

# New Gaussian Mixture State Estimation Schemes For Discrete Time Hybrid Gauss-Markov Systems

R. J. Elliott, F. Dufour and W. P. Malcolm

**Abstract**—In this article we compute state and mode estimation algorithms for discrete-time Gauss-Markov models whose parameter-sets switch according to a known Markov law. Our algorithms are distinct from extant methods, such as the so called Interacting Multiple Model algorithm (IMM) and sequential Monte Carlo methods, in that they are based on *exact* hybrid filter dynamics. The fundamental difficulty in estimation of jump Markov systems, is managing the geometrically growing history of candidate hypotheses. In our scheme, we address this issue by proposing an extension of an idea due to Viterbi. Our scheme maintains a fixed number of candidate paths in a history, each identified by an optimal subset of estimated mode probabilities. We compute a finite dimensional sub-optimal filter, which estimates the hidden state process and the mode probability. A computer simulation is provided.

**Index Terms**—Reference Probability, Jump Markov Systems, Hybrid Dynamics, Viterbi Algorithm, Filtering

## I. INTRODUCTION

In this article the reference probability method is used to compute state estimates for a discrete-time hybrid dynamical system. Using a general result, (see the reproducing Gaussian density Lemma in [6]), we propose a new suboptimal algorithm which provides an exact hypothesis management scheme, circumventing geometric growth in algorithmic complexity. Gaussian mixture densities are used to compute filters whose memory requirements remain fixed in time. In a related context, (see [16]), Gaussian mixtures have been applied to compute a truncated maximum likelihood method for parameters estimation. However, our contributions are to compute a new recursive state and mode estimation filter whose memory requirements remain fixed in time. A fundamental feature of our approach, is that state estimators are developed from the corresponding *exact* filter, which appears in [5]. Our approach is in contrast to ad-hoc schemes such as the IMM (due to Blom [3]) and sequential Monte Carlo methods, which are not based upon exact filter dynamics. We restrict our attention in this article to filtering only, however, a general form of the corresponding smoother is developed in [10].

Robert J. Elliott is with the Haskayne School of Business, Scurfie old Hall, University of Calgary, 2500 University Drive NW, Calgary, Alberta, Canada T2N 1N4. Email: relliott@ucalgary.ca

F. Dufour is with MAB (Mathematiques Appliquees de Bordeaux) Universite Bordeaux) 351, Cours de la Liberation, 33405 Talence Cedex, Email: dufour@math.u-bordeaux.fr

W. P. Malcolm with the National ICT Australia (NICTA), Systems Engineering and Complex Systems, Research School of Information Sciences and Engineering, The Australian National University Canberra ACT 0200 Australia, Email: paul.malcolm@nicta.com.au

## II. HYBRID STOCHASTIC DYNAMICS AND REFERENCE PROBABILITY

### A. Modulating Markov Chain Dynamics

To model parameter switching we consider a time homogenous discrete time discrete state Markov chain  $Z$ . We identify the state space of  $Z$  with an orthonormal basis indicator functions, which we denote by  $\mathcal{L}$ , that is,

$$\mathcal{L} = \{e_1, e_2, \dots, e_m\} = \left\{ \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \right\} \subseteq \mathbb{R}^m. \quad (1)$$

The dynamics for the process  $Z$  may be written as

$$Z_k = \Pi Z_{k-1} + L_k \in \mathbb{R}^m. \quad (2)$$

Here,  $\Pi = [\pi_{(j,i)}]_{\substack{1 \leq j \leq m \\ 1 \leq i \leq m}}$  is the transition matrix of  $Z$ , with elements

$$\pi_{(j,i)} \triangleq P(Z_k = e_j \mid Z_{k-1} = e_i) \quad (3)$$

for all  $k \in \mathbb{N}$ .

The process  $L$  is a sequence of  $(P, \sigma\{Z\})$ -martingale increments, and we suppose  $E[Z_0] = p_0$ .

