Available online at www.sciencedirect.com
SCIENCE DIRECT
stochastic processes and their applications
ELSEVIER Stochastic Processes and their Applications 112 (2004) 23–51
www.elsevier.com/locate/spa

Nonlinear Feynman–Kac formula and discrete-functional-type BSDEs with continuous coefficients

Ying Hu[a,*], Jin Ma[b,1]

[a] *IRMAR, U.M.R. C.N.R.S. 6625, Campus de Beaulieu, Université Rennes 1, 35042 Rennes Cedex, France*
[b] *Department of Mathematics, Purdue University, West Lafayette, IN 47907-1395, USA*

Received 30 September 2002; received in revised form 28 October 2003; accepted 2 February 2004

Abstract

In this paper, we study a class of multi-dimensional backward stochastic differential equations (BSDEs, for short) in which the terminal values and the generators are allowed to be "discrete-functionals" of a forward diffusion. We first establish some new types of Feynman–Kac formulas related to such BSDEs under various regularity conditions, and then we prove that under only bounded continuous assumptions on the generators, the adapted solution to such BSDEs does exist. Our result on the existence of the solutions to higher-dimensional BSDEs is new, and our representation theorem is the first step towards the long-standing "functional-type" Feynman–Kac formula.

MSC: primary: 60H10; secondary: 34F05; 93E03

Keywords: Backward SDEs with continuous coefficients; Discrete-functionals; Nonlinear Feynman–Kac formulae; Measurable selections

1. Introduction

Let $(\Omega, \mathscr{F}, P; \mathbf{F})$ be a complete, filtered probability space, where $\mathbf{F} \triangleq \{\mathscr{F}_t\}_{t\geqslant 0}$ is assumed to be the filtration generated by a standard, d-dimensional Brownian motion

* Corresponding author. Tel.: +33-223235820; fax: +33-223236790.
E-mail addresses: Ying.Hu@univ-rennes1.fr (Ying Hu), majin@math.purdue.edu (Jin Ma).
[1] This author is supported in part by NSF Grant #9971720. Part of this work was completed when this author was visiting IRMAR, Université de Rennes 1, France, whose hospitality is greatly appreciated.

doi:10.1016/j.spa.2004.02.002

$W=\{W_t; t\geqslant 0\}$. A *backward stochastic differential equation* (BSDE for short) is of the following form:

$$Y_t = \xi + \int_t^T F(r,Y_r,Z_r)\,\mathrm{d}r - \int_t^T Z_r\,\mathrm{d}W_r, \qquad t\in[0,T], \tag{1.1}$$

where $F:\Omega\times[0,T]\times\mathbb{R}^m\times\mathbb{R}^{m\times d}\mapsto\mathbb{R}^m$ is some appropriate measurable function, called the *generator* of the BSDE. An *adapted solution* to the BSDE (1.1) is a pair of $\mathbf{F}$-adapted, $\mathbb{R}^m\times\mathbb{R}^{m\times d}$-valued processes (Y,Z) that satisfies (1.1) almost surely.

The theory of Backward SDEs, initiated by Bismut (1976) and later developed by Pardoux and Peng (1990), has seen significant development during the past decade. We refer the readers to the books of El Karoui and Mazliak (1997), Ma and Yong (1999), as well as the well-known survey paper of El Karoui et al. (1997) for all detailed accounts of both theory and application (especially in mathematical finance and stochastic control) for such equations.

In this paper, we are interested in the following two long-standing problems in the theory of BSDEs:

(i) *Suppose $m>1$, and that the generator f is only bounded and continuous (in all variables). Do we still have the existence of the (strong) adapted solution to the BSDE (1.2)?*

(ii) *Suppose that the terminal value ξ and the generator f are of the form $\xi=g(X)_T$ and $f(\omega,t,y,z)=f(t,X_\cdot(\omega),y,z)$, where X is a forward diffusion, and $g(\cdot)$ and $f(t,\cdot,y,z)$ are functionals of X. Then to what extent we can still have the "nonlinear Feynman–Kac" formula? That is, we can represent an adapted solution of BSDE, whenever exists, as some function or functional of the forward diffusion via a solution of a system of partial differential equations (PDEs)?*

To better illustrate these two problems let us be more specific. Consider the following BSDE:

$$Y_t = g(X)_T + \int_t^T f(r,X,Y_r,Z_r)\,\mathrm{d}r - \int_t^T Z_r\,\mathrm{d}W_r, \qquad t\in[0,T], \tag{1.2}$$

where X is an n-dimensional diffusion satisfying the SDE

$$X_t = x + \int_0^t b(r,X_r)\,\mathrm{d}r + \int_0^t \sigma(r,X_r)\,\mathrm{d}W_r, \qquad t\in[0,T], \tag{1.3}$$

in which $b:[0,T]\times\mathbb{R}^n\mapsto\mathbb{R}^n$, $\sigma:[0,T]\times\mathbb{R}^n\mapsto\mathbb{R}^{n\times d}$ are some measurable functions, $f:[0,T]\times C([0,T];\mathbb{R}^n)\times\mathbb{R}^m\times\mathbb{R}^{m\times d}\mapsto\mathbb{R}^m$ is a "non-anticipative functional" with respect to X, and g is some functional defined on the path space $C([0,T];\mathbb{R}^n)$. We should note that while BSDE (1.2) is still "non-Markovian", it has more structure than (1.1). In what follows, our discussion will be mainly focus on such BSDEs.

