# The Determinability of Classes of Noisy Channels

By LEONARD J. FORYS

(Manuscript received June 30, 1969)

*This paper is concerned with the identification of a fairly general class of nonlinear operators using corrupted measurements. A precise mathematical definition of identification is presented and the relationship between a priori information and identification is studied. The a priori information is represented as a subset of a metric space of nonlinear operators. Necessary and sufficient conditions are developed to answer the question "When is identification possible?"*

## I. INTRODUCTION

A large body of literature already exists for the problem of identifying a control system or communication channel with noisy measurements. In the usual identification problems, a certain structure is assumed at the outset in order to reduce the identification problem to one of parameter estimation. The absence of such parametrization increases the difficulty of the problem substantially. It is often not clear if identification is even possible.

In this paper we are concerned with the determinability (identifiability) of quite general nonlinear operators whose outputs are corrupted by additive gaussian noise. We introduce a norm on this space of nonlinear operators and define precisely what we mean by determinability. Loosely speaking, we say that we can determine an operator $H$ if we can choose a finite observation interval $[0, T]$, a test signal with constrained peak value over this interval, a finite set of linear measurements over $[0, T]$, and an estimate $\hat{H}$ of $H$ which is a continuous function of our measurements such that $\hat{H}$ is close to $H$ in norm with high probability.

The question of determinability is of course intimately related to the kind of *a priori* knowledge one has of the operator. We represent this *a priori* information by saying that the operator $H$ belongs to a subset

$\mathfrak{D}$ of possible operators. We derive conditions on $\mathfrak{D}$ which are sufficient for determinability. We also show that most of these conditions are in fact necessary for the determination of $H$.

Our results are motivated by the work on the determinability of noiseless channels done by Root, Prosser, and Varaiya.[1-4] They derive necessary and sufficient conditions to estimate a noiseless channel closely with a "one-shot" experiment. These conditions are similar to those presented here. Some work on the noisy problem has been done by Root.[5] His approach and results are fundamentally different than those presented in this paper. Root investigated a class of stochastic nonlinear operators represented by a Volterra series whose kernels are gaussian random variables. He derived necessary and sufficient conditions for the second moments of the kernels to be determinable.

## II. PRELIMINARIES

The types of channels to be considered can be described as follows. The input signal $x$ and observed signal $w$ are related via the operator equation

$$w(t) = [Hx](t) + z(t) \qquad t \in [0, \infty) \tag{1}$$

where $H$ is an operator and $z$ is zero mean white gaussian noise† with covariance $Ez(t)z(\tau) = \delta(t - \tau)$. (The colored noise case will be treated separately in Section V.)

We constrain our input functions $x$ to have peak value less than $s$,

† The noise term $z(t)$ in equation (1) must be interpreted symbolically since white noise cannot be parametrized with a time variable, but must properly be parametrized with an element of a space of "testing functions." However, we deal only with functionals of $w(t)$ of the form

$$\int_0^b w(t)\phi(t)\,dt,$$

where $\phi \in L_2(0, b)$, or with quantities derivable from these functionals. Hence we can define

$$\int_0^b z(t)\phi(t)\,dt$$

to mean

$$\int_0^b \phi(t)\,d\zeta(t)$$

where $\zeta(t)$ is Brownian motion and the operations to be performed are readily justified.

that is,

$$x \in L_\infty(s) = \{x \mid x \text{ is a real valued measurable function on } [0, \infty) \text{ and } |x(t)| \leqq s \text{ for all } t \in [0, \infty)\}.$$

If we let $\| \ \|_2$ denote the norm on $L_2[0, \infty)$ and define the projection operator $P_T$ by

$$\begin{aligned}[P_T x](t) &= x(t) \quad \text{for} \quad t \leqq T \\ &= 0 \quad \text{for} \quad t > T\end{aligned}$$

then

$$\| P_T x \|_2 = \left( \int_0^T x^2(t)\, dt \right)^{\frac{1}{2}} \leqq s(T)^{\frac{1}{2}} \quad \text{for all} \quad x \in L_\infty(s).$$

The types of operators which we consider are assumed to belong to the space $\mathcal{H}$. The space $\mathcal{H}$ is defined: if $H \in \mathcal{H}$ then

(*i*) $H : L_\infty(s) \to L_{2e}$

where $L_{2e} = \{y \mid y$ is a real valued, measurable function on $[0, \infty)$, $\| P_T y \|_2 < \infty$ for all $T > 0\}$,

(*ii*) $H$ is causal; that is, for all $T > 0$, $x \in L_\infty(s)$, $P_T H x = P_T H P_T x$,

(*iii*) $\| H \| < \infty$.

Using the usual definitions of addition of operators and multiplication by scalars, the norm of $H$, $\| H \|$ is defined as:

$$\| H \| = \sup_{\substack{T>0 \\ x \in L_\infty(s) \\ \| P_T x \|_2 \neq 0}} \frac{\| P_T H x \|_2}{\| P_T x \|_2}.$$

We consider $H$ to be the zero operator† if $\| H \| = 0$. It is then easy to show that $\| \ \|$ satisfies the norm axioms. Obviously $\| H \| \geqq 0$ for all $H \in \mathcal{H}$ and $\| \lambda H \| = |\lambda| \, \| H \|$ for all scalars $\lambda$. The triangle inequality is also satisfied since

$$\| H + K \| = \sup \frac{\| P_T(Hx + Kx) \|_2}{\| P_T x \|_2} = \sup \frac{\| P_T H x + P_T K x \|_2}{\| P_T x \|_2}$$

† The equivalence classes defined in this manner are not unreasonable. In fact, if $\| H \| = 0$ then $\| P_T H x \|_2 = 0$ for all $x \in L_\infty(s)$, $\| P_T x \|_2 \neq 0$ and all $T > 0$. As far as we are concerned this is the zero operator since $Hx$ is then the zero function in the $L_2(0, \infty)$ sense.

$$\leqq \sup \left[ \frac{\| P_T H x \|_2 + \| P_T K x \|_2}{\| P_T x \|_2} \right]$$

$$\leqq \sup \frac{\| P_T H x \|_2}{\| P_T x \|_2} + \sup \frac{\| P_T K x \|_2}{\| P_T x \|_2} = \| H \| + \| K \|$$

where the supremums are taken over all $T > 0$, $x \in L_\infty(s)$, $\| P_T x \|_2 \neq 0$.

If we consider the metric induced by the norm $\| \ \|$ then $\mathcal{H}$ is a complete metric space. The proof of this proposition is contained in the appendix. The completeness property is crucial to Theorem 2 of this paper.

The space $\mathcal{H}$ includes many types of operators familiar to those in communication and control theory. Linear time invariant convolution operators whose kernels are either in $L_1(0, \infty)$ or $L_2(0, \infty)$ are in $\mathcal{H}$. If these operators are cascaded with a memoryless nonlinearity having bounded slope, the composite operators are also in $\mathcal{H}$. Operators described by certain nonlinear dynamical systems are also in $\mathcal{H}$. Let $x \in L_\infty(s)$ be the input to the following dynamical system and let $y$ be the output:

$$\dot{\underline{q}}(t) = \underline{f}(\underline{q}(t), x(t), t), \qquad \underline{q}(0) = 0$$

$$\underline{f} : R^n \times R \times R \rightarrow R^n$$

$$y(t) = g(\underline{q}(t))$$

$$g : R^n \rightarrow R$$

with

$$| g(\underline{q}) | \leqq K_1 | \underline{q} |, \qquad | \underline{f}(\underline{q}, x, t) | \leqq K_2 | \underline{q} | + K_3 | x |$$

for all $\underline{q} \in R^n$, $| x | < s$, $t > 0$. Assume also that for each $x \in L_\infty(s)$ there exists a solution to the differential equation. Then, via the Bellman-Gronwall inequality we see that

$$| \underline{q}(t) | \leqq K_3 \int_0^t e^{K_2(t-\tau)} | x(\tau) | \, d\tau.$$

Hence,

$$\left( \int_0^T | \underline{q}(t) |^2 \, dt \right)^{\frac{1}{2}} \leqq K_2 K_3 \left( \int_0^T x^2(t) \, dt \right)^{\frac{1}{2}}$$

and

$$\left( \int_0^T | y(t) |^2 \, dt \right)^{\frac{1}{2}} \leqq K_1 K_2 K_3 \| P_T x \|_2 .$$

Thus, the operator described is in $\mathcal{H}$ with norm bounded by $K_1 K_2 K_3$ .