### B. Hidden State Process Dynamics

We suppose the indirectly observed state vector  $x \in \mathbb{R}^n$ , has dynamics

$$x_k = \sum_{j=1}^m \langle Z_k, e_j \rangle A_j x_{k-1} + \sum_{j=1}^m \langle Z_k, e_j \rangle B_j w_k. \quad (4)$$

Here  $w$  is a vector-valued Gaussian process with  $w \sim N(0, I_n)$ .  $A_j$  and  $B_j$  are  $n \times n$  matrices and for each  $j \in \{1, 2, \dots, m\}$ , the  $B_j$  are nonsingular. This condition can be relaxed, (see [11]).

### C. System Observation Process Dynamics

Consider a vector-valued observation process with values in  $\mathbb{R}^d$  and dynamics

$$y_k = \sum_{j=1}^m \langle Z_k, e_j \rangle C_j x_k + \sum_{j=1}^m \langle Z_k, e_j \rangle D_j v_k. \quad (5)$$

Here  $v$  is a vector-valued Gaussian process with  $v \sim N(0, I_d)$ . We suppose the matrices  $D_j \in \mathbb{R}^{d \times d}$ , for each  $j \in \{1, 2, \dots, m\}$ , are nonsingular. The systems we shall

consider in this article are described by the dynamics (2), (4) and (5). We define our sigma-algebras as follows:

$$\mathcal{F}_k = \sigma\{x_\ell, 0 \leq \ell \leq k\}, \quad (6)$$

$$\mathcal{Z}_k = \sigma\{Z_\ell, 0 \leq \ell \leq k\}, \quad (7)$$

$$\mathcal{Y}_k = \sigma\{y_\ell, 0 \leq \ell \leq k\}, \quad (8)$$

$$\mathcal{G}_k = \sigma\{Z_\ell, x_\ell, y_\ell, 0 \leq \ell \leq k\}. \quad (9)$$

#### D. Reference Probability Details

The dynamics given at (2), (4) and (5), are each defined on a measurable space  $(\Omega, \mathcal{F})$ , under a measure  $P$ . However, consider a new measure  $P^\dagger$ , under which the dynamics for the processes  $Z$ ,  $x$  and  $y$ , are, respectively

$$P^\dagger \begin{cases} Z_k &= \Pi Z_{k-1} + L_k, \\ x_k &\text{are iid and } N(0, \mathbf{I}_n), \\ y_k &\text{are iid and } N(0, \mathbf{I}_d). \end{cases} \quad (10)$$

#### Notation:

Suppose  $\xi \in \mathbb{R}^d$ , is a random vector with Gaussian distribution  $N(0, \mathbf{I}_n)$ . The symbol  $\Phi(\cdot)$  will be used to denote the zero mean normal density on  $\mathbb{R}^n$ :

$$\Phi(\xi) = (2\pi)^{-d/2} \exp(-\frac{1}{2}\xi'\xi). \quad (11)$$

Similarly, if  $\kappa \in \mathbb{R}^n$ , is a random vector with Gaussian distribution  $N(0, \mathbf{I}_d)$ , the symbol  $\Psi(\cdot)$  will be used to denote the zero mean normal density on  $\mathbb{R}^d$ :

$$\Psi(\kappa) = (2\pi)^{-n/2} \exp(-\frac{1}{2}\kappa'\kappa). \quad (12)$$

To compute the filter dynamics we now define the measure  $P$  by setting the restriction of its Radon-Nikodym derivative to  $\mathcal{G}_k$  to

$$\Lambda_{0,k} \triangleq \frac{dP}{dP^\dagger} |_{\mathcal{G}_k} = \prod_{\ell=0}^k \lambda_\ell,$$

where

$$\lambda_1 = \sum_{j=1}^m \langle Z_\ell, e_j \rangle \frac{\Phi(D_j^{-1}(y_\ell - C_j x_\ell))}{|D_j| \Phi(y_\ell)} \times \frac{\Psi(B_j^{-1}(x_\ell - A_j x_{\ell-1}))}{|B_j| \Psi(x_\ell)}. \quad (13)$$

Here and throughout,  $|A|$  denotes the determinant of a matrix  $A$ . The existence of  $P$  follows from the Kolmogorov extension Theorem, (see Shiryaev, Theorem 4, p. 166 [15]). We quote the following form of Bayes' Theorem (see [8]).

