A fast algorithm for the computation of an upper bound on the μ-norm*

Craig T. Lawrence,†
André L. Tits†
Department of Electrical Engineering and Institute for Systems Research, University of Maryland, College Park, MD 20742, USA

Paul Van Dooren‡
Université Catholique de Louvain, CESAME, Bâtiment Euler, B-1348 Louvain-la-Neuve, Belgium.

June 10, 1999

Abstract

A fast algorithm for the computation of the optimally frequency-dependent scaled $\mathcal{H}_\infty$-norm of a finite dimensional LTI system is presented. It is well known that this quantity is an upper bound to the "μ-norm"; furthermore, it was recently shown to play a special role in the context of slowly time-varying uncertainty. Numerical experimentation suggests that the algorithm generally converges quadratically.

1 Introduction

In the context of robust control analysis and synthesis a quantity of great interest is the *structured singular value norm*, or μ-norm, of the system. Consider a feedback connection of a continuous-time system with real coefficients as in Figure 1. Let $P(s) = C(sI - A)^{-1}B$ be an $m \times m$ stable transfer

*This paper has already appeared, in slightly different form, in the Proceedings of the 1996 IFAC World Congress [13].

†Research supported in part by NSF's Engineering Research Center Program, under grant NSFD-CDR-88-03012.

‡Supported by the Belgian Programme on Interuniversity Poles of Attraction, initiated

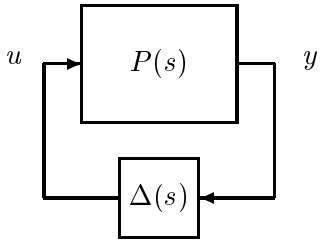

Figure 1: Standard framework for robustness analysis

function matrix and let $\Delta(s)$ be a *structured* perturbation constrained to lie in the set

$$\mathcal{R}(\mathbf{\Delta}) \stackrel{\Delta}{=} \{\ \Delta \in \mathcal{R}_S\ :\ \Delta(s_0) \in \mathbf{\Delta}\ \forall s_0 \in \overline{\mathbb{C}}_+\}$$

where $\mathcal{R}_S$ is the set of all real-rational, proper, stable, $m \times m$ transfer matrices and the uncertainty set

$$\mathbf{\Delta} \stackrel{\Delta}{=} \{\mathrm{diag}[\Delta_1, \ldots, \Delta_F]\ :\ \Delta_j \in \mathbb{C}^{m_j \times m_j}\}$$

is defined for integers $m_1, \ldots, m_F$. Let $\sigma_1(\cdot)$ denote the maximum singular value of its matrix argument and define the complex *structured singular value* for a constant matrix $M \in \mathbb{C}^{m \times m}$ as (see [24] for a complete discussion of the structured singular value)

$$\mu_{\mathbf{\Delta}}(M) \stackrel{\Delta}{=} \frac{1}{\min\{\sigma_1(\Delta)\ :\ \Delta \in \mathbf{\Delta},\ \det(I - M\Delta) = 0\}}$$

unless $\det(I - M\Delta) \neq 0$ for all $\Delta \in \mathbf{\Delta}$, in which case $\mu_{\mathbf{\Delta}}(M) \stackrel{\Delta}{=} 0$. Finally, define the "$\mu$-norm"[1] as:

$$\|P\|_{\mathbf{\Delta}} \stackrel{\Delta}{=} \sup_{\omega \in \mathbb{R}} \mu_{\mathbf{\Delta}}\left(P(\mathrm{j}\omega)\right).$$

It has been shown that the computation of $\mu_{\mathbf{\Delta}}(M)$ is NP-hard (see [22]), in consequence no efficient algorithms are likely to exist for its computation. In practice, a standard upper bound is used in its place. This upper bound is computed as follows. Define

$$\mathcal{D} \stackrel{\Delta}{=} \{\mathrm{diag}[d_1 I_{m_1}, \ldots, d_{F-1} I_{m_{F-1}}, I_{m_F}] : 0 < d_j \in \mathbb{R}\}.$$

by the Belgian State, Prime Minister's Office for Science, Technology and Culture. The scientific responsibility rests with the authors.

[1] Actually, it is not a norm.

Then, for a constant complex matrix M, an upper bound for $\mu_{\mathbf{\Delta}}(M)$ is

$$\hat{\mu}_{\mathbf{\Delta}}(M) \stackrel{\Delta}{=} \inf_{D \in \mathcal{D}} \sigma_1(DMD^{-1}).$$

Substituting this into the expression for $\|P\|_{\mathbf{\Delta}}$, an upper bound on the μ-norm of the system $P(s)$ is obtained as the optimally frequency-dependent scaled $\mathcal{H}_\infty$-norm[2]

$$\|P\|_{\hat{\mu}} \stackrel{\Delta}{=} \sup_{\omega \in \mathbb{R}} \hat{\mu}_{\mathbf{\Delta}}(P(\mathrm{j}\omega)).$$

It is well-known that if $P(s)$ is stable, then $\|P\|_{\mathbf{\Delta}} < 1$ is necessary and sufficient (thus $\|P\|_{\hat{\mu}} < 1$ is sufficient) for uniform robust stability of the $P - \Delta$ loop (Figure 1) for any linear time-invariant structured $\Delta(s)$ of L_2-gain no greater than one (see, e.g., Corollary 3 in [20]). Further, it has recently been shown that $\|P\|_{\hat{\mu}} < 1$ is necessary and sufficient for uniform robust stability of the $P - \Delta$ loop for any linear, arbitrarily slowly time-varying structured Δ of L_2-gain no greater than one [16].

Algorithms for the efficient computation of $\hat{\mu}_{\mathbf{\Delta}}(P(\mathrm{j}\omega))$, for given ω, have long been available. In fact, given ω,

$$\hat{\mu}_{\mathbf{\Delta}}(P(\mathrm{j}\omega))^2 = \inf_{\alpha \geq 0, D \in \mathcal{D}} \{ \alpha \ : \ P(\mathrm{j}\omega) D P(\mathrm{j}\omega)^* - \alpha D < 0 \}, \tag{1}$$

which is a linear matrix inequality (LMI) problem. Efficient algorithms exist for obtaining global solutions to such problems, e.g., [5]. Note that minimizers (or approximate minimizers) D for $\sigma_1(DMD^{-1})$ are related to the minimizers (or approximate minimizers) $\tilde{D}$ for (1) by $D = \tilde{D}^{1/2}$. $\|P\|_{\hat{\mu}}$ is usually computed via a "frequency sweep", i.e. choose a set of frequencies $\{\omega_1, \ldots, \omega_N\}$ and use the approximation

$$\|P\|_{\hat{\mu}} \approx \max\{\hat{\mu}_{\mathbf{\Delta}}(P(\mathrm{j}\omega_1)), \ldots, \hat{\mu}_{\mathbf{\Delta}}(P(\mathrm{j}\omega_N))\}.$$

The drawbacks of this approach are obvious. First, a large number of $\hat{\mu}_{\mathbf{\Delta}}(\cdot)$ computations are required. Second, an upper bound on $\|P\|_{\mathbf{\Delta}}$ is not necessarily obtained. Finally, the result can be arbitrarily bad, i.e. it is difficult to bound the error.

