Robust Synthesis in ℓ_1: A Globally Optimal Solution

Mustafa Khammash, Murti V. Salapaka, and Tim Van Voorhis

***Abstract*—In this paper, a method to synthesize controllers that achieve *globally optimal* robust performance within any prespecified tolerance, against structured norm-bounded time-varying and/or nonlinear uncertainty is developed. The performance measure considered is the infinity to infinity induced norm of a system's transfer function. The method developed utilizes linear relaxation techniques to solve the infinite dimensional nonconvex problem via finite dimensional linear programming problems.**

***Index Terms*—ℓ_1, optimal, robustness, uncertainty.**

I. Introduction

SYSTEM robustness has been studied extensively since the early 1980s. The reason this topic has received so much attention by many researchers stems from the fact that effective control designs must rely either on very accurate models, or on inherent robustness in the feedback system. Accurate models for the system to be controlled are quite rare and frequently obtaining accurate models is difficult, time-consuming, and expensive. Even in the cases when accurate models are available, inevitable variations in the parameters of such models render ineffective any design which lacks robustness. Therefore it is evident that "robustness" should be a central design goal.

Various results analyzing the stability and robustness of performance for several types of uncertainty models have emerged in the literature. These results depend on the way uncertainty in the model is characterized. For model uncertainty characterized by time-invariant norm bounded structured perturbations, the structured singular value (SSV) theory [1], [2] provides a nonconservative measure for stability. The multivariable robustness margin [3] provides an equivalent measure. While exact calculation of these measures have proved difficult in the general case, much work has been devoted to the computation of effective bounds that are used to provide robustness estimates (see, for example, [4]–[6]). For time-varying norm bounded structured perturbations, robustness analysis turns out to be surprisingly easier. In the work of [7]–[9], easily computable exact conditions for robust stability and performance were derived in the case when the induced ℓ_∞ norm is used to characterize allowable perturbations and exogenous disturbances. Likewise, when the induced-two norm is used instead, [10], [11] provide computable conditions as well.

The robust synthesis problem, however, is significantly more challenging than the analysis problem. The problem of synthesis of controllers that minimize the SSV function remains largely unsolved. However, controllers can be designed using the $D-K$ iterations method (see, e.g., [2]) which find locally optimal solutions. When perturbations are characterized by an induced ℓ_∞ norm bound, a $D-K$ type iteration procedure is also available [9]. However, as in the case of the SSV, this iteration scheme does not guarantee that a global minimum is achieved. The problem appears to be an inherently nonconvex one in either norm, and is difficult to solve in general. There are exceptions in some special cases. For instance, in the full information and state-feedback cases, the synthesis problem for H_∞ has been solved for any number of uncertainty blocks [12]. In the presence of one uncertainty block that is characterized by an induced ℓ_∞ norm bound, the robust performance synthesis problem can be solved by using sensitivity analysis of linear programming [13]. Likewise the nominal H_2 with robust ℓ_1 performance for a single uncertainty case can be solved by sensitivity analysis of quadratic programming problems [14]. Special cases notwithstanding, a general solution to the robust synthesis problem using any norm requires addressing the underlying nonconvex optimization problem directly. This is the approach taken in [15] which provides an algorithm for approximately finding a global solution to the constantly scaled H_∞ control synthesis problem. Such problem arises, for example, when structured time-varying uncertainty with an ℓ_2 induced-norm bound is assumed. In [16] controllers achieving globally optimal performance in the presence of uncertainty described in the H_∞ setting for process control applications have been synthesized.

In this article, we address the output feedback synthesis problem when the signal norm is the ℓ_∞ norm and the perturbations are structured time-varying and/or nonlinear systems with an induced ∞-norm bound. A *globally optimal* solution to the robust synthesis problem is obtained. It is shown that the solution involves only solving certain linear programming problems resulting from relaxations of the original nonconvex problem. The optimal solution can be obtained to any prescribed accuracy by following a proposed branch and bound scheme. This presents the first instance of a global solution to the general robust synthesis problem in the ℓ_1 setting.

II. Notation

R	The real number system.
R^n	The n dimensional Euclidean space.
$\lvert x\rvert_\infty$	The ∞-norm of the vector $x=(x_1,\ldots,x_n)$ defined as $\lvert x\rvert_\infty := \max_{1\le k\le n} \lvert x_k\rvert$.

Manuscript received July 7, 2000; revised January 10, 2001 and April 8, 2001. Recommended by Associate Editor K. M. Grigoriadis. This was was supported by the National Science Foundation under Grants ECS-9733802, ECS-9110764, ECS-9457485, and by Iowa State University under Grant SPRIGS-7041747.

M. Khammash and M. V. Salapaka are with the Department of Electrical and Computer Engineering, Iowa State University, Ames, IA 50011 USA.

T. Van Voorhis is with the Department of Industrial and Manufacturing Systems Engineering, Iowa State University, Ames, IA 50011 USA.

Publisher Item Identifier S 0018-9286(01)10345-4.

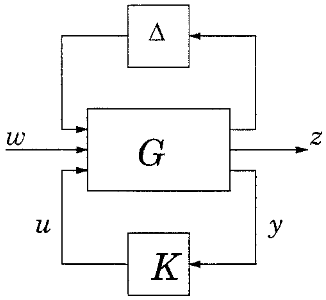

Fig. 1. In the robust performance synthesis problem a controller K that achieves stability and performance in the presence uncertainty Δ is sought. The uncertainty Δ is restricted to some prespecified set $\underline{\Delta}$.

ℓ_∞ — The space of bounded sequences of real numbers. The norm of an element x in ℓ_∞ is given by $\|x\|_\infty = \sup_k |x(k)|$.

ℓ_1 — The Banach space of right sided absolutely summable real sequences with the norm given by $\|x\|_1 := \sum_{k=0}^{\infty} |x(k)|$.

$\ell_1^{m\times n}$ — The Banach space of matrix valued right sided real sequences with the norm $\|x\|_1 := \max_{1\le i\le m} \sum_{j=1}^{n} \|x_{ij}\|_1$ where $x \in \ell_1^{m\times n}$ is the matrix (x_{ij}) and each x_{ij} is in ℓ_1.

$\geq$ — On the set of $p \times q$ matrices we define the order $\geq$ by $A \geq B$ if and only if $A_{ij} \geq B_{ij}$ for all $i = 1, \ldots, p$ and $j = 1, \ldots, q$ with A_{ij} and B_{ij} denoting the ijth element of the matrix A and B, respectively.

$|.|_1$ — For any matrix A in $R^{p\times q}$, $|A|_1 := \sup_{|w|_\infty = 1} |Aw|_\infty$ denotes the infinity induced operator norm.

P_N — Given a nonnegative integer N, P_N is the truncation operator mapping the set of right sided real sequences into itself such that for any sequence x, $(P_N x)(k) = x(k)$ when $k \leq N$, and $(P_N x)(k) = 0$ when $k > N$.

$\mathcal{S}$ — $\mathcal{S}$ is the shift operator which maps the set of right real sequences into itself such that for any sequence x, $(\mathcal{S}x)(k) = x(k-1)$ when $k \geq 1$, and $(\mathcal{S}x)(0) = 0$.

$\mathcal{I}$ — $\mathcal{I}$ is a $n \times n$ matrix such that all its elements are equal to one that is $\mathcal{I}_{ij} = 1$ for all $i = 1, \ldots, n$ and $j = 1, \ldots n$.

$\mathcal{D}$ — $\mathcal{D}$ is the set of $n \times n$ diagonal matrices with strictly positive diagonal elements that is $\mathcal{D} := \{\mathrm{diag}(d_1, \ldots d_n): d_i > 0\}$.

$[L, U]$ — Given two vectors $L = (L_1, \ldots, L_n)$ and $U = (U_1, \ldots, U_n)$ in R^n such that $L_i \leq U_i$, $[L, U]$ denotes the set $\{x \in R^n: L_i \leq x_i \leq U_i\}$.

III. Problem Statement

Consider the system described in Fig. 1, where $G, K, u, w, z,$ and y are the generalized linear time-invariant (LTI) plant (given), an LTI controller (to be designed), the control input, the exogenous disturbance, the regulated output, and the measured output respectively. In this article we will restrict the study to discrete-time systems. Model uncertainty is reflected by the perturbation block Δ which is restricted to lie in the following set of admissible perturbations:

$$\underline{\Delta} := \left\{ \Delta = \mathrm{diag}(\Delta_1, \ldots, \Delta_n) \colon \Delta_i \colon \ell_\infty \to \ell_\infty \text{ is causal and } \|\Delta_i\| := \sup_{u \neq 0} \frac{\|\Delta_i u\|_\infty}{\|u\|_\infty} \leq 1 \right\}.$$

The perturbation may therefore be nonlinear or time-varying. The system in Fig. 1 is said to be robustly stable if it is ℓ_∞-stable for all admissible perturbations, i.e., for all $\Delta \in \underline{\Delta}$. The problem we shall address is as follows.

Problem Statement: Find a linear finite-dimensional controller K such that:

1) the system achieves robust stability;
2) the system achieves robust performance, i.e.,

$$\|\mathcal{T}_{zw}(K, \Delta)\| < 1 \qquad \forall\, \Delta \in \underline{\Delta}$$

where $\mathcal{T}_{zw}$ is the map from w to z, and the norm used above is the induced ℓ_∞ operator norm.

