2021 60th IEEE Conference on Decision and Control (CDC)
December 13-15, 2021. Austin, Texas

2021 60th IEEE Conference on Decision and Control (CDC) | 978-1-6654-3659-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/CDC45484.2021.9683675

Linear Quadratic Tracking Control of Hidden Markov Jump Linear Systems Subject to Ambiguity

Ioannis Tzortzis, Christoforos N. Hadjicostis, and Charalambos D. Charalambous

***Abstract*— The linear quadratic tracking control problem is studied for a class of discrete-time uncertain Markov jump linear systems with time-varying conditional distributions. The controller is designed under the assumption that it has no access to the true states of the Markov chain, but rather it depends on the Markov chain state estimates. To deal with uncertainty, the transition probabilities of Markov state estimates between the different operating modes of the system are considered to belong in an ambiguity set of some nominal transition probabilities. The estimation problem is solved via the one-step forward Viterbi algorithm, while the stochastic control problem is solved via minimax optimization theory. An optimal control policy with some desired robustness properties is designed, and a maximizing time-varying transition probability distribution is obtained. A numerical example is given to illustrate the applicability and effectiveness of the proposed approach.**

I. Introduction

Markov jump linear systems (MJLS) form a class of linear dynamical systems in which the jumps between the system's different operating modes are modulated according to an underlying Markov chain (MC). Such systems have received increased attention due to their ability to characterize and model a variety of practical situations such as tracking control systems, fault-tolerant control systems, networked control systems, and others [1]–[3]. In practice, the designer of MJLS faces two different issues. The first issue concerns the information structure available to the controller, that is, whether or not the system states (both Markov and linear states) are accessible to the controller. This accessibility is a problem of state estimation [4]–[7]. The second issue concerns uncertainty (both model and Markov uncertainty). This uncertainty is a problem of robust control [8]–[13]. When not dealt with, both issues might lead to inconsistencies, and compromise the utility of the whole system.

In this paper, our focus is on linear quadratic (LQ) tracking control problems for the scenario in which: (i) the Markov states are not accessible to the controller, and (ii) the underlying MC is uncertain. The main objective is to develop an LQ optimization approach that applies to hidden MJLS with time-varying conditional distributions, and that is capable of reducing the influence of uncertainty on the performance of the optimal controller. To achieve this both the problems of estimation and control are addressed. In particular, the state estimation problem is addressed by employing the one-step forward Viterbi algorithm, which is implemented as soon as a new observation of the state is obtained. On the other hand, the robust control problem is addressed via minimax optimization using a dynamic programming approach. To capture and reduce the influence of uncertainty, an ambiguity set formed with all probability distributions within a total variation distance from a nominal estimate is considered.

This work was partially funded by the European Regional Development Fund and the Republic of Cyprus through the Cyprus Research and Innovation Foundation (Project: POST-DOC/0916/0139).

The authors are with the Department of Electrical and Computer Engineering, University of Cyprus, Nicosia, Cyprus. E-mails: {tzortzis.ioannis,chadjic,chadcha}@ucy.ac.cy

The rest of the paper is organised as follows. In Section II the LQ tracking problem is formulated, and the one-step forward Viterbi algorithm is introduced. In Section III the solution of the problem is derived, along with a maximizing, time-varying, conditional distribution. In Section IV the implementation of the proposed approach is discussed and an algorithm is proposed. In Section V an example is solved to illustrate the effectiveness of the proposed approach. Finally, in Section VI we give some concluding remarks.

II. Problem Formulation

A. Model description and the Viterbi algorithm

Consider a discrete-time control system with linear state dynamics defined by

$$x_{k+1} = A_k(\theta_k)x_k + B_k(\theta_k)u_k, \quad x_1 = x \tag{1}$$

and modulated by a finite-state non-homogeneous MC $\{\theta_k, k \geq 1\} \in \Theta$. Here, $x_k \in \mathcal{X} \triangleq \mathbb{R}^n$ and $u_k \in \mathcal{U} \triangleq \mathbb{R}^m$ are the state and control processes, respectively. It is assumed that: (1) we have complete observations of the system's state, and (2) the dynamics matrices $A_k(\theta_k) \in \mathbb{R}^{n\times n}$ and the input matrices $B_k(\theta_k) \in \mathbb{R}^{m\times m}$ are known for each $\theta_k \in \Theta$ with bounded and measurable entries for each $k = 1, 2, \ldots, N-1$. Furthermore, we consider the case where the state of the MC $\{\theta_k, k \geq 1\}$ is not observable, and hence, both the control and the controlled processes will depend on the state estimate of the MC. To this end, next we introduce the definition of a hidden Markov model, and an estimation algorithm which will be used to estimate the state of the MC. A systematic review of hidden Markov models along with their applications can be found in [14].

Consider a finite-state, non-homogeneous, first-order, hidden Markov model (HMM) specified by the following components: (i) A set of hidden states $\{\theta_k, k \geq 1\}$, taking values from a finite alphabet set Θ with cardinality $|\Theta| = n_\theta$, (ii) a set of observed variables $\{z_k, k \geq 1\}$, taking values from a finite alphabet set Z with cardinality $|Z| = n_z$, (iii) a time-varying transition probability distribution

$$\begin{aligned} q_{ij}(k) &\triangleq P_k(\theta_{k+1} = j|\theta_1, \ldots, \theta_k = i) \\ &= P_k(\theta_{k+1} = j|\theta_k = i), \quad i, j \in \Theta \end{aligned}$$

where $q_{ij}(k) \geq 0$ represents the probability of moving from state i at time k to state j at time $k+1$ such that $\sum_{j \in \Theta} q_{ij}(k) = 1$, $\forall i \in \Theta$, (iv) a time-varying output/emission probability distribution

$$\begin{aligned} e_{is}(k) &\triangleq P_k(z_k = s|\theta_1, \ldots, \theta_k = i, z_1, \ldots, z_{k-1}) \\ &= P_k(z_k = s|\theta_k = i), \quad s \in Z,\ i \in \Theta \end{aligned}$$

where $e_{is}(\cdot)$ represents the probability that $z_k = s$ is observed at time k given the state $\theta_k = i$, and (v) an initial probability distribution $\pi(i) = P(\theta_1 = i)$, $i \in \Theta$.

