MAXIMALLY ROBUST CONTROLLERS FOR MULTIVARIABLE SYSTEMS*

S. K. GUNGAH[†], G. D. HALIKIAS[‡], AND I. M. JAIMOUKHA[§]

Abstract. The set of all optimal controllers which maximize a robust stability radius for unstructured additive perturbations may be obtained using standard Hankel-norm approximation methods. These controllers guarantee robust stability for *all* perturbations which lie inside an open ball in the uncertainty space (say, of radius r_1). Necessary and sufficient conditions are obtained for a perturbation lying on the boundary of this ball to be destabilizing for *all* maximally robust controllers. It is thus shown that a “worst-case direction” exists along which all boundary perturbations are destabilizing. By imposing a parametric constraint such that the permissible perturbations cannot have a “projection” of magnitude larger than $(1 - \delta)r_1$, $0 < \delta \leq 1$, in the most critical direction, the uncertainty region guaranteed to be stabilized by a *subset* of all maximally robust controllers can be *extended* beyond the ball of radius r_1 . The choice of the “*best*” maximally robust controller—in the sense that the uncertainty region guaranteed to be stabilized becomes as large as possible—is associated with the solution of a superoptimal approximation problem. Expressions for the improved stability radius are obtained and some interesting links with μ -analysis are pursued.

Key words. robust control, μ -analysis, stability radius, superoptimization, Nehari problem

AMS subject classifications. 93B28, 93B40, 93B36

PII. S0363012999350559

1. Notation. $\mathcal{R}$, $\mathcal{R}_+$, and $\mathcal{C}$ denote the sets of real, nonnegative, and complex numbers, respectively. $\mathcal{C}_+$ ($\bar{\mathcal{C}}_+$), $\mathcal{C}_-$ ($\bar{\mathcal{C}}_-$) denote the open (closed) right half-plane and the open (closed) left half-plane, respectively. For a complex matrix A , A^T denotes the transpose while A' denotes the complex-conjugate transpose. $\sigma_i(A)$ denotes the i th largest singular value. The smallest singular value is denoted by $\underline{\sigma}(A)$ and the largest singular value is denoted by $\bar{\sigma}(A)$. The norm of A is defined as $\|A\| = \bar{\sigma}(A)$. For a square A , $\lambda(A)$ is the spectrum of A and $\lambda_{\max}(A)$ is the largest eigenvalue.

$\mathcal{L}_{\infty}^{p \times m}$ denotes the space of all $p \times m$ matrix functions with entries uniformly bounded on the $j\omega$ -axis. $\mathcal{H}_{\infty}^{p \times m}$ and $\mathcal{H}_{\infty}^{-p \times m}$ denote the subspaces of $\mathcal{L}_{\infty}^{p \times m}$ consisting of all matrix functions whose elements are analytic in $\bar{\mathcal{C}}_+$ and $\bar{\mathcal{C}}_-$, respectively. $\|\cdot\|_{\infty}$ denotes the $\mathcal{L}_{\infty}$ norm of matrices in $\mathcal{L}_{\infty}$ or the $\mathcal{H}_{\infty}$ norm of matrices in $\mathcal{H}_{\infty}$, depending on context. $\gamma\mathcal{BH}_{\infty}^{p \times m} = \{G \in \mathcal{H}_{\infty}^{p \times m} : \|G\|_{\infty} \leq \gamma\}$ is the γ ball of $\mathcal{H}_{\infty}^{p \times m}$. The prefix $\mathcal{R}$ before a set symbol means that the elements of the set are restricted to be real-rational. Matrix dimensions of spaces will be occasionally suppressed.

$G(s)^{\sim} := G'(-\bar{s})$ denotes the *para-hermitian conjugate* of $G(s)$. $G(s)^{-\sim}$ stands for $(G(s)^{\sim})^{-1}$. The *Hankel operator* with symbol $G \in \mathcal{H}_{\infty}$ is denoted by Γ_G while $\sigma_i(\Gamma_G)$ denotes the i th largest Hankel singular value of G . The Hankel norm of G , $\sigma_1(\Gamma_G)$, is also written as $\|\Gamma_G\|$ and the smallest Hankel singular value as $\underline{\sigma}(\Gamma_G)$.

A real-rational function $G(s)$ is called *stable* if it has no poles in $\bar{\mathcal{C}}_+$. If $G(s)$ has no poles in $\bar{\mathcal{C}}_-$, it is called *antistable*. Matrix (scalar and vector) transfer functions will

*Received by the editors January 4, 1999; accepted for publication (in revised form) October 18, 1999; published electronically June 20, 2000.

<http://www.siam.org/journals/sicon/38-6/35055.html>

[†]Interdisciplinary Research Center for Process Systems Engineering, Imperial College of Science, Technology and Medicine, London SW7-2BY, UK (s.gungah@ps.ic.ac.uk).

[‡]Department of Electronic and Electrical Engineering, University of Leeds, Leeds LS2 9JT, UK (een6gdh@sun.leeds.ac.uk).

[§]Department of Electrical and Electronic Engineering, Imperial College of Science, Technology and Medicine, London SW7-2BT, UK (i.jaimouka@ic.ac.uk).

be represented by uppercase (lowercase) boldface letters and with the dependence on s mostly suppressed. If $\mathbf{G}^{-1} = \gamma^{-2} \mathbf{G}^\sim$, then $\mathbf{G}$ is called γ -allpass (or simply *allpass* if $\gamma = 1$) and satisfies $\mathbf{G}\mathbf{G}^\sim = \mathbf{G}^\sim\mathbf{G} = \gamma^2 I$. A matrix function $\mathbf{G} \in \mathcal{RH}_\infty$ which satisfies $\mathbf{G}^\sim\mathbf{G} = I$ is called *inner*. A matrix function $\mathbf{G}(s) \in \mathcal{RH}_\infty$ which has full column rank for all $s \in \bar{\mathcal{C}}_+$ is called *outer*. If $\mathbf{U} \in \mathcal{L}_\infty^{l \times q}$ and

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_{11} & \mathbf{H}_{12} \\ \mathbf{H}_{21} & \mathbf{H}_{22} \end{bmatrix} \in \mathcal{L}_\infty^{(p+q) \times (m+l)}$$

with $\mathbf{H}_{11} \in \mathcal{L}_\infty^{p \times m}$, we define the lower linear fractional map $\mathcal{F}_l(\mathbf{H}, \mathbf{U}) = \mathbf{H}_{11} + \mathbf{H}_{12}\mathbf{U}(I - \mathbf{H}_{22}\mathbf{U})^{-1}\mathbf{H}_{21}$, provided that $I - \mathbf{H}_{22}(\infty)\mathbf{U}(\infty)$ is invertible. If $\mathbf{U} \in \mathcal{L}_\infty^{m \times p}$, we define the upper linear fractional map $\mathcal{F}_u(\mathbf{H}, \mathbf{U}) = \mathbf{H}_{22} + \mathbf{H}_{21}\mathbf{U}(I - \mathbf{H}_{11}\mathbf{U})^{-1}\mathbf{H}_{12}$, provided that $I - \mathbf{H}_{11}(\infty)\mathbf{U}(\infty)$ is invertible. If $\mathcal{U}$ is a set, then $\mathcal{F}_l(\mathbf{H}, \mathcal{U})$ denotes the set $\{\mathcal{F}_l(\mathbf{H}, \mathbf{U}) : \mathbf{U} \in \mathcal{U}\}$ and if $\mathbf{G}_1, \mathbf{G}_2, \mathbf{G}_3 \in \mathcal{L}_\infty$ have appropriate dimensions, then $\mathbf{G}_1 + \mathbf{G}_2\mathcal{U}\mathbf{G}_3$ denotes the set $\{\mathbf{G}_1 + \mathbf{G}_2\mathbf{U}\mathbf{G}_3 : \mathbf{U} \in \mathcal{U}\}$.

If $\mathbf{G} \in \mathcal{L}_\infty$, we define, for each i , $s_i^\infty(\mathbf{G}) = \sup_{\omega \in \mathbb{R}} \sigma_i(\mathbf{G}(j\omega))$. Clearly, $s_1^\infty(\mathbf{G}) = \|\mathbf{G}\|_\infty$. Suppose that $\mathcal{T}$ is a set of matrix functions. $\mathbf{T} \in \mathcal{T}$ is called a k th level *superoptimal* function if it minimizes the sequence $\{s_1^\infty(\mathbf{T}), s_2^\infty(\mathbf{T}), \dots, s_k^\infty(\mathbf{T})\}$ with respect to lexicographic ordering among all $\mathbf{T} \in \mathcal{T}$. The minimized sequence is denoted by $\{s_1(\mathcal{T}), \dots, s_k(\mathcal{T})\}$, and the $s_i(\mathcal{T})$ s are called the superoptimal levels of $\mathcal{T}$.

2. Introduction. The work presented in this paper is related to the problem of maximizing the robust stability radius for systems subject to unstructured additive perturbations [25], [6], [23], [24]. In [6] it was shown that this problem is equivalent to a Nehari approximation. Moreover, an explicit state-space parametrization was obtained for all controllers which guarantee a robust stabilization radius $r < r_1$. A parametrization of all maximally robust controllers ($r = r_1$) is implicit in [6] and may be obtained from [4], [5]. The theory of optimal interpolation is used in [23] to give a solution for single input/single output systems.

In the multi-input/single output or single input/multi-output case, the optimal controller is unique. In the matrix case, however, a continuum of optimal controllers typically exists. It is therefore natural to ask whether a subset of these controllers offers improved robust stability properties, in the sense that it guarantees closed-loop stability for a larger class of uncertainties, compared to those offered by the optimal solution set considered in total. More specifically, we seek to identify the set of all controllers which guarantees robust stability for the largest possible region of the uncertainty space containing the open ball of radius r_1 as a subset. Clearly, this can only be achieved by imposing a structure on the set of admissible uncertainties.

Our approach is as follows: From the work in [25], [24], and [6] the maximum robust stability radius r_1 is the inverse of the smallest achievable $\mathcal{H}_\infty$ norm among all interpolating functions $\mathcal{T} = \{\mathbf{K}(I - \mathbf{G}\mathbf{K})^{-1}\}$, as $\mathbf{K}$ varies over the set of all internally stabilizing compensators of $\mathbf{G}$. Using an allpass dilation technique, the set of all optimal interpolating functions $\mathcal{T}_1 = \{\mathbf{T} \in \mathcal{T} : \|\mathbf{T}\|_\infty = r_1^{-1}\} \subseteq \mathcal{T}$ has the form $\mathcal{T}_1 = \mathbf{Y} \text{diag}(r_1^{-1}\mathbf{a}, \hat{\mathbf{R}} + \mathcal{Q})\mathbf{X}$, where $\hat{\mathbf{R}} \in \mathcal{RH}_\infty^-$, $\mathbf{X}$ and $\mathbf{Y}$ are square inner matrices, $\mathbf{a}$ is a scalar allpass function, and $\mathcal{Q}$ is the set of all r_1^{-1} suboptimal Nehari extensions of $\hat{\mathbf{R}}$, i.e., $\mathcal{Q} = \{\mathbf{Q} \in \mathcal{H}_\infty : \|\hat{\mathbf{R}} + \mathbf{Q}\|_\infty \leq r_1^{-1}\}$.

Every optimal controller corresponding to an interpolating function in $\mathcal{T}_1$ stabilizes *all* perturbations which lie inside the open ball $\mathcal{D}_{r_1} = \{\Delta \in \mathcal{L}_\infty : \|\Delta\|_\infty < r_1, \eta(\mathbf{G} + \Delta) = \eta(\mathbf{G})\}$, where $\eta(\cdot)$ denotes number of poles in $\mathcal{C}_+$. Next, we consider perturbations Δ which lie on the boundary of $\mathcal{D}_{r_1}$. It is shown that such boundary

FIG. 1. Closed-loop system.

perturbations are *uniformly* destabilizing (i.e., they destabilize the closed-loop system for every optimal controller) if and only if $|\mathbf{x}^T(j\omega)\mathbf{\Delta}(j\omega)\mathbf{y}(j\omega)| = r_1$ for some frequency ω , and where $\mathbf{x}^T$ and $\mathbf{y}$ are the first row and column of $\mathbf{X}$ and $\mathbf{Y}$, respectively. Moreover, all frequencies ω are equally critical, in the sense that destabilizing boundary perturbations can be constructed for every $\omega \in \mathcal{R}$. This shows that it is futile to attempt to extend the uncertainty set guaranteed to be stabilized by a subset of all optimal controllers in the (frequency-dependent) direction defined by vectors $\mathbf{x}^T$ and $\mathbf{y}$. By imposing a parametric constraint (uniform in ω) such that the permissible perturbations cannot have a “projection” of magnitude larger than $(1 - \delta)r_1$ ($0 < \delta \leq 1$) in this direction, the uncertainty region guaranteed to be stabilized by a subset of all optimal controllers can be extended beyond $\mathcal{D}_{r_1}$. Using a result in [15], it is shown that for each $\delta \in (0, 1]$ the corresponding constrained robust stability radius is maximized by the set of controllers which minimize the first two superoptimal levels of $\mathcal{T}$. A closed-form expression of the improved stability radius is also obtained which involves δ and the first two superoptimal levels of $\mathcal{T}$. This work is related to the results presented in [18] which also uses superoptimization concepts to give an extension of the allowable perturbation set.

An alternative interpretation of our results leads to interesting connections with the problem of robust stabilization of systems subject to structured perturbations and μ -synthesis in general [19]. By suitably defining δ , robust stabilization problems for a number of uncertainty structures can be formulated in our setting, and bounds on the achievable robust-stability radius can be obtained. An upper bound on μ for the constant complex case is derived in the last section.

The layout of the paper is as follows: Section 3 outlines a number of known results in the area of robust stabilization of systems subject to unstructured additive perturbations. The maximum robust stability radius is obtained by solving a Nehari approximation problem [6] and leads to a parametrization of all optimal interpolating functions, using the results of [4], [5]. An alternative parametrization of this set is obtained in section 4, using an allpass dilation technique [7], [5]. A recursive application of this method leads to the solution of the superoptimal approximation problem [26], [22], [16], [13], [14], [12], [20], [21]. In our context, this parametrization of the set of all optimal interpolating functions is used to characterize all uniformly destabilizing boundary perturbations. This analysis is carried out in section 4, which also includes our main result (Theorem 4.8) whose proof is based on a result from [15] (see Theorems 4.6 and 4.7). In section 5 a new (upper) bound on the structured singular value μ is obtained (for the constant problem). Finally, section 6 contains the conclusions.

3. Robust stabilization for unstructured additive perturbations. Let $\mathbf{G} \in \mathcal{RL}_\infty$. When $\mathbf{\Delta} = 0$, the closed-loop system in Figure 1 is internally stable

if and only if it is well-posed, i.e., $\det(I - \mathbf{G}(\infty)\mathbf{K}(\infty)) \neq 0$ and the four transfer functions $(u_1, u_2) \rightarrow (y_1, y_2)$ are in $\mathcal{H}_\infty$. In this case, we write $(\mathbf{G}, \mathbf{K}) \in \mathcal{S}$ and $\mathbf{K} \in \mathcal{K}$. Consider the set of additively perturbed systems $\mathbf{G} + \Delta, \Delta \in \mathcal{D}_r(\mathbf{G}, \mathbf{w})$, where $\mathbf{w}$ is a scalar outer $\mathcal{RH}_\infty$ (weighting) function and

$$\mathcal{D}_r(\mathbf{G}, \mathbf{w}) = \{ \Delta \in \mathcal{L}_\infty : \|\mathbf{w}^{-1}\Delta\|_\infty < r, \eta(\mathbf{G}) = \eta(\mathbf{G} + \Delta) \},$$

in which $\eta(\cdot)$ denotes the number of poles in $\mathcal{C}_+$, counted in a MacMillan degree sense. The system $(\mathbf{G}, \mathbf{K})$ is said to be $(r, \mathbf{w})$ robustly stable if $(\mathbf{G} + \Delta, \mathbf{K}) \in \mathcal{S}$ for all $\Delta \in \mathcal{D}_r(\mathbf{G}, \mathbf{w})$. $(\mathbf{G}, \mathbf{K})$ is said to be maximally robustly stable if (i) $(\mathbf{G}, \mathbf{K})$ is $(r_1, \mathbf{w})$ robustly stable and (ii) there exists $\Delta \in \partial\mathcal{D}_{r_1}(\mathbf{G}, \mathbf{w}) = \{ \Delta \in \mathcal{L}_\infty : \|\mathbf{w}^{-1}\Delta\|_\infty = r_1, \eta(\mathbf{G}) = \eta(\mathbf{G} + \Delta) \}$, such that $(\mathbf{G} + \Delta, \mathbf{K}) \notin \mathcal{S}$. The following theorem [25] gives necessary and sufficient conditions for robust stabilization in the presence of additive unstructured perturbations.

