Computing Optimal Uncertainty Models from Frequency Domain Data[1]

Haitham Hindi   Chang-Yun Seong   Stephen Boyd

Abstract

Uncertainty models are an essential ingredient in robust control design. In addition, because of the tradeoff between uncertainty and performance, the uncertainty model should be as "tight" as possible. Given a set of multivariable frequency response measurements, we show that the computation of multivariable nonparametric uncertainty models which are consistent with the data (i.e. not invalidated), reduces to a linear matrix inequality (LMI) feasibility problem. Our method simultaneously searches for the responses of both the nominal system and the uncertainty weights that give an optimal uncertainty model. We then show that computing the optimal or least conservative model for the data can be done using semidefinite programming (SDP). Noise and fitting errors are explicitly factored into the computation using a bounded set approach. A state space uncertainty model can then be obtained from the optimal nonparametric model using frequency domain subspace identification techniques. The proposed technique is demonstrated on a generic MIMO example, where it outperforms the average-based approach by almost a factor of two (5dB), in the frequency range with largest uncertainty.

Notation: For a matrix $A \in \mathbf{C}^{m\times n}$, $\|A\|_2$ and $\bar{\sigma}(A)$ denote the maximum singular value. We use $[a_{rs}]$ to denote a matrix whose elements are given by $a_{rs},\ r = 1,\ldots,m,\ s = 1,\ldots,n$. The elementwise or Hadamard product of two matrices $A, B \in \mathbf{C}^{m\times n}$, $A \circ B = [a_{rs}\cdot b_{rs}]$. For simplicity of notation, when referring to the value of a transfer function at the frequency f_j, we will write $G(f_j)$ rather than $G(j2\pi f_j)$ or $G(e^{j2\pi f_j})$. If $\mathcal{S}$ is a subset of $\mathbf{C}^{m\times n}$, and A, B, C are matrices of appropriate dimensions, then $A + B\mathcal{S}C$ denotes the set $\{A + BXC \mid X \in \mathcal{S}\}$. We will use I to denote the identity matrix, and $\mathcal{B} = \{X \in \mathbf{C}^{m\times n} \mid \|X\| \le 1\}$ to denote a unit ball of matrices; the dimensions of I and $\mathcal{B}$ will be clear from the context. For matrices $A_1,\ldots,A_n$ we use $\mathrm{diag}(A_1,\ldots,A_n)$ to denote the associated block diagonal matrix.

[1]The authors are with the Information Systems Lab, Stanford University, hhindi@stanford.edu, seong@chinook.stanford.edu, boyd@stanford.edu. This work is supported in part by DARPA SEC F33615-99-C-3014.

1 Introduction

1.1 Basic Problem Statement:

This paper is concerned with the following problem, which arises frequently in real applications:

> Given a set of (possibly noisy, multivariable) frequency response measurements, compute a "tight" uncertainty model that is consistent with this data (i.e. not invalidated [19, 23]), which could be used for robust control design.

For example, these frequency response measurements could be several independent measurements of the frequency response matrix of a single multivariable plant G, taken at different operating conditions, or they could be frequency responses of several different plants G^i, taken for a multi-model control design.

We denote the data by $\{G^i(f_j)\}$, where each measurement $G^i(f_j)$ is a complex matrix, $i = 1,\ldots,M$ is the index of the entire frequency response measurement, and $j = 1,\ldots,N$ is the frequency index.

Our goal is to compute transfer matrices G_0, W_1 and W_2 for the *matrix additive uncertainty model* [25, 24, 28]:

$$\mathcal{M}(G_0, W_1, W_2)(f) \stackrel{\triangle}{=} \{G_0(f) + W_1(f)\Delta W_2(f) \mid \|\Delta\|_2 \le 1\}, \tag{1}$$

where G_0 is the nominal $m\times n$ transfer matrix, whose component transfer functions are denoted by $G_{0rs},\ r = 1,\ldots,m,\ s = 1,\ldots,n$, and W_1, W_2 are appropriate frequency weights. The uncertainty *radius* of model (1), at each frequency, is given by

$$\text{radius } \mathcal{M}(G_0, W_1, W_2)(f_j) \stackrel{\triangle}{=} \max_{\|\Delta\|_2\le 1} \|W_1(f_j)\Delta W_2(f_j)\|_2, = \|W_1(f_j)\|_2\|W_2(f_j)\|_2. \tag{2}$$

Thus the model (1) describes an uncertainty "tube" around the nominal model G_0, whose structure is given by the frequency weights W_i, and whose radius is $\bar{\sigma}(W_1(f))\bar{\sigma}(W_2(f))$. Because of the tradeoff between uncertainty and performance, an additional goal is that the model we compute should be as "tight" as possible, in the sense that its radius is minimized at each frequency f_j.

1.2 Nonparametric Computation

The focus in this paper is on the computation of optimal *nonparametric* uncertainty models from the given (possibly noisy) data. This means that we will compute the complex frequency responses of G_0, W_1 and W_2 at each frequency f_j. Our approach is based on techniques from robust convex optimization [5, 3, 9, 10].

Once the nonparametric model has been computed, one can either design optimal controllers directly from the nonparametric model, e.g. using the technique described in [27]; or one can use powerful subspace frequency domain identification methods [13, 17] to obtain a parametric state space model, which can then be used for standard state space control design.

We will compute MIMO matrix additive uncertainty models that are quite general, with no restrictions on the structure of G_0, W_1 and W_2. This generality allows us to compute matrix additive models that can capture certain LFT uncertainties as well.

Other approaches to uncertainty model synthesis are described in the literature, see for example [1, 6, 12, 25, 18, 26, 15]. Some of these techniques attempt to compute the parametric state space models of the nominal model G_0 and frequency weights W_1 and W_2 directly from the data. However, these methods are either time domain based, or they are not able compute G_0, W_1 and W_2 in such generality.

