Near-Optimal Controls of Random-Switching LQ Problems with Indefinite Control Weight Costs

Yuanjin Liu,* G. Yin,† Xun Yu Zhou‡

Abstract

In this paper, we consider hybrid controls for a class of linear quadratic problems with white noise perturbation and Markov regime switching, where the regime switching is modeled by a continuous-time Markov chain with a large state space and the control weights are indefinite. The use of large state space enables us to take various factors of uncertain environment into consideration, yet it creates computational overhead and adds difficulties. Aiming at reduction of complexity, we demonstrate how to construct near-optimal controls. First, in the model, we introduce a small parameter to highlight the contrast of the weak and strong interactions and fast and slow motions. This results in a two-time-scale formulation. In view of the recent developments on LQ problems with indefinite control weights and two-time-scale Markov chains, we then establish the convergence of the system of Riccati equations associated with the hybrid LQ problem. Based on the optimal feedback control of the limit system obtained using the system of Riccati equations, we construct controls for the original problem and show that such controls are near-optimal. A numerical demonstration of a simple system is presented.

Key words. Hybrid system, LQ problem, indefinite weight, near optimality.

*Department of Mathematics, Wayne State University, Detroit, MI 48202, `yuanjin@math.wayne.edu`. The research of this author was supported in part by Wayne State University Research Enhancement program.

†Department of Mathematics, Wayne State University, Detroit, MI 48202, `gyin@math.wayne.edu`. The research of this author was supported in part by the National Science Foundation and in part by Wayne State University Research Enhancement program.

‡Department of Systems Engineering and Engineering Management, the Chinese University of Hong Kong, Shatin, Hong Kong, `xyzhou@se.cuhk.edu.hk`. The research of this author was supported in part by the RGC Earmarked Grant CUHK 4234/01E.

1 Introduction

How can we deal with optimal controls for hybrid linear quadratic systems with white noise perturbation, in which the control weights are indefinite and the total number of elements in the switching set is large? More specifically, how can we reduce the complexity for handling such systems when the random process governing the regime switching is a Markov chain with large state space? To answer these questions is the objective of this paper. The models considered and the methods we use feature in several distinct aspects, including regime switching, indefinite controls, multi-scale modeling, and comparison controls.

Indefinite LQ Problems. LQ problems have a long and illustrated history. Nowadays, LQ control methodology is studied in almost every standard textbook in control theory, and is a "must" in undergraduate and graduate curriculum. Nevertheless, until very recently, the main focus has been on the case when the control weight (the matrix associated with the control action) is positive definite. In fact, for deterministic systems, if the control weight is not positive definite, the problem is not well posed. It has been shown recently that for stochastic systems, LQ with indefinite control weights could make sense if certain balance is reached [3, 8, 20]. The control weights can be indefinite or even negative definite as long as they are not "too negative." The stochastic influence, to some extent, compensates the negative control weights to make the problem well posed. The reference [3, p. 1686] contains a couple of simple although illuminating examples to explain this point. Owing to greatly many applications, especially optimal controls in financial engineering, LQ with indefinite control weights have drawn increasing attention lately.

Hybrid Systems. Along another line, realizing that many optimal control problems are hybrid in nature. Not only do they include the usual analog dynamics, but also they involve discrete events. In response to such demand, much effort has been devoted to the study of such systems. Since uncertainty is ubiquitous in daily life, it is necessary to take random environment into consideration. One of possible ways of handling the random environment is to use a regime switching model. It has been found that such models can be used in many applications. To mention just a few, they include tracking time-varying parameters, filtering [17], and mean-variance portfolio selection [22]. Mean-variance model was originally proposed in [11] for portfolio selection in a single period. It enables an investor to seek highest return after specifying the risk level (given by the variance of the return) that is acceptable to him/her. Observing that the environment does influence the market significantly resulting in the variations of key system parameters, such as appreciation and volatility rates, a viable alternative is to consider a regime switching model. Then the corresponding control system in the mean-variance portfolio selection problem becomes a LQ problem modulated by a continuous-time Markov chain. The inclusion of random discrete events opens up the possibilities for more realistic modeling and enables the consideration of various random environment factors. But the underlying system may be qualitatively different from one without the switching process. There is a vast literature on LQ controls involving Markov chains. For jump linear systems, coupled Riccati systems, and singular perturbation approaches in such problems, we refer the reader to [7, 10, 12, 13, 14, 15, 17, 18] and references therein.

Large-scale Systems. Focusing on the Markov modulated regime switching systems, from a modeling point of view, when we take many factors into consideration, the state space of the Markov chain becomes inevitably large. The resulting control problem becomes a large-scale one. Although, one may proceed with the dynamic programming approach, and obtain the associated Hamilton-Jacobi-Bellman (HJB) equations, the total number of equations (equal to the number of states of the Markov chain) is large. Computational complexity is of practical concerns. To resolve this issue, we realize that in many large-scale systems, not all components, parts, and subsystems evolve at the same rate. Some of them change in a fast pace, and others vary slowly. To take advantage of the intrinsic time-scale separation, one can introduce a small parameter $\varepsilon > 0$ to bring out the hierarchical structures of the underlying systems; see [1, 6, 15, 18, 21] among others. We remark that in practical systems, ε is just a fixed parameter and it separates different scales in the sense of order of magnitude in the generators. It does not need to go to 0. For further explanation and a specific example, see [18, Section 3.6: Interpretation of Time-Scale Separation]. In this paper, we use a two-time-scale formulation for the Markovian hybrid system having indefinite control weights. We obtain a system of Riccati equations associated with the optimal control problem, prove the convergence of the system of the Riccati equations, construct near-optimal controls based on the limit system, and establish near optimality.

Organization of the Paper. The rest of the paper is arranged as follows. Section 2 begins with the formulation of the problem. Section 3 discusses the optimal control, which is of theoretical, rather than computational, interest. In Section 4, we study the near-optimal controls by examining two cases. In the first one, all the states of the Markov chain are recurrent belonging to l recurrent classes, and in the second one, in addition to recurrent states, some transient states are included. As an illustration, Section 5 presents a numerical example for demonstration. Section 6 concludes the paper with further remarks.

2 Problem Formulation

Let $\alpha(t)$ be a homogeneous continuous-time Markov chain with state space $\mathcal{M} = \{1, \dots, m\}$ and $w(t)$ be a standard 1-dimensional Brownian motion on a complete probability space

$(\Omega, \mathcal{F}, P)$. Suppose that the generator of the Markov chain is given by $Q = (q_{ij}) \in \mathbb{R}^{m\times m}$. That is, $q_{ij} \geq 0$ for $i \neq j$ and $\sum_j q_{ij} = 0$ for each $i \in \mathcal{M}$. Let $\mathcal{F}_t = \sigma\{w(s), \alpha(s) : s \leq t\}$. We will work with a finite horizon $[0, T]$ for some $T > 0$. Our objective is: Find the optimal control $u(\cdot)$ to minimize the expected quadratic cost function

$$J(s,x,\alpha,u(\cdot)) = E_s\Big[\int_s^T [x'(t)M(\alpha(t))x(t) + u'(t)N(\alpha(t))u(t)]dt + x'(T)D(\alpha(T))x(T)\Big], \tag{2.1}$$

subject to the system constraint

$$\begin{aligned} &dx(t) = [A(\alpha(t))x(t) + B(\alpha(t))u(t)]dt + C(\alpha(t))u(t)dw(t), \\ &x(s) = x, \quad \alpha(s) = \alpha, \qquad \text{for} \quad s \leq t \leq T, \end{aligned} \tag{2.2}$$

where $x(t) \in \mathbb{R}^{n_1}$ is the state, $u(t) \in \mathbb{R}^{n_2}$ is the control, and E_s the expectation given $\alpha(s) = \alpha$ and $x(s) = x$. A control is admissible if it is an $\mathcal{F}_t$-adapted measurable process. The collection of all admissible controls is denoted by U_{ad}. In (2.2), $A(i), M(i), D(i) \in \mathbb{R}^{n_1\times n_1}$, $B(i), C(i) \in \mathbb{R}^{n_1\times n_2}$, and $N(i) \in \mathbb{R}^{n_2\times n_2}$ for each $i \in \mathcal{M}$. $i = 1, \ldots, m$. Define the value function

$$v(s,x,\alpha) = \inf_{u(\cdot)\in U_{ad}} J(s,x,\alpha,u(\cdot)). \tag{2.3}$$

We use the following conditions throughout this paper.

