Detection and Estimation with Fixed Lag for Abruptly Changing Systems

V. JOHN MATHEWS
JITENDRA K. TUGNAIT
University of Iowa

The problem of state estimation and system structure detection for discrete-time stochastic systems with parameters which may switch among a finite set of values is considered. The switchings are modeled by a semi-Markov, or Markov, chain with known transition statistics. A fixed time delay (lag) is allowed in estimation (smoothing) and detection. The optimal solutions require geometrically increasing computations and storage with time. Suboptimal solutions are proposed to alleviate this problem and simulation results are presented to illustrate the effectiveness of the proposed algorithms and the advantages of introducing a delay in processing of the observations.

Manuscript received April 20, 1982; revised March 31, 1983.

This work was supported in part by the National Science Foundation under Grant ECS-80-05956.

Authors' current addresses: V.J. Mathews, Department of Electrical and Computer Engineering, University of Iowa, Iowa City, IA 52242; J.K. Tugnait, Exxon Production Research Company, P.O. Box 2189, Houston, TX 77001.

0018-9251/83/0900-0730 $1.00 © 1983 IEEE

I. INTRODUCTION

This paper is concerned with state estimation and system structure detection for linear discrete-time stochastic systems with abruptly changing parameters. It is assumed that a fixed time delay (lag) is allowed in estimation and detection, i.e., the processes of estimation and detection at time t utilize measurements till time $t+N$ where the lag N is a nonnegative integer. The abruptly changing parameters are modeled as a finite-state semi-Markov, or Markov, chain with known transition statistics.

Motivation for considering system models with jumps stems from applicability of such models to a large class of realistic problems. Failures in components or subsystems of a dynamical system can be represented by abrupt changes in the system parameters [13–15]. Similarly, repairs of the failed components and system reconfigurations also cause abrupt changes in the system parameters [14]. Approximation of nonlinear systems by a set of linearized models to cover the entire dynamic range can also lead to linear systems with jump parameters [16–18]. Modeling with semi-Markov or Markov jump parameters may also be thought of as an extension of the multiple model partitioning approach to estimation, detection, and control of systems with unknown time-invariant parameters [19–22] to include systems with (abruptly) time-varying parameters [4, 7]. For some other motivations, see [1, 3, 23].

The problem of state estimation and system structure detection with zero lag, for abruptly changing systems, has received considerable attention [1–7, 16]. A recent survey concerning linear discrete-time systems with Markov jump parameters may be found in [7]. In [1, 2, 4–7], the abruptly changing parameters have been modeled as a finite-state Markov chain whereas, in [3, 16–18, 23, 25], a semi-Markov chain model has been employed. The semi-Markov chain model is more general in that it includes the Markov chain model [9–11]. The optimal solutions to the filtering and the detection problems are intractable; therefore, in [1–7, 16], efforts are directed toward finding suboptimal solutions.

For linear systems with completely known parameters, it is well known that state estimation with fixed time delay (fixed-lag smoothing) leads to an improvement in the performance at the cost of increased complexity when compared with the zero-lag case [8, 12]. In this paper we examine the consequences of introducing a fixed lag for systems with jump parameters. It is assumed that this time delay is of little consequence in the intended application. This is certainly true if the intended application is off-line analysis of a maneuvering target tracking problem [3, 4, 16].

The paper is organized as follows. In Section II a formal statement of the problem is presented. In Section III, we discuss the optimal solution to the state estimation and the system structure detection problems. Since the optimal solutions are computationally intractable, suboptimal algorithms are proposed in Section IV. Several simulation examples are presented in Section V to illustrate the effectiveness of the proposed approximations.

II. PROBLEM STATEMENT

Let $r(t) \in S = \{1, 2, ..., s\}$, $t \in \{0, 1, 2, ...\}$, denote a finite-state, discrete-time, stochastic process (either a Markov chain or a semi-Markov chain) with completely known probability laws. The process $r(t)$ governs the structure of a stochastic dynamical system. The system state equation is given by

$$x(t+1) = A(r(t+1))x(t) + B(r(t+1))w(t) + b(r(t+1)) \quad (1)^1$$

where $x(t) \in R^n$ is the system state, $w(t) \in R^l$ is a zero-mean white Gaussian noise sequence with covariance Q, and $b(r(t+1) \in R^n$ is a "bias" input which cannot be measured (e.g., it may be inaccessible). The matrices $A(r(t))$ and $B(r(t))$ are functions of the chain $r(t)$ and so is the vector $b(r(t))$. The observation equation associated with (1) is modeled by

$$z(t) = C(r(t))x(t) + D(r(t))v(t) + g(r(t)) \quad (2)$$

where $z \in R^m$ is the observation vector, $v \in R^m$ is a zero-mean white Gaussian measurement noise sequence with covariance R such that[2] $D_i R D_i^T > 0$ $(\forall i \in S)$ where $D(r(t)) \in \{D_i, i = 1, 2, ..., s\}$ and $g(r(t))$ is an unmeasurable "bias" input. The initial state $x(0)$ is assumed to be Gaussian with mean $\bar{x}_0$ and covariance matrix P_0. Finally, $x(0)$, $v(t)$, $w(t)$, and $r(t)$ are mutually independent.

The objective is to find the minimum mean-square error (MMSE) smoothed estimate $\hat{x}(t-N|t)$ of system state $x(t-N)$ given the observations $Z_t \triangleq \{z(k), 0 \le k \le t\}$ and to decide on the value of $r(t-N)$ (system structure detection), given Z_t, minimizing the probability of error. Here N (a fixed positive integer) is the fixed lag in estimation and detection.

III. OPTIMAL SOLUTION

A. Fixed-Lag Smoothing

It is well known that the MMSE smoothed state estimate $\hat{x}(t-N|t)$ is given by the conditional mean

$$\hat{x}(t-N|t) = E\{x(t-N)|Z_t\} . \quad (3)$$

Define a system structural state sequence $I(t)$ as

$$I(t) = \{r(0), r(1), .., r(t)\} . \quad (4)$$

[1]In the right-hand side of (1) we have used time index $t+1$ for $r(\cdot)$ to indicate that the quantities involved influence the system state $x(t+1)$ at time $t+1$; hence, they influence observation $z(t+1)$.

[2]$R>0$ denotes that the matrix R is positive definite.