THEOREM 3.1 (see [25], [6], [24]). *Let $\mathbf{G} \in \mathcal{RL}_\infty$ and suppose that $(\mathbf{G}, \mathbf{K}) \in \mathcal{S}$. Then $(\mathbf{G}, \mathbf{K})$ is $(r, \mathbf{w})$ robustly stable if and only if $\|\mathbf{w}\mathbf{K}(I - \mathbf{G}\mathbf{K})^{-1}\|_\infty \leq r^{-1}$.*

The following result [6], [1] shows that, without loss of generality, $\mathbf{G}$ can be assumed to be antistable and $\mathbf{w}$ can be taken to be equal to one.

THEOREM 3.2 (see [6], [1]). *Let $\mathbf{w}^{-1}\mathbf{G}$ have a decomposition $\mathbf{w}^{-1}\mathbf{G} = \mathbf{G}_1 + \mathbf{G}_2$ with $\mathbf{G}_1^\sim, \mathbf{G}_2 \in \mathcal{RH}_\infty$, and $\mathbf{G}_1(\infty) = 0$. Then there exists a $\mathbf{K} \in \mathcal{L}_\infty$ such that $(\mathbf{G}, \mathbf{K})$ is $(r, \mathbf{w})$ robustly stable if and only if $\mathbf{K} = \mathbf{w}^{-1}\mathbf{K}_1(I + \mathbf{G}_2\mathbf{K}_1)^{-1}$ for some $\mathbf{K}_1$ such that $(\mathbf{G}_1, \mathbf{K}_1)$ is $(r, 1)$ robustly stable and $\det\{(I + \mathbf{G}_2\mathbf{K}_1)(\infty)\} \neq 0$.*

Remark 3.1. We assume for simplicity that $\mathbf{G}^\sim \in \mathcal{RH}_\infty$, $\mathbf{G}(\infty) = 0$, and $\mathbf{w} = 1$. We also use the simplified notation

$$(1) \quad \mathcal{D}_r(\mathbf{G}) := \mathcal{D}_r(\mathbf{G}, 1) = \{ \Delta \in \mathcal{L}_\infty : \|\Delta\|_\infty < r, \eta(\mathbf{G}) = \eta(\mathbf{G} + \Delta) \},$$

$$(2) \quad \partial\mathcal{D}_r(\mathbf{G}) := \partial\mathcal{D}_r(\mathbf{G}, 1) = \{ \Delta \in \mathcal{L}_\infty : \|\Delta\|_\infty = r, \eta(\mathbf{G}) = \eta(\mathbf{G} + \Delta) \}.$$

Let $\mathbf{G}$ have left and right coprime factorizations $\mathbf{G} = \mathbf{N}\mathbf{M}^{-1} = \tilde{\mathbf{M}}^{-1}\tilde{\mathbf{N}}$, respectively, with $\mathbf{N}, \mathbf{M}, \tilde{\mathbf{N}}, \tilde{\mathbf{M}} \in \mathcal{RH}_\infty$ and let $\mathbf{U}, \mathbf{V}, \tilde{\mathbf{U}}, \tilde{\mathbf{V}} \in \mathcal{RH}_\infty$ satisfy the Bezout identities $\tilde{\mathbf{V}}\mathbf{M} - \tilde{\mathbf{U}}\mathbf{N} = I$ and $\tilde{\mathbf{M}}\mathbf{V} - \tilde{\mathbf{N}}\mathbf{U} = I$. Then the set of all stabilizing controllers of $\mathbf{G}$ is

$$(3) \quad \mathcal{K} = \{ (\mathbf{U} + \mathbf{M}\mathbf{Q})(\mathbf{V} + \mathbf{N}\mathbf{Q})^{-1} : \mathbf{Q} \in \mathcal{H}_\infty \}.$$

Let $\mathcal{T} = \{ \mathbf{K}(I - \mathbf{G}\mathbf{K})^{-1} : \mathbf{K} \in \mathcal{K} \}$. We refer to $\mathcal{T}$ as the set of all *interpolating functions*. Using the parametrization of $\mathcal{K}$ in (3) gives the alternative characterization of $\mathcal{T}$ as

$$(4) \quad \mathcal{T} = \{ (\mathbf{U} + \mathbf{M}\mathbf{Q})\tilde{\mathbf{M}} : \mathbf{Q} \in \mathcal{H}_\infty \}.$$

Let $\mathbf{G}$ have a minimal balanced realization $\mathbf{G}(s) = C(sI - A)^{-1}B + D$, such that $\operatorname{Re} \lambda_i(A) > 0$ for all i and with $A\Sigma + \Sigma A' = BB'$, $A'\Sigma + \Sigma A = C'C$, $\Sigma = \operatorname{diag}(\sigma_1, \sigma_2, \dots, \sigma_n)$, $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n > 0$. Let $F = B'\Sigma^{-1}$ and $H = \Sigma^{-1}C'$. The coprime factors $\mathbf{N}, \mathbf{M}, \tilde{\mathbf{N}}, \tilde{\mathbf{M}}$ can now be defined (together with $\mathbf{U}, \mathbf{V}, \tilde{\mathbf{U}}, \tilde{\mathbf{V}}$) as

$$\begin{bmatrix} \mathbf{M}(s) & \mathbf{U}(s) \\ \mathbf{N}(s) & \mathbf{V}(s) \end{bmatrix} = \begin{bmatrix} -F \\ C \end{bmatrix} (sI - A + BF)^{-1} \begin{bmatrix} B & H \end{bmatrix} + \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix},$$

$$\begin{bmatrix} \tilde{\mathbf{V}}(s) & -\tilde{\mathbf{U}}(s) \\ -\tilde{\mathbf{N}}(s) & \tilde{\mathbf{M}}(s) \end{bmatrix} = \begin{bmatrix} F \\ -C \end{bmatrix} (sI - A + HC)^{-1} \begin{bmatrix} B & H \end{bmatrix} + \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix},$$

with $\mathbf{M}$ and $\tilde{\mathbf{M}}$ inner [3]. The next result shows that the maximal robust stability radius is equal to the smallest Hankel singular value of $\mathbf{G}(-s)$.

THEOREM 3.3 (see [6], [24]). *Let $\mathbf{G}^\sim \in \mathcal{RH}_\infty$, $\mathbf{G}(\infty) = 0$. Then the maximum stability radius r_1 for which $(\mathbf{G}, \mathbf{K})$ is $(r_1, 1)$ robustly stable for some $\mathbf{K} \in \mathcal{K}$ is given by $r_1 = \underline{\sigma}(\Gamma \mathbf{G}_{(-s)})$.*

Proof. From Theorem 3.1 $(\mathbf{G}, \mathbf{K})$ is $(r, 1)$ robustly stable if (i) $\mathbf{K}$ stabilizes $\mathbf{G}$ internally, and (ii) $\|\mathbf{K}(\mathbf{I} - \mathbf{G}\mathbf{K})^{-1}\|_\infty \leq r^{-1}$. Hence,

$$r_1^{-1} = \inf \{ \|\mathbf{K}(\mathbf{I} - \mathbf{G}\mathbf{K})^{-1}\|_\infty : \mathbf{K} \in \mathcal{K} \},$$

where $\mathcal{K}$ is the set of all internally stabilizing controllers of $\mathbf{G}$. From (4),

$$(5) \quad r_1^{-1} = \inf \{ \|(U + \mathbf{M}\mathbf{Q})\tilde{\mathbf{M}}\|_\infty : \mathbf{Q} \in \mathcal{H}_\infty \} = \inf \{ \|\mathbf{M}^\sim U + \mathbf{Q}\|_\infty : \mathbf{Q} \in \mathcal{H}_\infty \}$$

since $\mathbf{M}, \tilde{\mathbf{M}}$ are inner. A straightforward state-space calculation shows that

$$(6) \quad \mathbf{M}^\sim U(s) = -B'\Sigma^{-1}(s\mathbf{I} - A)^{-1}\Sigma^{-1}C' \in \mathcal{RH}_\infty^-$$

in previously defined notation. From Nehari's theorem the infimum in (5) is attained and is given by the Hankel norm of $\mathbf{M}^\sim U(-s)$. It is also straightforward to verify that the realization in (6) is balanced with grammians equal to $-\Sigma^{-1}$. Thus, the realization in (6) is minimal, and

$$r_1^{-2} = \|\Gamma \mathbf{M}^\sim U_{(-s)}\|^2 = \lambda_{\max}(\Sigma^{-2}) = \underline{\sigma}^{-2}(\Gamma \mathbf{G}_{(-s)})$$

from which it follows that $r_1 = \underline{\sigma}(\Gamma \mathbf{G}_{(-s)})$, as required. $\square$

Remark 3.2. Let $\mathbf{G}$ satisfy the assumptions of Theorem 3.3 and assume that the MacMillan degree of $\mathbf{G}$ is n . Let the Hankel singular values of $\mathbf{G}(-s)$ and $\mathbf{M}^\sim U(-s)$ be $\{\sigma_i(\Gamma \mathbf{G}_{(-s)})\}$ and $\{\sigma_i(\Gamma \mathbf{M}^\sim U_{(-s)})\}$, respectively, ordered in nonincreasing order of magnitude. Then $\sigma_n(\Gamma \mathbf{G}_{(-s)}) > 0$. Further, a slight adaptation of Theorem 3.3 shows that $\sigma_i(\Gamma \mathbf{G}_{(-s)}) = \sigma_{n-i+1}^{-1}(\Gamma \mathbf{M}^\sim U_{(-s)})$ for each $i = 1, 2, \dots, n$.

4. Main results. The set of all maximally robust controllers may be characterized in terms of the set of all optimal Nehari extensions of $\mathbf{M}^\sim U$, i.e., the set of all $\mathbf{Q} \in \mathcal{H}_\infty$ which achieve

$$(7) \quad \|\mathbf{M}^\sim U + \mathbf{Q}\|_\infty = r_1^{-1}.$$

This set can be parametrized as a linear fractional map of the set of all r_1 stable contractions [4], [5]. This parametrization is outlined next.

Remark 4.1. To avoid a messy indexing system we assume hereafter that the largest Hankel singular values of $\mathbf{R}(-s)$ and $\hat{\mathbf{R}}(-s)$ defined below in Theorems 4.1 and 4.2, respectively, are nonrepeated. These conditions are equivalent to the assumption that the first two superoptimal levels $s_1(\mathcal{T})$ and $s_2(\mathcal{T})$ are nonrepeated.

THEOREM 4.1 (see [14], [5], [7]). *Let $\mathbf{R} = \mathbf{M}^\sim U \in \mathcal{RH}_\infty^{-p \times m}$ and define $r_1 = \bar{\sigma}^{-1}(\Gamma \mathbf{R}_{(-s)})$ (see Theorem 3.3). Then there exists an embedding of $\mathbf{R}$ of the form*

$$\begin{aligned} \mathbf{H} &= \begin{bmatrix} \mathbf{H}_{11} & \mathbf{H}_{12} \\ \mathbf{H}_{21} & \mathbf{H}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{R} + \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{bmatrix} \\ &:= \begin{bmatrix} \mathbf{R} & 0 \\ 0 & 0 \end{bmatrix} + \mathbf{Q}_a \in \mathcal{RL}_\infty^{(p+m-1) \times (m+p-1)} \end{aligned}$$

with $\mathbf{Q}_a \in \mathcal{RH}_\infty$, such that $\mathbf{H}\mathbf{H}^\sim = \mathbf{H}^\sim\mathbf{H} = r_1^{-2}I_{p+m-1}$ and $\|\mathbf{H}_{22}\|_\infty = \|\mathbf{Q}_{22}\|_\infty < r_1^{-1}$. Further, the set of all (optimal) $\mathbf{Q} \in \mathcal{H}_\infty^{p \times m}$ such that $\|\mathbf{R} + \mathbf{Q}\|_\infty = r_1^{-1}$ is given by

$$(8) \quad \mathcal{S}_1 = \mathcal{F}_l(\mathbf{Q}_a, r_1 \mathcal{BH}_\infty^{(p-1) \times (m-1)}).$$

Let $\mathcal{K}_1$ denote the set of all maximally robust (r_1 -robust) controllers of $\mathbf{G}$, and let $\mathcal{T}_1 = \{\mathbf{K}(I - \mathbf{G}\mathbf{K})^{-1} : \mathbf{K} \in \mathcal{K}_1\} \subseteq \mathcal{T}$ denote the set of all optimal interpolating functions. In view of (3) and (4), together with Theorems 3.3 and 4.1, these sets may be parametrized as $\mathcal{K}_1 = \{(\mathbf{U} + \mathbf{M}\mathbf{Q})(\mathbf{V} + \mathbf{N}\mathbf{Q})^{-1} : \mathbf{Q} \in \mathcal{F}_l(\mathbf{Q}_a, r_1 \mathcal{BH}_\infty^{(p-1) \times (m-1)})\}$ and

$$(9) \quad \mathcal{T}_1 = \{(\mathbf{U} + \mathbf{M}\mathbf{Q})\tilde{\mathbf{M}} : \mathbf{Q} \in \mathcal{F}_l(\mathbf{Q}_a, r_1 \mathcal{BH}_\infty^{(p-1) \times (m-1)})\},$$

respectively. The next theorem gives an alternative parametrization of the set of all optimal solutions of (7), and therefore of $\mathcal{T}_1$ in (9). The result shows that $\mathcal{T}_1$ can be diagonalized by rational allpass transformations and is used extensively in this work.

THEOREM 4.2 (see [10]). *Let all variables be as defined in Theorem 4.1. Then, the following hold:*

1. *There exists an r_1^{-1} -allpass completion of $\mathbf{H}_{22} = \mathbf{Q}_{22}$ of the form*

$$\bar{\mathbf{H}} = \begin{bmatrix} \bar{\mathbf{H}}_{11} & \bar{\mathbf{H}}_{12} \\ \bar{\mathbf{H}}_{21} & \bar{\mathbf{H}}_{22} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{R}} + \bar{\mathbf{Q}}_{11} & \bar{\mathbf{Q}}_{12} \\ \bar{\mathbf{Q}}_{21} & \bar{\mathbf{Q}}_{22} \end{bmatrix} := \begin{bmatrix} \hat{\mathbf{R}} & 0 \\ 0 & 0 \end{bmatrix} + \bar{\mathbf{Q}}_a$$

with $\bar{\mathbf{Q}}_a \in \mathcal{RH}_\infty$ such that $\bar{\mathbf{H}}\bar{\mathbf{H}}^\sim = \bar{\mathbf{H}}^\sim\bar{\mathbf{H}} = r_1^{-2}I_{p+m-2}$, $\hat{\mathbf{R}} \in \mathcal{RH}_\infty^{-(p-1) \times (m-1)}$ and $\bar{\mathbf{Q}}_{12}^{-1}, \bar{\mathbf{Q}}_{21}^{-1} \in \mathcal{RH}_\infty$.

2. *The set of all $\bar{\mathbf{Q}} \in \mathcal{H}_\infty^{(p-1) \times (m-1)}$ such that $\|\hat{\mathbf{R}} + \bar{\mathbf{Q}}\|_\infty \leq r_1^{-1}$ is given by*

$$\bar{\mathcal{S}}_1 = \mathcal{F}_l(\bar{\mathbf{Q}}_a, r_1 \mathcal{BH}_\infty^{(p-1) \times (m-1)}).$$

3. *There exist inner matrices $\mathbf{X}$ and $\mathbf{Y}$ and a scalar allpass function $\mathbf{a}$ such that*

$$(10) \quad \mathcal{T}_1 = \mathbf{Y} \operatorname{diag}(r_1^{-1}\mathbf{a}, \mathcal{F}_l(\bar{\mathbf{H}}, r_1 \mathcal{BH}_\infty^{(p-1) \times (m-1)}))\mathbf{X}.$$

Further,

$$(11) \quad \mathcal{F}_l(\bar{\mathbf{H}}, r_1 \mathcal{BH}_\infty^{(p-1) \times (m-1)}) = \{\hat{\mathbf{R}} + \bar{\mathbf{Q}} : \bar{\mathbf{Q}} \in \mathcal{H}_\infty^{(p-1) \times (m-1)}, \|\hat{\mathbf{R}} + \bar{\mathbf{Q}}\|_\infty \leq r_1^{-1}\}.$$

Proof.

