Pergamon

PII: S0005-1098(98)00083-1

Automatica, Vol. 34, No. 11, pp. 1375–1389, 1998

Printed in Great Britain
0005-1098/98 $—see front matter

Mixed Time/Frequency-Domain Based Robust Identification*

P. A. PARRILO,† M. SZNAIER,‡ R. S. SÁNCHEZ PEÑA§ and T. INANC‡

A new robust identification framework that incorporates both time and frequency domain data is proposed. This framework avoids situations where a good data fit in one domain leads to poor fitting in the other.

Key Words—Robust identification; control oriented identification; interpolation; convex optimization; Linear Matrix Inequalities.

Abstract—In this paper we propose a new robust identification framework that combines both frequency and time-domain experimental data. The main result of the paper shows that the problem of obtaining a nominal model consistent with the experimental data and bounds on the identification error can be recast as a constrained finite-dimensional convex optimization problem that can be efficiently solved using Linear Matrix Inequalities techniques. This approach, based upon a generalized interpolation theory, contains as special cases the Carathéodory–Fejér (purely time-domain) and Nevanlinna–Pick (purely frequency-domain) problems. The proposed procedure interpolates the frequency and time domain experimental data while restricting the identified system to be in an *a priori* given class of models, resulting in a nominal model consistent with both sources of data. Thus, it is convergent and optimal up to a factor of two (with respect to central algorithms). © 1998 Elsevier Science Ltd. All rights reserved.

1. INTRODUCTION

During the past few years a large research effort has been devoted to the problem of developing deterministic identification procedures that, starting from experimental data and an *a priori* class of models, generate a nominal model and bounds on identification errors. These models and bounds can then be combined with standard robust control synthesis methods (such as $\mathcal{H}_\infty$, μ or ℓ_1) to obtain robust systems. This problem, termed the Robust Identification problem was originally posed in Helmicki *et al.* (1991) and has since attracted considerable attention (Chen *et al.*, 1992; Gu and Khargonekar, 1992; Hakvoort, 1992; Jacobson *et al.*, 1992; Mäkilä, 1991a, b, 1992; Mäkilä and Partington, 1992; Milanese, 1994; Parrilo *et al.*, 1994; Partington, 1992; Sánchez Peña and Galarza, 1994; Smith and Doyle, 1992).

Classical parameter identification methods (Ljung, 1987), consider mainly *a priori* sets of parametric models and are based on a stochastic approach. The outcome of these identification procedures consists of a nominal model and confidence bounds on the parameters. They have been used extensively in relation to adaptive control methods. On the other hand, robust identification is based on non-parametric mathematical models and a deterministic worst case criterion. The identification procedures use the experimental data (*a posteriori* information) and the *a priori* assumptions on the class of systems to be identified. They generate both a nominal model and *a worst case bound* over the set of systems considered, which fits the robust control framework. See Mäkilä *et al.* (1995) for an excellent survey of the robust identification framework.

Usual *a priori* assumptions are that systems under consideration are linear time invariant (LTI), exponentially stable, with known bounds on the frequency response magnitude, stability margin and measurement noise. The case where the experimental data available is generated by frequency-domain experiments leads to $\mathcal{H}_\infty$-based identification procedures. In this context the main effort has been directed towards establishing robust convergence of the algorithms and analyzing their untuned characteristics. A complete class of robustly convergent algorithms was presented in Gu and Khargonekar (1992). The case of strongly

*Received 14 December 1995; revised 12 June 1997; received in final form 28 April 1998. This paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor Matt James under the direction of Editor Tamer Başar. *Corresponding author* Professor M. Sznaier. Tel. +1 814 865-7667; E-mail msznaier@frodo.ee.psu.edu.

† Control & Dynamical Systems, California Institute of Technology, Pasadena, CA 91125, U.S.A.

‡ Dept. of Electrical Engineering, Penn State University, University Park, PA 16802, U.S.A.

§ Comisión Nacional de Actividades Espaciales (CONAE), Argentina and Depto. de Electrónica, Fac. de Ingeniería, Univ. de Buenos Aires, Argentina.

stabilizable systems has been considered in Mäkilä (1991a) and Mäkilä and Partington (1992).

On the other hand, the case where the experimental data available originates from time-domain experiments leads to ℓ_1 identification, addressed in Hakvoort (1992), Jacobson et al. (1992), Mäkilä (1991b, 1992), Milanese (1994) and Parrilo et al. (1994). Algorithms based on time series are strongly dependent on the input sequence (Mäkilä, 1991b). In fact it can be shown that there is no untuned algorithm capable of identifying a system using only impulse response measurements (Jacobson et al., 1992).

Finally, recent papers (Chen and Nett, 1995; Zhou and Kimura, 1993) proposed interpolatory algorithms that use data obtained from time domain experiments to generate a nominal model together with an $\mathscr{H}_\infty$ bound on the identification error.

In this paper we propose a new robust identification framework that takes into account both time and frequency domain experiments. With this approach, the problem where "good" frequency response fitting (small $\mathscr{H}_\infty$ error norm) leads to "poor" fitting in the time-domain is avoided. Additionally, from an information theoretic viewpoint, more experiments produce a smaller consistency set of indistinguishable models, and as a consequence a smaller worst case error. From a more practical standpoint, robust identification algorithms are applied in many cases to systems that may not be exactly LTI. Under these circumstances it is desirable to perform both time and frequency response experiments (Sánchez Peña and Galarza, 1994) to assess the validity of these assumptions. Thus, in these cases the proposed algorithm takes advantage of the additional data available. It is worth mentioning that the Chebyshev algorithm (Mäkilä et al., 1995) also allows one to consider time and frequency-domain data simultaneously.

The paper is motivated by our earlier results (Sánchez Peña and Sznaier, 1995) furnishing necessary and sufficient conditions for the consistency of mixed time/frequency experimental data for finite impulse response (FIR) systems and proposing an identification algorithm based upon solving a finite-dimensional convex optimization problem. In this paper we extend these results to the general case of infinite impulse response (IIR) systems.*

The main result of the paper shows that the problems of establishing consistency of the data and of obtaining a nominal model and bounds on the identification error can be recast as a constrained finite-dimensional convex optimization problem that can be efficiently solved using Linear Matrix Inequalities techniques. Additional results include an analysis of the conditioning of the problem as the amount of experimental data increases and an analysis of the effects of variations in the data points. Our approach, based upon a generalized Nevanlinna–Pick interpolation theory, includes as special cases the frequency based approach of Chen et al. (1992) and the time domain approach of Chen and Nett (1995) and Zhou and Kimura (1993).

The paper is organized as follows. In Section 2 we introduce a robust identification framework using both time and frequency experiments and some background material, including a generalized Nevanlinna–Pick theory developed in Rotstein (1996) that contains as special cases the classical Carathéodory–Fejér and Nevanlinna–Pick problems. Section 3 contains the main theoretical results. Here we show that the problems of establishing consistency of the experimental data and the *a priori* information and of determining a nominal model can be recast into a finite-dimensional Linear Matrix Inequality (LMI) optimization form. This optimization generates a model that interpolates the frequency domain data points and, at the same time, is consistent with the data obtained from time-domain experiments. Since the proposed algorithm is interpolatory, it is optimal up to a factor of two with respect to strongly optimal *central* algorithms (Mäkilä, 1991b, 1992), i.e., its worst case identification error is at most twice the minimal error over all possible experiments and algorithms. Moreover, it is convergent in the sense that the modelling error tends to zero as the information is completed. Section 4 deals with some computational considerations. Section 5 illustrates the results with two examples that highlight the importance of taking into account both frequency and time domain experimental data. Finally, Section 6 contains some concluding remarks and directions for future research.

2. PRELIMINARIES

2.1. *Notation*

$\mathscr{L}_\infty$ denotes the Lebesgue space of complex valued functions essentially bounded on the unit circle, equipped with the norm

$$\|G(z)\|_\infty \triangleq \operatorname{ess}\sup_{|z|=1} |G(z)|.$$

By $\mathscr{H}_\infty$ we denote the subspace of functions in $\mathscr{L}_\infty$ with a bounded analytic continuation inside the unit disk $\mathscr{D}$, equipped with the norm $\|G(z)\|_\infty \triangleq \operatorname{ess}\sup_{|z|<1} |G(z)|$. Also of interest is the space $\mathscr{H}_{\infty,\rho}$ of transfer functions in $\mathscr{H}_\infty$ which have bounded analytic continuation inside the disk of

*Preliminary versions of these results have appeared in Parrilo and Sánchez Peña (1995) and Parrilo et al. (1996).

radius $\rho > 1$, i.e. the space of exponentially stable systems with a stability margin of $(\rho - 1)$. When equipped with the norm $\|G(z)\|_{\infty,\rho} \triangleq \sup_{z<\rho}|G(z)|$, $\mathscr{H}_{\infty,\rho}$ becomes a Banach space. $\mathscr{B}\mathscr{H}_\infty \triangleq \{F \in \mathscr{H}_\infty, \|F_\infty\| \leq 1\}$ denotes the closed unit-ball in $\mathscr{H}_\infty$. Similarly $\mathscr{B}\mathscr{H}_{\infty,\rho}$ denotes the closed unit-ball in $\mathscr{H}_{\infty,\rho}$.

