Available online at www.sciencedirect.com

SCIENCE DIRECT®

Stochastic Processes and their Applications 116 (2006) 1319–1339

stochastic processes and their applications

www.elsevier.com/locate/spa

Regularity and representation of viscosity solutions of partial differential equations via backward stochastic differential equations

Modeste N'Zi[a], Youssef Ouknine[b], Agnès Sulem[c,*]

[a] *Université de Cocody, UFR de Mathématiques et Informatique, BP 582 Abidjan 22, Côte d'Ivoire*
[b] *Dépt. de Math., Faculté des Sciences Semlalia, Université Cadi Ayyad, B.P S15 Marrakech, Maroc*
[c] *INRIA, Domaine de Voluceau, BP 105 Rocquencourt, 78153 Le Chesnay Cedex, France*

Received 26 November 2002; received in revised form 28 February 2006; accepted 1 March 2006
Available online 22 March 2006

Abstract

We study the regularity of the viscosity solution of a quasilinear parabolic partial differential equation with Lipschitz coefficients by using its connection with a forward backward stochastic differential equation (in short FBSDE) and we give a probabilistic representation of the generalized gradient (derivative in the distribution sense) of the viscosity solution. This representation is a kind of nonlinear Feynman–Kac formula. The main idea is to show that the FBSDE admits a unique linearized version interpreted as its distributional derivative with respect to the initial condition. If the diffusion coefficient of the forward equation is uniformly elliptic, we approximate the FBSDE by smooth ones and use Krylov's estimate to prove the convergence of the derivatives. In the degenerate case, we use techniques of Bouleau–Hirsch on absolute continuity of probability measures.

Keywords: Stochastic integrals; Brownian motion; Stochastic differential equations; Distributional derivative; Forward backward stochastic differential equations

1. Introduction

Backward stochastic differential equations (in short BSDEs) have been introduced in the linear case by Bismut in [4,5] when he was studying the adjoint equations associated with the

* Corresponding author. Tel.: +33 1 39 63 55 69; fax: +33 1 39 63 57 86.
E-mail addresses: nzim@ucocody.ci (M. N'Zi), ouknine@ucam.ac.ma (Y. Ouknine), agnes.sulem@inria.fr (A. Sulem).

0304-4149/$ - see front matter
doi:10.1016/j.spa.2006.03.001

stochastic maximum principle in optimal stochastic control. The nonlinear form was initiated by Pardoux–Peng [16,17] and found numerous applications, especially in optimal stochastic control (see, e.g., [10]) and mathematical finance (see [9]). In [3], Barles and Lesigne present the connections between SDEs, BSDEs and PDEs from an analytical point of view and in [2], Bally and Matoussi consider stochastic BSDEs.

The original motivation for the study of BSDEs was to give a probabilistic interpretation of the solutions of parabolic quasilinear partial differential equations (in short PDEs) of the form:

$$\begin{aligned} &\frac{\partial u}{\partial t} + Lu(t,x) + f(t,x,u(t,x),\partial_x u(t,x)\sigma(t,x)) = 0 \quad \text{in } [0,T)\times\mathbb{R}^d \\ &u(T,x) = g(x) \quad \text{in } \mathbb{R}^d \end{aligned} \tag{1}$$

where

$$L = \frac{1}{2}\sum_{i,j=1}^{d}(\sigma\sigma^T)_{i,j}(t,x)\frac{\partial^2}{\partial x_i\partial x_j} + \sum_{i=1}^{d} b_i(t,x)\frac{\partial}{\partial x_i}.$$

If f, g and the coefficients of the second order differential operator L are sufficiently smooth (e.g. of class C^3) in their spatial variables, then the PDE (1) has a classical solution which can be interpreted via the FBSDE[1]: for all $t \le s \le T$

$$\begin{cases} X_s^{t,x} = x + \int_t^s b(r, X_r^{t,x})\mathrm{d}r + \int_t^s \sigma(r, X_r^{t,x})\mathrm{d}W_r \\ Y_s^{t,x} = g(X_T^{t,x}) + \int_s^T f(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x})\mathrm{d}r - \int_s^T Z_r^{t,x}\mathrm{d}W_r. \end{cases} \tag{2}$$

More precisely, it is proved in [17] that

$$u(t,x) = Y_t^{t,x} = \mathbb{E}\left(g(X_T^{t,x}) + \int_t^T f(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x})\mathrm{d}r\right). \tag{3}$$

This formula can be seen as a generalization of the classical Feynman–Kac formula. Moreover, the following explicit representation of the solution of the BSDE in (2) was obtained by Ma–Protter–Yong [13]:

$$Y_s^{t,x} = u(s, X_s^{t,x}) \quad \text{and} \quad Z_s^{t,x} = \partial_x u(s, X_s^{t,x})\sigma(s, X_s^{t,x}), \quad \forall s \in [t,T], \forall x \in \mathbb{R}^d. \tag{4}$$

Recently, the smoothness conditions on the coefficients have been weakened by Ma–Zhang [14]: they proved that (3) and (4) remain true when the coefficients are only C^1 and when the diffusion coefficient of the forward equation is uniformly elliptic. They also obtain two representations of the gradient of the viscosity solution u of the PDE (1). Let $(\nabla X^{t,x}, \nabla Y^{t,x}, \nabla Z^{t,x})$ be the solution of the variational equation of (2):

[1] Observe that Eq. (2) is a special (decoupled) case of a FBSDE which consists of a forward SDE and a Markovian backward SDE.

$$\begin{cases} \text{for all } t \le s \le T \text{ and } i = 1, \dots, d \\ \nabla_i X_s^{t,x} = e_i + \int_t^s \partial_x b(r, X_r^{t,x}) \nabla_i X_r^{t,x} \mathrm{d}r + \sum_{j=1}^d \int_t^s \partial_x \sigma^j(r, X_r^{t,x}) \nabla_i X_r^{t,x} \mathrm{d}W_r^j \\ \nabla_i Y_s^{t,x} = \partial_x g(X_T^{t,x}) \nabla_i X_T^{t,x} + \int_s^T [\partial_x f(r, \Theta^{t,x}(r)) \nabla_i X_r^{t,x} + \partial_y f(r, \Theta^{t,x}(r)) \nabla_i Y_r^{t,x} \\ \qquad + \langle \partial_z f(r, \Theta^{t,x}(r)), \nabla_i Z_r^{t,x} \rangle] \mathrm{d}r - \int_s^T \nabla_i Z_r^{t,x} \mathrm{d}W_r, \end{cases}$$

where $e_i = (0, \dots, \overset{i}{1}, \dots, 0)$ is the i-th coordinate vector of $\mathbb{R}^d$; σ^j is the j-th column of the matrix σ; $\Theta^{t,x}(r)$ denotes $(X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x})$ and

$$\nabla X^{t,x} = (\nabla_1 X^{t,x}, \dots, \nabla_d X^{t,x}), \qquad \nabla Y^{t,x} = (\nabla_1 Y^{t,x}, \dots, \nabla_d Y^{t,x}),$$

$$\nabla Z^{t,x} = \begin{bmatrix} \nabla_1 Z^{t,x} \\ \vdots \\ \nabla_d Z^{t,x} \end{bmatrix}^* .$$

Ma–Zhang [14] proved that for all $(t, x) \in [0, T] \times \mathbb{R}^d$, we have

$$\partial_{x_i} u(t, x) = \mathbb{E} \left\{ \partial_x g(X_T^{t,x}) \nabla_i X_T^{t,x} + \int_t^T [\partial_x f(r, \Theta^{t,x}(r)) \nabla_i X_r^{t,x} \right. \\ \left. + \partial_y f(r, \Theta^{t,x}(r)) \nabla_i Y_r^{t,x} + \langle \partial_z f(r, \Theta^{t,x}(r)), \nabla_i Z_r^{t,x} \rangle] \mathrm{d}r \right\} \tag{5}$$

and

$$\partial_x u(t, x) = \mathbb{E} \left\{ g(X_T^{t,x}) N_T^{t,x} + \int_t^T f(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}) N_r^{t,x} \mathrm{d}r \right\}, \tag{6}$$

where $N_\cdot^{t,x}$ is some process defined on $[t, T]$, depending only on the forward diffusion and the solution of its variational equation. Eq. (6) can be thought of as a new type of a nonlinear Feynman–Kac formula for derivatives of solutions of PDEs. The advantage of the representation (6) comes from the fact that it does not depend on the derivatives of the coefficients of the BSDE.

On the other hand, if f, g are only Lipschitz continuous and the coefficients of the diffusion process are continuously differentiable with bounded derivatives then Ma–Zhang [14] established that

$$Z_s^{t,x} = \mathbb{E} \left\{ g(X_T^{s,x}) N_T^{s,x} + \int_s^T f(r, X_r^x, Y_r^x, Z_r^x) N_r^{s,x} \mathrm{d}r | \mathcal{F}_s^t \right\} \sigma(s, X_s^{t,x}) \tag{7}$$

where $\mathcal{F}_s^t = \sigma\{W_u - W_t : t \le u \le s\}$. This formula leads to path regularity of the process $Z^{t,x}$ (see [10,11]).

The objective of this paper is to extend the above results of Ma–Zhang [14] to the case where the coefficients of the diffusion process are only Lipschitz continuous. In addition to the uniformly elliptic case, we also consider the case where the diffusion coefficient can be degenerate. First, if g, b, and σ are Lipschitz continuous and f is of class C^1, we prove that the analogue of (4) and (5) holds, provided that the classical derivatives are replaced by the generalized one (in the distribution sense). In the nondegenerate case, the proof is essentially based on Krylov estimate for the diffusion process $X^{t,x}$, whereas the degenerate case is treated by using techniques introduced by Bouleau–Hirsch [7,8]. The nondegenerate case has an intrinsic interest and we shall restrict to it for stating a representation theorem. Second, we drop the smoothness condition on the coefficients of the diffusion process and establish (6) with $N^{s,x}$ replaced by a process depending only on the forward diffusion and its variational equation (in the distribution sense).

The superscript ^{t,x} indicates the dependence of the solution on the initial data (t, x), and will be omitted when the context is clear.

The paper is organized as follows. In Section 2, we set the assumptions and recall some results on SDEs. Section 3 deals with the regularity of the viscosity solution of the PDE (1) and its connection with (2). In Section 4, we establish a probabilistic representation for the generalized derivative of u via BSDEs.