1.3 Connection to Other Uncertainty Models

- **LFT uncertainty models** of the form $\{G_0(f) + W_1(f)\Delta(I - W_3\Delta)^{-1}W_2(f)) \mid \|\Delta\|_2 \leq 1\}$ may appear more general than the matrix additive counterparts. However, it turns out that in the frequency domain approach adopted here, matrix additive uncertainty models are general enough to include uncertainty sets generated by LFTs with unstructured perturbations, as shown in [7, Lem. 10.2.4, p.347]. However, this equivalence is true only if we allow G_0, W_1 and W_2 in the matrix additive model to be arbitrary matrices with no constraints imposed on their structure, eg: symmetry, diagonal form, realness, etc. Hence, in this paper, we allow G_0, W_1 and W_2 to be general arbitrary complex matrices.
- **Elementwise (or Hadamard) additive models** [8] $\{G_0(f) + W_0(f) \circ \Delta \mid |\Delta_{rs}| \leq 1\}$, which specify, at each frequency f, a disk of radius $|W_{0rs}(f)|$ around each component $G_{0rs}(f)$ of the nominal transfer matrix $G_0(f)$, can be easily handled in our framework, by treating each component as a 1×1-matrix additive uncertainty model. This approach produces $m \times n$ frequency weights, and when used for robust control, would require solving an $(mn + 1)$-block robust performance μ-synthesis problem. In contrast, using the matrix additive uncertainty model could produce upto $m^2 + n^2$ weights, but has the benefit of leading to a two-block robust performance μ-synthesis problem [8, 25, 24].

2 Bounded Set Modelling of Errors

2.1 Prior Assumptions on Data

We make the following standard prior assumptions on the plant, the uncertainty and the data: The true underlying systems $\{G^i_{\text{true}}\}$, which generated the data $\{G^i(f_j)\}$, are stable linear time invariant (LTI) systems. Each data point $G^i(f_j)$ is corrupted by a bounded additive noise, which is known to lie in a given set. The uncertainty is LTI and stable, and can be captured by an additive uncertainty model.

In the literature, it is common to also assume that the impulse responses of each of the underlying true systems $\{G^i_{\text{true}}\}$ is bounded by some known exponentially decaying sequence [12, 2]. This ensures that the true underlying frequency response has a bounded derivative between the data points, and allows the bounding of the error in the frequencies between the data points. However, we will not use this explicitly in our development, since our objective is strictly to produce a model which is consistent with (i.e, not invalidated by) the given data.

2.2 Modeling Noise and Fitting Error

In practice, there will always be noise on the data. Furthermore, if the data is to be fitted with a state space model, this fit will not be achieved exactly. These effects should be considered by any serious method for computing uncertainty models. We use the set bounded noise approach to deal with these issues [1, 9, 10, 3, 14, 18].

We assume that each data point $G^i(f_j)$ is corrupted by a bounded additive noise, which is known to lie in a set $U^i_1(f_j)\mathcal{B}U^i_2(f_j)$, where $\mathcal{B}$ is the unit ball of matrices of appropriate dimension, and $\left(U^i_1(f_j), U^i_2(f_j)\right)$ are known complex matrices that specify the magnitude, and possibly directional information about the measurement noise. In other words, we assume that the true value of each frequency response measurement $G^i_{\text{true}}(f_j)$ satisfies

$$G^i_{\text{true}}(f_j) \quad \in \quad G^i(f_j) + U^i_1(f_j)\mathcal{B}U^i_2(f_j), \quad \forall\, i, j. \quad (3)$$

We assume that $\left(U^i_1(f_j), U^i_2(f_j)\right)$ are given at each frequency, for example from the experimental setup, or the instrument which takes the data. Often, the estimates of the noise matrices are given in terms of means and variances, see [16, 2]. As a simple example, suppose that the noise were known to have a variance of $\sigma^2 I$ at all frequencies. Then one could set

$U_1(f_j) = U_2(f_j) = \sqrt{3\sigma} I$, for all f_j. In any case, the estimation of these matrices is not the subject of this paper.

We will also add in margins for fitting error in our computation. As with the noise modeling, we will assume that matrices $\left(V_1^i(f_j), V_2^i(f_j)\right)$ are given, which reflect the expected size and direction of bounds on the fitting error.

Thus our technique will "inflate" each point data point $G^i(f_j)$ by an amount $U_1^i(f_j)\mathcal{B}U_2^i(f_j)$ to account for the noise, and an amount $V_1^i(f_j)\mathcal{B}V_2^i(f_j)$ to factor in some tolerance to fitting errors, and require that the nonparametric uncertainty model contain these inflated points. This is illustrated in fig. 1.

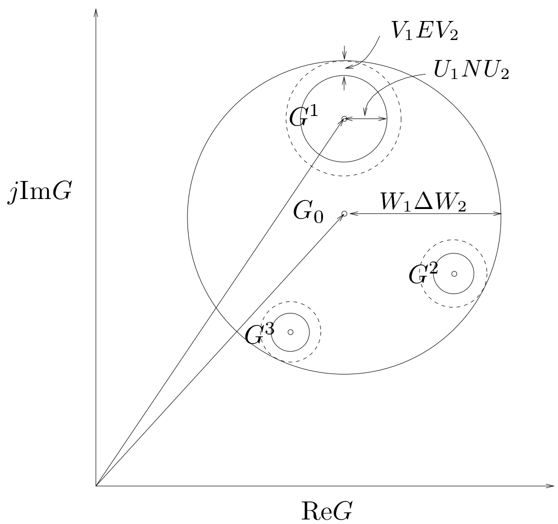

Figure 1: Robust covering of data. Here Δ, N and E are martices with unit norm.