1. The construction of the r_1^{-1} -allpass completion $\bar{\mathbf{H}}$ is an exercise in standard factorization theory which can be performed using either transfer function or state-space techniques [3]. The details can be found in [14], [11], [12], [10].
2. The fact that r_1^{-1} is a suboptimal level of $\hat{\mathbf{R}}$ follows from part 1 since $\bar{\mathbf{Q}}_{11} \in \mathcal{RH}_\infty$ and $\|\hat{\mathbf{R}} + \bar{\mathbf{Q}}_{11}\|_\infty \leq \|\bar{\mathbf{H}}\|_\infty = r_1^{-1}$. Part 2 now follows from part 1 and [5, Theorem 3.2] since $\bar{\mathbf{Q}}_a, \bar{\mathbf{Q}}_{12}^{-1}, \bar{\mathbf{Q}}_{21}^{-1}, \hat{\mathbf{R}}^\sim \in \mathcal{RH}_\infty$ by construction.
3. Define $\mathbf{V}_\perp = \mathbf{H}_{12}\bar{\mathbf{H}}_{12}^{-1} \in \mathcal{RH}_\infty$ and $\mathbf{W}_\perp = \mathbf{H}_{21}^\sim\bar{\mathbf{H}}_{21}^\sim \in \mathcal{RH}_\infty$. Since $r_1\mathbf{H}$ and $r_1\bar{\mathbf{H}}$ are allpass, a manipulation will verify that $\mathbf{V}_\perp^\sim\mathbf{H}_{11}\mathbf{W}_\perp = \bar{\mathbf{H}}_{11}$ and

$$\bar{\mathbf{H}}_{21}\bar{\mathbf{H}}_{21}^\sim = r_1^{-2}I_{m-1} - \mathbf{H}_{22}\mathbf{H}_{22}^\sim = \mathbf{H}_{21}\mathbf{H}_{21}^\sim \implies \mathbf{W}_\perp^\sim\mathbf{W}_\perp = I_{m-1}.$$

Similarly,

$$\bar{H}_{12}^{\sim} \bar{H}_{12} = r_1^{-2} I_{p-1} - H_{22}^{\sim} H_{22} = H_{12}^{\sim} H_{12} \implies V_{\perp}^{\sim} V_{\perp} = I_{p-1}.$$

Hence there exist $\mathbf{v} \in \mathcal{RH}_{\infty}^{p \times 1}$, $\mathbf{w}^{\sim} \in \mathcal{RH}_{\infty}^{1 \times m}$ such that

$$(12) \quad \mathbf{V} = \begin{bmatrix} \mathbf{v} & \mathbf{V}_{\perp} \end{bmatrix}, \quad \mathbf{W} = \begin{bmatrix} \mathbf{w} & \mathbf{W}_{\perp} \end{bmatrix}$$

are square inner. Now consider the product

$$(13) \quad \begin{bmatrix} \mathbf{V}^{\sim} & 0 \\ 0 & I \end{bmatrix} \mathbf{H} \begin{bmatrix} \mathbf{W} & 0 \\ 0 & I \end{bmatrix} = \begin{bmatrix} \mathbf{v}^{\sim} \mathbf{H}_{11} \mathbf{w} & \mathbf{v}^{\sim} \mathbf{H}_{11} \mathbf{W}_{\perp} & \mathbf{v}^{\sim} \mathbf{H}_{12} \\ \mathbf{V}_{\perp}^{\sim} \mathbf{H}_{11} \mathbf{w} & \mathbf{V}_{\perp}^{\sim} \mathbf{H}_{11} \mathbf{W}_{\perp} & \mathbf{V}_{\perp}^{\sim} \mathbf{H}_{12} \\ \bar{\mathbf{H}}_{21} \mathbf{w} & \bar{\mathbf{H}}_{21} \mathbf{W}_{\perp} & \bar{\mathbf{H}}_{22} \end{bmatrix} \\ = \begin{bmatrix} \mathbf{v}^{\sim} \mathbf{H}_{11} \mathbf{w} & \mathbf{Y}_{12} & \mathbf{Y}_{13} \\ \mathbf{Y}_{21} & \bar{\mathbf{H}}_{11} & \bar{\mathbf{H}}_{12} \\ \mathbf{Y}_{31} & \bar{\mathbf{H}}_{21} & \bar{\mathbf{H}}_{22} \end{bmatrix}.$$

Since $\mathbf{V}, \mathbf{W}, r_1 \mathbf{H}$, and $r_1 \bar{\mathbf{H}}$ are allpass, all $\mathbf{Y}_{ij}$ terms in (13) are equal to zero, and $\mathbf{v}^{\sim} \mathbf{H}_{11} \mathbf{w}$ is r_1^{-1} -allpass, i.e., $\mathbf{v}^{\sim} \mathbf{H}_{11} \mathbf{w} = r_1^{-1} \mathbf{a}$ for some scalar allpass rational function $\mathbf{a}$. A simple manipulation using (13) verifies that

$$(14) \quad \mathcal{F}_l(\mathbf{H}, r_1 \mathcal{BH}_{\infty}^{(p-1) \times (m-1)}) = \mathbf{V} \operatorname{diag}(r_1^{-1} \mathbf{a}, \mathcal{F}_l(\bar{\mathbf{H}}, r_1 \mathcal{BH}_{\infty}^{(p-1) \times (m-1)})) \mathbf{W}^{\sim}.$$

Since

$$\mathcal{F}_l(\mathbf{H}, r_1 \mathcal{BH}_{\infty}^{(p-1) \times (m-1)}) = \{\mathbf{R} + \mathbf{Q} : \mathbf{Q} \in \mathcal{H}_{\infty}, \|\mathbf{R} + \mathbf{Q}\|_{\infty} = r_1^{-1}\}$$

we may write from (8), (9), and (14)

$$\begin{aligned} \mathcal{T}_1 &= (\mathbf{U} + \mathbf{M} \mathcal{F}_l(\mathbf{Q}_a, r_1 \mathcal{BH}_{\infty}^{(p-1) \times (m-1)})) \tilde{\mathbf{M}} \\ &= \mathbf{M}(\mathbf{M}^{\sim} \mathbf{U} + \mathcal{F}_l(\mathbf{Q}_a, r_1 \mathcal{BH}_{\infty}^{(p-1) \times (m-1)})) \tilde{\mathbf{M}} \\ &= \mathbf{M} \mathcal{F}_l(\mathbf{H}, r_1 \mathcal{BH}_{\infty}^{(p-1) \times (m-1)}) \tilde{\mathbf{M}} \\ &= \mathbf{Y} \operatorname{diag}(r_1^{-1} \mathbf{a}, \mathcal{F}_l(\bar{\mathbf{H}}, r_1 \mathcal{BH}_{\infty}^{(p-1) \times (m-1)})) \mathbf{X} \end{aligned}$$

as required, with $\mathbf{Y} := \mathbf{M} \mathbf{V}$ and $\mathbf{X} := \mathbf{W}^{\sim} \tilde{\mathbf{M}}$ square inner.

This proves the theorem. $\square$

Remark 4.2. The theorem shows that every optimal interpolating function $\mathbf{T} \in \mathcal{T}_1$ has a partial pseudosingular value decomposition with largest “singular value” r_1^{-1} and corresponding left and right “singular vectors” $\mathbf{M} \mathbf{v}$ and $\mathbf{w}^{\sim} \tilde{\mathbf{M}}$, respectively. The two “singular vectors” corresponding to the largest “singular value” are real-rational.

In what follows we develop improved robust stability properties for the set of controllers which minimize the pair $\{s_1(\mathcal{T}), s_2(\mathcal{T})\}$ with respect to lexicographic ordering. We denote the set of interpolating functions with this property by $\mathcal{T}_2$ and the corresponding set of controllers by $\mathcal{K}_2$. Clearly, $\mathcal{T}_2 \subseteq \mathcal{T}_1 \subseteq \mathcal{T}$ and $\mathcal{K}_2 \subseteq \mathcal{K}_1 \subseteq \mathcal{K}$. We refer to $\mathcal{T}_2$ ($\mathcal{K}_2$) as the superoptimal set of interpolating functions (controllers) with respect to the first two levels. The following lemma gives a parametrization of the set $\mathcal{T}_2$.

LEMMA 4.3. $\mathcal{T}_2$ may be parametrized as

$$(15) \quad \mathcal{T}_2 = \mathbf{Y}_1 \operatorname{diag}(s_1 \mathbf{a}, s_2 \mathbf{b}, \check{\mathbf{R}} + \mathcal{S}_2) \mathbf{X}_1$$

in which s_1 and s_2 denote the first two superoptimal levels of $\mathcal{T}$ with $s_1 = r_1^{-1}$, $\mathbf{Y}_1$ and $\mathbf{X}_1$ are square inner matrices, $\mathbf{a}$ and $\mathbf{b}$ are scalar rational allpass functions, $\hat{\mathbf{R}} \in \mathcal{RH}_\infty^{-(p-2) \times (m-2)}$, and $\mathcal{S}_2 = \{\hat{\mathbf{Q}} \in \mathcal{H}_\infty : \|\hat{\mathbf{R}} + \hat{\mathbf{Q}}\|_\infty \leq s_2\}$. Further, the first column (row) of $\mathbf{Y}_1$ ($\mathbf{X}_1$) is identical to the first column (row) of $\mathbf{Y}$ ($\mathbf{X}$) defined in Theorem 4.2.

Proof. From (10), since $\mathbf{X}$ and $\mathbf{Y}$ are square inner and $\mathbf{a}$ is allpass, we have

$$s_2 = \inf \{ \|\mathcal{F}_l(\bar{\mathbf{H}}, \mathbf{U})\|_\infty : \mathbf{U} \in r_1 \mathcal{BH}_\infty^{(p-1) \times (m-1)} \}.$$

Using (11) this is equivalent to

$$s_2 = \inf \{ \|\hat{\mathbf{R}} + \hat{\mathbf{Q}}\|_\infty : \hat{\mathbf{Q}} \in \mathcal{H}_\infty^{(p-1) \times (m-1)} \} = \|\Gamma_{\hat{\mathbf{R}}(-s)}\|,$$

where the second equality follows from Nehari's theorem. Clearly, $\mathcal{T}_2$ may be obtained by replacing $\mathcal{F}_l(\bar{\mathbf{H}}, r_1 \mathcal{BH}_\infty^{(p-1) \times (m-1)})$ by the set

$$(16) \quad \{\hat{\mathbf{R}} + \hat{\mathbf{Q}} : \hat{\mathbf{Q}} \in \mathcal{H}_\infty^{(p-1) \times (m-1)}, \|\hat{\mathbf{R}} + \hat{\mathbf{Q}}\|_\infty = s_2\},$$

which reveals the recursive character of the problem. Since $\hat{\mathbf{R}} \in \mathcal{RH}_\infty^{-(p-1) \times (m-1)}$, a parametrization of (16) may be obtained from Theorem 4.1 with $\mathbf{R}$ replaced by $\hat{\mathbf{R}}$ and s_1 replaced by s_2 . That is, Theorem 4.1 guarantees that there exists an s_2 -allpass embedding of $\hat{\mathbf{R}}$ of the form

$$\hat{\mathbf{H}} = \begin{bmatrix} \hat{\mathbf{H}}_{11} & \hat{\mathbf{H}}_{12} \\ \hat{\mathbf{H}}_{21} & \hat{\mathbf{H}}_{22} \end{bmatrix} := \begin{bmatrix} \hat{\mathbf{R}} + \hat{\mathbf{Q}}_{11} & \hat{\mathbf{Q}}_{12} \\ \hat{\mathbf{Q}}_{21} & \hat{\mathbf{Q}}_{22} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{R}} & 0 \\ 0 & 0 \end{bmatrix} + \hat{\mathbf{Q}}_a$$

in which $\hat{\mathbf{Q}}_{11} \in \mathcal{RH}_\infty^{(p-1) \times (m-1)}$, $\hat{\mathbf{Q}}_{12} \in \mathcal{RH}_\infty^{(p-1) \times (p-2)}$, $\hat{\mathbf{Q}}_{21} \in \mathcal{RH}_\infty^{(m-2) \times (m-1)}$, and $\hat{\mathbf{Q}}_{22} \in \mathcal{RH}_\infty^{(m-2) \times (p-2)}$ and such that $\hat{\mathbf{H}}\hat{\mathbf{H}}^\sim = \hat{\mathbf{H}}^\sim\hat{\mathbf{H}} = s_2^2 I_{p+m-3}$ and $\|\hat{\mathbf{H}}_{22}\|_\infty = \|\hat{\mathbf{Q}}_{22}\|_\infty < s_2$. Moreover, the set of all $\hat{\mathbf{Q}}$ in (16) is generated by $\mathcal{F}_l(\hat{\mathbf{Q}}_a, s_2^{-1} \mathcal{BH}_\infty^{(p-2) \times (m-2)})$. Next, we apply Theorem 4.2 to obtain an s_2 -allpass embedding of $\hat{\mathbf{H}}_{22} = \hat{\mathbf{Q}}_{22}$ of the form

$$\check{\mathbf{H}} = \begin{bmatrix} \check{\mathbf{H}}_{11} & \check{\mathbf{H}}_{12} \\ \check{\mathbf{H}}_{21} & \check{\mathbf{H}}_{22} \end{bmatrix} = \begin{bmatrix} \check{\mathbf{R}} + \check{\mathbf{Q}}_{11} & \check{\mathbf{Q}}_{12} \\ \check{\mathbf{Q}}_{21} & \check{\mathbf{Q}}_{22} \end{bmatrix} = \begin{bmatrix} \check{\mathbf{R}} & 0 \\ 0 & 0 \end{bmatrix} + \check{\mathbf{Q}}_a$$

with $\check{\mathbf{Q}}_{11} \in \mathcal{RH}_\infty^{(p-2) \times (m-2)}$, $\check{\mathbf{Q}}_{12} \in \mathcal{RH}_\infty^{(p-2) \times (p-2)}$, and $\check{\mathbf{Q}}_{21} \in \mathcal{RH}_\infty^{(m-2) \times (m-2)}$. Also, $\check{\mathbf{R}} \in \mathcal{RH}_\infty^{-(p-2) \times (m-2)}$, $\check{\mathbf{Q}}_{12}^{-1}, \check{\mathbf{Q}}_{21}^{-1} \in \mathcal{RH}_\infty$, and $\check{\mathbf{H}}\check{\mathbf{H}}^\sim = \check{\mathbf{H}}^\sim\check{\mathbf{H}} = s_2^2 I_{p+m-4}$. Further, Theorem 4.2 shows that the set of all $\check{\mathbf{Q}} \in \mathcal{H}_\infty$ such that $\|\check{\mathbf{R}} + \check{\mathbf{Q}}\|_\infty \leq s_2$ is given by $\mathcal{S}_2 = \mathcal{F}_l(\check{\mathbf{Q}}_a, s_2^{-1} \mathcal{BH}_\infty^{(p-2) \times (m-2)})$. The proof of Theorem 4.2(3) may now be repeated step by step to show that there exist square inner matrices $\mathbf{V}_1$ and $\mathbf{W}_1^\sim$ and a scalar rational allpass function $\mathbf{b}$ such that

$$\mathcal{F}_l(\hat{\mathbf{H}}, s_2^{-1} \mathcal{BH}_\infty^{(p-2) \times (m-2)}) = \mathbf{V}_1 \text{diag}(s_2 \mathbf{b}, \mathcal{F}_l(\check{\mathbf{H}}, s_2^{-1} \mathcal{BH}_\infty^{(p-2) \times (m-2)})) \mathbf{W}_1^\sim.$$

Hence from (10), $\mathcal{T}_2 = \mathbf{Y}_1 \text{diag}(s_1 \mathbf{a}, s_2 \mathbf{b}, \mathcal{F}_l(\check{\mathbf{H}}, s_2^{-1} \mathcal{BH}_\infty^{(p-2) \times (m-2)})) \mathbf{X}_1$, where we have defined the square inner matrix functions

$$\mathbf{Y}_1 = \mathbf{Y} \text{diag}(1, \mathbf{V}_1), \quad \mathbf{X}_1 = \text{diag}(1, \mathbf{W}_1^\sim) \mathbf{X}.$$

Equation (10) agrees with the parametrization in (15). Note further that the first column (row) of $\mathbf{Y}_1$ ($\mathbf{X}_1$) is identical with the first column (row) of $\mathbf{Y}$ ($\mathbf{X}$) and that for every $\hat{\mathbf{Q}} \in \mathcal{S}_2$, $\|\hat{\mathbf{R}} + \hat{\mathbf{Q}}\|_\infty \leq s_2$, as required. $\square$

In the next part of the section we identify the set of all $\Delta \in \partial\mathcal{D}_{r_1}(\mathbf{G})$ which destabilize $(\mathbf{G}, \mathbf{K})$ for *every* $\mathbf{K} \in \mathcal{K}_1$. We refer to such Δ 's as *uniformly destabilizing*.

