[8] A. J. Laub, "Algebraic aspects of generalized eigenvalue problems for solving Riccati equations," in *Com. and Comb. Methods in System Theory*, Byrnes and Lindguist (Eds.). North Holland: Elsevier, 1986, pp. 213–227.
[9] P. Lancaster and L. Rodman, "Solutions of the continuous and discrete time algebraic Riccati equations: A review," in *The Riccati Equation*, S. Bittanti, A. J. Laub and J. C. Willems (Eds.). New York: Springer-Verlag, 1991.
[10] W.-W. Lin, "A new method for computing the closed-loop eigenvalues of a discrete-time algebraic Riccati equation," *Lin. Alg. Appl.*, vol. 96, pp. 157–180, 1987.
[11] V. L. Mehrmann, *The Autonomous Linear Quadratic Control Problem: Theory and Numerical Solution* (Lecture Notes in Control and Information Sciences, no. 163). Berlin: Springer-Verlag, 1991.
[12] T. Pappas, A. J. Laub and N. R. Sandell, Jr., "On the numerical solution of the discrete-time algebraic Riccati equation," *IEEE Trans. Automat. Contr.*, vol. AC-25, no. 4, pp. 631–641, 1980.
[13] G. W. Stewart, *Introduction to Matrix Computations*. New York: Academic, 1973.
[14] C. F. Van Loan, "A symplectic method for approximating all the eigenvalues of a Hamiltonian matrix," *Lin. Alg. Appl.*, vol. 61, pp. 233–251, 1984.
[15] H. J. Wilkinson, *The Algebraic Eigenvalue Problem*. Oxford: Clarendo Press, 1965, p. 87.

Linear Minimum Mean Square Error Estimation for Discrete-Time Markovian Jump Linear Systems

O. L. V. Costa

***Abstract*—The linear minimum mean square error estimator (LMMSE) for discrete-time linear systems subject to abrupt changes in the parameters modeled by a Markov chain $\theta(k) \in \{1, \cdots, N\}$ is considered. The filter equations are derived from geometric arguments in a recursive form, resulting in an on-line algorithm suitable for computer implementation. Our approach is based on estimating $x(k)1_{\{\theta(k)=i\}}$ instead of estimating directly $x(k)$. The uncertainty introduced by the Markovian jumps increases the dimension of the filter to $N(n+1)$, where n is the dimension of the state variable. An example where the dimension of the filter can be reduced to n is presented, as well as a numerical comparison with the IMM filter [2].**

I. Introduction

Several dynamic systems are inherently vulnerable to abrupt changes in their structures due, for instance, to component and/or interconnections failures, abrupt environment changes, etc., and have been considered nowadays by several authors (see, for instance, [4], [5], [10], [11], [14]). In this paper we shall be examining the linear minimum mean square error estimator (LMMSE) for the class of discrete-time linear systems which are subject to abrupt changes in their parameters. We assume that the abrupt changes can be modeled by a finite state-space Markov chain $\theta(k) \in \{1, \cdots, N\}$. In particular, the problem of optimal and suboptimal state estimation for this class of systems has been addressed in [1]–[3], [7], [12], [13], under the hypothesis of Gaussian distribution for the disturbances influencing the system. Suboptimal state estimation had to be considered to limit the computational requirements, since the optimal algorithm requires exponentially increasing memory and computations with time. These suboptimal algorithms, however, lack good computational performance at modest computational load in many situations. Besides, they heavily rely on the Gaussian assumption for the noise.

Manuscript received December 22, 1992; revised April 12, 1993 and October 8, 1993. This work was supported in part by the National Research Council-CNP$_q$, Brazil under Grant 305173/88-0/EE/FV and by FAPESP (Research Council of the State of São Paulo), under Grant 91/0508-3.

The author is with the Departamento de Engenharia Eletrônica, Escola Politécnica da Universidade de São Paulo, CEP: 05508 900-São Paulo-SP-BRAZIL.

IEEE Log Number 9402705.

The problem of obtaining optimal linear recursive least-square state estimators for linear systems with uncertain observation has been addressed in [8] and [9]. The uncertainty enters in the observation equation as

$$y(k) = \beta(k)Hx(k) + G\nu(k) \tag{1}$$

where $\beta(k)$ is a binary switching sequence. Conditions are established in [8] which leads to a recursive filter for $x(k)$ of the form

$$\hat{x}(k|k) = F_1(k)\hat{x}(k-1|k-1) + F_2(k)y(k). \tag{2}$$

In particular for $\{\beta(k)\}$ a general Markov chain, it was shown in [8] that a recursive structure of the form in (2) does not in general exists.

In this paper we shall obtain the optimal linear state estimation for the class of models mentioned above. The novelty in our approach, in comparison with the one in [8] and [9], is that we do not make an estimation for $x(k)$ directly but instead, we estimate the random vector $x(k)1_{\{\theta(k)=i\}}$, where $1_{\{\cdot\}}$ stands for the Dirac measure. By doing this we obtain a recursive filter of dimension Nn (where n stands for the dimension of $x(k)$). If a known input sequence $\{u(k)\}$ is also applied to the system, then an estimate of $1_{\{\theta(k)=i\}}$ is also required, increasing the dimension of the filter to $N(n+1)$. The advantage of this formulation is that it can be applied to a broader class of systems than the linear models with output uncertainty given by (1) studied in [8] and [9], and the filter equations obtained are suitable for computer implementation via an on-line recursive algorithm (all the heavy calculations can be made before we start collecting the data, since the gain matrices of the filter are not data-dependent; note, however, that they depend on the deterministic input sequence $\{u(k)\}$, which is assumed to be known *a priori*).