Subsets of $\mathcal{H}$ will be used to represent the a priori information in an identification problem. We call a subset $\mathfrak{D}$ of $\mathcal{H}$ determinable if every member of $\mathfrak{D}$ can be identified. The determinability of a subset $\mathfrak{D}$ depends of course on our definition of identification. We would like to consider only those identification procedures which could theoretically be implemented in real time. The identification procedures which we are concerned with must have the following properties. To identify $H$ we must be able to

(*i*) choose a finite observation interval,
(*ii*) select an input function with constrained peak value,
(*iii*) perform linear measurements on the noisy observations generated by this input, and
(*iv*) operate on these measurements to yield an estimate of $H$ which is a continuous function of these measurements,

so that our estimate of $H$ is close to $H$ with high probability.

The properties of such an identification procedure are physically very appealing. We obviously must be able to identify within a finite period of time. The peak value restriction is the usual kind of input constraint used in communication theory. Linear measurements are easily implemented and tend to reduce the sensitivity to unknown biases as does the continuity requirement on the estimate. Finally, we are usually satisfied to identify to within a small tolerance.

For $H \in \mathcal{H}$ and channel model given by equation (1) we may specify our definition of identification even further. A *linear measurement* over the time interval $[0, T]$ is a finite collection of bounded linear functionals† $(p_i, w)$, $i = 1, 2, \cdots N$, $p_i \in L_2[0, T]$ defined when $P_T w \in L_2[0, T]$ and

$$w(t) = [Hx](t) + z(t), \qquad 0 \leqq t \leqq T$$

is the received waveform with $H \in \mathcal{H}$, $x \in L_\infty(s)$. We say that a class $\mathfrak{D} \subset \mathcal{H}$ of channel operators is *determinable* if given arbitrary positive constants $\epsilon$ and $\eta$, there exists a finite observation interval $[0, T]$, an input (test) signal $x \in L_\infty(s)$, a linear measurement $[(p_1, w), (p_2, w), \cdots, (p_N, w)]$ over $[0, T]$, and a continuous function $g : R^N \to \mathcal{H}$ such that for each $H \in \mathfrak{D}$,

$$\text{Probability } (\| H - \hat{H} \| > \epsilon) < \eta$$

where $\hat{H} = g[(p_1, w), (p_2, w), \cdots, (p_N, w)]$. Thus, if $\mathfrak{D}$ is determinable,

† The symbol $(f, h)$ is used to represent the inner product in $L_2[0, T]$; that is, $(f, h) = \int_0^T f(t)h(t)\, dt$.

we can "identify" any element of $\mathfrak{D}$ to within any specified accuracy with sufficient processing and long enough observation time.

The bulk of this paper is related to answering the following question. What structure must $\mathfrak{D}$ have in order to be determinable? Theorem 1 derives sufficient conditions on $\mathfrak{D}$ in order to be determinable. The key condition is compactness. Theorem 2 indicates that this condition is in fact necessary for determinability. A number of corollaries are given which interpret these results for the case where $\mathfrak{D}$ is composed of linear convolution operators.

## III. SUFFICIENT DETERMINABILITY CONDITIONS

Despite the generality of our class of operators and the rather rigid nature of allowable identification schemes only two conditions guarantee the determinability of a subset of operators. Both conditions are somewhat obvious. One condition insures that the class may be approximated closely by a finite number of elements; the other insures that a test signal exists which will produce sufficiently dissimilar responses for dissimilar channels. These conditions are rigorously stated in Theorem 1.

*Theorem 1*: *Let $\mathfrak{D}$ be a subset of $\mathfrak{K}$ having the following properties*:

(*i*) *the closure of $\mathfrak{D}$ is compact* (*thus $\bar{\mathfrak{D}}$ is also bounded*; *that is*, *there exists a constant* $R > 0$, *such that* $\| H - K \| < R$ *for all* $H, K \in \bar{\mathfrak{D}}$)

(*ii*) *given any* $\delta > 0$ *there exists an unbounded sequence* $\{T_i\}$, *a sequence of inputs* $x_i \in L_\infty(s)$ *and a positive number* $r$ *such that*

$$\| P_{T_i}(Hx_i - Kx_i) \|_2^2 > rT_i$$

*for all pairs* $H, K \in \bar{\mathfrak{D}}$ *for which* $\| H - K \| \geqq \delta$. *Then $\mathfrak{D}$ is a determinable subset of $\mathfrak{K}$.*

*Proof*: Since the proof of this theorem is lengthy, we give here a brief, rough description of the key steps involved which the reader may use as a guide through the mathematical details.

(*i*) Using (*ii*) of Theorem 1 we select an input $x_i$ to give sufficient separation of outputs over $[0, T_i]$ for sufficiently dissimilar channels.

(*ii*) We then approximate the class $\mathfrak{D}$ to within a judiciously chosen accuracy by a finite number of elements.

(*iii*) The actual received signal due to the input selected in (*i*) of this proof is correlated over $[0, T_i]$ with the calculated outputs of the channels selected in step (*ii*) of this proof.

(*iv*) If one of these correlations is larger than the others by some amount we select as our estimate the corresponding element of the approximating class that yielded this correlation. If there is no such correlation we assign an arbitrary rule so as to make the identification procedure a continuous function of the correlated values.

(*v*) We finally show that as $i$ (and hence $T_i$) increases, the probability that there will not be a correlation larger than the others by some prescribed amount goes to zero. In addition, we show that the probability that our identification procedure yields an estimate which is further apart in norm from the actual channel than is desired is vanishingly small as $i$ increases.
The formal statement of the proof follows below.

We may assume that $\mathfrak{D}$ is closed, since subsets of a determinable set of channels are determinable. Using assumption (*ii*) of Theorem 1 with $\delta = 3\epsilon/4$ we have that there exists an unbounded sequence $\{T_i\}$, a positive number $r$, and for each $i$ an input signal $x_i \in L_\infty(x)$ such that for all pairs $H, K \in \mathfrak{D}$ with $\| H - K \| > 3\epsilon/4$

$$\| P_{T_i}(Hx_i - Kx_i) \|_2^2 \geqq rT_i \, . \tag{2}$$

In what follows we will denote the operator which we wish to identify by $H$. Since $\mathfrak{D}$ is closed, by assumption (*i*) of Theorem 1, it is also compact and hence totally bounded (see for example Ref. 6, p. 22). Therefore, given any $T_i \in \{T_i\}$ we can choose a finite number of balls of radius $r_0 = \min \{r^{\frac{1}{2}}/2s, \epsilon/4\}$ with centers $H_\alpha \in \mathfrak{D}$, $\alpha = 1, 2, \cdots, M$ to cover $\mathfrak{D}$. There may be operators $H_l, H_k \in \{H_\alpha\}$ for which $\| P_{T_i}(H_l x_i - H_k x_i) \|_2 = 0$, in which case retain only the $H_\alpha$'s with the lowest subscript. Thus we have a subset of $\{H_\alpha\}$ which we label $\{H_\beta\}$ for which $\| P_{T_i}(H_l x_i - H_k x_i) \|_2 > \theta_i > 0$ for some $\theta_i$ and all $H_l, H_k \in \{H_\beta\}$. For convenience order the $\{H_\beta\}$ so that $\| H - H_\beta \| \leqq 3\epsilon/4$ for $\beta = 1, 2, \cdots, N_0 - 1$ and $\| H - H_\beta \| > 3\epsilon/4$ for $\beta = N_0, N_0 + 1, \cdots, N, N \leqq M$.