Given a vector $\mathbf{x} \in \mathbb{R}^n$ its infinity norm is defined as $\|\mathbf{x}_\infty\| \triangleq \max_i|\mathbf{x}_i|$. ℓ_1 denotes the space of absolutely summable sequences $h = \{h(i)\}$ equipped with the norm $\|h\|_{\ell_1} \triangleq \sum_{i=0}^{\infty}|h(i)| < \infty$. Similarly, ℓ_∞ denotes the space of bounded sequences $h = \{h(i)\}$ equipped with the norm $\|h\|_{\ell_\infty} \triangleq \sup_{i \geq 0}|h(i)| < \infty$, and $\ell_\infty(\varepsilon)$ denotes the subset $\{h \in \ell_\infty, |h(i)| \leq \varepsilon\}$. Given a sequence $h \in \ell_1$, its $\mathscr{Z}$ transform is defined* as $H(z) = \sum_{i=0}^{\infty} h(i)\, z^i$. It is a standard result that the series converges uniformly on the unit disk and that $\|H(z)\|_\infty \leq \|h\|_{\ell_1} < \infty$. In the sequel, by a slight abuse of notation, we shall use indistinctly both notations (transfer function or sequence of coefficients) to refer to the same object. We will also associate with the sequence h a vector $\mathbf{h}$ whose components are the first n elements of the sequence (n will be clear from the context).

In this paper we will consider both continuous and discrete time systems. To unify the treatment, the results in the sequel are stated for the discrete time case, but can easily be extended to the continuous time by defining $H(z) = H_c(\lambda(1-z)/(1+z))$, $\lambda > 0$, where H_c denotes the continuous time transfer function. Finally, for simplicity we consider SISO models, although all results can be applied to MIMO systems, following Chen *et al.* (1994).

2.2. *The robust identification framework*

In this paper we consider the case where the *a posteriori* experimental data originates from two different sources: (i) frequency and (ii) time domain experiments. The first type of information consists of a set of N_f samples of the frequency response of the system: $y^f(k) = \hat{h}(k) + \eta^f(k)$, $k = 0, \dots, N_f - 1$, where $\hat{h}(k) = H(e^{j\Omega_k})$, $k = 0, \dots, N_f - 1$, Ω_k denotes the sampling frequencies; and where $\eta^f(k)$ represents complex additive noise, bounded by ε_f (i.e. $|\eta^f(k)| < \varepsilon_f$).

The time domain data are the first N_t samples of the time response corresponding to a known but otherwise arbitrary input, also corrupted by additive noise $y^t(n) = (U\mathbf{h})(n) + \eta^t(n)$, $n = 0, \dots, N_t - 1$, where

$$U = \begin{bmatrix} u(0) & 0 & \dots & 0 \\ u(1) & u(0) & \dots & 0 \\ \vdots & \ddots & \ddots & 0 \\ u(N_t - 1) & \dots & u(1) & u(0) \end{bmatrix}$$

is the Toeplitz matrix corresponding to the input sequence and where the noise $\eta^t(n)$ is real and belongs to $\ell_\infty(\varepsilon_t)$. In the sequel, for notational simplicity we will collect the samples $y^f(k)$ and $y^t(n)$ in the vectors $\mathbf{y}^f \in \mathbb{C}^{N_f}$ and $\mathbf{y}^t \in \mathbb{R}^{N_t}$ respectively.

The *a priori* information available is that the system H under consideration belongs to the following classes of models:

(1)

$$H \in \mathscr{H}_\infty(\rho, K) \triangleq \{H \in \mathscr{H}_{\infty,\rho}\colon \|H\|_{\infty,\rho} \leq K\}$$

i.e. the class of models considered in the frequency domain corresponds to exponentially stable systems (finite or infinite dimensional) having a stability margin of $(\rho - 1)$ and a peak response to complex exponential inputs of K. Thus the impulse response of these systems satisfies the time-domain bound

$$|h(k)| \leq K\rho^{-k}. \tag{1}$$

(2) Additionally, the system H is known to belong to a class Φ of models satisfying a time-domain bound of the form

$$\Phi \triangleq \{h(\cdot) | \phi_l(k) \leq h(k) \leq \phi_u(k),\ k = 0, \dots, N_\phi - 1, N_\phi \text{ given}\}.$$

Note that this class includes the systems described by equation (1) in the special case when $\phi_l(k) = -K\rho^{-k}$ and $\phi_u(k) = K\rho^{-k}$.

To combine both classes of models we define the *a priori* set of systems

$$\mathscr{S} \triangleq \Phi \cap \mathscr{H}_\infty(\rho, K).$$

The *a priori* information we have considered simply adds to the usual $\mathscr{H}_\infty$ identification procedures a bound on the first N_ϕ samples of the impulse response. In other words, the time domain bound provided by the set Φ tightens the bound in equation (1) due to the set $\mathscr{H}_\infty(\rho, K)$. Trivially, if there is no time domain *a priori* information, the class of models $\mathscr{S}$ reduces to $\mathscr{H}_\infty(\rho, K)$.

To recap, the *a priori* information and the *a posteriori* experimental input data are:

$$\begin{aligned} \mathscr{S} &= \Phi \cap \mathscr{H}_\infty(\rho, K), (\rho > 1, K < \infty) \\ \mathscr{N}_f &= \ell_\infty(\varepsilon_f) = \{\eta^f \in \mathbb{C}^{N_f}, |\eta^f(k)| \leq \varepsilon_f\} \\ \mathscr{N}_t &= \ell_\infty(\varepsilon_t) = \{\eta^t \in \mathbb{R}^{N_t}, |\eta^t(k)| \leq \varepsilon_t\} \\ \mathbf{y}^f &= \{\hat{\mathbf{h}} + \eta^f \in \mathbb{C}^{N_f}\} \\ \mathbf{y}^t &= \{(U\mathbf{h}) + \eta^t \in \mathbb{R}^{N_t}\}. \end{aligned} \tag{2}$$

By using these definitions the robust identification problem with mixed data can be precisely stated as:

Problem 1 (*Mixed $\ell_1/\mathscr{H}_\infty$ robust identification problem*). Given the experiments $(\mathbf{y}^f, \mathbf{y}^t)$ and the *a priori*

*Note that this is the inverse of the usual $\mathscr{Z}$ transform. Therefore for causal, stable systems $H(z)$ is analytic in $|z| < 1$.

sets $(\mathcal{S}, \mathcal{N}_f, \mathcal{N}_t)$, determine:

(1) If the *a priori* and *a posteriori* information are consistent, i.e. the consistency set

$$\mathcal{S}(\mathbf{y}^f, \mathbf{y}^t) \triangleq \left\{ \mathbf{H} \in \mathcal{S} \,\middle|\, \begin{matrix} (\mathbf{y}^f - \hat{\mathbf{h}}) \in \mathcal{N}_f \\ (\mathbf{y}^t - U\mathbf{h}) \in \mathcal{N}_t \end{matrix} \right\} \quad (3)$$

is nonempty.

(2) If equation (3) holds, find a nominal model which belongs to the consistency set $\mathcal{S}(\mathbf{y}^f, \mathbf{y}^t)$, and an error bound.

2.3. *Generalized interpolation framework*

In this section we briefly present a generalized interpolation framework developed in Ball *et al.* (1990) and applied to $\mathcal{H}_\infty$ control in Rotstein (1996). This framework will be used in Section 3 to solve Problem 1.

Theorem 1. There exists a transfer function $F(z) \in \mathcal{BH}_\infty(\bar{\mathcal{B}}\mathcal{H}_\infty)$ such that

$$\sum_{z_0 \in D} \mathrm{Res}_{z=z_0} F(z) C_-(zI - A)^{-1} = C_+ \quad (4)$$

if and only if the following discrete time Lyapunov equation has a unique positive (semi)definite solution

$$M = A^*MA + C_-^*C_- - C_+^*C_+, \quad (5)$$

where A, C_- and C_+ are constant complex matrices of appropriate dimensions. If $M > 0$ then the solution $F(z)$ is non-unique and the set of solutions can be parameterized as a linear fractional transformation (LFT) in terms of $Q(z)$, an arbitrary element of $\bar{\mathcal{B}}\mathcal{H}_\infty$, as follows:

$$F(z) = \frac{T_{11}(z)Q(z) + T_{12}(z)}{T_{21}(z)Q(z) + T_{22}(z)}, \quad (6)$$

$$T(z) = \begin{bmatrix} T_{11}(z) & T_{12}(z) \\ T_{21}(z) & T_{22}(z) \end{bmatrix}, \quad (7)$$

where $T(z)$ is the J-lossless* matrix

$$T(z) \equiv \left[\begin{array}{c|c} A_T & B_T \\ \hline C_T & D_T \end{array} \right],$$

$$A_T = A,$$

$$B_T = M^{-1}(A^* - I)^{-1}[-C_+^* \quad C_-^*],$$

$$C_T = \begin{bmatrix} C_+ \\ C_- \end{bmatrix}(A - I),$$

$$D_T = I + \begin{bmatrix} C_+ \\ C_- \end{bmatrix} M^{-1}(A^* - I)^{-1}[-C_+^* \quad C_-^*].$$

*A transfer function $H(z)$ is said to be J-lossless if $H^*(1/z)\,JH(z) = J$ when $|z| = 1$, and $H^*(1/z)JH(z) < J$ when $|z| < 1$. Here $J = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}$.

Proof. See Ball *et al.* (1990) and Rotstein (1996). □

Note that the matrices A and C_- provide the structure of the interpolation problem while C_+ provides the interpolation values. The following corollaries show that both the Nevanlinna–Pick and the Carathéodory–Fejér problems are special cases of this theorem, corresponding to an appropriate choice of the matrices A and C_-.

