Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

Optimal Linear Quadratic Regulator for Markovian Jump Linear Systems, in the presence of one time-step delayed mode observations ⋆

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

Department of Electrical and Computer Engineering, University of Maryland, College Park, 20742, MD, USA; e-mail: imatei@umd.edu, nmartins@umd.edu, baras@isr.umd.edu

Abstract: In this paper, we provide the solution to the optimal Linear Quadratic Regulator (LQR) paradigm for Markovian Jump linear Systems, when the continuous state is available at the controller instantaneously, but the mode is available only after a delay of one time step. This paper is the first to investigate the LQR paradigm in the presence of such mismatch between the delay in observing the mode and the continuous state at the controller. We show that the optimal LQR policy is a time-varying matrix gain multiplied by the continuous component of the state, where the gain is indexed in time by the one-step delayed mode. The solution to the LQR is expressed as a collection of coupled Riccati iterations and equations, for the finite and the infinite horizon cases respectively. In the infinite horizon case the solution of the coupled Riccati equations or a certificate of infeasibility is obtained by solving a set of linear matrix inequalities. We also explain the difficulties of solving the LQR problem when the mode is observed by the controller with a delay of more than one step. We show that, with delays of more than one time-step, the optimal control will be a time-varying nonlinear function of the continuous state and of the control input, without presenting an exact solution.

1. INTRODUCTION

Networked control systems often 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 at least one time-step. Under the packet drop model, a wide class of networked control systems, in the presence of unreliable links and delayed observations of the mode at the controller, can be modeled as a Markovian jump linear system.

Markovian jump linear systems represents an important class of stochastic time-variant due to their ability to model random abrupt changes that occur in a linear plant structure. Motivated by a wide spectrum of applications there has been active research in the analysis Blair [1975], Ji [1991] and in the design of controllers Chizeck [1986] of controllers for Markovian jump linear systems. More specifically, in the last fifteen, the classical paradigms of optimal control for Markovian jump linear systems (see Costa [2005] for a more detailed survey of existing work). Of particular interest is the Linear Quadratic Regulator problem addressed in Abou [1994], Rami [1996], Ji [1990] where the solution is obtained in terms of a set of coupled Riccati equations. Other relevant results concerning the LQR problem can be found in Boukas [1998, 1999].

In this paper we provide the solution to the optimal finite and infinite horizon Linear Quadratic Regulator paradigm for Markovian jump linear systems, when the mode observations are delayed by one time step. We show the optimal control is a matrix gain multiplied by of the continuous component of the state, where the gain is selected according to the one-step delayed mode observation. The optimal solution is obtained by solving a set of coupled algebraic Riccati equations generated by applying a dynamic programming approach on the optimization problem. The existence of infinite horizon optimal control gains is tested by formulating and equivalent LMI optimization problem.

Notations and abbreviations: Consider a general random process Z_t. Denote by Z_0^t the history of the process from 0 up to time time t as $Z_0^t = \{Z_0, Z_1, ..., Z_t\}$. Refer a realization of Z_0^t by $z_0^t = \{z_0, z_1, ..., z_t\}$. For simplicity we will use the abbreviations "MJLS" for denoting a Markovian jump linear system and "LQR^τ" to abbreviate the Linear Quadratic Regulator paradigm with τ steps delayed mode observations. We will use "CARE" to refer to a set of coupled Riccati equations.

Definition 1. (*Markovian jump linear system*) Consider n, m and s to be given positive integers together with a transition probability matrix $P \in [0,1]^{s\times s}$ satisfying $\sum_{j=1}^{s} p_{ij} = 1$, for each i in the set $\mathcal{S} = \{1, \ldots, s\}$, where p_{ij}, is the (i,j) element of the matrix P. Consider also a given set of matrices $\{A_i\}_{i=1}^s$ and $\{B_i\}_{i=1}^s$, with $A_i \in \mathbb{R}^{n\times n}$ and $B_i \in \mathbb{R}^{n\times m}$ for i belonging to the set $\mathcal{S}$. In addition consider the random vector X_0 and the random variable M_0 assumed independent which take values in $\mathbb{R}^n$

⋆ The material is based upon work supported by National Aeronautics and Space Administration under award No NCC8235.

