Taylor & Francis
Taylor & Francis Group

Stochastic Analysis and Applications

ISSN: 0736-2994 (Print) 1532-9356 (Online) Journal homepage: http://www.tandfonline.com/loi/lsaa20

Forward-backward SDEs with discontinuous coefficients

Jianfu Chen, Jin Ma & Hong Yin

To cite this article: Jianfu Chen, Jin Ma & Hong Yin (2017): Forward-backward SDEs with discontinuous coefficients, Stochastic Analysis and Applications, DOI: 10.1080/07362994.2017.1399799

To link to this article: https://doi.org/10.1080/07362994.2017.1399799

Published online: 07 Dec 2017.

STOCHASTIC ANALYSIS AND APPLICATIONS
2017, VOL. 0, NO. 0, 1–21
https://doi.org/10.1080/07362994.2017.1399799

Forward-backward SDEs with discontinuous coefficients

Jianfu Chen[a], Jin Ma[b], and Hong Yin[c]

[a]HSBC, New York, NY; [b]Department of Mathematics, University of Southern California, Los Angeles, CA; [c]Department of Mathematics, State University of New York, Brockport, NY

ABSTRACT

In this paper, we are interested in the well-posedness of a class of fully coupled forward-backward SDE (FBSDE) in which the forward drift coefficient is allowed to be discontinuous with respect to the backward component of the solution. Such an FBSDE is motivated by a practical issue in regime-switching term structure interest rate models, and the discontinuity makes it beyond any existing framework of FBSDEs. In a Markovian setting with non-degenerate forward diffusion, we show that a decoupling function can still be constructed and that it is a Sobolev solution to the corresponding quasilinear PDE. As a consequence we can then argue that the FBSDE admits a weak solution in the sense of [1, 2]. In the one-dimensional case, we further prove that the weak solution of the FBSDE is actually strong, and it is pathwisely unique. Our approach does not use the well-known Yamada–Watanabe Theorem, but instead follows the idea of Krylov for SDEs with measurable coefficients.

ARTICLE HISTORY
Received 8 May 2017
Accepted 30 October 2017

KEYWORDS
Forward-backward SDEs; decoupling function; discontinuous coefficients; distribution solution; weak solution; Krylov estimates; strong solution

MATHEMATICS SUBJECT CLASSIFICATION
60H10; 60H30; 35K58; 91G30

1. Introduction

In this paper, we are interested in the well-posedness of the following Markovian-type "regime-switching" forward-backward stochastic differential equation (FBSDE):

$$\begin{cases} X_t = x + \int_0^t b(s, X_s, Y_s)ds + \int_0^t \sigma(s, X_s, Y_s)dW_s \\ Y_t = g(X_T) - \int_t^T h(s, X_s, Y_s, Z_s)ds - \int_t^T Z_s dW_s, \\ X_0 = x, \quad Y_T = g(X_T). \end{cases} \tag{1.1}$$

We assume that the coefficients σ, h, and g are deterministic Lipschitz functions, but the drift coefficient b takes the following form:

$$b(t, x, y) = \sum_{i=1}^{m} b^i(t, x)\mathbf{1}_{[a_i, a_{i+1})}(y), \qquad (t, x, y) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}, \tag{1.2}$$

where $-\infty < a_1 < a_2 < \cdots < \infty$ is a finite partition of $\mathbb{R}$, and b^i's are deterministic Lipschitz functions. The main feature of this FBSDE is that the coefficient b has, albeit finitely many, jumps in the variable Y. Thus, it is beyond the scope of any existing literature of FBSDEs, for which the weakest assumption this far is "Lipschitz" in all spatial variables (see, e.g., [3]), as far as the strong solution is concerned.

CONTACT Jin Ma jinma@usc.edu Department of Mathematics, University of Southern California, Los Angeles, CA 90089, USA.

Our problem is motivated by the following "regime-switching" term structure model that is often seen in practice. Consider, for example, the Black-Karasinski short rate model that is currently popular in the industry: let $r = \{r_t : t \geq 0\}$ be the short rate process, and $X_t = \ln r_t$, $t \geq 0$. Then X satisfies the following SDE:

$$dX_t = k(\theta_t - X_t)dt + \sigma dW_t, \tag{1.3}$$

where W is a standard Brownian motion. A simple "regime-switching" version of (1.3) is that the mean reversion level θ shifts between two values $\theta_t \in \{b_1, b_2\}$. The existence of such structural shift was supported by empirical evidence (see, e.g., [4, 5]), and many dynamic models of the short rate have been proposed, and some of them are hidden Markovian in nature, that is, the switch is triggered by an exogenous factor (diffusion) process Y so that $\theta_t = b(Y_t)$, where $b(y) \in \{b_1, b_2\}$ (see, e.g., [6–9]). In particular, if we consider the case in which the triggering process is the long-term rate, then following the argument of a term structure model (see, for example, Duffie-Ma-Yong [10]), and assuming the triggering level to be $\alpha > 0$, we can derive an FBSDE with discontinuous coefficient:

$$\begin{cases} dX_t = [b(Y_t) - \beta X_t]dt + \sigma dW_t \\ dY_t = [e^{X_t} Y_t - 1]dt + Z_t dW_t \\ X_0 = x, \quad Y_T = g(X_T), \end{cases} \tag{1.4}$$

where $X_t = \ln r_t$, Y_t is the long-term treasury bond price; $b(y) = b_1 \mathbf{1}_{\{y \leq \alpha\}} + b_2 \mathbf{1}_{\{y > \alpha\}}$, $b_1 \neq b_2$; and α, β, σ are constants. Clearly, this is a special case of the FBSDE (1.1), and we are aiming at finding its strong solution in the sense of SDEs.

We should note that the FBSDEs (1.1) or (1.4) has discontinuities in its drift, in the component Y. Thus, they fall outside most of the existing works in the literature where at least the Lipschitz continuity of the coefficients on the component Y is assumed (see [11, 12], and recently [3]). In a recent work [13], the fully coupled FBSDEs with discontinuous terminal coefficient was studied, and the well-posedness was proved in the case when the FBSDE takes a simple form, and the forward diffusion is non-degenerate. The FBSDE considered in this paper is more general, and the singularity could appear in any time $t \in [0, T]$. Moreover, it should be emphasized that, unlike the cases of FBSDE with discontinuous coefficients studied previously (see, e.g., [14]), the discontinuity of the coefficients considered in this paper appear on the variable Y, which, in view of the corresponding PDE, is the "solution" variable. Thus, the singularity is much more malignant, and to our best knowledge, it has not been studied in any form.

Our plan of attack is quite standard. Since the FBSDE is Markovian, we shall first mollify the coefficients so that the FBSDE can be solved by Four Step Scheme (see [12, 15]). Then, with the help of some uniform estimates on the solution to the decoupling PDEs, as well as its derivatives, we can find a sequence of solutions that converges relatively nicely to a limit function. The main task is then to argue that this function is exactly a desired "decoupling" function. We shall carry out this task by first showing that the limit function is indeed a solution to the corresponding quasilinear PDE in the "distribution" sense. It then follows from the argument of Kim–Krylov [16] that the regularity of this distribution solution can be raised to $W_2^{1,2}((0, T) \times \mathbb{R})$, so that the Itô–Krylov formula can be applied. This then easily leads to the existence of a weak solution of the FBSDE (1.1) in the sense of [1] and [2]. Furthermore, in light of the well-known Yamada–Watanabe Theorem (or the FBSDE version by [1] and [17]), to obtain the strong existence of the solution we need only show the pathwise uniqueness of the FBSDE (1.1). But this becomes quite challenging with the presence of the discontinuity in the coefficients. We therefore take a different route. Namely, we prove directly that the weak existence and uniqueness will actually guarantee the strong existence and uniqueness. The key trick is a weak convergence argument using the so-called Krylov-type estimate, and the comparison theorems for SDE with regular coefficients. We note that such an argument has been frequently employed for PDEs and/or SDEs with measurable coefficients (see, e.g., [18], [19], or [20]), and it works quite effectively in the current situation.

The rest of the paper is organized as follows. In Section 2, we give the necessary preparations and introduce notations. In Section 3, we construct the decoupling function and prove some uniform estimates for the approximating PDEs, which leads to the facts that the limit function is a distribution solution of the FBSDE, and that it belongs to $W_2^{1,2}([0,T]\times\mathbb{R})$. In Section 4, we prove the weak-well-posedness. In Section 5, we show directly that the weak solution of FBSDE (1.1) is actually strong, and is pathwisely unique.

2. Preliminary

Throughout this paper we consider a complete, filtered probability space $(\Omega, \mathscr{F}, \mathbb{P}; \mathbb{F})$ on which is defined a standard Brownian motion W. We shall assume that the filtration $\mathbb{F} = \{\mathscr{F}_t\}_{t\geq 0}$ is Brownian, that is, $\mathbb{F} = \{\mathscr{F}_t^W\}_{t\geq 0}$, and is augmented by all the P-null sets of $\mathcal{F}$ so that it satisfies the usual hypotheses (see [21]).

In this paper, we shall content ourselves with one-dimensional case. Let $[0,T]$ be a given finite time interval, and denote $R_T \triangleq \{(t,x) : t\in[0,T], x\in\mathbb{R}\}$. The spaces $L^2(R_T)$ and $C^{1,2}(R_T)$ are defined in obvious ways. For a function u defined on R_T, we denote its generalized derivatives by $\partial_t u = u_t$, $\partial_x u = u_x$, and $\partial^2_{xx} u = u_{xx}$, and the Sobolev spaces $H^1(R_T) \triangleq \{u\in L^2(R_T) : u_x \in L^2(R_T)\}$, $W_2^{1,2}(R_T) \triangleq \{u \in L^2(R_T) : u_t, u_x, u_{xx} \in L^2(R_T)\}$ will be used frequently in the sequel.

The main objective of this paper is the following (Markovian) Forward-Backward SDE (FBSDE): for $t\in[0,T]$,

$$\begin{cases} X_t = x + \int_0^t b(s,X_s,Y_s)ds + \int_0^t \sigma(s,X_s,Y_s)dW_s, \\ Y_t = g(X_T) + \int_t^T h(s,X_s,Y_s,Z_s)ds - \int_t^T Z_s dW_s, \end{cases} \tag{2.1}$$

where

$$b(t,x,y) = \sum_{i=1}^N b^i(t,x)\mathbf{1}_{[y_{i-1},y_i)}(y), \qquad \text{for } (t,x)\in[0,T]\times\mathbb{R}^n, \tag{2.2}$$

and $-\infty < y_1 < \cdots < y_N < \infty$. For simplicity, we shall assume that all the processes involved are one-dimensional, but as we shall see later, in most of the situation the higher dimensional cases, especially for the forward component X and the Brownian motion W, can be argued in an identical way without substantial difficulties.

Our objective is to understand the well-posedness of the solution to (2.1). To this end, we first recall the two definitions of the solutions (see [1] and [2]).

Definition 2.1. For a given standard set-up $(\Omega, \mathscr{F}, \mathbb{P}, \mathbf{F}, W)$, a triplet of process (X,Y,Z), defined on the set-up, is called a strong solution to (2.1) if

1. (X,Y,Z) are $\mathscr{F}_t$-adapted process, and X, Y are continuous, such that

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

2. (X,Y,Z) satisfies (2.1) $\mathbb{P}$-almost surely.

