J Syst Sci Complex (2010) 23: 546–571

# FOUR STEP SCHEME FOR GENERAL MARKOVIAN FORWARD-BACKWARD SDES*

**Jin MA · Jiongmin YONG · Yanhong ZHAO**

DOI: 10.1007/s11424-010-0145-8
Received: 2 September 2009 / Revised: 4 February 2010


**Abstract** This paper studies a class of forward-backward stochastic differential equations (FBSDE) in a general Markovian framework. The forward SDE represents a large class of strong Markov semi-martingales, and the backward generator requires only mild regularity assumptions. The authors show that the Four Step Scheme introduced by Ma, et al. (1994) is still effective in this case. Namely, the authors show that the adapted solution of the FBSDE exists and is unique over any prescribed time duration; and the backward components can be determined explicitly by the forward component via the classical solution to a system of parabolic integro-partial differential equations. An important consequence the authors would like to draw from this fact is that, contrary to the general belief, in a Markovian set-up the martingale representation theorem is no longer the reason for the well-posedness of the FBSDE, but rather a consequence of the existence of the solution of the decoupling integral-partial differential equation. Finally, the authors briefly discuss the possibility of reducing the regularity requirements of the coefficients by using a scheme proposed by F. Delarue (2002) to the current case.

**Key words** Forward-backward stochastic differential equations, Four Step Scheme, parabolic integro-partial differential equation, strong Markov semi-martingales.

## 1 Introduction

A fully coupled forward-backward stochastic differential equation (FBSDE, for short) can be written in the following general form:

$$\begin{cases} X_t = x + \int_0^t b(s, X_{s-}, Y_{s-}, Z_{s-}) * d\xi_s, \\ Y_t = g(X_T) + \int_t^T h(s, X_s, Y_s, Z_s)ds - \int_t^T Z_{s-} * d\xi_s, \end{cases} \tag{1}$$

Jin MA
*Department of Mathematics, University of Southern California, Los Angeles, CA 90089, USA.*
Email: jinma@usc.edu.
Jiongmin YONG
*Department of Mathematics, University of Central Florida, Orlando, FL 32816, USA.*
Email: jyong@mail.ucf.edu.
Yanhong ZHAO
*Department of Mathematics, Purdue University, West Lafayette, IN 47907, USA.*
Email: yhzhao@math.purdue.edu.
*This research is supported by the National Science Foundation under Grant Nos. #DMS 0505472, 0806017, and #DMS 0604309.

where $\xi$ is a (multidimensional) semimartingale, and the coefficients $b$, $h$, $g$ are given deterministic functions with suitable dimensions. The notation "$*$" here is to emphasize that the semimartingale may contain a random measure, and in that case stochastic integral should be understood accordingly (this will be made precise in Section 2). The solution to (1) is a triplet $(X,Y,Z)$ of processes/random fields which is predictable with respect to a given filtration $\mathbb{F}\equiv\{\mathcal{F}_t\}_{t\ge0}$, usually assumed to be generated by the driving semimartingale $\xi$.

Since early 1990s, tremendous efforts have been made towards the understanding of the existence and uniqueness of adapted solutions to (1). While in general it is not clear whether (1) would even have a solution, special cases have been explored quite extensively. For example, if $\xi_t=(t,W_t)^{\mathrm{T}}$, with $W_t$ being a standard Brownian motion on $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\ge0},\mathbb{P})$, several methods have proved to be effective (see, for example, [1–6]). Among others, the Four Step Scheme proposed in [4] extends the nonlinear Feynman-Kac formula to the quasilinear case, and has been widely used as a method for solving FBSDEs in applications. Another case that has been well-understood is when $\xi_t=(t,W_t,\widetilde{N}_p(t,\cdot))^{\mathrm{T}}$, with $W_t$ being a standard Brownian motion and $\widetilde{N}_p$ being a compensated Poisson random measure, in such a case the well-posedness was established under a certain monotonicity conditions (cf. e.g, [7–8]), without using the Four Step Scheme.

In this paper we would like to look at the FBSDE from a slightly different perspective, by revisiting the Four Step Scheme for solving such equations. First of all, it has been long understood that a fundamental building block for the theory of backward SDE (BSDE) is the Martingale Representation Theorem. Therefore when a FBSDE of the form (1) is formulated, one often assumes that the filtration is Brownian. However, the Four Step Scheme seems to tell us a slightly different story. Recall that the essence of that scheme is to assume that the backward components $Y$ and $Z$ are represented as certain deterministic functions of the forward component $X$, and these deterministic functions can be obtained by solving a system of quasilinear parabolic PDEs. Once such "decoupling" functions are found, the forward component $X$ is then a diffusion process, and in the whole procedure the martingale representation is not needed, instead the (strong) Markovian property of the process $X$ plays an essential role. Consequently the requirement on the filtration seems to be irrelevant.

We now try to extend the idea further. Let us first look at the forward equation in (1), without the component $(Y,Z)$ (this can be done by simply assuming that $Y$ and $Z$ are some functions of $X$). Clearly, to serve our purpose it is natural to assume that it is both a Markov process and a semimartingale. It is known (see, e.g., [9–10]) that essentially all "reasonable" strong Markov semimartingales (with quasi-left-continuous path, or Hunt processes, for example) are the solutions to the SDEs of the form

$$X_t=x+\int_0^t b(X_r)dr+\int_0^t\sigma(X_r)dW_r+\int_0^t\int_{\mathbb{R}}\alpha(X_{r-},z)\widetilde{\Gamma}(drdz),\tag{2}$$

where $X_t\in\mathbb{R}^n$, $W$ is a $d$-dimensional Brownian motion, and $\widetilde{\Gamma}$ is a martingale random measure on $[0,\infty)\times\mathbb{R}^d$. A more workable form of (2) is the following SDE that contains "evolving intensities", that is, the jump intensity depends on the solution (see, e.g, [11–12]):

$$X_t=x+\int_0^t b(X_r)dr+\int_0^t\sigma(X_r)dW_r+\int_{[0,t]\times\mathbb{R}\times\mathbb{R}_+}\alpha(X_{r-},z)\mathbf{1}_{[0,\lambda(X_s)]}(u)\widetilde{N}(drdudz),\tag{3}$$

where $\widetilde{N}$ is a compensated Poisson random measure on $[0,T]\times\mathbb{R}_+\times\mathbb{R}^n$ with mean measure $m\times\nu\times m$, where $m$ is the Lebesgue measure, and $\nu$ is a Lévy measure. We note that a slightly more general version of (3) can also be considered, but we prefer this form for now as it is closer to the standard set-up (see discussions in Sections 2 and 3).

The above discussion leads us to consider, naturally, FBSDE (1) in a more extended form, within the Markovian paradigm. For notational clarity, from now on, we shall separate the martingale integrands for the continuous part and jump part.

$$
\begin{cases}
X_t = x + \int_0^t b(s, X_s, Y_s, Z_s, \widehat{Z}_s(\cdot))ds + \int_0^t \sigma(s, X_s, Y_s)dW_s \\
\qquad + \int_0^t \int_{\mathbb{R}_+ \times \mathbb{R}^n} \theta(s, X_{s-}, Y_{s-}, \lambda)\mathbf{1}_{[0,\eta(s,X_{s-},\lambda)]}(r)\widetilde{N}(drd\lambda ds), \\
Y_t = g(X_T) + \int_t^T h(s, X_s, Y_s, Z_s, \widehat{Z}_s(\cdot))ds - \int_t^T Z_s dW_s \\
\qquad - \int_t^T \int_{\mathbb{R}_+ \times \mathbb{R}^n} \widehat{Z}_{s-}(\lambda)\mathbf{1}_{[0,\eta(s,X_{s-},\lambda)]}(r)\widetilde{N}(drd\lambda ds),
\end{cases}
\tag{4}
$$

where $X_t \in \mathbb{R}^n$, $Y_t \in \mathbb{R}^m$, $W_t \in \mathbb{R}^d$ (the dimensions of $Z$ and $\widehat{Z}$ are then obvious); the coefficients $b$, $\sigma$, $h$, $\theta$, and $\eta$ are all deterministic functions with appropriate dimensions. In light of the Four Step Scheme of [4], we see that the problem will then be converted to solving the following system of quasilinear integro-partial differential equations (IPDEs):

$$
\begin{cases}
0 = u_t^k(t,x) + \frac{1}{2}\mathrm{tr}\left\{u_{xx}^k \sigma(t,x,u)\sigma(t,x,u)^{\mathrm{T}}\right\} \\
\quad + \langle u_x^k, b\big(t,x,u,\sigma(t,x,u)^{\mathrm{T}}u_x, u(t,x+\theta(t,x,u,\cdot)) - u(t,x)\big)\rangle \\
\quad + h^k(t,x,u,\sigma(t,x,u)^{\mathrm{T}}u_x, u(t,x+\theta(t,x,u,\cdot)) - u(t,x)) \\
\quad + \int_{\mathbb{R}^n} \eta(t,x,\lambda)\big\{u^k(t,x+\theta(t,x,u,\lambda)) - u^k(t,x) - \langle u_x^k, \theta(t,x,u,\lambda)\rangle\big\}\nu(d\lambda), \\
\qquad\qquad (t,x) \in [0,T] \times \mathbb{R}^n, \quad k = 1, 2, \cdots, m, \\
u(T,x) = g(x), \qquad x \in \mathbb{R}^n.
\end{cases}
\tag{5}
$$

Here, unless otherwise specified, all functions $u$, $u_x$, and $u_{xx}$ are evaluated at $(t,x)$. Our main purpose is then to show that as long as the above equation admits a classical solution, FBSDE (4) will have an adapted solution. Such a solution will in turn provide a probabilistic representation in terms of the solution $u$ of (5). However, the issue of well-posedness of the IPDE (5) is by no means trivial, since it is a system of parabolic IPDEs that contains spatial non-local terms. We shall prove that, under certain conditions on the coefficients, (5) does admit a unique classical solution, hence our solution scheme is validated.

At this point we should point out that all the previous results involving FBSDEs with jumps (e.g., [7–8]) gave only the existence and uniqueness of the adapted solutions, and they are not quite constructive. The advantage of our solution scheme here inherited the features as in the continuous case, that is, it provides an explicit representation for the adapted solution to FBSDE (4), and it can be considered as a generalized non-linear Feynman-Kac formula.

The rest of this paper is organized as follows. In Section 2, we formulate the problem, introducing standing assumptions, and give some preliminaries on SDEs under general Markovian assumptions. In Section 3, we use the idea of Four Step Scheme to derive the solution scheme applicable to FBSDE (4). Section 4 is devoted to the solvability of system (5), which will lead to the unique solvability of FBSDE (4) in Section 5. Some possible extensions will be carried out in Section 6.

## 2 Preliminaries

In this section we give a detailed formulation of our problem. To begin with, let us first make some notational convention. For any $n \geq 1$, let $|\cdot|$ and $\langle \cdot, \cdot \rangle$ be the usual Euclidean norm and inner product in $\mathbb{R}^n$, respectively. For any $m, n \geq 1$, $z, \widetilde{z} \in \mathbb{R}^{m\times n}$, let $\langle z, \widetilde{z} \rangle = \operatorname{tr}(z\widetilde{z}^{\mathrm{T}})$ (where $\widetilde{z}^{\mathrm{T}}$ is the transpose of $\widetilde{z}$). Then $\langle \cdot, \cdot \rangle$ is an inner product under which $\mathbb{R}^{m\times n}$ is a Hilbert space whose induced norm is denoted by $|\cdot|$. We will also let $L^p([0,T] \times \mathbb{R}^n; H)$ be the set of all $H$-valued measurable functions $\varphi(\cdot,\cdot)$ defined on $[0,T]\times\mathbb{R}^n$ such that $\int_0^T \int_{\mathbb{R}^n} |\varphi(t,x)|^p dxdt < \infty$, for any $p \in [1,\infty)$ with $H$ being any Euclidean space ($\mathbb{R}^m, \mathbb{R}^{m\times n}$, etc.). The case $p = \infty$ is defined in an obvious way. More function spaces will be introduced when they are needed.

### 2.1 Stochastic Differential Equations for Markov Processes

Let $(U, \mathscr{U})$ be a measurable space and $\nu$ be a $\sigma$-finite measure on $(U, \mathscr{U})$. Let

$$\theta : [0,\infty) \times \mathbb{R}^d \times U \mapsto \mathbb{R}^n \quad \text{and} \quad \eta : [0,\infty) \times \mathbb{R}^d \times U \mapsto [0,\infty)$$

be two continuous functions. Further, we assume that $\eta(t,x,\lambda) > 0$ for all $(t,x,u) \in [0,\infty) \times \mathbb{R}^d \times U$.