**Theorem 1** Suppose  $\gamma = \{\gamma_\ell, 0 \leq \ell \leq k\}$  is an integrable  $\mathcal{G}$ -adapted process. then

$$E[\gamma_k | \mathcal{Y}_k] = \frac{E^\dagger[\Lambda_{0,k} \gamma_k | \mathcal{Y}_k]}{E^\dagger[\Lambda_{0,k} | \mathcal{Y}_k]}. \quad (14)$$

### III. HYBRID FILTER DYNAMICS

#### A. Exact Filter Dynamics

To compute a filter jointly estimating the density of the state vector  $x$ , and the state of the chain  $Z$ , consider the expectation:

$$E[\langle Z_k, e_j \rangle] = \frac{E^\dagger[\Lambda_{0,k} \langle Z_k, e_j \rangle f(x_k) | \mathcal{Y}_k]}{E^\dagger[\Lambda_{0,k} | \mathcal{Y}_k]}. \quad (15)$$

Here the function  $f(\cdot)$  is an arbitrary bounded measurable real-valued test function. Write the numerator as:

$$\sigma(\langle Z_k, e_j \rangle f(x_k)) \triangleq E^\dagger[\Lambda_{0,k} \langle Z_k, e_j \rangle f(x_k) | \mathcal{Y}_k]. \quad (16)$$

Suppose we choose the test function  $f(x) = 1$ , then

$$\begin{aligned} \sum_{j=1}^m \sigma(\langle Z_k, e_j \rangle) &= \sigma\left(\sum_{j=1}^m \langle Z_k, e_j \rangle\right) \\ &= \sigma(1) \\ &= E^\dagger[\Lambda_{0,k} | \mathcal{Y}_k]. \end{aligned} \quad (17)$$

Therefore, if the numerator (16) can be evaluated for any such  $f$  and all  $j$ , the denominator of (15) can be found. For each  $j \in \{1, 2, \dots, m\}$ , the quantity  $\sigma(\langle Z_k, e_j \rangle f(x_k))$ , defined at (16), is a continuous linear functional on the space of continuous functions and so defines a unique measure (see Rudin, Theorem 2.14 p. 40, [13]). Further, suppose that there exists a corresponding unique, un-normalised density function  $q_k^j(x)$ , such that

$$\sigma(\langle Z_k, e_j \rangle f(x_k)) \triangleq \int_{\mathbb{R}^n} f(\eta) q_k^j(\eta) d\eta. \quad (18)$$

Then the normalised conditional density is:

$$P(x \in dx, Z_k = e_j | \mathcal{Y}_k) = \frac{q_k^j(x) dx}{\int_{\mathbb{R}^n} q_k^j(\xi) d\xi}. \quad (19)$$

**Theorem 2 (Elliott, Dufour, Sworder, 1996)** The un-normalised probability density  $q_k^j(x)$ , satisfies the following integral-equation recursion,

$$q_k^j(x) = \frac{\Phi(D_j^{-1}(y_k - C_j x))}{\Phi(y_k) |D_j| |B_j|} \times \sum_{r=1}^m \pi_{(j,r)} \int_{\mathbb{R}^n} \Psi(B_j^{-1}(x - A_j \xi)) q_{k-1}^r(\xi) d\xi. \quad (20)$$

The recursion given at (20) is an exact filter; it is expressed as a density which is in general infinite dimensional. However, Gaussian densities are determined by their mean and variance. Consider a finite Gaussian mixture representation of  $q_{k-1}^r(x)$ . We can then compute the integrals in (20) by using the reproducing density Lemma [6].