Definition 1 *An uncertainty model $\mathcal{M}(G_0, W_1, W_2)$* robustly covers *the data $\{G^i(f_j)\}$, with respect to the matrices $\left(U_{1,k}^i(f_j), U_{2,k}^i(f_j)\right), k = 1, \ldots, K$, if the following condition holds:*

$$\begin{aligned} G^i(f_j) + U_{1,1}^i(f_j)\mathcal{B}U_{2,1}^i(f_j) + \cdots + U_{1,K}^i(f_j)\mathcal{B}U_{2,K}^i(f_j) \\ \subseteq \quad \mathcal{M}(G_0, W_1, W_2)(f_j), \quad \forall i,\ j. \end{aligned} \tag{4}$$

If $\left(U_{1,k}^i(f_j), U_{2,k}^i(f_j)\right), k = 1, \ldots, K$ are all zero, then we say the model merely covers *the data (nonrobustly).*

Definition 2 *A model $\mathcal{M}(G_0^\star, W_1^\star, W_2^\star)$, is said to be a* tight *or* optimal *robust uncertainty model for the data with respect to given matrices $\left(U_{1,k}^i(f_j), U_{2,k}^i(f_j)\right), k = 1, \ldots, K$, if its radius is less than or equal to the radius of any other model which robustly covers the data, with respect to the same matrices. Note that tight robust uncertainty models may not necessarily be unique.*

From the bounded noise model (3) it follows that if a model $\mathcal{M}(G_0, W_1, W_2)$ robustly covers the data $\{G^i(f_j)\}$, with respect to $\left(U_1^i(f_j), U_2^i(f_j)\right)$ only, then it covers the true uncorrupted data $\{G_{\mathrm{true}}^i(f_j)\}$. Violations within $V_1^i(f_j)\mathcal{B}V_2^i(f_j)$ are allowed, since they are within the prescribed tolerance.

The main goal of this paper is to compute an optimal nonparametric matrix additive uncertainty model $\mathcal{M}(G_0, W_1, W_2)$, which is not invalidated, *given* the data $\{G^i(f_j)\}$ and the noise bounds $\left(U_1^i(f_j), U_2^i(f_j)\right)$ and fitting error tolerances $\left(V_1^i(f_j), V_2^i(f_j)\right)$.

3 Computing Optimal Nonparametric Matrix Additive Uncertainty Models

In this section, we will show that optimal nonparametric matrix additive uncertainty models can be computed reliably and efficiently using tools from convex optimization, and robust control. Throughout this section, the frequency index f_j will be dropped to avoid clutter, since in computing the nonparametric model, all computations are done for each frequency individually.

The following two lemmas will be useful. The first comes from the model validation literature [19, 23], the second from robust convex optimization [3, 9, 10] and has its roots in early literature absolute stability theory and μ-analysis, see [11, 22] and the references therein.

Lemma 1 *Given the complex matrices A, B and C of compatible dimensions, there exists a solution Δ to the linear matrix equation $A = B\Delta C$ with $\bar{\sigma}(\Delta) \leq 1$ if and only if*

$$\begin{bmatrix} BB^* & A \\ A^* & C^*C \end{bmatrix} \geq 0. \tag{5}$$

Proof: This lemma is proved in [23, 7]. ■

Lemma 2 *Let $F = F^*$, L, R, D be* complex *matrices of appropriate sizes, with $\|D\| \leq 1$. Let Δ be a complex perturbation that has the block diagonal structure $\Delta = \mathrm{diag}(\Delta_1, \ldots, \Delta_N)$. Partition L, R, and D conformally with Δ as: $L = [L_1, \ldots, L_N]$, $R = [R_1^*, \ldots, R_N^*]^*$, $D = [D_1, \ldots, D_N]$. Then a sufficient condition for*

$$F + L\Delta(I - D\Delta)^{-1}R + R^*(I - D\Delta)^{-*}\Delta^* L^* \geq 0, \quad \forall \|\Delta\| \leq 1 \tag{6}$$

is that there exists a matrix $\Sigma = \mathrm{diag}(\sigma_1 I, \ldots, \sigma_N I)$, partitioned conformally with Δ, with σ_i real scalars, such that

$$\Sigma \geq 0, \qquad \begin{bmatrix} F - L\Sigma L^* & R^* - L\Sigma D^* \\ R - D\Sigma L^* & \Sigma - D\Sigma D^* \end{bmatrix} \geq 0. \tag{7}$$

Moreover, when $L_i \neq 0$ and $N \leq 2$, this condition is also necessary.

Proof: This lemma is proved in [9, 10], for the case of real matrices using the S-procedure. Since the S-procedure is necessary and sufficient for complex spaces [11] for both 1- and 2-block structured matrices, we obtain our result for $N \leq 2$. The condition $L_i \neq 0$ is a simple means for constraint qualification, see [11, 9]. ■

Note that the requirement that the matrix inequality in (6) should hold for all $\|\Delta\|_2 \leq 1$, makes this a *robust* semi-infinite feasibility problem. In general, such semi-infinite constraints can be difficult to handle. However, for the particular case in (6), the lemma shows that the condition is equivalent to an LMI condition for $N \leq 2$.