2. Assumptions and preliminaries

Let $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ be a filtered, complete probability space satisfying the usual conditions, on which is defined a d-dimensional standard Brownian motion $\{W_t; 0 \le t \le T\}$; $\mathbf{F} \triangleq (\mathcal{F}_t)_{0\le t\le T}$ is the natural filtration generated by W_t augmented with $\mathbb{P}$-null sets. We denote by E a generic Euclidean space (or $E_1, E_2, \ldots$, if different spaces are used simultaneously). Regardless of their dimensions we denote by $\langle\cdot,\cdot\rangle$ and $|\cdot|$ the inner product and norm in all E's, respectively. We put $\partial x = (\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_d})$. Note that if $\psi = (\psi^1, \ldots, \psi^d) : \mathbb{R}^d \longmapsto \mathbb{R}^d$, then $\partial_x \psi \triangleq (\partial_{x_j} \psi^i)_{i,j=1}^d$ is a matrix. Let χ denote a generic Banach space. We consider the following spaces:

- for $t \in [0, T]$, $L^0([t, T]; \chi)$ is the space of all measurable functions $\varphi : [t, T] \longmapsto \chi$;
- for $t \in [0, T]$, $C([t, T]; \chi)$ is the space of all continuous functions $\varphi : [t, T] \longmapsto \chi$. For $p > 0$ we denote $|\varphi|_{t,T}^{*,p} \triangleq \sup_{t\le s\le T} |\varphi(s)|_\chi^p$;
- for integers k and l, $C^{k,l}([0, T] \times E; E_1)$ is the space of all E_1-valued functions $\varphi(t, e)$, $(t, e) \in [0, T] \times E$, which are k times continuously differentiable in t and l times continuously differentiable in e;
- $C_b^{k,l}([0, T] \times E; E_1)$ is the space of functions φ in $C^{k,l}([0, T] \times E; E_1)$ such that all the partial derivatives are uniformly bounded;
- $W^{1,\infty}(E, E_1)$ is the space of all measurable functions $\varphi : E \longmapsto E_1$, such that for some constant $K > 0$ it holds that $|\varphi(x) - \varphi(y)|_{E_1} \le K|x - y|_E$, $\forall x, y \in E$;
- for any sub-σ-field $\mathcal{G} \subseteq \mathcal{F}_T$ and $0 \le p < \infty$, $L^p(\mathcal{G}; E)$ denotes all E-valued, $\mathcal{G}$-measurable random variable ξ such that $\mathbb{E}|\xi|^p < \infty$. Moreover, $\xi \in L^\infty(\mathcal{G}; E)$ means it is $\mathcal{G}$-measurable and bounded;
- for $0 \le p < \infty$, $L^p(\mathbf{F}, [0, T]; \chi)$ is the space of all χ-valued, $\mathbf{F}$-adapted processes ξ satisfying $\mathbb{E}\int_0^T |\xi_t|_\chi^p \mathrm{d}t < \infty$. Moreover $\xi \in L^\infty(\mathbf{F}, [0, T]; \mathbb{R}^d)$ means it is a $\mathbf{F}$-adapted process uniformly bounded in (t, ω);
- $C(\mathbf{F}, [0, T] \times E; E_1)$ is the space of all E_1-valued, continuous random fields $\varphi : \Omega \times [0, T] \times E \longmapsto E_1$, such that for fixed $e \in E$, $\varphi(\cdot, \cdot, e)$ is an $\mathbf{F}$-adapted process.

To simplify the notation, we often denote $C([0, T] \times E; E_1)$ for $C^{0,0}([0, T] \times E; E_1)$. Moreover, if $E_1 = \mathbb{R}$, we suppress E_1 (e.g., $C^{k,l}([0, T] \times E; \mathbb{R}) = C^{k,l}([0, T] \times E)$, $C(\mathbf{F}, [0, T] \times E; \mathbb{R}) = C(\mathbf{F}, [0, T] \times E), \ldots$ etc.). Finally, unless otherwise specified (such as process Z), all vectors are regarded as column vectors.

Throughout this paper we make the following assumptions (except **(A2)** in Section 3.2).

(A1) The functions $\sigma \in C([0,T]\times\mathbb{R}^d;\mathbb{R}^{d\times d})\cap L^0([0,T];W^{1,\infty}(\mathbb{R}^d;\mathbb{R}^{d\times d}))$, $b \in C([0,T]\times\mathbb{R}^d;\mathbb{R}^d)\cap L^0([0,T];W^{1,\infty}(\mathbb{R}^d;\mathbb{R}^d))$ with a common Lipschitz constant $K > 0$ independent of t.

(A2) There exists a constant $c > 0$ such that

$$\xi^*\sigma(t,x)\sigma^*(t,x)\xi \geq c|\xi|^2 \quad \forall (x,\xi)\in\mathbb{R}^d\times\mathbb{R}^d,\ \forall t\in[0,T]$$

where the transpose of any matrix B is denoted by B^*.

(A3) The functions $f \in C([0,T]\times\mathbb{R}^d\times\mathbb{R}\times\mathbf{R}^d)\cap L^0([0,T];W^{1,\infty}(\mathbb{R}^d\times\mathbb{R}\times\mathbb{R}^d))$; and $g\in W^{1,\infty}(\mathbb{R}^d)$. We denote the Lipschitz constants of f and g by a common one $K > 0$ as in **(A1)** and we assume that

$$\sup_{0\leq t\leq T}\{|b(t,0)|+|\sigma(t,0)|+|f(t,0,0,0)|+|g(0)|\}\leq K.$$

The following lemmas are standard or slight variations of well-known results on SDEs and BSDEs (see, e.g. [11,15]).

Lemma 2.1. *Suppose $b \in C(\mathbf{F},[0,T]\times\mathbb{R}^d;\mathbb{R}^d)\cap L^0(\mathbf{F},[0,T];W^{1,\infty}(\mathbb{R}^d;\mathbb{R}^d))$, $\sigma\in C(\mathbf{F},[0,T]\times\mathbb{R}^d;\mathbb{R}^{d\times d})\cap L^0(\mathbf{F},[0,T];W^{1,\infty}(\mathbb{R}^d;\mathbb{R}^{d\times d}))$, with common Lipschitz constant $K > 0$. Let X be the solution of the following SDE:*

$$X_t = x + \int_0^t b(s,X_s)\mathrm{d}s + \int_0^t \sigma(s,X_s)\mathrm{d}W_s.$$

Then for any $p\geq 2$, there exists a constant $C > 0$ depending only on p, T, and K, such that

$$\mathbb{E}|X|_{t,T}^{*,p} \leq C\left(|x|^p + \mathbb{E}\int_0^T [|b(t,0)|^p + |\sigma(t,0)|^p]\mathrm{d}t\right).$$

Lemma 2.2. *Suppose that $f\in C(\mathbf{F},[0,T]\times\mathbb{R}\times\mathbb{R}^d)\cap L^0(\mathbf{F},[0,T];W^{1,\infty}(\mathbb{R}\times\mathbb{R}^d))$ with a uniform Lipschitz constant $K > 0$. For any $\xi\in L^2(\mathcal{F}_T,\mathbb{R})$, let (Y,Z) be the adapted solution of the BSDE*

$$Y_t = \xi + \int_t^T f(s,Y_s,Z_s)\mathrm{d}s - \int_t^T Z_s\mathrm{d}W_s.$$

Then there exists a constant $C > 0$ depending only on T and K such that

$$\mathbb{E}\int_0^T |Z_t|^2\mathrm{d}t \leq C\mathbb{E}\left(|\xi|^2 + \int_0^T |f(t,0,0)|^2\mathrm{d}t\right).$$

Moreover, for all $p\geq 2$, there exists a constant $C_p > 0$ such that

$$\mathbb{E}|Y|_{t,T}^{*,p} \leq C_p\mathbb{E}\left(|\xi|^p + \int_0^T |f(t,0,0)|^p\mathrm{d}t\right).$$

3. Regularity of viscosity solutions of PDEs

Let h be a continuous positive function on $\mathbb{R}^d$ such that

$$\int_{\mathbb{R}^d} h(x)\mathrm{d}x = 1 \qquad \text{and} \qquad \int_{\mathbb{R}^d} |x|^2 h(x)\mathrm{d}x < +\infty.$$

We set $D = \left\{ f \in L^2(h\mathrm{d}x), \text{ such that } \frac{\partial f}{\partial x_j} \in L^2(h\mathrm{d}x) \right\}$, where $\frac{\partial f}{\partial x_j}$ denotes the derivative in the distribution sense. Equipped with the norm

$$\|f\|_D = \left[\int_{\mathbb{R}^d} f^2 h\mathrm{d}x + \sum_{1\le j\le d} \int_{\mathbb{R}^d} \left(\frac{\partial f}{\partial x_j} \right)^2 h\mathrm{d}x \right]^{1/2},$$

D is a Hilbert space, which is a classical Dirichlet space (see [6]). Moreover D is a subset of the Sobolev space $H^1_{\mathrm{loc}}(\mathbb{R}^d)$.

3.1. The nondegenerate case

Let φ be a nonnegative smooth function defined on $\mathbb{R}^d$, with support in the unit ball such that $\int_{\mathbb{R}^d} \varphi(y)\,\mathrm{d}y = 1$. Define the following smooth functions by convolution

$$\begin{aligned} b^n(t,x) &= n^d \int_{\mathbb{R}^d} b(t, x-y)\varphi(ny)\,\mathrm{d}y \\ \sigma^{j,n}(t,x) &= n^d \int_{\mathbb{R}^d} \sigma^j(t, x-y)\varphi(ny)\,\mathrm{d}y \\ g^n(x) &= n^d \int_{\mathbb{R}^d} g(x-y)\varphi(ny)\,\mathrm{d}y. \end{aligned} \tag{8}$$

It is well known that the functions $b^n(t,x)$, $\sigma^{j,n}(t,x)$ and $g^n(x)$ are Borel measurable functions and Lipschitz continuous with constant K in x such that:

$$\left| b^n(t,x) - b(t,x) \right| + \left| \sigma^{j,n}(t,x) - \sigma^j(t,x) \right| + \left| g^n(x) - g(x) \right| \le \frac{C}{n},$$

where $C > 0$ is a constant (independent of t, x and n).

Since b, σ^j and g are Lipschitz continuous functions in the state variable they are differentiable almost everywhere in the sense of Lebesgue measure. Let us denote by b_x, σ^j_x and g_x any Borel measurable functions such that

$$\begin{aligned} &\partial_x b(t,x) = b_x(t,x)\mathrm{d}x \text{ a.e.} \\ &\partial_x \sigma^j(t,x) = \sigma^j_x(t,x)\mathrm{d}x \text{ a.e.} \\ &\partial_x g(x) = g_x(x)\mathrm{d}x \text{ a.e.} \end{aligned}$$

It is clear that the generalized derivatives are bounded by the Lipschitz constant K. The functions $b^n(t,x), \sigma^{j,n}(t,x)$ and $g^n(x)$ are C^∞-functions in x, and for all $t \in [0,T]$, we have

$$\begin{aligned} &\lim_{n\to+\infty} \partial_x b^n(t,x) = b_x(t,x)\mathrm{d}x \text{ a.e.} \\ &\lim_{n\to+\infty} \partial_x \sigma^{j,n}(t,x) = \sigma^j_x(t,x)\mathrm{d}x \text{ a.e.} \\ &\lim_{n\to+\infty} \partial_x g^n(x) = g_x(x)\mathrm{d}x \text{ a.e.} \end{aligned}$$

Let us consider the sequence of FBSDEs

$$\begin{cases} X^{t,x,n}_s = x + \int_t^s b^n(r, X^{t,x,n}_r)\mathrm{d}r + \int_t^s \sigma^n(r, X^{t,x,n}_r)\mathrm{d}W_r \\ Y^{t,x,n}_s = g^n(X^{t,x,n}_T) + \int_s^T f(r, X^{t,x,n}_r, Y^{t,x,n}_r, Z^{t,x,n}_r)\mathrm{d}r - \int_s^T Z^{t,x,n}_r \mathrm{d}W_r. \end{cases} \tag{9}$$

The approximating coefficients $b^n(t,x)$, $\sigma^{j,n}(t,x)$ satisfy the conditions **(A1)**, **(A2)**, moreover they are smooth in x with bounded derivatives. We recall Krylov's estimate for diffusion processes which play a key role in this subsection.