This paper is organized as follows. In Section II we present the problem formulation, some assumptions, notation, and auxiliary results. In Section III the recursive filter equations are derived using the geometric arguments presented in Section II. Two example are presented in Section IV. Example 1 shows that, when the uncertainty affects only the output of the system and the plant noise through a sequence of independent and identically distributed binary sequence, the dimension of the filter can be reduced to n, in agreement with [8]. In Example 2 some numerical comparisons of the LMMSE filter with the IMM algorithm [2] are presented. These comparisons indicate that the LMMSE can be a good alternative to situations where on-line calculations are time-critical and precomputed gain matrices are required. The paper is concluded in Section V with some final comments.

II. Preliminaries

Fix an underlying probability space $(\Omega, \mathcal{F}, \mathcal{P})$ and denote by $1_{\{\cdot\}}$ the Dirac measure (that is, for any $\mathcal{U} \in \mathcal{F}$ and $\omega \in \Omega$, $1_{\{\mathcal{U}\}}(\omega) = 1$ if $\omega \in \mathcal{U}$, 0 otherwise). Consider the following discrete-time Markovian jump linear system

$$\begin{aligned} x(k+1) &= A_{\theta(k)}x(k) + B_{\theta(k)}u(k) + C_{\theta(k)}\xi(k) \\ y(k) &= H_{\theta(k)}x(k) + G_{\theta(k)}\nu(k), \quad k = 0, 1, 2 \cdots. \end{aligned} \tag{3}$$

Here $\{x(k)\}$ denotes the $\mathbb{R}^n$-valued state sequence, $\{\xi(k)\}$ and $\{\nu(k)\}$ are random disturbances in $\mathbb{R}^{q_1}$ and $\mathbb{R}^{q_2}$ respectively, $\{u(k)\}$ is a known input sequence in $\mathbb{R}^p$, $\{y(k)\}$ is the $\mathbb{R}^m$-valued output sequence, $\{\theta(k)\}$ is a discrete-time Markov chain with finite state space $\{1, \cdots, N\}$, and transition probability matrix $\mathbb{P} = [p_{ij}]$ and A_i, B_i, C_i, H_i, G_i, $i = 1, \cdots, N$ are matrices of appropriate dimension. We shall make the following assumptions:

A1) $G_i G_i' > 0$ (where $'$ stands for transpose) for all $i = 1, \cdots, N$.

A2) $\{\xi(k)\}$ and $\{\nu(k)\}$ are null mean second-order, independent wide sense stationary sequences mutually independent with covariance matrices equal to the identity.

A3) $x(0)1_{\{\theta(0)=i\}}$, $i = 1, \cdots, N$ are second order random vectors with $E(x(0)1_{\{\theta(0)=i\}}) = \mu_i$ and $E(x(0)x(0)'1_{\{\theta(0)=i\}}) = V_i$, $i = 1, \cdots, N$.

A4) $x(0)$ and $\{\theta(k)\}$ are independent of $\{\xi(k)\}$ and $\{\nu(k)\}$.

Remark 1: All matrices $\mathbb{P}$ and A_i, B_i, C_i, H_i, G_i could be functions of k and the argument has been dropped here for simplicity in notation.

It is clear from the assumptions above that $\{x(k)\}$ and $\{y(k)\}$ are sequences of second-order random vectors, and we shall denote by $\mathcal{F}_k$ the σ-field generated by the random vectors and variables $\{x(t), y(t), \theta(t);\ t = 0, \cdots, k\}$. For any sequence of second-order random vectors $\{r(k)\}$ we shall define the "centered" random vector $r^c(k)$ as $r^c(k) := r(k) - E(r(k))$, $\hat{r}(k|\ell)$ as the best affine estimator of $r(k)$ given $y(0), \cdots, y(\ell)$, $\ell \le k$, $\tilde{r}(k|\ell)$ as $\tilde{r}(k|\ell) := r(k) - \hat{r}(k|\ell)$, $\hat{r}^c(k|\ell)$ as the best linear estimator of $r^c(k)$ given $y^c(0), \cdots, y^c(\ell)$ and $\tilde{r}^c(k|\ell)$ as $\tilde{r}^c(k|\ell) := r^c(k) - \hat{r}^c(k|\ell)$. It is well known (cf, [6, pp. 109]) that

$$\hat{r}(k|\ell) = \hat{r}^c(k|\ell) + E(r(k)) \tag{4}$$

and, in particular, $\tilde{r}^c(k|\ell) = \tilde{r}(k|\ell)$. We shall denote by $\mathcal{L}((y^c)^k)$ the linear subspace spanned by $(y^c)^k := (y^c(k)', \cdots, y^c(0)')'$, that is, a random variable $r \in \mathcal{L}((y^c)^k)$ if $r = \sum_{i=0}^k \alpha_i' y^c(i)$ for some $\alpha_i \in \mathbb{R}^m$, $i = 0, \cdots, k$. For $r, s \in \mathcal{L}((y^c)^k)$ the inner product $\langle \cdot; \cdot \rangle$ in $\mathcal{L}((y^c)^k)$ is given by $\langle s; r \rangle := E(sr)$ and therefore s and r are orthogonal if $\langle s; r \rangle = 0$. The best linear estimator $\hat{r}^c(k|\ell) = (\hat{r}_1^c(k|\ell), \cdots, \hat{r}_\iota^c(k|\ell))'$ of the random vector $r^c(k) = (r_1^c(k), \cdots, r_\iota^c(k))'$ in $\mathbb{R}^\iota$ is the projection of $r^c(k)$ onto the subspace $\mathcal{L}((y^c)^{\mathcal{L}})$ and satisfies the following properties (cf [6, pp. 108, 113]):