We can now choose an appropriate linear measurement over the interval $[0, T_i]$. We define the linear measurement $\underline{m}(w) = \{f(w, 1), f(w, 2), \cdots f(w, N)\}: f(w, \beta) = (w, 2H_\beta x_i)$, $\beta = 1, 2, \cdots N$ where the inner product is defined over the interval $[0, T_i]$. Thus for each received waveform $w(t)$, the linear measurement gives us a point in $R^N$. From this measurement we will determine an estimator function $g : R^N \to \mathfrak{H}$. We first partition $R^N$ into $N + 1$ disjoint subsets: $A_1, A_2, \cdots, A_N, B$, with

$$A_j = \{\underline{a} = (a_1, a_2, \cdots a_N): a_j - a_k > (H_j x_i, H_j x_i) - (H_k x_i, H_k x_i) + \theta_i/T_i, \quad k = 1, 2, \cdots N, \quad k \neq j\}$$

and $B$ the remainder of $R^N$,

$$B = \left( \bigcup_{j=1}^{N} A_j \right)^c .$$

The disjointness of the above subsets of $R^N$ is easily verified by making use of the fact that $\theta_i/T_i > 0$. The estimator function is defined in terms of this partition:

$$g(\underline{m}) = H_j \qquad \text{if} \quad \underline{m}(w) \in A_j$$

$$g(\underline{m}) = \sum_{i=1}^{N} \alpha_i(w) H_i \quad \text{if} \quad \underline{m}(w) \in B$$

where†

$$\alpha_i(w) = \frac{\prod_{j \neq i} d(\underline{m}(w), A_j)}{d(\underline{m}(w), A_i) + \prod_{j \neq i} d(\underline{m}(w), A_j)}$$

and

$$d(x, A) = \inf_{y \in A} | x - y |.$$

It is not difficult to show that $g$ is a continuous mapping from $R^N$ into $\mathcal{H}$. Having given the identification scheme we now show that for any $H \in \mathcal{H}$, $\epsilon > 0$

$$P\{\| H - \hat{H} \| > \epsilon\} \underset{T_i \to \infty}{\longrightarrow} 0.$$

Recalling the definition of $B$, $A_i$ and the labeling convention we have used, we see that

$$P\{\| H - g(\underline{m}(w)) \| > \epsilon\} \leqq P\{\underline{m}(w) \in B\} + P\left\{\underline{m}(w) \in \bigcup_{j=N_0}^{N} A_j\right\}$$

$$= P\left\{\underline{m}(w) \in \bigcap_{j=1}^{N} A_j^c\right\} + P\left\{\underline{m}(w) \in \bigcup_{j=N_0}^{N} A_j\right\}. \tag{3}$$

Let us first concentrate on obtaining bounds for the first term on the right side of equation (3). We rewrite $A_j$ as $A_j = (\bigcup_{k \neq j} F_{jk})^c$ where

$$F_{jk} = \{\underline{a} = (a_1, a_2, \cdots, a_N) : a_j - a_k \leqq (H_j x_i, H_j x_i) - (H_k x_i, H_k x_i) + \theta_i/T_i\}.$$

Thus

† It turns out that the form of $\alpha_i(w)$ is irrelevant since we show that $P[m(w) \in B]$ vanishes as $T_i$ increases. It is merely included to make the estimator function continuous.

$$\bigcap_{j=1}^{N} A_j^c = \bigcap_{j=1}^{N} \Big(\bigcup_{k \neq j} F_{jk}\Big). \tag{4}$$

Applying DeMorgan's rules to equation (4), and after some thought, we see that

$$\bigcap_{j=1}^{N} A_j^c = \bigcup_{l=1}^{(N-1)^N} D_l \tag{5}$$

where $D_l$ has the form

$$D_l = F_{1l_1} \cap F_{2l_2} \cap \cdots \cap F_{Nl_N}$$

with $l_j \neq j$ for all $j$. We can upper bound $P\{\underline{m}(w) \varepsilon D_l\}$ by

$$\sup_{\substack{k,j \\ k \neq j}} P\left\{-\frac{N\theta_i}{T_i} + (H_jx_i, H_jx_i) - (H_kx_i, H_kx_i) \leqq f(w, k) - f(w, j) \right.$$
$$\left. \leqq \frac{N\theta_i}{T_i} + (H_kx_i, H_kx_i) - (H_jx_i, H_jx_i)\right\}. \tag{6}$$

To see this, define $q(w, k) = f(w, k) - (H_kx_i, H_kx_i)$. Then $P\{\underline{m}(w) \varepsilon D_l\}$ is the probability of the $N$ events $q(w, 1) - q(w, l_1) \leqq \theta_i/T_i$, $q(w, 2) - q(w, l_2) \leqq \theta_i/T_i \cdots q(w, N) - q(w, l_N) \leqq \theta_i/T_i$ occurring simultaneously. Suppose $l_1 = k$. Then consider the two events $q(w, 1) - q(w_1, l_1) = q(w, 1) - q(w, k) \leqq \theta_i/T_i$ and $q(w, k) - q(w, l_k) \leqq \theta_i/T_i$. If $l_k = 1$, then these two events are contained in the event $-\theta_i/T_i \leqq q(w, 1) - q(w, k) \leqq \theta_i/T_i$. If $l_k = j \neq 1$ then consider the three events

$$q(w, 1) - q(w, k) \leqq \theta_i/T_i,$$
$$q(w, k) - q(w, j) \leqq \theta_i/T_i,$$
$$q(w, j) - q(w, l_j) \leqq \theta_i/T_i.$$

If $l_j = 1$, then these three simultaneous events are contained in the event $-\theta_i/T \leqq q(w, 1) - q(w, j) \leqq 2\theta_i/T_i$. If $l_j = k$, then these three simultaneous events are contained in the event $-\theta_i/T \leqq q(w, k) - q(w, j) \leqq \theta_i/T_i$. Continuing in this fashion we obtain the bound in equation (6).

Since $q(w, k) - q(w, j)$ is gaussian, we can bound the value of the expression in equation (6) quite easily.

Let

$$\begin{aligned} a_{kj} &= E[q(w, k) - q(w, j)] \\ &= \| P_{T_i}(Hx_i - H_jx_i) \|_2^2 - \| P_{T_i}(Hx_i - H_kx_i) \|_2^2 \end{aligned} \tag{7}$$

and

$$\sigma_{kj}^2 = \text{Var}\,[q(w, k) - q(w, j)] = 4\,\|P_{T_i}(H_k x_i - H_j x_i)\|_2^2 > 4\theta_i^2 . \tag{8}$$

Hence,

$$P\left\{-\frac{N\theta_i}{T_i} \leqq q(w, k) - q(w, j) \leqq \frac{N\theta_i}{T_i}\right\}$$

$$= (2\pi)^{-\frac{1}{2}} \int_{-(N\theta_i/T_i\sigma_{kj})-a_{kj}/\sigma_{kj}}^{N\theta_i/T_i\sigma_{kj}-a_{kj}/\sigma_{kj}} \exp\left(-\frac{z^2}{2}\right) dz$$

$$\leqq (2\pi)^{-\frac{1}{2}} \int_{-(N\theta_i/T\sigma_{kj})}^{N\theta_i/T_i\sigma_{kj}} \exp\left(-\frac{z^2}{2}\right) dz$$

$$\leqq (2\pi)^{-\frac{1}{2}} \int_{-(N/2T_i)}^{N/2T_i} \exp\left(-\frac{z^2}{2}\right) dz \leqq (2\pi)^{-\frac{1}{2}} \int_{-(M/2T_i)}^{M/2T_i} \exp\left(-\frac{z^2}{2}\right) dz \tag{9}$$

(recall that $N \leqq M$).