Corollary 1 (*Nevanlinna–Pick*). Let $\Gamma = \mathrm{diag}\{z_i\} \in \mathbb{C}^{r \times r}$ and take

$$A = \Gamma, \quad (8)$$

$$C_- = [1 \quad 1 \quad \dots \quad 1] \in \mathbb{R}^r, \quad (9)$$

$$C_+ = [w_1 \quad w_2 \quad \dots \quad w_r], \quad (10)$$

then equation (4) is equivalent to

$$F(z_i) = w_i, \quad i = 1, \dots, r,$$

and the solution to equation (5) is the standard Pick matrix

$$P = \left[\frac{1 - w_i^* w_j}{1 - z_i^* z_j} \right]_{ij}.$$

Proof. Substitute A, C_-, C_+ in equation (4). See Rotstein (1996) for details. □

Corollary 2 (*Carathéodory–Fejér*). Let $I_{n \times n}$ denote the identity matrix, and $A_f \in \mathbb{R}^{(n+1)\times(n+1)}$,

$$A_f = \begin{bmatrix} 0 & I_{n \times n} \\ 0 & 0 \end{bmatrix}.$$

Take

$$A = A_f, \quad (11)$$

$$C_- = [1 \quad 0 \quad \dots \quad 0] \in \mathbb{R}^{n+1}, \quad (12)$$

$$C_+ = [f(0) \quad f(1) \quad \dots \quad f(n)] \quad (13)$$

then equation (4) is satisfied if and only if $F(z)$ can be written as

$$F(z) = f(0) + f(1)z + f(2)z^2 + \cdots + f(n)z^n + \cdots,$$

and the solution to equation (5) is the matrix: $M_c = I - \mathcal{F}^*\mathcal{F}$ where

$$\mathcal{F} = \begin{bmatrix} f(0) & f(1) & \dots & f(n) \\ 0 & f(0) & \dots & f(n-1) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & f(0) \end{bmatrix}.$$

Note that $M_c > 0$ if and only if $\bar{\sigma}(\mathcal{F}) < 1$.

Proof. Substitute A, C_-, C_+ in equation (4). See Rotstein (1996) for details. □

3. MAIN RESULTS

Nevanlinna–Pick based identification algorithms address the case where the experimental data available is purely frequency domain, while Carathéodory–Fejér based identification deals only with time domain data. In this section we exploit the generalized interpolation framework introduced in the previous section to solve Problem 1, obtaining a robust identification algorithm that combines both sources of data. To this effect, we will divide Problem 1 into two subproblems: (i) consistency and (ii) identification. The first consists of determining the existence of a candidate model $H \in \mathscr{S}$ which may have produced both, the time and frequency domain experimental data. Clearly, this is a prerequisite to the second stage, the identification step, consisting of the computation of the nominal model itself and a bound on the identification error.

3.1. *Consistency*

From equation (3) it follows that the problem of determining consistency of the *a posteriori* and *a priori* information reduces to establishing whether or not there exists a model $H \in \mathscr{S}$ that interpolates the frequency experimental data

$$\hat{\mathbf{h}} = y^f - \eta^f, \quad \eta^f \in \mathscr{N}_f \tag{14}$$

and has an impulse response that satisfies the constraints

$$U\mathbf{h} = y^t - \eta^t, \quad \eta^t \in \mathscr{N}_t, \tag{15}$$

where the noiseless output $U\mathbf{h}$ is the convolution of the input sequence $\mathbf{u} = [u(0)\ u(1)\ \ldots\ u(N_{t-1})]$ and the system $H(z)$.

The main result of this section shows that consistency can be established by solving a finite-dimensional convex optimization problem. To establish this result we will first obtain an equivalent condition for consistency (Lemma 1). This condition, based upon the relationship between both admissible experimental noises $\eta_f \in \mathscr{N}_f$ and $\eta_t \in \mathscr{N}_t$, has the form of a linearly constrained generalized interpolation problem. In Theorems 2 and 3 we will show that this generalized problem can be recast in terms of an LMI optimization.

Lemma 1. The *a priori* and *a posteriori* information are consistent if and only if there exists a function $H \in \mathscr{H}_\infty(\rho, K)$ such that

$$\hat{\mathbf{h}} = y^f - \eta^f, \qquad \eta^f \in \mathscr{N}_f \tag{16}$$

$$y_L \le \begin{bmatrix} U \\ I \end{bmatrix} \mathbf{h} \le y_U, \tag{17}$$

where

$$\mathbf{y}_L = \begin{bmatrix} \mathbf{y}^t(1) - \varepsilon_t \\ \vdots \\ \mathbf{y}^t(N_{t-1}) - \varepsilon_t \\ \\ \phi_l(0) \\ \vdots \\ \phi_l(N_\phi - 1) \end{bmatrix}, \quad \mathbf{y}_U = \begin{bmatrix} \mathbf{y}^t(1) + \varepsilon_t \\ \vdots \\ \mathbf{y}^t(N_{t-1}) + \varepsilon_t \\ \\ \phi_u(0) \\ \vdots \\ \phi_u(N_{\phi-1}) \end{bmatrix}. \tag{18}$$

Proof. The proof follows immediately by substituting equations (14) and (15) in equation (3). □

The next theorem provides necessary and sufficient conditions for the existence of a function $H \in \mathscr{H}_\infty(\rho, K)$ which interpolates *fixed* frequency domain experimental data while, at the same time, satisfying a time-domain constraint.

Theorem 2. Given N_f frequency-domain data points $(z_i, w_i), |z_i| < \rho, i = 0, \ldots, N_f - 1$, and N_t time-domain data points $h(k), k = 0, \ldots, N_{t-1}$, there exists $H \in \mathscr{H}_\infty(\rho, K)$ that interpolates the frequency domain data (i.e., $H(z_i) = w_i$) and such that $H(z) = h(0) + h(1)z + h(2)z^2 + \cdots + h(N_{t-1}) z^{N_{t-1}} + \cdots$ if and only if

$$M_R(\mathbf{w}, \mathbf{h}) = \begin{bmatrix} Q - \frac{1}{K^2}\mathscr{W}_f^* Q \mathscr{W}_f & M_X \\ M_X^* & R^{-2} - \frac{1}{K^2}\mathscr{F}_t^* R^{-2}\mathscr{F}_t \end{bmatrix} > 0, \tag{19}$$

where

$$M_X = S_0 R^{-2} - \frac{1}{K^2}\mathscr{W}_f^* S_0 R^{-2}\mathscr{F}_t, \tag{20}$$

$$R = \mathrm{diag}[1\ \ \rho\ \ \rho^2\ \ \ldots\ \ \rho^{N_t - 1}\ ], \tag{21}$$

$$Q = \left[\frac{\rho^2}{\rho^2 - z_{i-1}^* z_{j-1}}\right]_{ij}, \quad i, j = 1, \ldots, N_f,$$

$$S_0 = [(z_{i-1}^{j-1})^*]_{ij}, \quad i = 1, \ldots, N_f,\ j = 1, \ldots, N_t \tag{23}$$

$$\mathscr{W}_f = \mathrm{diag}[w_0\ \ \ldots\ \ w_{N_f - 1}], \tag{24}$$

$$\mathscr{F}_t = \begin{bmatrix} h(0) & h(1) & \ldots & h(N_t - 1) \\ 0 & h(0) & \ldots & h(N_t - 2) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & h(0) \end{bmatrix}. \tag{25}$$

Proof. First we scale the desired class of functions to transform the generalized interpolation problem from $\mathscr{H}_{\infty,\rho}$ to $\mathscr{H}_\infty$, i.e., $KF(z/\rho) \in \mathscr{H}_\infty(\rho, K) \Leftrightarrow F(z) \in \mathscr{BH}_\infty$. Equivalently we can scale the data points which amounts to modifying the matrices in problem (4) as follows:

$$A = \begin{bmatrix} \frac{1}{\rho}\Gamma & 0 \\ 0 & A_f \end{bmatrix},$$

$$C_- = \begin{bmatrix} \underbrace{1 \; \ldots \; 1}_{N_f} & \overbrace{1 \; 0 \; \ldots \; 0}^{N_f} \end{bmatrix},$$

$$C_+ = \frac{1}{K}[w^* \quad \mathbf{h}^* R],$$

where Γ and A_f have been defined in Corollaries 1 and 2, respectively, and where $w \in \mathbb{C}^{N_f}$ is the vector of interpolation points. The above is simply a generalized interpolation problem combining the time and frequency domain constraints (Rtstein, 1996). Solving the Lyapunov equation (5) and pre-post multiplying the solution by the diagonal matrix:

$$T \triangleq \begin{bmatrix} I & 0 \\ 0 & R^{-1} \end{bmatrix}$$

yields M_R. Since T is non-singular the positiveness of the solution of equation (5) is not modified, i.e. $M > 0 \Leftrightarrow M_R > 0$. □

Remark 1. The (1, 1) block of M_R is the Pick matrix corresponding to the frequency domain consistency problem solved in Chen *et al.* (1992) via the classical Nevanlinna–Pick interpolation. Block (2, 2) is the Caratheodory–Fejer matrix corresponding to the time domain consistency problem solved in Chen and Nett (1995) and Chen *et al.* (1994). M_X is a cross-coupling term due to the existence of both types of experimental data.