a) $\hat{r}_j^c(k|\ell) \in \mathcal{L}((y^c)^\ell)$, $j = 1, \cdots, \iota$ (5)

b) $\tilde{r}_j(k|\ell)$ is orthogonal to $\mathcal{L}((y^c)^\ell)$, $j = 1, \cdots, \iota$ (6)

c) if $\operatorname{cov}((y^c)^\ell)$ is nonsingular then

$$\hat{r}^c(k|\ell) = E(r^c(k)((y^c)^\ell)')\operatorname{cov}((y^c)^\ell)^{-1}(y^c)^\ell \tag{7}$$

$$\begin{aligned}\hat{r}^c(k|k) = \hat{r}^c(k|k-1) &+ E(r^c(k)\tilde{y}(k|k-1)')E(\tilde{y}(k|k-1) \\ &\cdot \tilde{y}(k|k-1)')^{-1}(y^c(k) - \hat{y}^c(k|k-1)).\end{aligned} \tag{8}$$

We shall consider the following augmented state-space equations

$$\begin{aligned}\overline{x}(k+1) &= \overline{A}_{\theta(k)}\overline{x}(k) + \overline{B}_{\theta(k)}u(k) + \overline{C}_{\theta(k)}\xi(k),\\ \overline{x}(0) &= (1, x(0)')' \in \mathbb{R}^{(n+1)}\\ y(k) &= \overline{H}_{\theta(x)}\overline{x}(k) + G_{\theta(k)}\nu(k), \quad k = 0, 1, 2, \cdots\end{aligned} \tag{9}$$

where for $i = 1, \cdots, N$

$$\overline{A}_i := \begin{bmatrix} 1 & O_n' \\ O_n & A_i \end{bmatrix} \in \mathbb{R}^{(n+1)\times(n+1)}, \quad \overline{B}_i := \begin{bmatrix} O_p' \\ B_i \end{bmatrix} \in \mathbb{R}^{(n+1)\times p},$$

$$\overline{C}_i := \begin{bmatrix} O_{q_1}' \\ C_i \end{bmatrix} \in \mathbb{R}^{(n+1)\times q_1},$$

$$\overline{H}_i := [O_m \quad H_i] \in \mathbb{R}^{m\times(n+1)}$$

with O_j standing for the null vector in $\mathbb{R}^j$. It is easy to verify that $\overline{x}(k) = (1, x(k)')' \in \mathbb{R}^{n+1}$ for all $k \ge 0$. For $k \ge 0$ and $i \in \{1, \cdots, N\}$ we define $z(k, i) := \overline{x}(k)1_{\{\theta(k)=i\}} = (1_{\{\theta(k)=i\}}, x(k)'1_{\{\theta(k)=i\}})' \in \mathbb{R}^{(n+1)}$, $z(k) := (z(k, 1)', \cdots, z(k, N)')' \in \mathbb{R}^{N(n+1)}$. The reason for considering this augmented state space and making these definitions is that we shall obtain a recursive filter of dimension $N(n+1)$ in terms of $\hat{z}(k|k)$ and not in terms of $\hat{x}(k|k)$ as in (2). From the identity

$$\begin{aligned}\overline{B}_{\theta(k)}u(k) &= \sum_{i=1}^N \overline{B}_i u(k)1_{\{\theta(k)=i\}}\\ &= [\overline{B}_1 u(k)e_1' \cdots \overline{B}_N u(k)e_1']z(k)\end{aligned} \tag{10}$$

where e_1 stands for the vector in $\mathbb{R}^{n+1}$ formed by one in the first element and zero elsewhere, we can see that the control sequence could be regarded as part of state matrix for the new state sequence $\{z(k)\}$, and this is why an estimation of $1_{\{\theta(k)=i\}}$, $i = 1, \cdots, N$, is also required. Once we calculate $\hat{z}(k|k)$ it is straightforward to obtain $\hat{x}(k|k)$ since that $(1, x(k)')' = \sum_{i=1}^N z(k, i)$ and thus $(1, \hat{x}(k|k)')' = \sum_{i=1}^N \hat{z}(k, i|k)$. Note that although the estimator for $1_{\{\theta(k)=i\}}$ could be any real number, their sum over i from one to N must be equal to one. For the case with no input sequence ($u(k) = 0$, $k \ge 0$) the estimation for $1_{\{\theta(k)=i\}}$ will no longer be required and the order of the filter can be reduced to Nn (see Remark 5 below).