Denote the smoothed state estimate conditional on a specific sequence as

$$\hat{x}_j(t-N|t) \triangleq E\{x(t-N)|Z_t, I_j(t)\} \quad (5)$$

where $I_j(t)$ denotes a specific sequence from the space of all possible sequences $I(t)$. This leads to

$$\hat{x}(t-N|t) = \sum_{j=1}^{s^{t+1}} \hat{x}_j(t-N|t)P(I_j(t)|Z_t) \quad (6)$$

where $P(I_j(t)|Z_t)$ is the conditional probability of the sequence $I_j(t)$ given the observations Z_t. By the Bayes' rule it follows that

$$P(I_j(t+1)|Z_{t+1}) = \frac{f(z_{t+1}|Z_t, I_j(t+1))P(I_j(t+1)|Z_t)}{\sum_{n=1}^{s^{t+2}} f(z_{t+1}|Z_t, I_n(t+1))P(I_n(t+1)|Z_t)} \quad (7)$$

where $f(z_{t+1}|Z_t, I_j(t+1))$ is the conditional probability density of the observation z_{t+1} given the past observations Z_t and the particular structural state sequence $I_j(t+1)$. Furthermore, we have

$$P(I_j(t+1)|Z_t) = P(I_m(t)|Z_t)P(r(t+1)|Z_t, I_m(t)) \quad (8)$$

where $I_j(t+1) = \{I_m(t), r(t+1)\}$ with $r(t+1) \in S$, i.e., $I_m(t)$ is a subsequence of $I_j(t+1)$ and $r(t+1)$ is the "last" or "most recent" element of $I_j(t+1)$. Computation of the second quantity of the right side of (8) depends upon the nature of the probability law of $\{r(t)\}$ and is discussed later in the section.

Now $f(z(t+1)|Z_t, I_j(t+1))$ and $\hat{x}_j(t-N|t)$ can be calculated recursively by applying Kalman filtering/smoothing methods to the system model (1)–(2). The initial condition $x(0)$ has a Gaussian distribution. Given the sequence $I_j(t+1)$, the system (1)–(2) is now a linear system with Gaussian noise and known parameters. Therefore, equations for $\hat{x}_j(t-N|t)$ are those for a Kalman smoother [8, sec. 7.3] matched to the sequence $I_j(t+1)$; the equations may be found in the appendix. Finally, it is easy to show that (see, e.g., [1]) $f(z(t+1)|Z_t, I_j(t+1))$ is Gaussian, where the mean and covariance can be calculated by using the various quantities associated with the Kalman filter/smoother matched to the sequence $I_j(t+1)$; the details are given in the appendix.

It remains to calculate $P(r(t+1)|Z_t, I_m(t))$ needed in (8). This depends upon the nature of the process $r(t)$. We discuss two cases.

1. *Markov Jump Parameters*

In this case the probability law for the Markov chain $r(t)$ is completely specified [11] by specifying the initial probabilities $P\{r(0) = i\}$, $i \in S$, and the transition probabilities $p_{ij} = P\{r(t+1) = j|r(t) = i\}$, $i,j \in S$; for simplicity, we assume that the Markov chain is homogenous, i.e., p_{ij} is not a function of time t. It is assumed that these probabilities are known. With $r(t)$ modeled as a Markov chain, we have

$$P(r(t+1)|Z_t, I_m(t)) = P(r(t+1)|I_m(t)) = P(r(t+1)|r(t) = i) \quad (9)$$

where the first step in (9) follows from conditional independence of $\{r(t)\}$ and $\{z(t)\}$ and the second step in (9) follows from the Markovian nature of $r(t)$ [11]. In (9), $r(t) = i$ is the "last" element of sequence $I_m(t)$.

2. *Semi-Markov Jump Parameters*

Suppose the process $\{r(t), t \geq 0\}$ is modeled as a homogeneous semi-Markov chain [9–11]. The semi-Markov process is a renewal process that makes its transitions according to the transition probability matrix of a Markov process, but whose time between transitions can be an arbitrary random variable. The Markov chains are special cases of the semi-Markov chains. By conditional independence of $\{r(t)\}$ and $\{z(t)\}$, it follows that

$$P(r(t+1)|Z_t, I_m(t)) = P(r(t+1)|I_m(t)) . \quad (10)$$

To compute the above quantity, we need to define the following probabilities. With $l \in S$, let $(n \geq 0)$

$$P_l(n) \triangleq P\{r(t) \neq r(t-1)|r(t-1) = r(t-2) = \cdots = r(t-n+1) = l \neq r(t-n)\} \quad (11)$$

i.e., $P_l(n)$ is the probability that a jump (transition) occurs at time t given that the last jump occurred at time $t-n+1$ to state $l \in S$. By homogeneity of the semi-Markov chain, $P_l(n)$ is not a function of time t. Also, we must have $\Sigma_{n=0}^{\infty} P_l(n) = 1$ for each $l \in S$. It should be noted that $P_l(n)$ defined in (11) is a function of n ("holding" or "sojourn" time) and l (the "current" state) only. Furthermore, we have [9–11]

$$\beta_{li}(n) = P_l(n) p_{li} \quad (12)$$

where $\beta_{li}(n)$ denotes the probability that a jump occurs to state i at time t given that the last jump occurred to state l at time $t-n+1$, and p_{li} is the transition probability of the imbedded Markov chain as defined in Section IIIA-1. It is assumed that $P_l(n)$, $\forall l \in S$ and $\forall n \geq 0$, and p_{ij}, $\forall i,j \in S$, are completely known.

To rewrite (1) in terms of $P_l(n)$ and p_{ij}, let $I_m(t) = \{r(0), r(1), \ldots, r(t) : r(t) = r(t-1) = \cdots = r(t-j+1) = l \neq r(t-j), j \geq 0\}$, i.e., $I_m(t)$ is such that the last transition occurs at time $t-j+1$ to state $l \in S$. Then we have

$$P(r(t+1)|I_m(t)) = \begin{cases} \beta_{li}(j+1) & \text{if } r(t+1) = i \neq l \\ 1 - P_l(j+1) & \text{if } r(t+1) = l . \end{cases} \quad (13)$$

It is easy to see from (6)–(10), the appendix, and (13) that $\hat{x}(t-N|t)$ can be computed recursively by running s^{t+1} "elemental" smoothers in parallel. For large t, this is clearly impractical. Therefore, one has to resort to suboptimal schemes. In the next section we propose a suboptimal algorithm based on the detection-estimation approach of [6, 7].

