Advances in Difference Equations
a SpringerOpen Journal

Backward stochastic differential equations with Markov chains and related asymptotic properties

Huaibin Tang[1] and Zhen Wu[2*]

*Correspondence:
wuzhen@sdu.edu.cn
[2]School of Mathematics, Shandong
University, Jinan, 250100, P.R. China
Full list of author information is
available at the end of the article

Abstract

This paper is concerned with the solvability of a new kind of backward stochastic differential equations whose generator f is affected by a finite-state Markov chain. We also present the asymptotic property of backward stochastic differential equations involving a singularly perturbed Markov chain with weak and strong interactions and then apply this result to the homogenization of a system of semilinear parabolic partial differential equations.

MSC: 60H10; 35R60

Keywords: BSDE; Markov chain; weak convergence; homogenization

1 Introduction

Pioneered by the works of Pardoux and Peng [1] and Duffie and Epstein [2] about the nonlinear backward stochastic differential equations (BSDEs) driven by a Brownian motion, BSDEs have been extensively studied because of their deep connections with mathematical finance, stochastic control, and partial differential equations (PDEs), such as [3, 4], *etc.* Simultaneously, many researchers have been devoted to more general cases in the framework of continuous time diffusions or jump diffusions, such as BSDE driven by a Lévy process (see [5, 6]), BSDE with respect to both a Brownian motion and a Poisson random measure (see [7, 8]). For BSDEs driven by Markov chains, [9] considered the following kind of BSDE:

$$Y_t = \xi + \int_t^T f(s, Y_{s-}, Z_s) \, ds - \int_t^T Z_s \, dM_s,$$

where M_s is a martingale related to a continuous time, finite state Markov chain. After that, [10–12] further studied such kind of BSDE.

In this paper, we study the BSDE whose generator f is directly affected by a continuous time, finite-state Markov chain. Consider the following BSDE in the probability space $(\Omega, \mathcal{F}, P)$:

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s, \alpha_s) \, ds - \int_t^T Z_s \, dB_s. \tag{1}$$

Here $B = \{B_t; 0 \le t \le T\}$ is a d-dimensional Brownian motion, and $\alpha = \{\alpha_t; 0 \le t \le T\}$ is a continuous-time finite-state Markov chain independent of B with the state space

$\mathcal{M} = \{1, 2, \ldots, m\}$ and the generator $Q = (q_{ij})_{m \times m}$. The generator f of BSDE (1) can be considered as disturbed by random environment and takes a set of discrete values described by the Markov chain α. One major objective of this paper is to find a pair of $\{\mathcal{F}_t^B \vee \mathcal{F}_{t,T}^\alpha\}_{0 \leq t \leq T}$-adapted processes $(Y, Z) = \{(Y_t, Z_t); 0 \leq t \leq T\}$ as its solution. Here, for any process $\{\eta_t; 0 \leq t \leq T\}$, $\mathcal{F}_t^\eta = \mathcal{F}_{0,t}^\eta$ and $\mathcal{F}_{t,T}^\eta = \sigma\{\eta_r; t \leq r \leq T\}$.

When studying the solvability for BSDE (1) with the Markov chain α, since $\{\mathcal{F}_t^B \vee \mathcal{F}_{t,T}^\alpha\}_{0 \leq t \leq T}$ is not a filtration, the conventional approach [1, 3] to use the Itô representation theorem and contraction mapping cannot be applied straightforwardly. To tackle this problem, inspired by the method dealing with backward doubly stochastic differential equations [13], we construct an enlarged filtration generated by the Brownian motion B and the Markov chain α and propose a corresponding extended Itô representation theorem. By virtue of this theoretical tool, we prove the existence of the solution adapted to the enlarged filtration and then verify the desired adaptability.

When various factors in random environment are considered, the underlying Markov chain inevitably has a large state space, and the corresponding BSDE (1) becomes increasingly complicated. In many practical situations, the states of the underlying Markov chain can be divided into several classes such that the Markov chain fluctuates rapidly among different states in the same class and jumps less frequently among different classes. One hierarchical approach to reduce the complexity, explored by Yin and Zhang [14], introduces a small parameter $\varepsilon > 0$ and leads to a singularly perturbed Markov chain α^ε. The asymptotic property of α^ε can be studied by a Markov chain with a considerably smaller state space. There exists an extensive literature on singularly perturbed Markov chain and its applications. We suggest to the interested reader that they should consult [14, 15] to get a rather complete view on this research field.

In this paper, we study the following BSDE with the singularly perturbed Markov chain α^ε:

$$Y_t^\varepsilon = \xi + \int_t^T f\left(s, Y_s^\varepsilon, \alpha_s^\varepsilon\right) ds - \int_t^T Z_s^\varepsilon \, dB_s.$$

Combining with the asymptotic property of the singularly perturbed Markov chain α^ε, we show that the solution sequence $\{(Y_t^\varepsilon, \int_0^t Z_s^\varepsilon \, dB_s); 0 \leq t \leq T\}$ converges weakly with the limit formed by the solution of a simpler BSDE which involves the limit aggregated Markov chain. As far as we know, such asymptotic property of BSDE with a singularly perturbed Markov chain is the first study in BSDEs theory. As a straightforward application of our asymptotic result, we show the homogenization of a system of semi-linear parabolic PDEs with the singularly perturbed Markov chain α^ε.

This paper is organized as follows. In Section 2, we study the solvability of BSDE (1) with a Markov chain. Section 3 gives the asymptotic property of BSDE with a singularly perturbed Markov chain. Two interpretation examples are also presented. In Section 4, we apply the asymptotic result to the homogenization of a system of semi-linear parabolic PDEs with the singularly perturbed Markov chain. Finally, we give some concluding remarks in Section 5.

Throughout this paper, we introduce the following notations: $L^p(\mathcal{F}_t; R^n)$ is the space of R^n-valued $\mathcal{F}_t$-adapted random variable ξ satisfying $E(|\xi|^p) < \infty$; $M_{\mathcal{F}_t}^2(0, T; R^n)$ denotes the space of R^n-valued $\mathcal{F}_t$-adapted stochastic processes $\varphi = \{\varphi_t; t \in [0, T]\}$ satisfying

$E \int_0^T |\varphi_t|^2 \, dt < \infty$; $S^2_{\mathcal{F}_t}(0, T; R^n)$ is the space of R^n-valued $\mathcal{F}_t$-adapted continuous stochastic processes $\varphi = \{\varphi_t; t \in [0, T]\}$ satisfying $E(\sup_{0 \le t \le T} |\varphi_t|^2) < \infty$.

2 BSDEs with Markov chains

In this section, we study the solvability of BSDE (1) with the Markov chain α:

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s, \alpha_s) \, ds - \int_t^T Z_s \, dB_s.$$

For each $t \in [0, T]$, define $\mathcal{F}_t = \mathcal{F}_t^B \vee \mathcal{F}_{t,T}^\alpha$. We make the following assumptions for the coefficients of BSDE (1).

(H1) (i) $\xi \in L^2(\mathcal{F}_T; R^k)$; (ii) $f : \Omega \times [0, T] \times R^k \times R^{k \times d} \times \mathcal{M} \to R^k$ satisfies that
$\forall (y, z) \in R^k \times R^{k \times d}$, $i \in \mathcal{M}$, $f(\cdot, y, z, i) \in M^2_{\mathcal{F}_t^B}(0, T; R^k)$. Also, there exists a constant
$\mu > 0$ such that $\forall (\omega, t) \in \Omega \times [0, T]$, $i \in \mathcal{M}$, $(y_1, z_1), (y_2, z_2) \in R^k \times R^{k \times d}$,

$$\left| f(t, y_1, z_1, i) - f(t, y_2, z_2, i) \right| \le \mu \left(|y_1 - y_2| + |z_1 - z_2| \right).$$

The main purpose of this section is to prove the following solvability of BSDE (1).

Theorem 2.1 *Under hypothesis* (H1), *there exists a unique solution pair* $(Y, Z) \in S^2_{\mathcal{F}_t}(0, T; R^k) \times M^2_{\mathcal{F}_t}(0, T; R^{k \times d})$ *for BSDE* (1).

To obtain the solvability of BSDE (1), we borrow the idea from the method handling the backward doubly stochastic differential equations (see [13]). Firstly, we construct an enlarged filtration $(\mathcal{G}_t)_{0 \le t \le T}$: $\mathcal{G}_t \triangleq \mathcal{F}_t^B \vee \mathcal{F}_T^\alpha \vee \mathcal{N}$, $\forall t \in [0, T]$, which is generated by the Brownian motion B and the Markov chain α. Then we propose a related extended Itô representation theorem. After that, for the special case that f is independent of y and z, we show the existence of the solution of BSDE (1) and verify the adaptability of the solution with respect to $\{\mathcal{F}_t\}_{0 \le t \le T}$. Finally, the solvability of BSDE (1) for the general case is proved.

