Journal of Intelligent and Robotic Systems 7: 95–112, 1993.

Generalized Pseudo-Bayes Estimation and Detection for Abruptly Changing Systems

KEIGO WATANABE
Department of Mechanical Engineering, Faculty of Science and Engineering, Saga University, Honjomachi-1, Saga 840, Japan

and

SPYROS G. TZAFESTAS
Computer Engineering Division, National Technical University of Athens, Zografou, Athens 15773, Greece

(Received: 25 January 1991)

Abstract: The problem of state estimation and system-structure detection for linear discrete-time systems with unknown parameters which may switch among a finite set of values is considered. The switching parameters are modeled by a Markov chain with known transition probabilities. Since the optimal solutions require exponentially growing storage and computations with time, a new method of generalized pseudo-Bayes algorithm (GPBA) is proposed to circumvent this problem by using a multi-stage measurement update technique. A minor modification is also presented to correct a defect of the Jaffer and Gupta method. Some simulation comparisons are included to illustrate the effectiveness of the proposed algorithms. It is then shown that, as compared with other GPBAs, a feature of the present GPBA is that it noticeably decreases the size of the required memory when the number of states in the Markov chain is large. The cost to be paid is a slight increase in the computing time.

Key words. Bayes methods, nonlinear filtering, Kalman filters, Markov processes, stochastic systems, failure detection.

1. Introduction

An estimation policy for the state of a linear stochastic system with abruptly changing parameters can be determined as the solution of a Kalman filtering problem only if the jump parameters are observed. The motivation for considering system models with jumps arises from the applicability of such models to a large class of realistic problems:

(i) fault detection for a dynamic system with failures in components or subsystems (Willsky, 1976; Gustafson *et al.*, 1978; Willsky, 1980; Watanabe, 1989a,b),

(ii) target tracking for a moving vehicle with sudden maneuvers (Ricker and Williams, 1978; Moose, *et al.*, 1979; Chang and Tabaczynski, 1984) and

(iii) approximation of a nonlinear system by a set of linearized models to cover the entire dynamic range (Moose *et al.*, 1978).

However, the jump parameters are generally unknown and, therefore, we cannot solve the above problems by using a single Kalman filter. One approach is to use the multiple model adaptive filtering (MMAF) approach (Lainiotis, 1971, 1976) developed for systems having unknown constant parameters, setting an upper (or lower) bound on the *a-posteriori* probabilities (Gustafson *et al.*, 1978; Willsky, 1980; Watanbe, 1988a,b). Another approach is to extend the MMAF approach to stochastic systems with possibly unknown, time-varying parameters, which are modeled as a finite-state Markov (or semi-Markov) chain state with known transition statistics. The latter approach seems to be more natural than the former approach, if the transition probabilities are known.

It is well known (Ackerson and Fu, 1970) that to evaluate the minimum mean-squared error (MMSE) estimate of the system in switching environment, the computational and storage requirements increase exponentially with time, which renders the optimal solution impractical. To circumvent this problem, there are some suboptimal algorithms, e.g. (a) random sampling algorithm (RSA) (Akashi and Kumamoto, 1977), (b) detection-estimation algorithm (DEA) (Tugnait and Haddad, 1979; Tugnait, 1982a; Mathews and Tugnait, 1983), (c) generalized pseudo-Bayes algorithm (GPBA) (Ackerson and Fu, 1970; Jaffer and Gupta, 1971; Chang and Athans, 1978; Sugimoto and Ishizuka, 1983), and (d) interacting multiple model algorithm (IMMA) (Blom, 1984, 1985; Blom and Bar-Shalom, 1988). It is worth noting here that the IMMA performs nearly as well as the second-order GPBA method with notably less computation (see also Chang and Bar-Shalom, 1987).

It is interesting to note that several simulations results presented by Tugnait (1982a) indicates that, in general, the GPBA is to be preferred, compared with the RSA and DEA methods, though the performance of the various algorithms is very much dependent upon the system model under consideration. However, when considering the suboptimal estimation problem by applying the GPBA, the major two questions are naturally arisen:

(A) how to cut the exponentially growing trees of Markov chain state, and
(B) how to embed the elemental Kalman filters into the cut trees.

For the first question, the GPBA due to Jaffer and Gupta (1971) is the most general solution. That is, they assume that the probability density function of the system state at time k conditioned on the past observations depends upon at most the last $n + 1$ Markov chain states, though the true density function depends upon the all past k Markov chain states, where $n \geqslant 1$ is a fixed preselected number. The GPBA due to Ackerson and Fu (1970) and Chang and Athans (1978) are the special case. The result of Sugimoto and Ishizuka (1983) is entirely the same as that of Jaffer and Gupta (1971).For the second question, Jaffer and Gupta (1971) also proposed a method for obtaining the s^{n+1} conditional estimates at time k by updating the one-step past s^n conditional estimates at time $k - 1$, which will be called here the 'one-step measurement update method', where s denotes the number of Markov chain states. It is

naturally noted that another measurement update method will exist so that the s^{n+1} conditional estimates at time k can be computed by updating the n-step past s conditional estimates at time $k - n$. This means that, in the framework of GPBA approach, $n - 1$ measurement data can be further reprocessed for a case when $n \geqslant 2$.

In this paper, we propose a new GP BA in the problem of state estimation and system structure detection for discrete-time stochastic systems with unknown jump parameters, which may be modeled as a finite-state Markov chain with known transition statistics. The present method is based on using the concept of Jaffer and Gupta (1971) as a solution to the first question, but using a new algorithm to the second question, which will be called the 'n-step measurement update method'. This method is very similar to an algorithm for fixed-lag smoothing for a lag of $n - 1$ units of time, but it simply consists of reprocessing $n - 1$ measurement data with s parallel filtered estimates as starting conditions at time $k - n$. Some interesting features of the new GPBA are (1) it requires arithmetically growing storage with s and n; (2) it can directly use the *a-posteriori* probability in the estimation problem as the *a-posteriori* probability for the unknown structure in the detection problem; and (3) it gives an improvement in the estimation performance at the transient stages for a certain case.