**Theorem 3** Suppose the un-normalised probability density  $q_{k-1}^r(\xi)$ , (as it appears under the integral in equation (20)), can be written as a finite weighted Gaussian mixture with

$M^q \in \mathbb{N}$  components. That is, for  $k \in \{1, 2, \dots\}$ , we suppose

$$q_{k-1}^r(\xi) = \sum_{s=1}^{M^q} p_{k-1}^{r,s} \frac{1}{(2\pi)^{n/2} |\Sigma_{k-1|k-1}^{r,s}|^{\frac{1}{2}}} \times \exp\left\{-\frac{1}{2}(\xi - \alpha_{k-1|k-1}^{r,s})' \text{inv}(\Sigma_{k-1|k-1}^{r,s})(\xi - \alpha_{k-1|k-1}^{r,s})\right\}. \quad (21)$$

Here  $\Sigma_{k-1|k-1}^{r,s} \in \mathbb{M}^{n \times n}$ , and  $\alpha_{k-1|k-1}^{r,s} \in \mathbb{R}^n$ , are both  $\mathcal{Y}_{k-1}$ -measurable functions for all pairs  $(r, s) \in \{1, 2, \dots, m\} \times \{1, 2, \dots, M^q\}$ . Using this Gaussian mixture (21), the recursion for the optimal un-normalised density process has the form

$$q_k^j(x) \triangleq \frac{1}{(2\pi)^{(d+n)/2} \Phi(y_k)} \sum_{r=1}^m \sum_{s=1}^{M^q} K_{k,k-1}^q(j, r, s) \times \exp\left\{-\frac{1}{2}\left(x - \text{inv}(\sigma_{k-1}^{j,r,s})\delta_{k,k-1}^{j,r,s}\right)' \sigma_{k-1}^{j,r,s} \left(x - \text{inv}(\sigma_{k-1}^{j,r,s})\delta_{k,k-1}^{j,r,s}\right)\right\}.$$

Here

$$\bar{\Sigma}_{k-1|k-1}^{j,r,s} \triangleq B_j B_j' + A_j \Sigma_{k-1|k-1}^{r,s} A_j' \in \mathbb{R}^{n \times n}, \quad (22)$$

$$\tilde{u}_{k-1|k-1}^{j,r,s} \triangleq A_j \alpha_{k-1|k-1}^{r,s} \in \mathbb{R}^n \quad (23)$$

$$\sigma_{k-1}^{j,r,s} \triangleq C_r' \text{inv}(D_r D_r') C_r + \text{inv}(\bar{\Sigma}_{k-1|k-1}^{j,r,s}) \quad (24)$$

$$\delta_{k,k-1}^{j,r,s} \triangleq \text{inv}(\bar{\Sigma}_{k-1|k-1}^{j,r,s}) \tilde{u}_{k-1|k-1}^{j,r,s} + C_r' \text{inv}(D_r D_r') y_k \quad (25)$$

$$K_{k,k-1}^q(j, r, s) \triangleq \frac{\pi(j, r) p_{k-1}^{r,s}}{|\bar{\Sigma}_{k-1|k-1}^{j,r,s}|^{\frac{1}{2}} |D_j|} \times \exp\left\{\frac{1}{2}(\delta_{k,k-1}^{j,r,s})' \text{inv}(\sigma_{k-1}^{j,r,s}) \delta_{k,k-1}^{j,r,s}\right\} \times \exp\left\{-\frac{1}{2}\left[y_k' \text{inv}(D_r D_r') y_k + (\tilde{u}_{k-1|k-1}^{j,r,s})' \text{inv}(\bar{\Sigma}_{k-1|k-1}^{j,r,s}) \tilde{u}_{k-1|k-1}^{j,r,s}\right]\right\} \quad (26)$$

To prove Theorem 3, we apply the reproducing Gaussian density Lemma [6] to the recursion at (20). In the first application of this Lemma, we eliminate the integral in equation (20). Using the finite-mixture representation (21), we then write the recursion (20) for the function  $q^j(x)$  as a weighted Gaussian mixture.

**Corollary 1** *The estimated conditional mode probability, for model  $j$  is given by*

$$p_k^j \triangleq P(Z_k = e_j | \mathcal{Y}_k) = \frac{\int_{\mathbb{R}^n} q_k^j(\xi) d\xi}{\sum_{\ell=1}^m \int_{\mathbb{R}^n} q_k^\ell(\xi) d\xi} = \frac{q_k^j}{\sum_{\ell=1}^m q_k^\ell}. \quad (27)$$