We define now the following matrices associated to the second moment of the above variables.

$$Z(k) := E(z(k)z(k)') \in \mathbb{R}^{N(n+1)\times N(n+1)}$$

$$Z_i(k) := E(z(k, i)z(k, i)') \in \mathbb{R}^{(n+1)\times(n+1)}, i = 1, \cdots, N$$

$$\hat{Z}(k|\ell) := E(\hat{z}(k|\ell)\hat{z}(k|\ell)') \in \mathbb{R}^{N(n+1)\times N(n+1)}, \ell \le k$$

$$\tilde{Z}(k|\ell) := E(\tilde{z}(k|\ell)\tilde{z}(k|\ell)') \in \mathbb{R}^{N(n+1)\times N(n+1)}, \ell \le k.$$

Note that $Z(k) = \operatorname{diag}(Z_i(k))$ (the square matrix formed by $Z_i(k)$, $i = 1, \cdots, N$ in the diagonal and zero elsewhere). Indeed, $E(z(k, i)z(k, j)') = E(\overline{x}(k)\overline{x}(k)'1_{\{\theta(k)=i\}}1_{\{\theta(k)=j\}}) = 0$ if $i \ne j$. Finally the recursive filter equations, to be deduced in the next section, will be written in terms of the matrices

$$A := \begin{bmatrix} p_{11}\overline{A}_1 & \cdots & p_{N1}\overline{A}_N \\ \vdots & \dots & \vdots \\ p_{1N}\overline{A}_1 & \cdots & p_{NN}\overline{A}_N \end{bmatrix},$$

$$B(k) := \begin{bmatrix} p_{11}\overline{B}_1 u(k)e_1' & \cdots & p_{N1}\overline{B}_N u(k)e_1' \\ \vdots & \dots & \vdots \\ p_{1N}\overline{B}_1 u(k)e_1' & \cdots & p_{NN}\overline{B}_N u(k)e_1' \end{bmatrix}$$

$$\mathcal{A}(k) := A + B(k) \in \mathbb{R}^{N(n+1)\times N(n+1)}$$

$$G(k) := [G_1\pi_1(k)^{1/2} \cdots G_N\pi_N(k)^{1/2}] \in \mathbb{R}^{m\times Nq_2},$$

$$H := [\overline{H}_1 \quad \cdots \quad \overline{H}_N] \in \mathbb{R}^{m\times N(n+1)}$$

where $\pi_i(k) := \mathcal{P}(\theta(k) = i)$ and recall that e_1 stands for the vector in $\mathbb{R}^{n+1}$ formed by one in the first component, zero elsewhere.

III. Linear Least Square Estimator

Let M be the $n \times (n+1)$ matrix formed by zero in the first column and the $n \times n$ identity. We have the following theorem.

Theorem 1: For the system (3) with assumptions A1)–A4) the estimator $\hat{x}(k|k)$ is given by

$$\hat{x}(k|k) = \sum_{i=1}^{N} M\hat{z}(k, i|k) \tag{11}$$

where $\hat{z}(k|k)$ satisfies the recursive equation

$$\hat{z}(k|k) = \hat{z}(k|k-1) + \tilde{Z}(k|k-1)H'(H\tilde{Z}(k|k-1)H' + G(k)G(k)')^{-1}(y(k) - H\hat{z}(k|k-1)), \tag{12}$$

$$\hat{z}(k|k-1) = \mathcal{A}(k-1)\hat{z}(k-1|k-1), k \geq 1, \tag{13.a}$$

$$\hat{z}(0|-1) = \zeta(0) \text{ where } \zeta(k) = E(z(k)),$$
$$\zeta(0) = (\zeta_1', \cdots, \zeta_N')', \; \zeta_i = (\pi_i(0), \mu_i')'. \tag{13.b}$$

Moreover

$$\tilde{Z}(k|k-1) = Z(k) - \hat{Z}(k|k-1), \; \tilde{Z}(k|k) = Z(k) - \hat{Z}(k|k) \tag{14}$$

where $Z(k) = \operatorname{diag}(Z_j(k))$ with

$$Z_j(k+1) = \sum_{i=1}^{N} p_{ij}(\overline{A}_i + \overline{B}_i u(k)e_1')Z_i(k)(\overline{A}_i + \overline{B}_i u(k)e_1')' + \sum_{i=1}^{N} p_{ij}\pi_i(k)\overline{C}_i\overline{C}_i', \tag{15.a}$$

$$Z_j(0) = \begin{bmatrix} \pi_j(0) & \mu_j' \\ \mu_j & V_j \end{bmatrix}, \quad j \in \{1, \cdots, N\} \tag{15.b}$$

and

$$\hat{Z}(k|k) = \hat{Z}(k|k-1) + \tilde{Z}(k|k-1)H'(H\tilde{Z}(k|k-1)H' + G(k)G(k)')^{-1}H\tilde{Z}(k|k-1), \tag{16}$$

$$\hat{Z}(k|k-1) = \mathcal{A}(k-1)\hat{Z}(k-1|k-1)\mathcal{A}(k-1)', \tag{17.a}$$

$$\hat{Z}(0|-1) = \zeta(0)\zeta(0)'. \tag{17.b}$$

Proof: Since

$$y(k) = Hz(k) + G_{\theta(k)}\nu(k), \quad E(y(k)) = H\zeta(k) \tag{18}$$

we get that

$$y^c(k) = Hz^c(k) + G_{\theta(k)}\nu(k). \tag{19}$$

From independence of $\nu(k)$ and $\{\theta(k), (y^c)^{k-1}\}$ it follows that $G_{\theta(k)}\nu(k)$ is orthogonal to $\mathcal{L}((y^c)^{k-1})$. Indeed, $\langle \gamma' G_{\theta(k)}\nu(k); \alpha'(y^c)^{k-1}\rangle = E(\nu(k)')E(G_{\theta(k)}'\gamma\alpha'(y^c)^{k-1}) = 0$ for any $\gamma \in \mathbb{R}^m$, $\alpha \in \mathbb{R}^{km}$; similar reasoning shows orthogonality between $\tilde{z}(k|k-1)$ and $G_{\theta(k)}\nu(k)$. Recalling from (5) and (6) that $\hat{z}^c(k|k-1) \in \mathcal{L}((y^c)^{k-1})$ and $\tilde{z}(k|k-1)$ is orthogonal to $\mathcal{L}((y^c)^{k-1})$, we obtain from (18) and (19) that (note that $G(k)G(k)' > 0$ from A1)