Theorem 3.1 (*Krylov [12]*). *Let $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ be a filtered probability space, $(W_t)_{t\geq 0}$ a d-dimensional Brownian motion, $b : \mathbb{R}_+ \times \Omega \to \mathbb{R}^d$, $\sigma : \mathbb{R}_+ \times \Omega \to \mathbb{R}^d \otimes \mathbb{R}^d$ bounded adapted processes such that:*

$$\exists c > 0, \forall \xi \in \mathbb{R}^d, \forall (t,x) \in [0,T] \times \mathbb{R}^d, \quad \xi^* \sigma(t,\omega)\sigma^*(t,\omega)\xi \geq c\,|\xi|^2 .$$

Let $X_t = x + \int_0^t b(t,\omega)\mathrm{d}t + \int_0^t \sigma(t,\omega)\,\mathrm{d}W_t$ be an Itô process. Then for every Borel function $f : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ with support in $[0,T] \times B(0,M)$, the following inequality holds:

$$\mathbb{E}\left[\int_0^T |f(t,X_t)|\,\mathrm{d}t\right] \leq K \left[\int_0^T \int_{B(0,M)} |f(t,x)|^{d+1}\,\mathrm{d}t\,\mathrm{d}x\right]^{\frac{1}{d+1}}$$

where K is a constant and $B(0,M)$ is the ball of center 0 and radius M.

Now we state some preliminary lemmas which are needed later.

Lemma 3.1. (i) *For all* $0 \leq t \leq T, x \in \mathbb{R}^d$

$$\lim_{n\to\infty} \mathbb{E}\left(|X^{t,x,n} - X^{t,x}|^{*,2}_{t,T} + |Y^{t,x,n} - Y^{t,x}|^{*,2}_{t,T} + \int_0^T |Z_r^{t,x,n} - Z_r^{t,x}|^2 \mathrm{d}r\right) = 0.$$

(ii) *For all* $0 \leq t \leq T$,

$$\lim_{n\to\infty} \mathbb{E}\left[\int_{\mathbb{R}^d} \left(|X^{t,x,n} - X^{t,x}|^{*,2}_{t,T} + |Y^{t,x,n} - Y^{t,x}|^{*,2}_{t,T} + \int_0^T |Z_r^{t,x,n} - Z_r^{t,x}|^2 \mathrm{d}r\right) h(x)\mathrm{d}x\right] = 0.$$

Proof. This lemma follows from Lemmas 2.1 and 2.2 and the Lebesgue Dominated Convergence Theorem. $\square$

For $i = 1, \ldots, d$, let us denote formally $\left\{(\Phi_i^{t,x}(s), \Psi_i^{t,x}(s), \Gamma_i^{t,x}(s)) : t \leq s \leq T\right\}$ the solution of:

$$\begin{cases} \Phi_i^{t,x}(s) = e_i + \displaystyle\int_s^t b_x(r, X_r^{t,x})\,\Phi_i^{t,x}(r)\mathrm{d}r + \sum_{j=1}^d \int_s^t \sigma_x^j(r, X_r^{t,x})\,\Phi_i^{t,x}(r)\mathrm{d}W_r^j \\ \Psi_i^{t,x}(s) = g_x(X_T^{t,x})\,\Phi_i^{t,x}(T) + \displaystyle\int_s^T [\partial_x f(r,\Theta(r))\,\Phi_i^{t,x}(r) + \partial_y f(r,\Theta(r))\,\Psi_i^{t,x}(r) \\ \qquad + \langle \partial_z f(r,\Theta(r)), \Gamma_i^{t,x}(r)\rangle]\mathrm{d}r - \displaystyle\int_s^T \Gamma_i^{t,x}(r)\mathrm{d}W_r \end{cases} \tag{10}$$

where $e_i = (0, \ldots, \overset{i}{1}, \ldots, 0)$ is the i-th coordinate vector of $\mathbb{R}^d$, σ^j is the j-th column of the matrix σ, $\Theta(r) \equiv (X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x})$ and b_x, σ_x, and g_x are generalized derivatives of b, σ and g with respect to x. We denote

$$\Phi = (\Phi_1, \ldots, \Phi_d), \qquad \Psi = (\Psi_1, \ldots, \Psi_d) \qquad \text{and} \qquad \Gamma = \begin{bmatrix} \Gamma_1 \\ \vdots \\ \Gamma_d \end{bmatrix}^* .$$

(Φ, Ψ, Γ) is formally the solution to the first variation of Eq. (2). We prove that this process is well defined.

Since $b^n \in \mathcal{C}_b^{0,\infty}([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$, $\sigma^n \in \mathcal{C}_b^{0,\infty}([0,T] \times \mathbb{R}^d; \mathbb{R}^{d\times d})$ and $g^n \in C_b^{\infty}(\mathbb{R}^d)$, if $f \in C_b^{0,1}([0,T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbf{R}^d)$ then by virtue of Ma–Zhang [14] or Pardoux–Peng [16], there exists a process $\left\{(\Phi^{t,n,x}(s),\ \Psi^{t,n,x}(s),\ \Gamma^{t,n,x}(s));\ t \le s \le T\right\}$ solution of the following FBSDE of first variation of Eq. (9):

$$\begin{cases} \Phi_i^{t,x,n}(s) = e_i + \displaystyle\int_s^t \partial_x b^n(r, X_r^{t,x,n})\,\Phi_i^{t,x,n}(r)\mathrm{d}r \\ \qquad\qquad + \displaystyle\sum_{j=1}^d \int_s^t \partial_x \sigma^{n,j}(r, X_r^{t,x,n})\,\Phi_i^{t,x,n}(r)\mathrm{d}W_r^j \\ \Psi_i^{t,n,x}(s) = \partial_x g^n(X_T^{t,x,n})\,\Phi_i^{t,x,n}(T) + \displaystyle\int_s^T [\partial_x f(r, \Theta^n(r))\,\Phi_i^{t,x,n}(r) \\ \qquad\qquad + \partial_y f(r, \Theta^n(r))\,\Psi_i^{t,x,n}(r) + \langle \partial_z f(r, \Theta^n(r)), \Gamma_i^{t,x,n}(r)\rangle]\mathrm{d}r \\ \qquad\qquad - \displaystyle\int_s^T \Gamma_i^{t,x,n}(r)\mathrm{d}W_r \end{cases} \tag{11}$$

where $\Theta^n(r)$ denotes $(X_r^{t,x,n}, Y_r^{t,x,n}, Z_r^{t,x,n})$. Let

$$\Phi^{t,x,n} = (\Phi_1^{t,x,n}, \ldots, \Phi_d^{t,x,n}), \qquad \Psi^{t,x,n} = (\Psi_1^{t,x,n}, \ldots, \Psi_d^{t,x,n}) \qquad \text{and}$$
$$\Gamma^{t,x,n} = \begin{bmatrix} \Gamma_1^{t,x,n} \\ \Gamma_d^{t,x,n} \end{bmatrix}^*.$$

In the following, we denote by C a positive constant which may vary from line to line.

Lemma 3.2. *Assume (A1)–(A3) and suppose that $f \in C_b^{0,1}([0,T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d)$. Then $\left\{(\Phi^{t,x}(s),\ \Psi^{t,x}(s),\ \Gamma^{t,x}(s));\ t \le s \le T\right\}$ is a well defined process, that is it does not depend on Borel versions of the generalized derivatives of b, σ and g up to $\widetilde{\mathbb{P}}$-almost sure equality.*

Proof. Let b_x^1, b_x^2 be two Borel versions of the derivative of b at x, that is for all $t \in [0,T]$, $b_x^1(t,x) = b_x^2(t,x)\,\mathrm{d}x$ a.e. Let $\sigma_x^{j,1}, \sigma_x^{j,2}$ and g_x^1, g_x^2 defined in a likewise manner. Define $(\Phi^1(s), \Psi^1(s), \Gamma^1(s))$, (resp. $(\Phi^2(s), \Psi^2(s), \Gamma^2(s))$) the solution of Eq. (10) corresponding to $b_x^1, \sigma_x^{j,1}, g_x^1$ (resp. $b_x^2, \sigma_x^{j,2}, g_x^2$).Then by using Gronwall's inequality, we have

$$\begin{aligned} \mathbb{E}\left(\sup_{t\le s\le T}\left|\Phi^1(s) - \Phi^2(s)\right|^2\right) \le\ & C\left\{\mathbb{E}\left[\int_0^T \left|b_x^1\left(s, X_s^{t,x}\right) - b_x^2\left(s, X_s^{t,x}\right)\right|^2 \mathrm{d}s\right]\right. \\ & \left. + \sum_{1\le j\le d} \mathbb{E}\left[\int_0^T \left|\sigma_x^{j,1}\left(s, X_s^{t,x}\right) - \sigma_x^{j,2}\left(s, X_s^{t,x}\right)\right|^2 \mathrm{d}s\right]\right\} \\ =:\ & C\left\{I_1 + I_2\right\}. \end{aligned}$$

For each $p > 0$, $\mathbb{E}\left(\left|X^{t,x}\right|_{t,T}^{*,p}\right) < +\infty$. Thus,

$$\lim_{M\to+\infty} \mathbb{P}\left(\sup_{t\le s\le T}\left|X_s^{t,x}\right| > M\right) = 0. \tag{12}$$

Therefore without loss of generality, we may suppose that $b_x^1, b_x^2, \sigma_x^{j,1}, \sigma_x^{j,2}$ (resp. g_x^1, g_x^2) have compact support $[0, T] \times B(0, M)$ (resp. $B(0, M)$). By applying Krylov's inequality (thanks to condition **(A2)**), we obtain

$$I_1 \le C \left\| b_x^1 - b_x^2 \right\|_{d+1,M} = 0,$$

where for every function $v(t, x)$ with compact support $[0, T] \times B(0, M)$

$$\|v\|_{d+1,M} = \left[\int_0^T \int_{B(0,M)} |v(t, x)|^{d+1} \mathrm{d}t\mathrm{d}x \right]^{\frac{1}{d+1}}.$$

The fact that $I_2 = 0$ can be obtained similarly.

