

## CONTROL OF SINGULARLY PERTURBED MARKOV CHAINS: A NUMERICAL STUDY

H. YANG<sup>1</sup>, G. YIN<sup>2</sup>, K. YIN<sup>3</sup> and Q. ZHANG<sup>4</sup>

(Received 26 May, 2000; revised 13 March, 2001)

### Abstract

This work is devoted to numerical studies of nearly optimal controls of systems driven by singularly perturbed Markov chains. Our approach is based on the ideas of hierarchical controls applicable to many large-scale systems. A discrete-time linear quadratic control problem is examined. Its corresponding limit system is derived. The associated asymptotic properties and near optimality are demonstrated by numerical examples. Numerical experiments for a continuous-time hybrid linear quadratic regulator with Gaussian disturbances and a discrete-time Markov decision process are also presented. The numerical results have not only supported our theoretical findings but also provided insights for further applications.

### 1. Introduction

This work is concerned with nearly optimal controls of systems driven by singularly perturbed Markov chains. We consider both discrete-time and continuous-time problems. The main objectives are to study the issues related to reduction of complexity of the Markovian systems and to demonstrate the asymptotic properties numerically. Our numerical experiments have not only supported the theoretical findings, but also provided insights for further applications.

In many real-world problems, a common practice in quantifying the dynamic relationships of random events and uncertainties is to use stochastic processes in modelling and formulation. For systems having jump sample paths, such as those often encountered in communication, reliability, manufacturing and queueing networks, Markovian jump models have become popular. Many optimisation and control problems can be

<sup>1</sup>Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455, USA; e-mail: hyang@ece.umn.edu.

<sup>2</sup>Department of Mathematics, Wayne State University, Detroit, MI 48202, USA; e-mail: gyin@math.wayne.edu.

<sup>3</sup>Department of Wood and Paper Science, University of Minnesota, St. Paul, MN 55108, USA.

<sup>4</sup>Department of Mathematics University of Georgia, Athens, GA 30602, USA.

© Australian Mathematical Society 2003, Serial-fee code 1446-1811/03

formulated as stochastic control problems driven by Markovian noise. Typically the underlying Markov chains have large state spaces, which results in complex structures and leads to serious obstacles in obtaining optimal controls. Although optimal control problems can usually be solved by a dynamic programming approach, direct implementation of the dynamic programming principles only works well for those systems with moderate dimensions because large dimensionality often renders the computation infeasible. This is the so-called “curse of dimensionality” phenomenon. To overcome this difficulty requires reducing the complexity of the underlying problems.

In our continuing effort to study such large-scale systems [10, 11, 24–26, 28], we have adopted the hierarchical decomposition approach (see [4, 19, 20] among others), which leads to the formulation of singularly perturbed Markov chains. Related work on singularly perturbed systems can be found in [1, 3, 6, 9, 15–17, 22] and the references therein; the formulation of piecewise deterministic processes is in [5]. One of the crucial observations is that for the large number of states involved, their rates of changes are not the same. Some of them change very rapidly, while others may vary at rates orders of magnitude slower. According to singular perturbation theory, such a high contrast of rates of changes can be reflected by introducing a small parameter and by using time-scale separation. This yields two time scales, a fast changing one versus a slowly varying one. With different time scales, we can lump the states (in each recurrent class) together in accordance with their rates of change. This aggregation allows us to form a reduced system having fewer states, in which the optimal controls for the reduced problem are easier to obtain. Based on the solution of the reduced problem, we can then construct control policies that will lead to near optimality.

In this paper, we deal with both discrete-time and continuous-time cases, including discrete-time linear quadratic control (LQ) problems, discrete-time Markov decision processes (MDP), and continuous-time hybrid linear quadratic regulator (LQG) problems. Numerical experiments are carried out to verify the asymptotic optimality. The rest of the paper is arranged as follows. Section 2 studies an LQ problem in discrete time. Using weak convergence methods, we derive convergence of the dynamic systems and the value functions under suitable scaling. Section 3 treats hybrid LQG problems in continuous time. Section 4 is concerned with a discrete-time MDP. Section 5 concludes the paper with a couple of remarks. An appendix is furnished to cover issues in the simulation of Markov chains and other matters in computation.

**Notation.** This paper deals with finite-state Markov chains regardless of the problems being in discrete time or in continuous time. Denote the state space by  $\mathcal{M}$  and write  $\mathcal{M} = \{1, \dots, m\}$ . For a generator  $Q = (q_{ij})$  of a continuous-time Markov chain and a suitable function  $f(\cdot)$  defined on  $\mathcal{M}$ ,  $Qf(\cdot)(i)$  is meant to be

$$Qf(\cdot)(i) = \sum_{j \neq i} q_{ij} (f(j) - f(i)). \quad (1.1)$$

For simplicity, in this paper, we concentrate on the cases where the dominating parts of the Markov chains contain only recurrent states. In this case,  $\mathcal{M}$  is decomposable into  $l$  subspaces

$$\mathcal{M} = \mathcal{M}_1 \cup \dots \cup \mathcal{M}_l = \{s_{11}, \dots, s_{1m_1}\} \cup \dots \cup \{s_{l1}, \dots, s_{lm_l}\}. \tag{1.2}$$

Throughout the paper, we use the convention  $i \in \mathcal{M}$  to denote an element  $i \in \{1, \dots, m\}$  and use  $s_{i\ell}$  to denote an element in the  $\ell$ th subspace  $\mathcal{M}_\ell$ .

### 2. A discrete-time LQ problem

We examine a discrete-time LQ regulator problem. The design of traditional linear methods feedback controllers is based on a plant model with fixed parameters. Although it provides a way to manage for many dynamic systems, it cannot handle the situations in which the actual system is different from the nominal model. Much effort has been directed to the design of more robust methods in recent years. To take into consideration discrete shifts in regime across which the behaviour of the corresponding dynamic systems are markedly different, we present an LQ model with Markovian switches in this section. Rather than using a fixed system configuration, we take random jumps into consideration, propose a hybrid system, derive its limit, and demonstrate that its optimal control can be approximated by that of a continuous-time hybrid LQG problem and that the nearly optimal controls of the original system can be constructed using the limit system.

Let  $\varepsilon > 0$  be a small parameter and  $\alpha_n^\varepsilon$  a discrete-time Markov chain with a finite state space  $\mathcal{M}$  having  $m$  elements and transition probability matrix  $P^\varepsilon = (p_{ij}^\varepsilon)$  given by

$$P^\varepsilon = P + \varepsilon Q, \tag{2.1}$$

where  $P$  is a transition probability matrix and  $Q$  is a generator (the precise condition is given in (A1)). For any  $z \in \mathbb{R}$ , use  $[z]$  to denote the greatest integer function that gives the greatest integer less than or equal to  $z$ . For a finite real number  $T > 0$  and  $0 \leq n \leq [T/\varepsilon]$ , the dynamic system is given by

$$\begin{cases} x_{n+1} = x_n + \varepsilon A(\alpha_n^\varepsilon)x_n + \varepsilon B(\alpha_n^\varepsilon)u_n + \sqrt{\varepsilon}w_n, \\ x_0 = x, \quad \text{a deterministic vector,} \end{cases} \tag{2.2}$$

where  $x_n \in \mathbb{R}^{n_1}$  is the state,  $u_n \in \mathbb{R}^{n_2}$  is the control,  $A(i) \in \mathbb{R}^{n_1 \times n_1}$  and  $B(i) \in \mathbb{R}^{n_1 \times n_2}$  are well-defined and have finite values for each  $i \in \mathcal{M}$ , and  $\{w_n\}$  is a sequence of random variables with zero mean. Define a sequence of cost functions by

$$J_n^\varepsilon(x, \alpha, u(\cdot)) = \varepsilon E \left\{ \sum_{k=0}^{n-1} [x'_k M(\alpha_k^\varepsilon)x_k + u'_k N(\alpha_k^\varepsilon)u_k] + x'_n D x_n \right\}, \tag{2.3}$$

where  $E = E_{x,\alpha}$  is the expectation given  $\alpha_0^\varepsilon = \alpha$  and  $x_0 = x$ . Our objective is to find an optimal control  $u$  to minimise the expected quadratic cost function  $J_{[T/\varepsilon]}^\varepsilon(x, \alpha, u(\cdot))$ . To obtain the desired asymptotic results, we make several assumptions.

(A1) The following conditions hold:

- $\{w_n\}$  is a sequence of independent and identically distributed (i.i.d.) random variables with  $Ew_n = 0$  and  $E|w_n|^2 < \infty$ .
- $\Sigma = E(w_0w_0')$  is a symmetric positive definite matrix.
- $P$  in (2.1) has the form

$$P = \text{diag}(P_1, \dots, P_l) \tag{2.4}$$

such that each  $P_i$  is irreducible and aperiodic and  $Q = (q_{ij})$  is a generator, that is,  $q_{ij} \geq 0$  for  $i \neq j$  and  $q_{ii} = -\sum_{j \neq i} q_{ij}$ . In the above,  $\text{diag}(H_1, \dots, H_l)$  denotes a block diagonal matrix having matrix entries  $H_1, \dots, H_l$ .

- The Markov chain  $\alpha_n^\varepsilon$  and the random disturbance  $w_n$  are independent.
- For each  $i \in \mathcal{M}$ ,  $M(i)$  are symmetric nonnegative definite matrices, and  $N(i)$  and  $D$  are symmetric positive definite matrices.

**REMARK 2.1.** We concentrate on the cases of singularly perturbed Markov chains with recurrent states. The cases of inclusion of transient states can be treated in a similar way. As far as the optimal controls are concerned (see [28]), however, the transient states are asymptotically unimportant. We consider “white noise” only. The i.i.d. assumption on the sequence  $\{w_n\}$  is mainly for convenience. The limit results for more complex noise processes such as  $\phi$ -mixing processes can be obtained. The essence is that a central limit theorem holds for the noise, which yields a Brownian motion limit. For simplicity, we treat the cases where the variance of the noise (as a result of the diffusion term in the limit) does not depend on the singularly perturbed Markov chain. The results obtained can be extended to  $\alpha_k^\varepsilon$ -dependent variance. However, more complex averaging schemes are needed.

Denote the value functions by  $v_n^\varepsilon(x, \alpha) = \inf_{u(\cdot)} J_n^\varepsilon(x, \alpha, u(\cdot))$ . For each  $0 \leq n \leq N = \lfloor T/\varepsilon \rfloor$ , applying the dynamic programming principle with a slight modification of the argument in [2, p. 70], yields a system of dynamic programming equations:

$$\begin{cases} v_N^\varepsilon(x_N, \alpha_N^\varepsilon) = x_N' D x_N, \\ v_n^\varepsilon(x_n, \alpha_n^\varepsilon) = \min_u \{ \varepsilon x_n' M(\alpha_n^\varepsilon) x_n + \varepsilon u_n' N(\alpha_n^\varepsilon) u_n + v_{n+1}^\varepsilon(x_{n+1}, \alpha_{n+1}^\varepsilon) \}. \end{cases} \tag{2.5}$$

Define  $\widehat{F}(i) = \sum_{j=1}^m p_{ij}^\varepsilon F(j)$  for an appropriate function  $F(\cdot)$  defined on  $\mathcal{M}$ ,  $\widetilde{A}(i) = I + \varepsilon A(i)$ ,  $\widetilde{B}(i) = \varepsilon B(i)$ . Using a dynamic programming approach as in [8, pp. 165–166], assuming the value function to be of the form

$$v_n^\varepsilon(x, i) = x' K_n^\varepsilon(i) x + r_n^\varepsilon(i), \tag{2.6}$$

we proceed to determine  $K_n^\epsilon$  and  $r_n^\epsilon$ . For each  $i \in \mathcal{M}$ , substituting (2.6) into (2.5) yields the following system of Riccati equations:

$$\begin{cases} K_n^\epsilon(i) = \tilde{A}'(i)\hat{K}_{n+1}^\epsilon(i)\tilde{A}(i) + \epsilon M(i) - \tilde{A}'(i)\hat{K}_{n+1}^\epsilon(i)\tilde{B}(i)(\epsilon N(i) \\ \quad + \tilde{B}'(i)\hat{K}_{n+1}^\epsilon(i)\tilde{B}(i))^{-1}\tilde{B}'(i)\hat{K}_{n+1}^\epsilon(i)\tilde{A}(i), \\ K_N^\epsilon(i) = D \end{cases} \tag{2.7}$$

and

$$\begin{cases} r_n^\epsilon(i) = \hat{r}_{n+1}^\epsilon(i) + \epsilon \operatorname{tr}(\hat{K}_{n+1}^\epsilon(i)\Sigma), \\ r_N^\epsilon(i) = 0. \end{cases} \tag{2.8}$$

REMARK 2.2. Instead of solving one equation, we must solve a system of  $m$  equations. When the state space  $\mathcal{M}$  is large, the required computation is intensive. Using a singularly perturbed Markov chain approach, however, reduces the complexity and computational burden.

