WA08 11:20

Proceedings of the 34th
Conference on Decision & Control
New Orleans, LA - December 1995

Robust Identification with Mixed time/frequency experiments: Consistency and Interpolation algorithms

Ricardo S. Sánchez Peña [1] Mario Sznaier[2]

Abstract

In this paper we present a robust identification framework that combines both frequency and time response experiments. The main result of the paper provides necessary and sufficient conditions for consistency with both sources of experimental data in terms of the solution to a finite dimensional constrained convex optimization problem.

1. Introduction

One of the important research areas of the 90's is Robust Identification: deterministic identification procedures which can be used as a first step of a Robust control design method. The Robust Identification problem has been posed in [6] and has deserved considerable attention since then ([2, 5, 7]).

These new identification procedures are based not only on the data (*a posteriori* information), but also on the *a priori* assumptions on the class of systems to be identified. The algorithm produces a nominal model based on the experimental data and a *worst case* bound over the set of systems defined by the *a priori* information.

Starting from frequency response measurements, identification of exponentially stable LTI systems in the context of an $\mathcal{H}_\infty$ framework has been considered in [2, 6] and a complete class of robustly convergent algorithms has been presented in [5]. From time response data, ℓ_1 identification has been presented in [7] and references therein. Finally, interpolatory algorithms with time domain experiments and $\mathcal{H}_\infty$ identification errors have been developed recently [3, 11], although both types of experimental information, time and frequency, have not been considered.

In this paper we propose a robust identification framework that takes into account *both* time and frequency domain experiments. Thus, the problem were "good" frequency response fitting (small $\mathcal{H}_\infty$ error norm) has a "poor" fitting in the time domain, is avoided. Additionally, from an information theoretic point of view, more experiments produce a smaller consistency set of undistinguishable models, and as a consequence a smaller worst case error ([7]). The main result of the paper provides a necessary and sufficient condition for the consistency of the *a priori* information with the mixed time/frequency experimental data. This condition is given in terms of the solution to a finite–dimensional constrained convex optimization problem. Additionally, we propose an $\mathcal{H}_\infty$ interpolation algorithm which restricts the time response in the ℓ_∞ sense and produces a nominal model in the consistency set. Because this algorithm is interpolatory, it is optimal up to a factor of 2 with respect to strongly optimal *central* algorithms ([7]) and convergent. The problem is restricted to FIR systems or, in practical cases, to IIR systems when the time response experiment duration is longer than the settling time of the system. The general case has been solved in [10].

[1] CONAE, Av. Paseo Colón 751, (1063) Buenos Aires, Argentina and Depto. de Electrónica, Facultad de Ingeniería, Universidad de Buenos Aires, Argentina. email: ricardo@orbis.conae.gov.ar

[2] Dept. of Electrical Engineering, Penn State University, University Park, PA 16802. email: msznaier@frodo.ee.psu.edu.
Supported in part by NSF grant ECS-9211169

2. Background

The class of systems considered are continuous or discrete time, causal, linear and stable. To unify the treatment we denote them as $H(z) = H_c(\lambda\frac{1-z}{1+z})$, $\lambda > 0$ in the continuous case or $H(z) = H_d(\frac{1}{z})$ in the discrete time case, with $z \in \mathbb{C}$. Note that the above definition for continuous systems is the inverse of the usual bilinear transformation. Therefore, causal stable systems $H(z)$ will be analytic inside the unitary circle, with time and frequency representations related by $H(z) = \sum_{k=0}^{\infty} h_k z^k$. For simplicity we consider SISO models, although all results can be applied to MIMO systems, following [4].

The *a posteriori* experimental data can be obtained from two different sources. The set of N_f samples of the frequency response of the system measured with additive bounded noise $y_k^f = H(e^{j\Omega_k}) + \eta_k^f$, $k = 0, \ldots, N_f - 1$ represent the components of vector $\mathbf{y}^f \in \mathbb{C}^{N_f}$. The noise is complex and bounded by ϵ_f in the ℓ_∞ norm, therefore it belongs to the class $\mathcal{N}_f \triangleq \ell_\infty(\epsilon_f)$. The time domain data is the set of first N_t impulse response samples also corrupted by additive noise $y_n^t = h_n + \eta_n^t$, $n = 0, \ldots, N_t - 1$ which represent the components of vector $\mathbf{y}^t \in \mathbb{R}^{N_t}$. The noise is real and belongs to $\mathcal{N}_t \triangleq \ell_\infty(\epsilon_t)$.