Now, since the coefficients b_x and σ_x^j are bounded, the forward part in Eq. (10) satisfies the Itô conditions. Therefore it has a unique strong solution, which implies that the process $\{\Phi^{t,x}(s); t \le s \le T\}$ is well defined. In view of Lemma 2.2, we have

$$\mathbb{E}\left(\left| \Psi^1 - \Psi^2 \right|_{t,T}^{*,2} + \int_0^T \left| \Gamma^1(r) - \Gamma^2(r) \right|^2 \mathrm{d}r \right)$$
$$\le C\mathbb{E}\left(\left| g_x^1(X_T)\Phi^1(T) - g_x^2(X_T)\Phi^2(T) \right| \right),$$

hence

$$\Psi^1 = \Psi^2 \qquad \text{and} \qquad \Gamma^1 = \Gamma^2.$$

Since the BSDE part of Eq. (10) has an unique solution, we conclude that the processes $\Psi^{t,x}$ and $\Gamma^{t,x}$ are well defined. □

Lemma 3.3. *Assume (A1)–(A3) and suppose that* $f \in C_b^{0,1}([0, T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d)$. *Then,*

(i) *for all* $0 \le t \le T$, $x \in \mathbb{R}^d$

$$\lim_{n\to\infty} \mathbb{E}\left(|\Phi^{t,x,n} - \Phi^{t,x}|_{t,T}^{*,2} + |\Psi^{t,x,n} - \Psi^{t,x}|_{t,T}^{*,2} + \int_0^T |\Gamma_r^{t,x,n} - \Gamma_r^{t,x}|^2 \mathrm{d}r \right) = 0$$

(ii) *for all* $0 \le t \le T$

$$\lim_{n\to\infty} \mathbb{E}\left[\int_{\mathbb{R}^d} \left(|\Phi^{t,x,n} - \Phi^{t,x}|_{t,T}^{*,2} + |\Psi^{t,x,n} - \Psi^{t,x}|_{t,T}^{*,2} + \int_0^T |\Gamma_r^{t,x,n} - \Gamma_r^{t,x}|^2 \mathrm{d}r \right) h(x)\mathrm{d}x \right] = 0.$$

Proof. Applying the Burkholder–Davis–Gundy and Schwartz inequalities and the Gronwall lemma, we obtain for all $n \in \mathbb{N}$, $x \in \mathbb{R}^d$,

$$\mathbb{E}\left[\left| \Phi^{t,x,n} - \Phi^{t,x} \right|_{t,T}^{*,2} \right] \le C\mathbb{E}\left[\left| \Phi^{t,x,n} \right|_{t,T}^{*,4} \right]^{1/2} \times \left\{ \mathbb{E}\left[\int_0^T \left| \partial_x b^n\left(s, X_s^{t,x,n}\right) - b_x\left(s, X_s^{t,x}\right) \right|^4 \mathrm{d}s \right]^{1/2} + \sum_{1\le j\le d} \mathbb{E}\left[\int_0^T \left| \partial_x \sigma^{j,n}\left(s, X_s^{t,x,n}\right) - \sigma_x^j\left(s, X_s^{t,x}\right) \right|^4 \mathrm{d}s \right]^{1/2} \right\}.$$

Since the coefficients in the forward part of the linear FBSDE (11) are bounded by the Lipschitz constant, we have

$$\sup_n \mathbb{E}\left(\left| \Phi^{t,x,n} \right|_{t,T}^{*,4} \right) < +\infty.$$

Set

$$I_1^n := \mathbb{E}\left[\int_0^T \left| \partial_x b^n \left(s, X_s^{t,x,n}\right) - b_x \left(s, X_s^{t,x}\right) \right|^4 \mathrm{d}s \right]$$
$$I_2^{j,n} := \mathbb{E}\left[\int_0^T \left| \partial_x \sigma^{j,n} \left(s, X_s^{t,x,n}\right) - \sigma_x^j \left(s, X_s^{t,x}\right) \right|^4 \mathrm{d}s \right], \quad j = 1, 2, \ldots, d.$$

Let $n_0 \geq 1$ be a fixed integer, then it holds that

$$\begin{aligned} \lim_{n\to+\infty} I_1^n \leq\ & \limsup_{n\to+\infty} C \left\{ \mathbb{E}\left[\int_0^T \left| \partial_x b^n \left(s, X_s^{t,x,n}\right) - \partial_x b^{n_0} \left(s, X_s^{t,x,n}\right) \right|^4 \mathrm{d}s \right] \right. \\ & + \mathbb{E}\left[\int_0^T \left| \partial_x b^{n_0} \left(s, X_s^{t,x,n}\right) - \partial_x b^{n_0} \left(s, X_s^{t,x}\right) \right|^4 \mathrm{d}s \right] \\ & \left. + \mathbb{E}\left[\int_0^T \left| \partial_x b_x^{n_0} \left(s, X_s^{t,x}\right) - b_x \left(s, X_s^{t,x}\right) \right|^4 \mathrm{d}s \right] \right\} \\ =:\ & C \left(J_1^n + J_2^n + J_3^n \right). \end{aligned}$$

As in [12] page 87, let $w(t,x)$ be a continuous function such that $w(t,x) = 0$ if $t^2 + x^2 \geq 1$ and $w(0,0) = 1$. Then for $M > 0$, we have

$$\begin{aligned} \limsup_{n\to+\infty} J_1^n \leq\ & C \left\{ \mathbb{E}\left[\int_0^T \left(1 - w\left(\frac{s}{M}, \frac{X_s^{t,x}}{M} \right) \right) \mathrm{d}s \right] \right. \\ & \left. + \limsup_{n\to+\infty} \mathbb{E}\left[\int_0^T w\left(\frac{s}{M}, \frac{X_s^{t,x}}{M} \right) \left| \partial_x b^n \left(s, X_s^{t,x,n}\right) - \partial_x b^{n_0} \left(s, X_s^{t,x,n}\right) \right|^4 \mathrm{d}s \right] \right\}. \end{aligned}$$

By applying Krylov's inequality, we obtain

$$\begin{aligned} \limsup_{n\to+\infty} J_1^n \leq\ & C \left\{ \mathbb{E}\left[\int_0^T \left(1 - w\left(\frac{s}{M}, \frac{X_s^{t,x}}{M} \right) \right) \mathrm{d}s \right] \right. \\ & \left. + \limsup_{n\to+\infty} \left\| \left| \partial_x b^n - \partial_x b^{n_0} \right|^4 \right\|_{d+1,M} \right\}. \end{aligned}$$

Note that we have used the fact that the diffusion matrix $\sigma^n(t,x)$ satisfies the nondegeneracy condition with the same constant c as $\sigma(t,x)$. Since $\partial_x b^n$ converges to $b_x \mathrm{d}x$-a.e, the last expression in the right hand side of the above inequality tends to 0 as n_0 tends to $+\infty$. Next, let M goes to $+\infty$, then from the properties of the function $w(t,x)$ we conclude that $\limsup_{n\to+\infty} J_1^n = 0$.

Estimating J_3^n by a similar argument, we obtain that $\limsup_{n\to+\infty} J_3^n = 0$.

Finally, we use the continuity of $b_x^{n_0}$ in x and the convergence in probability (uniformly in s) of $X_s^{t,x,n}$ to $X_s^{t,x}$ to deduce that $b_x^{n_0}\left(s, X_s^{t,x,n}\right) \to b_x^{n_0}\left(s, X_s^{t,x}\right)$ in probability as $n \to +\infty$ and to infer by using the Dominated Convergence Theorem that $\limsup_{n\to+\infty} J_2^n = 0$. Hence

$\lim_{n\to+\infty} I_1^n = 0$. One proves similarly that $\lim_{n\to+\infty} I_2^{j,n} = 0$. It follows that

$$\lim_{n\to+\infty} \mathbb{E}\left[\left|\Phi^{t,x,n} - \Phi^{t,x}\right|_{t,T}^{*,2}\right] = 0.$$

Now, since the derivatives of the coefficients in Eqs. (10) and (11) are bounded, we have by Lemma 2.2

$$\mathbb{E}\left(\left|\Psi^{t,x,n} - \Psi^{t,x}\right|_{t,T}^{*,2} + \int_0^T \left|\Gamma^{t,x,n}(r) - \Gamma^{t,x}(r)\right|^2 \mathrm{d}r\right) \le C\mathbb{E}\left(\left|\zeta^n\right|^2 + \int_0^T \left|h^n(r)\right|^2 \mathrm{d}r\right)$$

where

$$\begin{aligned}
\zeta^n &= \partial_x g^n(X_T^{t,x,n})\Phi^{t,x,n}(T) - g_x(X_T^{t,x})\Phi^{t,x}(T),\\
h^n(s) &= \left(\partial_x f(s,\Theta^n(s)) - \partial_x f(s,\Theta(s))\right)\Phi^{t,x,n}(s)\\
&\quad + \left(\partial_y f(s,\Theta^n(s)) - \partial_y f(s,\Theta(s))\right)\Psi^{t,x,n}(s)\\
&\quad + \langle\left(\partial_z f(s,\Theta^n(s)) - \partial_z f(s,\Theta(s))\right), \Gamma^{t,x,n}(s)\rangle.
\end{aligned}$$

By combining Lemma 3.1(i) and the Dominated Convergence Theorem, we obtain

$$\lim_{n\longrightarrow\infty} \mathbb{E}\left(\left|\zeta^n\right|^2 + \int_0^T \left|h^n(r)\right|^2 \mathrm{d}r\right) = 0$$

which completes the proof of part (i). Part (ii) of Lemma 3.3 can be treated similarly. □

Theorem 3.2. *Assume that* ***(A1)–(A3)*** *and suppose that* $f \in C_b^{0,1}([0,T]\times\mathbb{R}^d\times\mathbb{R}\times\mathbb{R}^d)$. *Then,*

(i) *for every* $s \le t \le T$, *the function* $x \longmapsto (X_s^{t,x}, Y_s^{t,x})$ *belongs* $\mathbb{P}$*-almost surely to* $D^d \times D$;
(ii) *For every* $t \le s \le T$, $\mathbb{P}$*-almost surely*

$$\partial_x X_s^{t,x} = \Phi^{t,x}(s), \qquad \partial_x Y_s^{t,x} = \Psi^{t,x}(s)\mathrm{d}x \ a.e.,$$

where the derivatives are taken in the distribution sense.

Proof. By virtue of Lemma 2.2, there exists a constant $C > 0$ such that for all $t \le s \le T$, $x \in \mathbb{R}^d$, $n \in \mathbb{N}$, we have

$$\mathbb{E}\left(|\Psi^{t,n,x}|_{t,T}^{*,2} + \int_t^T \left|\Gamma^{t,n,x}(r)\right|^2 \mathrm{d}r\right) \le C(1 + \mathbb{E}(|\Phi^{t,n,x}(T)|^2)).$$

In view of Lemma 2.1, for all $t \in [0,T]$, $x \in \mathbb{R}^d$, we have

$$\sup_n \mathbb{E}\left(|\Phi^{t,n,x}|^{*,2}\right) \le C(1+|x|^2).$$

It follows that

$$\sup_n \int_{\mathbb{R}^d} \mathbb{E}\left(\left|\Phi^{t,x,n}\right|_{t,T}^{*,2} + |\Psi^{t,n,x}|_{t,T}^{*,2} + \int_t^T \left|\Gamma^{t,n,x}(r)\right|^2 \mathrm{d}r\right) h(x)\mathrm{d}x < \infty.$$

Therefore by using Lemma 3.1(ii) and a result of Bouleau–Hirch [7], we deduce that the function $x \longmapsto (X_s^{t,x}, Y_s^{t,x})$ belongs $\mathbb{P}$-almost surely to $D^d \times D$.