LEMMA 4.4. *There exists $\Delta \in \partial\mathcal{D}_{r_1}(\mathbf{G})$ such that $(\mathbf{G} + \Delta, \mathbf{K}) \notin \mathcal{S}$ for every $\mathbf{K} \in \mathcal{K}_1$. Furthermore, Δ can be chosen to be a stable real-rational matrix function.*

Proof. Pick any $\omega_o \in \mathcal{R}$ and define

$$\Delta_o = \mathbf{X}^\sim(j\omega_o) \operatorname{diag}(r_1/\mathbf{a}(j\omega_o), 0) \mathbf{Y}^\sim(j\omega_o) \in \mathcal{C}^{m \times p}.$$

Then $\eta(\mathbf{G}) = \eta(\mathbf{G} + \Delta_o)$, $\|\Delta_o\|_\infty = r_1$ and so $\Delta_o \in \partial\mathcal{D}_{r_1}(\mathbf{G})$. Let $\mathbf{T} \in \mathcal{T}_1$ be any optimal interpolating function so that $\mathbf{T} = \mathbf{Y} \operatorname{diag}(r_1^{-1}\mathbf{a}, \Phi)\mathbf{X}$, where $\Phi = \mathcal{F}_l(\bar{\mathbf{H}}, r_1\Psi)$ for some $\Psi \in \mathcal{BH}_\infty^{(p-1) \times (m-1)}$ from Theorem 4.2. Then it is simple to show that $\det[I - \Delta_o\mathbf{T}(j\omega_o)] = 0$ for any Ψ and so $\det[I - \Delta_o\mathbf{T}(j\omega_o)] = 0$ for all $\mathbf{T} \in \mathcal{T}_1$. Since

$$\det\{I - (\mathbf{G} + \Delta_o)\mathbf{K}(j\omega_o)\} = \det\{I - \mathbf{G}\mathbf{K}(j\omega_o)\} \det\{I - \Delta_o\mathbf{T}(j\omega_o)\} = 0,$$

$(\mathbf{G} + \Delta_o, \mathbf{K}) \notin \mathcal{S}$ for every $\mathbf{K} \in \mathcal{K}_1$ by the generalized Nyquist theorem [25].

In the final part of the proof, we construct a stable real-rational Δ such that $\Delta \in \partial\mathcal{D}_{r_1}(\mathbf{G})$ and $\Delta(j\omega_o) = \Delta_o$. Define the unit vectors

$$y_1 = \mathbf{a}(-j\omega_o)\mathbf{v}'(j\omega_o)\mathbf{M}'(j\omega_o) \in \mathcal{C}^{1 \times p}, \quad x_1 = \tilde{\mathbf{M}}'(j\omega_o)\mathbf{w}(j\omega_o) \in \mathcal{C}^{m \times 1}$$

so that x_1 is the first column of $\mathbf{X}^\sim(j\omega_o)$ and $y_1/\mathbf{a}(-j\omega_o)$ is the first row of $\mathbf{Y}^\sim(j\omega_o)$. Next, express each component of x_1 and y_1 as

$$(x_1)_i = \hat{x}_i e^{j\phi_i}, \quad (y_1)_i = \hat{y}_i e^{j\theta_i},$$

where $\hat{y}_i, \hat{x}_i \in \mathcal{R}$ and $\phi_i, \theta_i \in [0, \pi)$. (This fixes the signs of $\hat{y}_i, \hat{x}_i$.) Define the inner vector functions

$$\hat{\mathbf{y}}(s) = \begin{bmatrix} \hat{y}_1 \frac{s-\alpha_1}{s+\alpha_1} \\ \vdots \\ \hat{y}_p \frac{s-\alpha_p}{s+\alpha_p} \end{bmatrix}, \quad \hat{\mathbf{x}}(s) = \begin{bmatrix} \hat{x}_1 \frac{s-\beta_1}{s+\beta_1} \\ \vdots \\ \hat{x}_m \frac{s-\beta_m}{s+\beta_m} \end{bmatrix}$$

so that

$$\arg \left\{ \frac{j\omega_o - \alpha_i}{j\omega_o + \alpha_i} \right\} = \theta_i, \quad i = 1, 2, \dots, p, \quad \arg \left\{ \frac{j\omega_o - \beta_i}{j\omega_o + \beta_i} \right\} = \phi_i, \quad i = 1, 2, \dots, m,$$

with $\alpha_i, \beta_i \in \mathcal{R}_+$ for all i . (If $\theta_i = 0$ or $\phi_i = 0$, we simply replace the i th entry of $\hat{\mathbf{y}}$ or $\hat{\mathbf{x}}$ by $\hat{x}_i$ or $\hat{y}_i$, respectively.) Next, define the $\mathcal{RH}_\infty$ function $\Delta = r_1 \hat{\mathbf{x}}\hat{\mathbf{y}}^T \in \mathcal{RH}_\infty^{m \times p}$ and note that $\Delta(j\omega_o) = \Delta_o = r_1 x_1 y_1^T$ by construction. Since $\Delta \in \mathcal{RH}_\infty$, $\eta(\mathbf{G}) = \eta(\mathbf{G} + \Delta)$. Finally, note that $\|\Delta\|_\infty = r_1$, which implies that $\Delta \in \partial\mathcal{D}_{r_1}(\mathbf{G})$. $\square$

Remark 4.3. The proof is an adaptation of a result in [25]. Indeed, it is not surprising that (real-rational) destabilizing perturbations exist on $\partial\mathcal{D}_{r_1}(\mathbf{G})$. The new information supplied by Lemma 4.4 is that (real-rational) boundary perturbations exist which are destabilizing for *every* maximally robust controller $\mathbf{K} \in \mathcal{K}_1$.

Denote by $\mathbf{x}^T$ and $\mathbf{y}$ the first row and column of $\mathbf{X}$ and $\mathbf{Y}$, respectively, defined in Theorem 4.2. Then, all uniformly destabilizing perturbations constructed in Lemma 4.4 have the property that $|\mathbf{x}^T \Delta \mathbf{y}(j\omega)| = r_1$ for some $\omega \in \mathcal{R}$. Moreover, such perturbations can be constructed for *every* $\omega \in \mathcal{R}$. The next result shows that

condition $|\mathbf{x}^T \Delta \mathbf{y}(j\omega)| = r_1$ is necessary for a $\Delta \in \partial \mathcal{D}_{r_1}(\mathbf{G})$ to be destabilizing for every $\mathbf{K} \in \mathcal{K}_1$.

LEMMA 4.5. *Let $\Delta \in \partial \mathcal{D}_{r_1}(\mathbf{G})$ be a destabilizing perturbation of $\mathbf{G}$ for every $\mathbf{K} \in \mathcal{K}_1$. Then there exists an $\omega \in \mathcal{R}$, such that*

$$(17) \quad |\mathbf{x}^T(j\omega) \Delta(j\omega) \mathbf{y}(j\omega)| = r_1.$$

Proof. The set $\mathcal{T}_1$ is given by (10). Pick any $\Psi \in r_1 \mathcal{BH}_\infty^{(p-1) \times (m-1)}$ so that $\|\Psi\|_\infty < r_1$. Since $\bar{\mathbf{H}}$ is r_1^{-1} -allpass, $\Phi = \mathcal{F}_l(\bar{\mathbf{H}}, \Psi)$ satisfies $\|\Phi\|_\infty < r_1^{-1}$. Since Δ destabilizes $\mathbf{G}$ for every $\mathbf{K} \in \mathcal{K}_1$, it is also destabilizing for

$$\mathbf{K} = (\mathbf{U} + \mathbf{M}\mathcal{F}_l(\mathbf{Q}_a, \Psi))(\mathbf{V} + \mathbf{N}\mathcal{F}_l(\mathbf{Q}_a, \Psi))^{-1} \in \mathcal{K}_1,$$

corresponding to the interpolation function $\mathbf{T} = \mathbf{Y} \operatorname{diag}(r_1^{-1} \mathbf{a}, \Phi) \mathbf{X}$. Since $\alpha \Delta$ is a stabilizing perturbation of $\mathbf{G}$ for every $\alpha \in [0, 1)$, there exists $\omega_o \in \mathcal{R}$ such that $\det \{I_m - \Delta(j\omega_o) \mathbf{T}(j\omega_o)\} = 0$ or equivalently that

$$\det \left\{ I_m - \Delta(j\omega_o) \mathbf{Y}(j\omega_o) \begin{bmatrix} r_1^{-1} \mathbf{a}(j\omega_o) & 0 \\ 0 & \Phi(j\omega_o) \end{bmatrix} \mathbf{X}(j\omega_o) \right\} = 0,$$

which implies that

$$(18) \quad \det \left\{ I_m - \mathbf{X}(j\omega_o) \Delta(j\omega_o) \mathbf{Y}(j\omega_o) \begin{bmatrix} r_1^{-1} \mathbf{a}(j\omega_o) & 0 \\ 0 & \Phi(j\omega_o) \end{bmatrix} \right\} = 0.$$

Define

$$\begin{bmatrix} x_1^T \\ X_\perp^T \end{bmatrix} = \mathbf{X}(j\omega_o),$$

$$\begin{bmatrix} y_1 & Y_\perp \end{bmatrix} = \mathbf{Y}(j\omega_o), \text{ and}$$

$$\tilde{\Delta} = \begin{bmatrix} \tilde{\delta}_{11} & \tilde{\Delta}_{12} \\ \tilde{\Delta}_{21} & \tilde{\Delta}_{22} \end{bmatrix} = \begin{bmatrix} x_1^T \\ X_\perp^T \end{bmatrix} \Delta(j\omega_o) \begin{bmatrix} y_1 & \mathbf{a}(j\omega_o) & Y_\perp \end{bmatrix}.$$

Then (18) may be written as

$$(19) \quad \det \begin{bmatrix} 1 - r_1^{-1} \tilde{\delta}_{11} & -\tilde{\Delta}_{12} \Phi(j\omega_o) \\ -r_1^{-1} \tilde{\Delta}_{21} & I_{m-1} - \tilde{\Delta}_{22} \Phi(j\omega_o) \end{bmatrix} = 0.$$

Next, we show that $\tilde{\delta}_{11} = r_1$. Suppose for contradiction that

$$(20) \quad 1 - r_1^{-1} \tilde{\delta}_{11} \neq 0.$$

Then (19) implies that

$$\begin{aligned} & (1 - r_1^{-1} \tilde{\delta}_{11}) \det \left\{ I_{m-1} - \tilde{\Delta}_{22} \Phi - r_1^{-1} \tilde{\Delta}_{21} \left(1 - r_1^{-1} \tilde{\delta}_{11} \right)^{-1} \tilde{\Delta}_{12} \Phi(j\omega_o) \right\} = 0, \\ \Rightarrow & \det \left\{ I_{m-1} - \left(\tilde{\Delta}_{22} + r_1^{-1} \tilde{\Delta}_{21} \left(1 - r_1^{-1} \tilde{\delta}_{11} \right)^{-1} \tilde{\Delta}_{12} \right) \Phi(j\omega_o) \right\} = 0, \\ \Rightarrow & \det \left\{ I_{m-1} - \mathcal{F}_u(\tilde{\Delta}, r_1^{-1}) \Phi(j\omega_o) \right\} = 0. \end{aligned}$$

Since $1 - r_1^{-1}\tilde{\delta}_{11} \neq 0$ by assumption, the upper linear fractional map is well-posed; moreover, $\bar{\sigma}(\tilde{\Delta}) = r_1$ which implies that $\bar{\sigma}(\mathcal{F}_u(\tilde{\Delta}, r_1^{-1})) \leq r_1$ [5]. Also, since $\bar{\sigma}(\Phi(j\omega_o)) < r_1^{-1}$, we have that $\bar{\sigma}(\mathcal{F}_u(\tilde{\Delta}, r_1^{-1})\Phi(j\omega_o)) < 1$. Hence,

$$\underline{\sigma}\left(I - \mathcal{F}_u(\tilde{\Delta}, r_1^{-1})\Phi(j\omega_o)\right) \geq 1 - \bar{\sigma}\left(\mathcal{F}_u(\tilde{\Delta}, r_1^{-1})\Phi(j\omega_o)\right) > 0,$$

and thus $\det\{I - \mathcal{F}_u(\tilde{\Delta}, r_1^{-1})\Phi(j\omega_o)\} \neq 0$, contradicting (20). Hence,

$$\tilde{\delta}_{11} = r_1 \Rightarrow |\mathbf{x}^T(j\omega_o)\mathbf{\Delta}(j\omega_o)\mathbf{y}(j\omega_o)| = r_1$$

since $|\mathbf{a}(j\omega_o)| = 1$. $\square$

Remark 4.4. Lemma 4.5 above shows that every $\mathbf{\Delta} \in \partial\mathcal{D}_{r_1}(\mathbf{G})$ which is destabilizing for all $\mathbf{K} \in \mathcal{K}_1$ satisfies $|\mathbf{x}^T(j\omega_o)\mathbf{\Delta}(j\omega_o)\mathbf{y}(j\omega_o)| = r_1$ for some $\omega_o \in \mathcal{R}$. Define the inner product of two matrices of compatible dimensions A and B as $\langle A, B \rangle = \text{trace}(A'B)$. Then, (17) says that every $\mathbf{\Delta} \in \partial\mathcal{D}_{r_1}(\mathbf{G})$ which is destabilizing for all $\mathbf{K} \in \mathcal{K}_1$ satisfies $|\langle \mathbf{y}(j\omega_o)\mathbf{x}^T(j\omega_o), \mathbf{\Delta}(j\omega_o) \rangle| = r_1$, i.e., that it has projection of magnitude r_1 in the “most critical direction” $\mathbf{y}(j\omega_o)\mathbf{x}^T(j\omega_o)$ for some $\omega_o \in \mathcal{R}$. Moreover, the proof of Lemma 4.4 shows that all frequencies $\omega \in \mathcal{R}$ are “equally critical,” in the sense that the generalized Nyquist criterion can be violated at *any* $\omega \in \mathcal{R}$. This implies that it is futile to attempt to extend the uncertainty set guaranteed to be stabilized by a subset of $\mathcal{K}_1$ in the (frequency-dependent) direction $\mathbf{y}(j\omega)\mathbf{x}^T(j\omega)$, $\omega \in \mathcal{R}$. Suppose now that we impose a “structure” on the perturbation set of the form

$$|\mathbf{x}^T(j\omega)\mathbf{\Delta}(j\omega)\mathbf{y}(j\omega)| \leq r_1(1 - \delta) \quad \text{for all } \omega \in \mathcal{R}$$

for some (fixed) $\delta \in [0, 1)$. Note in view of Lemmas 4.4 and 4.5 that this bound is assumed to be uniform in ω . In other words, we constrain the perturbation set so that $\mathbf{\Delta}$ cannot have a projection of magnitude larger than $r_1(1 - \delta)$ in the most critical direction for all $\omega \in \mathcal{R}$. Formally, define the set

$$(21) \quad \mathcal{E}(\delta, \mu) = \{\mathbf{\Delta} \in \mathcal{D}_\mu(\mathbf{G}) : \|\mathbf{x}^T\mathbf{\Delta}\mathbf{y}\|_\infty \leq r_1(1 - \delta)\},$$

where $\mathcal{D}_\mu(\mathbf{G})$ is defined in (1). Then, for each $\delta \in (0, 1]$ we want to find the set of controllers $\mathcal{K}_\delta \subseteq \mathcal{K}_1$ which maximize $\mu = \mu(\delta)$ under the constraint that $\mathbf{G} + \mathbf{\Delta}$ is stable for all $\mathbf{\Delta} \in \mathcal{D}_{r_1}(\mathbf{G}) \cup \mathcal{E}(\delta, \mu)$. Suppose that the maximum μ is attained and is given by $\mu^*(\delta)$. It is clear that $\mu^*(\delta)$ is a nondecreasing function of $\delta \in (0, 1]$. It is shown below that the sets $\mathcal{K}_\delta$ are identical for every $\delta \in (0, 1]$ and equal to $\mathcal{K}_2$. A closed-form expression of $\mu^*(\delta)$ is also obtained which involves the first two superoptimal levels of $\mathcal{T}$.

