

RESEARCH

Open Access



Fully coupled forward-backward stochastic differential equations on Markov chains

Shaolin Ji¹, Haodong Liu¹ and Xinling Xiao^{2*}

*Correspondence:

xinlingxiao@mail.sdu.edu.cn

²School of Mathematical Sciences, Shandong Normal University, Jinan, 250014, P.R. China

Full list of author information is available at the end of the article

Abstract

We define fully coupled forward-backward stochastic differential equations on spaces related to continuous-time finite-state Markov chains. Existence and uniqueness results of the fully coupled forward-backward stochastic differential equations on Markov chains are obtained.

Keywords: forward-backward stochastic differential equations; monotone assumption; Markov chain

1 Introduction

Since the first introduction by Pardoux and Peng [1] in 1990, the theory of nonlinear backward stochastic differential equations (BSDEs) driven by a Brownian motion has been intensively researched by many researchers and has achieved abundant theoretical results. Now this theory is a powerful tool in stochastic analysis. It also has many important applications, namely in stochastic control, stochastic differential games, finance, and the theory of partial differential equations (PDEs).

In the classic BSDE theory, we consider a Brownian motion as the driver, but a Brownian motion is a kind of very idealized stochastic model, which limits greatly the applications of the classic BSDEs. There are many results about BSDEs associated with jump process. Tang and Li [2] first discussed BSDEs driven by a Brownian motion and Poisson process; Nualart and Schoutens [3] considered BSDEs driven by a Brownian motion and Lévy process. Furthermore, there are also results where the Brownian motion in the diffusion term of the BSDE is replaced by another process. For example, Cohen and Elliott [4] studied BSDEs driven by a continuous-time finite-state Markov chain. After that, many results, such as a comparison theorem about this kind of BSDEs, nonlinear expected results [5, 6], and so on, appeared.

Along with the rapid development of the BSDE theory, the theory of fully coupled forward-backward stochastic differential equations (FBSDEs), closely related to BSDEs, has been developed rapidly. Fully coupled FBSDEs with Brownian motion can be encountered in the optimization problem when applying stochastic maximum principle (see [7]) and in mathematical finance when considering a large investor in security market (see [8]). Such FBSDEs are also used in the potential theory (see [9]). As we know now, to get the

existence and uniqueness results of fully coupled FBSDEs solutions, there are mainly three methods: the method of contraction mappings [10, 11], the four-step scheme [12], and the method of continuation [13, 14]. For more details on fully coupled FBSDEs, we refer to Yong [15], Ma *et al.* [16], or Briand and Hu [17] and the references therein.

In this paper, we study fully coupled FBSDEs driven by a martingale generated by a continuous-time finite-state Markov chain. Inspired by Peng and Wu [14], we introduce an $m \times n$ full-rank matrix G to overcome the problem caused by the different dimensions of SDE and BSDE. Using the method of continuation, the Itô product rule of semimartingales, and the fixed point principle, based on the theory of BSDEs driven by a continuous-time finite-state Markov chain, we obtain existence and uniqueness results of the FBSDEs on a Markov chain. It is worth pointing out that due to the property of the martingale generated by a finite-state Markov chain, the form of the monotone assumptions we employed here is different from that in Peng and Wu [14].

Recently, many works have been done on BSDEs with Markov chains. In fact, there are two major formulations of the state space of the Markov chain in the literature. One is the set of unit vectors in some Euclidean space [18]. The other one is the set of positive integers [19]. Each of the formulations was used by many researchers when they studied BSDEs driven by Markov chains. When a martingale representation was needed, the second formulation was necessary. The first one facilitates the mathematics, that is, the symbolism is nicer; the second one makes the mathematics more rigorous. Given a continuous-time Markov chain, assuming its state space be the set of positive integers, there exists an integer-valued random measure that counts the jumps of a Markov chain. Crepey and Matoussi [19] and Crepey [20] considered BSDEs driven by both a Brownian motion and the compensated martingale of a random measure. Tao *et al.* [21] discussed BSDEs with a singular perturbed Markov chain; Tao *et al.* [22, 23] considered the regime-switching system modulated by a Markov chain based on Crepey and Matoussi [19] and Crepey [20]. When the state space of a continuous-time Markov chain is described by the set of unit vectors in R^d , Zhang *et al.* [18] considered a system of Markov regime-switching model with Poisson jumps. Different from the above works, based on Cohen and Elliott [4], we discussed FBSDEs purely driven by a Markov chain rather than by a Brownian motion. Our model depends only on the state space of a Markov chain and leads to a special structure.

In Section 2, we formulate our problem; in Section 3, the preliminaries are given; in Section 4, we give our main result on the existence and uniqueness of the fully coupled FBSDE on a Markov chain and present the proofs; in Section 5, we discuss the existence and uniqueness result of the FBSDE under other monotone assumptions; in the last section, we give an application of these theoretical results to stochastic optimal control problem through an example.

2 Formulation of the problem

Consider a continuous-time finite-state Markov chain $m = \{m_t, t \in [0, T]\}$. Following the convention of Elliott *et al.* [24], we assume that it takes values in unit vectors e_i in R^d , where d is the number of states of the chain.

We consider stochastic processes defined on the filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, where $\{\mathcal{F}_t\}$ is the completed natural filtration generated by the σ -fields $\mathcal{F}_t = \sigma(\{m_s, s \leq t\}, F \in \mathcal{F}_T : \mathbb{P}(F) = 0)$, and $\mathcal{F} = \mathcal{F}_T$. Note that if m is right-continuous, then this filtration is right-continuous. If \mathcal{A}_t denotes the rate matrix for m at time t , then this chain has

the following representation:

$$m_t = m_0 + \int_0^t \mathcal{A}_s m_s ds + M_t,$$

where M_t is a martingale (see [24]).

We consider the following forward-backward stochastic differential equations:

$$\begin{cases} X_t = x + \int_0^t b(s, X_s, Y_s, Z_s) ds + \int_0^t \sigma(s-, X_{s-}, Y_{s-}, Z_{s-}) dM_s, \\ Y_t = \Phi(X_T) + \int_t^T f(s, X_s, Y_s, Z_s) ds - \int_t^T Z_{s-} dM_s, \end{cases} \tag{1}$$

where X, Y, Z take values in $\mathbb{R}^n, \mathbb{R}^m, \mathbb{R}^{m \times d}$, $T > 0$ is an arbitrary fixed number, and b, σ, f, Φ are functions of appropriate dimensions:

$$\begin{aligned} b &: \Omega \times [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m \times d} \longrightarrow \mathbb{R}^n, \\ \sigma &: \Omega \times [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m \times d} \longrightarrow \mathbb{R}^{n \times d}, \\ f &: \Omega \times [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m \times d} \longrightarrow \mathbb{R}^m, \\ \Phi &: \Omega \times \mathbb{R}^n \longrightarrow \mathbb{R}^m. \end{aligned}$$

We shall seek an \mathcal{F}_t -adapted triple (X, Y, Z) that satisfies the forward-backward stochastic differential equation (1) on $[0, T]$ \mathbb{P} -almost surely. That is, our aim is to find an \mathcal{F}_t -adapted solution of (1).

Noting that only the s -left limit $\sigma(\omega, s-, X_{s-}, Y_{s-}, Z_{s-})$ enters into FBSDE (1), without loss of generality, we assume that $\sigma(s, X_s, Y_s, Z_s)$ is left-continuous in s for all ω and X, Y . Since M is a semimartingale, X, Y is *càdlàg* and adapted. We suppose the existence of the left limits of Z . So Z must have at most a countable number of discontinuities, and then it must be left-continuous at each t except possibly on a dt -null set. Hence, if Z_s satisfies FBSDE (1), then so does Z_{s-} :

$$\begin{cases} X_t = x + \int_0^t b(s, X_s, Y_s, Z_{s-}) ds + \int_0^t \sigma(s-, X_{s-}, Y_{s-}, Z_{s-}) dM_s, \\ Y_t = \Phi(X_T) + \int_t^T f(s, X_s, Y_s, Z_{s-}) ds - \int_t^T Z_{s-} dM_s. \end{cases}$$

Setting $Z_t^{**} := Z_{t-}$, we have a left-continuous process Z^{**} , which also satisfies the desired equations. Therefore, writing the left limits Z_{t-} is unnecessary since we simply assume that our solution is left-continuous.