Here the estimated un-normalised mode probability  $q_k^j$ , is computed by the double summation

$$q_k^j = \frac{1}{(2\pi)^{d/2} \Phi(y_k)} \sum_{r=1}^m \sum_{s=1}^{M^q} \zeta_{k,k-1}^q(j, r, s), \quad (28)$$

where

$$\zeta_{k,k-1}^q(j, r, s) \triangleq K_{k,k-1}^q(j, r, s) |\sigma_{k-1}^{j,r,s}|^{-\frac{1}{2}} \quad (29)$$

## B. Fixed Memory Filter Dynamics

In this section we develop a sub-optimal recursive filter by extending an idea due to Viterbi [17]. The motivation to develop a sub-optimal filter is immediate from the dynamics of the un-normalised density. Suppose at time  $k = 1$ , these dynamics involve  $m \times M^q$  densities. Then the next time,  $k = 2$ , these dynamics require  $m \times m \times M^q$  densities. It is clear, that the demand on memory requirements is exponential in time, with the number of densities required at time  $k$  being  $m^k \times M^q$ . What we wish to do is to circumvent this growth by identifying a subset of candidate mixture densities, from which we construct a sub-optimal density with fixed, (in time), memory requirements.

1) *Hypothesis Management:* Write

$$\Gamma^q \triangleq \{1, 2, \dots, m\} \times \{1, 2, \dots, M^q\}, \quad (30)$$

$$\tilde{S}_{k,k-1}^q(j, r, s) \triangleq \{\zeta_{k,k-1}^q(j, r, s)\}_{(r,s) \in \Gamma^q}. \quad (31)$$

To remove the growth in memory requirements, we propose to identify, at each time  $k$ , the  $M^q$ -best candidate densities for each suboptimal density  $q_k^j(x)$ , using the corresponding estimated mode probabilities. The key to this idea is to identify  $M^q$  optimal densities, that is, the  $M^q$  components in the Gaussian mixture, through their corresponding set of estimated mode probabilities,  $q^j$ ,  $j \in \{1, 2, \dots, m\}$ .

Since the estimated mode probabilities, given by equation (28), are formed by a summation over non-negative quantities, we can identify the  $M^q$  largest contributors to this sum and then use the corresponding indices to identify the  $M^q$ -best Gaussian densities. This maximisation procedure is as follows:

$$\zeta_{k,k-1}^q(j, r_{k,1}^*, s_{k,1}^*) \triangleq \max_{(r,s) \in \Gamma^q} \{\tilde{S}_{k,k-1}^q(j, r, s)\}, \quad (32)$$

⋮

$$\zeta_{k,k-1}^q(j, r_{k,M^q}^*, s_{k,M^q}^*) \triangleq \max_{(r,s) \in \Gamma^q \setminus \{(r_{k,1}^*, s_{k,1}^*), \dots\}} \{\tilde{S}_{k,k-1}^q(j, r, s)\} \quad (33)$$

$$\left\{ \zeta_{k,k-1}^q(j, r_{k,1}^*, s_{k,1}^*), \dots, \zeta_{k,k-1}^q(j, r_{k,M^q-1}^*, s_{k,M^q-1}^*) \right\}. \quad (34)$$

$$\zeta_{k,k-1}^q(j, r_{k,M^q}^*, s_{k,M^q}^*) \left. \right\}. \quad (35)$$

Note that we are not directly interested in the quantities  $\zeta_{k,k-1}^q(j, r_{k,\ell}^*, s_{k,\ell}^*)$ , rather, the indices that locate these quantities. The optimal index set, for the density of model  $j$ , is:

$$\mathcal{I}_k^q(j, M^q) \triangleq \{(j, r_{k,1}^*, s_{k,1}^*), (j, r_{k,2}^*, s_{k,2}^*), \dots, (j, r_{k,M^q}^*, s_{k,M^q}^*)\}. \quad (36)$$