It can be seen that when Δ is taken to be zero, the robust stability condition becomes a condition on nominal stability, while the robust performance condition becomes equivalent to a norm requirement on the induced ℓ_∞ norm of the nominal (LTI) system. It can easily be shown that in this case, this induced norm is equal to the ℓ_1 norm of the impulse response for the system mapping w to z. Therefore, when Δ is forced to be zero, the problem stated here reduces to the so-called ℓ_1 optimal control problem which has been addressed in the literature. In this sense, we are seeking to solve the "robust ℓ_1" synthesis problem, although the terminology is not accurate since for a given time-varying Δ, the operator $\mathcal{T}_{zw}(K, \Delta)$ is in general time-varying, and its induced ℓ_∞ norm is not generally equal to the ℓ_1 norm of its impulse response. Despite this fact, the term is convenient and will be used.

IV. Conditions for Robustness

Before addressing the robust ℓ_1 synthesis problem stated in the previous section we look at the relevant robust analysis conditions. The robust performance problem can be converted to a robust stability problem, by adding a fictitious perturbation block Δ_p connecting the output z into the input w (see [8]). Thus there is no loss of generality if we address the robust stability problem alone.

The robust stability conditions will be stated with the aid of Fig. 2, where G and any stabilizing controller K have been lumped into one system M. Let Φ be the impulse response matrix of the map M. Since M is LTI, causal, and stable, Φ belongs to $\ell_1^{n\times n}$. Each Φ_{ij} has a norm which can be computed arbitrarily accurately. In particular, $\|\Phi_{ij}\|_1 = |E_{ij}| + \sum_{k=0}^{\infty} |C_i A^k B_j|$, where A, B_i, C_j, E_{ij} are the constant matrices in the state-space description of M_{ij} [that is the transfer function matrix associated with M_{ij} is $C_j(sI - A)^{-1} B_i + E_{ij}$]. We can, therefore, compute the following nonnegative matrix:

$$\hat{\Phi} = \begin{bmatrix} \|\Phi_{11}\|_1 & \ldots & \|\Phi_{1n}\|_1 \\ \vdots & & \vdots \\ \|\Phi_{n1}\|_1 & \ldots & \|\Phi_{nn}\|_1 \end{bmatrix}.$$

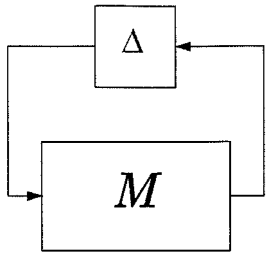

Fig. 2. In the robust stability analysis problem, a nonconservative condition is sought to assess the stability of the interconnection of lumped plant-controller system M and the uncertainty Δ that can lie in a prespecified set $\underline{\Delta}$.

The following robustness conditions will be the basis for the proposed synthesis method.

Theorem 1 [9]: The system in Fig. 2 achieves robust stability if and only if the following condition holds: $\inf_{D\in\mathcal{D}} \|D^{-1}\Phi D\|_1 < 1$ where $\mathcal{D} := \{\mathrm{diag}(d_1, \ldots, d_n): d_i > 0\}$.

V. Formulation as an Optimization Problem

The dependence of Φ on the controller K can be captured through the Youla parameter Q which provides a parameterization of all possible Φ's that can be obtained with a stabilizing controller (see [17]). Hence, the robust *synthesis* problem can be stated as follows:

$$\inf_{D\in\mathcal{D}} \inf_{Q\in\ell_1} \|D^{-1}\Phi(Q)D\|_1 =: \gamma_* \tag{1}$$

where $\Phi(Q) = H - U * Q * V$ is the standard parameterization of the closed-loop system using the Youla parameter Q (see [18]). For each fixed $D = \mathrm{diag}(d_1, \ldots, d_n)$, problem (1) is a standard ℓ_1 norm-minimization problem.

The formulation of problem (1) does not constrain the size of the Q which in principle may be arbitrarily large. This does not pose any difficulties if an optimal Q is known to exist. However, when such an optimal Q does not exist, as may happen when the z-transform of U or V has zeros on the unit circle, having an unconstrained Q in the optimization problem poses difficulties. In this case, the infimum in (1) may not be achieved by a single Q, and can be approached by a sequence Q_k whose norms $\|Q_k\|_1$ grow without bound. This situation arises even in the nominal ℓ_1 optimization problem, and is considered in [19]–[21] as one of main motivations for regularizing the underlying optimization problem through an *a priori* bound on Q. We will adopt this approach here by introducing a regularizing bound on Q of the form: $\|Q\|_1 \le \alpha$. We define

$$\begin{aligned} \gamma_{opt} := \quad & \inf \|D^{-1}\Phi_\epsilon D\|_1 \\ \text{subject to} \quad & \\ & \|Q\|_1 \le \alpha \\ & \Phi = H - U * Q * V \\ & D \in \mathcal{D}. \end{aligned} \tag{2}$$

Clearly, if the optimization problem (1) has a solution Q_{opt}, then for any $\alpha \ge \|Q_{opt}\|_1$ the two problems (1) and (2) are equivalent. At the same time, if the optimization problem (1) has no solution, it can be shown that the constrained problem (2) always has a solution, and the Q constraint offers a practical way to eliminate suboptimal solutions that result from Q parameters with unacceptably large gain. Hence, for the remainder of the work we will focus on problem (2) above.

A. Perturbing the Optimization Problem

In this section, we explore the effect of perturbing the problem data associated with the optimization problem (2). In particular, we are interested in studying how small perturbations in Φ affect the optimal solution of the resulting problem. As we will show below, if the perturbations are introduced in an appropriate way, the resulting optimization problem becomes easier to solve. In our case, the introduced perturbations will allow us to reduce the D parameter search space from the first orthant, an unbounded set, to a compact subset. Such a reduction in the search space will make it possible to use powerful relaxation methods that will yield the solution to the problem.

We begin by perturbing problem to obtain (2) the following problem:

$$\begin{aligned} \gamma_{opt}^\epsilon := \quad & \inf \|D^{-1}\Phi_\epsilon D\|_1 \\ \text{subject to} \quad & \\ & \|Q\|_1 \le \alpha \\ & \Phi_\epsilon = (\epsilon\mathcal{I} + \mathcal{S}H) - \mathcal{S}U * Q * V \\ & D \in \mathcal{D}. \end{aligned} \tag{3}$$

Note that the problem (3) is the same as problem (2) with H and U replaced by $\epsilon\mathcal{I} + \mathcal{S}H$ and $\mathcal{S}U$ respectively where $\mathcal{S}$ represents the shift operator.

We shall show in Theorem 2 below that the solution to the above perturbed problem can be made arbitrarily close to that of problem (2), provided ϵ is sufficiently small. Specifically, we show that given any $\delta > 0$ one can find $\epsilon > 0$ such that the $|\gamma_{opt}^\epsilon - \gamma_{opt}| < \delta$. Furthermore, the computation of ϵ does not require knowing the solution of either optimization problem and can therefore be done *a priori*. In establishing this theorem, we need to utilize a matrix perturbation result. Essentially, this results states that the change in the spectral radius of a nonnegative matrix upon perturbation can be bounded by the size of the perturbation, the norm of the matrix and its size.

Lemma 1: Let M and ϵ^* be positive numbers. Let

$$\mathcal{A} := \{A \in R^{n\times n} | A \ge 0 \text{ and } |A|_1 \le M\}.$$

Then for all $A \in \mathcal{A}$ and $0 < \epsilon \le \epsilon^*$

$$\rho(A + \epsilon\mathcal{I}) \le \rho(A) + \epsilon^{1/n} f(n, M, \epsilon^*)$$

where f is independent of A and ϵ.

Proof: The proof is a modification of the proof of [22, Lemma 5.6.10]. Note that for any matrix A, there exists a unitary matrix U_A such that $A = U_A^* \Lambda_A U_A$ where Λ_A is upper-triangular and has the form

$$\Lambda_A = \begin{pmatrix} \lambda_1(A) & d_{12}(A) & d_{13}(A) & \ldots & d_{1n}(A) \\ 0 & \lambda_2(A) & d_{23}(A) & \ldots & d_{2n}(A) \\ 0 & 0 & \ddots & \cdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 & \lambda_n(A) \end{pmatrix}$$

(see [22]). Note that $|\Lambda_A|_1 \le |U_A|_1 |A|_1 |U_A^*|_1 \le nM$ where we have used the fact that for any $n \times n$ unitary matrix U, $|U|_1 \le \sqrt{n}$. Therefore $|d_{ij}(A)| \le nM$ for all relevant indices i, j. Let

$t = (nM/\epsilon) + 1$ and let $D_t := \mathrm{diag}(t, t^2, \ldots t^n)$. It follows that $D_t \Lambda_A D_t^{-1}$ is given by the matrix:

$$\begin{pmatrix} \lambda_1(A) & t^{-1}d_{12}(A) & t^{-2}d_{13}(A) & \ldots & t^{-n+1}d_{1n}(A) \\ 0 & \lambda_2(A) & t^{-2}d_{23}(A) & \ldots & t^{-n+2}d_{2n}(A) \\ 0 & 0 & \ddots & \cdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 & \lambda_n(A) \end{pmatrix}.$$

Note that

$$\sum_{j>i} t^{-j+1}|d_{ij}(A)| \le \sum_{j>i} t^{-j+1} nM \le nM \frac{1}{t-1} = \epsilon$$

for all $A \in \mathcal{A}$ and for all $i = 1, \ldots n$. Thus the sum of the absolute values of the off-diagonal terms in $D_t \Lambda_A D_t^{-1}$ in each of its row is less than ϵ. Thus $|D_t \Lambda_A D_t^{-1}|_1 \le \rho(A) + \epsilon$ for all $A \in \mathcal{A}$.