Based on these arguments, we rewrite FBSDE (1) as

$$\begin{cases} X_t = x + \int_0^t b(s, X_s, Y_s, Z_s) ds + \int_0^t \sigma(s, X_{s-}, Y_{s-}, Z_s) dM_s, \\ Y_t = \Phi(X_T) + \int_t^T f(s, X_s, Y_s, Z_s) ds - \int_t^T Z_s dM_s. \end{cases}$$

Recall that FBSDE (1) is equivalent to

$$\begin{cases} X_t = x + \int_0^t b^{**}(s, X_s, Y_s, Z_s) ds + \int_0^t \sigma(s, X_{s-}, Y_{s-}, Z_s) dm_s, \\ Y_t = \Phi(X_T) + \int_t^T f^{**}(s, X_s, Y_s, Z_s) ds - \int_t^T Z_s dm_s, \end{cases} \tag{2}$$

where $b^{**}(s, X_s, Y_s, Z_s) = b(s, X_s, Y_s, Z_s) - \sigma(s, X_s, Y_s, Z_s)\mathcal{A}_s m_s$ and $f^{**}(s, X_s, Y_s, Z_s) = f(s, X_s, Y_s, Z_s) + Z_s \mathcal{A}_s m_s$.

Obviously, FBSDE (2) is driven by the Markov chain $m = \{m_t, t \in [0, T]\}$.

3 Preliminaries

3.1 Preliminary notation

Elliott *et al.* [24] obtained the optional quadratic variation of M_t as the matrix process

$$[M, M]_t = \sum_{s=0}^t \Delta M_s \Delta M_s^*,$$

where ‘*’ means transpose.

Observing that \mathcal{A} is the rate matrix of the Markov chain m , the predictable quadratic variation is

$$\langle M, M \rangle_t = \int_0^t [\text{diag}(\mathcal{A}_s m_s) - \text{diag}(m_s) \mathcal{A}_s^* - \mathcal{A}_s \text{diag}(m_s)] ds.$$

This can be seen through considering

$$\text{diag}(m_t) = \text{diag}(m_0) + \int_0^t \text{diag}(\mathcal{A}_s m_s) ds + \int_0^t \text{diag}(dM_s)$$

and

$$\begin{aligned} \text{diag}(m_t) &= m_t m_t^* \\ &= m_0 m_0^* + \int_0^t m_s m_s^* \mathcal{A}_s^* ds + \int_0^t m_{s-} dM_s^* + \int_0^t \mathcal{A}_s m_s m_s^* ds \\ &\quad + \int_0^t dM_s m_{s-}^* + \sum_{s=0}^t \Delta M_s \Delta M_s^*. \end{aligned}$$

Equating these two, we get

$$[M, M]_t = L_t + \int_0^t [\text{diag}(\mathcal{A}_s m_s) - \text{diag}(m_s) \mathcal{A}_s^* - \mathcal{A}_s \text{diag}(m_s)] ds,$$

where L is some martingale. This in turn suggests that

$$\langle M, M \rangle_t = \int_0^t [\text{diag}(\mathcal{A}_s m_s) - \text{diag}(m_s) \mathcal{A}_s^* - \mathcal{A}_s \text{diag}(m_s)] ds.$$

We will use the following notation:

By (\cdot, \cdot) we denote the usual inner product in \mathbb{R}^n or \mathbb{R}^m ; we use the usual Euclidean norm in \mathbb{R}^n and \mathbb{R}^m ; and for $z \in \mathbb{R}^{m \times d}$, we define $|z| = \{\text{tr}(zz^*)\}^{\frac{1}{2}}$.

For $z^1 \in \mathbb{R}^{m \times d}$, $z^2 \in \mathbb{R}^{m \times d}$,

$$((z^1, z^2)) = \text{tr}(z^1 (z^2)^*),$$

and for $u^1 = (x^1, y^1, z^1) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m \times d}$, $u^2 = (x^2, y^2, z^2) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m \times d}$,

$$[u^1, u^2] = (x^1, x^2) + (y^1, y^2) + ((z^1, z^2)).$$

We are given an $m \times n$ full-rank matrix G . For $u = (x, y, z) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m \times d}$, let

$$F(t, u) = (-G^*f(t, u), Gb(t, u), 0), \quad H(t, u) = (0, 0, G\sigma(t, u)),$$

where $G\sigma = (G\sigma_1 \cdots G\sigma_d)$.

Furthermore, we define

$$\langle C, D \rangle_V = \text{tr}(C[\text{diag}(\mathcal{A}_s v) - \text{diag}(v)\mathcal{A}_s^* - \mathcal{A}_s \text{diag}(v)]D^*), \quad \|C\|_V^2 = \langle C, C \rangle_V,$$

where v is a basis vector in \mathbb{R}^d .

3.2 BSDEs on Markov chains

We recall the existence and uniqueness results for a solution to the following BSDE on a Markov chain:

$$Y_t = \Phi(X_T) - \int_t^T F_1(\omega, s, Y_s, Z_s) ds - \int_t^T [F_2(s, Y_{s-}) + Z_s] dM_s$$

for functions $F_1 : \Omega \times [0, T] \times \mathbb{R}^n \times \mathbb{R}^{n \times d}$ and $F_2 : \Omega \times [0, T] \times \mathbb{R}^n$; these functions are assumed to be progressively measurable, that is, $F_1(\cdot, s, Y_s, Z_s)$ and $F_2(\cdot, s, Y_s)$ are \mathcal{F}_t -measurable for all $t \in [0, T]$.

Cohen and Elliott [4] gave the following result about martingale representation.

Lemma 1 *Any \mathbb{R}^d -valued martingale L defined on $(\Omega, \mathcal{F}_t, \mathbb{P})$ can be represented as a stochastic (in this case, Stieltjes) integral with respect to the martingale process M up to equality \mathbb{P} -a.s. This representation is unique up to a $d\langle M, M \rangle_t \times \mathbb{P}$ -null set. That is,*

$$L_t = L_0 + \int_0^t Z_s dM_s,$$

where Z_s is a predictable $\mathbb{R}^{n \times d}$ -valued matrix process.

Cohen and Elliott [4] also obtained the existence and uniqueness of a solution to the above BSDE.

Theorem 1 *Assume the Lipschitz continuity on the generators F_1 and F_2 , that is, that there exists $c \in \mathbb{R}$ such that, for all $s \in [0, T]$,*

$$E|F_1(s, Y_s^1, Z_s^1) - F_1(s, Y_s^2, Z_s^2)|^2 \leq c^2 E|Y_s^1 - Y_s^2|^2 + c^2 E\|Z_s^1 - Z_s^2\|_{M_s}^2,$$

$$E\|F_2(s, Y_{s-}^1) - F_2(s, Y_{s-}^2)\|_{M_s}^2 \leq c^2 E|Y_s^1 - Y_s^2|^2.$$

Under this Lipschitz condition, the above equation has at most one solution up to indistinguishability for Y and equality $d\langle M, M \rangle_t \times \mathbb{P}$ -a.s. for Y .

4 Existence and uniqueness for FBSDE (1)

We give two definitions.

Definition 1 We denote by $\mathcal{M}^2(0, T; \mathbb{R}^n)$ the set of all \mathbb{R}^n -valued \mathcal{F}_t -adapted processes such that

$$E \int_0^T |v(s)|^2 ds < +\infty.$$

Definition 2 A triple of processes $(X, Y, Z) : \Omega \times [0, T] \rightarrow \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m \times d}$ is called an adapted solution of the FBSDE (1) if $(X, Y, Z) \in \mathcal{M}^2(0, T; \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m \times d})$ and it satisfies FBSDE (1) \mathbb{P} -almost surely.

The adaptedness of the solution enables us to rewrite FBSDE (1) in a differential form:

$$\begin{cases} dX_t = b(t, X_t, Y_t, Z_t) dt + \sigma(t, X_t, Y_t, Z_t) dM_t, \\ -dY_t = f(t, X_t, Y_t, Z_t) dt - Z_t dM_t, \\ X_0 = x, \quad Y_T = \Phi(X_T). \end{cases}$$

Now we give the main assumptions of our paper.