Combining the previous result (that considers only noiseless data points) with Lemma 1 yields the following necessary and sufficient condition for consistency:

Lemma 2. The *a priori* and *a posteriori* information are consistent if and only if there exists two vectors

$$\mathbf{w} = \begin{bmatrix} w_0 \\ w_1 \\ \cdots \\ w_{N_f-1} \end{bmatrix}, \quad \mathbf{h} = \begin{bmatrix} h(0) \\ h(1) \\ \cdots \\ h(N_t-1) \end{bmatrix}$$

such that

$$M_R(\mathbf{w}, \mathbf{h}) > 0, \tag{26}$$

$$(\mathbf{y}^f - \mathbf{w}) \in \mathcal{N}_f, \quad (\mathbf{y}^t - U\mathbf{h}) \in \mathcal{N}_t. \tag{27}$$

Note that the components of $\mathbf{w}$ and $\mathbf{h}$ are elements of the matrices $\mathcal{W}_f$ and $\mathcal{F}_t$, respectively.

Proof. The proof follows immediately by combining Lemma 1 and Theorem 2. □

From Lemma 2 it follows that the consistency problem can be reduced to solving a feasibility problem in terms of the time and frequency domain vectors $\mathbf{h}$ and $\mathbf{w}$. In the next Theorem we will show that this feasibility problem is a convex problem that can be recast in terms of LMIs and thus efficiently solved, using for instance interior–point methods (Nesterov and Nemirovski, 1994; Boyd *et al.*, 1994).

Theorem 3. The consistency problem with mixed time/frequency-domain data is equivalent to a LMI feasibility problem.

Proof. The matrix M_R can be written as

$$M_R = M_0 - \frac{1}{K^2} X^* M_0 X, \tag{28}$$

with

$$M_0 = \begin{bmatrix} Q & S_0 R^{-2} \\ R^{-2} S_0^* & R^{-2} \end{bmatrix}, \tag{29}$$

$$X = \begin{bmatrix} \mathcal{W}_f & 0 \\ 0 & \mathcal{F}_t \end{bmatrix}. \tag{30}$$

Positiveness of the matrix M_0 is equivalent to consistency in the case where both frequency and time domain data are zero, i.e. $w_i = 0, i = 0, \ldots, N_f - 1$ and $h(j) = 0, j = 0, \ldots, N_t - 1$. This is a homogeneous interpolation problem, that always has solutions (in particular, the trivial solution $H(z) = 0$, and the Blaschke product). From Theorem 1 it follows that $M_0 > 0$, and thus, using Schur complements we have

$$M_R > 0 \Leftrightarrow Z \triangleq \begin{bmatrix} M_0^{-1} & \frac{1}{K} X \\ \frac{1}{K} X^* & M_0 \end{bmatrix} > 0. \tag{31}$$

Clearly this is a LMI in X. Define now $\mathcal{Y}_f \triangleq \mathrm{diag}\{\mathbf{y}^f\}$. Using Schur complements it can be easily shown that $(\mathbf{y}^f - \mathbf{w}) \in \mathcal{N}_f$ is equivalent to

$$\begin{bmatrix} \varepsilon_f I & \mathcal{Y}_f - \mathcal{W}_f \\ (\mathcal{Y}_f - \mathcal{W}_f)^* & \varepsilon_f I \end{bmatrix} > 0. \tag{32}$$

Finally, the constraint $(\mathbf{y}^t - U\mathbf{h}) \in \mathcal{N}_t$ is equivalent to the LMI

$$-\varepsilon_t < \mathbf{y}^t - \mathcal{F}_t^* \mathbf{u} < \varepsilon_t, \tag{33}$$

where this last inequality should be understood in the componentwise sense. Thus, the consistency problem is equivalent to finding a feasible solution to the set of LMIs (31)–(33). Alternatively, it can be recast as the following standard generalized eigenvalue minimization problem (Nesterov and Nemirovski, 1990; Overton, 1988; Vandenberghe and Boyd, 1996):

$$\min_{X \in \mathcal{X}} \quad \bar{\lambda}[-Z(X)] < 0$$

$$\mathcal{X} \triangleq \left\{ \begin{bmatrix} \mathscr{W}_f & 0 \\ 0 & \mathscr{F}_t \end{bmatrix} \text{subject to (32), (33)} \right\}. \qquad \square$$

It is interesting to analyze the extreme conditions on the *a priori* elements K and ρ. Note that for $K \to \infty$, $M_R \to M_0$, which is positive definite. Therefore in the limit the problem is always consistent. This seems reasonable considering that any unbounded function may be a candidate nominal model. On the other hand, as $K \to 0$, M_R tends to be negative definite and therefore establishes an empty consistency set in the case of nonzero *a posteriori* information.

3.2. *Identification*

Once consistency is established, the second step towards solving Problem 1 consists of generating a nominal model in the consistency set $\mathscr{S}(\mathbf{y}^f, \mathbf{y}^t)$. The identification algorithm that we propose is based on the parameterization of all solutions of the generalized Nevanlinna–Pick interpolation problem (Ball *et al.*, 1990) presented in Theorem 1. For simplicity we consider the case where the matrix M_R is strictly positive definite and therefore the solution is non-unique. Details for the degenerate case where there exists a unique solution can be found in (Ball *et al.*, 1990). The algorithm can be summarized as follows:

(1) Find feasible data vectors $\mathbf{w}, \mathbf{h}$ for the consistency problem (26), (27) by solving the LMI feasibility problem given by equations (31)–(33). Note that there is no need of any kind of optimality in the search for the feasible vectors. Instead, any pair $\mathbf{w}, \mathbf{h}$ in the admissible set will suffice.
(2) Compute the generalized Pick matrix M_R in equation (19) (which should be positive definite), corresponding to the vectors found in step (1).
(3) Use Theorem 1 to compute a model from the consistency set $\mathscr{S}$. Recall that all the models in $\mathscr{S}$ (i.e. all the solutions to the generalized interpolation problem) can be parameterized as a Linear Fractional Transformation (LFT) of a free parameter $Q(z) \in \mathscr{B}\mathscr{H}_\infty$ as follows:

$$H(z) = F_l[L(z), Q(z)], \qquad (34)$$

$$L(z) = \begin{bmatrix} T_{12}T_{22}^{-1} & T_{11} - T_{12}T_{22}^{-1}T_{21} \\ & \\ T_{22}^{-1} & -T_{22}^{-1}T_{21} \end{bmatrix}. \qquad (35)$$

In particular, if the free parameter $Q(z)$ is chosen as a constant, then the model order is less than or equal to $N_f + N_t$.

Remark 2. Note that $T(z)$ depends on the choice of vectors $\mathbf{w}, \mathbf{h}$. Thus, there are additional degrees of freedom available in the problem (choices of $\mathbf{w}, \mathbf{h}$ and $Q(z)$) that could be used to optimize additional performance criteria (e.g. model order).

Since the proposed algorithm is interpolatory, it has several advantages over the usual "two step" algorithms sometimes used in the context of robust identification (Gu and Khargonekar, 1992; Helmicki *et al.*, 1991). In particular, since the identified model is in the set $\mathscr{S}(\mathbf{y}^f, \mathbf{y}^t)$, its distance to the Chebyshev center of this set is within the diameter of information (Mäkilä, 1992). As a consequence the algorithm is optimal up to a factor of two as compared with central strongly optimal procedures. For the same reasons, it is also convergent and therefore the modeling error tends to zero as the information is completed.

3.3. *Analysis of the identification error*

In this section we derive upper and lower bounds for the worst-case identification error. Since these bounds are given in terms of the radius and diameter of information (Helmicki *et al.*, 1991; Chen *et al.*, 1992), they are valid for *all* interpolatory algorithms taking as inputs the available *a priori* and *a posteriori* information.

Lemma 3. Assume that $\Phi_u(k) = -\Phi_l(k) = \Phi(k) \geq 0$, $k = 0, \ldots, N_\Phi - 1$ (symmetric time domain *a priori* information), and let $\hat{b} = \min[\varepsilon_f, \Phi(0), \varepsilon_t/\|\mathbf{u}\|_\infty]$, where $\mathbf{u}$ is the vector associated with the input signal sequence. Then, the radius of information $\mathscr{R}_\mathscr{I}$ satisfies:

$$\text{If} \qquad \hat{\beta} \geq K, \quad \mathscr{R}_\mathscr{I} \geq K \qquad (36)$$

$$\text{If} \qquad \hat{\beta} < K, \mathscr{R}_\mathscr{I} \geq \frac{K\|B(z)\|_\infty + \hat{\beta}}{1 + \|B(z)\|_\infty \hat{\beta}/K} \qquad (37)$$

Proof. Let $N \triangleq \max[N_t - 1, N_\Phi - 1]$ and define

$$B(z) \triangleq \left(\frac{z}{\rho}\right)^N \prod_{k=0}^{N_f - 1} \frac{\rho(z - z_k)}{\rho^2 - z_k^* z}$$

Using the properties of the Blaschke product it is easily seen that $B(z)$ is in $\mathscr{H}_\infty(\rho, 1)$. Therefore

$$H(z) \triangleq K \frac{B(z) + \beta/K}{1 + B(z)\beta^*/K}$$

is in $\mathcal{H}_\infty(\rho, K)$, provided that $|\beta| < K$. Additionally, $H(z)$ is consistent with the *a posteriori* frequency information if $|\beta| \le \varepsilon_f$. Furthermore, if $|\beta| \le \min[\varepsilon_t/\|\mathbf{u}\|_\infty, \Phi(0)]$, $H(z)$ will also be consistent with the time domain *a priori* and *a posteriori* information. The desired result follows from the fact that $\mathcal{R}_\mathcal{I} \ge \|H(z)\|_\infty \ge |H(z_*)|$ by taking $\beta = \hat{\beta} e^{j\alpha}$, where $\alpha = \angle B(z_*)$ and $|B(z_*)| = \|B(z)\|_\infty$. □

Lemma 4. Assume the same *a priori* information as in the previous lemma. Then the radius of information $\mathcal{R}_\mathcal{I}$ can be bounded above by

$$\mathcal{R}_\mathcal{I} \le \sum_{i=0}^{M} v_i + \frac{K}{\rho^M(\rho - 1)}$$

where $M = N_t + N_f - 1$ and v_i are a function of the *a priori* information only.