Definition 2.2. A triplet of process (X, Y, Z) along with a standard set-up $(\Omega, \mathscr{F}, \mathbb{P}, \mathbf{F}, W)$ on which X, Y, Z are defined, is called a weak solution to (2.1) if

1. (X, Y, Z) are $\mathscr{F}_t$-adapted process, and X, Y are continuous;
2. denoting $f_t = f(t, X_t, Y_t)$ for $f = b, \sigma, h$, it holds that

$$\mathbb{P}\Big\{ \int_0^T (|b_t| + |\sigma_t|^2 + |h_t| + |Z_t|^2)ds + |g(X_T)| < \infty \Big\}$$

3. (X, Y, Z) satisfies (2.1) $\mathbb{P}$-almost surely.

In this paper, we shall focus on the Markovian case, namely, we assume that the coefficients $b^i : R_T \to \mathbb{R}, \sigma : R_T \times \mathbb{R} \to \mathbb{R}, h : R_T \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, and $g : \mathbb{R} \to \mathbb{R}$ are all *deterministic* measurable functions, such that the following *Standing Assumptions* hold throughout the paper.

(H1) Each $b^i(t, x)$ is bounded, continuous in t, and uniformly Lipschitz in x, with Lipschitz constant $K > 0$;

(H2) The function $\sigma(t, x, y)$ is continuous, and there exist constants $0 < \underline{\sigma} < \overline{\sigma}$ such that

$$\underline{\sigma} \leq \sigma(t, x, y) \leq \overline{\sigma}.$$

Furthermore, for fixed t, $\sigma(t, \cdot, \cdot)$, along with its spatial derivatives, σ_x and σ_y are all uniformly Lipschitz in (x, y) with Lipschitz constant $\Upsilon > 0$;

(H3) The function $h(t, x, y, z)$ is bounded, continuous in t and uniformly Lipschitz in (x, y, z) with Lipschitz constant $K > 0$;

(H4) The function $g(x)$ is bounded and uniform Lipschitz continuous with Lipschitz constant $\Upsilon > 0$.

Remark 2.1.

(i) We should note that although each b^i is regular, it is "piecewise continuous" in the variable y (see (2.2)), which is the main difficulty in this paper.

(ii) Because of the discontinuity of the drift coefficient b in the variable y, we need the non-degeneracy condition (H2). The following simple example shows that otherwise the pathwise uniqueness fails, even under the well-known "monotonicity condition" (see, e.g., [22, 23]).

Example 2.1. Let $\sigma = 0$, $g(x) = x$, $X_0 = 0$. Define $H(x) \stackrel{\triangle}{=} \mathbf{1}_{(0,\infty)}(x)$, $x \in \mathbb{R}$, and let $b(t, x, y) = H(x)$, $h(t, x, y) = H(y)$, $(t, x, y) \in R_T \times \mathbb{R}$. That is, we consider the following FBSDE:

$$\begin{cases} X_t = \int_0^t H(Y_s)ds \\ Y_t = X_T - \int_t^T H(X_s)ds - \int_t^T Z_s dW_s. \end{cases} \tag{2.3}$$

Then, it is easily to check that, for any $a \in (0, T)$, let $X_t^a \stackrel{\triangle}{=} (t - a)\mathbf{1}_{[a,T]}(t)$, $Y^a \equiv X^a$, and $Z^a \equiv 0$, then $\{(X^a, Y^a, Z^a) : a \in (0, T)\}$ are infinitely many solutions to (2.3).

It is by now well-understood that, in order to solve a fully coupled FBSDE one should look for a "decoupling random field" $u(t, x, \omega)$, such that $Y_t \equiv u(t, X_t)$, for all $t \in [0, T]$, $\mathbb{P}$-a.s. (see, e.g., [3]). In the Markovian case, the decoupling field should be a deterministic function $u : R_T \mapsto \mathbb{R}$, and in light of the Four Step Scheme (see [12] and [15]), we expect that such a function u should solve the following quasilinear PDEs, in a certain sense:

$$\begin{cases} 0 = u_t + \frac{1}{2}\sigma^2(t,x,u)u_{xx} + b(t,x,u)u_x + h(t,x,u,\sigma(t,x,u)u_x), & (t,x) \in [0,T) \times \mathbb{R}; \\ u(T,x) = g(x). \end{cases} \tag{2.4}$$

We should note that the main difficulty in solving the PDE (2.4) is again the discontinuity of the coefficient b in the variable u. In fact, to our best knowledge there has been no result in the literature that treats quasilinear PDEs with discontinuous coefficients in such a manner. A rather natural plan of attack, however, is to mollify the coefficient b to obtain a family of smooth coefficients $\{b^\varepsilon\}$, such that for each ε the PDE (2.4) has a classical solution, and then analyze the limit as $\varepsilon \to 0$. We shall follow such a route.

To end this section, we shall list some existing a priori estimates for the solution to the "regulated" version of PDE (2.4) (i.e., with b being mollified). These results are mainly based on the works [24], [14], [25], [26], [27], and [28], with slight modifications to fit our needs. We should note that the key points here is that these estimates are independent of the Lipschitz constant of b, whence ε(!). Let us denote the uniform bounded in Assumption (**H**) by Λ and recall that the Lipschitz constant of σ and g are denoted by Υ. In what follows we shall denote $C_u > 0$ to be a generic constant depending only on Υ, $\underline{\sigma}$, $\overline{\sigma}$, Λ, and T, but independent of the Lipschitz constant of b. We shall allow C_u vary from line to line for notational simplicity.

The following result is well-understood (see, e.g., [14, Theorem 3.2]).

Theorem 2.1. *Assume (H2)–(H4), and assume that b is bounded and smooth. Let u be the (classical) solution to (2.4). Then there exists a constant C_u, depending only on Υ, $\underline{\sigma}$, $\overline{\sigma}$, Λ, and T, and a constant $0 < \alpha \le 1$, depending only on $\underline{\sigma}$, $\overline{\sigma}$, and Λ, such that for any (t,x) and (s,y) in $[0,T] \times \mathbb{R}$,*

$$|u(t,x)| \le C_u, \tag{2.5}$$

$$|u(t,x) - u(s,y)| \le C_u(|x-y|^\alpha + |s-t|^{\alpha/2}). \tag{2.6}$$

Our argument for the a priori estimates follows closely that of [14], in which the following linear PDE plays an important role:

$$\begin{cases} v_t + \frac{1}{2}\sigma(t)v_{xx} + \varphi(t,x) = 0, & (t,x) \in [0,T) \times \mathbb{R} \\ v(T,x) = \rho(x), & x \in \mathbb{R}, \end{cases} \tag{2.7}$$

where σ (with a slight abuse of notation) satisfies (H2) and φ is a continuous function. In what follows we collect some results regarding the solution to the linear PDE (2.7) from [14]. We shall denote

$$\Gamma(t,s) \triangleq \int_t^s c(u)du, \text{ for all } 0 \le t < s \le T. \tag{2.8}$$

Lemma 2.1. *Assume (H2). Then, there exists a constant $C > 0$, depending on $\underline{\sigma}$, $\overline{\sigma}$, and p only, such that for every $\eta \in L^p(\mathbb{R})$ and any $(t,s,x) \in [0,T) \times [0,T] \times \mathbb{R}$, $t < s$,*

$$\left| \int_{\mathbb{R}} \eta\left(x + \Gamma^{1/2}(t,s)z\right) |z| \exp(-z^2/2)dz \right| \le C(s-t)^{-\frac{1}{2p}} \left[\int_{\mathbb{R}} |\eta(z)|^p dz \right]^{1/p}. \tag{2.9}$$

Theorem 2.2. *Assume (H2). Assume also that σ, φ, ρ are all bounded, such that $t \mapsto \sigma(t)$ is Hölder continuous; and for some $\beta > 0$, $\varphi \in \mathcal{C}^{\beta/2,\beta}([0,T] \times \mathbb{R}, \mathbb{R})$ and $\rho \in \mathcal{C}^{2+\beta}(\mathbb{R}, \mathbb{R})$. Then*

(i) there is a unique bounded solution $v \in \mathcal{C}^{1+\beta/2, 2+\beta}([0,T] \times \mathbb{R}, \mathbb{R})$;

(ii) *there exists a constant C_v, depending on $\underline{\sigma}$ and $\overline{\sigma}$ only, such that for all $(t,x) \in [0,T) \times \mathbb{R}$,*

$$
\begin{aligned}
|v_x(t,x)| \le C_v \Bigg\{ (T-t)^{-1/2} \int_{\mathbb{R}} \Big[\big|\rho\big(x + \Gamma^{1/2}(t,T)z\big) - \rho(x)\big| |z| \exp(-z^2/2) \Big] dz \\
+ \int_t^T \int_{\mathbb{R}} \Big[(s-t)^{-1/2} \big|\varphi\big(s, x + \Gamma^{1/2}(t,s)z\big)\big| |z| \exp(-z^2/2) \Big] dz ds \Bigg\};
\end{aligned}
\tag{2.10}
$$

(iii) *furthermore, if $\varphi(t,0) \equiv 0$, $t \in [0,T]$, then for any $t \in [0,T)$,*

$$
\begin{aligned}
\big|v_{xx}(t,0)\big| \le C_v \Bigg\{ (T-t)^{-1} \int_{\mathbb{R}} \Big[\big|\rho\big(\Gamma^{1/2}(t,T)z\big) - \rho(0)\big| (1+z^2) \exp(-z^2/2) \Big] dz \\
+ \int_t^T \int_{\mathbb{R}} \Big[(s-t)^{-1} \big|\varphi\big(s, \Gamma^{1/2}(t,s)z\big)\big| (1+z^2) \exp(-z^2/2) \Big] dz ds \Bigg\};
\end{aligned}
\tag{2.11}
$$

(iv) *if $\rho \equiv 0$, then for any $\theta \in (0,1)$, set $\theta' = (1+\theta)/2$. Then for any $p \ge 1$, $z \in \mathbb{R}$ and $R > 0$, the solution v satisfies*

$$
\begin{aligned}
&(1-\theta)^{2p} R^{2p} \int_0^T \int_{z-\theta R}^{z+\theta R} \big|v_{xx}(t,x)\big|^p dx dt \\
&\le C_v \left[(1-\theta')^{2p} R^{2p} \int_0^T \int_{z-\theta' R}^{z+\theta' R} \big|\varphi(t,x)\big|^p dx dt + \int_0^T \int_{z-\theta' R}^{z+\theta' R} \big|v(t,x)\big|^p dx dt \right] \\
&\quad + \frac{1}{2} (1-\theta')^{2p} R^{2p} \int_0^T \int_{z-\theta' R}^{z+\theta' R} \big|v_{xx}(t,x)\big|^p dx dt,
\end{aligned}
$$

where C_v^p depends only on $\underline{\sigma}$, $\overline{\sigma}$, and p;

(v) *if $\varphi \equiv 0$, then the linear PDE*

$$
\begin{cases}
w_t + \frac{1}{2}\sigma^2(t,x,u(t,x)) w_{xx} = 0, & (t,x) \in [0,T) \times \mathbb{R} \\
w(T,x) = g(x), & x \in \mathbb{R}
\end{cases}
\tag{2.12}
$$

admits a unique bounded strong solution $w \in \mathcal{C}^{1,2}([0,T] \times \mathbb{R}, \mathbb{R})$, and there exist two constants C_w and r_w, depending on Υ, $\underline{\sigma}$, $\overline{\sigma}$, Λ, and T, such that for all $(t,x) \in [0,T) \times \mathbb{R}$,

$$
\big|w_{xx}(t,x)\big| \le C_w (T-t)^{-1+r_w}. \tag{2.13}
$$

Finally, we recall Gagliardo–Nirenberg inequality.

