}(M) = \rho(M) = \rho_R(M)$.

Next we consider positive matrices, i.e., matrices whose elements are positive real numbers. For these matrices there is a wealth of results from Perron–Frobenius theory (see [11], for example) regarding eigenvalues and singular values, and these lead to the following result for μ.

Lemma 2: For any positive matrix $M \in \mathbf{R}^{n\times n}$ and any compatible block structure $\mathcal{K}$ with $m_C = 0$ (i.e., only scalar uncertainties), then $\mu_{\mathcal{K}}(M) = \min_{D\in\mathcal{D}_{\mathcal{K}}} \bar{\sigma}(DMD^{-1}) = \rho(M) = \rho_R(M)$.

These two cases are not of too much interest in themselves, since they are rarely encountered in practice. They may be of some interest, however, in providing crude bounds for more general problems. One example of such an application for the complex μ problem is provided in [12] where the optimal scalings for the positive matrix case are used to approximate the optimal scalings for a more general μ problem.

B. The Rank-One Case and "Kharitonov-Type" Results

Consider the problem of computing mixed μ for a matrix of rank one. This situation arises in a number of different settings, including perturbed single-input–single-output (SISO) coprime factor models and interval plants. One particular case which gives rise to a rank-one problem is the so-called "affine parameter variation" problem for a polynomial with perturbed coefficients. This problem has been examined in detail in the literature, and several celebrated "Kharitonov-type" results have been proven for variations of this problem (see [13], for example). We refer the reader to [10], [14], and [15] for examples of classes of problems that lead to rank-one μ problems.

The rank-one case is studied in detail in [10], where the following result is proven.

Theorem 5: Suppose we have a rank-one matrix $M \in \mathbf{C}^{n\times n}$, then $\mu_{\mathcal{K}}(M)$ equals its upper bound from Theorem 2.

This theorem says that for rank-one problems, μ equals its upper bound and is hence equivalent to a convex problem. This reinforces the results of [15] and offers some insight into why the problem becomes so much more difficult when we move away from the "affine parameter variation" case to the "multilinear" or "polynomial" cases (see [16]). These correspond to μ problems where M is not necessarily rank one, and hence may no longer be equal to the upper bound, and so may no longer be equivalent to a convex problem (note that there exist rank two matrices for which μ does not equal its upper bound).

C. Real Matrices

As we noted earlier, it is always possible to obtain an upper bound for a mixed μ problem simply by treating the real parameters as complex and using the standard complex μ upper bound. The upper bound from Theorem 2. however, is frequently much better than the complex μ upper bound because of the extra degrees-of-freedom we have in choosing the G scaling matrix (note that if we restrict ourselves to $G = 0_n$, we recover the complex μ upper bound). Unfortunately it is not always possible to improve upon the complex μ upper bound via the G scaling matrix as is illustrated by the following results.

Theorem 6: Given a matrix $M \in \mathbf{R}^{n\times n}$ and any block structure $\mathcal{K}$, define the following subsets of $\mathcal{D}_{\mathcal{K}}$ and $\mathcal{G}_{\mathcal{K}}$

$$\begin{aligned}\mathcal{D}_{R\mathcal{K}} &= \{D \in \mathcal{D}_{\mathcal{K}} : D \in \mathbf{R}^{n\times n}\}\\ \mathcal{G}_{R\mathcal{K}} &= \{G \in \mathcal{G}_{\mathcal{K}} : \mathbf{j}G \in \mathbf{R}^{n\times n}\}\end{aligned} \tag{16}$$

and the following optimization problems

$$\begin{aligned}\alpha_* &= \inf_{D\in\mathcal{D}_{\mathcal{K}}, G\in\mathcal{G}_{\mathcal{K}}} \left[\min_{\alpha\in\mathbf{R}}\{\alpha : M^*DM + \mathbf{j}(GM - M^*G) - \alpha D \leq 0\}\right]\\ \hat{\alpha} &= \inf_{D\in\mathcal{D}_{R\mathcal{K}}, G\in\mathcal{G}_{R\mathcal{K}}} \left[\min_{\alpha\in\mathbf{R}}\{\alpha : M^*DM + \mathbf{j}(GM - M^*G) - \alpha D \leq 0\}\right]\end{aligned} \tag{17}$$

then $\alpha_* = \hat{\alpha}$.

Proof: Clearly we have that $\alpha_* \leq \hat{\alpha}$. Now suppose we have $D \in \mathcal{D}_{\mathcal{K}}, G \in \mathcal{G}_{\mathcal{K}}$ and $\alpha \in \mathbf{R}$ such that $M^*DM + \mathbf{j}(GM - M^*G) - \alpha D \leq 0$. Split D and G into their real and imaginary parts as $D = D_R + \mathbf{j}D_I$, $G = G_R + \mathbf{j}G_I$ with $D_R, D_I, G_R, G_I \in \mathbf{R}^{n\times n}$. Then it is easy to show that D_R, G_R are real symmetric, and D_I, G_I are real skew symmetric. Now we have that

$$\begin{aligned}&M^*DM + \mathbf{j}(GM - M^*G) - \alpha D \leq 0\\ \Longrightarrow\ &x^*(M^*DM + \mathbf{j}(GM - M^*G) - \alpha D)x \leq 0 \quad \forall x \in \mathbf{C}^n\\ \Longrightarrow\ &x^T(M^TDM + \mathbf{j}(GM - M^TG) - \alpha D)x \leq 0 \quad \forall x \in \mathbf{R}^n.\end{aligned}$$