(A1) For each $i \in \mathcal{M}$, $M(i)$ is a symmetric positive semidefinite, $N(i)$ is symmetric, and $D(i)$ is symmetric positive semidefinite.

(A2) The process $\alpha(\cdot)$ and $w(\cdot)$ are independent.

Remark 2.1. We do not require the control weights $N(i)$ for $i \in \mathcal{M}$ be positive definite. This relaxation enables us to treat applications that do not verify the usual positive definiteness conditions. The regime switching allows us to take many factors with discrete shifts into consideration. The condition on the independence of $w(\cdot)$ and $\alpha(\cdot)$ is natural.

Notation. For $G \in \mathbb{R}^{n_1\times n_2}$, G' denotes its transpose and $|G|$ denotes its norm when G is a square matrix or a vector. We use S^n to denote the space of all $n \times n$ symmetric matrices, S^n_+ the subspace of all positive semidefinite matrices of S^n, and $\widehat{S}^n_+$ the subspace of all positive definite matrices of S^n. For a matrix $G \in S^n$, $G \geq 0$ means that $G \in S^n_+$, and $G > 0$ indicates that $G \in \widehat{S}^n_+$. The notation $C([0,T];X)$ is for the Banach space of X-valued continuous functions on $[0,T]$ endowed with the maximum norm for a given Hilbert space X. Given a probability space $(\Omega, \mathcal{F}, P)$ with a filtration $\mathcal{F}_t : a \leq t \leq b$ $(-\infty \leq a \leq b \leq \infty)$, a Hilbert space X with the norm $|\cdot|_X$, and p $(1 \leq p \leq \infty)$, define the Banach space

$L^p_{\mathcal{F}}([a,b];X)$ to be a collection of $\mathcal{F}_t$-adapted, X-valued measurable process $\varphi(\cdot)$ on $[a,b]$ such that $E\int_a^b |\varphi(t)|^p_X dt < \infty$, with the norm $|\varphi(\cdot)|_{\mathcal{F},p} = \left(E\int_a^b |\varphi(t)|^p_X dt\right)^{\frac{1}{p}}$. Suppose that Q is the generator of a continuous-time Markov chain. For a suitable $h(\cdot)$,

$$Qh(\cdot)(i) = \sum_{j=1}^{m} q_{ij}h(j) = \sum_{j\neq i} q_{ij}(h(j)-h(i)). \tag{2.4}$$

3 Optimal Controls

In this section, we highlight the optimal control for the LQG problem (2.1)–(2.2). Since most of the results in this section are extensions of those without Markovian switching in [3], we will only provide detailed argument of proofs that are distinct from previous work.

Introduce a system of Riccati equations as follows:

$$\begin{cases} \dot{K}(t,i) = -K(t,i)A(i) - A'(i)K(t,i) - M(i) - \displaystyle\sum_{j=1}^{m} q_{ij}K(t,j) \\ \qquad\qquad +K(t,i)B(i)(N(i)+C'(i)K(t,i)C(i))^{-1}B'(i)K(t,i), \quad s \le t \le T, \\ K(T,i) = D(i), \quad N(i)+C'(i)K(t,i)C(i) > 0, \quad i = 1,2,\cdots,m. \end{cases} \tag{3.1}$$

The following result is proved in [8, Corollary 3.1].

Theorem 3.1. *If $\{K(\cdot,i)\colon i \in \mathcal{M}\}$ is a solution of the system of Riccati equations* (3.1), *then the optimal feedback control for* (2.1)–(2.2) *is given by*

$$u(t) = -\sum_{i=1}^{m}(N(i)+C'(i)K(t,i)C(i))^{-1}B'(i)K(t,i)I_{\{\alpha(t)=i\}}x(t), \tag{3.2}$$

where the value function is given by $v(s,x,i) = x'K(s,i)x$.

Next, we derive a sufficient and necessary condition for the unique solvability of the system of Riccati equations (3.1). First, we have the following lemma.

Lemma 3.2. *Assume that $N(i) > 0$. Then* (3.1) *has a solution $\{K(\cdot,i) : i \in \mathcal{M}\}$ on $[0,T]$ with $K(t,i) \ge 0$, $\forall t \in [0,T]$, $i \in \mathcal{M}$.*

Proof. Since $N(i) + C'(i)D(i)C(i) > 0$, it follows from the classical theory of ordinary differential equations (see, e.g., [16, p. 66]) that the Riccati equations (3.1) have a *local* solution. Let $(\bar{t},T] \subset [0,T]$ be the maximal interval on which a local solution $\{K(\cdot,i) : i \in \mathcal{M}\}$ exists with $N(i)+C'(i)K(t,i)C(i) > 0$. In order to prove that the existence is in fact global on $[0,T]$, it suffices to show that there is no escape time, or each $K(\cdot,i)$ is uniformly bounded on $(\bar{t},T]$. To this end, we will show that there exists a positive scalar $\beta>0$ independent of $\bar{t}$ such that $0\le K(t,i)\le \beta I$, for all $t\in(\bar{t},T]$ and $i\in\mathcal{M}$.

First, let $x_0\in\mathbb{R}^{n_1}$ be an arbitrary initial state of the system (2.2) starting at a time $t\in(\bar{t},T]$, and $\alpha(t)=i$. Then Theorem 3.1 implies that $x_0'K(t,i)x_0=\inf_{u(\cdot)\in U_{ad}}J(t,x_0,i,u(\cdot))\ge 0$ owing to the positive (semi)definiteness assumption of the cost weighting matrices. Next, let $x(\cdot)$ be a solution to (2.2) corresponding to the initial $(x(t),\alpha(t))=(x_0,i)$ and the admissible control $u(\cdot)\equiv 0$. Then Theorem 3.1 implies

$$x_0'K(t,i)x_0\le E\Big[\int_t^T x'(s)M(\alpha(s))x(s)ds+x(T)'D(\alpha(T))x(T)\Big],\quad\text{for all}\quad i\in\mathcal{M}.$$

From the above inequality and the fact that $x(\cdot)$ satisfies a homogeneous linear equation, it follows that there exists a scalar $\beta>0$ such that $x_0'K(t,i)x_0\le\beta x_0'x_0$. The proof is completed. $\square$

Denote $\mathcal{W}=\{(W(1),\cdots,W(m))\in[L^\infty(0,T;\hat{S}_+^{n_2})]^m\;\Big|\;W(i)^{-1}\in L^\infty(0,T;\hat{S}_+^{n_2}),\quad\forall i\in\mathcal{M}\}$. It is easy to check that $[C(0,T;\hat{S}_+^{n_2})]^m\subset\mathcal{W}$. By virtue of Lemma 3.2, for any given $W=(W(1),\cdots,W(m))\in\mathcal{W}$ the following equation

$$\begin{cases}\dot{K}(t,i)=-K(t,i)A(i)-A'(i)K(t,i)-M(i)-\displaystyle\sum_{j=1}^m q_{ij}K(t,j)\\ \qquad\qquad +K(t,i)B(i)W(i)^{-1}B'(i)K(t,i),\quad s\le t\le T,\\ K(T,i)=D(i),\quad i=1,2,\cdots,m,\end{cases}\tag{3.3}$$

admits a unique solution $W=(W(\cdot,1),\cdots,W(\cdot,m))\in[C(0,T;S_+^{n_2})]^m$. Thus we can define a mapping $\Psi:\mathcal{W}\to[C(0,T;S_+^{n_2})]^m$ as $K\equiv(K(\cdot,1),\cdots,K(\cdot,m))=\Psi(W)\equiv(\Psi_1(W),\cdots,\Psi_m(W))$. The following theorem can be proved in exactly the same way as [3, Theorem 4.2].

Theorem 3.3. *The system of Riccati equations* (3.1) *admits a unique solution if and only if there exists* $W=(W(1),\cdots,W(m))\in[C(0,T;\hat{S}_+^{n_2})]^m$ *such that* $N(i)+C'(i)\Psi_i(W)C(i)\ge W(i),\ \ \forall i\in\mathcal{M}$.