Note that

$$\begin{aligned} K_n^\epsilon(i) &= \hat{K}_{n+1}^\epsilon(i) + \epsilon A'(i)\hat{K}_{n+1}^\epsilon(i) + \epsilon \hat{K}_{n+1}^\epsilon(i)A(i) \\ &\quad + \epsilon M(i) - \epsilon \hat{K}_{n+1}^\epsilon(i)B(i)(N(i) \\ &\quad + \epsilon B'(i)\hat{K}_{n+1}^\epsilon(i)B(i))^{-1}B'(i)\hat{K}_{n+1}^\epsilon(i) + O(\epsilon^2). \end{aligned} \tag{2.9}$$

Furthermore

$$\hat{K}_{n+1}^\epsilon(i) = K_{n+1}^\epsilon(i) + \epsilon \sum_{j=1}^m q_{ij}^\epsilon K_{n+1}^\epsilon(j), \tag{2.10}$$

where

$$Q^\epsilon = \frac{1}{\epsilon}(P - I) + Q = (q_{ij}^\epsilon) = \left( \frac{1}{\epsilon}(p_{ij}^\epsilon - \delta_{ij}) + q_{ij} \right) \tag{2.11}$$

and where  $\delta_{ij} = 1$  if  $i = j$  and is 0 otherwise. Using this notation, rewrite (2.9) as

$$\begin{aligned} K_{n+1}^\epsilon(i) &= K_n^\epsilon(i) - \epsilon A'(i)\hat{K}_{n+1}^\epsilon(i) - \epsilon \hat{K}_{n+1}^\epsilon(i)A(i) - \epsilon M(i) + \epsilon \hat{K}_{n+1}^\epsilon(i)B(i)(N(i) \\ &\quad + \epsilon B'(i)\hat{K}_{n+1}^\epsilon(i)B(i))^{-1}B'(i)\hat{K}_{n+1}^\epsilon(i) - \epsilon \sum_{j=1}^m q_{ij}^\epsilon K_{n+1}^\epsilon(j) + O(\epsilon^2). \end{aligned}$$

This equivalent form will be useful in the subsequent analysis.

For each  $i$ , define piecewise constant interpolated processes  $K^\epsilon(\cdot, i)$  and  $r^\epsilon(\cdot, i)$  as  $K^\epsilon(t, i) = K_n^\epsilon(i)$  and  $r^\epsilon(t, i) = r_n^\epsilon(i)$ , for  $t \in [n\epsilon, n\epsilon + \epsilon)$ . With a modification of the argument of the proofs of Lemmas 1 and 2 in [28], we obtain the following lemma.

LEMMA 2.1. *The following assertions hold:*

(i)  $K_n^\varepsilon(i)$  and  $K^\varepsilon(t, i)$  are positive definite for each  $0 \leq n \leq \lfloor T/\varepsilon \rfloor$  and each  $t \in [0, T]$ , respectively.

(ii) For each  $i$  and for some  $\kappa_T > 0$ ,

$$\begin{aligned} \sup_{0 \leq n \leq \lfloor T/\varepsilon \rfloor} |K_n^\varepsilon(i)| \leq \kappa_T, \quad \sup_{t \in [0, T]} |K^\varepsilon(t, i)| \leq \kappa_T, \\ \sup_{0 \leq n \leq \lfloor T/\varepsilon \rfloor} |r_n^\varepsilon(i)| \leq \kappa_T, \quad \sup_{t \in [0, T]} |r^\varepsilon(t, i)| \leq \kappa_T. \end{aligned}$$

It can be shown similarly as in [2, p. 73] or [8, p. 166] that the optimal feedback control for the LQ problem is linear in the state variable, and

$$u_n = - [\tilde{B}'(i) \hat{K}_{n+1}^\varepsilon(i) \tilde{B}(i) + \varepsilon N(i)]^{-1} \tilde{B}'(i) \hat{K}_{n+1}^\varepsilon(i) \tilde{A}(i) x_n \stackrel{\text{def}}{=} -\Phi_n(i) x_n.$$

Consequently the dynamic system can be written as

$$x_{n+1} = x_n + \varepsilon \Theta_n(\alpha_n^\varepsilon) x_n + \sqrt{\varepsilon} w_n, \quad \text{where } \Theta_n(i) = A(i) - B(i) \Phi_n(i). \quad (2.12)$$

Henceforth, we use  $\kappa$  to denote a generic positive constant with possible different values for different appearances and with the understanding of the notation  $\kappa + \kappa = \kappa$  and  $\kappa \kappa = \kappa$ . To proceed, we first obtain a bound on the second moment of  $x_n$ .

LEMMA 2.2. *Under (A1),  $\sup_{0 \leq n \leq \lfloor T/\varepsilon \rfloor} E|x_n|^2 < \infty$ .*

PROOF. For any  $0 \leq n \leq \lfloor T/\varepsilon \rfloor$ , from (2.12), it is easily seen that

$$E|x_n|^2 \leq \kappa \left( |x_0|^2 + \varepsilon^2 \sum_{k=0}^{n-1} \sum_{k_1=0}^{n-1} E \operatorname{tr} [x_{k_1} x_k' \Theta_{k_1}(\alpha_{k_1}^\varepsilon) \Theta_k'(\alpha_k^\varepsilon)] + \varepsilon E \sum_{k=0}^{n-1} \sum_{k_1=0}^{n-1} E w_{k_1}' w_k \right).$$

An application of the discrete version of Gronwall's inequality then yields

$$E|x_n|^2 \leq \kappa (|x_0|^2 + \operatorname{tr} \Sigma) \exp(\kappa T) < \infty.$$

Moreover the bounds hold uniformly for  $0 \leq n \leq \lfloor T/\varepsilon \rfloor$ .

To effectively reduce the complexity, we use the idea of aggregation. In accordance with the form of  $P$  given in (2.4), the state space of the underlying Markov chain can be written as (1.2) to reflect the fact that the state space can be decomposed into  $l$  recurrent classes. We take a continuous-time interpolation as  $x^\varepsilon(t) = x_n$  for  $t \in [n\varepsilon, (n+1)\varepsilon)$ . Define  $\bar{\alpha}_n^\varepsilon = i$  if  $\alpha_n^\varepsilon \in \mathcal{M}_i$  and define  $\bar{\alpha}^\varepsilon(t) = \bar{\alpha}_n^\varepsilon$  for  $t \in [n\varepsilon, (n+1)\varepsilon)$ . Working with the interpolated pair  $(x^\varepsilon(\cdot), \bar{\alpha}^\varepsilon(\cdot))$ , we will show that it converges weakly to  $(x(\cdot), \bar{\alpha}(\cdot))$ , that is, a solution of a hybrid system in continuous time.

LEMMA 2.3. Under (A1), the following assertions hold.

(i)  $\bar{\alpha}^\varepsilon(\cdot)$  converges weakly to  $\bar{\alpha}(\cdot)$  as  $\varepsilon \rightarrow 0$ , which is a continuous-time Markov chain generated by

$$\bar{Q} = \text{diag}(v^1, \dots, v^l) \widehat{Q} \text{diag}(\mathbf{1}_{m_1}, \dots, \mathbf{1}_{m_l}), \tag{2.13}$$

where  $v^i$  is the stationary distribution corresponding to  $P_i$  for each  $i = 1, \dots, l$ , and  $\mathbf{1}_{\ell_0}$  denotes an  $\ell_0$ -dimensional column vector with all entries being 1.

(ii)  $E\left(\varepsilon \sum_{k=0}^{\lfloor T/\varepsilon \rfloor} [I_{\{\alpha_k^\varepsilon = s, \iota\}} - v_i^\varepsilon I_{\{\alpha_k^\varepsilon \in \mathcal{A}_i\}}]\right)^2 \rightarrow 0$  as  $\varepsilon \rightarrow 0$ .

This result has been proved in [27]. We omit the proof here. For corresponding results for singularly perturbed continuous-time Markov chains, see [24, Chapter 7].

LEMMA 2.4. Under (A1),  $\{x^\varepsilon(\cdot)\}$  is tight in  $D^{n_1}[0, T]$ ; the space of  $\mathbb{R}^{n_1}$ -valued functions that are right continuous have left limits endowed with the Skorohod topology.

PROOF. We use the tightness criteria [7, Section 3.8, p. 132], and [12, Theorem 3, p. 47], to resolve the problem. Denote the conditional expectation on the  $\sigma$ -algebra  $\mathcal{F}_t^\varepsilon = \sigma\{\alpha^\varepsilon(u), u \leq t\}$  by  $E_t^\varepsilon$ . For simplicity and with a slight abuse of notation, in lieu of  $\lfloor t/\varepsilon \rfloor$  and  $\lfloor (t+s)/\varepsilon \rfloor$ , we often use  $t/\varepsilon$  and  $(t+s)/\varepsilon$  to denote integers in what follows. It should be clear from the context. For any  $\eta > 0$ ,  $t > 0$  and  $0 < s < \eta$ , using (2.12), we obtain

$$\begin{aligned} E_t^\varepsilon |x^\varepsilon(t+s) - x^\varepsilon(t)|^2 &= \varepsilon^2 E_t^\varepsilon \sum_{k_1=t/\varepsilon}^{(t+s)/\varepsilon-1} \sum_{k=t/\varepsilon}^{(t+s)/\varepsilon-1} \text{tr}[x_{k_1} x'_k \Theta_{k_1}(\alpha_{k_1}^\varepsilon) \Theta'_k(\alpha_k^\varepsilon)] \\ &\quad + 2\varepsilon^{3/2} E_t^\varepsilon \sum_{k_1=t/\varepsilon}^{(t+s)/\varepsilon-1} \sum_{k=t/\varepsilon}^{(t+s)/\varepsilon-1} w'_k \Theta_{k_1}(\alpha_{k_1}^\varepsilon) x_{k_1} \\ &\quad + \varepsilon E_t^\varepsilon \sum_{k_1=t/\varepsilon}^{(t+s)/\varepsilon-1} \sum_{k=t/\varepsilon}^{(t+s)/\varepsilon-1} w'_{k_1} w_k. \end{aligned}$$

Using Lemma 2.2,

$$\varepsilon^2 E E_t^\varepsilon \sum_{k_1=t/\varepsilon}^{(t+s)/\varepsilon-1} \sum_{k=t/\varepsilon}^{(t+s)/\varepsilon-1} \text{tr}[x_{k_1} x'_k \Theta_{k_1}(\alpha_{k_1}^\varepsilon) \Theta'_k(\alpha_k^\varepsilon)] = O(s^2).$$

Using independence and  $E w_k = 0$ , we obtain

$$\varepsilon E E_t^\varepsilon \sum_{k_1=t/\varepsilon}^{(t+s)/\varepsilon-1} \sum_{k=t/\varepsilon}^{(t+s)/\varepsilon-1} w'_{k_1} w_k = O(s).$$

By the Cauchy-Schwartz inequality,

$$\begin{aligned}
 & E \left| \varepsilon^{3/2} \sum_{k_1=t/\varepsilon}^{(t+s)/\varepsilon-1} \sum_{k=t/\varepsilon}^{(t+s)/\varepsilon-1} w'_k \Theta_{k_1}(\alpha_{k_1}^\varepsilon) x_{k_1} \right| \\
 & \leq E^{1/2} \left| \varepsilon \sum_{k_1=t/\varepsilon}^{(t+s)/\varepsilon-1} \Theta_{k_1}(\alpha_{k_1}^\varepsilon) x_{k_1} \right|^2 E^{1/2} \left| \varepsilon^{1/2} \sum_{k=t/\varepsilon}^{(t+s)/\varepsilon-1} w_k \right|^2 = O(s).
 \end{aligned}$$

Thus  $\lim_{\eta \rightarrow 0} \limsup_{\varepsilon \rightarrow 0} E|x^\varepsilon(t+s) - x^\varepsilon(t)|^2 \rightarrow 0$ . The desired tightness thus follows.