B. Structure Detection

Find $\hat{r}(t-N)$ such that

$$\hat{r}(t-N) = \arg\left\{\min_{i \in s} P(i \neq r(t-N)|Z_t)\right\} . \quad (14)$$

It is well known [12] that in such case we must have

$$\hat{r}(t-N) = \arg\left\{\max_{i \in S} \quad P(r(t-N) = i|Z_t)\right\} . \quad (15)$$

Now we have

$$P(r(t-N) = i|Z_t) = \sum_{j=1}^{s^t} P(I_j(t)|Z_t) \quad (16)$$

where $I_j(t)$ is a specific sequence from the space of all sequences $I(t)$ as defined in (4) such that its $(t-N+1)$th element $r(t-N) = i$. Note that out of s^{t+1} possible sequences at time t, s^t sequences satisfy this criterion.

Now (16), $\forall i \in S$, can be computed recursively using (7)–(10), the appendix, and (13), by running s^{t+1} elemental smoothers in parallel, as in Section IIIA. The approximation used to alleviate this difficulty is the same as for the smoothing problem; we discuss it in the next section.

IV. SUBOPTIMAL SOLUTION

In this section a suboptimal solution to the optimal smoothing problem is proposed. It is an extension of the detection-estimation approach to filtering proposed in [6]. The same approximation is used for structure detection also. The scheme consists of running at most M (instead of s^{t+1}) elemental smoothers in parallel at any time t where M is a fixed positive integer dictated by the computation and storage capabilities of the processor. A maximum of M "most likely" structural state sequences are selected based on their a posteriori probabilities $P(I(t)|Z_t)$ and then the state estimate is based on the selected sequences only.

Suppose that, at time $t-1$, we have selected L_{t-1} system structural state sequences where $L_{t-1} \leq M$. At time t consider all possible extensions of these L_{t-1} state sequences. By an extension $I(t)$ of $I_j(t-1)$ we mean

$$I(t) = \{I_j(t-1, r(t)\}, \qquad r(t) \in S .$$

Clearly, in all s extensions of $I_j(t-1)$ are possible. Let $I_l(t)$ denote an "extended" state sequence where $l = 1, 2, \ldots, (L_{t-1} \times s)$. Now compute the a posteriori probability $P(I_l(t)|Z_t)$ and choose L_t (where $L_t = M$ if $L_{t-1} \times s > M$ and $L_t = L_{t-1} \times s$ otherwise) most likely sequences out of the $L_{t-1} \times s$ extended sequences using the maximum a posteriori probability criterion which leads to minimum probability of error [12]. In other words, arrange $P(I_l(t)|Z_t)$ in a decreasing order of magnitude and select the first L_t sequences. (It should be noted that since $P(I_l(t)|Z_t) = [f(Z_t|I_l(t))P(I_l(t))/f(Z_t)]$ where the denominator is common for all values of l, only the numerator need be considered.) The approximate smoothed state estimate is then based on the selected structural state sequences. Therefore, (6) is modified as

$$\hat{x}(t-N|t) = \sum_{j=1}^{L_t} \hat{x}_j(t-N|t)P(I_j(t)|Z_t) \quad (17)$$

where

$$P(I_j(t)|Z_t) = \frac{f(Z_t|I_j(t))P(I_j(t))}{\sum_{l=1}^{L_t} f(Z_t|I_l(t))P(I_l(t))} \quad (18)$$

and subscripts j and l now index the selected L_t sequences.

Remark 1. Further refinements of the algorithm to reduce the computational requirements may be carried out following [6]. The number M can be viewed as a design parameter which is selected based on the computational and storage constraints.

Structure Detection. Equation (16) is modified in that the sequences on the right side of (16) are restricted to belong to the set of $L_{t-1} \times s$ extended sequences discussed earlier in this section.

Remark 2. The computational loads of the proposed approximations to the structure detection and the state smoothing problems are proportional to that of $M \times s$ elemental filters and M elemental smoothers, respectively.

Other Possible Approximations. The various approximations to state estimation and structure detection with zero lag (and Markov jump parameters) available in the literature can be classified into three broad categories [7]: (1) detection-estimation algorithm (DEA), (2) random sampling algorithm (RSA), (3) generalized pseudo-Bayes algorithm (GPBA). The suboptimal approach proposed earlier in this paper is based on the DEA. Approximations to smoothing based on RSA and GPBA can also be devised. It was noted in [7] that, in general, being a Monte Carlo technique, the RSA needs a fairly large number of elemental filters (smoothers in this paper) in parallel to be effective, i.e., it is computationally demanding. Therefore, we have not considered the RSA in this paper. The GPBA also has some drawbacks when applied to the smoothing problem. The essential assumption here is that the probability density of $x(t)$ conditioned on the past observations Z_t and the structural state sequence

$$I_j(t,t-n) \triangleq \{r(k), t-n \leq k \leq t\}, \qquad j = 1, 2, \ldots, s^{n+1}; \quad n \geq 0$$

is Gaussian, whereas, in truth, it is a Gaussian sum. Under this assumption, the smoothed state estimate is approximated as

$$\hat{x}(t-N|t) = \sum_{j=1}^{s^{n+1}} \hat{x}_j(t-N|t)P(I_j(t,t-n)|Z_t) \,. \quad (19)$$

In order to compute $\hat{x}_j(t-N|t)$ "matched" to $I_j(t,t-n)$, one needs filtered estimate $\hat{x}_j(t-N|t-N)$ and system parameters corresponding to $I_j(t,t-n)$; therefore, it is easy to see that one must have $n \geq N-1$ (for $N=0$, this amounts to $n \geq -1$ where $n = -1$ corresponds to assuming that the conditional density $f(x(t)|Z_t)$ is Gaussian, as is done in [1]). In other words, for given lag N, one has to run at least $s^N \times s$ elemental smoothers in parallel (which corresponds to s^N terms on the right side of (19)). This renders the GPBA impractical for large N. It was noted in [7] in the context of filtering that the performance of the GPBA improves considerably with increasing n. Therefore, in practice, it may be advisable to use $n \geq 0$. Further details regarding the GPBA may be found in [7] and references therein. We note that, for filtering, the GPBA with $n = 0$ has been used in [3, 4, 16–18, 25].

V. SIMULATION EXAMPLES

In this section several simulation examples are presented to illustrate the improvement in the estimation and the detection performances due to a (fixed) lag in the estimator and the detector. For reasons of implementability, the suboptimal algorithm proposed in Section IV (based on the detection-estimation approach) with only a "small" number of smoothers in parallel is considered.