The problem formulation in the above remark is motivated by a related problem in [15]: Suppose that $A \in \mathcal{C}^{n \times n}$ is nonsingular. We know that if $\bar{\sigma}(E) = \underline{\sigma}(A)$, then $A - E$ is singular if and only if $\langle u_n v'_n, E \rangle = u'_n E v_n = \underline{\sigma}(A)$, where u_n and v_n denote the singular vectors of A corresponding to $\underline{\sigma}(A)$. Also, if $\bar{\sigma}(E) < \underline{\sigma}(A)$, then $A - E$ is nonsingular. Suppose that $\bar{\sigma}(E) = \underline{\sigma}(A)$ and E is constrained to have a projection of magnitude (strictly) less than $\underline{\sigma}(A)$ in the direction $u_n v'_n$. This means that $A - E$ cannot become singular, and therefore, $\bar{\sigma}(E)$ must increase for $A - E$ to lose rank. To find how much $\bar{\sigma}(E)$ can increase before singularity occurs, we formulate the problem

$$(22) \quad d(\phi) = \min \{ \|E\| : \det(A - E) = 0, |\langle u_n v'_n, E \rangle| \leq \phi \}$$

for $\phi < \underline{\sigma}(A) := \sigma_n(A)$. The solution to this problem is provided by the next theorem.

THEOREM 4.6. Let A be a square nonsingular complex matrix which has a singular value decomposition $A = U\Sigma V'$, where $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_{n-1}, \sigma_n)$ with $\sigma_1 \geq \dots \geq \sigma_{n-2} \geq \sigma_{n-1} > \sigma_n > 0$ and denote by u_n and v_n the last columns of U and V , respectively. Then all E which minimize (22) are given by

$$E = U \begin{bmatrix} P_s & 0 & 0 \\ 0 & -\phi & \nu \\ 0 & \nu' & \phi \end{bmatrix} V',$$

where P_s is arbitrary except for the constraint

$$(23) \quad \|P_s\| \leq \sqrt{\sigma_n \sigma_{n-1} + \phi(\sigma_n - \sigma_{n-1})}$$

and ν is given by

$$\nu = \sqrt{(\phi + \sigma_{n-1})(\sigma_n - \phi)} e^{j\theta}, \quad \theta \in [0, 2\pi).$$

The minimum value of $d(\phi)$ in (22) is given by the right-hand side (RHS) of (23).

Proof. See [15]. In the original statement of the theorem [15], all singular values of A are assumed to be distinct. This assumption can be relaxed to the condition $\sigma_n \geq \dots \geq \sigma_{n-2} \geq \sigma_{n-1} > \sigma_n > 0$ used here. $\square$

Remark 4.5. Theorem 4.6 says that, provided $|\langle u_n v'_n, E \rangle| \leq \phi < \sigma_n$, $\|E\|$ can increase from σ_n to $d(\phi) = \sqrt{\sigma_n \sigma_{n-1} + \phi(\sigma_n - \sigma_{n-1})}$ before $A - E$ becomes singular. In [15] this is exploited to derive robust-stability bounds for a class of additive, multiplicative, and inverse-multiplicative perturbations. Note that these results are a posteriori, i.e., they can be applied to assess the robust stability of a design only *after* a compensator has been designed. In our case, the results in [15] can be applied a priori in the sense that they can be used to characterize directly the subset of all maximally robust controllers which maximize the “radius” $\mu(\delta)$ of the uncertainty set $\mathcal{E}(\mu, \delta)$ defined in (21). The a priori character of these results in our case is a consequence of the alternative parametrization of the set of all optimal interpolation functions given in Theorem 3.2, which shows that there exists a (frequency-dependent) worst-case direction (defined by the vectors $\mathbf{y} = \mathbf{M}\mathbf{v}$ and $\mathbf{x}^T = \mathbf{w}^T \tilde{\mathbf{M}}$ in (12)) which is *identical* for *all* maximally robust controllers $\mathbf{K} \in \mathcal{K}_1$. The vectors $\mathbf{v}$ and $\mathbf{w}$ are associated with the maximal Schmidt pair of the Hankel operator $\Gamma_{\mathbf{R}(-s)}$ (see [16] for details).

In what follows, we use Theorem 4.6 to characterize the subset of all optimal controllers $\mathcal{K}_1$ which maximize $\mu^*(\delta)$. We first need a slightly different version of Theorem 4.6 which also allows us to treat the nonsquare and the singular cases.

THEOREM 4.7. Suppose that $T \in \mathcal{C}^{p \times m}$ has a singular value decomposition, $T = U \text{diag}(\Sigma, 0) V'$, with $\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_t)$, $\sigma_1 > \sigma_2 \geq \sigma_3 \geq \dots \geq \sigma_t > 0$. Let v and u be the first columns of V and U , respectively, and let $\phi < \sigma_1^{-1}$ be given. Define

$$\mathcal{B}_d^{m \times p} = \{E \in \mathcal{C}^{m \times p} : \|E\| < d\},$$

$$(24) \quad \mathcal{P}(\phi) = \{E \in \mathcal{C}^{m \times p} : |v' E u| \leq \phi\},$$

and

$$(25) \quad d(\phi) = \sup \{d : \det(I_m - ET) \neq 0 \text{ for all } E \in \mathcal{B}_d^{m \times p} \cap \mathcal{P}(\phi)\}.$$

Then

1. $d(\phi)$ is given by

$$(26) \quad d(\phi) = \sqrt{\frac{1}{\sigma_1 \sigma_2} - \phi \left(\frac{1}{\sigma_2} - \frac{1}{\sigma_1} \right)};$$

2. all $E \in \mathcal{P}(\phi)$ such that $\det(I_m - ET) = 0$ and $\|E\| = d(\phi)$ are given by

$$(27) \quad E = V \begin{bmatrix} \phi & \nu & 0 \\ \nu' & -\phi & 0 \\ 0 & 0 & P_s \end{bmatrix} U', \quad \nu = e^{j\theta} \sqrt{\left(\frac{1}{\sigma_2} + \phi \right) \left(\frac{1}{\sigma_1} - \phi \right)},$$

where $\theta \in [0, 2\pi)$ and P_s is arbitrary except from the constraint $\|P_s\| \leq d(\phi)$.

Proof. Introduce the partitions $U = [U_1 \ U_2]$ and $V = [V_1 \ V_2]$, where $U_1 \in \mathcal{C}^{p \times t}$ and $V_1' \in \mathcal{C}^{t \times m}$. For $E \in \mathcal{C}^{m \times p}$,

$$\det(I_m - ET) \neq 0 \Leftrightarrow \det(I_m - V'EU \operatorname{diag}(\Sigma, 0)) \neq 0$$

which is also equivalent to

$$\det \begin{bmatrix} I_t - V_1'EU_1\Sigma & 0 \\ -V_2'EU_1\Sigma & I_{m-t} \end{bmatrix} \neq 0 \Leftrightarrow \det(\Sigma^{-1} - V_1'EU_1) \neq 0.$$

Let $r = \sigma_1^{-1}$. The transformation

$$(28) \quad \mathcal{C}^{m \times p} \rightarrow \mathcal{C}^{t \times t} : E \rightarrow \tilde{E} = V_1'EU_1$$

maps $\mathcal{B}_r^{m \times p}$ onto $V_1'\mathcal{B}_r^{m \times p}U_1 = \mathcal{B}_r^{t \times t}$. Clearly, for any $E \in \mathcal{B}_r^{m \times p}$, $\|\tilde{E}\| \leq \|E\|$ and hence $\tilde{E} \in \mathcal{B}_r^{t \times t}$. Conversely, since all solutions to the equation $\tilde{E} = V_1'EU_1$ are given by

$$(29) \quad E = \begin{bmatrix} V_1 & V_2 \end{bmatrix} \begin{bmatrix} \tilde{E} & \tilde{E}_{12} \\ \tilde{E}_{21} & \tilde{E}_{22} \end{bmatrix} \begin{bmatrix} U_1' \\ U_2' \end{bmatrix},$$

where $\tilde{E}_{21} \in \mathcal{C}^{(m-t) \times t}$, $\tilde{E}_{12} \in \mathcal{C}^{t \times (p-t)}$, and $\tilde{E}_{22} \in \mathcal{C}^{(m-t) \times (p-t)}$ are arbitrary matrices of the specified dimensions, every $\tilde{E} \in \mathcal{B}_r^{t \times t}$ is the image of the set

$$\left\{ E = \begin{bmatrix} V_1 & V_2 \end{bmatrix} \begin{bmatrix} \tilde{E} & \tilde{E}_{12} \\ \tilde{E}_{21} & \tilde{E}_{22} \end{bmatrix} \begin{bmatrix} U_1' \\ U_2' \end{bmatrix} : \left\| \begin{bmatrix} \tilde{E} & \tilde{E}_{12} \\ \tilde{E}_{21} & \tilde{E}_{22} \end{bmatrix} \right\| < d \right\} \subseteq \mathcal{B}_r^{m \times p}$$

under (28). Moreover, since $|v_1'Eu_1| \leq \phi \Leftrightarrow |\tilde{E}_{11}| \leq \phi$, (25) is equivalent to

$$d(\phi) = \sup \{d : \det(\Sigma^{-1} - \tilde{E}) \neq 0 \text{ for all } \tilde{E} \in \{\tilde{E} \in \mathcal{B}_r^{t \times t} : |\tilde{E}_{11}| < \phi\}\},$$

where $\tilde{E}_{11}$ denotes the (1,1) element of $\tilde{E}$. By introducing suitable permutations, $d(\phi)$ may be obtained by applying Theorem 4.6 and is given by the RHS of (26) which proves part 1. The set of all $\tilde{E}$ with $\|\tilde{E}\| = d(\phi)$ such that $\det(\Sigma^{-1} - \tilde{E}) = 0$ under the constraint $|\tilde{E}_{11}| \leq \phi < \sigma_1^{-1}$ is given via a slight adaptation of Theorem 4.6 as

$$(30) \quad \tilde{E} = \begin{bmatrix} \phi & \nu & 0 \\ \nu' & -\phi & 0 \\ 0 & 0 & \tilde{P}_s \end{bmatrix},$$

in which ν is defined in (27) and $\tilde{P}_s \in \mathcal{C}^{(t-2) \times (t-2)}$ is arbitrary except from the constraint $\|\tilde{P}_s\| \leq d(\phi)$. It then follows from (29) and the properties of (28) that all $E \in \mathcal{P}(\phi, d)$ such that $\det(I_m - ET) = 0$ and $\|E\| = d(\phi)$ are of the form given in (29) with $\tilde{E}$ given by (30) subject to the constraint $\|E\| \leq d(\phi)$. Thus all such E 's are of the form

$$(31) \quad E = V \begin{bmatrix} \phi & \nu & 0 & E_{14} \\ \nu' & -\phi & 0 & E_{24} \\ 0 & 0 & \tilde{P}_s & E_{34} \\ E_{41} & E_{42} & E_{43} & E_{44} \end{bmatrix} U'$$

subject to the constraint $\|E\| \leq d(\phi)$. Since

$$\begin{bmatrix} \phi & \nu \\ \nu' & -\phi \end{bmatrix} \begin{bmatrix} \phi & \nu \\ \nu' & -\phi \end{bmatrix} = d(\phi)^2 I_2$$

we have that $E_{14} = 0, E_{24} = 0, E_{41} = 0, E_{42} = 0$, and

$$\left\| \begin{bmatrix} \tilde{P}_s & E_{34} \\ E_{43} & E_{44} \end{bmatrix} \right\| \leq d(\phi).$$

Thus (31) agrees with the parametrization of part 2. $\square$

Remark 4.6. Note that $d(\phi)$ depends only on the two largest singular values of T , σ_1 and σ_2 , and on ϕ (and hence on u and v , the left and right singular vectors corresponding to σ_1). Note also that $d(\phi)$ is a decreasing function of σ_2 . Since all optimal interpolating functions have the same largest singular value s_1 (for all frequencies), and furthermore, share the same left and right singular vectors corresponding to s_1 , Theorem 4.7 suggests a link between the maximization of $\mu^*(\delta)$ and the minimization of the second largest singular value of the elements of $\mathcal{T}_1$.

The next theorem, which is our main result, shows that $\mu^*(\delta)$ is maximized uniquely by the set of all superoptimal controllers with respect to the first two levels.

THEOREM 4.8. *Let $\mathcal{T}_1 \subseteq \mathcal{H}_\infty^{p \times m}$ be as defined in (10). Let $\mathbf{x}^T$ and $\mathbf{y}$ be the first row and column of $\mathbf{X}$ and $\mathbf{Y}$, respectively, and define $\mathcal{D}_r(\mathbf{G})$ and $\mathcal{E}(\delta, \mu)$ as in (1) and (21), respectively, for some (fixed) $\delta \in [0, 1]$. Let $\mu^*(\delta)$ be the supremum of μ such that there exists a $\mathbf{K}$ for which $(\mathbf{G} + \Delta, \mathbf{K}) \in \mathcal{S}$ for every $\Delta \in \mathcal{D}_{r_1}(\mathbf{G}) \cup \mathcal{E}(\delta, \mu)$. Then the following hold:*

1. *For each δ ,*

$$\mu^*(\delta) = \sqrt{\frac{1}{s_1} \left(\frac{\delta}{s_2} + \frac{1-\delta}{s_1} \right)} \geq r_1,$$

where s_1 and s_2 are the two superoptimal levels of $\mathcal{T}$ with $s_1 = r_1^{-1}$.

2. *For each $0 < \delta \leq 1$ the following two statements are equivalent:*
 - (a) *$(\mathbf{G} + \Delta, \mathbf{K}) \in \mathcal{S}$ for every $\Delta \in \mathcal{D}_{r_1}(\mathbf{G}) \cup \mathcal{E}(\delta, \mu^*(\delta))$,*
 - (b) *$\mathbf{K} \in \mathcal{K}_2$.*
3. (a) *$\mathcal{E}(0, \mu^*(0)) = \mathcal{D}_{r_1}(\mathbf{G})$,*
 (b) *for each $\mathbf{K} \in \mathcal{K}_2$, $(\mathbf{G} + \Delta, \mathbf{K}) \in \mathcal{S}$ for every $\Delta \in \bigcup_{\delta \in [0, 1]} \mathcal{E}(\delta, \mu^*(\delta))$.*
4. *Let σ_n and σ_{n-1} denote the two smallest Hankel singular values of $\mathbf{G}(-s)$ with $\sigma_{n-1} > \sigma_n$. Then,*

$$\mu^*(\delta) \geq \sqrt{\delta \sigma_n \sigma_{n-1} + (1-\delta) \sigma_n^2}.$$

Proof. Let $\mathbf{K} \in \mathcal{K}_2$ and define $\mathbf{T} = \mathbf{K}(\mathbf{I} - \mathbf{G}\mathbf{K})^{-1} \in \mathcal{H}_\infty$. Fix $\delta \in (0, 1]$ and define

$$\mu_1^*(\delta, \mathbf{K}) = \sup \{ \mu : (\mathbf{G} + \mathbf{\Delta}, \mathbf{K}) \in \mathcal{S} \text{ for all } \mathbf{\Delta} \in \mathcal{D}_{r_1}(\mathbf{G}) \cup \mathcal{E}(\delta, \mu) \}$$

and $\mu_1^*(\delta) = \sup \{ \mu_1^*(\delta, \mathbf{K}) : \mathbf{K} \in \mathcal{K}_2 \}$. Clearly $\mu_1^*(\delta, \mathbf{K}) \leq \mu_1^*(\delta) \leq \mu^*(\delta)$. We show that

$$(32) \quad r_1 \geq \mu_1^*(\delta, \mathbf{K}) \geq \sqrt{\frac{1}{s_1} \left(\frac{\delta}{s_2} + \frac{1-\delta}{s_1} \right)}.$$

Since the largest Hankel singular value of $\mathbf{R}(-s)$ is assumed to be simple, we have from [16] that $s_2 \leq \sigma_2(\Gamma_{\mathbf{R}(-s)}) < \sigma_1(\Gamma_{\mathbf{R}(-s)}) = s_1$. Set μ_1 equal to the RHS of (32) and suppose for contradiction that there exists a $\mathbf{\Delta} \in \mathcal{D}_{r_1}(\mathbf{G}) \cup \mathcal{E}(\delta, \mu_1)$ such that $(\mathbf{G} + \mathbf{\Delta}, \mathbf{K}) \notin \mathcal{S}$. Clearly, if $\mathbf{\Delta} \in \mathcal{D}_{r_1}(\mathbf{G})$ it cannot be destabilizing and hence $\mathbf{\Delta} \in \mathcal{E}(\delta, \mu_1) \setminus \mathcal{D}_{r_1}(\mathbf{G})$. Thus $r_1 \leq \|\mathbf{\Delta}\|_\infty < \mu_1$ and $\|\mathbf{x}^T \mathbf{\Delta} \mathbf{y}\|_\infty \leq r_1(1-\delta)$. Since $\mathbf{K}$ stabilizes $\mathbf{G}$, it follows from the generalized Nyquist theorem that

$$(33) \quad \det(\mathbf{I} - \mathbf{G}\mathbf{K}(j\omega)) \neq 0 \text{ for all } \omega \in \mathcal{R}.$$

