Backward Stochastic Differential Equations with Markov Chains and Associated PDEs

Ning Ma[a], Zhen Wu[a,*], Huaizhong Zhao[b,c]

[a] *School of Mathematics, Shandong University, Jinan, 250100, China*
[b] *Department of Mathematical Sciences, Durham University, DH1 3LE, UK*
[c] *Research Centre for Mathematics and Interdisciplinary Sciences, Shandong University, Qingdao, 266237, China*

Abstract

In this paper we study the regularity of the solutions for backward stochastic differential equations (BSDEs) with finite state Markov chains and establish its link with associated partial differential equations (PDEs) in classical sense. Moreover, we study the existence and uniqueness of solutions for such BSDEs under Lipschitz conditions on f in the space $L^2_\rho(\mathbb{R}^d;\mathbb{R}^k)\otimes L^2_\rho(\mathbb{R}^d;\mathbb{R}^{d\times k})\otimes L^2_\rho(\mathbb{R}^d\times I;\mathbb{R}^k)$. In this way, we establish a new connection between $L^2_\rho(\mathbb{R}^d;\mathbb{R}^k)\otimes L^2_\rho(\mathbb{R}^d;\mathbb{R}^{d\times k})\otimes L^2_\rho(\mathbb{R}^d\times I;\mathbb{R}^k)$ valued solutions of BSDEs and the solutions of PDEs in a Sobolev space.

Keywords: Backward stochastic differential equations, Partial differential equations, Markov chains, Sobolev weak solutions

---

1. Introduction

Backward stochastic differential equations (BSDEs) were first introduced by Pardoux and Peng [1]. Since then many progresses have been made in fundamental research, for example, Peng and Wu [2], Feng, Wang and Zhao [3] studied the fully coupled forward and backward stochastic differential equations (FBSDEs). Furthermore, BSDEs have deep connections with PDEs and much influence in stochastic controls (e.g.:[4, 5]) and mathematical finance (e.g.:[6]).

*Corresponding author
Email address: `wuzhen@sdu.edu.cn` (Zhen Wu)
URL: `maning1990sdu@163.com` (Ning Ma), `huaizhong.zhao@durham.ac.uk` (Huaizhong Zhao)

It is remarkable that, BSDEs can give probabilistic interpretation for the quasi-linear and semi-linear partial differential equations (PDEs). For PDEs with smooth coefficients, Peng [7] gave a probabilistic interpretation for systems of quasilinear PDEs in classical sense and then Pardoux and Peng [8] studied both the classical and the viscosity solutions of such PDEs. Barles et al. [9] studied the viscosity solution of a kind of integral-partial differential equations by introducing BSDEs with jumps. After that, Barles and Lesigne [10] proved that the same probabilistic interpretation holds for variational formulation of the PDEs. Following the same line, Bally and Matoussi [11] studied the weak solution of Backward Doubly Stochastic Differential Equations (BDSDEs) in Sobolev space, and gave the probability interpretation for the corresponding PDEs in Sobolev space. Since then, there has been a lot of research on the weak solution of PDEs in Sobolev space. Feng, Wang and Zhao [3] studied the weak solutions of coupled forward-backward stochastic differential equations and the associated quasi-linear PDEs. For the first time, Wei, Wu and Zhao [5] studied the Sobolev weak solutions of Hamilton-Jacobi-Bellman equations corresponding to stochastic recursive control problems.

However, Brownian motion alone can not provide a good description of random phenomena in reality, such as the jump phenomenon in financial markets. In order to satisfy the need of more realistic models, we introduce Markov chains in the study of BSDEs, which can better reflect random environment and has a strong application significance. For example, the applications of the regime-switching model in finance have received significant attention in recent years. It modulates the system with a continuous-time finite-state Markov chain with each state representing a regime of the system or level of economic indicator, which depends on the market mode that switches among finite number states. The market mode could reflect the state of the underlying economy, the general mood of investors in the market, and other economic factors.

There are also a lot of research on the FBSDEs with Markov chains. Cohen [12, 13] studied BSDEs driven only by Markov chains. Wu and Tao [14] studied BSDEs driven by Markov chains and Brownian motion (the coefficient f does not depend on Z), and the viscosity solution to the associated PDEs. This article mainly studies the backward stochastic differential equations driven by Brownian motion and Markov chains (f depend on Z), and the probabilistic interpretation of their corresponding PDEs. When f depends on Z, due to the addition of the Markov chain, the traditional Malliavin analysis is invalid to obtain the representation of Z. We approached this difficulty innovatively through an approximation method. As far as we know, it is the first time to study the smoothness of solutions to the BSDEs with Markov chains, and give the classical solutions of PDEs with smooth coefficients. In the studies of weak solutions in Sobolev space, the main method is based on the stochastic flow theory established by Kunita [15] where the flow is generated by SDEs with smooth coefficient. In our problem, the coefficients of SDEs contain Markov chains, so we generalize the stochastic flow theory, and based on this, we study the BSDEs with Markov chains and the weak solution of the associated PDEs in Sobolev space.

The paper is organized as follows. In Section 2, we discuss the smoothness of the solution of SDE with Markov chains, and prove that the solution is a C^1-diffeomorphism, and give the general equivalent norm theorem. Moreover, we study the smoothness of the solution of BSDEs with Markov chains. In Section 3, the classical solutions of PDEs under the smooth coefficients are studied. In Section 4, we prove the existence and uniqueness of $L^2_\rho(\mathbb{R}^d;\mathbb{R}^k)\otimes L^2_\rho(\mathbb{R}^d;\mathbb{R}^{d\times k})\otimes L^2_\rho(\mathbb{R}^d\times I;\mathbb{R}^k)$ valued solutions of BSDEs with Markov chains under a functional Lipschitz condition. In Section 5, we study the weak solutions to the associated PDEs in Sobolev space.

2. Preliminaries

2.1. SDEs with finite-state Markov chains

Let $(\Omega,\mathscr{F},\mathbb{P})$ be a probability space, $\{B_t, 0\le t\le T\}$ be a d-dimensional Brownian motion and $\{\alpha_t, 0\le t\le T\}$ a finite-state Markov chain with the state space $I=\{1,2,\dots,m\}$, for some positive integer m. The transition intensities are $\lambda_{ij}(t)$ for $i\neq j$ with λ_{ij} are nonnegative and bounded and $\lambda_{ii}=-\sum_{j\in I\setminus\{i\}}\lambda_{ij}$. Let $\mathscr{F}=(\mathscr{F}_t)_{t\in[0,T]}$ be the filtration generated by $\{B_s,\alpha_s; 0\le s\le T\}$ and $\mathscr{F}_0$ contains all $\mathbb{P}$-null elements of $\mathscr{F}$.

For any $0\le t\le s\le T$, $p\ge 1, k\in\mathbb{N}$, we denote by $M^p([t,s];\mathbb{R}^k)$(resp. $S^p([t,s];\mathbb{R}^k)$) the set of $\mathscr{F}_s$-progressively measurable (resp. predictable) process such that for any

$\varphi_s \in M^p([t,s];\mathbb{R}^k)$(resp. $\in S^p([t,s];\mathbb{R}^k)$),

$$E\Big[\Big(\int_t^s |\varphi_r|^2 dr\Big)^{p/2}\Big] < +\infty \quad (\text{resp. } E \sup_{t\le r\le s} |\varphi_r|^p < +\infty).$$

By $C^k(\mathbb{R}^p;\mathbb{R}^q)$, $C^k_{l,b}(\mathbb{R}^p;\mathbb{R}^q)$, $C^k_p(\mathbb{R}^p;\mathbb{R}^q)$, we denote respectively the set of functions C^k from $\mathbb{R}^p$ into $\mathbb{R}^q$, the set of those functions of class C^k whose partial derivatives of order less than or equal to k are bounded (and hence the function itself growths at most like a linear function of the variable x at infinity), and the set of those 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. Moreover, by $C^k_b(\mathbb{R}^p;\mathbb{R}^q)$ we denote the set of those functions in $C^k_{l,b}(\mathbb{R}^p;\mathbb{R}^q)$ which are bounded.

Let $\mathscr{V}_t(j)$ denote the number of jumps of $\{\alpha_s, 0\le s\le T\}$ from any state in I to state j between time 0 and t and let $\mathscr{V}$ denote the corresponding integer-valued random measure on $([0,T]\times I, \mathscr{B}([0,T]\otimes\mathscr{B}_I))$. The compensator of $\mathscr{V}_t(j)$ is given by $1_{\alpha_{t-}\neq j}\lambda_{\alpha_{t-},j}dt$, i.e.,

$$d\widetilde{\mathscr{V}}_t(j) \triangleq d\mathscr{V}_t(j) - 1_{\alpha_{t-}\neq j}\lambda_{\alpha_{t-},j}dt$$

is a martingale (compensated measure). We set $\lambda_t(j) = 1_{\alpha_{t-}\neq j}\lambda_{\alpha_{t-},j}$. Then the canonical special semimartingale representation for α (see [16, 17]) is given by

$$d\alpha_t = \sum_{j\in I}\lambda_{\alpha_{t-},j}(t)(j-\alpha_{t-})dt + \sum_{j\in I}(j-\alpha_{t-})d\widetilde{\mathscr{V}}_t(j).$$

We will study the following SDEs and then give the well-known result of existence and uniqueness of solution:

$$\begin{cases} dX_s^{t,x,i} = b(s, X_s^{t,x,i}, \alpha_s^{t,i})ds + \sigma(s, X_s^{t,x,i}, \alpha_s^{t,i})dB_s, \quad s\in[t,T], \\ X_t^{t,x,i} = x, \quad \alpha_t^{t,i} = i, \quad x\in\mathbb{R}^d, \quad i\in I, \end{cases} \tag{2.1}$$

where $b(s,\cdot,i)\in C^3_{l,b}(\mathbb{R}^d;\mathbb{R}^d)$ and $\sigma(s,\cdot,i)\in C^3_{l,b}(\mathbb{R}^d;\mathbb{R}^{d\times d})$ for any $s\in[t,T]$ and $i\in I$. Moreover $b(s,x,i)$ and $\sigma(s,x,i)$ are continuous with $s\in[t,T]$ for any $x\in\mathbb{R}^d$ and $i\in I$. We denote by $\{X_s^{t,x,i}, t\le s\le T\}$ the unique strong solution of the following SDE: From [18] we know that the SDE (2.1) has a unique continuous solution in $S^2([t,T];\mathbb{R}^d)$. Next, we will study the regularity of the random field defined by $\{X_s^{t,x,i}; 0\le t\le s\le T, x\in\mathbb{R}^d\}$ in (t,s,x), for any $i\in I$.

By virtue of the same procedure as in the proof of Theorem 2.1 in Chapter 2 of Kunita [15], we give the following lemma without proof.

Lemma 2.1. *For any $p \geq 2$ and $q \in \mathbb{R}$, there exists a constant c such that for any $t \in \mathbb{R}$, x, $x' \in \mathbb{R}^d$, $i \in I$, $\varepsilon > 0$,*

$$E(\sup_{t \leq s \leq T} |X_s^{t,x,i}|^p) \leq c(1+|x|^p), \tag{2.2}$$

$$E(|X_s^{t,x,i} - X_{s'}^{t,x,i}|^p) \leq c(1+|x|^p)|s-s'|^{\frac{p}{2}}, \tag{2.3}$$

$$E[(\varepsilon + |X_s^{t,x,i}|^2)^q] \leq c(\varepsilon + |x|^2)^q, \tag{2.4}$$

$$E[(\varepsilon + |X_s^{t,x,i} - X_s^{t,y,i}|^2)^q] \leq c(\varepsilon + |x-y|^2)^q.$$

Theorem 2.2. *For any $p \geq 2$, there is a positive constant C such that*

$$E|X_s^{t,x,i} - X_{s'}^{t',x',i}|^p \leq C\Big[|x-x'|^p + (1+|x|^p+|x'|^p)(|t-t'| + o(|t-t'|) + |s-s'|^{\frac{p}{2}})\Big].$$

Proof. The proof is a combination of Theorem 2.1 in [15] and Lemma 3.3 in [18], which is easy to verify, so we omitted. □

We get the following corollary immediately from Kolmogorov's continuous Lemma:

Corollary 2.3. *The random field defined by $\{X_s^{t,x,i}, t \leq s \leq T, x \in \mathbb{R}^d\}$ is continuous in (s,x), and the function $\phi_i(t,s,x) := E[X_s^{t,x,i}]$ is continuous in (t,s,x).*

Lemma 2.4. *For any $p > 2$, there is a positive constant c such that*

$$E\Big(\sup_{t \leq s \leq T} |\Delta_h^l X_s^{t,x,i}|^p\Big) \leq c, \tag{2.5}$$

and

$$\begin{aligned} &E|\Delta_h^l X_s^{t,x,i} - \Delta_{h'}^l X_{s'}^{t',x',i}|^p \\ &\leq c\Big[|x-x'|^p + |h-h'|^p + (1+|x|+|x'|)^p(|s-s'|^{\frac{p}{2}} + |t-t'|^{\frac{1}{2}} + o(t-t')^{\frac{1}{2}})\Big], \end{aligned} \tag{2.6}$$

where $h \in \mathbb{R} - \{0\}$ and $\Delta_h^l g(x) := h^{-1}[g(x+he_l) - g(x)]$, $1 \leq l \leq d$, e_l denotes the l-th vector of an arbitrary orthonormal basis of $\mathbb{R}^d$.

Proof. We first show the boundedness of $E|\Delta_h^l X_s^{t,x,i}|^p$. By the mean value theorem, it holds that

$$\begin{aligned}\Delta_h^l X_s^{t,x,i} =& e_l + \int_t^s \int_0^1 b'(r, X_r^{t,x,i} + v(X_r^{t,x+he_l,i} - X_r^{t,x,i}), \alpha_r^{t,i}) \Delta_h^l X_r^{t,x,i} dv dr \\ &+ \int_t^s \int_0^1 \sigma'(r, X_r^{t,x,i} + v(X_r^{t,x+he_l,i} - X_r^{t,x,i}), \alpha_r^{t,i}) \Delta_h^l X_r^{t,x,i} dv dB_r.\end{aligned} \tag{2.7}$$

where b' and σ' are the first order partial derivatives in x of b and σ respectively. By the boundedness of b', σ' and Burkholder's inequality, we have

$$\begin{aligned}E|\Delta_h^l X_s^{t,x,i}|^p \leq& CE\Big[1 + \int_t^s |\int_0^1 b'(r, X_r^{t,x,i} + v(X_r^{t,x+he_l,i} - X_r^{t,x,i}), \alpha_r^{t,i}) dv \Delta_h^l X_r^{t,x,i}|^p dr \\ &+ \int_t^s |\int_0^1 \sigma'(r, X_r^{t,x,i} + v(X_r^{t,x+he_l,i} - X_r^{t,x,i}), \alpha_r^{t,i}) dv \Delta_h^l X_r^{t,x,i}|^p dr\Big] \\ \leq& C + C\int_t^s E|\Delta_h^l X_r^{t,x,i}|^p dr.\end{aligned}$$

Therefore by Gronwall's inequality, we see that $E|\Delta_h^l X_s^{t,x,i}|^p$ is bounded. By a standard argument of Burkholder's inequality, we get (2.5).

To prove (2.6) we discuss firstly the case when $s = s'$. Without loss of generality, we assume $t < t' < s$. Let $g = b; \sigma$, set

$$\begin{aligned}g_1'(r) &= g'(r, X_r^{t,x,i} + v(X_r^{t,x+he_l,i} - X_r^{t,x,i}), \alpha_r^{t,i}), \\ g_2'(r) &= g'(r, X_r^{t',x',i} + v(X_r^{t',x'+h'e_l,i} - X_r^{t',x',i}), \alpha_r^{t',i}), \\ g_3'(r) &= g'(r, X_r^{t',x',i} + v(X_r^{t',x'+h'e_l,i} - X_r^{t',x',i}), \alpha_r^{t,i}).\end{aligned}$$

Then we have

$$\begin{aligned}&E|\Delta_h^l X_s^{t,x,i} - \Delta_{h'}^l X_s^{t',x',i}|^p \\ &\leq E|\int_t^{t'} \int_0^1 b_1'(r) dv \Delta_h^l X_r^{t,x,i} dr + \int_t^{t'} \int_0^1 \sigma_1'(r) dv \Delta_h^l X_r^{t,x,i} dB_r|^p \\ &+ E|\int_{t'}^s |\int_0^1 b_1'(r) dv \Delta_h^l X_r^{t,x,i} - \int_0^1 b_2'(r) dv \Delta_{h'}^l X_r^{t',x',i}| dr|^p \\ &+ E|\int_{t'}^s (\int_0^1 \sigma_1'(r) dv \Delta_h^l X_r^{t,x,i} - \int_0^1 \sigma_2'(r) dv \Delta_{h'}^l X_r^{t',x',i}) dB_r|^p \\ &= I + II + III.\end{aligned}$$

From the boundedness of b' and σ' and $E|\Delta_h^l X_s^{t,x,i}|^p$, it is easy to get

$$I \leq C|t - t'|^{\frac{p}{2}}.$$

With the continuity and boundedness of b' and σ', by Burkholder's inequality,

$$
\begin{aligned}
III \leq & CE\Big|\int_{t'}^{s}\Big|\int_0^1 \sigma_1'(r)dv\Delta_h^l X_r^{t,x,i}-\int_0^1 \sigma_2'(r)dv\Delta_{h'}^l X_r^{t',x',i}\Big|^2 dr\Big|^{p/2}\\
\leq & CE\Big|\int_{t'}^{s}\Big[\int_0^1|\sigma_1'(r)|dv|\Delta_h^l X_r^{t,x,i}-\Delta_{h'}^l X_r^{t',x',i}|\\
&+\int_0^1|\sigma_1'(r)-\sigma_3'(r)|dv|\Delta_{h'}^l X_r^{t',x',i}|\Big]|^p dr\Big|\\
&+CE\Big|\int_{t'}^{s}\Big[\int_0^1|\sigma_2'(r)-\sigma_3'(r)|dv|\Delta_{h'}^l X_r^{t',x',i}|\Big]|^p dr\Big|\\
\leq & CE\int_{t'}^{s}|\Delta_h^l X_r^{t,x,i}-\Delta_{h'}^l X_r^{t',x',i}|^p dr\\
&+C\int_{t'}^{s}E[|X_r^{t,x,i}-X_r^{t',x',i}|^{2p}]^{\frac{1}{2}}E[|\Delta_{h'}^l X_r^{t',x',i}|^{2p}]^{\frac{1}{2}}dr\\
&+C\int_{t'}^{s}E[|X_r^{t,x+he_l,i}-X_r^{t',x'+h'e_l,i}|^{2p}]^{\frac{1}{2}}E[|\Delta_{h'}^l X_r^{t',x',i}|^{2p}]^{\frac{1}{2}}dr\\
&+C\int_{t'}^{s}E[(\int_0^1|\sigma_2'(r)|+|\sigma_3'(r)|dv)^{2p}I_{\{\alpha_r^{t,i}\neq\alpha_r^{t',i}\}}]^{\frac{1}{2}}E[|\Delta_{h'}^l X_r^{t',x',i}|^{2p}]^{\frac{1}{2}}dr.
\end{aligned}
$$