$$\hat{y}^c(k|k-1) = H\hat{z}^c(k|k-1), \tag{20}$$

$$\tilde{y}(k|k-1) = H\tilde{z}(k|k-1) + G_{\theta(k)}\nu(k), \tag{21}$$

$$E(\tilde{y}(k|k-1)\tilde{y}(k|k-1)') = H\tilde{Z}(k|k-1)H' + G(k)G(k)' > 0 \tag{22}$$

$$E(z^c(k)\tilde{y}(k|k-1)') = E((\tilde{z}(k|k-1) + \hat{z}^c(k|k-1))\tilde{y}(k|k-1)') = \tilde{Z}(k|k-1)H'. \tag{23}$$

We have from (7), (9), and (10) that

$$\begin{aligned}\hat{z}^c(k, j|k-1) &= E(E(z(k, j)((y^c)^{k-1})'|\mathcal{F}_{k-1})) \cdot (\operatorname{cov}((y^c)^{k-1}))^{-1}(y^c)^{k-1} \\ &= E\left(\sum_{i=1}^{N} E((\overline{A}_i + \overline{B}_i u(k-1)e_1')z(k-1, i)) \cdot 1_{\{\theta(k-1)=i\}}1_{\{\theta(k)=j\}}((y^c)^{k-1})'|\mathcal{F}_{k-1})\right) \\ &\quad \cdot (\operatorname{cov}((y^c)^{k-1}))^{-1}(y^c)^{k-1} \\ &= \sum_{i=1}^{N} p_{ij}(\overline{A}_i + \overline{B}_i u(k-1)e_1')\hat{z}^c(k-1, i|k-1)\end{aligned} \tag{24}$$

where the above inverse follows from assumption A1), and the second equality follows from the fact that $\overline{C}_{\theta(k-1)}\xi(k-1)$ and $\mathcal{L}((y^c)^{k-1})$ are orthogonal. From (4), (8), (20)–(24) we obtain (12) and (13), after noting that $\zeta(k+1) = \mathcal{A}(k)\zeta(k)$. Equation (14) easily follows from $\tilde{z}(k|\ell) = z^c(k) - \hat{z}^c(k|\ell)$ and the fact that $\tilde{z}(k|\ell)$ is orthogonal to $\mathcal{L}((y^c)^\ell)$ (see (6)) and therefore (see (5)) orthogonal to $\hat{z}^c(k|\ell)$. Equation (15) is readily derived from (9). Equation (16) follows from (12), after noting that $y(k) - H\hat{z}(k|k-1) = H\tilde{z}(k|k-1) + G_{\theta(k)}\nu(k)$ and recalling that $\hat{z}^c(k|k-1)$, $\tilde{z}(k|k-1)$, and $G_{\theta(k)}\nu(k)$ are all orthogonal among themselves. Finally (17) follows immediately from (13) and (11) from the identity $(1, x(k)')' = \sum_{i=1}^{N} z(k, i)$.

Remark 2: Note that to implement the above equations the sequences $\tilde{Z}(k|k-1)$ and $\tilde{Z}(k|k)$ can be computed off-line using (14)–(17), since they do not depend on the observations $y(k)$ (however, they depend on the deterministic input sequence $u(k)$ which is assumed to be known *a priori*). Calculation of $\hat{z}(k|k)$ can now be performed recursively using (12) and (13) as successive observations become available.

Remark 3: For the case with no jumps, the above equations reduce to the Kalman filter. Indeed, if A_1, B_1, C_1, H_1, and G_1 denote the parameters of a linear plant as in (3) with no jumps ($N = 1$) and $\tilde{X}(k|\ell)$ the associated covariance error matrix, then it can be easily shown that (12), (13), and (15) become

$$\begin{bmatrix} 1 \\ \hat{x}(k|k) \end{bmatrix} = \begin{bmatrix} 1 \\ \hat{x}(k|k-1) \end{bmatrix} + \begin{bmatrix} O_m' \\ \tilde{X}(k|k-1)H_1' \end{bmatrix} \cdot (H_1\tilde{X}(k|k-1)H_1' + G_1G_1')^{-1} \cdot (y(k) - H_1\hat{x}(k|k-1))$$

$$\begin{bmatrix} 1 \\ \hat{x}(k|k-1) \end{bmatrix} = \begin{bmatrix} 1 & O_n' \\ B_1u(k-1) & A_1 \end{bmatrix}\begin{bmatrix} 1 \\ \hat{x}(k-1|k-1) \end{bmatrix},$$

$$\tilde{Z}(k|\ell) = \begin{bmatrix} O & O_n' \\ O_n & \tilde{X}(k|\ell) \end{bmatrix}.$$

Remark 4: From the above algorithm it is easy to verify that the covariance error $\tilde{X}(k|k) = E(\tilde{x}(k|k)\tilde{x}(k|k)')$ is given by $\tilde{X}(k|k) = \sum_{i=1}^{N}\sum_{j=1}^{N}\tilde{X}_{ij}(k|k)$ where $\tilde{X}_{ij}(k|k) = E(\tilde{\varphi}(k, i|k)\tilde{\varphi}(k, j|k)')$, $\varphi(k, i|k) = x(k)1_{\{\theta(k)=i\}}$, is a submatrix of $\tilde{Z}(k|k)$.