Lemma 2.2 (Lemma 6.3, [14]). *Consider a triple (p, q_1, q_2), where $q_1, q_2 \in [1,\infty]$, and $\frac{1}{p} = \frac{1}{2q_1} + \frac{1}{2q_2}$. Then there exists a constant $C > 0$, depending on p, q_1, and q_2, such that for every smooth function $\varphi : \overline{B(0,1)} \to \mathbb{R}$,*

$$
\int_{B(0,1)} \big|\varphi_x(x)\big|^p dx \le C \left[\int_{B(0,1)} \left[\sum_{\ell=0}^{2} |\partial_x^{(\ell)} \varphi(x)|^{q_1} \right] dx \right]^{\frac{p}{2q_1}} \left[\int_{B(0,1)} |\varphi(x)|^{q_2} dx \right]^{\frac{p}{2q_2}}.
$$

3. Construction of the decoupling function

In this section, we begin to construct the desired decoupling function u. The idea is straightforward: we first mollify the coefficient $b(t, x, y)$ in the variable y by defining

$$b^{\varepsilon}(t, x, y) = \int b(t, x, z)m_{\varepsilon}(y - z)dz, \tag{3.1}$$

where $m_{\varepsilon} \in C^{\infty}$ with compact support, such that $\int m_{\varepsilon}(z)dz = 1$. We note that each b^{ε} is smooth in y, and satisfies (H1) in the variables (t, x), but $b^{\varepsilon}(t, x, y) \to b(t, x, y)$ only pointwisely due to the discontinuity of b in y.

Let us now consider the mollified version of (2.4):

$$\begin{cases} u_t + \frac{1}{2}\sigma^2(t, x, u)u_{xx} + b^{\varepsilon}(t, x, u)u_x + h(t, x, u, \sigma(t, x, u)u_x) = 0, \\ u(T, x) = g(x). \end{cases} \tag{3.2}$$

By Theorem 2.1, for each $\varepsilon > 0$ the PDE (3.2) admits a unique solution $u^{\varepsilon} \in C^{1,2}$, such that the sequence $\{u^{\varepsilon}\}$ is uniformly bounded and equi-continuous in (t, x). Therefore, applying Arzela–Ascoli Theorem we see that u^{ε} converges, uniformly on compacta, to a continuous function $\hat{u} \in C([0, T] \times \mathbb{R})$. We shall argue that this function is exactly the decoupling function that we are looking for.

It is worth noting that because the coefficient b is discontinuous, the stability result of viscosity solution does not apply here, and at this point it is by no means clear that $\hat{u}$ is a "solution" to the limiting equation (2.4) in any sense. The following theorem plays an important role in our study of the uniform regularity of $\{u^{\varepsilon}\}$, whence that of $\hat{u}$.

Theorem 3.1. *Assume (H1)–(H4), and let b^{ε} be defined by (3.1). Then, there exists a constant $r \in (0, 1]$, depending only on Υ, $\underline{\sigma}$, $\overline{\sigma}$, Λ, and T, such that for any $\varepsilon > 0$ and any $p \geq 1$, $R \geq 1$, $\delta \in (0, T]$, and $z \in \mathbb{R}$,*

$$\int_{T-\delta}^{T}\int_{z-R}^{z+R}\Big[(T - s)^{1-r}\big(|u_t^{\varepsilon}(s, y)| + |u_x^{\varepsilon}(s, y)|^2 + |u_{xx}^{\varepsilon}(s, y)|\big)\Big]^p dyds \leq C_{u,p}\delta R, \tag{3.3}$$

where $C_{u,p}$ is a constant depending only on p, Υ, $\underline{\sigma}$, $\overline{\sigma}$, Λ, and T.

Proof. The argument follows closely that of [29], we provide only a sketch of the proof for the completeness. Without loss of generality let us assume $z = 0$. For notational simplicity, in what follows we fix an $\varepsilon > 0$, and denote $u = u^{\varepsilon}$. Consider the following linear PDE:

$$\begin{cases} w_t(t, x) + \frac{1}{2}\sigma^2(t, x, u(t, x))w_{xx}(t, x) + \varphi(t, x) = 0, & (t, x) \in [0, T) \times \mathbb{R} \\ w(T, x) = 0, & x \in \mathbb{R}, \end{cases} \tag{3.4}$$

where $\varphi(t, x) = -h(t, x, u(t, x)) + b(t, x, u(t, x))u_x(t, x)$, for all $(t, x) \in R_T$. By Theorem 2.1 and Assumption (**H**), clearly that the linear PDE (3.4) admits a unique bounded solution $w^0 \in \mathbb{C}^{1,2}(R_T)$ with bounded and uniformly Hölder continuous partial derivatives of order one in t and of order one and two in x. Let $\bar{w} = u - w^0$. It can be checked that $\bar{w}$ is a solution to the linear PDE (2.12). Thus by Theorem 2.2 and Theorem 2.1, there exist $r, C > 0$, depending only on Υ, $\underline{\sigma}$, $\overline{\sigma}$, Λ, and T, such that any $R > 0$, $\delta \in (0, T]$,

$$\int_{T-\delta}^{T}\int_{-R}^{R}\Big[(T - s)^{1-r}\big|\bar{w}_{xx}(s, y)\big|\Big]^p dsdy \leq C\delta R. \tag{3.5}$$

We now denote, for any $\alpha > 0$, and $\varphi \in \mathbb{C}([0,T]\times\mathbb{R})$, $[\![\varphi]\!]^\alpha(t,x) \triangleq (T-t)^{1-\alpha}\varphi(t,x)$. Since w^0 is Hölder continuous on $[0,T]\times\mathbb{R}$, so is $[\![w^0]\!]^r$. Define two operators $\mathcal{L}_t$ and $\mathcal{L}_t^0$,

$$\mathcal{L}_t = \frac{1}{2}\sigma^2(t,x,u(t,x))\partial^2_{xx}, \quad \mathcal{L}_t^0 = \frac{1}{2}\sigma^2(t,0,u(t,0))\partial^2_{xx}. \tag{3.6}$$

Then it is straightforward to check from (3.4) that $[\![w^0]\!]^r$ satisfies the following PDE:

$$[\![w]\!]^r_t(t,x) + \mathcal{L}_t^0[w](t,x) = -[\![\varphi]\!]^r(t,x) + \left[\mathcal{L}_t^0 - \mathcal{L}_t\right][\![w^0]\!]^r(t,x) + (1-r)[\![w^0]\!]^{r+1}(t,x). \tag{3.7}$$

Now applying Theorem 2.2-(iv) to (3.7) on interval $(T-\delta,T)$ (noting the boundedness and uniform Hölder continuity of the right-hand side in (3.7)), we obtain that

$$\begin{aligned}
&(1-\theta)^{2p}R^{2p}\int_{T-\delta}^T\int_{-\theta R}^{\theta R}\left|[\![w^0_{xx}]\!]^r(t,x)\right|^p dxdt \\
&\quad\le C(1-\theta')^{2p}R^{2p}\int_{T-\delta}^T\int_{-\theta' R}^{\theta' R}\left|[\![\varphi]\!]^r(t,x)\right|^p dxdt \\
&\qquad + C(1-\theta')^{2p}R^{2p}\int_{T-\delta}^T\int_{-\theta' R}^{\theta' R}\left[\left|[\![[\sigma^2(\cdot,\cdot,u(\cdot,\cdot)) - \sigma^2(0,\cdot,u(0,\cdot))]w^0_{xx}]\!]^r(t,x)\right|^p\right]dxdt \\
&\qquad + C(1-\theta')^{2p}R^{2p}\int_{T-\delta}^T\int_{-\theta' R}^{\theta' R}\left|[\![w^0]\!]^{r+1}(t,x)\right|^p dxdt + C\int_{T-\delta}^T\int_{-\theta' R}^{\theta' R}\left|[\![w^0]\!]^r(t,x)\right|^p dxdt \\
&\qquad + \frac{1}{2}(1-\theta')^{2p}R^{2p}\int_{T-\delta}^T\int_{-\theta' R}^{\theta' R}\left|[\![w^0_{xx}]\!]^r(t,x)\right|^p dxdt \\
&\quad\le C(1-\theta')^{2p}R^{(2+r)p}\int_{T-\delta}^T\int_{-\theta' R}^{\theta' R}\left|[\![u_{xx}]\!]^r(t,x)\right|^p dxdt \\
&\qquad + (1-\theta')^{2p}R^{2p}(CR^{rp}+1/2)\int_{T-\delta}^T\int_{-\theta' R}^{\theta' R}\left|[\![w^0_{xx}]\!]^r(t,x)\right|^p dxdt + C\delta,
\end{aligned} \tag{3.8}$$

for any $p \ge 1, 0 < R \le 1, \theta \in (0,1)$, and $\theta' = (1+\theta)/2$. In the last inequality above we have used Assumption (**H**), Lemma 2.2, a scaling argument to the triple $(2p,p,\infty)$, and the fact that u and $\sigma^2(\cdot,\cdot,u(\cdot,\cdot))$ are uniformly Hölder continuous. Since $u = \bar{w} + w^0$, it then follows from (3.5) and (3.8) that

$$\begin{aligned}
&(1-\theta)^{2p}R^{2p}\int_{T-\delta}^T\int_{-\theta R}^{\theta R}\left|[\![u_{xx}]\!]^r(t,x)\right|^p dxdt \\
&\quad\le (1-\theta')^{2p}R^{2p}(CR^{rp}+1/2)\int_{T-\delta}^T\int_{-\theta' R}^{\theta' R}\left|[\![u_{xx}]\!]^r(t,x)\right|^p dxdt + C\delta.
\end{aligned}$$

Let R_0 be the number such that $CR_0^{rp} + 1/2 = 3/4$, and let $R_1 = \min(R_0, 1)$. Thus,

$$R_1^{2p}\sup_{\theta\in(0,1)}\left\{(1-\theta)^{2p}\int_{T-\delta}^T\int_{-\theta R_1}^{\theta R_1}\left|[\![u_{xx}]\!]^r(t,x)\right|^p dxdt\right\} \le 4C\delta.$$

Taking $\theta = 1/2$, we have

$$\int_{T-\delta}^T\int_{-R_1/2}^{R_1/2}\left|[\![u_{xx}]\!]^r(t,x)\right|^p dxdt \le 4C\delta R_1^{-2p} \triangleq C\delta.$$