Define $|\cdot|_A$: $R^{n\times n} \to R$ by

$$|B|_A = |(D_t U_A) B (D_t U_A)^{-1}|_1.$$

$|\cdot|_A$ is an induced norm on $R^{n\times n}$. This follows from the fact that if $|\cdot|$ is a induced norm then $|\cdot|_S$: $R^{n\times n} \to R$ defined by $|B|_T := |TBT^{-1}|$ for an invertible matrix T is also a induced norm on the set of $n \times n$ matrices (see [22]).

Also

$$\begin{aligned} |\mathcal{I}|_A &= |(D_t U_A)\mathcal{I}(D_t U_A)^{-1}|_1 \\ &\le |D_t|_1 |D_t^{-1}|_1 |U_A|_1 |U_A^*|_1 |\mathcal{I}|_1 \\ &\le t^n \frac{1}{t} n^2 = n^2 t^{n-1}. \end{aligned}$$

Thus for all A in $\mathcal{A}$, $|\cdot|_A$ is a induced norm with the following properties:

- $|\mathcal{I}|_A \le n^2 t^{n-1}$ with $t = (nM/\epsilon) + 1$;
- $|A|_A \le \rho(A) + \epsilon$;

Consider any A in $\mathcal{A}$ then it follows that:

$$\begin{aligned} \rho(\epsilon^n \mathcal{I} + A) &\le |\epsilon^n \mathcal{I} + A|_A \\ &\le \epsilon^n |\mathcal{I}|_A + |A|_A \\ &\le \epsilon^n n^2 \left(\frac{nM}{\epsilon} + 1\right)^{n-1} + \rho(A) + \epsilon \\ &\le \rho(A) + \epsilon(n^2(nM + \epsilon^*)^{n-1} + 1) \end{aligned}$$

where the first inequality follows from the fact that the spectral radius of a matrix is less than or equal to any induced norm of the matrix (see [22]). The second inequality follows because $|\cdot|_A$ is a induced norm. This proves the lemma. ∎

We now relate the solution of problem (3) to that of the perturbed problem (2).

Theorem 2: Let ϵ^* be any positive number. For any $0 < \epsilon \le \epsilon^*$,

$$|\gamma_{opt}^{\epsilon} - \gamma_{opt}| \le \epsilon^{1/n} f(n, M, \epsilon^*)$$

where $M = \|H\|_1 + \alpha \|U\|_1 \|V\|_1$.

Proof: For a nonnegative matrix P, $\rho(P) = \inf_{D\in\mathcal{D}} |D^{-1}PD|_1$. Using this fact along with the definition of the ℓ_1 norm, it can be easily seen that

$$\gamma_{opt}^{\epsilon} = \inf_{\|Q\|_1 \le \alpha} \rho\left(\epsilon\mathcal{I} + \hat{\Phi}(Q)\right),$$

$$\gamma_{opt} = \inf_{\|Q\|_1 \le \alpha} \rho\left(\hat{\Phi}(Q)\right).$$

Since $(\epsilon\mathcal{I} + \hat{\Phi}(Q)) \ge \hat{\Phi}(Q)$, we have that

$$\inf_{\|Q\|_1 \le \alpha} \rho\left(\epsilon\mathcal{I} + \hat{\Phi}(Q)\right) - \inf_{\|Q\|_1 \le \alpha} \rho\left(\hat{\Phi}(Q)\right) \ge 0. \quad (4)$$

Let $\delta > 0$ be arbitrary. Thus, there exists a $\underline{Q}$ with $\|\underline{Q}\|_1 \le \alpha$ and $\rho(\hat{\Phi}(\underline{Q})) \le \gamma_{opt} + \delta$. Note that

$$\begin{aligned} \gamma_{opt}^{\epsilon} - \gamma_{opt} &\le \rho\left(\epsilon\mathcal{I} + \hat{\Phi}\left(\underline{Q}\right)\right) - \inf_{\|Q\|_1 \le \alpha} \rho\left(\hat{\Phi}(Q)\right) \\ &\le \rho\left(\epsilon\mathcal{I} + \hat{\Phi}\left(\underline{Q}\right)\right) - \rho\left(\hat{\Phi}\left(\underline{Q}\right)\right) + \delta \\ &\le \epsilon^{1/n} f(n, M, \epsilon^*) + \delta. \end{aligned}$$

The last inequality follows from Lemma 1. Indeed, for any Q which satisfies $\|Q\|_1 \le \alpha$, $\Phi(Q)$ satisfies $\|\Phi(Q)\|_1 \le \|H\|_1 + \alpha\|U\|_1\|V\|_1$. Thus, $\hat{\Phi}(\underline{Q}) \in \{A | A \ge 0, \|A\|_1 \le M\}$ with $M = \|H\|_1 + \alpha\|U\|_1\|V\|_1$. Thus, all the assumptions of Lemma 1 are satisfied. Since δ is arbitrary, the result follows. ∎

In the following theorem, we show that the D variables in the optimization problem of γ_{opt}^{ϵ} can be restricted to a compact set, setting the stage for the use of relaxation methods in the next section.

Theorem 3: The following identity holds:

$$\gamma_{opt}^{\epsilon} = \inf_{\|Q\|_1 \le \alpha} \inf_{D \in \mathcal{D}_{feas}} \|D^{-1}(\epsilon\mathcal{I} + \mathcal{S}\Phi(Q))D\|_1$$

where $\mathcal{D}_{feas} = \{D \in \mathcal{D}\colon |D|_1 \le 1, d_j \ge \epsilon/(\|H\|_1 + \epsilon n - \epsilon)$ for all $j\}$.

Proof: From the definition in (3), we know that

$$\gamma_{opt}^{\epsilon} = \inf_{\|Q\|_1 \le \alpha} \inf_{\substack{D\in\mathcal{D} \\ |D|_1 = 1}} \|D^{-1}(\epsilon\mathcal{I} + \mathcal{S}\Phi(Q))D\|_1.$$

Note that an upper bound on γ_{opt}^{ϵ} is given by $\|H\|_1 + \epsilon n$ which is obtained by evaluating the objective with $Q = 0$ and $D = I$. Thus, if we define

$$\mathcal{D}' := \{D \in \mathcal{D}\colon |D|_1 = 1, |D^{-1}\epsilon\mathcal{I}D|_1 \le \|H\|_1 + \epsilon n\}$$

then we have

$$\gamma_{opt}^{\epsilon} = \inf_{\|Q\|_1 \le \alpha} \inf_{D\in\mathcal{D}'} \|D^{-1}(\epsilon\mathcal{I} + \mathcal{S}\Phi(Q))D\|_1.$$

Let $D \in \mathcal{D}'$. As D is diagonal and $|D|_1 = 1$, we know that there exists an index k such that $d_k = 1$. Also, $|D^{-1}\mathcal{I}D|_1 = \max_i \sum_{j=1}^{n}(d_j/d_i)$ and, therefore, the condition $\epsilon|D^{-1}\mathcal{I}D|_1 \le \|H\|_1 + \epsilon n$ implies that for all relevant i, $\epsilon\sum_{j=1}^{n}(d_j/d_i) \le \|H\|_1 + \epsilon n$. Thus for all i, $\epsilon(d_k/d_i) \le \|H\|_1 + \epsilon n - \epsilon$ from which it follows that for all i, $d_i \ge \epsilon/(\|H\|_1 + \epsilon n - \epsilon)$. Thus

$$\mathcal{D}' \subset \mathcal{D}_{feas}.$$

Since $\mathcal{D}' \subset \mathcal{D}_{feas} \subset \mathcal{D}$, we have

$$\gamma_{opt}^{\epsilon} = \inf_{\|Q\|_1 \le \alpha} \inf_{D \in \mathcal{D}_{feas}} \|D^{-1}(\epsilon \mathcal{I} + \mathcal{S}\Phi(Q))D\|_1.$$

This completes the proof. ∎

In the discussion above we have demonstrated that associated with the robust performance problem (2) we can define a perturbed problem (3) which attains an optimal value within any prespecified tolerance of the original problem's optimal value. Note that the perturbed problem (3) is a special case of the robust performance problem (2) with H and U in (2) replaced by $\epsilon\mathcal{I} + \mathcal{S}H$ and $\mathcal{S}U$ respectively. The main advantage of working with the perturbed problem (3) is that one can restrict the search for the optimal D to a compact set. With this in mind, in the subsequent discussion we shall focus solely on the optimization problem

$$\begin{aligned} \gamma_{opt} := \quad & \inf \|D^{-1}\Phi D\|_1 \\ \text{subject to} \quad & \\ & \|Q\|_1 \le \alpha \\ & \Phi = H - U * Q * V \\ & D \in \mathcal{D}_{feas} \end{aligned} \tag{5}$$

where D_{feas} will be redefined more generally to denote the compact set $D_{feas} := [L_d, U_d]$, where $L_d = (L_{d_1}, \ldots, L_{d_n})$ and $U_d = (U_{d_1}, \ldots, U_{d_n})$ are vectors in R^n satisfying $L_d \le U_d$.