For a hidden Markov model and an observation sequence $z_{1:N} \triangleq \{z_1, z_2, \ldots, z_N\}$, we wish to compute the most probable sequence of hidden states, that is,

$$\theta^*_{1:N} = \arg\max_{\theta_{1:N}} P(\theta_{1:N}|z_{1:N}) \tag{2}$$

or, equivalently, $\theta^*_{1:N} = \arg\max_{\theta_{1:N}} P(\theta_{1:N}, z_{1:N})$, where $P(\theta_{1:N}, z_{1:N}) = P(\theta_{1:N}|z_{1:N})P(z_{1:N})$, and $\theta_{1:N} \triangleq \{\theta_1, \theta_2, \ldots, \theta_N\}$. The solution of the above problem is obtained by defining the recursions

$$\begin{aligned} \mu_{k+1}(\theta_{k+1}) &\triangleq \max_{\theta_{1:k}} P(\theta_{1:k+1}, z_{1:k+1}) \\ &= \max_{\theta_k} P(z_{k+1}|\theta_{k+1})P(\theta_{k+1}|\theta_k)\mu_k(\theta_k) \quad \text{(3a)} \\ \mu_1(\theta_1) &\triangleq P(\theta_1, z_1) = P(z_1|\theta_1)P(\theta_1) \quad \text{(3b)} \end{aligned}$$

for $k = 1, 2, \ldots, N-1$. Then, the total maximizing probability is obtained by performing one last maximization, that is, $\max_{\theta_N} \mu_N(\theta_N) \triangleq \max_{\theta_{1:N}} P(\theta_{1:N}, z_{1:N})$, and the sequence of hidden states $\theta^*_{1:N}$ is obtained by simply backtracing the sequence which led to the total maximizing probability. The above algorithm, which is referred to as the Viterbi algorithm [15], is implemented as soon as a new observation of the hidden state is obtained.

The state estimate $\hat{\theta}_t$ for each time-step $t = 1, 2, \ldots, N$, is obtained by defining

$$\hat{\theta}_t \triangleq \arg\max_{\theta_t} \mu_t(\theta_t). \tag{4}$$

By backtracing the sequence of states which led to $\hat{\theta}_t$ then the most probable sequence of hidden states given the observations up to and including time t can be obtained, and will be denoted by $\theta^{*,t}_{1:t} = \{\theta^{*,t}_1, \ldots, \theta^{*,t}_t\}$, $t = 1, 2, \ldots, N$. Note that, although $\theta^{*,t}_s = \hat{\theta}_s$ for $s = t$, it is not necessarily true that $\theta^{*,t}_s = \hat{\theta}_s$ for $s \neq t$. In Section IV, we further discuss the implementation of the estimation algorithm in practice. For the rest of the paper, we will refer to the above procedure as the one-step forward Viterbi algorithm.

B. Set of nominal control policies

Assume that for the construction of u_k at any time $k = 1, 2, \ldots, N-1$, the controller has complete state information about x_k, and (using the one-step forward Viterbi algorithm) it has available information about $\hat{\theta}_k \in \Theta$. The set G of control policies is the set of measurable Markov feedback control policies $g : \mathcal{X} \times \Theta \longmapsto \mathcal{U}$, $g \in G$. Then, associated with the open-loop system (1), and the control policy $g \in G$, the nominal closed-loop system is given by the recursion

$$x^g_{k+1} = A_k(\hat{\theta}_k)x^g_k + B_k(\hat{\theta}_k)g_k(x_k, \hat{\theta}_k). \tag{5}$$

The control process is defined by $u^g_k \triangleq g_k(x_k, \hat{\theta}_k)$, where $u^g \in \mathcal{U}_{[1,N-1]}$ with $\mathcal{U}_{[1,N-1]} = \{u^g_k \in \mathcal{U} : \mathbb{E}\sum_{k=1}^{N-1} |u_k|^2 < \infty\}$. For simplicity, the dependence of the controlled process x^g, and the control process u^g, on the policy $g \in G$ will not be explicitly captured in our notation for the rest of the paper.

As an approximation of the true transition probability distribution $P_k(\hat{\theta}_{k+1} = j|\hat{\theta}_1, \ldots, \hat{\theta}_k = i)$ of the estimator, we consider a Markov, time-invariant transition probability distribution

$$\begin{aligned} p^0_{ij} &\triangleq P(\hat{\theta}_{k+1} = j|\hat{\theta}_1, \ldots, \hat{\theta}_k = i) \\ &= P(\hat{\theta}_{k+1} = j|\hat{\theta}_k = i), \quad i, j \in \Theta. \end{aligned} \tag{6}$$

We will refer to p^0_{ij} as the "*nominal*" transition probability distribution. Given $\hat{\theta}_{1:N} \triangleq \{\hat{\theta}_1, \ldots, \hat{\theta}_N\}$, the maximum likelihood estimates for the nominal transition probability distribution [16], are

$$p^0_{ij} = \frac{\sum_{t=2}^N n_{ij}(t)}{\sum_{m=1}^{n_\theta}\sum_{t=2}^N n_{im}(t)}, \quad i, j \in \Theta \tag{7}$$

where $n_{ij}(t)$ denotes the number of occurrences in state $\hat{\theta}_{t-1} = i$ at time $t-1$ and state $\hat{\theta}_t = j$ at time t.

C. Ambiguity class

Any uncertainty on the conditional distribution of the underlying MC, either on the transition probability distribution or the emission probability distribution or in both, will typically affect the MC state estimate and, consequently, the optimality of the controller. To reduce the influence of uncertainty on the performance of the optimal controller, the proposed method relies on the availability of a nominal transition probability distribution, as defined by (6). In particular, the ambiguity set is formed with all time-varying, transition probability distributions whose distance from the nominal distribution is within a radius R_{TV}. Parameter $R_{TV} \in [0, 2]$ controls the size of the ambiguity set, and its value is preselected by the decision maker based on its own belief on the nominal probability distribution.