Remark 5: For model (3) with no input sequence ($u(k) = 0$ for all $k \geq 0$), no estimation of $1_{\{\theta(k)=i\}}$ will be required and the dimension of the filter can be reduced to Nn with the equations rewritten in the following way. Redefine $\mathcal{A}$ and H replacing respectively $\overline{A}_i$ and $\overline{H}_i$ by A_i and H_i in the original definition. Thus $\mathcal{A}(k) = \mathcal{A} \in \mathbb{R}^{Nn \times Nn}$ and $H \in \mathbb{R}^{m \times Nn}$. Redefine $z(k, i)$ as $z(k, i) = x(k)1_{\{\theta(k)=i\}}$ so that $z(k) = (z(k, 1)', \cdots, z(k, n)')' \in \mathbb{R}^{Nn}$. In this case the filter equations are as in Theorem 1 above, replacing M, $\overline{A}_i$, and $\overline{C}_i$

TABLE I
SET OF PARAMETERS CONSIDERED IN THE SIMULATIONS

cases	p_{11}	p_{22}	a_1	a_2	c_1	c_2	h_1	h_2	g_1	g_2
01	0.975	0.95	0.995	0.99	1.0	1.0	1.0	1.0	1.0	1.0
02	0.975	0.95	0.995	0.99	0.1	0.1	1.0	1.0	5.0	5.0
03	0.995	0.99	0.995	0.99	0.1	0.1	1.0	1.0	5.0	5.0
04	0.975	0.95	0.995	0.99	8.0	8.0	1.0	1.0	1.0	1.0
05	0.975	0.95	0.995	0.75	1.0	1.0	1.0	1.0	0.6	0.6
06	0.975	0.95	0.995	0.995	2.0	2.0	1.0	0.95	0.5	0.5
07	0.975	0.95	0.995	0.995	1.0	1.0	1.0	0.8	0.2	0.2
08	0.75	0.5	0.995	0.995	0.5	0.5	1.0	0.8	0.8	0.8
09	0.995	0.99	0.995	0.995	0.5	0.5	1.0	0.8	0.8	0.8
10	0.975	0.95	0.995	0.995	0.1	5.0	1.0	1.0	1.0	1.0
11	0.975	0.95	0.995	0.995	1.0	1.0	1.0	1.0	0.1	5.0
12	0.995	0.8	0.950	0.0	1.0	1.0	1.0	1.0	1.0	1.0
13	0.9	0.5	0.95	0.95	0.5	0.5	1.0	1.0	1.0	40.0
14	0.975	0.95	0.995	0.25	1.0	1.0	1.0	1.0	1.0	1.0
15	0.98	0.8	0.95	1.1	1.0	1.0	1.0	1.0	1.0	1.0

by respectively I_n, A_i, and C_i and redefining the initial conditions $\hat{z}(0|-1)$, $Z_j(0)$ and $\hat{Z}(0|-1)$ appropriately.

IV. EXAMPLES

In this section we present two examples to illustrate the above technique. In Example 1 we show that the above equations can be rewritten as in (2) for linear systems with uncertain observations an uncertainty in the plant noise, generalizing some results derived in [8] and [9]. In Example 2 we compare the performances of the LMMSE filter and the IMM algorithm, which was introduced in [2].

Example 1: Consider model (3) with $u(k) = 0$ for all $k \geq 0$, $N = 2$, $A_1 = A_2$, $H_2 = 0$, $\{\theta(k)\}$ an independent and identically distributed random sequence with $p_{ij} = p_j$ for all $i, j = 1, 2$, $x(0)$ and $\theta(0)$ independent, and $\mathcal{P}(\theta(0) = i) = p_i$, $i = 1, 2$. Using the notation in Remark 5 we get, from independence between $\{y^{k-1}, x(k)\}$ and $\theta(k)$, that $\hat{z}(k, i|k-1) = p_i \hat{x}(k|k-1)$, and from (12)

$$\begin{aligned} x(k|k) &= \hat{z}(k, 1|k) + \hat{z}(k, 2|k) \\ &= \hat{x}(k|k-1) + (\tilde{Z}_{11}(k|k-1)H_1')\mathcal{I}(k)^{-1}(y(k) \\ &\quad - H_1 \hat{x}(k|k-1)p_1) \\ &\quad + (\tilde{Z}_{21}(k|k-1)H_1')\mathcal{I}(k)^{-1}(y(k) \\ &\quad - H_1 \hat{x}(k|k-1)p_1) \end{aligned}$$

where $\tilde{Z}_{ij}(k|k-1) = E(\tilde{z}(k, i|k-1)\tilde{z}(k, j|k-1)')$, $i, j = 1, 2$, and $\mathcal{I}(k) = (H\tilde{Z}(k|k-1)H' + G(k)G(k)') = (H_1\tilde{Z}_{11}(k|k-1)H_1' + p_1 G_1 G_1' + p_2 G_2 G_2')$. Defining $F_2(k) = (\tilde{Z}_{11}(k|k-1) + \tilde{Z}_{21}(k|k-1))H_1'\mathcal{I}(k)^{-1}$ and $F_1(k) = (I_n - p_1 F_2(k) H_1)A_1$, we get that

$$\begin{aligned} \hat{x}(k|k) &= \hat{x}(k|k-1) + F_2(k)(y(k) - p_1 H_1 \hat{x}(k|k-1)) \\ &= A_1 \hat{x}(k-1|k-1) + F_2(k)(y(k) \\ &\quad - p_1 H_1 A_1 \hat{x}(k-1|k-1)) \\ &= F_1(k)\hat{x}(k-1|k-1) + F_2(k)y(k). \end{aligned}$$

