Proceedings of the 36th
Conference on Decision & Control
San Diego, California USA • December 1997

System Identification of Complex Systems: Problem Formulation and Results [1]

Saligrama R. Venkatesh Munther A. Dahleh
Laboratory for Information and Decision Systems
Massachusetts Institute of Technology

Abstract

We are motivated by the need to derive simple control-oriented models of complex-systems. Briefly, complex-systems are a collection of systems that are not approximable (arbitrarily closely) by a finitely parameterized collection of systems. The study of system identification for such classes of systems has received considerable attention in the past few years. In this paper a new formulation of the system identification problem, which balances between the set-membership and probabilistic approaches is presented. The salient feature of the formulation is that we distinguish between the two principal sources of error encountered in the input-output data—noise and unmodeled dynamics. Unmodeled dynamics arises from the fact that the finitely parameterized model that we seek does not truly characterize the real system. Therefore, unmodeled dynamics is "modeled" as the residual error between the parametric model and the real system. This viewpoint leads to a decomposition between the parametric model class and unmodeled dynamics. In contrast noise is modeled so that it is uncorrelated (in a deterministic or a stochastic sense) from the input. The identification problem deals with obtaining the appropriate finite parametric model from input-output data. The identification problem is studied for several different norms including ℓ_1 and $\mathcal{H}_\infty$. One of the chief outcomes is a new notion of a persistent input and showing that there are both deterministic and stochastic inputs which meet the new criterion.

1 Introduction

The research reported in this paper is motivated by the need for deriving simple control-oriented models of complex systems. This area of research has received a lot of attention in the past few years and work along these lines has provided alternative formulations to the system identification problem (e.g. [4, 1, 10, 9, 6, 7, 5, 3, 8] and references therein).

[1]This research was supported by NSF Grant number 9157306-ECS, AFOSR F49620-95-0219, and, Siemens AG. The first author is affiliated with the United Technologies Research Center

In this paper, we will give an account of a new formulation for the system identification problem for complex systems. Complex systems ($\mathcal{T}$) are ones that cannot be uniformly approximated by a finite dimensional space. Nevertheless, we represent our *prejudice* by selecting a finitely parameterized set of models ($\mathcal{G}$) from which an estimate of the original system will ultimately be drawn. We will assume that an estimate of the distance (in some norm) between the actual process and this set is available as part of the prior information whenever we want to determine the sample-complexity.

If the actual process is known, then selecting a model in $\mathcal{G}$ that best approximates $T \in \mathcal{T}$ in some norm is a straightforward convex optimization problem. Hence, given any $T_0 \in \mathcal{T}$, we can write $T = G_0 + \Delta_0$ where

$$G_0 = \arg\min_{G \in \mathcal{G}} \|T_0 - G\|$$

(for simplicity, assume the above minimization is unique).

In system identification, however, the process is not known and only a finite set of input-output data is available. We will assume that this set of data is generated as:

$$y(k) = T_0 u(k) + w(k) = G_0 u(k) + \Delta_0 u(k) + w(k), \quad k \leq N \tag{1}$$

where u is the input (experiment) and w is a noise set. The objective of this development is to show, for rich classes of noise sets (either stochastic or deterministic that include white noise with high probability), how to select an input experiment u and an algorithm that picks an estimate $\hat{G}_N \in \mathcal{G}$ such that $\|G_0 - \hat{G}_N\|$ approaches zero in a reasonable length of time (hopefully with polynomial sample complexity). In other words, the derived algorithm used with the derived input provides a method for solving the actual approximation problem only from input-output data.

To contrast this with classical formulations, recall that in classical stochastic identification formulations the real-physical system is modeled in a suitably, finitely, parameterized model class and the measurements are assumed to be corrupted by noise uncorrelated with the input. Therefore, the estimation errors are assessed as though the the real system belongs to the chosen parametric class. Consequently, the discrepancy between the real-physical system and the model is hypothesized to be parametric. Under-modeling is not incorporated in the problem formulation.

In set-membership identification (see [2, 4, 7, 9, 10] and references there in), the assumption is that the real system is in a known set which is characterized by unknown parameters and unknown but bounded non-parametric (unmodeled) dynamics. In addition it is assumed that the measured data is corrupted by noise belonging to a known set. The set of all parameters consistent with prior information and experimental data is then computed. Owing to the fact that such a general identification problem is difficult the problem is simplified by modeling the effect of unmodeled dynamics and noise by imposing temporal constraint on input-output data. However, such an analysis leads to conservative set estimates for the parameters and even in the limit of infinite data the parameters can only be guaranteed to lie in some set of diameter at least the size of prior bound on the unmodeled dynamics.