4 Near-optimal Controls

By using the results obtained in the last section, we are able to derive optimal LQ controls. At a first glance, it seems that the problem is largely solved. Nevertheless, it barely begins. Unlike the usual LQ problem, in lieu of one Riccati equation, we have to solve a system of

Riccati equations. The number of Riccati equations in the system is exactly the same as that of $|\mathcal{M}|$, the cardinality of the set $\mathcal{M}$ or the total number of states of the Markov chain. Note that in the LQ formulation with regime switching, the factor process $\alpha(t)$ often has a large state space. The large dimensionality of the Markovian state presents us with opportunities of modeling various aspects of uncertainty, but it also creates computational burden. If $|\mathcal{M}|$ is very large, the computational task could be infeasible. Therefore, not only is it desirable, but also it is necessary to look into better alternatives. Our main concern here is to treat such large-scale systems and obtain the desired "good" controls with reduced computational effort.

To reduce the complexity, we introduce a small parameter $\varepsilon > 0$ into the system under consideration. The purpose is to highlight the contrast of the fast and slow transitions among different Markovian states and to bring out the hierarchical structure of the underlying system. This task is accomplished by letting the generator of the Markov chain depend on a scale parameter ε. That is, $Q = Q^\varepsilon$ with

$$Q^\varepsilon = \frac{1}{\varepsilon}\widetilde{Q} + \widehat{Q}, \tag{4.1}$$

where both $\widetilde{Q}$ and $\widehat{Q}$ are generators of continuous-time Markov chains. Thus the generator Q^ε consists of two parts, a rapidly changing part and a slowly varying part, and the resulting Markov chain can be written as $\alpha(t) = \alpha^\varepsilon(t)$.

In the rest of the paper, we consider (2.1) subject to (2.2) in which the process $\alpha(\cdot)$ is replaced by $\alpha^\varepsilon(\cdot)$. The resulting cost function is denoted by $J^\varepsilon(s, x, \alpha, u(\cdot))$. Our objective is to find the optimal control $u(\cdot)$ so as to minimize the expected quadratic cost

$$J^\varepsilon(s,x,\alpha,u(\cdot)) = E_s\Big[\int_s^T [x'(t)M(\alpha^\varepsilon(t))x(t) + u'(t)N(\alpha^\varepsilon(t))u(t)]dt + x'(T)D(\alpha^\varepsilon(T))x(T)\Big], \tag{4.2}$$

subject to the system constraint

$$\begin{aligned} &dx(t) = [A(\alpha^\varepsilon(t))x(t) + B(\alpha^\varepsilon(t))u(t)]dt + C(\alpha^\varepsilon(t))u(t)dw(t), \\ &x(s) = x, \ \alpha^\varepsilon(s) = \alpha, \quad \text{for} \quad s \le t \le T. \end{aligned} \tag{4.3}$$

Note that in the above, $x(t)$ should really have been written as $x^\varepsilon(t)$. We suppressed ε-dependence for notational simplicity. In the rest of the paper, whenever there is a need to emphasize the dependence of ε, we will retain it. It will become clear from the context.

It follows that the value function is $v^\varepsilon(s,x,\alpha) = \inf_{u(\cdot)\in U_{ad}} J^\varepsilon(s,x,\alpha,u(\cdot))$. We retain the assumptions (A1) and (A2). Then by the generalized Ito lemma for Markov-modulated processes in [2], the optimal feedback control provided in Theorem 3.1, the solvability of the

Riccati system Lemma 3.2, and Theorem 3.3 all hold. In fact, we have the following system of Riccati equations

$$\begin{cases} \dot{K}^\varepsilon(t,i) = -K^\varepsilon(t,i)A(i) - A'(i)K^\varepsilon(t,i) - M(i) - \displaystyle\sum_{j=1}^{m} q_{ij}^\varepsilon K^\varepsilon(t,j) \\ \qquad\qquad +K^\varepsilon(t,i)B(i)(N(i)+C'(i)K^\varepsilon(t,i)C(i))^{-1}B'(i)K^\varepsilon(t,i), \;\; s \le t \le T, \\ K^\varepsilon(T,i) = D(i), \quad N(i)+C'(i)K^\varepsilon(t,i)C(i) > 0, \quad i = 1,2,\ldots,m. \end{cases} \tag{4.4}$$

The corresponding optimal control is

$$u(t) = -\sum_{i=1}^{m}(N(i)+C'(i)K^\varepsilon(t,i)C(i))^{-1}B'(i)K^\varepsilon(t,i)I_{\{\alpha^\varepsilon(t)=i\}}x^\varepsilon(t), \;\; i \in \mathcal{M}. \tag{4.5}$$

The essence of our approach is that in lieu of solving the optimal control problem directly, we construct approximations, which are much simpler to solve than that of the original problem, leading to near optimality. Using the hierarchical structure of the two-time scales, we decompose the state space naturally so that within each subspace, the transition frequency is of the same order of magnitude, and transitions among different subspaces evolve in a much slower rate. For a finite-state Markov chain, there is at least one state that is recurrent. Either all states are recurrent, or in addition to recurrent states, there are some transient states. Taking advantage of the time-scale separation, we lump all the states in each ergodic class (each subspace of the recurrent states) into one state and obtain an aggregated process. Then the entire problem is recast into one, in which the total number of aggregated states is l. If $l \ll m$, a substantial reduction of complexity will be achieved. We show that as the small parameter goes to zero, the value functions converge to that of the limit problem. Using the optimal control from the limit problem, we then build controls leading to approximate optimality. We divide the rest of the section into two parts, namely, recurrent states only and inclusion of transient states.

4.1 Recurrent State Case

Here, suppose that the underlying Markov chain is divisible to a number of recurrent groups (or ergodic classes) such that it fluctuates rapidly among different states within a group consisting of recurrent states, but jump less frequently from one group to another.

Let the generator of the Markov chain be given by (4.1). Assume that $\widetilde{Q}$ has the form $\widetilde{Q} = \mathrm{diag}(\widetilde{Q^1},\ldots,\widetilde{Q^l})$, where $\mathrm{diag}(\widetilde{Q}^1,\ldots,\widetilde{Q}^l)$ denotes a block diagonal matrix having matrix entries $Q^1,\ldots,Q^l$, and $\widetilde{Q}^k \in \mathbb{R}^{m_k\times m_k}$ are irreducible, for $k=1,\ldots,l$, and $\sum_{k=1}^{l} m_k = m$. Let $\mathcal{M}_k = \{\zeta_{k1},\ldots,\zeta_{km_k}\}$ for $k=1,\ldots,l$ denotes the state space corresponding to $\widetilde{Q}^k$. Then the state space of the Markov chain admits the decomposition

$$\mathcal{M} = \mathcal{M}_1 \cup \cdots \cup \mathcal{M}_l = \{\zeta_{11}, \dots \zeta_{1m_1}\} \cup \cdots \cup \{\zeta_{l1}, \dots, \zeta_{lm_l}\}.$$