Straightforward calculations lead to $\tilde{Z}_{ii}(k|k-1) = p_1 p_2 X(k) + p_i^2 \tilde{X}(k|k-1)$ and $\tilde{Z}_{ij}(k|k-1) = p_1 p_2 \tilde{X}(k|k-1) - p_1 p_2 X(k)$, $i \neq j$, where $X(k) = E(x(k)x(k)')$ and $\tilde{X}(k|\ell) = E(\tilde{x}(k|\ell)\tilde{x}(k|\ell)')$.

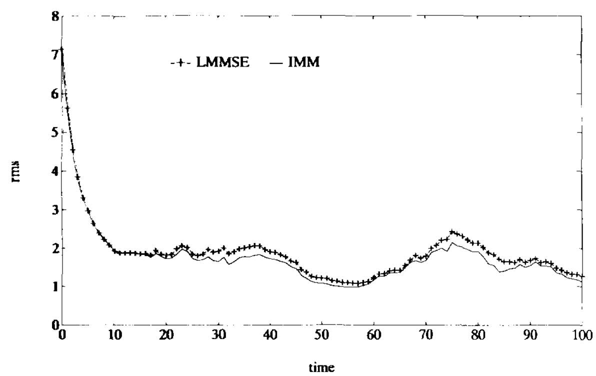

Fig. 1. Case 2, representative of Cases 1, 4, 6, 11.

Thus $\tilde{Z}_{11}(k|k-1) + \tilde{Z}_{21}(k|k-1) = p_1 \tilde{X}(k|k-1)$ and

$$\begin{aligned} F_2(k) = (p_1 \tilde{X}(k|k-1)H_1')(H_1(p_1 p_2 X(k) \\ + p_1^2 \tilde{X}(k|k-1))H_1' + p_1 G_1 G_1' + p_2 G_2 G_2')^{-1}. \end{aligned}$$

Moreover it is easy to verify that $X(k+1) = A_1 X(k) A_1' + p_1 C_1 C_1' + p_2 C_2 C_2'$, $\tilde{X}(k+1|k) = A_1 \tilde{X}(k|k) A_1' + p_1 C_1 C_1' + p_2 C_2 C_2'$ and $\tilde{X}(k|k) = (I_n - p_1 F_2(k) H_1)\tilde{X}(k|k-1)$, in agreement with [8] and [9]. Finally the above result could have been easily generalized to any positive integer N and any matrices $H_1, \cdots, H_N$, $G_1, \cdots, G_N$ and $C_1, \cdots, C_N$.

Example 2: In this example we compare the LMMSE with the IMM. We considered a scalar dynamical system described by the equation

$$\begin{aligned} x(k+1) &= a_{\theta(k)} x(k) + b_{\theta(k)} u(k) + c_{\theta(k)} \xi(k) \\ y(k) &= h_{\theta(k)} x(k) + g_{\theta(k)} \nu(k) \end{aligned}$$

where $\theta(k) \in \{1, 2\}$, $x(0)$ is Gaussian with mean 10 and variance 10, $u(k) = 10 \cos\{2\pi k/100\}$, $\xi(k)$ and $\nu(k)$ are independent white

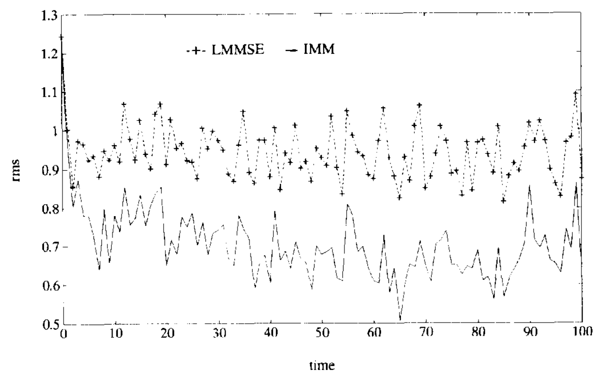

Fig. 2. Case 10, representative of Cases 3, 8, 12, 13.

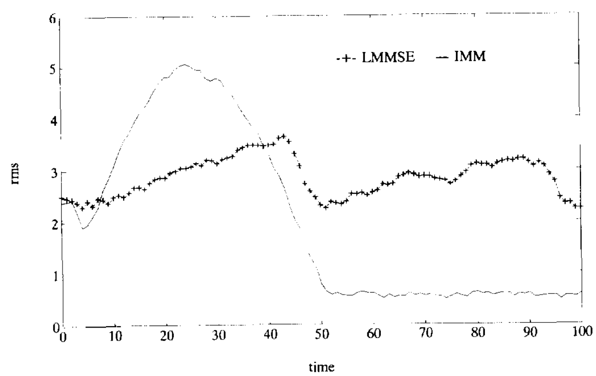

Fig. 3. Case 9, representative of Case 7.

noise sequences. We fixed $b_1 = b_2 = 1$, $\pi_1(0) = \pi_2(0) = 0.5$ and considered the set of 15 cases for the parameters a, c, h, g, and p_{ij}, which are presented on Table I.