Let $N$ be a Poisson random measure on $[0,\infty)\times U \times [0,\infty)$ with mean measure $m \times \nu \times m$, and $\widetilde{N}(A) = N(A) - (m\times\nu\times m)(A)$, for $A \in \mathcal{B}([0,\infty)\times U\times[0,\infty))$. Further, we assume that there exists a fixed set $U_1 \in \mathscr{U}$ such that, with $U_2 \triangleq U \setminus U_1$,

$$\Psi_1(t,x) \triangleq \int_{U_1} \eta(t,x,\lambda)|\theta(t,x,\lambda)|^2 \nu(d\lambda) < \infty, \tag{6}$$

$$\Psi_2(t,x) \triangleq \int_{U_2} \eta(t,x,\lambda)|\theta(t,x,\lambda)|\nu(d\lambda) < \infty; \tag{7}$$

or a stronger version that combines both (6) and (7):

$$\int_U \eta(t,x,\lambda)\Big\{|\theta(t,x,\lambda)|^2 + I_{U_2}(\lambda)\Big\}\nu(d\lambda) < \infty. \tag{8}$$

We also assume that

$$\int_U \eta(t,x,\lambda)|\theta(t,x,\lambda)|[I_{U_1}(\lambda) - I_{B_1}(\theta(t,x,\lambda))]\nu(d\lambda) < \infty, \tag{9}$$

where $B_1$ is the ball of radius 1 centered at the origin.

We shall make use of the following standing assumptions.

**H1** The functions $\sigma, b, \Psi_1, \Psi_2$ are all deterministic bounded continuous.

**H2** There exist a nonnegative function $\eta_0(\lambda)$ and a constant $M > 0$ such that for all $(t,x) \in [0,\infty)\times\mathbb{R}^d$, it holds that

$$\int_{U_1} |\theta(t,x,u)|^2 \eta_0(\lambda)\nu(d\lambda) + \int_{U_1} |\eta(t,x,\lambda) - \eta_0(\lambda)||\theta(t,x,\lambda)|\nu(d\lambda) \leq M.$$

**H3** There exists a constant $L > 0$ such that

$$
\begin{cases}
|b(t,x)-b(t,y)|+|\sigma(t,x)-\sigma(t,y)|\le L|x-y|, \\
\displaystyle\int_{U_1}|\theta(t,x,\lambda)-\theta(t,y,\lambda)|^2\eta_0(\lambda)\nu(d\lambda)\le L|x-y|^2, \\
\displaystyle\int_{U_1}|\theta(t,x,\lambda)-\theta(t,y,\lambda)|\,|\eta(t,x,\lambda)-\eta_0(\lambda)|\nu(d\lambda)\le L|x-y|, \\
\displaystyle\int_{U_2}|\theta(t,x,\lambda)-\theta(t,y,\lambda)|\eta(t,x,\lambda)\nu(d\lambda)\le L|x-y|, \\
\displaystyle\int_{U}|\theta(t,y,\lambda)|\,|\eta(t,x,\lambda)-\eta(t,y,\lambda)|\nu(d\lambda)\le L|x-y|, \quad t\in[0,T],\ x,y\in\mathbb{R}^n.
\end{cases}
\tag{10}
$$

We now consider the following general SDE with jumps:

$$
\begin{aligned}
X_t = X_0 &+ \int_0^t \sigma(s,X_s)dW_s + \int_0^t b(s,X_s)ds \\
&+ \int_{[0,\infty)\times U_1\times[0,t]} \theta(s,X_{s-},\lambda)\mathbf{1}_{[0,\eta(s,X_{s-},\lambda)]}(r)\widetilde{N}(drd\lambda ds) \\
&+ \int_{[0,\infty)\times U_2\times[0,t]} \theta(s,X_{s-},\lambda)\mathbf{1}_{[0,\eta(s,X_{s-},\lambda)]}(r)N(drd\lambda ds).
\end{aligned}
\tag{11}
$$

The following result can be found in [12].

**Theorem 1**[12] *Assume* H1–H3. *Then SDE* (11) *has a unique solution. Furthermore, the solution $X$ is a Markov process with generator:*

$$
\begin{aligned}
&A(t)f(x) \\
&= \frac{1}{2}\sum_{i,j=1}^n a_{ij}(t,x)\frac{\partial^2}{\partial_{x_i}\partial_{x_j}}f(x) + \langle \widehat{b}(t,x),\nabla f(x)\rangle \\
&\quad + \int_U \eta(t,x,\lambda)\big\{f(x+\theta(t,x,\lambda))-f(x)-\mathbf{1}_{U_1}(\lambda)\langle\theta(t,x,\lambda),\nabla f(x)\rangle\big\}\nu(d\lambda),
\end{aligned}
$$

*where*

$$
\begin{aligned}
a_{ij}(t,x) &= \sum_{k=1}^d \sigma_{ik}(t,x)\sigma_{jk}(t,x), \quad i,j=1,2,\cdots,n; \\
\widehat{b}(t,x) &= b(t,x) + \int_U \eta(t,x,\lambda)\theta(t,x,\lambda)(\mathbf{1}_{U_1}(\lambda)-\mathbf{1}_{B_1}(\theta(t,x,\lambda)))\nu(d\lambda).
\end{aligned}
$$

### 2.2 Formulation of FBSDEs

The idea of SDE of a Markov process can be naturally extended to the case of FBSDEs. Let us consider the forward-backward version of the above equation:

$$
\begin{cases}
X_t = x + \int_0^t b(s, X_s, Y_s, Z_s, \widehat{Z}_s(\cdot))ds + \int_0^t \sigma(s, X_s, Y_s)dW_s \\
\qquad + \int_0^t \int_{\mathbb{R}_+ \times U_1} \theta(s, X_{s-}, Y_{s-}, \lambda)\mathbf{1}_{[0,\eta(s,X_{s-},\lambda)]}(r)\widetilde{N}(drd\lambda ds) \\
\qquad + \int_0^t \int_{\mathbb{R}_+ \times U_2} \theta(s, X_{s-}, Y_{s-}, \lambda)\mathbf{1}_{[0,\eta(s,X_{s-},\lambda)]}(r)N(drd\lambda ds), \\
Y_t = g(X_T) + \int_t^T h(s, X_s, Y_s, Z_s, \widehat{Z}_s(\cdot))ds - \int_t^T Z_s dW_s \\
\qquad - \int_t^T \int_{\mathbb{R}^+ \times U_1} \widehat{Z}_{s-}(\lambda)\mathbf{1}_{[0,\eta(s,X_{s-},\lambda)]}(r)\widetilde{N}(drd\lambda ds) \\
\qquad - \int_t^T \int_{\mathbb{R}^+ \times U_2} \widehat{Z}_{s-}(\lambda)\mathbf{1}_{[0,\eta(s,X_{s-},\lambda)]}(r)N(drd\lambda ds),
\end{cases}
\tag{12}
$$

where the processes $X$, $Y$, $Z$, and $\widehat{Z}(\cdot)$ take values in $\mathbb{R}^n$, $\mathbb{R}^m$, $\mathbb{R}^{m\times d}$, and $\mathbb{R}^m$, respectively; $T > 0$ is an arbitrary prescribed time duration; and the coefficients $b$, $\sigma$, $\theta$, $h$, and $g$ are the mappings of the following forms:

$$
\begin{cases}
b : [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \times L^2(U; d\nu) \mapsto \mathbb{R}^n, \\
\sigma : [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}^{n\times n}, \\
\theta : [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \mapsto L^1(U; d\nu) \bigcap L^2(U; d\nu), \\
h : [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \times L^2(U; d\nu) \mapsto \mathbb{R}^m,
\end{cases}
\tag{13}
$$

Our objective is to find a quadruple $(X, Y, Z, \widehat{Z})$ which is $\{\mathcal{F}_t\}_{t\geq 0}$-adapted, square integrable, such that the Equation (12) is satisfied on $[0,T]$, $P$-almost surely. More precisely, we have the following definition.

**Definition 1** A quartuple of $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \times (L^1 \bigcap L^2)(d\nu)$-valued processes $(X, Y, Z, \widehat{Z})$ is called an adapted solution of FBSDE (12) if it is $\{\mathcal{F}_t\}_{t\geq 0}$-adapted and square-integrable, such that (12) holds $P$-almost surely.

We shall make use of the following more standing assumptions.

**H4** $d = n$.

**H5** The functions $b$, $\sigma$, $\theta$, $h$, and $g$ are all continuous in all variables, smooth in $(x, y, z)$, and are Frechét differentiable in $\widehat{z}(\cdot)$, whenever applicable. Furthermore, the first order derivatives in $x, y, z$ and the Frechét derivative in $\widehat{z}(\cdot)$ are all bounded by a generic constant $L > 0$.

**H6** For any $(t, x, y, z, \widehat{z}(\cdot)) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \times (L^1 \bigcap L^2)(U; d\nu)$, and $p = 1, 2$, it holds that

$$
|b(t, x, 0, 0, 0)| + |\sigma(t, x, y)| + \|\theta(t, 0, 0, \cdot)\|_{L^p} + |h(t, x, 0, z, \widehat{z}(\cdot))| \leq L, \quad t \in [0,T].
$$

**H7** The function $\sigma(\cdot, \cdot, \cdot)$ satisfies the following ellipticity condition:

$$
\sigma(t, x, y)\sigma(t, x, y)^{\mathrm{T}} \geq \sigma_0(|y|)I, \qquad \forall (t, x, y) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^m, \tag{14}
$$

for some positive continuous non-increasing function $\sigma_0 : [0, \infty) \to (0, \infty)$.

**H8** The function $g(\cdot)$ is bounded in $C^{2+\alpha}(\mathbb{R}^m)$ for some $\alpha \in (0,1)$.

**Remark 1** i) It is not hard to check that, H5 implies that for each $(t,y,z,\widehat{z}) \in [0,T] \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \times L^2(U;d\nu)$, the mapping $x \mapsto b(t,x,y,z,\widehat{z}(\cdot))$ and $\sigma(t,\cdot,y)$ satisfy H1. Also, for all $(t,x) \in [0,T] \times \mathbb{R}^n$, the mapping $(y,z,\widehat{z}) \mapsto b(t,x,y,z,\widehat{z}(\cdot))$ and $y \mapsto \sigma(t,x,y)$, $(x,y) \mapsto \theta(t,x,y,\lambda)$, $(x,y,z,\widehat{z}(\cdot)) \mapsto h(t,x,y,z,\widehat{z}(\cdot))$ are all uniformly Lipschitz continuous, and we shall denote all the Lipschitz constants by a generic one $L > 0$.

ii) We should note that in the above the notation $\|\cdot\|_{L^p}$ should be understood as the norm of $L^p(U;d\nu)$, which can be easily identified from the context.

### 2.3 Some Useful Function Spaces

In this paper we shall frequently use the following spaces of functions. First, let $C([0,T]\times\mathbb{R}^n)$ and $C(\mathbb{R}^n)$ be the sets of continuous functions defined on $[0,T]\times\mathbb{R}^n$ and $\mathbb{R}^n$, respectively. For any $u(\cdot\,,\cdot) \in C([0,T]\times\mathbb{R}^n)$, let us define

$$\left\{\begin{array}{l} \|u\|_0 \equiv |u|_0 \stackrel{\Delta}{=} \displaystyle\sup_{(t,x)\in[0,T]\times\mathbb{R}^n} |u(t,x)|, \\ |u|_\alpha \stackrel{\Delta}{=} \displaystyle\sup_{(t,x)\neq(s,y),|x-y|\le 1} \frac{|u(t,x)-u(s,y)|}{|t-s|^{\frac{\alpha}{2}} + |x-y|^\alpha}, \\ \|u\|_2 \stackrel{\Delta}{=} |u|_0 + |u_t|_0 + \displaystyle\sum_{i=1}^n |u_{x_i}|_0 + \sum_{i,j=1}^n |u_{x_ix_j}|_0, \\ \|u\|_{2+\alpha} \stackrel{\Delta}{=} \|u\|_2 + |u_t|_\alpha + \displaystyle\sum_{i,j=1}^n |u_{x_ix_j}|_\alpha, \end{array}\right. \tag{15}$$

where $\alpha \in (0,1)$. Similarly, for any $\varphi(\cdot) \in C(\mathbb{R}^n)$, we define

$$\left\{\begin{array}{l} |\varphi|_0 \stackrel{\Delta}{=} \displaystyle\sup_{x\in\mathbb{R}^n} |\varphi(x)|, \\ |\varphi|_\alpha \stackrel{\Delta}{=} \displaystyle\sup_{0<|x-y|\le 1} \frac{|\varphi(x)-\varphi(y)|}{|x-y|^\alpha}, \\ \|\varphi\|_2 \stackrel{\Delta}{=} |\varphi|_0 + \displaystyle\sum_{i=1}^n |\varphi_{x_i}|_0 + \sum_{i,j=1}^n |\varphi_{x_ix_j}|_0, \\ \|\varphi\|_{2+\alpha} \stackrel{\Delta}{=} \|\varphi\|_2 + |\varphi_{x_ix_j}|_\alpha. \end{array}\right. \tag{16}$$

Now, we define, for any $\alpha \in [0,1)$,

$$\left\{\begin{array}{l} C^{1+\frac{\alpha}{2},2+\alpha}([0,T]\times\mathbb{R}^n) \stackrel{\Delta}{=} \{u \in C([0,T]\times\mathbb{R}^n) \bigm| \|u\|_{2+\alpha} < \infty\}, \\ C^{2+\alpha}(\mathbb{R}) \stackrel{\Delta}{=} \{\varphi \in C(\mathbb{R}^n) \bigm| \|\varphi\|_{2+\alpha} < \infty\}. \end{array}\right. \tag{17}$$

Note that $\alpha$ in (17) is allowed to be 0. For $m \ge 1$, we denote

$$\left\{\begin{array}{l} C^{1+\frac{\alpha}{2},2+\alpha}([0,T]\times\mathbb{R}^n;\mathbb{R}^m) = C^{1+\frac{\alpha}{2},2+\alpha}([0,T]\times\mathbb{R}^n)^m, \\ C^{2+\alpha}(\mathbb{R}^n;\mathbb{R}^m) = C^{2+\alpha}(\mathbb{R}^n)^m. \end{array}\right. \tag{18}$$

Similarly, for any sub-region $G \subseteq \mathbb{R}^n$, we can define $C^{1,\frac{\alpha}{2},2+\alpha}([0,T]\times\overline{G};\mathbb{R}^m)$ and $C^{2+\alpha}(\overline{G};\mathbb{R}^m)$ in an obvious way.