Note that in the above, we have relabeled the states by use of ζ_{kj} with $k = 1, \dots, l$ and $j = 1, \dots, m_k$. Because $\widetilde{Q}^k = (\widetilde{q}^k_{ij})_{m_k \times m_k}$ and $\widehat{Q} = (\widehat{q}_{ij})_{m\times m}$ are generators, for $k = 1, \dots, l$, $\sum_{j=1}^{m_k} \widetilde{q}^k_{ij} = 0$ for $i = 1, \dots, m_k$, and $\sum_{j=1}^{m} \widehat{q}_{ij} = 0$, for $i = 1, \dots, m$. In the above, $\widetilde{Q}$ represents the rapidly changing part and $\widehat{Q}$ describes the slowly varying components. 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 $\mathcal{M}_k$ and varies less frequently among groups $\mathcal{M}_k$ and $\mathcal{M}_j$ for $k \neq j$. Aggregating the states in $\mathcal{M}_k$ as a single super-state, all such states are coupled by the matrix $\widehat{Q}$. We obtain an aggregated process $\overline{\alpha}^\varepsilon(\cdot)$ defined by $\overline{\alpha}^\varepsilon(t) = k$ if $\alpha^\varepsilon(t) \in \mathcal{M}_k$. The process $\overline{\alpha}^\varepsilon(\cdot)$ is not necessarily Markovian; see [18, Example 7.5]. Nevertheless, by a probabilistic argument, it is shown in [18, Theorem 7.4, pp. 172-174] that $\overline{\alpha}^\varepsilon(\cdot)$ converges weakly to a continuous-time Markov chain $\overline{\alpha}(\cdot)$ generated by

$$\overline{Q} = \mathrm{diag}(\nu^1, \dots, \nu^l)\widehat{Q}\mathrm{diag}(\mathbb{1}_{m_1}, \dots, \mathbb{1}_{m_l}), \tag{4.6}$$

where ν^k is the stationary distribution of $\widetilde{Q}^k$, for $k = 1, \dots, l$, and $\mathbb{1}_\iota = (1, \dots, 1)' \in \mathbb{R}^\iota$. Moreover, for any bounded deterministic $\beta(\cdot)$ (see [18, Theorem 7.2, pp. 170-171]),

$$E\left(\int_s^T [I_{\{\alpha^\varepsilon(t)=\zeta_{kj}\}} - \nu^k_j I_{\{\overline{\alpha}^\varepsilon(t)=k\}}]\beta(t)dt\right)^2 = O(\varepsilon). \tag{4.7}$$

For preparation of subsequent study, we establish two lemmas concerning the *a priori* estimate of K^ε and its Lipschitz continuity.

Lemma 4.1. *There exist constant κ_1 and κ_A such that for all $s \in [0,T]$ and $i \in \mathcal{M}$, we have $|K^\varepsilon(s,i)| \leq \kappa_1(T+1)e^{\kappa_A T}$.*

Proof. We claim that for each $i \in \mathcal{M}$,

$$0 \leq v^\varepsilon(s,i,x) \leq \kappa_1 e^{\kappa_A T}|x|^2(T+1). \tag{4.8}$$

Clearly $v^\varepsilon(s,i,x) \geq 0$ because $J^\varepsilon(s,i,x,u(\cdot)) \geq 0$ for all admissible $u(\cdot)$. To derive the upper bound, set $u_0(t) = 0$. Then, under such a control, $x(t) = x + \int_s^t A(\alpha^\varepsilon(r))x(r)dr$. Using the Gronwall inequality, we have $|x(t)| \leq |x|e^{\kappa_A T}$, where $\kappa_A = \max_i |A(i)|$. This inequality holds for all $t \in [0,T]$. For $0 \leq s \leq T$,

$$\begin{aligned} v^\varepsilon(s,i,x) &\leq J^\varepsilon(s,i,x,u_0(\cdot)) \leq E\{\int_s^T x'(t)M(\alpha^\varepsilon(t))x(t) + x'(T)D(\alpha^\varepsilon(T))x(T)\} \\ &\leq \kappa_1 e^{\kappa_A T}|x|^2(T+1). \end{aligned}$$

Owing to Theorem 3.3, $K^\varepsilon(s,i)$ is symmetric and positive semi-definite. By the definition of the matrix norm $|K^\varepsilon(s,i)| = \max\{\text{eigenvalues of } K^\varepsilon(s,i)\}$, it suffices to verify that for any unit vector ξ, $\xi' K^\varepsilon(s,i)\xi \le \kappa_1 e^{\kappa_A T}(T+1)$. In view of (4.8), by taking $x = a\xi$ with a being a scalar $a^2\xi' K^\varepsilon(s,i)\xi = v^\varepsilon(s,i,a\xi) \le \kappa_1 e^{\kappa_A T} a^2 (T+1)$. Dividing both sides of the above inequality by a^2, we complete the proof. $\square$

Lemma 4.2. *For $i \in \mathcal{M}$, the solutions to (4.4), namely $K^\varepsilon(\cdot,i)$ are uniformly (in ε) Lipschitz continuous on $[0,T]$.*

Proof. We first prove that $v^\varepsilon(s,i,x)$ is uniformly Lipschitz. Then, we prove the Lipschitz property of K^ε.

Step 1. There exists a constant κ_2 that may depend on T such that for any $\delta > 0$ and $(s,i,x) \in [0,T] \times \mathcal{M} \times \mathbb{R}^{n_1}$, $|v^\varepsilon(s+\delta,i,x) - v^\varepsilon(s,i,x)| \le \kappa_2(|x|^2+1)\delta$. For given (s,i,x), write the value function $v^\varepsilon(s,i,x)$ as

$$v^\varepsilon(s,i,x) = E\Big[\int_s^T [(x^{o,\prime}(t)M(\alpha^\varepsilon(t))x^o(t) + u^{o,\prime}(t)N(\alpha^\varepsilon(t))u^o(t))]dt + x^{o,\prime}(T)D(\alpha^\varepsilon(T))x^o(T)\Big], \tag{4.9}$$

where $u^o(t) = u^o(t,\alpha^\varepsilon(t),x^o(t))$ is the optimal control defined in (3.2), $x^o(t)$ is the corresponding trajectory, and E is the conditional expectation given $(\alpha^\varepsilon(s), x^o(s)) = (i,x)$. By changing variable $t \to t+\delta$ in (4.9), using similar argument as in [21], we obtain $v^\varepsilon(s+\delta,i,x) - v^\varepsilon(s,i,x) \le \kappa_2(|x|^2+1)\delta$. Moreover, we can also establish the reverse inequality $v^\varepsilon(s+\delta,i,x) - v^\varepsilon(s,i,x) \ge -\kappa_2(|x|^2+1)\delta$.

Step 2. The $K^\varepsilon(\cdot,i)$ are Lipschitz. As in the proof of lemma 4.1, taking $x = a\xi$ with ξ being a unit vector and sending $a \to \infty$, we obtain $|K^\varepsilon(s+\delta,i) - K^\varepsilon(s,i)| \le \kappa_2\delta$. $\square$

4.1.1 Limit Riccati Equations

To proceed, we obtain a result concerning the limit Riccati equations. For any function F on $\mathcal{M}$, we define $\overline{F}(k) = \sum_{j=1}^{m_k} \nu_j^k F(\zeta_{kj})$ for $k = 1,\dots,l$, for F_1 and F_2, define $\overline{F_1F_2}(k) = \sum_{j=1}^{m_k} \nu_j^k F_1(\zeta_{kj})F_2(\zeta_{kj})$.

Theorem 4.3. *For $k = 1,\dots,l$ and $j = 1,\dots,m_k$, $K^\varepsilon(s,\zeta_{kj}) \to \overline{K}(s,k)$ uniformly on $[0,T]$ as $\varepsilon \to 0$, where $\overline{K}(s,k)$ is the unique solution to*

$$\begin{aligned}\dot{\overline{K}}(s,k) &= -\overline{K}(s,k)\overline{A}(k) - \overline{A}'(k)\overline{K}(s,k) - \overline{M}(k)\\ &\quad + \overline{K}(s,k)\overline{B(k)(N(k)+C'(k)K(s,k)C(k))^{-1}B'(k)}\,\overline{K}(s,k) - \overline{Q}\,\overline{K}(s,\cdot)(k),\end{aligned} \tag{4.10}$$

where $\overline{K}(T,k) = \overline{D}(k) \overset{\text{def}}{=} \sum_{j=1}^{m_k} \nu_j^k D(\zeta_{kj})$.