Since  $\{\bar{\alpha}^\varepsilon(\cdot)\}$  is tight due to its weak convergence (see Lemma 2.3), we can further show  $(x^\varepsilon(\cdot), \bar{\alpha}^\varepsilon(\cdot))$  is tight. Choose a weakly convergent subsequence and for notational simplicity, still use  $\varepsilon$  as its index. We proceed to characterise the limit process and show that the following assertion holds.

**THEOREM 2.1.** *Under (A1),  $(x^\varepsilon(\cdot), \bar{\alpha}^\varepsilon(\cdot))$  converges weakly to  $(x(\cdot), \bar{\alpha}(\cdot))$  such that  $\bar{\alpha}(\cdot)$  is given in Lemma 2.3 and  $x(\cdot)$  is the solution of the hybrid system*

$$dx(t) = \Xi(\bar{\alpha}(t))x(t) + \sigma dw(t),$$

where  $\sigma\sigma' = \Sigma$ ,  $\Xi(t) = \overline{A(t) - B(t)N^{-1}(t)B'(t)K(t)A(t)}$ ,  $t \in \overline{\mathcal{M}} = \{1, \dots, l\}$ , and for a suitable function  $F(\cdot)$ ,  $\overline{F}(t) = \sum_{\ell=1}^{m_i} \nu_\ell^i F(s_{t\ell})$ .

**PROOF.** Using (2.12), we have

$$\begin{aligned}
 x^\varepsilon(t) &= x_0 + \varepsilon \sum_{i=1}^l \sum_{\ell=1}^{m_i} \sum_{k=0}^{t/\varepsilon-1} \Theta_k(s_{t\ell}) x_k I_{\{\alpha_k^\varepsilon = s_{t\ell}\}} + \sqrt{\varepsilon} \sum_{k=0}^{t/\varepsilon-1} w_k \\
 &= x_0 + \varepsilon \sum_{i=1}^l \sum_{\ell=1}^{m_i} \sum_{k=0}^{t/\varepsilon-1} \Theta_k(s_{t\ell}) x_k \nu_\ell^i I_{\{\alpha_k^\varepsilon \in \mathcal{A}_i\}} + \sqrt{\varepsilon} \sum_{k=0}^{t/\varepsilon-1} w_k \\
 &\quad + \varepsilon \sum_{i=1}^l \sum_{\ell=1}^{m_i} \sum_{k=0}^{t/\varepsilon-1} \Theta_k(s_{t\ell}) x_k [I_{\{\alpha_k^\varepsilon = s_{t\ell}\}} - \nu_\ell^i I_{\{\alpha_k^\varepsilon \in \mathcal{A}_i\}}].
 \end{aligned}$$

By virtue of the boundedness of  $\Theta_n$ , using Lemmas 2.2 and 2.3, a partial summation leads to

$$E \left| \varepsilon \sum_{i=1}^l \sum_{\ell=1}^{m_i} \sum_{k=0}^{t/\varepsilon-1} \Theta_k(s_{t\ell}) x_k [I_{\{\alpha_k^\varepsilon = s_{t\ell}\}} - \nu_\ell^i I_{\{\alpha_k^\varepsilon \in \mathcal{A}_i\}}] \right|^2 \rightarrow 0 \quad \text{as } \varepsilon \rightarrow 0.$$

The well-known Donsker invariance theorem implies that  $\sqrt{\varepsilon} \sum_{k=0}^{t/\varepsilon-1} w_k \Rightarrow \sigma w(\cdot)$ , where  $w(\cdot)$  is a standard Brownian motion and the symbol  $\Rightarrow$  means weak convergence. Using an argument similar to that of [24, Section 9.6], we obtain

$$\varepsilon \sum_{i=1}^l \sum_{t=1}^{m_i} \sum_{k=0}^{t/\varepsilon-1} \Theta_k(s_{i,t}) x_k v_t' I_{\{\alpha_t^i \in \mathcal{A}_i\}} \Rightarrow \int_0^t \Xi(\bar{\alpha}(s)) x(s) ds.$$

Putting all the above estimates together, the desired result is obtained.

To proceed, define piecewise constant interpolation  $v^\varepsilon(x, t, \alpha) = v_n^\varepsilon(x, \alpha)$  for  $t \in [n\varepsilon, (n+1)\varepsilon)$ . We derive the convergence of  $v^\varepsilon(x, t, \alpha)$  and derive the limit Riccati equations.

**THEOREM 2.2.** *Under the conditions of Theorem 2.1, as  $\varepsilon \rightarrow 0$ , the sequence of value functions also converges. In fact, we have the following limits for the Riccati equations (2.7) and (2.8). For each  $i = 1, \dots, l$ ,*

$$\begin{cases} \dot{\bar{K}}(t, i) = -\bar{K}(t, i)\bar{A}(i) - \bar{A}'(i)\bar{K}(t, i) - \bar{M}(i) \\ \quad + \bar{K}(t, i)\bar{B}N^{-1}B'(i)\bar{K}(t, i) - \bar{Q}\bar{K}(t, \cdot)(i), \\ \bar{K}(T, i) = D \end{cases} \tag{2.14}$$

and

$$\begin{cases} \dot{\bar{r}}(t, i) = -\text{tr}(\bar{K}(t, i)\Sigma) - \bar{Q}\bar{r}(t, \cdot)(i), \\ \bar{r}(T, i) = 0, \end{cases} \tag{2.15}$$

where  $\bar{Q}\bar{K}(t, \cdot)(i)$  and  $\bar{Q}\bar{r}(t, \cdot)(i)$  are defined in (1.1).

**PROOF.** For any  $t, s \in [0, T]$  satisfying  $t + s \leq T$ , again, for simplicity, we use  $(t + s)/\varepsilon$  and  $t/\varepsilon$  in lieu of  $\lfloor (t + s)/\varepsilon \rfloor$  and  $\lfloor t/\varepsilon \rfloor$ . Similar to the proof of Lemma 3 in [28], since  $v^\varepsilon(x, t + s, \alpha)$  and  $v^\varepsilon(x, t, \alpha)$  are the minimal costs, we have

$$v^\varepsilon(x, t + s, \alpha) - v^\varepsilon(x, t, \alpha) = \varepsilon E \sum_{k=t/\varepsilon}^{(t+s)/\varepsilon-1} (x_k' M(\alpha_k) x_k + x_k' \Phi_k'(\alpha_k) N(\alpha_k) \Phi_k(\alpha_k) x_k).$$

By the boundedness of  $M(\cdot)$ ,  $N(\cdot)$ , and  $K_{k+1}^\varepsilon$  and using Lemma 2.2, we can show that for any  $\eta > 0$ , there is a  $\Delta > 0$  such that

$$\limsup_{\varepsilon \rightarrow 0} \sup_{\substack{0 \leq t \leq \Delta \\ 0 \leq t+s \leq T}} |v^\varepsilon(x, t + s, \alpha) - v^\varepsilon(x, t, \alpha)| \leq \eta.$$

Thus  $v^\epsilon(\cdot)$  is equicontinuous in the extended sense (see [13, p. 73] for a definition). Note that  $v^\epsilon(x, t, \alpha) = x'K^\epsilon(t, \alpha)x + r^\epsilon(t, \alpha)$ , taking  $x = 0$  in the above yields the equicontinuity of  $r^\epsilon(\cdot)$  in the extended sense. Since  $x'K^\epsilon(t, \alpha)x = v^\epsilon(x, t, \alpha) - r^\epsilon(t, \alpha)$ , the quadratic form is equicontinuous in the extended sense. By repeatedly choosing appropriate vector  $x$ 's, we can show all the entries of  $K^\epsilon(t, \alpha)$  are equicontinuous in the extended sense. Thus  $K^\epsilon(t, \alpha)$  is equicontinuous in the extended sense.

Take  $i = s_{i\ell} \in \mathcal{M}_i$ . Since  $\{K^\epsilon(\cdot, i)\}$  is equicontinuous in the extended sense and is uniformly bounded, Theorem 4.2.2 in [13] implies that it has a subsequence which converges uniformly to a continuous limit  $K(\cdot, i)$ . We next characterise the limit. Using (2.10), we can write

$$\epsilon A'(i)\widehat{K}_{n+1}(i) = \epsilon A'(i)K_{n+1}(i) + g_n^\epsilon, \quad \text{where} \quad g_n^\epsilon = \epsilon^2 A'(i) \sum_{j=1}^m q_{ij}^\epsilon K_{n+1}(j).$$

Similarly, we can treat the other terms involving  $K_{n+1}^\epsilon(i)$ . Note that  $Q^\epsilon$  defined in (2.11) is a generator of a singularly perturbed Markov chain in continuous time with  $Q^\epsilon = \widetilde{Q}/\epsilon + Q$  and  $\widetilde{Q} = P - I = \text{diag}(P_1 - I_1, \dots, P_l - I_l)$ , where  $I_\iota$  denotes an identity matrix of dimension  $\iota \times \iota$ . (The asymptotic properties of such singularly perturbed Markov chains have been studied extensively in [24].) Since  $i = s_{i\ell}$ , we can write (2.9) as

$$\begin{aligned} K_{n+1}^\epsilon(s_{i\ell}) &= K^\epsilon(s_{i\ell}) - \epsilon A'(s_{i\ell})K_{n+1}^\epsilon(s_{i\ell}) - \epsilon K_{n+1}^\epsilon(s_{i\ell}) - \epsilon M(s_{i\ell}) \\ &\quad + K_{n+1}^\epsilon(s_{i\ell})B(s_{i\ell})(N(s_{i\ell}) + O(\epsilon))^{-1}B'(s_{i\ell})K_{n+1}^\epsilon(s_{i\ell}) \\ &\quad - \epsilon Q^\epsilon K_{n+1}^\epsilon(\cdot)(s_{i\ell}) + g_n^\epsilon. \end{aligned}$$

Consequently,

$$\begin{aligned} &K^\epsilon(t + s, s_{i\ell}) - K^\epsilon(t, s_{i\ell}) \\ &= -\epsilon \sum_{k=t/\epsilon}^{(t+s)/\epsilon-1} [A'(s_{i\ell})K_{k+1}^\epsilon(s_{i\ell}) + K_{k+1}^\epsilon(s_{i\ell})A(s_{i\ell}) + M(s_{i\ell})] \\ &\quad + \epsilon \sum_{k=t/\epsilon}^{(t+s)/\epsilon-1} K_{k+1}^\epsilon(s_{i\ell})B(s_{i\ell})(N(s_{i\ell}) + O(\epsilon))^{-1}B'(s_{i\ell})K_{k+1}^\epsilon(s_{i\ell}) \\ &\quad - \epsilon \sum_{k=t/\epsilon}^{(t+s)/\epsilon-1} \dot{Q}^\epsilon K_{k+1}^\epsilon(\cdot)(s_{i\ell}) + G^\epsilon(t), \end{aligned} \tag{2.16}$$

where  $G^\epsilon(t) = \sum_{k=t/\epsilon}^{(t+s)/\epsilon-1} g_k^\epsilon$ . It can be shown that the terms in  $\widetilde{G}^\epsilon(t)$  converge to 0 uniformly in  $t \in [0, T]$ . Thus we need only examine the rest of the terms in (2.16).