In summary, the above standard formulations occupy two extreme positions: the first assumes crisp prior information and the latter assumes overly coarse prior information. Our formulation provides a compromise between these two and incorporates the under-modeling as part of the problem formulation.

Our work continues along the lines of [14] and distinguishes between the two sources of error—unmodeled dynamics and noise. We define a natural notion of separation between the parametric part and unmodeled dynamics. This notion arises naturally if the parametric part is a subspace of a linear space of systems. As mentioned earlier, the idea behind the separation is that the real system belongs to some large space of systems $\mathcal{T}$ and if the model is to be obtained in some subspace, $\mathcal{G}$, then the uncertainty is isometrically isomorphic to the quotient $\mathcal{T}/\mathcal{G}$ (a mathematical consequence of the optimization problem mentioned earlier). The noise is assumed to belong to a set of signals that are uncorrelated with the input (in either a deterministic or stochastic sense) is also rich enough to contain white noise sequences. Equipped with the triple—space of systems, parametric representation, and noise we study consistency, sample complexity and algorithms. We emphasize that consistency is studied with respect to aforementioned decomposition.

We first delineate the problem setup. We do not go into the reasons for the need for a new formulation here other than what has already been mentioned. The interested reader is referred to [11] for an elaborate discussion. The formulation has two aspects—consistency and sample-complexity. Consistency is achieved if estimates converge in the model class. The sample-complexity is said to be polynomial if the length of the input is a polynomial function of the prior error between the estimate and the true model and the size of the model class. The second aspect concerns the study of solutions to the formulated identification problem for LTI stable systems equipped with $\mathcal{H}_2$, $\mathcal{H}_\infty$, ℓ_1 topologies. An important issue that arises in this context is input design. We will construct an input that we term a "robust" input with the property that its n^{th} autocorrelation terms decay uniformly in n, with a polynomial rate of decay. Equipped with this input we focus on the solution to the identification problem for stable systems belonging to $\mathcal{H}_2$ and show that consistency in the finitely parameterized model class can be achieved at a polynomial rate in the presence of noise. In addition, we delineate a systematic solution that can be applied to all systems belonging to a Hilbert-Space.

We next take up $\mathcal{H}_\infty$ identification for which no satisfactory solution is shown to exist. In particular, it is shown that identification of finite-dimensional models for systems in the $\mathcal{H}_\infty$ space has infinite sample-complexity. This, in a sense, implies that the estimate of the distance between an element of this space and the finitely parameterized set of models is too coarse if given in terms of the $\mathcal{H}_\infty$ norm. To this end we formulate the problem on a Hardy-Sobelev space which is closely related to $\mathcal{H}_\infty$ space. This space happily is a Hilbert-Space and so our identification solution can be readily applied.

Finally, we solve the ℓ_1 identification problem and show that it has polynomial sample-complexity.

2 Problem Setup

Before we present the problem certain assumptions about the system and the data generation has to be made. An elaborate discussion on these topics can be found in [11, 15]. We will discuss these issues very briefly here for the sake of completion.

We assume that the system belongs to a complex collection of systems, $\mathcal{I}(\gamma)$, which is a subset a normed linear vector space $\mathcal{T}$. We assume the following structure for the prior, $\mathcal{I}(g)$:

$$\mathcal{I}(\gamma) = \{T \in \mathcal{T} \mid \|T - G\| \leq \gamma, \; G \in \mathcal{G}\} \qquad (2)$$

where $\mathcal{G}$ is some subspace of $\mathcal{T}$. The prior is based on the belief that the unmodeled dynamics, when the system is sought to be modeled in a space $\mathcal{G}$, doesn't exceed γ.

In practice, it is possible to estimate a lower bound for γ from input-output data. Figure 1 illustrates the relationship between the various entities we have discussed

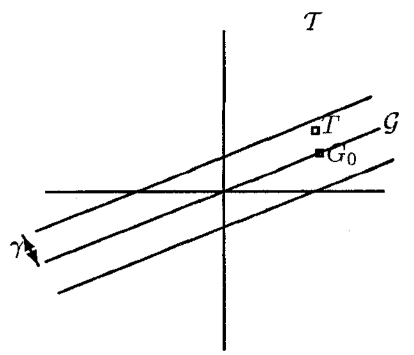

Figure 1: Relationship between the real-system, T, and the model-space $\mathcal{G}$

so far. This type of prior, $\mathcal{I}(\gamma)$ facilitates a decomposition between the real system and the model space $\mathcal{G}$ as shown in the following theorem.