Another approach that has been used to compute an upper bound to the μ-norm is based on the Main Loop Theorem (see, e.g., [15]) and the extension to μ of the Maximum Modulus Theorem (see, e.g., top of p.1201 in [21]). In the discrete-time case, given $P(z) = C(zI - A)^{-1}B + D$, with A

[2] Again, this is not a norm.

stable, and a positive scalar γ, it can be seen that the μ-norm of P is less than γ if and only if $\mu_{\boldsymbol{\Delta_p}}(M(\gamma)) < 1$, where

$$M(\gamma) \triangleq \begin{bmatrix} A & B \\ \frac{1}{\gamma}C & \frac{1}{\gamma}D \end{bmatrix}$$

and $\boldsymbol{\Delta_p}$ is an "augmented" uncertainty structure given by

$$\boldsymbol{\Delta_p} \triangleq \{\mathrm{diag}[\delta I_n, \Delta] : \delta \in \mathbb{C}, \Delta \in \boldsymbol{\Delta}\}.$$

The original idea is due to Doyle and Packard [7]. The continuous-time case can be reduced to the discrete-time case by means of a bilinear transformation (see Section 10.2 in [24] for details). Repeated evaluation of the upper bound $\hat{\mu}_{\boldsymbol{\Delta_p}}(\cdot)$ according to a bisection search over γ yields an upper bound on the μ-norm of P. On the down side, note that this upper bound is generally less tight than that obtained by gridding, because the augmented uncertainty structure involves an additional block which, to make things worse, is of the "repeated" type. The bisection search can be done away with, as shown by Ferreres and Fromion [10], by invoking the "skewed μ" proposed and studied by Fan and Tits [8].[3] Computation then entails the solution of a single LMI constrained quasi-convex generalized eigenvalue minimization problem (GEVP; see [5]).[4] The approach proposed in this paper improves on the scheme of [10] in that it computes the same (tighter) upper bound as the gridding approach. As far as computational cost is concerned, the tradeoff is that of a sequence of LMIs (μ upper bound computations with respect to the original, non-augmented uncertainty structure) versus a single GEVP with a larger number of variables and constraints. Due to the "repeated" block, the complexity of the latter is strongly affected by the dimension of the state space.

Recently, an efficient algorithm has been proposed [4, 6] for the computation of the $\mathcal{H}_\infty$-norm

$$\|P\|_\infty = \sup_{\omega \in \mathbb{R}} \sigma_1(P(\mathrm{j}\omega)). \tag{2}$$

This algorithm makes use of the well-known fact that a given scalar $\xi > 0$ is a singular value of $C(\mathrm{j}\omega - A)^{-1}B$ if and only if $\mathrm{j}\omega$ is an eigenvalue of the

[3] In [10] the continuous-time case is dealt with directly, by limiting the frequency range to a compact interval and using a "repeated real scalar" block in the augmentation. This idea is due to Sideris [17].

[4] To make the optimization problem of [8] into a standard GEVP, a symmetric $n \times n$ matrix of slack variables must be introduced; see p. 8-33 in [11].

related Hamiltonian matrix

$$H(\xi, A, B, C) \triangleq \begin{bmatrix} A & \xi^{-1}BB^* \\ -\xi^{-1}C^*C & -A^* \end{bmatrix}.$$

Given any $\xi \in (0, \|P\|_\infty)$, a set of frequency intervals may thus be computed where maximizers for (2) are known to lie. At step k, ξ is selected as $\sigma_1(P(\mathrm{j}\omega_{k-1}))$ and ω_k is chosen as the mid-point of the largest among the intervals thus determined. A quadratic rate of convergence ensues.

In this paper, using the idea just outlined as a stepping stone, an algorithm is constructed for the fast computation of $\|P\|_{\hat{\mu}}$. (A similar algorithm can be used for the computation of the real stability radius; see [18].)

2 Key Ideas

Let

$$\hat{\mu}(\omega) \triangleq \inf_{D \in \mathcal{D}} \sigma_1(DP(\mathrm{j}\omega)D^{-1}).$$

The goal of the algorithm is to maximize $\hat{\mu}(\omega)$ over $\omega \in \mathbb{R}$. Define

$$\hat{\mu}^* \triangleq \sup_{\omega \in \mathbb{R}} \hat{\mu}(\omega)$$

and

$$\omega^* \triangleq \arg\max_{\omega \in \mathbb{R}} \hat{\mu}(\omega)$$

assuming for now that such a unique maximizer exists. Finally, for a fixed $D \in \mathcal{D}$, define the curve

$$\nu_D(\omega) \triangleq \sigma_1(DP(\mathrm{j}\omega)D^{-1}).$$

While reading the following, it may be helpful to refer to Figure 2. The main idea of the algorithm is as follows (more details are given in Section 3). At iteration k, suppose ξ_k is the best known lower bound to $\hat{\mu}^*$ thus far, and let ω_k be the current trial frequency. Suppose further that ω^* is known to lie in a certain open set Ω_k.[5] Compute

$$D_{k+1} = \arg\min_{D \in \mathcal{D}} \sigma_1(DP(\mathrm{j}\omega_k)D^{-1}),$$

[5]Below, Ω_k is restricted to be a subset of $(0, \infty)$ for all k. Thus, in the case that $\omega^* = 0$, clearly ω^* will not be in Ω_k. It is readily checked that this will not affect any of the results.