Proof. By virtue of Lemma 4.1 and Lemma 4.2, $\{K^\varepsilon(s,\zeta_{kj})\}$ is equicontinuous and uniformly bounded. By Arzela-Ascoli theorem, for each sequence of $(\varepsilon \to 0)$, extract a further subsequence (still indexed by ε for simplicity) such that $K^\varepsilon(s,\zeta_{kj})$ converges uniformly on $[0,T]$ to a continuous function $K(s,\zeta_{kj})$. The Riccati equation in its integral form with time running backward and terminal condition $K^\varepsilon(T,\zeta_{kj}) = D(\zeta_{kj})$ is:

$$
\begin{aligned}
K^\varepsilon(s,\zeta_{kj}) = D(\zeta_{kj}) + \int_s^T & [K^\varepsilon(r,\zeta_{kj})A(\zeta_{kj}) + A'(\zeta_{kj})K^\varepsilon(r,\zeta_{kj}) + M(\zeta_{kj}) \\
& -K^\varepsilon(r,\zeta_{kj})B(\zeta_{kj})P^{-1}(r,\zeta_{kj})B'(\zeta_{kj})K^\varepsilon(r,\zeta_{kj}) + Q^\varepsilon K^\varepsilon(r,\cdot)(\zeta_{kj})]dr,
\end{aligned}
\tag{4.11}
$$

where $P^{-1}(r,\zeta_{kj}) = (N(\zeta_{kj}) + C'(\zeta_{kj})K(r,\zeta_{kj})C(\zeta_{kj}))^{-1}$ and noting that Q^ε is given by (4.1) and the uniform boundedness of K^ε, multiplying both sides of (4.11) by ε and sending $\varepsilon \to 0$ yield $\int_s^T \widetilde{Q}^k K(r,\cdot)(\zeta_{kj})dr = \lim_{\varepsilon\to 0}\int_s^T \widetilde{Q}^k K^\varepsilon(r,\cdot)(\zeta_{kj})dr = 0$, for $s \in [0,T]$. Owing to the continuity of $K(s,\zeta_{kj})$, we obtain

$$
\widetilde{Q}^k K(s,\cdot)(\zeta_{kj}) = 0, \quad \text{for} \quad s \in [0,T]. \tag{4.12}
$$

The irreducibility of $\widetilde{Q}^k$ implies that the null space of $\widetilde{Q}^k$ is one-dimensional spanned by $\mathbb{1}_{m_k}$. Eq. (4.12) yields that the vector $(K(s,\zeta_{k1}), K(s,\zeta_{k2}), \ldots, K(s,\zeta_{km_k}))'$ is in the null space of $\widetilde{Q}^k$ and as a result $K(s,\zeta_{k1}) = K(s,\zeta_{k2}) = \cdots = K(s,\zeta_{km_k}) := K(s,k)$ (see also [18, Lemma A.39, pp. 327–328] for further details). We proceed to show $K(s,k) = \overline{K}(s,k)$. For each $k = 1,\ldots,l$, multiplying $K^\varepsilon(s,\zeta_{kj})$ by ν_j^k, summing over the index j and sending $\varepsilon \to 0$, we obtain

$$
\begin{aligned}
K(s,k) = \overline{D}(k) + \int_s^T & [K(r,k)\overline{A}(k) + \overline{A}'(k)K(r,k) + \overline{M}(k) \\
& -K(r,k)\overline{BP^{-1}B'}(r,k)K(r,k) + \overline{Q}K(r,\cdot)(k)]dr.
\end{aligned}
$$

[Recall the notation (2.4).] Thus, the uniqueness of the Riccati equation implies $K(s,k) = \overline{K}(s,k)$. As a result, $K^\varepsilon(s,\zeta_{kj}) \to \overline{K}(s,k)$. □

4.1.2 Near-optimal Controls

The convergence of $K^\varepsilon(s,\zeta_{kj})$ leads to that of $v^\varepsilon(s,\zeta_{kj},x) = x'K^\varepsilon(s,\zeta_{kj})x$. It follows that as $\varepsilon \to 0$, $v^\varepsilon(s,\zeta_{kj},x) \to v(s,k,x)$ for $j = 1,\ldots,m_k$, where $v(s,k,x) = x'\overline{K}(s,k)x$ corresponds to the value function of a limit problem. Let $\mathcal{U}$ denote the control set of the limit problem:

$\mathcal{U} = \{U = (U^1, \dots, U^l) : U^k = (u^{k1}, \dots, u^{km_k},\ u^{kj} \in \mathbb{R}^{n_2})\}$. Define

$$\begin{aligned}
f(s,k,x,U) &= \overline{A}(k)x + \sum_{j=1}^{m_k} \nu_j^k B(\zeta_{kj}) u^{kj}, \\
\widetilde{N}(k,U) &= \sum_{j=1}^{m_k} \nu_j^k (u^{kj,\prime} N(\zeta_{kj}) u^{kj}), \\
\widetilde{C}(k,U)\widetilde{C}'(k,U) &= \sum_{j=1}^{m_k} \nu_j^k C(\zeta_{kj}) u^{kj} u^{kj,\prime} C'(\zeta_{kj}).
\end{aligned} \tag{4.13}$$

The control problem for the limit system is:

$$\begin{cases}
\min J(s,k,x,U(\cdot)) = E\Big[\int_s^T [x'(t)\overline{M}(\overline{\alpha}(t))x(t) + \widetilde{N}(\overline{\alpha}(t), U(t))]dt + x'(T)\overline{D}(k)x(T)\Big], \\
\text{s.t. } dx(t) = f(t, \overline{\alpha}(t), x(t), U(t))dt + \widetilde{C}(k,U)dw(t), \quad x(s) = x,
\end{cases}$$

where $\overline{\alpha}(\cdot) \in \{1, \dots, l\}$ is a Markov chain generated by $\overline{Q}$. The optimal control for this limit problem is $U^o(s,k,x) = (U^{o,1}(s,x), \dots, U^{o,l}(s,x))$ with $U^{o,k}(s,x) = (u^{o,k1}(s,x), \dots, u^{o,km_k}(s,x))$, and $u^{o,kj}(s,x) = -P^{-1}(s,\zeta_{kj})B'(\zeta_{kj})\overline{K}(s,k)x$. Using such controls (as in [15] for manufacturing systems), we construct

$$u^\varepsilon(s,\alpha,x) = \sum_{k=1}^{l}\sum_{j=1}^{m_k} I_{\{\alpha=\zeta_{kj}\}} u^{o,kj}(s,x) \tag{4.14}$$

for the original problem. This control can also be written as if $\alpha \in \mathcal{M}_k$, $u^\varepsilon(s,\alpha,x) = -P^{-1}(s,\alpha)B'(\alpha)\overline{K}(s,k)x$. We use

$$u^\varepsilon(t) = u^\varepsilon(t, \alpha^\varepsilon(t), x(t)) \tag{4.15}$$

for the original problem. It will be shown that this is a near-optimal control. If $A(\zeta_{kj}) = A(k)$, $B(\zeta_{kj}) = B(k)$, $C(\zeta_{kj}) = C(k)$, $M(\zeta_{kj}) = M(k)$, and $N(\zeta_{kj}) = N(k)$ are independent of j, owing to (4.7), we may replace $I_{\{\alpha^\varepsilon(t)=\zeta_{kj}\}}$ by $I_{\{\overline{\alpha}^\varepsilon(t)=k\}}\nu_j^k$ and consider

$$\overline{u}^\varepsilon(s,\alpha,x) = \sum_{k=1}^{l}\sum_{j=1}^{m_k} I_{\{\alpha\in\mathcal{M}_k\}}\nu_j^k u^{o,kj}(s,x) = -P^{-1}(s,k)B'(k)\overline{K}(s,k)x. \tag{4.16}$$

Therefore, we write $\overline{u}^\varepsilon(s,\alpha,x) = \overline{u}^\varepsilon(s,k,x)$. The control $\overline{u}^\varepsilon$ only needs the information $\alpha^\varepsilon(t) \in \mathcal{M}_k$. As a result, we use

$$\overline{u}^\varepsilon(t) = \overline{u}^\varepsilon(t, \overline{\alpha}^\varepsilon(t), x(t)). \tag{4.17}$$

Theorem 4.4. *The following assertions hold.*

1) *The* $u^\varepsilon(t)$ *defined in* (4.15) *with the use of* (4.14) *is near-optimal in that* $|J^\varepsilon(s,\alpha,x,u^\varepsilon)-v^\varepsilon(s,\alpha,x)|\to 0$ *as* $\varepsilon\to 0$.