Return now to the two questions. It is clear that the first one is simply a question of existence of adapted solution under merely continuous assumption on the coefficients. Such problems have been studied by many authors (see, for example, Hamadene, 1996; Lepeltier and San Martin, 1997, 1998; Kobylanski, 2000, to mention a few). However, most of the existing results are restricted to the one-dimensional case (i.e., $m=1$), due to the comparison-theorem-technique used in the proofs. The case when $m > 1$ was studied by Hamadene et al. (1997), but only in the "Markovian" case when both g and f are "functions" of the forward diffusion. As a matter of fact, to our best knowledge, to date there has been no result on the existence of the adapted solution for higher dimensional, functional-type of BSDEs with only continuous coefficients.

The second question is more subtle. It is known that if g and f are both "*functions*" of X, then the nonlinear Feynman–Kac formula (see, e.g., Pardoux and Peng, 1992 or Ma and Yong, 1999), modulo some regularity assumptions, provides us the following representations for the adapted solution (Y,Z):

$$Y_t = u(t,X_t), \quad Z_t = \partial_x u(t,X_t)\sigma(t,X_t), \quad t \in [0,T], \tag{1.4}$$

where u is the solution to a semilinear/quasilinear parabolic PDE, in a certain sense. In fact, a result by El Karoui et al. (1997) indicated that one can always represent the components of the adapted solution in terms of the forward diffusion:

$$Y_t = u(t,X_t), \quad Z_t = v(t,X_t)\sigma(t,X_t), \quad t \in [0,T], \tag{1.5}$$

where u and v are only measurable functions, based on a deep result in semimartingale theory by Çinlar et al. (1980). These nice features, however, will lose their grounds in the functional BSDE case. For example, while it might still be conceivable that the relations in (1.5) could be modified to

$$Y_t = u(t,X_\cdot), \quad Z_t = v(t,X_\cdot)\sigma(t,X_t), \quad t \in [0,T], \tag{1.6}$$

where $u(\cdot,\cdot)$ and $v(\cdot,\cdot)$ are two progressively measurable functionals defined on $[0,T] \times C([0,T];\mathbb{R}^n)$, the form of a Feynman–Kac formula, if it exists at all, is by no means clear. In fact, one of our motivations of studying such problem comes from finance: for example, can we generalize the Black–Scholes PDE to general path-dependent exotic options in any form?

This paper is a first attempt to answer these two questions. To be more precise, we shall consider the case where the functionals g and f are of the following "*discrete-functional type*":

$$g(X) = g(X_{t_1},\dots,X_{t_N}), \tag{1.7}$$

$$f(t,X,Y_t,Z_t) = f(t,X_{t_1\wedge t},\dots,X_{t_N\wedge t},Y_t,Z_t), \tag{1.8}$$

where $0 = t_0 < t_1 < \cdots < t_N = T$ is a given partition of $[0,T]$. We shall first prove that the Feynman–Kac formula still holds in this case and derive the corresponding PDEs, in both classical sense and viscosity sense. It is worth noting that in this "piecewise Markovian" case, we can show that the following representation holds:

$$Y_t = u(t,X_{t_1\wedge t},\dots,X_{t_N\wedge t}); \quad Z_t = v(t,X_{t_1\wedge t},\dots,X_{t_N\wedge t})\sigma(t,X_t), \quad t \in [0,T], \tag{1.9}$$

where u is a solution (in a certain sense) of a system of semilinear PDEs, partially justifying our conjecture (1.6). Using the Feynman–Kac formula and borrowing some ideas from Hamadene et al. (1997), we can then prove the existence of the adapted solution to BSDE (1.2) with continuous coefficients in this piecewise Markovian case.

Finally, we remark that the relations between BSDEs with discrete-functional-type terminal have been discussed also by Zhang and Zheng (2002) and Ma and Zhang

(2002). But in those cases, the generator f were assumed to be of "Markovian" type, that is, it is the function with respect to the forward diffusion, rather than a functional as is proposed in this paper. It is our hope that our result has the potential to be developed into some completely non-Markovian cases. We also remark that the technical assumptions made in this paper are by no means the sharpest. Some of them can be improved with some extra effort. But since this is not the main purpose of the paper, we prefer not to over stress these technical points so as to make this already complicated subject a little easier to bear.

The rest of the paper is organized as follows. In Section 2, we give all the necessary preparations. In Section 3, we present the Feynman–Kac formula with strong conditions on the coefficients; we then establish a less trivial version of the Feynman–Kac formula, in terms of the viscosity solution of the corresponding PDEs, in Section 4. In Section 5, we prove an important measurable selection theorem, with which we prove the second main result: the existence of adapted solution to BSDEs with continuous coefficients, in Section 6.

2. Preliminaries

Throughout this paper, we assume that $(\Omega,\mathscr{F},P)$ is a complete probability space on which is defined a d-dimensional Brownian motion $W=(W_t)_{t\geqslant 0}$. Let $\mathbf{F}\triangleq\{\mathscr{F}_t\}_{t\geqslant 0}$ denote the natural filtration generated by W, augmented by the P-null sets of $\mathscr{F}$; and let $\mathscr{F}=\mathscr{F}_\infty$. We note here that if necessary we may assume that $(\Omega,\mathscr{F},P;\{\mathscr{F}_t\})$ is the *canonical space*. Thus for any sub-σ-field $\mathscr{G}\subseteq\mathscr{F}$, the regular conditional probabilities $P^{\omega}_{\mathscr{G}}(\cdot)\triangleq P\{\cdot|\mathscr{G}\}(\omega)$ exist, for a.e. $\omega\in\Omega$.