We define Θ as the last term on the right hand side of the above inequality. Then by the same discussion of (3.74) on page 107 in [18]:

$$
\begin{aligned}
\Theta &\leq C\int_{t'}^{s}E\Big[\big[(\int_0^1|^2\sigma'(r)|+|^3\sigma'(r)|dv)^{2p}\big]I_{\{\alpha_r^{t,i}\neq\alpha_r^{t',i}\}}\Big]^{\frac{1}{2}}dr\\
&\leq CE\big[E[I_{\{\alpha_r^{t,i}\neq\alpha_r^{t',i}\}}|\alpha_r^{t,i}]\big]^{\frac{1}{2}}\leq C|t-t'|^{\frac{1}{2}}+o(|t-t'|)^{\frac{1}{2}}.
\end{aligned}
$$

By Theorem 2.2, we have

$$
\begin{aligned}
III \leq & C\big[|x-x'|^p+|h-h'|^p+(1+|x|^p+|x'|^p)(|t-t'|^{\frac{1}{2}}+o(|t-t'|)^{\frac{1}{2}}\big]\\
&+C\int_{t'}^{s}E|\Delta_h^l X_r^{t,x,i}-\Delta_{h'}^l X_r^{t',x',i}|^p dr.
\end{aligned}
$$

By the same discussion of III, we have

$$
\begin{aligned}
II \leq & C\big[|x-x'|^p+|h-h'|^p+(1+|x|^p+|x'|^p)(|t-t'|^{\frac{1}{2}}+o(|t-t'|)^{\frac{1}{2}}\big]\\
&+C\int_{t'}^{s}E|\Delta_h^l X_r^{t,x,i}-\Delta_{h'}^l X_r^{t',x',i}|^p dr.
\end{aligned}
$$

Combining I,II and III, by Gronwall's inequality, we have

$$
\begin{aligned}
&E|\Delta_h^l X_s^{t,x,i}-\Delta_{h'}^l X_s^{t',x',i}|^p\\
&\leq c\Big[|x-x'|^p+|h-h'|^p+(1+|x|+|x'|)^p(|t-t'|^{\frac{1}{2}}+o(|t-t'|)^{\frac{1}{2}})\Big].
\end{aligned}
$$

It remains to prove (2.6) in case $s \neq s'$. Assuming $s < s'$, we have

$$\begin{aligned}
&\Delta_h^l X_s^{t,x,i} - \Delta_{h'}^l X_{s'}^{t',x',i} = \Delta_h^l X_s^{t,x,i} - \Delta_{h'}^l X_s^{t',x',i} \\
&- \int_s^{s'} \int_0^1 {}^2 b'(r) dv \Delta_h^l X_r^{t,x,i} dr - \int_t^{t'} \int_0^1 {}^2 \sigma'(r) dv \Delta_h^l X_r^{t,x,i} dB_r.
\end{aligned}$$

Using (2.4) and standard arguments, together with the result we already proved when $s = s'$, we can easily see (2.6) holds. $\square$

From Lemma 2.4, let $h \to 0$ and using Kolmogorov's Lemma, we obtain the next theorem immediately.

Theorem 2.5. *For any $t \in [0,T]$, $x \in \mathbb{R}^d$ and $i \in I$ the mapping $x \mapsto X_s^{t,x,i}$ is a.s. differentiable, and the matrix of partial derivatives $\nabla X_s^{t,x,i} (\nabla X_s^{t,x,i} = (\frac{\partial X_s}{\partial x^j})_{1 \leq i,j \leq d}; t \leq s \leq T)$ possesses a version which is a.s. continuous in (s,x). Moreover the process solves the following SDE:*

$$\nabla X_s^{t,x,i} = I + \int_t^s b'(r, X_r^{t,x,i}, \alpha_r^{t,i}) \nabla X_r^{t,x,i} dr + \int_t^s \sigma'(r, X_r^{t,x,i}, \alpha_r^{t,i}) \nabla X_r^{t,x,i} dB_r, \tag{2.8}$$

where b' and σ' are matrix valued function $(\frac{\partial b_i}{\partial x_j})_{1 \leq i,j \leq d}$.

Now we consider the solution of the SDEs (2.1) as a stochastic flow from $\mathbb{R}^d$ to $\mathbb{R}^d$. According to Section 4, Chapter II in [15], Lemma 2.1 and Theorem 2.2, we have the following theorem.

Theorem 2.6. *The map $X_s^{t,\cdot,i} : \mathbb{R}^d \to \mathbb{R}^d$ is a homeomorphism a.s. This is to say that the map $X_s^{t,\cdot,i}$ is one-to-one and onto, so its inverse map exists. Moreover, the inverse map, denoted by $\hat{X}_s^{t,\cdot,i} : \mathbb{R}^d \to \mathbb{R}^d$, is also continuous in s a.s.*

Lemma 2.7. *Let $g(r,x,i)$ be a continuous function of (r,x), $f(r,x,i)$ be a C^1 function of x for any $i \in I$. Then*

$$\int_t^s g(r, X_r^{t,y,i}, \alpha_r^{t,i}) dr |_{y = \hat{X}_s^{t,x,i}} = \int_t^s g(r, \hat{X}_s^{r,x,\alpha_r^{t,i}}, \alpha_r^{t,i}) dr, \tag{2.9}$$

$$\begin{aligned}
\int_t^s f(r, X_r^{t,y,i}, \alpha_r^{t,i}) dB_r |_{y = \hat{X}_s^{t,x,i}} = &\int_t^s f(r, \hat{X}_s^{r,x,\alpha_r^{t,i}}, \alpha_r^{t,i}) d\hat{B}_r \\
&- \int_t^s (\sigma^\intercal f')(r, \hat{X}_s^{r,x,\alpha_r^{t,i}}, \alpha_r^{t,i}).
\end{aligned} \tag{2.10}$$

Proof. The equation (2.9) is obviously. Now let's prove (2.10). We first assume that $f(s,x,i)$ is a C^1 function of s and a C^2 function of x for any $i \in I$. Define $\Delta^n = \{t = t_1^n < t_2^n < \cdots < t_n^n = s\}$ as a sequence of partitions of $[t,s]$. Then for any $r \in [\tau_{k-1}, \tau_k)$, we have by Corollary 7.8, Chapter I in [15], there is a sequence of partitions $\{\Delta_n\}$ with $|\Delta_n| \to 0$ such that

$$\int_{\tau_{k-1}}^{\tau_k} f(r, X_r^{t,y,i}, \alpha_r^{t,i}) dB_r = \lim_{n\to+\infty} \int_{\tau_{k-1}}^{\tau_k} f^{\Delta_n}(r, X_r^{t,y,i}, \alpha_r^{t,i}) I_{[\tau_{k-1},\tau_k)} dB_r \quad \text{a.s.},$$

where

$$f_r^{\Delta_n} = \sum_{k=1}^{n-1} f_{t_k^n} 1_{[t_k^n, t_{k+1}^n)}.$$

Therefore

$$\int_t^s f(r, X_r^{t,y,i}, \alpha_r^{t,i}) dB_r = \sum_{k=1}^{n-1} \int_{\tau_{k-1}}^{\tau_k} f(r, X_r^{t,y,i}, \alpha_r^{t,i}) dB_r = \lim_{n\to+\infty} \int_t^s f^{\Delta_n}(r, X_r^{t,y,i}, \alpha_r^{t,i}) dB_r \quad \text{a.s.}$$

Then it holds that

$$\begin{aligned}
&\int_t^s f(r, X_r^{t,y,i}, \alpha_r^{t,i}) dB_r |_{y=\hat{X}_s^{t,x,i}} \\
&= \lim_{n\to+\infty} \sum_{j=1}^{n-1} f(t_j^n, \hat{X}_s^{t_j^n, x, \alpha_{t_j^n}^{t,i}}, \alpha_{\tau_{k-1}}^{t,i})(B_{t_{j+1}^n} - B_{t_j^n}) \\
&= \lim_{n\to+\infty} \sum_{j=1}^{n-1} f(t_{j+1}^n, \hat{X}_s^{t_{j+1}^n, x, \alpha_{t_{j+1}^n}^{t,i}}, \alpha_{\tau_{k-1}}^{t,i})(B_{t_{j+1}^n} - B_{t_j^n}) \\
&\quad - \lim_{n\to+\infty} \sum_{j=1}^{n-1} \left[f(t_{j+1}^n, \hat{X}_s^{t_{j+1}^n, x, \alpha_{t_{j+1}^n}^{t,i}}, \alpha_{\tau_{k-1}}^{t,i}) - f(t_j^n, \hat{X}_s^{t_j^n, x, \alpha_{t_j^n}^{t,i}}, \alpha_{\tau_{k-1}}^{t,i}) \right] (B_{t_{j+1}^n} - B_{t_j^n}).
\end{aligned} \tag{2.11}$$

The first limit on the right hand side exists and equals to the backward Itô integral $\int_t^s g(r, \hat{X}_s^{r,x,\alpha_r^{t,i}}, \alpha_r^{t,i}) d\hat{B}_r$. The second limit equals to

$$< f(\cdot, X_\cdot^{t,y,i}, \alpha_\cdot^{t,i}), B_\cdot - B_t >_s |_{y=\hat{X}_s^{t,x,i}},$$

where the symbol $< \cdot, \cdot >$ is the joint quadratic variation. By Itô's formula, we have

$$\begin{aligned}
f(s, X_s^{t,y,i}, \alpha_s^{t,i}) =& f(t,y,i) + N_s + \int_t^s (\sigma^\intercal f')(r, X_r^{t,y,i}, \alpha_r^{t,i}) dB_r \\
&+ \sum_{j=1}^N \int_t^s (f(r, X_r^{t,y,i}, j) - f(r, X_r^{t,y,i}, \alpha_{r-}^{t,i})) d\widetilde{V}_r(j),
\end{aligned}$$

where

$$N_s = \int_t^s \Big[\frac{\partial f}{\partial r}(r, X_r^{t,y,i}, \alpha_r^{t,i}) + (bf')(r, X_r^{t,y,i}, \alpha_r^{t,i}) + \frac{1}{2}(\sigma^\intercal f''\sigma)(r, X_r^{t,y,i}, \alpha_r^{t,i})$$
$$+ \sum_{j=1}^N \gamma_{\alpha_r^{t,x},j}\sigma(r, X_r^{t,y,i}, \alpha_r^{t,i})\Big]dr$$

is a process of bounded variation. Therefore, we have

$$< f(\cdot, X_\cdot^{t,y,i}, \alpha_\cdot^{t,i}), B_\cdot - B_t >_s = \int_t^s (\sigma^\intercal f')(r, X_r^{t,y,i}, \alpha_r^{t,i})dr.$$

Substitute $y = \hat{X}_s^{t,x,i}$ to the above and apply (2.9), we get that the second term of the right hand side of (2.11) equals to

$$-\int_t^s (\sigma^\intercal f')(r, \hat{X}_s^{r,x,\alpha_r^{t,i}}, \alpha_r^{t,i})dr.$$

It remains to prove (2.10) for general f. For each $i \in I$, choose a sequence of smooth functions $f_n(r,x,i)$ such that $f_n(r,x,i) \to f(r,x,i)$ and $f_n'(r,x,i) \to f'(r,x,i)$ locally uniformly. Then Eq. (2.10) is valid to all f_n. Let $n \to +\infty$, we get Eq. (2.10) for f by virtue of Theorem 7.7 in Chapter I in [15]. $\square$

Remark 2.8. *In the above proof, we use the backward Itô integral* $\int_t^s g(r, \hat{X}_s^{r,x,\alpha_r^{t,i}}, \alpha_r^{t,i})d\hat{B}_r$ *and it is well defined. If we set* $G_r = F_{r,s}^B \vee F_{t,r}^\alpha$*, then for any* $r \in [t,s]$, $g(r, \hat{X}_s^{r,x,\alpha_r^{t,i}}, \alpha_r^{t,i})$ *is* G_r*-adapted and the following still holds for any* G_r*-adapted square integrable process* $g(t)$*,*

$$E[\int_t^s g(r)d\hat{B}_r] = 0, \quad E[|\int_t^s g(r)d\hat{B}_r|^2] = E[\int_t^s g(r)^2 dr].$$

Remark 2.9. *By Lemma 2.7, it's easy to get the inverse of the flow* $\{\hat{X}_s^{t,x,i}, t \le s \le T\}$ *satisfying the following backward SDEs:*

$$\hat{X}_s^{t,x,i} = x - \int_t^s \sigma(r, \hat{X}_s^{r,x,\alpha_r^{t,i}}, \alpha_r^{t,i})d\hat{B}_r + \int_t^s (\sigma^\intercal \sigma')(r, \hat{X}_s^{r,x,\alpha_r^{t,i}}, \alpha_r^{t,i})dr$$
$$- \int_t^s b(r, \hat{X}_s^{r,x,\alpha_r^{t,i}}, \alpha_r^{t,i})dr.$$

Using the same procedure as in the proof of Lemma 2.4, we can prove that the inverse flow $\hat{X}_s^{t,\cdot,i}$ *is differentiable, and the derivative* $\nabla\hat{X}_s^{t,\cdot,i}$ *is continuous in s. So* $X_s^{t,\cdot,i}$ *defines a* C^1*-diffeomorphism. We denote by* $J(\hat{X}_s^{t,\cdot,i})$ *the determinant of Jacobian matrix of* $\hat{X}_s^{t,\cdot,i}$*, which is positive because it is a continuous function of* $s \in [t,T]$*, which does not vanish and* $J(\hat{X}_t^{t,\cdot,i}) = 1$*.*

Lemma 2.10 (Generalized equivalence of norm principle)**.** *We take $\rho(x) := e^{F(x)}$ as the weight function, where $F : \mathbb{R}^d \to \mathbb{R}$ is a continuous function. Moreover, we assume that there exists a constant $R > 0$ such that for $|x| > R$, $F \in C^2_{l,b}(\mathbb{R}^d;\mathbb{R})$ and $\sup_{x\in\mathbb{R}^d} |F'(x)x| < +\infty$. For instance, we can take $\rho(x) = (1+|x|)^q$, with $q \in \mathbb{R}$ or $\rho(x) = e^{\frac{\alpha}{1+|x|}}$ with $\alpha \in \mathbb{R}$. If $\varphi\rho^{-1} \in L^1(\mathbb{R}^d)$. Then there exist two constants $c > 0$ and $C > 0$ such that*

$$c\int_{\mathbb{R}^d} |\varphi(x)|\rho^{-1}(x)dx \le E\Big[\int_{\mathbb{R}^d} |\varphi(X_s^{t,x,i})|\rho^{-1}(x)dx\Big] \le C\int_{\mathbb{R}^d} |\varphi(x)|\rho^{-1}(x)dx. \quad (2.12)$$

Moreover if $\Psi : \Omega \times [t,T] \times \mathbb{R}^d \to \mathbb{R}$, $\Psi(s,\cdot)$ is F_s^α measurable for $s \in [t,T]$ and $\Psi\rho^{-1} \in L^1(\Omega \times [0,T] \times \mathbb{R}^d)$, then there exist two constants $c > 0$ and $C > 0$ such that

$$\begin{aligned} cE\int_t^T \int_{\mathbb{R}^d} |\Psi(s,x)|\rho(x)^{-1}dx \le E\int_t^T \int_{\mathbb{R}^d} |\Psi(s,X_s^{t,x,i})|\rho(x)^{-1}dx \\ \le CE\int_t^T \int_{\mathbb{R}^d} |\Psi(s,x)|\rho(x)^{-1}dx. \end{aligned} \quad (2.13)$$

The constants c and C depend on T, ρ, the bounds of σ and the bounds of the first (resp. first and second) derivatives of b (resp. of σ).

Proof. We only need to verify the following inequality:

$$c \le E\Big[\frac{J(\hat{X}_s^{t,x,i})\rho(\hat{X}_s^{t,x,i})}{\rho(x)}|F_s^\alpha\Big] \le C \quad a.s.. \quad (2.14)$$

In fact, using a change of variable $y = X_s^{t,x,i}$, we get

$$\begin{aligned} \int_{\mathbb{R}^d} E(|\Psi(s,X_s^{t,x,i})|)\rho(x)dx &= \int_{\mathbb{R}^d} E\Big[E\big[|\Psi(s,y)|J(\hat{X}_s^{t,y,i})\rho(\hat{X}_s^{t,y,i})|F_s^\alpha\big]\Big]dy \\ &= \int_{\mathbb{R}^d} E\Big[|\Psi(s,y)|\rho(y)E\big[\frac{J(\hat{X}_s^{t,y,i})\rho(\hat{X}_s^{t,y,i})}{\rho(y)}|F_s^\alpha\big]\Big]dy \end{aligned}$$

and if (2.14) holds, integrating with respect to $s \in [t,T]$, we get (2.13). Eq. (2.12) is a special case of (2.13) so Lemma 2.10 is proved. Now we devote to the proof of (2.14).