Proposition 2.1 *Let* $N \in L^2(\mathcal{G}_T; R^k)$. *Then there exists a unique stochastic process* $Z = \{Z_t; 0 \le t \le T\} \in M^2_{\mathcal{G}_t}(0, T; R^{k \times d})$ *such that*

$$N = E(N | \mathcal{F}_T^\alpha) + \int_0^T Z_t \, dB_t. \tag{2}$$

This result can be derived from the martingale representation theorem for initially enlarged filtration (see Theorem 4.2 in [16]). For the readers' convenience, we also give a more specific proof in the Appendix.

Similarly, for all $t \in [0, T]$, we can introduce the filtration $\{\mathcal{N}_s\}_{t \le s \le T}$ defined by $\mathcal{N}_s = \mathcal{F}_s^B \vee \mathcal{F}_{t,T}^\alpha$. And we have the following propositions.

Proposition 2.2 *Let* $N \in L^2(\mathcal{N}_T; R^k)$. *Then there exists a unique stochastic process* $Z = \{Z_s; t \le s \le T\} \in M^2_{\mathcal{N}_s}(t, T; R^{k \times d})$ *such that*

$$N = E(N | \mathcal{N}_t) + \int_t^T Z_s \, dB_s.$$

Proposition 2.3 *Under hypothesis* (H1), *the following BSDE*

$$Y_t = \xi + \int_t^T f(s, \alpha_s) \, ds - \int_t^T Z_s \, dB_s, \quad 0 \leq t \leq T, \tag{3}$$

has a solution pair $(Y, Z) \in S^2_{\mathcal{F}_t}(0, T; R^k) \times M^2_{\mathcal{F}_t}(0, T; R^{k \times d})$.

Proof Applying Hölder's inequality, we have

$$E \left(\int_0^T f(s, \alpha_s) \, ds \right)^2 \leq T E \int_0^T |f(s, \alpha_s)|^2 \, ds \leq T \sum_{i \in \mathcal{M}} E \int_0^T |f(s, i)|^2 \, ds.$$

Combining the assumptions of ξ and f in hypothesis (H1), we can conclude that $\xi + \int_0^T f(s, \alpha_s) \, ds \in L^2(\mathcal{G}_T; R^k)$. $\forall t \in [0, T]$, define $N_t = E(\xi + \int_0^T f(s, \alpha_s) \, ds | \mathcal{G}_t)$. Applying Proposition 2.1 yields that there exists a unique stochastic process $Z = \{Z_s; 0 \leq s \leq t\} \in M^2_{\mathcal{G}_s}(0, t; R^{k \times d})$ such that

$$N_t = E \left(\xi + \int_0^T f(s, \alpha_s) \, ds \Big| \mathcal{F}_T^\alpha \right) + \int_0^t Z_s \, dB_s.$$

$\forall t \in [0, T]$, letting

$$Y_t = N_t - \int_0^t f(s, \alpha_s) \, ds,$$

it can be verified that $Y = \{Y_t; 0 \leq t \leq T\} \in M^2_{\mathcal{G}_t}(0, T; R^k)$ and (Y, Z) satisfies BSDE (3). Besides, combining the Burkholder-Davis-Gundy inequality and the form of BSDE (3), we can conclude that Y is continuous and satisfies $E(\sup_{0 \leq t \leq T} |Y_t|^2) < \infty$.

The remaining work is to show that the $\{\mathcal{G}_t\}_{0 \leq t \leq T}$-adapted process (Y, Z) is also adapted to $\{\mathcal{F}_t\}_{0 \leq t \leq T}$. Recall that $\{\mathcal{F}_t\}_{0 \leq t \leq T}$ is defined by $\mathcal{F}_t = \mathcal{F}_t^B \vee \mathcal{F}_{t,T}^\alpha$. We firstly show that $\forall t \in [0, T]$, Y_t is $\mathcal{F}_t^B \vee \mathcal{F}_{t,T}^\alpha$-measurable. Combining the definitions of N_t and Y_t yields $Y_t = E(\xi + \int_t^T f(s, \alpha_s) \, ds | \mathcal{G}_t)$. For ease of notation, we denote $\vartheta \triangleq \xi + \int_t^T f(s, \alpha_s) \, ds$. Obviously, ϑ is $\mathcal{F}_T^B \vee \mathcal{F}_{t,T}^\alpha$-measurable. Let $\{t_k^*\}_{k \geq 0}$ be a dense subset of $[t, T]$ with $t_0^* = t$. For each integer $l \geq 0$, denote $\alpha_{t_0^*}^{t_l^*} = (\alpha_{t_0^*}, \alpha_{t_1^*}, \ldots, \alpha_{t_l^*})$, $\mathcal{G}_l^* = \mathcal{F}_T^B \vee \sigma\{\alpha_{t_0^*}^{t_l^*}\}$, and $\mathcal{M}^{l+1} = \underbrace{\mathcal{M} \times \cdots \times \mathcal{M}}_{l+1}$.

It is clear that $\mathcal{G}_l^* \subset \mathcal{G}_{l+1}^*$ and $\sigma\{\bigcup_{l \geq 0} \mathcal{G}_l^*\} = \mathcal{F}_T^B \vee \mathcal{F}_{t,T}^\alpha$. Applying the Doob-Dynkin lemma (Lemma A.1) and the Itô representation theorem to $E(\vartheta | \mathcal{G}_l^*)$, we have

$$E(\vartheta | \mathcal{G}_l^*) = \sum_{i_0^l \in \mathcal{M}^{l+1}} I_{\{\alpha_{t_0^*}^{t_l^*} = i_0^l\}} \left(\vartheta_l(i_0^l) + \int_0^T Z_{\vartheta_l}(s, i_0^l) \, dB_s \right)$$

with $\vartheta_l(i_0^l) \in R^k$ and $\{Z_{\vartheta_l}(s, i_0^l); 0 \leq s \leq T\} \in \mathcal{M}^2_{\mathcal{F}_s^B}(0, T; R^{k \times d})$. Therefore

$$\begin{aligned}
E(E(\vartheta | \mathcal{G}_l^*) | \mathcal{G}_t) &= E(E(\vartheta | \mathcal{G}_l^*) | \mathcal{F}_t^B \vee \mathcal{F}_T^\alpha) \\
&= \sum_{i_0^l \in \mathcal{M}^{l+1}} I_{\{\alpha_{t_0^*}^{t_l^*} = i_0^l\}} \left(\vartheta_l(i_0^l) + \int_0^t Z_{\vartheta_l}(s, i_0^l) \, dB_s \right) \\
&\triangleq \vartheta_l(\alpha_{t_0^*}^{t_l^*}) + \int_0^t Z_{\vartheta_l}(s, \alpha_{t_0^*}^{t_l^*}) \, dB_s.
\end{aligned} \tag{4}$$

Since $[t_0^*, t_l^*] \subset [t, T]$, we have $\vartheta_l(\alpha_{t_0^*}^{t_l^*}) \in L^2(\mathcal{F}_{t,T}^\alpha)$ and $\{Z_{\vartheta_l}(s, \alpha_{t_0^*}^{t_l^*}); 0 \le s \le t\} \in \mathcal{M}^2_{\mathcal{F}_s^B \vee \mathcal{F}_{t,T}^\alpha}(0, t; R^{k \times d})$.

Now, let us show the convergence of equation (4) as l tends to ∞. For its left-hand side, by Doob's martingale convergence theorem (Lemma A.2), $E(\vartheta | \mathcal{G}_l^*) = \vartheta$ as $l \to \infty$. Thus

$$E\big(E(\vartheta | \mathcal{G}_l^*) | \mathcal{F}_t^B \vee \mathcal{F}_T^\alpha\big) \to E\big(\vartheta | \mathcal{F}_t^B \vee \mathcal{F}_T^\alpha\big) = Y_t.$$

For the right-hand side of equation (4),

$$\vartheta_l\big(\alpha_{t_0^*}^{t_l^*}\big) = E\big(E\big(E(\vartheta | \mathcal{G}_l^*) | \mathcal{F}_t^B \vee \mathcal{F}_T^\alpha\big) | \mathcal{F}_{t,T}^\alpha\big).$$

Thus $\vartheta_l(\alpha_{t_0^*}^{t_l^*})$ is a Cauchy sequence indexed by l in $L^2(\mathcal{F}_{t,T}^\alpha)$. Thus, by the Itô isometry, $\{Z_{\vartheta_l}(s, \alpha_{t_0^*}^{t_l^*}); 0 \le s \le t\}$ is a Cauchy sequence indexed by l in $\mathcal{M}^2_{\mathcal{F}_s^B \vee \mathcal{F}_{t,T}^\alpha}(0, t; R^{k \times d})$. Hence the right-hand side of equation (4) converges to some $\mathcal{F}_t^B \vee \mathcal{F}_{t,T}^\alpha$-measurable random variable, which yields that Y_t is also $\mathcal{F}_t^B \vee \mathcal{F}_{t,T}^\alpha$-measurable.