In what follows, we denote $\mathbb{E}$ to be a generic Euclidean space (or $\mathbb{E}_1$, $\mathbb{E}_2,\ldots$, if different spaces are used simultaneously); and regardless of their dimensions we denote $\langle\cdot,\cdot\rangle$ and $|\cdot|$ to be the inner product and norm in all $\mathbb{E}$'s, respectively. Furthermore, we use the notations $\partial_t=\partial/\partial t$, $\partial_x=(\partial/\partial x_1,\ldots,\partial/\partial x_n)$, and $\partial^2=\partial_{xx}=(\partial^2_{x_ix_j})$, for $(t,x)\in[0,T]\times\mathbb{R}^n$. Note that if $\psi=(\psi^1,\ldots,\psi^m):\mathbb{R}^n\mapsto\mathbb{R}^m$, then $\partial_x\psi\triangleq(\partial_{x_j}\psi^i)$ is an $m\times n$ matrix.

Now let $\mathscr{X}$ be a generic Banach space, whose topological Borel field is denoted by $\mathscr{B}(\mathscr{X})$. If $\mathscr{X}$ and $\mathscr{Y}$ are two such spaces, we shall denote $L^0(\mathscr{X};\mathscr{Y})$ to be the space of all $\mathscr{B}(\mathscr{X})/\mathscr{B}(\mathscr{Y})$-measurable functions. The following spaces will be frequently used in the sequel:

- for any sub-σ-field $\mathscr{G}\subseteq\mathscr{F}_T$, $L^0(\mathscr{G};\mathscr{X})$ denotes the space of all $\mathscr{X}$-valued, $\mathscr{G}$-measurable random variables.
- for any sub-σ-field $\mathscr{G}\subseteq\mathscr{F}_T$ and $1\leqslant p<\infty$, $L^p(\mathscr{G};\mathscr{X})$ denotes the space of all $\mathscr{X}$-valued, $\mathscr{G}$-measurable random variables ξ such that $E\|\xi\|^p_{\mathscr{X}}<\infty$. Moreover, $\xi\in L^\infty(\mathscr{G};\mathscr{X})$ means that it is $\mathscr{G}$-measurable and bounded under $\|\cdot\|_{\mathscr{X}}$;
- for $1\leqslant p<\infty$, $L^p(\mathbf{F},[0,T];\mathbb{E})$ is the space of all $\mathbb{E}$-valued, $\mathbf{F}$-progressively measurable processes ξ satisfying $E\int_0^T|\xi_t|^p\,\mathrm{d}t<\infty$; and $L^\infty(\mathbf{F},[0,T];\mathbb{E})$ is the space of all $\mathbb{E}$-valued, $\mathbf{F}$-progressively measurable processes uniformly bounded in (t,ω).
- for $1\leqslant p<\infty$ and $\mathscr{X}=C([0,T];\mathbb{E})$ or $L^p([0,T];\mathbb{E})$, $L^p(\mathbf{F},\Omega;\mathscr{X})$ denotes the space consisting of all $\xi\in L^p(\mathbf{F},[0,T];\mathbb{E})\cap L^p(\mathscr{F}_T;\mathscr{X})$ such that the paths $\xi_\cdot\in\mathscr{X}$.

Our main objective of this paper is to study the following (discrete) functional-type (forward–)backward SDEs defined on an arbitrary interval $[t,T]\subseteq[0,T]$: for $s\in[t,T]$,

$$X_s = x + \int_t^s b(r,X_r)\,\mathrm{d}r + \int_t^s \sigma(r,X_r)\,\mathrm{d}W_r,$$
$$Y_s = g(X_{t_1},\dots,X_{t_N}) + \int_s^T f(r,X_{t_1\wedge r},\dots,X_{t_N\wedge r},Y_r,Z_r)\,\mathrm{d}r - \int_s^T Z_r\,\mathrm{d}W_r, \tag{2.1}$$

where $\pi: 0 = t_0 < t_1 < \cdots < t_N = T$ is a given partition on $[0,T]$. We denote any solution to (2.1), whenever exists, by $(X^{t,x},Y^{t,x},Z^{t,x})$ to indicate its dependence on the initial data (t,x). We should note that in general the solution to (2.1) is not unique, but we nevertheless use the same notation when the context is clear.

Let $\pi: 0=t_0<t_1<t_2<\cdots<t_N=T$ be a given partition. For any $\mathbf{x}=(x_1,\dots,x_N)\in\mathbb{R}^{Nn}$ and $k=1,2,\dots,N$, we denote

$$\mathbf{x}^{(k)} = (x_1,\dots,x_k)\in\mathbb{R}^{kn}, \quad \mathbf{x}^{(k,N)} = (x_k,\dots,x_N)\in\mathbb{R}^{(N-k+1)n}.$$

Further, let $X^{0,x}$ be the solution to the forward SDE in (2.1), we define

$$\mathbf{X}_t^{(k)} \triangleq (X^{0,x}_{t_1\wedge t},\dots,X^{0,x}_{t_k\wedge t}), \quad \mathbf{X}_t^{(k,N)} \triangleq (X^{0,x}_{t_k\wedge t},\dots,X^{0,x}_{t_N\wedge t}), \quad k=1,2,\dots,N. \tag{2.2}$$

In particular, we denote

$$\mathbf{X}^{(k)} \triangleq (X^{0,x}_{t_1},\dots,X^{0,x}_{t_k}), \quad \mathbf{X}^{(k,N)} \triangleq (X^{0,x}_{t_k},\dots,X^{0,x}_{t_N}). \tag{2.3}$$