Example 1

Consider a scalar dynamical system described by the following equations

$$x(t+1) = 1.04\,x(t) + w(t)$$

$$z(t) = C(r(t))x(t) + D(r(t))v(t),$$

$$t = 0, 1, 2, \ldots$$

$$s = 2, \qquad \text{i.e., } r(t) \in \{1, 2\} \,.$$

The initial conditions are as follows: $x(0) \sim N(30,400)$, $P(r(0) = 1) = P(r(0) = 2) = 0.5$. In the actual (true) system we use $x(0) = 1$ for all simulation runs to generate $z(t)$. Furthermore, $\{w(t)\}$ and $\{v(t)\}$ are mutually independent zero-mean white Gaussian noise sequences with covariances $Q = 0.1$ and $R = 1.0$, respectively. The process $r(t)$ is modeled by a Markov chain with transition probabilities:

$$p_{11} = P(r(t+1) = 1|r(t) = 1) = 0.85$$

$$p_{12} = P(r(t+1) = 2|r(t) = 1) = 0.15$$

$p_{22} = P(r(t+1) = 2 | r(t) = 2) = 0.3$

$p_{21} = P(r(t+1) = 1 | r(t) = 2) = 0.7$.

Finally, we take $D(1) = 40$, $D(2) = 1$, $C(1) = C(2) = 1$.

This example has been considered in [2, 6, 7]. The system is unstable and the state uncertainty increases with time t unless the measurements are processed. The choice of the transition probabilities implies that accurate measurement data occur only rarely.

The suboptimal algorithm was simulated for various design parameters (lag N and number of smoothers in parallel M) and operated on the same set of data. The structure detection and the state estimation performances were evaluated by averaging over 50 Monte Carlo runs. In Fig. 1, rms errors in state estimation are compared for various lags ($N = 0, 1, 2, 5$) with number of smoothers/

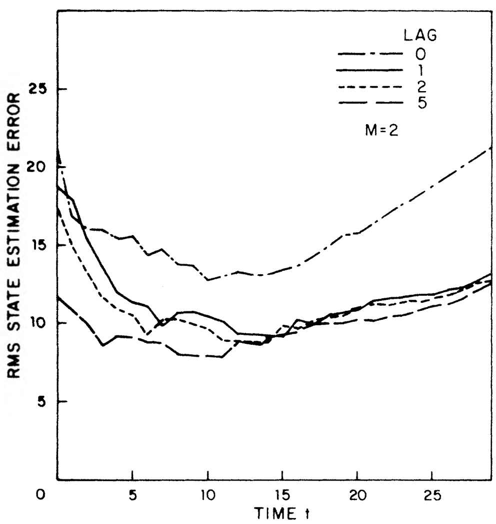


Fig. 1. RMS errors in state estimation for Example 1 for $M = 2$ and $N = 0, 1, 2, 5$.

filters in parallel fixed at 2, i.e., $M = 2$. Fig. 2 shows the probability of error in structure detection for $M = 2$ and $N = 0, 1, 2, 5$. The information in Figs. 1 and 2 is summarized in Table I after averaging over 30 time stages also. It is seen that, on the average, an increase in lag N results in improved performance. However, this is

TABLE I
Combined Ensemble and Time Averages of RMS State Estimation Error and Probability of Error in Structure Detection for Example 1

Number of Smoothers M	Lag N	Average RMS Error in State Estimation	Average Probability of Error in Structure Detection
2	0	16.02	0.15
2	1	11.68	0.12
2	2	11.10	0.11
2	5	9.84	0.10
4	0	12.75	0.13
8	0	8.07	0.13

not true for every time instant. The reason for this is that we are using a suboptimal algorithm and if an error is committed in selecting the M most likely structural state sequences, the error is compounded in that it first causes an error in filtering which in turn contributes to smoothing errors since filtering is a prerequisite to smoothing (see the appendix). It is also seen that the first few lags (in this case $N = 1$) lead to substantial performance improvement, whereas, for larger lags, any further (incremental) improvement is only marginal; this is a well-known fact [12].

In Figs. 3 and 4, the estimation and the detection performances, respectively, are shown for $N = 0$ (filtering or zero lag) and $M = 2, 4, 8$; see also Table I for time averages. It is seen that increasing the number M of smoothers/filters in parallel leads to improved performance (although not necessarily at every time instant, due to the reasons discussed earlier). However, the detection performance does not appear to be improved as much by increasing M with fixed $N = 0$. Note also

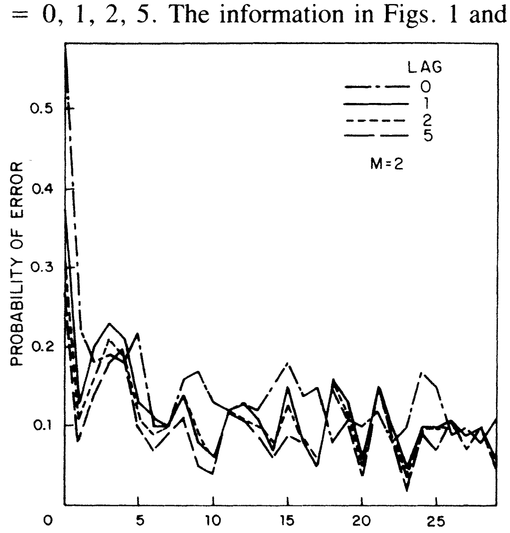


Fig. 2. Probability of error in system structure detection for Example 1 for $M = 2$ and $N = 0, 1, 2, 5$.

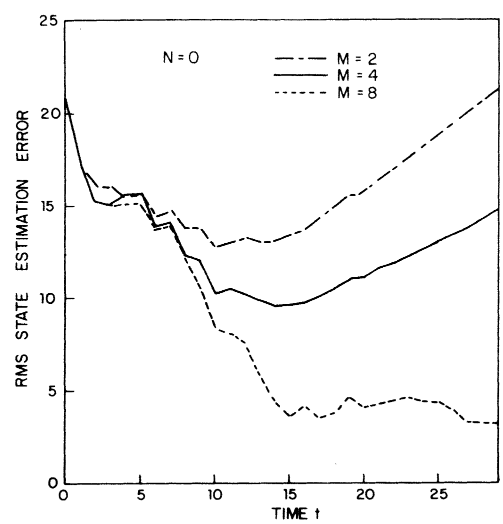


Fig. 3. RMS errors in state estimation for Example 1 for $N = 0$ and $M = 2, 4, 8$.