Toward this end, let us define the set of transition probability distributions on Θ by

$$\mathbb{P}_k(\Theta|i) \triangleq \{p_{i\bullet}(k) : p_{ij}(k) \geq 0, \forall j \in \Theta, \sum_{j \in \Theta} p_{ij}(k) = 1\}$$

for all $i \in \Theta$, and $k = 1, 2, \ldots, N-1$. The ambiguity set of all possible, time-varying, transition probability distributions is defined by

$$\begin{aligned} \mathbb{B}(i,k) \triangleq \{&p_{i\bullet}(k) \in \mathbb{P}_k(\Theta|i) : \\ &\sum_{j \in \Theta} |p_{ij}(k) - p^0_{ij}| \leq R_{TV}(i,k)\} \end{aligned} \tag{8}$$

where $R_{TV}(i,k) \in [0, 2]$, $i \in \Theta$, $k = 1, 2, \ldots, N-1$. In addition, we define

$$\mathbb{B}(k) \triangleq \{p_{i\bullet}(k) \in \mathbb{B}(i,k), i = 1, 2, \ldots, |\Theta|\}. \tag{9}$$

We note that, (8) allows different levels of ambiguity to be assigned between different operating modes of the system, a feature which can be proved useful in many applications.

D. Optimal stochastic control problem

Define the N-stage expected cost by

$$J_N(g,p) \triangleq \mathbb{E}^{g,p}_{x,i}\big[(x_N-\bar{x}_N)^T Q_N(\hat{\theta}_N)(x_N-\bar{x}_N) + \sum_{k=1}^{N-1}\Big((x_k-\bar{x}_k)^T Q_k(\hat{\theta}_k)(x_k-\bar{x}_k)+u_k^T R_k(\hat{\theta}_k)u_k\Big)\big]$$

where $\bar{x}_k$ denotes the desired or reference value of the state vector, $\mathbb{E}^{g,p}_{x,i}[\cdot]$ indicates the dependence of the expectation operation on feedback Markov control policy $g \in G$, and induced by a transition probability distribution $p_{i\bullet}(k) \in \mathbb{B}(i,k)$, for fixed initial data $x_1 = x$ and $\hat{\theta}_1 = i$. Here, $Q_k(\cdot) \succeq 0$ are $n \times n$ real symmetric positive semi-definite matrices, and $R_k(\cdot) \succ 0$ are $m \times m$ real symmetric positive definite matrices.

Minimax stochastic control problem: Find an optimal control policy $g^* \in G$, and a maximizing transition probability distribution $p^*_{i\bullet}(k) \in \mathbb{B}(i,k)$, that causes the closed-loop system (5) to follow a reference trajectory signal, by solving the optimization problem

$$J^* = J_N(g^*,p^*) \triangleq \min_{g\in G} \max_{\substack{p(k)\in\mathbb{B}(k)\\ k=1,2,\dots,N}} J_N(g,p). \tag{10}$$

In the next section the solution of the minimax stochastic control problem is provided.

III. Problem Solution

A. Dynamic programming

For $(k,x,i) \in \{1,2,\dots,N\}\times\mathcal{X}\times\Theta$ let $V_k(x,i)$ denote the minimal cost-to-go or value function on the time horizon $\{k,k+1,\dots,N\}$, given an optimal policy $g_t^*(\cdot)$, $t=1,2,\dots,k-1$, and worst case transition probability distribution $p^*_{i\bullet}(t)\in\mathbb{B}(i,t)$, $t=1,2,\dots,k-1$, defined by

$$V_k(x,i) = \min_{u\in\mathcal{U}_{[k,N-1]}} \max_{\substack{p(t)\in\mathbb{B}(t)\\ t=k,k+1,\dots,N}} \mathbb{E}^{g,p}_{x,i}\Big[\sum_{t=k}^{N-1}\big((x_t-\bar{x}_t)^T Q_t(\hat{\theta}_t)(x_t-\bar{x}_t)+u_t^T R_t(\hat{\theta}_t)u_t\big) + (x_N-\bar{x}_N)^T Q_N(\hat{\theta}_N)(x_N-\bar{x}_N)\Big] \tag{11}$$

where $\mathbb{E}^{g,p}_{x,i}[\cdot]$ denotes conditional expectation given that $x_k^g = x$ and $\hat{\theta}_k = i$ with x,i fixed. The dynamic programming algorithm gives [17]

$$V_N(x_N,\hat{\theta}_N) = (x_N-\bar{x}_N)^T Q_N(\hat{\theta}_N)(x_N-\bar{x}_N) \tag{12a}$$

$$V_k(x,i) = \min_{u_k\in\mathcal{U}} \max_{p_{i\bullet}(k)\in\mathbb{B}(i,k)} \mathbb{E}^{g,p}_{x,i}\big[(x-\bar{x}_k)^T Q_k(i)(x-\bar{x}_k) + u_k^T R_k(i)u_k + V_{k+1}(x_{k+1},\hat{\theta}_{k+1})\big], \quad x\in\mathcal{X}. \tag{12b}$$

Dynamic programming equation (12b) generates the value function $V_k(\cdot)$ for all $k=N-1,N-2,\dots,1$ by backward recursion. Next, we will show by backward induction that the solution is of the following form

$$V_k(x,i) = x^T P_k(i)x + x^T f_k(i) + r_k(i) \tag{13}$$

for $k=1,\dots,N$, $\hat{\theta}_k = i\in\Theta$, $f_k(i)\in\mathbb{R}^n$, $r_k(i)\in\mathbb{R}$, and some matrices $P_k(i)\succeq 0$.