Note that this inequality holds for $(z - R_1/2, z + R_1/2)$, any $z \in \mathbb{R}$. For any $R > 1$ and $z \in \mathbb{R}$, since $(z - R, z + R)$ can be covered by $[\frac{2R}{R_1} + 1]$ intervals of length R_1, we see that

$$\int_{T-\delta}^{T} \int_{z-R}^{z+R} \left| [\![u_{xx}]\!]^r(t,x) \right|^p dxdt \leq C\delta R.$$

The estimate on u_x can be obtained by applying Lemma 2.2 once more. Finally, the estimate on u_t can be obtained because of the uniform boundedness of the coefficients and the estimates on u_x and u_{xx}. □

Theorem 3.2. *In the same setting as Theorem 3.1, there exist constants $r, C > 0$, depending on Υ, $\underline{\sigma}$, $\overline{\sigma}$, Λ, and T, such that for any $\varepsilon > 0$ and any $(t, x) \in [0, T) \times \mathbb{R}$,*

$$|u_x^\varepsilon(t,x)| \leq C(T-t)^{-(1-r)/2}. \tag{3.9}$$

Proof. Again we fix $\varepsilon > 0$, and denote $u = u^\varepsilon$. Let $\eta : \mathbb{R} \to [0, 1]$ be a smooth function such that $\eta(x) = 1$ when $x \in (0, 1)$ and $\eta(x) = 0$ when $x \notin (-2, 2)$. For all (t, x), define $\tilde{u}(t,x) \triangleq u(t,x)\eta(x)$, $\tilde{g}(x) \triangleq g(x)\eta(x)$, and

$$\begin{aligned}\tilde{\varphi}(t,x) \triangleq (\mathcal{L}_t - \mathcal{L}_t^0)[\tilde{u}](t,x) + \eta(x)\varphi(t,x) \\ - \sigma^2(t,x,u(t,x))\eta'(x)u_x(t,x) - u(t,x)\mathcal{L}_t[\eta](x),\end{aligned}$$

where $\varphi(t, x)$ is defined as in the proof of Theorem 3.1. From Assumption (**H**) and Theorem 2.1, we know that for any (t, x),

$$|\tilde{\varphi}(t,x)| \leq C_1\big(1 + |u_x(t,x)|^2 + |u_{xx}(t,x)|\big)\mathbf{1}_{\{|x|\leq 2\}},$$

where C_1 is a constant depending only on $\overline{\sigma}$, Λ, and T. By Theorem 2.2,

$$\begin{aligned}|u_x(t,0)| &= |\tilde{u}_x(t,0)| \\ &\leq C_2\Bigg\{(T-t)^{-1/2}\int_{\mathbb{R}}\Big[\,\big|\tilde{g}\big(\Gamma^{1/2}(t,T)z\big) - \tilde{g}(0)\big||z|\exp(-z^2/2)\Big]dz \\ &\quad + \int_t^T\int_{\mathbb{R}}\Big[(s-t)^{-1/2}\big|\tilde{\varphi}\big(s,\Gamma^{1/2}(t,s)z\big)\big||z|\exp(-z^2/2)\Big]dzds\Bigg\} \triangleq T(1) + T(2),\end{aligned} \tag{3.10}$$

where C_2 is a constant depending on $\underline{\sigma}$ and $\overline{\sigma}$. By the boundedness of coefficients and Lipschitz continuity of g, we know $T(1) \leq C_3$, where C_3 is a constant depending on Υ, $\underline{\sigma}$, $\overline{\sigma}$, Λ, and T. Let r be that in Theorem 3.1, and let $p = 1/r$ in Lemma 2.1. Then by the estimate of $\tilde{\varphi}$ and Theorem 3.1, one gets

$$\begin{aligned}T(2) &\leq C\int_t^T (s-t)^{-(1+r)/2}\left[\int_{-2}^{2} |\tilde{\varphi}(s,z)|^{\frac{1}{r}}dz\right]^r ds \\ &= C\int_t^T (s-t)^{-(1+r)/2}(T-s)^{-(1-r)}\left[\int_{-2}^{2} |[\![\tilde{\varphi}]\!]^r(s,z)|^{\frac{1}{r}}dz\right]^r ds \\ &\leq C\left[\int_t^T (s-t)^{-\frac{1+r}{2(1-r)}}(T-s)^{-1}ds\right]^{1-r}\left[\int_t^T\int_{-2}^{2} |[\![\tilde{\varphi}]\!]^r|^{\frac{1}{r}}dzds\right]^r \\ &\leq C(T-t)^{-\frac{1+r}{2}-(1-r)+1} = C(T-t)^{-\frac{(1-r)}{2}},\end{aligned} \tag{3.11}$$

where C_ε is a constant depending on ε, Υ, $\underline{\sigma}$, $\overline{\sigma}$, Λ, and T, proving the theorem. □

The following theorem shows the regularity of $\{u_x^\varepsilon\}$.

Theorem 3.3. *Assume that the assumptions of Theorem 3.2 are in force. Then there exist constants* $r, C > 0$*, depending on* $\Upsilon, \underline{\sigma}, \overline{\sigma}, \Lambda$*, and* T*, such that for any* $\varepsilon > 0$ *and any* $(t, x),\ (s, y) \in [0, T) \times \mathbb{R}$*,* $t \leq s$*,*

$$\left|u_x^\varepsilon(t, x) - u_x^\varepsilon(s, y)\right| \leq C(T-s)^{-(1-r)/2}\left(|x-y|^r + |t-s|^{r/2}\right).$$

Proof. Let us first fix t and ε, and denote $u = u^\varepsilon$. We shall prove the following inequality

$$|u_x(t, x) - u_x(t, y)| \leq C(T-t)^{-(1-r)/2}|x-y|^r, \tag{3.12}$$

for some $r, C > 0$. In light of Theorem 3.2, without loss of generality, we assume $y = 0$ and $|x| \leq 1$. Let η, $\tilde{u}$, $\tilde{g}$, and $\tilde{\varphi}$ be defined as in Theorem 3.2. Thus by Theorem 2.2, we have

$$\begin{aligned}\tilde{u}_x(t, x) &= \int_{\mathbb{R}} \tilde{g}(z)\psi_x^{(c)}(t, x; T, z)dz + \int_t^T \int_{\mathbb{R}} \tilde{\varphi}(s, z)\psi_x^{(c)}(t, x; s, z)dzds \\ &\triangleq R_1(t, x) + R_2(t, x), \qquad \forall (t, x) \in [0, T) \times \mathbb{R}.\end{aligned} \tag{3.13}$$

Here $\psi^{(c)}$ is the kernel defined by

$$\psi^{(c)}(t, x; s, y) \triangleq (2\pi)^{-1/2}(\Gamma(t, s))^{-1/2}\exp\left(\frac{-(x-y)^2}{2\Gamma(t, s)}\right), \quad \forall\, x, y \in \mathbb{R}, \ \ 0 \leq t < s \leq T.$$

Clearly, $R_1(t, x)$ is differentiable with respect to x. Moreover, since g is bounded and Lipschitz, slightly modifying the proof of Theorem 2.2 and by the definition of $\psi^{(c)}$, one obtains that for some $r, C > 0$, it holds that

$$|\partial_x R_1(t, x)| \leq C(T-t)^{-(1-r)}, \ \text{ and } \ |R_1(t, x)| \leq C(T-t)^{-(1-r)/2}, \quad \forall x \in \mathbb{R}.$$

Hence for every $\varepsilon \in (0, 1]$ and $(t, x) \in [0, T] \times \overline{B(0, 1)}$,

$$\begin{aligned}|R_1(t, x) - R_1(t, 0)| &= |R_1(t, x) - R_1(t, 0)|^{1-\varepsilon}|R_1(t, x) - R_1(t, 0)|^\varepsilon \\ &\leq 2\sup_{y \in \mathbb{R}}|R_1(t, y)|^{1-\varepsilon}\sup_{y \in \mathbb{R}}|\nabla R_1(t, y)|^\varepsilon|x|^\varepsilon \\ &\leq C(T-t)^{-(1-r)(1-\varepsilon)/2-(1-r)\varepsilon}|x|^\varepsilon \leq C(T-t)^{-(1-r/2)/2}|x|^r.\end{aligned} \tag{3.14}$$

The last step in the above inequality is achieved by choosing appropriate ε and r.

To estimate $R_2(t, x)$, let us denote

$$r_2(s, x) \triangleq \int_{\mathbb{R}} \tilde{\varphi}(s, z)\nabla\psi^{(c)}(t, x; s, z)dz$$

so that $R_2(t, x) = \int_t^T r_2(s, x)ds$. By a change of variable and applying Lemma 2.1, one deduce easily that

$$\begin{aligned}|r_2(s, x)| &\leq \frac{C}{\sqrt{\Gamma(t, s)}}\int_{\mathbb{R}} \tilde{\varphi}(s, x + z\Gamma^{1/2}(t, s))|z|\exp\left(-\frac{|z|^2}{2}\right)dz \\ &\leq C(s-t)^{-(1+\varepsilon)/2}\left\{\int_{\mathbb{R}}|\tilde{\varphi}(s, z)|^{1/\varepsilon}dz\right\}^\varepsilon,\end{aligned}$$

for any $0 < \varepsilon < 1$. Here $C > 0$ is a generic constant depending only on $\Upsilon, \underline{\sigma}, \overline{\sigma}, \Lambda$, and T, and is allowed to vary from line to line. Applying the same technique again, one obtains

$$|\partial_x r_2(s, x)| \leq C(s-t)^{-(1+\varepsilon/2)}\left\{\int_{\mathbb{R}}|\tilde{\varphi}(s, z)|^{1/\varepsilon}dz\right\}^\varepsilon.$$

Thus similar to (3.14), we obtain

$$|r_2(s,x) - r_2(s,0)| \leq C(s-t)^{-1/2-\varepsilon} \left\{ \int_{\mathbb{R}} |\tilde{\varphi}(s,z)|^{1/\varepsilon} dz \right\}^{\varepsilon} |x|^{\varepsilon}.$$

Finally, by definition of R_2 we obtain

$$|R_2(t,x) - R_2(t,0)| \leq C|x|^{\varepsilon} \int_t^T (s-t)^{-1/2-\varepsilon} \left\{ \int_{\mathbb{R}} |\tilde{\varphi}(s,z)|^{1/\varepsilon} dz \right\}^{\varepsilon} ds.$$

Similar to (3.11), we can conclude that

$$|R_2(t,x) - R_2(t,0)| \leq C(T-t)^{-(1-r/2)/2} |x|^r, \quad \forall x \in \overline{B(0,1)}. \tag{3.15}$$

Combining (3.14) and (3.15) we derive (3.12).