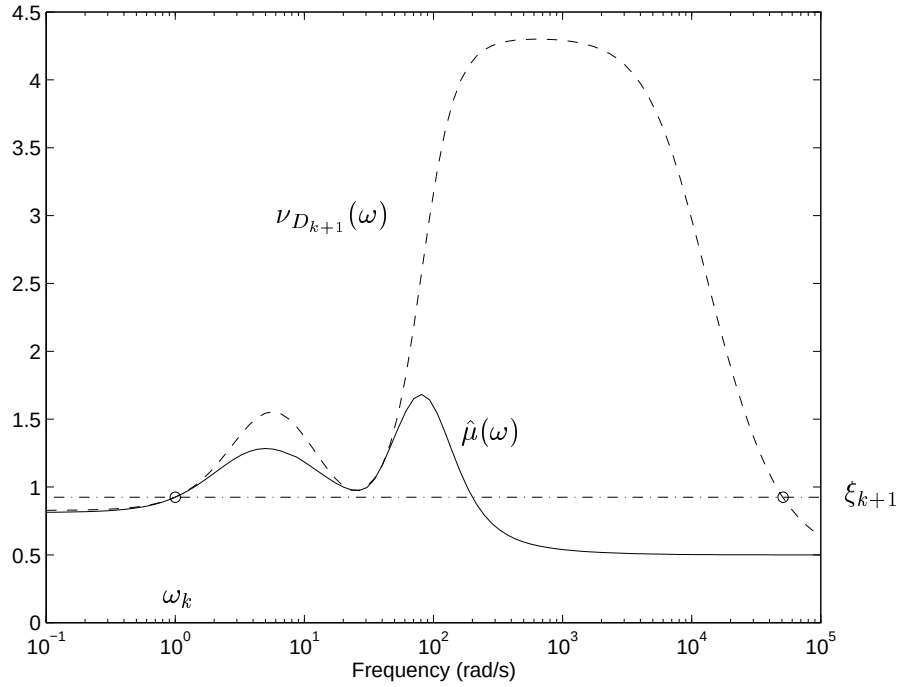

Figure 2: An example of the various definitions.

(assume for simplicity that the minimum is achieved), and let

$$\xi'_{k+1} \stackrel{\Delta}{=} \hat{\mu}(\omega_k) = \nu_{D_{k+1}}(\omega_k).$$

If $\xi'_{k+1} > \xi_k$, take $\xi_{k+1} = \xi'_{k+1}$ as the new estimate of $\hat{\mu}^*$ (as is the case in Figure 2), otherwise keep the old estimate (i.e. $\xi_{k+1} = \xi_k$). Note that

$$\hat{\mu}(\omega) \leq \nu_{D_{k+1}}(\omega) \quad \forall \omega \in \mathbb{R}, \tag{3}$$

with equality at $\omega = \omega_k$, since by definition $\hat{\mu}(\omega)$ is the *lower envelope* of the family of curves $\{\nu_D(\omega) : D \in \mathcal{D}\}$. Now note that $\nu_{D_{k+1}}(\omega)$ is the maximum singular value curve for the transfer function $D_{k+1}P(s)D_{k+1}^{-1}$ at $s = \mathrm{j}\omega$. Thus, as suggested in Section 1, the ideas developed for the computation of the $\mathcal{H}_\infty$-norm may be applied to locate the open set of frequencies, call it Ω'_{k+1}, in which $\nu_{D_{k+1}}(\omega) > \xi_{k+1}$ (in Figure 2, $\Omega'_{k+1} \approx (1, 5.1 \times 10^4)$). In view of (3) and since ξ_{k+1} is a lower bound to $\hat{\mu}^*$, it is clear that (i) if $\Omega'_{k+1} = \emptyset$, then $\hat{\mu}^* = \xi_{k+1}$ and (ii) if $\Omega'_{k+1} \neq \emptyset$, then $\omega^* \in \Omega'_{k+1}$. In the latter case it follows that $\omega^* \in \Omega_{k+1} \stackrel{\Delta}{=} \Omega_k \cap \Omega'_{k+1}$ (in Figure 2, it is assumed

that $\Omega_k \supset \Omega'_{k+1}$, thus $\Omega_{k+1} = \Omega'_{k+1}$). Now, the next frequency ω_{k+1} must be selected. Several possibilities come to mind. For example, ω_{k+1} could simply be chosen as the mid-point of the largest interval contained in Ω_{k+1}. Since, for all k, $\Omega_{k+1} \not\ni \omega_k$, the size of the largest interval in Ω_k would go to zero as $k \to \infty$ and convergence of ξ_k to $\hat{\mu}^*$ would ensue (see Section 4 below). This, however, results in a relatively slow rate of convergence. More sophisticated alternatives are considered in the next section.

A rough outline of the algorithm described above is as follows.

Algorithm
Data: $P(s) = C(sI - A)^{-1}B$, $\mathbf{\Delta}$, $\epsilon > 0$.
Initialization: $k = 0$, pick $\omega_0 \geq 0$, $\xi_0 = 0$, $\Omega_0 = (0, \infty)$.
Step 1: Compute D_{k+1}, ξ_{k+1}.
Step 2: Compute Ω_{k+1}.
Step 3: Compute ω_{k+1}.
Step 4: $k \leftarrow k + 1$. If an appropriate stopping criterion is satisfied, stop. Otherwise, go back to Step 1.

3 Details of the Algorithm

The computation of Ω'_{k+1} is now considered in greater detail. The pure imaginary eigenvalues of the Hamiltonian matrix

$$H_k = H(\xi_{k+1}, A, BD_{k+1}^{-1}, D_{k+1}C)$$

tell us the frequencies at which *one* of the singular values of $D_{k+1}P(\mathrm{j}\omega)D_{k+1}^{-1}$ takes the value ξ_{k+1}. That is, if $\mathrm{j}\tilde{\omega}$ is an imaginary eigenvalue of H_k, then

$$\sigma_r(D_{k+1}P(\mathrm{j}\tilde{\omega})D_{k+1}^{-1}) = \xi_{k+1}$$

for some $r \in \{1, \ldots, m\}$ (see Figure 3). Of course, the frequencies of interest are those for which $r = 1$. Suppose the Schur decomposition of the Hamiltonian matrix H_k has been performed and the $\tilde{\omega}$'s have been identified. Let $\tilde{\omega}_1, \ldots, \tilde{\omega}_q$ be the nonnegative $\tilde{\omega}$'s, sorted in increasing order, and let $\lambda_\ell = \mathrm{j}\tilde{\omega}_\ell$, $\ell = 1, \ldots, q$. In [4], the authors suggest determining those frequencies corresponding to the maximum singular value by eliminating the frequencies for which

$$\xi_{k+1}^2 I - D_{k+1}^{-1}P^*(\mathrm{j}\omega)D_{k+1}^2P(\mathrm{j}\omega)D_{k+1}^{-1}$$

is not positive semi-definite. This is a rather costly computation, though. An alternative scheme is as follows. The pure imaginary simple eigenvalues of