We assume first that $T-h \le t \le T$ for some small $h > 0$ and $F \in C^2_{l,b}(\mathbb{R}^d)$. Applying Lemma 2.7 and Itô's formula, we get the inverse of the flow $\{\hat{X}_s^{t,x,i}, t \le s \le T\}$ satisfies the following backward SDEs:

$$\begin{aligned} \hat{X}_s^{t,x,i} = x - \int_t^s \sigma(r,\hat{X}_s^{r,x,\alpha_r^{t,i}},\alpha_r^{t,i})d\hat{B}_r + \int_t^s (\sigma^T\sigma')(r,\hat{X}_s^{r,x,\alpha_r^{t,i}},\alpha_r^{t,x,i})dr \\ - \int_t^s b(r,\hat{X}_s^{r,x,\alpha_r^{t,i}},\alpha_r^{t,i})dr, \end{aligned}$$

and

$$F(\hat{X}_s^{t,x,\alpha_r^{t,i}}) = F(x) - \int_t^s \sigma(r,\hat{X}_s^{r,x,\alpha_r^{t,i}},\alpha_r^{t,i})F'(\hat{X}_s^{r,x,\alpha_r^{t,i}})d\hat{B}_r + \int_t^s \hat{L}F(\hat{X}_s^{r,x,\alpha_r^{t,i}})dr,$$

where

$$\hat{L}F(\hat{X}_s^{r,x,\alpha_r^{t,i}}) = \frac{1}{2}(\sigma^T\sigma)(r,\hat{X}_s^{r,x,\alpha_r^{t,i}},\alpha_r^{t,x,i})F''(\hat{X}_s^{r,x,\alpha_r^{t,i}}) - b(r,\hat{X}_s^{r,x,i},\alpha_r^{t,i})F'(\hat{X}_s^{r,x,i}) \\ + (\sigma\nabla\sigma)(r,\hat{X}_s^{r,x,\alpha_r^{t,i}},\alpha_r^{t,i})F'(\hat{X}_s^{r,x,i}).$$

It follows that

$$\begin{aligned}
\frac{\rho(\hat{X}_s^{r,x,\alpha_r^{t,i}})}{\rho(x)} &= \exp(F(\hat{X}_s^{t,x,i}) - F(x)) \\
&= \exp\Big(-\int_t^s \sigma(r,\hat{X}_s^{r,x,i},\alpha_r^{t,i})F'(\hat{X}_s^{r,x,i})d\hat{B}_r + \int_t^s \hat{L}F(\hat{X}_s^{r,x,i})dr\Big) \\
&= \exp\Big(-\int_t^s \sigma(r,\hat{X}_s^{r,x,i},\alpha_r^{t,i})F'(\hat{X}_s^{r,x,i})d\hat{B}_r \\
&\qquad + \frac{1}{2}\int_t^s |\sigma(r,\hat{X}_s^{r,x,i},\alpha_r^{t,i})F'(\hat{X}_s^{r,x,i})|^2 dr\Big) \\
&\quad \times \exp\Big(-\frac{1}{2}\int_t^s |\sigma(r,\hat{X}_s^{r,x,i},\alpha_r^{t,i})F'(\hat{X}_s^{r,x,i})|^2 dr + \int_t^s \hat{L}F(\hat{X}_s^{r,x,i})dr\Big) \\
&:= M_t^s(x,i)N_t^s(x,i).
\end{aligned}$$

Since the first and second order derivatives of F, σ, and $|F'(x)x|$ are bounded, we have

$$|b(r,\hat{X}_s^{r,x,i},\alpha_r^{t,i})F'(\hat{X}_s^{r,x,i})| \leq c_1|(1+|\hat{X}_s^{r,x,i}|)F'(\hat{X}_s^{r,x,i})| \leq c_2.$$

So, there exists two constants $r > 0$ and $R > 0$ such that $r \leq N_t^s(x,i) \leq R$. On the other hand, M_t^s satisfies the following linear backward SDE

$$M_t^s(t,i) = 1 - \int_t^s M_t^r(x,i)\sigma(r,\hat{X}_s^{r,x,i},\alpha_r^{t,i})F'(\hat{X}_s^{r,x,i})d\hat{B}_r := 1 + I_t^s(x,i),$$

and therefore

$$E[|M_t^s(x,i)|^2|F_s^\alpha] \leq 2(1 + E\big[|\int_t^s M_t^r(x,i)\sigma(r,\hat{X}_s^{r,x,i},\alpha_r^{t,i})F'(\hat{X}_s^{r,x,i})d\hat{B}_r|^2|F_s^\alpha\big]).$$

Due to the independence of the Brownin Motion B_t and the Markov chains α_t,

$$\begin{aligned}
E[|M_t^s(x,i)|^2|F_s^\alpha] &\leq C(1 + E\big[\int_t^s |M_t^r(x,i)\sigma(r,\hat{X}_s^{r,x,i},\alpha_r^{t,i})F'(\hat{X}_s^{r,x,i})|^2 dr|F_s^\alpha\big]) \\
&\leq C(1 + \int_t^s E[|M_t^r(x,i)|^2|F_s^\alpha]).
\end{aligned}$$

By the Gronwall inequality, we have

$$E[|M_t^s(x,i)|^2|F_s^{\alpha}] \le C \quad \text{a.s.},$$

and by a similar method,

$$E[|I_t^s(x,i)|^2|F_s^{\alpha}] \le C(s-t) \quad \text{a.s.}$$

By virtue of Remark 2.9, Proposition 5.1 in [11] and the same discussion as above, we get the result. $\square$

2.2. BSDEs with finite-state Markov chains

Let $f : [0,T]\times\mathbb{R}^d\times\mathbb{R}^k\times\mathbb{R}^{k\times d}\times I\to\mathbb{R}^k$ and $h:\mathbb{R}^d\to\mathbb{R}^k$. Moreover, we assume

(A.1) $h\in C_p^3(\mathbb{R}^d;\mathbb{R}^k)$;

(A.2) for any $s\in[t,T]$, $i\in I$, $(x,y,z)\to f(s,x,y,x,i)$ is of class C^3, $f(s,\cdot,0,0,i)\in C_p^3(\mathbb{R}^d;\mathbb{R}^k)$, for $s\in[t,T]$ and $i\in I$;

(A.3) the first order partial derivatives in y and z of f are bounded on $[t,T]\times\mathbb{R}^d\times\mathbb{R}^k\times\mathbb{R}^{k\times d}\times I$, as well as their derivatives of order one and two with respect to x,y,z.

From now on, we will denote by $N^p([t,T]\times I;\mathbb{R}^k)$, $p\ge 1$, the space consisting of $\mathscr{F}_s$-progressively measurable process, such that for any $\varphi\in N^p([t,T]\times I;\mathbb{R}^k)$,

$$E\Big[\Big(\sum_{j\in I}\int_t^T|\varphi_r(j)|^2\lambda_r(j)dr\Big)^{p/2}\Big]<+\infty.$$

From conditions (A.1)-(A.3), we know that $f(s,X_s^{t,x,i},y,z,i)$ is $\mathscr{F}_s$-adapted, $h(X_T^{t,x,i})$ is $\mathscr{F}_T$-measure and $E|g(X_T^{t,x,i})|^2<+\infty$. So, under condition (A.1)-(A.3), from [17] we know that there exists a unique tripe $(Y,Z,W)\in S^2([t,T];\mathbb{R}^k)\times M^2([t,T];\mathbb{R}^k)\times N^2([t,T]\times I;\mathbb{R}^k)$ which solves:

$$\begin{cases} dY_s^{t,x,i} = -f(s,X_s^{t,x,i},Y_s^{t,x,i},Z_s^{t,x,i},\alpha_s^{t,i})ds + Z_s^{t,x,i}dB_s + \sum\limits_{j\in I}W_s^{t,x,i}(j)d\tilde{\mathscr{V}}_s(j), \\ Y_T^{t,x,i} = h(X_T^{t,x,i}), \quad \alpha_t^{t,i}=i. \end{cases} \tag{2.15}$$

Proposition 2.11. *For any $0\le t\le T$, $x\in\mathbb{R}^d$, $p>1$, we have*

$$(Y_\cdot^{t,\cdot,i},Z_\cdot^{t,\cdot,i},W_\cdot^{t,\cdot,i}(\cdot))\in S^p([t,T];\mathbb{R}^k)\times M^p([t,T];\mathbb{R}^k)\times N^p([t,T]\times I;\mathbb{R}^k).$$

Proof. Let $\varphi_{n,p}(x) = (x\wedge n)^p + pn^{p-1}(x-n)^+$. Then $\varphi_{n,p} \in C^1(\mathbb{R}_+)$,

$$\varphi'_{n,p}(x) = p(x\wedge n)^{p-1}$$

is bounded and absolutely continuous in x and $\varphi''_{n,p}(x) = p(p-1)(x\wedge n)^{p-2}1_{[0,n]}(x)$.

Applying Itô's formula for semi-martingale (cf Theorem 5.1 in Chapter II, [19]) to $\varphi_{n,p}(|Y_s|^2)$, we get

$$\begin{aligned}
&\varphi_{n,p}(|Y_T|^2) + 2\int_s^T \varphi'_{n,p}(|Y_r|^2)Y_r f(r,X_r,Y_r,Z_r,\alpha_r)dr \\
=&\varphi_{n,p}(|Y_s|^2) + 2\int_s^T \varphi'_{n,p}(|Y_r|^2)Y_r Z_r dB_r + 2\int_s^T \varphi''_{n,p}(|Y_r|^2)|Y_r Z_r|^2 dr \\
&+\int_s^T \varphi'_{n,p} Tr(Z_r Z_r^T)dr + \sum_{j\in I}\int_s^T \Big[\varphi_{n,p}(|Y_{r-}+W_r(j)|^2) - \varphi_{n,p}(|Y_{r-}|^2)\Big] d\mathscr{V}_r(j) \\
&+\sum_{j\in I}\int_s^T \Big[\varphi_{n,p}(|Y_{r-}+W_r(j)|^2) - \varphi_{n,p}(|Y_{r-}|^2) - 2\varphi'_{n,p}(|Y_r|^2)Y_r W_r(j)\Big]\lambda_r(j)dr \\
=&\varphi_{n,p}(|Y_s|^2) + 2\int_s^T \varphi'_{n,p}(|Y_r|^2)Y_r Z_r dB_r + 2\int_s^T \varphi''_{n,p}(|Y_r|^2)|Y_r Z_r|^2 dr \\
&+\int_s^T \varphi'_{n,p} Tr(Z_r Z_r^T)dr + \sum_{j\in I}\int_s^T \Big[\varphi_{n,p}(|Y_{r-}+W_r(j)|^2) - \varphi_{n,p}(|Y_{r-}|^2)\Big] d\mathscr{V}_r(j) \\
&+\sum_{j\in I}\int_s^T \Big[\varphi_{n,p}(|Y_{r-}+W_r(j)|^2) - \varphi_{n,p}(|Y_{r-}|^2) \\
&-\varphi'_{n,p}(|Y_r|^2)(|Y_{r-}+W_r(j)|^2 - (|Y_{r-}|^2))\Big]\lambda_r(j)dr \\
&+\sum_{j\in I}\int_s^T \varphi'_{n,p}(|Y_r|^2)|W_r(j)|^2\lambda_r(j)dr.
\end{aligned} \tag{2.16}$$

Since $(Y_s, Z_s, W_s(j)) \in S^2([t,T];\mathbb{R}^k)\times M^2([t,T];\mathbb{R}^k)\times N^2([t,T]\times I;\mathbb{R}^k)$ and the boundedness of $\varphi'_{n,p}$, it follows from the B-D-G inequality, that

$$E\Big[\sup_{0\le s\le T}|\int_0^s \varphi'_{n,p}(|Y_r|^2)Y_r Z_r dB_r|\Big] \le Cpn^{p-1}||Y||_{S^2}||Z||_{M^2}.$$

Moreover, from the boundedness of $\varphi_{n,p}$, there exists a bounded and predictable process $\phi_s(j)$ such that

$$\varphi_{n,p}(|Y_{r-}+W_r(j)|^2) - \varphi_{n,p}(|Y_{r-}|^2) = \phi_r(j)(Y_{r-}^{\intercal}W_r(j) + |W_r(j)|^2).$$

By the B-D-G inequality, we have

$$
\begin{aligned}
&E\Big[\sup_{0\le t\le T}\sum_{j\in I}\int_s^T |\varphi_{n,p}(|Y_{r-}+W_r(j)|^2)-\varphi_{n,p}(|Y_{r-}|^2)|d\tilde{\mathscr{V}}_r(j)\Big]\\
=&E\Big[\sup_{0\le t\le T}\sum_{j\in I}\int_s^T |\phi_r(j)(Y_{r-}W_r(j)+|W_r(j)|^2)|d\tilde{\mathscr{V}}_r(j)\Big]\\
\le&Cpn^{p-1}(||Y||_{S^2}||W||_{N^2}+||W||^2_{N^2}).
\end{aligned}
$$

So, the dB integral and the $d\tilde{\mathscr{V}}$ integral are integrable with respect to the probability measure $\mathbb{P}$, hence they are martingales with zero expection.

From Taylor's expansion of $\varphi_{n,p}$ and the positivity of $\varphi''_{n,p}$, we conclude

$$
\begin{aligned}
\sum_{j\in I}\int_s^T \Big[&\varphi_{n,p}(|Y_{r-}+W_r(j)|^2)-\varphi_{n,p}(|Y_{r-}|^2)\\
&-\varphi'_{n,p}(|Y_r|^2)(|Y_{r-}+W_r(j)|^2-\varphi_{n,p}(|Y_{r-}|^2))\Big]\lambda_r(j)dr\ge 0.
\end{aligned}
\tag{2.17}
$$

Taking expection on (2.16) and from Assumption (A.2), we get that there exists a constant $q>0$ such that

$$
\begin{aligned}
&E\varphi_{n,p}(|Y_s|^2)+E\int_s^T\varphi'_{n,p}(|Y_r|^2)[|Z_r|^2+\sum_{j\in I}|W_r(j)|^2\lambda_r(j)]dr\\
\le&E\varphi_{n,p}(|h(X_T)|^2)+2E\int_s^T|\varphi'_{n,p}(|Y_r|^2)Y_sf(r,X_r,Y_r,Z_r,\alpha_r)|dr\\
\le&E\varphi_{n,p}(|h(X_T)|^2)+2CE\int_s^T\varphi'_{n,p}(|Y_r|^2)Y_r(1+|X_r|^q+|Y_r|+|Z_r|+|f(r,0,0,0,\alpha_r)|)dr.
\end{aligned}
$$

Let $n\to+\infty$, by monotone convergence theorem,

$$
\begin{aligned}
&E|Y_s|^{2p}+pE\int_s^T(|Y_r|^{2p-2})[|Z_r|^2+\sum_{j\in I}|W_r(j)|^2\lambda_r(j)]dr\\
\le&E(|h(X_T)|^{2p})+CE\int_s^T(|Y_r|^{2p-2})|Y_r|(1+|X_r|^q+|Y_r|+|Z_r|+\sup_{i\in I}|f(r,0,0,0,i)|)dr\\
\le&E(|h(X_T)|^{2p})+CE\int_s^T(1+|Y_r|^{2p})dr+\frac{p}{2}E\int_s^T(|Y_r|^{2p-2}|Z_r|^2)dr.
\end{aligned}
$$

We have used condition (A.3) that $\sup_{i\in I}f(t,0,0,0,i)<+\infty$ for $t\in[0,T]$ a.e. and the result that $E|X_s|^p<+\infty$.

It then follows from Gronwall's inequality and Condition (A.1) that there exist constants $C_{p,T}$ and q such that

$$
\sup_{0\le s\le T}E(|Y_s|^{2p})\le C_{p,T}(T+E(|h(X_T)|^{2p}))\le C(1+|x|^{2pq}),\tag{2.18}
$$

and also

$$E\int_s^T (|Y_r|^{2p-2})[|Z_r|^2+\sum_{j\in I}|W_r(j)|^2\lambda_r(j)]dr \le C(1+|x|^{2pq}). \tag{2.19}$$

Taking expection on both sides of (2.16) again, letting $n\to+\infty$, combining (2.18) and (2.19), we get that

$$\begin{aligned} &E\sum_{j\in I}\int_s^T \Big[|Y_{r-}+W_r(j)|^{2p}-|Y_{r-}|^{2p} \\ &\quad -p(|Y_{r-}+W_r(j)|^2-|Y_{r-}|^2)|Y_r|^{2(p-1)}\Big]\lambda_r(j)ds \le C(1+|x|^q). \end{aligned} \tag{2.20}$$

From (2.18), (2.19) and (2.20), we get

$$E\sum_{j\in I}\int_s^T \Big[|Y_{r-}+W_r(j)|^{2p}-|Y_{r-}|^{2p}\Big]\lambda_r(j)ds \le C(1+|x|^q). \tag{2.21}$$

Now, again from Ito's formula and (2.17),

$$\begin{aligned} |Y_s|^{2p} &\le |Y_T|^{2p}+2p\int_s^T |Y_r|^{2(p-1)}|Y_r||f(r,X_r,Y_r,Z_r,\alpha_r)|dr \\ &\quad -p\int_s^T |Y_r|^{2(p-1)}|Z_r|^2dr-p\sum_{j\in I}\int_s^T |Y_r|^{2(p-1)}|W_r(j)|^2\lambda_r(j)dr \\ &\quad -2p\int_s^T |Y_r|^{2(p-1)}Y_rZ_rdB_r-\sum_{j\in I}\int_s^T \Big[|Y_{r-}+W_r(j)|^{2p}-|Y_{r-}|^{2p}\Big]d\mathscr{V}_r(j). \end{aligned}$$

It follows from (2.18) and (2.19) that the above dB integral is uniformly integrable, and from (2.21) that the $d\tilde{\mathscr{V}}$ integral is a uniformly integrable martingale. It is then easy to conclude that

$$E(\sup_{t\le s\le T}|Y_s|^{2p}) \le C(1+|x|^q).$$

Finally, for any $t\le a\le s\le b\le T$,

$$\int_a^s Z_rdB_r+\sum_{j\in I}\int_a^s W_r(j)d\tilde{\mathscr{V}}_r(j) = Y_s-Y_a+\int_a^s f(r,X_r,Y_r,Z_r,\alpha_r)dr.$$

Hence, from Burkholder-Davis-Gundy's inequality, for any $p\ge 2$, there exists a constant C_p such that

$$\begin{aligned} &E\Big[\Big(\int_a^b (|Z_r|^2+\sum_{j\in I}|W_r(j)|^2\lambda_r(j))dr\Big)^{p/2}\Big] \\ &\le C_pE\Big[\sup_{a\le s\le b}|\int_a^s Z_rdB_r+\sum_{j\in I}\int_a^s W_r(j)d\tilde{\mathscr{V}}_r(j)|^p\Big] \\ &\le C_p\Big(1+(b-a)^{p/2}E\Big[\big(\int_a^b (|Z_r|^2+\sum_{j\in I}|W_r(j)|^2\lambda_r(j))dr\big)^{p/2}\Big]\Big). \end{aligned}$$