By virtue of the definition of (2.11) and the uniform boundedness of  $K^\varepsilon(\cdot)$ , multiplying both sides of (2.16) by  $\varepsilon$  and sending  $\varepsilon \rightarrow 0$  leads to

$$\int_t^{t+s} (P_t - I_t)K(\tau, \cdot)(s_{i\ell}) d\tau = \lim_{\varepsilon \rightarrow 0} \varepsilon \sum_{k=t/\varepsilon}^{(t+s)/\varepsilon-1} (P_t - I_t)K_{k+1}^\varepsilon(\cdot)(s_{i\ell}) = 0.$$

The continuity of  $K(\cdot, t)$  then implies  $(P_t - I_t)K(t, \cdot)(s_{i\ell}) = 0$  for all  $t \in [0, T]$ . Irreducibility in turn implies that  $K(t, s_{i\ell}) = K(t, t)$ , independent of  $\ell$ , and that

$$\varepsilon \sum_{k=t/\varepsilon}^{(t+s)/\varepsilon-1} Q^\varepsilon K_{k+1}^\varepsilon(\cdot)(s_{i\ell}) \rightarrow \int_t^{t+s} QK(\tau, \cdot)(t) d\tau.$$

Using this fact, multiplying (2.16) by  $v_t^i$  and taking summation over  $\ell$  yields

$$\begin{aligned} & \sum_{\ell=1}^{m_t} v_t^i [K^\varepsilon(t+s, s_{i\ell}) - K^\varepsilon(t, s_{i\ell})] \\ &= - \int_t^{t+s} \sum_{\ell=1}^{m_t} v_t^i (A'(s_{i\ell})K^\varepsilon(\tau, s_{i\ell}) + K^\varepsilon(\tau, s_{i\ell})A(s_{i\ell})) + M(s_{i\ell}) d\tau \\ &+ \int_t^{t+s} \sum_{\ell=1}^{m_t} v_t^i [K^\varepsilon(\tau, s_{i\ell})B(s_{i\ell})N^{-1}(s_{i\ell})B'(s_{i\ell})K^\varepsilon(\tau, s_{i\ell})] d\tau \\ &- \int_t^{t+s} \left( \sum_{\ell=1}^{m_t} v_t^i Q\mathbf{1}_{m_t} \right) K(\tau, \cdot)(t) d\tau + o(1), \end{aligned}$$

where  $o(1) \rightarrow 0$  as  $\varepsilon \rightarrow 0$  uniformly in  $t$ .

Letting  $\varepsilon \rightarrow 0$  and using the uniform convergence of  $K^\varepsilon(t, s_{i\ell}) \rightarrow K(t, t)$ , we will show for each  $t \in \overline{\mathcal{M}}$ ,  $K(t, t) = \overline{K}(t, t)$ . In fact,

$$\left( \sum_{\ell=1}^{m_t} v_t^i Q\mathbf{1}_{m_t} \right) K(t, \cdot)(t) = \overline{Q}K(t, \cdot)(t).$$

Then  $\sum_{\ell=1}^{m_t} v_t^i = 1$  leads to

$$\begin{aligned} K(t+s, t) - K(t, t) &= - \int_t^{t+s} [\overline{A}'(t)K(\tau, t) + K(\tau, t)\overline{A}(t) + \overline{M}(t) \\ &- K(\tau, t)\overline{BN^{-1}B'}(t)K(\tau, t) + \overline{Q}K(\tau, \cdot)(t)] d\tau. \end{aligned}$$

Finally, the uniqueness of the Riccati equation (see [8, Chapter VI]) implies that  $K(t, t) = \overline{K}(t, t)$ . As a result,  $K^\varepsilon(t, s_{i\ell}) \rightarrow \overline{K}(t, t)$ . This yields the desired limit.

**2.1. Numerical results** This section presents a numerical example of a four-state Markov chain  $\alpha_k^\varepsilon \in \mathcal{M} = 1, 2, 3, 4$ , with transition probability matrix  $P^\varepsilon = P + \varepsilon Q$ , where

$$P = \begin{pmatrix} 0.50 & 0.50 & 0 & 0 \\ 0.55 & 0.45 & 0 & 0 \\ 0 & 0 & 0.4 & 0.6 \\ 0 & 0 & 0.5 & 0.5 \end{pmatrix}, \quad Q = \begin{pmatrix} -0.6 & 0 & 0.3 & 0.3 \\ 0 & -0.3 & 0.1 & 0.2 \\ 0.2 & 0.3 & -0.5 & 0 \\ 0.1 & 0.3 & 0 & -0.4 \end{pmatrix}.$$

For a two-dimensional dynamic system (2.2) and the cost function (2.3), let  $x_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ ,  $\Sigma = \begin{pmatrix} 1.5 & 0.5 \\ 0.5 & 2.0 \end{pmatrix}$ ,  $D = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ ,  $A(1) = \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix}$ ,  $A(2) = \begin{pmatrix} -2 & -1 \\ -1 & 1 \end{pmatrix}$ ,  $A(3) = \begin{pmatrix} -3 & -2 \\ -2 & 0 \end{pmatrix}$ ,  $A(4) = \begin{pmatrix} -4 & -3 \\ -3 & -1 \end{pmatrix}$ ,  $B(1) = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ ,  $B(2) = \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}$ ,  $B(3) = \begin{pmatrix} 3 & 4 \\ 4 & 6 \end{pmatrix}$ ,  $B(4) = \begin{pmatrix} 4 & 5 \\ 5 & 7 \end{pmatrix}$ ,  $M(1) = \begin{pmatrix} 5 & 3 \\ 3 & 7 \end{pmatrix}$ ,  $M(2) = \begin{pmatrix} 4 & 3/2 \\ 3/2 & 5 \end{pmatrix}$ ,  $M(3) = \begin{pmatrix} 11/3 & 1 \\ 1 & 13/3 \end{pmatrix}$ ,  $M(4) = \begin{pmatrix} 7/2 & 3/4 \\ 3/4 & 4 \end{pmatrix}$ ,  $N(1) = \begin{pmatrix} 8 & 3 \\ 3 & 10 \end{pmatrix}$ ,  $N(2) = \begin{pmatrix} 10 & 6 \\ 6 & 14 \end{pmatrix}$ ,  $N(3) = \begin{pmatrix} 12 & 9 \\ 9 & 18 \end{pmatrix}$ ,  $N(4) = \begin{pmatrix} 14 & 12 \\ 12 & 22 \end{pmatrix}$ . The time horizon for this discrete-time model is  $0 \leq n \leq \lfloor T/\varepsilon \rfloor$  with  $T = 5$ . We use step size  $h = 0.01$  to discretise the limit Riccati equations.

Take  $\alpha_0^\varepsilon = 1$ . The trajectories of  $r_n^\varepsilon(i)$  versus  $\bar{r}(t)$ ,  $K_n^\varepsilon(i)$  versus  $\bar{K}(t, \cdot)$  and  $v_n^\varepsilon(x, i)$  versus  $v(\cdot)$  are given in Figure 1 for  $\varepsilon = 0.01$ . The simulation results show that the discrete-time LQ regulator problem is closely approximated by the corresponding continuous-time hybrid LQG problem, which allows us to further construct nearly optimal controls for the original system.

### 3. A continuous-time hybrid LQG problem

As was mentioned in the previous section, much effort has been directed to the design in recent years of more “robust” controllers. In various applications, it is often necessary to develop models involving disturbances of discrete-event type in addition to the additive white noise. Introducing a Markovian jump model and extending the original “state space” model to cover both the original state variables of the LQG problem and those variables of the Markov chain will result in a new system that displays both continuous and discrete characteristics and is thus termed a hybrid system.

**3.1. Formulation** Again, for reduction of complexity, we use a singularly perturbed Markov chain  $\alpha^\varepsilon(t)$  whose state space is  $\mathcal{M} = \{1, \dots, m\}$ . For further motivation, the reader is referred to [28]. Let us work with a finite horizon for some finite  $T > 0$ . Consider the linear system

$$\begin{aligned} dx(t) &= [A(\alpha^\varepsilon(t))x(t) + B(\alpha^\varepsilon(t))u(t)]dt + \sigma dw(t), \\ x(s) &= x, \quad \text{for } s \leq t \leq T, \end{aligned} \tag{3.1}$$

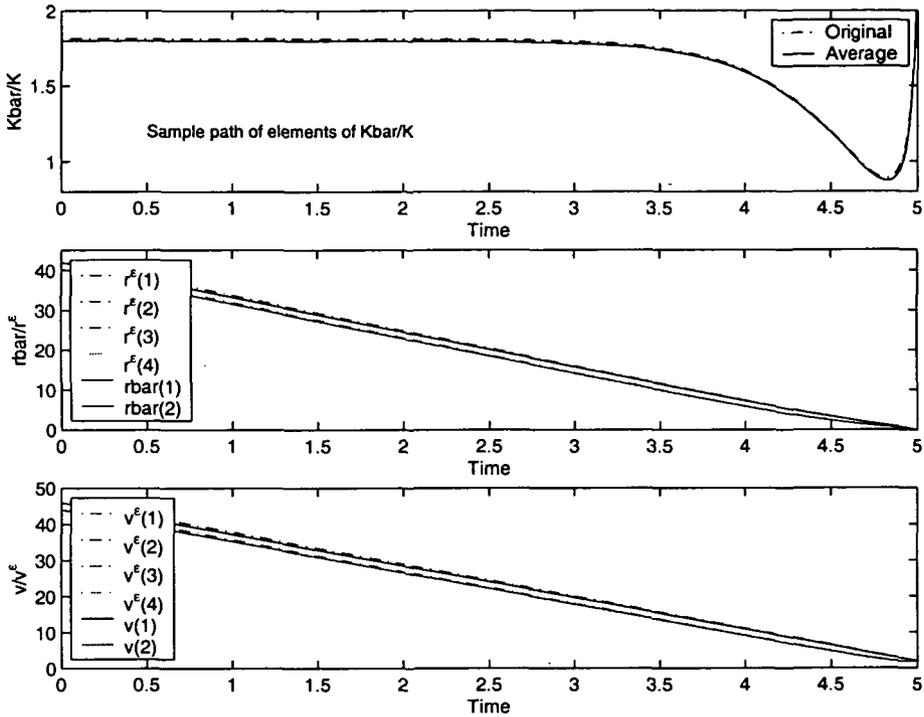


FIGURE 1. A discrete-time LQG.

where  $x(t) \in \mathbb{R}^{n_1}$  denotes state variables,  $u(t) \in \mathbb{R}^{n_2}$  denotes control variables,  $A(i) \in \mathbb{R}^{n_1 \times n_1}$  and  $B(i) \in \mathbb{R}^{n_1 \times n_2}$  are well-defined and have finite values for any  $i \in \mathcal{M}$  and  $w(\cdot)$  is a standard Brownian motion. Our objective is to find the optimal control  $u(\cdot)$  to minimise the expected quadratic cost function

$$\begin{aligned}
 & J(s, x, \alpha, u(\cdot)) \\
 &= E \left\{ \int_s^T [x'(t)M(\alpha^\epsilon(t))x(t) + u'(t)N(\alpha^\epsilon(t))u(t)] dt + x'(T)Dx(T) \right\}, \quad (3.2)
 \end{aligned}$$

where  $E = E_{x,\alpha}$  is the expectation given  $\alpha(s) = \alpha$  and  $x(s) = x$ ;  $M(i)$ ,  $i = 1, \dots, m$ , are symmetric nonnegative definite matrices;  $N(i)$ ,  $i = 1, \dots, m$ , and  $D$  are symmetric positive definite matrices;  $\alpha^\epsilon(\cdot)$  and  $w(\cdot)$  are independent.