Throughout this paper, by "smooth" we mean that the involved functions possess partial derivatives of all necessary orders. We prefer not to indicate the exact order of smoothness for the simplicity of presentation. We should note that some of the conditions mentioned above can be substantially relaxed, and we will discuss this issue a little later.

To conclude this section, we state an Itô formula corresponding the solution of (11), which will be useful in our future discussion. The argument is rather straightforward, follows directly from the well-known Itô-Meyer formula. We state it here because of its special form, and for ready reference.

**Proposition 1** *Let $X$ be the solution of the forward SDE* (11). *Then, for any $\psi \in C_b^{1,2}([0,\infty)\times\mathbb{R}^d)$ and $t \geq 0$, it holds that*

$$
\begin{aligned}
\psi(t,X_t) = \psi(0,X_0) &+ \int_0^t \langle \nabla\psi(s,X_s), \sigma(s,X_s)dW_s\rangle \\
&+ \int_0^t \psi_t(s,X_s)ds + \int_0^t \langle \nabla\psi(s,X_s), b(s,X_s)\rangle ds \\
&+ \int_0^t \frac{1}{2}\mathrm{tr}\left\{D^2\psi(s,X_s)\sigma(s,X_s)\sigma(s,X_s)^{\mathrm{T}}\right\}ds \\
&+ \int_0^t \int_{[0,\infty)\times U} \Delta_\theta\psi(s,X_{s-},\lambda)\mathbf{1}_{[0,\eta(s,X_{s-},\lambda)]}(r)\widetilde{N}(drd\lambda ds) \\
&+ \int_0^t \int_U \eta(s,X_{s-},\lambda)\big\{[\Delta_\theta\psi](s,X_{s-},\lambda) \\
&\qquad -\mathbf{1}_{U_1}(\lambda)\langle\nabla\psi(s,X_{s-},\lambda), \theta(s,X_{s-},\lambda)\rangle\big\}\nu(d\lambda)ds,
\end{aligned}
$$

*where*

$$
[\Delta_\theta\psi](t,x,\lambda) \stackrel{\triangle}{=} \psi(t,x+\theta(t,x,\lambda)) - \psi(t,x).
$$

## 3 The Solution Scheme

In this section, we will follow the idea of "Four Step Scheme" ([4]) to approach FBSDE (12). One should note that with the special form of the FBSDE, there are in fact only three main steps in the decoupling process.

We begin by assuming a priori that there exists an adapted solution $\Theta \stackrel{\triangle}{=} (X,Y,Z,\widehat{Z})$ to (12) and that the following relation holds:

$$
Y_t = u(t,X_t), \qquad \forall t \in [0,T],\ \text{a.s. } P, \tag{19}
$$

where $u \in C^{1,2}([0,T]\times\mathbb{R}^n;\mathbb{R}^m)$ is some deterministic function. To simplify the arguments in what follows we shall only argue for the case when $U_1 = U$. The general case is similar, modulo some notational complexity. Note that in this case the uncompensated stochastic integrals over $U_2$ in (12) will vanish. We denote $\theta_\Theta(s,\lambda) \stackrel{\triangle}{=} \theta(s,X_{s-},Y_{s-},\lambda)$ for simplicity. For each

$k=1,2,\cdots,m$, we apply Itô's formula to $u^k(t,X_t)$ over $[t,T]$ to get

$$\begin{aligned} u^k(T,X_T) &= u^k(t,X_t)+\int_t^T\langle u^k_x(s,X_s),\sigma(s,X_s,Y_s)dW_s\rangle \\ &\quad+\int_t^T u^k_t(s,X_s)ds+\int_t^T\langle u^k_x(s,X_s),b(s,\Theta_s)\rangle ds \\ &\quad+\int_t^T\frac{1}{2}\mathrm{tr}\left\{u^k_{xx}(s,X_s)\sigma(s,X_s,Y_s)\sigma(s,X_s,Y_s)^{\mathrm{T}}\right\}ds \\ &\quad+\int_t^T\int_{[0,\infty)\times U}[\Delta_{\theta_\Theta}u^k](s,X_{s-},\lambda)\mathbf{1}_{[0,\eta(s,X_{s-},\lambda)]}(r)\widetilde N(drd\lambda ds)\rangle \\ &\quad+\int_t^T\int_U\eta(s,X_{s-},\lambda)\big\{[\Delta_{\theta_\Theta}u^k](s,X_{s-},\lambda) \\ &\qquad-\langle u^k_x(s,X_{s-}),\theta_\Theta(s,\lambda)\rangle\big\}\nu(d\lambda)ds, \end{aligned} \tag{20}$$

where $\Delta_{\theta_\Theta}u^k(s,X_{s-},\lambda)\stackrel{\Delta}{=}u^k(s,X_{s-}+\theta_\Theta(s,\lambda))-u^k(s,X_{s-})$. Comparing (12) and (20), and noting that neither $dW$ nor $dt$ charges a countable sets, we obtain that

$$\begin{cases} Z_t=\sigma^{\mathrm{T}}(t,X_t,Y_t)u_x(t,X_t), \\ \widehat Z_t(\lambda)=u(t,X_t+\theta_\Theta(t,\lambda))-u(t,X_t), \end{cases} \tag{21}$$

and $u=(u^1,u^2,\cdots,u^m)$ satisfies, for $k=1,2,\cdots,m$,

$$\begin{aligned} 0&=u^k_t(t,X_t)+\frac{1}{2}\mathrm{tr}\left\{u^k_{xx}(t,X_t)\sigma(t,X_t,Y_{t-})\sigma(t,X_t,Y_t)^{\mathrm{T}}\right\} \\ &\quad+\langle u_x(t,X_t),b(t,X_t,Y_t,Z_t,\widehat Z_{t-}(\cdot))\rangle+h^k(t,X_t,Y_t,Z_t,\widehat Z_t(\cdot)) \\ &\quad+\int_U\eta(t,X_t,\lambda)\{u^k(t,X_t+\theta_\Theta(t,\lambda))-u^k(t,X_t) \\ &\qquad-\langle u_x(t,X_t),\theta_\Theta(t,\lambda)\rangle\}\nu(d\lambda). \end{aligned} \tag{22}$$

The above discussion suggests that a solution scheme can be defined as follows:

**Solution Scheme**

**Step 1** Solve the following system of integro-partial differential equations (IPDE, for short) for $k=1,2,\cdots,m$,

$$\begin{cases} 0=u^k_t(t,x)+\frac{1}{2}\mathrm{tr}\left\{u^k_{xx}(t,x)\sigma(t,x,u)\sigma(t,x,u)^{\mathrm{T}}\right\} \\ \quad+\langle u^k_x(t,x),b\big(t,x,u,\sigma(t,x,u)^{\mathrm{T}}u_x(t,x),u(t,x+\theta(t,x,u,\cdot))-u(t,x)\big)\rangle \\ \quad+h^k(t,x,u(t,x),\sigma(t,x,u)^{\mathrm{T}}u_x(t,x),u(t,x+\theta(t,x,u,\cdot))-u(t,x)) \\ \quad+\int_U\eta(t,x,\lambda)\big\{u^k(t,x+\theta(t,x,u,\lambda))-u^k(t,x) \\ \qquad-\langle u^k_x(t,x),\theta(t,x,u,\lambda)\rangle\big\}\nu(d\lambda),\qquad (t,x)\in[0,T]\times\mathbb{R}^n, \\ u(T,x)=g(x),\qquad x\in\mathbb{R}^n. \end{cases} \tag{23}$$

**Step 2** Suppose that the IPDE (23) possesses a sufficiently regular solution $u$, use it to solve the following SDE:

$$\begin{aligned} X_t = x &+ \int_0^t \widetilde{b}(s, X_s)ds + \int_0^t \widetilde{\sigma}(s, X_s)dW_s \\ &+ \int_0^t \int_{\mathbb{R}_+ \times U} \theta(s, X_{s-}, u(s, X_{s-}), \lambda)\mathbf{1}_{[0,\eta(s,X_{s-},\lambda)]}\widetilde{N}(drd\lambda ds), \end{aligned} \tag{24}$$

where

$$\widetilde{b}(s, x) \overset{\triangle}{=} b\big(s, x, u, \sigma(s, x, u)^{\mathrm{T}} u_x(s, x), u(s, x + \theta(s, x, u, \cdot)) - u(s, x)\big), \tag{25}$$

$$\widetilde{\sigma}(s, x) \overset{\triangle}{=} \sigma(s, X_s, u(s, x)). \tag{26}$$

**Step 3** Define

$$\begin{cases} Y_t = u(t, X_t), \qquad \big(\text{or} \quad Y_{t-} = u(t, X_{t-})\big); \\ Z_t = \sigma(t, X_t, u(t, X_t))^{\mathrm{T}} u_x(t, X_t), \qquad\qquad t \in [0, T]. \\ \widehat{Z}_t(\lambda) = u(t, X_t + \theta(t, X_t, u(t, X_t), \lambda)) - u(t, X_t), \end{cases} \tag{27}$$

Then $(X, Y, Z, \widehat{Z})$ is an adapted solution to (12).

## 4 Solvability of the IPDE

In this section, we are going to establish the existence and uniqueness of classical solution to system (23). To begin with, let us first recall some standard results. Consider the following partial differential inequalities:

$$w_t + \sum_{i,j=1}^n a_{ij}(t, x)w_{x_i x_j} + \sum_{i=1}^n b_i(t, x)w_{x_i} + c(t, x)w + f(t, x) \geq 0, \quad (t, x) \in \mathbb{R}^n_T, \tag{28}$$

and

$$\begin{cases} w_t + \sum_{i,j=1}^n a_{ij}(t, x)w_{x_i x_j} + \sum_{i=1}^n b_i(t, x)w_{x_i} + c(t, x)w + f(t, x) \geq 0, \quad (t, x) \in B^r_T, \\ w(t, x) \leq 0, \qquad (t, x) \in [0, T] \times \partial B^r, \end{cases} \tag{29}$$

where $\mathbb{R}^n_T = (0, T) \times \mathbb{R}^n$, and $B^r_T = (0, T) \times B^r$ with $B^r \subseteq \mathbb{R}^n$ being the open ball with radius $r > 0$ centered at the origin. The following result is a modification of Theorems 2.1 and 2.5 from pp. 13 and 18 of [13].

**Lemma 1** *Suppose*

$$\begin{cases} a_{ij}, b_i, c, f \in L^\infty(\mathbb{R}^n_T), \quad 1 \leq i, j \leq n, \\ \sum_{i,j=1}^n a_{ij}(t, x)\xi_i\xi_j \geq 0, \qquad \forall(\xi_1, \xi_2, \cdots, \xi_n) \in \mathbb{R}^n, \ (t, x) \in \mathbb{R}^n_T. \end{cases} \tag{30}$$

i) *Let* $w(\cdot\,,\cdot)\in C^{1,2}(\mathbb{R}^n_T)$ *satisfy* (28). *Then*

$$\max_{(t,x)\in\mathbb{R}^n_T} w(t,x)\le\Big\{\max_{x\in\mathbb{R}^n} w(T,x)+T\operatorname*{esssup}_{(t,x)\in\mathbb{R}^n_T} f(t,x)\Big\}\mathrm{e}^{T\,\mathrm{esssup}_{(t,x)\in\mathbb{R}^n_T}\,c(t,x)}\,. \tag{31}$$

ii) *Let* $w(\cdot\,,\cdot)\in C^{1,2}(B^r_T)$ *satisfy* (29). *Then*

$$\max_{(t,x)\in B^r_T} w(t,x)\le\Big\{\max_{x\in B^r_T} w(T,x)+T\operatorname*{esssup}_{(t,x)\in B^r_T} f(t,x)\Big\}\mathrm{e}^{1+T\,\mathrm{esssup}_{(t,x)\in B^r_T}\,c(t,x)}\,. \tag{32}$$

The above results are often referred to as the maximum principle. It is very important that the right hand sides of (31) and (32) do not depend on the bounds of $a_{ij}$ and $b_i$.

Next, we consider the following system of quasilinear parabolic equation:

$$\begin{cases} u^k_t+\displaystyle\sum_{i,j=1}^n a_{ij}(t,x,u)u^k_{x_ix_j}+\sum_{i=1}^n b_i(t,x,u,u_x)u^k_{x_i}+h^k(t,x,u,u_x)=0,\\ \qquad\qquad\qquad\qquad (t,x)\in B^r_T,\ 1\le k\le m,\\ u(t,x)=0,\qquad (t,x)\in[0,T]\times\partial B^r,\\ u(T,x)=g(x),\qquad x\in B^r, \end{cases} \tag{33}$$

The following result is concerned with the existence and uniqueness of classical solution to system (33), which is a modification of [13].