Hence, provided $b-a \leq C_p^{-4/p}$, then

$$E\Big[\Big(\int_a^b (|Z_r|^2 + \sum_{j\in I} |W_r(j)|^2 \lambda_r(j))\Big)^{p/2} dr\Big] \leq C(1+|x|^q)$$

and we finish the proof. $\square$

Proposition 2.12. *For any $p \geq 2$, there exist reals C_p and c_p such that for any $0 \leq t, t' \leq s \leq T$, x, $x' \in \mathbb{R}^d$, $h, h' \in R\backslash\{0\}$, $1 \leq i \leq d$,*

$$\begin{aligned}
E\Big[\sup_{s\in[0,T]} |Y_s^{t,x,i} - Y_s^{t,x',i}|^p\Big] + E\Big[\Big(\int_{t\wedge t'}^T |Z_s^{t,x,i} - Z_s^{t',x',i}|^2 ds\Big)^{p/2}\Big] \\
+ E\Big[\Big(\sum_{j\in I}\int_{t\wedge t'}^T |W_s(j)^{t,x,i} - W_s(j)^{t',x',i}|^2 \lambda_s(j) ds\Big)^{p/2}\Big] \\
\leq C_p(1+|x|+|x'|)^{c_p}(|x-x'|^p + |t-t'|^{1/2} + o(|t-t'|^{1/2}))
\end{aligned}$$

and

$$\begin{aligned}
E\Big[\sup_{s\in[0,T]} |\Delta_h^l Y_s^{t,x,i} - \Delta_{h'}^l Y_s^{t,x',i}|^p\Big] + E\Big[\Big(\int_{t\wedge t'}^T |\Delta_h^l Z_s^{t,x,i} - \Delta_{h'}^l Z_s^{t',x',i}|^2 ds\Big)^{p/2}\Big] \\
+ E\Big[\Big(\sum_{j\in I}\int_{t\wedge t'}^T |\Delta_h^l W_s(j)^{t,x,i} - \Delta_{h'}^l W_s(j)^{t',x',i}|^2 \lambda_s(j) ds\Big)^{p/2}\Big] \\
\leq C_p(1+|x|+|x'|)^{c_p}(|x-x'|^p + |t-t'|^{1/2} + |h-h'|^p + o(|t-t'|^{1/2})).
\end{aligned}$$

Proof. We first treat the case $t=t'$. For $x, x' \in \mathbb{R}^d$,

$$\begin{aligned}
Y_s^{t,x,i} - Y_s^{t,x',i} =& \Big(\int_0^1 h'(X_T^{t,x',i} + \theta(X_T^{t,x,i} - X_T^{t,x',i}))d\theta\Big)[X_T^{t,x,i} - X_T^{t,x',i})] \\
&+ \int_s^T \Big(\varphi_r(x,x')[X_r^{t,x,i} - X_r^{t,x',i})] + \psi_r(x,x')[Y_r^{t,x,i} - Y_r^{t,x',i}] \\
&\phi_r(x,x')\Big) dr - \int_s^T (Z_r^{t,x,i} - Z_r^{t,x',i}) dB_r \\
&- \sum_{j\in I}\int_s^T (W_r^{t,x,i}(j) - W_r^{t,x',i}(j)) d\widetilde{\mathscr{V}}_s(j),
\end{aligned}$$

where

$$\varphi_r(x,x') = \int_0^1 f_x'(\Sigma_{r,\theta}^{x,x'})d\theta, \psi_r(x,x') = \int_0^1 f_y'(\Sigma_{r,\theta}^{x,x'})d\theta, \phi_r(x,x') = \int_0^1 f_z'(\Sigma_{r,\theta}^{x,x'})d\theta$$

and

$$\Sigma_{r,\theta}^{x,x'} = (r, X_r^{t,x',i} + \theta(X_r^{t,x,i} - X_r^{t,x',i}), Y_r^{t,x',i} + \theta(Y_r^{t,x,i} - Y_r^{t,x',i}), Z_r^{t,x',i} + \theta(Z_r^{t,x,i} - Z_r^{t,x',i}), \alpha_r^{t,i}).$$

Combining the arguments of Proposition 2.11 and Theorem 2.2, we obtain:

$$E\Big[\sup_{0\le s\le T}|Y_s^{t,x,i}-Y_s^{t,x',i}|^p\Big]\le C_p(1+|x|^q+|x'|^q)|x-x'|^p. \tag{2.22}$$

In fact we should restrict the sup to $t\le s\le T$, but (2.22) follows easily from that restricted results. We have moreover

$$E\Big[\Big(\int_s^T ||Z_s^{t,x,i}-Z_s^{t,x',i}||^2ds\Big)^{p/2}\Big]\le C(1+|x|^q+|x'|^q)|x-x'|^p \tag{2.23}$$

and

$$E\Big[\Big(\sum_{j\in I}\int_t^T ||W_s(j)^{t,x,i}-W_s(j)^{t,x',i}||^2ds\Big)^{p/2}\Big]\le C(1+|x|^q+|x'|^q)|x-x'|^p.$$

We next have

$$\begin{aligned}
&\Delta_h^l Y_s^{t,x,i}=\int_0^1 h'(X_T^{t,x,i}+\theta h\Delta_h^l X_T^{t,x,i})\Delta_h^l X_T^{t,x,i}d\theta\\
&+\int_s^T\int_0^1\big[f'_x(\Xi_{r,\theta}^{t,x,h,i})\Delta_h^l X_r^{t,x,i}+f'_y(\Xi_{r,\theta}^{t,x,h,i})\Delta_h^l Y_r^{t,x,i}+f'_z(\Xi_{r,\theta}^{t,x,h,i})\Delta_h^l Z_r^{t,x,i}\big]d\theta dr\\
&-\int_s^T\Delta_h^l Z_r^{t,x,i}dB_r-\sum_{j\in I}\int_s^T\Delta_h^l W_r^{t,x,i}(j)d\widetilde{\mathscr{V}}_s(j),
\end{aligned}$$

where $\Xi_{r,\theta}^{t,x,h,i}=(r,X_r^{t,x,i}+\theta h\Delta_h^l X_r^{t,x,i},Y_r^{t,x,i}+\theta h\Delta_h^l Y_r^{t,x,i})$.

It is easy to deduce by (2.22) and (2.23) that

$$\begin{aligned}
E[\sup_{0\le s\le T}|\Delta_h^l Y_r^{t,x,i}|^p]+E[\big(\int_s^T||\Delta_h^l Z_r^{t,x,i}||^2\big)^{p/2}ds]&\\
+\sum_{j\in I}E[\big(\int_s^T||\Delta_h^l Z_r^{t,x,i}||^2\lambda_s(j)\big)^{p/2}]<+\infty.&
\end{aligned}$$

Using Itô's formula to $|\Delta_h^l Y_s^{t,x,i}|^p$ and the similar arguments in Proposition 2.11 and Lemma 2.4, we obtain that there exists C_p and q such that

$$\begin{aligned}
&E[\sup_{0\le s\le T}|\Delta_h^l Y_r^{t,x,i}|^p]+E[\big(\int_s^T||\Delta_h^l Z_r^{t,x,i}||^2\big)^{p/2}ds]\\
&+\sum_{j\in I}E[\big(\int_s^T||\Delta_h^l Z_r^{t,x,i}||^2\lambda_s(j)\big)^{p/2}]\le C_p(1+|x|^q+|x'|^q+|h|^q)).
\end{aligned} \tag{2.24}$$

Finally,

$$\begin{aligned}
\Delta_h^l Y_s^{t,x,i} - \Delta_{h'}^l Y_s^{t,x',i} = & \int_0^1 h'(X_T^{t,x,i} + \theta h \Delta_h^l X_T^{t,x,i}) \Delta_h^l X_T^{t,x,i} d\theta \\
& - \int_0^1 h'(X_T^{t,x',i} + \theta h' \Delta_{h'}^l X_T^{t,x',i}) \Delta_{h'}^l X_T^{t,x',i} d\theta \\
& + \int_s^T \int_0^1 \big[f_x'(\Xi_{r,\theta}^{t,x,h,i}) \Delta_h^l X_r^{t,x,i} - f_x'(\Xi_{r,\theta}^{t,x',h',i}) \Delta_{h'}^l X_r^{t,x',i} \big] d\theta dr \\
& + \int_s^T \int_0^1 \big[f_y'(\Xi_{r,\theta}^{t,x,h,i}) \Delta_h^l Y_r^{t,x,i} - f_y'(\Xi_{r,\theta}^{t,x',h',i}) \Delta_{h'}^l Y_r^{t,x',i} \big] d\theta dr \\
& + \int_s^T \int_0^1 \big[f_z'(\Xi_{r,\theta}^{t,x,h,i}) \Delta_h^l Z_r^{t,x,i} - f_z'(\Xi_{r,\theta}^{t,x',h',i}) \Delta_{h'}^l Z_r^{t,x',i} \big] d\theta dr \\
& - \int_s^T [\Delta_h^l Z_r^{t,x,i} - \Delta_{h'}^l Z_r^{t,x',i}] dB_r \\
& - \sum_{j \in I} \int_s^T [\Delta_h^l W_r^{t,x,i}(j) - \Delta_h^l W_r^{t,x',i}(j)] d\widetilde{\mathscr{V}}_s(j),
\end{aligned}$$

where

$$\begin{aligned}
A_s(x,h;x',h') = & \int_0^1 h'(X_T^{t,x,i} + \theta h \Delta_h^l X_T^{t,x,i}) \Delta_h^l X_T^{t,x,i} d\theta \\
& - \int_0^1 h'(X_T^{t,x',i} + \theta h' \Delta_{h'}^l X_T^{t,x',i}) \Delta_{h'}^l X_T^{t,x',i} d\theta \\
& + \int_s^T \int_0^1 \big[f_x'(\Xi_{r,\theta}^{t,x,h,i}) \Delta_h^l X_r^{t,x,i} - f_x'(\Xi_{r,\theta}^{t,x',h',i}) \Delta_{h'}^l X_r^{t,x',i} \big] d\theta dr \\
& + \int_s^T \int_0^1 \big[f_y'(\Xi_{r,\theta}^{t,x,h,i}) - f_y'(\Xi_{r,\theta}^{t,x',h',i}) \big] \Delta_{h'}^l Y_r^{t,x',i} d\theta dr \\
& + \int_s^T \int_0^1 \big[f_z'(\Xi_{r,\theta}^{t,x,h,i}) - f_z'(\Xi_{r,\theta}^{t,x',h',i}) \big] \Delta_{h'}^l Z_r^{t,x',i} d\theta dr,
\end{aligned}$$

and $\overline{f_y'}(x,h;x',h') = \int_0^1 f_x'(\Xi_{r,\theta}^{t,x,h,i})$. Again by the procedure of Proposition 2.11 and Lemma 2.4, using the properties of f and (2.22), we deduce that

$$\begin{aligned}
E\Big[\sup_{0\le s\le T} |\Delta_h^l Y_s^{t,x,i} - \Delta_{h'}^l Y_s^{t,x',i}|^p\Big] \le & C_p(1+|x|^q+|x'|^q+|h|^q+|h'|^q) \\
& \times (|x-x'|^p + |h-h'|^p),
\end{aligned}$$

$$\begin{aligned}
E\Big[\Big(\int_s^T ||\Delta_h^l Z_s^{t,x,i} - \Delta_{h'}^l Z_s^{t,x',i}||^2 ds\Big)^{p/2}\Big] \le & C_p(1+|x|^q+|x'|^q+|h|^q+|h'|^q) \\
& \times (|x-x'|^p + |h-h'|^p)
\end{aligned}$$

and

$$\begin{aligned}
& \sum_{j\in I} E\Big[\Big(\int_s^T ||\Delta_h^l W_s^{t,x,i}(j) - \Delta_{h'}^l W_s^{t,x',i}(j)||^2 \lambda_s(j) ds\Big)^{p/2}\Big] \\
& \le C_p(1+|x|^q+|x'|^q+|h|^q+|h'|^q) \times (|x-x'|^p + |h-h'|^p).
\end{aligned}$$

Indeed, take Itô's formula to $|\Delta_h^l Y_s^{t,x,i} - \Delta_{h'}^l Y_s^{t,x',i}|^{2p}$ with $p \geq 1$ and $t \leq a \leq b \leq T$,

$$
\begin{aligned}
|\overline{\nabla Y_b^t}|^{2p} - |\overline{\nabla Y_a^t}|^{2p} &= -\int_a^b 2p|\overline{\nabla Y_r^t}|^{2p-1}(\overline{f_x'}(r) + \overline{f_y'}(r) + \overline{f_z'}(r))dr \\
&\quad + \int_a^b 2p|\overline{\nabla Y_r^t}|^{2p-1}|\overline{\nabla Z_r^t}|dB_r + \int_a^b p(2p-1)|\overline{\nabla Y_r^t}|^{2p-2}|\overline{\nabla Z_r^t}|^2 dr \\
&\quad + \sum_{j\in I}\int_a^b \left[|\overline{\nabla Y_{r-}^t} + \overline{\nabla W_r^t}(j)|^{2p} - |\overline{\nabla Y_{r-}^t}|^{2p}\right] d\tilde{V}_r(j) \\
&\quad + \sum_{j\in I}\int_a^b p|\overline{\nabla Y_r^t}|^{2p-2}|\overline{\nabla W_r^t}(j)|^2 \lambda_r(j)dr \\
&\quad + \sum_{j\in I}\int_a^b \left[|\overline{\nabla Y_r^t} + \overline{\nabla W_r^t}(j)|^{2p} - |\overline{\nabla Y_r^t}|^{2p}\right. \\
&\qquad \left. - p(|\overline{\nabla Y_r^t} + \overline{\nabla W_r^t}(j)|^2 - |\overline{\nabla Y_r^t}|^2)|\overline{\nabla Y_r^t}|^{2p-2}\right]\lambda_r(j)dr,
\end{aligned}
$$

where $\overline{\nabla Y_r^t} = \Delta_h^l Y_r^{t,x,i} - \Delta_{h'}^l Y_r^{t,x',i}$, $\overline{\nabla Z_r^t} = \Delta_h^l Z_r^{t,x,i} - \Delta_{h'}^l Z_r^{t,x',i}$, $\overline{\nabla W_r^t} = \Delta_h^l W_r^{t,x,i} - \Delta_{h'}^l W_r^{t,x',i}$, and

$$
\begin{aligned}
\overline{f_x'}(r) &= \int_0^1 f_x'(\Xi_{r,\theta}^{t,x,h,i})\Delta_h^l X_r^{t,x,i} - f_x'(\Xi_{r,\theta}^{t,x',h',i})\Delta_{h'}^l X_r^{t,x',i} d\theta, \\
\overline{f_y'}(r) &= \int_0^1 f_y'(\Xi_{r,\theta}^{t,x,h,i})\Delta_h^l Y_r^{t,x,i} - f_y'(\Xi_{r,\theta}^{t,x',h',i})\Delta_{h'}^l Y_r^{t,x',i} d\theta, \\
\overline{f_z'}(r) &= \int_0^1 f_z'(\Xi_{r,\theta}^{t,x,h,i})\Delta_h^l Z_r^{t,x,i} - f_z'(\Xi_{r,\theta}^{t,x',h',i})\Delta_{h'}^l Z_r^{t,x',i} d\theta.
\end{aligned}
$$

It is easy to deduce that the last term on the right hand side of the last equation is positive. Let us only show that how to deal with the "hardest" term comparing with the proof of Theorem 2.9 in [8]:

$$
\begin{aligned}
&|E\int_a^b \overline{f_z'}(r)|\overline{Y_r^t}|^{2p-1}dr| \\
&= |E\int_a^b \Big(\int_0^1 [f_z'(\Xi_{r,\theta}^{t,x,h,i})\Delta_h^l Z_r^{t,x,i} - f_z'(\Xi_{r,\theta}^{t,x',h',i})\Delta_{h'}^l Z_r^{t,x',i}]d\theta\Big)|\overline{Y_r^t}|^{2p-1}dr| \\
&\leq CE\int_a^b ||\Delta_h^l Z_r^{t,x,i} - \Delta_{h'}^l Z_r^{t,x',i}|| \times |\overline{Y_r^t}|^{2p-1}dr| \\
&\quad + CE\int_a^b ||\Delta_h^l Z_r^{t,x,i}||\Big(\int_0^1 |\Xi_{r,\theta}^{t,x,h,i} - \Xi_{r,\theta}^{t,x',h',i}|d\theta\Big)|\overline{Y_r^t}|^{2p-1}dr| \\
&\leq \frac{1}{2}E\Big(\sup_{a\leq r\leq b}|\Delta_h^l Y_r^{t,x,i} - \Delta_{h'}^l Y_r^{t,x',i}|^{2p}\Big) + \bar{c}(b-a)E\Big[\Big(\int_a^b ||\Delta_h^l Z_r^{t,x,i} - \Delta_{h'}^l Z_r^{t,x',i}||^2 dr\Big)^p\Big] \\
&\quad + \bar{c}\sqrt{E\Big[\Big(\int_a^b ||\Delta_h^l Z_r^{t,x,i}||^2 dr\Big)^p\Big]}\sqrt{E\Big[\Big(\int_a^b\int_0^1 |\Xi_{r,\theta}^{t,x,h,i} - \Xi_{r,\theta}^{t,x',h',i}|^2 d\theta dr\Big)^p\Big]}.
\end{aligned}
$$

We note that the first two terms on the right hand side are subtracted from the left terms of the full inequality, with $(b-a)$ small enough, and the last term is estimated with the help of (2.24). Note also that we choose first $b=T$, $a=T-\alpha$, then $b=T-\alpha$, $a=T-2\alpha$, etc.