Having a different interpretation than that of the discrete-time counterpart, the generator of  $\alpha^\epsilon(t)$  consists of two parts, a rapidly changing part and a slowly varying one, and is given by

$$Q^\epsilon = \tilde{Q}/\epsilon + \hat{Q}. \quad (3.3)$$

Note that  $\tilde{Q}/\varepsilon$  represents the fast changing part and  $\hat{Q}$  represents the slowly varying part. As a result, the form of (3.3) is similar in spirit to that of (2.1), where the transition matrix  $P^\varepsilon$  consists of  $P$  and a “slowly varying” part  $\varepsilon Q$ . A small parameter  $\varepsilon > 0$  makes the system under consideration display a two-time-scale behaviour [6, 17].

**3.2. Optimal control** Let  $v^\varepsilon(s, \alpha, x) = \inf_{u(\cdot)} J^\varepsilon(s, \alpha, x, u(\cdot))$  be the value function. Then  $v^\varepsilon$  satisfies the following system of Hamilton-Jacobi-Bellman (HJB) equations: for  $0 \leq s \leq T$  and  $i \in \mathcal{M}$ ,

$$0 = \frac{\partial v^\varepsilon(s, i, x)}{\partial s} + \min_u \left\{ (A(i)x + B(i)u)' \frac{\partial v^\varepsilon(s, i, x)}{\partial x} + x' M(i)x + u' N(i)u + \frac{1}{2} \text{tr} \left( \sigma \sigma' \frac{\partial^2 v^\varepsilon(s, i, x)}{\partial x^2} \right) + Q^\varepsilon v^\varepsilon(s, \cdot, x)(i) \right\}, \tag{3.4}$$

with the boundary condition  $v^\varepsilon(T, i, x) = x' D x$ , where  $Q^\varepsilon v^\varepsilon(s, \cdot, x)(i)$  is defined in (1.1).

Assume  $\tilde{Q}$  has a block-diagonal form  $\tilde{Q} = \text{diag}(\tilde{Q}^1, \dots, \tilde{Q}^l)$ , where  $\tilde{Q}^\iota \in \mathbb{R}^{m_\iota \times m_\iota}$  are weakly irreducible (see [24, p. 21], for a definition), for  $\iota = 1, \dots, l$ , and  $\sum_{\iota=1}^l m_\iota = m$ . Let  $\mathcal{M}_\iota = \{s_{i,1}, \dots, s_{i,m_\iota}\}$  for  $\iota = 1, \dots, l$  be the state space corresponding to  $\tilde{Q}^\iota$  with decomposition of the form (1.2). The slow and fast components are coupled through weak and strong interactions in the sense that the underlying Markov chain fluctuates rapidly within a single group  $\mathcal{M}_\iota$  and jumps less frequently among different groups. If we consider the states in  $\mathcal{M}_\iota$  as a single “state,” then all such “states” are coupled through the matrix  $\hat{Q}$ .

Following the approach in [8, pp. 165–166] (see also [24, pp. 309–325]), let

$$v^\varepsilon(s, i, x) = x' K^\varepsilon(s, i)x + r^\varepsilon(s, i). \tag{3.5}$$

Again, the  $m \times m$  matrix-valued functions  $K^\varepsilon(\cdot)$  and real-valued functions  $r^\varepsilon(\cdot)$  are to be determined. Substituting (3.5) into (3.4) and comparing the coefficients of  $x$  leads to the following Riccati equations for  $K^\varepsilon(s, i)$ :

$$\begin{cases} \dot{K}^\varepsilon(s, i) = -K^\varepsilon(s, i)A(i) - A'(i)K^\varepsilon(s, i) - M(i) \\ \quad + K^\varepsilon(s, i)B(i)N^{-1}(i)B'(i)K^\varepsilon(s, i) - Q^\varepsilon K^\varepsilon(s, \cdot)(i), \\ K^\varepsilon(T, i) = D, \end{cases} \tag{3.6}$$

and

$$\begin{cases} \dot{r}^\varepsilon(s, i) = -\text{tr}(\sigma \sigma' K^\varepsilon(s, i)) - Q^\varepsilon r^\varepsilon(s, \cdot)(i), \\ r^\varepsilon(T, i) = 0, \end{cases} \tag{3.7}$$

where  $Q^\varepsilon K^\varepsilon(s, \cdot)(i)$  is as defined in (1.1). Moreover, similar to [8, Chapter VI] (see also [24, Appendices A.4 and A.5]), it is easy to show that these equations have a

unique solution. In view of the positive definiteness of  $K^\epsilon$ , the optimal control  $u^{\epsilon,*}$  has the form

$$u^{\epsilon,*}(s, i, x) = -N^{-1}(i)B'(i)K^\epsilon(s, i)x. \tag{3.8}$$

By aggregating the states in  $\mathcal{M}_i$  into one state  $\iota$ , we obtain an aggregated process  $\{\bar{\alpha}^\epsilon(\cdot)\}$  defined by  $\bar{\alpha}^\epsilon(t) = \iota$  when  $\alpha^\epsilon(t) \in \mathcal{M}_i$ . The process  $\bar{\alpha}^\epsilon(\cdot)$  is not necessarily Markovian. However, using probabilistic arguments, we have shown in [24, Section 7.2] that as  $\epsilon \rightarrow 0$ ,  $\bar{\alpha}^\epsilon(\cdot)$  converges weakly to  $\bar{\alpha}(\cdot)$  generated by  $\bar{Q}$  that has the form (2.13) with  $\nu^\iota$  being the stationary distribution of  $\tilde{Q}^\iota$  (for each  $\iota = 1, \dots, l$ ). Moreover, for any bounded and measurable deterministic function  $\beta(\cdot)$ ,

$$E \left( \int_s^T [I_{\{\alpha^\epsilon(t)=s_{i\ell}\}} - \nu^\iota I_{\{\bar{\alpha}^\epsilon(t)=\iota\}}] \beta(t) dt \right)^2 = O(\epsilon). \tag{3.9}$$

The following theorem is concerned with the convergence of  $K^\epsilon$  and  $r^\epsilon$ , whose proof is in [28]. For appropriate functions  $F(\cdot)$  and  $F_i(\cdot)$  on  $\mathcal{M}$ , define

$$\bar{F} = \sum_{\ell=1}^{m_i} \nu^\iota F(s_{i\ell}), \quad \bar{F}_1 \bar{F}_2 = \sum_{\ell=1}^{m_i} \nu^\iota F_1(s_{i\ell}) F_2(s_{i\ell}), \quad \text{for } \iota = 1, \dots, l.$$

**THEOREM 3.1.** For  $\iota = 1, \dots, l$  and  $\ell = 1, \dots, m_i$ ,  $K^\epsilon(s, s_{i\ell}) \rightarrow \bar{K}(s, \iota)$  and  $r^\epsilon(s, s_{i\ell}) \rightarrow \bar{r}(s, \iota)$ , uniformly on  $[0, T]$  as  $\epsilon \rightarrow 0$ , where  $\bar{K}(s, \iota)$  and  $\bar{r}(s, \iota)$  are the unique solutions to the differential equations (2.14) and (2.15), respectively, with  $\Sigma = \sigma\sigma'$ .

The convergence of  $K^\epsilon(s, i)$  and  $r^\epsilon(s, i)$  leads to that of  $v^\epsilon(s, i, x)$  given by (3.5) where  $K^\epsilon(s, i)$  and  $r^\epsilon(s, i)$  are the solutions to the differential equations (3.6) and (3.7), respectively. It follows that for  $\ell = 1, \dots, m_i$ , as  $\epsilon \rightarrow 0$ ,  $v^\epsilon(s, s_{i\ell}, x) \rightarrow v(s, \iota, x) = x' \bar{K}(s, \iota)x + \bar{r}(s, \iota)$  corresponds to the value function of a limit problem. Let  $\mathcal{U}$  denote the control set for the limit problem:

$$\mathcal{U} = \{ U = (U^1, \dots, U^l) : U^\iota = (u^{\iota 1}, \dots, u^{\iota m_i}), u^{\iota \ell} \in \mathbb{R}^{n_2} \}.$$

Define

$$f(s, \iota, x, U) = \bar{A}(s, \iota)x + \sum_{j=1}^{m_i} \nu^\iota B(s_{i\ell})u^{\iota j}, \quad \tilde{N}(s, U) = \sum_{j=1}^{m_i} \nu^\iota (u^{\iota j})' N(s_{i\ell})u^{\iota j}.$$

Then  $v(s, \iota, x)$  satisfies the following HJB equations:

$$0 = \frac{\partial v(s, \iota, x)}{\partial s} + \min_{U \in \mathcal{U}} \left\{ f(s, \iota, x, U) \frac{\partial v(s, \iota, x)}{\partial x} + x' \bar{M}(s, \iota)x + \tilde{N}(s, U) + \frac{1}{2} \text{tr} \left( \sigma \sigma' \frac{\partial^2 v(s, \iota, x)}{\partial x^2} \right) + \bar{Q}v(s, \cdot, x)(s) \right\},$$

$$v(T, \iota, x) = x' D x.$$

The corresponding control problem is

$$\begin{aligned} \text{minimise } J(s, \iota, x, U(\cdot)) &= E \left\{ \int_s^T [x'(t)\overline{M}(\overline{\alpha}(t))x(t) + \tilde{N}(\overline{\alpha}(t), U(t))] dt \right. \\ &\quad \left. + x'(T)Dx(T) \right\} \\ \text{s.t. } dx(t) &= f(t, \overline{\alpha}(t), x(t), U(t)) dt + \sigma dw(t), \quad x(s) = x, \end{aligned}$$

where  $\overline{\alpha}(\cdot) \in \{1, \dots, l\}$  is a Markov chain generated by  $\overline{Q}$ .

The optimal control for this limit problem is:

$$U^o(s, \iota, x) = (U^{1o}(s, x), \dots, U^{lo}(s, x))$$

with

$$\begin{aligned} U^{io}(s, x) &= (u^{i1,o}(s, x), \dots, u^{im_i,o}(s, x)) \quad \text{and} \\ u^{i\ell,o}(s, x) &= -N^{-1}(s_{i\ell})B'(s_{i\ell})\overline{K}(s, \iota)x. \end{aligned}$$

Using such controls (as in [19], see also [24, Chapter 9]), we construct

$$u^\epsilon(s, \alpha, x) = \sum_{i=1}^l \sum_{\ell=1}^{m_i} I_{\{\alpha=s_{i\ell}\}} u^{i\ell,o}(s, x). \tag{3.10}$$

for the original problem. Although (3.10) involves a summation, at any given instance, it has only one term. Equivalently, this control can also be written as if  $\alpha \in \mathcal{M}_i$ ,

$$u^\epsilon(s, \alpha, x) = -N^{-1}(\alpha)B'(\alpha)\overline{K}(s, \iota)x.$$

It is clear that this control is identical to the optimal control in (3.8) except  $K^\epsilon$  is replaced by  $\overline{K}$ . We use  $u^\epsilon(t) = u^\epsilon(t, \alpha^\epsilon(t), x(t))$  for the original problem, which is nearly optimal.

