2008 American Control Conference
Westin Seattle Hotel, Seattle, Washington, USA
June 11-13, 2008

Optimal State Estimation for Discrete-Time Markovian Jump Linear Systems, in the Presence of Delayed Mode Observations

Ion Matei, Nuno C. Martins and John S. Baras

Abstract— **In this paper, we investigate an optimal state estimation problem for Markovian Jump Linear Systems. We consider that the state has two components: the first component of the state is finite valued and is denoted as mode, while the second (continuous) component is in a finite dimensional Euclidean space. The continuous state is driven by a deterministic control input and a zero mean, white and Gaussian process noise. The observable output has two components: the first is the mode delayed by a fixed amount and the second is a linear combination of the continuous state observed in zero mean white Gaussian noise. Our paradigm is to design optimal estimators for the current state, given the current output observation. We provide a solution to this paradigm by giving a recursive estimator of the continuous state, in the minimum mean square sense, and a finitely parameterized recursive scheme for computing the probability mass function of the current mode conditional on the observed output. We show that the optimal estimator is nonlinear on the observed output and on the control input. In addition, we show that the computation complexity of our recursive schemes is polynomial in the number of modes and exponential in the mode observation delay.**

I. INTRODUCTION

Markovian jump linear systems (MJLS) can be used to model plants with structural changes, such as in networked control [11], where communication networks/channels are used to interconnect remote sensors, actuators and processors. Moreover, linear plants with random time-delays [12] can also be modeled as Markovian jump systems. Motivated by this wide spectrum of applications, for the last three decades, there has been active research in the analysis [2], [6], controllers and estimators design [5], [6], [8], [9] for Markovian jump linear systems.

A MJLS is characterized by a a state with two components: the first component is finite valued and is denoted as mode, while the second (continuous) component is in a finite dimensional Euclidean space. The continuous state is driven by a deterministic control input and by some process noise. The observation output has two components as well: the first is the mode and the second is a linear combination of the continuous state and some measurement noise. Existing results solve the problem of state estimation for MJLS for two main cases. In the first case, the two components of the observation output are assumed known up until the current time and the Minimum Mean Square Error (MMSE) estimator is derived from the Kalman filter for time varying systems [6], [9]. Off-line computation of the filter is inadvisable due to the path dependence of the filter's gain. An alternative estimator filter, whose gain depends only on the current mode and for which off-line computations are feasible, is given in [7]. In the second case, only the continuous output observation component is known, without any observation of the mode, and the optimal nonlinear filter is obtained by a bank of Kalman filters which require exponentially increasing memory and computation with time [3]. To limit the computational requirements suboptimal estimators have been proposed in the literature [1], [3], [10]. A linear MMSE estimator, for which the gain matrices can be calculated off-line is described in [8].

In this paper we address the problem of state estimation for MJLS with delayed mode observations. The motivation behind considering such setup comes from many practical applications. For example the delayed mode observation setup could model networked systems which rely on acknowledgments as a way to deal with unreliable network links. In real applications, these acknowledgments are not received at the controller instantaneously; instead they are delayed by one or more time-steps.

Notations and abbreviations: Consider a general random process Z_t. By $Z_0^t = \{Z_0, Z_1, ..., Z_t\}$, we denote the history of the process Z_t from 0 up to time t. A realization of Z_0^t is referred to by $z_0^t = \{z_0, z_1, ..., z_t\}$. Let $\{X_t | Y_0^t = y_0^t, M_0^{t-h} = m_0^{t-h}\}$ be a vector valued random process. We denote by $f_{X_t|Y_0^{t-h}M_0^{t-h}}$ its probability density function (p.d.f.). By $\mu^X_{t|(t,t-h)}$ and $\Sigma^X_{t|(t,t-h)}$ we will refer its mean and covariance matrix respectively. For notational simplicity, we will make an abuse of notation and denote by $f_{M_t|Y_0^t M_0^{t-h}}(m_t|y_0^t m_0^{t-h})$ the probability mass function $prob(M_t = m_t | Y_0^t = y_0^t M_0^{t-h} = m_0^{t-h})$. We will compactly write the sum $\sum_{m_0=1}^s \sum_{m_1=1}^s \cdots \sum_{m_t=1}^s$ as $\sum_{m_0^t}$. Assuming that x is a vector in $\mathbb{R}^n$, by the integral $\int f(x)dx$ we understand $\int ... \int f(x_1, ..., x_n) dx_1 ... dx_n$, where x_i are entries of vector x and f is a function defined on $\mathbb{R}^n$ with values in $\mathbb{R}$.

Paper organization: This paper has five more sections besides the introduction. After the formulation of the problem in Section II, in Section III we presents the main results of this paper. Section IV provides the proofs for the results stated in Section III. We end the paper with a simulation section and some conclusion and comments on our solution.