For $k=N$, the induction hypothesis in (13) is true with $P_N(i)=Q_N(i)$, $f_N(i)=-2Q_N(i)\bar{x}_N$, and $r_N(i)=\bar{x}_N^T Q_N(i)\bar{x}_N$, $i\in\Theta$. Then, $P_N(i)=P_N(i)^T\succeq 0$ and $V_N(x,i)=x^TP_N(i)x+x^Tf_N(i)+r_N(i)$. Suppose that for $t=k+1,k+2,\dots,N$, $P_t(i)=P_t(i)^T\succeq 0$ and $V_t(x,i)=x^TP_t(i)x+x^Tf_t(i)+r_t(i)$. It will be shown that then $V_k(x,i)=x^TP_k(i)x+x^Tf_k(i)+r_k(i)$, where $P_k(i)=P_k(i)^T\succeq 0$. To this end, we re-write (12b) as follows

$$V_k(x,i) = \min_{u_k\in\mathcal{U}}\Big\{(x-\bar{x}_k)^TQ_k(i)(x-\bar{x}_k)+u_k^TR_k(i)u_k + \max_{p_{i\bullet}\in\mathbb{B}(i,k)}\sum_{j\in\Theta}\Big(\int_{\mathcal{X}_{k+1}}V_{k+1}(x_{k+1},j)P^g(x_{k+1}|x,i)\Big)p_{ij}(k)\Big\} \tag{14}$$

and we define the sequence

$$\begin{aligned}\ell_k(x_k,\hat{\theta}_k,\hat{\theta}_{k+1},u_k) &\triangleq \int_{\mathcal{X}_{k+1}}V_{k+1}(x_{k+1},\hat{\theta}_{k+1})P^g(x_{k+1}|x_k,\hat{\theta}_k)\\ &\overset{(a)}{=}\Big(A_k(\hat{\theta}_k)x_k+B_k(\hat{\theta}_k)u_k\Big)^TP_{k+1}(\hat{\theta}_{k+1})\Big(A_k(\hat{\theta}_k)x_k + B_k(\hat{\theta}_k)u_k\Big)+\Big(A_k(\hat{\theta}_k)x_k+B_k(\hat{\theta}_k)u_k\Big)^Tf_{k+1}(\hat{\theta}_{k+1})\\ &\quad + r_{k+1}(\hat{\theta}_{k+1}),\end{aligned} \tag{15}$$

where (a) follows by the induction hypothesis and (5). Then, (14) becomes

$$V_k(x,i) = \min_{u_k\in\mathcal{U}}\Big((x-\bar{x}_k)^TQ_k(i)(x-\bar{x}_k)+u_k^TR_k(i)u_k + \max_{p_{i\bullet}\in\mathbb{B}(i,k)}\sum_{j\in\Theta}\ell_k(x,i,j,u_k)p_{ij}(k)\Big), \tag{16}$$

where $p_{ij}(k)\triangleq P(\hat{\theta}_{k+1}=j|\hat{\theta}_k=i)$. Next, we provide the solution of the inner optimization in (16).

B. Worst case transition probability distribution

First, let us define the maximum and minimum values of (15) with respect to $\hat{\theta}_{k+1}\in\Theta$, by

$$\ell_{\max,k}(x_k,\hat{\theta}_k,u_k) \triangleq \max_{\hat{\theta}_{k+1}\in\Theta}\ell_k(x_k,\hat{\theta}_k,\hat{\theta}_{k+1},u_k) \tag{17}$$

$$\ell_{\min,k}(x_k,\hat{\theta}_k,u_k) \triangleq \min_{\hat{\theta}_{k+1}\in\Theta}\ell_k(x_k,\hat{\theta}_k,\hat{\theta}_{k+1},u_k) \tag{18}$$

and its corresponding sets of states that achieve the maximum and minimum values of $\ell_k(x_k,\hat{\theta}_k,\hat{\theta}_{k+1},u_k)$ by

$$\Theta^0(k,\hat{\theta}_k) \triangleq \{\hat{\theta}_{k+1}\in\Theta : \ell_k(x_k,\hat{\theta}_k,\hat{\theta}_{k+1},u_k) = \ell_{\max,k}(x_k,\hat{\theta}_k,u_k)\} \tag{19}$$

$$\Theta_0(k,\hat{\theta}_k) \triangleq \{\hat{\theta}_{k+1}\in\Theta : \ell_k(x_k,\hat{\theta}_k,\hat{\theta}_{k+1},u_k) = \ell_{\min,k}(x_k,\hat{\theta}_k,u_k)\}. \tag{20}$$

For all remaining sequences, such that $\Theta^0(k,\hat{\theta}_k) \cup \Theta_0(k,\hat{\theta}_k) \subset \Theta$, and for $1 \le r \le |\Theta \setminus \{\Theta^0(k,\hat{\theta}_k) \cup \Theta_0(k,\hat{\theta}_k)\}|$, define recursively the set of states for which (15) achieves its $(j+1)$st smallest value by

$$\Theta_j(k,\hat{\theta}_k) \triangleq \Big\{\hat{\theta}_{k+1} \in \Theta : \ell_k(x_k,\hat{\theta}_k,\hat{\theta}_{k+1},u_k) = \min\big\{ \ell_k(x_k,\hat{\theta}_k,\alpha_k,u_k) : \alpha_k \in \Theta \backslash (\Theta^0(k,\hat{\theta}_k) \cup \{ (\cup_{i=1}^{j}\Theta_{j-1}(k,\hat{\theta}_k))\})\big\}\Big\}, \quad j \in \{1,2,\ldots,r\} \tag{21}$$

till all the elements of Θ are exhausted. Further, we define the corresponding values of the sequence on these sets by

$$\ell_{\Theta_j,k}(x_k,\hat{\theta}_k,u_k) \triangleq \min_{\hat{\theta}_{k+1} \in \Theta \backslash (\Theta^0(k,\hat{\theta}_k) \cup (\cup_{i=1}^{j}\Theta_{j-1}(k,\hat{\theta}_k)))} \ell_k(x_k,\hat{\theta}_k,\hat{\theta}_{k+1},u_k).$$

The notation $\mathbb{E}_p^{\mathcal{P}}[\cdot]$ will be used to denote expectation with respect to the conditional distribution $p_{ij}(k)$ over the identified partition $\mathcal{P}(k,\hat{\theta}_k) \triangleq \{\Theta^0(k,\hat{\theta}_k), \Theta_0(k,\hat{\theta}_k), \ldots, \Theta_r(k,\hat{\theta}_k)\}$. In addition, we define $(x)^+ \triangleq \max\{0,x\}$. The next theorem characterizes the maximization in (16).