To conclude the proof we note that, by the similar argument we can show that for fixed x,

$$|u_x(t,x) - u_x(s,x)| \leq C(T-s)^{-(1-r)/2}(s-t)^{r/2}, \qquad \forall 0 \leq t < s < T.$$

This proves the theorem. □

We are now ready to prove the main result of this section. First, note that by (H1), we can rewrite (2.4) in a divergence form:

$$u_t - \partial_x \left\{ \frac{1}{2}\sigma^2(t,x,u)u_x \right\} + \tilde{b}(t,x,u,u_x)u_x + h(t,x,u,\sigma(t,x,u)u_x) = 0, \tag{3.16}$$

where $\tilde{b}(t,x,u,u_x) \stackrel{\triangle}{=} \sigma(t,x,u)\sigma_x(t,x,u) + \sigma(t,x,u)\sigma_y(t,x,u)u_x - b(t,x,u)$. We say that a continuous function u is a *distribution solution* of (2.4) if for any $\varphi \in \mathbb{C}^\infty(\overline{(0,T) \times \mathbb{R}})$ such that $\varphi(t,\cdot) \in C_0^\infty(\mathbb{R})$ for all $t \in (0,T]$, and that for all $0 < s < t < T$ it holds that

$$\begin{aligned} &\int_{\mathbb{R}} u(r,x)\varphi(r,x)dx\Big|_s^t + \int_{(s,t)\times\mathbb{R}} \Big[-u\varphi_t + \frac{1}{2}\sigma^2(\cdot,\cdot,u)u_x\varphi_x \\ &\quad + \tilde{b}(\cdot,\cdot,u,u_x)u_x\varphi + h(r,x,u,\sigma(\cdot,\cdot,u)u_x)\varphi \Big](r,x)dxdr = 0. \end{aligned} \tag{3.17}$$

We have the following result.

Theorem 3.4. *Assume (H1)–(H4). Then the limit function $\hat{u}(t,x)$ is a distributional solution to (2.4).*

Proof. First, we note that Theorem 3.2 and Theorem 3.3 guarantee that the sequence $\{u_x^\varepsilon\}$ is also uniformly bounded and equi-continuous on every compact subset of $[0,T) \times \mathbb{R}$. Thus by Arzela–Ascoli theorem again, we can further extract a subsequence, may assume itself, converging uniformly on compacta in $[0,T) \times \mathbb{R}$ to a limit $\hat{v}$. It is readily seen that $\hat{v} = \hat{u}_x$.

Furthermore, Theorem 3.1 guarantees that there are subsequences, denoted by $\{u_{xx}^{\varepsilon_k}\}$ and $\{u_t^{\varepsilon_k}\}$, converging weakly in $L^p([0,T(1-\delta)] \times B(0,1/\delta))$ for every $0 < \delta < 1$ and $p > 1$. Again, it is clear that these limits are $\hat{u}_{xx}$ and $\hat{u}_t$, respectively, in the distribution sense.

We now argue that $\hat{u}$ is a distribution solution to (3.16). For notational simplicity, we denote $u^k = u^{\varepsilon_k}$, $b^k(t,x,y) = b^{\varepsilon_k}(t,x,y)$, and

$$\sigma^k(t,x) = \sigma(t,x,u^k(t,x)), \qquad h^k(t,x) = h(t,x,u^k(t,x),[\sigma^k u_x^k](t,x)).$$

Note that for each k, u^k is a classical solution to (3.16). For $\varphi \in \mathbb{C}^\infty(\overline{(0,T)\times\mathbb{R}})$, multiplying φ to both sides of (3.16), and integrating by parts, we get (suppressing variables)

$$\int_{\mathbb{R}} u^k(r,x)\varphi(r,x)dx\Big|_s^t + \int_s^t \int_{\mathbb{R}} \Big[-u^k\varphi_t + \frac{1}{2}(\sigma^k)^2 u^k_x\varphi_x + u^k_x\big(\sigma^k\sigma^k_x + \sigma^k\sigma^k_u u^k_x\big)\varphi - u^k_x b^k(t,x,u^k)\varphi + h^k\varphi\Big]dxdr = 0. \tag{3.18}$$

By the uniform convergence of u^k and u^k_x and that σ, σ_x, σ_u, h are all bounded continuous functions, we see that

$$\sigma^k \to \sigma(t,x,\hat{u}), \text{ and } h^k \to h(t,x,\hat{u},\sigma(t,x,\hat{u})\hat{u}_x), \quad \text{as } k\to\infty,$$

uniformly on compacta in $[0,T)\times\mathbb{R}$. Therefore by Dominated Convergence Theorem, all the term in (3.18) converges to the right limit (replacing u^k by $\hat{u}$), except for the term $\int_s^t\int_{\mathbb{R}} u^k_x b^k(t,x,u^k)\varphi dxdr$, due to the discontinuity of b in the variable y. To analyze this term, we define

$$B^k(t,x,y) = \int_0^y b^k(t,x,z)dz, \qquad B(t,x,y) = \int_0^y b(t,x,z)dz.$$

Clearly, B^k and B are both continuous, and B^k converges to B pointwisely. Next, note that b^k, u^k, and b are all bounded, say, by a constant $M > 0$. We have

$$\begin{aligned}\Big|B^k(t,x,u^k) - B(t,x,\hat{u})\Big| &= \Big|\int_0^{u^k} b^k(t,x,y)dy - \int_0^{\hat{u}} b(t,x,y)dy\Big| \\ &\le \int_0^{u^k} |b^k(t,x,y) - b(t,x,y)|dy + \Big|\int_{u^k}^{\hat{u}} b(t,x,y)dy\Big| \\ &\le \int_{-M}^{M} |b^k(t,x,y) - b(t,x,y)|dy + M|\hat{u}(t,x) - u^k(t,x)|. \end{aligned} \tag{3.19}$$

Now by Dominated Convergence theorem, and that $u^k \to \hat{u}$, we see that $B^k(t,x,u^k)$ converges to $B(t,x,\hat{u})$, and the limit is uniform in (t,x) on compacts in $[0,T)\times\mathbb{R}$. Similarly, we can show that, uniformly in t on any compact subset of $[0,T)$,

$$\int_{\mathbb{R}}\int_0^{u^k} b^k_x(t,x,y)\varphi(t,x)dydx \to \int_{\mathbb{R}}\int_0^{\hat{u}} b_x(t,x,y)\varphi(t,x)dydx.$$

Finally, note that

$$\begin{aligned} B^k_x(t,x,u^k) &= \int_0^{u^k} b^k_x(t,x,y)dy + b^k(t,x,u^k)u^k_x, \ B_x(t,x,\hat{u}) \\ &= \int_0^{\hat{u}} b_x(t,x,y)dy + b(t,x,\hat{u})\hat{u}_x. \end{aligned}$$

Thus, integrating by parts and applying Dominated Convergence Theorem again, we obtain that, as $k\to\infty$, for any $0 \le s < t < T$, it holds

$$\int_s^t\int_{\mathbb{R}} u^k_x b^k(r,x,u^k)\varphi(r,x)dxdr = \int_s^t\int_{\mathbb{R}}\Big[B^k_x(r,x,u^k) - \int_0^{u^k} b^k_x(r,x,y)dy\Big]\varphi(r,x)dxdr$$

$$
\begin{aligned}
&= -\int_s^t \int_{\mathbb{R}} B^k(r,x,u^k)\varphi_x(r,x)dxdr \\
&\quad - \int_s^t \int_{\mathbb{R}} \int_0^{u^k} b_x^k(r,x,y)\varphi(r,x)dydxdr \\
&\longrightarrow -\int_s^t \int_{\mathbb{R}} B(r,x,\hat{u})\varphi_x(r,x)dxdr \\
&\quad - \int_s^t \int_{\mathbb{R}} \int_0^{\hat{u}} b_x(r,x,y)\varphi(r,x)dydxdr \\
&= \int_s^t \int_{\mathbb{R}} \left[B_x(r,x,\hat{u}) - \int_s^t \int_0^{\hat{u}} b_x(r,x,y)dy \right] \varphi(r,x)dxdr \\
&= \int_s^t \int_{\mathbb{R}} \hat{u}_x b(r,x,\hat{u})\varphi(r,x)dxdr.
\end{aligned}
$$

This shows that $\hat{u}$ is a distribution solution of (3.16), proving the theorem. □

4. Weak solution of FBSDE

Although we have identified a distribution solution $\hat{u}$ of the decoupling PDE (2.4), we still need to argue that the function $\hat{u}$ can be used as the desired decoupling function. More precisely, we shall argue that it is regular enough for us to apply a certain form of Itô's formula, a crucial step in the Four-Step Scheme.

We begin by recalling the identify (3.16). One can easily check that it can be rewritten as

$$\int_{Q_T} u\varphi_t dxdt = \int_{Q_T} (F\nabla\varphi + g\varphi)dxdt, \tag{4.1}$$

where $F = \frac{1}{2}\sigma^2\hat{u}_x$ and $g = \hat{u}_x\sigma\sigma_x + (\hat{u}_x)^2\sigma\sigma_u - b\hat{u}_x + h)$, both are in $L^2(R_T)$, and $\varphi \in C_0^\infty(Q_T)$. Or in other words, by definition (see, e.g., [30]) $\hat{u} \in W_2^{1,2}((0,T)\times\mathbb{R})$, with

$$\|u\|_{W_2^{1,2}} = \|u_t\|_{L^2} + \inf\left\{ \left(\int_{[0,T]\times\mathbb{R}} (|F|^2 + |g|^2)dxdt \right)^{1/2} \right\}. \tag{4.2}$$

Furthermore, since σ is uniformly non-degenerate and all the coefficients are bounded, it follows from [16] that the non-divergence form PDE (2.4) has a unique $W_2^{1,2}$-solution. Thus it must coincide with $\hat{u}$. It is now clear that the function $\hat{u}$ is a good candidate for the desired decoupling function.

We now consider the following FBSDE on $[t,T]$:

$$\begin{cases} X_s^{t,x} = x + \int_t^s b(r, X_r^{t,x}, Y_r^{t,x})dr + \int_t^s \sigma(r, X_r^{t,x}, Y_r^{t,x})dW_r; \\ Y_s^{t,x} = g(X_T^{t,x}) - \int_s^T h(r, X_r^{t,x}, Y_r^{t,x})dr - \int_s^T Z_r^{t,x}dW_r, \end{cases} \tag{4.3}$$

where b, σ, h, and g satisfy (H1)–(H4). We first show that the FBSDE (4.3) possesses at least a weak solution $(X_s^{t,x}, Y_s^{t,x}, Z_s^{t,x})$ in the sense of Definition 2.2, such that the following "decoupling relation" holds

$$Y_s^{t,x} = u(s, X_s^{t,x}), \qquad Z_s^{t,x} = \sigma(s, X_s^{t,x}, u(s, X_s^{t,x}))u_x(s, X_s^{t,x}). \tag{4.4}$$

To begin with, let us define

$$\bar{b}(t,x) = b(t,x,\hat{u}(t,x)), \quad \bar{\sigma}(t,x) = \sigma(t,x,\hat{u}(t,x)), \tag{4.5}$$

and consider the following forward SDE:

$$X_s^{t,x} = x + \int_t^s \bar{b}(r, X_r^{t,x})dr + \int_t^s \bar{\sigma}(r, X_r^{t,x})dW_r. \tag{4.6}$$

Without loss of generality in what follows we assume $t = 0$. We claim that this SDE possesses a weak solution. Indeed, on any given probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which is defined a standard Brownian motion W, consider the following SDE:

$$X_t = x + \int_0^t \bar{\sigma}(r, X_r)dW_r. \tag{4.7}$$

We note that, by the construction in the previous section, the function $\hat{u} \in W_2^{1,2}(R_T)$ actually has a bounded spatial derivative $\hat{u}_x$. Combining with (H1), it is readily seen that the coefficient $\sigma(t, x)$ is uniformly Lipschitz in x. Thus, the SDE (4.7) admits a unique strong solution, denoted by $X = X^{0,x}$.