Proof. Consider any $H(z) \in \mathcal{S}(0, 0)$ and partition it as follows:

$$H(z) = \underbrace{\sum_{k=0}^{M} h(k) z^k}_{F(z)} + \overbrace{\sum_{k=M+1}^{\infty} h(k) z^k}^{G(z)}.$$

Consider now the first portion of the expansion. The vector $\mathbf{f}$ (the coefficients of $F(z)$) satisfies:

$$V^{-1}\mathbf{f} = \begin{bmatrix} \eta^t \\ \eta^f \end{bmatrix},$$

$$V^{-1} = \begin{bmatrix} & U & & 0 & \\ 1 & z_0 & z_0^2 & \cdots & z_0^M \\ 1 & z_1 & z_1^2 & \cdots & z_1^M \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & z_{N_f-1} & z_{N_f-1}^2 & \cdots & z_{N_f-1}^M \end{bmatrix},$$

where $\eta^t \in \mathcal{N}_t$, $\eta^f \in \ell_\infty(\bar{\varepsilon})$, and where $\bar{\varepsilon} \triangleq \varepsilon_f + K/(\rho^M(\rho - 1))$. Note that V^{-1} is nonsingular as long as $u_0 \ne 0$. Next partition V as

$$V = \begin{bmatrix} V_{1t} & V_{1f} \\ V_{2t} & V_{2f} \\ \vdots & \vdots \\ V_{(N_f+N_t)t} & V_{(N_f+N_t)f} \end{bmatrix}.$$

It is well known that $\|g\|_{\ell_1} \le K/(\rho^M(\rho - 1))$. It follows that

$$\mathcal{R}_\mathcal{I} \le \sup_{F(z) \in \mathcal{S}(0,0)} \|F(z)\|_\infty + \frac{K}{\rho^M(\rho - 1)}$$

$$\le \sup_{F(z) \in \mathcal{S}(0,0)} \|f\|_{\ell_1} + \frac{K}{\rho^M(\rho - 1)}. \tag{38}$$

Thus, in order to compute a bound on $\mathcal{R}_\mathcal{I}$ we need to compute an upper bound on $\|f\|_{\ell_1}$. This bound can be computed as follows:

$$\|f\|_1 = \sum_{i=0}^{M} |V_{it}\eta^t + V_{if}\eta^f| \tag{39}$$

$$\le \sum_{i=0}^{M} \varepsilon_t \|V_{it}\|_1 + \bar{\varepsilon}\|V_{if}\|_1. \tag{40}$$

Define

$$v_i \triangleq \min[\phi(i), K\rho^{-i}, \varepsilon_t\|V_{it}\|_1 + \bar{\varepsilon}\|V_{if}\|_1].$$

From equation (1) and the definition of Φ, we have that

$$\|f\|_{\ell_1} \le \sum_{i=0}^{M} v_i \tag{41}$$

$$\Rightarrow \mathcal{R}_\mathcal{I} \le \sum_{i=0}^{M} v_i + \frac{K}{\rho^M(\rho - 1)}. \quad □ \tag{42}$$

4. COMPUTATIONAL CONSIDERATIONS

In this section we address some issues concerning the practical implementation of the algorithm outlined in Section 3.2. In particular we analyze the conditioning of the problem as the cardinality of the data grows.

The consistency question, as we saw in Section 2.3, can be reduced to establishing positiveness of a generalized Pick matrix that depends quadratically on the optimization variables (28). To reduce the problem to one affine in X, an explicit inversion of M_0 is needed. However, the matrix M_0, while always being positive definite, is asymptotically singular, with its condition number growing without bound as the number of data points increases. This is certainly reasonable, because as the amount of data available increases, the solution (if one exists) will tend to be unique. It is also consistent with the fact that the condition number does not decrease when data points are added (this last fact can be easily established using the well-known singular values inclusion property).

The following lemma gives an estimate on the growth of M_0's condition number in the most favorable case, i.e. when the z_i are equidistant* (roots of the unity). It provides a lower bound on the conditioning of matrix M_0.

Lemma 5. Let $z_k = e^{j(2\pi k/N_f)}$, $k = 0, \ldots, N_f - 1$ (the N_fth roots of the unity). In this case, the singular values and condition number of M_0 are bounded by

$$\sigma_1(M_0) \ge \sigma_i(Q) = \frac{1}{\rho^{2(i-1)}} \frac{N_f \rho^{2N_f}}{\rho^{2N_f} - 1} \ge \sigma_{N_f}(M_0),$$

$$i = 1, \ldots, N_f, \tag{43}$$

$$\kappa(M_0) \ge \kappa(Q) = \frac{\sigma_1(Q)}{\sigma_{N_f}(Q)} = \rho^{2(N_f - 1)}. \tag{44}$$

*If for some i, j, $z_i - z_j < \varepsilon$, as $\varepsilon \to 0$, M_0 tends to singularity.

Proof. When the z_k are chosen as the roots of unity, the Pick matrix is a circulant matrix, i.e. $Q_{ij} = c_{(i-j)\bmod N_f}$. Since Q is normal (it is hermitian), its singular values are the absolute value of its eigenvalues. Since the eigenvalues of a circulant matrix can be obtained as the Discrete Fourier Transform of the elements of the first row, it follows that the singular values of Q can be obtained from the identity:

$$\sum_{k=0}^{N_f-1} \frac{\gamma e^{(2\pi/N_f)km}}{\gamma - e^{j(2\pi/N_f)k}} = \frac{N_f \gamma^m}{\gamma^{N_f} - 1}, \quad m = 1, \ldots, N_f.$$

The desired result follows now from setting $\gamma = \rho^2$, and using the interlacing property of the eigenvalues of a symmetric matrix and its diagonal submatrix (Q in this case). □

Thus, we see that the condition number of the generalized Pick matrix, has at least an exponential growth with the number of frequency data samples.

Next we show that the structure of the consistency set allows for taking into account variations in the experimental data points, and that this set can be characterized as an LFT of the experimental data. This is a generalization of a result proved in Zhou and Kimura (1995) for the case where only time domain experimental data is available. From the methodological point of view this justifies using the proposed robust identification procedure in a robust control framework. It provides a natural structure (Zhou and Kimura, 1995) for the set of unfalsified models. Furthermore it allows a direct connection between the identification and design procedures. As a consequence, it is possible to provide an algorithm that takes as inputs the time and frequency domain experimental data and produces a controller guaranteed to provide robust performance of the physical system.

For simplicity and without loss of generality, in the sequel we consider the special case where the interpolation functions are in $\mathscr{BH}_\infty$, i.e. $\rho = K = 1$, which implies $R = I$. The general case can be obtained by simply using the scaling in Theorem 2.

Lemma 6. Consider the consistency set described in equation (6), with

$$T(z) = I + \begin{bmatrix} C_+ \\ C_- \end{bmatrix} W(z)$$
$$\times M_I^{-1}(A^* - I)^{-1}[-C_+^* \quad C_-^*], \qquad (45)$$
$$W(z) = I + (A - I)(zI - A)^{-1}$$
$$= (z - 1)(zI - A)^{-1}, \qquad (46)$$

where $M_I = M_R|_{R=I}$. If M_I is positive definite for any possible experimental noise compatible with the *a priori* information then $\mathscr{S}$ can also be described by a LFT of the form $F_u(\mathscr{T}(z), \Delta)$, where $\mathscr{T}(z)$ is a constant transfer matrix and the structured uncertainty is described by

$$\Delta = \mathrm{diag}[X \quad X \quad X^* \quad X^*]. \qquad (47)$$

Furthermore, this uncertainty block can be presented in the standard form, with 4 repeated blocks of N_f complex diagonal matrices, and $4N_t$ repeated blocks of N_t real diagonal matrices, as follows:

$$\tilde{\Delta} = \mathrm{diag}[\Delta_c, \Delta_c^*, \Delta_r], \qquad (48)$$

$$\Delta_c = \mathrm{diag}[\delta_1, \delta_1, \ldots, \ldots, \delta_{N_f}, \delta_{N_f}], \quad \delta_i \in \mathbb{C}, \qquad (49)$$

$$\Delta_r = \mathrm{diag}\, \underbrace{\delta_{N_f+1}, \ldots, \delta_{N_f+1}}_{4N_t}, \ldots,$$
$$\overbrace{\delta_{N_f+N_t}, \ldots, \delta_{N_f+N_t}}^{4N_t}], \quad \delta_i \in \mathbb{R}. \qquad (50)$$

Proof. The expressions (45) and (46) follow directly from Theorem 1. Note that $C_+ = C_- X$, where X is defined in equation (30). Next we show that the term M^{-1} in $T(z)$ can be written as a LFT of the data. To see this, note that using Schur's formula for the inverse of a partitioned matrix, we have

$$M^{-1} = [0 \quad I] \begin{bmatrix} M_0^{-1} & X \\ X^* & M_0 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ I \end{bmatrix}.$$