Assumption 1 For each $u = (x, y, z) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m \times d}$, $F(\cdot, u), H(\cdot, u) \in \mathcal{M}^2(0, T; \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m \times d})$, and for each $x \in \mathbb{R}^n$, $\Phi(x) \in L^2(\Omega, \mathcal{F}_T; \mathbb{R}^n)$. Moreover,

- (i) $F(t, u)$ is uniformly Lipschitz with respect to u ;
- (ii) $H(t, u)$ is uniformly Lipschitz with respect to u ;
- (iii) $\Phi(x)$ is uniformly Lipschitz with respect to x .

Assumption 2 There exists constants c_2, c'_2, c_3 such that

$$\begin{aligned} [F(t, u^1) - F(t, u^2), u^1 - u^2] &\leq -c_2 |G(x^1 - x^2)|^2 - c'_2 (|G^*(y^1 - y^2)|^2 + |G^*(z^1 - z^2)|^2), \\ [H(t, u^1) - H(t, u^2), u^1 - u^2] &\leq -c_2 |G(x^1 - x^2)|^2 - c'_2 (|G^*(y^1 - y^2)|^2 + |G^*(z^1 - z^2)|^2), \\ \forall u^1 = (x^1, y^1, z^1), u^2 = (x^2, y^2, z^2) &\in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m \times d}, \mathbb{P}\text{-a.s., a.e. } t \in \mathbb{R}^+, \end{aligned}$$

and

$$(\Phi(x^1) - \Phi(x^2), G(x^1 - x^2)) \geq c_3 |G(x^1 - x^2)|^2, \quad \forall x^1 \in \mathbb{R}^n, \forall x^2 \in \mathbb{R}^n,$$

where c_2, c'_2 , and c_3 are given positive constants.

Now we give the main result of our paper.

Theorem 2 Let Assumptions 1 and 2 hold. Then there exists a unique adapted solution (X, Y, Z) of FBSDE (1).

Proof of uniqueness Let $U^1 = (X^1, Y^1, Z^1)$ and $U^2 = (X^2, Y^2, Z^2)$ be two adapted solutions of FBSDE (1). We set

$$\begin{aligned} (\hat{X}, \hat{Y}, \hat{Z}) &= (X^1 - X^2, Y^1 - Y^2, Z^1 - Z^2), \\ \begin{cases} \hat{b}(t) = b(t, U_t^1) - b(t, U_t^2), \\ \hat{\sigma}(t) = \sigma(t, U_t^1) - \sigma(t, U_t^2), \\ \hat{f}(t) = f(t, U_t^1) - f(t, U_t^2). \end{cases} \end{aligned}$$

Then we have

$$\begin{cases} d\hat{X}_t = \hat{b}(t, X_t, Y_t, Z_t) dt + \hat{\sigma}(t, X_{t-}, Y_{t-}, Z_t) dM_t, \\ -d\hat{Y}_t = \hat{f}(t, X_t, Y_t, Z_t) dt - \hat{Z}_t dM_t. \end{cases}$$

From Assumption 1 it follows that

$$E(\sup_{t \in [0, T]} |\hat{X}_t|^2) + E(\sup_{t \in [0, T]} |\hat{Y}_t|^2) < +\infty.$$

Using the Itô product rule of semimartingales, we have

$$d(G\hat{X}_s, \hat{Y}_s) = \hat{Y}_{s-} d(G\hat{X}_s) + (G\hat{X}_{s-}) d\hat{Y}_s + d(G\hat{X}_s) d\hat{Y}_s,$$

where $d(G\hat{X}_s) d\hat{Y}_s = d[G\hat{X}_s, \hat{Y}_s]_t$, and hence, taking expectation and evaluating at $t = T$, by Assumption 2 we get

$$\begin{aligned} E(\hat{Y}_T, G\hat{X}_T) &= E(\Phi(X_T^1) - \Phi(X_T^2), G(X_T^1 - X_T^2)) \\ &= E \int_0^T [(-G^* \hat{f}, \hat{X}_s) + (G\hat{b}, \hat{Y}_s)] ds + E \sum_{s=0}^T ((G\hat{\sigma}, \hat{Z}_s) \Delta M_s \Delta M_s^*) \\ &= E \int_0^T [F(s, U^1) - F(s, U^2), U^1 - U^2] ds \\ &\quad + E \sum_{s=0}^T [H(s, U^1) - H(s, U^2), U^1 - U^2] \Delta M_s \Delta M_s^* \\ &= E \int_0^T [F(s, U^1) - F(s, U^2), U^1 - U^2] ds \\ &\quad + E \int_0^T [H(s, U^1) - H(s, U^2), U^1 - U^2] d\langle M, M \rangle_s \\ &= E \int_0^T [F(s, U^1) - F(s, U^2), U^1 - U^2] ds \\ &\quad + E \int_0^T [H(s, U^1) - H(s, U^2), U^1 - U^2] d\langle M, M \rangle_s \\ &\leq -c_2 E \int_0^T |G\hat{X}|^2 ds - c'_2 E \int_0^T (|G^* \hat{Y}|^2 + |G^* \hat{Z}|^2) ds \\ &\quad - c_2 E \int_0^T |G\hat{X}|^2 d\langle M, M \rangle_s - c'_2 E \int_0^T (|G^* \hat{Y}|^2 + |G^* \hat{Z}|^2) d\langle M, M \rangle_s. \end{aligned}$$

Then we deduce that

$$\begin{aligned} &c_2 \left(E \int_0^T |G\hat{X}|^2 ds + E \int_0^T |G\hat{X}|^2 d\langle M, M \rangle_s \right) \\ &\quad + c'_2 \left[E \int_0^T (|G^* \hat{Y}|^2 + |G^* \hat{Z}|^2) ds + E \int_0^T (|G^* \hat{Y}|^2 + |G^* \hat{Z}|^2) d\langle M, M \rangle_s \right] + c_3 |G\hat{X}|^2 \\ &\leq 0. \end{aligned}$$

Since $c_2, c'_2,$ and c_3 are given positive constants, we have $|G\hat{X}|^2 \equiv 0, |G^*\hat{Y}|^2 \equiv 0, |G^*\hat{Z}|^2 \equiv 0$. So $\hat{X}_s^1 \equiv \hat{X}_s^2, \hat{Y}_s^1 \equiv \hat{Y}_s^2, \hat{Z}_s^1 \equiv \hat{Z}_s^2, d\langle M, M \rangle_t \times P$ -a.s. \square

We now consider the following family of FBSDEs parameterized by $l \in [0, 1]$:

$$\begin{cases} dX_t^l = [(1-l)c'_2(-G^*Y_t^l) + lb(t, U_t^l) + \phi_t] dt \\ \quad + [(1-l)c'_2(-G^*Z_t^l) + l\sigma(t, U_t^l) + \psi_t] dM_t, \\ -dY_t^l = [(1-l)c_2GX_t^l + lf(t, U_t^l) + \gamma_t] dt - Z_t^l dM_t, \\ X_0^l = x, \quad Y_T^l = l\Phi(X_T^l) + (1-l)GX_T^l + \xi, \end{cases} \tag{3}$$

where $\phi, \psi,$ and γ are given processes in $M^2(0, T)$ with values in $R^n, R^{n \times d},$ and $R^m,$ respectively. Clearly, when $l = 1, \xi \equiv 0,$ the existence of a solution of FBSDE (3) implies the existence of that of FBSDE (1). When $l = 0$ in FBSDE (3), it is

$$\begin{cases} dX_t^0 = [-c'_2G^*Y_t^0 + \phi_t] dt + [-c'_2G^*Z_t^0 + \psi_t] dM_t, \\ -dY_t^0 = [c_2GX_t^0 + \gamma_t] dt - Z_t^0 dM_t, \\ X_0^0 = x, \quad Y_T^0 = GX_T^0 + \xi. \end{cases} \tag{4}$$

For proving the existence part of the theorem, we first need the following two lemmas.

Lemma 2 *The following equation has a unique solution:*

$$\begin{cases} dX_t = [-c'_2G^*Y_t + \phi_t] dt + [-c'_2G^*Z_t + \psi_t] dM_t, \\ -dY_t = [c_2GX_t + \gamma_t] dt - Z_t dM_t, \\ X_0 = x, \quad Y_T = \lambda GX_T + \xi, \end{cases} \tag{5}$$

where λ is a nonnegative constant.