**Lemma 2** *Suppose that all the functions* $a_{ij}$, $b_i$, $h^k$ *are twice continuously differentiable. Suppose also that for all* $(t,x,u,z)\in[0,T]\times B^r\times\mathbb{R}^m\times\mathbb{R}^{m\times n}$, *it holds that*

$$\begin{cases} \mu_0(|u|)I\le\big(a_{ij}(t,x,u)\big)\le\mu(|u|)I,\\ \big|a_{ijx_\ell}(t,x,u)\big|+\big|a_{iju^k}(t,x,u)\big|\le\mu(|u|),\quad 1\le i,j,k,\ell\le n,\\ |b(t,x,u,z)|\le\mu(|u|)(1+|z|),\\ \langle u,h(t,x,u,z)\rangle\le L(1+|u|^2),\\ |h(t,x,u,z)|\le[\varepsilon(|u|)+P(|u|,|z|)](1+|z|^2), \end{cases} \tag{34}$$

*for some constant* $L>0$; *continuous functions* $\mu_0,\mu,\varepsilon:(0,\infty)\to(0,\infty)$; *and* $P:\mathbb{R}_+\times\mathbb{R}_+\to\mathbb{R}_+$ *with*

$$\begin{cases} \displaystyle\lim_{|z|\to\infty}P(|u|,|z|)=0,\\ \displaystyle\sup_{M>0}\frac{(2M+1)\varepsilon(M)}{\mu_0(M)}<1. \end{cases} \tag{35}$$

*Finally, suppose that* $g\in C^{2+\alpha}(B^r;\mathbb{R}^m)$ *for some* $\alpha\in(0,1)$.

*Then,* (33) *admits a unique classical solution* $u(\cdot\,,\cdot)\in C^{1+\frac{\alpha}{2},2+\alpha}(B^r_T;\mathbb{R}^m)$ *such that the following estimates hold:*

$$ll\|u\|_{L^\infty(B^r_T;\mathbb{R}^m)}\le\mathrm{e}^{2LT}\big(\|g\|_{L^\infty(B^r;\mathbb{R}^m)}+1\big)\equiv M, \tag{36}$$

$$\|u_x\|_{L^\infty(B^r_T;\mathbb{R}^{n\times m})}+\|u\|_{C^{\frac{\gamma}{2},\gamma}(B^r_T;\mathbb{R}^m)}\le M_1, \tag{37}$$

*where $M_1 > 0$ and $\gamma \in (0,1)$ depend only on $L$, $M$, $\mu_0(M)$, $\mu(M)$, $\varepsilon(M)$, and $P(\cdot\,,\cdot)$;*

$$\|u_x\|_{C^{\frac{\gamma}{2},\gamma}(B_T^r;\mathbb{R}^{n\times m})} \le M_2, \tag{38}$$

*where $M_2 > 0$ and $\gamma \in (0,1)$ depend only on $L$, $M$, $\mu_0(M)$, $\mu(M)$, $\varepsilon(M)$, $P(\cdot\,,\cdot)$, $M_1$, and the bounds of the first order partial derivatives of $a_{ij}$ in $(t,x,u,z)$; and*

$$\|u\|_{C^{1+\frac{\alpha}{2},2+\alpha}(B_T^r;\mathbb{R}^m)} \le C, \tag{39}$$

*where $C > 0$ depends only on $L$, $\mu_0(M)$, $\mu(M)$, $\varepsilon(M)$, $P(\cdot\,,\cdot)$, $M_1$, and the bounds of the first order partial derivatives of $a_{ij}$, $b_i, h^k$ in $(t,x,u,z)$.*

*Proof* For any $r > 0$, consider the following terminal-boundary value problem:

$$\begin{cases} u_t^k + \displaystyle\sum_{i,j=1}^n a_{ij}(t,x,u)u^k_{x_ix_j} + \sum_{i=1}^n b_i(t,x,u,u_x)u^k_{x_i} + h^k(t,x,u,u_x) = 0, \\ \qquad\qquad\qquad\qquad\qquad 1 \le k \le m,\ (t,x) \in [0,T]\times\mathbb{R}^n, \\ u(T,x) = g(x)\rho_r(x), \qquad x \in \mathbb{R}^n, \end{cases} \tag{40}$$

where $B_r \subseteq \mathbb{R}^n$ is the open ball centered at the origin with radius $r > 0$, and $\rho_r(\cdot) \in C_0^\infty(B_r)$.

Under the assumptions we can apply Theorem 7.1 of [13] to conclude that (40) admits a unique classical solution $u^r(\cdot\,,\cdot) \in C^{1+\frac{\alpha}{2},2+\alpha}([0,T]\times\overline{B}_r)$. Since $g$ is bounded in $C^{2+\alpha}(\mathbb{R}^n)$, the solution $u^r(\cdot,\cdot)$ of (40) satisfies

$$\|u^{\widehat{r}}\|_{C^{1+\frac{\alpha}{2},2+\alpha}([0,T]\times\overline{B}_r;\mathbb{R}^m)} \le C, \qquad \forall\, \widehat{r} \ge r. \tag{41}$$

This is a consequence of interior type Schauder estimate for the parabolic equations. Letting $r \to \infty$ we see that (33) admits a classical solution, and the estimates (37)–(38) follows easily (see [13]).

Now, let us suppose that (33) admits another classical solution $v(\cdot\,,\cdot)$ which is bounded in $C^{1+\frac{\alpha}{2},2+\alpha}([0,T]\times\mathbb{R}^n;\mathbb{R}^m)$. Then

$$\begin{aligned} 0 &= (u^k - v^k)_t + \sum_{i,j=1}^n a_{ij}(t,x,u)(u^k - v^k)_{x_ix_j} + \sum_{i=1}^n b_i(t,x,u,u_x)(u^k - v^k)_{x_i} \\ &\quad + \sum_{i,j=1}^n [a_{ij}(t,x,u) - a_{ij}(t,x,v)]v^k_{x_ix_j} + \sum_{i=1}^n [b_i(t,x,u,u_x) - b_i(t,x,v,v_x)]v^k_{x_i} \\ &\quad + h^k(t,x,u,u_x) - h^k(t,x,v,v_x) \\ &= (u^k - v^k)_t + \sum_{i,j=1}^n a_{ij}(t,x,u)(u^k - v^k)_{x_ix_j} + \sum_{i=1}^n b_i(t,x,u,u_x)(u^k - v^k)_{x_i} \\ &\quad + \mathrm{tr}\left[\widehat{B}^k(t,x)(u-v)_x\right] + \langle \widehat{b}^k(t,x), u - v\rangle, \end{aligned} \tag{42}$$

where, denoting $U^{u,v,\beta}(t,x) \stackrel{\triangle}{=} (1-\beta)v(t,x) + \beta u(t,x)$,

$$\begin{aligned} \widehat{B}^k(t,x) &= \sum_{i=1}^n v^k_{x_i}(t,x) \int_0^1 (b_i)_{u_x}(t,x,U^{u,v,\beta}(t,x),U_x^{u,v,\beta}(t,x))d\beta \\ &\quad + \int_0^1 h^k_{u_x}(t,x,U^{u,v,\beta}(t,x),U_x^{u,v,\beta}(t,x))d\beta, \end{aligned} \tag{43}$$

and

$$
\begin{aligned}
\widehat{b}^k(t,x) = &\sum_{i,j=1}^n v^k_{x_ix_j}(t,x)\int_0^1 (a_{ij})_u(t,x,U^{u,v,\beta}(t,x))d\beta \\
&+\sum_{i=1}^n v^k_{x_i}(t,x)\int_0^1 (b_i)_u(t,x,U^{u,v,\beta}(t,x),U^{u,v,\beta}_x(t,x))d\beta \\
&+\int_0^1 h^k_u(t,x,U^{u,v,\beta}(t,x),U^{u,v,\beta}_x(t,x))d\beta. \qquad (44)
\end{aligned}
$$

Since both $u$ and $v$ are bounded in $C^{1+\frac{\alpha}{2},2+\alpha}([0,T]\times\mathbb{R}^n;\mathbb{R}^m)$, we see that $\widehat{B}^k(\cdot\,,\cdot)$ and $\widehat{b}^k(\cdot\,,\cdot)$ are continuous and bounded. Now, we let

$$
w(t,x)=\frac{1}{2}|u(t,x)-v(t,x)|^2, \qquad (t,x)\in[0,T]\times\mathbb{R}^n. \qquad (45)
$$

Then, suppressing variables, we have

$$
\begin{aligned}
-w_t &= -\langle u-v,(u-v)_t\rangle \\
&= \sum_{k=1}^m (u^k-v^k)\Big\{\sum_{i,j=1}^n a_{ij}(t,x,u)(u^k-v^k)_{x_ix_j} \\
&\qquad +\sum_{i=1}^n b_i(t,x,u,u_x)(u^k-v^k)_{x_i}+\mathrm{tr}\,\big[\widehat{B}^k(t,x)(u-v)_x\big]+\langle \widehat{b}^k(t,x),u-v\rangle\Big\} \\
&= \sum_{i,j=1}^n a_{ij}(t,x,u)\Big[w_{x_ix_j}-\sum_{k=1}^m (u^k-v^k)_{x_i}(u^k-v^k)_{x_j}\Big]+\sum_{i=1}^n b_i(t,x,u,u_x)w_{x_i} \\
&\quad +\sum_{k=1}^m (u^k-v^k)\big\{\mathrm{tr}\,\big[\widehat{B}^k(t,x)(u-v)_x\big]+\langle \widehat{b}^k(t,x),u-v\rangle\big\} \\
&\le \sum_{i,j=1}^n a_{ij}(t,x,u)w_{x_ix_j}-\sum_{k=1}^m \mu_0(|u|)|(u^k-v^k)_x|^2+\sum_{i=1}^n b_i(t,x,u,u_x)w_{x_i} \\
&\quad +C\sqrt{w}\sum_{k=1}^m |(u^k-v^k)_x|+Cw \\
&\le \sum_{i,j=1}^n a_{ij}(t,x,u)w_{x_ix_j}+\sum_{i=1}^n b_i(t,x,u,u_x)w_{x_i}+Cw. \qquad (46)
\end{aligned}
$$

Since $w(T,x)=u(T,x)-v(T,x)=0$, $x\in\mathbb{R}^n$, if follows that $w\equiv 0$, thanks to Lemma 1. The uniqueness follows. ∎

We remark that in [13], it was claimed that the constant $C>0$ in (39) depends also on the bounds of second order partial derivatives of $a_{ij}$ in its arguments. But this can be relaxed to our claim by first using $L^p$-estimate (after establishing the boundedness of $u_x$), and followed by Schauder estimates for linear equations (after having the Hölder continuity of $u$ and $u_x$). We leave the detailed argument to interested reader.

We now establish the existence and uniqueness of classical solution to (23). To this end, let

$r > 0$, and consider the following system of quasilinear parabolic equations:

$$\begin{cases} 0 = u_t^k(t,x) + \frac{1}{2}\mathrm{tr}\left\{u_{xx}^k(t,x)\sigma(t,x,u)\sigma(t,x,u)^{\mathrm{T}}\right\} \\ \quad +\langle u_x^k(t,x), b\big(t,x,u,\sigma(t,x,u)^{\mathrm{T}}u_x(t,x), u(t,x+\theta(t,x,u,\cdot)) - u(t,x)\big)\rangle \\ \quad +h^k(t,x,u,\sigma(t,x,u)^{\mathrm{T}}u_x(t,x), u(t,x+\theta(t,x,u,\cdot)) - u(t,x)) \\ \quad + \int_U \eta(t,x,\lambda)\big\{[\Delta_\theta u^k](t,x,\lambda) - \langle u_x^k(t,x), \theta(t,x,u,\lambda)\rangle\big\}\nu(d\lambda), \\ \qquad\qquad (t,x) \in [0,T]\times B^r, \quad k = 1,2,\cdots,m, \\ u(t,x) = 0, \qquad \forall (t,x) \in [0,T]\times\partial B_r, \\ u(T,x) = g(x), \\ x \in B^r. \end{cases} \tag{47}$$