Proposition 1 Suppose $\mathcal{T}$ is a normed linear vector space, $\mathcal{G}$, a subspace of $\mathcal{T}$, and $\mathcal{G}^{\perp}$ the annihilator of $\mathcal{G}$ in the dual $\mathcal{T}^*$ of $\mathcal{T}$. Then the following statements are equivalent:

1. $G \in \text{argmin}_{G' \in \mathcal{G}} \|T - G'\|$

2. $T = G + \Delta$ for some $G \in \mathcal{G}$ and $< \phi, \Delta >= \|\Delta\|$ for some $\phi \in \mathcal{G}^{\perp}$ and $\|\phi\| \leq 1$.

■

Thus the unmodeled dynamics is characterized by its alignment to the annihilators of $\mathcal{G}$ in the special case when $\mathcal{G}$ is the subspace of $\mathcal{T}$.

We now move on to the second aspect—that of generation of input-output data. We assume that the input-output data is generated as in Equation 1 and, to keep the treatment simple, we assume that the noise enters additively at the output of the system. Noise is modeled by imposing temporal constraints on sample paths. This is the case for instance in the probabilistic setting. It has been found to be particularly fruitful in the deterministic setting as evidenced from [11, 15, 12, 14]. For simplicity we deal with the deterministic case alone here. Herein sample paths for noise are assumed to belong to an a priori known deterministic set, $\mathcal{W}$. The main problem with coming up with useful descriptions for noise is that if the set is too large, for instance, persistent bounded set like $\mathcal{B}\ell_\infty$, noise and input can be completely correlated. If the set is too small, for instance, $\mathcal{B}\ell_2$, we sacrifice persistency. Thus any set-based description for noise has to balance these the two extremes. For a further discussion on this topic the reader is referred to [14]. One model that balances these extremes is the following:

$$\mathcal{W}_N = \left\{ w \in \mathbb{R}^N \mid \sup_{P \in \mathcal{P}^m} \left| \frac{1}{\sqrt{N}\log(N)} \sum_{t=0}^{N} w(t) e^{iP(t)} \right| \leq 1 \right\} \tag{3}$$

where $\mathcal{P}^m$ is the class of polynomials in t of order m over the field of reals. Note that the above model is characterized by linear constraints on finite truncations of noise, w, and it follows that such a class is convex. The following theorem verifies that such a noise model is rich enough to contain typical sample paths generated by an i.i.d. process.

Theorem 1 [Richness]
Suppose $x(0)$, $x(1)$, $x(2)$, ... is a discrete time random process (white gaussian process or bernoulli process) with mean 0 and bounded variance. Then:

$$\mathcal{P}(P_N x(t) \in \mathcal{W}_N) \overset{N \to \infty}{\longrightarrow} 1 \tag{4}$$

where $\mathcal{W}_N$ is given by Equation 3.

Note that an alternate description in the form of linear constraints can also be given. In other words the set $\mathcal{W}$ contains all elements that satisfy:

$$\mathcal{F}_n w \leq 1 \tag{5}$$

for some linear operator $\mathcal{F}_n$ that maps $\mathbb{R}^n$ to $\mathcal{L}_\infty$.

We are now ready to define the identification problem.

Problem 1 (Identification) *Suppose the output satisfies Equation 1 for every input u and the noise w belongs to the set $\mathcal{W}$ given by Equation 3. Determine an input u and an algorithm $\hat{G}^n$ so:*

1. *$\hat{G}^n$ is consistent:*

$$\limsup_{n} \sup_{w \in \mathcal{W}_n} \|T - \hat{G}^n\| \leq \|T - G'\|, \ \forall G' \in \mathcal{G}, \ \forall T \in \mathcal{I} \tag{6}$$

2. *Sample complexity: Given any $\epsilon, \gamma > 0$ obtain $\mathcal{N}(\epsilon, \gamma) \in Z^+$ such that:*

$$\sup_{w \in \mathcal{W}_n} \sup_{T \in \mathcal{I}(\gamma)} \left| \|T - \hat{G}^n\| - \|T - G(T)\| \right| \leq \epsilon, \ \forall n \geq \mathcal{N}(\epsilon, \gamma) \tag{7}$$

where $\mathcal{I}(\gamma)$ is given by Equation 2.