For point (ii), let us note that in view of Theorem 3.1 in Ma–Zhang [14] or Pardoux–Peng [16], we have

$$\partial_x X_s^{t,n,x} = \Phi^{t,n,x}(s), \qquad \partial_x Y_s^{t,n,x} = \Psi^{t,n,x}(s), \qquad \partial_x Z_s^{t,n,x} = \Gamma^{t,n,x}(s).$$

By using again the Bouleau–Hirch result and Lemmas 3.1 and 3.3, we conclude. □

Corollary 3.1. *Assume (A1)–(A3) and suppose that $f \in C_b^{0,1}([0,T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d)$. Let $\left(X^{t,x}, Y^{t,x}, Z^{t,x}\right)$ be the adapted solution of (2) and define $u(t,x) = Y_t^{t,x}$. Then,*

(i) *for every $0 \le t \le T$, the function $x \longmapsto u(t,x)$ belongs to D and for each t and $i = 1, \ldots, d$, the following representation holds:*

$$\partial_{x_i} u(t,\cdot) = \mathbb{E}\left\{\partial_{x_i} g(X_T^{t,\cdot})\Phi_i^{t,\cdot}(T) + \int_t^T [\partial_x f(r, \Theta^{t,\cdot}(r))\Phi_i^{t,\cdot}(r) + \partial_y f(r, \Theta^{t,\cdot}(r))\Psi_i^{t,\cdot}(r) + \langle \partial_x f(r, \Theta^{t,\cdot}(r)), \Gamma_i^{t,\cdot}(r)\rangle]\mathrm{d}r\right\} \mathrm{d}x \ a.e.$$

where $\Theta^{t,x}(r) \equiv \left(X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}\right)$ and $\left(\Phi^{t,x}(r), \Psi^{t,x}(r), \Gamma^{t,x}(r)\right)$ is the solution of the variational equation (10);

(ii) *for every $t \in [0,T]$, we have $Z_s^{t,x} = \partial_x u(t, X_s^{t,x})\sigma(s, X_s^{t,x})\mathrm{d}s \otimes \mathrm{d}x \otimes \mathrm{d}\mathbb{P}$ a.e. where the derivative of u is taken in the distribution sense.*

Proof. Since $u(t,x) = Y_t^{t,x}$, we have $\partial_{x_i} u(t,x) = \Psi_i^{t,x}\mathrm{d}x$ a.e. Taking the expectation in the BSDE part of Eq. (10) and letting $s = t$, we obtain (i).

Now, for every $(t,x) \in [0,T] \times \mathbb{R}^d$:

$$\partial_{x_i} u^n(t,x) = \mathbb{E}\left\{\partial_{x_i} g^n(X_T^{t,x,n})\Phi_i^{t,x,n}(T) + \int_t^T [\partial_x f(r, \Theta^{t,x,n}(r))\Phi_i^{t,x,n}(r) + \partial_y f(r, \Theta^{t,x,n}(r))\Psi_i^{t,x,n}(r) + \langle \partial_x f(r, \Theta^{t,x,n}(r)), \Gamma_i^{t,x,n}(r)\rangle]\mathrm{d}r\right\},$$

where $\left(X^{t,x,n}, Y^{t,x,n}, Z^{t,x,n}\right)$ is the solution of Eq. (9), $\left(\Phi^{t,x,n}(.), \Psi^{t,x,n}(.), \Gamma^{t,x,n}(.)\right)$ is the solution of the corresponding first variation equation (11), $\Theta^{t,x,n}(r) = \left(X_r^{t,x,n}, Y_r^{t,x,n}, Z_r^{t,x,n}\right)$ and $u^n(t,x) = Y_t^{t,x,n}$. By using Lemmas 3.1 and 3.3, we deduce that

$$\begin{aligned}\lim_{n\to\infty} \partial_{x_i} u^n(t,x) &= \lim_{n\to\infty} \mathbb{E}\left\{\partial_{x_i} g^n(X_T^{t,x,n})\Phi_i^{t,x,n}(T) + \int_t^T [\partial_x f(r, \Theta^{t,x,n}(r))\Phi_i^{t,x,n}(r) + \partial_y f(r, \Theta^{t,x,n}(r))\Psi_i^{t,x,n}(r) + \langle \partial_x f(r, \Theta^{t,x,n}(r)), \Gamma_i^{t,x,n}(r)\rangle]\mathrm{d}r\right\} \\ &= \mathbb{E}\left\{\partial_{x_i} g(X_T^{t,x})\Phi_i^{t,x}(T) + \int_t^T [\partial_x f(r, \Theta^{t,x}(r))\Phi_i^{t,x}(r) + \partial_y f(r, \Theta^{t,x}(r))\Psi_i^{t,x}(r) + \langle \partial_x f(r, \Theta^{t,x}(r)), \Gamma_i^{t,x}(r)\rangle]\mathrm{d}r\right\}, \mathrm{d}x \text{ a.e.} \\ &= \partial_{x_i} u(t,x), \mathrm{d}x \text{ a.e.}\end{aligned}$$

It follows that along a subsequence

$$\begin{aligned} Z_s^{t,x} &= \lim_{n\to\infty} Z_s^{t,x,n} \\ &= \lim_{n\to\infty} \partial_x u^n(t, X_s^{t,x,n})\sigma^n\left(s, X_s^{t,x,n}\right) \\ &= \partial_x u(t, X_s^{t,x})\sigma\left(s, X_s^{t,x}\right)\mathrm{d}s \otimes \mathrm{d}x \otimes \mathrm{d}\mathbb{P} \text{ a.e.} \quad \square \end{aligned}$$

3.2. The degenerate case

The method performed in the previous section is intimately linked to the Krylov estimate. In some sense, this inequality says that the law of the random variable X_s is absolutely continuous with respect to Lebesgue measure for $s > 0$. This property was the key fact to define a unique linearized version of the stochastic differential equation (2). That is, if we choose two versions of the generalized derivatives of b, σ and g then the corresponding solutions are equal. In this section we drop the uniform ellipticity condition on the diffusion matrix $\sigma(t,x)\sigma^*(t,x)$. It is clear that the method used earlier will no longer be valid, and the kind of derivative (with respect to the initial condition) defined will have no sense.

The idea is then to define a slightly different stochastic differential equation defined on an enlarged probability space, where the initial condition x will be taken as a random element. This allows us to perform operations outside negligible sets (in x), which are not possible for the initial equation. The method is inspired from a result of Bouleau and Hirsch [7] where the authors have proved an absolute continuity result extending the well known Malliavin calculus method.

Let us recall some preliminaries and notation of the Bouleau–Hirsch method which will be applied in this section to establish the regularity of the viscosity solution of the PDE (1). See [7] for details and proofs.

From now on, we let $\Omega = \mathcal{C}_0(\mathbb{R}_+, \mathbb{R}^d)$ be the space of continuous functions ω such that $\omega(0) = 0$, endowed with the topology of uniform convergence on compact subsets of $\mathbb{R}_+$.

$\mathcal{F}$ is the Borel σ-field over Ω.

$\mathbb{P}$ is the Wiener measure on $(\Omega, \mathcal{F})$.

$(\mathcal{F}_t)_{t\geq 0}$ is the filtration of coordinates augmented with $\mathbb{P}$-null sets of $\mathcal{F}$.

We define the canonical process $W_t(\omega) = \omega(t)$, for all $t \geq 0$.

$\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}, W_t\right)$ is a Brownian motion.

Let $\widetilde{\Omega} = \mathbb{R}^d \times \Omega$, and $\widetilde{\mathcal{F}}$ the Borel σ-field over $\tilde{\Omega}$ and $\widetilde{\mathbb{P}} = h\mathrm{d}x \otimes \mathrm{d}\mathbb{P}$.

Let $\widetilde{W}_t(x,\omega) = W_t(\omega)$ and $\widetilde{\mathcal{F}}_t$ the natural filtration of $\widetilde{W}_t$ augmented with $\widetilde{\mathbb{P}}$-negligible sets of $\widetilde{\mathcal{F}}$. It is clear that $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, (\widetilde{\mathcal{F}}_t)_{t\geq 0}, \widetilde{\mathbb{P}}, \widetilde{W}_t)$ is a Brownian motion starting from 0.

Let $(e_1, \ldots, e_d)$ be the canonical basis of $\mathbb{R}^d$.

We define the Hilbert space $\widetilde{D}_i$ which is a general Dirichlet space by

$$\widetilde{D}_i = \left\{ \begin{array}{l} u : \widetilde{\Omega} \to \mathbb{R}, \exists \widetilde{u} : \widetilde{\Omega} \to \mathbb{R} \text{ Borel measurable such that } u = \widetilde{u}\,, \widetilde{\mathbb{P}}\text{-a.e. and} \\ \forall\,(x,\omega) \in \widetilde{\Omega}, t \to \widetilde{u}(x + te_i, \omega) \text{ is locally absolutely continuous.} \end{array} \right\}$$

$\widetilde{D}_i$ is considered as a set of classes (with respect to the $\widetilde{\mathbb{P}}$-a.e. equality). If u is in $\widetilde{D}_i$ and $\widetilde{u}$ is associated with it according to the above definition, we can write

$$\nabla_i u(x,\omega) = \varliminf_{t\to 0} \frac{\widetilde{u}(x + te_i, \omega) - \widetilde{u}(x,\omega)}{t}.$$

We denote by $\widetilde{D}$ the space

$$\widetilde{D} = \left\{ u \in L^2(\widetilde{\mathbb{P}}) \bigcap \left(\bigcap_{i=1}^{n} \widetilde{D}_i \right); \forall 1 \leq i \leq d, \nabla_i u \in L^2(\widetilde{\mathbb{P}}) \right\}.$$

The space $\widetilde{D}$ equipped with the norm

$$\|u\|_{\widetilde{D}} = \left(\int_{\mathbb{R}^d \times \Omega} u^2 \mathrm{d}\widetilde{\mathbb{P}} + \sum_{i=1}^{d} \int_{\mathbb{R}^d \times \Omega} (\nabla_i u)^2 \mathrm{d}\widetilde{\mathbb{P}} \right)^{1/2}$$

is a Hilbert space which is a general Dirichlet space.

We introduce the process $\left\{\left(\widetilde{X}_s^t, \widetilde{Y}_s^t, \widetilde{Z}_s^t\right); t \leq s \leq T\right\}$ defined on the enlarged space $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, (\widetilde{\mathcal{F}}_t)_{t\geq 0}, \widetilde{\mathbb{P}}, \widetilde{W}_t)$, solution of the forward backward stochastic differential equation

$$\begin{cases} \widetilde{X}_s^t = x + \int_t^s b(r, \widetilde{X}_r^t)\mathrm{d}s + \int_t^s \sigma\left(r, \widetilde{X}_r^t\right)\mathrm{d}\widetilde{W}_s \\ \widetilde{Y}_s^t = g(\widetilde{X}_T^t) + \int_s^T f(r, \widetilde{X}_r^t, \widetilde{Y}_r^t, \widetilde{Z}_r^t) - \int_s^T \widetilde{Z}_r^t \mathrm{d}\widetilde{W}_r. \end{cases} \tag{13}$$

Since the coefficients are Lipschitz continuous and grow at most linearly, Eq. (13) has a unique $\widetilde{\mathcal{F}}_t$-adapted solution with continuous trajectories. Eqs. (2) and (13) are almost the same except that uniqueness for (13) is slightly weaker. One can easily prove that the uniqueness implies that for each $t \leq s \leq T$, $\left(\widetilde{X}_s^t, \widetilde{Y}_s^t, \widetilde{Z}_s^t\right) = \left(X_s^{t,\cdot}, Y_s^{t,\cdot}, Z_s^{t,\cdot}\right)$, $\widetilde{\mathbb{P}}$-a.s.