Proof We note that the matrix G is of full rank. The proof of the existence for FBSDE (5) can be divided into two cases, $n \leq m$ and $n > m$.

For the first case, the matrix G^*G is of full rank. Let

$$\begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix} = \begin{pmatrix} X \\ G^*Y \\ G^*Z \end{pmatrix}, \quad \begin{pmatrix} Y'' \\ Z'' \end{pmatrix} = \begin{pmatrix} (I_m - G(G^*G)^{-1}G^*)Y \\ (I_m - G(G^*G)^{-1}G^*)Z \end{pmatrix}.$$

Multiplying G^* for (Y, Z) on both sides of the BSDE yields

$$\begin{cases} dX_t' = [-c'_2Y_t' + \phi_t] dt + [-c'_2Z_t' + \psi_t] dM_t, \\ -dY_t' = [c_2G^*GX_t' + G^*\gamma_t] dt - Z_t' dM_t, \\ X_0' = x, \quad Y_T' = \lambda G^*GX_T' + G^*\xi. \end{cases} \tag{6}$$

Similarly, multiplying $(I_m - G(G^*G)^{-1}G^*)$ on both sides of the same equation gives

$$\begin{cases} -dY_t'' = (I_m - G(G^*G)^{-1}G^*)\gamma_t dt - Z_t'' dM_t, \\ Y_T'' = (I_m - G(G^*G)^{-1}G^*)\xi. \end{cases}$$

Obviously, the pair (Y'', Z'') is uniquely determined (see Cohen and Elliott [4]). The uniqueness of (X', Y', Z') follows from Theorem 2. In order to solve FBSDE (6), we introduce the following $n \times n$ -symmetric matrix-valued ODE, which is known as the matrix-Riccati equation:

$$\begin{cases} -\dot{K}(t) = -c'_2 K^2 + c_2 G^* G, & t \in [0, T], \\ K_T = \lambda G^* G. \end{cases}$$

It is well known that this equation has a unique nonnegative solution $K \in C^1([0, T]; S^n)$, where S^n stands for the space of all $n \times n$ -symmetric matrices. We then study the solution $(p, q) \in M^2(0, T; R^{n+n \times d})$ of the following linear simple BSDE:

$$\begin{cases} -dp_t = [-c'_2 K(t)p_t + K(t)\phi_t + G^* \gamma_t] dt + [K(t)\psi_t - (I_n + c'_2 K(t))q_t] dM_t, & t \in [0, T], \\ p_T = G^* \xi. \end{cases}$$

We now consider X'_t as the solution of the SDE

$$\begin{cases} dX'_t = [-c'_2 (K(t)X'_t + p_t) + \phi_t] dt + [\psi_t - c'_2 q_t] dM_t, \\ X'_0 = x. \end{cases}$$

Then we can easily check that $(X'_t, Y'_t, Z'_t) = (X'_t, K(t)X'_t + p_t, q_t)$ is a solution of FBSDE (6). Once (X', Y', Z') and (Y'', Z'') are resolved, we can obtain the triple (X, Y, Z) uniquely by

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} X' \\ G(G^*G)^{-1}Y' + Y'' \\ G(G^*G)^{-1}Z' + Z'' \end{pmatrix}.$$

For the second case, the matrix GG^* is of full rank. We set

$$\begin{pmatrix} X' \\ Y' \\ Z' \\ X'' \end{pmatrix} = \begin{pmatrix} GX \\ Y \\ Z \\ (I_n - G^*(GG^*)^{-1}G)X \end{pmatrix},$$

where X'' is the unique solution of the following linear SDE:

$$\begin{cases} dX''_t = (I_n - G^*(GG^*)^{-1}G)\phi_t dt + (I_n - G^*(GG^*)^{-1}G)\varphi_t dM_t, \\ X''_0 = (I_n - G^*(GG^*)^{-1}G)x. \end{cases}$$

The triple (X', Y', Z') solves the FBSDE

$$\begin{cases} dX'_t = [-c'_2 GG^* Y'_t + G\phi_t] dt + [-c'_2 GG^* Z'_t + G\psi_t] dM_t, \\ -dY'_t = [c_2 X'_t + \gamma_t] dt - Z'_t dM_t, \\ X'_0 = Gx, \quad Y'_T = \lambda X'_T + \xi. \end{cases} \tag{7}$$

In order to solve this equation, we introduce the following $m \times m$ -symmetric matrix-valued matrix-Raccati equation:

$$\begin{cases} -\dot{K}(t) = c_2 I_m - c_2' K G G^* K, & t \in [0, T], \\ K(T) = \lambda I_m. \end{cases}$$

It is well known that this equation has a unique nonnegative solution $K \in C^1([0, T]; S^m)$, where S^m stands for the space of all $m \times m$ -symmetric matrices. We then look for the solution $(p, q) \in M^2(0, T; R^{n+m \times d})$ of the following linear simple BSDE:

$$\begin{cases} -dp_t = [-c_2' K(t) G G^* p_t + K(t) G \phi_t + \gamma_t] dt + (K(t) G \psi_t - (I_m + c_2' K(t) G G^*) q_t) dM_t, \\ t \in [0, T], \\ p_T = \xi. \end{cases}$$

We now let X'_t be the solution of the following SDE:

$$\begin{cases} dX'_t = [-c_2' G G^* (K(t) X'_t + \gamma_t) + G \phi_t] dt + [G \psi_t - c_2' G G^* q_t] dM_t, \\ X'_0 = Gx. \end{cases}$$

Then we can easily check that $(X'_t, Y'_t, Z'_t) = (X'_t, K(t) X'_t + p_t, q_t)$ is a solution of FBSDE (7). Once (X', X'', Y', Z') are resolved, then by the definition we can obtain the triple (X, Y, Z) uniquely as

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} G^* (G G^*)^{-1} X' + X'' \\ Y' \\ Z' \end{pmatrix}.$$

The proof is complete. □

Lemma 3 *We assume Assumptions 1 and 2. If for an $l_0 \in [0, 1)$, there exists a solution $(X^{l_0}, Y^{l_0}, Z^{l_0})$ of FBSDE (3), then there exists a positive constant δ_0 such that for each $\delta \in [0, \delta_0]$, there exists a solution $(X^{l_0+\delta}, Y^{l_0+\delta}, Z^{l_0+\delta})$ of FBSDE (3) for $l = l_0 + \delta$.*

Proof Since for all $\phi \in M^2(0, T; \mathbb{R}^n)$, $\gamma \in M^2(0, T; \mathbb{R}^m)$, $\psi \in M^2(0, T; \mathbb{R}^{n \times d})$, $\xi \in L^2(\Omega, \mathcal{F}_T, \mathbb{P})$, $x \in \mathbb{R}^n$, $l_0 \in [0, 1)$, there exists a unique solution of (3), for every $U_s = (X_s, Y_s, Z_s) \in M^2(0, T; \mathbb{R}^{n+m \times d})$, there exists a unique triple $u_s = (x_s, y_s, z_s) \in M^2(0, T; \mathbb{R}^{n+m \times d})$ satisfying the following FBSDE:

$$\begin{cases} dx_t = [(1 - l_0)c_2' (-G^* y_t) + l_0 b(t, u_t) + \delta b(t, U_t) + \delta c_2' (G^* Y_t) + \phi_t] dt \\ \quad + [(1 - l_0)c_2' (-G^* z_t) + l_0 \sigma(t, u_t) + \delta \sigma(t, U_t) + \delta c_2' (G^* Z_t) + \psi_t] dM_t, \\ -dy_t = [(1 - l_0)c_2 G x_t + l_0 f(t, u_t) + \delta (-c_2 G X_t) + \delta f(t, U_t) + \gamma_t] dt - z_t dM_t, \\ x_0 = x, \quad y_T = l_0 \Phi(x_T) + (1 - l_0) G x_T + \delta (\Phi(X_T) - G X_T) + \xi. \end{cases}$$

We want to prove that the mapping defined by

$$\begin{aligned} I_{l_0+\delta}(U \times X_T) &= u \times x_T : M^2(0, T; \mathbb{R}^{n+m \times d}) \times L^2(\Omega, \mathcal{F}_T, \mathbb{P}) \\ &\rightarrow M^2(0, T; \mathbb{R}^{n+m \times d}) \times L^2(\Omega, \mathcal{F}_T, \mathbb{P}) \end{aligned}$$

is a contraction.