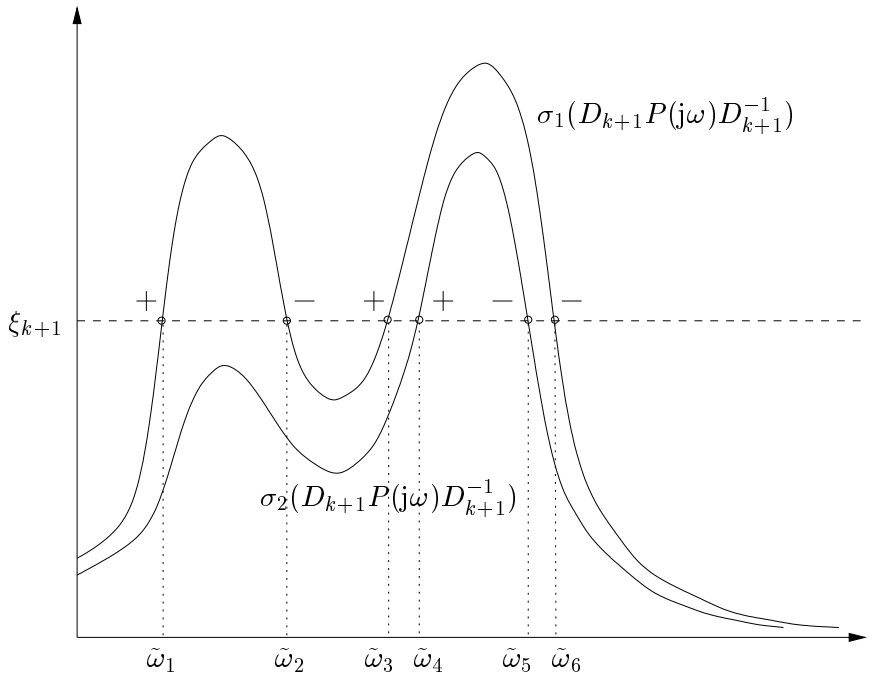

Figure 3: Sample singular value plot.

the Hamiltonian may be differentiated with respect to the scalar parameter ξ via the expression

$$\frac{\partial \lambda_\ell}{\partial \xi}(\xi_{k+1}) = \frac{u_\ell^* \dfrac{\partial H}{\partial \xi}(\xi_{k+1}, A, BD_{k+1}^{-1}, D_{k+1}C)v_\ell}{u_\ell^* v_\ell}, \quad \ell = 1, \ldots, q,$$

(see, e.g., Theorem IV.2.3 in [19]) where v_ℓ and u_ℓ^* are respectively right and left eigenvectors of H_k corresponding to the imaginary eigenvalue λ_ℓ; these few vectors can be computed from the Schur decomposition at little extra cost, e.g., using "inverse iteration". Since $\lambda_\ell = \mathrm{j}\tilde{\omega}_\ell$,

$$\frac{\partial \xi}{\partial \omega}(\tilde{\omega}_\ell) = \left(-\mathrm{j}\frac{\partial \lambda_\ell}{\partial \xi}(\xi_{k+1})\right)^{-1},$$

which is the slope of the corresponding singular value curve for the matrix $D_{k+1}P(\mathrm{j}\omega)D_{k+1}^{-1}$ at the frequency $\omega = \tilde{\omega}_\ell$ and is a real number for every simple eigenvalue on the imaginary axis. This is all the information needed to compute Ω'_{k+1}.

Indeed, define

$$s_\ell \triangleq \frac{\partial \xi}{\partial \omega}(\tilde{\omega}_\ell), \quad \ell = 1, \ldots, q.$$

For the example in Figure 3, the sign of s_ℓ is marked at each intersection point. As $P(s)$ is strictly proper, $\nu_{D_{k+1}}(\omega) \to 0$ as $\omega \to \infty$. Thus, as long as $\xi_{k+1} > 0$,[6] $\tilde{\omega}_q$ must correspond to the largest singular value, i.e. $\nu_{D_{k+1}}(\tilde{\omega}_q) = \xi_{k+1}$, and Ω_{k+1} will be bounded. Let c_ℓ, $\ell = 1, \ldots, q$, $c_\ell \in \{1, \ldots, m\}$, denote the number of the singular value curve corresponding to $\tilde{\omega}_\ell$. That is,

$$\sigma_{c_\ell}(D_{k+1} P(\mathrm{j}\tilde{\omega}_\ell) D_{k+1}^{-1}) = \xi_{k+1}.$$

Thus, $c_q = 1$. The rest of the c_ℓ's may be assigned by sweeping to the left from $\tilde{\omega}_{q-1}$ and using the following rules:[7]

- If $s_{\ell+1} < 0$ and $s_\ell < 0$ then $c_\ell = c_{\ell+1} + 1$.
- If $s_{\ell+1} < 0$ and $s_\ell > 0$ then $c_\ell = c_{\ell+1}$.
- If $s_{\ell+1} > 0$ and $s_\ell < 0$ then $c_\ell = c_{\ell+1}$.
- If $s_{\ell+1} > 0$ and $s_\ell > 0$ then $c_\ell = c_{\ell+1} - 1$.

Now define

$$\{\hat{\omega}_1, \ldots, \hat{\omega}_r\} \triangleq \{\tilde{\omega}_\ell : c_\ell = 1\},$$

ordered so that $\hat{\omega}_1 < \cdots < \hat{\omega}_r$. Clearly, $r \leq q$. In the obvious manner define $\hat{s}_\ell$, $\ell = 1, \ldots, r$ as the slopes of the maximum singular value curve computed at $\hat{\omega}_\ell$. If $\hat{s}_1 < 0$, then

$$\Omega'_{k+1} = (0, \hat{\omega}_1) \cup (\hat{\omega}_2, \hat{\omega}_3) \cup \cdots \cup (\hat{\omega}_{r-1}, \hat{\omega}_r).$$

If $\hat{s}_1 > 0$, then

$$\Omega'_{k+1} = (\hat{\omega}_1, \hat{\omega}_2) \cup \cdots \cup (\hat{\omega}_{r-1}, \hat{\omega}_r).$$

To help clarify the ideas, consider the example in Figure 3 where $q = 6$. Letting $c_6 = 1$, we begin the left sweep. As $s_6 < 0$ and $s_5 < 0$, $c_5 = c_6 + 1 = 2$. Then, since $s_5 < 0$ and $s_4 > 0$, $c_4 = c_5 = 2$. Next, since $s_4 > 0$ and $s_3 > 0$, $c_3 = c_4 - 1 = 1$. Proceeding in this manner, we find $\{c_1, \ldots, c_6\} =$

[6] Certainly true if $\hat{\mu}(\omega_0) > 0$, which will hold in all but a few pathological cases.