B. Converging Upper and Lower Bounds

The optimization problem (5) which we seek to solve is an infinite-dimensional problem, having in general an infinite number of variables arising from Q and Φ. The approach we take in arriving at a solution consists of solving sequences of finite dimensional problems whose optimal solutions yield converging upper and lower bounds for γ_{opt}. This is not unlike the approach taken in solving the nominal ℓ_1 optimal control problem. Of course the main difference between the two is that here the finite dimensional problems that result are nonconvex owing to the D scales, whereas they are convex (and in fact linear) in the case of the nominal ℓ_1 optimal control.

To define the subproblems of interest, we start by defining for any $D \in \mathcal{D}_{feas}$

$$\begin{aligned} \gamma(D) := \quad & \inf \|D^{-1}\Phi D\|_1 \\ \text{subject to} \quad & \\ & \|Q\|_1 \le \alpha \\ & \Phi = H - U * Q * V \end{aligned} \tag{6}$$

$$\begin{aligned} \overline{\gamma}^N(D) := \quad & \inf \|D^{-1}\Phi D\|_1 \\ \text{subject to} \quad & \\ & \|Q\|_1 \le \alpha \\ & \Phi = H - U * Q * V \\ & Q(k) = 0 \text{ for all } k \ge N \end{aligned} \tag{7}$$

$$\begin{aligned} \underline{\gamma}^N(D) := \quad & \inf \|D^{-1}(P_N\Phi)D\|_1 \\ \text{subject to} \quad & \\ & \|Q\|_1 \le \alpha \\ & \Phi = H - U * Q * V. \end{aligned} \tag{8}$$

We will assume that H, U, and V have finite impulse response sequences. Note that the optimization problems arising in the definitions in (7) and (8) can be stated as finite dimensional linear programming problems (see [19]). Using the above definitions, we further define

$$\gamma_{opt} := \inf_{D \in \mathcal{D}_{feas}} \gamma(D), \tag{9}$$

$$\overline{\gamma}_{opt}^N := \inf_{D \in \mathcal{D}_{feas}} \overline{\gamma}^N(D), \tag{10}$$

$$\underline{\gamma}_{opt}^N := \inf_{D \in \mathcal{D}_{feas}} \underline{\gamma}^N(D). \tag{11}$$

The following results can be established.

Lemma 2: For any $D \in \mathcal{D}_{feas}$, we have

$$\overline{\gamma}^N(D) \searrow \gamma(D) \quad \text{and} \quad \underline{\gamma}^N(D) \nearrow \gamma(D), \text{as } N \to \infty.$$

Proof: See [19] and [20]. ∎

While the previous lemma relates the finite dimensional upper and lower bounds to $\gamma(D)$ for any fixed D, the next lemma relates the infimum over all feasible D scales of the finite dimensional upper and lower bounds to γ_{opt}.

Theorem 4: $\overline{\gamma}_{opt}^N \searrow \gamma_{opt}$ and $\underline{\gamma}_{opt}^N \nearrow \gamma_{opt}$ as $N \to \infty$.

Proof: Due to the continuity of the ℓ_1 norm, $\underline{\gamma}^N(D)$, $\overline{\gamma}^N(D)$, and $\gamma(D)$ are all continuous functions of D on the compact set $\mathcal{D}_{feas}$. Furthermore, from Lemma 2, $\overline{\gamma}^N(D) \searrow \gamma(D)$ and $\underline{\gamma}^N(D) \nearrow \gamma(D)$ for each $D \in \mathcal{D}_{feas}$. It follows (see [23, Th. 7.13]) that $\underline{\gamma}^N(D) \to \gamma(D)$ and $\overline{\gamma}^N(D) \to \gamma(D)$ *uniformly* on $\mathcal{D}_{feas}$. Hence, $\inf_{D\in\mathcal{D}_{feas}} \underline{\gamma}^N(D) \nearrow \inf_{D\in\mathcal{D}_{feas}} \gamma(D)$ and $\inf_{D\in\mathcal{D}_{feas}} \overline{\gamma}^N(D) \searrow \inf_{D\in\mathcal{D}_{feas}} \gamma(D)$ as $N \to \infty$. This completes the proof. ∎

The previous lemma suggests that if an effective solution procedure exists to solve for $\overline{\gamma}_{opt}^N$ and $\underline{\gamma}_{opt}^N$ then we can obtain converging upper and lower bounds to γ_{opt}. The rest of the development will be focused on deriving a method to obtain $\overline{\gamma}_{opt}^N$ for any given N. The methodology for finding $\underline{\gamma}_{opt}^N$ can be identically developed.

VI. Problem Solution

As mentioned earlier, the optimization associated with (7) can be cast into a finite dimensional linear program. More explicitly, we can write

$$\begin{aligned} \overline{\gamma}^N(D) := \; & \inf \gamma \\ & \text{subject to} \\ & \sum_j \sum_{k=1}^{N'} d_j \left(\Phi_{ij}^+(k) + \Phi_{ij}^-(k)\right) \le d_i \gamma \\ & \sum_j \sum_{k=1}^{N} \left(Q_{ij}^+(k) + Q_{ij}^-(k)\right) \le \alpha \\ & \mathcal{A}^N(Q^+ - Q^-) = \Phi^+ - \Phi^- \\ & \Phi^+ \ge 0,\ \Phi^- \ge 0,\ Q^+ \ge 0,\ Q^- \ge 0,\ \gamma \ge 0. \end{aligned} \tag{12}$$

This can be seen by standard arguments where positive variables x^+ and x^- are used to replace the variable x. The constraint $\Phi^+ - \Phi^- = \mathcal{A}^N(Q^+ - Q^-)$ results from the condition that $\Phi = H - U * Q * V$ whereas the constraint

$\sum_j \sum_{k=1}^{N} (Q_{ij}^{+}(k) + Q_{ij}^{-})(k) \le \alpha$ results from the condition $\|Q\|_1 \le \alpha$.

The problem to be solved for obtaining converging upper bounds is given by (10). Thus, the corresponding finite-dimensional optimization problem has the following structure:

$$\begin{aligned} \mu := \ & \inf \gamma \\ & \text{subject to} \\ & \sum_j d_j p_{ij} \le d_i \gamma \\ & p_{ij} = \sum_{k=1}^{N'} \Phi_{ij}^{+}(k) + \Phi_{ij}^{-}(k) \\ & \sum_j \sum_{k=1}^{N} \left(Q_{ij}^{+}(k) + Q_{ij}^{-}(k)\right) \le \alpha \\ & \mathcal{A}^N (Q^{+} - Q^{-}) = \Phi^{+} - \Phi^{-} \\ & \Phi^{+} \ge 0,\ \Phi^{-} \ge 0,\ Q^{+} \ge 0,\ Q^{-} \ge 0 \\ & L_{ij} \le p_{ij} \le U_{ij},\ L_{d_j} \le d_j \le U_{d_j}. \end{aligned} \tag{13}$$

Note that the upper and lower bounds on p_{ij} can be obtained from the one norm bound α on the Q parameter.

Next, we consider problem (13) posed on a smaller grid on the feasible region for the d variables, i.e.,

$$\begin{aligned} \mu(\ell_d, u_d) := \ & \inf \gamma \\ & \text{subject to} \\ & \sum_j d_j p_{ij} \le d_i \gamma \\ & p_{ij} = \sum_{k=1}^{N'} \Phi_{ij}^{+}(k) + \Phi_{ij}^{-}(k) \\ & \sum_j \sum_{k=1}^{N} \left(Q_{ij}^{+}(k) + Q_{ij}^{-}(k)\right) \le \alpha \\ & \mathcal{A}^N (Q^{+} - Q^{-}) = \Phi^{+} - \Phi^{-} \\ & \Phi^{+} \ge 0,\ \Phi^{-} \ge 0,\ Q^{+} \ge 0,\ Q^{-} \ge 0 \\ & L_{ij} \le p_{ij} \le U_{ij},\ \ell_{d_j} \le d_j \le u_{d_j} \end{aligned} \tag{14}$$

where $\ell_d = (\ell_{d_1}, \ldots, \ell_{d_n})$, $u_d = (u_{d_1}, \ldots, u_{d_n})$. For our purposes the set $[\ell_d, u_d]$ is a subset of the set $[L_d, U_d]$. Note that since the variables in the statement of $\mu(\ell_d, u_d)$ are confined to a smaller region than in μ, the above problem is being solved for a subproblem on a grid of the d variables' feasible region. We will also assume that an upper bound U_γ and a lower bound L_γ are available for the variable γ. Note that any upper and lower bounds on $\mu(\ell_d, u_d)$ yield particular values of L_γ and U_γ. For notational convenience we use p to denote the vector formed by stacking the variables p_{ij}. We define the set Ω to be the set of three-tuple's $(\gamma, p, D) \in R \times R^{n^2} \times \mathcal{D}$ such that the constraint $L_\gamma \le \gamma \le U_\gamma$ and all the constraints of (14) except the nonlinear constraint $\sum_j d_j p_{ij} \le d_i \gamma$ are satisfied. With this notation, we have

$$\begin{aligned} \mu(\ell_d, u_d) := \ & \inf \gamma \\ & \text{subject to} \\ & \sum_j d_j p_{ij} \le d_i \gamma \\ & (\gamma, p, D) \in \Omega. \end{aligned} \tag{15}$$

A. A Relaxation Result

In solving the subproblem stated above, a relaxation scheme (see [24]) will be employed. The following is a key result used for that purpose.