Now, let ξ vary continuously in the interval $[0, 1]$ and consider the resulting deformation of the Nyquist plot of $\det(\mathbf{I} - (\mathbf{G} + \xi \mathbf{\Delta})\mathbf{K}(j\omega))$. Since $\mathbf{\Delta}$ is destabilizing, there exist an $\omega_o \in \mathcal{R}$ and a $\xi_o \in (0, 1]$ such that

$$\det(\mathbf{I} - (\mathbf{G}(j\omega_o) + \xi_o \mathbf{\Delta}(j\omega_o))\mathbf{K}(j\omega_o)) = 0$$

which implies that

$$\det(\mathbf{I} - \mathbf{G}(j\omega_o)\mathbf{K}(j\omega_o)) \det(\mathbf{I} - \xi_o \mathbf{\Delta}(j\omega_o)\mathbf{T}(j\omega_o)) = 0$$

or equivalently that

$$(34) \quad \det(\mathbf{I} - \xi_o \mathbf{\Delta}(j\omega_o)\mathbf{T}(j\omega_o)) = 0$$

from (33). Now $\|\mathbf{x}^T \mathbf{\Delta} \mathbf{y}\|_\infty \leq r_1(1-\delta)$ implies that

$$(35) \quad \xi_o |\mathbf{x}^T(j\omega_o) \mathbf{\Delta}(j\omega_o) \mathbf{y}(j\omega_o)| \leq \frac{1-\delta}{s_1} := \phi$$

since $0 < \xi_o \leq 1$. Since the two largest singular values of $\mathbf{T}(j\omega_o)$ are s_1 and s_2 , respectively, Theorem 4.7 guarantees that $\det(\mathbf{I} - \xi_o E \mathbf{T}(j\omega_o)) \neq 0$ for all E such that

$$(36) \quad \|E\| < \frac{1}{\xi_o} \sqrt{\frac{1}{s_1 s_2} - \phi \left(\frac{1}{s_2} - \frac{1}{s_1} \right)} = \frac{1}{\xi_o} \sqrt{\frac{1}{s_1} \left(\frac{\delta}{s_2} + \frac{1-\delta}{s_2} \right)} = \frac{\mu_1}{\xi_o}$$

provided that

$$(37) \quad \xi_o |\mathbf{x}^T(j\omega_o) E \mathbf{y}(j\omega_o)| \leq \frac{1-\delta}{s_1} = \phi.$$

Thus, Theorem 4.7 also guarantees the nonsingularity of $\mathbf{I} - \xi_o \mathbf{\Delta}(j\omega_o)\mathbf{T}(j\omega_o)$ from (35), (36), and the fact that $\|\mathbf{\Delta}(j\omega_o)\| \leq \|\mathbf{\Delta}\|_\infty < \mu_1 \leq \mu_1 \xi_o^{-1}$. This contradicts (34) and hence shows that

$$(38) \quad \mu^*(\delta) \geq \mu_1^*(\delta) \geq \mu_1^*(\delta, \mathbf{K}) \geq \sqrt{\frac{1}{s_1} \left(\frac{\delta}{s_2} + \frac{1-\delta}{s_1} \right)}.$$

Next, it is shown that the second and third inequalities in (38) are in fact equalities. To establish this fact, it suffices to construct a $\Delta \in \mathcal{RH}_\infty$ ($\Rightarrow \eta(\mathbf{G}) = \eta(\mathbf{G} + \Delta)$) such that (i) $\|\Delta\|_\infty$ is equal to the RHS of (38), (ii) $\|\mathbf{x}^T \Delta \mathbf{y}\|_\infty \leq r_1(1 - \delta)$, and (iii) $(\mathbf{G} + \Delta, \mathbf{K})$ is unstable for every $\mathbf{K} \in \mathcal{K}_2$.

Using Lemma 4.3 the interpolating function $\mathbf{T} \in \mathcal{T}_2$ corresponding to any $\mathbf{K} \in \mathcal{K}_2$ can be written in the form

$$\mathbf{T} = \mathbf{Y}_1 \text{diag}(s_1 \mathbf{a}, s_2 \mathbf{b}, \check{\mathbf{R}} + \check{\mathbf{Q}}) \mathbf{X}_1,$$

in which $\mathbf{Y}_1$ and $\mathbf{X}_1$ are square inner matrices, $\mathbf{a}$ and $\mathbf{b}$ are scalar allpass functions, and $\|\check{\mathbf{R}} + \check{\mathbf{Q}}\|_\infty \leq s_2$. In addition, also from Lemma 4.3, the first column (row) of $\mathbf{Y}_1$ ($\mathbf{X}_1$) is identical to the first column (row) of $\mathbf{Y}$ ($\mathbf{X}$); these are denoted by $\mathbf{y}$ and $\mathbf{x}^T$, respectively. Define the allpass matrix function $\mathbf{Y}_2 = \mathbf{Y}_1 \text{diag}(\mathbf{a}, \mathbf{b}, I_{p-2})$. Then $\mathbf{T} = \mathbf{Y}_2 \text{diag}(s_1, s_2, \check{\mathbf{R}} + \check{\mathbf{Q}}) \mathbf{X}_1$. Factor $\mathbf{X}_1^\sim$ and $\mathbf{Y}_2^\sim$ as $\mathbf{X}_1^\sim = \mathbf{N}_1 \text{diag}(\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_m)$ and $\mathbf{Y}_2^\sim = \text{diag}(\tilde{\mathbf{d}}_1, \tilde{\mathbf{d}}_2, \dots, \tilde{\mathbf{d}}_p) \mathbf{N}_2$ where the functions $\mathbf{N}_1, \mathbf{N}_2, \mathbf{d}_i^{-1}, \tilde{\mathbf{d}}_i^{-1}$ are in $\mathcal{RH}_\infty$, $\mathbf{N}_1, \mathbf{N}_2$ are square inner, and the $\mathbf{d}_i$'s, $\tilde{\mathbf{d}}_i$'s are scalar allpass functions; these are left and right coprime factorizations of the columns of $\mathbf{X}_1^\sim$ and the rows of $\mathbf{Y}_2^\sim$ with inner denominators. Next pick any $\omega_o \in \mathcal{R}$, and write for each $i = 1, 2$, $\mathbf{d}_i(j\omega_o) = \exp(j\phi_i)$, $\tilde{\mathbf{d}}_i(j\omega_o) = \exp(j\tilde{\phi}_i)$, where $-\pi \leq \phi_i, \tilde{\phi}_i < \pi$. Define two diagonal inner matrices $\mathbf{A}_1 = \text{diag}(\alpha_1, \alpha_2)$ and $\mathbf{A}_2 = \text{diag}(\tilde{\alpha}_1, \tilde{\alpha}_2)$ as follows: For each $i \in \{1, 2\}$, if $0 < \phi_i < \pi$ ($0 < \tilde{\phi}_i < \pi$), set $\alpha_i(s) = (s - \beta_i)(s + \beta_i)^{-1}$ ($\tilde{\alpha}_i(s) = (s - \tilde{\beta}_i)(s + \tilde{\beta}_i)^{-1}$), where $\arg(j\omega_o - \beta_i)(j\omega_o + \beta_i)^{-1} = \phi_i > 0$ ($\arg(j\omega_o - \tilde{\beta}_i)(j\omega_o + \tilde{\beta}_i)^{-1} = \tilde{\phi}_i > 0$). In the case that $-\pi < \phi_i < 0$, set $\alpha_i(s) = -(s - \beta_i)(s + \beta_i)^{-1}$ ($\tilde{\alpha}_i(s) = -(s - \tilde{\beta}_i)(s + \tilde{\beta}_i)^{-1}$), where $\arg(j\omega_o - \beta_i)(j\omega_o + \beta_i)^{-1} = \pi + \phi_i > 0$ ($\arg(j\omega_o - \tilde{\beta}_i)(j\omega_o + \tilde{\beta}_i)^{-1} = \pi + \tilde{\phi}_i > 0$). Finally, if $\phi_i = 0$ ($\tilde{\phi}_i = 0$) or $\phi_i = -\pi$ ($\tilde{\phi}_i = -\pi$), set α_i ($\tilde{\alpha}_i$) to 1 or -1 , respectively.

Next, let $\mathbf{N}_{11}$ ($\mathbf{N}_{21}$) denote the matrix consisting of the first two columns (rows) of $\mathbf{N}_1$ ($\mathbf{N}_2$), and define $\Delta \in \mathcal{RH}_\infty$ as

$$\Delta = \mathbf{N}_{11} \mathbf{A}_1 \begin{bmatrix} \phi & \nu_o \\ \nu_o & -\phi \end{bmatrix} \mathbf{A}_2 \mathbf{N}_{21},$$

where ϕ is defined in (35) and

$$(39) \quad \nu_o = \sqrt{\left(\frac{1}{s_2} + \phi\right) \left(\frac{1}{s_1} - \phi\right)} = \sqrt{\frac{\delta}{s_1} \left(\frac{1}{s_2} + \frac{1 - \delta}{s_1}\right)},$$

where the second equality in (39) follows by using the definition of ϕ in (35). Since $\mathbf{N}_{11}$, $\mathbf{N}_{21}^T$, $\mathbf{A}_1$, and $\mathbf{A}_2$ are inner matrices,

$$\|\Delta\|_\infty = \left\| \begin{bmatrix} \phi & \nu_o \\ \nu_o & -\phi \end{bmatrix} \right\| = \sqrt{\phi^2 + \nu_o^2},$$

which is equal to the RHS of (38) after some simple algebra.

Since, $\mathbf{X}_1 \mathbf{X}_1^\sim = I_m$ and $\mathbf{X}$ and $\mathbf{X}_1$ have the same first row ($\mathbf{x}^T$), we have

$$\mathbf{x}^T \mathbf{X}_1^\sim = [1 \ 0 \ \dots \ 0] \Rightarrow \mathbf{x}^T \mathbf{N}_1 \text{diag}(\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_m) = [1 \ 0 \ \dots \ 0],$$

and hence

$$(40) \quad \mathbf{x}^T \mathbf{N}_{11} = [\mathbf{d}_1^{-1} \ 0].$$

Similarly, since $\mathbf{Y}_1^\sim \mathbf{Y}_1 = I_p$ and matrices $\mathbf{Y}$ and $\mathbf{Y}_1$ share their first column ($\mathbf{y}$),

$$\mathbf{Y}_1^\sim \mathbf{y} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} \mathbf{a} & 0 & 0 \\ 0 & \mathbf{b} & 0 \\ 0 & 0 & I_{p-2} \end{bmatrix} \mathbf{Y}_2^\sim \mathbf{y} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

or, equivalently, that

$$(41) \quad \mathbf{N}_2 \mathbf{y} = \begin{bmatrix} \tilde{\mathbf{d}}_1^{-1} \mathbf{a}^{-1} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \Rightarrow \mathbf{N}_{21} \mathbf{y} = \begin{bmatrix} \tilde{\mathbf{d}}_1^{-1} \mathbf{a}^{-1} \\ 0 \end{bmatrix}.$$

Using (40) and (41) we conclude that

$$\|\mathbf{x}^T \boldsymbol{\Delta} \mathbf{y}\|_\infty = \left\| \mathbf{x}^T \mathbf{N}_{11} \mathbf{A}_1 \begin{bmatrix} \phi & \nu_o \\ \nu_o & -\phi \end{bmatrix} \mathbf{A}_2 \mathbf{N}_{21} \mathbf{y} \right\|_\infty = \|\phi \mathbf{a}^{-1} \tilde{\mathbf{d}}_1^{-1} \tilde{\mathbf{d}}_1^{-1} \boldsymbol{\alpha}_1 \tilde{\boldsymbol{\alpha}}_1\|_\infty = \phi,$$

using (37) and the fact that $\mathbf{a}, \mathbf{d}_1, \mathbf{d}_3, \boldsymbol{\alpha}_1$, and $\tilde{\boldsymbol{\alpha}}_1$ are all scalar allpass.

Using the identity $\det(I - AB) = \det(I - BA)$, $\det(I_m - \boldsymbol{\Delta}(j\omega_o) \mathbf{T}(j\omega_o))$ can be written as

$$\det \left\{ I - \mathbf{X}_1(j\omega_o) \mathbf{N}_{11}(j\omega_o) \mathbf{A}_1(j\omega_o) \begin{bmatrix} \phi & \nu_o \\ \nu_o & -\phi \end{bmatrix} \mathbf{A}_2(j\omega_o) \mathbf{N}_{21}(j\omega_o) \mathbf{Y}_2(j\omega_o) \Pi \right\},$$

where we have defined $\Pi = \text{diag}(s_1, s_2, \check{\mathbf{R}}(j\omega_o) + \check{\mathbf{Q}}(j\omega_o))$. It is now easy to verify from the construction of $\mathbf{A}_1$ and $\mathbf{A}_2$ above that

$$\mathbf{X}_1(j\omega_o) \mathbf{N}_{11}(j\omega_o) \mathbf{A}_1(j\omega_o) = \begin{bmatrix} I_2 \\ 0_{m-2,2} \end{bmatrix}, \quad \mathbf{A}_2(j\omega_o) \mathbf{N}_{21}(j\omega_o) \mathbf{Y}_2(j\omega_o) = [I_2 \ 0_{2,p-2}],$$

and hence

$$\det(I_m - \boldsymbol{\Delta}(j\omega_o) \mathbf{T}(j\omega_o)) = \det \begin{bmatrix} 1 - \phi s_1 & -\nu_o s_2 & 0 \\ -\nu_o s_1 & 1 + \phi s_2 & 0 \\ 0 & 0 & I_{m-2} \end{bmatrix}$$

or

$$\det(I_m - \boldsymbol{\Delta}(j\omega_o) \mathbf{T}(j\omega_o)) = (1 - \phi s_1)(1 + \phi s_2) - \nu_o^2 s_1 s_2 = 0,$$

after some simple algebra using (39). This implies that

$$\det(I_m - (\mathbf{G}(j\omega_o) + \boldsymbol{\Delta}(j\omega_o)) \mathbf{K}(j\omega_o)) = 0,$$

and hence $\boldsymbol{\Delta}$ is destabilizing from the generalized Nyquist theorem [25]. This shows that the third inequality in (38) is indeed an equality as claimed. Since $\boldsymbol{\Delta}$ is destabilizing for every $\mathbf{K} \in \mathcal{K}_2$, the second equality in (38) also follows.

To establish that the first inequality in (38) is an equality it suffices to construct for each $\mathbf{K} \in \mathcal{K}_1 \setminus \mathcal{K}_2$ a $\boldsymbol{\Delta} \in \mathcal{RH}_\infty^{m \times p}$ such that (i) $\|\boldsymbol{\Delta}\|_\infty$ is (strictly) less than the RHS of (38), (ii) $\|\mathbf{x}^T \boldsymbol{\Delta} \mathbf{y}\|_\infty \leq r_1(1 - \delta)$, and (iii) $(\mathbf{G} + \boldsymbol{\Delta}, \mathbf{K}) \notin \mathcal{S}$. Take any $\mathbf{K} \in \mathcal{K}_1$, $\mathbf{K} \notin \mathcal{K}_2$ and let $\mathbf{T} = \mathbf{K}(I - \mathbf{G}\mathbf{K})^{-1}$. From Theorem 4.2, $\mathbf{T} \in \mathcal{T}_1$ has the form

$\mathbf{T} = \mathbf{Y} \operatorname{diag}(s_1 \mathbf{a}, \hat{\mathbf{R}} + \bar{\mathbf{Q}}) \mathbf{X}$, where $\|\hat{\mathbf{R}} + \bar{\mathbf{Q}}\|_\infty \leq s_1$. Since $\mathbf{T} \notin \mathcal{T}_2$, there exists an $\omega_o \in \mathcal{R}_+$ such that

$$s_2 < \bar{\sigma} \left(\hat{\mathbf{R}}(j\omega_o) + \bar{\mathbf{Q}}(j\omega_o) \right) \leq s_1.$$

Let $\hat{\mathbf{R}}(j\omega_o) + \bar{\mathbf{Q}}(j\omega_o)$ have a singular value decomposition

$$\hat{\mathbf{R}}(j\omega_o) + \bar{\mathbf{Q}}(j\omega_o) = U \operatorname{diag}(\Sigma, 0_{p-t-1, m-t-1}) V',$$

where $\Sigma = \operatorname{diag}(\sigma_1, \sigma_2, \dots, \sigma_t)$ with $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_t > 0$. Then, $s_2 < \sigma_1 \leq s_1$. Denote by $u = [u_1 \ u_2 \ \dots \ u_{p-1}]^T$ and $v = [v_1 \ v_2 \ \dots \ v_{m-1}]^T$ the first column of U and V , respectively. Define an inner vector $\boldsymbol{\psi} = [\psi_1 \ \psi_2 \ \dots \ \psi_{p-1}]^T \in \mathcal{RH}_\infty^{(p-1) \times 1}$ as follows: Write $u_i = \rho_i \exp(j\theta_i)$ for each $i = 1, 2, \dots, p$, where the ρ_i 's are real and $0 \leq \theta_i < \pi$; if $\theta_i \neq 0$, set $\psi_i(s) = \rho_i(s - \beta_i)(s + \beta_i)^{-1}$, where $\beta_i > 0$ and is such that $\arg((j\omega_o - \beta_i)(j\omega_o + \beta_i)^{-1}) = \theta_i$; if $\theta_i = 0$, set $\psi_i = \rho_i$. Clearly, $\boldsymbol{\psi} \in \mathcal{RH}_\infty^{(p-1) \times 1}$, $\boldsymbol{\psi}^* \boldsymbol{\psi} = 1$, and $\boldsymbol{\psi}(j\omega_o) = u$. In a similar way, construct an $\mathcal{RH}_\infty^{(m-1) \times 1}$ inner vector $\boldsymbol{\xi}$ which "interpolates" v at $s = j\omega_o$, i.e., $\boldsymbol{\xi}(j\omega_o) = v$.