Let $U' \times X'_T = (X', Y', Z') \times X'_T \in M^2(0, T; \mathbb{R}^{n+m+m \times d}) \times L^2(\Omega, \mathcal{F}_T, \mathbb{P})$ and $u' \times x'_T = I_{l_0+\delta}(U' \times X'_T)$. We set

$$\begin{aligned} \hat{U} &= (\hat{X}, \hat{Y}, \hat{Z}) = (X - X', Y - Y', Z - Z'), \\ \hat{u} &= (\hat{x}, \hat{y}, \hat{z}) = (x - x', y - y', z - z'), \\ \hat{f}_s &= f(s, U_s) - f(s, U'_s), \quad \hat{b}_s = b(s, U_s) - b(s, U'_s), \quad \hat{\sigma}_s = \sigma(s, U_s) - \sigma(s, U'_s), \\ \bar{f}_s &= f(s, u_s) - f(s, u'_s), \quad \bar{b}_s = b(s, u_s) - b(s, u'_s), \quad \bar{\sigma}_s = \sigma(s, u_s) - \sigma(s, u'_s). \end{aligned}$$

Then

$$\begin{cases} d\hat{x}_s = [(1 - l_0)c'_2(-G^*\hat{y}_s) + l_0\bar{b}_s + \delta\hat{b}_s + \delta c'_2 G^*\hat{Y}_s] ds \\ \quad + [(1 - l_0)c'_2(-G^*\hat{z}_s) + l_0\bar{\sigma}_s + \delta\hat{\sigma}_s + \delta c'_2 G^*\hat{Z}_s] dM_s, \\ -d\hat{y}_s = [(1 - l_0)c_2 G\hat{x}_s + l_0\bar{f}_s + \delta(-c_2 G\hat{X}_s) + \delta\hat{f}_s] ds - \hat{z}_s dM_s, \\ \hat{x}_0 = 0, \quad \hat{y}_T = l_0(\Phi(x_T) - \Phi(x'_T)) + (1 - l_0)G\hat{x}_T + \delta(\Phi(x_T) - \Phi(x'_T) - G\hat{x}_T). \end{cases}$$

Using the Itô product rule of semimartingales for the product $(G\hat{x}_s, \hat{y}_s)$, taking expectations, and evaluating at $t = T$ yield

$$\begin{aligned} & l_0 E(\Phi(x_T) - \Phi(x'_T), G\hat{x}_T) + (1 - l_0)E(G\hat{x}_T, G\hat{x}_T) + \delta E(\Phi(x_T) - \Phi(x'_T) - G\hat{x}_T, G\hat{x}_T) \\ &= l_0 E \int_0^T [F(s, u_s) - F(s, u'_s), \hat{u}_s] ds + l_0 E \int_0^T [H(s, u_s) - H(s, u'_s), \hat{u}_s] d\langle M, M \rangle_s \\ &\quad - (1 - l_0)E \int_0^T [c_2 |G\hat{x}_s|^2 + c'_2 |G^*\hat{y}_s|^2] ds - (1 - l_0)E \int_0^T [c'_2 |G^*\hat{z}_s|^2] d\langle M, M \rangle_s \\ &\quad + \delta E \int_0^T [c_2(G\hat{x}_s, G\hat{X}_s) + c'_2(G^*\hat{y}_s, G^*\hat{Y}_s) + (\hat{x}_s, -G^*\hat{f}_s) + (G^*\hat{y}_s, \hat{b}_s)] ds \\ &\quad + \delta E \int_0^T [c'_2(G^*\hat{Z}_s, G^*\hat{z}_s) + (\hat{z}_s, G\hat{\sigma}_s)] d\langle M, M \rangle_s. \end{aligned}$$

By Assumptions 1 and 2, since $c_2 > 0$ and $c_3 > 0$, we obtain

$$\begin{aligned} & (l_0 c_3 + (1 - l_0))E|G\hat{x}_T|^2 + c_2 E \int_0^T |G\hat{x}_s|^2 ds \\ &\quad + c'_2 E \int_0^T |G^*\hat{y}_s|^2 ds + l_0 c'_2 E \int_0^T |G^*\hat{z}_s|^2 ds \\ &\quad + l_0 c'_2 E \int_0^T |G^*\hat{y}_s|^2 d\langle M, M \rangle_s + l_0 c_2 E \int_0^T |G\hat{x}_s|^2 d\langle M, M \rangle_s \\ &\quad + c'_2 E \int_0^T |G^*\hat{z}_s|^2 d\langle M, M \rangle_s \\ &\leq \delta c_4 E \int_0^T (|\hat{U}_s|^2 + |\hat{u}_s|^2) ds + \delta c_4 E \int_0^T (|\hat{U}_s|^2 + |\hat{u}_s|^2) d\langle M, M \rangle_s \\ &\quad + \delta c_4 E |\hat{X}_T|^2 + \delta c_4 E |\hat{x}_T|^2. \end{aligned}$$

For the difference of the solutions $(\hat{y}, \hat{z}) = (y - y', z - z')$, we apply the usual technique to the BSDE and derive that

$$\begin{aligned} & E \int_0^T |\hat{y}_s|^2 ds + E \int_0^T |\hat{z}_s|^2 d\langle M, M \rangle_s \\ & \leq c_4 \delta E \int_0^T |\hat{U}_s|^2 ds + c_4 \delta E |\hat{X}_T|^2 + c_4 E \int_0^T |\hat{x}_s|^2 ds + c_4 E |\hat{x}_T|^2, \\ & E \int_0^T |\hat{y}_s|^2 ds + E \int_0^T |\hat{z}_s|^2 d\langle M, M \rangle_s \\ & \leq c_4 \delta E \int_0^T |\hat{U}_s|^2 ds + c_4 \delta E |\hat{X}_T|^2 + c_4 E \int_0^T |\hat{x}_s|^2 ds + c_4 E |\hat{x}_T|^2. \end{aligned}$$

Similarly, for the difference of the solution $\hat{x} = x - x'$, we apply the usual technique to the forward part:

$$\begin{aligned} \sup_{0 \leq s \leq T} |\hat{x}_s|^2 & \leq c_4 E \int_0^T (|\hat{y}_s|^2 + |\hat{z}_s|^2) ds + c_4 E \int_0^T |\hat{u}_s|^2 d\langle M, M \rangle_s \\ & \quad + c_4 \delta E \int_0^T |\hat{U}_s|^2 ds + c_4 \delta E \int_0^T |\hat{U}_s|^2 d\langle M, M \rangle_s, \\ E \int_0^T |\hat{x}_s|^2 ds & \leq c_4 T E \int_0^T (|\hat{y}_s|^2 + |\hat{z}_s|^2) ds + c_4 T E \int_0^T |\hat{u}_s|^2 d\langle M, M \rangle_s \\ & \quad + c_4 \delta T E \int_0^T |\hat{U}_s|^2 ds + c_4 \delta T E \int_0^T |\hat{U}_s|^2 d\langle M, M \rangle_s. \end{aligned}$$

Here the constant c_4 depends on the Lipschitz constants, and so do G, c_2, c'_2 , and T . Combing the above estimates, we get

$$\begin{aligned} & E \int_0^T |\hat{u}_s|^2 ds + E \int_0^T |\hat{u}_s|^2 d\langle M, M \rangle_s + E |\hat{x}_T|^2 \\ & \leq c_5 \delta \left\{ E \int_0^T |\hat{U}_s|^2 ds + E \int_0^T |\hat{U}_s|^2 d\langle M, M \rangle_s + E |\hat{X}_T|^2 \right\}, \end{aligned}$$

where the constant c_5 depends on c_4, G, c_2 , and c_3 . If we choose $\delta_0 = \frac{1}{2c_5}$, then clearly for each fixed $\delta \in [0, \delta_0]$, the mapping $I_{\alpha+\delta}$ is a contraction in the sense that

$$\begin{aligned} & E \int_0^T |\hat{u}_s|^2 ds + E \int_0^T |\hat{u}_s|^2 d\langle M, M \rangle_s + E |\hat{x}_T|^2 \\ & \leq \frac{1}{2} \left\{ E \int_0^T |\hat{U}_s|^2 ds + E \int_0^T |\hat{U}_s|^2 d\langle M, M \rangle_s + E |\hat{X}_T|^2 \right\}. \end{aligned}$$

This indicates that this mapping has a unique fixed point $(u^{l_0+\delta}) = (x^{l_0+\delta}, y^{l_0+\delta}, z^{l_0+\delta})$, which is a solution of FBSDE (3) for $l = l_0 + \delta$. The proof is complete. \square

Proof of existence From Lemma 2 we immediately see that, when $\lambda = 1$ in FBSDE (5), FBSDEs (3) for $l = 0$ (that is, FBSDE (4)) has a unique solution. It then follows from Lemma 3 that there exists a positive constant δ_0 depending on Lipschitz constants, c_2, c'_2, c_3 , and T

such that, for each $\delta \in [0, \delta_0]$, FBSDEs (3) for $l = l_0 + \delta$ has a unique solution. We can repeat this process for N times with $1 \leq N\delta_0 < 1 + \delta_0$. It then follows that, in particular, FBSDE (3) for $l = 1$ with $\xi \equiv 0$ has a unique solution. The proof is complete. \square

5 Another existence and uniqueness theorem

Next, we give another existence and uniqueness theorem. First, we give another assumption.