2) *Assume* $A(\zeta_{kj})=A(k)$, $B(\zeta_{kj})=B(k)$, $C(\zeta_{kj})=C(k)$, $M(\zeta_{kj})=M(k)$ *and* $N(\zeta_{kj})=N(k)$ *independent of* j. *Then* $\overline{u}^\varepsilon(t)$ *defined in* (4.17) *is near-optimal in that* $|J^\varepsilon(s,\alpha,x,\overline{u}^\varepsilon)-v^\varepsilon(s,\alpha,x)|\to 0$ *as* $\varepsilon\to 0$.

Proof. Recall $\overline{\alpha}^\varepsilon(t)=k$ if $\alpha^\varepsilon(t)\in\mathcal{M}_k$. The constructed control u^ε can be written as $u^\varepsilon(t)=-P^{-1}(t,\alpha^\varepsilon(t))B'(\alpha^\varepsilon(t))\overline{K}(t,\overline{\alpha}^\varepsilon(t))x^\varepsilon(t)$, where $x^\varepsilon(t)$ is the corresponding trajectory governed by the differential equation

$$\begin{aligned}dx^\varepsilon(t)&=(A(\alpha^\varepsilon(t))-B(\alpha^\varepsilon(t))P^{-1}(t,\alpha^\varepsilon(t))B'(\alpha^\varepsilon(t))\\&\quad\times\overline{K}(t,\alpha^\varepsilon(t)))x^\varepsilon(t)dt-C(\alpha^\varepsilon(t))P^{-1}(t,\alpha^\varepsilon(t))B'(\alpha^\varepsilon(t))\overline{K}(t,\overline{\alpha}^\varepsilon(t))x^\varepsilon(t)dw(t)\end{aligned}$$

with $x^\varepsilon(s)=x$. Let $x(t)$ denote the optimal trajectory of the limit problem. Then

$$dx(t)=f(t,\overline{\alpha}(t),x(t),U^o(t))dt+\widetilde{C}(k,U^o)dw(t) \tag{4.18}$$

with $x(s)=x$, and $f(\cdot)$ and $\widetilde{C}(\cdot)$ defined by (4.13). Using the convergence of $\overline{\alpha}^\varepsilon(\cdot)\to\overline{\alpha}(\cdot)$ and (4.7), as in [18, Chapter 9], by using the Skorohod representation,

$$E|x^\varepsilon(t)-x(t)|\to 0,\qquad\text{so}\quad |J^\varepsilon(s,\alpha,x,u^\varepsilon)-v(s,k,x)|\to 0. \tag{4.19}$$

This together with $v^\varepsilon\to v$ leads to $\lim_{\varepsilon\to 0}|J^\varepsilon(s,\alpha,x,u^\varepsilon)-v^\varepsilon(s,\alpha,x)|\to 0$, as desired.

To obtain the second part of the theorem, under the condition $A(\zeta_{kj})=A(k)$, $B(\zeta_{kj})=B(k)$, $C(\zeta_{kj})=C(k)$, $M(\zeta_{kj})=M(k)$, and $N(\zeta_{kj})=N(k)$, we have

$$\overline{u}^\varepsilon(t)=-P^{-1}(t,\overline{\alpha}(t))B'(\overline{\alpha}(t))\overline{K}(t,\overline{\alpha}(t))x^\varepsilon(t).$$

The corresponding trajectory is

$$\begin{aligned}dx^\varepsilon(t)&=(A(\alpha^\varepsilon(t))-B(\overline{\alpha}^\varepsilon(t))P^{-1}(t,\overline{\alpha}^\varepsilon(t))B'(\overline{\alpha}^\varepsilon(t))\overline{K}(t,\overline{\alpha}(t))x^\varepsilon(t)dt\\&\quad-C(\overline{\alpha}^\varepsilon(t))P^{-1}(t,\overline{\alpha}^\varepsilon(t))B'(\overline{\alpha}^\varepsilon(t))\overline{K}(t,\overline{\alpha}(t))x^\varepsilon(t)dw(t)\end{aligned}$$

with $x^\varepsilon(s)=x$. The optimal trajectory $x(t)$ for the limit problem is

$$dx(t)=f(t,\overline{\alpha}(t),x(t),U^0(t))dt+\widetilde{C}(k,U^0)dw(t)$$

with $x(s)=x$, and $f(\cdot)$ and $\widetilde{C}(\cdot)$ being defined in (4.13). Then (4.19) can be verified in a similar way as in the previous case. □

4.2 Inclusion of Transient States

In this section, we consider the case in which the Markov chain includes transient states. To incorporate the transient states, we assume $\widetilde{Q} = \begin{pmatrix} \widetilde{Q}_r & 0 \\ \widetilde{Q}_0 & \widetilde{Q}_* \end{pmatrix}$ where $\widetilde{Q}_r = \mathrm{diag}(\widetilde{Q}^1, \ldots, \widetilde{Q}^l)$, $\widetilde{Q}_0 = (\widetilde{Q}_*^1, \ldots, \widetilde{Q}_*^l)$ such that for each $k = 1, \ldots, l$, $\widetilde{Q}^k$ is a generator with dimension $m_k \times m_k$, $\widetilde{Q}_* \in \mathbb{R}^{m_* \times m_*}$ matrix, $\widetilde{Q}_*^k \in \mathbb{R}^{m_* \times m_k}$, and $m_1 + \cdots m_l + m_* = m$. Consequently, the state space of the underlying Markov chain is

$$\mathcal{M} = \mathcal{M}_1 \cup \cdots \cup \mathcal{M}_l \cup \mathcal{M}_* = \{\zeta_{11}, \ldots, \zeta_{1m_1}\} \cup \cdots \cup \{\zeta_{l1}, \ldots, \zeta_{lm_l}\} \cup \{\zeta_{*1}, \ldots, \zeta_{*m_*}\},$$

where $\mathcal{M}_*=\{\zeta_{*1}, \ldots, \zeta_{*m_*}\}$ consists of the transient states. Suppose for $k = 1, \ldots, l$, $\widetilde{Q}^k$ are irreducible, and $\widetilde{Q}_*$ has eigenvalues with negative real parts. Write $\widehat{Q} = \begin{pmatrix} \widehat{Q}^{11} & \widehat{Q}^{12} \\ \widehat{Q}^{21} & \widehat{Q}^{22} \end{pmatrix}$, where $\widehat{Q}^{11} \in \mathbb{R}^{(m-m_*)\times(m-m_*)}$, $\widehat{Q}^{12} \in \mathbb{R}^{(m-m_*)\times m_*}$, $\widehat{Q}^{21} \in \mathbb{R}^{m_* \times (m-m_*)}$, and $\widehat{Q}^{22} \in \mathbb{R}^{m_* \times m_*}$. Define

$$\overline{Q}_* = \mathrm{diag}(\nu^1, \ldots, \nu^l)(\widehat{Q}^{11}\widetilde{\mathbb{1}} + \widehat{Q}^{12}(a_{m_1}, \ldots, a_{m_l})) \tag{4.20}$$

with $\widetilde{\mathbb{1}} = \mathrm{diag}(\mathbb{1}_{m_1}, \ldots, \mathbb{1}_{m_l})$, $\mathbb{1}_{m_j} = (1, \ldots, 1)' \in \mathbb{R}^{m_j \times 1}$ and, for $j = 1, \ldots, l$,

$$a_{m_j} = (a_{m_j}(1), \ldots, a_{m_j}(l)) = -\widetilde{Q}_*^{-1}\widetilde{Q}_*^j \mathbb{1}_{m_j}. \tag{4.21}$$

As in [19], it can be shown that $a_{m_j} \geq 0$ and $\sum_{j=1}^l a_{m_j} = \mathbb{1}_{m_*}$. Let ξ_j denote a random variable such that $P(\xi_j = i | \alpha^\varepsilon(t) = \zeta_{*j}) = a_{m_j}(i)$. Define $\overline{\alpha}^\varepsilon(t) = \begin{cases} k, & \text{if } \alpha^\varepsilon(t) \in \mathcal{M}_k \\ \xi_j, & \text{if } \alpha^\varepsilon(t) = \zeta_{*j}. \end{cases}$ It can be shown as in [19], that $\overline{\alpha}^\varepsilon(\cdot)$ converges weakly to $\overline{\alpha}(\cdot)$, where $\overline{\alpha}(\cdot) \in \{1, \ldots, l\}$ is a Markov chain generated by $\overline{Q}_*$. Moreover, for $k = 1, \ldots, l$,

$$E\left(\int_0^T [I_{\{\alpha^\varepsilon(t)=\zeta_{k_j}\}} - \nu_j^k I_{\{\overline{\alpha}^\varepsilon(t)=k\}}]\beta(t)dt\right)^2 = O(\varepsilon), \quad E\left(\int_0^T I_{\{\alpha^\varepsilon(t)=\zeta_{*j}\}}dt\right)^2 = O(\varepsilon^2). \tag{4.22}$$