Theorem 3.3 (*Bouleau–Hirsch [7,8]*). *For $\mathbb{P}$-almost every ω*

(i) *For all $t \leq s \leq T \geq 0$, $X_s^{t,\cdot}(\omega) \in D^d \subset \left(H_{\mathrm{loc}}^1(\mathbb{R}^d)\right)^d$*

(ii) *There exists a $\widetilde{\mathcal{F}}_t$-adapted $GL_d(\mathbb{R})$-valued continuous process $(\widetilde{\Phi}^t)$ such that for $\widetilde{\mathbb{P}}$-almost every ω:*

$$\forall t \leq s \leq T, \quad \frac{\partial}{\partial x}\left(X_s^x(\omega)\right) = \widetilde{\Phi}_s^t(x, \omega)\mathrm{d}x \ a.e.$$

where $\frac{\partial}{\partial x}$ denotes the derivative in the distribution sense.

Remark 3.1. It is proved in [7] that the image measure of $\widetilde{\mathbb{P}}$ by the map $\widetilde{X}_s^t$ is absolutely continuous with respect to the Lebesgue measure.

Lemma 3.4. *The distributional derivative $\widetilde{\Phi}^t$ is the unique solution of the linear stochastic differential equation*

$$\widetilde{\Phi}_i^t(s) = e_i + \int_s^t b_x(r, \widetilde{X}_r^t)\widetilde{\Phi}_i^t(r)\mathrm{d}r + \sum_{j=1}^d \int_s^t \sigma_x^j(r, \widetilde{X}_r^t)\widetilde{\Phi}_i^t(r)\mathrm{d}\widetilde{W}_r^j \tag{14}$$

where b_x and σ_x^j are versions of the almost everywhere derivatives of b and σ^j.

Proof. First of all, we observe that since the law of $\widetilde{X}_s^t$ is absolutely continuous with respect to the Lebesgue measure, $\widetilde{\Phi}^t$ is well defined and does not depend on the possible choices of the Borel derivatives b_x, σ_x^j. Moreover the coefficients $b_x(s, \widetilde{X}_s^t)$ and $\sigma_x^j(s, \widetilde{X}_s^t)$ are bounded, therefore Eq. (14) satisfies the classical Itô conditions and has a unique $\widetilde{\mathcal{F}}_t$-adapted continuous solution.

The fact that $\widetilde{\Phi}^t$ satisfies Eq. (14) is based on the absolute continuity of the law of $\widetilde{X}_s^t$ and on approximations of the coefficients b and σ by smooth ones (see [7] for details). $\square$

Let us consider formally $\left\{(\widetilde{\Phi}^t(s), \widetilde{\Psi}^t(s), \widetilde{\Gamma}^t(s)), t \le s \le T\right\}$ the solution of the FBSDE of first variation associated to $\left\{(\widetilde{X}^t_s, \widetilde{Y}^t_s, \widetilde{Z}^t_s), t \le s \le T\right\}$:

$$\begin{cases} \widetilde{\Phi}^t_i(s) = e_i + \displaystyle\int_s^t b_x(r, \widetilde{X}^t_r)\widetilde{\Phi}^t_i(r)\mathrm{d}r + \sum_{j=1}^d \int_s^t \sigma^j_x(r, \widetilde{X}^t_r)\widetilde{\Phi}^t_i(r)\mathrm{d}\widetilde{W}^j_r \\ \widetilde{\Psi}^t_i(r) = g_x(\widetilde{X}^t_T)\widetilde{\Phi}^t_i(T) + \displaystyle\int_s^T [\partial_x f(r, \widetilde{\Theta}(r))\widetilde{\Phi}^t_i(r) + \partial_y f(r, \widetilde{\Theta}(r))\widetilde{\Psi}^t_i(r) \\ \qquad + \langle \partial_z f(r, \widetilde{\Theta}(r)), \widetilde{\Gamma}^t_i(r)\rangle]\mathrm{d}r - \displaystyle\int_s^T \widetilde{\Gamma}^t_i(r)\mathrm{d}\widetilde{W}_r. \end{cases} \tag{15}$$

Lemma 3.5. *Assume (A1), (A3) and suppose that* $f \in C_b^{0,1}([0,T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d)$. *Then* $\left\{(\widetilde{\Phi}^t(s), \widetilde{\Psi}^t(s), \widetilde{\Gamma}^t(s)); t \le s \le T\right\}$ *is a well defined process, that is, it does not depend on Borel versions of the generalized derivatives of* b, σ, g *up to* $\widetilde{\mathbb{P}}$*-almost sure equality.*

Proof. Let b^1_x, b^2_x be two Borel versions of the derivative of b at x, that is for each $t \in [0, T]$, $b^1_x(t, \cdot) = b^2_x(t, \cdot)$ dx-a.e. Let $\sigma^{j,1}_x, \sigma^{j,2}_x$ and g^1_x, g^2_x be defined in a likewise manner.

Define $(\widetilde{\Phi}^1(s), \widetilde{\Psi}^1(s), \widetilde{\Gamma}^1(s))$, (resp. $(\widetilde{\Phi}^2(s), \widetilde{\Psi}^2(s), \widetilde{\Gamma}^2(s))$) the solution of (15) corresponding to $b^1_x, \sigma^{j,1}_x, g^1_x$ (resp. $b^2_x, \sigma^{j,2}_x, g^2_x$). By virtue of Lemma 3.4, $\widetilde{\Phi}^1 = \widetilde{\Phi}^2 \widetilde{\mathbb{P}}$-a.e. In view of Lemma 2.2, we have

$$\mathbb{E}\left(\left|\widetilde{\Psi}^2 - \widetilde{\Psi}^1\right|^{*,2}_{t,T} + \int_0^T \left|\widetilde{\Gamma}^2(r) - \widetilde{\Gamma}^1(r)\right|^2 \mathrm{d}r\right) \\ \le C\mathbb{E}\left(\left|g^2_x(\widetilde{X}_T)\widetilde{\Phi}^2(T) - g^1_x(\widetilde{X}_T)\widetilde{\Phi}^1(T)\right|^2\right).$$

Using the absolute continuity of the law of the $\widetilde{X}_s$ and the fact that $\widetilde{\Phi}^1 = \widetilde{\Phi}^2\widetilde{\mathbb{P}}$-a.e., it is easy to see that the right hand side of the above inequality is null. It follows that

$$\left(\widetilde{\Phi}^1, \widetilde{\Psi}^1, \widetilde{\Gamma}^1\right) = \left(\widetilde{\Phi}^2, \widetilde{\Psi}^2, \widetilde{\Gamma}^2\right)\widetilde{\mathbb{P}}\text{-a.e.} \quad \square$$

Let b^n, σ^n, g^n be the regularized functions of b, σ, g as in (8). Let us define for $n \in \mathbb{N}$,

$$\widetilde{X}^{t,n}_s = x + \int_t^s b^n\left(r, \widetilde{X}^t_r\right)\mathrm{d}r + \int_t^s \sigma^n\left(r, \widetilde{X}^t_r\right)\mathrm{d}\widetilde{W}_r,$$

$$\widetilde{\Phi}^{t,n}_i(s) = e_i + \int_t^s \partial_x b^n\left(r, \widetilde{X}^t_r\right)\widetilde{\Phi}^t_i(r)\mathrm{d}r + \sum_{j=1}^d \int_t^s \partial_x \sigma^{n,j}\left(r, \widetilde{X}^t_r\right)\widetilde{\Phi}^t_i(r)\mathrm{d}\widetilde{W}^j_r \tag{16}$$

and consider the sequence of BSDEs

$$\widetilde{Y}^{t,n}_s = g^n(\widetilde{X}^t_T) + \int_s^T f(r, \widetilde{X}^t_r, \widetilde{Y}^{t,n}_r, \widetilde{Z}^{t,n}_r) - \int_s^T \widetilde{Z}^{t,n}_r\mathrm{d}\widetilde{W}_r,$$

$$\widetilde{\Psi}^{t,n}_i(s) = \partial_x g^n(\widetilde{X}^t_T)\widetilde{\Phi}^t_i(T) + \int_s^T [\partial_x f(r, \widetilde{\Theta}^n(r))\widetilde{\Phi}^t_i(r) + \partial_y f(r, \widetilde{\Theta}^n(r))\widetilde{\Psi}^{t,n}_i(r) \\ + \langle \partial_z f(r, \widetilde{\Theta}^n(r)), \widetilde{\Gamma}^{t,n}_i(r)\rangle]\mathrm{d}r - \int_s^T \widetilde{\Gamma}^{t,n}_i(r)\mathrm{d}\widetilde{W}_r$$

where $\widetilde{\Theta}^n \equiv \left(\widetilde{X}_r^t, \widetilde{Y}_r^{t,n}, \widetilde{Z}_r^{t,n}\right)$. Since the coefficient b^n, σ^n, g^n are C^∞-functions in the spatial variable and $f \in C_b^{0,1}([0,T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d)$, for all $t \le s \le T$, $n \in \mathbb{N}$, we have

$$\left(\widetilde{X}_s^{t,n}, \widetilde{Y}_s^{t,n}, \widetilde{Z}_s^{t,n}\right) \in \widetilde{D}^d \times \widetilde{D} \times \widetilde{D}^{d\times d},$$

with

$$\nabla_i \widetilde{X}_s^{t,n} = \widetilde{\Phi}_i^{t,n}(s), \qquad \nabla_i \widetilde{Y}_s^{t,n} = \widetilde{\Psi}_i^{t,n}(s), \qquad \text{and} \qquad \nabla_i \widetilde{Z}_s^{t,n} = \widetilde{\Gamma}_i^{t,n}(s). \tag{17}$$

Lemma 3.6. *Assume (A1), (A3) and suppose that* $f \in C_b^{0,1}([0,T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d)$. *Then, for all* $0 \le t \le T$

$$\lim_{n\to\infty} \widetilde{\mathbb{E}}\left(|\widetilde{X}^{t,n} - \widetilde{X}^t|_{t,T}^{*,2} + |\widetilde{Y}^{t,n} - \widetilde{Y}^t|_{t,T}^{*,2} + \int_0^T |\widetilde{Z}_s^{t,n} - \widetilde{Z}_s^t|^2 \mathrm{d}s\right) = 0.$$

Proof. This lemma is proved by combining Lemmas 2.1 and 2.2, and the Dominated Convergence Theorem. □

Lemma 3.7. *Assume (A1), (A3) and suppose that* $f \in C_b^{0,1}([0,T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d)$. *Then, for all* $0 \le t \le T$,

$$\lim_{n\to\infty} \widetilde{\mathbb{E}}\left(|\widetilde{\Phi}^{t,n} - \widetilde{\Phi}^t|_{t,T}^{*,2} + |\widetilde{\Psi}^{t,n} - \widetilde{\Psi}^t|_{t,T}^{*,2} + \int_0^T |\widetilde{\Gamma}_s^{t,n} - \widetilde{\Gamma}_s^t|^2 \mathrm{d}s\right) = 0.$$