Remark: There are several comments that are to be made in reference to the formulation.

- A minimizer $G(T)$ will exist for every T on account of finite dimensionality, however, it need not be unique and the problem has been posed accordingly.

- The sample-complexity is parameterized by γ, an a priori known upper-bound of the distance between the system, T, and the subspace $\mathcal{G}$. As noted earlier it is possible to estimate γ from input-output data.

3 Results

An important aspect of identification problem is the choice of input. In statistical identification, for the case of LTI systems, any input which satisfies a persistency of excitation condition is sufficient to guarantee consistency. However, in our case, since we need to robustly estimate the parameters in the face of unmodeled dynamics this property will not suffice. A complete study of input selection for identification of complex-systems is carried out in [13]. We only touch upon this topic here for the sake of completion. To motivate the discussion consider the following example.

Example 1 :

Suppose the real system, T, is an arbitrary LTI stable system, the model space is the FIR space of length m, and an experiment and model space are given respectively as:

$$\begin{aligned} y &= Tu + w, \quad T \in \{(t(k)) \in \ell_1\} \quad (8) \\ \mathcal{G} &= \{(g(k))_{k=0}^{M}\} \end{aligned}$$

Consider identification of the FIR model in the ℓ_1 norm. The appropriate model and unmodeled dynamics is:

$$G(T) = \text{argmin}_{G \in \mathcal{G}} \|T - G\|_1 = (t(k))_{k=0}^{M} \quad (9)$$

Thus, from Proposition 1 it follows that Δ is the tail of the impulse response of T, i.e.

$$\Delta(T) = (t(k))_{k>M} \quad (10)$$

and so we have:

$$y(s) = \sum_{k=0}^{M} t(k)u(s-k) + \sum_{k=M+1}^{s} t(k)u(s-k) + w(s) \quad (11)$$

the annihilators, v_k, $k \leq M$, for Δ are:

$$v_k = \begin{pmatrix} 0 & \ldots & \underset{k}{1} & 0 & \ldots \end{pmatrix} \quad (12)$$

Notice the following dichotomy in the input selection strategy. In the absence of noise a pulse input of unit-amplitude is sufficient. For, we may "filter" the output y resulting from the pulse input through the annihilators. This results in identification of the model because:

$$\hat{g}(k) = < v_k, y > = t(k), \; k \leq M \quad (13)$$

However, in the presence of noise a long persistently exciting input is necessary to average out the noise and this results in exciting the unmodeled dynamics. A case in point is a periodic input u. Suppose, u, with $u(s+l) = u(s)$, is one such input, the experiment results in aliasing unmodeled dynamics with the model rendering identification of the model impossible. In other words for $s \geq 0$:

$$\begin{aligned} y(s) &= \sum_{k=0}^{s} t(k)u(s-k) + w(s) \quad (14) \\ &= (t(0) + t(l) + \ldots)u(0) \\ &\quad + (t(1) + t(l+1) + \ldots)u(1) + \ldots + w(s) \end{aligned}$$

Thus one obtains information only on linear combinations of the $t(k)$'s and not individual coefficients. We don't comment on this example any further and revisit it at the end of the section where we will comment on a systematic way to cancel unmodeled dynamics and noise. At this point all that the reader is asked to do is to make a note of the fact that separating unmodeled dynamics and noise is non-trivial.

■

Remark: It is worth noting that in the set-membership literature, often, the unmodeled dynamics is modeled by explicitly assuming a decay rate for the unmodeled part. Such an assumption with our model for noise will reduce the problem to one where the real system can be modeled as an FIR and therefore the system will no longer be complex.

The example above can be generalized further to the case of identification of limited complexity models for space of LTI systems and we can derive necessary and sufficient conditions on the inputs required which we state here in the form of the following theorem.