Theorem 4.5. *As $\varepsilon \to 0$, $K^\varepsilon(s, \zeta_{kj}) \to \overline{K}(s, k)$, for $k = 1, \ldots, l$, $j = 1, \ldots, m_k$, $K^\varepsilon(s, \zeta_{*j}) \to \overline{K}_*(s, j)$ for $j = 1, \ldots, m_*$ uniformly on $[0, T]$, where $\overline{K}_*(s, j) = a_{m_1}(j)\overline{K}(s, 1) + \cdots + a_{m_l}(j)\overline{K}(s, l)$ and $\overline{K}(s, k)$ is the unique solution to*

$$\begin{aligned}&\dot{\overline{K}}(s, k) = -\overline{K}(s, k)\overline{A}(k) - \overline{A}'(k)\overline{K}(s, k) - \overline{M}(k) + \overline{K}(s, k)\overline{BP^{-1}B'}(s, k)\overline{K}(s, k) - \overline{Q}_*\overline{K}(s, \cdot)(k),\\ &\overline{K}(T, k) = \overline{D}(k).\end{aligned} \tag{4.23}$$

Proof. For notational simplicity, consider only a scalar case (K^ε is a scalar function). Following the proof of Theorem 4.3 up to (4.12), for $s \in [0,T]$, $\widetilde{Q}^k K(s,\cdot)(\zeta_{kj}) = 0$, for $k=1,\ldots,l, j=1,\ldots,m_k$, and

$$(\widetilde{Q}^1_*,\ldots,\widetilde{Q}^l_*,\widetilde{Q}_*)(K(s,\zeta_{11}),\ldots,K(s,\zeta_{1m_1}),\ldots,K(s,\zeta_{l1}),\\ \ldots,K(s,\zeta_{lml}),K(s,\zeta_{*1}),\ldots,K^0(s,\zeta_{*m*}))' = 0.$$

The irreducibility of $\widetilde{Q}^k$ implies $(K(s,\zeta_{k1}),\ldots,K^0(s,\zeta_{km_k}))' = K(s,k)\mathbb{1}_{m_k}$. Let $K_*(s) = (K(s,\zeta_{*1}),\ldots,K^{0,\prime}(s,\zeta_{*m_*}))'$. Then, $\widetilde{Q}^1_*\mathbb{1}_{m_1}K(s,1)+\cdots+\widetilde{Q}^l_*\mathbb{1}_{m_l}K(s,l)+\widetilde{Q}_*K_*(s) = 0$. Hence, $K_*(s) = a_{m_1}K(s,1)+\cdots+a_{m_l}K(s,l)$. The rest of except replacing $\overline{Q}$ by $\overline{Q}_*$. □

The convergence of K^ε leads to $v^\varepsilon(s,\zeta_{kj},x)\to v(s,k,x)$, for $k=1,\ldots,l$, $j=1,\ldots,m_k$, $v^\varepsilon(s,\zeta_{*j},x)\to v_*(s,j,x)$, for $j=1,\ldots,m_*$, where $v_*(s,j,x) = a_{m_1}(j)v(s,1,x)+\cdots+a_{m_l}(j)v(s,l,x)$ and $v(s,k,x) = x'\overline{K}(s,k)x$. The control set for the limit problem is the same as that for the recurrent case $\mathcal{U} = \{U=(U^1,\ldots,U^l) : U^k = (U^{k1},\ldots,u^{km_k}),\ u^{kj}\in\mathbb{R}^{n_2}\}$.

$$\begin{cases} \min J(s,k,x,U(\cdot)) = E\{\int_s^T[x'(t)\overline{M}(\overline{\alpha}(t))x(t)+\widetilde{N}(\overline{\alpha}(t),U(t))]dt + x'(T)\overline{D}(k)x(T)\} \\ \text{s.t.}\quad dx(t) = f(t,\overline{\alpha}(t),x(t),U(t))dt+\widetilde{C}(k,U)dw(t),\quad x(s)=x, \end{cases}$$

where $f(t,\overline{\alpha},x,U)$ and $\widetilde{C}(k,U)$ are defined in (4.13).

The optimal control for this limit problem is $U^o(s,k,x) = (U^{o,1}(s,x),\ldots,U^{o,l}(s,x))$ with $U^{o,k}(s,x) = (u^{o,k1}(s,x),\ldots,u^{o,km_k}(s,x))$ and $u^{o,kj} = -P^{-1}(s,\zeta_{kj})B'(\zeta_{kj})\overline{K}(s,k)x$. Similar to that of the recurrent case, construct

$$u^\varepsilon(s,\alpha,x) = \sum_{k=1}^{l}\sum_{j=1}^{m_k} I_{\{\alpha=\zeta_{kj}\}}u^{o,kj}(s,x) + \sum_{j=1}^{m_*} I_{\{\alpha=\zeta_{*j}\}}u^{o,*j}(s,x) \tag{4.24}$$

for the original problem, where $u^{o,*j}(s,x) = -P^{-1}(s,\zeta_{*j})B'(\zeta_{*j})\overline{K}_*(s,j)x$. Assume that $A(\zeta_{kj}) = A(k)$, $B(\zeta_{kj}) = B(k)$, $C(\zeta_{kj}) = C(k)$, $D(\zeta_{kj}) = D(k)$, $M(\zeta_{kj}) = M(k)$, and $N(\zeta_{kj}) = N(k)$ are independent of j. We may consider

$$\overline{u}^\varepsilon(s,\alpha,x) = \sum_{k=1}^{l}\sum_{j=1}^{m_k} I_{\{\alpha\in\mathcal{M}_k\}}\nu^k_j u^{o,kj}(s,x) + \sum_{j=1}^{m*} I_{\{\alpha=\zeta_{*j}\}}u^{o,*j}(s,x). \tag{4.25}$$

Note that $\overline{u}^\varepsilon$ only needs the information on if $\alpha^\varepsilon(t)\in\mathcal{M}_k$ for $k=1,\ldots,l$ and $\alpha^\varepsilon(t) = \zeta_{*j}$.

Theorem 4.6. *The following assertions hold.*

1) *The control* $u^\varepsilon(t)$ *defined in* (4.24) *is near-optimal in that* $\lim_{\varepsilon\to 0}|J^\varepsilon(s,\alpha,x,u^\varepsilon) - v^\varepsilon(s,\alpha,x)| = 0$.

2) *If* $A(\zeta_{kj}) = A(k)$, $B(\zeta_{kj}) = B(k)$, $C(\zeta_{kj}) = C(k)$, $D(\zeta_{kj}) = D(k)$, $M(\zeta_{kj}) = M(k)$, *and* $N(\zeta_{kj}) = N(k)$ *are independent of* j, $\overline{u}^{\varepsilon}(t)$ *defined in* (4.25) *is near-optimal in that* $\lim_{\varepsilon\to 0} |J^{\varepsilon}(s,\alpha,x,\overline{u}^{\varepsilon}) - v^{\varepsilon}(s,\alpha,x)| = 0$.