To prove the adaptability of Z, we rewrite BSDE (3) as

$$\int_t^T Z_s \, dB_s = -Y_t + \xi + \int_t^T f(s, \alpha_s) \, ds$$

with the right-hand side being $\mathcal{F}_T^B \vee \mathcal{F}_{t,T}^\alpha$-measurable. Applying Proposition 2.2 yields $\{Z_s; t \le s \le T\}$ is $\{\mathcal{F}_s^B \vee \mathcal{F}_{t,T}^\alpha\}_{t \le s \le T}$-adapted. Then, guaranteed by the continuous property of the Markov chain α, we can conclude $\{Z_s; t \le s \le T\}$ is also adapted to $\{\mathcal{F}_s^B \vee \mathcal{F}_{s,T}^\alpha\}_{t \le s \le T}$, and the proof is completed. □

Now let us consider the solvability of BSDE (1) for the general case.

Proof For $(y, z) \in M^2_{\mathcal{F}_t}(0, T; R^k \times R^{k \times d})$, using the assumptions in hypothesis (H1) and Hölder's inequality, we have

$$E\left(\int_0^T f(s, y_s, z_s, \alpha_s) \, ds\right)^2$$
$$\le 2E\left(\int_0^T \big(f(s, y_s, z_s, \alpha_s) - f(s, 0, 0, \alpha_s)\big) \, ds\right)^2 + 2E\left(\int_0^T f(s, 0, 0, \alpha_s) \, ds\right)^2$$
$$\le 2T\left(\mu^2 E \int_0^T \big(|y_s|^2 + |z_s|^2\big) \, ds + \sum_{i \in \mathcal{M}} E \int_0^T \big|f(s, 0, 0, i)\big|^2 \, ds\right) < \infty.$$

Thus we have $\xi + \int_0^T f(s, y_s, z_s, \alpha_s) \, ds \in L^2(\mathcal{G}_T; R^k)$. Setting

$$Y_t = \xi + \int_t^T f(s, y_s, z_s, \alpha_s) \, ds - \int_t^T Z_s \, dB_s,$$

and following the proof of Proposition 2.3, we can conclude that the mapping

$$I\big((y, z)\big) = (Y, Z) : M^2_{\mathcal{F}_t}\big(0, T; R^k \times R^{k \times d}\big) \to M^2_{\mathcal{F}_t}\big(0, T; R^k \times R^{k \times d}\big)$$

is well defined. Notice $\alpha = \{\alpha_t; 0 \le t \le T\}$ is a continuous-time finite-state Markov chain with the state space $\mathcal{M} = \{1, 2, \ldots, m\}$. As α varies with t, $\alpha_t \in \mathcal{M}$, $\forall t \in [0, T]$. Under hypothesis (H1), we have $\forall (\omega, t) \in \Omega \times [0, T]$, $(y_1, z_1), (y_2, z_2) \in R^k \times R^{k \times d}$,

$$\left| f(t, y_1, z_1, \alpha_t) - f(t, y_2, z_2, \alpha_t) \right| \le \mu \left(|y_1 - y_2| + |z_1 - z_2| \right).$$

Then, using the conventional approach (see [1, 3]), we can prove that this mapping is a contraction mapping. The details are omitted here for the compactness of the paper. Thus BSDE (1) has a unique solution pair in $M_{\mathcal{F}_t}^2(0, T; R^k \times R^{k \times d})$.

Together with the form of BSDE (1) and the Burkholder-Davis-Gundy inequality, we can conclude that Y is continuous and $Y \in S_{\mathcal{F}_t}^2(0, T; R^k)$. □

3 BSDEs with singularly perturbed Markov chains

Consider the case that the Markov chain has a large state space, one natural method is to adopt a hierarchical approach to reduce the complexity, and thus a singularly perturbed Markov chain is involved (see [14]). In this section we study the asymptotic property of a BSDE with the singularly perturbed Markov chain.

3.1 Preliminary for singularly perturbed Markov chains

Consider a continuous-time ε-dependent singularly perturbed Markov chain $\alpha^\varepsilon = \{\alpha_t^\varepsilon; 0 \le t \le T\}$ with a time-invariant generator $Q^\varepsilon = (1/\varepsilon)\tilde{Q} + \hat{Q}$, $\tilde{Q} = \operatorname{diag}(\tilde{Q}^1, \ldots, \tilde{Q}^l)$. Then its state space $\mathcal{M} = \{1, 2, \ldots, m\}$ can be written as $\mathcal{M} = \mathcal{M}_1 \cup \cdots \cup \mathcal{M}_l$ such that $\forall k \in \{1, \ldots, l\}$, $\tilde{Q}^k$ is a weakly irreducible[a] generator corresponding to the states in $\mathcal{M}_k$. When the magnitudes of generators $\tilde{Q}$ and $\hat{Q}$ have the same order, the singularly perturbed Markov chain α^ε fluctuates rapidly in a single group $\mathcal{M}_k$ and jumps less frequently among groups $\mathcal{M}_k$ and $\mathcal{M}_j$ for $k \ne j$. For notational convenience, for all $k \in \{1, \ldots, l\}$, we denote the states in $\mathcal{M}_k$ as $\{s_{k1}, \ldots, s_{km_k}\}$. For more details of this singularly perturbed Markov chain, we refer the interested reader to [14]. Here, we only recall the following asymptotic property of singularly perturbed Markov chains.

Proposition 3.1 [14] *Define the aggregated process* $\bar{\alpha}^\varepsilon = \{\bar{\alpha}_t^\varepsilon; 0 \le t \le T\}$ *as follows*: $\forall k \in \{1, \ldots, l\}$, $\bar{\alpha}_t^\varepsilon = k$, *when* $\alpha_t^\varepsilon \in \mathcal{M}_k$. *Then*

(1) *as* $\varepsilon \to 0$, $\bar{\alpha}^\varepsilon$ *converges weakly to a continuous-time Markov chain* $\bar{\alpha}$ *with the generator*

$$\bar{Q} = \operatorname{diag}(\nu^1, \ldots, \nu^l) \hat{Q} \operatorname{diag}(\mathbb{I}_{m_1}, \ldots, \mathbb{I}_{m_l}).$$

Here $\forall k \in \{1, \ldots, l\}$, ν^k *is the quasi-stationary distribution of* $\tilde{Q}^k$ *and* $\mathbb{I}_{m_k} = (1, \ldots, 1)' \in R^{m_k}$;

(2) *for any bounded deterministic function* $\beta(\cdot)$, $\forall k \in \{1, \ldots, l\}$,

$$E\left(\int_t^T \left(I_{\{\alpha_r^\varepsilon = s_{kj}\}} - \nu_j^k I_{\{\bar{\alpha}_r^\varepsilon = k\}} \right) \beta(r) \, dr \right)^2 = O(\varepsilon), \quad \forall j \in \{m_1, \ldots, m_k\}.$$

Here I_A *is the indicator function of the set* A.

3.2 Asymptotic property of BSDEs with singularly perturbed Markov chains

Denote $D(0, T; R^{2k})$ as the Skorohod space of *càdlàg* functions equipped with the Jakubowski S-topology. On this space, we consider the asymptotic property of the following BSDE with a singularly perturbed Markov chain with a generator f that does not depend on Z^ε:

$$Y_t^\varepsilon = \xi + \int_t^T f\left(s, Y_s^\varepsilon, \alpha_s^\varepsilon\right) ds - \int_t^T Z_s^\varepsilon \, dB_s. \tag{5}$$

In addition to hypothesis (H1), we make the following assumption for the coefficients in BSDE (5).

(H2) (i) $\xi \in L^2(\mathcal{F}_T^B; R^k)$; (ii) There exists a constant $C > 0$ such that, $\forall i \in \mathcal{M}$,
$\sup_{0 \le t \le T} |f(t, 0, i)| \le C$.

Theorem 3.1 *Under hypothesis* (H1) *and hypothesis* (H2), *the sequence of processes* $\{(Y_t^\varepsilon, \int_0^t Z_s^\varepsilon \, dB_s); 0 \le t \le T\}$ *converges weakly to the process* $\{(\bar{Y}_t, \int_0^t \bar{Z}_s d\bar{B}_s); 0 \le t \le T\}$ *as* $\varepsilon \to 0$ *on the space* $D(0, T; R^{2k})$. *Here* $(\bar{Y}, \bar{Z})$ *is the solution pair to the following BSDE with the limit Markov chain* $\bar{\alpha}$:

$$\bar{Y}_t = \xi + \int_t^T \bar{f}(s, \bar{Y}_s, \bar{\alpha}_s) \, ds - \int_t^T \bar{Z}_s \, d\bar{B}_s, \tag{6}$$

where $\bar{B}$ *is a d-dimensional Brownian motion satisfying* $\xi \in L^2(\mathcal{F}_T^{\bar{B}}; R^k)$, *and* $\bar{f}(s, y, i) = \sum_{j=1}^{m_i} v_j^i f(s, y, s_{ij})$ *for* $i \in \{1, \dots, l\}$.