Theorem 2 A necessary condition for an input to satisfy the consistency condition in Problem 1 is if there exists a sequence $q_n \in \ell_\infty$ satisfying:

$$\sup_{w \in \mathcal{W}_n} |\langle q^n, w\rangle| \leq \epsilon, \; \forall n \geq N(\epsilon, \gamma) \quad (15)$$

$$\|q^n\|_\infty \leq C/\sqrt{n}\log(n) \quad (16)$$

$$\langle q^n, u\rangle = 1 \quad (17)$$

$$\max_{0<k\leq n} |\langle q^n, \lambda^k u\rangle| \leq \epsilon/\gamma, \; \forall n \geq N(\epsilon, \gamma) \quad (18)$$

where C is some constant and the notation $< x, y >$ is given by (when the limit exists):

$$< x, y >= \lim_{n \to \infty} \sum_{k=0}^{n} x(t)y(t)$$

We term any input that meets the necessary condition a robust input. It follows from the above theorem that a robust input be persistently exciting of infinite order which we state in the form of a corollary below:

Corollary 1 *Robust inputs are persistently exciting of infinite order.*

Based on these conditions we classify commonly used inputs based on the criteria of robustness.

Bernoulli	robust
Periodic	not robust
$\exp(i\alpha t^2)$	not robust
PRBS	not robust
$\exp(i\alpha t^3)$	robust

The fact that the higher order chirp, $\exp(i\alpha t^3)$ is robust depends on a theorem by Hardy-Littlewood and is derived in [13, 11]. The cross-correlations of the higher-order chirp vanish uniformly at a polynomial rate in the length of the input. This is all we need to average out the transients and noise. The convolution operation being commutative allows us to rewrite Equation 11 as follows.

$$\hat{g}^n(j) = \frac{1}{n}\sum_{k=0}^{n} z(k)u(k) =< v_j, \frac{1}{n}\sum_{k=1}^{n}(\lambda^{-k}P_n(y))u^*(k) > \quad (19)$$

where the operator $\lambda^{-k}P_n$ is given by

$$\lambda^{-k}P_N(y) = (y(k),\, y(k+1),\, \ldots,\, y(n),\, 0,\, \ldots)' \quad (20)$$

Simplifying Equation 19 we obtain contributions from unmodeled dynamics and noise (as in Equation 3), i.e., the estimate can be verified to satisfy:

$$\hat{g}^n(j) = t(j) + \sum_{\tau \geq 1} t(\tau)r^n_{uy}(\tau) + r^n_{uw}(0) \quad (21)$$

It is now straightforward to prove that:

$$|\hat{g}^n(j) - t(j)| \leq \gamma\frac{\log(n)}{\sqrt{n}} \quad (22)$$

In general we have the following theorem:

Theorem 3 For the experiment (Equation 1) with noise belonging to the set of Equation 3 and input being the higher order chirp it follows that identification of FIR models in ℓ_1 norm has polynomial sample-complexity. Alternatively, if $\mathcal{G}$ is the space of FIR models of length m, and $\min_{G \in \mathcal{G}} \|T - G\| \leq \gamma$ then

$$\|\hat{G}^n - G(T)\|_1 \leq \epsilon;\ \forall n \geq m(\gamma/\epsilon)^2 \quad (23)$$

∎

The steps involved here can be summarized as follows.

- Obtain a decomposition of the system in to a model and unmodeled dynamics.
- Determine the annihilator (filter) of the unmodeled part.
- Pass the output through filter.
- Cross-correlate the filtered output with the input.

We have applied these steps in more general contexts and met with much success. We describe these in the sequel.

Suppose the real system, T, belongs to the class of MIMO linear shift-invariant operators, $\mathcal{T}^{p\times q}$. Such operators take inputs, u, in $\ell^q_{\infty,e}$, to outputs, y_T, in $\ell^p_{\infty,e}$. They are characterized by the following equation:

$$y_T(t) = (Tu)(t) = \sum_{k=0}^{t} t(k)u(t-k) \quad (24)$$

where $t(k) \in \mathbb{R}^{p\times q}$ is called the kernel of T. The model space $\mathcal{G}$ is a mq dimensional subspace of $\mathcal{T}^{p\times q}$ given by:

$$\mathcal{G} = \left\{ G \in \mathcal{T}^{p\times q} \mid G \equiv \left[\begin{array}{c|c} \mathrm{A} & \mathrm{B} \\ \hline \mathrm{C} & 0 \end{array}\right],\ B \in \mathbb{R}^{mq} \right\} \quad (25)$$

where A and C are fixed a priori. We solve the identification problem outlined earlier in Section 2 for a variety of norms/spaces. The solutions have been tabulated in the table below and for proofs the reader is referred to [11, 12].