Proof. The proof is similar to that of Theorem 4.4, except the use of (4.22) lieu of (4.7) to verify the convergence of the trajectories. $\square$

5 A Numerical Example

As a demonstration, we consider a numerical example with strictly negative definite control weights $N(i)$ for all $i \in \mathcal{M}$. Suppose that the Markov chain $\alpha^{\varepsilon}(t)$ have three states, $\mathcal{M} = \{1,2,3\}$, and generator (4.1) with

$$\widetilde{Q} = \begin{pmatrix} -0.7 & 0.3 & 0.4 \\ 0.4 & -0.7 & 0.3 \\ 0.3 & 0.4 & -0.7 \end{pmatrix}, \text{ and } \widehat{Q} = \begin{pmatrix} -0.2 & 0.1 & 0.1 \\ 0.2 & -0.4 & 0.2 \\ 0.1 & 0.2 & -0.3 \end{pmatrix}. \tag{5.1}$$

Consider the control objective (4.2) subject to the system constraint (4.3), in which $x(0) = 3$, $\alpha^{\varepsilon}(0) = 2$, for $0 \le t \le T = 10$, $A(1) = 0.2$, $A(2) = 0.3$, $A(3) = 0.4, B(1) = 18$, $B(2) = 19$, $B(3) = 22, C(1) = 10$, $C(2) = 11$, $C(3) = 12, M(1) = 1$, $M(2) = 2$, $M(3) = 3, N(1) = -1$, $N(2) = -2$, $N(3) = -3$, $D(i) = 4$, $\nu = (1/3, 1/3, 1/3)$. The limit system of Riccati equations is given by (4.10) with $\overline{Q} = 0$.

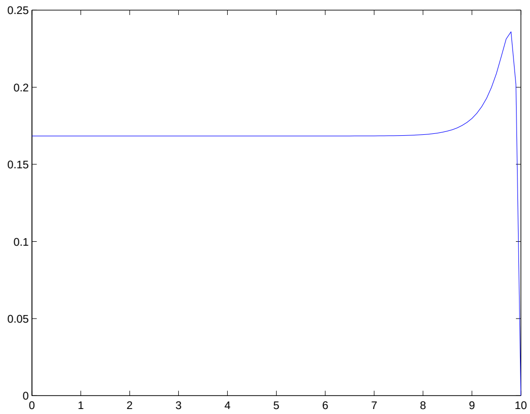
By using an explicit Runge-Kutta solver with relative accuracy tolerance 10^{-5} and absolute error tolerance 10^{-8}, we numerically solve the Riccati equations to obtain $K^{\varepsilon}(\cdot,i)$ and $\overline{K}(\cdot)$. The integration step size is chosen to be $h = 0.01$. Define the following norm to measure the average absolute errors and maximal relative errors:

$$|K^{\varepsilon} - \overline{K}| = \frac{h}{T}\sum_{k=1}^{3}\sum_{j=1}^{\lfloor \frac{T}{h} \rfloor} |K^{\varepsilon}(jh,k) - \overline{K}(jh)|, \tag{5.2}$$

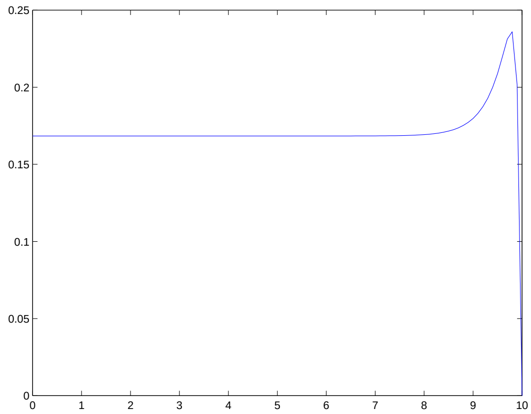
where $\lfloor z \rfloor$ denote the integer part of $z \in \mathbb{R}$. The calculated error bounds are provided in Table 1. It shows the dependence of the bounds on ε. To further demonstrate, we plot the trajectories of the difference of $K^{\varepsilon}(s,i)$ and $\overline{K}(s)$ by using the norm defined above. Figure 1 displays the results corresponding to $\varepsilon = 0.1$ and $\varepsilon = 0.01$, respectively. It is easily seen that the smaller the ε, the better approximation one obtains.

ε	$\lvert K^{\varepsilon} - \bar{K} \rvert$	$\lvert v^{\varepsilon} - v \rvert$
0.1	1.7234ε	0.3394ε
0.01	2.1477ε	0.3553ε
0.001	2.2003ε	0.3549ε
0.0001	2.2065ε	0.3618ε

Table 1: Error Bounds



(a) $\varepsilon = 0.1$, CPU times: 0.125 second.



(b) $\varepsilon = 0.01$, CPU times: 0.421 second.

Figure 1: Horizontal axes denote time t, and vertical axes denote $\sum_{i=1}^{3} |K^{\varepsilon}(t,i) - \overline{K}(t)|$; $0 \le t \le T$ and $T = 10$.

6 Concluding Remark

Near-optimal controls of LQ with regime switching and indefinite control weights are treated in this paper. Currently, the formulation is under the assumption that a scalar Brownian motion is used. The techniques used and the near-optimal control construction can be extended to multidimensional diffusion with regime switching. Although the asymptotic results are obtained with the small parameter $\varepsilon \to 0$, in practical systems, ε is a fixed parameter that may present different order of magnitude of the elements of the generators. The near-optimal control presented here gives an effective approximation using comparison controls, which indicates that instead of solving the complex systems with the Markov chain having large state space, we can use a much simplified system.

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] T. Björk, Finite dimensional optimal filters for a class of Ito processes with jumping parameters, *Stochastics*, **4** (1980), 167–183.

[3] S. Chen, X. Li, and X.Y. Zhou, Stochastic linear quadratic regulators with indefinite control weight costs, *SIAM J. Control Optim.* **36** (1998), 1685-1702.

[4] W. H. Fleming and R. W. Rishel, *Deterministic and Stochastic Optimal Control*, Springer-Verlag, New York, NY, 1975.

[5] J. K. Hale, *Ordinary Differential Equations*, R.E. Krieger Publishing Co., 2nd Ed., Malabar, 1980.

[6] F. Hoppenstead, H. Salehi, and A. Skorohod, An averaging priciple for dynamic systems in Hilber space with Markov random perturbations, *Stochastic Process Appl.*, **61** (1996), 85-108.

[7] Y. Ji and H. J. Chizeck, Jump linear quadratic Gaussian control in continuous-time, *IEEE Trans. Automat. Control*, **37** (1992), 1884–1892.

[8] X. Li and X.Y. Zhou, Indefinite stochastic LQ controls with Markovian jumps in a finite time horizon, *Comm. Inform. & Syst.*, **2** (2002), 265–282.

[9] D.G. Luenberger, *Investment Science*, Oxford University Press, New York, 1998.

[10] M. Mariton and P. Bertrand, Robust jump linear quadratic control: A mode stabilizing solution, *IEEE Trans. Automat. Control*, **30** (1985), 1145–1147.

[11] H. Markowitz, Portfolio selection, *J. Finance* **7** (1952), 77-91.

[12] Z. G. Pan and T. Başar, H^∞-control of Markovian jump linear systems and solutions to associated piecewise-deterministic differential games, in *New Trends in Dynamic Games and Applications*, G. J. Olsder (Ed.), 61-94, Birkhäuser, Boston, 1995.

[13] A. A. Pervozvanskii and V. G. Gaitsgori, *Theory of Suboptimal Decisions: Decomposition and Aggregation*, Kluwer, Dordrecht, 1988.

[14] 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.

[15] S. P. Sethi and Q. Zhang, *Hierarchical Decision Making in Stochastic Manufacturing Systems*, Birkhäuser, Boston, MA, 1994.

[16] W. Walter, *Ordinary Differential Equations*, Springer, New York, 1998.

[17] G. Yin and S. Dey, Weak convergence of hybrid filtering problems involving nearly completely decomposable hidden Markov chains, *SIAM J. Control Optim.*, **41** (2003), 1820-1842.

[18] G. Yin and Q. Zhang, *Continuous-Time Markov Chains and Applications: A Singular Perturbation Approach*, Springer-Verlag, New York, 1998.

[19] 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.

[20] J. Yong and X.Y. Zhou, *Stochastic Controls: Hamiltonian Systems and HJB Equations*, Springer-Verlag, New York, 1999.

[21] Q. Zhang and G. Yin, On nearly optimal controls of hybrid LQG problems, *IEEE Trans. Automat. Control*, **44** (1999) 2271-2282.

[22] X.Y. Zhou and G. Yin, Markowitz's mean-variance portfolio selection with regime switching: A Continuous-time model, *SIAM J. Control Optim*, **42** (2003), 1466-1482.