Now we note that $(M^TDM + \mathbf{j}(GM - M^TG) - \alpha D) = S + \mathbf{j}W$ where

$$\begin{aligned}S &= M^TD_RM + \mathbf{j}((\mathbf{j}G_I)M - M^T(\mathbf{j}G_I)) - \alpha D_R\\ W &= M^TD_IM + G_RM - M^TG_R - \alpha D_I.\end{aligned}$$

It is easy to check that S is real symmetric, and W is real skew symmetric, so that $x^T(S + \mathbf{j}W)x = x^TSx\ \forall x \in \mathbf{R}^n$. Thus we have

$$\begin{aligned}&x^T(M^TD_RM + \mathbf{j}((\mathbf{j}G_I)M - M^T(\mathbf{j}G_I)) - \alpha D_R)x \leq 0 \quad \forall x \in \mathbf{R}^n\\ \Longrightarrow\ &(M^*D_RM + \mathbf{j}((\mathbf{j}G_I)M - M^*(\mathbf{j}G_I)) - \alpha D_R) \leq 0.\end{aligned}$$

Similarly we can show that $D > 0 \Longrightarrow D_R > 0$ and so $D_R \in \mathcal{D}_{R\mathcal{K}}, (\mathbf{j}G_I) \in \mathcal{G}_{R\mathcal{K}}$ which gives $\hat{\alpha} \leq \alpha_*$ and hence $\alpha_* = \hat{\alpha}$. □

Basically Theorem 6 says that when computing the upper bound for real matrices, we may restrict our attention to purely real $D \in \mathcal{D}_{\mathcal{K}}$ (i.e., $D \in \mathcal{D}_{\mathcal{K}}$ is real symmetric) and purely imaginary $G \in \mathcal{G}_{\mathcal{K}}$ (i.e., $G \in \mathcal{G}_{\mathcal{K}}$ is of the form $G = \mathbf{j}\hat{G}$ where $\hat{G}$ is real skew symmetric). As a consequence of this we immediately obtain the following theorem.

Theorem 7: Suppose we have a real matrix $M \in \mathbf{R}^{n\times n}$ and a block structure $\mathcal{K}$ with $k_i = 1$ for $i = 1, \cdots, m_r$ (i.e., none of the real scalars are repeated), then

$$\inf_{D\in\mathcal{D}_{\mathcal{K}}, G\in\mathcal{G}_{\mathcal{K}}} \left[\min_{\alpha\in\mathbf{R}}\{\alpha\colon M^*DM + \mathbf{j}(GM - M^*G) - \alpha D \le 0\}\right] = \inf_{D\in\mathcal{D}_{R\mathcal{K}}} \left[\min_{\alpha\in\mathbf{R}}\{\alpha\colon M^*DM - \alpha D \le 0\}\right] \tag{18}$$

where $\mathcal{D}_{R\mathcal{K}}$ is defined as in Theorem 6.

Proof: Apply Theorem 6 to conclude that we may restrict our attention in the left-hand side of (18) to $D \in \mathcal{D}_{R\mathcal{K}}$, $G \in \mathcal{G}_{R\mathcal{K}}$. Now note that for this block structure (none of the real scalars are repeated) G is diagonal (and Hermitian) and hence purely real. Thus we have $\mathcal{G}_{R\mathcal{K}} = \{0_n\}$. □

Note that Theorem 7 says that for μ problems involving real matrices where none of the real scalars are repeated, the choice $G = 0_n$ in the upper bound is optimal, or in other words the mixed μ upper bound equals the complex μ upper bound. This result has important implications for the analysis of linear systems subject to real parameter uncertainty. For the remainder of this section we will assume that none of the real parameters are repeated, so that Theorem 7 applies.

Note first of all that for a (nominally) stable linear system, then the transfer matrix approaches a real matrix at high and low frequencies. Thus we immediately find that the mixed and complex μ upper bounds approach each other at the extreme frequency limits. This means that if the peak value of the μ plot (across frequency) is occurring at one of these limits, then we will not be able to differentiate between real and complex parameters in the upper bounds. For such problems, we may refine the bounds via more powerful computational tools, such as Branch-and-Bound [17]. Note however that we will not be able to guarantee reasonable computation times on large problems for such schemes, since the problem is NP hard. Recall that the only large problems we can (currently) guarantee to compute in reasonable time are those where μ equals its upper bound.