To present the *a priori* information, consider the set $\mathcal{H}_\infty(\rho, K)$ defined in [6] which corresponds to exponentially stable systems (finite or infinite dimensional). From a practical viewpoint, these systems have a stability margin of $(\rho - 1)$ and a peak response to complex exponential inputs of K, and satisfy the bound $|h(k)| \leq K\rho^{-k}$.

In the time domain, a general class of models is defined as follows:

$$\Phi \triangleq \{h(\cdot) \mid |h(k)| \leq \phi(k), \ \ \phi(\cdot) \in \ell_1, \text{ nonincreasing}\} \quad (1)$$

which includes $\mathcal{H}_\infty(\rho, K)$ as a particular case, when $\phi(k) = K\rho^{-k}$. Therefore to combine both classes of models we define the set $\mathcal{S} \triangleq \mathcal{H}_\infty(\rho, K) \cap \Phi = \mathcal{H}_\infty(\rho, K)$.

For the experimental *a posteriori* information we consider, without loss of generality, the impulse as the input to the time domain experiment. This paper is restricted to the case of FIR systems or, from a practical viewpoint, a time response experimental duration $N_t \geq T_s$, with T_s the settling time of the system (see [10] for the general IIR case). If $\phi(N_t) \leq \epsilon_t$ then we have the number of data points which provides *all* the useful (time) experimental information ([9]), according to the *a priori* assumptions. For computational reasons (but without loss of generality) we consider the number of data points of both experiments as $N = \max(N_f, N_t)$ and the frequency points equispaced, i.e. $\Omega_k = \frac{2\pi}{N}k$. With these assumptions we can use the FFT algorithm and consider $H(e^{j\Omega_k}) = \sum_{n=0}^{N-1} h_n e^{jn\Omega_k}$ for the FIR case, or approximately in the case of IIR systems when we take N_t as mentioned above.

Therefore the model $H(z)$ in the frequency domain can be represented by the complex vector $\hat{h} = \left[\ H(e^{j\Omega_0}) \ \ \ldots \ \ H(e^{j\Omega_{N-1}})\ \right]^T$ and in the time domain by the real vector $h = \left[\ h_0 \ \ \ldots \ \ h_{N-1}\ \right]^T$. Both are related by the linear operations $\hat{h} = FFT(h)$ and $h = IFFT(\hat{h})$.

Here we approach $\mathcal{H}_\infty$ robust identification with a restricted behavior in the time domain. Specifically the time response of the model (step, impulse, etc.) should be limited in a weighted ℓ_∞ sense, according to the consistency in the time domain. The consistency and identification procedures will be attempted as follows.

Given the experiments $\mathbf{y}^f$ and $\mathbf{y}^t$ and the *a priori* sets $\mathcal{S}$, $\mathcal{N}_f$ and $\mathcal{N}_t$, determine:

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

$$\mathcal{S}(\mathbf{y}^f, \mathbf{y}^t) \triangleq \left\{ H \in \mathcal{S} \;\middle|\; \begin{array}{l} y_k^f - H(e^{j\Omega_k}) \in \mathcal{N}_f, \\ y_n^t - h_n \in \mathcal{N}_t, \end{array} \right\} \quad (2)$$

   is nonempty. Here $k = 0, \ldots, N_f - 1$ and $n = 0, \ldots, N_t - 1$.

2. Compute a nominal model which belongs to the consistency set by means of an interpolation algorithm.

3. Consistency

The consistency problem stated above can be recast into the following constrained interpolation form: Find a model $H \in \mathcal{S}$ which interpolates the frequency experimental data:

$$\hat{h} = \mathbf{y}^f + \eta^f, \qquad \eta^f \in \mathcal{N}_f \quad (3)$$

and has an impulse response restricted by

$$\mathbf{y}_L \leq h \leq \mathbf{y}_U \quad (4)$$

$$(\mathbf{y}_L)_j = \max\left[y_j^t - \epsilon_t, -K\rho^{-j}\right] \quad (5)$$