Proof. First, let us prove that

$$\lim_{n\to\infty} \widetilde{\mathbb{E}}\left[\left|\widetilde{\Phi}^{t,n}(s) - \widetilde{\Phi}^t(s)\right|_{t,T}^{*,2}\right] = 0. \tag{18}$$

In view of the Burkholder–Gundy and Schwartz inequalities and the Gronwall lemma, we have

$$\begin{aligned}
\widetilde{\mathbb{E}}\left[\left|\widetilde{\Phi}^{t,n}(s) - \widetilde{\Phi}^t(s)\right|_{t,T}^{*,2}\right] \le{}& M\widetilde{\mathbb{E}}\left[\left|\widetilde{\Phi}^t\right|_{t,T}^{*,4}\right]^{1/2} \\
&\times \left\{\widetilde{\mathbb{E}}\left[\int_0^T \left|\partial_x b^n\left(s, \widetilde{X}_s^t\right) - b_x\left(s, \widetilde{X}_s^t\right)\right|^4 \mathrm{d}t\right]^{1/2}\right. \\
&\left. + \sum_{1\le j\le d} \widetilde{\mathbb{E}}\left[\int_0^T \left|\partial_x \sigma^{j,n}\left(s, \widetilde{X}_s^t\right) - \sigma_x^j\left(s, \widetilde{X}_s^t\right)\right|^4 \mathrm{d}t\right]^{1/2}\right\}.
\end{aligned}$$

Since the coefficients in the linear stochastic differential equation (16) are bounded, we have

$$\sup_n \widetilde{\mathbb{E}}\left[\left|\widetilde{\Phi}^{t,n}\right|_{t,T}^{*,4}\right] < +\infty.$$

To derive (18), it is sufficient to prove the following:

$$\widetilde{\mathbb{E}}\left[\int_0^T \left|\partial_x b^n\left(s, \widetilde{X}_s^t\right) - b_x\left(s, \widetilde{X}_s^t\right)\right|^4 \mathrm{d}s\right] \to 0 \quad \text{as } n \to +\infty$$

and

$$\widetilde{\mathbb{E}}\left[\int_0^T \left|\partial_x \sigma^{j,n}\left(s, \widetilde{X}_s^{t,n}\right) - \sigma_x^j\left(s, \widetilde{X}_s^t\right)\right|^4 \mathrm{d}s\right] \to 0 \quad \text{as } n \to +\infty,\ j = 1, 2, \ldots, d.$$

Let us prove the first limit. Since the law of $\widetilde{X}_s^t$ is absolutely continuous with respect to the Lebesgue measure, let $\widetilde{p}^t(s,y)$ its density. Then

$$\widetilde{\mathbb{E}}\left[\int_0^T \left|b_x^n\left(s,\widetilde{X}_s^t\right)-b_x\left(s,\widetilde{X}_s^t\right)\right|^4\mathrm{d}t\right]=\int_0^T\int_{\mathbb{R}^d}\left|b_x^n(s,y)-b_x(s,y)\right|^4\widetilde{p}^t(s,y)\mathrm{d}y\mathrm{d}s.$$

Since $\partial_x b^n$, b_x are bounded by the Lipschitz constant and $\partial_x b^n$ converges to b_x, we conclude by the Dominated Convergence Theorem. The case of the second limit can be treated by the same technique.

Now, since the derivatives of the coefficients are bounded, we have, by Lemma 2.2

$$\widetilde{\mathbb{E}}\left(|\widetilde{\Psi}^{t,n}-\widetilde{\Psi}^t|_{t,T}^{*,2}+\int_0^T|\widetilde{\Gamma}^{t,n}(r)-\widetilde{\Gamma}(r)|^2\mathrm{d}r\right)\le C\mathbb{E}\left(|\widetilde{\zeta}^n|^2+\int_0^T|\widetilde{h}^n(r)|^2\mathrm{d}r\right)$$

where

$$\begin{aligned}
\widetilde{\zeta}^n &= \partial_x g^n(\widetilde{X}_T^t)\widetilde{\Phi}^t(T)-g_x(\widetilde{X}_T^t)\widetilde{\Phi}^t(T)\\
\widetilde{h}^n(s) &= \left(\partial_x f(s,\widetilde{\Theta}^n(s))-\partial_x f(s,\widetilde{\Theta}(s))\right)\widetilde{\Phi}^t(s)\\
&\quad+\left(\partial_y f(s,\widetilde{\Theta}^n(s))-\partial_y f(s,\widetilde{\Theta}(s))\right)\widetilde{\Psi}^t(s),\\
&\quad+\langle\left(\partial_z f(s,\widetilde{\Theta}^n(s))-\partial_z f(s,\widetilde{\Theta}(s))\right),\widetilde{\Gamma}^t(s)\rangle.
\end{aligned}$$

We have

$$\mathbb{E}|\widetilde{\zeta}^n|^2\le\left(\mathbb{E}|\widetilde{\Phi}^t(T)|^2\right)^{1/2}\left(\mathbb{E}|\partial_x g^n(\widetilde{X}_T^t)-g_x(\widetilde{X}_T^t)|^2\right)^{1/2}.$$

By using that $\partial_x g^n$, g_x are bounded by the Lipschitz constant, the convergence of $\partial_x g^n$ to g_x, the absolute continuity of the law of $\widetilde{X}_T^t$ with respect to the Lebesgue measure and the Dominated Convergence Theorem, we obtain $\lim_{n\to 0}\mathbb{E}|\widetilde{\zeta}^n|^2=0$. Combining Lemma 3.6 and the Dominated Convergence Theorem, one can prove that

$$\lim_{n\longrightarrow\infty}\mathbb{E}\int_0^T|\widetilde{h}^n(r)|^2\mathrm{d}r=0.\quad\square$$

Theorem 3.4. *Assume **(A1)**, **(A3)** and suppose that* $f\in C_b^{0,1}([0,T]\times\mathbb{R}^d\times\mathbb{R}\times\mathbb{R}^d)$. *Then, for* $\mathbb{P}$*-almost every* ω*:*

(i) *for every* $s\le t\le T$*, the function* $x\longmapsto(X_s^{t,x}(\omega),Y_s^{t,x}(\omega))$ *belongs to* $D^d\times D$ $\mathbb{P}$*-almost surely*
(ii) *for every* $t\le s\le T$*,* $\mathbb{P}$*-almost surely* $\partial_x X_s^{t,x}(\omega)=\widetilde{\Phi}_s^t(x,\omega)$, $\partial_x Y_s^{t,x}(\omega)=\widetilde{\Psi}_s^t(x,\omega)\mathrm{d}x$ *a.e.*

Proof. By Lemma 2.2, there exists a constant $C>0$ such that for all t,n,s, it holds

$$\widetilde{\mathbb{E}}\left(|\widetilde{\Psi}^{t,n}|_{t,T}^{*,2}+\int_t^T|\widetilde{\Gamma}^{t,n}(r)|^2\mathrm{d}r\right)\le C(1+\widetilde{\mathbb{E}})\left(|\widetilde{\Phi}^{t,n}(T)|^2\right).$$

By Lemma 2.1, we have for all t,

$$\sup_n\widetilde{\mathbb{E}}\left(|\widetilde{\Phi}^{t,n}(T)|^2\right)\le C.$$

It follows that

$$\sup_n\widetilde{\mathbb{E}}\left(|\widetilde{\Psi}^{t,n}|_{t,T}^{*,2}+\int_t^T|\widetilde{\Gamma}^{t,n}(r)|^2\mathrm{d}r\right)<\infty.$$

Using (17), Lemmas 3.2 and 3.3 and the Bouleau–Hirch result, we deduce that the function $(\widetilde{X}_s^t, \widetilde{Y}_s^t)$belongs $\widetilde{\mathbb{P}}$-almost surely to $\widetilde{D}^d \times \widetilde{D}$ with

$$\nabla_i \widetilde{X}_s^t = \widetilde{\Phi}_i^t(s), \qquad \nabla_i \widetilde{Y}_s^t = \widetilde{\Psi}_i^t(s) \quad \widetilde{\mathbb{P}}\text{-almost surely.}$$

We deduce that for every $t \leq s \leq T, \mathbb{P}$-almost every $\omega \in \Omega$, the function $x \longmapsto (X_s^{t,x}(\omega), Y_s^{t,x}(\omega))$ belongs to $D^d \times D$ and

$$\partial_x X_s^{t,x}(\omega) = \widetilde{\Phi}_s^t(x,\omega), \qquad \partial_x Y_s^{t,x}(\omega) = \widetilde{\Psi}_s^t(x,\omega)\mathrm{d}x \text{ a.e.} \quad \square$$

Corollary 3.2. *Assume (A1), (A3) and suppose that* $f \in C_b^{0,1}([0,T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d)$. *Let* $\left(X^{t,x}, Y^{t,x}, Z^{t,x}\right)$ *be the adapted solution of* (2) *and define* $u(t,x) = Y_t^{t,x}$. *Then,*

(i) *for every* $0 \leq t \leq T$, *the function* $x \longmapsto u(t,x)$ *belongs to* D *and for each* t *and* $i = 1, \ldots, d$, *the following representation holds:*

$$\partial_{x_i} u(t,\cdot) = \mathbb{E}\left\{\partial_{x_i} \widetilde{X}_T^t(\cdot,\cdot)\widetilde{\Phi}_{i,T}^t(\cdot,\cdot) + \int_t^T [\partial_x f(r, \widetilde{\Theta}_r^t(\cdot,\cdot))\widetilde{\Phi}_{i,r}^t(\cdot,\cdot) \right.$$
$$\left. + \ \partial_y f(r, \widetilde{\Theta}_r^t(\cdot,\cdot))\widetilde{\Psi}_{i,r}^t(\cdot,\cdot) + \langle \partial_z f(r, \widetilde{\Theta}_r^t(\cdot,\cdot)), \widetilde{\Gamma}_{i,r}^{t,\cdot}(\cdot,\cdot)\rangle]\mathrm{d}r\right\} \mathrm{d}x \ a.e.$$

where $\widetilde{\Theta}_r^t \equiv \left(\widetilde{X}_r^t, \widetilde{Y}_r^t, \widetilde{Z}_r^t\right)$ *is the solution of the FBSDE* (13) *and* $\left(\widetilde{\Phi}^t, \widetilde{\Psi}^t, \widetilde{\Gamma}^t\right)$ *is the solution of the variational equation Eq.* (15);

(ii) *For every* $t \leq s \leq T$, $\widetilde{\mathbb{P}}$*-almost surely, we have* $Z_s^{t,\cdot} = \partial_{x_i} u(s, X_s^{t,\cdot})\sigma(s, X_s^{t,\cdot})$.

4. The representation theorem

Our aim is now to give a probabilistic representation of the gradient of the viscosity solution of the quasilinear PDE (1). More precisely, we prove an extension of the nonlinear Feynman–Kac formula of Pardoux–Peng [16] and Ma–Zhang [14]. We restrict ourselves to the nondegenerate case.