Clearly, using this notation the BSDE in (2.1) can be rewritten as

$$Y_t = g(\mathbf{X}^{(N)}) + \int_t^T f(s,\mathbf{X}_s^{(N)},Y_s,Z_s)\,\mathrm{d}s - \int_t^T Z_s\,\mathrm{d}W_s. \tag{2.4}$$

We shall make use of the following *Standing Assumptions*:

(A1) The functions $b:[0,T]\times\mathbb{R}^n\mapsto\mathbb{R}^n$ and $\sigma:[0,T]\times\mathbb{R}^n\mapsto\mathbb{R}^{n\times d}$ are continuous. Moreover, there exists a constant $L_1>0$, such that

$$|b(t,x_1)-b(t,x_2)| + |\sigma(t,x_1)-\sigma(t,x_2)| \leqslant L_1|x_1-x_2|$$

for any $(t,x_1,x_2)\in[0,T]\times\mathbb{R}^n\times\mathbb{R}^n$; and

$$|b(t,x)|+|\sigma(t,x)| \leqslant L_1(1+|x|) \quad \forall (t,x)\in[0,T]\times\mathbb{R}^n.$$

(A2) The function f belongs to the space $C_b([0,T]\times\mathbb{R}^{Nn}\times\mathbb{R}^m\times\mathbb{R}^{m\times d};\mathbb{R}^m)$ and the function g belongs to the space $L^\infty(\mathbb{R}^{Nn};\mathbb{R}^m)$.

In some of our discussions assumptions (A1) and (A2) need to be strengthened. We list the possible extra assumptions for convenience.

(A3) The functions b and σ are Lipschitz in (t,x), and $\exists 0<c<C$, such that

$$\begin{aligned} &|b(t,x)|+|\sigma(t,x)| \leqslant C \quad \forall (t,x)\in[0,T]\times\mathbb{R}^n,\\ &\xi^T\sigma(t,x)\sigma^T(t,x)\xi \geqslant c|\xi|^2 \quad \forall x,\xi\in\mathbb{R}^n,\ t\in[0,T]. \end{aligned} \tag{2.5}$$

(A4) There exists a constant $L_2 > 0$, such that

$$|f(t, \mathbf{x}_1, y_1, z_1) - f(t, \mathbf{x}_2, y_2, z_2)| \leqslant L_2(|\mathbf{x}_1 - \mathbf{x}_2| + |y_1 - y_2| + |z_1 - z_2|)$$

for all $(t, \mathbf{x}_1, y_1, z_1, \mathbf{x}_2, y_2, z_2) \in [0, T] \times \mathbb{R}^{Nn} \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \times \mathbb{R}^{Nn} \times \mathbb{R}^m \times \mathbb{R}^{m\times d}$, and

$$|g(\mathbf{x})| + |f(t, \mathbf{x}, y, z)| \leqslant L_2(1 + |\mathbf{x}| + |y| + |z|)$$

for all $(t, \mathbf{x}, y, z) \in [0, T] \times \mathbb{R}^{Nn} \times \mathbb{R}^m \times \mathbb{R}^{m\times d}$.

To end this section, we give a lemma concerning the transition probability densities of the diffusion process X, the solution to the forward SDE (1.3). Note that under assumption (A3), X is a strong Markov process with positive transition density. The following result will be useful in our future discussion. Its proof can be found in Aronson (1967) and/or Aronson (1968), we only state it here for ready reference.

Lemma 2.1. *Assume* (A3). *For* $(t, x) \in [0, T] \times \mathbb{R}^n$, *denote* $\mu(t, x; s, \mathrm{d}y) \triangleq P\{X_s^{t,x} \in \mathrm{d}y\}$ *to be the transition probability of* $X^{t,x}$ *and* $p(t, x; s, y)$, $(t < s \leqslant T)$ *to be its density. Then, the mapping* $(t, x) \mapsto p(t, x; s, y)$ *is continuous, for fixed* (s, y); *and there exist constants* $m, \lambda, M, \Lambda > 0$, *such that the density function* $p(t, x; s, y)$ *satisfies the following estimation: for* $0 \leqslant t < s \leqslant T$,

$$m(s-t)^{-n/2} \exp\left\{\frac{-\lambda|y-x|^2}{s-t}\right\} \leqslant p(t, x; s, y)$$

$$\leqslant M(s-t)^{-n/2} \exp\left\{\frac{-\Lambda|y-x|^2}{s-t}\right\}.$$

3. Nonlinear Feynman–Kac formula via classical solutions

In this section, we take a first look at the possible nonlinear Feynman–Kac formula in the case where the BSDEs are of "discrete functional" form. To be more precise, recall BSDE (2.4):

$$Y_t = g(\mathbf{X}^{(N)}) + \int_t^T f(s, \mathbf{X}_s^{(N)}, Y_s, Z_s)\,\mathrm{d}s - \int_t^T Z_s\,\mathrm{d}W_s. \tag{3.1}$$

We shall assume that the corresponding system of (semilinear) PDEs has a classical solution (say, u) and prove that in such a case the adapted solution of (3.1) (Y, Z) is related to the forward component X via a pair of functions $u : [0, T] \times \mathbb{R}^{Nn} \mapsto \mathbb{R}^m$ and $v : [0, T] \times \mathbb{R}^{Nn} \mapsto \mathbb{R}^{m\times n}$, such that

$$Y_t = u(t, \mathbf{X}_t^{(N)}); \quad Z_t = v(t, \mathbf{X}_t^{(N)})\sigma(t, X_t^{0,x}), \tag{3.2}$$

as we predicted. We should note that various assumptions can be made to guarantee the existence and uniqueness of the classical solution to the system of PDEs, as well as the adapted solution to the BSDE (3.1). For example, if b and σ satisfy assumption (A3), and f and g are bounded and smooth with bounded derivatives (hence satisfy (A4)), then the resulting system of PDEs will have a classical solution with bounded derivatives (cf. e.g., Ladyzenskaja et al., 1968), and BSDE (3.1) will have a unique adapted solution. But at this point we would rather not to concentrate on the particular assumptions.