and $\mathcal{S}$, respectively. Given an i.i.d. vector valued random process W_t taking values in $\mathbb{R}^n$, the following dynamic equation describes a discrete-time Markovian jump linear system:

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

where M_t is a Markovian jump process taking values in $\mathcal{S}$ with conditional probabilities given by $pr(M_{t+1} = j|M_t = i) = p_{ij}$. In equation (1), $X_t \in \mathbb{R}^n$ is the state vector, $U_t \in \mathbb{R}^m$ represents the control input and $W_t \in \mathbb{R}^n$ with zero mean and covariance matrix Σ_W is the process noise. The initial condition vector valued random variable X_0 with covariance matrix Σ_{X_0}, together with the Markovian process M_t and the noise W_t are assumed independent for all values of time t. We also make the assumption that the Markovian process M_t is ergodic with the steady state distribution denoted by $\pi_i = \lim_{t\to\infty} pr(M_t = i)$, for $i \in \mathcal{S}$.

As it can be noticed, the Markovian jump linear system described by (1) 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) up to (A_s, B_s). The Markovian process M_t (which will also be called mode process), determines which mode of operation is active at each time instant.

Paper structure: The paper is organized in five sections. Section 2 introduces the problem formulation addressed in this paper. The main results concerning the optimal solution for LQR^1 paradigm are presented in Section 3. Section 4 contains the proof of the main result together with an analysis of the more general LQR^τ, where the mode observations are delayed by an arbitrary number of steps.

2. PROBLEM FORMULATION

In this section we formulate the optimal LQR paradigm under full state and one step delayed mode feedback and present our main results concerning this problem.

Let us first define the class of admissible controllers:

Definition 2. (*admissible controllers*) Given the sequence of state observations up to time t and the sequence of mode observations up to time $t-1$, the class of admissible controllers $\mathfrak{U}^{LQR^1}$ consists of all feedback policies $\mathcal{U}^{LQR^1}$ with the following structure:

$$U_t = \mathcal{U}^{LQR^1}(t, X_0^t, M_0^{t-1}). \tag{2}$$

Using *Definition 2* we can now formulate LQR^1 problem for MJLS.

Problem 1. (*problem statement*) Consider a Markovian jump linear system as in *Definition* 1. Given an admissible controller $\mathcal{U} \in \mathfrak{U}^{LQR^1}$, a time horizon $N \in \mathbb{N} \cup \{\infty\}$, a positive definite matrix $Q \in \mathbb{R}^{n\times n}$ and a positive definite matrix $R \in \mathbb{R}^{m\times m}$, consider the following quadratic cost function:

$$\mathcal{J}_N(X_0, M_0, \mathcal{U}) =$$

$$= E[\sum_{t=0}^{N-1} (X_t^T Q X_t + U_t^T R U_t) + X_N^T Q X_N | X_0 M_0], \tag{3}$$

where $N < \infty$ and $U_t = \mathcal{U}(t, X_0^t, M_0^{t-1})$. The optimal LQR^1 paradigm is defined by the following optimization problem:

$$\mathcal{U}^{*,LQR^1,N} = arg \min_{\mathcal{U}\in\mathfrak{U}^{LQR^1}} \mathcal{J}_N(X_0, M_0, \mathcal{U}). \tag{4}$$

Consider the long run average cost function $\mathcal{J}_{av}(X_0, M_0, \mathcal{U})$ defined for some arbitrary $\mathcal{U} \in \mathfrak{U}^{LQR^1}$

$$\mathcal{J}_{av}(X_0, M_0, \mathcal{U}) = \limsup_{N\to\infty} \frac{1}{N} \mathcal{J}_N(X_0, M_0, \mathcal{U}). \tag{5}$$

The solution for the infinite-horizon optimal LQR^1 paradigm is given by:

$$\mathcal{U}^{*,LQR^1,\infty} = arg \min_{\mathcal{U}\in\mathfrak{U}^{LQR^1}} \mathcal{J}_{av}(X_0, M_0, \mathcal{U}). \tag{6}$$

3. MAIN RESULT

In this section we solve the finite-horizon and infinite-horizon LQR^1 problem for MJLS. Our main result consists in two theorem whose proof are deferred for the following section. We end Section 3 with a proposition concerning the more general LQR^τ problem.