If  $B(s_{i\ell}) = B(i)$  and  $N(s_{i\ell}) = N(i)$  are independent of  $\ell$ , then, in view of (3.9), we may replace  $I_{\{\alpha^\epsilon(t)=s_{i\ell}\}}$  by  $I_{\{\overline{\alpha}^\epsilon(t)=i\}} v_i^\ell$  and consider

$$\overline{u}^\epsilon(s, \alpha, x) = \sum_{i=1}^l \sum_{\ell=1}^{m_i} I_{\{\alpha \in \mathcal{M}_i\}} v_i^\ell u^{i\ell,o}(s, x) = -N^{-1}(i)B'(i)\overline{K}(s, \iota)x, \quad \text{if } \alpha \in \mathcal{M}_i.$$

Therefore we can write  $\overline{u}^\epsilon(s, \alpha, x) = \overline{u}^\epsilon(s, \iota, x)$ . Since  $\overline{u}^\epsilon$  only requires the information  $\alpha^\epsilon(t) \in \mathcal{M}_i$ , we can use

$$\overline{u}^\epsilon(t) = \overline{u}^\epsilon(t, \overline{\alpha}^\epsilon(t), x(t)). \tag{3.11}$$

**THEOREM 3.2.** *The following assertions hold:*

- (1) *The control  $u^\epsilon(t)$  defined in (3.10) is nearly optimal, that is,*

$$\lim_{\epsilon \rightarrow 0} |J^\epsilon(s, \alpha, x, u^\epsilon(\cdot)) - v^\epsilon(s, \alpha, x)| = 0.$$

- (2) *Assume  $B(s_{i\ell}) = B(t)$  and  $N(s_{i\ell}) = N(t)$  independent of  $\ell$ . Then  $\bar{u}^\epsilon(t)$  defined in (3.11) is nearly optimal, that is,  $\lim_{\epsilon \rightarrow 0} |J^\epsilon(s, \alpha, x, \bar{u}^\epsilon(\cdot)) - v^\epsilon(s, \alpha, x)| = 0$ .*

**3.3. Numerical results** This section presents a numerical example of a four-state Markov chain  $\alpha^\epsilon(t) \in \mathcal{M} = 1, 2, 3, 4, t \geq 0$ , generated by

$$Q^\epsilon = \frac{1}{\epsilon} \begin{pmatrix} -0.10 & 0.10 & 0 & 0 \\ 0.12 & -0.12 & 0 & 0 \\ 0 & 0 & -0.04 & 0.04 \\ 0 & 0 & 0.06 & -0.06 \end{pmatrix} + \begin{pmatrix} -0.8 & 0 & 0.8 & 0 \\ 0 & -0.7 & 0 & 0.7 \\ 0.5 & 0 & -0.5 & 0 \\ 0 & 0.1 & 0 & -0.1 \end{pmatrix}.$$

For a two-dimensional dynamic system (3.1) and cost function (3.2), let  $x(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ ,  $\sigma = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$ ,  $D = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ ,  $A(1) = \begin{pmatrix} 0 & 1 \\ 1 & 3 \end{pmatrix}$ ,  $A(2) = \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix}$ ,  $A(3) = \begin{pmatrix} -2 & -1 \\ -1 & 1 \end{pmatrix}$ ,  $A(4) = \begin{pmatrix} -3 & -2 \\ -2 & 0 \end{pmatrix}$ ,  $B(1) = \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}$ ,  $B(2) = \begin{pmatrix} 3 & 4 \\ 4 & 6 \end{pmatrix}$ ,  $B(3) = \begin{pmatrix} 4 & 5 \\ 5 & 7 \end{pmatrix}$ ,  $B(4) = \begin{pmatrix} 5 & 6 \\ 6 & 8 \end{pmatrix}$ ,  $M(1) = \begin{pmatrix} 4 & 3 \\ 3 & 6 \end{pmatrix}$ ,  $M(2) = \begin{pmatrix} 3 & 1.5 \\ 1.5 & 4 \end{pmatrix}$ ,  $M(3) = \begin{pmatrix} 8/3 & 1.0 \\ 1.0 & 10/3 \end{pmatrix}$ ,  $M(4) = \begin{pmatrix} 2.5 & 0.75 \\ 0.75 & 3.0 \end{pmatrix}$ ,  $N(1) = \begin{pmatrix} 6 & 3 \\ 3 & 8 \end{pmatrix}$ ,  $N(2) = \begin{pmatrix} 8 & 6 \\ 6 & 12 \end{pmatrix}$ ,  $N(3) = \begin{pmatrix} 10 & 9 \\ 9 & 16 \end{pmatrix}$ ,  $N(4) = \begin{pmatrix} 12 & 12 \\ 12 & 20 \end{pmatrix}$ . The time horizon for the continuous-time model is  $[0, T]$  with  $T = 5$ . We discretise the system equations with step size  $h = 0.01$ .

Take  $s = 0$  and  $\alpha(0) = 1$ . Sample paths of  $\alpha^\epsilon(\cdot)$  are given in Figure 2 for  $\epsilon = 0.01$  and  $\epsilon = 0.1$ , respectively. It shows that a smaller  $\epsilon$  leads to more rapid jumps. The trajectories of  $r^\epsilon(\cdot)$  versus  $\bar{r}(\cdot)$ ,  $K^\epsilon(\cdot)$  versus  $\bar{K}(\cdot)$ , and  $v^\epsilon(\cdot)$  versus  $v(\cdot)$  are given in Figure 3. The simulation results demonstrate that our algorithm approximates the optimal solutions well with less computational effort.

#### 4. A discrete-time MDP problem

Discrete-time MDPs are of interest in various applications such as resource allocation, communication channels and queueing networks. Classical treatments of discrete-time MDP models can be found in [18] and [23] (see also [21] for applications to reliability models) among others. It is important to consider discrete-time systems since system measurements are frequently recorded in discrete time. In addition, success in obtaining optimal controls for the underlying system relies on efficient computation procedures, which in turn depend on the corresponding discretised dynamic systems. A common practice in dealing with MDPs is to use a dynamic programming (DP) approach, which requires solving a set of DP equations and finding

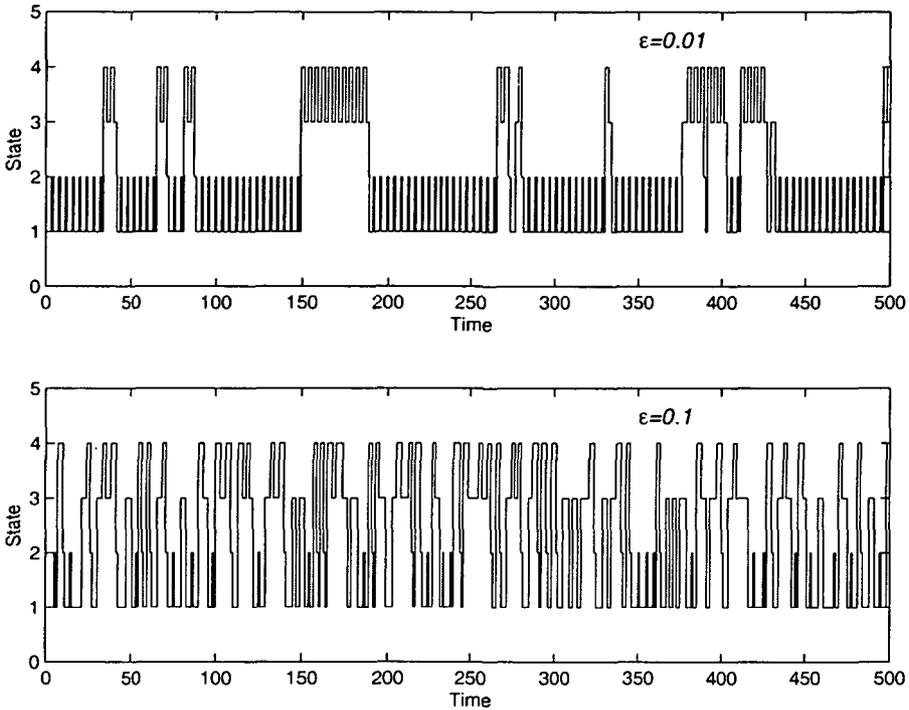


FIGURE 2. Continuous-time MC sample paths,  $\epsilon = 0.1$  and  $\epsilon = 0.01$ .

the optimal solutions. Such an approach becomes infeasible computationally when the dimension of the underlying system is very large, and therefore other alternatives are needed. In [14], a hierarchical control approach was proposed, and nearly optimal control strategies were developed. The essences are state space decomposition and aggregation. The asymptotic result is based on the asymptotic expansions obtained in [25]. We provide numerical results herein to further elaborate our theoretical findings and to offer insight for practical applications.

**4.1. Formulation and results** Consider a discrete-time Markov chain  $\{\alpha_n^\epsilon, n = 0, 1, \dots\}$  with finite state space  $\mathcal{M} = \{1, \dots, m\}$ , where  $\epsilon > 0$  is a small parameter. Let the control space  $\Gamma$  be a compact subset of an Euclidean space. Consider feedback control  $u_n = u(\alpha_n^\epsilon)$  such that  $u_n \in \Gamma, n = 0, 1, \dots$ . Let  $P^\epsilon(u_n) = (p_{ij}^\epsilon(u_n))_{m \times m}$  be the probability transition matrix of  $\alpha_n^\epsilon$  given by (2.1) with both  $P = P(u)$  and  $Q = Q(u)$  depending on  $u$ . In view of (2.1), it is clear that the dominating factor is given by the transition matrix  $P(u)$ . Since we are focusing on finite-state Markov chains, the MDP corresponding to  $P(u)$  can either consist of a number of recurrent (ergodic) classes of states or also include transient states.

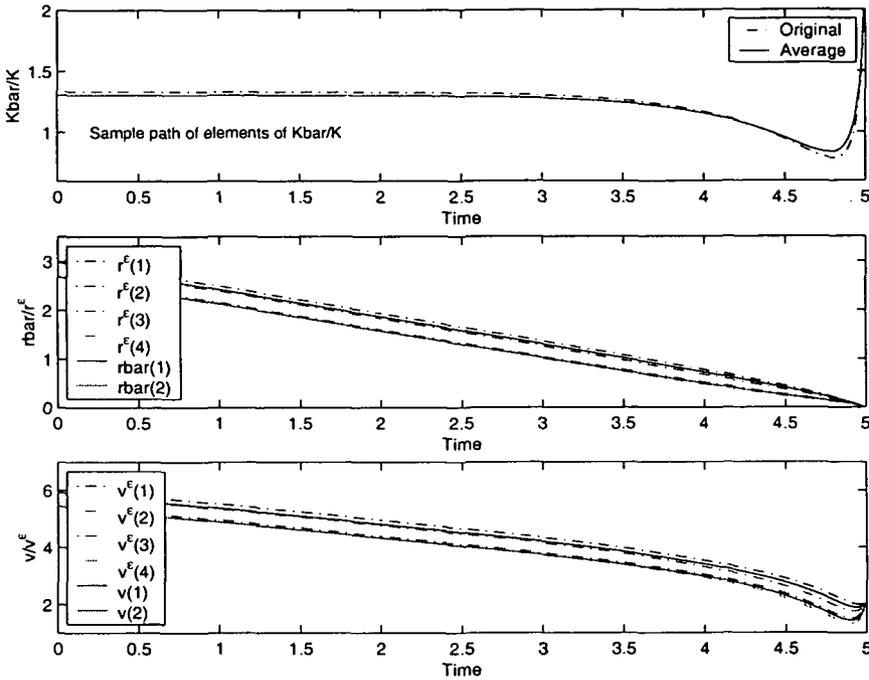


FIGURE 3. A continuous-time LQG.

For simplicity, we concentrate on the recurrent cases here. That is, the transition matrix  $P(u)$  consists only of  $l$  recurrent classes, and  $P(u)$  has the form (2.4) with each  $P_i = P_i(u)$ . For  $i = 1, \dots, l$ , let  $\mathcal{M}_i = \{s_{i,1}, \dots, s_{i,m_i}\}$  denote the recurrent sub-state space of  $\alpha_n^\epsilon$  corresponding to the block  $P_i(u)$ . The entire state space of  $\alpha_n^\epsilon$  can be decomposed as in (1.2).