To begin with, for each $k = N, N-1, \ldots, 1$, we consider a sequence of semilinear PDEs with parameters, defined recursively in a "backward" manner as follows: first fix $\mathbf{x}^{(N-1)}$ as a parameter, and define

$$u^{N+1}(T, \mathbf{x}^{(N-1)}, x, x) = g(\mathbf{x}^{(N-1)}, x), \quad x \in \mathbb{R}^n.$$

Next, for each $k = N, N-1, \ldots, 1$, we fix $\mathbf{x}^{(k-1)}$ as a parameter, and consider the following system of PDEs: for $(t,x) \in [t_{k-1}, t_k] \times \mathbb{R}^n$,

$$\begin{aligned} &\frac{\partial u_i^k}{\partial t}(t, \mathbf{x}^{(k-1)}, x) + \mathscr{L} u_i^k(t, \mathbf{x}^{(k-1)}, x) \\ &\quad + f_i(t, \mathbf{x}^{(k-1)}, \underbrace{x, \ldots, x}_{N-k+1}, u^k(t, \mathbf{x}^{(k-1)}, x), \partial_x u^k(t, \mathbf{x}^{(k-1)}, x)\sigma(t,x)) = 0, \\ &u_i^k(t_k, \mathbf{x}^{(k-1)}, x) = u_i^{k+1}(t_k, \mathbf{x}^{(k-1)}, x, x), \quad i = 1, \ldots, m. \end{aligned} \tag{3.3}$$

Here, for $\varphi \in C^2(\mathbb{R}^n)$ the operator $\mathscr{L}$ is given by

$$[\mathscr{L}\varphi](t,x) = \tfrac{1}{2}\,\mathrm{tr}\{\sigma\sigma^T(t,x)\partial_x^2\varphi(x)\} + \langle b(t,x), (\partial_x\varphi)^T(x)\rangle. \tag{3.4}$$

Now let us suppose that all PDEs in (3.3) have classical solutions, and we denote them by u_i^k, $i = 1, \ldots, m$, $k = N, N-1, \ldots, 1$. Let us denote also that

$$v_{ij}^k(t, \mathbf{x}^{(k-1)}, x) = \partial_{x_j} u_i^k(t, \mathbf{x}^{(k-1)}, x), \quad k = N, N-1, \ldots, 1; \; i, j = 1, \ldots, m$$

(or simply $v^k(\cdots) = \partial_x u^k(\cdots)$). We consider the following processes:

$$\begin{aligned} Y_t^k &= u^k(t, \mathbf{X}_t^{(k-1)}(\omega), X_t), \\ Z_t^k &= v^k(t, \mathbf{X}_t^{(k-1)}(\omega), X_t)\sigma(t, X_t) = \partial_x u^k(t, \mathbf{X}_t^{(k-1)}(\omega), X_t)\sigma(t, X_t). \end{aligned} \tag{3.5}$$

Furthermore, let us "patch" the functions u^k's and v^k's together by defining the following functions u and v:

$$u(T, \mathbf{x}^{(N)}) \triangleq u^N(T, \mathbf{x}^{(N)}), \quad v(T, \mathbf{x}^{(N)}) \triangleq v^N(T, \mathbf{x}^{(N)})$$

and for $t \in [0, T)$,

$$\left.\begin{aligned} u(t, \mathbf{x}^{(N)}) &\equiv u(t, \mathbf{x}^{(k-1)}, \underbrace{x_k, \ldots, x_k}_{N-k+1}) \triangleq u^k(t, \mathbf{x}^{(k)}) \\ v(t, \mathbf{x}^{(N)}) &\equiv v(t, \mathbf{x}^{(k-1)}, \underbrace{x_k, \ldots, x_k}_{N-k+1}) \triangleq v^k(t, \mathbf{x}^{(k)}) \end{aligned}\right\} \quad \text{if } t \in [t_{k-1}, t_k) \tag{3.6}$$

and finally define two processes

$$Y_t = u(t, \mathbf{X}_t^{(N)}), \quad Z_t = v(t, \mathbf{X}_t^{(N)})\sigma(t, X_t), \quad t \in [0, T]. \tag{3.7}$$

Then, we have the following version of "*nonlinear Feynman–Kac*" formula.

Theorem 3.1. *Assume that all PDEs in system* (3.3) *have classical solutions with bounded derivatives. Then, the processes* (Y,Z) *defined by* (3.7) *solves BSDE* (2.4) *on* $[0,T]$.

Proof. We shall check only the case when $t\in[t_{N-1},t_N]=[t_{N-1},T]$, the other cases can be argued in the same way.