Note that matrix Z (defined in equation (31) with $K = 1$) is affine in X and its conjugate. Therefore from M^{-1} we obtain a LFT with an "uncertainty" block $\Delta_X \triangleq \mathrm{diag}[X, X^*]$ and from C_+ and C_+^* in equation (45) we obtain another uncertainty block Δ_X. Since the combination of LFTs is another LFT, it follows that $T(z)$ can be written as $T(z) = F_u(\mathscr{T}(z), \Delta)$, where Δ is given by equation (47). Furthermore, this uncertainty block can be represented in the standard form by writing $X = \sum_{i=1}^{N_f+N_t} C_i \delta_i$. Here $C_i \in \mathbb{R}^{N_t \times N_t}$ are constant matrices and δ_i are complex scalars for $i = 1, \ldots, N_f$ and real scalars for $i = N_f + 1, \ldots, N_f + N_t$. □

The assumption on the positiveness of matrix M_I is related to the magnitude of the experimental noise, which may be explained as follows. Assume that for a certain measurement noise in $(\mathscr{N}_t, \mathscr{N}_f)$, the *a priori* class $\mathscr{S}$ and the *a posteriori* information are consistent, i.e., $M_I > 0$. This provides an experimental set of data which we interpolate to obtain a nominal model. It is clear that for sufficiently large noise bounds ε_t^*, ε_f^* there will exist noise elements in the new classes $(\mathscr{N}_t^*, \mathscr{N}_f^*)$, such that when they are added to the experimental set defined above, produce a new possible data set which is inconsistent with $\mathscr{S}$, i.e., M is not positive definite. By "continuity" arguments we note that as the noise sets $(\mathscr{N}_t, \mathscr{N}_f)$

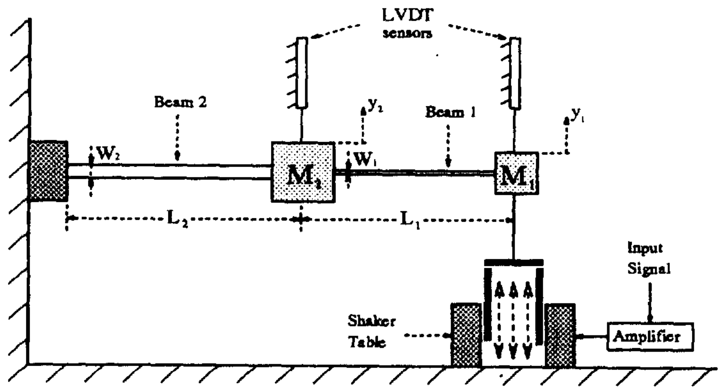


Fig. 1. The flexible testbed.

grow, i.e., ε_t (ε_f) approach ε_f^* (ε_f^*); there is a noise class size for which M is singular. Therefore it is clear that for small enough noise bounds, we can assume the positiveness of M over all noise elements.

On the other hand, the current state of the control design procedures prevents its application to system descriptions which include a large number of uncertainty blocks. Furthermore, in the case of real (Poljak and Rohn, 1993; Braatz, *et al.*, 1994) and mixed (real + complex) uncertainty the computational complexity is NP-hard. As a consequence, from a purely numerical point of view, there is still research to be done which should produce algorithms which can cope with systems described by many real and complex Δ blocks.

5. ILLUSTRATIVE EXAMPLES

In this section, we present two examples that illustrate the importance of considering both time and frequency experimental information. The first one, very simple and mainly of conceptual value, shows that consistency with time or frequency data does not imply consistency with both. The second example is more practical, and deals with the identification of a flexible structure, using both time and frequency domain experimental data.

Example 1. For the first example, consider the following data:

- *A priori* information: $K = 10$, $\rho = 5$. For simplicity, we will initially consider $\varepsilon_f = \varepsilon_t = 0$ (noiseless sampling).
- *A posteriori* information:
  - Frequency data: $F(1) = 1$, $F(j) = 1$, $F(-1) = 1$.
  - Time domain data: $f(0) = 1, f(1) = 0.1$, $f(2) = 0.01$.

We will see that the *a priori* assumptions are consistent with the time domain or frequency domain *a posteriori* information, but not with both simultaneously. To this end, note that

$$G(z) = 1$$

belongs to $\mathscr{H}_\infty(\rho, K)$, and interpolates exactly the frequency data. On the other hand,

$$G(z) = \frac{10}{10 - z}$$

also belongs to $\mathscr{H}_\infty(\rho, K)$, and interpolates exactly the time domain data.

However, the generalized Pick matrix M_R corresponding to this data is not positive definite, and therefore *there is no function* in $\mathscr{H}_\infty(\rho, K)$, that interpolates simultaneously both sets of data. A direct proof of the above fact follows. Using equation (1), consider the norm of one of the frequency experiments as a function of the time domain data:

$$F(1) = \left|\sum_{k=0}^{\infty} f(k)\right| = \left|\sum_{k=0}^{2} f(k) + \sum_{k=3}^{\infty} f(k)\right| \tag{51}$$

$$\geq 1.11 - \sum_{k=3}^{\infty} |f(k)| \tag{52}$$

$$\geq 1.11 - \frac{K}{\rho^2(\rho - 1)} = 1.01. \tag{53}$$

Instead, the experiment outcome is $F(1) = 1$, which clearly denotes the inconsistency between both experiments.

In fact, in the noiseless case, it is not necessary to use the generalized theory, as we can always find the solution to the "pure" Carathéodory–Fejér problem, and then find interpolation constraints on the free parameter $Q(z)$. The real advantage of our procedure appears in practical cases with the presence of both, time and frequency measurements errors.

To see this, we will use our algorithm to compute the smallest noise bound* that makes the

*For simplicity, we consider the time and frequency noise bounds to be equal. There is no difficulty in removing this assumption.

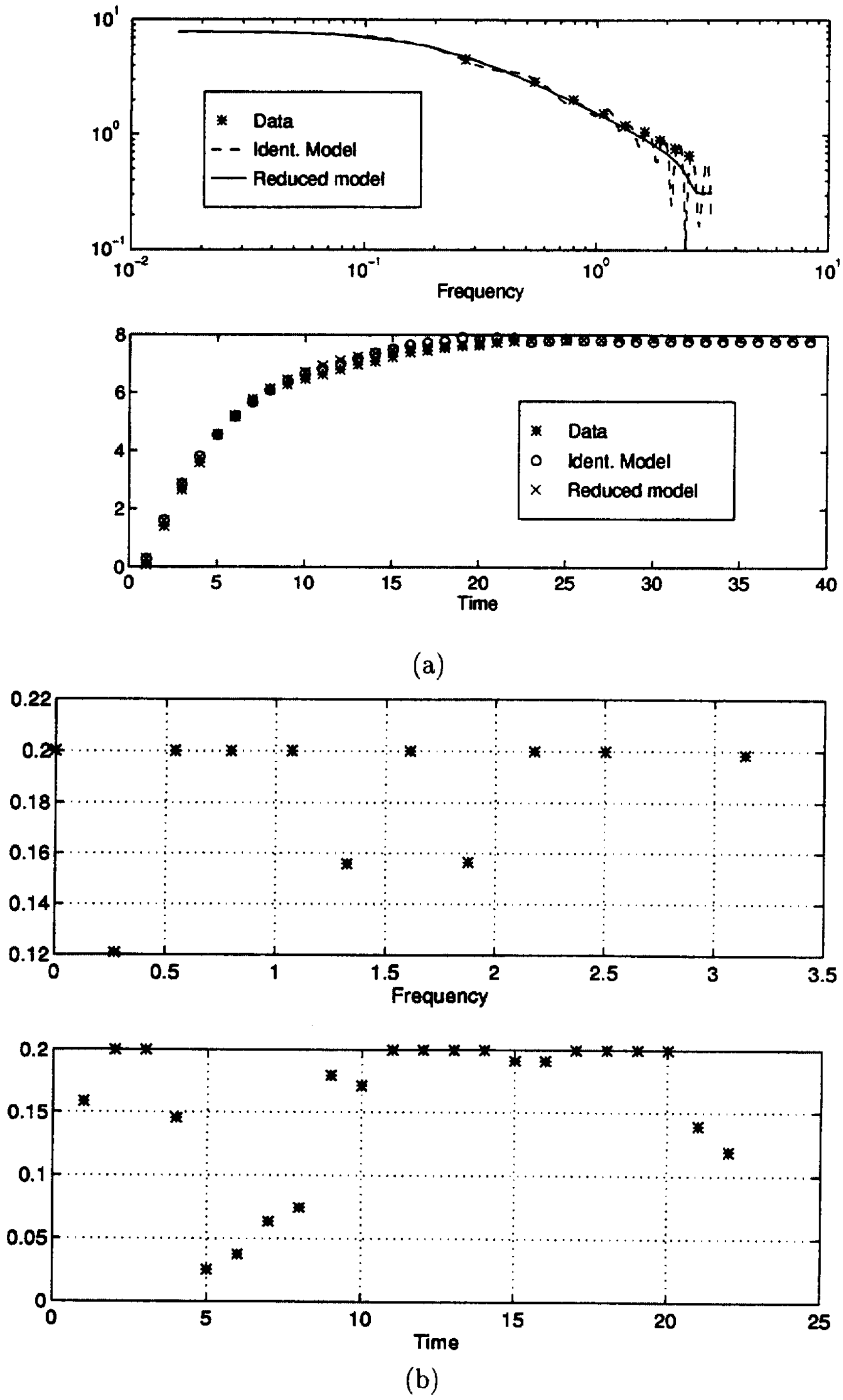


Fig. 2. (a) Experimental data and identified system, (b) Identification error.

experimental data consistent with the *a priori* information. This search can be cast as a convex optimization problem, and solved by using a minor modification of the consistency algorithm. In this example the smallest noise bound necessary for consistency satisfies $0.0484 < \varepsilon_{\min} < 0.0485$. This means that if the (time and frequency) noise is below 0.0484 then the *a posteriori* and *a priori* information are inconsistent. On the other hand, if both (time and frequency) noises are above 0.0485, there always exists an interpolating function for both types of data.