We define $u(t,x) = 0$, for all $x \in \mathbb{R}^n \setminus B_r$. Now suppose that $u$ is a classical solution of (20). We observe the following identity:

$$\begin{aligned} -[|u|^2]_t &= -2\sum_{k=1}^m u^k u_t^k \\ &= \sum_{k=1}^m u^k \Big\{ \mathrm{tr}\left[u_{xx}^k \sigma(t,x,u)\sigma(t,x,u)^{\mathrm{T}}\right] \\ &\qquad +2\Big\langle u_x^k, b\Big(t,x,u,\sigma(t,x,u)^{\mathrm{T}}u_x, u(t,x+\theta(t,x,u,\cdot)) - u(t,x)\Big)\Big\rangle \\ &\qquad +2h^k\Big(t,x,u,\sigma(t,x,u)^{\mathrm{T}}u_x, u(t,x+\theta(t,x,u,\cdot)) - u(t,x)\Big) \\ &\qquad +2\int_U \eta(t,x,\lambda)\Big[u^k(t,x+\theta(t,x,u,\lambda)) - u^k - \langle u_x^k, \theta(t,x,u,\lambda)\rangle\Big]\nu(d\lambda)\Big\} \\ &\le \frac{1}{2}\mathrm{tr}\left\{\left[|u|^2\right]_{xx}\sigma(t,x,u)\sigma(t,x,u)^{\mathrm{T}}\right\} - \sum_{i=1}^n |\sigma(t,x,u)^{\mathrm{T}}u_{x_i}|^2 \\ &\quad +\Big\langle \left[|u|^2\right]_x, b(t,x,u,\sigma(t,x,u)^{\mathrm{T}}u_x, u(t,x+\theta(t,x,u,\cdot)) - u) \\ &\qquad - \int_U \eta(t,x,\lambda)\theta(t,x,u,\lambda)\nu(d\lambda)\Big\rangle \\ &\quad +2L(1+|u|^2) + \int_U \left[|u(t,x+\theta(t,x,\theta,\lambda))|^2 - |u|^2\right]\eta(t,x,\lambda)\nu(d\lambda) \\ &\le \frac{1}{2}\mathrm{tr}\left\{\left[|u|^2\right]_{xx}\sigma(t,x,u)\sigma(t,x,u)^{\mathrm{T}}\right\} + \langle \left[|u|^2\right]_x, \widehat{b}(t,x)\rangle + 2L|u|^2 + f(t,x), \end{aligned} \tag{48}$$

where (note that $\eta(t,x,\lambda) > 0$ and $\eta_0 = \sup_{(t,x)\in[0,T]\times\mathbb{R}^n} \int_U \eta(t,x,\lambda)\nu(d\lambda)$).

$$\begin{cases} \widehat{b}(t,x) \triangleq b(t,x,u,\sigma(t,x,u)^{\mathrm{T}}u_x, u(t,x+\theta(t,x,u,\cdot)) - u) \\ \qquad\qquad - \int_U \eta(t,x,\lambda)\theta(t,x,u,\lambda)\nu(d\lambda)\rangle \\ f(t,x) \triangleq 2L + \eta_0\|u(t,\cdot)\|^2_{L^\infty(B^r)}. \end{cases} \tag{49}$$

Thus, by Lemma 4.1, for each $t \in [0,T]$,

$$\|u\|^2_{L^\infty([t,T]\times B^r)} \le \mathrm{e}^{2LT+1}\big\{\|g\|^2_{L^\infty(B^r)} + 2LT + (T-t)\eta_0\|u\|^2_{L^\infty([t,T]\times B^r)}]\big\}. \tag{50}$$

Then, by induction one can easily deduce that

$$\|u\|^2_{L^\infty([0,T]\times B^r)} \le M, \tag{51}$$

with $M > 0$ depending only on $\eta_0$, $L$, $T$, and $\|g\|_{L^\infty(B^r)}$. Next, by H2, we have for any $(t,x,u,z,\widehat{z}) \in [0,T]\times\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^{n\times m}\times L^2(d\nu)$,

$$\begin{cases} \mu_0(|u|)I \le \sigma(t,x,u)\sigma(t,x,u)^{\mathrm{T}} \le \mu(|u|)I, \\ |\sigma_{x_\ell}(t,x,u)| + |\sigma_u(t,x,u)| \le L, \\ |b(t,x,u,z,\widehat{z})| \le \mu(|u|)(1+|z|), \\ |h(t,x,u,z,\widehat{z})| \le [\varepsilon(|u|) + P(|u|,|z|)](1+|z|^2). \end{cases} \tag{52}$$

Then it follows from Lemma 2 that

$$\|u_x\|_{L^\infty([T-\delta,T]\times B^r)} + \|u\|_{C^{\frac{\gamma}{2},\gamma}([T-\delta,T]\times B^r)} \le M_1, \tag{53}$$

with $M_1 > 0$ and $\gamma \in (0,1)$ depending only on $M$, $\mu_0(M)$, $\mu(M)$, $P(\cdot)$, $\varepsilon(M)$, and $L$. Consequently, the mapping $(t,x)\mapsto \sigma(t,x,u(t,x))\sigma(t,x,u(t,x))^{\mathrm{T}}$ is Hölder continuous.

We can now regard (47) as a linear equation with Hölder continuous leading coefficient, and with other coefficients all being bounded. It then follows from the $L^p$-estimate[13] that

$$\|u,u_t,u_x,u_{xx}\|_{L^p([T-\delta,T]\times B^r)} \le C_p, \qquad \forall p \in [1,\infty), \tag{54}$$

where the constant $C > 0$ depends only on $M$ and $M_1$. It then follows from Lemma 2 that $u_x$ is Hölder continuous. Hence, so are all the coefficients of (47) (still regard it as a linear equation). Based on this fact, and applying the Schauder estimate[13], the following estimate holds:

$$\|u\|_{C^{1+\frac{\alpha}{2},2+\alpha}([T-\delta,T]\times B^r)} \le C, \tag{55}$$

with $C > 0$ being an absolute constant depending only on $M$, $M_1$, the bounds of partial derivatives of $\sigma$, $b$, $h$, and the norm $\|g\|_{C^{2+\alpha}(B^r)}$. Then similar to the arguments used in [13] involving the Leray-Schauder fixed point theorem, we can obtain the existence of a classical solution $u$ to (47) on $[T-\delta,T]\times B^r$. Since $\delta > 0$ is an absolute constant, a simple argument of induction would lead to the existence of a classical solution $u$ on $[0,T]\times B^r$.

Summarizing, we now present our first main result of the paper.

**Theorem 1** *Assume* H1–H8. *Then the system of quasilinear PDEs* (47) *possesses a unique classical solution* $u \in C^{1+\frac{\alpha}{2},2+\alpha}([0,T]\times\mathbb{R}^m)$.

*Proof* The existence follows from the arguments preceding the theorem. We need only note that the existence and uniqueness of the solution on $[0,T]\times\mathbb{R}^m$ is equivalent to the existence and uniqueness of the solution over $[0,T]\times B^r$, for all $r > 0$.

We now turn to the uniqueness. Let $v$ be another (classical) solution of (47). Then (recall that $[\Delta_\theta u](t,x,\lambda) \stackrel{\triangle}{=} u(t,x+\theta(t,x,u,\lambda)) - u(t,x))$,

$$\begin{aligned}
0 &= (u^k - v^k)_t + \frac{1}{2}\mathrm{tr}\,[(u^k - v^k)_{xx}\sigma(t,x,u)\sigma(t,x,u)^{\mathrm{T}}] \\
&\quad + \frac{1}{2}\mathrm{tr}\left\{ v^k_{xx}\Big[\sigma(t,x,u)\sigma(t,x,u)^{\mathrm{T}} - \sigma(t,x,v)\sigma(t,x,v)^{\mathrm{T}}\Big]\right\} \\
&\quad + \Big\langle (u^k - v^k)_x, b\Big(t,x,u,\sigma(t,x,u)^{\mathrm{T}}u_x, [\Delta_\theta u](t,x,\cdot)\Big)\Big\rangle \\
&\quad + \Big\langle v^k_x, b(t,x,u,\sigma(t,x,u)^{\mathrm{T}}u_x, [\Delta_\theta u](t,x,\cdot)) - b(t,x,v,\sigma(t,x,v)^{\mathrm{T}}v_x, [\Delta_\theta v](t,x,\cdot)\Big\rangle \\
&\quad + h^k\Big(t,x,u,\sigma(t,x,u)^{\mathrm{T}}u_x, [\Delta_\theta u](t,x,\cdot)\Big) - h^k\Big(t,x,v,\sigma(t,x,v)^{\mathrm{T}}v_x, [\Delta_\theta v](t,x,\cdot)\Big) \\
&\quad + \int_U \eta(t,x,\lambda)\Big\{\Big[\Delta_\theta(u^k - v^k)\Big](t,x,\lambda) - \Big\langle u^k_x - v^k_x, \theta(t,x,u,\lambda)\Big\rangle\Big\}\nu(d\lambda).
\end{aligned} \tag{56}$$

Note that

$$\begin{aligned}
&b\Big(t,x,u,\sigma(t,x,u)^{\mathrm{T}}u_x, [\Delta_\theta u](t,x,\cdot)\Big) - b\Big(t,x,u,\sigma(t,x,u)^{\mathrm{T}}v_x, [\Delta_\theta u](t,x,\cdot)\Big) \\
&= \mathrm{tr}\,\Big[\int_0^1 b_z\Big(t,x,u,\sigma(t,x,u)^{\mathrm{T}}[(1-\beta)v_x + \beta u_x], [\Delta_\theta u](t,x,\cdot)\Big)d\beta\sigma(t,x,u)^{\mathrm{T}}(u-v)_x\Big] \\
&\stackrel{\triangle}{=} \mathrm{tr}\,[B(t,x,u,u_x,v_x,\cdot)(u-v)_x].
\end{aligned} \tag{57}$$

Similarly,

$$\begin{aligned}
&h^k(t,x,u,\sigma(t,x,u)^{\mathrm{T}}u_x, [\Delta_\theta u](t,x,\cdot)) - h^k(t,x,u,\sigma(t,x,u)^{\mathrm{T}}v_x, [\Delta_\theta u](t,x,\cdot)) \\
&= \mathrm{tr}\,\Big[\int_0^1 h^k_z(t,x,u,\sigma(t,x,u)^{\mathrm{T}}[(1-\beta)v_x + \beta u_x], [\Delta_\theta u](t,x,\cdot))d\beta\sigma(t,x,u)^{\mathrm{T}}(u-v)_x\Big] \\
&\stackrel{\triangle}{=} \mathrm{tr}\,[H^k(t,x,u,u_x,v_x,\cdot)(u-v)_x].
\end{aligned} \tag{58}$$

Then (56) becomes

$$\begin{aligned}
0 &= (u^k - v^k)_t + \frac{1}{2}\mathrm{tr}\,\Big[(u^k - v^k)_{xx}(\sigma\sigma^{\mathrm{T}})(t,x,u)\Big] + \Big\langle (u^k - v^k)_x, \widetilde{b}(t,x,u,u_x,\cdot)\Big\rangle \\
&\quad + \sum_{i=1}^n \mathrm{tr}\left\{\Big[u^k_{x_i}B_i(t,x,u,u_x,v_x;u) + H^k(t,x,u,u_x,v_x;u)\Big](u-v)_x\right\} + F^k,
\end{aligned} \tag{59}$$

where

$$\widetilde{b}(t,x,u,u_x,\cdot) \stackrel{\triangle}{=} b(t,x,u,\sigma(t,x,u)^{\mathrm{T}}u_x, [\Delta_\theta u](t,x,\cdot)) - \int_U \eta(t,x,\lambda)\theta(t,x,u,\lambda)\nu(d\lambda),$$

and

$$\begin{aligned}
F^k &= \Big\langle v^k_x, b\Big(t,x,u,\sigma(t,x,u)^{\mathrm{T}}v_x, [\Delta_\theta u](t,x,\cdot)\Big) - b\Big(t,x,v,\sigma(t,x,v)^{\mathrm{T}}v_x, [\Delta_\theta v](t,x,\cdot)\Big)\Big\rangle \\
&\quad + \frac{1}{2}\mathrm{tr}\left\{ v^k_{xx}\Big[\sigma(t,x,u)\sigma(t,x,u)^{\mathrm{T}} - \sigma(t,x,v)\sigma(t,x,v)^{\mathrm{T}}\Big]\right\} \\
&\quad + h^k\Big(t,x,u,\sigma(t,x,u)^{\mathrm{T}}v_x, [\Delta_\theta u](t,x,\cdot)\Big) - h^k\Big(t,x,v,\sigma(t,x,v)^{\mathrm{T}}v_x, [\Delta_\theta v](t,x,\cdot)\Big) \\
&\quad + \int_U \eta(t,x,\lambda)\Big[[\Delta_\theta u^k](t,x,\lambda) - [\Delta_\theta v^k](t,x,\lambda)\Big]\nu(d\lambda).
\end{aligned} \tag{60}$$