It follows from Theorem 2.1 that both BSDEs (5) and (6) have unique solutions $(Y^\varepsilon, Z^\varepsilon)$ and $(\bar{Y}, \bar{Z}) \in S_{\mathcal{F}_t}^2(0, T; R^k) \times M_{\mathcal{F}_t}^2(0, T; R^{k \times d})$. To prove the expected convergence in Theorem 3.1, we firstly show the tightness property for $\{(Y_t^\varepsilon, \int_0^t Z_s^\varepsilon \, dB_s); 0 \le t \le T\}$ and then identify the limit.

Step 1: *tightness of* $\{(Y_t^\varepsilon, \int_0^t Z_s^\varepsilon \, dB_s); 0 \le t \le T\}$

Proposition 3.2 *There exists a positive constant* $C > 0$ *such that*

$$\sup_{\varepsilon > 0} E\left(\sup_{0 \le t \le T} \left|Y_t^\varepsilon\right|^2 + \int_0^T \left|Z_s^\varepsilon\right|^2 ds \right) \le C.$$

This result can be easily obtained by the standard estimation methods. We omit the proof here and put the details in our technical report (Proposition 3.2) on the website: http://arxiv.org/abs/1009.5074.

For notational convenience, we set $M_t^\varepsilon = \int_0^t Z_s^\varepsilon \, dB_s$. Thus BSDE (5) can be rewritten as

$$Y_t^\varepsilon = \xi + \int_t^T f\left(s, Y_s^\varepsilon, \alpha_s^\varepsilon\right) ds - \left(M_T^\varepsilon - M_t^\varepsilon\right). \tag{7}$$

Proposition 3.3 *The sequence of processes* $(Y^\varepsilon, M^\varepsilon) = \{(Y_t^\varepsilon, \int_0^t Z_s^\varepsilon \, dB_s), 0 \le t \le T\}$ *indexed by* $\varepsilon > 0$ *is tight on the space* $D(0, T; R^{2k})$ *equipped with the Jakubowski S-topology.*

Before giving the proof, let us firstly recall the Meyer-Zheng tightness criteria (see [17, 18] for more details): on the filtered probability space $\{\Omega; \mathcal{F}, \{\mathcal{F}_t\}_{0 \le t \le T}, P\}$, the sequence of

quasi-martingales $\{U_t^n; 0 \le t \le T\}$ indexed by n is tight whenever

$$\sup_n \left(\sup_{0 \le t \le T} E|U_t^n| + CV_T^0(U^n) \right) < \infty.$$

Here

$$CV_T^0(U^n) = \sup E\left(\sum_i \left| E(U_{t_{i+1}}^n - U_{t_i}^n | \mathcal{F}_{t_i}) \right| \right)$$

with the supreme taken over all partitions of the interval $[0, T]$.

Proof On the space $D(0, T; R^{2k})$, for fixed $\varepsilon > 0$, define the filtration $\{\mathcal{G}_t^\varepsilon\}_{0 \le t \le T} = \{\mathcal{F}_t^B \vee \mathcal{F}_T^{\alpha^\varepsilon} \vee \mathcal{N}; 0 \le t \le T\}$. We firstly prove the process M^ε is a $\{\mathcal{G}_t^\varepsilon\}_{0 \le t \le T}$-martingale. Let $\{t_i\}_{i \ge 0}$ be a dense subset of $[0, T]$ with $t_0 = 0$. For $n \ge 0$ and fixed $t \in [0, T]$, we denote $\alpha_{t_0}^{\varepsilon, t_n} = (\alpha_{t_0}^\varepsilon, \alpha_{t_1}^\varepsilon, \ldots, \alpha_{t_n}^\varepsilon)$ and $\mathcal{H}_n^\varepsilon = \mathcal{F}_{t,T}^B \vee \sigma\{\alpha_{t_0}^{\varepsilon, t_n}\}$. It is clear that $\mathcal{H}_n^\varepsilon \subset \mathcal{H}_{n+1}^\varepsilon$ and $\sigma(\bigcup_{n=1}^\infty \mathcal{H}_n^\varepsilon) = \mathcal{F}_{t,T}^B \vee \mathcal{F}_T^{\alpha^\varepsilon}$. Since $\int_t^T Z_s^\varepsilon \, dB_s$ is $\mathcal{F}_{t,T}^B \vee \mathcal{F}_T^{\alpha^\varepsilon}$-measurable, by Lemmas A.1 and A.2, we have

$$E\left(\int_t^T Z_s^\varepsilon \, dB_s \Big| \mathcal{H}_n^\varepsilon \right) \to E\left(\int_t^T Z_s^\varepsilon \, dB_s \Big| \mathcal{F}_{t,T}^B \vee \mathcal{F}_T^{\alpha^\varepsilon} \right) = \int_t^T Z_s^\varepsilon \, dB_s \quad \text{as } n \to \infty$$

and

$$E\left(\int_t^T Z_s^\varepsilon \, dB_s \Big| \mathcal{H}_n^\varepsilon \right) = H_n^\varepsilon(\alpha_{t_0}^{\varepsilon, t_n}) = \sum_{i_0^n \in \mathcal{M}^{m+1}} I_{\{\alpha_{t_0}^{\varepsilon, t_n} = i_0^n\}} H_n^\varepsilon(i_0^n),$$

where each $H_n^\varepsilon(i_0^n)$ is $\mathcal{F}_{t,T}^B$-measurable. Thus

$$E\left(E\left(\int_t^T Z_s^\varepsilon \, dB_s \Big| \mathcal{H}_n^\varepsilon \right) \Big| \mathcal{F}_t^B \vee \mathcal{F}_T^{\alpha^\varepsilon} \right) = \sum_{i_0^n \in \mathcal{M}^{m+1}} I_{\{\alpha_{t_0}^{\varepsilon, t_n} = i_0^n\}} E(H_n^\varepsilon(i_0^n) | \mathcal{F}_t^B \vee \mathcal{F}_T^{\alpha^\varepsilon}) = 0.$$

Letting $n \to \infty$, we have $E(\int_t^T Z_s^\varepsilon \, dB_s | \mathcal{F}_t^B \vee \mathcal{F}_T^{\alpha^\varepsilon}) = 0$. Thus $M^\varepsilon = \int_0^t Z_s^\varepsilon \, dB_s$ is a $\{\mathcal{G}_t^\varepsilon\}_{0 \le t \le T}$-martingale and it gives $CV_T^0(M^\varepsilon) = 0$. Using the Burkholder-Davis-Gundy inequality, for any $\varepsilon > 0$,

$$\sup_{0 \le t \le T} E|M_t^\varepsilon| \le \frac{1}{2}\left(1 + E\left(\sup_{0 \le t \le T} |M_t^\varepsilon|^2 \right) \right) \le \frac{1}{2}\left(1 + C \sup_{\varepsilon > 0} E[M^\varepsilon]_T \right)$$

$$= \frac{1}{2}\left(1 + C \sup_{\varepsilon > 0} E \int_0^T |Z_s^\varepsilon|^2 \, ds \right).$$

From Proposition 3.2, we deduce that $\sup_{\varepsilon > 0}(\sup_{0 \le t \le T} E|M_t^\varepsilon|) < \infty$.

$\forall \varepsilon > 0$, for any partition of $[0, T]$, we have

$$E\left(\sum_i \left| E(Y_{t_{i+1}}^\varepsilon - Y_{t_i}^\varepsilon | \mathcal{G}_{t_i}^\varepsilon) \right| \right) \le E \int_0^T \left| f(s, Y_s^\varepsilon, \alpha_s^\varepsilon) \right| ds$$

$$\le E \int_0^T \left(\left| f(s, 0, \alpha_s^\varepsilon) \right| + \mu |Y_s^\varepsilon| \right) ds$$

$$\le E \int_0^T \left| f(s, 0, \alpha_s^\varepsilon) \right| ds + \mu T E\left(\sup_{0 \le s \le T} |Y_s^\varepsilon| \right).$$

Combining hypothesis (H1) and Proposition 3.2, we have $\sup_{\varepsilon>0}(CV_T^0(Y^\varepsilon)) < \infty$. Thus the 'Meyer-Zheng tightness criteria' are fully satisfied, *i.e.*,

$$\sup_{\varepsilon>0}\left(CV_T^0(Y^\varepsilon) + \sup_{0\leq t\leq T} E|Y_t^\varepsilon| + \sup_{0\leq t\leq T} E|M_t^\varepsilon|\right) < \infty,$$

and the tightness of $(Y^\varepsilon, M^\varepsilon)$ is obtained. $\qquad\square$

Step 2: *identification of the limit*

Since $(Y^\varepsilon, M^\varepsilon)$ is tight in $D(0, T; R^{2k})$, there exists a subsequence of $(Y^\varepsilon, M^\varepsilon)$, which we still denote by $(Y^\varepsilon, M^\varepsilon)$, converging weakly toward a càdlàg process $(\tilde{Y}, \tilde{M})$.