Using equations (9) and (5) we see that

$$P\{\underline{m}(w) \,\varepsilon \bigcap_{i=1} A_i^c\} \leqq \sum_{l=1}^{(N-1)^N} P\{\underline{m}(w) \,\varepsilon\, D_l\}$$

$$< (M - 1)^M (2\pi)^{-\frac{1}{2}} \int_{-(M/T_i)}^{M/T_i} \exp\left(-\frac{z^2}{2}\right) dz. \tag{10}$$

Since the right side of equation (10) goes to zero as $T_i$ increases we can choose a $T \,\varepsilon\, \{T_i\}$ large enough so that this term is less than $\eta/2$. We now bound the second term on the right side of equation (3):

$$P\left\{\underline{m}(w) \,\varepsilon \bigcup_{j=N_0}^{N} A_j\right\} = \sum_{j=N_0}^{N} P\{\underline{m}(w) \,\varepsilon\, A_j\}. \tag{11}$$

Recall that

$$A_j = (\bigcup_{k\neq j} F_{jk})^c = \bigcap_{k\neq j} F_{jk}^c .$$

Hence

$$P\left\{\underline{m}(w) \,\varepsilon \bigcup_{j=N_0}^{N} A_j\right\} = \sum_{j=N_0}^{N} P\{\underline{m}(w) \,\varepsilon \bigcap_{k\neq j} F_{jk}^c\}. \tag{12}$$

Observe that for all $k \neq j$

$$P\{\underline{m}(w) \,\varepsilon \bigcap_{k\neq j} F_{jk}^c\} \leqq P\{\underline{m}(w) \,\varepsilon\, F_{jk}^c\} \tag{13}$$

$$= P\left\{q(w, j) - q(w, k) > \frac{\theta_i}{T_i}\right\} \tag{14}$$

$$= \int_{\theta_i/T_i - a_{jk}/\sigma_{jk}}^{\infty} \frac{1}{(2\pi)^{\frac{1}{2}}} \exp\left(-\frac{z^2}{2}\right) dz. \tag{15}$$

Since $\mathfrak{D}$ was covered by balls of radius $r_0$, there exists at least one integer $\tilde{k} < N_0$ such that $\| H - H_{\tilde{k}} \| < r_0 \leqq \epsilon/4$ and hence $\| P_{T_i}(Hx_i - H_{\tilde{k}}x_i)\|^2 < r_0^2 s^2 T_i$. Note also that since $\| H_j - H \| > 3\epsilon/4$ for $j \geqq N_0$, $\| P_{T_i}(Hx_i - H_jx_i) \|^2 > rT_i$. Hence,

$$\begin{aligned} -a_{j\tilde{k}} &= \| P_{T_i}(Hx_i - H_jx_i) \|_2^2 - \| P_{T_i}(Hx_i - H_{\tilde{k}}x_i) \|_2^2 \\ &\geqq rT_i - r_0^2 s^2 T_i \geqq \left(r - \frac{r}{4}\right) T_i = \tfrac{3}{4} r T_i . \end{aligned} \tag{16}$$

Recalling that $\mathfrak{D}$ was bounded,

$$\sigma_{j\tilde{k}}^2 = 4 \| P_{T_i}(H_jx_i - H_{\tilde{k}}x_i) \|_2^2 \leqq 4R^2 s^2 T_i . \tag{17}$$

Using equations (16) and (17) in equation (15) we see that

$$P\{\underline{m}(w) \,\varepsilon \bigcap_{k \neq j} F_{jk}^c\} \leqq P\{\underline{m}(w) \,\varepsilon\, F_{j\tilde{k}}^c\} \leqq \int_{3rT_i^{\frac{1}{2}}/16Rs}^{\infty} (2\pi)^{-\frac{1}{2}} \exp\left(-\frac{z^2}{2}\right) dz. \tag{18}$$

Hence from equation (11) we see that

$$P\left\{\underline{m}(w) \,\varepsilon \bigcup_{j=N_0}^{N} A_j\right\} \leqq M \int_{3rT_i^{\frac{1}{2}}/16Rs}^{\infty} (2\pi)^{-\frac{1}{2}} \exp\left(-\frac{z^2}{2}\right) dz. \tag{19}$$

Thus we can select a $T \,\varepsilon\, \{T_i\}$ so that this term is less than $\eta/2$. This $T$ makes $P\{\| H - \hat{H} \| > \epsilon\} < \eta$ for all $H \,\varepsilon\, \mathfrak{D}$.

The identification technique proposed in the above proof is not necessarily a practical technique. Our intent is to indicate the possibility of identification rather than to derive easily implementable techniques. Notice, however, that since the measurements are linear functionals on $L_2(0, T)$ they are iterative in nature because of the integral representation of such functionals.

Theorem 1 gives sufficient conditions for determinability. Theorem 2 indicates that some of these conditions are in fact necessary for identification.

## IV. NECESSARY DETERMINABILITY CONDITIONS

In this section we show that the approximability condition given by condition (*i*) of Theorem 1 is in fact necessary. We also show that a type of separation property is necessary, although it is not as strong as that given by condition (*ii*) of Theorem 1.

*Theorem 2*: *Let* $\mathfrak{D}$ *be a bounded, determinable subset of* $\mathfrak{H}$, *then*

(*i*) *the closure of* $\mathfrak{D}$ *is compact*

(*ii*) *given any* $\delta > 0$ *there exists an* $\hat{x} \varepsilon L_\infty(s)$, $\hat{T} > 0$ *and a positive number* $r(\delta)$ *such that*

$$\| P_{\hat{T}}(H\hat{x} - K\hat{x}) \|_2^2 > r(\delta) \quad \textit{for all} \quad H, K \varepsilon \bar{\mathfrak{D}}$$

*satisfying* $\| H - K \| \geqq \delta$.

*Proof*: (*i*) Given $\epsilon > 0$, choose $\hat{T}$, $\hat{x} \varepsilon L_\infty(s)$, $\hat{N}$ linear measurements and an estimator $g(\underline{m}(H, \omega))$ so that†

$$P\{\| H - g(\underline{m}(H, \omega)) \| < \epsilon/2\} > \tfrac{3}{4} \quad \text{for all} \quad H \varepsilon \mathfrak{D}.$$

Since $\mathfrak{D}$ is bounded and the measurements are linear, there exists a compact ball $B_{\hat{T}} \varepsilon R^{\hat{N}}$ so that

$$P\{\underline{m}(H, \omega) \varepsilon B_{\hat{T}}^c\} < \tfrac{1}{4} \quad \text{for all} \quad H \varepsilon \mathfrak{D}.$$

Thus, since $g$ is continuous, $g(B_{\hat{T}})$ is compact. We can therefore cover $B_{\hat{T}}$ by a finite number of balls of radius $\epsilon/2$. If $g(B_{\hat{T}}) \supset \mathfrak{D}$ we could also cover $\mathfrak{D}$ by the balls. We don't have enough information to verify that $g(B_{\hat{T}}) \supset \mathfrak{D}$. Notice however that

$$\begin{aligned} P\{[\omega: \| H - g(\underline{m}(H, \omega)) \| > \epsilon/2] \cap [\omega: \underline{m}(H, \omega) \varepsilon B_{\hat{T}}]\} \\ = P\{\omega: \| H - g(\underline{m}(H, \omega)) \| > \epsilon/2\} + P\{\omega: \underline{m}(H, \omega) \varepsilon B_{\hat{T}}\} \\ - P\{[\omega: \| H - g(\underline{m}(H, \omega)) \| > \epsilon/2] \cup [\omega: \underline{m}(H, \omega) \varepsilon B_{\hat{T}}]\} \\ \geqq \tfrac{3}{4} + \tfrac{3}{4} - 1 = \tfrac{1}{2}. \end{aligned} \tag{20}$$