Lemma 3 [24]: If the variables $x_j \in R$ satisfy the conditions $\ell_j \le x_j \le u_j$ and $t_{ij} := x_i x_j$ then

$$t_{ij} \ge u_j x_i + u_i x_j - u_i u_j \tag{16}$$

$$t_{ij} \le \ell_j x_i + u_i x_j - u_i \ell_j \tag{17}$$

$$t_{ij} \le u_j x_i + \ell_i x_j - \ell_i u_j \tag{18}$$

$$t_{ij} \ge \ell_j x_i + \ell_i x_j - \ell_i \ell_j. \tag{19}$$

Furthermore, if variables $t_{ij} \in R$ satisfy (16)–(19) and x_k satisfy $\ell_k \le x_k \le u_k$ then

$$|t_{ij} - x_i x_j| \le \tfrac{1}{4}(u_i - \ell_i)(u_j - \ell_j). \tag{20}$$

Proof: $(\Rightarrow)$ Follows immediately from the conditions:

$$(u_i - x_i)(u_j - x_j) \ge 0$$

$$(u_i - x_i)(x_j - \ell_j) \ge 0$$

$$(x_i - \ell_i)(u_j - x_j) \ge 0$$

$$(x_i - \ell_i)(x_j - \ell_j) \ge 0.$$

$(\Leftarrow)$ Let $t_{ij} := x_i x_j$, where $\ell_i \le x_i \le u_i$ and $\ell_j \le x_j \le u_j$. Let S a subset of R^3 be the bounded region defined by inequalities (16)–(19). The absolute difference $|t_{ij} - x_i x_j| \le (1/4)(u_i - \ell_i)(u_j - \ell_j)$ for all $(t_{ij}, x_i, x_j) \in S$ if and only if $\max_S \{|t_{ij} - x_i x_j|\} = \max\{\max_S \{t_{ij} - x_i x_j\}, \max_S \{x_i x_j - t_{ij}\}\} \le (1/4)(u_i - \ell_i)(u_j - \ell_j)$.

Here we show that $\max_S \{t_{ij} - x_i x_j\} = (1/4)(u_i - \ell_i)(u_j - \ell_j)$. To do this, we utilize the Karush–Kuhn–Tucker necessary conditions for optimality: Let $\overline{x}$ be a feasible solution to the minimization problem

$$\min f(x), \qquad \text{subject to } \ g_i(x) \le 0,\ i = 1, \ldots, m \tag{21}$$

and let $I = \{i\colon\ g_i(\overline{x}) = 0\}$. If functions f and g_i are differentiable, $\nabla g_i(\overline{x})$ for $i \in I$ are linearly independent, and if $\overline{x}$ is a local solution to (21), then there exist scalars $\theta_i \ge 0$, $i = 1, \ldots, m$ such that $\nabla f(\overline{x}) + \sum_{i=1}^{m} \theta_i \nabla g_i(\overline{x}) = 0$ and $\theta_i g_i(\overline{x}) = 0$, $i = 1, \ldots, m$ (see [25]).

Assuming that $\ell_i \ne u_i$ and $\ell_j \ne u_j$, it is not difficult to verify that the gradients of the active constraints are linearly independent for any $x \in S$. Thus, any local solution must satisfy the optimality conditions. Transforming the problem $\max_S \{t_{ij} - x_i x_j\}$ into the form of (21) allows the optimality condition $\nabla f(x) + \sum_{i=1}^{m} \theta_i \nabla g_i(x) = 0$ to be expressed as

$$-1 - \theta_1 + \theta_2 + \theta_3 - \theta_4 = 0 \tag{22}$$

$$x_j + u_j\theta_1 - \ell_j\theta_2 - u_j\theta_3 + \ell_j\theta_4 = 0 \tag{23}$$

$$x_i + u_i\theta_1 - u_i\theta_2 - \ell_i\theta_3 + \ell_i\theta_4 = 0 \tag{24}$$

where $\theta_1, \ldots, \theta_4$ are the scalars corresponding to constraints (16)–(19), respectively. Since $\theta \geq 0$, (22) implies that $\theta_2 + \theta_3 \geq 1 > 0$. Constraints (17) and (18) are symmetric, so, without loss of generality, we can assume that $\theta_2 > 0$. Given this assumption, the complementary slackness condition $\theta_i g_i(x) = 0$ implies that any local solution must satisfy (17) at equality.

To complete the proof, we consider three possible cases.

1) If $\theta_1 = \theta_3 = \theta_4 = 0$, then (22)–(24) imply that $\theta_2 = 1$, $\overline{x}_i = u_i$ and $\overline{x}_j = \ell_j$. Solving (17) for $\overline{t}_{ij}$ yields $\overline{t}_{ij} = u_i\ell_j = \overline{x}_i\overline{x}_j$. Thus, the only point that satisfies optimality conditions yields an objective function value of $\overline{t}_{ij} - \overline{x}_i\overline{x}_j = 0$.
2) If $\theta_1 + \theta_4 > 0$, then either (16) or (19) (or both) must hold at equality. If (16) holds at equality along with (17), then $u_j\overline{x}_i - u_iu_j = \ell_j\overline{x}_i - u_i\ell_j$, or $\overline{x}_i = u_i$. This implies that $\overline{t}_{ij} = u_i\overline{x}_j = \overline{x}_i\overline{x}_j$. Similarly, if both (19) and (17) are tight, $\overline{x}_j = \ell_j$ and $\overline{t}_{ij} = \ell_j\overline{x}_i = \overline{x}_i\overline{x}_j$. Thus, if $\theta_1 + \theta_4 > 0$, any point that satisfies optimality conditions yields an objective function value of zero.
3) If $\theta_1 = \theta_4 = 0$ and $\theta_3 > 0$, then optimality conditions require that the values of θ_2, θ_3, $\overline{t}_{ij}$, $\overline{x}_i$ and $\overline{x}_j$ must satisfy (22)–(24) and must satisfy (17) and (18) at equality. It is easily verified that the solution to this system of equations is $\overline{x}_i = (\ell_i + u_i)/2$, $\overline{x}_j = (\ell_j + u_j)/2$, $\overline{t}_{ij} = (1/2)(\ell_i\ell_j + u_iu_j)$, and $\theta_2 = \theta_3 = 1/2$. The objective function value associated with this point is $t_{ij} - x_ix_j = (1/2)(\ell_i\ell_j + u_iu_j) - (1/4)(u_i - \ell_i)(u_j - \ell_j) = (1/4)(\ell_i\ell_j + u_iu_j - \ell_iu_j - u_i\ell_j) = (1/4)(u_i - \ell_i)(u_j - \ell_j)$. Thus, this point maximizes $t_{ij} - x_ix_j$ over S and $\max_S \{t_{ij} - x_ix_j\} = (1/4)(u_i - \ell_i)(u_j - \ell_j)$.

The proof that $\max_S \{x_ix_j - t_{ij}\} = (1/4)(u_i - \ell_i)(u_j - \ell_j)$ is analogous to the preceding proof and is left to the reader. ■

B. Solving the Relaxed Problem

We now setup the relaxed problems that will utilize the key result in the previous subsection. First define

$$\begin{aligned} W_{ij} := \{(p_{ij}, d_j, w_{ij}) \in R^3 |\ &\text{(16), (17), (18), (19) are} \\ &\text{satisfied with } t_{ij}, x_i, x_j, u_i, \ell_i, u_j, \ell_j \text{ replaced} \\ &\text{by } w_{ij}, p_{ij}, d_j, U_{ij}, L_{ij}, u_{d_j}, \ell_{d_j} \text{ respectively}\} \end{aligned}$$

$$\begin{aligned} W_i := \{(\gamma, d_i, w_i) \in R^3 |\ &\text{(16), (17), (18), (19) are} \\ &\text{satisfied with } t_{ij}, x_i, x_j, u_i, \ell_i, u_j, \ell_j \text{ replaced} \\ &\text{by } w_i, \gamma, d_i, U_\gamma, L_\gamma, u_{d_j}, \ell_{d_j}, \text{ respectively}\}. \end{aligned}$$

Using the above notation, consider the linear programming problem

$$\begin{aligned} \mu_R(\ell_d, u_d) = \ &\inf \gamma \\ &\text{subject to} \\ &\sum_j w_{ij} \leq w_i \\ &(p_{ij}, d_j, w_{ij}) \in W_{ij} \\ &(\gamma, d_i, w_i) \in W_i \\ &(\gamma, p, D) \in \Omega. \end{aligned} \tag{25}$$

One of the reasons behind the usefulness of this linear optimization problem is that $\mu_R(\ell_d, u_d)$ offers lower bound for $\mu(\ell_d, u_d)$. Indeed, we have the following.

Lemma 4: If the problem $\mu_R(\ell_d, u_d)$ is infeasible then so is the problem $\mu(\ell_d, u_d)$. If $\mu_R(\ell_d, u_d)$ is feasible then $\mu_R(\ell_d, u_d) \leq \mu(\ell_d, u_d)$.