[7] The algorithm as stated does not handle the case when $s_i = 0$ or the eigenvalue curve is non-differentiable. This happens when multiple eigenvalues are found on the imaginary axis. A simple way to get around the difficulty in such cases is to increase the current level ξ_k slightly.

$\{1,1,1,2,2,1\}$. Thus, $r = 4$ and $\{\hat{\omega}_1, \ldots, \hat{\omega}_4\} = \{\tilde{\omega}_1, \tilde{\omega}_2, \tilde{\omega}_3, \tilde{\omega}_6\}$. Finally, since $\hat{s}_1 = s_1 > 0$, $\Omega'_{k+1} = (\tilde{\omega}_1, \tilde{\omega}_2) \cup (\tilde{\omega}_3, \tilde{\omega}_6)$.

As mentioned in Section 2, one way to select the next trial point ω_{k+1} is to pick the mid-point of the largest interval in Ω_{k+1}. A better approach, yielding faster convergence, is to use some form of interpolation scheme in the final iterations. In particular, suppose $(\omega_{k_1}, \xi_{k_1+1})$ and $(\omega_{k_2}, \xi_{k_2+1})$ are the two most recent trial points for which $\hat{\mu}(\omega_{k_i}) = \xi_{k_i+1}$, $i = 1, 2$. Then, for $k = k_i$, $i = 1, 2$, $\xi_{k+1} = \xi'_{k+1}$. Thus ω_k is one of the $\tilde{\omega}_\ell$s for the $(k+1)$st iteration, say $\omega_k = \tilde{\omega}_{\ell_k}$, and the corresponding c_{ℓ_k} is 1, i.e.,

$$\hat{\mu}(\omega_k) = \sigma_1(D_{k+1}P(\mathrm{j}\omega_k)D_{k+1}^{-1}) = \xi_{k+1}.$$

Since the $\hat{\mu}(\cdot)$ curve lies entirely below the $\sigma_1(D_{k+1}P(\mathrm{j}\cdot)D_{k+1}^{-1})$ curve, for all such k, whenever both curves are differentiable at ω_k,

$$\frac{\partial \hat{\mu}}{\partial \omega}(\omega_k) = \frac{\partial}{\partial \omega}\sigma_1(D_{k+1}P(\mathrm{j}\omega_k)D_{k+1}^{-1}) = \frac{\partial \xi}{\partial \omega}(\tilde{\omega}_{\ell_k}) = s_{\ell_k}.$$

Using the available values of $\hat{\mu}$ and of its derivative at ω_{k_1} and ω_{k_2}, a cubic or rational (specifically, linear fractional) function [2] may be passed through the two points and ω_{k+1} may be taken as the maximum of the interpolating function, subject to the constraint $\omega_{k+1} \in \overline{\Omega}_{k+1}$ (closure of Ω_{k+1}).

Finally, a stopping criterion which guarantees that the algorithm terminates with an estimate that is within prescribed $\epsilon > 0$ of $\|P\|_{\hat{\mu}}$ is as follows. At iteration k, given ξ_k, D_k, and Ω_k, define the set

$$S_k \triangleq \{\ \omega\ :\ \nu_{D_k}(\omega) > \xi_k + \epsilon\ \}.$$

Like Ω'_{k+1}, S_k may be computed via an appropriate Hamiltonian eigendecomposition. Since $\omega^* \in \Omega_k$, if $\Omega_k \cap S_k = \emptyset$, then $\omega^* \notin S_k$, hence

$$\xi_k < \|P\|_{\hat{\mu}} \leq \xi_k + \epsilon.$$

4 Convergence

Theorems 1 and 2 below are proved for the case when the mid-point rule is used throughout to select the next trial point ω_{k+1}.

Assumption 1. The choice of ω_0 is such that

$$\hat{\mu}(\omega_0) > 0.$$

Assumption 1 is necessary in order for the algorithm to be well-defined. Typically, $\omega_0 = 0$ will satisfy this assumption; if not, a random search will provide a suitable ω_0 as long as $P(s)$ is not identically 0. Under this assumption, Ω_1 is bounded and, in view of the fact that $\omega_k \notin \Omega_{k+1}$, a simple induction argument shows that the length of the largest interval in Ω_k tends to zero as $k \to \infty$.

Theorem 1. *The sequence $\{\xi_k\}_{k\in\mathbb{N}}$ converges to $\hat{\mu}^*$ as $k \to \infty$.*

Proof. By contradiction. Suppose $\xi_k \not\to \hat{\mu}^*$. Since the sequence $\{\xi_k\}_{k\in\mathbb{N}}$ is monotone non-decreasing, and since $\xi_k \leq \hat{\mu}^*$ for all k, $\xi_k \to \xi^*$ for some $\xi^* < \hat{\mu}^*$. Thus, $E \stackrel{\Delta}{=} \{\omega : \hat{\mu}(\omega) > \xi^*\}$ is not empty and, clearly, $E \subset \Omega_k$ for all k. Since $\hat{\mu}(\cdot)$ is continuous (see [3]), E contains a non-trivial interval S. This contradicts the fact that the size of the largest interval in Ω_k goes to 0. □

For this preliminary convergence analysis, consider now instead the problem of maximizing $\tilde{\mu}(\omega)$, where

$$\tilde{\mu}(\omega) \stackrel{\Delta}{=} \min\{\sigma_1(DP(\mathrm{j}\omega)D^{-1}) : D \in \mathcal{D}, \underline{b}I \leq D^2 \leq \bar{b}I\} \tag{4}$$

with $\bar{b} \geq \underline{b} > 0$ arbitrary prescribed numbers. Note that $\tilde{\mu}(\omega)$ is an upper bound to $\hat{\mu}(\omega)$, that it can be efficiently computed, and that it can be made arbitrarily close to $\hat{\mu}(\omega)$, uniformly over $\omega \in \mathbb{R}$, by selecting $\underline{b} > 0$ small enough and $\bar{b}$ large enough [14]. Let

$$\tilde{\mu}^* \stackrel{\Delta}{=} \sup_{\omega\in\mathbb{R}} \tilde{\mu}(\omega).$$

It is not difficult to show that Theorem 1 still holds when $\hat{\mu}(\omega)$ is replaced with $\tilde{\mu}(\omega)$ and $\hat{\mu}^*$ is replaced with $\tilde{\mu}^*$.