2. Problem Statement

Let $\alpha(k) \in S = \{1, 2, \ldots, s\}$, $k \in \{1, 2, \ldots\}$, denote a finite-state, discrete, Markov chain with completely known time-invariant transition probabilities

$$p_{ij} \triangleq \Pr\{a(k) = j \mid \alpha(k - 1) = i\}, \quad i, j \in S \tag{1}$$

and initial probability distribution $p_i = \Pr\{\alpha(0) = i\}$, $i \in S$. Let $\pi = [p_{ij}]$, an $s \times s$ matrix, denote the transition probability matrix. Now $\alpha(k)$ governs the structure of a stochastic dynamical system under the normal or failure mode. The system state equation is given by

$$x(k + 1) = A(\alpha(k + 1))x(k) + B(\alpha(k + 1))w(k) \tag{2}$$

where $x(k) \in \mathscr{R}^{n_1}$ is the system state, $w(k) \in \mathscr{R}^{p}$ is a zero-mean white Gaussian noise sequence with covariance Q. The observation equation associated with (2) is modeled by

$$z(k) = C(\alpha(k))x(k) + D(\alpha(k))v(k), \tag{3}$$

where $z(k) \in \mathscr{R}^{m_1}$ is the observation vector, $v(k) \in \mathscr{R}^{m_1}$ is a zero-mean white Gaussian measurement noise with covariance R such that $D_i R D_i^T > 0$ $(i \in S)$ where $D(\alpha(k)) \in \{D_i, i = 1, 2, \ldots, s\}$. The initial state is assumed to be subject to the following Gaussian distribution

$$x(0) \sim N(\bar{x}_0, P_0). \tag{4}$$

Finally, $x(0)$, $w(k)$, $v(k)$ and $\alpha(k)$ are mutually independent.

The objective is to find the minimum mean-squared error (MMSE) state estimate $\hat{x}(k \mid k)$ of $x(k)$ given the observations $Z_k = \{z(i), 1 \leqslant i \leqslant k\}$ and to decide on the value of $\alpha(k)$ (system fault detection), given Z_k, minimizing the probability of error.

3. Optimal Solution

3.1. STATE ESTIMATION

It is well-known that the MMSE filtered state estimate $\hat{x}(k \mid k)$ is given by the conditional mean

$$\hat{x}(k \mid k) = E[x(k) \mid Z_k]. \tag{5}$$

Define a Markov chain state sequence $I(k)$ as

$$I(k) \triangleq \{\alpha(1), \ldots, \alpha(k)\} \tag{6}$$

and let $I_j(k)$ denote a specific sequence from the space of all possible sequences $I(k)$ which contains s^k elements. If the state estimate conditioned on a specific sequence is written as

$$\hat{x}_j(k \mid k) = E[x(k) \mid I_j(k), Z_k], \tag{7}$$

then we have

$$\hat{x}(k \mid k) = \sum_{j=1}^{s^k} \hat{x}_j(k \mid k) p(I_j(k) \mid Z_k), \tag{8}$$

where $p(I_j(k) \mid Z_k)$ is the *a-posteriori* probability of $I_j(k)$ given Z_k, which is subject to

$$p(I_j(k) \mid Z_k) = \frac{f(z(k) \mid I_j(k), Z_{k-1}) p(I_j(k) \mid Z_{k-1})}{\sum_{l=1}^{s^k} f(z(k) \mid I_l(k), Z_{k-1}) p(I_l(k) \mid Z_{k-1})}, \tag{9}$$

where $f(\cdot \mid \cdot)$ is the conditional probability density of the observation $z(k)$ given the past observations Z_{k-1} and the particular state mode sequence $I_j(k)$. Furthermore, it is found (Mathews and Tugnait, 1983) that

$$p(I_j(k) \mid Z_{k-1}) = p(\alpha(k) \mid \alpha(k-1) = i) p(I_m(k-1) \mid Z_{k-1}), \quad i \in S, \tag{10}$$

because

$$\begin{aligned} &p(\alpha(k) \mid I_m(k-1), Z_{k-1}) \\ &\quad = p(\alpha(k) \mid I_m(k-1)) \\ &\quad = p(\alpha(k) \mid \alpha(k-1) = i), \quad i \in S, \end{aligned} \tag{11}$$

where the first equality in (11) follows from conditional independence of $\{\alpha(k)\}$ and $\{z(k)\}$, and the second equality in (11) follows from the Markovian nature of $\alpha(k)$. Here, $I_m(k-1)$ denotes a specific sequence from the space of all possible sequences $I(k-1)$ as defined in (6). The associated state estimation error covariance matrix

$$P(k \mid k) = E\{[x(k) - \hat{x}(k \mid k)][x(k) - \hat{x}(k \mid k)]^T \mid Z_k\} \tag{12}$$

is given (Lainiotis, 1971) by

$$P(k \mid k) = \sum_{j=1}^{s^k} \{P_j(k \mid k) + [\hat{x}_j(k \mid k) - \hat{x}(k \mid k)] \times [\hat{x}_j(k \mid k) - \hat{x}(k \mid k)]^T\} p(I_j(k) \mid Z_k) \tag{13}$$

where

$$P_j(k \mid k) = E\{[x(k) - \hat{x}_j(k \mid k)][x(k) - \hat{x}_j(k \mid k)]^T \mid I_j(k), Z_k\}. \tag{14}$$

Now, provided the initial information (4) and the sequence $I_j(k)$, we can obtain $\{\hat{x}_j(k \mid k), P_j(k \mid k)\}$ recursively by applying Kalman filters matched to sequences $I_j(k)$, $j = 1, \ldots, s^k$. Furthermore, the weighting probability $p(I_j(k) \mid Z_k)$ can also be computed from the information supplied by same Kalman filters, because $f(z(k) \mid I_j(k), Z_{k-1})$ is Gaussian. Thus, the optimal estimator (8) which is a weighted sum of s^k estimates $\hat{x}_j(k \mid k)$, requires an exponentially increasing memory and computational capability with time. Therefore, one has to resort to suboptimal schemes to circumvent this difficulty.

3.2. STRUCTURE DETECTION

In a structure (or fault) detection problem, we find $\alpha(k)$ such that

$$\hat{\alpha}(k) = \arg \left\{ \max_{\alpha \in S} p(\alpha(k) \mid Z_k) \right\}. \tag{15}$$

Now we have

$$p(\alpha(k) \mid Z_k) = \sum_{j=1}^{s^{k-1}} p(\alpha(k), I_j(k-1) \mid Z_k). \tag{16}$$

Thus, s^{k-1} Kalman filters in parallel are necessary to compute (16) for given $\alpha(k) \in S$. The approximation used to aleviate this difficulty is the same as for the state estimation problem.