Theorem 3.1: Consider the inner optimization problem in (16), with $\ell_k(x_k,\hat{\theta}_k,\hat{\theta}_{k+1},u_k)$ defined as in (15). Then, the solution is given by

$$\begin{aligned}
&\max_{p_{i\bullet} \in \mathbb{B}(i,k)} \mathbb{E}_p^{\mathcal{P}}[\ell_k(x_k = x, \hat{\theta}_k = i, \hat{\theta}_{k+1}, u_k)] \\
&= \ell_k(x,i,\hat{\theta}_{k+1} \in \Theta^0(k,i),u_k) \sum_{j \in \Theta^0(k,i)} p_{ij}^*(k) \\
&+ \ell_k(x,i,\hat{\theta}_{k+1} \in \Theta_0(k,i),u_k) \sum_{j \in \Theta_0(k,i)} p_{ij}^*(k) \\
&+ \sum_{s=1}^{r} \ell_k(x,i,\hat{\theta}_{k+1} \in \Theta_s(k,i),u_k) \sum_{j \in \Theta_s(k,i)} p_{ij}^*(k).
\end{aligned} \tag{22}$$

The maximizing, time-varying, transition probability distribution $p_{i\bullet}^*(k) \in \mathbb{B}(i,k)$, for each $\hat{\theta}_k = i \in \Theta$ and $k = 1,\ldots,N$, is given by

$$\sum_{j \in \Theta^0(k,i)} p_{ij}^*(k) = \sum_{j \in \Theta^0(k,i)} p_{ij}^0 + \frac{\alpha_i(k)}{2} \tag{23a}$$

$$\sum_{j \in \Theta_0(k,i)} p_{ij}^*(k) = \Big(\sum_{j \in \Theta_0(k,i)} p_{ij}^0 - \frac{\alpha_i(k)}{2}\Big)^+ \tag{23b}$$

$$\sum_{j \in \Theta_s(k,i)} p_{ij}^*(k) = \Big(\sum_{j \in \Theta_s(k,i)} p_{ij}^0 - \Big(\frac{\alpha_i(k)}{2} - \sum_{z=1}^{s} \sum_{j \in \Theta_{z-1}(k,i)} p_{ij}^0\Big)^+\Big)^+ \quad \text{for } s = 1,2,\ldots,r \tag{23c}$$

$$\alpha_i(k) = \min\Big(R_{TV}(i,k), 2\big(1 - \sum_{j \in \Theta^0(k,i)} p_{ij}^0\big)\Big). \tag{23d}$$

Proof: The proof is similar to that of the problem addressed in [18], [19], and hence, it is omitted. ■

The main feature of the time-varying maximizing transition probability distribution (23) is its characterization via a water-filling solution based on the nominal transition probability distribution, and the TV distance parameter R_{TV}.

C. Optimal control policy and optimal cost

Substituting (15) into (16), differentiating the right-side of (16) with respect to u_k and setting the derivative equal to zero, we obtain

$$u_k^* = -L_k(i)x - s_k(i), \quad \text{for } \hat{\theta}_k = i \tag{24}$$

with

$$L_k(i) = \Big(R_k(i) + B_k^T(i)\mathbb{E}_{p_{i\bullet}^*}^{\mathcal{P}}[P_{k+1}(\hat{\theta}_{k+1})]B_k(i)\Big)^{-1} \times B_k^T(i)\mathbb{E}_{p_{i\bullet}^*}^{\mathcal{P}}[P_{k+1}(\hat{\theta}_{k+1})]A_k(i) \tag{25a}$$

$$s_k(i) = \frac{1}{2}\Big(R_k(i) + B_k^T(i)\mathbb{E}_{p_{i\bullet}^*}^{\mathcal{P}}[P_{k+1}(\hat{\theta}_{k+1})]B_k(i)\Big)^{-1} \times B_k^T(i)\mathbb{E}_{p_{i\bullet}^*}^{\mathcal{P}}[f_{k+1}(\hat{\theta}_{k+1})] \tag{25b}$$

By our assumption on $P_k(\cdot)$ and $R_k(\cdot)$, namely, the facts that $P_k(\cdot)$ is a positive semi-definite matrix, and $R_k(\cdot)$ is a positive definite matrix, it follows that $R_k(\cdot) + B_k^T(\cdot)\mathbb{E}[P_{k+1}]B_k(\cdot)$ is positive definite, and hence, the inverse in both (25a) and (25b) exists. Substituting (13) and (24) into (16) it follows that $V_k(x,i) = x^T P_k(i)x + x^T f_k(i) + r_k(i)$ holds, with

$$P_k(i) = Q_k(i) - L_k^T(i)B_k^T(i)\mathbb{E}_{p_{i\bullet}^*}^{\mathcal{P}}[P_{k+1}(\hat{\theta}_{k+1})]A_k(i) + A_k^T(i)\mathbb{E}_{p_{i\bullet}^*}^{\mathcal{P}}[P_{k+1}(\hat{\theta}_{k+1})]A_k(i) \tag{26a}$$

$$f_k(i) = -2Q_k(i)\bar{x}_k + A_k^T(i)\mathbb{E}_{p_{i\bullet}^*}^{\mathcal{P}}[f_{k+1}(\hat{\theta}_{k+1})] - 2A_k^T(i)\mathbb{E}_{p_{i\bullet}^*}^{\mathcal{P}}[P_{k+1}(\hat{\theta}_{k+1})]B_k(i)s_k(i) \tag{26b}$$

$$r_k(i) = \bar{x}_k^T Q_k(i)\bar{x}_k - \frac{1}{2}s_k^T(i)B_k^T(i)\mathbb{E}_{p_{i\bullet}^*}^{\mathcal{P}}[f_{k+1}(\hat{\theta}_{k+1})] + \mathbb{E}_{p_{i\bullet}^*}^{\mathcal{P}}[r_{k+1}(\hat{\theta}_{k+1})]. \tag{26c}$$

In addition, the optimal cost for the minimax stochastic problem (10) is given by $J^* = \mathbb{E}^{g^*,p^*}[V_1(x_1,\hat{\theta}_1)]$.