Theorem 1. (*finite-horizon*) **Part I-Optimal Solution -** The optimal solution for the finite horizon LQR^1 paradigm is given by

$$U_t^* = -\mathcal{K}_{t,M_{t-1}} X_t \tag{7}$$

where

$$\mathcal{K}_{t,j} = \mathcal{B}_{t,j}^{-1} \mathcal{C}_{t,j}^T \tag{8}$$

with

$$\mathcal{B}_{t,j} = \sum_{i=1}^{s} p_{ji} B_i^T P_{t+1,i} B_i + R \tag{9}$$

$$\mathcal{C}_{t,j} = \sum_{i=1}^{s} p_{ji} A_i^T P_{t+1,i} B_i \tag{10}$$

for $t \in \{1, ..., N-1\}$, and

$$\mathcal{B}_{0,j} = B_j^T P_{1,j} B_j + R$$
$$\mathcal{C}_{0,j} = A_j^T P_{1,j} B_j$$

for $t = 0$, and where

$$P_{t,j} = \mathcal{A}_{t,j} - \mathcal{C}_{t,j} \mathcal{B}_{t,j}^{-1} \mathcal{C}_{t,j}^T, \tag{11}$$

where $P_{t,j}$ are $n \times n$ symmetric positive semidefinite matrices with $P_N = Q$,

$$\mathcal{A}_{t,j} = \sum_{i=1}^{s} p_{ji} A_i^T P_{t+1,i} A_i + Q \tag{12}$$

(**Part II-Optimal Cost**) The optimal cost function is

$$\mathfrak{J}_N^*(X_0, M_0) = X_0^T P_{0,M_0} X_0 + g_{0,M_0} \tag{13}$$

where P_{0,M_0} is computed backwards according to (11) and g_{0,M_0} is obtain from

$$g_{t,M_{t-1}} = E[trace(P_{t+1,M_t}\Sigma_W) + g_{t+1,M_t}|M_{t-1}] \tag{14}$$

with $g_N = 0$ (and by convention $M_{-1} = M_0$).

Theorem 2. (*infinite-horizon*) (**Part I-Existence**) An optimal solution for the infinite horizon LQR^1 described in *Problem 1* exists if and only if the following CARE has solution:

$$P_{\infty,j} = \mathcal{A}_j - \mathcal{C}_j\mathcal{B}_j^{-1}\mathcal{C}_j^T,\ P_{\infty,j} = P_{\infty,j}^T > 0 \tag{15}$$

where

$$\mathcal{A}_j = \sum_{i=1}^{s} p_{ji}(A_i^T P_{\infty,i} A_i + Q) \tag{16}$$

$$\mathcal{B}_j = \sum_{i=1}^{s} p_{ji}(B_i^T P_{\infty,i} B_i + R) \tag{17}$$

$$\mathcal{C}_j = \sum_{i=1}^{s} p_{ji} A_i^T P_{\infty,i} B_i \tag{18}$$

(Part II-Optimal solution) If the optimal solution exists then it is given by:

$$U_t^* = -\mathcal{K}_{\infty,M_{t-1}} X_t \tag{19}$$

where

$$\mathcal{K}_{\infty,j} = \mathcal{B}_j^{-1}\mathcal{C}_j^T \tag{20}$$

(Part III-Optimal Cost) The infinite horizon optimal cost is expressed as:

$$\mathcal{J}_{av}^*(X_0, M_0) = \sum_{i=1}^{S} \pi_i trace(\Sigma_W P_{\infty,i}) \tag{21}$$

Remark 1. Theorems 1 and *2* shows that the optimal policy for the LQR^1 paradigm, as intuitively may have been expected, is a matrix gain multiplied by of the continuous component of the state, where the gain is selected according to the one-step delayed mode observation.

Remark 2. The main difference between the solutions of the standard LQR problem for MJLS (Chizeck [1986]) and of our problem consists in the way the coupled Ricatti equations in (11) are formulated. Compared to the standard case, where the symmetric matrices $P_{t,i}$ are obtained from averaged versions of $P_{t+1,M_{t+1}}$, within the current setup the resulting coupled Riccati equations have a more complex expression. This is mainly due to the fact that since $P_{t,M_{t-1}}$ depends on M_{t-1} (rather then M_t as in the standard case) we need to average entire expressions of the form $A_{M_t}^T P_{t+1,M_t} A_{M_t}$ resulting in the terms $\mathcal{A}_{t,j}$, $\mathcal{B}_{t,j}$ and $\mathcal{C}_{t,j}$.