We now deal with the case that $t\le t'$. Without loss of generality, we set $t\le t'$ and by the uniqueness of solution of the BSDE, we have $Y_s^{t,x,i}=Y_s^{t',X_{t'}^{t,x,i},\alpha_{t'}^{t,i}}$,

$$
\begin{aligned}
&E\Big[\sup_{s\in[0,T]}|Y_s^{t,x,i}-Y_s^{t',x',i}|^p\Big]\\
&\le E\Big[\sup_{s\in[0,T]}|Y_s^{t',X_{t'}^{t,x,i},i}-Y_s^{t',x',i}|^p\Big]+E\Big[\sup_{s\in[0,T]}|Y_s^{t',X_{t'}^{t,x,i},i}-Y_s^{t',X_{t'}^{t,x,i},\alpha_{t'}^{t,i}}|^p\Big]\\
&\le CE(1+|x|^q+|X_{t'}^{t,x,i}|^q)|X_{t'}^{t,x,i}-x'|^p+CE\Big[\sup_{s\in[0,T]}\Big(|Y_s^{t',X_{t'}^{t,x,i},i}|^p+|Y_s^{t',X_{t'}^{t,x,i},\alpha_{t'}^{t,i}}|^p\Big)I_{\alpha_{t'}^{t,i}\neq i}\Big]\\
&\le C\sqrt{E(1+|x|^q+|X_{t'}^{t,x,i}|^{2q})}\Big(\sqrt{E|X_{t'}^{t,x,i}-x|^{2p}}+|x-x'|^p\Big)+C\sqrt{E\big[|X_{t'}^{t,x,i}|^{2q}\big]}\sqrt{E\big[I_{\alpha_{t'}^{t,i}\neq i}\big]}.
\end{aligned}
$$

Then by virtue of Theorem 2.2 and Markov property, and the same procedures to $Z_s^{t,x,i}$, $W_s^{t,x,i}(j)$, $\Delta_h^l Y_s^{t,x,i}$, $\Delta_h^l Z_s^{t,x,i}$ and $\Delta_h^l W_s^{t,x,i}(j)$, we get the result. □

Corollary 2.13. *The function*

$$(s,t,x)\mapsto E[Y_s^{t,x,i}]$$

belongs to $C^{0,0,2}([0,T]\times[0,T]\times\mathbb{R}^d;\mathbb{R}^k)$*, and in particular*

$$(t,x)\to E[Y_t^{t,x,i}]=Y_t^{t,x,i}\in C_p^{0,2}([0,T]\times\mathbb{R}^d;\mathbb{R}^k).$$

3. Related PDEs: classical sense

In this section, we will build the connection of our BSDEs with the following system of PDEs:

$$
\begin{cases}
\dfrac{\partial u}{\partial t}(t,x,i)=-\mathscr{L}u(t,x,i)-f(t,x,u(t,x),(\nabla u\sigma)(t,x,i),i)\\
\qquad\qquad -\displaystyle\sum_{j\neq i,j\in I}\lambda_{ij}(t)(u(t,x,j)-u(t,x,i)),\\
u(T,x,i)=h(x),\quad (t,x,i)\in[0,T]\times\mathbb{R}^n\times I,
\end{cases}
\tag{3.1}
$$

where

$$\mathscr{L}\varphi(t,x,i) = \frac{1}{2}\sum_{p,q=1}^{n}(\sigma\sigma^T)_{pq}(t,x,i)\frac{\partial^2}{\partial x_p \partial x_q}\varphi(t,x,i) + \sum_{p=1}^{n} b^p(t,x,i)\frac{\partial}{\partial x_p}\varphi(t,x,i).$$

In [14], the viscosity solution of this kind of PDEs has been discussed but without the dependence of f on z. In this section, we are going to study the classical solution of the above PDEs.

Theorem 3.1. *Let $b \in C^3_{l,b}$, $\sigma \in C^3_{l,b}$, f and h satisfy the assumptions (A.1)-(A.3). Then the function defined by*

$$u(t,x,i) =: Y_t^{t,x,i}, \quad (t,x,i) \in ([0,T]\times\mathbb{R}^d\times I)$$

is the unique solution of the partial differential equation (3.1) *in* $C^{1,2}([0,T]\times\mathbb{R}^d;\mathbb{R}^k)$

Before proceeding to the proof of Theorem 3.1, let us give the following two lemmas first.

Lemma 3.2. *Under the assumptions in Theorem 3.1, for all $(t,x,i) \in [0,T]\times\mathbb{R}^d\times I$, we have a.s.*

$$\begin{aligned} Y_s^{t,x,i} &= u(s, X_s^{t,x,i}, \alpha_s^{t,i}), \\ W_s^{t,x,i}(j) &= u(s, X_s^{t,x,i}, j) - u(s, X_s^{t,x,i}, \alpha_{s-}^{t,i}). \end{aligned} \tag{3.2}$$

where the function u was defined in Theorem 3.1. Moreover, for all $(x,i) \in \mathbb{R}^d\times I$, for a.e. $s \in [0,T]$, we have a.s.

$$Z_s^{t,x,i} = (\nabla u\sigma)(s, X_s^{t,x,i}, \alpha_s^{t,i}). \tag{3.3}$$

Proof. The proof of equation (3.2) can be obtained from the uniqueness of the solution of BSDE (2.15). For the details of the proof one can refer to Lemma 4.2 in [14]. Now we prove the equation (3.3). By virtue of the general Itô's formula [see Theorem 1.45

in [18]] and (3.2), we have

$$
\begin{aligned}
&u(t+h,x,i)-u(t,x,i)\\
=&u(t+h,x,i)-u(t+h,X^{t,x,i}_{t+h},\alpha^{t,i}_{t+h})+u(t+h,X^{t,x,i}_{t+h},\alpha^{t,i}_{t+h})-u(t,x,i)\\
=&-\int_t^{t+h}(\nabla u\sigma)(t+h,X^{t,x,i}_r,\alpha^{t,i}_r)dB_r-\int_t^{t+h}Lu(t+h,X^{t,x}_r,\alpha^{t,i}_r)dr\\
&-\sum_{j\in I}\int_t^{t+h}(u(t+h,X^{t,x,i}_r,j)-u(t+h,X^{t,x,i}_r,\alpha^{t,i}_{r-}))\lambda_t(j)dr\\
&-\sum_{j\in I}\int_t^{t+h}(u(t+h,X^{t,x,i}_r,j)-u(t+h,X^{t,x,i}_r,\alpha^{t,i}_{r-}))d\tilde{\mathscr{V}}_r(j)\\
&-\int_t^{t+h}f(r,X^{t,x,i}_r,Y^{t,x,i}_r,Z^{t,x,i}_r,\alpha^{t,i}_r)dr+\int_t^{t+h}Z^{t,x,i}_r dB_r+\sum_{j\in I,j\neq i}\int_t^{t+h}W^{t,x,i}_r(j)d\tilde{\mathscr{V}}_r(j).
\end{aligned}
\tag{3.4}
$$

Combining (2.2), (2.18) and (2.24), it is easy to deduce that the function f, $\nabla u\sigma$ and Lu grow at most like a polynomial function of the variable x at infinity. Moreover, the transition intensities $\lambda_{ij}(t)$ are bounded and positive. As $u(t+h,x,i)-u(t,x,i)$ is deterministic, we get by taking expectation in the preceding equality that

$$
|u(t+h,x,i)-u(t,x,i)|\leq C(1+|x|^q)|h|,\quad t,t+h\in[0,T]. \tag{3.5}
$$

From (3.4) and (3.5), by virtue of Burkholder's inequality, we have

$$
\begin{aligned}
&E\int_t^{t+h}|(\nabla u\sigma)(t+h,X^{t,x,i}_r,\alpha^{t,i}_r)-Z^{t,x,i}_r|^2dr\\
&+E\sum_{j\in I}\int_t^{t+h}|u(t+h,X^{t,x,i}_r,j)-u(t+h,X^{t,x,i}_r,\alpha^{t,i}_{r-})|^2\lambda_t(j)dr\\
&\leq 2E\int_t^{t+h}|\theta(r)|^2dr+2|u(t+h,x,i)-u(t,x,i)|^2,
\end{aligned}
$$

where

$$
\begin{aligned}
\theta(r)=&-Lu(t+h,X^{t,x}_r,\alpha^{t,i}_r)-\sum_{j\in I}\big(u(t+h,X^{t,x,i}_r,j)\\
&-u(t+h,X^{t,x,i}_r,\alpha^{t,i}_{r-})\big)\lambda_r(j)-f(r,X^{t,x,i}_r,Y^{t,x,i}_r,Z^{t,x,i}_r,\alpha^{t,i}_r).
\end{aligned}
$$

Then by a simple calculus, we have

$$
\begin{aligned}
&\frac{1}{2}E\int_t^{t+h}|(\nabla u\sigma)(r,X_r^{t,x,i},\alpha_r^{t,i})-Z_r^{t,x,i}|^2dr\\
&\leq 2E|\int_t^{t+h}\theta(r)dr|^2+2|u(t+h,x,i)-u(t,x,i)|^2\\
&+8E\int_t^{t+h}|(\nabla u\sigma)(t+h,X_r^{t,x,i},\alpha_r^{t,i})-(\nabla u\sigma)(r,X_r^{t,x,i},\alpha_r^{t,i})|^2dr\\
&\leq C(1+|x|^q)|h|^2.
\end{aligned}
$$

Consequently, considering a partition $t_k^n=t+hk2^{-n}, 0\leq k\leq 2^n$, we have from the preceding estimate applied to (t_k^n,t_{k+1}^n) instead of $(t,t+h)$ and Hölder's inequality:

$$
\begin{aligned}
&E\int_t^{t+h}|(\Delta u\sigma)(r,X_r^{t,x,i},\alpha_r^{t,i})-Z_r^{t,x,i}|^2dr\\
&=E[\sum_{k=0}^{2^n-1}E[\int_{t_k^n}^{t_{k+1}^n}|(\Delta u\sigma)(r,X_r^{t_k^n,y,j},\alpha_r^{t_k^n,j})-Z_r^{t_k^n,y,j}|^2dr|\mathscr{F}_{t_k^n}]|_{y=X_{t_k^n}^{t,x,i},j=\alpha_{t_k^n}^{t,i}}]\\
&\leq C(1+|x|^q)\sum_{k=0}^{2^n-1}(h2^{-n})^2\to 0,\quad \text{as}\quad n\to+\infty.
\end{aligned}
$$

Then (3.3) follows. □

Lemma 3.3. *Let $g\in C_p(\mathbb{R}^d)$. Then*

$$
\big((s,t,x)\to E[g(X_{s\vee t}^{t,x,i})]\big)\in C_p([t,T]\times[0,T]\times\mathbb{R}^d)\quad \textit{for any } i\in I.
$$

Proof. Fix $(s,t,x,i), t\leq s\leq T$, and let $0\leq t'\leq s'\leq T, x,x'\in\mathbb{R}^d$ with $|x-x'|\leq 1$. Then

$$
\begin{aligned}
&E[g(X_s^{t,x,i})]-E[g(X_{s'}^{t',x',i})]\leq E[|g(X_s^{t,x,i})-g(X_{s'}^{t',x',i})|I_{\{|X_s^{t,x,i}|\leq N,|X_s^{t,x,i}-X_{s'}^{t',x',i}|\leq\delta\}}]\\
&+E[|g(X_s^{t,x,i})-g(X_{s'}^{t',x',i})|I_{\{|X_s^{t,x,i}|>N\}}]+E[|g(X_s^{t,x,i})-g(X_{s'}^{t',x',i})|I_{\{|X_s^{t,x,i}-X_{s'}^{t',x',i}|>\delta\}}].
\end{aligned}
$$

Clearly, there exists a constant $N(\varepsilon,x)$ which is only dependent on ε and x, such that

$$
\begin{aligned}
&E[|g(X_s^{t,x,i})-g(X_{s'}^{t',x',i})|I_{\{|X_s^{t,x,i}|>N\}}]\\
&\leq E[(|g(X_s^{t,x,i})-g(X_{s'}^{t',x',i})|)^2]^{1/2}P\{|X_s^{t,x,i}|>N\}^{1/2}\\
&\leq CE(1+|X_s^{t,x,i}|+|X_{s'}^{t',x',i}|)^q\frac{1}{N}E[|X_s^{t,x,i}|^2]^{1/2}\\
&\leq\frac{C}{N}(1+|x|)^q\leq\frac{\varepsilon}{3},\ \text{if } N\geq N(\varepsilon,x).
\end{aligned}
$$

Fixing $N = N(\varepsilon, x)$, then there exists $\delta^\star(\varepsilon, N) > 0$ such that

$$
\begin{aligned}
& E[|g(X_s^{t,x,i}) - g(X_{s'}^{t',x',i})| I_{\{|X_s^{t,x,i}| \le N, |X_s^{t,x,i} - X_{s'}^{t',x',i}| \le \delta\}}] \\
& \le \sup_{|y| \le N, |y-y'| \le \delta^\star} |g(y) - g(y')| \le \frac{\varepsilon}{3}, \text{ if } \delta \le \delta^\star(\varepsilon, N).
\end{aligned}
$$

Now we fix $\delta^\star = \delta^\star(\varepsilon, N)$. Then there exists $\delta(\delta^\star, \varepsilon, x)$

$$
\begin{aligned}
& E[|g(X_s^{t,x,i}) - g(X_{s'}^{t',x',i})| I_{\{|X_s^{t,x,i} - X_{s'}^{t',x',i}| > \delta\}}] \\
& \le C(1+|x|)^q \frac{1}{\delta} E[|X_s^{t,x,i} - X_{s'}^{t',x',i}|^2]^{1/2} \\
& \le \frac{C}{\delta}(1+|x|)^q (|s-s'| + |t-t'|^{\frac{1}{2}} + o(t-t')^{\frac{1}{2}} + |x-x'|^2)^{1/2} \\
& \le \frac{\varepsilon}{3},
\end{aligned}
$$

if $|s-s'| + |t-t'|^{\frac{1}{2}} + o(t-t')^{\frac{1}{2}} + |x-x'|^2 \le \delta(\delta^\star, \varepsilon, x)$.

Hence, if $|s-s'| + |t-t'|^{\frac{1}{2}} + o(t-t')^{\frac{1}{2}} + |x-x'|^2 \le \delta(\delta^\star, \varepsilon, x)$,

$$
|E[g(X_s^{t,x,i})] - E[g(X_{s'}^{t',x',i})]| \le \varepsilon.
$$

□

Proof of Theorem 3.1. From Lemma 3.2 we have $Y_s^{t,x,i} = u(s, X_s^{t,x,i}, \alpha_s^{t,i})$ and $W_s^{t,x,i}(j) = u(s, X_s^{t,x,i}, j) - u(s, X_s^{t,x,i}, \alpha_{s-}^{t,i})$ when $j \neq \alpha_{s-}^{t,i}$. From Corollary 2.13, $u(\cdot,\cdot,i) \in C^{0,2}([0,T] \times \mathbb{R}^d)$. Let $h > 0$ be such that $t+h \le T$. Clearly, $Y_{t+h}^{t,x,i} = Y_{t+h}^{t+h, X_{t+h}^{t,x,i}, \alpha_{t+h}^{t,i}}$. Hence by virtue of Eq.(3.4) and Proposition 2.12 we know that $x \to f(s,x,u(s,x,i),(\triangledown u\sigma)(s,x,i),i)$ and $x \to \mathscr{L}u(s,x,i)$ are in $C_p(\mathbb{R}^d)$ and continuous in s. Then from Lemma 3.3, the functions

$$
(s,t,x) \to E[f(s, X_s^{t,x,i}, u(s, X_s^{t,x,i}, \alpha_s^{t,i}), (\triangledown u\sigma)(s, X_s^{t,x,i}, \alpha_s^{t,i}), \alpha_s^{t,i})]
$$

and

$$
(s,t,t',x) \to E[Lu(t', X_s^{t,x,i})]
$$

are continuous.

Let now $t = t_0 < t_1 < \cdots < t_n = T$, we have

$$
\begin{aligned}
h(x) - u(t,x,i) &= E[h(x) - u(t,x,i)] \\
= -\sum_{k=0}^{n-1} &\int_{t_k}^{t_{k+1}} E\left[Lu(t_{k+1}, X_r^{t_k,x,i}, \alpha_r^{t_k,i}) + f(r, \alpha_r^{t_k,i}, X_r^{t_k,x,i}, Y_r^{t_k,x,i})\right] dr \\
&- \sum_{j\in I, j\neq i} \int_{t_k}^{t_{k+1}} \lambda_{\alpha_{t-}^{t_k,i}, j} E[u(t_{k+1}, X_r^{t_k,x,i}, j) - u(t_{k+1}, X_r^{t_k,x,i}, \alpha_r^{t_k,i})] dr \\
&+ E\Theta,
\end{aligned}
$$

where Θ_t^n is the sum of all dB and $d\tilde{\mathscr{V}}_r(j)$ integral. It is easy to see that the above dB and $d\tilde{\mathscr{V}}_r(j)$ integrals are uniformly integrable. So Θ_t^n is a martingale with zero expectation. It follows from Corollary 2.13 that, if we take a sequence of partitions $t = t_0^n < t_1^n < \cdots t_n^n = T$ such that $\lim_{n\to+\infty} \sup_{k<n}(t_{k+1}^n - t_k^n) = 0$, we obtain in the limit that

$$
\begin{aligned}
h(x) - u(t,x,i) = &-\int_s^T \left[Lu(s,x,i) + f(r,i,x,u(t,x,i),(\triangledown u\sigma)(t,x,i))\right] dr \\
&- \sum_{j\in I, j\neq i} \int_t^T \lambda_{i,j}(s)(u(s,x,j) - u(s,x,i)) dr.
\end{aligned}
$$

So, we get that $u(\cdot,\cdot,i) \in C^{1,2}([0,T]\times\mathbb{R}^d)$ and solves equation (3.1).

It remains to prove the uniqueness of the solution. Let $u(\cdot,\cdot,i) \in C^{1,2}([0,T]\times\mathbb{R}^d;\mathbb{R}^k)$ be any solution of (3.1) and put

$$
\begin{aligned}
\hat{Y}_s^{t,x,i} &= u(s, X_s^{t,x,i}, \alpha_s^{t,i}), \\
\hat{Z}_s^{t,x,i} &= (\triangledown u\sigma)(s, X_s^{t,x,i}, \alpha_s^{t,i}), \\
\hat{W}_s^{t,x,i}(j) &= u(s, X_s^{t,x,i}, j) - u(s, X_s^{t,x,i}, \alpha_{s-}^{t,i}).
\end{aligned}
$$

From the general Ito's formula and equation (3.1) we have

$$
\hat{Y}_t = g(X_T) + \int_t^T f(s, X_s, \hat{Y}_s, \hat{Z}_s, \alpha_s) ds - \int_t^T \hat{Z}_s dB_s - \sum_{j\in I} \int_t^T \hat{W}_s(j) d\tilde{\mathscr{V}}_s(j).
$$

The uniqueness comes from the uniqueness of the solution of BSDE (2.15). □

4. BSDEs in integrable function space with Markov chains

In the previous discussion on the BSDEs with Markov chains, we have assumed that the coefficient f must satisfy the traditional Lipschitz condition in variables y and

z. However, in this section, we will introduce another type of Lipschitz condition on f, called the functional Lipschitz condition and will study the solutions of the BSDEs with Markov chains under such condition. Assume that.