Before we proceed with the implementation of the solution, first we redefine the sequence $\ell_k(x_k,\hat{\theta}_k,\hat{\theta}_{k+1},u_k)$. In particular, by utilizing the structure of the optimal control as given by (24), then $\ell_k(x_k,\hat{\theta}_k,\hat{\theta}_{k+1},u_k) = \ell_k(x_k,\hat{\theta}_k,\hat{\theta}_{k+1})$ where

$$\begin{aligned}
\ell_k(x_k,\hat{\theta}_k,\hat{\theta}_{k+1}) \triangleq\ & x_k^T\Big(\big(A_k(\hat{\theta}_k) - B_k(\hat{\theta}_k)L_k(\hat{\theta}_k)\big)^T P_{k+1}(\hat{\theta}_{k+1})\big(A_k(\hat{\theta}_k) - B_k(\hat{\theta}_k)L_k(\hat{\theta}_k)\big)\Big)x_k - x_k^T\Big(\\
& \big(A_k(\hat{\theta}_k) - B_k(\hat{\theta}_k)L_k(\hat{\theta}_k)\big)^T\Big(2P_{k+1}(\hat{\theta}_{k+1})B_k(\hat{\theta}_k)s_k(\hat{\theta}_k) - f_{k+1}(\hat{\theta}_{k+1})\Big)\Big) - s_k^T(\hat{\theta}_k)B_k^T(\hat{\theta}_k)\Big(f_{k+1}(\hat{\theta}_{k+1}) \\
& - P_{k+1}(\hat{\theta}_{k+1})B_k(\hat{\theta}_k)s_k(\hat{\theta}_k)\Big) + r_{k+1}(\hat{\theta}_{k+1}).
\end{aligned} \tag{27}$$

By (27), it follows that the results of Section III-B, including Theorem 3.1, can be expressed in terms of (27). In the next section, we discuss the implementation of the solution.

IV. Implementation of the Solution

Recall that the solution of the estimation problem is to be obtained by employing the one-step forward Viterbi algorithm. Specifically, the algorithm at each time step has to carry out the maximization in (3) for each $\theta_{k+1} \in \Theta$, and then choose among all θ_{k+1} the one which maximizes (4). The above procedure suggests that at every time step we need to remember for all $\theta_{k+1} \in \Theta$ the following: (i) the probability $\mu_{k+1}(\theta_{k+1})$; (ii) the previous state $\theta_k \in \Theta$, which achieves the maximum in (3); and (iii) the state estimate $\hat{\theta}_{k+1} \in \Theta$, which maximizes (4). On the other hand, the solution of the control problem requires a solution of the estimation problem. In particular, the solution of the control problem requires at each time step k the use of the MC state estimate $\hat{\theta}_k$. Note that, the estimation sequence $\hat{\theta}_{1:N}$ might be different than the most probable sequence of hidden states $\theta^*_{1:N}$. Next, an algorithm is proposed.

Algorithm 4.1: Input data: $A_k(\theta_k) \in \mathbb{R}^{n\times n}$, $B_k(\theta_k) \in \mathbb{R}^{n\times m}$, $Q_k(\theta_k) \in \mathbb{R}^{n\times n}$, and $R_k(\theta_k) \in \mathbb{R}^{m\times m}$. Choose: (i) the TV distance parameter $R_{TV}(\theta_k, k) \in [0, 2]$, (ii) the transition probability distribution $q_{ij}(k)$, (iii) the output/emission probability distribution $e_{is}(k)$, and the initial probability distribution $\pi(i)$. Set: (iv) the initial state x_1.

- As soon as a new observation is obtained apply the one-step forward Viterbi algorithm to find the state estimate, as given by (4).
- As soon as all observations $z_1, \ldots, z_N$ are obtained do:
  1: Evaluate p^0_{ij}, as given by (7).
  2: (Backward recursion) For $k = N-1, N-2, \ldots, 1$, calculate recursively:
    a) the feedback gain matrices $L_k(\hat{\theta}_k)$, and $s_k(\hat{\theta}_k)$ for $\hat{\theta}_k = i$ $(i = 1, 2, \ldots, n_\theta)$, as given by (25) (but with p^0_{ij} replacing $p^*_{ij}(k)$);
    b) the Riccati equations $P_k(\hat{\theta}_k)$, $f_k(\hat{\theta}_k)$, and $r_k(\hat{\theta}_k)$ for $\hat{\theta}_k = i$ $(i = 1, 2, \ldots, n_\theta)$, as given by (26) (but with p^0_{ij} replacing $p^*_{ij}(k)$);
  3: (Forward recursion) For $k = 1, 2, \ldots, N-1$ calculate recursively:
    a) the sequence $\ell_k(x_k, \hat{\theta}_k, \hat{\theta}_{k+1})$ and the partition $\mathcal{P}(k, \hat{\theta}_k)$, for $\hat{\theta}_k = i$ $(i = 1, 2, \ldots, n_\theta)$, and $\hat{\theta}_{k+1} = j$ $(j = 1, 2, \ldots, n_\theta)$;
    b) the maximizing probability distribution $p^*_{i\bullet}(k)$, as given by (23), for $\hat{\theta}_k = i$ $(i = 1, 2, \ldots, n_\theta)$;
    c) the control policy u^*_k and the state x_{k+1};
  4: (Update step)
    a) Repeat step 1, using the maximizing probability distribution $p^*_{i\bullet}(k)$ as obtained in step 2;
    b) update the solution of u^*_k and x_{k+1}, using the results of step 3a).