Assumption 3

There exist constants c_2, c'_2, c_3 such that

$$\begin{aligned} [F(t, u^1) - F(t, u^2), u^1 - u^2] &\geq c_2 |G(x^1 - x^2)|^2 + c'_2 (|G^*(y^1 - y^2)|^2 + |G^*(z^1 - z^2)|^2), \\ [H(t, u^1) - H(t, u^2), u^1 - u^2] &\geq c_2 |G(x^1 - x^2)|^2 + c'_2 (|G^*(y^1 - y^2)|^2 + |G^*(z^1 - z^2)|^2), \\ \forall u^1 = (x^1, y^1, z^1), u^2 = (x^2, y^2, z^2) &\in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m \times d}, \mathbb{P}\text{-a.s.}, \text{ a.e. } t \in \mathbb{R}^+, \end{aligned}$$

and

$$(\Phi(x^1) - \Phi(x^2), G(x^1 - x^2)) \leq -c_3 |G(x^1 - x^2)|^2, \quad \forall x^1 \in \mathbb{R}^n, \forall x^2 \in \mathbb{R}^n,$$

where $c_2, c'_2,$ and c_3 are given positive constants.

Theorem 3 *Let Assumptions 1 and 3 hold. Then there exists a unique adapted solution (X, Y, Z) for FBSDE (1).*

Proof of uniqueness Using the same procedure as in the proof of uniqueness of Theorem 2, by Assumption 3 we get

$$\begin{aligned} &E(\hat{Y}_T, G\hat{X}_T) \\ &= E(\Phi(X_T^1) - \Phi(X_T^2), G(X_T^1 - X_T^2)) \\ &= E \int_0^T [F(s, U^1) - F(s, U^2), U^1 - U^2] ds \\ &\quad + E \int_0^T [H(s, U^1) - H(s, U^2), U^1 - U^2] d\langle M, M \rangle_s \\ &\geq c_2 E \int_0^T |G\hat{X}|^2 ds + c'_2 E \int_0^T (|G^*\hat{Y}|^2 + |G^*\hat{Z}|^2) ds \\ &\quad + c_2 E \int_0^T |G\hat{X}|^2 d\langle M, M \rangle_s + c'_2 E \int_0^T (|G^*\hat{Y}|^2 + |G^*\hat{Z}|^2) d\langle M, M \rangle_s. \end{aligned}$$

Thus, we deduce that

$$\begin{aligned} &c_2 \left(E \int_0^T |G\hat{X}|^2 ds + E \int_0^T |G\hat{X}|^2 d\langle M, M \rangle_s \right) \\ &\quad + c'_2 \left[E \int_0^T (|G^*\hat{Y}|^2 + |G^*\hat{Z}|^2) ds + E \int_0^T (|G^*\hat{Y}|^2 + |G^*\hat{Z}|^2) d\langle M, M \rangle_s \right] + c_3 |G\hat{X}|^2 \\ &\leq 0. \end{aligned}$$

Since c_2, c'_2 , and c_3 are given positive constants, we have $|G\hat{X}|^2 \equiv 0, |G^*\hat{Y}|^2 \equiv 0, |G^*\hat{Z}|^2 \equiv 0$. So $\hat{X}_s^1 \equiv \hat{X}_s^2, \hat{Y}_s^1 \equiv \hat{Y}_s^2, \hat{Z}_s^1 \equiv \hat{Z}_s^2, d\langle M, M \rangle_t \times P$ -a.s. □

We now consider the following family of FBSDEs parameterized by $l \in [0, 1]$:

$$\begin{cases} dX_t^l = [(1-l)c'_2(G^*Y_t^l) + lb(t, U_t^l) + \phi_t] dt \\ \quad + [l\sigma(t, U_t^l) + (1-l)c'_2(G^*Z_t^l) + \psi_t] dM_t, \\ -dY_t^l = [-(1-l)c_2GX_t^l + lf(t, U_t^l) + \gamma_t] dt - Z_t^l dM_t, \\ X_0^l = x, \quad Y_T^l = l\Phi(X_T^l) + (1-l)GX_T^l + \xi, \end{cases} \tag{8}$$

where ϕ, ψ , and γ are given processes in $M^2(0, T)$ with values in $R^n, R^{n \times d}$, and R^m , respectively. Clearly, when $l = 1$ and $\xi \equiv 0$, the existence of a solution of FBSDE (8) implies that of FBSDE (1). When $l = 0$ in FBSDEs (8), it is

$$\begin{cases} dX_t^0 = [c'_2G^*Y_t^0 + \phi_t] dt + [c'_2G^*Z_t^0 + \psi_t] dM_t, \\ -dY_t^0 = [-c_2GX_t^0 + \gamma_t] dt - Z_t^0 dM_t, \\ X_0^0 = x, \quad Y_T^0 = GX_T^0 + \xi. \end{cases} \tag{9}$$

For proving the existence part of the theorem, we first need the following two lemmas.

Lemma 4 *The following equation has a unique solution:*

$$\begin{cases} dX_t = [c'_2G^*Y_t + \phi_t] dt + [c'_2G^*Z_t + \psi_t] dM_t, \\ -dY_t = [-c_2GX_t + \gamma_t] dt - Z_t dM_t, \\ X_0 = x, \quad Y_T = \lambda GX_T + \xi, \end{cases} \tag{10}$$

where λ is a nonnegative constant.

Lemma 5 *Let Assumptions 1 and 3 hold. If for an $l_0 \in [0, 1)$, there exists a solution $(X^{l_0}, Y^{l_0}, Z^{l_0})$ of FBSDE (8), then there exists a positive constant δ_0 such that for each $\delta \in [0, \delta_0]$, there exists a solution $(X^{l_0+\delta}, Y^{l_0+\delta}, Z^{l_0+\delta})$ of FBSDE (8) for $l = l_0 + \delta$.*

Remark 1 The proofs of Lemma 4 and Lemma 5 are similar to those of Lemma 2 and Lemma 3.

Proof of existence From Lemma 4 we immediately see that when $\lambda = 1$ in FBSDE (10), FBSDE (8) for $l = 0$ (that is, FBSDE (9)) has a unique solution. It then follows from Lemma 5 that there exists a positive constant δ_0 depending on the Lipschitz constants, c_2, c'_2, c_3 , and T such that, for each $\delta \in [0, \delta_0]$, FBSDE (8) for $l = l_0 + \delta$ has a unique solution. We can repeat this process for N times with $1 \leq N\delta_0 < 1 + \delta_0$. It then follows that, in particular, FBSDEs (8) for $l = 1$ with $\xi \equiv 0$ has a unique solution.

The proof is complete. □

6 Application to stochastic optimal control problems

In this section, we give an application of the theorem to the stochastic maximum principle in optimal control problems through an example of the linear quadratic case. Consider a continuous-time finite-state Markov chain $m = \{m_t, t \in [0, T]\}$. We identify the states of

the process with the unit vectors e_i in \mathbb{R}^k , where k is the number of states of the chain. The rate matrix of m_t is denoted by \mathcal{A}_t .