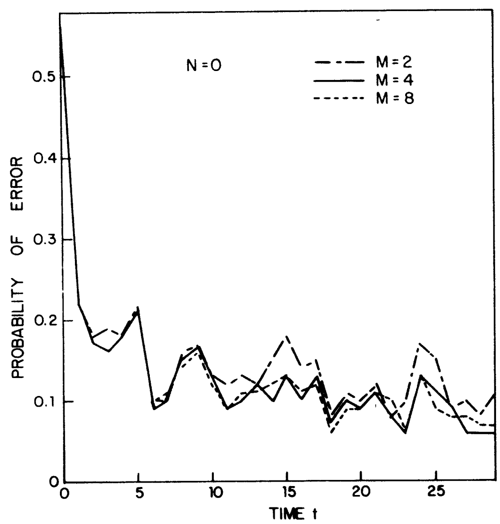


Fig. 4. Probability of error in system structure detection for Example 1 for $N = 0$ and $M = 2, 4, 8$.

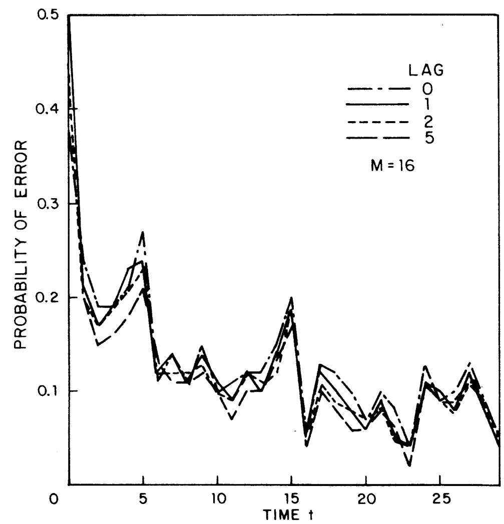


Fig. 6. Probability of error in system structue detection for Example 2 for $M = 16$ and $N = 0, 1, 2, 5$.

that, with $M = 2$, in Fig. 1, an increase in N for $N \geq 1$ affects primarily the initial uncertainty (for $t \leq 10$), whereas, from Fig. 3, an increase in M affects the state estimation accuracy primarily for $t \geq 10$. Therefore, provided the resources permit, increasing both M and N may be advantageous in a complementary sense.

Example 2

Again consider the system of Example 1 except that now $C(1) = 0$ and $C(2) = 1$; all other parameters remain unchanged. The detection and the estimation performances were evaluated by averaging over 50 Monte Carlo runs. Fig. 5 shows the rms error in state estimation

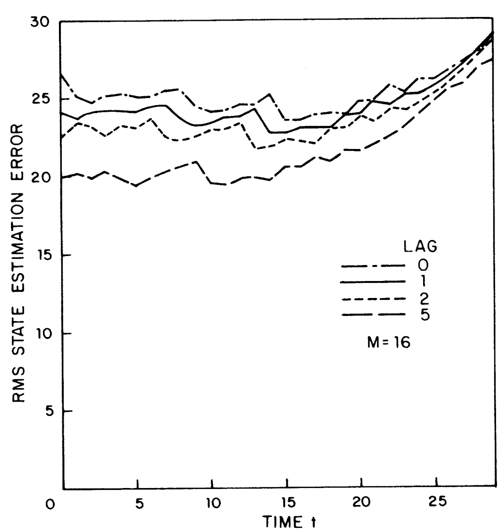


Fig. 5. RMS errors in state estimation for Example 2 for $M = 16$ and $N = 0, 1, 2, 5$.

for $M = 16$ and lag $N = 0, 1, 2, 5$, and Fig. 6 provides the corresponding information about probability of error in structure detection. Table II summarizes this information after averaging the Monte Carlo averages over 30 time stages. It is seen that an increase in lag N

TABLE II

Combined Ensemble and Time Averages of RMS State Estimation Error and Probability of Error in Structure Detection for Example 2

Number of Smoothers M	Lag N	Average RMS Error in State Estimation	Average Probability of Error in Structure Detection
16	0	25.23	0.15
16	1	24.52	0.14
16	2	23.73	0.13
16	5	21.66	0.12

leads to improvement in both estimation and detection performances. The choice of the transition probabilities (see Example 1) implies that most of the time the observations contain noise alone. Therefore, compared with Example 1, the estimation performance is much worse, as would be expected. The performance for $M < 16$ was still worse and has not been presented.

Example 3

Now we model $r(t)$ as a semi-Markov chain. The scalar system considered is described by

$$x(t+1) = 1.04\,x(t) + w(t)$$

$$z(t) = x(t) + D(r(t))v(t), \qquad t = 0, 1, 2, \ldots$$

$$s = 3, \qquad \text{i.e., } r(t) \in \{1, 2, 3\}\,.$$

We have $D(1) = 100$, $D(2) = 10$, and $D(3) = 1$. The initial conditions are $x(0) \sim N(30, 400)$, $P(r(0) = i) = 1/3$ for $i = 1, 2, 3$. In all simulation runs to generate $z(t)$ we use $x(0) = 1$ in the true system. Furthermore, $\{w(t)\}$ and $\{v(t)\}$ are mutually independent zero-mean white Gaussian noise sequences with covariances $Q = 0.1$ and $R = 1.0$, respectively. The process $r(t)$ is modeled by a semi-Markov chain with transition probabilities of the imbedded Markov chain given by $p_{11} = p_{22} = p_{33} = 0$, $p_{12} = 0.7$, $p_{13} = 0.3$, $p_{21} = 0.6$, $p_{23} = 0.4$, $p_{31} = 0.3$, and $p_{32} = 0.7$. The probability mass functions $P_i(n)$ defined in Section IIIA-2 are assumed to be $P_1(n) = a_1 \exp[-|n-3|]$, $P_2(n) = a_2 \exp[-|n-6|]$, and $P_3(n) = a_3 \exp[-|n-8|]$ for $n \geq 0$ with a_i such that $\Sigma_{n=0}^{\infty} P_i(n) = 1$, $i = 1, 2, 3$.

The results of 50 Monte Carlo runs average are shown in Figs. 7–10 for various values of N and M. The results of averaging over 30 time stages are displayed in Table III. The discussion of Example 1 pertaining to the effects of increasing N and M on the estimation and the detection performances applies here too.

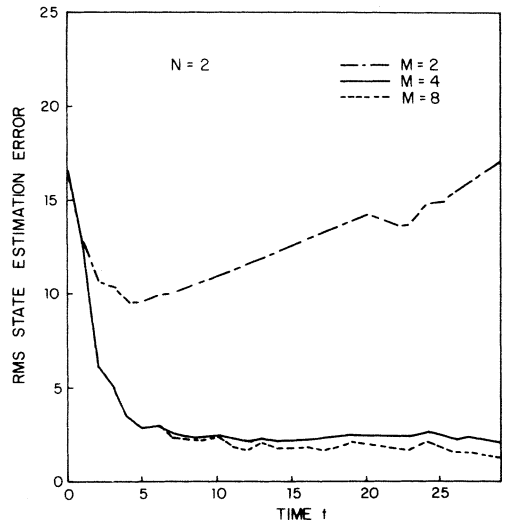


Fig. 7. RMS errors in state estimation for Example 3 for $N = 2$ and $M = 2, 4, 8$.