3.1 Analysis: Validation via LMIs

Applying condition (4) to the matrix additive model we see that $\mathcal{M}(G_0, W_1, W_2)$ *robustly covers* the data $\{G^i(f_j)\}$, with respect to $\big(U_1^i(f_j), U_2^i(f_j)\big)$ and $\big(V_1^i(f_j), V_2^i(f_j)\big)$, then for all the data points at frequency f_j, we must have

$$\begin{aligned} &\{G^i + U_1^i N^i U_2^i + V_1^i E^i V_2^i \mid \|N^i\|_2 \leq 1, \|E^i\|_2 \leq 1\} \\ &\subseteq \quad \{G_0 + W_1 \Delta W_2 \mid \|\Delta\|_2 \leq 1\}, \end{aligned} \tag{8}$$

We will now show that this condition is equivalent to a linear matrix inequality (LMI). The condition (8) is equivalent to requiring that for each frequency f_j: $\forall \,\|N^i\|_2 \leq 1,\ \forall \,\|E^i\|_2 \leq 1, \exists \,\|\Delta\|_2 \leq 1 \quad \text{s.t.}$

$$G^i - G_0 + U_1^i N^i U_2^i + V_1^i E^i V_2^i = W_1 \Delta W_2, \quad \forall i. \tag{9}$$

From Lemma 1, it follows that condition (9) can be written as: $\forall \,\|N^i\|_2 \leq 1,\ \forall \,\|E^i\|_2 \leq 1,$

$$\begin{bmatrix} T_1 & (G^i - G_0 + U_1^i N^i U_2^i + V_1^i E^i V_2^i) \\ (*)^* & T_2 \end{bmatrix} \geq 0 \quad \forall i, \tag{10}$$

where we have defined the variables $T_1 = W_1 W_1^*$ and $T_2 = W_2^* W_2$. Defining $L^i = [U_1^i \ V_1^i]$, and $R^i = [U_2^{i*} \ V_2^{i*}]^*$, (10) can be written as: $\forall \,\|N^i\|_2 \leq 1,\ \forall \,\|E^i\|_2 \leq 1,$

$$\begin{aligned} &\begin{bmatrix} T_1 & (G^i - G_0) \\ (*)^* & T_2 \end{bmatrix} + \begin{bmatrix} L^i \\ 0 \end{bmatrix} \begin{bmatrix} N^i & \\ & E^i \end{bmatrix} \begin{bmatrix} 0 & R^i \end{bmatrix} \\ &+ \begin{bmatrix} 0 \\ R^{i*} \end{bmatrix} \begin{bmatrix} N^{i*} & \\ & E^{i*} \end{bmatrix} \begin{bmatrix} L^{i*} & 0 \end{bmatrix} \geq 0 \quad \forall i. \end{aligned}$$

which , by Lemma 2, is equivalent to the following LMI condition in the real variable $\Sigma^i = \operatorname{diag}(\sigma_1^i I, \sigma_2^i I)$:

$$\Sigma^i \geq 0, \quad \begin{bmatrix} T_1 - L^i \Sigma L^{i*} & (G^i - G_0) & 0 \\ (G^i - G_0)^* & T_2 & R^{i*} \\ 0 & R^i & \Sigma^i \end{bmatrix} \geq 0 \quad \forall\, i. \tag{11}$$

Hence we arrive at the following result, which shows how the result above can be used to check that a *given* uncertainty model (G_0, W_1, W_2) robustly covers the data, i.e., *analysis* or model validation:

Theorem 1 *A given matrix additive uncertainty model* $\mathcal{M}(G_0, W_1, W_2)$ robustly covers *the data* $\{G^i(f_j)\}$*, with respect to the matrices* $\big(U_1^i(f_j), U_2^i(f_j)\big)$ *and* $\big(V_1^i(f_j), V_2^i(f_j)\big)$*, if and only if, at each frequency* f_j*, the LMI condition (11) holds in the real variables* $\Sigma^i = \operatorname{diag}(\sigma_1^i I, \sigma_2^i I)$*, with* G_0 *as in the model, and* $T_1 = W_1 W_1^*$*,* $T_2 = W_2^* W_2$*,* $L^i = [U_1^i \ V_1^i]$*, and* $R^i = [U_2^{i*} \ V_2^{i*}]^*$*.*

3.2 Synthesis: Computation via SDP

We now turn to the problem computing an optimal model, i.e., model *synthesis*. Specifically, given the data $\{G^i(f_j)\}$, we wish to find optimizers $(G_0^\star, W_1^\star, W_2^\star)$ which solve, at each f_j, the optimization problem:

$$\begin{array}{cl} \min\limits_{G_0, W_1, W_2} & \text{radius } \mathcal{M}(G_0, W_1, W_2) \\ \text{s.t.} & (8). \end{array} \tag{12}$$

Lemma 3 below, provides a means for finding a *feasible* robustly covering uncertainty model using convex optimization. Then problem of computing an *optimal* robustly covering model is solved in Theorem 2.

Lemma 3 *There exits a matrix additive uncertainty model* $\mathcal{M}(G_0, W_1, W_2)$ *which* robustly covers *the data* $\{G^i(f_j)\}$*, with respect to the matrices* $\big(U_1^i(f_j), U_2^i(f_j)\big)$ *and* $\big(V_1^i(f_j), V_2^i(f_j)\big)$*, if and only if, at each frequency* f_j*, the LMI condition (11) holds in the complex variables* G_0*,* $T_1 = T_1^*$*,* $T_2 = T_2^*$ *and real variable* $\Sigma^i = \operatorname{diag}(\sigma_1^i I, \sigma_2^i I)$*, where* $L^i = [U_1^i \ V_1^i]$*, and* $R^i = [U_2^{i*} \ V_2^{i*}]^*$*. Appropriate* W_1 *and* W_2 *are then given by the left and right Hermitian square roots of* $T_1^\star$ *and* $T_2^\star$*, respectively.*

Proof: Condition (11), when viewed as a constraint in the variables (G_0, T_1, T_2, Σ) rather than (G_0, W_1, W_2, Σ), (11) is actually an LMI. Furthermore, condition (11) ensures that T_1 and T_2 are both positive semidefinite. Hence, given any feasible T_1 and T_2, one can immediately compute the frequency weights as $W_1 = T_1^{1/2}$ and $W_2 = T_2^{*/2}$, i.e, as the left and right Hermitian square roots of T_1 and T_2, respectively. ■

Remark: Concerning the computation of the weights: First, note that we have chosen to use Hermitian square roots for uniqueness. Other choices of unique matrix square roots (e.g. Cholesky factors) should produce equally valid W_1 and W_2. Second, solving (11) will produce $W_1^\star,\ W_2^\star$ with no useful phase information. This is because of the squaring of $W_1,\ W_2$ in the definitions of $T_1,\ T_2$. However, an appropriate phase can be easily constructed using the Cepstrum technique [20] when $W_1^\star,\ W_2^\star$ are diagonal; if $W_1^\star,\ W_2^\star$ are not diagonal, the technique in [21] may be used.