For every  $u \in \Gamma$ , assume that for each  $i = 1, \dots, l$ ,  $P_i(u)$  is an  $m_i \times m_i$  transition matrix and is irreducible and aperiodic. Let  $u(\cdot) = \{u(i) : i \in \mathcal{M}\}$  be a function such that  $u(i) \in \Gamma$  for all states  $i \in \mathcal{M}$ . A function  $u(\cdot)$  is called an admissible control and the collection of all such functions is denoted by  $\mathcal{A}_f$ .

Consider the cost functional  $J^\epsilon(i, u(\cdot))$  defined on  $\mathcal{M} \times \mathcal{A}_f$ :

$$J^\epsilon(i, u(\cdot)) = E \left( \epsilon \sum_{k=0}^{\infty} (1 - \beta\epsilon)^k g(\alpha_k^\epsilon, u(\alpha_k^\epsilon)) \right), \tag{4.1}$$

where  $i = \alpha_0^\epsilon$  is the initial state of the chain,  $g(i, u)$  is the cost-to-go function, and  $\beta > 0$  is a given constant. The objective is to find a function  $u(\cdot) \in \mathcal{A}_f$  that minimises  $J^\epsilon(i, u(\cdot))$ . In the discrete-time setting, the discount factor  $(1 - \beta\epsilon)$  depends on  $\epsilon$ . Such a dependence leads to cancellation of  $\epsilon$  in the associated DP equations.

The original MDP problem, termed  $\mathcal{P}^\varepsilon$ , is of the form

$$\mathcal{P}^\varepsilon : \begin{cases} \text{minimise:} & J^\varepsilon(i, u(\cdot)) = E \left( \varepsilon \sum_{k=0}^{\infty} (1 - \beta\varepsilon)^k g(\alpha_k^\varepsilon, u(\alpha_k^\varepsilon)) \right), \\ \text{subject to:} & \alpha_k^\varepsilon \sim P^\varepsilon(u(\alpha_k^\varepsilon)), k = 0, 1, \dots, \alpha_0^\varepsilon = i, u(\cdot) \in \mathcal{A}_f, \\ \text{value function:} & v^\varepsilon(i) = \inf_{u(\cdot) \in \mathcal{A}_f} J^\varepsilon(i, u(\cdot)). \end{cases}$$

The notation  $\alpha_n^\varepsilon \sim P^\varepsilon(u(\alpha_n^\varepsilon))$  means that  $\alpha_n^\varepsilon$  is a discrete-time Markov chain whose probability transition matrix is  $P^\varepsilon(u(\alpha_n^\varepsilon))$ .

To solve the problem  $\mathcal{P}^\varepsilon$  by using the DP approach, for each  $i \in \mathcal{M}$ , the associated discrete-time DP equation is

$$v^\varepsilon(i) = \min_{u \in \Gamma} \left\{ \varepsilon g(i, u) + (1 - \beta\varepsilon) \sum_j p_{ij}^\varepsilon(u) v^\varepsilon(j) \right\}.$$

As was shown in [14], as  $\varepsilon \rightarrow 0$ , a reduced problem or limit problem in an appropriate sense is obtained. Although the original problem is a discrete-time one, the reduced problem becomes a continuous-time MDP.

To proceed, define  $\bar{g}(\iota, U^\iota) = \sum_{t=1}^{m_\iota} v_t^\iota(U^\iota) g(s_{\iota t}, u^{\iota t})$ ,  $\iota = 1, \dots, l$ . For each  $\iota = 1, \dots, l$ , define

$$\begin{aligned} \Gamma_\iota &= \{U^\iota := (u^{\iota 1}, \dots, u^{\iota m_\iota}) : u^{\iota \ell} \in \Gamma, \ell = 1, \dots, m_\iota\}, \\ \bar{\Gamma} &= \Gamma_1 \times \dots \times \Gamma_l = \{U = (U^1, \dots, U^l) : U^\iota \in \Gamma_\iota, \iota = 1, \dots, l\}. \end{aligned}$$

Let  $\mathcal{A}^0$  denote a class of functions  $U(\cdot) = \{U(\iota) : \iota \in \bar{\mathcal{M}}\}$  such that  $U(\iota) \in \Gamma_\iota$  for  $\iota = 1, \dots, l$ . For convenience, call  $U = (U(1), \dots, U(l)) \in \mathcal{A}^0$  an admissible control for the limit problem, termed  $\mathcal{P}^0$ . Use  $\tilde{x}(t) \sim \bar{Q}(U(\tilde{x}(t)))$  to denote that  $\tilde{x}(t)$  is a Markov chain generated by  $\bar{Q}(U(\tilde{x}(t)))$ .

Now we have a continuous-time MDP:

$$\mathcal{P}^0 : \begin{cases} \text{minimise:} & J^0(\iota, U) = E \int_0^\infty e^{-\beta t} \bar{g}(\tilde{x}(t), U(\tilde{x}(t))) dt, \\ \text{subject to:} & \tilde{x}(t) \sim \bar{Q}(U(\tilde{x}(t))), t \geq 0, \tilde{x}(0) = k, U \in \mathcal{A}^0, \\ \text{value function:} & v(\iota) = \inf_{U \in \mathcal{A}^0} J^0(\iota, U). \end{cases}$$

Problem  $\mathcal{P}^0$  is the limit problem in an appropriate sense.

The DP equation for the limit problem  $\mathcal{P}^0$  is

$$\beta v(\iota) = \min_{U^\iota \in \Gamma_\iota} \{ \bar{g}(\iota, U^\iota) + \bar{Q}(U^\iota)v(\cdot)(\iota) \}, \quad \text{for } \iota = 1, \dots, l. \tag{4.2}$$

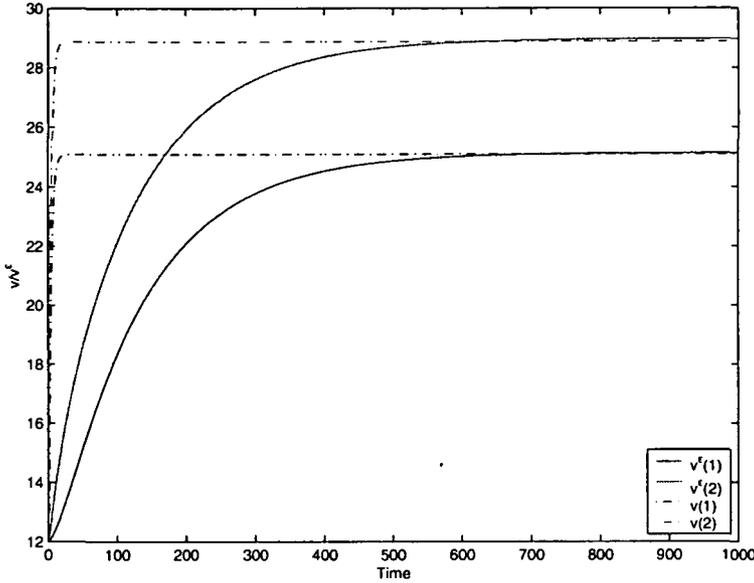


FIGURE 4. A discrete-time MDP.

Let  $U_o = (U_o^1, \dots, U_o^l) \in \bar{\Gamma}$  denote a minimiser of the right-hand side of (4.2). Then  $U_o \in \mathcal{A}^0$  is optimal for  $\mathcal{P}^0$ . We need the following assumptions on the transition matrix  $P^\epsilon(u)$  and the cost-to-go function  $g(i, u)$ :

(A2) For each  $\epsilon > 0$ ,  $P^\epsilon(u)$  is a continuous function of  $u$ . Moreover, for each  $U^i \in \Gamma_i, i = 1, \dots, l, P_i(U^i)$  is irreducible and aperiodic.

(A3) For each  $i \in \mathcal{M}, g(i, \cdot)$  is a continuous function on  $\Gamma$ .

For each  $U^i \in \Gamma_i, i = 1, \dots, l$ , let  $v^i(U^i) = (v_1^i(U^i), \dots, v_{m_i}^i(U^i))$  denote the stationary distribution of  $P_i(U^i)$ , that is,  $v^i(U^i)$  is the unique solution of the following system of equations:

$$\begin{cases} v^i(U^i)P_i(U^i) = v^i(U^i), \\ \sum_{j=1}^{m_i} v_j^i(U^i) = 1. \end{cases}$$

**THEOREM 4.1.** Assume (A2) and (A3). For each  $i = s_{i\ell} \in \mathcal{M}, i = 1, \dots, l, \lim_{\epsilon \rightarrow 0} v^\epsilon(i) = v(i)$ , where  $v(i)$  is the value function of the limit problem  $\mathcal{P}^0$ .

Let  $U_o = (U_o^1, \dots, U_o^l) \in \mathcal{A}^0$  be an optimal control for the limit problem  $\mathcal{P}^0$ . Define a control  $u_\epsilon(\cdot) = \{u_\epsilon(\alpha) : \alpha \in \mathcal{M}\}$  for the original problem  $\mathcal{P}^\epsilon$ :

$$u_\epsilon(\alpha) = \sum_{i=1}^l \sum_{\ell=1}^{m_i} I_{\{\alpha=s_{i\ell}\}} u_o^{i\ell}(\cdot). \tag{4.3}$$

Furthermore, the control  $u_\varepsilon(\cdot) = \{u_\varepsilon(\alpha)\}$  constructed in (4.3) is asymptotically optimal in that  $\lim_{\varepsilon \rightarrow 0} |J^\varepsilon(i, u_\varepsilon(\cdot)) - v^\varepsilon(i)| = 0$ , for  $i \in \mathcal{M}$ .

**4.2. Numerical results** This section presents a numerical example of a discrete-time MDP problem concerning a four-state Markov chain  $\alpha_n^\varepsilon \in \mathcal{M} = 1, 2, 3, 4$ ,  $n = 0, 1, \dots$ , with transition probability matrix given by (2.1). Note both  $P = P(u)$  and  $Q = Q(u)$  depend on  $u$ . In the meantime,  $u = u(\alpha_n^\varepsilon)$  are functions of states. We assume three possible control actions in each state and use the following specifications:

$$\begin{aligned}
 u(1) &= \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}, & u(2) &= \begin{pmatrix} 2 \\ 4 \\ 3 \end{pmatrix}, & u(3) &= \begin{pmatrix} 3 \\ 4 \\ 5/2 \end{pmatrix}, & u(4) &= \begin{pmatrix} 3/2 \\ 3 \\ 2 \end{pmatrix}, \\
 P(u) &= \begin{pmatrix} (u^{i_1}(1))^{-1} & 1 - (u^{i_1}(1))^{-1} & 0 & 0 \\ 1 - (u^{i_2}(2))^{-1} & (u^{i_2}(2))^{-1} & 0 & 0 \\ 0 & 0 & (u^{i_3}(3))^{-1} & 1 - (u^{i_3}(3))^{-1} \\ 0 & 0 & 1 - (u^{i_4}(4))^{-1} & (u^{i_4}(4))^{-1} \end{pmatrix}, \\
 Q(u) &= \begin{pmatrix} -u^{i_1}(1) & 0 & u^{i_1}(1) & 0 \\ 0 & -u^{i_2}(2) & 0 & u^{i_2}(2) \\ u^{i_3}(3) & 0 & -u^{i_3}(3) & 0 \\ 0 & u^{i_4}(4) & 0 & -u^{i_4}(4) \end{pmatrix},
 \end{aligned}$$

where  $u^r(i)$  denotes the  $r$ th control in state  $i$ , with all possible combinations of  $i_j = 1, 2, 3$  and  $j = 1, 2, 3, 4$ . To find the optimal control requires comparing the costs from all possible combinations of control actions. The cost-to-go function is given by (4.1) with  $g(i, u) = (i + u)^2$ .

The trajectories of  $v^\varepsilon(i)$ , the value function of the original problem, and  $v(i)$ , the value function of the limit problem, are displayed in Figure 4. The simulation results show that the original discrete-time problem can be approximated by a continuous-time MDP closely, which has confirmed our theoretical findings.