An important practical aspect of Algorithm 4.1 is that due to the separation that exists between the solution of the estimation problem and the control problem, alternative estimation algorithms, other than the Viterbi, may also be used. In the next section, a numerical example (drawn from [8] and modified to incorporate a two-state HMM) is solved to illustrate the capabilities of the proposed approach.

V. Numerical Example

Consider (1) with number of states $n = 2$, and number of different operating modes $n_\theta = 2$. The dynamics and input matrices, for each operating mode of (1), are given by

$$A_k(\theta_k{=}1) = \begin{pmatrix} 0.99 & 0.53 \\ -0.10 & 1.15 \end{pmatrix}, \quad B_k(\theta_k{=}1) = \begin{pmatrix} 0.0013 \\ 0.0539 \end{pmatrix}$$
$$A_k(\theta_k{=}2) = \begin{pmatrix} 1 & -0.59 \\ -0.025 & 0 \end{pmatrix}, \ B_k(\theta_k{=}2) = \begin{pmatrix} 0.02 \\ 0.10 \end{pmatrix}.$$

Note that, under both its operating modes the system is unstable but completely controllable. The state cost matrices are given by

$$Q_k(\theta_k = 1) = Q_k(\theta_k = 2) = \begin{pmatrix} 0.5 & 0 \\ 0 & 0.05 \end{pmatrix}$$

with $Q_N(\theta_N = 1) = Q_N(\theta_N = 2) = Q_k(\theta_k = 1)$, and input cost $R_k(\theta_k = 1) = R_k(\theta_k = 2) = 0.01$. The final time is set equal to $N = 200$, with initial conditions $x_0 = [2 \ \ 0]^T$, and a reference trajectory signal given by $\bar{x}_k = [10(1 - \exp(-0.05k)) \ \ 0]^T$ for $k = 1, 2, \ldots, N$. A two-state HMM is considered with

$$q_{ij}(k) = \begin{pmatrix} 0.45 & 0.55 \\ 0.6 & 0.4 \end{pmatrix}, \quad e_{is}(k) = \begin{pmatrix} 0.1 & 0.9 \\ 0.8 & 0.2 \end{pmatrix}$$

for $i, j \in \Theta$, $s \in Z$, and $\forall k$. The initial probability distribution of the HMM is set equal to $\pi = [0.5 \ \ 0.5]^T$.

For the sake of this example, the true transition probability distribution of the HMM is assumed to be known and given by:

$$q^{true}_{ij}(k) = \begin{pmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{pmatrix}, \quad \text{for } k = 1, 2, \ldots, N/2,$$
$$q^{true}_{ij}(k) = \begin{pmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{pmatrix}, \quad \text{for } k = N/2+1, \ldots, N-1.$$

The TV distance parameter is set equal to:

$$R_{TV}(1, k) = R_{TV}(2, k) = \alpha, \quad k = 1, 2, \ldots, N/2,$$
$$R_{TV}(1, k) = R_{TV}(2, k) = \alpha/2, \quad k = N/2+1, \ldots, N-1$$

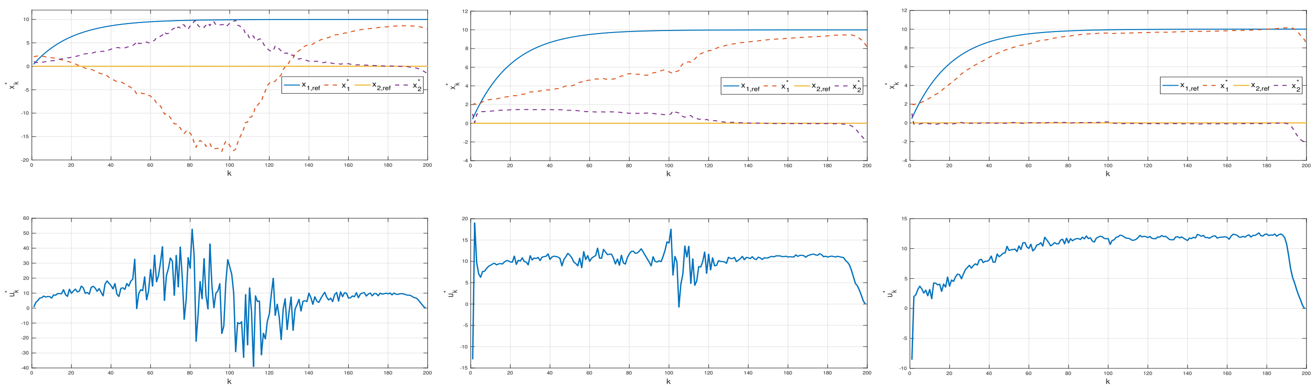
with $\alpha \in [0, 2]$. We will refer to the HMM under the true transition probability distribution as the "*true HMM*", and to the HMM under the nominal transition probability distribution as the "*nominal HMM*". For the solution of the problem two scenarios are considered:

(S1) Assume that the nominal HMM is correct, and hence, no uncertainty is present. This scenario is achieved by setting $\alpha = 0$.

(S2) Assume that the nominal HMM is uncertain. This scenario is achieved by setting $\alpha > 0$.

For the implementation of both scenarios, Algorithm 4.1 is employed, with the system's jumps/transitions between the different operating modes realized in simulations using the true HMM. The optimal results are obtained by performing a total of 10^3 Monte Carlo realizations of the HMM. The mean values of the optimal trajectories and control policies for scenarios (S1) – (S2), are as shown in Figure 1.