Proposition 3.4 *For the limit process $(\tilde{Y}, \tilde{M})$, we have*

(i) *there exists a countable subset D of $[0, T]$ such that for all $t \in [0, T] \setminus D$,*

$$\tilde{Y}_t = \xi + \int_t^T \bar{f}(s, \tilde{Y}_s, \bar{\alpha}_s)\, ds - (\tilde{M}_T - \tilde{M}_t). \tag{8}$$

Here the definition of the Markov chain $\bar{\alpha}$ and its relation to the singularly perturbed Markov chain α^ε are presented in Proposition 3.1;

(ii) *for a d-dimensional Brownian motion $\bar{B}$, denote the filtration $\{\mathcal{H}_t\}_{0\leq t\leq T} = \{\mathcal{F}_t^{\bar{B}} \vee \mathcal{F}_T^{\bar{\alpha}}; 0 \leq t \leq T\}$. If $\tilde{Y}$ is $\{\mathcal{H}_t\}_{0\leq t\leq T}$-adapted and ξ is $\mathcal{F}_T^{\bar{B}}$-measurable, then $\tilde{M}$ is an $\{\mathcal{H}_t\}_{0\leq t\leq T}$-martingale.*

Proof Consider BSDE (7) satisfied by the process $(Y^\varepsilon, M^\varepsilon)$

$$Y_t^\varepsilon = \xi + \int_t^T f(s, Y_s^\varepsilon, \alpha_s^\varepsilon)\, ds - (M_T^\varepsilon - M_t^\varepsilon),$$

and recall that $(Y^\varepsilon, M^\varepsilon)$ converges weakly to $(\tilde{Y}, \tilde{M})$. To obtain its weak limit as $\varepsilon \to 0$, we only need to consider the second item on the right-hand side

$$\int_t^T f(s, Y_s^\varepsilon, \alpha_s^\varepsilon)\, ds$$

$$= \int_t^T \sum_{i=1}^l \sum_{j=1}^{m_i} f(s, Y_s^\varepsilon, s_{ij}) I_{\{\alpha_s^\varepsilon = s_{ij}\}}\, ds$$

$$= \int_t^T \sum_{i=1}^l \sum_{j=1}^{m_i} f(s, Y_s^\varepsilon, s_{ij})\left(I_{\{\alpha_s^\varepsilon = s_{ij}\}} - \nu_j^i I_{\{\alpha_s^\varepsilon \in \mathcal{M}_i\}}\right) ds + \int_t^T \bar{f}(s, Y_s^\varepsilon, \bar{\alpha}_s^\varepsilon)\, ds.$$

Here $\bar{\alpha}^\varepsilon$ is the aggregated process for the singularly perturbed Markov chain α^ε given in Proposition 3.1, and $\bar{f}(s, y, i) = \sum_{j=1}^{m_i} \nu_j^i f(s, y, s_{ij})$ for $i \in \{1, \dots, l\}$. Since $(Y^\varepsilon, \bar{\alpha}^\varepsilon)$ converges weakly to $(\tilde{Y}, \bar{\alpha})$, we have

$$\int_t^T \bar{f}(s, Y_s^\varepsilon, \bar{\alpha}_s^\varepsilon)\, ds \quad \text{converges weakly to} \quad \int_t^T \bar{f}(s, \tilde{Y}_s, \bar{\alpha}_s)\, ds.$$

In addition, as a direct consequence of Proposition 3.1, Proposition 3.2 and Proposition 3.3,

$$\sup_{0 \le t \le T} E \left| \int_t^T f(s, Y_s^\varepsilon, s_{ij}) \left(I_{\{\alpha_s^\varepsilon = s_{ij}\}} - \nu_j^i I_{\{\alpha_s^\varepsilon \in \mathcal{M}_i\}} \right) ds \right| \to 0 \quad \text{as } \varepsilon \to 0.$$

The limit of BSDE (7) as $\varepsilon \to 0$ leads to assertion (i).

Now, let us prove assertion (ii). For any $t_1 \in [0, T]$, let Φ_{t_1} be a continuous mapping from $C(0, t_1; R^d) \times D(0, t_1; R^k) \times D(0, T; \bar{\mathcal{M}})$. Here $\bar{\mathcal{M}} = \{1, \ldots, l\}$ is the state space of aggregated process $\bar{\alpha}^\varepsilon$. $\forall \varepsilon > 0$, noticing that M^ε is a $\{\mathcal{G}_t^\varepsilon\}_{0 \le t \le T}$-martingale, Y^ε and $\bar{\alpha}^\varepsilon$ are $\{\mathcal{G}_t^\varepsilon\}_{0 \le t \le T}$-adapted, for all $t_2 \in [t_1, T]$, we have

$$E\left(\Phi_{t_1}\left(B, Y^\varepsilon, \bar{\alpha}^\varepsilon\right)\left(M_{t_2}^\varepsilon - M_{t_1}^\varepsilon\right)\right) = 0, \quad i.e.,$$

$$E\left(\Phi_{t_1}\left(B, Y^\varepsilon, \bar{\alpha}^\varepsilon\right)\left(Y_{t_2}^\varepsilon - Y_{t_1}^\varepsilon + \int_{t_1}^{t_2} f(s, Y_s^\varepsilon, \alpha_s^\varepsilon)\, ds \right)\right) = 0.$$

With the same analysis of assertion (i), we can prove that $\int_{t_1}^{t_2} f(s, Y_s^\varepsilon, \alpha_s^\varepsilon)\, ds$ converges weakly to $\int_{t_1}^{t_2} \bar{f}(s, \tilde{Y}_s, \bar{\alpha}_s)\, ds$ as $\varepsilon \to 0$. By virtue of the fact that $(Y^\varepsilon, \bar{\alpha}^\varepsilon)$ converges weakly to $(\tilde{Y}, \bar{\alpha})$ as $\varepsilon \to 0$, we have

$$E\left(\Phi_{t_1}(\bar{B}, \tilde{Y}, \bar{\alpha})\left(\tilde{Y}_{t_2} - \tilde{Y}_{t_1} + \int_{t_1}^{t_2} \bar{f}(s, \tilde{Y}_s, \bar{\alpha}_s)\, ds \right)\right) = 0, \quad i.e.,$$

$$E\left(\Phi_{t_1}(\bar{B}, \tilde{Y}, \bar{\alpha})(\tilde{M}_{t_2} - \tilde{M}_{t_1})\right) = 0.$$

Here $\bar{B}$ is a d-dimensional Brownian motion with $\bar{B}_0 = 0$. With the assumptions that $\tilde{Y}$ is $\{\mathcal{H}_t\}_{0 \le t \le T}$-adapted and ξ is $\mathcal{F}_T^{\bar{B}}$-measurable, the adaptness of $\tilde{M}$ with respect to $\{\mathcal{H}_t\}_{0 \le t \le T}$ follows. From the freedom choice of t_1, t_2 and Φ_{t_1}, we can conclude that $\tilde{M}$ is an $\{\mathcal{H}_t\}_{0 \le t \le T}$-martingale. □

We come back to finish the proof of Theorem 3.1. Denote $\bar{M} = \{\bar{M}_t; 0 \le t \le T\} = \{\int_0^t \bar{Z}_s\, dB_s; 0 \le t \le T\}$. For all $t \in [0, T] \setminus D$, using Itô's formula to $|\tilde{Y}_t - \bar{Y}_t|^2$ on $[t, T]$, we have

$$E\left| \tilde{Y}_t - \bar{Y}_t \right|^2 + E\left([\tilde{M} - \bar{M}]_T - [\tilde{M} - \bar{M}]_t\right)$$

$$= 2E \int_t^T \left(\bar{f}(s, \tilde{Y}_s, \bar{\alpha}_s) - \bar{f}(s, \bar{Y}_s, \bar{\alpha}_s) \right)(\tilde{Y}_s - \bar{Y}_s)\, ds$$

$$\le CE \int_t^T \left| \tilde{Y}_s - \bar{Y}_s \right|^2 ds.$$

From Gronwall's lemma, we obtain $E|\tilde{Y}_t - \bar{Y}_t|^2 = 0$ for all $t \in [0, T] \setminus D$. Since $\bar{Y}$ is continuous, $\tilde{Y}$ is *càdlàg*, and D is countable, we get $\tilde{Y}_t = \bar{Y}_t$, P-a.s. for all $t \in [0, T]$. And the identity between $\tilde{M}$ and $\bar{M}$ follows directly.

Remark 3.1 In this paper, we study that the asymptotic property of the solution to BSDE with the generator f only depends on Y^ε. For the general case that f depends on $(Y^\varepsilon, Z^\varepsilon)$, to prove the corresponding asymptotic property requires the tightness of the sequence

$\{(Y_t^\varepsilon, Z_t^\varepsilon, \int_0^t Z_s^\varepsilon \, dB_s); 0 \le t \le T\}$. To our knowledge, due to the lack of suitable estimates, it is hard to show $\{Z_t^\varepsilon; 0 \le t \le T\}$ to be tight. This will be one of our future research objects.