The following proposition characterize the more general LQR^τ paradigm where at time t the controller has available a sequence of mode observations up to time $t-\tau$.

Proposition 3. Consider the LQR^τ paradigm. For $\tau = 1$ the optimal solution is expressed in *Theorems 1* and *2*. For $\tau \geq 2$ the optimal solution is a nonlinear function of the state sequence $X_{t-\tau}^t$, optimal control sequence $U_{t-\tau}^{t-1}$ and the mode process $M_{t-\tau}$.

4. PROOF OF THE MAIN RESULT

In this section we address the proofs of the results introduced in the previous section. The proof of *Theorems 1* and *2* are based on the following lemma.

Lemma 4. Consider the LQR^1 paradigm presented in *Problem 1*. Then the optimal cost to go function

$$\mathcal{J}_t^*(X_0^t, M_0^{t-1}) = \min_{\mathcal{U}\in\mathfrak{U}^{LQR^1}} E[\sum_{k=t}^{N-1}(X_k^T Q X_k + $$
$$+ U_k^T R U_k) + X_N^T Q X_N | X_0^t, M_0^{t-1}]$$

for $t \in \{0, ..., N\}$ (and where by convention we take $M_0^{-1} = M_0$), can be written as:

$$\mathcal{J}_t^*(X_0^t, M_0^{t-1}) = X_t^T P_{t,M_{t-1}} X_t + g_{t,M_{t-1}} \tag{22}$$

where $P_{t,j},\ j \in \mathcal{S}$ is computed as in (11) and $g_{t,M_{t-1}}$ is given by the recurrence (14).

Proof: The proof uses an inductive technique and follows the steps of the proof of a similar lemma in the case of the standard LQR problem for MJLS. For $t = N$ we have that:

$$\mathcal{J}_N^*(X_0^N, M_0^{N-1}) = X_N^T P_N X_N$$

with $P_N = Q$. Assume the optimal cost to go at time $t+1$ has a form as in (22):

$$\mathcal{J}_{t+1}^*(X_0^{t+1}, M_0^t) = X_{t+1}^T P_{t+1,M_t} X_{t+1} + g_{t+1,M_t}$$

Then the optimal cost to go at time t can be written as:

$$\mathcal{J}_t^*(X_0^t, M_0^{t-1}) = \min_{\mathcal{U}\in\mathfrak{U}^{LQR^1}} E[X_t^T Q X_t + U_t^T R U_t + $$
$$+ \mathcal{J}_{t+1}^{*,LQR^1}(X_0^{t+1}, M_0^t) | X_0^t, M_0^{t-1}]$$

or,

$$\mathcal{J}_t^{*,LQR^1}(X_0^t, M_0^{t-1}) = \min_{\mathcal{U}\in\mathfrak{U}^{LQR^1}} E[X_t^T Q X_t + U_t^T R U_t + $$
$$+ X_{t+1}^T P_{t+1,M_t} X_{t+1} + g_{t+1,M_t} | X_0^t, M_0^{t-1}]$$

By further replacing X_{t+1} from equation (1) we get:

$$\mathcal{J}_t^{*,LQR^1}(X_0^t, M_0^{t-1}) = \min_{\mathcal{U}\in\mathfrak{U}^{LQR^1}} E[X_t^T Q X_t + U_t^T R U_t +$$
$$+ (A_{M_t} X_t + B_{M_t} U_t + W_t)^T P_{t+1,M_t}(A_{M_t} X_t + B_{M_t} U_t + W_t) +$$
$$+ g_{t+1,M_t} | X_0^t, M_0^{t-1}] = \min_{\mathcal{U}\in\mathfrak{U}^{LQR^1}} X_t^T E[A_{M_t}^T P_{t,M_t} A_{M_t} +$$
$$+ Q | X_0^t, M_0^{t-1}] X_t + 2X_t^T E[A_{M_t}^T P_{t,M_t} B_{M_t} | X_0^t, M_0^{t-1}] U_t +$$
$$+ U_t^T E[B_{M_t}^T P_{t,M_t} B_{M_t} + R | X_0^t, M_0^{t-1}] U_t +$$
$$+ E[trace(P_{t+1,M_t}\Sigma_W) + g_{t+1,M_t} | X_0^t, M_0^{t-1}]$$