Figure 1(a) depicts the optimal trajectory and the optimal LQ controller under (S1). We immediately notice the effect of uncertainty on the LQ trajectory tracking performance, which causes a considerable difference between the optimal



(a) $\alpha = 0$ (b) $\alpha = 0.3$ (c) $\alpha = 0.9$

Fig. 1: Solution of the LQ trajectory tracking problem for different values of the TV distance parameter. Top row graphs depict the optimal state trajectories. Bottom row graphs depict the optimal LQ controller.

trajectory x_k^* and the reference trajectory signal $\bar{x}_k$. This is to be expected since the optimal controller is evaluated by assuming that the nominal and the true transition probability distribution of the HMM are equal, and hence, any uncertainty affects the optimality of the LQ controller.

On the contrary, Figures 1(b) and 1(c) confirm that the proposed LQ trajectory tracking approach restricts the influence of uncertainty on the performance of the optimal LQ controller. In particular, by allowing the TV distance parameter to increase we observe that the difference between x_k^* and $\bar{x}_k$ is reduced. This behavior is expected since under scenario (S2) the optimal controller is evaluated using the worst-case transition probability distribution over all possible transition probability distributions of the HMM within the TV distance ambiguity set, and as a result, the optimal controller is more robust with respect to uncertainty and more effective in achieving its tracking objectives. As a final remark, we note that the "tail" which appears as $t \to 200$ occurs because the optimal controller anticipates the end of the control interval. Thus, the controller conserves control effort by allowing x^* to deviate from its reference trajectory.

VI. Conclusion

The LQ tracking control problem is analyzed for MJLS with hidden states, and uncertain conditional distributions. Both problems of state estimation and robust control are addressed, resulting in a maximizing, time-varying conditional distribution, and in an optimal control policy with some desired robustness properties. Results validate the capability of the approach on restricting the influence of uncertainty and on ensuring the optimal performance of the LQ controller.

References

[1] M. Mariton, *Jump Linear Systems in Automatic Control*. New York: Marcel Dekker, 1990.

[2] L. Zhang, B. Cai, and Y. Shi, "Stabilization of hidden semi-Markov jump systems: Emission probability approach," *Automatica*, vol. 101, pp. 87–95, 2019.

[3] S. Dong, Z.-G. Wu, Y.-J. Pan, H. Su, and Y. Liu, "Hidden-Markov-model-based asynchronous filter design of nonlinear Markov jump systems in continuous-time domain," *IEEE Transactions on Cybernetics*, vol. 49, no. 6, pp. 2294–2304, 2019.

[4] V. Gupta, R. M. Murray, and B. Hassibi, "On the control of jump linear Markov systems with Markov state estimation," in *Proceedings of the 2003 American Control Conference*, vol. 4, 2003, pp. 2893–2898.

[5] B. Cai, L. Zhang, and Y. Shi, "Observed-mode-dependent state estimation of hidden semi-Markov jump linear systems," *IEEE Transactions on Automatic Control*, vol. 65, no. 1, pp. 442–449, Jan 2020.

[6] A. R. P. Andriën and D. J. Antunes, "Near-optimal map estimation for Markov jump linear systems using relaxed dynamic programming," *IEEE Control Systems Letters*, vol. 4, no. 4, pp. 815–820, 2020.

[7] W. Liu, "State estimation for discrete-time Markov jump linear systems with time-correlated measurement noise," *Automatica*, vol. 76, pp. 266–276, 2017.

[8] I. Tzortzis, C. D. Charalambous, and C. N. Hadjicostis, "Jump LQR systems with unknown transition probabilities," *IEEE Transactions on Automatic Control*, vol. 66, no. 6, pp. 2693–2708, 2021.

[9] M. G. Todorov and M. D. Fragoso, "New methods for mode-independent robust control of Markov jump linear systems," *Systems & Control Letters*, vol. 90, pp. 38–44, 2016.

[10] I. Tzortzis, C. D. Charalambous, and C. N. Hadjicostis, "Robust LQG for Markov jump linear systems," in *58th IEEE Conference on Decision and Control (CDC)*, Dec. 2019, pp. 6760–6765.

[11] L. Zhang and J. Lam, "Necessary and sufficient conditions for analysis and synthesis of Markov jump linear systems with incomplete transition descriptions," *IEEE Transactions on Automatic Control*, vol. 55, no. 7, pp. 1695–1701, Jul. 2010.

[12] Y. Z. Lun, A. D'Innocenzo, and M. D. D. Benedetto, "Robust LQR for time-inhomogeneous Markov jump switched linear systems," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 2199–2204, 2017.

[13] A. M. de Oliveira, O. L. V. Costa, and J. Daafouz, "Design of stabilizing dynamic output feedback controllers for hidden Markov jump linear systems," *IEEE Control Systems Letters*, vol. 2, no. 2, pp. 278–283, 2018.

[14] B. Mor, S. Garhwal, and A. Kumar, "A systematic review of hidden Markov models and their applications," *Archives of Computational Methods in Engineering*, vol. 28, pp. 1429–1448, 2021.

[15] G. D. Forney, "The Viterbi algorithm," *Proceedings of the IEEE*, vol. 61, no. 3, pp. 268–278, 1973.

[16] T. W. Anderson and L. A. Goodman, "Statistical inference about Markov chains," *The Annals of Mathematical Statistics*, vol. 28, no. 1, pp. 89–110, 1957.

[17] I. Tzortzis, C. D. Charalambous, and T. Charalambous, "Dynamic programming subject to total variation distance ambiguity," *SIAM J. Control Optim.*, vol. 53, no. 4, pp. 2040–2075, Jul. 2015.

[18] C. D. Charalambous, I. Tzortzis, S. Loyka, and T. Charalambous, "Extremum problems with total variation distance and their applications," *IEEE Transactions on Automatic Control*, vol. 59, no. 9, pp. 2353–2368, Sep. 2014.

[19] I. Tzortzis, C. D. Charalambous, T. Charalambous, C. N. Hadjicostis, and M. Johansson, "Approximation of Markov processes by lower dimensional processes via total variation metrics," *IEEE Transactions on Automatic Control*, vol. 62, no. 3, pp. 1030–1045, Mar. 2017.