3.3 Examples

Comparing the limit Markov chain $\bar{\alpha}$ with the singularly perturbed Markov chain α^ε, the state space of the former is much smaller. Thus it follows from Theorem 3.1 that the asymptotic property of BSDE (5) with the singularly perturbed Markov chain α^ε can be characterized by the limit BSDE (6) which is simpler. This advantage is more clearly illustrated by the following two examples.

Example 3.1 Consider the case that BSDE (5) with a singularly perturbed Markov chain α^ε is a fast-varying noise process, *i.e.*, the generator $\tilde{Q}$ is weakly irreducible with the state space $\mathcal{M}$. From Theorem 3.1, as $\varepsilon \to 0$, the corresponding limit BSDE is

$$Y_t = \xi + \int_t^T \sum_{i=1}^m v_i f(s, Y_s, i) \, ds - \int_t^T Z_s \, d\bar{B}_s, \tag{9}$$

where $v = (v_1, \ldots, v_m)$ is the quasi-stationary distribution of $\tilde{Q}$. Thus its solution, an $\{\mathcal{F}_t^{\bar{B}}\}_{0 \le t \le T}$-adapted process Y, can be used to study the asymptotic distribution for the sequence of $\{\mathcal{F}_t^B \vee \mathcal{F}_{t,T}^{\alpha^\varepsilon}\}_{0 \le t \le T}$-adapted process Y^ε indexed by ε.

Example 3.2 Consider the following BSDE:

$$Y_t = \xi + \int_t^T f(s, Y_s, \alpha_s) \, ds - \int_t^T Z_s \, dB_s, \tag{10}$$

where the generator of the Markov chain α is

$$Q = \begin{pmatrix} -22 & 20 & 2 \\ 41 & -42 & 1 \\ 1 & 2 & -3 \end{pmatrix},$$

and the corresponding state space is $\mathcal{M} = \{s_1, s_2, s_3\}$. It is obvious that the transition rate between s_1 and s_2 is larger than transition rates between s_3 and other states. Due to this property, we can show some approximation properties for the solution of BSDE (10) by Theorem 3.1.

Firstly, we rewrite the generator Q in the form of the generator for a singularly perturbed Markov chain

$$Q = \frac{1}{0.05} \tilde{Q} + \hat{Q} = \frac{1}{0.05} \begin{pmatrix} -1 & 1 & 0 \\ 2 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} -2 & 0 & 2 \\ 1 & -2 & 1 \\ 1 & 2 & -3 \end{pmatrix}.$$

Here $\tilde{Q}$, $\hat{Q}$ and the number 0.05 are chosen to guarantee $\tilde{Q}$ and $\hat{Q}$ to be generators with the same order of magnitude.

Then, with the assigned $\tilde{Q}$ and $\hat{Q}$, we introduce an ε-dependent singularly perturbed Markov chain $\alpha^\varepsilon = \{\alpha_t^\varepsilon; 0 \le t \le T\}$ with the generator

$$Q^\varepsilon = \frac{1}{\varepsilon}\tilde{Q} + \hat{Q} = \frac{1}{\varepsilon}\begin{pmatrix} -1 & 1 & 0 \\ 2 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} -2 & 0 & 2 \\ 1 & -2 & 1 \\ 1 & 2 & -3 \end{pmatrix}.$$

The state space $\mathcal{M}$ can be divided into two groups with $\mathcal{M}_1 = \{s_1, s_2\}$ and $\mathcal{M}_2 = \{s_3\}$, and the related weakly irreducible generators are $\tilde{Q}^1 = \begin{pmatrix} -1 & 1 \\ 2 & -2 \end{pmatrix}$ and $\tilde{Q}^1 = (0)$.

After that, following the procedures in Sections 3.1 and 3.2, we define the aggregated process $\bar{\alpha}^\varepsilon = \{\bar{\alpha}_t^\varepsilon; 0 \le t \le T\}$, with

$$\bar{\alpha}_t^\varepsilon = \begin{cases} 1, & \text{when } \alpha_t^\varepsilon = s_1 \text{ or } s_2, \\ 2, & \text{when } \alpha_t^\varepsilon = s_3. \end{cases}$$

And we can calculate that the generator of the limit Markov chain $\bar{\alpha}$ is $\bar{Q} = \begin{pmatrix} -5/3 & 5/3 \\ 3 & -3 \end{pmatrix}$, and quasi-stationary distributions of $\tilde{Q}^1$ and $\tilde{Q}^2$ are $v^1 = (2/3, 1/3)$ and $v^2 = 1$. By Theorem 3.1, with the following limit BSDE:

$$Y_t = \xi + \int_t^T \bar{f}(s, Y_s, \bar{\alpha}_s)\, ds - \int_t^T Z_s\, d\bar{B}_s,$$

where $\bar{f}(t, y, 1) = (2/3)f(t, y, s_1) + (1/3)f(t, y, s_2)$ and $\bar{f}(t, y, 2) = f(t, y, s_3)$, thus we can use the distribution property of its solution as an approximation for the property of the solution to the original BSDE (10).

Noticing that the limit averaged Markov chain has two states and the original one has three states, we have reduced the complexity of model (10). This advantage will be much clearer when the state space of the original Markov chain is sufficiently larger.

4 Application: homogenization of a system of semilinear parabolic PDEs with singularly perturbed Markov chains

It is well known that BSDEs provide a probabilistic approach to study one class of semilinear parabolic PDEs (see [19–21]). In this section, as an application of our obtained asymptotic results, we show the homogenization property of a sequence of semi-linear parabolic PDEs with a singularly perturbed Markov chain α^ε.

For any $(t, x) \in [0, T] \times R^m$, consider the following semi-linear parabolic PDE with the singularly perturbed Markov chain α^ε:

$$u^\varepsilon(t, x) = h(x) + \int_t^T \left(\mathcal{L}u^\varepsilon(r, x) + f\big(r, x, u^\varepsilon(r, x), \alpha_r^\varepsilon\big) \right) dr. \tag{11}$$

Here $\mathcal{L}$ is the following second-order differential operator:

$$\mathcal{L}u^\varepsilon(r, x) = \left(\frac{1}{2} \mathrm{Tr}\big(\sigma\sigma' D^2 u^\varepsilon\big) + \langle Du^\varepsilon, b \rangle \right)(r, x).$$

As in [19], we adopt the following notations: $C^k(R^p; R^q)$ is the space of functions of class C^k from R^p to R^q, $C_{l,b}^k(R^p; R^q)$ is the space of functions of class C^k whose partial derivatives

of order less than or equal to k are bounded, and $C_p^k(R^p; R^q)$ is the space of functions of class C^k which, together with all their partial derivatives of order less than or equal to k, grow at most like a polynomial function of the variable x at infinity. Also, we make the following assumptions for the coefficients h, b, σ, and f.

(H3) (i) $b \in C_{l,b}^3(R^m; R^m)$, $\sigma \in C_{l,b}^3(R^m; R^{m \times d})$, and $h \in C_p^3(R^k)$; (ii) For
$f : [0,T] \times R^m \times R^k \times \mathcal{M} \to R^k$, $\forall (r,i) \in [0,T] \times \mathcal{M}$, $(x,u) \to f(r,x,u,i)$ is of class C^3. Moreover, $f(r,\cdot,0,i) \in C_p^3(R^m; R^k)$, and the first-order partial derivative in u is bounded on $[0,T] \times R^m \times R^k \times \mathcal{M}$, as well as their derivatives of order one and two with respect to x and u.

For the backward PDE (11), we consider the solution restricted to $C^{1,2}([0,T] \times R^m; R^k)$. For a fixed $\varepsilon > 0$, let the process $u^\varepsilon = \{u^\varepsilon(t,x); 0 \le t \le T, x \in R^m\}$ be the unique solution to equation (11). We have the following homogenization result.