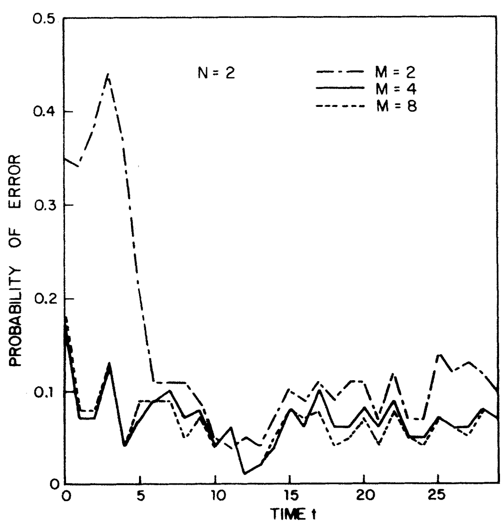


Fig. 8. Probability of error in system structure detection for Example 3 for $N = 2$ and $M = 2, 4, 8$.

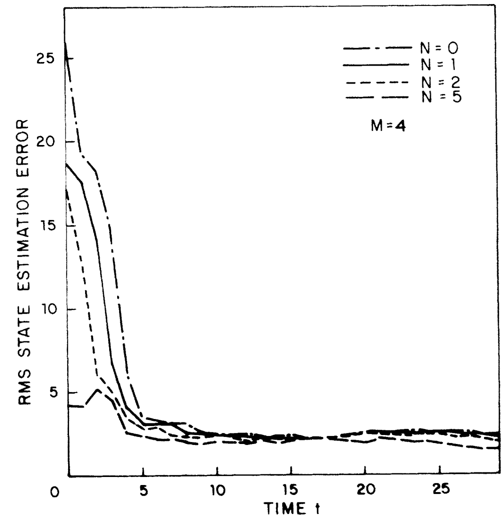


Fig. 9. RMS errors in state estimation for Example 3 for $M = 4$ and $N = 0, 1, 2, 5$.

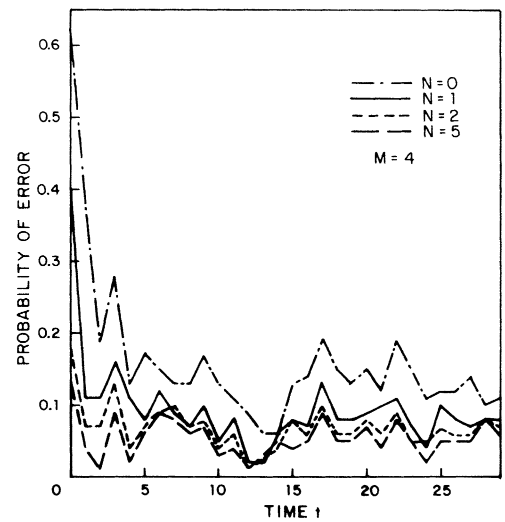


Fig. 10. Probability of error in system structure detection for Example 3 for $M = 4$ and $N = 0, 1, 2, 5$.

Table III
Combined Ensemble and Time Averages of RMS State Estimation Error and Probability of Error in Structure Detection for Example 3

Number of Smoothers M	Lag N	Average RMS Error in State Estimation	Average Probability of Error in Structure Detection
4	0	4.93	0.16
4	1	4.07	0.09
4	2	3.44	0.07
4	5	2.36	0.05
2	2	12.80	0.14
8	2	3.09	0.06

Example 4

Now we consider the systems of Examples 1 and 3, except that in the truth models (used for generating the data), we fix the structural state sequence $\{r(t), t \geq 0\}$. The suboptimal algorithm is designed assuming the switching models with transition statistics as used in Examples 1 and 3. The objective is to test the ability of the algorithm to identify the system structure if it is held constant for a sufficiently long length of time, even if the design transition statistics and the true transition statistics are mismatched.

The fixed sequence used for the truth model in Example 1 was selected as

$$r(t) = 1, \quad 0 \le t \le 9, \quad 20 \le t \le 29$$
$$= 2, \quad 10 \le t \le 19 .$$

The detector was designed assuming that $r(t)$ was a Markov chain as described in Example 1. In Fig. 11, the (approximate) a posteriori probabilities

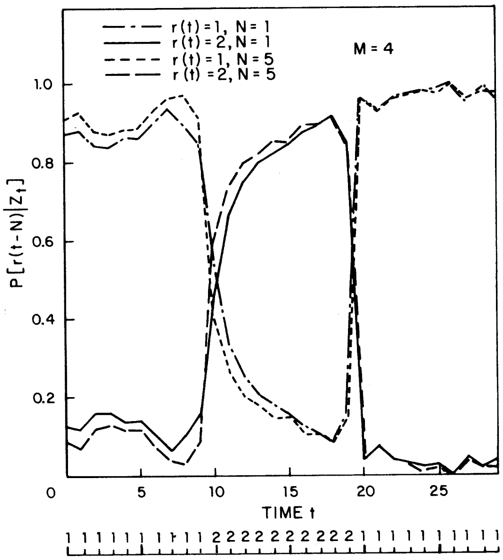


Fig. 11. A posteriori probabilities of system structural states for Example 4 assuming Markov switchings.

$P(r(t-N)|Z_t)$, $r(t) \in \{1, 2\}$, have been plotted for N = 1 and 5 and M = 4 after averaging over 50 Monte Carlo runs (the approximation results because of suboptimality of the detection-estimation algorithm). Ideally, one would like to have, for any N, $P(r(t-N)|Z_t) = 1$ if $r(t-N)$ = its true value and is equal to zero otherwise. It is seen from Fig. 11 that increasing N leads to improved discrimination between the hypothesized values of the Markov chain states.

The fixed sequence used for the truth model in Example 3 was selected as

$$r(t) = 1, \quad 0 \le t \le 9$$
$$= 2, \quad 10 \le t \le 19$$
$$= 3, \quad 20 \le t \le 29 .$$

The detector was designed assuming that $r(t)$ was a semi-Markov chain as described in Example 3. In Fig. 12, the (approximate) a posteriori probabilities $P(r(t-N)|Z_t)$, $r(t) \in \{1, 2, 3\}$, have been plotted for N = 1 and 5 and M = 4 after averaging over 50 Monte Carlo runs. Again, it is seen from Fig. 12 that increasing N leads to improved discrimination.