4. Suboptimal Algorithms

The various approximations to state estimation or structure detection available in the literature have been classified into some categories as indicated in the introduction. In this section we shall focus on the GPBA and propose a new GPBA which is different from the GPBA due to Jaffer and Gupta (1971) in the view point of computational implementation.

4.1. GENERALIZED PSEUDO-BAYES ALGORITHM

The essential assumption of GPBA is that the probability density of the system state at time k conditioned on Z_k and the Markov chain state sequence $I_j(k, k-n) \triangleq \{\alpha(i), k - n \leqslant i \leqslant k\}, j = 1, \ldots, s^{n+1}, n \geqslant 1$ is Gaussian, whereas, in truth, it is

a Gaussian sum. That is, it is assumed that

$$f(x(k) \mid I_j(k, k-n), Z_k) \sim N(\hat{x}_j(k \mid k), P_j(k \mid k)). \tag{17}$$

Under this assumption, the state estimate $\hat{x}(k \mid k)$ and the associated state estimation error covariance matrix $P(k \mid k)$ are approximated by

$$\hat{x}(k \mid k) = \sum_{j=1}^{s^{n+1}} \hat{x}_j(k \mid k) p(I_j(k, k-n) \mid Z_k) \tag{18}$$

and

$$P(k \mid k) = \sum_{j=1}^{s^{n+1}} \{P_j(k \mid k) + [\hat{x}_j(k \mid k) - \hat{x}(k \mid k)] \times [\hat{x}_j(k \mid k) - \hat{x}(k \mid k)]^T\} p(I_j(k, k-n) \mid Z_k). \tag{19}$$

It is here significant to note that, given an observation $z(k)$, we can consider two approaches to updating the conditional estimates which have been obtained up to time $k-1$. One is that the s^{n+1} conditional estimates $\hat{x}_j(k \mid k)$ can be obtained by updating the one-step past s^n conditional estimates at time $k-1$. This method is called here the one-step measurement update method'. Another is based on computing the s^{n+1} conditional estimates by updating the n-step past s conditional estimates at time $k-n$, but with a condition that $z(k-1), \ldots, z(k-n+1)$ have already been stored. This method is called in this paper the 'n-step measurement update method'.

Structure detection. The approximation used for the state estimation algorithm also leads to a suboptimal detection algorithm. Equation (16) is approximated by

$$p(\alpha(k) \mid Z_k) = \sum_{j=1}^{s^n} p(\alpha(k), I_j(k-1, k-n) \mid Z_k) \tag{20}$$

because $I_j(k, k-n) = \{\alpha(k), I_j(k-1, k-n)\}$.

4.2. ONE-STEP MEASUREMENT UPDATE METHOD

Almost all approaches that have appeared in the literature are in this category. The fundamental form was first proposed in Ackerson and Fu (1970), in which $f(x(k) \mid Z_k)$ was assumed to be Gaussian. For the detailed algorithm, see Jaffer and Gupta (1971). It should be noted that Chang and Athans (1978) and Tugnait (1982) use a refined equation to evaluate the covariance matrix $P(k \mid k)$, whereas in Ackerson and Fu (1970) the following approximation is used

$$P(k \mid k) = \sum_{i=1}^{s} P_{i_k}(k \mid k) p(H_{i_k} \mid Z_k), \tag{21}$$

where H_{i_k} denotes the following hypothesis

$$H_{i_k}: \alpha(k) = i_k, \quad i \in S, \tag{22}$$

and in the original paper of Jaffer and Gupta (1971), this covariance matrix is not referred to. Although Tugnait (1982a) claimed that this difference was minor and had

not shown any appreciable difference in computer simulations, a similar modification is crucial, as will be found in the later simulations, in evaluating the following equation:

$$P_{i_{k-n+1},\ldots,i_k}(k\,|\,k) = \sum_{i_{k-n}=1}^{s} \frac{P^*_{i_{k-n},\ldots,i_k}(k\,|\,k)p(H_{i_{k-n}}, \ldots, H_{i_k}\,|\,Z_k)}{p(H_{i_{k-n+1}}, \ldots, H_{i_k}\,|\,Z_k)}, \tag{23}$$

where

$$\begin{aligned} &P^*_{i_{k-n},\ldots,i_k}(k\,|\,k) \\ &\quad = P_{i_{k-n},\ldots,i_k}(k\,|\,k) + [\hat{x}_{i_{k-n},\ldots,i_k}(k\,|\,k) - \hat{x}_{i_{k-n},\ldots,i_k}(k\,|\,k)] \\ &\qquad \times [\hat{x}_{i_{k-n},\ldots,i_k}(k\,|\,k) - \hat{x}_{i_{k-n},\ldots,i_k}(k\,|\,k)]^T. \end{aligned} \tag{24}$$

The one-step measurement update algorithm requires the storage of s^n estimates (n_1-dimension), s^n covariances ($n_1(n_1 + 1)/2$-dimension) and s^n *a-posteriori* probabilities (scalar), regardless of the number of the storage the system runs. As $n \to k - 1$, of course, the method approaches the optimal algorithm. Note here that for $k \leqslant n + 1$ the suboptimal method coincides with the optimal one. Figure 1 shows the time evolution of elemental Kalman filters for this method with $s = n = 2$. Notice that if all evolution in Figure (b) except for two evolution $\hat{x}_{11} \to \hat{x}_{111}$ and $\hat{x}_{22} \to x_{222}$ are neglected, then we can get the algorithm of Ackerson and Fu (1970).

Structure detection. The suboptimal detection algorithm (20) can be rewritten as

$$p(H_{i_k}\,|\,Z_k) = \sum_{i_{k-n}=1}^{s} \cdots \sum_{i_{k-1}=1}^{s} p(H_{i_{k-n}}, \ldots, H_{i_{k-1}}, H_{i_k}\,|\,Z_k) \tag{25}$$

or

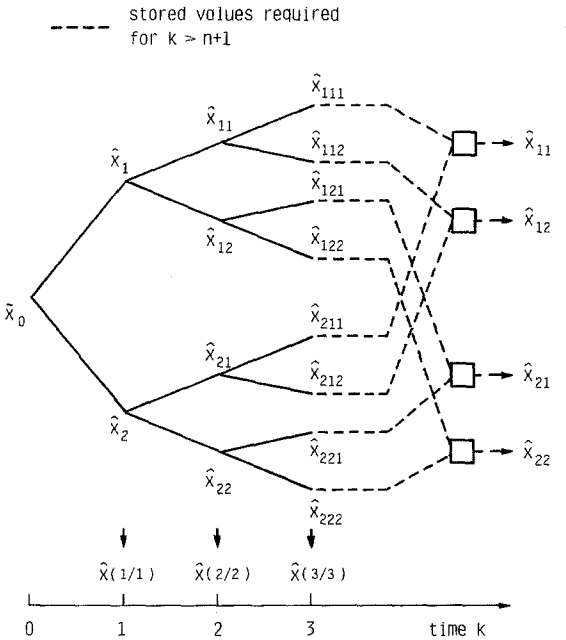
$$p(H_{i_k}\,|\,Z_k) = \sum_{i_{k-n+1}=1}^{s} \cdots \sum_{i_{k-1}=1}^{s} p(H_{i_{k-n+1}}, \ldots, H_{i_{k-1}}, H_{i_k}\,|\,Z_k) \tag{26}$$