We conclude that there exists an $\omega_0$ so that $\underline{m}(H, \omega_0) \varepsilon B_{\hat{T}}$ and $\| H - g(\underline{m}(H, \omega_0)) \| < \epsilon/2$. We can repeat this argument for each $H \varepsilon \mathfrak{D}$. Therefore, $\mathfrak{D}$ must lie within an $\epsilon/2$ neighborhood of $g(B_{\hat{T}})$. By expanding the balls of radius $\epsilon/2$ which cover $g(B_{\hat{T}})$ by a factor of two, the expanded balls will also cover $\mathfrak{D}$. Since this argument holds for any $\epsilon > 0$, $\mathfrak{D}$ is shown to be totally bounded. Since $\mathfrak{H}$ is complete, $\bar{\mathfrak{D}}$ is complete; and hence $\bar{\mathfrak{D}}$ is compact (see Ref. 6, p. 22).

(*ii*) If $\mathfrak{D}$ is determinable, then the closure of $\mathfrak{D}$, $\bar{\mathfrak{D}}$, is also determinable. This is easily shown by noting that any channel in $\bar{\mathfrak{D}}$ can be approximated arbitrarily closely by a channel in $\mathfrak{D}$. Hence the measurements will be arbitrarily close and because of the continuity of the estimate, the estimate will be close with high probability.

† Since the measurements are gaussian random variables we have included the dependence on the sample points $\omega$ of the corresponding sample space $\Omega$.

Since $\bar{\mathfrak{D}}$ is determinable, for every $\epsilon > 0$ there exists an observation interval $[0, \hat{T}]$, a test signal $\hat{x} \in L_\infty(s)$, and an estimator $g(\underline{m}(\cdot, \omega))$ so that

$$P\{\| H - g[\underline{m}(H, \omega)] \| < \delta/2\} > \tfrac{3}{4} \quad \text{for all} \quad H \in \mathfrak{D}. \tag{21}$$

Suppose that $\| P_{\hat{T}}(H\hat{x} - K\hat{x}) \|_2 = 0$. Then, the measurements obtained will be the same irrespective of whether $H$ or $K$ were used and therefore the estimates for $K$ and $H$ will be identical. Since

$$P\{\omega: \| H - g[\underline{m}(H, \omega)] \| < \delta/2\} > \tfrac{3}{4}$$

and

$$P\{\omega: \| K - g[\underline{m}(H, \omega)] \| < \delta/2\} > \tfrac{3}{4},$$

we see that

$$\begin{aligned} P\{[\omega: \| K - g[\underline{m}(H, \omega)] \| < \delta/2\} & \\ \cap \{\omega: \| H - g[\underline{m}(H, \omega)] \| < \delta/2]\} & \\ = P\{\omega: \| K - g[\underline{m}(H, \omega)] \| < \delta/2\} & \\ + P\{\omega: \| H - g[\underline{m}(H, \omega)] \| < \delta/2\} & \\ - P\{[\omega: \| K - g[\underline{m}(H, \omega)] \| < \delta/2\} & \\ \cup \{\omega: \| H - g[\underline{m}(H, \omega)] \| < \delta/2]\} & \\ \geqq \tfrac{3}{4} + \tfrac{3}{4} - 1 = \tfrac{1}{2}. & \end{aligned}$$

Thus there exists at least one sample point $\omega_0$ such that

$$\| K - g(\underline{m}(H, \omega_0)) \| < \delta/2$$

and

$$\| H - g(\underline{m}(H, \omega_0)) \| < \delta/2$$

which together imply that $\| H - K \| < \delta$. If $H, K \in \bar{\mathfrak{D}}$ and $\| H - K \| > \delta$ then $\| P_{\hat{T}}(H\hat{x} - K\hat{x}) \|_2 > 0$.

Note that $\bar{\mathfrak{D}} \times \bar{\mathfrak{D}}$ is compact in the product topology and hence $C(\delta) = \{(H, K): \| H - K \| \geqq \delta, H, K \in \bar{\mathfrak{D}}\}$ is also compact. The function $f(H, K) = \| P_{\hat{T}}(H\hat{x} - K\hat{x}) \|_2$ is a continuous map of $C(\delta)$ into the real line and hence it has a minimum value. This minimum value cannot be zero because we have already shown that $f(H, K) > 0$ for $(H, K) \in C(\delta)$. As a consequence, there exists a positive number $r(\delta)$ such that

$$\| P_{\hat{T}}(H\hat{x} - K\hat{x}) \|_2^2 > r(\delta) \quad \text{for all} \quad H, K \in \mathfrak{D}$$

satisfying $\| H - K \| \geqq \delta$.

## V. LINEAR CONVOLUTION OPERATORS

When we specialize the results of Theorems 1 and 2 to linear convolution operators, it is possible to obtain the characterization of the determinable sets in terms of the kernels of these operators. These results are given in Corollaries 1, 2 and 3 below. We note that the resulting conditions are similar to those obtained by Root and Prosser for the deterministic identification problem.[1]

*Corollary 1*: *If* $\mathcal{H}$ *is composed only of causal linear time invariant convolution operators* H, $[\mathrm{Hx}](t) = \int_0^t \mathrm{h}(\mathrm{t} - \tau)x(\tau)\, \mathrm{d}\tau$, $\mathrm{h} \in \mathrm{L}_1(0, \infty)$ *and if*

(*i*) $\tilde{\mathfrak{D}} = \{\mathrm{h} \mid \mathrm{h} \in \mathrm{L}_1(0, \infty), \mathrm{H} \in \mathfrak{D}\}$ *has a compact closure in* $\mathrm{L}_1(0, \infty)$, *and*

(*ii*) *for each* $\delta > 0$ *there exists an* $\mathrm{x} \in \mathrm{L}_\infty(\mathrm{s})$, $\mathrm{T} > 0$ *such that* $\| \mathrm{P_T Hx} - \mathrm{P_T Kx} \|_2 > 0$ *for all* $\mathrm{h, k} \in \bar{\tilde{\mathfrak{D}}}$ *for which* $\| \mathrm{h} - \mathrm{k} \|_1 = \int_0^\infty | \mathrm{h(t)} - \mathrm{k(t)} | \,\mathrm{dt} \geqq \delta$ *then* $\mathfrak{D}$ *is determinable.*

Necessary and sufficient conditions for $\tilde{\mathfrak{D}}$ to have a compact closure are (see Ref. 6, pp. 298–299):

(*i*) $\tilde{\mathfrak{D}}$ is a bounded subset of $L_1(0, \infty)$,
(*ii*) $\lim_{\tau \to 0} \int_0^\infty | h(t + \tau) - h(t) | \, dt = 0$ uniformly for $h \in \tilde{\mathfrak{D}}$, and
(*iii*) $\lim_{T \to \infty} \int_T^\infty | h(t) | \, dt = 0$ uniformly for $h \in \tilde{\mathfrak{D}}$.