Next, define $\theta(t, x) = \frac{\bar{b}(t,x)}{\bar{\sigma}(t,x)}$, which is bounded, thanks to (H1) and (H2). Thus,

$$M_t \triangleq \exp\left\{\int_0^t \theta(s, X_s)dW_s - \frac{1}{2}\int_0^t |\theta(s, X_s)|^2 ds\right\}, \quad t \geq 0, \tag{4.8}$$

is a martingale under $\mathbb{P}$. Now define $\frac{d\mathbb{Q}}{d\mathbb{P}}\Big|_{\mathcal{F}_t} = M_t$, $t \in [0, T]$. Then by the Girsanov theorem, under $\mathbb{Q}$ the process $\bar{W}_t = W_t - \int_0^t \theta(s, X_s^{0,x})ds$ is a Brownian motion, and X satisfies (4.6). In other words, $(\Omega, \mathcal{F}, \mathbb{P}; \bar{W}, X)$ is a weak solution of (4.6).

Let us now define $Y_t = Y_t^{0,x} \triangleq \hat{u}(t, X_t)$. Since $\hat{u}$ is a $W_2^{1,2}$-solution to (2.4), by the Itô–Krylov formula we have

$$\begin{aligned} Y_T - Y_t &= \hat{u}(T, X_T) - \hat{u}(t, X_t) \\ &= \int_t^T \left(\hat{u}_t(r, X_r) + \frac{1}{2}\bar{\sigma}_r^2 \hat{u}_{xx}(r, X_r) + \bar{b}_r \hat{u}_x(r, X_r)\right) dr + \int_t^\tau \bar{\sigma}_r \hat{u}_x(r, X_r) d\bar{W}_r \\ &= -\int_t^T h(r, X_r, \hat{u}(r, X_r), \bar{\sigma}(r, X_r)\hat{u}_x(r, X_r))dr + \int_t^T Z_r d\bar{W}_r, \end{aligned} \tag{4.9}$$

where $Z_t = Z_t^{0,x} \triangleq \sigma(r, X_r, \hat{u}(r, X_r))\hat{u}_x(r, X_r)$. That is, we have shown that the seven-tuple $(\Omega, \mathcal{F}, \mathbb{Q}, \bar{W}, X, Y, Z)$ is a weak solution to the FBSDE (4.3).

We now state our main result of the section.

Theorem 4.1. *Assume (H1)–(H4). There exists a unique weak solution to the FBSDEs (1.1), for any $(t, x) \in [0, T] \times \mathbb{R}$.*

Proof. The argument preceding the theorem proves the existence. Thus we shall only argue the uniqueness. Suppose there is another weak solution $(\hat{X}, \hat{Y}, \hat{Z}) = (\hat{X}^{t,x}, \hat{Y}^{t,x}, \hat{Z}^{t,x})$ of (4.3) on a filtered space $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}}, \hat{W})$. We first claim that $(\hat{X}, \hat{Y}, \hat{Z})$ must satisfy the decoupling relation as well. To see this, denoting

$$\tilde{Y}_t = \hat{u}(t, \hat{X}_t) - \hat{Y}_t, \quad \tilde{Z}_t = \bar{\sigma}(t, \hat{X}_t)\hat{u}_x(t, \hat{X}_t) - \hat{Z}_t$$

and applying the Itô–Krylov formula to $\hat{u}(t, \hat{X}_t^{t,x})$ we get

$$\begin{aligned} d\tilde{Y}_t &= d\left\{\hat{u}(s, \hat{X}_s) - \hat{Y}_s\right\} \\ &= \left\{\hat{u}_t(s, \hat{X}_s) + \frac{1}{2}\hat{u}_{xx}(s, \hat{X}_s)\sigma^2(s, \hat{X}_s, \hat{u}(s, \hat{X}_s)) + \hat{u}_x(s, \hat{X}_s)b(s, \hat{X}_s, \hat{u}(s, \hat{X}_s))\right\}ds \end{aligned}$$

$$
\begin{aligned}
&+ \hat{u}_x(s, \hat{X}_s)\sigma(s, \hat{X}_s, \hat{u}(s, \hat{X}_s))d\hat{W}_s - h(s, \hat{X}_s, \hat{Y}_s, \hat{Z}_s)ds - \hat{Z}_s d\hat{W}_s \\
&= \Big\{h(s, \hat{X}_s, \hat{u}(s, \hat{X}_s), \bar{\sigma}(r, X_r)\hat{u}_x(r, X_r))) - h(s, \hat{X}_s, \hat{Y}_s, \hat{Z}_s)\Big\}ds + \tilde{Z}_s d\hat{W}_s \\
&= [\alpha_s \tilde{Y}_s + \beta_s \tilde{Z}_s]ds + \tilde{Z}_s d\hat{W}_s,
\end{aligned}
$$

where α and β are two bounded processes, thanks to the Lipschitz property of h on (y, z). Now note that $\tilde{Y}_T = 0$, by a standard argument first applying Girsanov theorem and then exponentiating $\tilde{Y}$ we obtain that $\tilde{Y}_t \equiv 0$ for $t \in [0, T]$, and consequently,

$$
u(t, \hat{X}_t) = \hat{Y}_t \text{ and } \hat{Z}_t = u_x(t, \hat{X}_t)\sigma(s, \hat{X}_t, u(t, \hat{X}_t)), \; t \in [0, T], \; \hat{\mathbb{P}}\text{-a.s.} \tag{4.10}
$$

From (4.10) we see that to show $\mathbb{P} \circ (X, Y, Z, W)^{-1} = \hat{\mathbb{P}} \circ (\hat{X}, \hat{Y}, \hat{Z} \hat{W})^{-1}$ it suffices to show that $(\hat{X}, \hat{W})$ and (X, W) are identical in law. But from the existence argument we see that by the Girsanov Theorem using the same kernel $\theta(t, \xi_t) \triangleq \frac{\bar{b}(t,\xi_t)}{\bar{\sigma}(t,\xi_t)}$, where

$$
\bar{b}(s, x) \triangleq b(s, x, u(s, x)), \quad \bar{\sigma}(s, x) \triangleq \sigma(s, x, u(s, x)),
$$

and $\xi = X$ and $\hat{X}$, respectively, we can find two probability measures $\mathbb{Q}$ and $\hat{\mathbb{Q}}$, under which the processes

$$
W_s^1 \triangleq W_s + \int_t^s \theta(r, X_r^{t,x})dr \text{ and } W_s^2 \triangleq \hat{W}_s + \int_t^s \theta(r, \hat{X}_r^{t,x})dr
$$

are Brownian motions, respectively, and (X, W^1) and $(\hat{X}, W^2)$ satisfies the same SDE:

$$
\bar{X}_s = x + \int_t^s \sigma(r, \bar{X}_r, \hat{u}(r, \bar{X}_r))d\bar{W}_r, \qquad s \in [t, T]. \tag{4.11}
$$

But the SDE (4.11) is pathwise unique, whence unique in law, we have that $\mathbb{Q} \circ (X, W^1)^{-1} = \hat{\mathbb{Q}} \circ (\hat{X}, W^2)^{-1}$. It is then readily seen that, using the fact that the Girsanov kernel are the same, $\mathbb{P} \circ (X, W)^{-1} = \hat{\mathbb{P}} \circ (\hat{X}, \hat{W})^{-1}$. This proves the theorem. □

5. Strong solution of the FBSDE

We now turn our attention to the strong solution. Since we have proved the weak existence and uniqueness of the FBSDE (1.1), it is tempting to follow the Yamada–Watanabe Theorem (see, e.g., [1]) to argue only the pathwise uniqueness. However, it should be noted that in the current situation the degree of difficulty for proving the pathwise uniqueness is actually the same as proving the strong existence. We will therefore present a direct method for the existence and uniqueness of the strong solutions without invoking the Yamada–Watanabe type of results. The argument is similar to that of [19] (see also [20]).

We begin by recalling that, for SDE (4.6), the following comparison theorem. The proof is standard, we omit it.

Lemma 5.1. *Assume (H1) and (H2). Let $\bar{b}_1$ and $\bar{b}_2$ be measurable and bounded functions such that*

$$
\bar{b}_1(t, x) \le \bar{b}_2(t, x) \quad \textit{for all } t \textit{ and } x.
$$

Assume one of them is Lipschitz in x, uniformly in s. If X^1 and X^2 are the solutions to (4.6) with coefficients $\bar{b}_1$ and $\bar{b}_2$, respectively, then $X_t^1 \le X_t^2$, for all $t \in [0, T]$, a.s.

Next, we give a well-known estimate of the density of the solution for forward SDEs. Consider the SDE on $[t, T]$:

$$X = x + \int_t^s b(r, X_r)dr + \int_t^s \sigma(r, X_r)dW_r, \quad t \geq 0. \tag{5.1}$$

Assume that the coefficients b and σ are Lipschitz in (s, y), and that there exist $0 < c < C$, such that

$$|b(s, y)| + |\sigma(s, y)| \leq C, \ \forall (s, y) \in [t, T] \times \mathbb{R}, \tag{5.2}$$

$$c \leq \sigma(s, y) \leq C, \ \forall (s, y) \in [t, T] \times \mathbb{R}. \tag{5.3}$$

Clearly, under these assumptions the SDE (5.1) has a unique strong solution, denote it by $X^{t,x}$. Furthermore, denoting $\mu(s, dy; t, x) \triangleq \mathbb{P}\{X_s^{t,x} \in dy\} = \mathbb{P}\{X_s \in dy | X_t = x\}$ to be the transition probability of X_s starting from (t, x) and $p(s, y; t, x)$ to be its density, the following estimate is well-known (see [31] and [32]):

Lemma 5.2. *There exist constants $m, M, \lambda, \Lambda > 0$, so that the density function $p(s, y; t, x)$ satisfies the following estimation:*

$$m(s-t)^{-\frac{1}{2}} \exp\left\{\frac{-\lambda|y-x|^2}{s-t}\right\} \leq p(s, y; t, x) \leq M(s-t)^{-\frac{1}{2}} \exp\left\{\frac{-\Lambda|y-x|^2}{s-t}\right\}. \tag{5.4}$$

We now turn our attention to the case when b is only bounded and measurable, and consider only the weak solution of (5.1), still denote it by $X^{t,x}$. The following Krylov estimate will be crucial in proving the strong well-posedness of the decoupled SDE (4.6).