Example 2. Next we illustrate the proposed framework by applying it to the problem of identifying a flexible structure used as a damage mitigation testbed (Tangirala *et al.*, 1995). The structure is a two degree of freedom mass–beam system

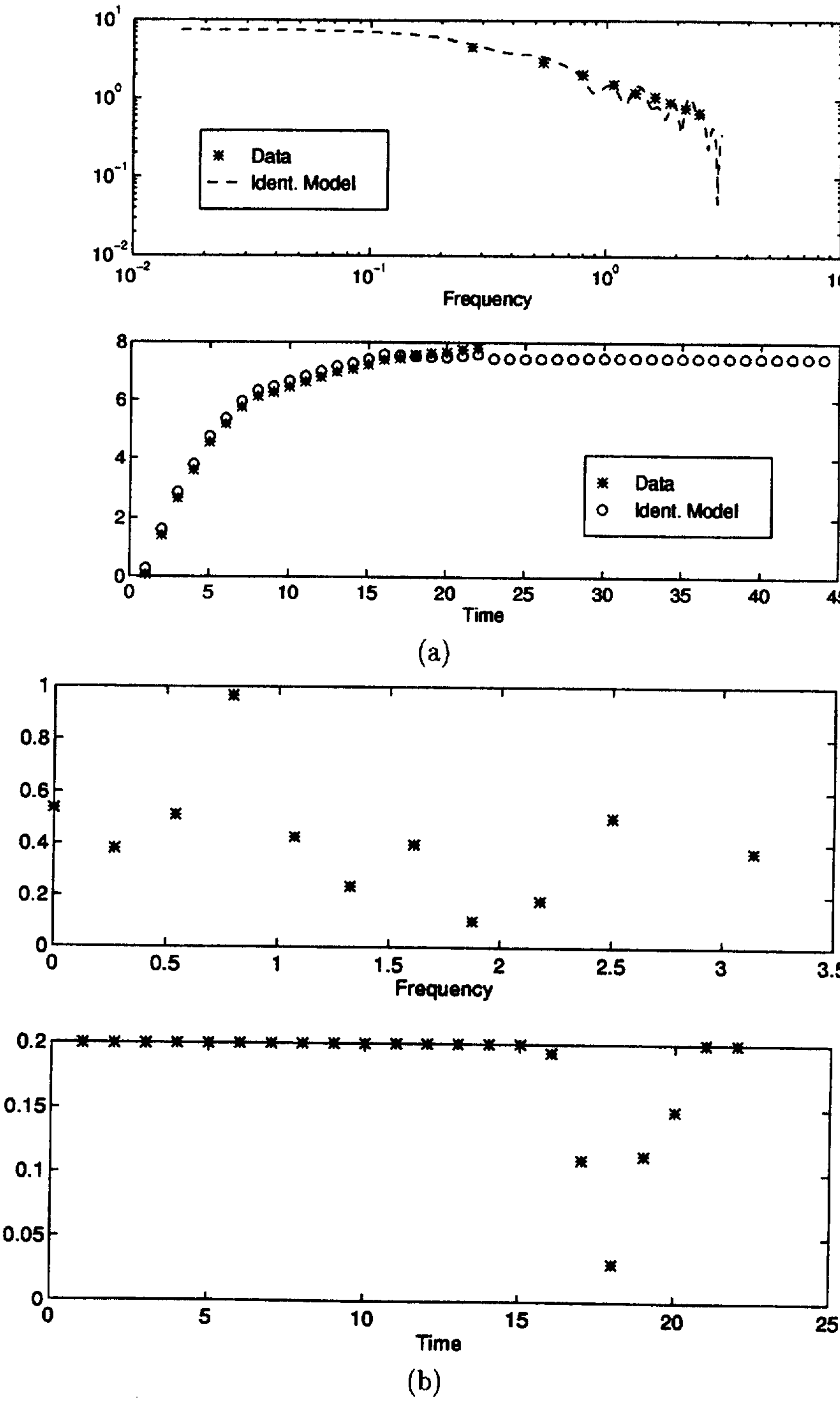


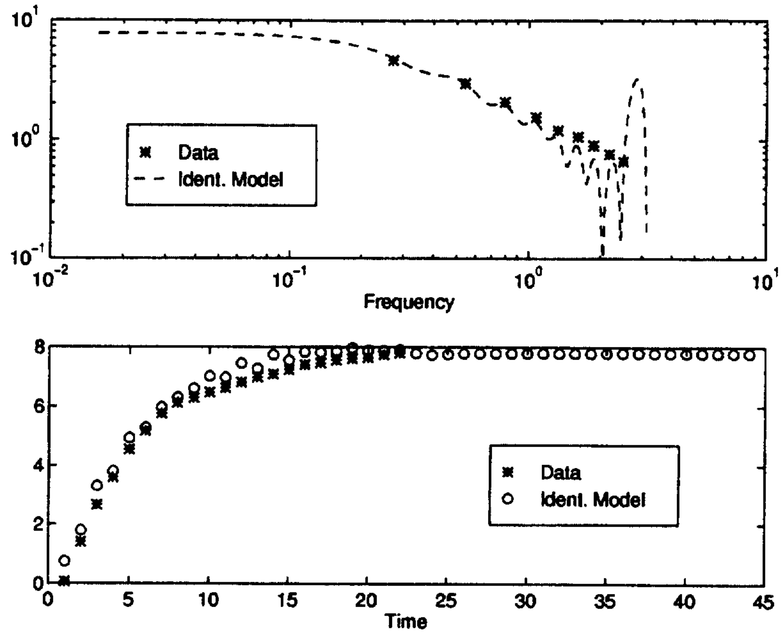
Fig. 3. ℓ^1 identification: (a) nominal model, (b) identification error.

consisting of two discrete masses supported by cantilever beams, excited by the vibratory motion of a shaker table as shown in Fig. 1.

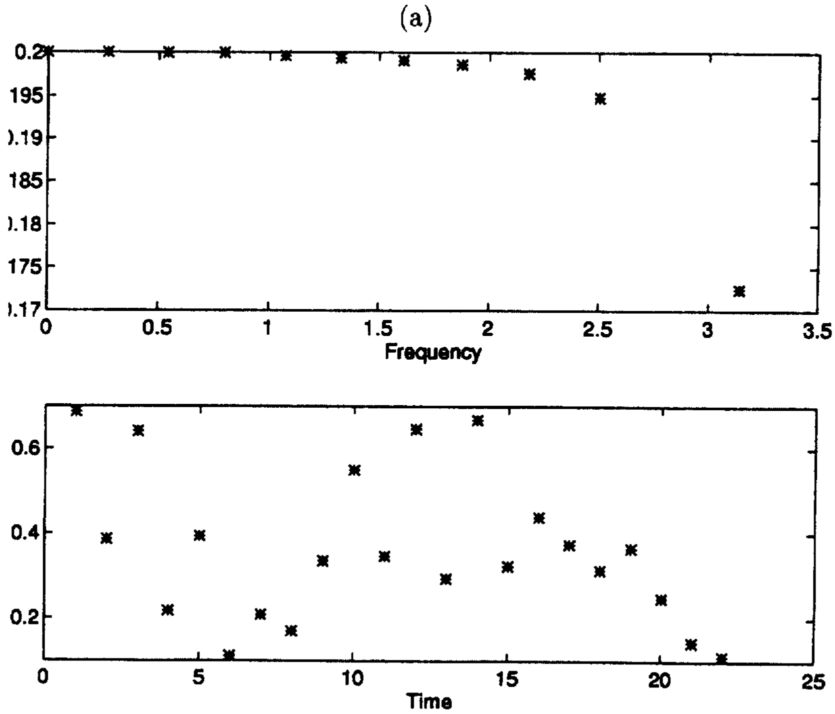
The first mass is connected to the shaker table, which excites the mechanical system by vibrating up and down, through a flexible pivot. The displacement y_1 caused by the shaker table is measured using a linear variable differential transformer (LVDT) sensor located at the midpoint of the mass M_1. The numerical values of the parameters are (Tangirala et al., 1995):

$$M_1 = 2.702 \text{ lbm} \qquad M_2 = 7.664 \text{ lbm}$$
$$L_1 = 8.5 \text{ inch} \qquad L_2 = 11.84 \text{ inch}$$
$$W_1 = 0.437 \text{ inch} \qquad W_2 = 0.87 \text{ inch}$$

This mass–beam system, intended to model a plant subjected to damage inducing stress, is being used to test the concepts of life-extending and damage mitigating control (Tangirala et al., 1995). Life



(a)



(b)

Fig. 4. $\mathscr{H}_\infty$ identification: (a) nominal model, (b) identification error.

extension is achieved by designing multiobjective controllers that keep the peak values of both the time and frequency responses below some prespecified thresholds. Thus, in this application is important to have models that accurately reproduce the behavior of the system in both domains.