II. PROBLEM FORMULATION

In this section we formulate the problem for the MMSE state estimation for MJLS in the presence of delayed mode

Ion Matei and John S. Baras were funded by National Aeronautics and Space Administration under award No NCC8235

Nuno C. Martins was funded by Award NSF EECS 0644764 CAREER: Distributed control of dynamic systems using a wireless communication medium: two new paradigms

978-1-4244-2079-7/08/$25.00 ©2008 AACC.

observations.

Let us first introduce the definition of a Markovian jump linear system:

Definition 2.1: (*Discrete-time Markovian jump linear system*) Consider n, m, q and s to be given positive integers together with a transition probability matrix $P \in [0,1]^{s\times s}$ satisfying $\sum_{j=1}^{s} p_{ij} = 1$, $p_{ij} \geq 0$, for each i in the set $\mathscr{S} = \{1,\ldots,s\}$, where p_{ij} is the (i,j) element of the matrix P. Consider also a given set of matrices $\{A_i\}_{i=1}^{s}$, $\{B_i\}_{i=1}^{s}$ and $\{C_i\}_{i=1}^{s}$ with $A_i \in \mathbb{R}^{n\times n}$, $B_i \in \mathbb{R}^{n\times m}$ and $C_i \in \mathbb{R}^{q\times n}$ for i belonging to the set $\mathscr{S}$. In addition consider two independent random variable X_0 and M_0 taking values in $\mathbb{R}^n$ and $\mathscr{S}$, respectively. Given the vector valued random processes W_t and V_t taking values in $\mathbb{R}^n$ and $\mathbb{R}^q$ respectively, the following dynamic equations describe a discrete-time Markovian jump linear system:

$$X_{t+1} = A_{M_t}X_t + B_{M_t}u_t + W_t \tag{1}$$

$$Y_t = C_{M_t}X_t + V_t. \tag{2}$$

The state of the system is represented by the doublet (X_t, M_t) where $X_t \in \mathbb{R}^n$ is the state continuous component and M_t is the discrete component. The process M_t is a Markovian jump process taking values in $\mathscr{S}$ with conditional probabilities given by $pr(M_{t+1} = j|M_t = i) = p_{ij}$. The vector $u_t \in \mathbb{R}^m$ is the control input assumed deterministic. The observation output is given by the doublet (Y_t, M_t), where $Y_t \in \mathbb{R}^q$ is the continuous component. Throughout this paper we will consider W_t and V_t to be independent identically distributed (i.i.d.) Gaussian noises with zero means and identity covariance matrices (the covariance matrices of the two noise process were assumed to be identity just for simplicity, every results presented in this paper being valid for any covariance matrix). The initial random vector X_0 has a Gaussian multivariate distribution with mean μ_{X_0} and covariance matrix Σ_{X_0} which, together with the Markovian process M_t and the noises W_t, V_t, are assumed independent for all time instants t.

As it can be noticed, the Markovian jump linear system described by (1)-(2) has a hybrid state with a continuous component X_t taking values on a finite dimensional Euclidean space and a discrete valued component M_t representing the mode of operation. The system has s mode of operations defined by the set of matrices (A_1, B_1, C_1) up to (A_s, B_s, C_s). The Markovian process M_t (called also mode process) determines which mode of operation is active at each time instant. For simplicity, throughout this paper we will differentiate among the different components of the MJLS state and observation output as following. We will refer to X_t as the state vector and to M_t as mode. If known, we will call Y_t as output observation and M_t and mode observation.

We can now proceed with the formulation of our problem of interest.

Problem 2.1: (*MMSE state estimation for MJLS with delayed mode observations*) Consider a Markovian jump linear system as in *Definition 2.1*. Let h be a positive integer representing how long the mode observations are delayed. Assuming that the state vector X_t and the mode M_t are not known, and that at the current time the data available consists in the output observations up to the current time t ($Y_0^t = y_0^t$) and mode observations up to time $t-h$ ($M_0^{t-h} = m_0^{t-h}$) we want to derive the MMSE estimators for the state vector X_t and the mode indicator function $\mathbb{1}_{\{M_t=m_t\}}, m_t \in \mathscr{S}$. More precisely, considering the optimal solution of the MMSE estimators ([13]) we want to compute the following:
MMSE state estimator:

$$\hat{X}_t^h = \mu_{t|(t,t-h)} = E[X_t|Y_0^t = y_0^t, M_0^{t-h} = m_0^{t-h}], \tag{3}$$

MMSE mode indicator function estimator:

$$\hat{\mathbb{1}}^h_{\{M_t=m_t\}} = E[\mathbb{1}_{\{M_t=m_t\}}|Y_0^t = y_0^t, M_0^{t-h} = m_0^{t-h}], \tag{4}$$

where the indicator function $\mathbb{1}_{\{M_t=m_t\}}$ is one if $M_t = m_t$ and zero otherwise.