*Proof*: We first show that if the closure of $\tilde{\mathfrak{D}}$ is compact then the closure of $\mathfrak{D}$ is compact in the respective topologies. Let $\| H \|^*$, $H \in \mathcal{H}$ denote the usual operator norm, that is,

$$\| H \|^* = \sup_{\substack{x \in L_2(0,\infty) \\ x \neq 0}} \frac{\| Hx \|_2}{\| x \|_2}.$$

Given any $\epsilon > 0$ there exists $T^* > 0$, $x^* \in L_\infty(s)$ such that

$$\| H \| \leqq \epsilon + \frac{\| P_{T^*} H x^* \|_2}{\| P_{T^*} x^* \|_2} \leqq \epsilon + \frac{\| P_{T^*} H P_{T^*} x^* \|_2}{\| P_T x^* \|_2}. \tag{23}$$

Note however that $P_{T^*} x^* \in L_2(0, \infty)$; hence $\| H \| \leqq \epsilon + \| H \|^*$ for arbitrary $\epsilon > 0$, so

$$\| H \| \leqq \| H \|^*. \tag{24}$$

Using the linearity of $H$ and Holder's inequality we see that

$$\| H \|^* = \sup_{x \in L_2(0,\infty)} \frac{\| Hx \|_2}{\| x \|_2} \leqq \sup_{x \in L_2(0,\infty)} \frac{\| h \|_1 \| x \|_2}{\| x \|_2} = \| h \|_1 . \tag{25}$$

Thus compactness in $L_1(0, \infty)$ implies compactness in $\mathcal{H}$ and condition $(i)$ of Theorem 1 is satisfied.

Given $\delta > 0$, choose $x_o$, $T^o$ so that condition $(ii)$ of Corollary 1 is satisfied. We have already used the fact that $\| HP_{T^o}x_o \|_2 \leqq \| h \|_1 \cdot \| P_{T^o}x_o \|_2$. Hence $HP_{T^o}x_o$ is a continuous linear mapping (that is, mapping the kernels into time functions) from $L_1(0, \infty)$ into $L_2(0, \infty)$. Thus the image of $\bar{\mathfrak{D}}$ under this mapping has a compact closure. We can therefore choose a number $\hat{T} > T^o$ so that

$$\int_{\hat{T}}^{\infty} (HP_{T^o}x_o - KP_{T^o}x_o)^2(t)\, dt < \tfrac{1}{4} \quad \text{for all} \quad H, K \in \bar{\mathfrak{D}}. \tag{26}$$

Define $\hat{x}$ as follows:

$$\begin{aligned} \hat{x}(t) &= x_0(t) && \text{for} \quad 0 < t \leqq T^o \\ &= 0 && \text{for} \quad T^o < t \leqq \hat{T} \\ &= x_0(t - \hat{T}) && \text{for} \quad \hat{T} < t \leqq \hat{T} + T^o \\ &= 0 && \text{for} \quad \hat{T} + T^o < t \leqq 2\hat{T} \\ &\ \vdots \\ &= x_0(t - n\hat{T}) && \text{for} \quad n\hat{T} < t \leqq n\hat{T} + T^o \\ &= 0 && \text{for} \quad n\hat{T} + T^o < t \leqq (n+1)\hat{T} \\ &\ \vdots \end{aligned} \tag{27}$$

Note that $\hat{x} \in L_\infty(s)$. Following the same line of reasoning as in the proof of condition $(ii)$ of Theorem 2 we can show that there exists an $r(\delta) > 0$ so that $\| P_{T^o}(Hx_0 - Kx_0) \|_2^2 > r(\delta)$ for all $H, K \in \bar{\mathfrak{D}}$ for which $\| h - k \|_1 > \delta$. We now proceed to show that

$$\| P_{n\hat{T}}(H\hat{x} - K\hat{x}) \|_2^2 \geq \tilde{r}(\delta) n\hat{T} \tag{28}$$

where $\tilde{r}(\delta) = r(\delta)/4\hat{T}$. Let $y_0(t) = [HP_{T^o}x_0 - KP_{T^o}x_0](t)$ and $y_i(t) = y_0(t - i\hat{T})$. Then, by linearity and time invariance,

$$\int_{(i-1)\hat{T}}^{i\hat{T}} (H\hat{x} - K\hat{x})^2(t)\, dt = \int_{(i-1)\hat{T}}^{i\hat{T}} [y_0(t) + y_1(t) + \cdots + y_{i-1}(t)]^2\, dt \tag{29}$$

and

$$\int_{i\hat{T}}^{(i+1)\hat{T}} y_j^2(t)\, dt = \int_{(i-j)\hat{T}}^{(i+1-j)\hat{T}} y_0^2(t)\, dt \quad \text{for} \quad j \leqq i. \tag{30}$$

Using these relationships we see that

$$\begin{aligned}\int_{i\hat{T}}^{(i+1)\hat{T}} & (y_0 + \cdots + y_i)^2 \, dt \\ & \geqq \int_{i\hat{T}}^{(i+1)\hat{T}} y_i^2 \, dt - 2 \int_{i\hat{T}}^{(i+1)\hat{T}} | y_i | (| y_0 | + \cdots + | y_{i-1} |) \, dt \\ & \geqq \int_{i\hat{T}}^{(i+1)\hat{T}} y_i^2 \, dt \left[ 1 - 2 \int_{i\hat{T}}^{(i+1)\hat{T}} (y_0^2 + \cdots + y_{i-1}^2) \, dt \right] \\ & = \int_{i\hat{T}}^{(i+1)\hat{T}} y_i^2 \, dt \left[ 1 - 2 \int_{\hat{T}}^{(i+1)\hat{T}} y_0^2 \, dt \right] \geqq r(\delta)/2. \end{aligned} \tag{31}$$

Hence

$$\begin{aligned} \| P_{n\hat{T}}(H\hat{x} - K\hat{x}) \|_2^2 & = \int_0^{\hat{T}} y_0^2 \, dt + \int_{\hat{T}}^{2\hat{T}} (y_0 + y_1)^2 \, dt + \cdots \\ & \quad + \int_{(n-1)\hat{T}}^{n\hat{T}} (y_0 + \cdots + y_{n-1})^2 \, dt \\ & \geqq nr(\delta)/4 = r'(\delta)n\hat{T}. \end{aligned} \tag{32}$$

We see that this relation implies that condition (*ii*) of Theorem 1 is satisfied; thus $\mathfrak{D}$ is determinable.

When $\mathfrak{H}$ is composed only of causal linear time invariant convolution operators we can also strengthen the conclusion of Theorem 2. This result is given in the following corollary.

*Corollary 2: If $\mathfrak{H}$ is composed only of causal linear time invariant convolution operators and if $\mathfrak{D}$ is a determinable subset of $\mathfrak{H}$ then*

(*i*) *given any $\delta > 0$ there exists an unbounded sequence $T_i$, a sequence of inputs $x_i \varepsilon L_\infty(s)$ and a positive number $r(\delta)$ such that $\| P_{T_i}(Hx_i - Kx_i)^2 \|_2 > r(\delta)T_i$ for all pairs $H, K \varepsilon \bar{\mathfrak{D}}$ for which $\| H - K \| \geq \delta$.*

*Proof*: As a consequence of Theorem 2 we know that for any $\delta > 0$ there exists an $\hat{x} \varepsilon L_\infty(s)$, $\hat{T} > 0$ and a positive number $r(\delta)$ such that $\| P_{\hat{T}}(H\hat{x} - K\hat{x})^2 \|_2 > r(\delta)$ for all $H, K \varepsilon \bar{\mathfrak{D}}$ satisfying $\| H - K \| \geq \delta$.

Obviously, $\| HP_{\hat{T}}\hat{x} \|_2 \leq \| H \| \, \| P_{\hat{T}}\hat{x} \|_2$. Hence $HP_{\hat{T}}\hat{x}$ is a continuous linear mapping from $\mathfrak{H}$ into $L_2(0, \infty)$. Thus the image of $\mathfrak{D}$ under this mapping has a compact closure. We can therefore choose a positive number $\tilde{T} > \hat{T}$ so that

$$\int_{\tilde{T}}^{\infty} (HP_{\hat{T}}\hat{x} - HP_{\hat{T}}\hat{x})^2(t) \, dt < \tfrac{1}{4} \quad \text{for all} \quad H, K \varepsilon \mathfrak{D}. \tag{33}$$

Proceeding as in the proof of corollary 1 we can easily establish (*i*) of Corollary 2.