Now using the Lipschitz conditions on the coefficients from assumptions H2, H3, H5, and H6 we deduce from (60) that

$$
\begin{aligned}
|F^k| \le L|v_x^k|\big[|u-v| + |v_x|\,|\sigma(t,x,u)-\sigma(t,x,v)| \\
+|u(t,x+\theta(t,x,u,\cdot))-v(t,x+\theta(t,x,v,\cdot))| + |u-v|\big] \\
+C|v_{xx}^k|\,|\sigma(t,x,u)\sigma(t,x,u)^{\mathrm{T}}-\sigma(t,x,v)\sigma(t,x,v)^{\mathrm{T}}| \\
+L\big(|u-v|+|v_x|\,|\sigma(t,x,u)-\sigma(t,x,v)| \\
+|u(t,x+\theta(t,x,u(t,x),\cdot))-v(t,x+\theta(t,x,v(t,x),\cdot))|+|u-v|\big) \\
+\int_U \eta(t,x,\lambda)\big|u^k(t,x+\theta(t,x,u,\lambda))-v^k(t,x+\theta(t,x,v,\lambda))-u^k+v^k\big|\nu(d\lambda) \\
\le C\|u(t,\cdot)-v(t,\cdot)\|_{L^\infty(B^r)},
\end{aligned}
\tag{61}
$$

where $C>0$ is a constant depending only on $M$ (and $M_1$). Hence,

$$
\begin{aligned}
-[|u-v|^2]_t &= -2\sum_{k=1}^m (u^k-v^k)(u^k-v^k)_t \\
&= \sum_{k=1}^m (u^k-v^k)\Big\{\mathrm{tr}\Big[(u^k-v^k)_{xx}(\sigma\sigma^{\mathrm{T}})(t,x,u)\Big]+2\Big\langle (u^k-v^k)_x, \widetilde{b}(t,x,u,u_x)\Big\rangle \\
&\quad +2\sum_{i=1}^n \mathrm{tr}\Big\{\big[v_{x_i}^k B_i(t,x,u,u_x,v_x,\cdot)+H^k(t,x,u,u_x,v_x,\cdot)\big](u-v)_x\Big\}+2F^k\Big\} \\
&\le \frac{1}{2}\mathrm{tr}\Big\{\big[|u-v|^2\big]_{xx}(\sigma\sigma^{\mathrm{T}})(t,x,u)\Big\}+\Big\langle \big[|u-v|^2\big]_x, \widetilde{b}(t,x,u,u_x)\Big\rangle \\
&\quad -\sum_{i=1}^n |\sigma^{\mathrm{T}}(u-v)_{x_i}|^2 + C|u-v|\Big(|(u-v)_x| + \|u(t,\cdot)-v(t,\cdot)\|_{L^\infty(\mathbb{R}^n)}\Big) \\
&\le \mathrm{tr}\Big\{\frac{1}{2}\big[|u-v|^2\big]_{xx}\sigma(t,x,u)\sigma(t,x,u)^{\mathrm{T}}\Big\}+\Big\langle \big[|u-v|^2\big]_x, \widetilde{b}(t,x,u,u_x)\Big\rangle \\
&\quad +C\|u(t,\cdot)-v(t,\cdot)\|^2_{L^\infty(B^r)}.
\end{aligned}
\tag{62}
$$

Applying Lemma 1, we obtain that

$$
\|u-v\|_{L^\infty([t,T]\times B^r)} \le (T-t)C_0\|u-v\|_{L^\infty([t,T]\times B^r)}, \qquad t\in[0,T] \tag{63}
$$

where the constant $C_0$ is independent of $t\in[0,T]$. In other words, $u$ and $v$ coincide on $[T-\delta,T]\times\mathbb{R}^n$, for some absolute constant $\delta\in(0,T)$. Applying the above arguments on $[T-2\delta,T-\delta]$, etc., we obtain the uniqueness of classical solution to (47).

Note that the solution $u\equiv u^r$ of (47) depends on $r>0$. By interior estimates, we know that for any fixed $r>0$, and any sequence $\widehat{r}\to\infty$, the family $\{u^{\widehat{r}}:\widehat{r}\ge r\}$ is pre-compact in $C^{1,2}(B_T^r)$. Thus, we may extract a subsequence so that

$$
\lim_{\widehat{r}\to\infty}\|u^{\widehat{r}}-u\|_{C^{1,2}(B_T^r)}=0. \tag{64}
$$

By a diagonalization argument, we see that a function $u\in C^{1,2}(\mathbb{R}_T^n)$ exists so that (64) holds for any $r>0$. Clearly, this $u$ is a classical solution of (23). The same as (62)–(63), we can prove the uniqueness of classical solution to (23). ∎

## 5 Well-Posedness of the FBSDE

We are now ready to present the main result of this paper.

**Theorem 2** *Assume* H1–H8. *Then, the FBSDE* (12) *admits a unique solution* $(X, Y, Z, \widehat{Z}(\cdot)) \in L^2_{\mathbb{F}}([0,T] \times \Omega; \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \times L^2(\mathbb{R}^m, d\nu))$*, such that,* $\mathbb{P}$*-almost surely, for all* $t \in [0,T]$ *and* $d\nu$*-a.e.* $\lambda \in \mathbb{R}^m$*, it holds that*

$$\begin{cases} Y_t = u(t, X_t) \\ Z_t = \sigma(t, X_t, u(t, X_t))u_x(t, X_t) \\ \widehat{Z}_t(\lambda) = u(t, X_t + \theta(t, X_t, u(t, X_t), \lambda) - u(t, X_t), \end{cases} \tag{65}$$

*where* $u : [0,T] \times \mathbb{R}^n \mapsto \mathbb{R}^m$ *is the unique classical solution to the IPDE* (5).

*Proof* In light of the discussion of the previous sections, we see that as long as we can find a classical solution to the IPDE (23), the Four Step Scheme goes through, and we obtain at least one solution. Thus, Theorem 1 implies the existence of the solution to FBSDE (12). We now argue that any adapted solution to the FBSDE (12) must coincide with the solution constructed by the Four Step Scheme. For notational simplicity, we assume that $n = m = 1$, and again, we argue only for the case when $U_1 = U$. The general, higher dimensional case can be proved in the same way without substantial difficulty.

Let $(X, Y, Z, \widehat{Z})$ be any solution of (12). Then, we define a "bench mark" solution via Four Step Scheme as follows:

$$\begin{cases} Y^0_t = u(t, X_t), \\ Z^0_t = \sigma(t, X_t, u(t, X_t))u_x(t, X_t), \\ \widehat{Z}^0_t(\lambda) = u(t, X_t + \theta(t, X_t, u(t, X_t), \lambda) - u(t, X_t). \end{cases} \tag{66}$$

We show that $(X, Y, Z, \widehat{Z})$ is identical with $(X, Y^0, Z^0, \widehat{Z}^0)$ in a pathwise sense. To simplify notations in what follows we denote $\overline{\theta}(t,\lambda) = \overline{\theta}^{X,Y}(t,\lambda) \stackrel{\triangle}{=} \theta(t, X_{t-}, Y_{t-}, \lambda)$, when the context is clear.

Applying Ito-Meyer formula to $u(t, X_t)$, we have (suppressing variables):

$$\begin{aligned} dY^0_t &= du(t, X_t) \\ &= \Big\{ u_t(t, X_t) + u_x(t, X_t) b(t, X_t, Y_t, Z_t, \widehat{Z}_t(\cdot)) + \frac{1}{2} u_{xx}(t, X_t)\sigma^2(t, X_t, Y_t) \Big\} dt \\ &\quad + u_x(t, X_t)\sigma(t, X_t, Y_t) dW_t \\ &\quad + \int_{\mathbb{R}_+ \times U} [\varDelta_{\overline{\theta}}] u(t, X_{t-}, \lambda) 1_{[0,\eta(t,X_{t-},\lambda)]}(r) \widetilde{N}(drd\lambda dt) \\ &\quad + \int_U \eta(t, X_{t-}, \lambda) \{ [\varDelta_{\overline{\theta}} u](t, X_{t-}, \lambda) - u_x(t, X_{t-})\overline{\theta}(t,\lambda) \} \nu(d\lambda) dt, \end{aligned}$$

where $[\varDelta_{\overline{\theta}} u](t, X_{t-}, \lambda) \stackrel{\triangle}{=} u(t, X_{t-} + \overline{\theta}(t,\lambda)) - u(t, X_{t-})$. Since $u(t,x)$ is the solution to the integro-partial differential equation (23) and $(Y^0, Z^0, \widehat{Z}^0)$ satisfies (66), by Itô's formula we have

$$
\begin{aligned}
dY_t^0 &= du(t,X_t)\\
&= \Big\{u_t(t,X_t)+u_x(t,X_t)b(t,X_t,Y_t,Z_t,\widehat{Z}_t(\cdot))+\frac{1}{2}u_{xx}(t,X_t)\sigma^2(t,X_t,Y_t)\Big\}dt\\
&\quad +u_x(t,X_t)\sigma(t,X_t,Y_t)dW_t\\
&\quad +\int_{\mathbb{R}_+\times U}[\Delta_\theta u](t,X_{t-},\lambda)1_{[0,\eta(t,X_{t-},\lambda)]}(r)\widetilde{N}(drd\lambda dt)\\
&\quad +\int_U \eta(t,X_{t-},\lambda)\{[\Delta_{\overline{\theta}}u](t,X_{t-},\lambda)-u_x(t,X_{t-})\overline{\theta}(t,\lambda)\}\nu(d\lambda)dt\\
&= \Big\{u_x(t,X_t)\big[b(t,X_t,Y_t,Z_t,\widehat{Z}_t)-b(t,X_t,Y_t^0,Z_t^0,\widehat{Z}_t^0)\big]\\
&\quad +\frac{1}{2}u_{xx}(t,X_t)\big[\sigma^2(t,X_t,Y_t)-\sigma^2(t,X_t,Y_t^0)\big]+h(t,X_t,Y_t^0,Z_t^0,\widehat{Z}_t^0)\Big\}dt\\
&\quad +u_x(t,X_t)\sigma(t,X_t,Y_t)dW_t\\
&\quad +\int_U \eta(t,X_{t-},\lambda)\Big\{u(t,X_{t-}+\theta(t,X_{t-},Y_{t-},\lambda))-u(t,X_{t-}+\theta(t,X_{t-},Y_{t-}^0,\lambda))\\
&\qquad -u_x(t,X_{t-})[\theta(t,X_{t-},Y_{t-},\lambda)-\theta(t,X_{t-},Y_{t-}^0,\lambda)]\Big\}\nu(d\lambda)dt\\
&\quad +\int_{\mathbb{R}_+\times U}[\Delta_{\overline{\theta}}u](t,X_{t-},\lambda)1_{[0,\eta(t,X_{t-},\lambda)]}(r)\widetilde{N}(drd\lambda dt).
\end{aligned}
$$

Here we used the fact that the Lebesgue measure does not charge the countable jumps, hence there is no difference between, say, $X_t$ and $X_{t-}$, in the integrands. Now, recalling (12) we write

$$
dY_t = h(t,X_t,Y_t,Z_t,\widehat{Z}_t)dt+Z_tdW_t+\int_{\mathbb{R}_+\times U}\widehat{Z}_{t-}(\lambda)1_{[0,\eta(t,X_{t-},\lambda)]}(r)\widetilde{N}(drd\lambda dt),
$$

Then, denoting $\overline{\theta}^0(t,\lambda)=\overline{\theta}^{X,Y^0}(t,\lambda)\stackrel{\triangle}{=}\theta(t,X_{t-},Y_{t-}^0,\lambda)$, we have

$$
\begin{aligned}
d(Y_t^0-Y_t) =& \Big\{u_x(t,X_t)\big(b(t,X_t,Y_t,Z_t,\widehat{Z}_t)-b(t,X_t,Y_t^0,Z_t^0,\widehat{Z}_t^0)\big)\\
&+\frac{1}{2}u_{xx}(t,X_t)\big(\sigma^2(t,X_t,Y_t)-\sigma^2(t,X_t,Y_t^0)\big)\\
&+h(t,X_{t-},Y_{t-}^0,Z_{t-}^0,\widehat{Z}_{t-}^0)-h(t,X_t,Y_t,Z_t,\widehat{Z}_t)\Big\}dt\\
&+\Big(u_x(t,X_t)\sigma(t,X_t,Y_t)-Z_t\Big)dW_t\\
&+\int_U \eta(t,X_{t-},\lambda)\Big\{u(t,X_{t-}+\overline{\theta}^{X,Y}(t,\lambda))-u(t,X_{t-}+\overline{\theta}^{X,Y^0}(t,\lambda))\\
&\qquad -u_x(t,X_{t-})[\overline{\theta}^{X,Y}(t,\lambda)-\overline{\theta}^{X,Y^0}(t,\lambda)]\Big\}\nu(d\lambda)dt\\
&+\int_{\mathbb{R}_+\times U}\Big([\Delta_{\overline{\theta}}u](t,X_{t-},\lambda)-\widehat{Z}_{t-}(\lambda)\Big)1_{[0,\eta(t,X_{t-},\lambda)]}(r)\widetilde{N}(drd\lambda dt).
\end{aligned}
$$

Since $\widehat{Z}^0_{t-}(\lambda) = \overline{\theta}^{X,Y^0}(t,\lambda)$, thanks to the Four Step Scheme, we have

$$\begin{aligned}\delta u(\Theta)(t,\lambda) &\stackrel{\triangle}{=} [\Delta_{\overline{\theta}} u](t, X_{t-}, \lambda) - \widehat{Z}_{t-}(\lambda)\\ &= u(t, X_{t-} + \overline{\theta}^{X,Y}(t,\lambda)) - u(t, X_{t-}) - \widehat{Z}_{t-}(\lambda)\\ &= u(t, X_{t-} + \overline{\theta}^{X,Y}(t,\lambda)) - u(t, X_{t-} + \overline{\theta}^{X,Y^0}(t,\lambda)) + \widehat{Z}^0_{t-}(\lambda) - \widehat{Z}_{t-}(\lambda). \end{aligned} \tag{67}$$