Remark 2.1: Obtaining an MMSE estimation of the mode indicator function allows us to replace any mode dependent function $g(M_t)$ by an estimation $\widehat{g(M_t)} = \sum_{i\in\mathscr{S}} g(i)\hat{\mathbb{1}}_{\{M_t=i\}}$. We are interested in an estimation of the indicator function rather than of the mode itself because the MMSE estimator of the mode can produce real values which may have limited usefulness;

Remark 2.2: Considering the definition of the indicator function, the MMSE mode indicator function estimation can be also written as: $\hat{\mathbb{1}}^h_{\{M_t=m_t\}} = pr(M_t = m_t|Y_0^t = y_0^t, M_0^{t-h} = m_0^{t-h})$. Then we can also produce a marginal maximal a posteriori mode estimation expressed in terms of the indicator function: $\hat{M}_t^h = arg\max_{m_t\in\mathscr{S}} pr(M_t = m_t|Y_0^t = y_0^t, M_0^{t-h} = m_0^{t-h}) = arg\max_{m_t\in\mathscr{S}} \hat{\mathbb{1}}^h_{\{M_t=m_t\}}$.

III. Main result

In this section we present the solution for *Problem 2.1*. We introduce here two corollaries describing the formulas for computing the state and mode indicator function estimations. An efficient online algorithm implementing the estimators is also given. The proofs of these corollaries are deferred for the next section. Let us first remind ourselves some properties of the Kalman filter for MJLS synthesized in the following theorem.

Theorem 3.1: Consider a discrete MJLS as in *Definition 2.1*. The random processes $\{X_t|Y_0^t = y_0^t M_0^t = m_0^t\}$, $\{X_t|Y_0^{t-1} = y_0^{t-1} M_0^{t-1} = m_0^{t-1}\}$ and $\{Y_t|Y_0^{t-1} = y_0^{t-1} M_0^t = m_0^t\}$ are Gaussian distributed with the means and covariance matrices calculated by the following recursive equations:

$$\begin{aligned}
{\Sigma^X_{t|(t,t)}}^{-1} &= {\Sigma^X_{t|(t-1,t-1)}}^{-1} + C^T_{m_t}C_{m_t} && (5)\\
\mu^X_{t|(t,t)} &= \Sigma^X_{t|(t,t)}\left[C^T_{m_t}y_t + {\Sigma^X_{t|(t-1,t-1)}}^{-1}\mu^X_{t|(t-1,t-1)}\right] && (6)\\
\mu^X_{t|(t-1,t-1)} &= A_{m_{t-1}}\mu^X_{t-1|(t-1,t-1)} + B_{m_{t-1}}u_{t-1} && (7)\\
\Sigma^X_{t|(t-1,t-1)} &= A_{m_{t-1}}\Sigma^X_{t-1|(t-1,t-1)}A^T_{m_{t-1}} + I_n && (8)\\
\mu^Y_{t|(t-1,t)} &= C_{m_t}\mu^X_{t|(t-1,t-1)} && (9)\\
\Sigma^Y_{t|(t-1,t)} &= C_{m_t}\Sigma^X_{t|(t-1,t-1)}C^T_{m_t} + I_q, && (10)
\end{aligned}$$

with initial conditions $\mu^X_{0|(-1,-1)} = \mu_{X_0}$ and $\Sigma^X_{0|(-1,-1)} = \Sigma_{X_0}$.

Remark 3.1: Equations (5)-(8) are a more compact representation of the standard recursive equations of the Kalman filter for MJLS. Equations (9) and (10) follow immediately from de derivation of the filter. We can recover the well known equations by applying the Matrix Inversion Lemma on (5)-(8). Derivation of the Kalman filter equation can be found in [6] for example.

Our main result consists in *Corollaries 3.1 and 3.2* which show the algorithmic steps necessary to compute the MMSE state and mode indicator function estimators for MJLS when the mode observations are affected by some arbitrary (but fixed) delay.

Corollary 3.1: Given a MJLS as in *Definition 2.1* and a positive integer h, the MMSE state estimator from Problem *2.1* is given by the following formula:

$$\mu_{t|(t,t-h)} = \sum_{m^t_{t-h+1}} c_t(m^t_{t-h+1})\mu^X_{t|(t,t)}(m^t_{t-h+1})$$

where $\mu^X_{t|(t,t)}$ is the estimation produced for the Kalman filter (5)-(8) for each of the missing mode path m^t_{t-h+1} and the coefficients $c_t(m^t_{t-h+1})$ are given by

$$c_t(m^t_{t-h+1}) =$$

$$= \frac{\prod_{k=0}^{h-1} p_{m_{t-k-1}m_{t-k}} f_{Y_{t-k}|Y_0^{t-k-1}M_0^{t-k}}(y_{t-k}|y_0^{t-k-1}m_0^{t-k})}{\sum_{m^t_{t-h+1}} \prod_{k=0}^{h-1} p_{m_{t-k-1}m_{t-k}} f_{Y_{t-k}|Y_0^{t-k-1}M_0^{t-k}}(y_{t-k}|y_0^{t-k-1}m_0^{t-k})}, \tag{11}$$

where $f_{Y_{t-k}|Y_0^{t-k-1}M_0^{t-k}}$ is the Gaussian p.d.f. of the process $\{Y_{t-k}|Y_0^{t-k-1} = y_0^{t-k-1}, M_0^{t-k} = m_0^{t-k}\}$ whose mean and covariance matrix are expressed recursively in (9) and (10).