Theorem 2. *Assume that ω^* is the unique global maximizer of $\tilde{\mu}(\omega)$. Then $\omega_k \to \omega^*$ as $k \to \infty$.*

Proof. By contradiction. Let $\mathcal{K}$ be an infinite index set and suppose that $\omega_k \to \tilde{\omega}$ on $\mathcal{K}$, with $\tilde{\omega} \neq \omega^*$ (since Ω_1 is bounded, such $\mathcal{K}$ exists for some $\tilde{\omega}$ in the closure of Ω_1). By continuity of $\tilde{\mu}(\cdot)$ and uniqueness of the global maximizer ω^*, there exists $\delta > 0$ such that, for k large enough, $k \in \mathcal{K}$,

$$\tilde{\mu}(\omega_k) < \tilde{\mu}^* - \delta.$$

Let D_k be the minimizer of (4) at ω_k. Since $\{D_k\}$ lies in a compact set, there exists an infinite index set $\mathcal{K}' \subset \mathcal{K}$ such that $D_k \to D^*$ on $\mathcal{K}'$. By uniform continuity on compact sets,

$$\nu_{D_k}(\omega) < \tilde{\mu}^* - \delta,$$

for all $k \in \mathcal{K}'$, in some fixed neighborhood V of $\tilde{\omega}$. On the other hand, in view of Theorem 1, $\xi_k > \tilde{\mu}^* - \delta$ for k large enough. Thus, for $k \in \mathcal{K}'$ large enough, V is taken out of Ω_k in Step 2 of the algorithm, which contradicts the fact that $\omega_k \to \tilde{\omega}$ on $\mathcal{K}$. □

5 Numerical Experiments

The algorithm was implemented[8] in MATLAB$^{\text{TM}}$. In the implementation, a mid-point rule is used until Ω_k is reduced to just one interval and enough information has been accumulated to compute an interpolating function. All numerical experiments were run on a Sun UltraSparc 10, 300MHz machine with 128Meg of RAM, running Solaris 2.5 operating system.

The performance of the new algorithm was first compared to that of the "skewed μ" approach mentioned in the introduction. For the former, function `mu` of the μ-Analysis and Synthesis Toolbox [1] was used for computing ξ_{k+1} and D_{k+1} in Step 1. The GEVP in the latter was solved using the LMI Control Toolbox Version 1.0.4 [11]. Tolerance parameters were set so as to obtain roughly the same accuracy with both approaches; in particular, option `C` was used in `mu`. On 100 randomly generated[9] stable systems with 10 variables, 5 inputs, 5 outputs and 4 complex uncertainty block problems, there were 5 instances where the new approach yielded a μ-norm upper bound 1% or more lower than the upper bound obtained with the skewed-μ approach; in one of these instances, the difference was about 33%. The computation time was always less with the new algorithm, often by more than an order of magnitude.

To illustrate the behavior of the new algorithm, we then applied it to an example taken from [1]. The system is a model of an experimental aircraft (the HIMAT vehicle) with an $\mathcal{H}_\infty$ sub-optimal pitch axis controller. Uncertainty and performance weightings are included in the model. The resultant system is not strictly proper, hence a state space realization will have a non-zero feed-through matrix E. Redefine the Hamiltonian matrix

[8] The implementation is available from the authors.

[9] MATLAB's function `sysrand` was used.

as

$$H(\xi, A, B, C, E) \triangleq \left[\begin{array}{cc} A & 0 \\ 0 & -A^* \end{array}\right] + \left[\begin{array}{cc} B & 0 \\ 0 & -C^* \end{array}\right] \left[\begin{array}{cc} -E & \xi I \\ \xi I & -E^* \end{array}\right]^{-1} \left[\begin{array}{cc} C & 0 \\ 0 & B^* \end{array}\right], \quad (5)$$

and make sure that the initial frequency ω_0 is such that $\hat{\mu}(\omega_0) > \sigma_1(D)$. Then the algorithm can handle such problems with no other changes. The uncertainty structure $\mathbf{\Delta}$ consists of two 2×2 full blocks (robust performance problem). Figure 4 shows a plot of $\hat{\mu}_{\mathbf{\Delta}}(P(\mathrm{j}\omega))$ (in this case, equal to $\mu_{\mathbf{\Delta}}(P(\mathrm{j}\omega))$) as a function of frequency.

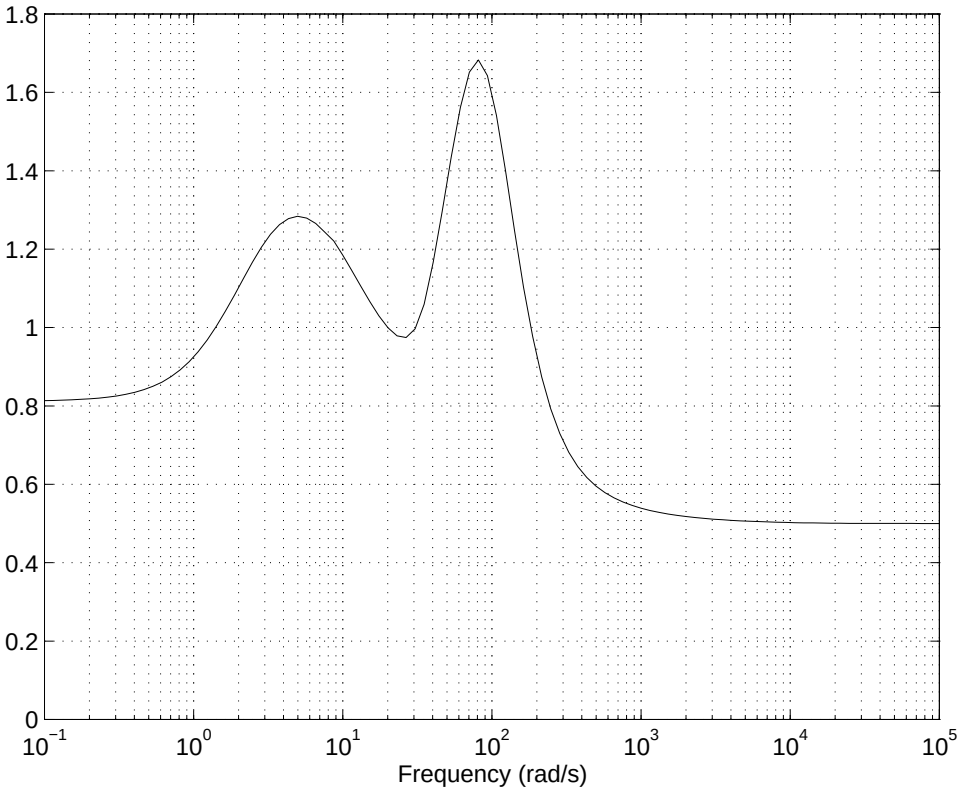

Figure 4: μ-plot for the closed-loop system.