(B.0) $b(s,\cdot,i) \in C^2_{l,b}(\mathbb{R}^d;\mathbb{R}^d)$, $\sigma(s,\cdot,i) \in C^3_b(\mathbb{R}^d;\mathbb{R}^{d\times d})$ for $s \in [t,T]$ and $i \in I$. Moreover, b and σ are continuous in s;

(B.1) $h : \mathbb{R}^d \to \mathbb{R}^k$ is Borel measurable and $\int_{\mathbb{R}^d} |h(x)|^2 \rho^{-1}(x)dx < +\infty$;

(B.2) $f : \Omega \times [0,T] \times \mathbb{R}^d \times \mathbb{R}^k \times \mathbb{R}^{d\times k} \times I \to \mathbb{R}$ is progressively measurable and functional Lipschitz, i.e. for all $t \in [0,T]$, $i \in I$ and any $Y_1, Y_2 \in L^2_\rho(\mathbb{R}^d;\mathbb{R}^k)$, $X \in L^2_\rho(\mathbb{R}^d;\mathbb{R}^d)$, $Z_1, Z_2 \in L^2_\rho(\mathbb{R}^d;\mathbb{R}^{d\times k})$, there exists a constant C, such that

$$
\begin{aligned}
&\int_{\mathbb{R}^d} |f(t,X(x),Y_1(x),Z_1(x),i) - f(t,X(x),Y_2(x),Z_2(x),i)|^2 \rho^{-1}(x)dx \\
&\leq C \int_{\mathbb{R}^d} (|Y_1(x) - Y_2(x)|^2 + |Z_1(x) - Z_2(x)|^2)\rho^{-1}(x)dx,
\end{aligned}
$$

and

$$
\sum_{j\in I} \int_t^T \int_{\mathbb{R}^d} |f(s,x,0,0,i)|^2 \rho^{-1}(x)dxds < +\infty,
$$

where ρ is the weighted function defined in Lemma 2.10.

Definition 4.1. *Let $\mathbb{S}$ be a Banach space with norm $||\cdot||_{\mathbb{S}}$ and Borel σ-field $\mathscr{S}$. We denote by $M^2([t,T];\mathbb{S})$ the set of $\mathscr{B}_{[t,T]} \otimes \mathscr{F}/\mathscr{S}$ measurable random processes $\{\phi(s)\}_{t\leq s\leq T}$ with values on $\mathbb{S}$ satisfying:*

(i) $\phi(s) : \Omega \to \mathbb{S}$ is $\mathscr{F}_s$-adapted for $t \leq s \leq T$;

(ii) $E[\int_t^T ||\phi(s)||^2_{\mathbb{S}} ds] < +\infty$.

We denote by $N^2([t,T]\times I;\mathbb{S})$ the set of $\mathscr{B}_{[t,T]} \otimes \mathscr{F}/\mathscr{S}$ measurable random processes for each $j \in I$ $\{\phi(s,j)\}_{t\leq s\leq T, j\in I}$ with values on $\mathbb{S}$ satisfying:

(i) $\phi(s,j) : \Omega \to \mathbb{S}$ is $\mathscr{F}_s$-adapted measurable for $t \leq s \leq T$, $j \in I$;

(ii) $E[\sum_{j\in I} \int_t^T ||\phi(s(j))||^2_{\mathbb{S}} \lambda_s(j)ds] < +\infty$.

We also denote by $S^2([t,T];\mathbb{S})$ the set of $\mathscr{B}_{[t,T]} \otimes \mathscr{F}/\mathscr{S}$ measurable random processes $\{\phi(s)\}_{t\leq s\leq T}$ with values on $\mathbb{S}$ satisfying:

(i) $\phi(s) : \Omega \to \mathbb{S}$ is $\mathscr{F}_s$-adapted measurable for $t \leq s \leq T$;

(ii) $E[\sup_{t\leq s\leq T}||\phi(s)||^2_{\mathbb{S}}ds] < +\infty$.

We denote by $L^2_\rho(\mathbb{R}^p;\mathbb{R}^q)$ *the* ρ*-weighted Hilbert space, with the normal*

$$||\phi||^2_{L^2_\rho} := \int_{\mathbb{R}^d} |\phi(x)|^2\rho(x)^{-1}dx.$$

Definition 4.2. *A triple of process* $(Y^{t,x,i}_s, Z^{t,x,i}_s, W^{t,x,i}_s(j))$ *is called a solution of BSDE* (2.15) *if* $(Y^{t,\cdot,i}_\cdot, Z^{t,\cdot,i}_\cdot, W^{t,\cdot,i}_\cdot) \in S^2([t,T];L^2_\rho(\mathbb{R}^d;\mathbb{R}^k)) \times M^2([t,T];L^2_\rho(\mathbb{R}^d;\mathbb{R}^{d\times k})) \times N^2([t,T]\times I;L^2_\rho(\mathbb{R}^d;\mathbb{R}^k))$ *and* $(Y^{t,x,i}_s, Z^{t,x,i}_s, W^{t,x,i}_s(j))$ *satisfies BSDE* (2.15) *for a.a. x with probability one. Due to the density of* $C^0_c(\mathbb{R}^d;\mathbb{R}^k)$ *in* $L^2_\rho(\mathbb{R}^d;\mathbb{R}^k)$*, it is equivalent to that for an arbitrary* $\varphi \in C^0_c(\mathbb{R}^d;\mathbb{R}^k)$*,* $(Y^{t,x,i}_s, Z^{t,x,i}_s, W^{t,x,i}_s(j))$

$$\begin{aligned}\int_{\mathbb{R}^d} Y^{t,x,i}_s\varphi(x)dx = &\int_{\mathbb{R}^d} h(X^{t,x,i}_T)\varphi(x)dx + \int_s^T\int_{\mathbb{R}^d} f(r,X^{t,x,i}_r,Y^{t,x,i}_r,Z^{t,x,i}_r,\alpha^{t,i}_r)\varphi(x)dxdr\\ &-\int_s^T\langle\int_{\mathbb{R}^d} Z^{t,x,i}_r\varphi(x)dx, dB_r\rangle - \sum_{j\in I}\int_t^T\int_{\mathbb{R}^d} W^{t,x,i}_r(j)\varphi(x)dxd\widetilde{\mathscr{V}}_r(j).\end{aligned} \tag{4.1}$$

First, we give a lemma, which is a straightforward extension of Lemma 3.3 in [20].

Lemma 4.3. *Under conditions (B.0)-(B.2), if there exists* $(Y_\cdot(\cdot), Z_\cdot(\cdot), W_\cdot(\cdot,\cdot)) \in M^2([t,T];L^2_\rho(\mathbb{R}^d;\mathbb{R}^k)) \times M^2([t,T];L^2_\rho(\mathbb{R}^d;\mathbb{R}^{d\times k})) \times N^2([t,T]\times I;L^2_\rho(\mathbb{R}^d;\mathbb{R}^k))$ *satisfying the spatial integral form of Eq.* (2.15)*, i.e.* (4.1) *for* $t\leq s\leq T$*, then* $Y_\cdot(\cdot)\in S^2([t,T];L^2_\rho(\mathbb{R}^d;\mathbb{R}^k))$ *and therefore* $(Y_\cdot(\cdot), Z_\cdot(\cdot), W_\cdot(\cdot,\cdot))$ *is a solution of Eq.* (2.15)*.*

Theorem 4.4. *Under condition (B.1)-(B.2), BSDE* (2.15) *has a unique solution* $(Y^{t,\cdot,i}_\cdot, Z^{t,\cdot,i}_\cdot, W^{t,\cdot,i}_\cdot)$ *in* $S^2([t,T];L^2_\rho(\mathbb{R}^d;\mathbb{R}^k)) \times M^2([t,T];L^2_\rho(\mathbb{R}^d;\mathbb{R}^{d\times k})) \times N^2([t,T]\times I;L^2_\rho(\mathbb{R}^d;\mathbb{R}^k))$.

Proof. First, we consider the linear BSDE

$$\begin{aligned}\widetilde{Y}^{t,\varphi,i}_s = &\int_{\mathbb{R}^d} h(X^{t,x,i}_T)\varphi(x)dx + \int_s^T\int_{\mathbb{R}^d}\widetilde{f}(r,X^{t,x,i}_r,\alpha^{t,i}_r)\varphi(x)dxds - \int_s^T\widetilde{Z}^{t,\varphi,i}_r dB_s\\ &-\sum_{j\in I}\int_s^T\widetilde{W}^{t,\varphi,i}_r(j)d\widetilde{\mathscr{V}}_s(j),\end{aligned} \tag{4.2}$$

where $\widetilde{f}\in M^2([t,T];L^2_\rho(\mathbb{R}^d;\mathbb{R}^1))$ and h satisfy (B.1). We denote by

$$A_r(j) := \{\omega\in\Omega|\alpha^{t,i}_r = j, j\in I; t\leq r\leq T\}.$$

Then for any $\varphi \in C_c^0(\mathbb{R}^d;\mathbb{R}^k)$, by virtue of the Hölder's inequality and Lemma 2.10, we have

$$
\begin{aligned}
E\Big[\int_{\mathbb{R}^d} h(X_T^{t,x,i})\varphi(x)dx\Big]^2 &\le E\Big[\int_{\mathbb{R}^d} |h(X_T^{t,x,i})|^2\rho^{-1}(x)dx\Big]E\Big[\int_{\mathbb{R}^d} \varphi(x)^2\rho(x)dx\Big] \\
&\le CE\Big[\int_{\mathbb{R}^d} |h(x)|^2\rho^{-1}(x)dx\Big] < +\infty.
\end{aligned}
$$

By similarly, we can get

$$
\begin{aligned}
E\Big[\int_{\mathbb{R}^d} \widetilde{f}(r,X_r^{t,x,i},\alpha_r^{t,i})\varphi(x)dx\Big]^2 &\le CE\Big[\int_{\mathbb{R}^d} |\widetilde{f}(r,X_r^{t,x,i},\alpha_r^{t,i})|^2\rho^{-1}(x)dx\Big] \\
&= CE\sum_{j\in I}\Big[\int_{\mathbb{R}^d} |\widetilde{f}(r,X_r^{t,x,i},j)|^2 I_{A_r(j)}\varphi(x)dx\Big]^2 \\
&\le C\sum_{j\in I}E\Big[\int_{\mathbb{R}^d} |\widetilde{f}(r,x,j)|^2\rho^{-1}(x)dx\Big] < +\infty.
\end{aligned}
$$

Then according to Tao, Wu and Zhang [14], BSDE (4.2) has a unique solution $\big(\widetilde{Y}_s^{t,\varphi,i},\widetilde{Z}_s^{t,\varphi,i},\widetilde{W}_s^{t,\varphi,i}(j)\big) \in S^2([t,T];\mathbb{R}^k)\otimes M^2([t,T];\mathbb{R}^{d\times k})\otimes N^2([t,T]\times I;\mathbb{R}^k)$.

The solution of our problem will be $(\widetilde{Y}_s^{t,\varphi,i},\widetilde{Z}_s^{t,\varphi,i},\widetilde{W}_s^{t,\varphi,i}(j))$ defined by $\varphi \mapsto (\widetilde{Y}_s^{t,\varphi,i},\widetilde{Z}_s^{t,\varphi,i},\widetilde{W}_s^{t,\varphi,i}(j))$ are linear functionals on $L^2_\rho(\mathbb{R}^d;\mathbb{R}^k)\times L^2_\rho(\mathbb{R}^d;\mathbb{R}^{d\times k})\times L^2_\rho(\mathbb{R}^d;\mathbb{R}^k)$.

We use the explicit form of $\widetilde{Y}_s^{t,\cdot,i}$

$$
\begin{aligned}
\widetilde{Y}_s^{t,\varphi,i} &= E^{\mathscr{F}_s}\Big[\int_{\mathbb{R}^d} h(X_T^{t,x,i})\varphi(x)dx + \int_s^T\int_{\mathbb{R}^d}\widetilde{f}(r,\alpha_r,X_r^{t,x,i})\varphi(x)dxds\Big] \\
&\le CE^{\mathscr{F}_s}\Big[\int_{\mathbb{R}^d} |h(X_T^{t,x,i})|^2\rho^{-1}(x)dx\Big]^{\frac{1}{2}}\Big[\int_{\mathbb{R}^d}|\varphi(x)|^2\rho(x)dx\Big]^{\frac{1}{2}} \\
&\quad + \int_s^T\Big[\int_{\mathbb{R}^d}|\widetilde{f}(r,\alpha_r^{t,i},X_r^{t,x,i})|^2\rho^{-1}(x)dx\Big]^{\frac{1}{2}}\Big[\int_{\mathbb{R}^d}|\varphi(x)|^2\rho(x)dx\Big]^{\frac{1}{2}}ds \\
&\le CE^{\mathscr{F}_s}\bigg\{\Big[\int_{\mathbb{R}^d} |h(X_T^{t,x,i})|^2\rho^{-1}(x)dx\Big]^{\frac{1}{2}} + \int_s^T\Big[\int_{\mathbb{R}^d}|\widetilde{f}(r,\alpha_r,X_r^{t,x,i})|^2\rho^{-1}(x)dx\Big]^{\frac{1}{2}}dr\bigg\} \\
&\qquad\qquad \times\Big[\int_{\mathbb{R}^d}|\varphi(x)|^2\rho(x)dx\Big]^{\frac{1}{2}}.
\end{aligned}
$$

Then by Lemma 2.10, we get

$$
\sup_{s\in[t,T]} E\big[||\widetilde{Y}_s^{t,\cdot,i}||_{L^2_\rho(\mathbb{R}^d;\mathbb{R}^k)}\big] = \sup_{s\in[t,T]} E \sup_{\varphi\in L^2_\rho(\mathbb{R}^d;\mathbb{R}^k)}\frac{|\widetilde{Y}_s^{t,\varphi,i}|}{||\varphi||_{L^2_\rho(\mathbb{R}^d;\mathbb{R}^k)}} < +\infty. \tag{4.3}
$$

Moreover, we can choose a version of $\widetilde{Y}_s^{t,\varphi,i}$ that is linear in φ. So, we get that $Y_s^{t,\cdot,i}$ is a functional and by Riesz representation theorem $\widetilde{Y}_\cdot^{t,\cdot,i} \in M^2([t,T];L^2_\rho(\mathbb{R}^d;\mathbb{R}^k))$. Next, we study $\widetilde{Z}_s^{t,\cdot,i}$ and $\widetilde{W}_s^{t,\cdot,i}(j)$ by the method of mollifiers.

We take h^m (resp. $\widetilde{f}^m$) as smooth functions which approximate h (resp. $\widetilde{f}$) in $L^2_\rho(\mathbb{R}^d)$ (resp. $L^2_\rho([t,T]\times\mathbb{R}^d)$). Denote by $(\widetilde{Y}^{t,x,i}_{s,m}, \widetilde{Z}^{t,x,i}_{s,m}, \widetilde{W}^{t,x,i}_{s,m})$ the solution of the following BSDE:

$$\widetilde{Y}^{t,x,i}_{s,m} = h^m(X^{t,x}_T) + \int_s^T \widetilde{f}^m(r, X^{t,x,i}_r, \alpha_r)ds - \int_s^T \widetilde{Z}^{t,x,i}_{r,m} dB_s - \sum_{j\in I}\int_s^T \widetilde{W}^{t,x,i}_{r,m}(j) d\widetilde{\mathscr{V}}_s(j). \tag{4.4}$$

We define $\widetilde{Y}^{t,\varphi,i}_{s,m} := \int_{\mathbb{R}^d} \widetilde{Y}^{t,x,i}_{s,m}\varphi(x)dx$, $\widetilde{Z}^{t,\varphi,i}_{s,m} := \int_{\mathbb{R}^d} \widetilde{Z}^{t,x,i}_{s,m}\varphi(x)dx$ and $\widetilde{W}^{t,x,i}_{s,m}(j) := \int_{\mathbb{R}^d} \widetilde{W}^{t,x,i}_{s,m}(j)$ $\varphi(x)dx$. By a standard calculus for BSDE with Markov chains (see e.g., [14] Theorem 3.2)

$$\begin{aligned}
&E\Big[\sup_{t\le s\le T}|\widetilde{Y}^{t,\varphi,i}_{s,m} - \widetilde{Y}^{t,\varphi,i}_s|^2 + \int_s^T |\widetilde{Z}^{t,\varphi,i}_{s,m} - \widetilde{Z}^{t,\varphi,i}_s|^2 + \sum_{j\in I}|\widetilde{W}^{t,\varphi,i}_{s,m}(j) - \widetilde{W}^{t,\varphi,i}_s(j)|^2 ds\Big]\\
&\le CE\Big[\int_{\mathbb{R}^d}|h(X^{t,x,i}_T) - h^m(X^{t,x,i}_T)|\varphi(x)dx\Big]^2\\
&\quad + CE\Big[\int_s^T\int_{\mathbb{R}^d}|\widetilde{f}(r, X^{t,x,i}_r, \alpha^{t,i}_r) - \widetilde{f}^m(r, X^{t,x,i}_r, \alpha^{t,i}_r)|\varphi(x)dx\Big]^2\\
&\le CE\Big[\int_{\mathbb{R}^d}|h(X^{t,x,i}_T) - h^m(X^{t,x,i}_T)|^2\rho^{-1}(x)dx\Big]\Big[\int_{\mathbb{R}^d}\varphi(x)^2\rho(x)dx\Big]\\
&\quad + CE\sum_{j\in I}\Big[\int_{\mathbb{R}^d}|\widetilde{f}(r, X^{t,x,i}_r, j) - \widetilde{f}^m(r, \alpha^{t,i}_r, X^{t,x,i}_r)|^2\rho^{-1}(x)dx\Big]\Big[\int_{\mathbb{R}^d}\varphi(x)^2\rho(x)dx\Big]\\
&\le C\int_{\mathbb{R}^d}|h(x) - h^m(x)|^2\rho^{-1}(x)dx + \sum_{j\in I}\int_s^T\int_{\mathbb{R}^d}|\widetilde{f}(r,j,x) - \widetilde{f}^m(r,j,x)|^2\rho^{-1}(x)dx\\
&\to 0, \quad \text{as} \quad m\to+\infty.
\end{aligned} \tag{4.5}$$