It is possible to avoid using a frequency sweep when computing μ, by noting that the state-space representation of a linear system can be rewritten as a linear fractional transformation. This state-space μ test is fully detailed in [6] (for the complex μ case), and the procedure essentially amounts to including the frequency parameter, z, as one of the uncertainties against which robustness is desired. In this way one obtains a one-shot test, involving a constant matrix μ problem, for the worst-case robustness margin across all frequencies. Note, however, that this matrix is real (since A, B, C, D are), so that this test will not differentiate at all (in the upper bound) between real and complex parameters. Again one may alleviate the situation, at additional computational expense, via Branch-and-Bound techniques. Note that in this setting the Branch-and-Bound procedure is also partitioning the frequency variable z, so that this may be thought of as doing an intelligent frequency sweep.

In fact, the complex μ upper bound for this state-space μ test is a robustness margin in its own right. It is shown in [18] and [19] that this test is exact, but with respect to linear time-varying uncertainties. The implications of Theorem 7 for robustness analysis with time-varying parametric and dynamic uncertainties is currently under investigation.

Note that by further restricting this class to purely real μ problems, we obtain the following well known "vertex result."

Lemma 3: Suppose we have a real matrix $M \in \mathbf{R}^{n\times n}$ and a block structure $\mathcal{K}$ with $m_r = n$, $m_c = m_C = 0$, and $k_i = 1$ for $i = 1, \cdots, m_r$ (i.e., none of the real scalars are repeated). Then it suffices to consider perturbations at the vertices of the allowed perturbation set.

This result follows immediately from the fact that $\det(I_n - \Delta M)$ for $\Delta \in X_{\mathcal{K}}$ is a real-valued multilinear function of the δ_i^r's. For these problems then we can compute μ exactly by checking a finite number of points. Note, however, that the required computation grows exponentially with problem size, so that this result is only applicable to small problems, and in fact, even this restricted class of the mixed μ problem is NP hard [20]. The extent to which assumptions about the structure of M can affect the computational complexity of the problem is evident on comparing Lemmas 2 and 3.

REFERENCES

[1] M. K. H. Fan, A. L. Tits, and J. C. Doyle, "Robustness in the presence of mixed parametric uncertainty and unmodeled dynamics," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 25–38, 1991.

[2] P. M. Young and J. C. Doyle, "Computation of μ with real and complex uncertainties," in *Proc. 29th Conf. Decision Contr.*, 1990, pp. 1230–1235.

[3] P. M. Young, M. P. Newlin, and J. C. Doyle, "Practical computation of the mixed μ problem," in *Proc. Am. Contr. Conf.*, 1992, pp. 2190–2194.

[4] G. J. Balas, J. C. Doyle, K. Glover, A. K. Packard, and R. S. Smith, *The μ Analysis and Synthesis Toolbox.* MathWorks and MUSYN, 1991.

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

[6] A. K. Packard and J. C. Doyle, "The complex structured singular value," *Automatica*, vol. 29, pp. 71–109, 1993.

[7] ——, "Structured singular value with repeated scalar blocks," in *Proc. Am. Contr. Conf.*, 1988, pp. 1213–1218.

[8] R. D. Braatz, P. M. Young, J. C. Doyle, and M. Morari, "Computational complexity of μ calculation," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 1000–1002, 1994.

[9] P. M. Young, "Robustness with parametric and dynamic uncertainty," Ph.D. dissertation, California Institute of Technology, 1993.

[10] ——, "The rank one mixed μ problem and 'Kharitonov-type' analysis," *Automatica*, vol. 30, pp. 1899–1911, 1994.

[11] R. A. Horn and C. R. Johnson, *Matrix Analysis.* New York: Cambridge Univ. Press, 1985.

[12] M. G. Safonov, "Stability margins for diagonally perturbed multivariable feedback systems," *IEE Proc., Pt D*, vol. 129, pp. 251–256, 1982.

[13] A. C. Bartlett, C. V. Hollot, and H. Lin, "Root locations of an entire polytope of polynomials: It suffices to check the edges," in *Mathematics of Control, Signals and Systems.* New York: Springer Verlag, 1988.

[14] A. Rantzer and A. Megretsky, "A convex parameterization of robustly stabilizing controllers," *Technical Report*, Department of Mathematics, Royal Institute of Technology, Stockholm, Sweden, 1992.

[15] J. Chen, M. K. H. Fan, and C. N. Nett, "The structured singular value and stability of uncertain polynomials: A missing link," *Contr. Syst, Inexact Dynamic Models, ASME*, pp. 15–23, 1991.

[16] A. Sideris and R. S. Sánchez Peña, "Fast computation of the multivariable stability margin for real interrelated uncertain parameters," *IEEE Trans. Automat. Contr.*, vol. 34, pp. 1272–1276, 1989.

[17] M. P. Newlin and P. M. Young, "Mixed μ problems and Branch and Bound techniques," in *Proc. 31st Conf. Decision Contr.*, 1992, pp. 3175–3180.

[18] J. S. Shamma, "Robustness analysis for time-varying systems," in *Proc. 31st Conf. Decision Contr.*, 1992, pp. 3163–3168.

[19] A. Megretsky, "Power distribution approach in robust control," *Technical Report*, Department of Mathematics, Royal Institute of Technology, Stockholm, Sweden, 1992.

[20] J. Rohn and S. Poljak, "Checking robust nonsingularity is NP-hard," *Math. Contr., Signals Syst.*, vol. 6, pp. 1–9, 1992.