First note that when $t\in[t_{N-1},T]$ we have

$$Y_t=Y_t^N=u^N(t,\mathbf{X}^{(N-1)},X_t),$$
$$Z_t=Z_t^N=v^N(t,\mathbf{X}^{(N-1)},X_t)\sigma(t,X_t)=\partial_x u^N(t,\mathbf{X}^{(N-1)},X_t)\sigma(t,X_t).$$

We need only to show that the set

$$A=\left\{\omega'\in\Omega\colon\ Y_t^N(\omega')=g(\mathbf{X}^{(N)}(\omega'))+\int_t^T f(s,\mathbf{X}^{(N-1)},X_s,Y_s^N,Z_s^N)(\omega')\,\mathrm{d}s\right.$$
$$\left.-\int_t^T Z_s^N\,\mathrm{d}W_s(\omega'),\ t\in[t_{N-1},T]\right\}$$

satisfies $P(A)=1$. To see this we consider the regular conditional probability

$$P_{N-1}^{\omega}(B)\triangleq P(B|\mathscr{F}_{t_{N-1}})(\omega)\quad\forall B\in\mathscr{F},\ P\text{-a.e. }\omega\in\Omega.$$

By the property of regular conditional probability, we know that for any $\mathscr{F}_{t_{N-1}}$-measurable random vector η, it holds that

$$P_{N-1}^{\omega}\{\eta(\omega')=\eta(\omega)\}=1\quad\text{ for }P\text{-a.e. }\omega\in\Omega$$

and furthermore,

$$P(A)=\int_\Omega P_{N-1}^{\omega}(A)P(\mathrm{d}\omega).\tag{3.8}$$

Note that for P-a.e. $\omega\in\Omega$, one has P_{N-1}^{ω}-a.s.,

$$g(\mathbf{X}^N(\cdot))=g(\mathbf{X}^{(N-1)}(\omega),X_T(\cdot)),$$
$$f(s,\mathbf{X}^{(N-1)},X_s,Y_s^N,Z_s^N)(\cdot)=f(s,\mathbf{X}^{(N-1)}(\omega),X_s(\cdot),Y_s^N(\cdot),Z_s^N(\cdot)).$$

Thus, we can apply the nonlinear Feynman–Kac formula (see, e.g., Pardoux and Peng, 1992; Ma et al., 1994) to obtain that for P-a.e. $\omega\in\Omega$, on the probability space $(\Omega,\mathscr{F},P_{N-1}^{\omega})$ the pair (Y^N,Z^N) solves the BSDE over $[t_{N-1},T]$,

$$Y_t^N=g(\mathbf{X}^{(N-1)}(\omega),X_T)+\int_t^T f(s,\mathbf{X}_s^{(N-1)}(\omega),X_s,Y_s^N,Z_s^N)\,\mathrm{d}s-\int_t^T Z_s^N\,\mathrm{d}W_s.\tag{3.9}$$

That is $P^{\omega}_{N-1}(A)=1$, for P-a.e. $\omega\in\Omega$, whence $P(A)=1$, thanks to (3.8). To complete the proof we note that at $t=t_{N-1}$ one has, for P-a.e. $\omega\in\Omega$,

$$\begin{aligned} Y^N_{t_{N-1}}(\omega) &= u^N(t_{N-1},\mathbf{X}^{(N-1)}_{t_{N-1}}(\omega),X_{t_{N-1}}(\omega)) \\ &= u^N(t_{N-1},\mathbf{X}^{(N-2)}(\omega),X_{t_{N-1}}(\omega),X_{t_{N-1}}(\omega)) \\ &= u^{N-1}(t_{N-1},\mathbf{X}^{(N-2)}(\omega),X_{t_{N-1}}(\omega)) = Y^{N-1}_{t_{N-1}}(\omega). \end{aligned}$$

Using the definition of the functions u^{N-1} and v^{N-1} we can similarly prove that $(Y,Z)=(Y^{N-1},Z^{N-1})$ solves BSDE (2.4) on $[t_{N-2},t_{N-1}]$. Continuing this way for N steps and noting that the consistency requirements: $Y^i_{t_{i-1}}=Y^{i-1}_{t_{i-1}}$ holds almost surely for all $i=1,2,\dots,N$, we obtain an adapted solution (Y,Z) on the whole interval $[0,T]$, proving the theorem. □

Next, we consider BSDEs defined on $[t,T]$, where $0<t<T$. Note that in this case no information is given on the interval $[0,t)$. However, for the sake of consistency we shall still assume that a partition is given on $[0,T]$ as before, but the initial time $t\in[t_{k-1},t_k)$ for some $1\leqslant k\leqslant N$. As usual let $X^{t,x}$ be the solution to the forward SDE on $[t,T]$:

$$X_s = x+\int_t^s b(r,X_r)\,\mathrm{d}r+\int_t^s \sigma(r,X_r)\,\mathrm{d}W_r,\qquad s\in[t,T]. \tag{3.10}$$

For $1\leqslant k<j\leqslant N$, we denote

$$\begin{aligned} &\mathbf{x}^{(k,j)}=(x_k,\dots,x_j)\in\mathbb{R}^{(j-k+1)n},\\ &\mathbf{X}^{(k,j),(t,x)}_s=(X^{t,x}_{t_k\wedge s},\dots,X^{t,x}_{t_j\wedge s}),\qquad s\in[t,T],\\ &\mathbf{X}^{(k,j),(t,x)}=(X^{t,x}_{t_k},\dots,X^{t,x}_{t_j}). \end{aligned} \tag{3.11}$$

If the context is clear, we shall denote simply $\mathbf{X}^{(k,j)}_s=\mathbf{X}^{(k,j),(t,x)}_s$, and $\mathbf{X}^{(k,j)}=\mathbf{X}^{(k,j),(t,x)}$. For any $\mathbf{x}^{(k-1)}\in\mathbb{R}^{(k-1)n}$, we denote $(Y^{t,x}(\mathbf{x}^{(k-1)}),Z^{t,x}(\mathbf{x}^{(k-1)}))$ to be the solution to the following BSDE:

$$Y_s=g(\mathbf{x}^{(k-1)},\mathbf{X}^{(k,N)})+\int_s^T f(r,\mathbf{x}^{(k-1)},\mathbf{X}^{(k,N)}_r,Y_r,Z_r)\,\mathrm{d}r-\int_s^T Z_r\,\mathrm{d}W_r. \tag{3.12}$$

We have the following variation of Theorem 3.1.