Corollary 3.2: Given a MJLS as in *Definition 2.1* and a positive integer h, the MMSE mode indicator function estimator from *Problem 2.1* is computed according to the next formula:

$$\hat{\mathbb{1}}^h_{\{M_t=m_t\}} = \sum_{m^{t-1}_{t-h+1}} c_t(m^t_{t-h+1})$$

where the coefficients $c_t(m^t_{t-h+1})$ are the same as in the previous corollary.

Remark 3.2: We will show later that the coefficients $c_t(m^t_{t-h+1})$ are the conditional probabilities $pr(M^t_{t-h+1} = m^t_{t-h+1}|Y^t_0 = y^t_0 M_0^{t-h} = m_0^{t-h})$ and therefore they sum up to one. Then, we can express the estimation error as following:

$$\varepsilon_t = X_t - \mu_{t|(t,t-h)} = \sum_{m^t_{t-h+1}} c_t(m^t_{t-h+1})[X_t - \mu_{t|(t,t)}(m^t_{t-h+1})] =$$

$$= \sum_{m^t_{t-h+1}} c_t(m^t_{t-h+1})\tilde{\varepsilon}_t(m^t_{t-h+1}) \leq \sum_{m^t_{t-h+1}} \tilde{\varepsilon}_t(m^t_{t-h+1})$$

where $\tilde{\varepsilon}_t(m^t_{t-h+1})$ is the estimation error of the Kalman filter for MJLS. Therefore the covariance matrix of the estimation error in the case of delayed mode observations is bounded if the covariance matrix of the estimation error produced by the Kalman filter is bounded as well, which implies the stability of the estimator. Note also from above, that the estimation error speed of convergence (in the mean square sense) can be expressed in terms of the speed on convergence of the estimation error produce by the Kalman filter.

These results can be regarded as a generalization of the estimation problem for MJLS. Since we assumed the delay to be fixed, the estimation formulas have a polynomial complexity in the number of modes. However the complexity increases exponentially with the delay which is in accord with the results concerning the Kalman filter for MJLS with no mode observations [3]. We can notice that through the coefficients $c_t(m^h_{t-h+1})$, the estimation introduced in the previous corollaries are nonlinear in the sequence of observed outputs y^t_{t-h+1} and control inputs u^{t-1}_{t-h+1}. This nonlinearity (especially the one in the inputs) makes the estimation error to depend on the inputs as well, and therefore an attempt to solve the optimal linear quadratic problem with partial information using MMSE state estimations becomes difficult since the separation principle can no longer be obtained.

A. Algorithms implementation

By *Corollary 3.1* we note that in order to calculate the optimal estimation $\hat{X}^h_t$ we need to compute a number of s^h Kalman filter estimations, corresponding to all possible paths of the Markov process M_t from $t-h+1$ up to t, plus an equal number of coefficients $c_t(m^t_{t-h+1})$. Hence for the MMSE state estimator for a MJLS with delayed mode observations the numerical complexity and required memory space increases exponentially with the delay h. However for a fixed h the estimators can be implemented (with polynomial complexity in the number of modes of operation) and an algorithm is presented in the following. A naive way to implement the estimators consists in using, at each time instant, the Kalman filter iterations to compute the estimations $\mu_{t|t,t}$ for all possible mode paths m^t_{t-h+1} for each time instant having has initial condition the Kalman filter estimate $\mu_{t-h|(t-h,t-h)}$. The numerical complexity in this case would be $s + s^2 + \ldots + s^h$. This is naive because it does not make use of information already available from past time instants. An implementation with a lower numerical complexity is presented in the following (see next page).

We start the algorithm having as initial information the mean and the covariance of the initial state vector X_0 together with the initial output and mode observations, y_0 and m_0 respectively, and the mode observation delay h. Before entering the infinite time loop at line 5 we compute the mean and covariance matrix for the process $\{X_0|Y_0 = y_0, M_0 = m_0\}$ which will constitute the initial value for the iterations in lines 6-21. In lines 7-13 we compute the mean and covariance matrix for the process $\{X_t|Y^t_0 = y^t_0, M^t_0 = m^t_0\}$ for all possible paths m^t_{t-h+1}. This provides also all the values necessary for evaluating the parameters c_t. After computing the desired estimations in line 16-17 we make use of the newly arrived mode observation m_{t-h+1} corresponding to time instant $t+1$

and keep only the means and the covariances matrices that match the newly updated mode path (lines 18-21).

Observe that at each time step t the number of computations is of order of s^h. In terms of memory requirements, the memory space needed is of order of $s+s^2+...+s^{h-1}$ mainly due to the need for storing the values of the functions $f_{Y_{t-k}|Y_0^{t-k-1}M_0^{t-k}}$ with $k \in \{h-1,...,2,1\}$.