Define $\hat{\mathbf{X}}_1 = \operatorname{diag}(1, \boldsymbol{\xi}^*) \in \mathcal{RL}_\infty^{2 \times m}$ and $\hat{\mathbf{Y}}_1 = \mathbf{Y} \operatorname{diag}(\mathbf{a}, \boldsymbol{\psi}) \in \mathcal{RL}_\infty^{p \times 2}$. Clearly, $\hat{\mathbf{X}}_1 \hat{\mathbf{X}}_1^* = \hat{\mathbf{Y}}_1^* \hat{\mathbf{Y}}_1 = I_2$. Define factorizations of the columns (rows) of $\hat{\mathbf{X}}_1$ ($\hat{\mathbf{Y}}_1$) of the form $\hat{\mathbf{X}}_1^* = \hat{\mathbf{N}}_1 \operatorname{diag}(\mathbf{d}_1, \mathbf{d}_2)$ and $\hat{\mathbf{Y}}_1^* = \operatorname{diag}(\tilde{\mathbf{d}}_1, \tilde{\mathbf{d}}_2) \hat{\mathbf{N}}_2$ such that $\hat{\mathbf{N}}_1$ and $\hat{\mathbf{N}}_2^T$ are inner and $\mathbf{d}_1, \mathbf{d}_2, \tilde{\mathbf{d}}_1, \tilde{\mathbf{d}}_2$ are scalar allpass. Similar to a previous part of the proof, define 2×2 inner matrices $\mathbf{A}_1 = \operatorname{diag}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2)$ and $\mathbf{A}_2 = \operatorname{diag}(\tilde{\boldsymbol{\alpha}}_1, \tilde{\boldsymbol{\alpha}}_2)$ such that $\mathbf{d}_1(j\omega_o) \boldsymbol{\alpha}_1(j\omega_o) = 1$, $\mathbf{d}_2(j\omega_o) \boldsymbol{\alpha}_2(j\omega_o) = 1$, $\tilde{\mathbf{d}}_1(j\omega_o) \tilde{\boldsymbol{\alpha}}_1(j\omega_o) = 1$, and $\tilde{\mathbf{d}}_2(j\omega_o) \tilde{\boldsymbol{\alpha}}_2(j\omega_o) = 1$.

Define $\boldsymbol{\Delta} \in \mathcal{RH}_\infty^{m \times p}$ as

$$\boldsymbol{\Delta} = \hat{\mathbf{N}}_1 \mathbf{A}_1 \begin{bmatrix} \phi & \nu_1 \\ \nu_1 & -\phi \end{bmatrix} \mathbf{A}_2 \hat{\mathbf{N}}_2, \quad \nu_1 = \sqrt{\left(\frac{1}{\sigma_1} + \phi\right) \left(\frac{1}{s_1} - \phi\right)}.$$

The inner character of $\hat{\mathbf{N}}_1, \hat{\mathbf{N}}_2^T, \mathbf{A}_1$, and $\mathbf{A}_2$ implies that $\|\boldsymbol{\Delta}\|_\infty = \sqrt{\phi^2 + \nu_1^2} < \sqrt{\phi^2 + \nu_o^2}$ since $s_2 < \sigma_1$. Thus, $\|\boldsymbol{\Delta}\|_\infty$ is strictly less than the RHS of (38). Moreover, it can be easily verified that $\|\mathbf{x}^T \boldsymbol{\Delta} \mathbf{y}\|_\infty = \phi$. Finally, using the identity $\det(I - AB) = \det(I - BA)$ we can write

$$\det(I_m - \boldsymbol{\Delta}(j\omega_o) \mathbf{T}(j\omega_o)) = \det \left(I_m - Z_1 \begin{bmatrix} \phi & \nu_1 \\ \nu_1 & -\phi \end{bmatrix} Z_2 \begin{bmatrix} s_1 & 0 & 0 \\ 0 & \sigma_1 & 0 \\ 0 & 0 & * \end{bmatrix} \right),$$

where

$$\begin{aligned} Z_1 &= \operatorname{diag}(1, V') \mathbf{X}(j\omega_o) \mathbf{N}_1(j\omega_o) \mathbf{A}_1(j\omega_o), \\ Z_2 &= \mathbf{A}_2(j\omega_o) \mathbf{N}_2(j\omega_o) \mathbf{Y}(j\omega_o) \operatorname{diag}(\mathbf{a}(j\omega_o), U), \end{aligned}$$

and $*$ denotes a matrix not relevant for our present purposes. It can be easily verified from the above construction that $Z_1^T = [I_2 \ 0_{2, m-1}]$ and $Z_2 = [I_2 \ 0_{2, p-1}]$, and hence

$$\det(I_m - \boldsymbol{\Delta}(j\omega_o) \mathbf{T}(j\omega_o)) = \det \begin{bmatrix} 1 - \phi s_1 & -\nu_1 \sigma_1 & 0 \\ -\nu_1 s_1 & 1 + \phi \sigma_1 & 0 \\ 0 & 0 & I_{m-2} \end{bmatrix} = 0,$$

FIG. 2. *Extended permissible uncertainty set.*

which implies that $(\mathbf{G} + \Delta, \mathbf{K})$ is unstable from the generalized Nyquist theorem. Since such destabilizing Δ 's with $\|\Delta\|_\infty < \mu_1^*(\delta)$ and $\|\mathbf{x}^T \Delta \mathbf{y}\|_\infty = r_1(1 - \delta)$ can be constructed for any $\mathbf{K} \in \mathcal{K}_1 \setminus \mathcal{K}_2$, we conclude that $\mu^*(\delta) = \mu_1^*(\delta)$ and part 1 is proved. It is also clear that $\mathcal{K}_2$ is the set of all controllers $\mathbf{K}$ such that $(\mathbf{G} + \Delta, \mathbf{K}) \in \mathcal{S}$ for every $\Delta \in \mathcal{D}_{r_1}(\mathbf{G}) \cup \mathcal{E}(\delta, \mu^*(\delta))$ and part 2 follows.

To prove part 3, note that setting $\delta = 0$ in (21) gives

$$\mathcal{E}(0, \mu) = \{\Delta \in \mathcal{L}_\infty : \|\Delta\|_\infty < \mu, \|\mathbf{x}^T \Delta \mathbf{y}\|_\infty \leq r_1, \eta(\mathbf{G}) = \eta(\mathbf{G} + \Delta)\}.$$

Since from Lemma 4.4 there exist uniformly destabilizing perturbations of $\mathbf{G}$ in $\partial\mathcal{D}(\mathbf{G})$, we have that $\mu^*(0) \leq r_1$. Now,

$$(42) \quad \mathcal{E}(0, r_1) = \{\Delta \in \mathcal{L}_\infty : \|\Delta\|_\infty < r_1, \|\mathbf{x}^T \Delta \mathbf{y}\|_\infty \leq r_1, \eta(\mathbf{G}) = \eta(\mathbf{G} + \Delta)\}.$$

In view of the condition $\|\Delta\|_\infty < r_1$ in (42) and the fact that $\|\mathbf{x}\|_\infty = \|\mathbf{y}\|_\infty = 1$, condition $\|\mathbf{x}^T \Delta \mathbf{y}\|_\infty \leq r_1$ in the characterization of $\mathcal{E}(0, r_1)$ in (42) is superfluous and thus $\mathcal{E}(0, r_1) = \mathcal{D}_{r_1}(\mathbf{G})$. Hence, $\mu^*(0) = r_1$ and the set of all $\mathbf{K}$ such that $(\mathbf{G} + \Delta, \mathbf{K}) \in \mathcal{S}$ for every $\Delta \in \mathcal{D}_{r_1}(\mathbf{G}) \cup \mathcal{E}(0, \mu^*(0)) = \mathcal{D}_{r_1}(\mathbf{G})$ is clearly $\mathcal{K}_1$. Since $\mathcal{K}_2 \subseteq \mathcal{K}_1$, part 3 follows immediately from part 2. Finally, part 4 follows from the relations $s_1 = \sigma_1(\Gamma \mathbf{R}_{(-s)}) = \sigma_n^{-1}$, $\sigma_2(\Gamma \mathbf{R}_{(-s)}) = \sigma_{n-1}^{-1}$ (see Theorem 3.3 and the subsequent remark), and the inequality $s_2 \leq \sigma_2(\Gamma \mathbf{R}_{(-s)})$ (see [16]). $\square$

Remark 4.7. Figure 2 is an illustration of the set $\bigcup_{\delta \in [0,1]} \mathcal{E}(\delta, \mu^*(\delta))$ in the two-dimensional case. Here, $s_1 = 1$ and $s_2 = 0.25$. The “worst direction” is assumed to be the horizontal axis. The (open) disc of radius one represents the set of uncertainties guaranteed to be stabilized by optimal controllers. The area bounded by the solid curve represents the set of uncertainties guaranteed to be stabilized by (second-level) superoptimal controllers. Note the increase in the stability radius in all directions other than the worst direction.

5. An upper bound on the structured-singular value. So far, our results have been restricted to the unstructured uncertainty case. Our overall aim has been to use the degrees of freedom in the set of all optimal (“maximally robust”) controllers $\mathcal{K}_1$ in order to extend as far as possible the region of the uncertainty space guaranteed to be stabilized by a subset of $\mathcal{K}_1$. The optimal subset of $\mathcal{K}_1$ has been identified as the set of superoptimal controllers with respect to the first two levels, $\mathcal{K}_2$.

A different interpretation of our method allows us to apply our results to structured uncertainty models as well. The crucial point is that the region of the uncertainty space which is nondestabilizing can be extended beyond $\mathcal{D}_{r_1}$ only *by imposing a structure* on the admissible set of uncertainties; this structure, in our case, is of the form of a projection (uniform in frequency) in the “worst-case direction” $\langle \mathbf{y}\mathbf{x}^T, \cdot \rangle$, along which all uniformly destabilizing perturbations in $\partial\mathcal{D}_{r_1}(\mathbf{G})$ have been shown to lie. Suppose that the uncertainty is known to have a (block) diagonal structure, Δ , say. The following general procedure can be used in principle to obtain a lower bound on r_Δ , the robust stability radius with respect to structure Δ .

- Maximize the robust stability radius for a class of unstructured perturbations; let the maximum (unstructured) robust stability radius be r_1 and the corresponding worst-case direction be $\langle \mathbf{y}\mathbf{x}^T, \cdot \rangle$.
- Given a specific uncertainty structure Δ , find the largest $\delta^* \in (0, 1]$ compatible with Δ , i.e., the maximum $\delta^* \in (0, 1]$ such that $|\langle \mathbf{y}\mathbf{x}^T, \Delta \rangle| \leq r_1(1 - \delta^*)$, uniformly in ω for every $\Delta \in \Delta$.
- Then $\mu^*(\delta^*)$ is a guaranteed lower bound of the robust stability radius of the system with respect to uncertainty structure Δ .

This general method for calculating bounds for the structured robust-stability radius, r_Δ (equivalently the structured singular value μ_Δ), will be developed in future work. In this section we present preliminary results for the constant μ problem and a simple example illustrating our method. A more complete development is given in [9].

We use the definitions and notation of [19]. Let $T \in \mathcal{C}^{n \times n}$ have a singular value decomposition

$$(43) \quad T = U\Sigma V', \quad \Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_n), \quad U, V \in \mathcal{C}^{n \times n}, \quad U'U = V'V = I_n$$

and assume that

$$(44) \quad \sigma_1 > \sigma_2 \geq \sigma_3 \geq \dots \geq \sigma_n > 0.$$

Define the structured uncertainty set

$$\Delta = \{\text{diag}(\delta_1 I_{r_1}, \dots, \delta_S I_{r_S}, \Delta_1, \dots, \Delta_F) : \delta_1, \dots, \delta_S \in \mathcal{C}, \Delta_j \in \mathcal{C}^{m_j \times m_j}, j = 1, \dots, F\}$$

with $\sum_{i=1}^S r_i + \sum_{j=1}^F m_j = n$ and let $\mathbf{B}_\Delta = \{\Delta \in \Delta : \|\Delta\| \leq 1\}$. Then the structured singular value of T is defined as

$$\mu_\Delta(T)^{-1} = \min_{\substack{\Delta \in \Delta \\ \det(I - \Delta T) = 0}} \|\Delta\| = \min_{\substack{\Delta \in \Delta \\ \det(\Sigma^{-1} - V' \Delta U) = 0}} \|\Delta\|$$

(if there exists no $\Delta \in \Delta$ such that $\det(I - \Delta T) = 0$, we define $\mu_\Delta(T) = 0$). Let $u, v \in \mathcal{C}^{n \times 1}$ be the first columns of U and V , respectively. Partition u and v compatibly with Δ as follows:

$$(45) \quad u = \begin{bmatrix} u_1 \\ \vdots \\ u_S \\ u_{S+1} \\ \vdots \\ u_{S+F} \end{bmatrix}, \quad v = \begin{bmatrix} v_1 \\ \vdots \\ v_S \\ v_{S+1} \\ \vdots \\ v_{S+F} \end{bmatrix}, \quad u_i, v_i \in \mathcal{C}^{r_i}, \quad u_{S+j}, v_{S+j} \in \mathcal{C}^{m_j}$$

for $i = 1, \dots, S$ and $j = 1, \dots, F$. Then it is straightforward to verify that

$$(46) \quad \alpha_0 := \max_{\Delta \in \mathbf{B}_\Delta} |v' \Delta u| = \sum_{i=1}^S |v'_i u_i| + \sum_{j=1}^F \|v_{S+j}\| \|u_{S+j}\| \leq 1.$$

Define the set $\mathbf{\Delta}_{\alpha_0} := \{\Delta \in \mathcal{C}^{n \times n} : |v' \Delta u| \leq \alpha_0 \|\Delta\|\}$. Clearly $\mathbf{\Delta} \subseteq \mathbf{\Delta}_{\alpha_0}$. It follows that

$$\mu_{\mathbf{\Delta}}(T)^{-1} = \min_{\substack{\Delta \in \mathbf{\Delta} \\ \det(\Sigma^{-1} - V' \Delta U) = 0}} \|\Delta\| \geq \min_{\substack{\Delta \in \mathbf{\Delta}_{\alpha_0} \\ \det(\Sigma^{-1} - V' \Delta U) = 0}} \|\Delta\| =: \bar{\mu}_{\mathbf{\Delta}}(T)^{-1}.$$

Thus $\bar{\mu}_{\mathbf{\Delta}}(T)$ is an upper bound on $\mu_{\mathbf{\Delta}}(T)$. The evaluation of $\bar{\mu}_{\mathbf{\Delta}}(T)$ is related to the results of [15] (see Theorem 4.7). The next result uses Theorem 4.7 to give an expression for $\bar{\mu}_{\mathbf{\Delta}}(T)$ that involves only σ_1, σ_2, u, v , and the uncertainty set $\mathbf{\Delta}$ and shows that $\bar{\mu}_{\mathbf{\Delta}}(T)$ is increasing in σ_2 .