The algorithm may now be applied to compute $\|P\|_{\hat{\mu}}$. In Table 1 the progress of the algorithm from iteration to iteration is shown. All quantities are as defined earlier. A * in the column labeled I indicates that rational

interpolation was used to generate ω_k. If Ω_k contains multiple intervals, or if the algorithm has not yet accumulated enough information, the mid-point rule is used to generate ω_k. It appears from the table that the algorithm exhibits quadratic convergence. Note that, in the process of computing $\hat{\mu}^*$, $\hat{\mu}$ has been evaluated at only a few points, which demonstrates great computational savings in comparison with a grid search. The total computation time was about 2.5 seconds. In comparison, on the same problem, the "skewed μ" approach took over 50 seconds. Interestingly, it yielded the same value for $\hat{\mu}^*$.

k	ω_{k-1}	ξ_k	$\lvert\omega_{k-1} - \omega^*\rvert$	$\lvert\xi_k - \hat{\mu}^*\rvert$	I
1	0	0.7959766	$8.0 \times 10^{+1}$	8.9×10^{-1}	
2	3790.8315	0.7959766	$3.7 \times 10^{+4}$	8.9×10^{-1}	
3	367.31532	0.7959766	$2.9 \times 10^{+2}$	8.9×10^{-1}	
4	129.91042	1.3439461	$5.0 \times 10^{+1}$	3.4×10^{-1}	
5	64.955208	1.6041191	$1.6 \times 10^{+1}$	7.9×10^{-2}	
6	76.939264	1.6791253	3.5×10^{0}	3.7×10^{-3}	*
7	80.240400	1.6827812	2.1×10^{-1}	1.3×10^{-5}	*
8	80.450844	1.6827940	5.6×10^{-4}	9.5×10^{-11}	*

Table 1: Results for example system.

6 Extension to mixed-μ

The algorithm is readily extended to compute an upper bound on the mixed-μ norm, that is, the μ-norm for systems with mixed dynamic and real parametric uncertainty. (See [9] for a sophisticated grid-based approach to this computation, accounting for possible discontinuities.) Two expressions for the widely used "D-G" upper bound $\hat{\mu}_{\mathbf{\Delta}}(M)$ are as follows (see, e.g., Chapter 18 in [23])

$$\hat{\mu}_{\mathbf{\Delta}}(M) = \inf_{\substack{D\in\mathcal{D},G\in\mathcal{G}\\ \xi>0}} \left\{ \xi \ : \ \bar{\sigma}\left(\left(\frac{DMD^{-1}}{\xi} - \mathrm{j}G\right)(I+G^2)^{-1/2}\right) \le 1 \right\}, \tag{6}$$

$$\hat{\mu}_{\mathbf{\Delta}}(M) = \inf_{\substack{D\in\mathcal{D},G\in\mathcal{G}\\ \alpha\ge 0}} \left\{ \alpha \ : \ M^*DM + \mathrm{j}(GM - M^*G) - \alpha^2 D \le 0 \right\}, \tag{7}$$

where $\mathcal{D}$ and $\mathcal{G}$ are determined by the structure. Expression (7) is an efficiently solvable LMI problem (see [5]) and, given α, $\tilde{D}$, $\tilde{G}$ satisfying the constraint in (7), the values $\xi = \alpha$, $D = \tilde{D}^{1/2}$, and $G = \frac{1}{\alpha}\tilde{D}^{-1/2}\tilde{G}\tilde{D}^{-1/2}$ satisfy the constraint in (6) for the same M [23]. Moreover, given ξ, D, and G, it follows from (6), with $M = P(\mathrm{j}\omega)$, that $\hat{\mu}_{\mathbf{\Delta}}(P(\mathrm{j}\omega)) \leq \xi$ for all ω such that

$$\bar{\sigma}(F(\mathrm{j}\omega)) \leq 1, \tag{8}$$

with

$$F(\mathrm{j}\omega) \triangleq \left(\frac{DP(\mathrm{j}\omega)D^{-1}}{\xi} - \mathrm{j}G\right)(I + G^2)^{-1/2}. \tag{9}$$

Note that, given a realization (A, B, C, E) for $P(\mathrm{j}\omega)$, $F(\mathrm{j}\omega)$ can be expressed as $C_F(\mathrm{j}\omega I - A)^{-1}B_F + E_F$, where $B_F = \xi^{-1/2}BD^{-1}(I + G)^{-1/2}$, $C_F = \xi^{-1/2}DC$, and $E_F = \xi^{-1}(DED^{-1} - \mathrm{j}G)(I + G^2)^{-1/2}$ are obtained by inspection from (9). Thus, the end-points of intervals where (8) holds correspond to imaginary eigenvalues of $H = H(1, A, B_F, C_F, E_F)$, where H is as in (5). When $E = 0$ (no feedthrough term in P), this reduces to[10]

$$H = \begin{bmatrix} A & 0 \\ 0 & -A^* \end{bmatrix} + \frac{1}{\xi}\begin{bmatrix} \mathrm{j}BD^{-1}GDC & BD^{-2}B^* \\ -C^*D(I + G^2)DC & \mathrm{j}C^*DGD^{-1}B^* \end{bmatrix}.$$

With this in hand, the algorithm discussed in the previous sections is readily extended to the mixed-μ case. However, the $\hat{\mu}(\cdot)$ curve does not enjoy the same regularity properties as in the purely complex case. As a result, the interpolation rules for updating ω_{k+1} may not be appropriate and quadratic convergence may be lost. In the extreme case where the supremum of $\hat{\mu}(\omega)$ is strictly greater than the supremum of its lower envelope (function whose epigraph is the closure of that of $\hat{\mu}(\omega)$), convergence of ξ_k to $\hat{\mu}^*$ (Theorem 1) may even be lost unless frequency points where $\hat{\mu}$ is discontinuous (including possibly $\omega = \infty$) are checked separately, and the initial value of ξ_k is set accordingly. This can happen even in the trivial case of "real μ" (real uncertainty) of a scalar transfer function such as $p(s) = 1/(s^3 + (3/2)s^2 + s + 1)$, since $\hat{\mu}(p(\mathrm{j}\omega)) = 0$ for $\omega \notin \{0, \pm 1\}$, $\hat{\mu}(p(\mathrm{j}0)) = 1$, $\mu(p(\pm\mathrm{j})) = 2$. Indeed, the real μ of a scalar is equal to its upper bound $\hat{\mu}$ and is the absolute value of the scalar if the scalar is real, and zero otherwise.