Let us define $u(t,\cdot,i) := \widetilde{Y}^{t,\cdot,i}_t$ and $u^m(t,\cdot,i) := \widetilde{Y}^{t,\cdot,i}_{t,m}$. By (4.5), we have as $m\to+\infty$

$$\int_{\mathbb{R}^d} u^m(t,x,i)\varphi(x)dx = \int_{\mathbb{R}^d}\widetilde{Y}^{t,x,i}_{t,m}\varphi(x)dx \to \int_{\mathbb{R}^d}\widetilde{Y}^{t,x,i}_t\varphi(x)dx = \int_{\mathbb{R}^d} u(t,x,i)\varphi(x)dx.$$

This implies that $u^m(t,\cdot,i)\to u(t,\cdot,i)$ weakly in the space $L^2_{\rho^{-1}}(\mathbb{R}^d)$. Moreover, using a change of variable $y = X^{t,x,i}_s$, we get

$$\widetilde{Y}^{t,\varphi,i}_{s,m} = \int_{\mathbb{R}^d} u^m(s, X^{t,x,i}_s, \alpha^{t,i}_s)\varphi(x)dx = \int_{\mathbb{R}^d} u^m(s,y,\alpha^{t,i}_s)\varphi^i_t(s,y)dy = \widetilde{Y}^{s,\Phi,\alpha^{t,i}_s}_{s,m},$$

where $\Phi(x) = \varphi^i_t(s,x)$ and $\varphi^i_t(s,x) = \varphi(\hat{X}^{t,x,i}_s)J(\hat{X}^{t,x,i}_s)$. A similar discussion to (4.5) gives

$$\begin{aligned}
&E\Big[\sup_{s\le s\le T}|\widetilde{Y}^{s,\Phi,\alpha^{t,i}_s}_{s,m} - \widetilde{Y}^{s,\Phi,\alpha^{t,i}_s}_s|^2 + \int_s^T |\widetilde{Z}^{s,\Phi,\alpha^{t,i}_s}_{s,m} - \widetilde{Z}^{s,\Phi,\alpha^{t,i}_s}_s|^2\\
&+ \sum_{j\in I}|\widetilde{W}^{s,\Phi,\alpha^{t,i}_s}_{s,m}(j) - \widetilde{W}^{s,\Phi,\alpha^{t,i}_s}_s(j)|^2 ds\Big] \to 0, \quad \text{as} \quad m\to+\infty.
\end{aligned}$$

So that

$$\widetilde{Y}_s^{t,\varphi,i} = \lim_{m\to+\infty} \widetilde{Y}_{s,m}^{t,\varphi,i} = \lim_{m\to+\infty} \widetilde{Y}_{s,m}^{s,\Phi,\alpha_s^{t,i}} = \widetilde{Y}_s^{s,\Phi,\alpha_s^{t,i}} = \int_{\mathbb{R}^d} u(s,x,\alpha_s^{t,i})\Phi(x)dx.$$

We get $\widetilde{Y}_s^{t,\cdot,i} = u(s,X_s^{t,\cdot,i},\alpha_s^{t,i})$. Furthermore, since $Z_{s,m}^{t,x,i} = (\nabla u^m\sigma)(s,X_s^{t,x,i},\alpha_s^{t,i})$, by using a change of variable and integration by parts formula, we have

$$\begin{aligned}
\widetilde{Z}_{s,m}^{t,\varphi,i} &= \int_{\mathbb{R}^d} (\nabla u\sigma)(s,X_s^{t,x,i},\alpha_s^{t,i})\varphi(x)dx \\
&= \int_{\mathbb{R}^d} (\nabla u\sigma)(s,x,\alpha_s^{t,i})\varphi_t^i(s,x)dx \\
&= \int_{\mathbb{R}^d} u(s,x,\alpha_s^{t,i})\big[-\nabla(\sigma(s,x,\alpha_s^{t,i})\varphi_t^i(s,x))\big]dx \\
&= \widetilde{Y}_{s,m}^{s,\Psi,\alpha_s^{t,i}},
\end{aligned}$$

where $\Psi(x) = -\nabla(\sigma(s,x,\alpha_s^{t,i})\varphi_t^i(s,x))$. Passing the limit in the above we get

$$\begin{aligned}
\widetilde{Z}_s^{t,\varphi,i} &= \widetilde{Y}_s^{s,\Psi,\alpha_s^{t,i}} \\
&= \int_{\mathbb{R}^d} u(s,x,\alpha_s^{t,i})\big[-\nabla(\sigma(s,x,\alpha_s^{t,i})\varphi_t^i(s,x))\big]dx \\
&= \int_{\mathbb{R}^d} (\nabla u\sigma)(s,X_s^{t,x,i},\alpha_s^{t,i})\varphi(x)dx.
\end{aligned}$$

That is $\widetilde{Z}_s^{t,\cdot,i} = (\nabla u\sigma)(s,X_s^{t,\cdot,i},\alpha_s^{t,i})$. Here ∇u is the weak derivative in Sobolev space. Moreover, from the boundness of σ and σ', we have

$$\begin{aligned}
&E\int_{\mathbb{R}^d} (\nabla u\sigma)(s,X_s^{t,x,i},\alpha_s^{t,i})\varphi(x)dx \\
&= -E\int_{\mathbb{R}^d} u(s,X_s^{t,x,i},\alpha_s^{t,i})\nabla(\sigma(s,X_s^{t,x,i},\alpha_s^{t,i})\varphi(x))dx \\
&\leq \left[E\int_{\mathbb{R}^d} u(s,X_s^{t,x,i},\alpha_s^{t,i})^2\rho^{-1}(x)dx\right]^{\frac{1}{2}} \left[E\int_{\mathbb{R}^d} [\sigma'(s,X_s^{t,x,i},\alpha_s^{t,i})\nabla X_s^{t,x,i}\varphi(x)]^2\rho(x)dx\right]^{\frac{1}{2}} \\
&\quad + \left[E\int_{\mathbb{R}^d} u(s,X_s^{t,x,i},\alpha_s^{t,i})^2\rho^{-1}(x)dx\right]^{\frac{1}{2}} \left[E\int_{\mathbb{R}^d} [\sigma(s,X_s^{t,x,i},\alpha_s^{t,i})\varphi'(x)]^2\rho(x)dx\right]^{\frac{1}{2}} \\
&\leq \Big[\sum_{j\in I} u(s,x,j)^2\rho^{-1}(x)dx\Big]^{\frac{1}{2}} \left[\int_{\mathbb{R}^d} \varphi(x)^2\rho(x)dx\right]^{\frac{1}{2}}.
\end{aligned}$$

This implies

$$E\big[||\widetilde{Z}_s^{t,\cdot,i}||_{L_\rho^2(\mathbb{R}^d;\mathbb{R}^k)}\big] = E\left[\sup_{\varphi\in L_\rho^2(\mathbb{R}^d;\mathbb{R}^k)} \frac{|\widetilde{Z}_s^{t,\varphi,i}|}{||\phi||_{L_\rho^2(\mathbb{R}^d;\mathbb{R}^k)}}\right] < +\infty.$$

So, $\widetilde{Z}_s^{t,\cdot,i}$ is also a distribution and $\widetilde{Z}_\cdot^{t,\cdot,i} \in M^2([t,T];L^2_\rho(\mathbb{R}^d;\mathbb{R}^{d\times k}))$. Furthermore,

$$\begin{aligned}\widetilde{W}_{s,m}^{t,\varphi,i}(j) &= \int_{\mathbb{R}^d} \left[u^m(s,X_s^{t,x,i},j) - u^m(s,X_s^{t,x,i},\alpha_{s-}^{t,i})\right]\varphi(x)dx \\ &= \int_{\mathbb{R}^d} \left[u(s,x,j) - u(s,x,\alpha_{s-}^{t,i})\right]\varphi_t^i(s,x)dx \\ &= \widetilde{Y}_{s,m}^{t,\Phi,j} - \widetilde{Y}_{s,m}^{t,\Phi,\alpha_{s-}^{t,i}}.\end{aligned}$$

Passing the limit as above we get

$$\begin{aligned}\widetilde{W}_s^{t,\varphi,i}(j) &= \widetilde{Y}_s^{s,\Phi,j} - \widetilde{Y}_s^{s,\Phi,\alpha_{s-}^{t,i}} \\ &= \int_{\mathbb{R}^d} \left[u(s,x,j) - u(s,x,\alpha_{s-}^{t,i})\right]\varphi_t^i(s,x)dx \\ &= \int_{\mathbb{R}^d} \left[u(s,X_s^{t,x,i},j) - u(s,X_s^{t,x,i},\alpha_{s-}^{t,i})\right]\varphi(x)dx.\end{aligned}$$

So, we get $\widetilde{W}_s^{t,\cdot,i}(j) = u(s,X_s^{t,\cdot,i},j) - u(s,X_s^{t,\cdot,i},\alpha_{s-}^{t,i})$ and it is clear that $\widetilde{W}_s^{t,\cdot,i}(j)$ is a linear functional as $\widetilde{Y}_s^{t,\cdot,i}$ is a linear functional. From (4.3) we get $\widetilde{W}_\cdot^{t,\cdot,i}(\cdot) \in N^2([t,T]\times I;L^2_\rho(\mathbb{R}^d;\mathbb{R}^k))$. So by Lemma 4.3, $\widetilde{Y}_\cdot^{t,\cdot,i} \in S^2([t,T];L^2_\rho(\mathbb{R}^d;\mathbb{R}^k))$.

Until now, we have proved that the linear BSDE (4.2) has a unique solution. We next discuss the nonlinear case by the Picard iteration.

Given $(Y_s^{t,x,i,N-1},Z_s^{t,x,i,N-1},W_s^{t,x,i,N-1}) \in M^2([t,T];\mathbb{R}^k)\times M^2([t,T];\mathbb{R}^k)\times N^2([t,T]\times I;\mathbb{R}^k)$, define $(Y_s^{t,x,i,N},Z_s^{t,x,i,N},W_s^{t,x,i,N})$ as follows:

$$\begin{aligned}Y_s^{t,x,i,N} =& h(X_T^{t,x,i}) + \int_s^T f(r,X_r^{t,x,i},Y_r^{t,x,i,N-1},Z_r^{t,x,i,N-1},\alpha_r^{t,i})dr - \int_s^T \langle Z_r^{t,x,i,N},dB_r\rangle \\ &- \sum_{j\in I}\int_s^T W_r^{t,x,i,N} d\widetilde{\mathscr{V}}_s(j).\end{aligned} \tag{4.6}$$

Let $(Y_s^{t,x,i,0},Z_s^{t,x,i,0},W_s^{t,x,i,0}) = (0,0,0)$. By condition (B.1) and (B.2) and Lemma 2.10, we know $h(\cdot)$ and $f(r,X_r^{t,x,i},0,0)$ satisfy the condition in Step 1. So equation (4.6) has a unique solution $(Y_\cdot^{t,\cdot,i,1},Z_\cdot^{t,\cdot,i,1},W_\cdot^{t,\cdot,i,1}) \in M^2([t,T];L^2_\rho(\mathbb{R}^d;\mathbb{R}^k))\times M^2([t,T];L^2_\rho(\mathbb{R}^d;\mathbb{R}^{d\times k}))\times N^2([t,T]\times I;L^2_\rho(\mathbb{R}^d;\mathbb{R}^k))$. Moreover, by virtue of Condition (B.2) and Lemma 2.10, we have

$$
\begin{aligned}
&E\int_s^T\int_{\mathbb{R}^d}|f(r,X_r^{t,x,i},Y_r^{t,x,i,1},Z_r^{t,x,i,1},\alpha_r^{t,i})|^2\rho^{-1}(x)dxdr\\
=&E\int_s^T\int_{\mathbb{R}^d}|f(r,X_r^{t,x,i},Y_r^{t,x,i,1},Z_r^{t,x,i,1},\alpha_r^{t,i})-f(r,X_r^{t,x,i},0,0,\alpha_r^{t,i})\\
&+f(r,X_r^{t,x,i},0,0,\alpha_r^{t,i})|^2\rho^{-1}(x)dxdr\\
\leq&2\sum_{j\in I}E\int_s^T\int_{\mathbb{R}^d}|f(r,X_r^{t,x,i},0,0,j)|^2\rho^{-1}(x)dxdr\\
&+CE\int_s^T\int_{\mathbb{R}^d}(|Y_r^{t,x,i,1}|^2+|Z_r^{t,x,i,1}|^2)\rho^{-1}(x)dxdr\\
<&+\infty.
\end{aligned}
$$

So both h and $f(r,X_r^{t,x,i},Y_r^{t,x,i,1},Z_r^{t,x,i,1})$ satisfy the condition in step 1. Following the same procedure, we obtain a sequence of $(Y_s^{t,x,i,N},Z_s^{t,x,i,N},W_s^{t,x,i,N})_{N=0,1,2,\cdots}$ which is from the iterated mapping (4.6) from $M^2([t,T];L^2_\rho(\mathbb{R}^d;\mathbb{R}^k))\times M^2([t,T];L^2_\rho(\mathbb{R}^d;\mathbb{R}^{d\times k}))\times N^2([t,T]\times I;L^2_\rho(\mathbb{R}^d;\mathbb{R}^k))$ to itself. Next we will prove that (4.6) is a contraction mapping. For this, define
$\overline{Y}_s^{t,x,i,N}=Y_s^{t,x,i,N}-Y_s^{t,x,i,N-1}$, $\overline{Z}_s^{t,x,i,N}=Z_s^{t,x,i,N}-Z_s^{t,x,i,N-1}$, $\overline{W}_s^{t,x,i,N}=W_s^{t,x,i,N}-W_s^{t,x,i,N-1}$,

$$
\overline{f}^N(s,x)=f(s,X_s^{t,x,i},Y_s^{t,x,i,N},Z_s^{t,x,i,N})-f(s,X_s^{t,x,i},Y_s^{t,x,i,N-1},Z_s^{t,x,i,N-1}),\quad N=1,2,3,\cdots
$$

$t\leq s\leq T$. Then, for a.e. $x\in\mathbb{R}^d$, we have

$$
\overline{Y}_s^{t,x,i,N}=\int_s^T\overline{f}^{N-1}(r,x)dr-\int_s^T\langle\overline{Z}_r^{t,x,i,N},dB_r\rangle-\sum_{j\in I}\int_s^T\overline{W}_r^{t,x,i,N}(j)d\widetilde{\mathscr{V}}_r(j).
$$

Applying Itô's formula to $e^{Ks}|\overline{Y}_s^{t,x,i,N}|^2$, by condition (B.2), we can deduce that

$$
\begin{aligned}
&\int_{\mathbb{R}^d}e^{Ks}|\overline{Y}_s^{t,x,i,N}|^2\rho^{-1}dx+K\int_s^T\int_{\mathbb{R}^d}e^{Kr}|\overline{Y}_r^{t,x,i,N}|^2\rho^{-1}(x)dxdr\\
&\quad+\int_s^T\int_{\mathbb{R}^d}e^{Kr}|\overline{Z}_r^{t,x,i,N}|^2\rho^{-1}(x)dxdr+\int_s^T\int_{\mathbb{R}^d}e^{Kr}\sum_{j\in I}|\overline{W}_r^{t,x,i,N}|^2\rho^{-1}(x)dxdr\\
\leq&\int_s^T\int_{\mathbb{R}^d}e^{Kr}(2C|\overline{Y}_s^{t,x,i,N}|^2+\frac{1}{2}|\overline{Y}_s^{t,x,i,N-1}|^2+\frac{1}{2}|\overline{Z}_s^{t,x,i,N-1}|^2)\rho^{-1}(x)dxdr \qquad (4.7)\\
&\quad-\int_s^T\int_{\mathbb{R}^d}2e^{Kr}|\overline{Y}_s^{t,x,i,N}||\overline{Z}_s^{t,x,i,N}|\rho^{-1}(x)dxdB_r\\
&\quad-\int_s^T\int_{\mathbb{R}^d}2e^{Kr}\sum_{j\in I}|\overline{W}_r^{t,x,i,N}||\overline{Y}_r^{t,x,i,N}|dxd\widetilde{\mathscr{V}}_r(j).
\end{aligned}
$$

Then we have

$$
\begin{aligned}
&(K-2C)E\Big[\int_s^T\int_{\mathbb{R}^d}e^{Kr}|\overline{Y}_r^{t,x,i,N}|^2\rho^{-1}(x)dxdr\Big]+E\Big[\int_s^T\int_{\mathbb{R}^d}e^{Kr}|\overline{Z}_r^{t,x,i,N}|^2\rho^{-1}(x)dxdr\Big]\\
&\quad+E\Big[\int_s^T\int_{\mathbb{R}^d}e^{Kr}\sum_{j\in I}|\overline{W}_r^{t,x,i,N}|^2\rho^{-1}(x)dxdr\Big]\\
&\le\frac{1}{2}E\Big[\int_s^T\int_{\mathbb{R}^d}|\overline{Y}_s^{t,x,i,N-1}|^2\rho^{-1}(x)dxdr\Big]+\frac{1}{2}E\Big[\int_s^T\int_{\mathbb{R}^d}|\overline{Z}_s^{t,x,i,N-1}|^2)\rho^{-1}(x)dxdr\Big]\\
&\le\frac{1}{2}E\Big[\int_s^T\int_{\mathbb{R}^d}|\overline{Y}_s^{t,x,i,N-1}|^2\rho^{-1}(x)dxdr\Big]+\frac{1}{2}E\Big[\int_s^T\int_{\mathbb{R}^d}|\overline{Z}_s^{t,x,i,N-1}|^2)\rho^{-1}(x)dxdr\Big]\\
&\quad+\frac{1}{2}E\Big[\int_s^T\int_{\mathbb{R}^d}e^{Kr}\sum_{j\in I}|\overline{W}_r^{t,x,i,N-1}|^2\rho^{-1}(x)dxdr\Big].
\end{aligned}
\tag{4.8}
$$