$$(\mathbf{y}_U)_j = \min\left[y_j^t + \epsilon_t, K\rho^{-j}\right] \quad (6)$$

In the above, we have considered without loss of generality, an impulse response experiment. The case of a generic input $u^T = \left[\ u_0 \ \ u_1 \ \ \ldots \ \ u_n\ \right]$ can be handled by simply replacing h by Uh in (4), were

$$U = \begin{bmatrix} u_0 & 0 & \ldots & 0 \\ u_1 & u_0 & \ldots & 0 \\ \vdots & \ddots & \ddots & 0 \\ u_n & \ldots & u_1 & u_0 \end{bmatrix} \quad (7)$$

Also in the general case were the experimental points $\mathbf{y}^t$ in (4) do not coincide with the N points of vector h to be determined, the necessary data points can be interpolated from the experiment. As mentioned before, we are considering the case were each h_k has a corresponding experimental datum y_k^t, for $k = 0, \ldots, N-1$.

Next we define an equivalent condition for consistency, based on the relation between both admissible experimental noises $\eta_f \in \mathcal{N}_f$ and $\eta_t \in \mathcal{N}_t$.

Lemma 3.1 *Both* a priori *and* a posteriori *information are consistent if and only if there exists a function $H \in \mathcal{S}$ with frequency points $\hat{h} = \mathbf{y}^f + \eta^f$, such that $\eta^f \in \Lambda$, the latter being defined by:*

$$\Lambda \triangleq \{\eta^f \in \mathcal{N}_f \mid \mathbf{y}_L - IDFT(\mathbf{y}^f) \leq IDFT(\eta^f) \leq \mathbf{y}_U - IDFT(\mathbf{y}^f)\}$$

a convex subset of $\mathcal{N}_f$.

The next theorem from [2] provides an equivalent condition for the existence of a function $H \in \mathcal{S}$ which satisfies the unconstrained problem (3).

Theorem 3.1 *There exists $H \in \mathcal{S}$ which interpolates the data points $\hat{h} = \mathbf{y}^f + \eta^f$ as in (3) if and only if $\bar{\sigma}\left(Q^{-\frac{1}{2}}\ diag[\hat{h}]Q^{\frac{1}{2}}\right) \leq K$ and $Q > 0$, where*

$$Q_{ij} = \frac{1}{1 - e^{j\Omega_i}e^{-j\Omega_j}/\rho^2} \quad i,j = 0, \ldots, N-1 \quad (8)$$

and where $diag[a]$ *denotes a diagonal matrix whose main diagonal is vector* a.

Using this result we can state now the main result:

Theorem 3.2 *The mixed time/frequency robust identification problem presented so far is consistent if and only if* $\Lambda \neq \emptyset$ *and,*

$$\inf_{\eta^f \in \Lambda} \bar{\sigma} \left(Q^{-\frac{1}{2}}\ diag[\mathrm{y}^f + \eta^f] Q^{\frac{1}{2}} \right) \leq K \qquad (9)$$

The above is a linearly constrained convex (nondifferentiable) optimization problem.

Corollary 3.1 *The mixed time/frequency robust identification problem is consistent if the following are satisfied:*

$$\mathrm{y}_L \leq IDFT(\mathrm{y}^f) \leq \mathrm{y}_U \qquad (10)$$

$$\bar{\sigma} \left(Q^{-\frac{1}{2}}\ diag[\mathrm{y}^f] Q^{\frac{1}{2}} \right) \leq K \qquad (11)$$

Furthermore, the models which belong to this subset of $\mathcal{S}(\mathrm{y}^f, \mathrm{y}^t)$ *interpolate exactly the experiment* y^f.

4. Robust Identification

Once consistency is established, the second step can be attempted. Using the above results, we obtain an identification procedure which produces a nominal model in the consistency set $\mathcal{S}(\mathrm{y}^f, \mathrm{y}^t)$. This type of algorithms are called interpolation algorithms and have several advantages over the usual "two step" ones ([5, 6]). Because the identified model is in set $\mathcal{S}(\mathrm{y}^f, \mathrm{y}^t)$, its distance to the Chebyshev center of this set is within the diameter of information ([7]). As a consequence the algorithm is optimal up to a factor of 2 as compared with central strongly optimal procedures. For the same reasons, it is also convergent and therefore the modelling error tends to zero as the information is completed.