*Corollary* 3: *If* $\mathcal{H}$ *is composed only of causal Hilbert-Schmidt operators* H, $[Hx](t) = \int_0^t h(t, \tau)x(\tau)dt$, $\int_0^\infty \int_0^\infty | h(t, \tau) |^2 \, dt \, d\tau < \infty$, $h(t, \tau) = 0$ *for* $\tau > t$ *and if*

(*i*) $\hat{\mathfrak{D}} = \{h \mid H \varepsilon \mathfrak{D}\}$ *has compact closure in the Hilbert-Schmidt metric* ($\| h - k \|_2^2 = \int_0^\infty \int_0^\infty | h(t, \tau) - k(t, \tau) |^2 \, dt \, d\tau$)

(*ii*) *for each* $\delta > 0$ *there exists an unbounded sequence* $T_i$ , *a sequence of* $x_i \varepsilon L_\infty(s)$ *and a positive constant* $r(\delta)$ *so that* $\| P_{T_i}(Hx_i - Kx_i) \|_2 \geq r(\delta)T_i$ *for all* h, k $\varepsilon$ $\bar{\hat{\mathfrak{D}}}$ *for which* $\| h - k \|_2 > \delta$.

*Then* $\mathfrak{D}$ *is determinable.*

*Proof*: As in the proof of Corollary 1 we can show that $\| H \| \leq \| H \|^*$ where

$$\| H \|^* = \sup_{x \varepsilon L_2(0,\infty)} \frac{\| Hx \|_2}{\| x \|_2}.$$

From the Schwartz inequality we see that

$$\begin{aligned} \| Hx \|_2^2 &= \int_0^\infty \left( \int_0^t h(t, \tau)x(\tau) \right)^2 dt = \int_0^\infty \left( \int_0^\infty h(t, \tau)x(\tau) \, d\tau \right)^2 dt \\ &\leqq \int_0^\infty \left[ \int_0^\infty | h(t, \tau) |^2 \, d\tau \int_0^\infty | x(\tau) |^2 \, d\tau \right] dt \\ &\leqq \| h \|_2^2 \, \| x \|_2^2 , \end{aligned} \tag{34}$$

which implies that

$$\| H \| \leq \| h \|_2 . \tag{35}$$

Hence, compactness of $\hat{\mathfrak{D}}$ implies that $\mathfrak{D}$ is compact and condition (*i*) and (*ii*) of Theorem 1 are easily verified to hold.

## VI. COLORED NOISE

Theorems 1 and 2 were derived for the case when $z(t)$ the additive noise was a zero mean white stochastic process. The situation when $Ez(t)z(\tau) = R(t, \tau)$ can be handled in a similar fashion. The only additional assumptions are:

(*i*) $R(t, \tau)$ is positive definite; that is,

$$\int_0^\infty \int_0^\infty R(t, \tau)w(t)w(\tau) \, dt \, d\tau > 0 \quad \text{for all} \quad w \, \varepsilon \, L_2(0, \infty)$$

satisfying $\int_0^\infty | w(t) |^2 \, dt > 0$, and either

(*ii*) $R(t, \tau)$ is Hilbert-Schmidt; that is,

$$\int_0^\infty \int_0^\infty | R(t, \tau) |^2 \, dt \, d\tau = C^2 < \infty , \quad \text{or}$$

(*iii*) if $R(t, \tau) = R_0(t - \tau)$ then

$$\int_{-\infty}^\infty | R_0(t) |^2 \, dt = C_0^2 < \infty .$$

Inspecting the proof of Theorem 1, one sees that the whiteness assumption was only used in equations (8) and (17). If $Ez(t)z(\tau) = R(t, \tau)$, then equation (8) becomes

$$\begin{aligned} \sigma_{kj}^2 &= \text{Var}\,[q(w, k) - q(w, j)] \\ &= 4 \int_0^{T_i} \int_0^{T_i} R(t, \tau)(H_k x_i - H_j x_i)(t) \cdot (H_k x_i - H_j x_i)(\tau) \, dt \, d\tau . \end{aligned} \tag{36}$$

Since $H_k$ and $H_j$ were chosen so that $\| P_{T_i}(H_k x_i - H_j x_i) \| > 0$, we see that since $R(t, \tau)$ is positive definite, $\sigma_{kj}^2 > 0$. If we choose $\theta_i$ to be less than $\min_{j,k} \sigma_{jk}^2$ instead of $\| P_{T_i}(H_k x_i - H_j x_i) \|_2^2$, inequality (9) will remain true.

Equation (17) is changed as follows. If $Ez(t)z(\tau) = R(t, \tau)$, then by the Schwartz inequality

$$\begin{aligned} \sigma_{j\bar{k}}^2 &= 4 \int_0^{T_i} \int_0^{T_i} R(t, \tau)(H_j x_i - H_{\bar{k}} x_i)(t)(H_j x_i - H_{\bar{k}} x_i)(\tau) \, dt \, d\tau \\ &\leqq 4 \int_0^{T_i} (H_j x_i - H_{\bar{k}} x_i)(\tau) \left\{ \int_0^{T_i} | R(t, \tau) |^2 \, dt \right\}^{\frac{1}{2}} \\ &\quad \cdot \left\{ \int_0^{T_i} (H_j x_i - H_{\bar{k}} x_i)^2(t) \, dt \right\}^{\frac{1}{2}} d\tau \\ &\leqq 4 \left\{ \int_0^{T_i} \int_0^{T_i} | R(t, \tau) |^2 \, dt \, d\tau \right\}^{\frac{1}{2}} \| P_{T_i}(H_j x_i - H_{\bar{k}} x_i) \|_2^2 \\ &\leqq 4CR^2 s^2 T_i . \end{aligned} \tag{37}$$

On the other hand, if $Ez(t)z(\tau) = R_0(t - \tau)$, equation (17) is changed as follows.

$$\begin{aligned} \sigma_{j\bar{k}}^2 &= 4 \int_0^{T_i} \int_0^{T_i} R(t - \tau)(H_j x_i - H_{\bar{k}} x_i)(t)(H_j x_i - H_{\bar{k}} x_i)(\tau) \, dt \, d\tau \\ &\leqq 4 \int_0^{T_i} | (H_j x_i - H_{\bar{k}} x_i)(t) | \left[ \int_0^{T_i} | R(t - \tau) |^2 \, d\tau \right]^{\frac{1}{2}} \end{aligned}$$

$$\cdot \left[ \int_0^{T_i} (H_j x_i - H_{\bar{k}} x_i)^2(\tau)\, d\tau \right]^{\frac{1}{2}} dt$$

$$\leqq 4C_0 Rs(T_i)^{\frac{1}{2}} \int_0^{T_i} | (H_j x_i - H_{\bar{k}} x_i)(t) |\, dt$$

$$\leqq 4C_0 Rs(T_i)^{\frac{1}{2}} \left( \int_0^{T_i} dt \right)^{\frac{1}{2}} \left[ \int_0^{T_i} (H_j x_i - H_{\bar{k}} x_i)^2(t)\, dt \right]^{\frac{1}{2}}$$

$$\leqq 4C_0 Rs(T_i)^{\frac{1}{2}} \cdot (T_i)^{\frac{1}{2}} \cdot Rs(T_i)^{\frac{1}{2}} = C_0 R^2 s^2 T_i^{\frac{3}{2}} \,. \tag{38}$$

From equation (37) we see that the limit of integration in equation (19) now becomes $3rT_i^{\frac{1}{2}}/4RsC^{\frac{1}{2}}$. If we use equation (38), this limit becomes $3rT_i^{\frac{1}{2}}/16RsC_0^{\frac{1}{2}}$. In either case this limit diverges as $i$ increases. Thus Theorem 1 is still correct if the noise is colored. One can also see that Theorem 2 is true without any modifications. The whiteness assumption does enter into the proof in any substantial manner.