[10] Once ξ in Section 5 of [13] is changed to ξ^2 (to match the notation used in the other sections of [13]), the G matrix here corresponds to $\frac{G}{\xi}$ in Section 5 of [13]. The Hamiltonian matrix obtained here is equivalent to that given in [13]; the latter may be obtained from the former by a similarity transformation $S^{-1}HS$ with $S = \begin{bmatrix} I & 0 \\ 0 & -\xi I \end{bmatrix}$.

7 Concluding Remarks

The algorithm just outlined in this note can be refined in various ways. For instance, in the process of minimizing $\nu_D(\omega_k)$ to evaluate $\hat{\mu}(\omega_k)$, if it is found that the minimum value is less than ξ_k, then it is unnecessary to compute it (or D_k) with great accuracy.

The new algorithm can be extended to discrete-time systems. The Hamiltonian eigenvalue problem $H(\xi, A, B, C)$ is replaced by a simplectic eigenvalue problem $S(\xi, A, B, C)$ which will have an eigenvalue $e^{\mathrm{j}\omega}$ on the unit circle if and only if ξ is a singular value of $C(e^{\mathrm{j}\omega}I - A)^{-1}B$. We refer the reader to [12] for the appropriate formulas.

References

[1] G. J. Balas, J. C. Doyle, Keith Glover, Andy Packard, and Roy Smith. *μ-Analysis and Synthesis Toolbox: User's Guide*. MUSYN Inc. and The MathWorks, Inc., July 1993.

[2] J. Barzilai and A. Ben-Tal. Nonpolynomial and inverse interpolation for line search: synthesis and convergence rates. *SIAM J. Numer. Anal.*, 19(6):1263–1277, December 1982.

[3] H. Bercovici, C. Foias, and A. Tannenbaum. Structured interpolation theory. In I. Gohberg, editor, *Extension and Interpolation of Linear Operators and Matrix Functions*, pages 195–220. Birkhauser Verlag, Boston, 1990. Operator Theory, Advances and Applications, vol. 47.

[4] S. Boyd and V. Balakrishnan. A regularity result for the singular values of a transfer matrix and a quadratically convergent algorithm for computing its L_∞-norm. *Systems and Control Letters*, 15:1–7, 1990.

[5] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*. SIAM, Philadelphia, 1994.

[6] N. A. Bruinsma and M. Steinbuch. A fast algorithm to compute the $\mathcal{H}_\infty$-norm of a transfer function matrix. *Systems and Control Letters*, 14(5):287–293, 1990.

[7] J. Doyle and A. Packard. Uncertain multivariable systems from a state space perspective. In *Proceedings of the 1987 American Control Conference (Minneappolis)*, pages 2147–2152, 1987.

[8] M. K. H. Fan and A. L. Tits. A measure of worst-case H_∞ performance and of largest acceptable uncertainty. *Systems and Control Letters*, 18:409–421, 1992.

[9] E. Feron. A more reliable robust stability indicator for linear systems subject to parametric uncertainties. *IEEE Trans. Automatic Control*, 42(9):1326–1330, 1997.

[10] G. Ferreres and V. Fromion. Computation of the robustness margin with the skewed μ tool. *Systems and Control Letters*, 32:193–202, 1997.

[11] P. Gahinet, A. Nemirovski, A.L. Laub, and M. Chilali. *LMI Control Toolbox – For Use with MATLAB*. The MathWorks, Inc., May 1995.

[12] D. Hinrichsen and N. K. Son. The complex stability radius of discrete-time systems and simplectic pencils. In *Proceedings of the 1989 Conference on Decision and Control (Tampa)*, pages 2265–2270, 1989.

[13] C. Lawrence, A. Tits, and P. Van Dooren. A fast algorithm for the computation of an upper bound on the μ-norm. In *Proceedings of the 1996 IFAC World Congress (San Francisco)*, volume H, pages 59–64, July 1996.

[14] L. Lee and A. L. Tits. On continuity/discontinuity in robustness indicators. *IEEE Transactions on Automatic Control*, AC-38(10):1551–1553, 1993.

[15] A. Packard and J. C. Doyle. The complex structured singular value. *Automatica*, 29:71–109, 1993.

[16] K. Poolla and A. Tikku. Robust performance against time-varying structured perturbations. *IEEE Transactions on Automatic Control*, AC-40(9):1589–1602, 1995.

[17] A. Sideris. Elimination of frequency search from robustness tests. *IEEE Transactions on Automatic Control*, AC-37(10):1635–1640, 1992.

[18] J. Sreedhar, P. Van Dooren, and A. L. Tits. A fast algorithm to compute the real structured stability radius. In Rolf Jeltsch and Mohamed Mansour, editors, *Stability Theory: Proceedings of Centenary Conference (Ticino, Switzerland, May 21–26, 1995)*, pages 219–230. Birkhäuser Verlag, 1996.

[19] G. W. Stewart and J. Sun. *Matrix Perturbation Theory.* Academic Press, New York, 1990.

[20] A. L. Tits and V. Balakrishnan. Small-μ theorems with frequency-dependent uncertainty bounds. *Mathematics of Control, Signals, and Systems*, 11(3):220–243, 1998.

[21] A.L. Tits and M.K.H. Fan. On the Small-μ Theorem. *Automatica*, 31(8):1199–1201, 1995.

[22] O. Toker and H. Özbay. On the NP-hardness of the purely complex μ computation, analysis/synthesis, and some related problems in multi-dimensional systems. In *Proceedings of the 1995 American Control Conference*, pages 447–451, June 1995.

[23] K. Zhou and J. Doyle. *Essentials of Robust Control.* Prentice Hall, Englewood Cliffs, NJ, 1998.

[24] K. Zhou, J. Doyle, and K. Glover. *Robust and Optimal Control.* Prentice Hall, Englewood Cliffs, NJ, 1995.