Theorem 3.2. *Suppose that all PDEs in system (3.3) have classical solutions with bounded derivatives. Assume that $t\in[t_{k-1},t_k)$, for some $1\leqslant k\leqslant N$, and denote $(Y^{t,x}(\mathbf{x}^{(k-1)}),Z^{t,x}(\mathbf{x}^{(k-1)}))$ to be the solution to (3.12), for any $\mathbf{x}^{(k-1)}\in\mathbb{R}^{(k-1)n}$. Then, it holds that*

$$\left.\begin{aligned} &Y^{t,x}_s(\mathbf{x}^{(k-1)})=u(s,\mathbf{x}^{(k-1)},\mathbf{X}^{(k,N)}_s),\\ &Z^{t,x}_s(\mathbf{x}^{(k-1)})=v(s,\mathbf{x}^{(k-1)},\mathbf{X}^{(k,N)}_s)\sigma(s,X_s), \end{aligned}\right.\quad s\in[t,T], \tag{3.13}$$

where $\mathbf{X}^{k,N}$ is defined by (3.11) and u,v are defined by (3.6).

Proof. The case when $k=N$ is trivial. We assume that $k<N$, and $s\in[t_{N-1},t_N]$. For notational simplicity, in what follows we denote $\xi_s=\xi_s^{t,x}(\mathbf{x}^{(k-1)})$, for $\xi=Y,Z$. That is, (Y,Z) satisfies the BSDE (3.12) on $[t,T]$.

Again let us consider the probability space $(\Omega,\mathscr{F},P^\omega_{N-1})$, for P-a.e. $\omega\in\Omega$. Applying the nonlinear Feynman–Kac formula, and using the Markovian property of X we have

$$Y_s=u^N(s,\mathbf{x}^{(k-1)},\mathbf{X}^{(k,N-1)}(\omega),X_s^{t,x}),$$

$$Z_s=v^N(s,\mathbf{x}^{(k-1)},\mathbf{X}^{(k,N-1)}(\omega),X_s^{t,x})\sigma(s,X_s^{t,x}).$$

Consequently, using the definition of the functions u, v, and the property of the regular conditional probability as before, we see that,

$$\begin{aligned}P\{\omega' : Y_s(\omega')&=u(s,\mathbf{x}^{(k-1)},\mathbf{X}_s^{(k,N)}(\omega'));\\ &Z_s(\omega')=v(s,\mathbf{x}^{(k-1)},\mathbf{X}_s^{(k,N)}(\omega'))\sigma(s,X_s^{t,x}(\omega'))\}=1.\end{aligned}$$

The case when s belongs to other intervals can be argued in the same way, proving the theorem. □

Remark 3.3. From (3.6), we see that the process Z is càdlàg.

4. Nonlinear Feynman–Kac formula via viscosity solutions

In the previous section, we proved one direction of the nonlinear Feynman–Kac formula, that is, we assume that the system of PDEs (3.3) has classical solutions, then it can produce the adapted solution to the "functional-type" BSDEs. In this section, we show that the reverse direction is also true. To be more precise, we shall prove that if the BSDE (3.12) has an adapted solution, then it will provide a probabilistic solution to the system of PDEs (3.3) in the sense of "viscosity". To this end, we assume that the standing assumptions (A1) and (A4) hold in the sequel. Also, for technical simplicity in this section we consider only the case $m=1$. We should note, however, that such a simplification by no means affects our future results.

To begin with, let us recall the BSDE (3.12):

$$Y_s=g(\mathbf{x}^{(k-1)},\mathbf{X}^{(k,N)})+\int_s^T f(r,\mathbf{x}^{(k-1)},\mathbf{X}_r^{(k,N)},Y_r,Z_r)\,\mathrm{d}r-\int_s^T Z_r\,\mathrm{d}W_r, \tag{4.1}$$

where $t\in[t_{k-1},t_k)$, $0\leqslant t\leqslant s\leqslant T$, $\mathbf{x}^{(k-1)}\in\mathbb{R}^{(k-1)n}$, and $\mathbf{X}^{(k,N)}=\mathbf{X}^{(k,N),(t,x)}$, as defined by (3.11). We note that under assumption (A4), BSDE (4.1) has a unique adapted solution on any interval $[t,T]$. We denote such solution by $(Y^{t,x}(\mathbf{x}^{(k-1)}),Z^{t,x}(\mathbf{x}^{(k-1)}))$.

The following lemma can be proved in a rather standard way. We shall only state it for ready reference, but omit the proof.

Lemma 4.1. *Suppose that assumptions* (A1) *and* (A4) *are in force. Then there exists a constant* $L>0$, *such that for all* $t\in[0,T]$, *it holds that*

$$E\left[\sup_{t\leqslant s\leqslant T}|Y_s^{t,x}(\mathbf{x}^{(k-1)})-Y_s^{t,y}(\mathbf{y}^{(k-1)})|^2\right]\leqslant L\{|\mathbf{x}^{(k-1)}-\mathbf{y}^{(k-1)}|^2+|x-y|^2\}.$$

Our main result of the section is the following theorem.