4.3. n-STEP MEASUREMENT UPDATE METHOD

As discussed in Section 4.1, this approach is first based on reprocessing the past observations $z(k-1), \ldots, z(k-n+1)$ with the starting conditions at time $k - n$. After that, when provided $z(k)$ we can produce s^{n+1} conditional estimates by applying the most recent estimates (i.e. estimates at time $k - 1$), which had already been generated by the reprocessing. The resulting algorithm is summarized in the following.

n-step measurement update algorithm. At the time $k - n$, $k > n + 1$, the following quantities are provided

$$\begin{aligned} &\hat{x}_{i_{k-n}}(k-n\,|\,k-n) \triangleq E[x(k-n)\,|\,H_{i_{k-n}}, Z_{k-n}], \\ &P_{i_{k-n}}(k-n\,|\,k-n) \\ &\quad \triangleq E\{[x(k-n) - \hat{x}_{i_{k-n}}(k-n\,|\,k-n)] \\ &\qquad \times [x(k-n) - \hat{x}_{i_{k-n}}(k-n\,|\,k-n)]^T\,|\,H_{i_{k-n}}, Z_{k-n}\} \end{aligned}$$



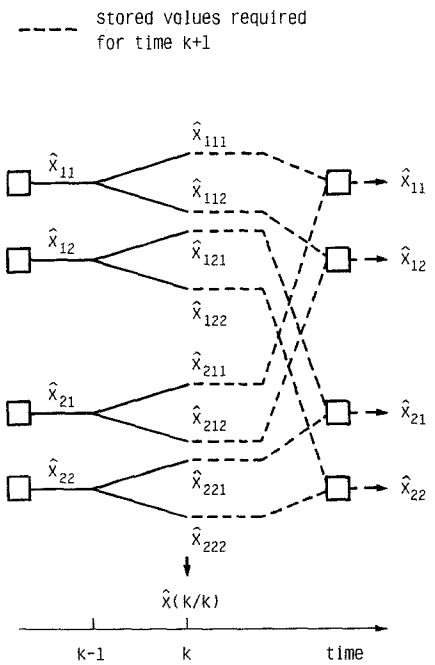


Fig. 1. Time evolution of elemental Kalman filters in one-step measurement update method; (a) for $k \leqslant n + 1$, (b) for $k > n + 1$, where $s = n = 2$.

and $p(H_{i_{k-n}} | Z_{k-n})$. Then, given $H_{i_{k-n+1}}$ and Z_{k-n+1} at time $k - n + 1$, we compute

$$\hat{x}_{i_{k-n},i_{k-n+1}}(k - n + 1 | k - n + 1) \triangleq E[x(k - n + 1) | H_{i_{k-n}}, H_{i_{k-n+1}}, Z_{k-n+1}]$$

and

$$\begin{aligned} &P_{i_{k-n},i_{k-n+1}}(k - n + 1 | k - n + 1) \\ &\quad \triangleq E\{[x(k - n + 1) - \hat{x}_{i_{k-n},i_{k-n+1}}(k - n + 1 | k - n + 1)] \\ &\qquad \times [x(k - n + 1) - \hat{x}_{i_{k-n},i_{k-n+1}}(k - n + 1 | k - n + 1)]^T | H_{i_{k-n}}, H_{i_{k-n+1}}, Z_{k-n+1}\} \end{aligned}$$

by applying the s^2 parallel Kalman filters. In addition, compute the probability density functions

$$\begin{aligned} f(z(k - n + 1) | H_{i_{k-n}}, H_{i_{k-n+1}}, Z_{k-n}) &\sim N[C(\alpha(k - n + 1) \\ &= i_{k-n+1})\hat{x}_{i_{k-n},i_{k-n+1}}(k - n + 1 | k - n), V_{i_{k-n},i_{k-n+1}}(k - n + 1 | k - n)] \end{aligned} \tag{27}$$

where

$$\begin{aligned} &V_{i_{k-n},i_{k-n+1}}(k - n + 1 | k - n) \\ &\quad = C(\alpha(k - n + 1) = i_{k-n+1})P_{i_{k-n},i_{k-n+1}}(k - n + 1 | k - n) \\ &\qquad \times C^T(\alpha(k - n + 1) = i_{k-n+1}) + D(\alpha(k - n + 1) = i_{k-n+1}) \\ &\qquad \times RD^T(\alpha(k - n + 1) = i_{k-n+1}). \end{aligned} \tag{28}$$

Applying the same manner for time stages $k - n + 2, \ldots, k$, we compute the parallel Kalman filters $s^3, \ldots, s^{n+1}$. Then, Equations (25) and (26) can be rewritten as

$$\hat{x}(k | k) = \sum_{i_k=1}^{s} \hat{x}_{i_k}(k | k)p(H_{i_k} | Z_k), \tag{29}$$

$$\begin{aligned} P(k | k) &= \sum_{i_k=1}^{s} \{P_{i_k}(k | k) + [\hat{x}_{i_k}(k | k) - \hat{x}(k | k)][\hat{x}_{i_k}(k | k) - \hat{x}(k | k)]^T\} \\ &\quad \times p(H_{i_k} | Z_k) \end{aligned} \tag{30}$$

where

$$\begin{aligned} \hat{x}_{i_k}(k | k) &= E[x(k) | H_{i_k}, Z_k], \\ P_{i_k}(k | k) &= E\{[x(k) - \hat{x}_{i_k}(k | k)][x(k) - \hat{x}_{i_k}(k | k)]^T | H_{i_k}, Z_k\} \end{aligned}$$

and $p(H_{i_k} | Z_k)$ are as follows:

$$\hat{x}_{i_k}(k | k) = \sum_{i_{k-n}=1}^{s} \cdots \sum_{i_{k-1}=1}^{s} \hat{x}_{i_{k-n},\ldots,i_k}(k | k)p(H_{i_{k-n}}, \ldots, H_{i_{k-1}} | H_{i_k}, Z_k), \tag{31}$$

$$\begin{aligned} P_{i_k}(k | k) &= \sum_{i_{k-n}=1}^{s} \cdots \sum_{i_{k-1}=1}^{s} \{P_{i_{k-n},\ldots,i_k}(k | k) + [\hat{x}_{i_{k-n},\ldots,i_k}(k | k) - \hat{x}_{i_k}(k | k)] \\ &\quad \times [\hat{x}_{i_{k-n},\ldots,i_k}(k | k) - \hat{x}_{i_k}(k | k)]^T\}p(H_{i_{k-n}}, \ldots, H_{i_{k-1}} | H_{i_k}, Z_k) \end{aligned} \tag{32}$$

$$p(H_{i_k} \mid Z_k) = \frac{\Sigma_{i_{k-n}=1}^{s} \cdots \Sigma_{i_{k-1}=1}^{s} [\Pi_{j=0}^{n-1} F(i, j, k, n)] p(H_{i_{k-n}} \mid Z_{k-n})}{\Sigma_{i_{k-n}=1}^{s} \cdots \Sigma_{i_{k-1}=1}^{s} \Sigma_{i_k=1}^{s} [\Pi_{j=0}^{n-1} F(i, j, k, n)] p(H_{i_{k-n}} \mid Z_{k-n})}, \tag{33}$$

where

$$F(i, j, k, n) \triangleq f(z(k-j) \mid H_{i_{k-n}}, \ldots, H_{i_{k-j}}, Z_{k-1-j}) p(H_{i_{k-j}} \mid H_{i_{k-1-j}}). \tag{34}$$

Here, the probability $p(H_{i_{k-n}}, \ldots, H_{i_{k-1}} \mid H_{i_k}, Z_k)$ can be given by

$$p(H_{i_{k-n}}, \ldots, H_{i_{k-1}} \mid H_{i_k}, Z_k) = \frac{[\Pi_{j=0}^{n-1} F(i, j, k, n)] p(H_{i_{k-n}} \mid Z_{k-n})}{\Sigma_{i_{k-n}=1}^{s} \cdots \Sigma_{i_{k-1}=1}^{s} [\Pi_{j=0}^{n-1} F(i, j, k, n)] p(H_{i_{k-n}} \mid Z_{k-n})}.$$

Note that $\hat{x}_{i_k}(k \mid k)$, $P_{i_k}(k \mid k)$ and $P(H_{i_k} \mid Z_k)$ must be stored as the starting conditions for the $k + n$ stage.

Structure detection. Equation (20) can now be directly provided by (33). Therefore, we need no additional computations to choose the normal or fault state corresponding to the largest value of (20).

Figure 2 shows the time evolution of elemental Kalman filters for the proposed method with $s = n = 2$. Note that $\hat{x}_{i_{k-1}}(k-1 \mid k-1)$, $P_{i_{k-1}}(k-1 \mid k-1)$ and $p(H_{i_k} \mid Z_{k-1})$ had already been stored as the starting conditions for the $k + 1$ stage. Thus, the present algorithm requires the storage of $s \times n$ estimates (n_1-dimension), $s \times n$ covariances ($n_1(n_1 + 1)/2$-dimension), $s \times n$ *a-posteriori* probabilities (scalar) and $n - 1$ observations (m_1-dimension). It is then interesting to note that the present method requires algebraically growing storage with s and n, whereas the one-step measurement update method described in Section 4.2 requires exponentially growing storage. Table I gives their storage requirements;

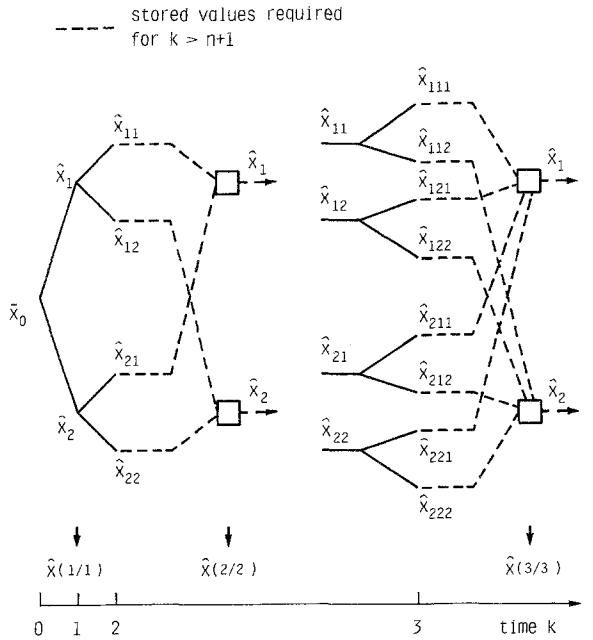
$$T_1 = s^n[n_1 + n_1(n_1 + 1)/2 + 1] \tag{36}$$

for the one-step measurement update method and

$$T_n = sn[n_1 + n_1(n_1 + 1)/2 + 1] + (n-1)m_1 \tag{37}$$

for the n-step measurement update method, where $n_1 = 10$ and $m_1 = 5$. It is seen from Table I that for storage requirements a remarkable difference between the two methods appears as s and n become large. Note, however, that the computational speed of the present algorithm is slightly slower than that of the one-step measurement update algorithm. This is, for example, confirmed by comparing the numbers of elemental Kalman filters required for one-step and n-step measurement updates, i.e.,

$$E_1 = s^{n+1} \quad \text{and} \quad E_n = \sum_{i=1}^{n} s^{i+1} \tag{38}$$



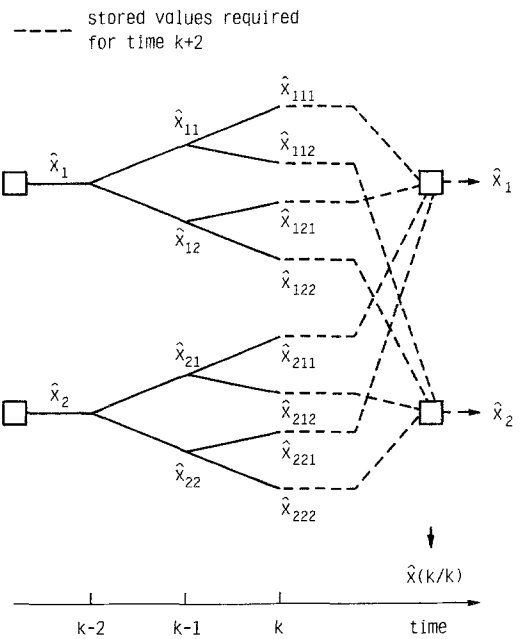


Fig. 2. Time evolution of elemental Kalman filters in n-step measurement update method; (a) for $k \leqslant n + 1$, (b) for $k > n + 1$, where $s = n = 2$.