VI. CONCLUDING REMARKS

The problem of state estimation and system structure detection with a fixed lag for discrete systems with abruptly changing structure was considered. The changes were modeled by a Markov or a semi-Markov chain with known transition statistics. Since the optimal solution was impractical, a suboptimal approach was proposed and its efficacy was demonstrated through several simulation examples. It was shown that (as expected) a delay in processing the observations could lead to improvements in both estimation and detection performances.

We considered only the suboptimal approach based on the detection-estimation algorithm of [6]. Other approximations are also possible following [7] and were briefly discussed in Section IV. As discussed there, it was felt that the proposed algorithm is the most efficient computationally, given limited resources. However, for small lags, the generalized pseudo-Bayes algorithm may offer a good alternative (see the results pertaining to detection and estimation with zero lag in [7]). Finally, it should also be noted that recursive *optimal linear* filters and smoothers (when they exist) may perform better than *suboptimal nonlinear* filter/smoothers. Note, however, that for the system model (1)–(2), the recursive optimal linear state estimator may not exist [24]. Moreover, use of linear state estimators does not lead to the solution of the detection problem.

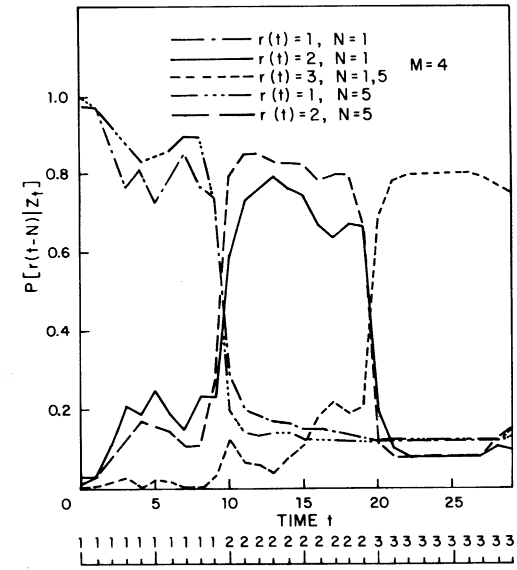


Fig. 12. A posteriori probabilities of system structural states for Example 4 assuming semi-Markov switchings.

APPENDIX

Equations are presented here for an elemental smoother match to a particular system structural state sequence. Equations for computing the density function $f(z(t+1)|Z_t, I_j(t+1))$ are also given. We follow [8, sec. 7.3] for the fixed-lag smoothing equations.

Let $I_j(t+1) \triangleq \{r(0), r(1), \ldots, r(t+1)\}$ and $\hat{x}_j(k|i) \triangleq E\{x(k)|Z_i, I_j(t+1)\}$ where $k, i \le t+1$. The equations for the smoothed estimate $\hat{x}_j(t-N|t)$ given Z_t and $I_j(t)$ are as follows:

$$\hat{x}_j(t-N|t) = \hat{x}_j(t-N|t-N) + \Sigma(t-N|t-N-1)\{\sum_{i=1}^{N} e_{i+1}(t+i-N)\} \quad (A1)$$

where

$$e_{i+1}(t) = [A(r(t-i+1)) - K(t-i) C(r(t-i))]e_i(t), \qquad i \geq 1 \quad (A2)$$

$$e_1(t) = C^T(r(t))[C(r(t))\Sigma(t|t-1)C^T(r(t)) + D(r(t))RD^T(r(t))]^{-1} \{z(t) - C(r(t))\hat{x}_j(t|t-1) - g(r(t))\} \quad (A3)$$

$$\Sigma(k|i) \triangleq E\{[x(k) - \hat{x}_j(k|i)] [x(k) - \hat{x}_j(k|i)]^T | I_j(t+1)\} \quad (A4)$$

and $k, i \leq t+1$. The smoother (A1)–(A3) uses the filtering solution given by

$$\hat{x}_j(t+1|t) = A(r(t+1))\hat{x}_j(t|t) + b(r(t+1)) \quad (A5)$$

$$\hat{x}_j(t|t) = \hat{x}_j(t|t-1) + K(t)\{z(t) - C(r(t))\hat{x}_j(t|t-1) - g(r(t))\} \quad (A6)$$

$$K(t) = \Sigma(t|t-1)C^T(r(t)) [C(r(t))\Sigma(t|t-1)C^T(r(t)) + D(r(t))RD^T(r(t))]^{-1} \quad (A7)$$

$$\Sigma(t|t-1) = A(r(t))\Sigma(t-1|t-1)A^T(r(t)) + B(r(t))AB^T(r(t)) \quad (A8)$$

$$\Sigma(t|t) = \Sigma(t|t-1) - K(t)C(r(t))\Sigma(t|t-1)\,. \quad (A9)$$

The initial conditions for the above equations are

$$\hat{x}_j(0|-1) = \bar{x}_0 \quad (A10)$$

$$\Sigma(0|-1) = P_0 \quad (A11)$$

since it has been assumed in Section II that $x(0) \sim N(\bar{x}_0, P_0)$.

Conditional Density Function

The conditional density function $f(z(t+1)|Z_t, I_j(t+1))$ is Gaussian; therefore, it suffices to calculate its mean $\hat{z}_j(t+1|t)$ and covariance $\Sigma_z(t+1|t)$. We have

$$\hat{z}_j(t+1|t) = E\{z(t+1)|Z_t, I_j(t+1)\} = C(r(t+1))\{A(r(t+1))\hat{x}_j(t|t) + b(r(t+1))\} + g(r(t+1)) \quad (A12)$$

$$\Sigma_z(t+1|t) = E\{[z(t+1) - \hat{z}_j(t+1|t)]\,[z(t+1) - \hat{z}_j(t+1|t)]^T | Z_t, I_j(t+1)\} = C(r(t+1))\Sigma(t+1|t)C^T(r(t+1)) + D(r(t+1))RD^T(r(t+1)) \quad (A13)$$

where $\Sigma(t+1|t)$ is given by equations similar to (A8) and (A9).

REFERENCES

[1] Ackerson, G.A., and Fu, K.S. (1970)
On state estimation in switching environments.
IEEE Transactions on Automatic Control, AC-15 (Feb. 1970), 10–17.

[2] Akashi, H., and Kumamoto, H. (1977)
Random sampling approach to state estimation in switching environments.
Automatica, 13 (July 1977), 429–434.

[3] Moose, R.L. (1975)
An adaptive state estimator solution to the maneuvering target problem.
IEEE Transactions on Automatic Conntrol, AC-20 (June 1975), 359–362.