Minimizing in terms of U_t (by a square completion technique) we obtain the results (22)-(10). Expressions (12)-(10) are calculated by observing that $p(M_t = i | X_0^t = x_0^t, M_0^{t-1} = m_0^{t-1}) = p(M_t = i | M_{t-1} = m_{t-1})$. Moreover the optimal control that minimizes the cost to go function at time t is given by:

$$U_t^* = -\mathcal{K}_{t,M_{t-1}} X_t \tag{23}$$

where

$$\mathcal{K}_{t,j} = \mathcal{B}_{t,j}^{-1}\mathcal{C}_{t,j}^T \tag{24}$$

□

Proof of Theorem 1

Using the Dynamic Programming algorithm (Bertsekas [1995], Whittle [1982]), the optimal cost $\mathcal{J}_N^*(X_0, M_0)$ is given by of the following algorithm which proceeds backwards in time

$$J_N(X_0^N, M_0^N) = X_N^T Q X_N$$

$$J_t(X_0^t, M_0^{t-1}) = \min_{\mathcal{U}\in\mathfrak{U}^{LQR^1}} E[X_t^T Q X_t + U_t^T R U_t + \\ + \min_{\mathcal{U}\in\mathfrak{U}^{LQR^1}} J_{t+1}(X_0^{t+1}, M_0^t)|X_0^t, M_0^{t-1}]$$

for $t = 0, 1, \ldots, N-1$. Furthermore, the control sequence U_t^*, $t = 0, 1, \ldots, N-1$ is the optimal solution for the LQR^1 problem where each U_t^* minimizes the right side of the above equation. We conclude the proof by invoking the results of *Lemma 4*.

□

Proof of Theorem 2

Is is immediate that if the CARE in (15) does not have a solution we can not obtain a solution for the optimal control problem. Let us assume that (15) does have a solution denoted by $P_{\infty,i}$, for $1 \leq i \leq s$. By using the control expressed in (19) it is not difficult to show that the closed loop system is stable. Indeed since

$$E[X_{t+1}^T P_{\infty,M_t} X_{t+1}] - E[X_t^T P_{\infty,M_{t-1}} X_t] < \infty$$

we obtain also that

$$E[X_t^T X_t] < \infty$$

which implies stability of the system.

Then since the system is stable $\tilde{J}_{av}(X_0, M_0) < \infty$ where $\tilde{J}_{av}(X_0, M_0)$ is a long run average cost function as in (5), with the control input chosen to be as in (19). Notice however that for every $N > 0$ the following inequality holds

$$\mathcal{J}_N^*(X_0, M_0) \leq \tilde{J}_N(X_0, M_0)$$

because $\mathcal{J}_N^*$ is the optimal cost. But since the optimal cost is also an increasing quantity as a function of N and it is upper bounded by $\tilde{J}_N$ we can conclude that $\mathcal{J}_{av}(X_0, M_0)$ is a bounded quantity. Therefore the matrices $P_{t,i}$ must converge and moreover they converge to $P_{\infty,i}$ the solution of (15). The second and third part of the theorem follow from part one together with the results of *Theorem 1* and from the ergodicity property of the Markov chain M_t.

□

In the following we establish a link between a LMI optimization problem and the (maximal) solution to the CARE (15). This approach lead us to an efficient way to compute the matrices $P_{\infty,i}$ (if they exist). Original results showing how solutions of CARE can be obtained as solutions of convex programming problem can be found in Rami [1996, 1994].

Consider the following optimization problem.

$$\max\ trace(\sum_{i=1}^{s} \mathcal{P}_{\infty,s}) \tag{25}$$

subject to

$$\begin{pmatrix} -\mathcal{P}_{\infty,j} + \sum_{i=1}^{s} p_{ji} A_i^T \mathcal{P}_{\infty,i} A_i + Q & \sum_{i=1}^{s} p_{ji} A_i^T \mathcal{P}_{\infty,i} B_i \\ \sum_{i=1}^{s} p_{ji} B_i^T \mathcal{P}_{\infty,i} A_i & \sum_{i=1}^{s} p_{ji} B_i^T \mathcal{P}_{\infty,i} B_i + R \end{pmatrix} \geq 0 \tag{26}$$

$$\sum_{i=1}^{s} p_{ji} B_i^T \mathcal{P}_{\infty,i} B_i + R > 0 \tag{27}$$

$$\mathcal{P}_{\infty,i} = \mathcal{P}_{\infty,i}^T$$

with $1 \leq i \leq s$.