All the stochastic processes are defined on the filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, where $\{\mathcal{F}_t\}$ is the completed natural filtration generated by the σ -fields $\sigma(\{m_s, s \leq t\}, F \in \mathcal{F}_T : \mathbb{P}(F) = 0)$, and $\mathcal{F} = \mathcal{F}_T$. As noted before, the martingale $M_t = m_t - m_0 - \int_0^t \mathcal{A}_s m_s ds$ is used to drive the controlled state process X_t . Next, we define the function $\Lambda : [0, T] \times \{e_i | i = 1, 2, \dots, k\} \rightarrow S^k$ by

$$\Lambda(s, e_i) := \text{diag}(\mathcal{A}_s e_i) - \text{diag}(e_i) \mathcal{A}_s^* - \mathcal{A}_s \text{diag}(e_i).$$

Note that

$$\begin{aligned} \langle M, M \rangle_t &= \int_0^t [\text{diag}(\mathcal{A}_s m_s) - \text{diag}(m_s) \mathcal{A}_s^* - \mathcal{A}_s \text{diag}(m_s)] ds \\ &= \int_0^t \Lambda(s, m_s) ds. \end{aligned}$$

Consider the one-dimensional linear quadratic optimal control problem

$$\begin{cases} dX_t = [AX_t + Bu_t] dt + [CX_{t-} + Du_t] dM_t, \\ X_0 = x, \end{cases} \tag{11}$$

and the cost function

$$J(u(\cdot)) = E \left[\frac{1}{2} \int_0^T (QX_t^2 + Ru_t^2) dt + \frac{1}{2} Gx_T^2 \right], \tag{12}$$

where the parameters A, B, Q, R, G are positive constant numbers, and C, D are $\mathbb{R}^{1 \times k}$ vectors.

Denote $\mathcal{U} = \{u : [0, T] \times \Omega \rightarrow \mathbb{R} \mid u \text{ is } \{\mathcal{F}_t\}\text{-predictable, } E \int_0^T |u(s)|^2 ds < +\infty\}$. The stochastic control problem is to find an optimal control $u^* \in \mathcal{U}$ such that

$$J(u^*) = \inf_{u \in \mathcal{U}} J(u).$$

In order to solve the problem, we define the Hamiltonian $\mathcal{H} : [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^{1 \times k} \times \{e_i | i = 1, 2, \dots, k\} \rightarrow \mathbb{R}$ by

$$\mathcal{H}(t, x, u, p, s, e_i) = \frac{1}{2} Qx^2 + \frac{1}{2} Ru^2 + (Ax + Bu)p + (Cx + Du)\Lambda(t, e_i)s^*.$$

The adjoint equation is given by the following BSDE:

$$\begin{cases} dp_t = -\mathcal{H}_x(t, X_t, u_t, p_t, s_t, m_t) dt + s_t dM_t \\ \quad = -(QX_t + Ap_t + C_t \Lambda(t, m_t) s_t^*) dt + s_t dM_t, \\ p_T = GX_T. \end{cases} \tag{13}$$

Then the sufficient stochastic maximum principle is stated as follows.

Theorem 4 *Given the assumptions on the parameters above, the coupled FBSDE*

$$\begin{cases} dX_t = [AX_t - BR^{-1}Bp_t - BR^{-1}D\Lambda(t, m_t)s_t^*] dt, \\ \quad + [CX_{t-} - DR^{-1}Bp_{t-} - DR^{-1}D\Lambda(t, m_t)s_t^*] dM_t, \\ dp_t = -(QX_t + Ap_t + C_t\Lambda(t, m_t)s_t^*) dt + s_t dM_t, \\ X_0 = x, \quad p_T = GX_T, \end{cases} \tag{14}$$

exists a unique solution $(\hat{X}_t, \hat{p}_t, \hat{s}_t)$. Moreover, the stochastic linear quadratic problem is solvable with the optimal controlled state process \hat{X}_t and the optimal control defined by

$$\hat{u}_t = -R^{-1}(B\hat{p}_{t-} + D\Lambda(t, m_t)\hat{s}_t^*). \tag{15}$$

Proof First, it is easy to check that FBSDE (14) satisfies Assumptions 1 and 3. So, according to Theorem 3, there exists a unique adapted solution $(\hat{X}_t, \hat{p}_t, \hat{s}_t) \in \mathcal{M}^2(0, T; \mathbb{R} \times \mathbb{R} \times \mathbb{R}^{1 \times k})$. So, according to (15), $\hat{u} \in \mathcal{U}$, and \hat{X} is the corresponding state process. Next, for any fixed $u \in \mathcal{U}$ with the corresponding state process X_t , consider

$$J(u) - J(\hat{u}) = E \left[\frac{1}{2} \int_0^T (QX_t^2 + Ru_t^2 - Q\hat{X}_t^2 - R\hat{u}_t^2) dt + \frac{1}{2} (GX_T^2 - G\hat{X}_T^2) \right].$$

Since $G > 0$, we have

$$\begin{aligned} & E \left[\frac{1}{2} (GX_T^2 - G\hat{X}_T^2) \right] \\ & \geq E[(X_T - \hat{X}_T)G\hat{X}_T] \\ & = E[(X_T - \hat{X}_T)\hat{p}_T] \\ & = E \left[\int_0^T (X_{t-} - \hat{X}_{t-}) d\hat{p}_t + \int_0^T \hat{p}_{t-} d(X_t - \hat{X}_t) + \sum_{0 < t \leq T} \Delta(X_t - \hat{X}_t) \Delta p_t \right] \\ & = E \left[\int_0^T (X_{t-} - \hat{X}_{t-}) (-\mathcal{H}_x(t, \hat{X}_{t-}, \hat{u}_t, \hat{p}_{t-}, \hat{s}_t, m_t)) dt + \int_0^T (X_{t-} - \hat{X}_{t-}) \hat{s}_t dM_t \right. \\ & \quad \left. + \int_0^T \hat{p}_{t-} [A(X_{t-} - \hat{X}_{t-}) + B(u_t - \hat{u}_t)] dt + \int_0^T \hat{p}_{t-} [C(X_{t-} - \hat{X}_{t-}) + D(u_t - \hat{u}_t)] dM_t \right. \\ & \quad \left. + \sum_{0 < t \leq T} [C(X_{t-} - \hat{X}_{t-}) + D(u_t - \hat{u}_t)] \Delta M_t s_t \Delta M_t \right]. \end{aligned}$$

Notice that

$$\begin{aligned} & \sum_{0 < t \leq T} [C(X_{t-} - \hat{X}_{t-}) + D(u_t - \hat{u}_t)] \Delta M_t s_t \Delta M_t \\ & = \int_0^T [C(X_{t-} - \hat{X}_{t-}) + D(u_t - \hat{u}_t)] d[M, M]_t s_t^* \\ & = L_t + \int_0^T [C(X_{t-} - \hat{X}_{t-}) + D(u_t - \hat{u}_t)] d\langle M, M \rangle_t s_t^* \\ & = L_t + \int_0^T [C(X_{t-} - \hat{X}_{t-}) + D(u_t - \hat{u}_t)] \Lambda(t, m_t) s_t^* dt \end{aligned}$$

for some local martingale L (see [5]). So

$$\begin{aligned} & E \left[\frac{1}{2} (GX_T^2 - G\hat{X}_T^2) \right] \\ & \geq E \left[\int_0^T (X_{t-} - \hat{X}_{t-}) (-\mathcal{H}_x(t, \hat{X}_{t-}, \hat{u}_t, \hat{p}_{t-}, \hat{s}_t, m_t)) dt \right. \\ & \quad + \int_0^T \hat{p}_{t-} [A(X_{t-} - \hat{X}_{t-}) + B(u_t - \hat{u}_t)] dt \\ & \quad \left. + \int_0^T [C(X_{t-} - \hat{X}_{t-}) + D(u_t - \hat{u}_t)] \Lambda(t, m_t) \hat{s}_t^* dt \right]. \end{aligned}$$