Using these indices, we approximate the suboptimal un-normalised density, of order  $M^q$ , corresponding to the density  $q_k^j(x)$  as

$$\begin{aligned} \hat{q}_k^j(x) \triangleq & \frac{1}{(2\pi)^{(d+n)/2} \Phi(y_k)} \sum_{\ell=1}^{M^q} K_{k,k-1}^q(j, r_{k,\ell}^*, s_{k,\ell}^*) \times \\ & \exp\left\{-\frac{1}{2}\left(x - \text{inv}(\sigma_{k-1}^{j, r_{k,\ell}^*, s_{k,\ell}^*}) \delta_{k,k-1}^{j, r_{k,\ell}^*, s_{k,\ell}^*}\right)'\right. \\ & \left. \sigma_{k-1}^{j, r_{k,\ell}^*, s_{k,\ell}^*} \left(x - \text{inv}(\sigma_{k-1}^{j, r_{k,\ell}^*, s_{k,\ell}^*}) \delta_{k,k-1}^{j, r_{k,\ell}^*, s_{k,\ell}^*}\right)\right\} \end{aligned} \quad (37)$$

Further, by Corollary 1, the un-normalised mode probability corresponding to the expectation  $E[Z_k = e_j \mid \mathcal{Y}_k]$ , is approximated by

$$\hat{q}_k^j \triangleq \frac{\sum_{\ell=1}^{M^q} \zeta_{k,k-1}^q(j, r_{k,\ell}^*, s_{k,\ell}^*)}{(2\pi)^{d/2} \Phi(y_k)}. \quad (38)$$

To normalise the function  $q_k^j(x)$ , we need the sum of all terms  $\{q_k^1, \dots, q_k^m\}$  and so write

$$\varphi_{k|k} \triangleq \sum_{j=1}^m \hat{q}_k^j. \quad (39)$$

2) *Filter State Statistics:* Densities such as  $q_k^j(x)$ , provide all the information available. However, what is often required in state estimation (filtering), is an expression for the state estimate of  $x_k$  at time  $k$  given  $\mathcal{Y}_k$ , that is:

$$\hat{x}_{k|k} \triangleq E[x_k \mid \mathcal{Y}_k]. \quad (40)$$

The expectation at (40) is exact, that is, it is computed under the measure corresponding to the exact filter, which is not suitable for implementation due to its exponential growth in complexity as a function of the discrete-time index. To circumvent this difficulty, we propose a suboptimal estimator

$$\hat{x}_{k|k} \triangleq E^{\hat{P}_{M^q}^*} [x_k \mid \mathcal{Y}_k]. \quad (41)$$

Here, the measure  $\hat{P}_{M^q}^*$  corresponds to the normalised form of the approximated Gaussian mixture give by equation (37).

**Proposition 1** *The sub-optimal state estimate  $\hat{x}_{k|k}$ , defined by the expectation given at (41), has the representation*

$$\begin{aligned} \hat{x}_{k|k} = & \frac{1}{\sum_{j=1}^m \sum_{\ell=1}^{M^q} \frac{K_{k,k-1}^q(j, r_{k,\ell}^*, s_{k,\ell}^*)}{|\sigma_{k-1}^{j, r_{k,\ell}^*, s_{k,\ell}^*}|^{\frac{1}{2}}}} \times \\ & \sum_{j=1}^m \sum_{\ell=1}^{M^q} \left\{ \frac{K_{k,k-1}^q(j, r_{k,\ell}^*, s_{k,\ell}^*)}{|\sigma_{k-1}^{j, r_{k,\ell}^*, s_{k,\ell}^*}|^{\frac{1}{2}}} \times \right. \\ & \left. \text{inv}\{\sigma_{k-1}^{j, r_{k,\ell}^*, s_{k,\ell}^*}\} \delta_{k,k-1}^{j, r_{k,\ell}^*, s_{k,\ell}^*} \right\}. \end{aligned} \quad (42)$$

**Remark III.1** A corresponding estimator for the state error covariance associated to (42) can also be calculated, (see [10]). This particular state error covariance has the form of the addition of two state error covariances; one standard error covariance arising from the basic state estimation task and one state error covariance related to the order of the Gaussian mixture.