Lemma 5.3. *Assume that the coefficient $b : [t, T] \times \mathbb{R} \to \mathbb{R}$ is only a bounded measurable function, and let X be (unique) weak solution of (5.1). Then for $\rho > \frac{3}{2}$, there exists a constant $K(T, \rho)$ depending on T, ρ, and the uniform bound of b, such that for any Borel measurable function $f : [t, T] \times \mathbb{R} \to \mathbb{R}$,*

$$\mathbb{E}\int_t^T |f(s, X)|ds \leq K(T, \rho)\left\{\int_t^T \int_{-\infty}^{\infty} |f(s, y)|^\rho dyds\right\}^{\frac{1}{\rho}}. \tag{5.5}$$

Proof. Similar to the arguments in the previous section, we start from the strong solution of the following SDE

$$X_s = x + \int_t^s \bar{\sigma}(r, X_r)d\tilde{W}_r, \quad t \geq 0, \tag{5.6}$$

where $\tilde{W}$ is a Brownian motion defined on a probability space $(\Omega, \mathcal{F}, \tilde{\mathbb{P}}, \mathbb{F})$. Next, denote $\theta(s, x) \triangleq \frac{b(s,y)}{\bar{\sigma}(s,y)}$. Define

$$M_s \triangleq \exp\left\{\int_t^s \theta(r, X_r)d\tilde{W}_r - \frac{1}{2}\int_t^s |\theta(r, X_r)|^2 dr\right\}, \qquad s \geq t,$$

and a new probability measure $d\mathbb{P} \triangleq M_T d\tilde{\mathbb{P}}$. Then under $\mathbb{P}$, $W_s \triangleq \tilde{W}_s - \int_t^s \theta(r, X_r)dr$, $s \geq t$, is a Brownian motion, and $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F}, X, W)$ is a weak solution to SDE (5.1).

Now let $p(s, y; t, x)$ be the transition density of X under $\tilde{\mathbb{P}}$. Then by Lemma 5.2, there exist constants M and Λ, such that

$$0 < p(s, y; t, x) \leq M(s-t)^{-\frac{1}{2}} \exp\left\{\frac{-\Lambda|y-x|^2}{s-t}\right\}.$$

On the other hand, for any measurable function $f : [t, T] \times \mathbb{R} \to \mathbb{R}$, by the Hölder inequality,

$$\begin{aligned}
\mathbb{E}\int_t^T |f(s, X_s)|ds &= \int_\Omega \int_t^T |f(s, X_s)|dsd\mathbb{P} = \int_\Omega \int_t^T |f(s, X_s)|dsM_T d\tilde{\mathbb{P}} \\
&\leq K(T)\left\{\mathbb{E}^{\tilde{\mathbb{P}}}(M_T^\alpha)\right\}^{\frac{1}{\alpha}} \left\{\mathbb{E}^{\tilde{\mathbb{P}}}\left(\left(\int_t^T |f(s, X_s)|^\beta ds\right)\right)\right\}^{\frac{1}{\beta}} \\
&\leq K(T, \alpha)\left\{\int_{-\infty}^\infty \int_t^T |f(s, y)|^\beta p(s, y)dsdy\right\}^{\frac{1}{\beta}} \\
&\leq K(T, \alpha)\left\{\int_{-\infty}^\infty \int_t^T |f(s, y)|^{\beta\gamma} dsdy\right\}^{\frac{1}{\beta\gamma}} \left\{\int_{-\infty}^\infty \int_t^T |p(s, y)|^\delta dsdy\right\}^{\frac{1}{\beta\delta}},
\end{aligned} \tag{5.7}$$

where $\alpha, \beta, \gamma > 1$, such that

$$\frac{1}{\alpha} + \frac{1}{\beta} = 1 \quad \text{and} \quad \frac{1}{\gamma} + \frac{1}{\delta} = 1.$$

Then for $3 > \delta > 1$, we have

$$\begin{aligned}
\left\{\int_{-\infty}^\infty \int_t^T |p(s, y)|^\delta dsdy\right\}^{\frac{1}{\beta\delta}} &\leq \left\{\int_t^T \int_{-\infty}^\infty M^\delta (s-t)^{-\frac{\delta}{2}} \exp\left\{\frac{-\delta\Lambda|y-x|^2}{s-t}\right\} dyds\right\}^{\frac{1}{\beta\delta}} \\
&= K(M, \Lambda, \delta)\left\{\int_t^T (s-t)^{\frac{1-\delta}{2}} ds\right\}^{\frac{1}{\beta\delta}} < \infty.
\end{aligned} \tag{5.8}$$

Now, let $\rho \triangleq \beta\gamma$ and denote $K(T, \rho)$ to be a generic constant depending on $T, \rho, M, \Lambda, \delta$, and the uniform bound of b, which could vary from line to line, then we deduce from (5.7) and (5.8) that

$$\mathbb{E}\int_t^T |f(s, X_s)|ds \leq K(T, \rho)\left\{\int_t^T \int_{-\infty}^\infty |f(s, y)|^\rho dyds\right\}^{\frac{1}{\rho}}. \tag{5.9}$$

Finally, note that $\delta < 3$ implies $\gamma > \frac{3}{2}$. By letting β close to 1 we have $\rho > \frac{3}{2}$, proving the lemma. □

The following lemma shows the stability of the Krylov estimate (5.5).

Lemma 5.4. *Suppose that $\{b^n(s, y)\}_{n=1}^\infty$ are measurable functions and bounded uniformly in n. Suppose that X^n is a solution of (5.1) with drift b^n. Suppose that there exists X such that for every t,*

$$\lim_{n\to\infty} X_t^n = X_t, \qquad P\text{-a.s.}$$

Then the estimate (5.5) holds for X.

Proof. By a Monotone Class argument we need only consider the case where f is a continuous, nonnegative, and bounded function. But since $\{b^n\}_{n=1}^\infty$ are bounded uniformly in n, the constant $K(T, \rho)$ is independent of n. The result then follows from an easy application of Bounded Convergence Theorem. □

The following result, which is a direct consequence of the Krylov estimate (5.5), will be instrumental in the main argument below.

Lemma 5.5. *Assume that the conditions in Lemma 5.4 are in force. In addition, suppose that there exists a measurable function b, such that*

$$\lim_{n\to\infty} b^n(s,y) = b(s,y), \quad \textit{for a.e. } (s,y) \in [t,T]\times\mathbb{R}.$$

Then it holds that

$$\lim_{n\to\infty} \mathbb{E}\int_t^T |b^n(s,X_s^n) - b(s,X_s)|ds = 0.$$

Proof. The argument is more or less standard (see, e.g., [18], [19], or [20]), we provide a sketch for completeness . First, we note that

$$\mathbb{E}\int_t^T |b^n(s,X_s^n) - b(s,X_s)|ds \le \sup_k J_1(n,k) + J_2(n),$$

where

$$J_1(n,k) \triangleq \mathbb{E}\int_t^T |b^k(s,X_s^n) - b^k(s,X_s)|ds; \quad J_2(n) \triangleq \mathbb{E}\int_t^T |b^n(s,X_S) - b(s,X_s)|ds.$$

Let $\kappa:\mathbb{R}\to\mathbb{R}$ be a smooth function such that $\kappa(0)=1$, $0\le\kappa(y)\le 1$, for $y\in(-1,1)$; and $\kappa(y)=0$, otherwise. Then, for any $\varepsilon>0$, there exists $R_0>0$, such that

$$\mathbb{E}\int_t^T |1-\kappa(X/R_0)|ds < \varepsilon. \tag{5.10}$$

Now for fixed R_0, the sequence $\{b^n\}$ converges in $L^2([t,T]\times[-R_0,R_0])$, we can find finitely many elements of $\{b^n\}$, denoting them as $b_1, b_2, \dots, b_N$, such that for every k, there is an $1\le i_k\le N$, such that $\|b^k - b_{i_k}\|_{L^2([t,T]\times[-R_0,R_0])} < \varepsilon$. Now let us write

$$\lim_{n\to\infty}\sup_k J_1(n,k) \le \lim_{n\to\infty}\sup_k I_1(n,k) + \lim_{n\to\infty} I_2(n) + \sup_k I_3(k),$$

where

$$\begin{cases} I_1(n,k) \triangleq \mathbb{E}\int_t^T |b^k(s,X_s^n) - b_{i_k}(s,X_s^n)|ds; \\ I_2(n) \triangleq \sum_{j=1}^N \mathbb{E}\int_t^T |b_j(s,X_s^n) - b_j(s,X_s)|ds; \\ I_3(k) \triangleq \mathbb{E}\int_t^T |b^k(s,X_s) - b_{i_k}(s,X_s)|ds. \end{cases}$$

Applying the Krylov estimate in Lemma 5.3, we see that

$$\begin{aligned} I_1(n,k) &= \mathbb{E}\int_t^T \kappa(X_n/R_0)|b^k(s,X_s^n) - b_{i_k}(s,X_s^n)|ds \\ &\quad + \mathbb{E}\int_t^T (1-\kappa(X_s^n/R_0))|b^k(s,X_s^n) - b_{i_k}(s,X_s^n)|ds \\ &\le K\|b^k - b_{i_k}\|_{L^2([t,T]\times[-R_0,R_0])} + K\mathbb{E}\int_t^T (1-\kappa(X_s^n/R_0))ds \le 2K\varepsilon, \end{aligned}$$

for some constant $K>0$, independent of k and n. Thus, taking the supreme in k and the limit in n, and then sending $\varepsilon\to 0$ we obtain that

$$\lim_{n\to\infty}\sup_k I_1(n,k) = 0.$$

Similarly, by Lemma 5.4, one can also show that $\sup_k I_3(k) \le L\varepsilon$ for some constant L, and $\lim_{n\to\infty} I_2(n) = 0$. Thus, letting $\varepsilon \to 0$ we get

$$\lim_{n\to\infty} \sup_k J_1(n,k) = 0.$$

Finally, it is easy to see that $\lim_{n\to\infty} J_2(n) = 0$, the proof is complete. □

Now we are ready to prove the strong well-posedness of the FBSDE (1.1). We proceed as follows. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a given probability space on which is defined a Brownian motion W. We assume that the filtration $\mathbb{F} = \mathbb{F}^W$ is Brownian. Let $\hat{u} \in W_2^{1,2}([0,T]\times\mathbb{R})$ be the decoupling function, and is the unique solution to the quasilinear PDE (2.4). Recall the functions $\bar{b}$ and $\bar{\sigma}$ defined in (4.5), and let $\{b^j\}_{j=1}^{\infty}$ be the usual mollifiers of $\bar{b}$ so that $b^j(t,x) \to \bar{b}(t,x)$ pointwisely as $j \to \infty$.

Remark 5.1. We should note that the mollifiers $\{b^j(t,x)\}$ defined above are different from the functions $b^\varepsilon(t,x,\hat{u}(t,x))$, where $b^\varepsilon(t,x,\cdot)$ are the mollifiers of $b(t,x,\cdot)$ (in the variable y) defined before. In fact, it is by no means clear that $b^\varepsilon(t,x,\hat{u}(t,x)) \to b(t,x,\hat{u}(t,x))$, a.e. (t,x), as $\varepsilon \to 0$, since the set $\{(t,x) : \hat{u}(t,x) \in \text{jump points of } b(t,x,\cdot)\}$ could have a positive measure.