Letting $K=2C+1$, from the contraction principle, the mapping (4.6) has a triple of fixed point $(Y_\cdot^{t,\cdot,i},Z_\cdot^{t,\cdot,i},W_\cdot^{t,\cdot,i})$ that is the limit of the Cauchy sequence $\{(Y_\cdot^{t,\cdot,i},Z_\cdot^{t,\cdot,i},W_\cdot^{t,\cdot,i})\}_{N=1}^{+\infty}$ in $M^2([t,T];L^2_\rho(\mathbb{R}^d;\mathbb{R}^k))\times M^2([t,T];L^2_\rho(\mathbb{R}^d;\mathbb{R}^{d\times k}))\times N^2([t,T]\times I;L^2_\rho(\mathbb{R}^d;\mathbb{R}^k))$. We then prove $Y_\cdot^{t,\cdot,i}$ is also a limit of $Y_\cdot^{t,\cdot,i,N}$ in $S^2([t,T];L^2_\rho(\mathbb{R}^d;\mathbb{R}^k))$ as $N\to+\infty$. For this, we only need to prove $\{Y_\cdot^{t,\cdot,i,N}\}_{N=1}^{+\infty}$ a Cauchy sequence in $S^2([t,T];L^2_\rho(\mathbb{R}^d;\mathbb{R}^k))$. By virtue of the Burkholder-Davis-Gundy inequality, from (4.7), we have

$$
\begin{aligned}
&E\Big[\sup_{t\le s\le T}\int_{\mathbb{R}^d}e^{Ks}|\overline{Y}_s^{t,x,i,N}|^2\rho^{-1}(x)dx\Big]\le C_1E\Big[\int_s^T\int_{\mathbb{R}^d}e^{Kr}(|\overline{Y}_s^{t,x,i,N-1}|^2+|\overline{Z}_s^{t,x,i,N-1}|^2\\
&\quad+|\overline{Y}_s^{t,x,i,N}|^2+|\overline{Z}_s^{t,x,i,N}|^2+|\overline{W}_s^{t,x,i,N}|^2)\rho^{-1}(x)dxdr\Big],
\end{aligned}
\tag{4.9}
$$

where C_1 is a constant which is independent of N. Without loss of any generality, assume that $M \geq N$. Combine (4.8) and (4.9), we have

$$
\begin{aligned}
& E\Big[\sup_{t\leq s\leq T}\int_{\mathbb{R}^d} e^{Ks}|Y_s^{t,x,i,N}-Y_s^{t,x,i,M}|^2\rho^{-1}(x)dx\Big] \\
& \leq \sum_{l=N+1}^{M} E\Big[\sup_{t\leq s\leq T}\int_{\mathbb{R}^d} e^{Ks}|\overline{Y}_s^{t,x,i,l}|^2\rho^{-1}(x)dx\Big] \\
& \leq \sum_{l=N+1}^{M} C_1E\Big[\int_s^T\int_{\mathbb{R}^d} e^{Kr}(|\overline{Y}_s^{t,x,i,l-1}|^2+|\overline{Z}_s^{t,x,i,l-1}|^2+|\overline{Y}_s^{t,x,i,l}|^2+|\overline{Z}_s^{t,x,i,l}|^2 \\
& \qquad\qquad +|\overline{W}_s^{t,x,i,l}|^2)\rho^{-1}(x)dxdr\Big] \\
& \leq \sum_{l=N+1}^{M} \frac{3}{2}C_1E\Big[\int_s^T\int_{\mathbb{R}^d} e^{Kr}(|\overline{Y}_s^{t,x,i,l-1}|^2+|\overline{Z}_s^{t,x,i,l-1}|^2)\rho^{-1}(x)dxdr\Big] \\
& \leq \sum_{l=N+1}^{+\infty} (\frac{1}{2})^{l-2}3C_1E\Big[\int_s^T\int_{\mathbb{R}^d} e^{Kr}(|\overline{Y}_s^{t,x,i,1}|^2+|\overline{Z}_s^{t,x,i,1}|^2)\rho^{-1}(x)dxdr\Big] \\
& \to 0 \quad \text{as} \quad M,N\to+\infty.
\end{aligned}
$$

$\square$

5. Related PDEs: weak sense

In Section 4, we proved the existence and uniqueness of BSDE (2.15) in integrable space. In this section, we will discuss the related PDE in the Sobolev space.

We define $\mathscr{H}$ the set of functions $u(s,x,i)$ such that $(u,\nabla u\sigma)\in L^2([0,T]\times\mathbb{R}^d;\mathbb{R}^k)\otimes L^2([0,T]\times\mathbb{R}^d;\mathbb{R}^{d\times k})$ for each $i\in I$ with the norm $\Big(\int_0^T\int_{\mathbb{R}^d}(|\omega(s,x,i)|^2+|(\nabla\omega\sigma)(s,x,i)|^2 +\sum_{j\in I}|(\omega(s,x,j)-\omega(s,x,i)|^2\lambda_{ij}(s)dxds\Big)^{\frac{1}{2}}$. Following a standard argument as in the proof of the completeness of the Sobolev spaces, we can prove $\mathscr{H}$ is complete.

Definition 5.1. *We say that $u\in\mathscr{H}$ is a weak solution of the PDE* (3.1) *if u satisfies*

$$
\begin{aligned}
& \int_t^T\int_{\mathbb{R}^d} u(s,x,i)\partial_s\varphi(s,x)dxds+\int_{\mathbb{R}^d} u(t,x,i)\varphi(t,x)dx-\int_{\mathbb{R}^d} h(x)\varphi(T,x)dx \\
& \quad +\frac{1}{2}\int_t^T\int_{\mathbb{R}^d}(\sigma^T\nabla u)(s,x,i)(\sigma^T(t,x,i)\nabla\varphi(s,x))dxds \\
& \quad +\int_t^T\int_{\mathbb{R}^d} u(s,x,i)div((b-A)\varphi(s,x)dxds \\
& =\int_t^T\int_{\mathbb{R}^d}\varphi(s,x)f(s,x,u(s,x),(\nabla u\sigma)(s,x,i),i)dxds \\
& \quad +\sum_{j\in I}\int_t^T\int_{\mathbb{R}^d}\lambda_{ij}(s)(u(s,x,j)-u(s,x,i))\varphi(s,x)dxds
\end{aligned}
\tag{5.1}
$$

for every $\varphi \in C_c^{1,+\infty}([0,T]\times\mathbb{R}^d)$, *where* $A_j = \sum_{i=1}^d \frac{\partial}{\partial x_i}(\sigma\sigma^T)_{ij}(s,x,i)$ *and* $A = (A_1, A_2, \cdots, A_d)^\intercal$.

Theorem 5.2. *Under conditions (B.1)-(B.2), if we define* $u(t,x,i) = Y_t^{t,x,i}$, *where* $Y_s^{t,x,i}$ *is the solution of BSDE* (2.15)*, then* $u^i(t,x)$ *is the unique weak solution of Eq.*(3.1) *with* $u(T,x,i) = h(x)$*. Moreover,* $u(s, X_s^{t,x,i}) = Y_s^{t,x,i}$ *for a.e.* $s \in [t,T]$, $x \in \mathbb{R}^d$ *and for all* $i \in I$ *a.s.*

Proof. Existence. From Theorem 3.1, we know that $u^m(t,x,i)$ defined in the proof of Theorem 4.4 is the unique classical solution of the following PDE:

$$\begin{aligned} u^m(t,x,i) =& h(x) + \int_t^T \mathscr{L}u^m(s,x,i)ds + \int_t^T \widetilde{f}^m(s,x,i)ds \\ &+ \sum_{j\in I, j\neq i} \lambda_{ij}(t)(u^m(t,x,j) - u^m(t,x,i)). \end{aligned} \tag{5.2}$$

Then by the formula of integration by parts,

$$\begin{aligned} &\int_{\mathbb{R}^d} u^m(t,x,i)\varphi(t,x)dx - \int_{\mathbb{R}^d} h(x)\varphi(T,x)dx + \int_t^T\int_{\mathbb{R}^d} u^m(s,x,i)\partial_s\varphi(s,x)dxds \\ &= -\frac{1}{2}\int_t^T\int_{\mathbb{R}^d} (\sigma^\intercal\nabla u^m)(s,x,i)\sigma^\intercal(s,x,i)\nabla\varphi^i(s,x)dxds \\ &\quad - \int_{\mathbb{R}^d} u(s,x,i) div((b-A)\varphi(s,x))dxds + \int_t^T\int_{\mathbb{R}^d} \tilde{f}^m(s,x,i)\varphi(s,x)dxds \\ &\quad + \sum_{j\in I, j\neq i}\int_t^T\int_{\mathbb{R}^d} \lambda_{ij}(s)(u^m(s,x,j) - u^m(s,x,i))dxds. \end{aligned} \tag{5.3}$$

But by standard estimates

$$\begin{aligned} &E\Big[\int_t^T\int_{\mathbb{R}^d} (|\tilde{Y}_{s,m}^{t,x,i} - \tilde{Y}_s^{t,x,i}|^2 + |\tilde{Z}_{s,m}^{t,x,i} - \tilde{Z}_s^{t,x,i}|^2 \\ &\quad + \sum_{j\in I} |\tilde{W}_{s,m}^{t,x,i}(j) - \tilde{W}_s^{t,x,i}(j)|^2\lambda_{ij}(s))\rho^{-1}(x)dxds\Big] \to 0, \quad \text{as } m \to +\infty. \end{aligned}$$

And as m_1, $m_2 \to +\infty$,

$$\begin{aligned} &E\Big[\int_t^T\int_{\mathbb{R}^d} (|\tilde{u}^{m_1}(s,X_s^{t,x,i},i) - \tilde{u}^{m_2}(s,X_s^{t,x,i},i)|^2 + |\tilde{u}^{m_1}(s,X_s^{t,x,i},i) - \tilde{u}^{m_2}(s,X_s^{t,x,i},i)|^2 \\ &\quad + \sum_{j\in I} |\tilde{u}^{m_1}(s,X_s^{t,x,i},j) - \tilde{u}^{m_1}(s,X_s^{t,x,i},\alpha_{s-}^{t,i}) + \tilde{u}^{m_2}(s,X_s^{t,x,i},\alpha_{s-}^{t,i}) - \tilde{u}^{m_2}(s,X_s^{t,x,i},j)|^2 \\ &\qquad \lambda_{ij}(s))\rho^{-1}(x)dxds\Big] \to 0. \end{aligned} \tag{5.4}$$

Now by Lemma 2.10 and Eq. (5.4), we can see that $\tilde{u}^m$ is a Cauchy sequence in $\mathscr{H}$. So there exists $\tilde{u} \in \mathscr{H}$ such that $(\tilde{u}^m, \nabla \tilde{u}^m \sigma) \to (\tilde{u}, \nabla \tilde{u} \sigma)$ in $L^2([0,T] \times \mathbb{R}^d; \mathbb{R}^k) \otimes L^2([0,T] \times \mathbb{R}^d; \mathbb{R}^{d\times k}))$. Moreover $\tilde{Y}_s^{t,x,i} = \tilde{u}(s, X_s^{t,i}, \alpha_s^{t,i})$, $\tilde{Z}_s^{t,x,i} = (\nabla \tilde{u} \sigma)(s, X_s^{t,i}, \alpha_s^{t,i})$ and $\tilde{W}_s^{t,x,i}(j) = \tilde{u}(s, X_s^{t,i}, j) - \tilde{u}(s, X_s^{t,i}, \alpha_{s-}^{t,i})$ for a.e. $s \in [t,T]$, $x \in \mathbb{R}^d$ a.s.. Now it is easy to pass the limit as $m \to +\infty$ in (5.3) to get $\tilde{u}$ which solves Eq. (3.1).

For the nonlinear case, we set $\widetilde{f}(t,x,i) = f(t,x,u(t,x,i),(\nabla u\sigma)(t,x,i),i)$. From (A.2) we know that f satisfies the condition of $\tilde{f}$. Then $v = \nabla u \sigma$ and u solves the PDE (5.2) with $\widetilde{f}(s,x,i) = f(s,x,u(s,x,i),(\nabla u\sigma)(s,x,i),i)$. Moreover it's easy to check that Eq. (5.3) coincides with the Eq. (5.1) with $\widetilde{f} = f$. So we have obtained a solution of the PDE (5.1).

Uniqueness. Let u be a weak solution of (3.1). From the proof of Theorem 4.4 we know, $Y_\cdot^{t,\cdot,i} := u(\cdot, X_\cdot^{t,\cdot,i}, \alpha_s^{t,i})$, $Z_\cdot^{t,\cdot,i} := (\nabla u \sigma)(\cdot, X_\cdot^{t,\cdot,i}, \alpha_s^{t,i})$ is the unique solution of the following BSDE:

$$
\begin{aligned}
Y_s^{t,x,i} =& h(X_T^{t,x,i}) + \int_s^T f(r, X_r^{t,x,i}, u(r, X_r^{t,x,i}, \alpha_r^{t,i}), (\nabla u \sigma)(r, X_r^{t,x,i}, \alpha_r^{t,i}))dr \\
&- \int_s^T Z_r^{t,x,i} dr - \sum_{j\in I} \int_s^T W_r^{t,x,i}(j) d\widetilde{\mathscr{V}}_r.
\end{aligned}
$$

Then the uniqueness of the PDE (5.1) follows from the uniqueness of the solution of the BSDE (2.15). □

6. Concluding remarks

In this paper, we studied the BSDEs with Markov chains in Sobolev space and the associated PDEs both in classical and weak sense. The existence and uniqueness results of the solutions to the BSDEs and associated PDEs were given. Our results provide a powerful tool to solve models with Markov chains in practice by virtue of the PDEs technique. For the studies of the BSDEs and PDEs in Sobolev space, we generalized the stochastic flow theory associated with SDEs driven by both Brownian motion and Markov chains, in which the coefficients b and σ are smooth enough. If b and σ only satisfy the Lipschitz condition and the BSDEs driven by Markov chains and doubly Brownian motions, it is difficult to prove the stochastic flow results and the existence and uniqueness of the weak solution to the associated stochastic PDEs. We shall come back to this case in the future work.

Acknowledgements

This work was supported by the Natural Science Foundation of China (11831010, 61961160732), and Shandong Provincial Natural Science Foundation (ZR2019ZD42) and The Taishan Scholars Climbing Program of Shandong (TSPD20210302). HZ would like to acknowledge the financial supports of a EPSRC Established Career Fellowship grant (EP/S005293/2)

Competing interests

The authors declare that they have no competing interests.

Declarations of interest: none

References

[1] E. Pardoux, S. Peng, Adapted solution of a backward stochastic differential equation, Systems and; Control Letters 14 (1) (1990) 55–61. doi:10.1016/0167-6911(90)90082-6.

[2] S. Peng, Z. Wu, Fully coupled forward-backward stochastic differential equations and applications to optimal control, SIAM Journal on Control and Optimization 37 (3) (1999) 825–843. doi:10.1137/S0363012996313549.

[3] C. Feng, X. Wang, H. Zhao, Quasi-linear pdes and forward–backward stochastic differential equations: Weak solutions, Journal of Differential Equations 264 (2) (2018) 959–1018. doi:10.1016/j.jde.2017.09.030.

[4] Z. Wu, Z. Yu, Dynamic programming principle for one kind of stochastic recursive optimal control problem and hamilton–jacobi–bellman equation, SIAM Journal on Control and Optimization 47 (5) (2008) 2616–2641. doi:10.1137/060671917.

[5] L. Wei, Z. Wu, H. Zhao, Sobolev weak solutions of the hamilton–jacobi–bellman equations, SIAM Journal on Control and Optimization 52 (3) (2014) 1499–1526. doi:10.1137/120889174.

[6] N. El Karoui, S. Peng, M. C. Quenez, Backward stochastic differential equations in finance, Mathematical finance 7 (1) (1997) 1–71. doi:10.1111/1467-9965.00022.

[7] S. Peng, Probabilistic interpretation for systems of quasilinear parabolic partial differential equations, Stochastics and stochastics reports 37 (1-2) (1991) 61–74. doi:10.1080/17442509108833727.

[8] E. Pardoux, S. Peng, Backward stochastic differential equations and quasilinear parabolic partial differential equations, in: B. L. Rozovskii, R. B. Sowers (Eds.), Stochastic Partial Differential Equations and Their Applications, Springer Berlin Heidelberg, Berlin, Heidelberg, 1992, pp. 200–217.

[9] G. Barles, R. Buckdahn, E. Pardoux, Backward stochastic differential equations and integral-partial differential equations, Stochastics: An International Journal of Probability and Stochastic Processes 60 (1-2) (1997) 57–83. doi:10.1080/17442509708834099.

[10] G. Barles, E. Lesigne, Sde, bsde and pde, Pitman Research Notes in Mathematics Series (1997) 47–82.

[11] V. Bally, A. Matoussi, Weak solutions for spdes and backward doubly stochastic differential equations, Journal of Theoretical Probability 14 (1) (2001) 125–164. doi:10.1023/A:1007825232513.

[12] S. N. Cohen, R. J. Elliott, et al., Comparisons for backward stochastic differential equations on markov chains and related no-arbitrage conditions, The Annals of Applied Probability 20 (1) (2010) 267–311. doi:10.1214/09-AAP619.

[13] S. N. Cohen, R. J. Elliott, Solutions of backward stochastic differential equations on markov chains, Communications on Stochastic Analysis 2 (2) (2008) 5. doi:10.31390/cosa.2.2.05.

[14] R. Tao, Z. Wu, Q. Zhang, Bsdes with regime switching: Weak convergence and applications, Journal of Mathematical Analysis and Applications 407 (1) (2013) 97–111. doi:10.1016/j.jmaa.2013.05.011.

[15] H. Kunita, Stochastic differential equations and stochastic flows of diffeomorphisms, in: Ecole d'été de probabilités de Saint-Flour XII-1982, Springer, 1984, pp. 143–303.

[16] S. Crépey, About the pricing equations in finance, in: Paris-Princeton Lectures on Mathematical Finance 2010, Springer, 2011, pp. 63–203.

[17] S. Crépey, A. Matoussi, et al., Reflected and doubly reflected bsdes with jumps: a priori estimates and comparison, The Annals of Applied Probability 18 (5) (2008) 2041–2069. `doi:10.1214/08-AAP517`.

[18] X. Mao, C. Yuan, Stochastic differential equations with Markovian switching, Imperial college press, 2006.

[19] N. Ikeda, S. Watanabe, Stochastic differential equations and diffusion processes, Vol. 24, Elsevier, 2014.

[20] Q. Zhang, H. Zhao, Stationary solutions of spdes and infinite horizon bdsdes, Journal of Functional Analysis 252 (1) (2007) 171–219. `doi:10.1016/j.jfa.2007.06.019`.