Suppose the above optimization problem has solution. Then the above optimization problem has solution if and only if a solution for (15) can be found. A proof of this claim can be mimicked after similar proof in Rami [1996]. The key idea comes from the fact that from the Schur complement Lemma, $\mathcal{P}_{\infty,i}$ satisfies (26-27) if and only if

$$-\mathcal{P}_{\infty,j} + \left(\sum_{i=1}^{s} p_{ji} A_i^T \mathcal{P}_{\infty,i} A_i + Q\right) - \left(\sum_{i=1}^{s} p_{ji} A_i^T \mathcal{P}_{\infty,i} B_i\right) \\ \left(\sum_{i=1}^{s} p_{ji} B_i^T \mathcal{P}_{\infty,i} B_i + R\right)^{-1} \left(\sum_{i=1}^{s} p_{ji} B_i^T \mathcal{P}_{\infty,i} A_i\right) \geq 0$$

We know turn the statement of *Proposition 3* regarding the more general LQR^τ problem. The following lemma expresses the probability of the current mode M_t to be in a certain state given the entire history of the state and a τ steps delayed history of the mode process. This lemma will be instrumental in proving the claim of *Proposition 3.*

Lemma 5. Within the context of *Definition 1* and *Problem 1*, the conditional probability distribution of the mode at time t given the history of the state X_0^t and a τ steps ($\tau > 1$) delayed history of the mode $M_0^{t-\tau}$ is given by:

$$pr(M_t = i|X_0^t = x_0^t, M_0^{t-\tau} = m_0^{t-\tau}) = \\ = \frac{\sum_{i_1^{\tau-1}=1}^{s} \prod_{k=0}^{\tau-1} p_{i_{\tau-k-1} i_{\tau-k}} f_W(\xi_k)}{\sum_{i_0^{\tau}=1}^{s} \prod_{k=0}^{\tau-1} p_{i_{\tau-k-1} i_{\tau-k}} f_W(\xi_k)} \tag{28}$$

where $\xi_k = x_{t-k} - A_{i_{\tau-k-1}} x_{t-k-1} - B_{i_{\tau-k-1}} u_{t-k-1}$, $f_W(\cdot)$ is the probability density function of the noise process W_t and by convention, in the denominator, we consider $i_\tau = i$ and $i_0 = m_{t-\tau}$ (for both the numerator and denominator).

Proof: Before presenting the proof we should notice that conditioning on X_0^t and $M_0^{t-\tau}$ implies conditioning on U_0^{t-1} as well, because of the chosen control policy. Hence we can write:

$$pr(M_t = i|X_0^t = x_0^t, M_0^{t-\tau} = m_0^{t-\tau}) = \\ = pr(M_t = i|X_0^t = x_0^t, M_0^{t-\tau} = m_0^{t-\tau}, U_0^{t-1} = u_0^{t-1})$$

The proof is based on the following recurrence:

$$pr(X_0^t = x_0^t, M_0^t = m_0^t) = p_{m_{t-1} m_t} \cdot f_W(\xi_t) \\ \cdot pr(X_0^{t-1} = x_0^{t-1}, M_0^{t-1} = m_0^{t-1}) \tag{29}$$

where $\xi_t = x_t - A_{m_{t-1}} x_{t-1} - B_{m_{t-1}} u_{t-1}$.

Using Bayes rule, the probability (28) can be expressed as:

$$pr(M_t = i|X_0^t = x_0^t, M_0^{t-\tau} = m_0^{t-\tau}) =$$
$$= \frac{pr(M_t = i, X_0^t = x_0^t, M_0^{t-\tau} = m_0^{t-\tau})}{pr(X_0^t = x_0^t, M_0^{t-\tau} = m_0^{t-\tau})}$$

with

$$pr(M_t = i, X_0^t = x_0^t, M_0^{t-\tau} = m_0^{t-\tau}) =$$
$$\sum_{i_1^{\tau-1}=1}^{s} pr(M_t = i, M_{t-\tau+1}^{t-1} = i_1^{\tau-1}, X_0^t = x_0^t, M_0^{t-\tau} = m_0^{t-\tau})$$

and

$$pr(X_0^t = x_0^t, M_0^{t-\tau} = m_0^{t-\tau}) =$$
$$\sum_{i_1^{\tau}=1}^{s} pr(M_{t-\tau+1}^{t} = i_1^{\tau}, X_0^t = x_0^t, M_0^{t-\tau} = m_0^{t-\tau})$$

Applying the recurrence (29), we obtain (28).