Algorithm 1: MMSE state estimation for MJLS with delayed mode observations

Input: μ_{X_0}, Σ_{X_0}, m_0, h

1 **begin**

2 $\quad \mu^X_{0|(-1,-1)} = \mu_{X_0}$, $\Sigma^X_{0|(-1,-1)} = \Sigma_{X_0}$;

3 $\quad {\Sigma^X_{0|(0,0)}}^{-1} = {\Sigma^X_{0|(-1,-1)}}^{-1} + C^T_{m_0} C_{m_0}$;

4 $\quad \mu^X_{0|(0,0)} = \Sigma^X_{0|(0,0)} [C^T_{m_0} y_0 + {\Sigma^X_{0|(-1,-1)}}^{-1} \mu^X_{0|(-1,-1)}]$;

5 $\quad$ **forall** $t \geq 1$ **do**

6 $\quad\quad$ **for** $m^t_{t-h+1} \in \mathscr{S}$ **do**

7 $\quad\quad\quad \mu^X_{t|(t-1,t-1)} = A_{m_{t-1}} \mu^X_{t-1|(t-1,t-1)} + B_{m_{t-1}} u_{t-1}$;

8 $\quad\quad\quad \Sigma^X_{t|(t-1,t-1)} = A_{m_{t-1}} \Sigma^X_{t-1|(t-1,t-1)} A^T_{m_{t-1}} + I_n$;

9 $\quad\quad\quad \mu^Y_{t|(t-1,t)} = C_{m_t} \mu^X_{t|(t-1,t-1)}$;

10 $\quad\quad\quad \Sigma^Y_{t|(t-1,t)} = C_{m_t} \Sigma^X_{t|(t-1,t-1)} C^T_{m_t} + I_q$;

11 $\quad\quad\quad f_{Y_t|Y_0^{t-1}M_0^t}$=Gaussian($\mu^Y_{t|(t-1,t)}$,$\Sigma^Y_{t|(t-1,t)}$);

12 $\quad\quad\quad {\Sigma^X_{t|(t,t)}}^{-1} = {\Sigma^X_{t|(t-1,t-1)}}^{-1} + C^T_{m_t} C_{m_t}$;

13 $\quad\quad\quad \mu^X_{t|(t,t)} = \Sigma^X_{t|(t,t)} [C^T_{m_t} y_t + {\Sigma^X_{t|(t-1,t-1)}}^{-1} \mu^X_{t|(t-1,t-1)}]$;

14 $\quad\quad\quad c_t(m^t_{t-h+1}) = \dfrac{\prod_{k=0}^{h-1} p_{m_{t-k-1}m_{t-k}} f_{Y_{t-k}|Y_0^{t-k-1}M_0^{t-k}}(y_{t-k}|y_0^{t-k-1}m_0^{t-k})}{\sum_{m^t_{t-h+1}} \prod_{k=0}^{h-1} p_{m_{t-k-1}m_{t-k}} f_{Y_{t-k}|Y_0^{t-k-1}M_0^{t-k}}(y_{t-k}|y_0^{t-k-1}m_0^{t-k})}$;

15 $\quad\quad$ **end**

16 $\quad\quad \mu^X_{t|(t,t-h)} = \sum_{m^t_{t-h+1}} c_t(m^t_{t-h+1}) \mu^X_{t|(t,t)}$;

17 $\quad\quad \hat{\mathbb{1}}^h_{\{M_t=m_t\}} = \sum_{m^{t-1}_{t-h+1}} c_t(m^t_{t-h+1})$;

18 $\quad\quad \mu^X_{t-1|(t-1,t-1)}(m^t_{t-h+2}) = \mu^X_{t|(t,t)}(m^t_{t-h+2})$;

19 $\quad\quad$ **for** $k \in \{h-1,\dots,1\}$ **do**

20 $\quad\quad\quad f_{Y_{t-k-1}|Y_0^{t-k-2}M_0^{t-k-1}}(\cdot) = f_{Y_{t-k}|Y_0^{t-k-1}|M_0^{t-k}}(\cdot)$;

21 $\quad\quad$ **end**

22 $\quad$ **end**

23 **end**

IV. PROOF OF THE MAIN RESULT

In this section we present the proof of our main results. *Corollaries 3.1 and 3.2*, are a direct consequence of the statistical properties of the random process $\{X_t|Y_0^t = y_0^t M_0^{t-h} = m_0^{t-h}\}$ introduced in *Theorem 4.1*. In this theorem we show that the p.d.f $f_{X_t|Y_0^t M_0^{t-h}}$ is a mixture of Gaussian p.d.f's with coefficients depending nonlinearly on output observations and control inputs. To simplify the proof of *Theorem 4.1* we introduce the following corollary in which we characterized the statistical properties of a linear combination of two Gaussian random vectors.