In all these cases we run 100 Monte Carlo simulations from $k = 0$ to 100. The paths of θ were generated randomly, and the filters were compared under the same conditions, that is, the same set of paths of θ, initial condition x_0 and noises ξ and ν. For the Cases 1, 2, 4, 6, and 11 both filters presented a similar performance. Fig. 1 presents Case 2, which is representative of these five cases. For the cases 3, 8, 10, 12, and 13 the LMMSE was outperformed by the IMM, but the difference of performances was not too great. Fig. 2 illustrates these situations through Case 10. For the set of parameters 7 and 9, the IMM started worse and finished better than the LMMSE. Fig. 3 shows Case 9, representative of these situations. Finally in Cases 5, 14, and 15 both filters had a similar performance but the IMM presented some peaks where it had a worse performance than the LMMSE. Fig. 4 illustrates these situations through Case 15.

V. Conclusions

In this paper the linear recursive least-square state estimator for linear systems subject to Markovian jumps was obtained. The novelty in our approach is that the estimation is not made directly for $x(k)$ but, instead, for $x(k)1_{\{\theta(k)=i\}}$. By doing this we obtain an optimal linear recursive filter of dimension $N(n+1)$ (Nn if no input deterministic sequence is applied), where the increase in the dimension of the filter is due to the uncertainty introduced by the Markov chain in the linear model. Next we

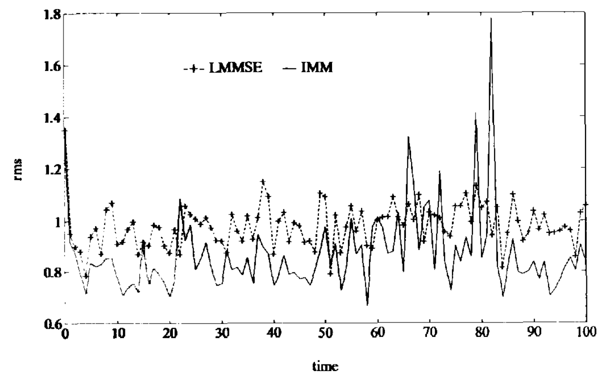

Fig. 4. Case 15, representative of Cases 5, 14.

show how the order of the filter can be reduced from Nn to n for a special case where the uncertainty affects only the output of the system and the plant noise through a sequence of independent and identically distributed binary random variables. The paper is concluded with a numerical comparison between the IMM algorithm and the LMMSE filter, indicating that the latter can be a good alternative to other more computationally expensive approaches.

Acknowledgment

The author would like to thank a reviewer for pointing out reference [8].

References

[1] G. A. Ackerson and K. S. Fu, "On state estimation in switching environments," *IEEE Trans. Automat. Contr.*, vol. 15, pp. 10–17, 1970.
[2] H. A. P. Blom and Y. Bar-Shalom, "The interacting multiple model algorithm for systems with Markovian switching coefficients," *IEEE Trans. Automat. Contr.*, vol. 33, pp. 780–783, 1988.
[3] C. G. Chang and M. Athans, "State estimation for discrete systems with switching parameters," *IEEE Trans. Aerospace Electron. Syst.*, vol. 14, pp. 418–424, 1978.
[4] H. J. Chizeck, A. S. Willsky, and D. Castano, "Discrete-time Markovian-jump linear quadratic optimal control," *Int. J. Contr.*, vol. 43, pp. 213–231, 1986.
[5] O. L. V. Costa and M. D. Fragoso, "Stability results for discrete-time Markovian jump linear systems," *J. Math. Anal. Appl.*, vol. 179, pp. 154–178, 1993.
[6] M. H. A. Davis and R. B. Vinter, *Stochastic Modeling and Control*. London: Chapman and Hall, 1985.
[7] B. E. Griffiths and K. A. Loparo, "Optimal control of jump-linear gaussian systems," *Int. J. Contr.*, vol. 42, pp. 791–819, 1985.
[8] M. T. Hadidi and C. S. Schwartz, "Linear recursive state estimators under uncertain observations," *IEEE Trans. Automat. Contr.*, vol. 24, pp. 944–948, 1979.
[9] N. E. Nahi, "Optimal recursive estimation with uncertain observation," *IEEE Trans. Inform. Theory*, vol. 15, pp. 457–462, 1969.
[10] D. D. Siljak, "Reliable control using multiple control systems," *Int. J. Contr.*, vol. 31, pp. 303–329, 1980.
[11] D. D. Sworder and R. O. Rogers, "An LQG solution to a control problem with a solar thermal receiver," *IEEE Trans. Automat. Contr.*, vol. 28, pp. 971–978, 1983.
[12] J. K. Tugnait, "Detection and estimation for abruptly changing systems," *Automatica*, vol. 18, pp. 607–615, 1982.
[13] J. K. Tugnait, "Adaptive estimation and identification for discrete systems with Markov jump parameters," *IEEE Trns. Automat. Contr.*, vol. 27, pp. 1054–1064, 1982.
[14] Y. Ji, and H. J. Chizeck, "Optimal quadratic control of jump linear systems with separately controlled transition probabilities," *Int. J. Contr.*, vol. 49, pp. 481–491, 1989.