**REMARK 4.1.** To solve the limit control problem numerically, a cautionary note is in order. If one proceeds with a direct implementation, the numerical procedure may be unsuitable, which was confirmed by our numerical experiments. This is due partially to the result of the minimisation operation. Since a generator is involved and since the diagonal elements of the generator are negative, the iterates of the value functions may become negative resulting in the lost of probabilistic meaning in the iterations. To overcome these difficulties, we use the same procedure as suggested in [24, p. 240]. Numerically solving the limit problem leads to an equivalent discrete DP problem. To exploit the equivalence, we need to identify the corresponding transition probabilities. In fact, it can be shown (see the appendix) that the DP equation (4.2) is

equivalent to

$$v(l) = \min_{U' \in \Gamma_l} \left\{ \frac{\bar{g}(l, U')}{\beta + |\bar{q}_{ll}(U')|} + \sum_{\ell \neq l} \frac{\bar{q}_{\ell l}(U')}{\beta + |\bar{q}_{\ell l}(U')|} v(\ell) \right\}. \tag{4.4}$$

### 5. Further remarks

This paper has focused on numerical solutions for nearly optimal controls of singularly perturbed systems driven by Markov chains. It complements our theoretical investigation of the asymptotic properties of such large-scale systems. A central theme here is the reduction of complexity. It should be noted that the original problem may not include a small parameter. However, the presence of different rates of changes will allow us to introduce a small parameter  $\varepsilon > 0$  into the system; see [24, pp. 47–49] for an illustrative example. In many cases  $\varepsilon > 0$  may not be tending to 0, but rather a “small” constant. The asymptotics we obtained (as  $\varepsilon > 0$ ) provide guidance for practical models. In the original problem, the Markov chain can take  $m$  possible values, whereas in the reduced system, the number of states becomes  $l$ . If  $l \ll m$ , the complexity is significantly reduced. In the numerical experiments, we have used a value iteration approach. A policy improvement method may also be used.

### 6. Appendix

This appendix illustrates the procedure of generating Markovian sample paths for both discrete and continuous times in simulation studies, and shows that (4.4) is equivalent to (4.2).

**6.1. Discrete-time chains** To simulate a Markov chain  $\alpha_k$  in discrete time requires first prescribing its transition probability matrix  $P = (p_{ij})$ , then building its sample paths. Suppose that  $\alpha_k \in \mathcal{M} = \{1, \dots, m\}$ , for  $k \geq 0$ . The sample paths are constructed via comparison determined by the transition probability matrix  $P$ . At any time  $k \geq 0$ , the chain’s next move is specified by

$$\alpha_{k+1} = \begin{cases} 1, & \text{if } U < p_{i1}, \\ 2, & \text{if } p_{i1} < U < p_{i1} + p_{i2}, \\ \dots & \dots \\ m, & \text{if } p_{i1} + \dots + p_{i,m-1} < U < 1, \end{cases} \tag{6.1}$$

where  $U$  is a random variable following a uniform distribution in  $(0, 1)$  (that is,  $U \sim U(0, 1)$ ).

**6.2. Continuous-time chains** Suppose that  $\alpha(t)$  is a continuous-time Markov chain with state space  $\mathcal{M} = \{1, \dots, m\}$  and generator  $Q = (q_{ij})$ . To simulate the sample paths of  $\alpha(t)$  amounts to determining its sojourn time at each state and its subsequent moves. The chain sojourns in any given state  $i$  for a random length of time,  $S_i$ , which has an exponential distribution with parameter  $(-q_{ii})$ . Subsequently the process will enter another state. Each state  $j$  ( $j = 1, \dots, m, j \neq i$ ) has a probability of  $q_{ij}/(-q_{ii})$  of being the chain's next residence. Let a discrete random variable  $X_i$  denote the post-jump location and take values in  $\{1, 2, \dots, i - 1, i + 1, \dots, m\}$ . Its value is specified by

$$X_i = \begin{cases} 1, & \text{if } U < q_{i1}/(-q_{ii}), \\ 2, & \text{if } q_{i1}/(-q_{ii}) < U < (q_{i1} + q_{i2})/(-q_{ii}), \\ \dots & \dots \\ m, & \text{if } \sum_{k \neq i, k < m} q_{ik}/(-q_{ii}) < U, \end{cases} \tag{6.2}$$

where  $U$  is a random variable uniformly distributed in  $(0, 1)$ . Thus the sample path of  $\alpha(t)$  is constructed by sampling from exponential and  $U(0, 1)$  random variables alternately.

**6.3. Computation procedure for a continuous-time MDP** For each  $U \in \bar{\Gamma}$ ,

$$\beta v(\iota) \leq \bar{g}(\iota, U^i) + \sum_{\ell \neq \iota} \bar{q}_{\ell\iota}(U^i)(v(\ell) - v(\iota)). \tag{6.3}$$

Since  $\sum_{\ell \neq \iota} \bar{q}_{\ell\iota}(U^i)v(\ell) = |\bar{q}_u(U^i)|v(\iota)$ , and  $\beta > 0$ , the inequality (6.3) is equivalent to

$$v(\iota) \leq \frac{\bar{g}(\iota, U^i)}{\beta + |\bar{q}_u(U^i)|} + \sum_{\ell \neq \iota} \frac{\bar{q}_{\ell\iota}(U^i)}{\beta + |\bar{q}_u(U^i)|} v(\ell).$$

It follows that

$$v(\iota) \leq \min_{U^i \in \bar{\Gamma}_i} \left\{ \frac{\bar{g}(\iota, U^i)}{\beta + |\bar{q}_u(U^i)|} + \sum_{\ell \neq \iota} \frac{\bar{q}_{\ell\iota}(U^i)}{\beta + |\bar{q}_u(U^i)|} v(\ell) \right\}.$$

The equality holds if and only if  $U$  is the minimiser of the right-hand side of (4.2). To show that (4.4) is equivalent to a DP equation of a discrete-time MDP, let

$$\begin{aligned} \tilde{g}(\iota, U^i) &= \frac{\bar{g}(\iota, U^i)}{\beta + |\bar{q}_u(U^i)|}, & \rho &= \max_{\iota=1, \dots, \iota, U^i \in \bar{\Gamma}} \frac{|\bar{q}_u(U^i)|}{\beta + |\bar{q}_u(U^i)|}, \\ \tilde{p}_{\ell\iota} &= \frac{\bar{q}_{\ell\iota}(U^i)}{\rho(\beta + |\bar{q}_u(U^i)|)} & \text{for } \ell \neq \iota & \text{ and } \tilde{p}_{u} = 0. \end{aligned}$$

Then  $0 < \rho < 1$ ,  $\sum_{\ell} \tilde{p}_{\ell\iota}(U^i) = 1$ . The corresponding discrete-time version of the DP equation is  $v(\iota) = \min_{U^i \in \bar{\Gamma}_i} \{ \tilde{g}(\iota, U^i) + \rho \sum_{\ell} \tilde{p}_{\ell\iota}(U^i)v(\ell) \}$ .

### Acknowledgements

The research of H. Yang and K. Yin was supported in part by the Minnesota Sea Grant College Program by the NOAA Office of Sea Grant, U.S. Department of Commerce, under Grant NA46-RG0101; G. Yin's research was supported in part by the National Science Foundation under Grant DMS-9877090; Q. Zhang's research was supported in part by the USAF Grant F30602-99-2-0548 and ONR Grant N00014-96-1-0263.

### References

- [1] M. Abbad, J. A. Filar and T. R. Bielecki, "Algorithms for singularly perturbed limiting average Markov control problems", *IEEE Trans. Automat. Control* **AC-37** (1992) 1421–1425.
- [2] D. Bertsekas, *Dynamic programming: deterministic and stochastic models* (Prentice-Hall, Englewood Cliffs, New Jersey, 1987).
- [3] G. Blankenship, "Singularly perturbed difference equations in optimal control problems", *IEEE Trans. Automat. Control* **T-AC 26** (1981) 911–917.
- [4] P. J. Courtois, *Decomposability: queuing and computer system applications* (Academic Press, New York, 1977).
- [5] M. H. A. Davis, *Markov models and optimization* (Chapman and Hall, London, 1993).
- [6] F. Delebecque and J. Quadrat, "Optimal control for Markov chains admitting strong and weak interactions", *Automatica* **17** (1981) 281–296.
- [7] S. N. Ethier and T. G. Kurtz, *Markov processes: characterization and convergence* (J. Wiley, New York, 1986).
- [8] W. H. Fleming and R. W. Rishel, *Deterministic and stochastic optimal control* (Springer, New York, 1975).
- [9] F. C. Hoppensteadt and W. L. Miranker, "Multitime methods for systems of difference equations", *Studies Appl. Math.* **56** (1977) 273–289.
- [10] R. Z. Khasminskii, G. Yin and Q. Zhang, "Asymptotic expansions of singularly perturbed systems involving rapidly fluctuating Markov chains", *SIAM J. Appl. Math.* **56** (1996) 277–293.
- [11] R. Z. Khasminskii, G. Yin and Q. Zhang, "Constructing asymptotic series for probability distribution of Markov chains with weak and strong interactions", *Quart. Appl. Math.* **55** (1997) 177–200.
- [12] H. J. Kushner, *Approximation and weak convergence methods for random processes, with applications to stochastic systems theory* (MIT Press, Cambridge, MA, 1984).
- [13] H. J. Kushner and G. Yin, *Stochastic approximation algorithms and applications* (Springer, New York, 1997).
- [14] R. H. Liu, Q. Zhang and G. Yin, "Nearly optimal control of singularly perturbed Markov decision processes in discrete time", *Appl. Math. Optim.* **44** (2001) 105–129.
- [15] Z. G. Pan and T. Başar, " $H^\infty$ -control of Markovian jump linear systems and solutions to associated piecewise-deterministic differential games", in *New trends in dynamic games and applications* (ed. G. J. Olsder), (Birkhäuser, Boston, 1995) 61–94.
- [16] A. A. Pervozvanskii and V. G. Gaitsgori, *Theory of suboptimal decisions: decomposition and aggregation* (Kluwer, Dordrecht, 1988).
- [17] R. G. Phillips and P. V. Kokotovic, "A singular perturbation approach to modelling and control of Markov chains", *IEEE Trans. Automat. Control* **26** (1981) 1087–1094.

- [18] S. Ross, *Introduction to stochastic dynamic programming* (Academic Press, New York, 1983).
- [19] S. P. Sethi and Q. Zhang, *Hierarchical decision making in stochastic manufacturing systems* (Birkhäuser, Boston, 1994).
- [20] H. A. Simon and A. Ando, "Aggregation of variables in dynamic systems", *Econometrica* **29** (1961) 111–138.
- [21] W. A. Thompson, Jr., *Point process models with applications to safety and reliability* (Chapman and Hall, New York, 1988).
- [22] D. N. C. Tse, R. G. Gallager and J. N. Tsitsiklis, "Statistical multiplexing of multiple time-scale Markov streams", *IEEE J. Selected Areas Comm.* **13** (1995) 1028–1038.
- [23] D. J. White, *Markov decision processes* (Wiley, New York, 1992).
- [24] G. Yin and Q. Zhang, *Continuous-time Markov chains and applications: a singular perturbation approach* (Springer, New York, 1998).
- [25] G. Yin and Q. Zhang, "Singularly perturbed discrete-time Markov chains", *SIAM J. Appl. Math.* **61** (2000) 834–854.
- [26] G. Yin, Q. Zhang and G. Badowski, "Asymptotic properties of a singularly perturbed Markov chain with inclusion of transient states", *Ann. Appl. Probab.* **10** (2000) 549–572.
- [27] G. Yin, Q. Zhang and G. Badowski, "Discrete-time singularly perturbed Markov chains: aggregation, occupation measures, and switching diffusion limit", to appear in *Adv. Appl. Probab.*
- [28] Q. Zhang and G. Yin, "On nearly optimal controls of hybrid LQG problems", *IEEE Trans. Automat. Control* **44** (1999) 2271–2282.