Remark: Regarding the generality of Lemma 3, the LMI (11) in Lemma 3 searches over *all* possible models with (G_0, W_1, W_2) of compatible dimensions. In particular, it makes no assumptions on the structure of W_1 and W_2. Hence in general, the resulting uncertainty model will have W_1 and W_2 square. However, if the only feasible models have a particular structure, say W_1 tall and W_2 fat, then these *will* be found, since they are feasible solutions of LMI (11), with $T_1 = W_1 W_1^*$, $T_2 = W_2^* W_2$. In this case, the tall or fat structure of W_1 and W_2 will be manifested in rank deficiency of T_1 and T_2. Singular value decomposition can then be used to construct W_1 and W_2. On the other hand, it is possible to impose certain constraints on T_1 and T_2 which restrict the search in (11) to matrices of a certain structure. For example, constraining T_1 and T_2 to be (block) diagonal results in (block) diagonal weights W_1 and W_2.

Theorem 2 *An optimal matrix additive uncertainty model $\mathcal{M}(G_0^\star, W_1^\star, W_2^\star)$ which* robustly covers *the data $\{G^i(f_j)\}$ with respect to the matrices $\left(U_1^i(f_j), U_2^i(f_j)\right)$ and $\left(V_1^i(f_j), V_2^i(f_j)\right)$ can be computed by solving the following semidefinite program (SDP) in the complex variables G_0, $T_1 = T_1^*$, $T_2 = T_2^*$ and real variable $\Sigma^i = \mathrm{diag}(\sigma_1^i I, \sigma_2^i I)$, at each frequency f_j:*

$$\begin{array}{ll} \min\limits_{t, G_0, T_1, T_2, \Sigma^i} & t \\ \text{s.t.} & T_1 \leq tI,\ T_2 \leq tI,\ (11) \end{array} \tag{13}$$

Then $W_1^\star = (T_1^\star)^{1/2}$ and $W_2^\star = (T_2^\star)^{*/2}$ are the left and right Hermitian square roots of $T_1^\star$ and $T_2^\star$, respectively.

Proof: At each f_j, we must solve (12), which can be written as

$$\begin{array}{ll} \min\limits_{G_0, W_1, W_2} & \bar{\sigma}(W_1) \cdot \bar{\sigma}(W_2), \\ \text{s.t.} & (8). \end{array} \tag{14}$$

In Lemma 3, we showed that (8) was equivalent to the LMI condition (11). Using $T_1 = W_1 W_1^*$ and $T_2 = W_2^* W_2$ our optimization problem becomes

$$\begin{array}{ll} \min & \bar{\sigma}^{1/2}(T_1) \cdot \bar{\sigma}^{1/2}(T_2), \\ \text{s.t.} & (11) \end{array} \tag{15}$$

$$\Leftrightarrow \quad \begin{array}{ll} \min & t_1^{1/2} \cdot t_2^{1/2} \\ \text{s.t.} & \bar{\sigma}(T_1) \leq t_1,\ \bar{\sigma}(T_2) \leq t_2,\ (11). \end{array} \tag{16}$$

The proof is completed by noting that in the product $W_1 \Delta W_2$, we can without loss of generality always take $\bar{\sigma}(W_1) = \bar{\sigma}(W_2)$. ■

Remark: Equations (11) and (13) are written as complex LMIs. These can be easily converted to real LMIs, using the following fact: For any Hermitian matrix $M \in \mathbf{C}^{n \times n}$, the following condition holds:

$$M \leq 0 \quad \Leftrightarrow \quad \begin{bmatrix} \mathrm{Re}\ M & \mathrm{Im}\ M \\ -\mathrm{Im}\ M & \mathrm{Re}\ M \end{bmatrix} \leq 0.$$

Remark: The optimal $G_0^\star,\ T_1^\star,\ T_2^\star$ obtained from solving (13) may not be unique, due to the singular value constraints. Thus in practice, to insure uniqueness of the optimal values at the different frequencies f_j, one may perform two optimizations at each frequency
1. Solve (13) for the optimal radius $t^\star$
2. Pick a small $\epsilon > 0$ and solve the following SDP problem to obtain the unique ϵ-suboptimal $G_0,\ T_1,\ T_2$:

$$\begin{array}{cl} \min\limits_{G_0, T_1, T_2, S_0, S_1, S_2, \Sigma^i} & \mathrm{Tr}(S_0) + \mathrm{Tr}(S_1) + \mathrm{Tr}(S_2) \\ \text{s.t.} & T_1 \leq (1+\epsilon) t^\star I,\ T_2 \leq (1+\epsilon) t^\star I, \end{array}$$
$$\begin{bmatrix} I & G_0 \\ G_0^* & S_0 \end{bmatrix} \geq 0,\ \begin{bmatrix} I & T_1 \\ T_1^* & S_1 \end{bmatrix} \geq 0,\ \begin{bmatrix} I & T_2 \\ T_2^* & S_2 \end{bmatrix} \geq 0,$$
$$(11). \tag{17}$$

Note that the objective in (17) is equivalent to the sum of the squared Frobenius norms of G_0, T_1 and T_2, which is *strictly* convex in those variables, hence the optimal value of (17) is unique [4]. The procedure above essentially searches the $(1+\epsilon)t^\star$-suboptimal set for the least norm solution. Note that in (17) the $t^\star$ is constant, which is obtained in Step 1.