Proof: It follows from Lemma 3 that:

$$\begin{aligned} \mu(\ell_d, u_d) = \ &\inf \gamma \\ &\text{subject to} \\ &w_{ij} = d_jp_{ij}, \ w_i = d_i\gamma \\ &\sum_j w_{ij} \leq w_i \\ &(p_{ij}, d_j, w_{ij}) \in W_{ij} \\ &(\gamma, d_i, w_i) \in W_i \\ &(\gamma, p, D) \in \Omega. \end{aligned} \tag{26}$$

Note that in the above equation all the constraints are linear except for the constraints $w_{ij} = p_{ij}d_j$ and $w_i = \gamma d_i$. Both assertions in the statement of the lemma follow from the fact that the optimization problem (25) is a *relaxation* of the above problem, i.e., the constraints in (26) contain all the constraints in (25). ■

In addition to providing a lower bound to $\mu(\ell_d, u_d)$ [problem (14)] as seen above, the solution to the optimization problem in $\mu_R(\ell_d, u_d)$ also offers an upper bound to μ [problem (13)]. We establish the following result, which relates the upper and lower bounds for μ and sets the stage for the global algorithm in the next subsection.

Theorem 5: Consider $\mu(\ell_d, u_d)$ defined by problem (14). Let $\mu_R(\ell_d, u_d)$ be the optimal objective of the relaxed problem as defined by (25). Suppose

$$(\gamma^*, p_{ij}^*, d_j^*, \Phi^{+*}, \Phi^{-*}, Q^{+*}, Q^{-*}, w_{ij}^*, w_i^*)$$

is the optimal solution of the relaxed problem (25) with $\gamma^* = \mu_R(\ell_d, u_d)$. Then

$$(\mu_F, p_{ij}^*, d_j^*, \Phi^{+*}, \Phi^{-*}, Q^{+*}, Q^{-*})$$

where

$$\mu_F = \mu_F(\ell_d, u_d) := \max_i \sum_j (p_{ij}^* d_j^*)/d_i^* \tag{27}$$

is feasible for $\mu(\ell_d, u_d)$ [as defined by (14)]. Furthermore

$$\mu_R(\ell_d, u_d) \leq \mu(\ell_d, u_d) \leq \mu_F(\ell_d, u_d) \tag{28}$$

and

$$\mu_F(\ell_d, u_d) - \mu_R(\ell_d, u_d) \leq C|u_d - \ell_d|_\infty \tag{29}$$

where

$$C = \frac{\frac{1}{4}\sum_j \max_i \{(U_{ij} - L_{ij})\} + \frac{1}{4}(U_\gamma - L_\gamma)}{L_{d_m}}$$

where $L_{d_m} = \min_{1\leq i\leq n} L_{d_i}$. Moreover

$$\overline{\gamma}(\ell_d, u_d) := \inf_{D\in\mathcal{D}} \left\{ \max_{1\leq i\leq n} \sum_{j=1}^{n} \frac{p_{ij}^* d_j}{d_i} \right\} \geq \mu. \tag{30}$$

Proof: It can be easily established that

$$\left(\mu_F,\, p_{ij}^*,\, d_j^*,\, \Phi^{+*},\, \Phi^{-*},\, Q^{+*},\, Q^{-*}\right)$$

is feasible for problem (14). This together with Lemma 4 shows that $\mu_F(\ell_d, u_d)$ is an upper bound on $\mu(\ell_d, u_d)$. This proves (28). Now, we will establish (29). Let i_0 be an index such that

$$\mu_F(\ell_d, u_d) = \max_i \sum_j (p_{ij}^* d_j^*)/d_i^* = \sum_j (p_{i_0 j}^* d_j^*)/d_{i_0}^*.$$

From Lemma 3 it follows that:

$$|w_{ij}^* - p_{ij}^* d_j^*| \leq \tfrac{1}{4}(U_{ij} - L_{ij})\,|u_d - \ell_d|_\infty,$$

and

$$|w_i^* - \gamma^* d_i^*| \leq \tfrac{1}{4}(U_\gamma - L_\gamma)\,|u_d - \ell_d|_\infty.$$

Thus, we have that

$$\begin{aligned}
&\mu_F(\ell_d, u_d) d_{i_0}^* - \gamma^* d_{i_0}^* \\
&\quad = \sum_j p_{i_0 j}^* d_j^* - \gamma^* d_{i_0}^* \\
&\quad = \left(\sum_j p_{i_0 j}^* d_j^* - \sum_j w_{i_0 j}^*\right) \\
&\qquad + \left(\sum_j w_{i_0 j}^* - w_{i_0}^*\right) + (w_{i_0}^* - \gamma^* d_{i_0}^*) \\
&\quad \leq \left(\tfrac{1}{4}\sum_j (U_{i_0 j} - L_{i_0 j})\,|u_d - \ell_d|_\infty\right) \\
&\qquad + \left(\tfrac{1}{4}(U_\gamma - L_\gamma)\,|u_d - \ell_d|_\infty\right).
\end{aligned}$$

This implies that

$$\mu_F(\ell_d, u_d) - \gamma^* \leq \frac{\frac{1}{4}\sum_j (U_{i_0 j} - L_{i_0 j})\,|u_d - \ell_d|_\infty}{d_{i_0}} + \frac{\frac{1}{4}(U_\gamma - L_\gamma)\,|u_d - \ell_d|_\infty}{d_{i_0}}.$$

As $d_{i_0}^* \geq L_{d_m}$ and $\gamma^* = \mu_R(\ell_d, u_d)$ it follows that:

$$\mu_F(\ell_d, u_d) - \mu_R(\ell_d, u_d) \leq |u_d - \ell_d|_\infty C$$

with

$$C = \frac{\frac{1}{4}\sum_j \max_i \{(U_{ij} - L_{ij})\} + \frac{1}{4}(U_\gamma - L_\gamma)}{L_{d_m}}.$$

The proof of the last assertion follows from the fact that p_{ij}^* are feasible for (13). This completes the proof. ∎

C. A Branch and Bound Algorithm

From the development in the previous section, it may be seen that $\mu_R(\ell_d, u_d)$ is a lower bound to $\mu(\ell_d, u_d)$ and $\mu_F(\ell_d, u_d)$ is an upper bound to $\mu(\ell_d, u_d)$, that is

$$\mu_R(\ell_d, u_d) \leq \mu(\ell_d, u_d) \leq \mu_F(\ell_d, u_d).$$

Theorem 5 gives an estimate on how good an approximation is given by $\mu_F(\ell_d, u_d)$ and $\mu_R(\ell_d, u_d)$. Note that for any optimization on a smaller grid the optimal value $\mu(\ell_d, u_d)$ is an upper bound on μ [because the optimization for μ is on the entire d variable space whereas it is only on a portion of the variable space for $\mu(\ell_d, u_d)$]. Thus, the upper bound obtained on $\mu(\ell_d, u_d)$ is indeed an upper bound on μ.

Theorem 6: Let ℓ_d^k and u_d^k (enumerated by k in some finite set Λ) be vectors in R^n such that $\ell_d^k \leq u_d^k$, $|u_d^k - \ell_d^k|_\infty \leq \epsilon$, and $\bigcup_{k\in\Lambda}[\ell_d^k, u_d^k] = [L_d, U_d]$. Let $\overline{k} = \arg\{\min_{k\in\Lambda} \mu(\ell_d^k, u_d^k)\}$ where $\mu(\ell_d, u_d)$ is defined in (13). Then

$$\mu_R\left(\ell_d^{\overline{k}}, u_d^{\overline{k}}\right) \leq \mu \leq \mu_R\left(\ell_d^{\overline{k}}, u_d^{\overline{k}}\right) + \epsilon C \tag{31}$$

where

$$C = \frac{\frac{1}{4}\sum_j \max_i \{(U_{ij} - L_{ij})\} + \frac{1}{4}(U_\gamma - L_\gamma)}{L_{d_m}}.$$

Proof: It is clear that $\mu = \mu(\ell_d^{\overline{k}}, u_d^{\overline{k}})$. If $\mu_R(\ell_d^{\overline{k}}, u_d^{\overline{k}})$ denotes the relaxation associated with $\mu(\ell_d^{\overline{k}}, u_d^{\overline{k}})$, and if $\mu_F(\ell_d^{\overline{k}}, u_d^{\overline{k}})$ denotes the upper bound as obtained in Theorem 5, then it follows that:

$$\mu_R\left(\ell_d^{\overline{k}}, u_d^{\overline{k}}\right) \leq \mu \leq \mu_F\left(\ell_d^{\overline{k}}, u_d^{\overline{k}}\right).$$

Also, it follows from Theorem 5 that:

$$\mu_F\left(\ell_d^{\overline{k}}, u_d^{\overline{k}}\right) - \mu_R\left(\ell_d^{\overline{k}}, u_d^{\overline{k}}\right) \leq \epsilon C.$$

This completes the proof. ∎

Thus to obtain a value with an ϵ tolerance the number of problems to be solved is no larger than the order of $1/\epsilon^n$ (Typical performance is much better). Also for any region $[\ell_d, u_d]$, if $\mu_R(\ell_d, u_d)$ is greater than any upper bound on any other region then $[\ell_d, u_d]$ can be removed from the list of regions where an optimal d may lie. This can be used as the basis for a branch and bound algorithm whereby branches corresponding to regions in the d-parameter space can be pruned as soon as the lower bound obtained from solving the relaxed problem for these regions becomes larger than the best available global upper bound for the problem. If this does not happen, further gridding on that region is performed, and hence new branches are formed. The process continues until all the branches which have not been pruned have lower bounds equal to (or up to a given tolerance of) the best available upper bound, which also will be the global optimal value. Note that when considering a given region $[\ell_d, u_d]$ for elimination from further consideration having a good upper bound is desirable. Clearly, $\mu(\ell_d, u_d)$ is one such upper bound. However, a considerably better upper bound is provided by $\overline{\gamma}(\ell_d, u_d)$. In fact given any set of regions $\{[\ell_d^{k_i}, u_d^{k_i}]\}_{i=1,\ldots,N}$ over which the relaxed problem has been solved then a global upper bound is given by

$$\overline{\gamma} := \min_{i=1,\ldots,N} \overline{\gamma}(\ell_d^{k_i}, u_d^{k_i}).$$

We now present an algorithm to solve (13) described by (13). In the algorithm each node corresponds to a subproblem posed on a smaller grid [see Problem (14)]. Each node has the fields: ℓ_d, u_d, and lowerBound which are indicated by $node.\ell_d$, $node.u_d$ and $node.lowerBound$.