Table I. Storage requirements for GPBAs, where the upper and lower denote the one- and n-step measurement update methods, respectively

	$s = 2$	$s = 3$	$s = 4$	$s = 5$
$n = 1$	132	198	264	330
	132	198	264	330
$n = 2$	264	594	1056	1650
	269	401	533	665
$n = 3$	528	1782	4224	8250
	406	604	802	1000
$n = 4$	1056	5346	16896	41250
	543	807	1071	1335
$n = 5$	2112	16038	67584	206250
	680	1010	1340	1670

Table II. Numbers of elemental Kalman filters for GPBAs, where the upper and lower denote the one- and n-step measurement update methods, respectively

	$s = 2$	$s = 3$	$s = 4$	$s = 5$
$n = 1$	4	9	16	25
	4	9	16	25
$n = 2$	8	27	64	125
	12	36	80	150
$n = 3$	16	81	256	625
	28	117	336	775
$n = 4$	32	243	1024	3125
	60	360	1360	3900
$n = 5$	64	729	4096	15625
	124	1089	5456	19525

It can be seen from Table II that a difference between two methods disappears as s becomes large for any n. From these facts, the proposed n-step measurement update method is effective for reducing the storage requirements of systems with a relatively large number of Markov chain states.

Note that for $n = 1$ both methods are completely identical in the implementation form; that is, in fact, the algorithm of Chang and Athans (1978).

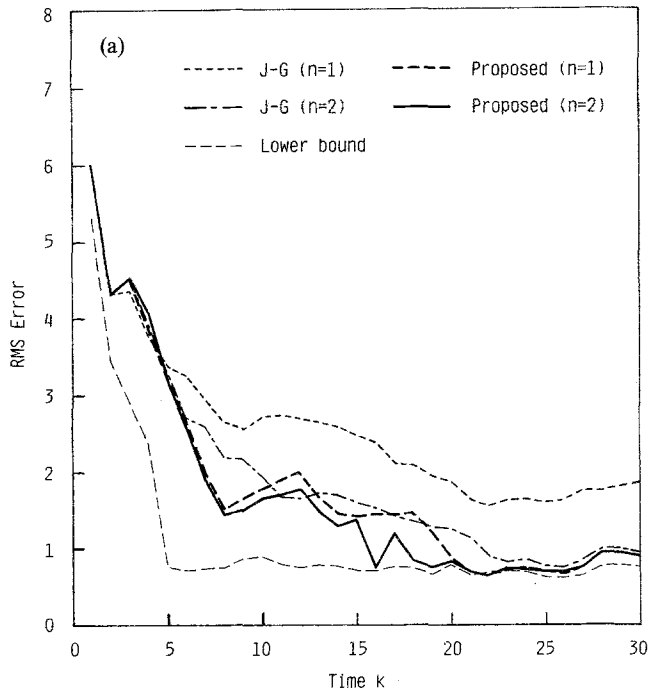
5. Simulation Examples

In this section, two simulation examples are presented to compare the estimation and detection performances of four suboptimal algorithms:

(a) Ackerson and Fu (A-F) method (Ackerson and Fu (1970)),
(b) Jaffer and Gupta (J-G) method (Jaffer and Gupta (1971)),
(c) modified J-G method using (23), (24),
(d) n-step measurement update method given by (27)–(35).

EXAMPLE 1. Consider a scalar dynamical system described by the following equations:

$$x(k + 1) = 1.04x(k) + w(k),$$



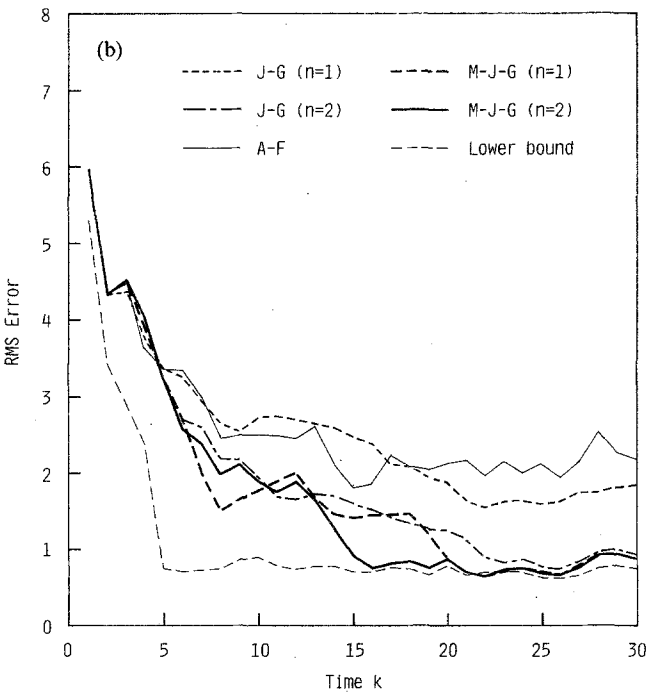


Fig. 3. Comparison of rms errors in state estimation due to suboptimal algorithms: (a) J-G and n-step measurement update methods; (b) A-F, J-G and modified J-G methods, for Example 1.

$$z(k) = x(k) + D(\alpha(k))v(k), \qquad k = 1, 2, \ldots,$$

$$s = 2, \quad \text{i.e. } \alpha(k) \in \{1, 2\}.$$

The initial conditions are as follows: $x(0) \sim N(30,400)$. The sequences $\{w(k)\}$ and $\{v(k)\}$ are mutually independent zero-mean white Gaussian noises with covariances $Q = 0.1$ and $R = 1.0$, respectively. The process $\alpha(k)$ is modeled by a Markov chain with transition probability matrix:

$$\pi = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}.$$

We take $p_1 = p_2 = 0.5$. Finally, we have $D(1) = 10$, $D(2) = 1$.