□

Remark 3. We can notices from *Lemma 5* that $pr(M_t = i|X_0^t = x_0^t, M_0^{t-\tau} = m_0^{t-\tau})$ depends nonlinearly on the sequence of state variable $X_{t-\tau}^t$, control input $U_{t-\tau-1}^{t-1}$. The nonlinearity is induced by the fact that the fore-mentioned sequences determine the value of function $f_W(\xi_k)$, a Gaussian p.d.f. of the noise W_k and thus nonlinear.

Let us know address the statement of *Proposition 3* and show that in the case of delays greater then two in mode observations, the optimal control is non-linear in the state vectors and previous control inputs.

Assume that $\tau = 2$, and let us apply the dynamic programming algorithm to solve the LQR^τ problem. Obviously, the cost to go function for $t = N$ is

$$\mathfrak{J}_N^{*,LQR^2}(X_0^N, M_0^{N-2}) = X_N^T P_N X_N,\ P_N = Q.$$

At time $t = N - 1$ the cost to go function becomes

$$\mathfrak{J}_{N-1}^{*,LQR^2}(X_0^{N-1}, M_0^{N-3}) =$$
$$= \min_{U_{N-1}\in\mathcal{U}} E[X_{N-1^T} Q X_{N-1} + U_{N-1}^T R U_{N-1} +$$
$$+ X_N^T P_N X_N | X_0^{N-1}, M_0^{N-3}]$$

Or by replacing X_N we further get:

$$\mathfrak{J}_{N-1}^{*,LQR^2}(X_0^{N-1}, M_0^{N-3}) =$$
$$= \min_{U_{N-1}\in\mathcal{U}} \{X_{N-1}^T \mathcal{A}_{N-1} X_{N-1} + U_{N-1}^T \mathcal{B}_{N-1} U_{N-1} +$$
$$+ 2X_{N-1}^T \mathcal{C}_{N-1} U_{N-1} + trace(P_N \Sigma_W)\}$$

where

$$\mathcal{A}_{N-1} = E[A_{M_{N-1}}^T P_N A_{M_{N-1}} | X_0^{N-1}, M_0^{N-3}] + Q$$
$$\mathcal{B}_{N-1} = E[B_{M_{N-1}}^T P_N B_{M_{N-1}} | X_0^{N-1}, M_0^{N-3}] + R$$
$$\mathcal{C}_{N-1} = E[A_{M_{N-1}}^T P_N B_{M_{N-1}} | X_0^{N-1}, M_0^{N-3}]$$

An explicit formula for $\mathcal{A}_{N-1}$ is

$$\mathcal{A}_{N-1} = \sum_{i=1}^{s} pr(M_{N-1} = i|X_0^{N-1}, M_0^{N-3}) A_i^T P_N A_i$$

Notice that by (28) the term $\mathcal{A}_{N-1}$ will depend nonlinearly on X_{N-1}, X_{N-2}, X_{N-3} U_{N-2}, U_{N-3} and M_{N_3}. Obviously the same is true for $\mathcal{B}_{N-1}$ and $\mathcal{C}_{N-1}$.

Then the optimal cost to go at time $t = N - 1$ can be expressed as:

$$\mathfrak{J}_{N-1}^{*,LQR^1}(X_0^{N-1}, M_0^{N-3}) = X_{N-1}^T P_{N-1,M_{N-3}} X_{N-3}$$
$$+ g_{N-1,M_{N-3}}$$

where the matrix $P_{N-1,M_{N-3}}$ is computed by an expression similar to the one in (11) and it will also be dependent nonlinearly on X_{N-3}^{N-1}, U_{N-3}^{N-2} and M_{N-3} respectively. The same is true for the optimal control U_{N-1}^* since at $t = N-1$ it is expressible by a formula similar to (7).