Corollary 4.1: Consider two Gaussian random vectors V and X of dimension m and n respectively, with means $\mu_V = 0$ and μ_X and covariance matrices $\Sigma_V = I_m$ and Σ_X respectively. Let Y be a Gaussian random vector resulted from a linear combination of X and V, $Y = CX + V$ where C is a matrix of appropriate dimensions. Then the following holds:

$$\int_{\mathbb{R}^n} f_V(y - Cx) f_X(x) dx = f_Y(y), \tag{12}$$

where $f_Y(y)$ is the multivariate Gaussian p.d.f. of Y with parameters $\mu_Y = C\mu_Y$ and $\Sigma_Y = C\Sigma_X C^T + I_m$. Also,

$$f_V(y - Cx) \cdot f_X(x) = \tilde{f}_{\tilde{X}}(x) \cdot f_Y(y), \tag{13}$$

where $\tilde{f}_{\tilde{X}}(x)$ is a Gaussian p.d.f. with parameters $\mu_{\tilde{X}} = \Sigma_{\tilde{X}}(C^T y + \Sigma_X^{-1}\mu_X)$, $\Sigma_{\tilde{X}}^{-1} = \Sigma_X^{-1} + C^T C$ and, $f_Y(y)$ being defined in (12).

The above corollary is just in generalization at the level of vectors for a well known results concerning the sum of two Gaussian random variables.

Theorem 4.1: Consider a discrete MJLS as in *Definition 2.1* and let h be a known positive integer value. Then the p.d.f. of the random process $\{X_t|Y_0^t = y_0^t M_0^{t-h} = m_0^{t-h}\}$ is a mixture of Gaussian probability densities. More precisely:

$$f_{X_t|Y_0^t M_0^{t-h}}(x|y_0^t m_0^{t-h}) = \sum_{m^t_{t-h+1}} c_t(m^t_{t-h+1}) f_{X_t|Y_0^t M_0^t}(x|y_0^t, m_0^t) \tag{14}$$

where $c_t(m^t_{t-h+1}) = f_{M^t_{t-h+1}|Y_0^t M_0^{t-h}}(m^t_{t-h+1}|y_0^t m_0^{t-h})$ are the (time varying) mixture coefficients and $f_{X_t|Y_0^t M_0^t}(x|y_0^t, m_0^t)$ is the gaussian p.d.f. of the process $\{X_t|Y_0^t = y_0^t, M_0^t = m_0^t\}$ whose statistics is computed according to the recursions (5)-(8). The coefficients $c_t(m^t_{t-h+1})$ are computed by the following formula:

$$c_t(m^t_{t-h+1}) = \frac{\prod_{k=0}^{h-1} p_{m_{t-k-1}m_{t-k}} f_{Y_{t-k}|Y_0^{t-k-1}M_0^{t-k}}(y_{t-k}|y_0^{t-k-1}m_0^{t-k})}{\sum_{m^t_{t-h+1}} \prod_{k=0}^{h-1} p_{m_{t-k-1}m_{t-k}} f_{Y_{t-k}|Y_0^{t-k-1}M_0^{t-k}}(y_{t-k}|y_0^{t-k-1}m_0^{t-k})} \tag{15}$$

where $f_{Y_{t-k}|Y_0^{t-k-1}M_0^{t-k}}$ is the Gaussian p.d.f. of the process $\{Y_{t-k}|Y_0^{t-k-1} = y_0^{t-k-1}, M_0^{t-k} = m_0^{t-k}\}$ whose mean and covariance matrix are expressed in (9) and (10).

Proof: Using the law of marginal probabilities we get:

$$f_{X_t|Y_0^t M_0^{t-h}}(x|y_0^t m_0^{t-h}) = \sum_{m^t_{t-h+1}} f_{X_t M^t_{t-h+1}|Y_0^t M_0^{t-h}}(x, m^t_{t-h+1}|y_0^t m_0^{t-h}) =$$

$$= \sum_{m^t_{t-h+1}} f_{X_t|Y_0^t M_0^t}(x|y_0^t, m_0^t) f_{M^t_{t-h+1}|Y_0^t M_0^{t-h}}(m^t_{t-h+1}|y_0^t m_0^{t-h}) =$$

$$= \sum_{m^t_{t-h+1}} c_t(m^t_{t-h+1}) f_{X_t|Y_0^t M_0^t}(x|y_0^t, m_0^t)$$

Thus we obtained (14). All you are left to do is to compute coefficients of this linear combination. By applying the Bayes rule we get:

$$f_{M_{t-h+1}^t|Y_0^t M_0^{t-h}}(m_{t-h+1}^t|y_0^t m_0^{t-h}) = \frac{f_{Y_0^t M_0^t}(y_0^t m_0^t)}{\sum_{m_{t-h+1}^t} f_{Y_0^t M_0^t}(y_0^t m_0^t)} \tag{16}$$