Algorithm to Solve (13)

- Initialize GUB to any available global upper bound.
- Initialize ActiveNodeList to have a node with its fields set as $lowerBound = 0$, $\ell_d = L_d$ and $u_d = U_d$. Initialize $bestD \in R^n$ to $(1, \ldots, 1)$.

Step 1

If ActiveNodeList is empty stop.

Step 2

For each active node if $node.lowerBound \geq GUB - \epsilon$ remove the node from ActiveNodeList.

Step 3

Identify the active node with the least lowerBound as the *parentNode*. Solve the relaxed problem $\mu_R(parentNode.\ell_d, parentNode.u_d)$ for the above identified node. Also obtain $\overline{\gamma}(parentNode.\ell_d, parentNode.u_d)$ as constructed in (27). If $\overline{\gamma}(parentNode.\ell_d, parentNode.u_d)) < GUB$ then update

$$GUB \leftarrow \min(GUB, \overline{\gamma}(parentNode.\ell_d, parentNode.u_d))$$

and $bestD$ corresponding to the solution of $\overline{\gamma}$ above. Update

$$parentNode.lowerBound \leftarrow \mu_R(parentNode.\ell_d, parentNode.u_d).$$

Check if $parentNode.lowerBound \geq GUB - \epsilon$. If it is remove *parentNode* from ActiveNodeList.

Otherwise, if $parentNode.lowerBound < GUB - \epsilon$ let *parentNode* have two children nodes which inherit the lowerBound of the *parentNode*. Assign the ℓ_d and the u_d values to the children node so that the union of regions specified by the childrens' ℓ_d and u_d values equal the region $[parentNode.\ell_d, parentNode.u_d]$. Add the two children nodes to the ActiveNodeList and remove the *parentNode* from the ActiveNodeList.

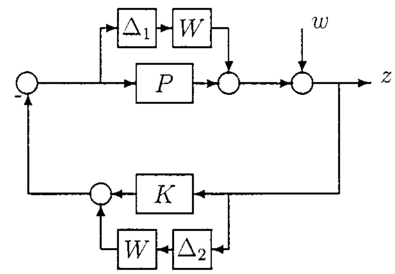

Fig. 3. Example system.

Step 4

Go to Step 1.

VII. Robust Control Example

Consider the system in Fig. 3 where

$$P = \frac{z+2}{z^2(z-0.5)(z+0.5)}$$

and

$$W = \frac{0.02z^2 - 0.04z + 0.02}{z^2 + 1.56z + 0.64}.$$

The objective is to design (if possible) a controller that makes the worst-case norm of the mapping from w to z less than one. This can posed as a robust stability problem by connecting z to w through a fictitious perturbation whose norm is allowed to be less than or equal to one, and then designing (if possible) a controller that makes the resulting system with three perturbation blocks robustly stable. The synthesis problem that must be solved to achieve that is:

$$\inf_{D \in \mathcal{D}} \inf_{Q \in \ell_1} \left\| D^{-1}\Phi(Q)D \right\|_1 =: \gamma_* \tag{32}$$

where Φ is the impulse response matrix of the system "seen" by the three uncertainty blocks.

Since both D and D^{-1} appear in the objective, one of the components of D may be fixed to one, and the search in the D parameter space becomes a two dimensional search. In this case we used $D = \mathrm{diag}(d_1, d_2, 1)$. The lower and upper bounds for the variables d_1 and d_2 were taken to be $L_d = (0.01, 0.01)$ and $U_d = (100, 100)$ respectively. It was determined that finite dimensional approximations corresponding to $N = 25$ were sufficiently close to the infinite dimensional optimal. Next, we present the results for solving Problem (12) with $N = 25$ using our proposed algorithm for finding the global optimal solution.

When using a tolerance of $\epsilon = 0.01$, the globally optimal value of the objective function was found after 71 iterations. Its value was found to be between 2.9370 and $2.9370 - \epsilon$, with $D = \mathrm{diag}(1.0, 0.37, 1.0)$ achieving 2.9370. After only four iterations a D scale that achieves an objective value of 2.939 was obtained. After 24 iterations a D scale achieving a value of 2.9372 was obtained. Clearly, either of these solutions would be considered close to the global optimal. Additional iterations (up to a total of 71) were needed only to verify that 2.9370 was indeed within ϵ away from the true global optimal value.

TABLE I
PROGRESS OF THE ALGORITHM FOR THE FIRST 35 ITERATIONS

It.	ℓ_d	u_d	$\mu_R(\ell_d, u_d)$	$\overline{\gamma}$	Prune
1	(0.010, 0.010)	(100.0, 100.0)	0.9353	8.3380	
2	(0.010, 1.000)	(100.0, 100.0)	1.5892	8.3380	
3	(0.010, 0.010)	(100.0, 1.000)	0.9353	8.3380	
4	(1.000, 0.010)	(100.0, 1.000)	2.5500	2.9390	
5	(0.010, 0.010)	(1.000, 1.000)	0.9354	2.9390	
6	(0.100, 0.010)	(1.000, 1.000)	1.3291	2.9390	
7	(0.010, 0.010)	(0.100, 1.000)	9.3534	2.9390	√
8	(0.100, 0.100)	(1.000, 1.000)	1.3978	2.9390	
9	(0.100, 0.010)	(1.000, 0.100)	8.0749	2.9390	√
10	(0.100, 0.316)	(1.000, 1.000)	1.5566	2.9390	
11	(0.100, 0.100)	(1.000, 0.316)	2.6425	2.9390	
12	(0.316, 0.316)	(1.000, 1.000)	1.9985	2.9390	
13	(0.100, 0.316)	(0.316, 1.000)	4.9223	2.9390	√
14	(1.000, 1.000)	(100.0, 100.0)	3.5647	2.9390	√
15	(0.010, 1.000)	(1.000, 100.0)	1.5893	2.9390	
16	(0.100, 1.000)	(1.000, 100.0)	2.0416	2.9390	
17	(0.010, 1.000)	(0.100, 100.0)	15.8922	2.9390	√
18	(0.562, 0.316)	(1.000, 1.000)	2.3673	2.9390	
19	(0.316, 0.316)	(0.562, 1.000)	3.5539	2.9390	√
20	(0.316, 1.000)	(1.000, 100.0)	2.5646	2.9390	
21	(0.100, 1.000)	(0.316, 100.0)	6.4558	2.9390	√
22	(0.750, 0.316)	(1.000, 1.000)	2.5963	2.9390	
23	(0.562, 0.316)	(0.750, 1.000)	3.1568	2.9390	√
24	(1.000, 0.100)	(100.0, 1.000)	2.6452	2.9372	
25	(1.000, 0.010)	(100.0, 0.100)	10.8202	2.9372	√
26	(0.562, 1.000)	(1.000, 100.0)	2.9784	2.9372	√
27	(0.316, 1.000)	(0.562, 100.0)	4.5606	2.9372	√
28	(0.866, 0.316)	(1.000, 1.000)	2.7289	2.9372	
29	(0.750, 0.316)	(0.866, 1.000)	2.9982	2.9372	√
30	(0.316, 0.100)	(1.000, 0.316)	3.0447	2.9372	√
31	(0.100, 0.100)	(0.316, 0.316)	4.4200	2.9372	√
32	(1.000, 0.316)	(100.0, 1.000)	2.8728	2.9372	
33	(1.000, 0.100)	(100.0, 0.316)	3.4345	2.9372	√
34	(0.931, 0.316)	(1.000, 1.000)	2.7989	2.9372	
35	(0.866, 0.316)	(0.931, 1.000)	2.9325	2.9372	√
⋮	⋮	⋮	⋮	⋮	

Table I shows the progress of the search for the first 35 iterations. Column 1 shows the iteration number; column 2 and column 3 show ℓ_d and u_d and thus define the region $[\ell_d, u_d]$ over which the relaxed problem for the give iteration is solved. Column 4 shows $\mu_R(\ell_d, u_d)$, the optimal solution of the relaxed problem, which serves as a lower bound for the optimal solution over the region $[\ell_d, u_d]$. Column 5 shows $\overline{\gamma}$ which is equal to the smallest value of $\overline{\gamma}(\ell_d, u_d)$ obtained over all prior iterations. At each iteration, $\overline{\gamma}$ serves as a global upper bound for the problem. A region $[\ell_d, u_d]$ is pruned when $\mu_R(\ell_d, u_d)$ for that region is larger than $\overline{\gamma} - \epsilon$ (here $\epsilon = 0.01$ was used). The iteration stops when the pruned regions cover the entire search space $[L_d, U_d]$. For this problem, $L_d = (0.01, 0.01)$ and $U_d = (100, 100)$ were used. Fig. 4 shows the regions in the search space that have been pruned after 35 iterations. Note that large regions in the search space have been eliminated in single iterations. Pruned regions are eliminated from further consideration. These are shown in gray in Fig. 4. Further branching on the white region allows us to prune the entire search space. At that point, the D scales corresponding to $\overline{\gamma}$ yield a ϵ-optimal global solution.