THEOREM 5.1. *Let $T \in \mathcal{C}^{n \times n}$ have a singular value decomposition as in (43) and assume that (44) is satisfied. Let u and v and the first columns of U and V , respectively, be partitioned as in (45) and define α_0 as in (46).*

1. *Let*

$$(47) \quad d = (\sigma_1 - \sigma_2)\alpha_0/2 + \sqrt{[(\sigma_1 - \sigma_2)\alpha_0/2]^2 + \sigma_1\sigma_2}.$$

Then

$$(48) \quad \bar{\mu}_{\mathbf{\Delta}}(T) := \left(\min_{\substack{\det(\Sigma^{-1} - V' \Delta U) = 0 \\ |v' \Delta u| \leq \alpha_0 \|\Delta\|}} \|\Delta\| \right)^{-1} = d.$$

2. *For all $\alpha_0 \in [0, 1]$ we have*

$$(49) \quad \mu_{\mathbf{\Delta}}(T) \leq \bar{\mu}_{\mathbf{\Delta}}(T) \leq \sigma_1.$$

3. *If $\alpha_0 = 1$, then*

$$(50) \quad \mu_{\mathbf{\Delta}}(T) = \bar{\mu}_{\mathbf{\Delta}}(T) = \sigma_1.$$

4. *If $\alpha_0 < 1$, then*

$$(51) \quad \mu_{\mathbf{\Delta}}(T) \leq \bar{\mu}_{\mathbf{\Delta}}(T) < \sigma_1$$

with $\mu_{\mathbf{\Delta}}(T) = \bar{\mu}_{\mathbf{\Delta}}(T)$ if and only if there exists $\Delta \in \mathbf{\Delta}$ such that

$$(52) \quad V' \Delta U = d^{-1} \begin{bmatrix} \alpha_0 & e^{j\theta} \sqrt{1 - \alpha_0^2} & 0 \\ e^{-j\theta} \sqrt{1 - \alpha_0^2} & -\alpha_0 & 0 \\ 0 & 0 & \Delta_{22} \end{bmatrix}$$

for arbitrary θ and any $\Delta_{22} \in \mathcal{C}^{(n-2) \times (n-2)}$ satisfying $\|\Delta_{22}\| \leq 1$.

Proof.

1. We first show that

$$(53) \quad \bar{\mu}_{\mathbf{\Delta}}(T)^{-1} \leq d^{-1}.$$

Let

$$\Delta = d^{-1}V \begin{bmatrix} \alpha_0 & \sqrt{1-\alpha_0^2} & 0 \\ \sqrt{1-\alpha_0^2} & -\alpha_0 & 0 \\ 0 & 0 & 0_{(n-2) \times (n-2)} \end{bmatrix} U'.$$

Then it is easy to verify that $\|\Delta\| = d^{-1}$, $|v'\Delta u| = \alpha_0 d^{-1} \leq \alpha_0 \|\Delta\|$, and

$$\det(\Sigma^{-1} - V'\Delta U) = \sigma_1^{-1}\sigma_2^{-1} - \alpha_0 d^{-1}(\sigma_2^{-1} - \sigma_1^{-1}) - \alpha_0^2 d^{-2} = 0$$

after some manipulation, and this proves (53). Thus we can restrict our search in (48) to the set

$$\{\Delta \in \mathcal{C}^{n \times n} : \|\Delta\| \leq d^{-1}, |v'\Delta u| \leq \alpha_0 d^{-1}\}$$

and so

$$\begin{aligned} \bar{\mu}_{\Delta}(T)^{-1} &= \min_{\substack{\det(\Sigma^{-1} - V'\Delta U) = 0 \\ |v'\Delta u| \leq \alpha_0 d^{-1} \\ \|\Delta\| \leq d^{-1}}} \|\Delta\| \\ (54) \quad &= \sqrt{\sigma_1^{-1}\sigma_2^{-1} - \alpha_0 d^{-1}(\sigma_2^{-1} - \sigma_1^{-1})} = d^{-1}, \end{aligned}$$

where the first equality in (54) follows from Theorem 4.7 and the second equality follows from a straightforward calculation using the definition of d (in fact, d is defined so that d^{-1} is the positive solution of (54)).

2. It is straightforward to verify that $d \leq \sigma_1$. The first inequality follows from the definitions of $\bar{\mu}_{\Delta}(T)$ and $\mu_{\Delta}(T)$.
3. Suppose that $\alpha_0 = 1$. Then a simple calculation verifies that $d = \sigma_1$ which proves the second equality in (50). To prove the first equality, define

$$\Delta = \sigma_1^{-1} \text{diag}(\delta_1 I_{r_1}, \dots, \delta_S I_{r_S}, \Delta_1, \dots, \Delta_F) \in \mathbf{\Delta},$$

where

$$\begin{aligned} \delta_i &= \frac{|v'_i u_i|}{v'_i u_i} \in \mathcal{C} \quad (\delta_i = 0 \text{ if } v'_i u_i = 0), \quad i = 1, \dots, S, \\ \Delta_j &= \frac{v_j u'_j}{\|v_j\| \|u_j\|} \in \mathcal{C}^{m_j \times m_j} \quad (\Delta_j = 0 \text{ if } \|v_j\| \|u_j\| = 0), \quad j = 1, \dots, F, \end{aligned}$$

where u_i, v_i are defined in (45). Then $\|\Delta\| = \sigma_1^{-1}$ and $v'\Delta u = \sigma_1^{-1}$. This implies that $V'\Delta U = \text{diag}(\sigma_1^{-1}, \Delta_{22})$ for some Δ_{22} with $\|\Delta_{22}\| \leq \sigma_1^{-1}$. It is easy to verify that $\det(\Sigma^{-1} - V'\Delta U) = 0$ and so $\mu_{\Delta}(T) = \sigma_1$ and the first equality in (50) is proved.

4. Suppose that $\alpha_0 < 1$. Then a simple verification shows that $d < \sigma_1$ and establishes the second inequality in (51). Part 2 of Theorem 4.7 gives all Δ such that $\|\Delta\| = d^{-1}$, $\det(\Sigma^{-1} - V'\Delta U) = 0$, and $|v'\Delta u| \leq \alpha_0 d^{-1}$ as in (52). Hence $\mu_{\Delta}(T) = \bar{\mu}_{\Delta}(T)$ if and only if there exists such a $\Delta \in \mathbf{\Delta}$.

This completes the proof. $\square$

Remark 5.1. Note that $\bar{\mu}_{\Delta}(T)$ depends only on σ_1, σ_2 , and α_0 (see (47) and (48)), and α_0 in turn depends only on u, v , and the structured uncertainty set $\mathbf{\Delta}$ (see (46)). In the context of the robust stabilization of systems with unstructured additive

FIG. 3. $\sigma_1(\mathbf{T})$ (dashed), D -iteration upper bound (dash-dot), and upper bound $\bar{\mu}_\Delta(\mathbf{T})$ (solid).

perturbations, all optimal interpolating functions share the same largest singular value s_1 at all frequencies. They also share the same (frequency-dependent) singular vectors corresponding to s_1 (the inner vectors $\mathbf{x}$ and $\mathbf{y}$). Thus the only free parameter that can be used to minimize $\bar{\mu}_\Delta(\cdot)$ within the set $\mathcal{T}_1$ is the second largest singular value. Noting that $\bar{\mu}_\Delta(\cdot)$ is a nonincreasing function of σ_2 (see (47)) suggests that within $\mathcal{T}_1$, $\bar{\mu}_\Delta(\cdot)$ is minimized by $\mathcal{T}_2$. This will be elaborated in a future work.

Remark 5.2. The bound $\bar{\mu}_\Delta$, although in general tighter than σ_1 , is less tight than the upper bound of the D -iteration [19]. In fact, it is shown in [19] that at the end of the D -iteration, either

1. $\sigma_1(T) = \sigma_2(T)$, in which case our results are not applicable (see (44)), or
2. $\sigma_1(T) > \sigma_2(T)$, in which case $\mu_\Delta(T) = \sigma_1(T)$. It can be shown that this corresponds to $\alpha_0 = 1$.

The main purpose in this work is to illustrate the improved robustness properties of superoptimal controllers, rather than attempting to improve the D -iteration bound.

Example 5.1. This example illustrates the upper bound $\bar{\mu}_\Delta(\mathbf{T})$, where

$$\mathbf{T} \stackrel{s}{=} \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \left[\begin{array}{ccc|ccc} -5.91 & -11.49 & 6.03 & -0.59 & -1.90 & 0.19 & 1.35 \\ -1.85 & -5.62 & 1.63 & -0.31 & 1.11 & 0.63 & 0.12 \\ -7.71 & -17.40 & 7.07 & 0.97 & 0.72 & -0.35 & -0.58 \\ \hline -0.37 & 0.49 & 0.52 & 0.01 & 0.22 & 0.71 & 0.97 \\ 1.43 & -0.09 & 1.36 & 0.60 & 0.70 & 0.23 & 0.36 \\ 0.07 & 0.37 & -0.41 & 0.82 & 0.52 & 0.45 & 0.05 \\ -0.23 & -0.15 & 0.66 & 0.98 & 0.93 & 0.17 & 0.76 \end{array} \right]$$

is chosen as random with A stable. The computation is carried out pointwise across the frequency grid, i.e., for each ω , $\sigma_1(\mathbf{T}(j\omega))$ and $\bar{\mu}_\Delta(\mathbf{T}(j\omega))$ are computed and compared with the D -iteration upper bound. The uncertainty structure Δ is taken to be diagonal, i.e., $S = 0, F = 4, m_j = 1$ for $j = 1, \dots, 4$. The plots are shown in Figure 3.

6. Conclusions.

By way of conclusion, we summarize our contribution.

- We have analyzed in detail the maximum robust stabilization problem subject to unstructured additive perturbations. We have shown that a critical direction exists in the uncertainty space, along which all maximum-norm boundary

perturbations are destabilizing for every optimal controller.

- We have shown that by imposing a parametric constraint in the most critical direction, the set of uncertainties guaranteed to be stabilized by a subset of all optimal controllers can be further extended. We have shown that the optimal solution to this problem is associated with the set of superoptimal controllers with respect to the first two levels, and we have obtained a closed-form expression for the improved robust stability radius which involves the first two superoptimal levels.
- By adapting our results to the structured uncertainty case, we have obtained an easily computable upper bound on the structured-singular value (which is tighter than the largest singular value), without the need to carry out a D -iteration. We have further shown that the minimization of this bound is equivalent to the minimization of the second largest singular value, which again motivates superoptimization.

There are a number of related research directions which we intend to pursue.

- For purposes of clarity, our technique has been restricted to unstructured additive uncertainty models. There is no conceptual difficulty, however, in extending our method to other types of unstructured uncertainty (multiplicative, inverse-multiplicative, coprime) or to include frequency weightings. Rather than analyzing each case individually, we intend to address the general problem involving linear fractional transformation uncertainty models [17]. This is likely to involve a general-distance superoptimal approximation problem, the solution of which is already in place [26], [8], [16], [13], [14], [12].
- Our method relies on Theorem 4.7 which generalizes a result in [15]. Section 5 suggests that generalizing this theorem should be useful in robust stability analysis of systems subject to structured uncertainty. We have derived some results in this direction which will be reported in a future publication.
- We intend to investigate the possibility of applying our method as an alternative to current μ -synthesis techniques. The main potential advantage of our approach is the possibility of avoiding the calculation of the optimal scaling “ D -matrix” in the $D - K$ iteration [2] (currently carried out pointwise over a discretized frequency grid) by using instead the directionality information provided by the two Schmidt vectors, which define the worst-case direction in our setting. The success of such an approach will ultimately rest on how tightly we can overbound the structured singular value. Although our computational experience so far is promising, this remains an open question.

REFERENCES

- [1] R. F. CURTAIN AND H. J. ZWART, *An Introduction to Infinite-Dimensional Linear Systems Theory*, Springer-Verlag, New York, 1991.
- [2] J. C. DOYLE, *Lecture Notes in Advances in Multivariable Control*, ONR/Honeywell Workshop, Minneapolis, MN, 1984.
- [3] B. A. FRANCIS, *A Course in $\mathcal{H}_\infty$ Control Theory*, Springer-Verlag, Berlin, 1987.
- [4] K. GLOVER, *All optimal Hankel-norm approximations of linear multivariable systems and their $\mathcal{L}_\infty$ error bounds*, Internat. J. Control, 39 (1984), pp. 1115–1193.
- [5] K. GLOVER, *A tutorial on Hankel-norm approximation*, in From Data to Model, J. C. Willems, ed., Springer-Verlag, Berlin, New York, 1989.
- [6] K. GLOVER, *Robust stabilization of linear multivariable systems: Relations to approximation*, Internat. J. Control, 43 (1986), pp. 741–746.
- [7] K. GLOVER, D. J. N. LIMBEER, J. C. DOYLE, E. M. KASENALLY, AND M. G. SAFONOV, *A characterization of all solutions to the four block general distance problem*, SIAM J.

- Control Optim., 29 (1991), pp. 283–324.
- [8] D. W. GU, M. C. TSAI, AND I. POSTLETHWAITE, *An algorithm for superoptimal $\mathcal{H}_\infty$ design: The 2-block case*, Automatica J. IFAC, 25 (1989), pp. 437–440.
 - [9] S. K. GUNGAH, *Maximally Robust Controllers for a Class of Unstructured Uncertainty*, Ph.D. thesis, London University, 1998.
 - [10] G. D. HALIKIAS, *An affine parametrization of all one-block $\mathcal{H}_\infty$ -optimal matrix interpolating functions*, Internat. J. Control, 57 (1993), pp. 1421–1441.
 - [11] G. D. HALIKIAS AND I. M. JAIMOUKHA, *Hierarchical optimization in $\mathcal{H}_\infty$* , IEEE Trans. Automat. Control, 43 (1998), pp. 1123–1128.
 - [12] G. D. HALIKIAS AND I. M. JAIMOUKHA, *The two-block superoptimal AAK problem*, Math. Control Signals Systems, 11 (1998), pp. 244–264.
 - [13] G. D. HALIKIAS, D. J. N. LIMEBEER, AND K. GLOVER, *A state-space analysis for the super-optimal Hankel-norm approximation problem*, SIAM J. Control Optim., 31 (1993), pp. 960–982.
 - [14] I. M. JAIMOUKHA AND D. J. N. LIMEBEER, *A state-space algorithm for the solution of the 2-block superoptimal distance problem*, SIAM J. Control Optim., 31 (1993), pp. 1115–1134.
 - [15] N. A. LEHTOMAKI, D. A. CASTANON, B. D. LEVY, G. STEIN, N. R. SANDEL, AND M. ATHANS, *Robustness and modeling error characterization*, IEEE Trans. Automat. Control, 29 (1984), pp. 212–220.
 - [16] D. J. N. LIMEBEER, G. D. HALIKIAS, AND K. GLOVER, *State-space algorithm for the computation of superoptimal matrix interpolating functions*, Internat. J. Control, 50 (1989), pp. 2431–2466.
 - [17] D. C. MCFARLANE AND K. GLOVER, *Robust Controller Design Using Normalized Coprime Factor Plant Descriptions*, Lecture Notes in Control and Inform. Sci. 138, Springer-Verlag, Berlin, New York, 1990.
 - [18] P.-O. NYMAN, *Improving robustness by superoptimization*, in Proceedings of the 3rd European Control Conference, Rome, Italy, 1995, pp. 1039–1044.
 - [19] A. PACKARD AND J. C. DOYLE, *The complex structured singular value*, Automatica J. IFAC, 21 (1993), pp. 79–109.
 - [20] S. TREIL, *On superoptimal approximation by analytic and meromorphic matrix valued functions*, J. Funct. Anal., 131 (1995), pp. 386–414.
 - [21] M. C. TSAI, D. W. GU, AND I. POSTLETHWAITE, *A state space approach to super-optimal $\mathcal{H}^\infty$ control problems*, IEEE Trans. Automat. Control, 33 (1988), pp. 833–843.
 - [22] M. C. TSAI, D. W. GU, I. POSTLETHWAITE, AND B. D. O. ANDERSON, *A Pseudo-Singular Value Decomposition and Inner Functions in Superoptimal $\mathcal{H}_\infty$ Control*, Report OUEL 1738/88, University of Oxford, 1988.
 - [23] M. S. VERMA, *Synthesis of Infinity-Norm Optimal Linear Feedback Systems*, Ph.D. thesis, University of Southern California, Los Angeles, CA, 1985.
 - [24] M. S. VERMA, J. W. HELTON, AND E. A. JONCKHEERE, *Robust stabilization of a family of plants with varying number of right half plane poles*, in Proceedings of the American Control Conference, Seattle, WA, IEEE, Piscataway, NJ, 1986, pp. 1827–1832.
 - [25] M. VIDYASAGAR, *Control System Synthesis: A Factorization Approach*, MIT Press, Cambridge, MA, 1985.
 - [26] N. J. YOUNG, *The Nevanlinna-Pick problem for matrix-valued functions*, J. Operator Theory, 15 (1986), pp. 239–265.