The p.d.f. $f_{Y_0^t M_0^t}$ can be expressed recursively as:

$$f_{Y_0^t M_0^t}(y_0^t m_0^t) = \int_{\mathbb{R}^n} f_{X_t Y_0^t M_0^t}(x_t y_0^t m_0^t) dx_t =$$

$$= \int_{\mathbb{R}^n} f_{Y_t|X_t M_t}(y_t|x_t m_t) f_{X_t|Y_0^{t-1} M_0^{t-1}}(x_t|y_0^{t-1} m_0^{t-1}) dx_t$$

$$p_{m_{t-1} m_t} f_{Y_0^{t-1} M_0^{t-1}}(y_0^{t-1} m_0^{t-1}).$$

Applying (12) we obtain:

$$f_{Y_0^t M_0^t}(y_0^t m_0^t) =$$

$$= f_{Y_t|Y_0^{t-1} M_0^t}(y_t|y_0^{t-1} m_0^t) p_{m_{t-1} m_t} f_{Y_0^{t-1} M_0^{t-1}}(y_0^{t-1} m_0^{t-1})$$

Using this recursive expression we get:

$$f_{Y_0^t M_0^t}(y_0^t m_0^t) =$$

$$= \prod_{k=0}^{h-1} p_{m_{t-k-1} m_{t-k}} f_{Y_{t-k}|Y_0^{t-k-1} M_0^{t-k}}(y_{t-k}|y_0^{t-k-1} m_0^{t-k}) f_{Y_0^{t-h} M_0^{t-h}}(y_0^{t-h} m_0^{t-h})$$

By replacing the previous expression in (16) we obtain the coefficients $c_t(m_{t-h+1}^t)$ expressed in (14). We can conclude de proof by making the observations that the p.d.f. $f_{Y_{t-k}|Y_0^{t-k-1} M_0^{t-k}}$ is completely characterized in *Theorem 3.1*, equation (9) and (10). ∎

Corollary 3.1 *Proof:* The result follow immediately from the linearity property of the expectation operator and from (14). ∎

Corollary 3.2 *Proof:* By the law of marginal probability we can write

$$f_{M_t|Y_0^t M_0^{t-h}}(m_t|y_0^t m_0^{t-h}) = \sum_{m_{t-h+1}^{t-1}} f_{M_{t-h+1}^t|Y_0^t M_0^{t-h}}(m_{t-h+1}^t|y_0^t m_0^{t-h})$$

Together with (15), the proof is concluded. ∎

V. SIMULATIONS

In this section we present a comparison between the simulation results obtained with our MMSE state estimator and with two heuristic estimation schemes which will be described in what follows. We consider a MJLS as in (1) with the state vector $X_t \in \mathbb{R}^2$ and the output $Y_t \in \mathbb{R}$. We assume that the system is being driven just by the Gaussian noise, and there are four modes of operations described by the matrices:

$$A_1 = \begin{pmatrix} 0.9 & 0.1 \\ 0.4 & 0.2 \end{pmatrix}, \; A_2 = \begin{pmatrix} 0.8 & 0.2 \\ 0.7 & 0.3 \end{pmatrix},$$

$$A_3 = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.1 \end{pmatrix}, \; A_4 = \begin{pmatrix} 1 & 0 \\ 0.5 & 0.2 \end{pmatrix},$$

$$C_1 = \begin{pmatrix} 0 & 0.5 \end{pmatrix}, \; C_2 = \begin{pmatrix} 0.5 & 1 \end{pmatrix},$$

$$C_3 = \begin{pmatrix} 1 & 0.5 \end{pmatrix}, \; C_4 = \begin{pmatrix} 0.5 & 0.5 \end{pmatrix}.$$

The Markov chain M_t has four states and the probability transition matrix is

$$P = \begin{pmatrix} 0.3 & 0.7 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0.3 & 0.4 & 0.3 \\ 0.5 & 0.5 & 0 & 0 \end{pmatrix}.$$

The covariance matrices of the two noise processes are chosen as $\Sigma_W = 0.1I$ and $\Sigma_V = 0.4I$. We assume that the modes observations are delayed by three time instants ($h = 3$). We use two heuristic estimation schemes for comparison. The first scheme (S1) consists in using the delayed mode observations as current ones and using the Kalman filter to determine the estimation. For example at time t, the available modes are m_{t-3} and m_{t-2} which will be used in stead of m_t and m_{t-1}, which are the modes needed in the Kalman iteration. The second scheme (S2) consists in using a rudimentary estimation of the mode provided by the following optimization:

$$m_t^* = \arg\max_{m_t} Prob(M_t = m_t | M_0^{t-3} = m_0^{t-3})$$

where $m_t \in \mathscr{S}$ and probability in the above optimization being computed using the probability transition matrix P. We solve a similar optimization problem to derive an estimate for the mode m_{t-1}. The initial condition X_0 is extracted from a Gaussian distribution with zero mean and covariance matrix $\Sigma_{X_0} = 0.1I$. The initial distribution of the mode was chosen to be $p_0 = \begin{pmatrix} 0.2 & 0.3 & 0.1 & 0.4 \end{pmatrix}$.

In the following we provide simulation results of the estimation schemes proposed above. The simulations were run from t=0 to 3000. The paths M_t were generated randomly and the filters were compared under the same conditions, that is, the same set of paths of M_t, initial conditions X_0 and noises W_t and V_t.