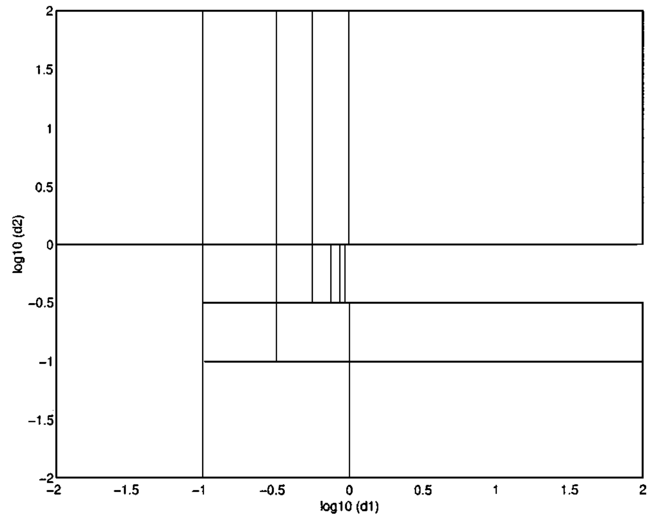

Fig. 4. Pruned regions (gray color) in the d parameter space after 35 iterations.

VIII. CONCLUSION

In this paper, we have proposed a solution to the ℓ_1 *robust synthesis problem*. The solution to this nonconvex infinite dimensional problem is obtained via relaxed problems which are linear programs.

Using the proposed solution, it is now possible to synthesize controllers that achieve global optimal performance against uncertainty which has prespecified structure, is uniformly norm-bounded and is time-varying and/or nonlinear. The performance measure considered is the infinity to infinity induced norm of a system transfer function.

REFERENCES

[1] J. C. Doyle, "Analysis of feedback systems with structured uncertainty," *Proc. IEE* , vol. 129-D, no. 6, pp. 242–250, Nov. 1982.
[2] J. C. Doyle, J. E. Wall, and G. Stein, "Performance and robustness analysis for structured uncertainty," in *Proc. 20th IEEE Conf. Decision Control*, 1982, pp. 629–636.
[3] M. G. Safonov, "Stability margins of diagonally perturbed multivariable feedback systems," *Proc. IEE*, pt. D, vol. 129, no. 6, pp. 251–256.
[4] A. Packard and J. C. Doyle, "The complex structured singular value," *Automatica, Special Issue Robust Control*, vol. 29, no. 1, pp. 71–109, 1993.
[5] R. R. E De Gaston and M. G. Safonov, "Exact calculation of the multiloop stability margin," *IEEE Trans. Automat. Contr.*, vol. 33, pp. 156–171, Feb. 1988.
[6] A. Sideris and R. R. E De Gaston, "Multivariable stability margin calculation with uncertain correlated parameters," in *Proc. IEEE Conf. Decision Control*, vol. 2, Athens, Greece, 1986, pp. 766–771.
[7] M. A. Dahleh and Y. Ohta, "A necessary and sufficient condition for robust bibo stability," *Syst. Control Lett.*, vol. 11, pp. 271–275, 1988.
[8] M. H. Khammash and J. B. Pearson, "Performance robustness of discrete-time systems with structured uncertainty," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 398–412, Apr. 1991.
[9] ——, "Analysis and design for robust performance with structured uncertainty," *Syst. Control Lett.*, vol. 20, no. 3, pp. 179–187, 1993.
[10] A. Megretsky, "Necessary and sufficient conditions of stability: A multiloop generalization of the circle criterion," *IEEE Trans. Automat. Contr.*, vol. 38, pp. 753–756, May 1993.

[11] J. Shamma, "Robust stability with time-varying structured uncertainty," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 714–724, Apr. 1994.
[12] A. Packard, K. Zhou, P. Pandey, J. Leonhardson, and G. Balas, "Optimal constant i/o similarity scaling for full-information and state-feedback control problems," *Syst. Control Lett.*, vol. 19, pp. 271–280, 1992.
[13] M. H. Khammash, "Synthesis of globally optimal controllers for robust performance with unstructured uncertainty," *IEEE Trans. Automat. Contr.*, vol. 41, pp. 189–198, Feb. 1996.
[14] M. V. Salapaka, M. Dahleh, A. Vicino, and A. Tesi, "Nominal $\mathcal{H}_2$ performance and ℓ_1 robust performance," in *Proc. IEEE Conference Decision Control*, Kobe, Japan, Dec. 1996, pp. 4034–4039.
[15] Y. Yamada and S. Hara, "Global optimization for h_∞ control with constant diagonal scaling," *IEEE Trans. Automat. Contr.*, vol. 43, pp. 191–203, Feb. 1999.
[16] J. G. VanAntwerp, R. D. Braatz, and N. V. Sahinidis, "Globally optimal robust process control," *J. Process Control*, vol. 9, pp. 375–383, 1999.
[17] D. C. Youla, H. A. Jabr, and J. J. Bongiorno, "Modern Wiener–Hopf design of optimal controllers—Part 2: The multivariable case," *IEEE Trans. Automat. Contr.*, vol. AC-21, pp. 319–338, Mar. 1976.
[18] M. A. Dahleh and I. J. Diaz-Bobillo, *Control of Uncertain Systems: A Linear Programming Approach*. Upper Saddle River, NJ: Prentice-Hall, 1995.
[19] M. H. Khammash, "A new approach to the solution of the ℓ_1 control problem: The scaled-q method," *IEEE Trans. Automat. Contr.*, vol. 45, pp. 180–187, Feb. 2000.
[20] M. V. Salapaka and M. Khammash, "Multi-objective mimo optimal control design without zero interpolation," in *Proc. Amer. Control Conf.*, vol. 1, Philadelphia, PA, June 1998, pp. 660–664.
[21] M. V. Salapaka, M. Khammash, and M. Dahleh, "Solution of mimo $\mathcal{H}_2/\ell_1$ problem without zero interpolation," *SIAM J. Control Optim.*, vol. 37, no. 6, pp. 1865–1873, 1999.
[22] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge University Press, 1985.
[23] W. Rudin, *Principles of Mathematical Analysis*. New York: McGraw-Hill, 1976.
[24] H. D. Sherali and C. H. Tuncbilek, "A global optimization algorithm for polynomial programming problems using a reformulation-linearization technique," *J. Global Optim.*, vol. 2, pp. 101–112, 1992.
[25] M. S. Bazarra, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms*. New York: Wiley, 1993.

Mustafa Khammash was born in Amman, Jordan, in 1964. He received the B.S. degree (*magna cum laude*) from Texas A&M University, College Station, TX, and the Ph.D. degree from Rice University, Houston, TX, in 1986 and 1990, respectively, both in electrical engineering.

In September 1990, he joined the Electrical and Computer Engineering Department at Iowa State University, Ames, IA, where he currently holds the position of Associate Professor. His research interests are in the areas of linear and nonlinear robust control, system theory, power systems stability and control, flight control, and feedback control in biological systems. He is an Associate Editor of *Systems and Control Letters*.

Dr. Khammash was the recipient of the National Science Foundation Young Investigator Award, the Iowa State University Foundation Early Achievement in Research and Scholarship Award, the Iowa State University College of Engineering Young Faculty Research Award and the Ralph Budd Best Engineering Ph.D. Thesis Award.

Murti V. Salapaka was born in Andhra Pradesh, India, in 1969. He received the B.Tech. degree in mechanical engineering from the Indian Institute of Technology, Madras, and the M.S. an Ph.D. degrees in mechanical engineering from the University of California, Santa Barbara, in 1991, 1993, and 1997, respectively.

Since 1997, he has been with the Electrical Engineering Department at Iowa State University, Ames, as an Assistant Professor.

Dr. Salapaka was the recipient of the 1997 National Science Foundation CAREER Award, and the 2001 Iowa State University Young Engineering Faculty Research Award.

Tim Van Voorhis received undergraduate degrees from Covenant College, Lookout Mountain, TN, and the Georgia Institute of Technology (Georgia Tech), Atlanta, in 1988. After working as a consultant, he returned to Georgia Tech and received the Ph.D. degree in 1997 from their School of Industrial and System Engineering.

Since then, he has worked in the Industrial and Manufacturing Systems Engineering Department at Iowa State University, Ames. His research area is global optimization, and he has worked on projects in sports scheduling, foundry scheduling, robust control, and power systems.