For every $n \in \mathbb{N}$, let (X^n, Y^n, Z^n) and $(\Phi^n, \Psi^n, \Gamma^n)$ be the solutions of FBSDEs (9) and (11) respectively. For every $t < r_1 < T$, we introduce the martingales $\left\{M_{r_2}^{n,r_1} : r_1 \leq r_2 \leq T\right\}$:

$$M_{r_2}^{x,n,r_1} = \int_{r_1}^{r_2} \left[\sigma_n^{-1}(v, X_v^{x,n})\Phi_v^{x,n}\right]^* \mathrm{d}W_v.$$

We also consider the martingale $\left\{M_{r_2}^{x,r_1} : r_1 \leq r_2 \leq T\right\}$:

$$M_{r_2}^{x,r_1} = \int_{r_1}^{r_2} \left[\sigma^{-1}(v, X_v^x)\Phi_v^x\right]^* \mathrm{d}W_v,$$

(X^x, Y^x, Z^x) and $(\Phi^x, \Psi^x, \Gamma^x)$ being the solutions of the FBSDEs (2) and (10) respectively. Set

$$N_r^{x,n,s} = \frac{1}{r-s}(M_r^{x,n,s})^T [\Phi_s^{x,n}]^{-1} \quad \text{and}$$
$$N_r^{x,s} = \frac{1}{r-s}(M_r^{x,s})^* [\Phi_s^x]^{-1}, \quad 0 \leq t \leq s < r \leq T.$$

Theorem 4.1. *Assume (A1)–(A3) and suppose that* $g \in C_b^1(\mathbb{R}^d)$. *Let* (X^x, Y^x, Z^x) *be the adapted solution of* (2). *Then,*

(i) *for every $s \in [t, T]$, we have $\mathbb{P}$-almost surely*

$$Z_s^x = \mathbb{E}\left\{g(X_T^x)N_T^{x,s} + \int_s^T f(r, X_r^x, Y_r^x, Z_r^x)N_r^{x,s}\mathrm{d}r \mid \mathcal{F}_s^t\right\}\sigma(s, X_s^x)\mathrm{d}x \ a.e.; \tag{19}$$

(ii) *for almost every $x \in \mathbb{R}^d$, there exists a version of Z^x such that for $\mathbb{P}$-almost every $\omega \in \Omega$, the mapping $s \longmapsto Z_s^x(\omega)$ is continuous;*

(iii) *for every $t \in [0, T]$ we have*

$$\partial_x u(t, x) = \mathbb{E}\left\{g(X_T^x)N_T^{x,t} + \int_t^T f(r, X_r^x, Y_r^x, Z_r^x)N_r^{x,t}\mathrm{d}r\right\}\mathrm{d}x \ a.e.$$

where $\partial_x u(t, x)$ denotes the derivative in the distribution sense of u with respect to x.

Proof. Let us note that, in view of Theorem 4.2 in Ma–Zhang [14], we have $\mathbb{P}$-almost surely, $\forall s \in [t, T] \forall x \in \mathbb{R}^d$:

$$Z_s^{x,n} = \mathbb{E}\left\{g(X_T^{x,n})N_T^{x,n,s} + \int_s^T f(r, X_r^{x,n}, Y_r^{x,n}, Z_r^{x,n})N_r^{x,n,s}\mathrm{d}r \mid \mathcal{F}_s^t\right\}\sigma_n(s, X_s^{x,n}). \tag{20}$$

Lemmas 2.1 and 2.2 imply that for all $p \geq 2$

$$\lim_{n\to\infty} \mathbb{E}\left(|X^{t,x,n} - X^{t,x}|_{t,T}^{*,p} + |Y^{t,x,n} - Y^{t,x}|_{t,T}^{*,p} + \int_0^T |Z_s^{t,x,n} - Z_s^{t,x}|^2\mathrm{d}s\right) = 0, \tag{21}$$

$$\lim_{n\to\infty} \mathbb{E}\left(|\Phi^{t,n,x} - \Phi^{t,x}|_{t,T}^{*,p} + |\Psi^{t,n,x} - \Psi^{t,x}|_{t,T}^{*,p} + \int_0^T |\Gamma_s^{t,n,x} - \Gamma_s^{t,x}|^2\mathrm{d}s\right) = 0. \tag{22}$$

It follows that for any $p \geq 1$,

$$\lim_{n\to\infty} \mathbb{E}\left|M^{s,n} - M^s\right|_{t,T}^{*,2p} = 0. \tag{23}$$

Therefore (see Bahlali–Mezerdi–Ouknine [1])

$$\lim_{n\to\infty} \mathbb{E}\left|\mathbb{E}\left\{g(X_T^{x,n})N_T^{x,n,s} \mid \mathcal{F}_s^t\right\} - \mathbb{E}\left\{g(X_T^x)N_T^{x,s} \mid \mathcal{F}_s^t\right\}\right| = 0.$$

Now, we have

$$\begin{aligned}
&\mathbb{E}\left|\mathbb{E}\left\{\int_s^T f(r, X_r^{x,n}, Y_r^{x,n}, Z_r^{x,n})N_r^{x,n,s}\mathrm{d}r \mid \mathcal{F}_s^t\right\} - \mathbb{E}\left\{\int_s^T f(r, X_r^x, Y_r^x, Z_r^x)N_r^{x,s}\mathrm{d}r \mid \mathcal{F}_s^t\right\}\right| \\
&\leq \mathbb{E}\int_s^T \left|f(r, X_r^{x,n}, Y_r^{x,n}, Z_r^{x,n})N_r^{x,n,s} - f(r, X_r^x, Y_r^x, Z_r^x)N_r^{x,s}\right| \\
&\leq \mathbb{E}\int_s^T \left|f(r, X_r^{x,n}, Y_r^{x,n}, Z_r^{x,n})\right|\left|N_r^{x,n,s} - N_r^{x,s}\right|\mathrm{d}r \\
&\quad + \mathbb{E}\int_s^T \left|f(r, X_r^{x,n}, Y_r^{x,n}, Z_r^{x,n}) - f(r, X_r^x, Y_r^x, Z_r^x)\right|\left|N_r^{x,s}\right|\mathrm{d}r \\
&=: I_1^n + I_2^n.
\end{aligned}$$

Since f is Lipschitz continuous, we have

$$I_2^n \leq K\mathbb{E}\int_s^T \left(\left|X_r^{x,n} - X_r^x\right| + \left|Y_r^{x,n} - Y_r^x\right|\right)\left|N_r^{x,s}\right|\mathrm{d}r + K\mathbb{E}\int_s^T \left|Z_r^{x,n} - Z_r^x\right|\left|N_r^{x,s}\right|\mathrm{d}r.$$

By using (21) and (23), one can prove that the first term in the right hand side converges to 0 as n goes to infinity. For the second term, we use Corollary 3.2 in Ma–Zhang [14] and the Dominated Convergence Theorem to show that it converges to 0 as n goes to infinity.

To prove that $\lim_{n\to\infty} I_1^n = 0$, it suffices to observe, by using Corollary 3.2 in Ma–Zhang [14] and Lemmas 2.1 and 2.2, that for any $p > 0$,

$$\sup_n \mathbb{E}\left(|X^{t,x,n}|_{t,T}^{*,p} + |Y^{t,x,n}|_{t,T}^{*,p} + |Z^{t,x,n}|^{*,p}\right) < \infty,$$

$$\sup_n \mathbb{E}\left(|\Phi^{t,x,n}|_{t,T}^{*,p} + |\Psi^{t,x,n}|_{t,T}^{*,p}\right) < \infty$$

and combine (22), (23) with the Dominated Convergence Theorem to conclude.

Thus, by letting $n \to \infty$ in (20), we obtain that (19) holds $\mathbb{P}$-almost surely, for each fixed $s \in [t, T]$. Now, since part (ii) of the theorem can be proved as in Ma–Zhang [14], one can prove that part (i) is satisfied.

To obtain part (iii) it suffices to let $s = t$ in (19). $\square$

Remark 4.1. In [1], a representation theorem for functionals of diffusion processes with Lipschitz coefficients is proved. Therefore it is natural to try to obtain this kind of result for $(Y^{t,x}, Z^{t,x})$ which can be seen as a functional of $X^{t,x}$. To this purpose, we have to prove that $(Y^{t,x}, Z^{t,x}) = L(X^{t,x})$ and show that the functional L is Frechet differentiable.

Acknowledgement

The second author's research was supported by TWAS Research Grant 99-231 RG/MATHS/AF/AC.

References

[1] K. Bahlali, B. Mezerdi, Y. Ouknine, Martingale representation of functionals of diffusion processes with Lipschitz coefficients, J. Appl. Math. Stoch. Anal. 15 (4) (2002) 371–383.
[2] V. Bally, A. Matoussi, Weak solutions for SPDEs and backward doubly stochastic differential equations, J. Theoret. Probab. 14 (1) (2001) 125–164.
[3] G. Barles, E. Lesigne, SDE, BSDE and PDE, in: Pitman Res. Notes Math. Ser., vol. 364, Longman, Harlow, 1997.
[4] J. Bismut, Martingales, the Malliavin calculus and hypoellipticity under general Hörmander's condition, Z. Wahrshein 56 (1981) 469–505.
[5] J. Bismut, An introductory approach to duality in stochastic control, SIAM Rev. 20 (1978) 62–78.
[6] N. Bouleau, F. Hirsch, Dirichlet Forms and Analysis on Wiener Space, in: De Gruyter Studies in Math., vol. 14, Berlin, New York, 1991.
[7] N. Bouleau, F. Hirsch, On the derivability with respect to the initial data of the solution of a stochastic differential equation with Lipschitz coefficients, Lecture Notes in Math. 1393 (1989) 39–57.
[8] N. Bouleau, F. Hirsch, Propriétés d'absolue continuité dans les espaces de Dirichlet et application aux équations différentielles stochastiques, in: Séminaire de Probabilités XX, 1984/85, in: Lecture Notes Math., vol. 1204, Springer, 1986, pp. 131–161.
[9] N. El Karoui, S. Peng, M.C. Quenez, BSDEs in finance, Math. Finance 7 (1997) 1–71.
[10] S. Hamadène, J.P. Lepeltier, Zero-sum stochastic differential games and BSDEs, Systems Control. Lett. 24 (1995) 259–263.
[11] I. Karatzas, S.E. Shreve, Brownian Motion and Stochastic Calculus, 2nd ed., Springer-Verlag, New York, 1991.
[12] N.V. Krylov, Controlled Diffusion Processes, Springer Verlag, 1980.
[13] J. Ma, P. Protter, J. Yong, Solving forward-backward stochastic differential equations explicitly a four step scheme, Probab. Theory Related Fields 98 (1998) 339–359.
[14] J. Ma, J. Zhang, Representation theorems for backward stochastic differential equations, Ann. Appl. Probab. 4 (2002) 1390–1418.

[15] E. Pardoux, Backward stochastic differential equations and viscosity solutions of systems of semilinear parabolic and elliptic PDEs of second order, in: Stochastic Analysis and Related Topics (Geilo, 1996), in: Progr. Probab., vol. 42, Birkhäuser Boston, Boston, MA, 1998, pp. 79–127.

[16] E. Pardoux, S. Peng, Backward stochastic differential equations and quasilinear parabolic partial differential equations, Lecture Notes in Control and Inform Sci. 176 (1992) 200–217.

[17] E. Pardoux, S. Peng, Adapted solution of a backward stochastic differential equation, Systems Control Lett. 14 (1990) 55–61.