Theorem 4.1 *Under hypothesis* (H3), *as $\varepsilon \to 0$, the sequence of u^ε converges weakly to a process $\bar{u}$, which solves the following PDE*:

$$\bar{u}(t,x) = h(x) + \int_t^T \left(\mathcal{L}\bar{u}(r,x) + \bar{f}\big(r,x,\bar{u}(r,x),\bar{\alpha}_r\big) \right) dr, \quad 0 \le t \le T. \tag{12}$$

Here $\bar{\alpha}$ is the limit averaged Markov chain with the state space $\bar{\mathcal{M}} = \{1,\ldots,l\}$ given in Proposition 3.1, *and for $i \in \bar{\mathcal{M}}$, $\bar{f}(t,x,\bar{u},i) = \sum_{j=1}^{m_i} v_j^i f(t,x,\bar{u},s_{ij})$.*

Before giving the proof, we first present a relation between backward PDE (11) and the following FBSDE with the singularly perturbed Markov chain α^ε:

$$X_s^{t,x} = x + \int_t^s b\big(X_r^{t,x}\big) dr + \int_t^s \sigma\big(X_r^{t,x}\big) dB_r, \tag{13}$$

$$Y_s^{\varepsilon,t,x} = h\big(X_T^{t,x}\big) + \int_s^T f\big(r, X_r^{t,x}, Y_r^{\varepsilon,t,x}, \alpha_r^\varepsilon\big) dr - \int_s^T Z_r^{\varepsilon,t,x} dB_r. \tag{14}$$

For $s \le t$, define $X_s^{t,x} = X_{s \vee t}^{t,x}$, $Y_s^{t,x} = Y_{s \vee t}^{t,x}$ and $Z_s^{t,x} = 0$. Then $(X,Y,Z) = (X_s^{t,x}, Y_s^{t,x}, Z_s^{t,x})$ is defined on $(s,t) \in [0,T]^2$.

Proposition 4.1 *For a fixed $\varepsilon > 0$, the process $\{Y_t^{\varepsilon,t,x}; 0 \le t \le T, x \in R^m\}$ solves the backward PDE* (11), *i.e., $\forall (t,x) \in [0,T] \times R^m$, $u^\varepsilon(t,x) = Y_t^{\varepsilon,t,x}$.*

The proof of this proposition, similar to Theorem 3.2 in [19], is omitted here, but it is fully presented in our technical report (Theorem 4.2) on the website: http://arxiv.org/abs/1009.5074. Now we are ready to prove Theorem 4.1.

Proof Noticing that the forward SDE (13) does not involve the singularly Markov chain α^ε, by hypothesis (H3), we can verify that the terminal value of BSDE (14), $h(X_T^{t,x})$, satisfies hypotheses (H1)-(H2). Denote $\tilde{f}(r,y,i) = f(r, X_r^{t,x}, y, i)$. Then $\tilde{f}$ also satisfies hypotheses (H1)-(H2). By Theorem 3.1, we obtain that for all $(t,x) \in [0,T] \times R^m$, $Y_t^{\varepsilon,t,x}$ converges weakly to $\bar{Y}_t^{t,x}$ as $\varepsilon \to 0$, where $\{(\bar{Y}_s^{t,x}, \bar{Z}_s^{t,x}); t \le s \le T\}$ satisfies

$$\bar{Y}_s^{t,x} = h\big(X_T^{t,x}\big) + \int_s^T \bar{f}\big(r, X_r^{t,x}, \bar{Y}_r^{t,x}, \bar{\alpha}_r\big) dr - \int_s^T \bar{Z}_r^{t,x} d\bar{B}_r. \tag{15}$$

For all $(t, x) \in [0, T] \times R^m$, denote $\bar{u}(t, x) = \bar{Y}_t^{t,x}$. From Proposition 4.1, we know that the process $\bar{u} = \{\bar{u}(t, x); 0 \le t \le T, x \in R^m\}$ solves backward PDE (12). Noticing that $\{u^\varepsilon(t, x) = Y_t^{\varepsilon,t,x}; 0 \le t \le T, x \in R^m\}$ solves PDE (11), the result follows. □

5 Conclusion

In this paper, the solvability of a new kind of BSDE with the generator involving a finite-state Markov chain is studied. To overcome the difficulty applying the conventional approach with the contraction mapping, an enlarged filtration is constructed and the related Itô representation result is proposed. Then, considering the case that a singularly perturbed Markov chain is involved, we present the asymptotic property of the corresponding BSDE by virtue of the related property of the singularly perturbed Markov chain. This result provides a probabilistic approach to study the homogenization property of the system of backward PDEs with the singularly perturbed Markov chain.

It is noted that in this paper we only give the homogenization result of PDEs system with Markov chains when there exists the classical solution under sufficiently regular assumptions. In the successive work, we will study the weak solution in a Sobolev space for the related PDEs system and homogenization problem by virtue of BSDEs with Markov chains.

As in [22], the Markov chain can be used to capture the market trends which are crucial factors that affect most investment decisions. We believe that our results have applications in such financial markets due to the deep connection between BSDEs and finance. Some applications of this kind of BSDEs in optimal control would also be interesting to investigate in our future research.

Appendix: A specific proof for Proposition 2.1

Let us first recall the Doob-Dynkin lemma and Doob's martingale convergence theorem before giving the proof.

Lemma A.1 (Doob-Dynkin lemma) *Let $X, Y : \Omega \to R^d$ be two given functions. Then Y is $\sigma(X)$-measurable if and only if $Y = g(X)$ for some Borel measurable function $g : R^d \to R^d$.*

Lemma A.2 (Doob's martingale convergence theorem) *Let $\{\mathcal{F}_t\}_{t \ge 0}$ be a filtration on the space $(\Omega, \mathcal{F}, P)$, $\mathcal{F}_\infty = \sigma(\bigcup_{t \ge 0} \mathcal{F}_t)$ and $X \in L^1(\mathcal{F}; R^d)$. Then*

$$E(X|\mathcal{F}_t) \to E(X|\mathcal{F}_\infty) \quad as\ t \to \infty,\ a.s.\ and\ in\ L^1\ as\ well.$$

Proof The uniqueness of Proposition 2.2 follows straightforwardly from the Itô isometry, so we only focus on the existence. Let $\{t_i\}_{i \ge 0}$, $\{t_j'\}_{j \ge 0}$ be two dense subsets of $[0, T]$ with $t_0 = t_0' = 0$. For each integer $n, m \ge 0$, we denote $B_{t_0}^{t_n} = (B_{t_0}, B_{t_1}, \ldots, B_{t_n})$, $\alpha_{t_0'}^{t_m'} = (\alpha_{t_0'}, \alpha_{t_1'}, \ldots, \alpha_{t_m'})$, and $\mathcal{G}_{n,m} = \sigma\{B_{t_0}^{t_n}, \alpha_{t_0'}^{t_m'}\}$. It is clear that $\mathcal{G}_{n,m} \subset \mathcal{G}_{n+1,m}$, $\mathcal{G}_{n,m} \subset \mathcal{G}_{n,m+1}$, $\mathcal{G}_{n,m} \subset \mathcal{G}_{n+1,m+1}$, $\sigma(\bigcup_{n=1}^\infty \mathcal{G}_{n,m}) = \mathcal{F}_T^B \vee \sigma\{\alpha_{t_0'}^{t_m'}\}$, and $\sigma(\bigcup_{m=0}^\infty (\mathcal{F}_T^B \vee \sigma\{\alpha_{t_0'}^{t_m'}\})) = \mathcal{G}_T$.

For $N \in L^2(\mathcal{G}_T; R^k)$, applying Lemma A.2, we get

$$E(N|\mathcal{G}_{n,m}) \to E(N|\mathcal{F}_T^B \vee \sigma\{\alpha_{t_0'}^{t_m'}\}) \quad as\ n \to \infty.$$

By Lemma A.1, $\forall n, m \geq 0$, there exists a Borel measurable function $N_{n,m} : R^{(n+1) \times d} \times \mathcal{M}^{m+1} \to R^k$ such that $E(N|\mathcal{G}_{n,m}) = N_{n,m}(B_{t_0}^{t_n}, \alpha_{t_0'}^{t_m'})$. Thus we can write $E(N|\mathcal{F}_T^B \vee \sigma\{\alpha_{t_0'}^{t_m'}\})$ as follows:

$$E\big(N|\mathcal{F}_T^B \vee \sigma\big\{\alpha_{t_0'}^{t_m'}\big\}\big) = \sum_{(i_0, i_1, \ldots, i_m) \in \mathcal{M}^{m+1}} I_{\{\alpha_{t_0'}^{t_m'} = (i_0, i_1, \ldots, i_m)\}} N_m(i_0, i_1, \ldots, i_m),$$

where $N_m(i_0, i_1, \ldots, i_m)$ is $\mathcal{F}_T^B$-measurable. For ease of notation, let $N_m(\alpha_{t_0'}^{t_m'}) \triangleq E(N|\mathcal{F}_T^B \vee \sigma\{\alpha_{t_0'}^{t_m'}\})$, $i_0^m \triangleq (i_0, i_1, \ldots, i_m)$, and $\mathcal{M}^{m+1} = \underbrace{\mathcal{M} \times \cdots \times \mathcal{M}}_{m+1}$. For the above equation, applying the Itô representation theorem to each $N_m(i_0^m)$ yields

$$N_m\big(\alpha_{t_0'}^{t_m'}\big) = \sum_{i_0^m \in \mathcal{M}^{m+1}} I_{\{\alpha_{t_0'}^{t_m'} = i_0^m\}} \left(E\big(N_m\big(i_0^m\big)\big) + \int_0^T Z_{N_m}\big(t, i_0^m\big) \, dB_t \right). \tag{16}$$