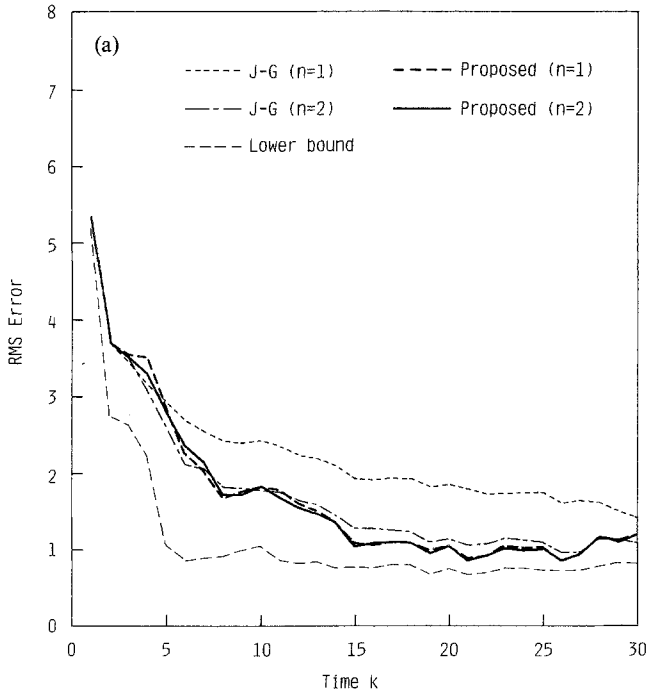
A similar example has been considered in Akashi and Kumamoto (1977), Tugnait and Haddad (1979), Tugnait (1982) and Mathews and Tugnait (1983). The system models a failure mode due to the abrupt change of measurement noise.

The four suboptimal algorithms were simulated and their state estimation performances are compared in Figure 3, rms errors in state estimation are compared for various n. The performances were evaluated by averaging over 50 Monte Carlo runs. A lower bound is obtained by a Kalman filter which knows the true values of the switching parameters. Note also that the modified J-G method with $n = 1$ and the n-step measurement update method with $n = 1$ are the same as the algorithm of Chang and Athans (1978). The information in Figure 3 is also summarized in Table III after averaging over 30 time stages, together with the average probability of error in structure detection.

From Figure 3 and Table III, it is observed that the performances of the modified J-G and the n-step measurement update methods are better than those of other

Table III. RMS state estimation error and probability of error in structure detection for Example 1

Algorithm	Average RMS error in state estimation	Average probability of error in structure detection
A-F method	2.629	0.3133
J-G method with		
$n = 1$	2.548	0.2133
$n = 2$	1.926	0.1973
$n = 3$	1.715	0.1900
Modified J-G method with		
$n = 1$	1.803	0.1887
$n = 2$	1.736	0.1867
$n = 3$	1.712	0.1893
n-step measurement update method with		
$n = 1$	1.803	0.1887
$n = 2$	1.694	0.1793
$n = 3$	1.672	0.1820
Lower bound	1.102	*



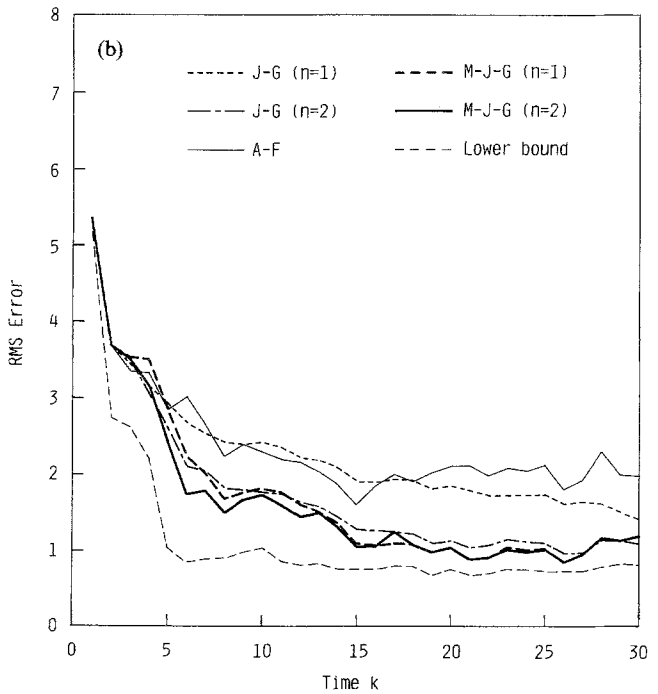


Fig. 4. Comparison of rms errors in state estimation due to suboptimal algorithms: (a) J-G and n-step measurement update methods; (b) A-F, J-G amd modified J-G methods, for Example 2.

methods in both estimation and detection problems; in particular the n-step measurement update method has the best estimation and detection performances. Note also that an increasing in n leads to improvement in estimation performance, but does not necessary lead to improvement in detection performance.

EXAMPLE 2. Consider a relatively high-dimensional Markov chain, i.e. the system of Example 1 except that $p_1 = 0.25$ for $i = 1, \ldots, 4$,

$$D(1) = 10, \quad D(2) = 7, \quad D(3) = 4, \quad D(4) = 1$$

and

$$\pi = \begin{bmatrix} 0.15 & 0.15 & 0.2 & 0.5 \\ 0.2 & 0.2 & 0.3 & 0.3 \\ 0.3 & 0.2 & 0.1 & 0.4 \\ 0.5 & 0.15 & 0.2 & 0.15 \end{bmatrix}.$$

The results of averaging over 50 Monte Carlo runs are shown in Figure 4 and Table IV. Observe that, for this case, the proposed method is comparable in the performances of estimation and detection to the modified J-G method. However, it is significant to note, from calculating (36) and (37) with $n_1 = m_1 = 1$ and $s = 4$, that the present method requires about 50% less storage requirement than the existing GPBA methods for the case of $n = 2$. Also, it is found from Table IV that the actual ratio of computational speed between the J-G or modified J-G method and the n-step measurement update for the case of $n = 2$ is nearly close to an ideal ratio (i.e. 1.25) presented in Table II, as would be expected.