Remark: In some situations, one is required to generate uncertainty models which cover data that comes from simulations or from finite element models. In this case, the data $\{G^i(f_j)\}$ might contain no noise. Of course, one could analyze and compute uncertainty models by simply applying Theorems 1 and 2 with the matrices $\left(U_1^i(f_j), U_2^i(f_j)\right)$ and $\left(V_1^i(f_j), V_2^i(f_j)\right)$ set to zero. However, in this case, significant savings in computation are possible. For example, the elementwise case reduces to a second order cone program (SOCP), which can be solved an order of magnitude faster than a semidefinite program.

4 Average-Based Uncertainty Models

There is a simple two-step procedure for computing suboptimal elementwise or matrix additive uncertainty models, using the average of the data. The idea is, at each frequency f_j, to take the average of the data to be the “center” of the set, and then compute the weights with respect to that fixed center. The procedure goes as follows:

- Compute the nominal model as $G_0^{\text{avg}} \triangleq \frac{1}{M}\sum_{i=1}^{M} G^i$
- Compute the weights using

$$(W_1^{\text{avg}}, W_2^{\text{avg}}) \triangleq \arg\min\{(13) \mid G_0 = G_0^{\text{ave}}\} \qquad (18)$$

In the second step, we obtain W_i^{avg}'s by solving (13) but with G_0 *fixed* at G_0^{avg}.

As a consequence of the averaging, this technique has the nice property that it smooths out the data. This has the benefit of reducing sensitivity to output error noise in the measurements, and usually requiring a lower order state space fits than the optimal approaches proposed here.

However, the averaging technique has the disadvantage of being sensitive to outliers. Thus it can produce uncertainty models which are considerably more conservative than the corresponding optimal uncertainty models. Specifically, in the elementwise case, the uncertainty radii can be up to 6dB larger than the optimal radii; in the matrix case, the difference might be even larger.

5 Example

This section demonstrates the computation the tight uncertainty models using the matrix additive approach, for MIMO model of a high-speed positioning mechanism,

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] = \left[\begin{array}{cccc|cc} 0 & 1 & 0 & 0 & 0 & 0 \\ -\frac{k}{M} & -\frac{\gamma}{M} & \frac{k}{M} & \frac{\gamma}{M} & \frac{1}{M} & -\frac{1}{M} \\ 0 & 0 & 1 & 0 & 0 & 0 \\ \frac{k}{m} & \frac{\gamma}{m} & -\frac{k}{m} & -\frac{\gamma}{m} & 0 & \frac{1}{m} \\ \hline 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{array}\right],$$

where the mass, damping and stiffness parameters are given in Table 1, together with their uncertainties. The plant was also scaled at the input and output by diagonal scaling matrices $\text{diag}(6.5e3, 0.8e5 * \beta)$ and $\text{diag}(\alpha, 1)$. Twenty frequency responses were collected for different instances of the parameters, within their 10% range[1]. These twenty MIMO frequency responses were used as the data for computing matrix uncertainty models. The frequency weights were constrained to be diagonal, as explained in the remarks

[1]Of course, in this simple example, with real parametric uncertainty, a linear fractional transformation (LFT) parametric uncertainty model would be more appropriate. But the point is that in many real applications, such a parametric model is either not available or very difficult to construct.

	Nominal	% Uncertainty
M	1.0e-4	±10%
m	1.0e-5	±10%
γ	1.2e-4	±10%
k	1.0e-4	±10%
α	10dB	0
β	5dB	0

Table 1: Parameters and uncertainties.

following Lemma 3. For comparison, we also computed uncertainty models using the average-based technique of Section 4. We considered the following two scenarios.

In the first case the data was assumed to be clean and we computed an optimal matrix additive uncertainty model which (nonrobustly) covers the data. The results for clean data case are shown in Fig. 2. The corresponding weights are shown in Fig. 3 and the overall uncertainty radius is shown in Fig. 4. There is a significant difference between the optimal model and the average-based model, especially in the frequency range where the uncertainty is large, where the radius of the optimal model is 1-5dB tighter than that of the average-based model. (Recall that the matrix uncertainty weights in Fig. 3 have the units of the square-root of the radius, and should therefore not be compared directly to radius.)

In the second case, we compute a matrix additive uncertainty model which robustly covers the data, with respect to the matrices $V_1^i(f_j) = V_2^i(f_j) = 0.75I, \forall\, i, j$ (in other words, we have added in a $(0.75)^2$-radius margin for fitting error, at all frequencies and data points). The results for robust covering are shown in Fig. 5, Fig. 6, and Fig. 7. The results are similar to the clean data case in the regions where the frequency responses are large. However, as expected, now the radius is increased by the tolerance value of $(0.75)^2$, and similarly for the weights. This difference is most pronounced around the frequency $10^{4.3}$rad/sec, where the radius is comparable to the fitting tolerance.

It is interesting to observe the low frequency roll off of the $G_{0,12}$ and $G_{0,22}$ in the case of the optimal. We suspect that this is because at low frequencies, the plant becomes almost a rank-one plant

$$G_0 \approx \begin{bmatrix} G_{0,11} \\ G_{0,21} \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix},$$

and the perturbations are large enough that the data from the $[G_{0,12}G_{0,22}]^T$ part of the transfer function is covered automatically by the uncertainty around $[G_{0,11}G_{0,21}]^T$.