The procedure is similar in spirit to the one in [2] and is based in the parameterization of all solutions of the Nevanlinna-Pick problem ([1]). Here we consider the generic case for which the Pick matrix is strictly positive definite and therefore the solution is nonunique. The degenerate case for which there is a unique solution can be found in [1, 2].

- Find a noise vector $\eta^f_* \in \Lambda$ such that

$$\bar{\sigma} \left(Q^{-\frac{1}{2}}\ \mathrm{diag} \left[\mathrm{y}^f + \eta^f_* \right] Q^{\frac{1}{2}} \right) < K \qquad (12)$$

Note that there is no need to find the optimal in (9), but any vector η^f in the admissible set Λ which achieves strict inequality.

- Compute the Pick matrix

$$P_{ij} = \frac{1 - \frac{1}{K^2} H(e^{j\Omega_i}) H(e^{-j\Omega_j})}{1 - \frac{1}{\rho^2} e^{j(\Omega_i - \Omega_j)}}, \qquad (13)$$

(which should be positive definite), replacing $H(e^{j\Omega_k}) = y^f_k + \eta^f_{*k}$.

- Compute the interpolating solution

$$H(z) = K \frac{T_{11}(\frac{z}{\rho}) q(z) + T_{12}(\frac{z}{\rho})}{T_{21}(\frac{z}{\rho}) q(z) + T_{22}(\frac{z}{\rho})} \qquad (14)$$

for $q \in \overline{\mathcal{B}}\mathcal{H}_\infty$ and $T(z)$ defined in [1]. If $q(z)$ is constant, $H(z)$ is of order N. The multiplicity of solutions depends on the choice of $q(z)$ and of $T(z)$, which in turn depends on $\eta^f \in \Lambda$. This suggests a further optimization step to select the optimal η^f and $q(z)$, which minimizes a certain criteria. This problem can also be solved via convex optimization.

References

[1] Ball J., Gohberg I., Rodman L., Interpolation of Rational Matrix Functions, Operator Theory: Advances and Applications, Vol. 45, Birkhäuser, 1990.

[2] J. Chen, C. Nett, M. Fan, "Worst-Case System Identification in $\mathcal{H}_\infty$: Validation of A Priori Information, Essentially Optimal Algorithms, and Error Bounds," *Proceedings 1992 ACC*, Chicago, IL, 1992.

[3] J. Chen, C. Nett, "The Carathéodory-Fejér problem and the $\mathcal{H}_\infty/\ell_1$ identification: A time domain approach", *IEEE Transactions Automatic Control*, Vol. 40, No. 4, 1995.

[4] J. Chen, J. Farrell, C. Nett, K. Zhou, "$\mathcal{H}_\infty$ Identification of Multivariable systems by tangential interpolation methods", *Proceedings of the Conference on Decision & Control*, 1994.

[5] G. Gu, P. Khargonekar, "A class of Algorithms for Identification in $\mathcal{H}_\infty$," *Automatica*, Vol. 28, No. 2, 1992.

[6] J. Helmicki, C. Jacobson, C. Nett, "Control Oriented System Identification: A worst case/deterministic approach in $\mathcal{H}_\infty$," *IEEE Transactions Automatic Control*, Vol. 36, No. 10, 1991.

[7] P. Mäkilä, "Worst-case input-output identification," *International Journal of Control*, Vol. 56, No. 3, 1992.

[8] Nesterov Y., Nemirovsky A., "Optimization over positive semidefinite matrices: Mathematical Background and User's Manual", USSR Acad. Science Central Econ. Mathematical Institute, 1990.

[9] P. Parrilo, R. Sánchez Peña, C. Galarza, "Tuned ℓ_1 Identification from Impulse Response Data: Application to a Fluid Dynamics Problem," *Proceedings 1994 ACC*, Baltimore, 1994.

[10] P. Parrilo, M. Sznaier, R. Sánchez Peña, "Mixed time/frequency Robust Identification," submitted.

[11] T. Zhou, H. Kimura, "Time domain identification for Robust control", *System & Control Letters*, Vol. 20, pp. 167-178, 1993.