Denote

$$b_{nk} \triangleq \bigwedge_{j=n}^{k} b^j, \;\; n \le k, \;\; \text{and } B_n \triangleq \bigwedge_{j=n}^{\infty} b^j.$$

Then for a.e. x and any t, it holds that,

$$b_{nk} \searrow B_n \quad \text{as } k \to \infty, \qquad \text{and} \qquad B_n \nearrow \bar{b} \quad \text{as } n \to \infty.$$

Since each b_{nk} is uniformly Lipschitz, the SDE (5.1) with $b = b_{nk}$ has a unique strong solution, denoted by X^{nk}. Now by the standard Comparison Theorem for (forward) SDEs we know that the sequence $\{X^{nk}\}$ is decreasing in k. The boundedness of the coefficients then renders the pathwise uniform boundedness of X^{nk}, and consequently $X_t^n \triangleq \lim_{k\to\infty} X_t^{nk}$ exists, for all $t \in [0,T]$, $\mathbb{P}$-a.s.

Now applying Lemma 5.5, it is easy to see that X^n is a solution of the SDE (5.1) with $b = B_n$. Furthermore, since for any $n \le m \le k$, $b^{nk} \le b^{mk}$, by comparison theorem, we know that $X_{nk} \le X_{mk}$ for every $k > 0$. Thus $X_n \le X_m$, and again the limit $X \triangleq \lim_{n\to\infty} X_n$ exists. Lemma 5.5 thus implies that X solves (5.1). In other words we have proved the existence of the strong solution (5.1).

Our main result of this section is the following theorem.

Theorem 5.1. *Assume (H1)–(H4). Then there exists a unique strong solution to FBSDE (1.1).*

Proof. The arguments preceding the theorem show the existence of the strong solution. We shall prove only the pathwise uniqueness. To this end, let (X,Y,Z) and $(\hat{X},\hat{Y},\hat{Z})$ be two (weak) solutions defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

By the weak uniqueness they must be identical in law, and the decoupling relationship

$$Y_t = \hat{u}(t,X_t),\; Z_t = \bar{\sigma}(t,X_t)\hat{u}_x(t,X_t); \quad \hat{Y}_t = \hat{u}(t,\hat{X}_t),\; \hat{Z}_t = \bar{\sigma}(t,\hat{X}_t)\hat{u}_x(t,\hat{X}_t)$$

hold for all $t \in [0,T]$, $\mathbb{P}$-a.s.

Let us denote $\varphi(t) \triangleq \bar{b}(t,\hat{X}_t)$, and consider the SDE (5.1) with $b = \tilde{b}^{nk} \triangleq b^{nk} \vee \varphi$. Since $|a\vee\varphi - b\vee\varphi| \le |a-b|$, we know that $b^{nk}\vee\varphi$, albeit random, is still uniformly Lipschitz in x, uniform in (t, ω). Thus, the SDE (5.1) with $\tilde{b}^{nk}$ admits a unique strong solution $\tilde{X}^{nk}$. Again, by comparison theorem of the SDEs, we know that $\{\tilde{X}^{nk}\}_{k=1}^{\infty}$ is a decreasing sequence as k increases. We then conclude that $\tilde{X}^n \triangleq \lim_{k\to\infty} \tilde{X}^{nk}$ exists, and that $\tilde{X}^n$ solves the SDE (5.1) with $b = B_n \vee \varphi$.

Now note that, by definition of B_n, we have $B_n(t, \hat{X}_t) \le \varphi(t) = \bar{b}(t, \hat{X}_t)$, $t \in [0, T]$, a.s. Thus $\hat{X}$ itself is also a solution to SDE (5.1) with $b = B_n \vee \varphi$. Since $b^{nk} \vee \varphi \ge B_n \vee \varphi$, applying the comparison theorem again we see that $\hat{X}_t \le \tilde{X}_t^{nk}$, $t \in [0, T]$, a.s., and consequently $\hat{X}_t \le \tilde{X}_t^n$, $t \in [0, T]$, a.s. On the other hand, similar to the proof of Theorem 4.1, we know that weak uniqueness holds for (5.1). Therefore $\tilde{X}_t^n$ and $\hat{X}$ have the same law. Thus, we must have

$$\mathbb{P}\left\{\sup_{0\le t\le T}(\tilde{X}_t^n - \hat{X}_t) = 0\right\} = 1. \tag{5.11}$$

Indeed, if not, then $\mathbb{P}\{\sup_{0\le t\le T}(\tilde{X}_t^n - \hat{X}_t) > 0\} > 0$, and there exists a rational number r and $t > 0$ such that $\mathbb{P}\{\tilde{X}_t^n > r > \hat{X}_t\} > 0$. But since $\{\tilde{X}_t^n > r\} = \{\hat{X}_t > r\} \cup \{\tilde{X}_t^n > r \ge \hat{X}_t\}$, we have

$$\mathbb{P}\{\tilde{X}_t^n > r\} = \mathbb{P}(\hat{X}_t > r\} + \mathbb{P}\{\tilde{X}_t^n > r > \hat{X}_t\} > \mathbb{P}\{\hat{X}_t > r\}.$$

This contradicts with the fact that $\tilde{X}_t^n$ and $\hat{X}_t$ have the same law. But (5.11) clearly implies that $\hat{X}_t = \tilde{X}_t^n$, $t \in [0, T]$, a.s. That is, $\hat{X}_t = \lim_{k\to\infty} \tilde{X}_t^{nk}$, $t \in [0, T]$, $\mathbb{P}$-a.s., for all n.

To conclude, we note that $b^{nk} \le b^{nk} \vee \varphi$. Thus by comparison, we know that $X^{nk} \le \tilde{X}^{nk}$, a.s. Thus $X^n \le \tilde{X}^n = \hat{X}$, a.s., which implies $X \le \hat{X}$. Again by the weak uniqueness, we know that X and $\hat{X}$ have the same distribution. Repeating the arguments before we obtain that $X_t = \hat{X}_t$, $t \in [0, T]$, $\mathbb{P}$-a.s., proving the pathwise uniqueness. □

Funding

This author is supported in part by NSF grants #0806017 and #1106853.

References

[1] Antonelli, F., Ma, J. (2003). Weak solutions of forward-backward SDE's. *Stochastic Analysis and Applications* 21(3):493–514.
[2] Ma, J., Zhang, J., Zheng, Z. (2008). Weak solutions for backward stochastic differential equations, A martingale approach. *The Annals of Probability* 36(6):2092–2125.
[3] Ma, J., Wu, Z., Zhang, D., Zhang, J. (2015). On well-posedness of forward-backward SDEs—A unified approach. *The Annals of Applied Probability* 25(4):2168–2214.
[4] Bali, T. (2003). Modeling the stochastic behavior of short-term interest rates: Pricing implications for discount bonds. *Journal of Banking and Finance* 27(2):201–228.
[5] Driffill, J., Kenc, T., Sola, M. (2003). An empirical examination of term structure models with regime shifts. *EFMA 2003 Helsinki Meetings*.
[6] Bansal, R., Zhou, H. (2002). Term structure of interest rates with regime shifts. *Journal of Finance* 571997–2043.
[7] Dai, Q., Singleton, K., Yang, W. (2007). Regime shifts in a dynamic term structure model of U.S. treasury bond yields. *Review of Financial Studies* 20(5):1669–1706
[8] Evans, M. (2003). Real risk, inflation risk, and the term structure. *The Economic Journal* 113(487):345–389
[9] Landen, C. (2000). Bond pricing in a hidden Markov model of the short rate. *Finance and Stochastics* 4(4):371–389.
[10] Duffie, D., Ma, J., Yong, J. (1995). Black's console rate conjecture. *The Annals of Applied Probability* 5(2):356–382.

[11] Antonelli, F. (1993). Backward-forward stochastic differential equations. *Annals of Applied Probability* 3(3):777–793.
[12] Ma, J., Protter, P., Yong, J. (1994). Solving forward-backward stochastic differential equations explicitly– A four step scheme. *Probability Theory and Related Fields* 98339–359.
[13] Carmona, R., Delarue, F., Espinosa, G.-E., Touzi, N. (2013). Singular forward-backward stochastic differential equations and emissions derivatives. *The Annals of Applied Probability*, 23(3): 1086–1128.
[14] Delarue, F., Guatteri, G. (2006). Weak existence and uniqueness for forward-backward SDEs. *Stochastic Processes and their Applications* 116(12):1712–1742
[15] Ma, J., Yong, J. (1999). *Forward-Backward Stochastic Differential Equations and Their Applications* (*Lecture Notes in Mathematics*). Vol. 1702. New York: Springer Verlag.
[16] Kim, D., Krylov, N. V. (2007). Parabolic equations with measurable coefficients. *Potential Analysis* 26(4):345–361.
[17] Buckdahn, R., Engelbert, H.-J., Rascanu, A. (2004). On weak solutions of backward stochastic differential equations. *Teor. Veroyatn. Primen*. 49(1):70–108; Theory of Probability and its Applications, 2005, 49(1):16–50.
[18] Gyöngy, I., Pardoux, E. (1993). On quasi-linear stochastic partial differential equations. *Probability Theory and Related Fields* 94413–425.
[19] Nualart, D., Ouknine, Y. (2002). Regularization of differential equations by fractional noise. *Stochastic Processes and their Applications* 102103–116.
[20] Bai, L., Ma, J. (2015). Stochastic differential equations driven by fractional Brownian motion and Poisson point processes. *Bernoulli* 21(1):303–334.
[21] Protter, P. (1990). *Stochastic Integration and Differential Equations*. New York: Springer.
[22] Hu, Y., Peng, S. (1995). Solution of forward-backward stochastic differential equations. *Probability Theory and Related Fields* 103273–283.
[23] Peng, S., Wu, Z. (1999). Fully coupled forward-backward stochastic differential equations and applications to optimal control. *SIAM Journal on Control and Optimization* 37(3):825–843.
[24] Delarue, F. (2003). *Estimates of the Solutions of a System of Quasi-Linear PDEs. A Probabilistic Scheme* (*Lecture Notes in Math.*). Vol. 1832. Séminaire de Probabilités XXXVII, Berlin: Springer, pp. 290–332.
[25] Friedman, A. (1964). *Partial Differential Equations of Parabolic Type*. Englewood Cliffs: Prentice-Hall, Inc.
[26] Kobylanski, M. (2000). Backward stochastic differential equations and partial differential equations with quadratic growth. *Annals of Probability* 28(2):558–602.
[27] Krylov, N. V. (1996). *Lectures on Elliptic and Parabolic Equations in Hölder Spaces*. Providence, RI: Graduate Studies in Mathematics. American Mathematical Society.
[28] Stroock, D. W., Varadhan, S. R. S. (1979). *Multidimensional Diffusion Processes*. New York: Springer-Verlag.
[29] Delarue, F. (2002). On the existence and uniqueness of solutions to FBSDEs in a non-degenerate case. *Stochastic Processes and their Applications* 99209–286.
[30] Byun, S., Wang, L. (2005). Parabolic equations in Reifenberg domains. *Archive for Rational Mechanics and Analysis* 176(2):271–301.
[31] Aronson, D. G. (1967). Bounds for the fundamental solution of a parabolic equation. *Bulletin of the American Mathematical Society* 73890–896.
[32] Aronson, D. G. (1968). Non-negative solutions of linear parabolic equations. *Annali Della Scuola Normale Superiore di Pisa* 22607–694.