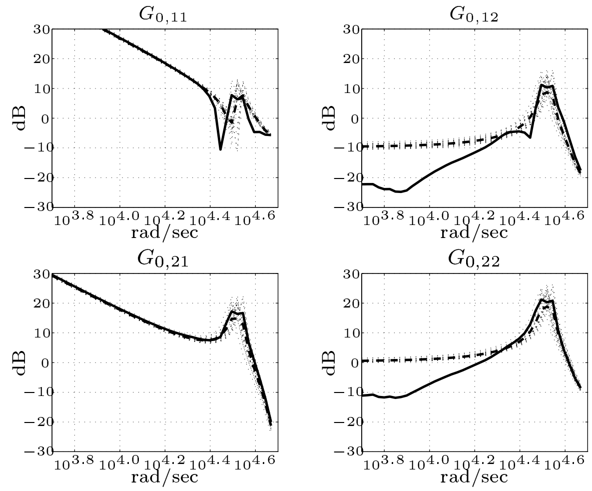

Figure 2: Nonrobust covering: nonparametric nominal models SDP (—), AVG (− −), and Measurements (· · ·).

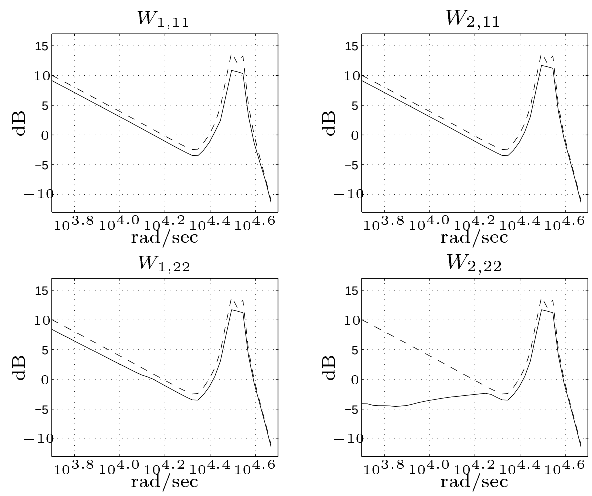

Figure 3: Nonrobust covering: nonparametric frequency weights: SDP (—) and AVG (− −).

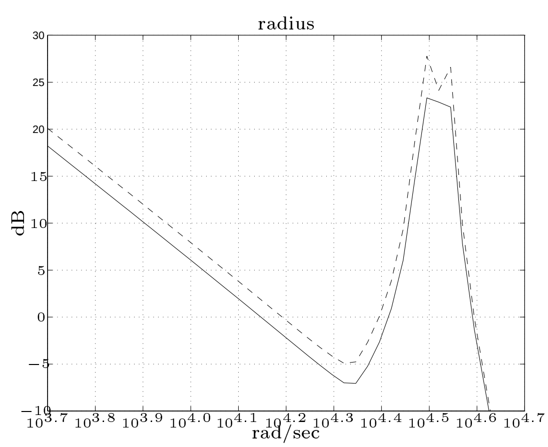

Figure 4: Nonrobust covering: matrix uncertainty radii: SDP (—) and AVG (− −).

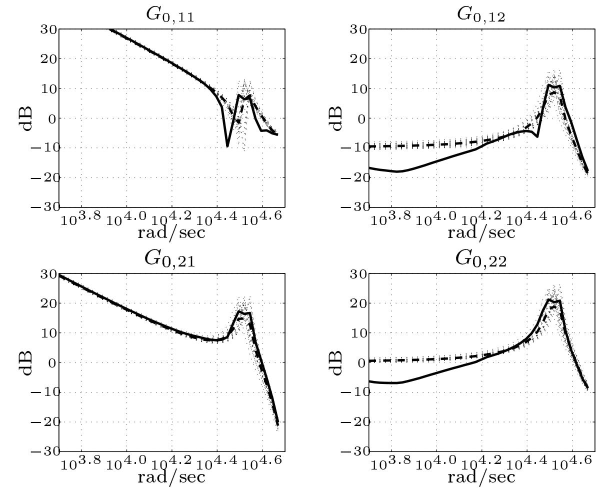

Figure 5: Robust covering: nonparametric nominal models SDP (—), AVG (− −), and Measurements (· · ·).

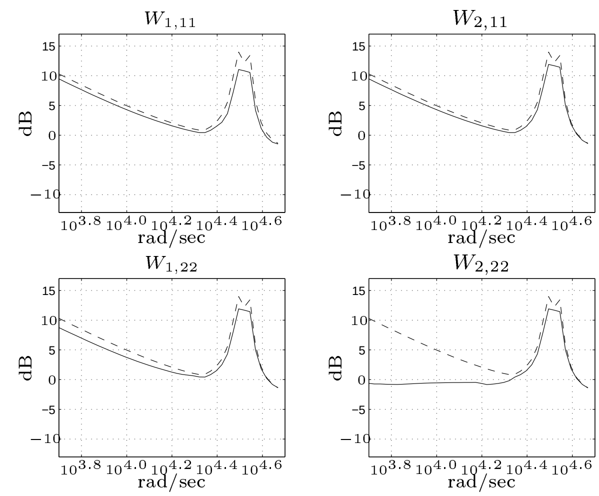

Figure 6: Robust covering: nonparametric frequency weights: SDP (—) and AVG (− −).

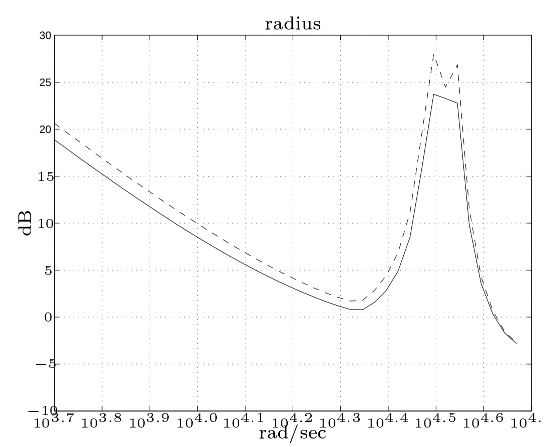

Figure 7: Robust covering: matrix uncertainty radii: SDP (—) and AVG (− −).