Noticing that $\forall i_0^m \in \mathcal{M}^{m+1}$, $\{Z_{N_m}(t, i_0^m); 0 \leq t \leq T\} \in M_{\mathcal{F}_t^B}^2(0, T; R^{k \times d})$, and α is independent of B, we have $E(\int_0^T Z_{N_m}(t, i_0^m) \, dB_t | \mathcal{F}_T^\alpha) = 0$. Taking the conditional expectation of equation (16) with respect to $\mathcal{F}_T^\alpha$, one can get

$$E\big(N_m\big(\alpha_{t_0'}^{t_m'}\big)|\mathcal{F}_T^\alpha\big) = \sum_{i_0^m \in \mathcal{M}^{m+1}} I_{\{\alpha_{t_0'}^{t_m'} = i_0^m\}} E\big(N_m\big(i_0^m\big)\big).$$

Thus equation (16) can be rewritten as

$$N_m\big(\alpha_{t_0'}^{t_m'}\big) = E\big(N_m\big(\alpha_{t_0'}^{t_m'}\big)|\mathcal{F}_T^\alpha\big) + \int_0^T Z_{N_m}\big(t, \alpha_{t_0'}^{t_m'}\big) \, dB_t, \tag{17}$$

where

$$Z_{N_m}\big(t, \alpha_{t_0'}^{t_m'}\big) = \sum_{i_0^m \in \mathcal{M}^{m+1}} I_{\{\alpha_{t_0'}^{t_m'} = i_0^m\}} Z_{N_m}\big(t, i_0^m\big).$$

We are going to show that as $m \to \infty$, equation (17) converges to $N = E(N|\mathcal{F}_T^\alpha) + \int_0^T Z_t \, dB_t$. For the left-hand side, applying Doob's martingale convergence theorem and using the fact that $N_m(\alpha_{t_0'}^{t_m'}) = E(N|\mathcal{F}_T^B \vee \sigma\{\alpha_{t_0'}^{t_m'}\})$ and $\sigma(\bigcup_{m=0}^\infty (\mathcal{F}_T^B \vee \sigma\{\alpha_{t_0'}^{t_m'}\})) = \mathcal{G}_T$ yields

$$N_m\big(\alpha_{t_0'}^{t_m'}\big) \to E(N|\mathcal{G}_T) = N \quad \text{as } m \to \infty.$$

Then

$$E\big(N_m\big(\alpha_{t_0'}^{t_m'}\big)|\mathcal{F}_T^\alpha\big) \to E\big(N|\mathcal{F}_T^\alpha\big) \quad \text{as } m \to \infty,$$

and the convergence of the first term on the right-hand side of equation (17) is guaranteed. The remaining work is to prove that $\{Z_{N_m}(t, \alpha_{t_0'}^{t_m'}), 0 \leq t \leq T\}$ is a Cauchy sequence.

Actually, by the Itô isometry, we have

$$E\big(\big(N_k\big(\alpha_{t_0'}^{t_k'}\big) - N_m\big(\alpha_{t_0'}^{t_m'}\big)\big)^2\big)$$

$$= E\bigg(E\big(N_k\big(\alpha_{t_0'}^{t_k'}\big)|\mathcal{F}_T^\alpha\big) - E\big(N_m\big(\alpha_{t_0'}^{t_m'}\big)|\mathcal{F}_T^\alpha\big) + \int_0^T \big(Z_{N_k}\big(t,\alpha_{t_0'}^{t_k'}\big) - Z_{N_m}\big(t,\alpha_{t_0'}^{t_m'}\big)\big)\,dB_t\bigg)^2$$

$$= E\big(E\big(N_k\big(\alpha_{t_0'}^{t_k'}|\mathcal{F}_T^\alpha\big)\big) - E\big(N_m\big(\alpha_{t_0'}^{t_m'}|\mathcal{F}_T^\alpha\big)\big)\big)^2 + \int_0^T E\big(Z_{N_k}\big(t,\alpha_{t_0'}^{t_k'}\big) - Z_{N_m}\big(t,\alpha_{t_0'}^{t_m'}\big)\big)^2\,dt$$

$$\to 0 \quad \text{as } k, m \to \infty.$$

Noticing that $\{Z_{N_m}(t,\alpha_{t_0'}^{t_m'}); 0 \le t \le T\} \in M_{\mathcal{G}_t}^2(0,T;R^{k\times d})$, hence it converges to some $Z \in M_{\mathcal{G}_t}^2(0,T;R^{k\times d})$, and the proof is completed. □

Abbreviations
BSDEs: backward stochastic differential equations; PDEs: partial differential equations.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
Authors contributed equally in writing this article. All authors read and approved the final manuscript.

Author details
[1] School of Information Science and Engineering, Shandong University, Jinan, 250100, P.R. China. [2] School of Mathematics, Shandong University, Jinan, 250100, P.R. China.

Endnote
[a] A generator Q is called weakly irreducible if the system of equations $\nu Q = 0$ and $\sum_{i=1}^m \nu_i = 1$ has a unique nonnegative solution. This nonnegative solution $\nu = (\nu_1, \cdots, \nu_m)$ is called the quasi-stationary distribution of Q.

References
1. Pardoux, E, Peng, S: Adapted solution of a backward stochastic differential equations. Syst. Control Lett. **14**, 55-61 (1990)
2. Duffie, D, Epstein, L: Stochastic differential utility. Econometrica **60**(2), 353-394 (1992)
3. El Karoui, N, Peng, S, Quenez, MC: Backward stochastic differential equations in finance. Math. Finance **7**, 1-71 (1997)
4. Eyraud-Loisel, A: Backward stochastic differential equations with enlarged filtration: option hedging of an insider trader in a financial market with jumps. Stoch. Process. Appl. **115**, 1745-1763 (2005)
5. Nualart, D, Schoutens, W: BSDE's and Feynman-Kac formula for Lévy processes with applications in finance. Bernoulli **7**(5), 761-776 (2001)
6. Bahlali, K, Eddahbi, M, Essaky, E: BSDE associated with Lévy processes and application to PDIE. J. Appl. Math. Stoch. Anal. **16**, 1-17 (2003)
7. Tang, S, Li, X: Necessary conditions for optimal control of stochastic systems with random jumps. SIAM J. Control Optim. **32**(5), 1447-1475 (1994)
8. Kharroubi, I, Ma, J, Pham, H, Zhang, J: Backward SDEs with constrained jumps and quasi-variational inequalities. Ann. Probab. **38**(2), 794-840 (2010)
9. Cohen, S, Elliott, R: Solutions of backward stochastic differential equations on Markov chains. Commun. Stoch. Anal. **2**(2), 251-262 (2008)
10. Cohen, S, Elliott, R: On Markovian solutions to Markov chain BSDEs. Numer. Algebra Control Optim. **2**, 257-267 (2012)
11. Cohen, S, Szpruch, L: Comparisons for backward stochastic differential equations on Markov chains and related no-arbitrage conditions. Ann. Appl. Probab. **20**, 267-311 (2010)
12. Lu, W, Ren, Y: Anticipated backward stochastic differential equations on Markov chains. Stat. Probab. Lett. **83**, 1711-1719 (2013)

13. Pardoux, E, Peng, S: Backward doubly stochastic differential equations and systems of quasilinear SPDEs. Probab. Theory Relat. Fields **98**(2), 209-227 (1994)
14. Yin, G, Zhang, Q: Continuous-Time Markov Chains and Applications: A Singular Perturbation Approach. Springer, Berlin (1997)
15. Yin, G, Zhang, Q: Continuous-Time Markov Chains and Applications: A Two-Time-Scale Approach. Springer, Berlin (2013)
16. Amendinger, J: Martingale representation theorems for initially enlarged filtrations. Stoch. Process. Appl. **89**, 101-116 (2000)
17. Meyer, P, Zheng, W: Tightness criteria for laws of semimartingales. Ann. Inst. Henri Poincaré Probab. Stat. **20**(4), 353-372 (1984)
18. Bahlali, K, Elouaflin, A, Pardoux, E: Homogenization of semilinear PDEs with discontinuous averaged coefficients. Electron. J. Probab. **14**(18), 477-499 (2009)
19. Pardoux, E, Peng, S: Backward stochastic differential equations and quasilinear parabolic partial differential equations. In: Stochastic Partial Differential Equations and Their Applications. Lecture Notes in Control and Inform. Sci., vol. 176, pp. 200-217 (1992)
20. Barles, G, Buckdahn, R, Pardoux, E: Backward stochastic differential equations and integral-partial differential equations. Stoch. Stoch. Rep. **60**, 57-83 (1997)
21. Pardoux, E, Sow, AB: Homogenization of a periodic degenerate semilinear elliptic PDE. Stoch. Dyn. **11**, 475-493 (2011)
22. Zhang, Q, Yin, G: Nearly-optimal asset allocation in hybrid stock investment models. J. Optim. Theory Appl. **121**(2), 419-444 (2004)