By the definition of \mathcal{H} we have

$$\begin{aligned} & E \left[\frac{1}{2} \int_0^T (QX_t^2 + Ru_t^2 - Q\hat{X}_t^2 - R\hat{u}_t^2) dt \right] \\ & = E \left[\frac{1}{2} \int_0^T (QX_{t-}^2 + Ru_t^2 - Q\hat{X}_{t-}^2 - R\hat{u}_t^2) dt \right] \\ & = E \left[\int_0^T \left\{ \mathcal{H}(t, X_{t-}, u_t, \hat{p}_{t-}, \hat{s}_t, m_t) - \mathcal{H}(t, \hat{X}_{t-}, \hat{u}_t, \hat{p}_{t-}, \hat{s}_t, m_t) \right. \right. \\ & \quad \left. \left. - [A(X_{t-} - \hat{X}_{t-}) + B(u_t - \hat{u}_t)] \hat{p}_{t-} - [C(X_{t-} - \hat{X}_{t-}) + D(u_t - \hat{u}_t)] \Lambda(t, m_t) \hat{s}_t^* \right\} dt \right]. \end{aligned}$$

Adding the two equations, we have

$$\begin{aligned} J(u) - J(\hat{u}) & \geq E \left[\int_0^T (\mathcal{H}(t, X_{t-}, u_t, \hat{p}_{t-}, \hat{s}_t, m_t) - \mathcal{H}(t, \hat{X}_{t-}, \hat{u}_t, \hat{p}_{t-}, \hat{s}_t, m_t) \right. \\ & \quad \left. - (X_{t-} - \hat{X}_{t-}) \mathcal{H}_x(t, \hat{X}_{t-}, \hat{u}_t, \hat{p}_{t-}, \hat{s}_t, m_t)) dt \right] \\ & = E \left\{ \int_0^T [\mathcal{H}(t, X_{t-}, u_t, \hat{p}_{t-}, \hat{s}_t, m_t) - \mathcal{H}(t, \hat{X}_{t-}, \hat{u}_t, \hat{p}_{t-}, \hat{s}_t, m_t) \right. \\ & \quad \left. - (X_{t-} - \hat{X}_{t-}) (Q\hat{X}_{t-} + A\hat{p}_{t-} + C\Lambda(t, m_t)\hat{s}_t)] dt \right\} \\ & = E \left[\int_0^T \left(\frac{1}{2} QX_{t-}^2 - \frac{1}{2} Q\hat{X}_{t-}^2 - (X_{t-} - \hat{X}_{t-}) Q\hat{X}_{t-} \right. \right. \\ & \quad \left. \left. + \frac{1}{2} Ru_t^2 + Bu_t\hat{p}_{t-} + Du_t\Lambda(t, m_t)\hat{s}_t^* \right. \right. \\ & \quad \left. \left. - \frac{1}{2} R\hat{u}_t^2 - B\hat{u}_t\hat{p}_{t-} - D\hat{u}_t\Lambda(t, m_t)\hat{s}_t^* \right) dt \right] \\ & \geq 0. \end{aligned}$$

By the convexity of $\frac{1}{2}Qx^2$ and the fact that $\hat{u}_t = -R^{-1}(B\hat{p}_{t-} + D\Lambda(t, m_t)\hat{s}_t^*)$ we get the minimum point of the function $F(u) = \frac{1}{2}Ru_t^2 + Bu_t\hat{p}_{t-} + Du_t\Lambda(t, m_t)\hat{s}_t^*$. Therefore, we conclude that

$$J(u) - J(\hat{u}) \geq 0,$$

which proves that \hat{u} is optimal. □

Remark 2 Differently from the Brownian motion case, the optimal control and the Hamiltonian additionally depend on the state of the Markov chain. This is because Markov chains are discontinuous stochastic processes with finite variation and the quadratic processes of the martingales generated by Markov chains are quite different from the continuous-time diffusions.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

All authors jointly worked on the results, and they read and approved the final manuscript.

Author details

¹Qilu Institute of Finance, Shandong University, Jinan, 250100, P.R. China. ²School of Mathematical Sciences, Shandong Normal University, Jinan, 250014, P.R. China.

Acknowledgements

This research are supported by School of Mathematical Sciences project (shxzhxxm201503) of Shandong Normal University, National Natural Science Foundation project (11301309) of China, National Natural Science Foundation project (11201271) of China, and Shandong Provincial Scientific Research Foundation for Excellent Young Scientists (BS2013SF003). We would like to thank the reviewers, whose careful reading, help suggestions, and valuable comments helped us to improve the manuscript.

Received: 11 May 2015 Accepted: 10 May 2016 Published online: 18 May 2016

References

- Pardoux, E, Peng, S: Adapted solution of a backward stochastic differential equations. *Syst. Control Lett.* **14**, 55-61 (1990)
- Tang, S, Li, X: Necessary conditions for optimal control of stochastic systems with random jumps. *SIAM J. Control Optim.* **32**(5), 1447-1475 (1994)
- Nualart, D, Schoutens, W: Backward stochastic differential equations and Feynman-Kac formula for Levy processes, with applications in finance. *Bernoulli* **7**(5), 761-776 (2001)
- Cohen, S, Elliott, R: Solutions of backward stochastic differential equations on Markov chains. *Commun. Stoch. Anal.* **2**(2), 251-262 (2008)
- Cohen, S, Elliott, R: Comparisons for backward stochastic differential equations on Markov chains and related no-arbitrage conditions. *Ann. Appl. Probab.* **20**(1), 267-311 (2010)
- Cohen, S, Hu, Y: Ergodic BSDEs driven by Markov chains. *SIAM J. Control Optim.* **51**(5), 4138-4168 (2013)
- Ma, J, Yong, J: *Forward-Backward Stochastic Differential Equations and Their Applications*. Lecture Notes in Mathematics. Springer, Berlin (1999)
- Buchdahn, R, Hu, Y: Hedging contingent claims for a large investor in an incomplete market. *Adv. Appl. Probab.* **30**(1), 239-255 (1998)
- Hu, Y: Potential kernels associated with a filtration and forward-backward SDEs. *Potential Anal.* **10**(2), 103-118 (1999)
- Antonelli, F: Backward-forward stochastic differential equations. *Ann. Appl. Probab.* **3**, 777-793 (1993)
- Pardoux, E, Tang, S: Forward-backward stochastic differential equations and quasilinear parabolic PDEs. *Probab. Theory Relat. Fields* **114**(2), 123-150 (1999)
- Ma, J, Protter, P, Yong, J: Solving forward-backward stochastic differential equations explicitly: a four step scheme. *Probab. Theory Relat. Fields* **98**(3), 339-359 (1994)
- Hu, Y, Peng, S: Solutions of forward-backward stochastic differential equations. *Probab. Theory Relat. Fields* **103**(2), 273-283 (1995)
- Peng, S, Wu, Z: Fully coupled forward-backward stochastic differential equations. *SIAM J. Control Optim.* **37**(3), 825-843 (1999)
- Yong, J: Forward-backward stochastic differential equations with mixed initial-terminal conditions. *Trans. Am. Math. Soc.* **362**(2), 1047-1096 (2010)
- Ma, J, Wu, Z, Zhang, D, Zhang, J: On wellposedness of forward-backward SDEs: a unified approach. *Ann. Appl. Probab.* **25**(4), 2168-2214 (2015)
- Briand, P, Hu, Y: Probabilistic approach to singular perturbations of semilinear and quasilinear parabolic PDEs. *Nonlinear Anal., Theory Methods Appl.* **35**(7), 815-831 (1999)
- Zhang, X, Elliott, R, Siu, T: A stochastic maximum principle for a Markov regime-switching jump-diffusion model and its application to finance. *SIAM J. Control Optim.* **50**, 964-990 (2012)
- Crepey, S, Matoussi, A: Reflected and doubly reflected BSDEs with jumps: a priori estimates and comparison. *Ann. Appl. Probab.* **18**, 2041-2069 (2008)
- Crepey, S: *About the Pricing Equations in Finance*. Paris-Princeton Lectures on Mathematical Finance. Springer, Berlin (2010)
- Tao, R, Wu, Z, Zhang, Q: BSDEs with regime switching: weak convergence and applications. *J. Math. Anal. Appl.* **407**, 97-111 (2013)
- Tao, R, Wu, Z: Maximum principle for optimal control problems of forward-backward regime-switching system and applications. *Syst. Control Lett.* **61**, 911-917 (2012)
- Tao, R, Wu, Z, Zhang, Q: Optimal switching under a regime-switching model with two-time-scale Markov chains. *Multiscale Model. Simul.* **13**, 99-131 (2015)
- Elliott, R, Aggoun, L, Moore, J: *Hidden Markov Models: Estimation and Control*. Springer, Berlin (1994)