The optimal cost at $t = N - 2$ is

$$\mathfrak{J}_{N-2}^{*,LQR^2}(X_0^{N-2}, M_0^{N-4}) =$$
$$= \min_{U_{N-2}\in\mathcal{U}} E[X_{N-2^T} Q X_{N-2} + U_{N-2}^T R U_{N-2} +$$
$$+ X_{N-1}^T P_{N-1,M_{N-3}} X_{N-1} | X_0^{N-2}, M_0^{N-4}]$$

Notice that at this stage the cost function is no longer quadratic in U_{N-2}. This is due to the fact the matrix $P_{N-1,M_{N-3}}$ from the previous step depends on U_{N-2} as well. Hence the optimal cost to go function can be generically written as:

$$\mathfrak{J}_{N-2}^{*,LQR^2}(X_0^{N-2}, M_0^{N-4}) =$$
$$= \min_{U_{N-2}\in\mathcal{U}} g(X_{N-4}^{N-2}, U_{N-4}^{N-2}, M_{N-4})$$

where $g(\cdot)$ is cost function which is no longer quadratically in U_{N-2}. Although we were not able to obtain an analytical expression for the optimal control it should be clear that since exponential terms depending on U_{N-2} are contained by $g(\cdot)$ the optimal control is not going to have a linear expression as in the standard LQR problem.

5. CONCLUSIONS

In this paper we solved the optimal Linear Quadratic Regulator paradigm for Markovian Jump linear System with one step delayed mode observations. This setup is useful is networked control applications which rely on acknowledgments not received at the controller instantaneously. The optimal control policy was shown to be a linear feedback of the current state with a gain dependent on the delayed mode. The optimal solution was provided in terms of a set of coupled Riccati equations. For the infinite horizon case the solution of the associated CARE was shown to be obtained as a solution of LMI optimization problem. An analysis of the LQR problem in case of arbitrarily delayed mode observations is presented. We demonstrate that the optimal policy depends nonlinearly on a sequence of state observations and control inputs whose length depends on the value of the delay. As future work, we intend to address the optimal Linear Quadratic Gaussian Regulator paradigm for Markovian Jump Linear Systems with delayed mode observations.

REFERENCES

H. Abou-Kandil, G. Freiling and G. Jank. Solution and asymptotic behavior of coupled Riccati equations in jump linear systems. *IEEE Trans. Automatic Control* 39, 1994, 1631-1636.

D.P. Bertsekas. Dynamic Programming and Optimal Control. *Athena Scientific*, 1995.

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.

E.K. Boukas, A. Swierniak and H. Yang. On the robustness of jump linear quadratic control. *International Journal of Robust and Nonlinear Control*, Vol 7, Issue 10, pages 899-910, 1998.

E.K. Boukas and H. Yang. Robust LQ Regulator for Jump Linear Systems with Uncertain Parameters. *Dynamics and Control*, Springer, Vol 9, No 2, pages 125-134, 1999.

H.J. Chizeck, A.S. Willsky and D. Castanon. Discrete-time Markovian-jump linear quadratic optimal control. *Intl. J. Control*, Vol 43, No 1, pages 213-231, 1986.

O.L.V. Costa, M.D. Fragoso and R.P. Marques. Discrete-Time Markovian Jump Linear Systems. *Springer*, 2005.

Y. Ji and H.J. Chizeck. Controllability, observability and continuous-time Markovian jump linear quadratic control. *IEEE Trans. Automat. Control* 35 (7), 1990, pages 777-788.

Y. Ji, H.J. Chizeck, X. Feng and K.A. Loparo. Stability and Control of discrete-time jump linear systems. *Control Theory Adv. Technol.*, Vol 7, No 2, pages 247-270, 1991.

M. Ait-Rami and L. El Ghaoui. LMI Optimization for Nonstandard Riccati Equations Arising in stochastic Control. *IEEE Transactions on Automatic Control*, Vol 41, Issue 11, pages 1666-1671, 1996.

M. Ait-Rami and L. El Ghaoui. Robust stabilization of jump linear systems using linear matrix inequalities. In *IFAC Symposium on Robust Control Design*, pages 148-151,1994.

P. Whittle Optimization over time, Volume I. *Wiley*, 1982.