Space	Sample-Complexity
$\mathcal{H}_2$	Polynomial
ℓ_1	Polynomial
$\mathcal{H}_\infty$	infinite
$\mathcal{H}^{2,1}$	Polynomial

The table calls for several remarks. First, in arriving at the polynomial complexity results it turns out that the recipe (algorithms) used was exactly the one outlined earlier. Second, we prove using a counterexample that the space $\mathcal{H}_\infty$ has infinite sample-complexity.

This is not surprising considering the fact that even in the case of non-noisy data the problem boils down to approximating an arbitrary continuous function with finite data. Finally, the symbol $\mathcal{H}^{2,1}$ stands for the Hardy-Sobelev space. These spaces are given by:

$$\mathcal{H}^{2,1} = \left\{ T \in \mathcal{T} \mid \sum_{k=0}^{\infty} (k^2+1)\text{trace}(t(k)t^*(k)) < \infty \right\} \tag{26}$$

It can be shown (see [11]) that for such spaces (unlike $\mathcal{H}_2$) a parallel robust control theory can be developed. In this way such spaces are unique in the respect that both identification and robust control synthesis is simple.

4 Conclusions

We have conducted a theoretical study of the identification problem ranging from set-membership id to classical identification. One outgrowth of this study has been the recognition that the current set-membership identification goals are too ambitious. Our focus has been to impose structure on the space of systems to obtain "high-fidelity" models with a short length of data. An outcome of imposing such structures is a natural definition of separation between a model for a system and the residual error between the model and the system. This plays a major role in all our identification schemes. It is important to stress here that we are interested in obtaining consistency in the model space as opposed to the real system. Our sample-complexity results pertain to how fast can one identify an "appropriate" model of a system in a finite-dimensional model space with particular cognizance to the fact that the system doesn't necessarily belong to the finite-dimensional space.

Acknowledgments

The authors are grateful to Sanjoy Mitter and Fernando Paganini for valuable discussions and comments.

References

[1] E. W. Bai, K. M. Nagpal, and R. Tempo. Bounded error parameter estimation: noise models, recursive algorithms and $\mathcal{H}_\infty$ optimality. *Proc. of Automatic Control Conference*, pages 3065–3069, 1995.

[2] E. Fogel and Y. F. Huang. On the value of information in system identification–bounded noise case. *Automatica*, vol.18, no.2:229–238, 1982.

[3] G. Gu and P. P. Khargonekar. Linear and nonlinear algorithms for identification in $\mathcal{H}_\infty$ with error bounds. *IEEE Trans. A-C*, Vol 37, No. 7, 1992.

[4] R. L. Kosut, M. K. Lau, and S. Boyd. Set membership identification of systems with parametric and nonparametric uncertainty. *IEEE Trans. A-C*, Vol 37, No. 7, 1992.

[5] P. M. Makila and J. R. Partington. Robust approximation and identification in $\mathcal{H}_\infty$. *Proceedings of American Control Conference*, 1991.

[6] M. Milanese and G. Belforte. Estimation theory and uncertainty intervals evaluation in the presence of unknown but bounded errors: linear family of models and estimators. *IEEE Trans. A-C*, 27:408–414, 1982.

[7] M. Milanese and A. Vicino. Optimal estimation theory for dynamic systems with set-membership uncertainty: an overview. *Automatica*, 27:997–1009, 1991.

[8] P. Poolla and A. Tikku. On the time complexity of worst–case system identification. *IEEE Trans. A-C*, vol. 39, no. 5:944–50, 1994.

[9] F. C. Schweppe. *Uncertain Dynamical Systems.* Prentice Hall, Englewood Cliffs, NJ, 1973.

[10] D. N. C. Tse, Dahleh M. A., and Tsitsiklis J. N. Optimal identification under bounded disturbances. *IEEE Transactions on A-C*, 38:1176–90, 1993.

[11] S. R. Venkatesh. System identification for complex systems. *Ph.D thesis, Massachusetts Institute of Technology*, 1997.

[12] S. R. Venkatesh and Dahleh M. A. Identification of complex-systems with limited complexity models. *IEEE Transactions on Automatic Control*, submitted.

[13] S. R. Venkatesh and Dahleh M. A. Inputs for robust identification of limited complexity models for complex-systems. *IEEE Transactions on Automatic Control*, to be submitted.

[14] S. R. Venkatesh and Dahleh M. A. Classical identification in a deterministic setting. *Proceedings of the 1995 Control and Decision Conference*, 1995.

[15] S. R. Venkatesh and Dahleh M. A. Identification in the presence of unmodeled dynamics and noise. *IEEE Transactions on Automatic Control*, August, 1997.