Applying Ito's formula, we then get

$$\begin{aligned}|Y^0_T - Y_T|^2 =& |Y^0_t - Y_t|^2\\ &+ \int_t^T 2(Y^0_s - Y_s)\cdot \Big\{ u_x(s,X_s)(b(s,X_s,Y_s,Z_s,\widehat{Z}_s) - b(s,X_s,Y^0_s,Z^0_s,\widehat{Z}^0_s))\\ &\qquad + \frac{1}{2}u_{xx}(s,X_s)\big(\sigma^2(s,X_s,Y_s) - \sigma^2(s,X_s,Y^0_s)\big)\\ &\qquad + h(s,X_s,Y^0_s,Z^0_s,\widehat{Z}^0_s) - h(s,X_s,Y_s,Z_s,\widehat{Z}_s)\Big\} ds\\ &+ \int_t^T \Big|(\sigma(s,X_s,Y_s) - \sigma(s,X_s,Y^0_s))u_x(s,X_s) + Z^0_s - Z_s\Big|^2 ds\\ &+ \int_t^T 2(Y^0_s - Y_s)\int_U \eta(t,X_{t-},\lambda)\Big\{ u(t,X_{t-} + \overline{\theta}(t,\lambda)) - u(t, X_{t-} + \overline{\theta}^0(t,\lambda))\\ &\qquad - u_x(t,X_{t-})[\overline{\theta}(t,\lambda) - \overline{\theta}^0(t,\lambda)]\Big\}\nu(d\lambda)ds\\ &+ \int_t^T \int_{\mathbb{R}_+\times U} \delta u(\Theta)(s,\lambda)\big(2X_{s-} + \delta u(\Theta)(s,\lambda)\big)\mathbf{1}_{[0,\eta(t,X_{s-},\lambda)]}(r)\widetilde{N}(drd\lambda ds)\\ &+ \int_t^T \int_U \eta(s,X_{s-},\lambda)(\delta u(\Theta)^2(s,\lambda)\nu(d\lambda)ds. \end{aligned} \tag{68}$$

Taking expectation on both sides we get

$$\begin{aligned}&\mathbb{E}|Y^0_t - Y_t|^2\\ &= -\mathbb{E}\int_t^T 2(Y^0_s - Y_s)\cdot \Big\{ u_x(s,X_s)\Big(b(s,X_s,Y_s,Z_s,\widehat{Z}_s) - b(s,X_s,Y^0_s,Z^0_s,\widehat{Z}^0_s)\Big)\\ &\qquad + \frac{1}{2}u_{xx}(s,X_s)\big(\sigma^2(s,X_s,Y_s) - \sigma^2(s,X_s,Y^0_s)\big)\\ &\qquad + h(s,X_s,Y^0_s,Z^0_s,\widehat{Z}^0_s) - h(s,X_s,Y_s,Z_s,\widehat{Z}_s)\Big\} ds\\ &\quad - \mathbb{E}\int_t^T \Big|(\sigma(s,X_s,Y_s) - \sigma(s,X_s,Y^0_s))u_x(s,X_s) + Z^0_s - Z_s\Big|^2 ds\\ &\quad - \mathbb{E}\int_t^T 2(Y^0_s - Y_s)\cdot \int_U \eta(t,X_{t-},\lambda)\Big\{ u(t,X_{t-} + \overline{\theta}(t,\lambda)) - u(t, X_{t-} + \overline{\theta}^0(t,\lambda))\\ &\qquad - u_x(t,X_{t-})[\overline{\theta}(t,\lambda) - \overline{\theta}^0(t,\lambda)]\Big\}\nu(d\lambda)ds\\ &\quad - \mathbb{E}\int_t^T \int_U \eta(s,X_{s-},\lambda)(\delta u(\Theta))^2(s,\lambda)\nu(d\lambda)ds. \end{aligned} \tag{69}$$

Let us now note that for any $a, b \in \mathbb{R}$ and $0 < \beta < 1$ , it holds that $(a+b)^2 \geq (1-\beta)a^2 - \frac{1}{\beta}b^2$, thus by the boundedness of $u_x$ and the uniform Lipschitz continuity of $\sigma$, we have

$$\begin{aligned}&\Big|(\sigma(s,X,Y)-\sigma(s,X,Y^0))u_x(s,X)+Z^0-Z\Big|^2\\ \geq\ &(1-\beta)|Z^0-Z|^2-\frac{1}{\beta}|(\sigma(s,X,Y)-\sigma(s,X,Y^0))u_x(s,X)|^2\\ \geq\ &(1-\beta)|Z^0-Z|^2-C|Y^o-Y|^2,\end{aligned}$$

here and in the sequel $C>0$ is again a generic constant which may vary from line to line. By Assumptions H5, H6, and boundedness of $u_x$, $u_{xx}$, we have

$$\begin{aligned}&\big|u_x(s,X_s)(b(s,X_s,Y_s,Z_s,\widehat{Z}_s)-b(s,X_s,Y_s^0,Z_s^0,\widehat{Z}_s^0))\big|\\ \leq\ &C|Y_s-Y_s^0|+|Z_s-Z_s^0|+\|\widehat{Z}_s(\cdot)-\widehat{Z}_s^0(\cdot)\|_{L^2},\\ &\big|h(s,X_s,Y_s,Z_s,\widehat{Z}_s)-h(s,X_s,Y_s^0,Z_s^0,\widehat{Z}_s^0))\big|\\ \leq\ &C|Y_s-Y_s^0|+|Z_s-Z_s^0|+\|\widehat{Z}_s(\cdot)-\widehat{Z}_s^0(\cdot)\|_{L^2};\\ &\big|u_{xx}(s,X_s)\big(\sigma^2(s,X_s,Y_s)-\sigma^2(s,X_s,Y_s^0)\big)\big|\leq C|Y_s-Y_s^0|.\end{aligned}$$

Similarly, by H5 for $p=1$ and the boundedness of $u_x$,

$$\int_U\big|\eta(s,X_{s-},\lambda)u_x(s,X_{s-})[\overline{\theta}(s,\lambda)-\overline{\theta}^0(s,\lambda)]\big|\nu(d\lambda)\leq C|Y_{s-}-Y_{s-}^0|.$$

Also, the boundedness of $u_x$ also implies that

$$\int_U\eta(s,X_{s-},\lambda)\Big\{u(s,X_{s-}+\overline{\theta}(s,\lambda))-u(s,X_{s-}+\overline{\theta}^0(s,\lambda))\Big\}\nu(d\lambda)\leq C|Y_{s-}-Y_{s-}^0|.$$

It remains to estimate the last term of right side of (69). By the definition of $\delta u(\Theta)$ and by the boundedness of $u_x$, we have

$$\begin{aligned}|\delta u(\Theta)(s,\lambda)|^2\geq\ &(1-\beta)\big(\widehat{Z}_{s-}^0(\lambda)-\widehat{Z}_{s-}(\lambda)\big)^2\\ &-\frac{1}{\beta}|u(s,X_{s-}+\overline{\theta}(t,\lambda))-u(s,X_{s-}+\overline{\theta}^0(s,\lambda))|^2\\ \geq\ &(1-\beta)\big|\widehat{Z}_{s-}^0(\lambda)-\widehat{Z}_{s-}(\lambda)\big|^2-C\big|\overline{\theta}(s,\lambda)-\overline{\theta}^0(s,\lambda)|^2.\end{aligned}$$

Thus, the last term in the right hand side of (69) reads

$$\begin{aligned}&\int_U\eta(s,X_{s-},\lambda)|\delta u(\Theta)|^2(s,\lambda)v(d\lambda)\\ \geq\ &(1-\beta)\int_U\eta(s,X_{s-},\lambda)\big|\widehat{Z}_{s-}^0(\lambda)-\widehat{Z}_{s-}(\lambda)\big|^2\nu(d\lambda)\\ &-C\int_U\eta(s,X_{s-},\lambda)|\overline{\theta}(s,\lambda)-\overline{\theta}^0(s,\lambda)|^2\nu(d\lambda)\\ \geq\ &(1-\beta)\int_U\eta(s,X_{s-},\lambda)\big|\widehat{Z}_{s-}^0(\lambda)-\widehat{Z}_{s-}(\lambda)\big|^2\nu(d\lambda)-C|Y_{s-}-Y_{s-}^0|^2.\end{aligned}$$

This leads to that

$$
\begin{aligned}
&\mathbb{E}|Y_t^0 - Y_t|^2 + (1-\beta)\int_t^T \mathbb{E}|Z_s - Z_s^0|^2 ds \\
&+(1-\beta)\int_t^T \mathbb{E}\int_{\mathbb{R}} \eta(s, X_{s-}, \lambda)\big|\widehat{Z}_{s-}^0(\lambda) - \widehat{Z}_{s-}(\lambda)\big|^2 \nu(d\lambda)ds \\
\le\ & C\int_t^{\mathrm{T}} \mathbb{E}\Big\{|Y_s^0 - Y_s|^2 + |Y_s^0 - Y_s|\big(|Z_s^0 - Z_s| + \|\widehat{Z}_s^0(\cdot) - \widehat{Z}_s(\cdot)\|_{L^2}\big)\Big\}ds \\
\le\ & C_\varepsilon \int_t^T \mathbb{E}|Y_s^0 - Y_s|^2 ds + \varepsilon\int_t^T \mathbb{E}|Z_s^0 - Z_s|^2 ds + \varepsilon\int_t^T \mathbb{E}\|\widehat{Z}_s^0(\cdot) - \widehat{Z}_s(\cdot)\|_{L^2}^2 ds,
\end{aligned}
$$

where $\varepsilon > 0$ is arbitrary and $C_\varepsilon$ depends on $\varepsilon$. Since $\beta < 1$, choosing $\varepsilon < 1-\beta$ and applying Gronwall's inequality, we conclude that

$$
Y_t = Y_t^0, \quad Z_t = Z_t^0, \quad \|\widehat{Z}_t^0(\cdot) - \widehat{Z}_t(\cdot)\|_{L^2} = 0 \quad a.s. \quad t \in [0,T].
$$

Thus, any solution of (12) must have the form that we have constructed. This, together with Theorem 1, proves the thoerem. ∎

A direct consequence of the well-posedness of FBSDE is the following Martingale representation Theorem. We should note that the underlying filtration is not an issue here as long as $\mathcal{F}_0 = \overline{\{\Omega, \emptyset\}}^P$, so that $Y_0 = \mathbb{E}[g(X_T)]$.

**Corollary 1** *Let $(\Omega, \mathcal{F}, P, \mathbb{F})$ be a filtered probability space with $\mathcal{F}_0 = \overline{\{\Omega, \emptyset\}}^P$. Let $W$ be an $n$-dimensional, $\mathbb{F}$-Brownian motion, and $\widetilde{N}$ be an $\mathbb{F}$-Poisson martingale measure on $[0,T] \times U \times \mathbb{R}_+$, independent of $W$. Assume* H1–H8*, with $b$ and $\sigma$ depending only on $(t,x)$, and $h \equiv 0$; and let $X$ be the solution to the (decoupled) SDE* (11)*. Then it holds that*

$$
\begin{aligned}
g(X_T) = \mathbb{E}[g(X_T)] &+ \int_0^T \langle \nabla u(s, X_s), \sigma(s, X_s) dW_s\rangle \\
&+ \int_0^T \int_{[0,\infty)\times U} \langle \Delta_\theta u(s, X_{s-}) \mathbf{1}_{[0,\eta(s,X_{s-},\lambda)]}(r) \widetilde{N}(drd\lambda ds)\rangle,
\end{aligned} \tag{70}
$$

*where $u \in C^{1+\alpha/2, 2+\alpha}([0,T] \times \mathbb{R}^n)$ is the classical solution of the following IPDE*:

$$
\begin{cases}
0 = u_t^k(t,x) + \frac{1}{2}\mathrm{tr}\left\{u_{xx}^k(t,x)\sigma(t,x)\sigma(t,x)^{\mathrm{T}}\right\} + \langle u_x^k(t,x), b(t,x)\rangle \\
\qquad + \int_U \eta(t,x,\lambda)\big\{u^k(t, x+\theta(t,x,\lambda)) - u^k(t,x) - \mathbf{1}_{U_1}(\lambda)\langle u_x^k(t,x), \theta(t,x,\lambda)\rangle\big\}\nu(d\lambda), \\
\qquad\qquad (t,x) \in [0,T] \times \mathbb{R}^n, \quad k = 1,2,\cdots,m, \\
u(T,x) = g(x), \qquad x \in \mathbb{R}^n.
\end{cases}
$$

## 6 Possible Extensions

The Four Step Scheme proposed could be extended to cases where the regularity of the coefficients are reduced, provided that filtration is generated by $(W, N)$, the driving Brownian motion and the Poisson random measure. The arguments can follow exactly as those of [2], which we now describe briefly.

### 6.1 FBSDEs on a Small Duration

We first consider the following FBSDE on a given filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{F})$:

$$\begin{cases} X_t = \xi + \displaystyle\int_0^t b(s, X_s, Y_s, Z_s, \widehat{Z}_s(\cdot))ds + \int_0^t \sigma(s, X_s, Y_s)dW_s \\ \qquad + \displaystyle\int_0^t \int_{\mathbb{R}_+ \times U} \theta(s, X_{s-}, Y_{s-}, \lambda)\mathbf{1}_{[0,\eta(s,X_{s-},\lambda)]}(r)\widetilde{N}(drd\lambda ds), \\ Y_t = g(X_T) + \displaystyle\int_t^T h(s, X_s, Y_s, Z_s, \widehat{Z}_s(\cdot))ds - \int_t^T Z_s dW_s \\ \qquad - \displaystyle\int_t^T \int_{\mathbb{R}_+ \times U} \widehat{Z}_{s-}(\lambda)\mathbf{1}_{[0,\eta(s,X_{s-},\lambda)]}(r)\widetilde{N}(drd\lambda ds), \end{cases} \tag{71}$$

where we assume that $W$ is an $\mathbb{F}$-BM, and $N$ is an $\mathbb{F}$-Poisson random measure on $[0,T] \times U \times \mathbb{R}_+$. Denote $\mathcal{F}_t^0 = \mathcal{F}_0 \vee \mathcal{F}_t^{W,N}$, $\forall t$, and assume that $\xi \in \mathcal{F}_0$.