To obtain the frequency-domain data points required by $\mathscr{H}_\infty$ identification the system was driven by a peak-to-peak 0.5 V sinusoidal signal, with frequency ranging from 1 to 21 Hz. The time domain data points where obtained by exciting the structure with a peak-to-peak 0.5 V square wave with frequency 2 Hz. In both cases the outputs were sampled with a sampling time $T_s = 0.0215$. Finally, ρ was estimated to be $\rho = 1.25$ and by measuring the output in the absence of a driving signal it was determined that the measurement noise was bounded by $\varepsilon_f = \varepsilon_t = 0.2$. A total of 33 samples were used, 22 from time domain data and the rest from frequency (not counting the ones that are obtained by the complex conjugate symmetry of the transfer function, otherwise the total is 42). The limited number of samples is mainly due to numerical problems with the optimization software, that currently does not exploit the structure available in the problem. The identification stage was followed by a model reduction stage (via balancing and truncation), resulting in a final identified model of third order. Figure 2 shows the time and frequency responses of this model versus experimental data. As it can be seen there, the model interpolates both sources of data within the experimental error.

To illustrate the advantage of our approach, we also identified the system using "pure" time (i.e., ℓ_1) and frequency ($\mathscr{H}_\infty$) domain methods. Figures 3 and 4 show the responses of the resulting models as well as the identification error. As predicted, methods using only one source of data achieved a good fit in the corresponding domain, but incurred in large errors in the order.

6. CONCLUSIONS AND DIRECTIONS FOR FUTURE RESEARCH

In this work we presented a new generalized robust identification framework that combines both frequency and time-domain experimental data, thus avoiding situations where a "good" fit of the data provided by one class of experiments (such as frequency domain) leads to poor fitting of the data provided by the other experiments, as illustrated with the experimental example of Section 5.

The main result of the paper shows that the problems of establishing consistency of the data and of obtaining a nominal model and bounds on the identification error can be recast as a LMI feasibility problem that can be efficiently solved.

Additionally, we have shown that in this context the set of models consistent with both the a priori and a posteriori information can be parameterized as a LFT of the experimental data, thus justifying the combination of the proposed algorithm with standard robust control synthesis techniques.

Finally, we have explored the conditioning of the problem as the amount of experimental data available increases. From this analysis it follows that it will be desirable to carry further research addressing the following issues:

- The step which takes M_R to Z (to make the problem affine in X) duplicates the size of the matrix and explicitly inverts M_0. While this step is desirable from a theoretical standpoint (to establish convexity of the problem), it could possibly be avoided by using a quadratic convex optimization program.
- The algorithm should not destroy the theoretical properties of the matrices involved. This is important when inverting or constructing complex hermitian matrices.
- A reliable algorithm which transforms complex realizations of real systems to real ones should be included in the general procedure.
- For the case of interior point optimization procedures, a procedure which "cheaply" computes good initial points by taking advantage of the problem structure should be introduced. In this regard, the procedure of using as a starting point the solution of a similar problem but with fewer interpolation points seems specially promising.
- Finally, as we indicated in Section 3, there are still degrees of freedom available in the problem. This raises the interesting possibility of using these degrees of freedom to optimize an additional performance criteria, for instance minimizing the order of the nominal model.

Acknowledgements—The work of the first and third authors was supported by UBACyT projects IN018 and IN049, University of Buenos Aires. Mario Sznaier was supported in part by NSF under grant ECS-9211169 and ECS-9625920.

REFERENCES

Ball, J. A., I. Gohberg and L. Rodman (1990). *Interpolation of Rational Matrix Functions*, Operator Theory: Advances and Applications, Vol. 45. Birkhäuser, Basel.

Boyd, S., L. El Ghaoui, E. Feron and V. Balakrishnan (1994). *Linear Matrix Inequalities in Systems and Control Theory*, SIAM Studies in Applied Mathematics, Vol 15. SIAM, Philadelphia.

Braatz, R. D., P. M. Young, J. C. Doyle and M. Morari (1994). Computational complexity of μ calculation. *IEEE Trans. Automat. Control*, **39**(5), 1000–1002.

Chen, J., J. A. Farrell, C. N. Nett and K. Zhou (1994). $\mathscr{H}_\infty$ Identification of multivariable systems by tangential interpolation methods. *Proc. Conf. on Decision and Control*, Orlando, FL.

Chen, J. and C. N. Nett (1995). The Carathéodory-Fejér problem and/identification: A time domain approach. *IEEE Trans. Automat. Control*, **40**(4), 729–735.

Chen, J., C. N. Nett and M. K. H. Fan (1992). Worst-case system identification in $\mathscr{H}_\infty$: validation of Apriori information, essentially optimal algorithms, and error bounds. *Proc. Amer. Control Conf.*, Chicago, IL.

Gu, G. and P. P. Khargonekar (1992). A class of algorithms for identification in $\mathscr{H}_\infty$. *Automatica*, **28**(2), 299–312.

Gu, G., P. P. Khargonekar and E. B. Lee (1989). Approximation of Infinite Dimensional Systems. *IEEE Trans. Automat. Control*, **34**(6), 610–618.

Hakvoort, R. (1992). Worst-case system identification in ℓ_1: error bounds, optimal models and model reduction. *Proc. Conf. on Decision and Control*, Tucson, AZ.

Helmicki, A. J., C. A. Jacobson and C. N. Nett (1991). Control oriented system identification: a worst case/deterministic approach in $\mathscr{H}_\infty$. *IEEE Trans. Automat. Control*, **36**(10), 1163–1176.

Hille, E. (1962). *Analytic Function Theory*. Vol. I. Chelsea, New York.

Jacobson, C. A., C. N. Nett and J. R. Partington (1992). Worst-case system identification in ℓ_1: optimal algorithms and error bounds. *System Control Lett.*, **19**(6), 419–424.

Ljung, L. (1987). *System Identification: Theory for the User*. Prentice-Hall, Englewood Cliffs, NJ.

Mäkilä, P. M. (1991a). Identification of stabilizable systems: closed-loop approximation. *Int. J. Control*, **54**(3), 577–592.

Mäkilä, P. M. (1991b). Robust identification and Galois sequences. *Int. J. Control*, **54**(5), 1189–1200.

Mäkilä, P. M. (1992). Worst-case input–output identification. *Int. J. Control*, **56**(3), 673–689.

Mäkilä, P. M. and J. R. Partington (1992). Robust identification of strongly stabilizable systems. *IEEE Trans. Automatic Control*, **37**(11), 1709–1716.

Mäkilä, P. M., J. R. Partington and T. K. Gustafsson (1995). Worst case control-relevant identification. *Automatica*, **31**, 1799–1819.

Milanese, M. (1994). Worst-case ℓ_1 identification. In *Bounding Approaches to System Identification*, eds Milanese, Norton, Piet-Lahanier and Walter. Plenum Press, New York.

Nesterov, Y. and A. Nemirovski (1990). Optimization over Positive Semidefinite Matrices: Mathematical Background and User's Manual. Central Econ. Mathematical Institute, USSR Acad. Science.

Nesterov, Y. and A. Nemirovski (1994). *Interior Point Polynomial Methods in Convex Programming: Theory and Applications.* SIAM Studies in Applied Mathematics, Vol. 13. SIAM, Philadelphia.

Overton, M. L. (1988). On minimizing the maximum eigenvalue of a symmetric matrix. *SIAM J. Matrix Anal. Appl.*, **9**(2), 256–268.

Parrilo, P. A., R. S. Sánchez Peña and C. G. Galarza (1994). Tuned ℓ_1 identification from impulse response data: application to a fluid dynamics problem. *Proc. American Control Conf.*, Baltimore, MD.

Parrilo, P. A. and R. S. Sánchez Peña (1995). Convex optimization based algorithms for Mixed time/frequency robust identification. *Proc. 2nd Latin American Seminar on Advanced Control,* Santiago, Chile.

Parrilo, P. A., M. Sznaier and R. S. Sánchez Peña (1996). Mixed time/frequency-domain based robust identification, *Proc. 35th Conf. Decision and Control*, Kobe, Japan.

Partington, J. R. (1992). Robust identification in $\mathscr{H}_\infty$. *J. Math. Anal. Appl.*, **166**, 428–441.

Poljak, S. and J. Rohn (1993). Checking robust nonsingularity is NP-hard. *Math. Control Signals Systems*, **6**(1), 1–9.

Rotstein, H. (1996). A Nevanlinna–Pick approach to time-domain constrained $\mathscr{H}_\infty$ control. *SIAM J. Control Optim.*, **34**(4), 1329–1341.

Sánchez Peña, R. S. and C. G. Galarza (1994). Practical issues in robust identification. *IEEE Trans. Control Systems Technol.*, **2**(1), 54–56.

Sánchez Peña, R. S. and M. Sznaier (1995). Robust identification with mixed time/frequency experiments: consistency and interpolation algorithms. *Proc. Conf. on Decision and Control*, New Orleans, LA.

Smith, R. S. and J. C. Doyle (1992). Model validation: a connection between robust control and identification. *IEEE Trans. Automat. Control*, **37**(7), 942–952.

Tangirala, S., A. Ray and M. Carpino (1995). Damage mitigating control of mechanical structures: experimental verification of the concept. *Proc. 1995 ACC*, Seattle, WA, 21–23 June, 1995, pp. 4106–4110.

Vandenberghe, L. and S. Boyd (1994). SP—Software for Semidefinite Programming. November 1994.

Vandenberghe, L. and S. Boyd (1996). Semidefinite programming. *SIAM Rev.* **38**, 49–95.

Wu, N. and G. Gu (1990). Discrete fourier transform and approximations. *IEEE Trans. Automat. Control*, **35**(9), 1044–1046.

Zhou, T. and H. Kimura (1993). Time domain identification for robust control. *System Control Lett.*, **20**, 167–178.

Zhou, T. and H. Kimura (1995). Structure of model uncertainty for a weakly corrupted plant. *IEEE Trans. Automat. Control*, **40**(4), 639–655.