6 Conclusion

Given a set of multivariable frequency response measurements, we have shown that the computation of matrix additive (and elementwise additive) nonparametric uncertainty models which are consistent with the data (i.e. not invalidated), reduces to a linear matrix inequality (LMI) feasibility problem. This LMI condition can be used either to analyze a given uncertainty model, i.e. check its consistency with the data, or it can be used to synthesize an uncertainty model which is consistent with the data. We then showed that computing the optimal or least conservative model for the data can be done using semidefinite programming (SDP). Our method simultaneously searches for the responses of both the nominal system and the uncertainty weights that give an optimal uncertainty model. Noise and fitting errors were explicitly factored into the computation using a bounded set approach. The proposed technique was demonstrated on a generic MIMO example, where it outperformed the average-based approach by almost a factor of two (5dB), in the frequency range with largest uncertainty.

References

[1] T. Asai, S. Hara, and T. Iwasaki. Simultaneous parametric uncertainty modeling and robust control synthesis by LFT scaling. *Automatica*, 36:1457–1467, 2000.

[2] D. Bayard and Y. Yam. Frequency domain identification for robust control design. In R. Smith and M. Dahleh, editors, *The Modeling of Uncertainty in Control Systems*, pages 303–336. Springer-Verlag, 1994.

[3] A. Ben-Tal and A. Nemirovski. Robust convex optimization. *Mathematics of Operations Research*, 23(4):769–805, 1998.

[4] D. Bertsekas. *Nonlinear Programming*. Athena Scientific Press, 1999.

[5] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*, volume 15 of *Studies in Applied Mathematics*. SIAM, Philadelphia, PA, June 1994.

[6] M. E. Campbell and S. Grocott. Parametric uncertainty model for control design and analysis. *IEEE Transactions on Control Systems Technology*, 7:85 –96, 1999.

[7] J. Chen and G. Gu. *Control-Oriented System Identification, An $\mathcal{H}_\infty$ Approach*. Wiley-Interscience, 2000.

[8] R. de Callafon. *Feedback oriented identification for enhanced and robust control*. PhD thesis, Delft University of Technology, Mechanical Engineering Systems and Control Group, Netherlands, 1998.

[9] L. L. El Ghaoui and H. Lebret. Robust solutions to least-squares problems with uncertain data. *SIAM J. Matrix Anal. Appl.*, 18(4):1035–1064, 1997.

[10] L. L. El Ghaoui, F. Oustry, and H. Lebret. Robust solutions to uncertain semidefinite programs. *SIAM J. Optim.*, 9(1):33–52, 1998.

[11] A. L. Fradkov and V. A. Yakubovich. The $\mathcal{S}$-procedure and duality relations in nonconvex problems of quadratic programming. *Vestnik Leningrad Univ. Math.*, 6(1):101–109, 1979.

[12] R.G. Hakvoort and M.J. Van den Hof. Identification of probabilistic system uncertainty regions by explicit evaluation of bias and variance errors. *IEEE Transactions on Automatic Control*, 42:1516–1528, 1997.

[13] R.N. Jacques. *On-line system identification and control design for flexible structures*. PhD thesis, MIT, Department of Aeronautics and Astronautics, 1994.

[14] R.L. Kosut, M.K. Lau, and S.P. Boyd. Set-membership identification of systems with parametric and nonparametric uncertainty. *IEEE Transactions on Automatic Control*, 37(7):929 –941, 1992.

[15] T. L. Link, J. S. Lew, E. Garcia, and L. H. Keel. Interval model identification and robustness analysis for uncertain flexible structures. *IEEE Transactions on Control Systems Technology*, 4:411–418, 1996.

[16] L. Ljung. *System Identification, theory for the user*. Prentice Hall, 1987.

[17] T. McKelvey, H. Akcay, and L. Ljung. Subspace-based multivariable system identification from frequency response data. *IEEE Transactions on Automatic Control*, 41:960–, 1996.

[18] M. Milanese, J. Norton, H. Piet-Lahanier, and E. Walter, editors. *Bounding approaches to system identification*. Plenum Press, New York, 1996.

[19] M. Newlin and R. S. Smith. A generalization of the structured singular value and its application to model validation. *IEEE Transactions on Automatic Control*, 43(7):901–907, 1998.

[20] A. V. Oppenheim and R. W. Schafer. *Discrete-Time Signal Processing*. Prentice Hall Signal Processing Series. Prentice-Hall, Englewood Cliffs, New Jersey, 1989.

[21] P. Van Overschee, B. De Moor, W. Dehandschutter, and J. Swevers. A subspace algorithm for the identification of discrete time frequency domain power spectra. *Automatica*, 33:2147–, 1997.

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

[23] K. Poolla, P. Khargonekar, A. Tikku, J. Krause, and K. Nagpal. A time-domain approach to model validation. *IEEE Transactions on Automatic Control*, 39(5):951–959, 1994.

[24] S. Skogestad and I. Postlethwaite. *Multivariable Feedback Control*. Wiley, New York, 1996.

[25] R. Smith and M. Dahleh, editors. *The Modeling of Uncertainty in Control Systems*. Springer-Verlag, 1994.

[26] M. Taragna. Uncertainty model identification for H_∞ robust control. In *Proceedings on the 37th IEEE Conference on Decision & Control*, Tampa, Florida, USA, December 1998.

[27] K.-Y. Tsai and H. Hindi. DQIT: μ-synthesis without D-scale fitting. In *Proc. Am. Contr. Conf.*, pages 493–498, 2002.

[28] K. Zhou, J. Doyle, and K. Glover. *Robust and Optimal Control*. Prentice Hall, 1996.