#### IV. EXAMPLE

In the example described here, we consider a scalar-valued Gauss-Marvov system and a three-state Markov chain. Our three dynamical systems are each defined by the model parameter-sets:

$$\{A_1, B_1, C_1, D_1\} \triangleq \{1, 0, 2, 1\} \quad (43)$$

$$\{A_2, B_2, C_2, D_2\} \triangleq \{0.8, 1, 2, 1\} \quad (44)$$

$$\{A_3, B_3, C_3, D_3\} \triangleq \{-0.8, 1, 2, 1\}. \quad (45)$$

Our transition matrix and initial density for the modulating Markov chain, are, respectively,

$$\Pi \triangleq \begin{bmatrix} 0.9 & 0.05 & 0.05 \\ 0.05 & 0.9 & 0.05 \\ 0.05 & 0.05 & 0.9 \end{bmatrix} \quad (46)$$

and

$$p_0 \triangleq [1/3 \quad 1/3 \quad 1/3]. \quad (47)$$

The fixed realisation of the Markov chain used throughout is shown in the first subfigure of Figure 1. In the second subfigure of Figure 1, we show the corresponding hidden state process  $x$ . The regions where this particular hidden state process remain constant, correspond to the Markov chain being in state  $e_1$ , that is, the model  $\{A_1, B_1, C_1, D_1\}$  is in effect. This model holds the previous state value until the Markov chain jumps to either  $e_2$  or  $e_3$ . Using the Markov chain and state process shown in Figure 1, twenty realisations of the observation process were generated with independent additive noise processes  $v$ . Figure 2 shows ensemble averages of the estimated sample paths for the hidden state process. In the first subfigure of Figure 2, we show the ensemble average of the exact filter, that is, the Kalman filter supplied with the exact realisation of the model parameters generated by the Markov chain. In the second subfigure of Figure 2, we show the ensemble

average generated by the new Gaussian mixture estimator defined by equation (42). Here the Gaussian mixture was of order  $M^q = 3$ . Again in Figure 2, in the final subfigure, we show the absolute value of the difference computed from the previous two subfigures. In Figure 3, we show the true Markov chain and the estimated mode probabilities generated by the Gaussian-mixture scheme. The simulation

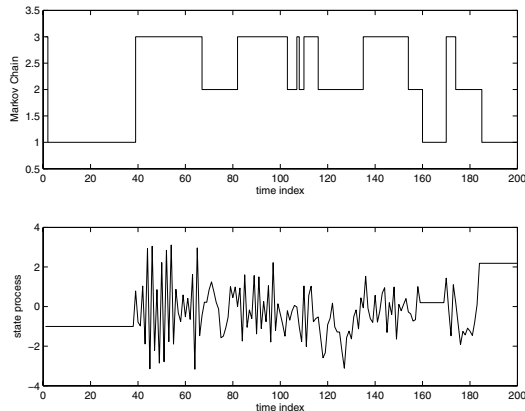


Fig. 1. Realisations for the modulating Markov chain and the hidden state process  $x$ .

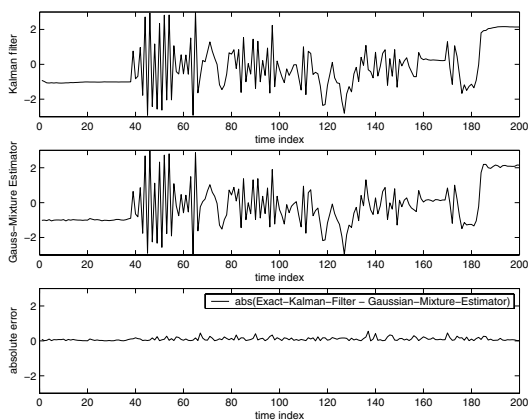


Fig. 2. Comparison of the state estimator performance between Kalman filter (supplied with the exact model parameter set) and the new Gaussian mixture estimator with  $M^q = 3$ .

study given here demonstrates excellent performance for the Gaussian mixture scheme, using only three components in the weighted mixtures.

**Remark IV.1** It is worth noting that sequential Monte Carlo schemes, (particle filters), will fail in the simulation scenario considered in this article. This is due to the presence of a constant state process model. Further, in object/target tracking, where stochastic hybrid systems such as the collection of dynamics at (2), (4) and (5) arise quite naturally, a constant scenario could correspond to a maneuvering object that stops, such as a helicopter or a jump jet. The Gaussian-mixture scheme presented here

is unaffected by the inclusion of a constant state process model.