## VII. CONCLUSIONS

In this paper we have attempted to formalize the notion of identification and examined conditions under which the *a priori* information would guarantee that the problem of identification was well formulated. Our purpose has been to indicate when identification was possible and not to specify a given identification procedure. It is hoped that the conditions derived here may motivate researchers to consider larger classes of identification problems than have hitherto been examined and also to indicate for what classes of problems identification is not possible.

## VIII. ACKNOWLEDGMENTS

The author wishes to thank J. M. Holtzman and P. P. Varaiya for their helpful discussions and criticisms. He also wishes to acknowledge support for the initial phases of this work from NASA under Grant NsG-354, Supplement 4 while at the University of California at Berkeley.

## APPENDIX

### *Proof that the Space $\mathcal{H}$ Is Complete*

In this appendix we show that the space $\mathcal{H}$ with the metric induced by its norm is a complete space. If $\{H_n\}$ is a Cauchy sentence in $\mathcal{H}$, we show that there exists an element $\tilde{H} \in \mathcal{H}$ such that $\lim_{n\to\infty} \| \tilde{H} - H_n \| = 0$.

Let $\{H_n\}$ be a Cauchy sequence in $\mathcal{H}$. Then given any $\epsilon > 0$ there exists a number $N(\epsilon)$ such that if $m, n > N(\epsilon)$, $\| H_n - H_m \| < \epsilon$. From the definition of the metric,

$$\| H_n - H_m \| \geqq \frac{\| P_T H_n x - P_T H_m x \|_2}{\| P_T x \|_2} \tag{39}$$

for all $T > 0$, $x \in L_\infty(s)$, $\| P_T x \|_2 \neq 0$. Using the definition of $L_\infty(s)$,

$$\epsilon s(T)^{\frac{1}{2}} \geq \epsilon \| P_T x \|_2 \geq \| P_T(H_n x - H_m x) \|_2 \tag{40}$$

for all $n, m > N(\epsilon)$, $T > 0$, $x \in L_\infty(s)$, $\| P_T x \|_2 \neq 0$. Thus, for each $T > 0$, $x \in L_\infty(s)$, $\| P_T x \|_2 \neq 0$, $\{H_n x\}$ is a sequence of functions in $L_{2e}$ and for each $T > 0$, $P_T H_n x$ is a Cauchy sequence in $L_2[0, T]$. Hence, for each $T$ there exists at least one time function $y_T \in L_{2e}$ such that $P_T y_T \in L_2(0, \infty)$ and $\lim_{n \to \infty} \| P_T H_n x - P_T y_T \|_2 = 0$. Furthermore, $y_T$ is uniquely (except for a set of measure zero) specified over $[0, T]$. Because of this uniqueness, if $T_1 < T_2$, then $P_{T_1} y_{T_2} = P_{T_1} y_{T_1}$. Hence there exists a unique function $\tilde{y} \in L_{2e}$ such that $P_T \tilde{y} = P_T y_T$ for each $T > 0$. This function can be constructed:

$$\begin{aligned} \tilde{y}(t) &= y_1(t) \quad \text{for} \quad 0 \leqq t < 1 \\ &= y_2(t) \quad \text{for} \quad 1 \leqq t < 2 \\ &\vdots \qquad\qquad\quad \vdots \\ &= y_n(t) \quad \text{for} \quad n - 1 \leqq t < n \\ &\vdots \qquad\qquad\quad \vdots \end{aligned} \qquad . \tag{41}$$

For each $x \in L_\infty(s)$, $x \neq 0$ we have uniquely specified a function $\tilde{y} \in L_{2e}$. For $x = 0$ we arbitrarily put $\tilde{y} = 0$. Call the operator defined by this association $\tilde{H}$; that is, $\tilde{H}x = \tilde{y}$. We now show that $\lim_{n \to \infty} \| \tilde{H} - H_n \| = 0$.

For each $T > 0$, $x \in L_\infty(s)$, $\| P_T x \|_2 \neq 0$ we can use the triangle inequality to show that

$$\frac{\| P_T(\tilde{H}x - H_n x) \|_2}{\| P_T x \|_2} \leqq \frac{\| P_T \tilde{H} x - P_T H_m x \|_2}{\| P_T x \|_2} + \frac{\| P_T(H_n x - H_m x) \|_2}{\| P_T x \|_2}. \tag{42}$$

If $H_n$, $H_m$ are members of the Cauchy sequence, from our previous development we know that there exists a number $N(\epsilon/2)$ independent of $x$ and $T$ such that

$$\frac{\| P_T(H_n x - H_m x) \|_2}{\| P_T x \|_2} < \epsilon/2 \quad \text{for} \quad m, n > N(\epsilon/2). \tag{43}$$

Since $\lim_{m\to\infty} \| P_T(\tilde{H}x - H_m x) \|_2 = 0$ we can find another number $N^*(\epsilon/2, x, T) > N(\epsilon/2)$ such that

$$\frac{\| P_T(\tilde{H}x - H_m x) \|_2}{\| P_T x \|_2} < \epsilon/2 \quad \text{for} \quad m > N^*(\epsilon/2, x, T). \tag{44}$$

Hence for all $T > 0$, $P_T x \neq 0$

$$\frac{\| P_T(\tilde{H}x - H_n x) \|_2}{\| P_T x \|_2} < \epsilon \quad \text{for} \quad n > N(\epsilon/2), \tag{45}$$

and if $\tilde{H}$ were causal it follows that $\tilde{H} \in \mathcal{H}$ with $\lim \| \tilde{H} - H_n \| = 0$. The causality of $\tilde{H}$ is easily established. For each $x \in L_\infty(s)$, $T > 0$:

$$\| P_T\tilde{H}x - P_T\tilde{H}P_T x \|_2$$

$$\leqq \| P_T\tilde{H}x - P_T H_n x \|_2 + \| P_T\tilde{H}P_T x - P_T H_n x \|_2 \tag{46}$$

$$= \| P_T(\tilde{H}x - H_n x) \|_2 + \| P_T(\tilde{H}P_T x - H_n P_T x) \|_2 . \tag{47}$$

For $n$ sufficiently large each term on the right side may be arbitrarily small, hence $\| P_T\tilde{H}x - P_T\tilde{H}P_T x \|_2 = 0$ for all $x \in L_\infty(s)$, $T > 0$.

If $\mathcal{H}$ is composed only of linear operators the completeness proof follows as above except to additionally observe that $\tilde{H}$ is linear.

## REFERENCES

1. Prosser, R. T., and Root, W. L., "Determinable Classes of Channels," J. Math. and Mechanics, *16,* No. 4 (October 1966), pp. 365–398.
2. Root, W. L., "On System Measurement and Identification," *Proc. Symp. on System Theory,* Polytechnic Institute of Brooklyn, April 1965, pp. 133–157.
3. Root, W. L., "On the Measurement and Use of Time-Varying Communication Channels," Inform. and Control, *8,* No. 4 (August 1965), pp. 390–422.
4. Varaiya, P. P., "On Determinable Classes of Signals and Linear Channels," SIAM J. Appl. Math., *15,* No. 2 (March 1967), pp. 440–449.
5. Root, W. L., unpublished work.
6. Dunford, N., and Schwartz, J. T., *Linear Operators-Part I,* New York: Interscience Publishers, 1964.