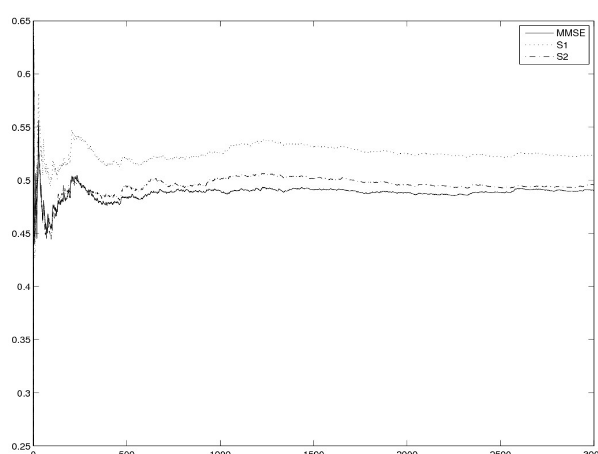

Fig. 1. Mean square estimation error for MMSE state estimation with delayed mode observations and for schemes S1 and S2.

In Fig. 1 we present the simulation results for a realization of the mode path, initial condition and noise. We plot the mean square error estimations obtained with the MMSE with delayed mode observations, S1 and S2 estimations schemes, respectively. As expected the first scheme give the smallest error and, as intuitively may have been expected, S1 scheme behave the worst. Notice that due to the presence of the (Gaussian) noise, the estimation error does not converge to zero but rather it stabilizes to some value depending of the covariances matrices of the noises.

VI. Conclusions

In this paper we considered the problem of state estimation for a MJLS when the discrete component of the output observation (namely the mode) is affected by an arbitrary but fixed delay. We introduced formulas for MMSE estimators for both the continuous and discrete components of the state of the MJLS. These formulas admit recursive implementation and have polynomial complexity in terms of the number of modes of operation and therefore are feasible for practical implementation. We showed that the MMSE state estimation with delayed mode observations depends nonlinearly on a sequence of output observations and control inputs, sequence whose length is determined by the value of the delay and that the same property remain valid for the estimation error as well. We also provided an efficient algorithm for computing the optimal state estimation which admits an online implementation. Although the estimators provided in this paper may prove difficult to use in solving optimal linear quadratic control problems with partial information, they are however useful for deriving sub-optimal control strategies or in tracking problems where an accurate state estimation is desired.

References

[1] G.A. Ackerson and K.S. Fu, On the state estimation in switching environments, *IEEE Transaction on Automatic Control*, 15:10-17, 1970.

[2] W.P. Blair and D.D. Sworder, "Feedback control of a class of linear discrete systems with jump parameters and quadratic cost criteria", *International Journal of Control*, 21:833-844,1975.

[3] Y. Bar-Shalom and X.R. Li, "Estimation and Tracking. Principles, Techniques and Software", Artech House, 1993.

[4] Y. Ji, H.J. Chizeck, X. Feng and K.A. Loparo, "Stability and control of dicrete-time jump linear systems", *Control Theory Adv. Technol.*, vol. 7, no. 2, 1991, pages 247-270.

[5] H.J. Chizeck, A.S. Willsky and D. Castanon, "Discrete-time Markovian-jump linear quadratic optimal control", *Int. J. Control*, vol. 43, no. 1, 1986, pages 213-231.

[6] H.J. Chizeck and Y. Ji, "Optimal quadratic control of jump linear systems with Gaussian noise in discrete-time", *Proceedings of the 27th IEEE Conference on Decision and Control*, pages 1989-1992, 1988.

[7] O.L.V. Costa, M.D. Fragoso and R.P. Marques, "Discrete-Time Markov Jump Linear Systems", Springer, 2005.

[8] O.L.C. Costa, "Linear Minimum Mean Square Error Estimation for Discrete-Time Markovian Jump Linear Systems", *IEEE Transactions on Automatic Control*, vol. 39, no. 8, 1994.

[9] M.H.A. Davis and R.B. Vinter, "Stochastic Modelling and Control", Chapman and Hall, 1985.

[10] F. Dufour and R.J. Elliot, Adaptive control of linear systems with Markov perturbations, *IEEE Transaction on Automatic Control*, 43:351- 372,1997.

[11] J. P. Hespanha, P. Naghshtabrizi and Xu Yonggang , "A Survey of Recent Results in Networked Control Systems", *Proceedings of the IEEE*, Volume 95, Issue 1, Jan. 2007 Page(s):138 - 162

[12] J. Nilsson, B. Bernhardsson and B. Wittenmark, "Stochastic analysis and control of real-time systems with random time delays", *Automatica*, Volume 34, Issue 1, January 1998, Pages 57-64

[13] I.B. Rhodes, "A Tutorial Introduction to Estimation and Filtering", *IEEE Transactions on Automatic Control*, vol. AC-16, no. 6, 1971.

[14] L. Aggoun and R. Elliot, "Measure Theory and Filtering", Cambridge University Press, 2004.