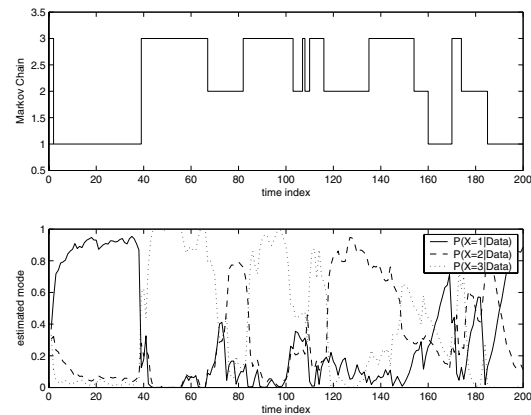


Fig. 3. Estimated mode probability computed by the Gaussian mixture scheme.

## V. ACKNOWLEDGEMENT

R. J. Elliott, F. Dufour and W. P. Malcolm, would like to acknowledge the support of the National Sciences and Engineering Research Council of Canada through operating grant number 69-0672.

W. P. Malcolm would like to acknowledge the support of the National ICT Australia. The National ICT Australia is funded by the Australian Government's Department of Communications, Information Technology and the Arts and the Australian Research Council through Backing Australia's Ability and the ICT Centre of Excellence program.

## REFERENCES

- [1] L. Aggoun and R. J. Elliott, *Measure Theory and Filtering*, Cambridge University Press, 2004.
- [2] G. Ackerson and K. Fu, On state estimation in switching environments, *IEEE Transactions on Automatic Control* 15(1), pp. 10-17, 1970.
- [3] H. Blom, An efficient filter for abruptly changing systems, *23rd IEEE Conference on Decision and Control*, Las Vegas USA, November 1984.
- [4] J. Cloutier, J. Evers and J. Freely, Assessment of air-to-air missile guidance and control technology, *IEEE Control systems magazine*, 9, pp. 27-34, 1989
- [5] R. J. Elliott, F. Dufour and D. D. Sworner, Exact Hybrid Filters in Discrete Time, *IEEE Transactions on Automatic Control*, 41, 1996, pp. 1807-1810.
- [6] R. J. Elliott and W. P. Malcolm, Reproducing Gaussian Densities and Linear Gaussian Detection, *Systems and Control Letters*, 40 (2000), pp. 133-138.
- [7] R. J. Elliott, *Stochastic Calculus and its Applications*, Springer Verlag 1982.
- [8] R. J. Elliott, L. Aggoun and J. B. Moore, *Hidden Markov Models Estimation and Control*, Springer Verlag Applications of Mathematics Series 29, 1995.
- [9] Elliott R. J., A general recursive discrete time filter, *Journal of Applied Probability*, 30, pp. 575-588, 1993.
- [10] R. J. Elliott, F. Dufour and W. P. Malcolm, State and Mode Estimation for Discrete-Time Jump Markov Systems, *SIAM Journal of Optimization and Control*, to appear.

- [11] R. J. Elliott and V. Krishnamurthy, New finite-dimensional filters for parameter estimation of discrete-time linear Gaussian models, IEEE Transactions on Automatic Control, Volume 44, Number 5, May, 1999.
- [12] M. Mariton, *Jump linear systems in automatic control*, Marcel Dekker, New York, 1990.
- [13] W. Rudin, *Real and Complex Analysis*, McGraw-Hill, 1966.
- [14] H. W. Sorenson and Alspach, D. L., Recursive Bayesian Estimation Using Gaussian Sums, Automatica, Volume 7, Number 4, July 1971, pp. 465-479.
- [15] A. N. Shiryaev, *Probability*, Second Edition, Springer Verlag 1996.
- [16] J. K. Tugnait, Adaptive Estimation and Identification for Discrete Systems with Markov Jump Parameters, IEEE Transactions on Automatic Control, Volume AC-27, Number 5, October 1982.
- [17] A. J. Viterbi, Error Bounds for the White Gaussian and Other Very Noisy Memoryless Channels with Generalised Decision Regions, IEEE Transactions on Information Theory, Volume IT, Number 2, March 1969.