We shall assume that all coefficients are deterministic, are of linear growth, and are uniformly Lipschitz in the variables $(x, y, z, \widehat{z})$. It is then not hard to prove, using the contraction mapping theorem, as well as the Martingale representation theorem (cf. e.g., [14]) that there exists a $T_0 > 0$, depending only on the Lipschitze constant of the coefficients, such that for every $\xi \in L^2(\mathcal{F}_0; \mathbb{R}^d)$, the solution $(X, Y, Z, \widehat{Z})$ to FBSDE (71) exists and is unique, over any $[0,T]$, whenever $T \le T_0$. Furthermore, the process $(X, Y)$ has càdlàg paths; and

$$\mathbb{E}\Big\{ \sup_{t\in[0,T]} |X_t|^2 + \sup_{t\in[0,T]} |Y_t|^2 \Big\} < \infty.$$

### 6.2 Construction of Decoupling Function on a Small Duration

We consider a slightly modified form of the FBSDE (71), defined on $[t,T]$, and $\xi \in \mathcal{F}_t$. Then if $T - t \le T^0$, there exists a unique solution to the FBSDE (71), over $[t,T]$. Denote the solution by $(X_s^{t,x}, Y_s^{t,x}, Z_s^{t,x}, \widehat{Z}_s^{t,x})$, for $s \in [t,T]$. We extend it to $[0,T]$ by setting

$$X_s^{t,x} = x, \quad Y_s^{t,x} = Y_t^{t,x}, \quad Z_s^{t,x} = 0, \quad \widehat{Z}_s^{t,x} = 0, \qquad s \in [0,t].$$

Finally, we define the (deterministic) mapping $(t,x) \mapsto Y_t^{t,x}$ by $u(t,x)$.

Next, we can follow the similar argument as in the uniqueness proof to obtain the continuous dependence of the solution on the initial date $(t,\xi)$, and in particular the following estimates for the function $\theta$: for some constants $C_1$, $C_2$, $C_3 > 0$, depending only on the Lipschitz constant of the coefficients, such that

$$\begin{cases} |u(t,x)|^2 \le C_1(1 + |x|^2); \\ |u(t',x') - u(t,x)| \le C_2|x - x'|^2 + C_3(1 + |x|^2)|t - t'|; \\ \forall t \in [0,T], \forall \xi \in L^2(\mathcal{F}_t; \mathbb{R}^n), \exists N \in \mathcal{F}_0, \mathbb{P}(N) = 0, \text{ such that} \\ Y_s^{t,\xi}(\omega) = u(s, X_s^{t,\xi}(\omega)), \qquad \forall s \in [t,T], \ \forall \omega \notin N. \end{cases} \tag{72}$$

### 6.3 Mollification and Stability Reuslt

In order to use the function $\theta$ for decoupling, one would hope that it is the solution to the quasilinear IPDE. But this is impossible unless the coefficients are smooth. The main idea here is that to combine the decoupling scheme with the "stability" result. Namely, we mollify the coefficients $(b, \sigma, h, g)$ to the smooth functions $\{b^n, \sigma^n, h^n, g^n)$ and denote the corresponding solutions of FBSDEs by $(X^n, Y^n, Z^n, \widehat{Z}^n(\cdot))$, $n = 1, 2, \cdots$, and the decoupling function by $u^n(\cdot,\cdot)$. Since each $u^n$ satisfies (72) with constants independent of $n(!)$, we can conclude that they converge to a function $\widetilde{u}$, which also satisfies (72). But on the other hand, a stability result will show that $(X^n, Y^n, Z^n, \widehat{Z}^n)$ converges to the true solution $(X, Y, Z, \widehat{Z})$ (which only need the argument as in the uniqueness proof. Therefore, we conclude by the uniqueness of the solution that $\widetilde{u}(s, x) \equiv u(s, x)$, for all $(s, x) \in [t, T] \times \mathbb{R}^n$, whenever the duration $T - t < T_0$.

### 6.4 Construction of "Global" Decoupling Function

To construct the global solution one can use a "running down" induction: Partition the interval $[0, T]$ into $0 = t_0 < t_1 < \cdots < t_N = T$, s.t. $t_{i+1} - t_i = T/N < T^0$. Consider the following FBSDEs on $[t, t_{i+1}]$, $i = N-1, \cdots, 1$:

$$\begin{cases} X_s = \xi + \displaystyle\int_t^s b(r, X_r, Y_r, Z_r, \widehat{Z}_r(\cdot))dr + \int_t^s \sigma(r, X_r, Y_r)dW_r \\ \qquad + \displaystyle\int_t^s \int_{\mathbb{R}_+ \times U} \theta(s, X_{s-}, Y_{s-}, \lambda)\mathbf{1}_{[0,\eta(s,X_{s-},\lambda)]}(r)\widetilde{N}(drd\lambda ds), \\ Y_t = u(t_{i+1}, X_{t_{i+1}}) + \displaystyle\int_s^{t_{i+1}} h(s, X_s, Y_s, Z_s, \widehat{Z}_s(\cdot))ds - \int_s^{t_{i+1}} Z_s dW_s \\ \qquad - \displaystyle\int_s^{t_{i+1}} \int_{\mathbb{R}_+ \times U} \widehat{Z}_{s-}(\lambda)\mathbf{1}_{[0,\eta(s,X_{s-},\lambda)]}(r)\widetilde{N}(drd\lambda ds). \end{cases} \tag{73}$$

Denote the solution to each (local) FBSDE by $(X^{t,\xi,i}, Y^{t,\xi,i}, Z^{t,\xi,i}, \widehat{Z}^{t,\xi,i}(\cdot))$. Then we define $u(t, x) = Y_t^{t,x,i}$, for $t \in [t_i, t_{i+1}]$ is the desired global "decoupling" function.

### 6.5 Construction of Global Solution

Once the "decoupling machine" $u$ is defined on $[0, T] \times \mathbb{R}^n$, then the following "running-up" induction gives the desired solution on $[0, T]$: First, for $0 \le s \le t_1$, let $\Theta^{(0)} \triangleq (X^{(0)}, Y^{(0)}, Z^{(0)}, \widehat{Z}^{(0)})$ solve the FBSDE:

$$X_s^{(0)} = x + \int_0^s b(r, \Theta_r^{(0)})dr + \int_t^s \sigma(r, X_r^{(0)}, Y_r^{(0)})dW_r$$

$$+ \int_0^s \int_{\mathbb{R}_+ \times U} \theta(r, X_{r-}^{(0)}, Y_{r-}^{(0)}, \lambda)\mathbf{1}_{[0,\eta(r,X_{r-}^{(0)},\lambda)]}(v)\widetilde{N}(dvd\lambda dr),$$

$$Y_s^{(0)} = \theta(t_1, X_{t_1}^{(0)}) + \int_s^{t_{i+1}} h(r, \Theta^{(0)})ds - \int_s^{t_1} Z_r^{(0)} dW_r$$

$$- \int_s^{t_{i+1}} \int_{\mathbb{R}_+ \times U} \widehat{Z}_{r-}^{(0)}(\lambda)\mathbf{1}_{[0,\eta(r,X_{r-}^{(0)},\lambda)]}(v)\widetilde{N}(dvd\lambda dr).$$

Suppose that the solution have been constructed on $[0,t_{k-1}]$, for $t_{k-1}\le s\le t_k$, let $\Theta^{(k)}\triangleq(X^{(k)},Y^{(k)},Z^{(k)},\widehat{Z}^{(k)}(\cdot))$ solve the FBSDE:

$$
\begin{aligned}
X_s^{(k)} &= X_{t_{k-1}}^{(k-1)} + \int_{t_{k-1}}^s b(r,\Theta_r^{(k)})dr + \int_{t_{k-1}}^s \sigma(r,X_r^{(k)},Y_r^{(k)})dW_r\\
&\quad + \int_{t_{k-1}}^s \int_{\mathbb{R}_+\times U} \theta(s,X_{r-}^{(k)},Y_{r-}^{(k)},\lambda)\mathbf{1}_{[0,\eta(r,X_{r-}^{(k)},\lambda)]}(v)\widetilde{N}(dvd\lambda dr),\\
Y_s^{(k)} &= \theta(t_k,X_{t_k}^{(k)}) + \int_s^{t_k} h(r,\Theta^{(k)})ds - \int_s^{t_k} Z_r^{(k)}dW_r\\
&\quad - \int_s^{t_{i+1}}\int_{\mathbb{R}_+\times U} \widehat{Z}_{r-}^{(k)}(\lambda)\mathbf{1}_{[0,\eta(r,X_{r-}^{(k)},\lambda)]}(v)\widetilde{N}(dvd\lambda dr).
\end{aligned}
$$

It remains to verify that the solutions on each sub-intervals will actually "patch-up". But it would suffice to observe the following facts from the construction:

$$
X_{t_k}^{(k-1)} = X_{t_k}^{(k)} \quad \text{and} \quad Y_{t_k}^{(k)} = \theta(t,X_{t_k}^{(k)}) = \theta(t,X_{t_k}^{(k-1)}) = Y_{t_k}^{(k-1)}.
$$

This completes the construction of a global solution to the FBSDE (1).

## References


[1] F. Antonelli, Backward-forward stochastic differential equations, *Ann. Applied Probab.*, 1993, **3**: 777–793.

[2] F. Delarue, On the Existence and Uniqueness of Solutions to FBSDEs in a Non-Degenerate case, *Stoch. Proc. Appl.*, 2002, **99**: 209–286.

[3] Y. Hu and S. Peng, Solution of forward-backward stochastic differential equations, *Probab. Theory Related Fields*, 1995, **103**: 273–283.

[4] J. Ma, P. Protter, and J. Yong, Solving forward-backward stochastic differential equations explicitly—a four step scheme, *Prob. Theory & Rel. Fields*, 1994, **98**: 339–359.

[5] J. Yong, Finding adapted solutions of forward-backward stochastic differential equations: Method of continuation, *Probab. Theory Related Fields*, 1997, **107**: 537–572.

[6] J. Ma and J. Yong, Forward-backward stochastic differential equations and their applications, *Lecture Notes in Mathematics*, Springer-Verlag, 1999, **1702**.

[7] Z. Wu, Fully coupled FBSDE with Brownian motion and Poisson process in stopping time duration, *J. Aust. Math. Soc.*, 2003, **74**(2): 249–266.

[8] J. Yin and R. Situ, On solutions of forward-backward stochastic differential equations with Poisson jumps, *Stoch. Anal. Appl.*, 2003, **21**(6): 1419–1448.

[9] E. Çinlar and J. Jacod, *Representation of Semimartingale Markov Processes in Terms of Wiener Processes and Poisson Random Measures*, Seminar on Stochastic Processes, 1981 (Evanston, Ill., 1981, 159–242, Progr. Prob. Statist., 1, Birkhäuser, Boston, Mass., 1981.

[10] E. Ç inlar, J. Jacod, P. Protter, and M. J. Sharpe, *Semimartingales and Markov processes*, ed. by Z. Wahrsch. Verw. Gebiete, 1980, **54**(2): 161–219.

[11] P. Protter, *Stochastic Integration and Differential Equations*, Springer-Verlag, Berlin, 1990.

[12] T. G. Kurtz and P. Protter, Weak convergence of stochastic integrals and differential equations II: Infinite dimensional case, Probabilistic models for nonlinear partial differential equations, *Lecture Notes in Math.*, Springer, Berlin, 1996.

[13] O. A. Ladyzenskaja, V. A. Solonnikov, and N. N. Ural'ceva, Linear and quasilinear equations of paralic type, *Amer. Math. Soc.*, R. I. Providence, 1968.

[14] J. Jacod and A. N. Shiryaev, Limit Theorems for Stochastic Processes, Second edition, Grundlehren der Mathematischen Wissenschaften, Fundamental Principles of Mathematical Sciences, Springer-Verlag, Berlin, 2003, **288**.

[15] U. G. Haussmann, On the integral representation of functionals of Itô's processes, *Stochastics*, 1979, **3**: 17–27.

[16] I. Karatzas, D. L. Ocone, and J. Li, An extension of Clark's formula, Stochastics and Stochastic Reports, 1991, **37**: 127–131.

[17] J. Ma and J. Yong, Solvability of forward-backward SDEs and the nodal set of Hamilton-Jacobi-Bellman equations, *Chinese Ann. Math. Ser. B*, 1995, **16**: 279–298.

[18] D. Nualart, Noncausal stochastic integrals and calculus, *Stochastic Analysis and Related Topics, Lecture Notes in Math.*, 1988, **1316**: 80–129.

[19] D. Ocone, Malliavin's calculus and stochastic integral representations of functionals of diffusion processes, *Stochastics*, 1984, **12**: 161–185.

[20] E. Pardoux and S. Peng, Adapted solution of a backward stochastic differential equation, *Systems & Control Letters*, 1990, **14**: 55–61.