[4] Chang, C.B., and Athans, M. (1978)
State estimation for discrete systems with switching parameters.
IEEE Transactions on Aerospace and Electronic Systems, AES-14 (May 1978), 418–425.

[5] Tugnait, J.K. (1979)
Comments on "State estimation for discrete systems with switching parameters."
IEEE Transactions on Aerospace and Electronic Systems, AES-15 (May 1979), 464.

[6] Tugnait, J.K., and Haddad, A.H. (1979)
A detection-estimation scheme for state estimation in switching environments.
Automatica, 15 (July 1979), 477–481.

[7] Tugnait, J.K. (1981)
Detection and estimation for abruptly changing systems.
In *Proceedings of the 20th IEEE Conference on Decision and Control,* San Diego, Calif., Dec. 1981, pp. 1357–1362; final version in *Automatica* (Sept. 1982).

[8] Anderson, B.D.O., and Moore, J.B. (1979)
Optimal Filtering.
Englewood Cliffs, N.J.: Prentice-Hall, 1979.

[9] Howard, R.A. (1964)
System analysis of semi-Markov processes.
IEEE Transactions on Military Electronics, MIL-8 (Apr. 1964), 114–124.

[10] Howard, R.A. (1971)
Dynamic Probabilistic Systems, vol. 2: Semi-Markov and Decision Processes.
New York: Wiley, 1971.

[11] Ross, S.M. (1970)
Applied Probability Models with Optimization Applications.
San Francisco, Calif.: Holden-Day, 1970.

[12] Van Trees, H.L. (1968)
Detection, Estimation, and Modulation Theory, Part I.
New York: Wiley, 1968.

[13] Sworder, D.D. (1976)
Control of systems subject to sudden change in character.
Proc. IEEE, 64 (Aug. 1976), 1219–1225.

[14] Birdwell, J.D., Castanon, D.A., and Athans, M. (1979)
On reliable control system designs with and without feedback reconfigurations.

In *Proceedings of the 17th IEEE Conference on Decision and Control*, San Diego, Calif., Jan. 1979, pp. 709–715.

[15] Willsky, A.S. (1976)
A survey of design methods for failure detection in dynamic systems.
Automatica, 12 (Nov. 1976), 601–611.

[16] Van Landingham, H.F., and Moose, R.L. (1977)
Digital control of high performance aircraft using adaptive estimation techniques.
IEEE Transactions on Aerospace and Electronic Systems, AES-13 (Mar. 1977), 112–120.

[17] Moose, R.L., Van Landingham, H.F., and Zwicke, P.E. (1978)
Digital set point control of nonlinear stochastic systems.
IEEE Transactions on Industrial Electronics and Control Instrumentation, IECI-25 (Feb. 1978), 39–45.

[18] Van Landingham, H.F., Moose, R.L., and Lucas, W.H. (1979)
Modelling and control of nonlinear plants.
In *Proceedings of the 17th IEEE Conference on Decision and Control*, San Diego, Calif., Jan. 1979, pp. 337–341.

[19] Lainiotis, D.G. (1976)
Partitioning: A unifying framework for adaptive systems, I: Estimation.
Proceedings of the IEEE, 64 (Aug. 1976), 1126–1142.

[20] Lainiotis, D.G. (1976)
Partitioning: A unifying framework for adaptive systems, II: Control.
Proceedings of the IEEE, 64 (Aug. 1976), 1182–1198.

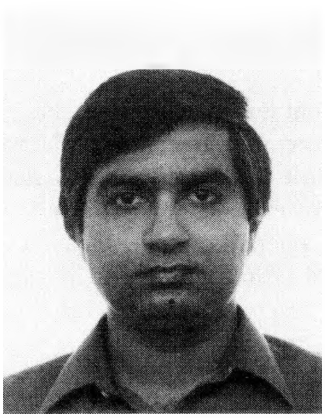
[21] Gustafson, D.E., et al. (1978)
ECG/VCG rhythm diagnosis using statistical signal analysis, I: Identification of persistent rhythms.
IEEE Transactions on Biomedical E ngineering, BME-25 (July 1978), 344–353.

[22] Willsky, A.S., et al. (1980)
Dynamic model-based techniques for detection of incidents on freeways.
IEEE Transactions on Automatic Control, AC-25 (June 1980), 347–360.

[23] Sanderson, A.C., Segen, J., and Richey, E. (1980)
Hierarchical modeling of EEG signals.
IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-2 (Sept. 1980), 405–415.

[24] Hadidi, M.T., and Schwartz, S.C. (1979)
Linear recursive state estimators under uncertain observations.
IEEE Transactions on Automatic Control, AC-24 (Dec. 1979), 944–948.

[25] Gholsun, N.H., and Moose, R.L. (1977)
Maneuvering target tracking using adaptive state estimation.
IEEE Transactions on Aerospace and Electronic System, AES-13 (May 1977), 310–317.



Jitendra K. Tugnait was born in Jabalpur, M.P., India, on December 3, 1950. He received the B.Sc. degree in electronics and electrical communication engineering (Hons.) from the Punjab Engineering College, Chandigarh, India, the M.S. and electrical engineering degrees from Syracuse University, Syracuse, N.Y., and the Ph.D. degree from the University of Illinois, Urbana (all in electrical engineering) in 1971, 1973, 1974, and 1978, respectively.

He was the recipient of a National Merit Scholarship from 1965 to 1971 in India and of a Government of India National Scholarship for Study Abroad from 1972 to 1974 at Syracuse University. From 1974 to 1976 he was in India, where he was associated first with the School of Radar Studies, Indian Institute of Technology, New Delhi, and later with the Digital Communications Techniques Division of the Space Applications Center, Indian Space Research Organization, Ahmedabad. He was a Research Assistant at the Coordinated Science Laboratory of the University of Illinois, Urbana, from 1976 to 1978. From 1978 to 1982 he was on the faculty of the Division of Information Engineering, University of Iowa, Iowa City. Since 1982 he has been with the Long Range Research Division, Exxon Production Research Company, Houston, Tex., as a Research Specialist.

V. John Mathews was born in Nedungadappally, India, in 1958. He received the B.E. (Hons.) degree from the University of Madras, India, in 1980 and M.S. degree from the University of Iowa, Iowa City, in 1981, all in electrical engineering. He is currently working toward his Ph.D. degree in electrical engineering at the University of Iowa.

Since 1980 he has held a teaching and research fellowship at the University of Iowa. His research interests include spectral estimation, adaptive filtering, and time delay estimation.