Table IV. RMS state estimation error and probability of error in structure detection for Example 2

Algorithm	Average RMS error in state estimation	Average probability of error in structure detection	Average computation time (sec)
A-F method	2.379	0.5733	1.664
J-G method with			
$n = 1$	2.258	0.5147	3.322
$n = 2$	1.739	0.4973	10.442
$n = 3$	1.727	0.4867	39.480
Modified J-G method with			
$n = 1$	1.704	0.4860	3.650
$n = 2$	1.633	0.4840	11.722
$n = 3$	1.676	0.4827	44.422
n-step measurement update method with			
$n = 1$	1.704	0.4860	3.770
$n = 2$	1.693	0.4853	12.886
$n = 3$	1.688	0.4853	49.100
Lower bound	1.127	*	0.646

6. Conclusions

The problem of state estimation and system structure detection has been considered for linear discrete-time systems with abruptly changing parameters. The changes were modeled by a Markov chain with known transition statistics. Since the optimal solution was impractical, a suboptimal approach based on the 'n-step measurement update method' and a modified J-G approach were proposed and their efficiencies were demonstrated through several simulation studies. It has been shown that a new GPBA requires arithmetically growing storage with s and n, whereas the modified or original J-G algorithm based on the 'one-step measurement update method' requires exponentially growing storage. Furthermore, the simulation results indicate that, in general the present GPBA performs better than the J-G algorithm. Note, however, that since the n-step measurement update method is essentially a type of reprocessing the observations data, the proposed algorithm necessitates a slightly extra time of computation than the J-G algorithm.

No discussion was provided in this paper, nor was any referenced, to establish: (i) the feasibility of knowing the transition probabilities for practical problems, (ii) the robustness of the Markov approach to errors in the transition probability assignments, and (iii) the feasibility of estimating the transition probabilities. We may be very interested in enlightenment concerning these issues. For a reference of (iii), see, for example, Tugnait (1982b). In addition, for some references to control related work on the same class of models, for example, see Blair and Sworder (1975), Griffiths and Loparo (1985), Chizeck *et al.* (1986) and Birdwell *et al.* (1986).

References

Ackerson, G.A. and Fu, K.S., (1970), On the state estimation in switching environments, *IEEE Trans. Automat. Control* **AC-15**(1), 10–17.

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

Birdwell, J.D., Castanon, D., and Athans, M., (1986), On reliable control system design, *IEEE Trans. Systems Man Cybernet.* **SMC-16**(5), 703–711.

Blair, W.P. and Sworder, D.D., (1975), Feedback control of a class of linear discrete systems with jump parameters and quadratic cost criteria, *Internat. J. Control* **21**(5), 833–841.

Blom, H.A.P., (1984), An efficient filter for abruptly changing systems, *Proc. 23rd IEEE Conf. on Decision and Control*, pp. 656–658.

Blom, H.A.P., (1985), An efficient decision-making-free filter for processes with abrupt changes, *Proc. IFAC Identification and System Parameter Estimation 1985*, pp. 631–636.

Blom, H.A.P. and Bar-Shalom, Y., (1988), The interacting multiple model algorithm for systems with Markovian switching coefficients, *IEEE Trans. Automat. Control* **AC-33**(8), 780–783.

Chang, C.B. and Athans, M., (1978), State Estimation for Discrete Systems with Switching Parameters, *IEEE Trans. Aerospace Electron. Systems* **AES-14**(3), 418–425.

Chang, C.B. and Tabaczynski, J.A., (1984), Application of state estimation to target tracking, *IEEE Trans. Automat. Control*, **AC-29**(2) 98–109.

Chizeck, H.J., Willsky, A.S., and Castanon, D., (1986), Discrete-time Markovian-jump linear quadratic optimal control, *Internat. J. Control* **43**(1), 213–231.

Gustafson, D.E., Willsky, A.S., Wang, J.Y., Lancaster, M.C., and Triebwasser, J.H., (1978), ECG/VCG rhythm diagnosis using statistical signal analysis, I. Identification of persistent rhythms, *IEEE Trans. Biomed. Engng.* **BME-25**(4), 344–353.

Griffiths, B.E. and Loparo, K.A., (1985), Optimal control of jump-linear gaussian systems, *Internat. J. Control* **42**(4), 791–819.

Jaffer, A.G. and Gupta, S.C., (1971), On estimation of discrete processes under multiplicative and noise conditions, *Inform. Sci.* **3**, 267–276.

Lainiotis, D.G., (1971), Optimal Adaptive Estimation: Structure and Parameter Adaptation, *IEEE Trans. Automat. Control* **AC-16**, 160–170.

Lainiotis, D.G., (1976), Partitioning: A unifying framework for adaptive systems, I: Estimation. *Proc. IEEE* **64**, 1126–1142.

Mathews, V.J. and Tugnait, J.K., (1988), Detection and estimation with fixed lag for abruptly changing systems, *IEEE Trans. Aerospace Electron. Systems* **AES-19**(5), 730–739.

Moose, R.L., VanLandingham, H.F., and Zwicke, P.E., (1978), Digital set point control of nonlinear stochastic systems, *IEEE Trans. Indust. Electron. Contr. Instrum.* **IECI-25**(1), 39–45.

Moose, R.L., VanLandingham, H.F., and McCabe, D.H., (1979), Modeling and estimation for tracking maneuvering targets, *IEEE Trans. Aerospace Electron. Systems* **AES-15**(3), 448–456.

Ricker, G.G. and Williams, J.R., (1978), Adaptive tracking filter for maneuvering targets, *IEEE Trans Aerospace Electron. Systems* **AES-14**(1), 185–193.

Sugimoto, S. and Ishizuka, I., (1983), Identification and estimation algorithms for a Markov chain plus AR process, *Proc. IEEE Internat. Conf. Acoustics, Speech and Signal Processing, ICASSP 83*, pp. 247–250.

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

Tugnait, J.K., (1982a), Detection and estimation for abruptly changing systems, *Automatica*, **18**(5), 607–615.

Tugnait, J.K., (1982b), Adaptive estimation and identification for discrete systems with Markov jump parameters, *IEEE Trans. Automat. Control* **AC-27**(5), 1054–1065.

Watanabe, K., (1989a), A multiple model adaptive filtering approach to fault diagnosis in stochastic systems, in R.J. Patton, P.M. Frank and R.N. Clark (eds.), *Fault Diagnosis in Dynamic Systems* Prentice-Hall, London, pp. 411–438.

Watanabe, K., (1989b), A decentralized multiple model adaptive filtering for discrete-time stochastic systems, *ASME Trans. J. Dynamic Systems Measure. Control* **111**, 371–377.

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

Willsky, A.S., (1980), Failure detection in dynamic systems, *AGARD* No. 109.