Theorem 4.2. *Assume* (A1) *and* (A4), *and that* $m=1$. *Let* $t\in[t_{k-1},t_k)$, $\mathbf{x}^{(k-1)}\in\mathbb{R}^{(k-1)n}$, *and denote* $(Y^{t,x}(\mathbf{x}^{(k-1)}),Z^{t,x}(\mathbf{x}^{(k-1)}))$ *to be the solution of* (4.1). *Define* $u^k(t,\mathbf{x}^{(k-1)},x)=Y_t^{t,x}(\mathbf{x}^{(k-1)})$. *Then, the function* $(t,x)\mapsto u^k(t,\mathbf{x}^{(k-1)},x)$ *is the unique viscosity solution of the following semilinear PDE*

$$\begin{aligned}
&\frac{\partial u^k}{\partial t}(t,\mathbf{x}^{(k-1)},x)+\mathscr{L}u^k(t,\mathbf{x}^{(k-1)},x)\\
&\quad+f(t,\mathbf{x}^{(k-1)},\underbrace{x,\ldots,x}_{N-k+1},u^k(t,\mathbf{x}^{(k-1)},x),u_x^k(t,\mathbf{x}^{(k-1)},x)\sigma(t,x))=0,\\
&u^k(t_k,\mathbf{x}^{(k-1)},x)=u^{k+1}(t_k,\mathbf{x}^{(k-1)},x,x),
\end{aligned}\tag{4.2}$$

where

$$u^{N+1}(T,\mathbf{x}^{(N-1)},x,x)\triangleq g(T,\mathbf{x}^{(N-1)},x).$$

Proof. To prove the theorem, we first consider the case when $t\in[t_{N-1},t_N]$. In this case, FSDE (3.10) and BSDE (3.12) become a usual Markovian (decoupled) FBSDE (with parameter $\mathbf{x}^{(N-1)}$):

$$\begin{aligned}
X_s&=x+\int_t^s b(r,X_r)\,\mathrm{d}r+\int_t^s\sigma(r,X_r)\,\mathrm{d}W_r,\\
Y_s&=g(\mathbf{x}^{(N-1)},X_T)+\int_s^T f(r,\mathbf{x}^{(N-1)},X_r,Y_r,Z_r)\,\mathrm{d}r-\int_s^T Z_r\,\mathrm{d}W_r.
\end{aligned}\tag{4.3}$$

The conclusion then follows from the well-known result of Pardoux and Peng (1992) and moreover,

$$Y_s^{t,x}(\mathbf{x}^{(N-1)})=u^N(s,\mathbf{x}^{(N-1)},X_s^{t,x}),\qquad s\in[t,t_N].$$

Next, we assume that $t\in[t_{N-2},t_{N-1}]$, and $s\in[t,t_{N-1}]$. We note that this step is the key, as all other steps can be argued inductively in a similar way.

Let us first fix $\mathbf{x}^{(N-2)}$, and write BSDE (4.1) as

$$\begin{aligned}
Y_s^{t,x}(\mathbf{x}^{(N-2)})&=Y_{t_{N-1}}^{t,x}(\mathbf{x}^{(N-2)})\\
&\quad+\int_s^{t_{N-1}} f(r,\mathbf{x}^{(N-2)},X_r^{t,x},X_r^{t,x},Y_r^{t,x}(\mathbf{x}^{(N-2)}),Z_r^{t,x}(\mathbf{x}^{(N-2)}))\,\mathrm{d}r\\
&\quad-\int_s^{t_{N-1}} Z_r^{t,x}(\mathbf{x}^{(N-2)})\,\mathrm{d}W_r.
\end{aligned}$$

We shall prove the following two assertions:

(i) P-almost surely,

$$Y^{t,x}_{t_{N-1}}(\mathbf{x}^{(N-2)}) = u^N(t_{N-1},\mathbf{x}^{(N-2)},X^{t,x}_{t_{N-1}},X^{t,x}_{t_{N-1}}); \tag{4.4}$$

(ii) the function $(\mathbf{x}^{(N-2)},x)\mapsto u^N(t_{N-1},\mathbf{x}^{(N-2)},x,x)$ is Lipschitz.

It is clear that (ii) follows directly from Lemma 4.1. To see (i), let us recall that $(Y^{t,x}_s(\mathbf{x}^{(N-2)}),Z^{t,x}_s(\mathbf{x}^{(N-2)}))$ satisfies the following BSDE on $[t_{N-1},t_N]$:

$$\begin{aligned}Y_s &= g(\mathbf{x}^{(N-2)},X^{t,x}_{t_{N-1}},X^{t,x}_T)\\ &\quad+\int_s^T f(r,\mathbf{x}^{(N-2)},X^{t,x}_{t_{N-1}},X^{t,x}_r,Y_r,Z_r)\,\mathrm{d}r-\int_s^T Z_r\,\mathrm{d}W_r.\end{aligned} \tag{4.5}$$

Using the Markovian property of the process X, this BSDE can be rewritten as

$$\begin{aligned}Y_s &= g(\mathbf{x}^{(N-2)},X^{t,x}_{t_{N-1}},X_T^{t_{N-1},X^{t,x}_{t_{N-1}}})\\ &\quad+\int_s^T f(r,\mathbf{x}^{(N-2)},X^{t,x}_{t_{N-1}},X_r^{t_{N-1},X^{t,x}_{t_{N-1}}},Y_r,Z_r)\,\mathrm{d}r-\int_s^T Z_r\,\mathrm{d}W_r.\end{aligned} \tag{4.6}$$

Once again, consider for P-a.e. $\omega\in\Omega$, the probability space $(\Omega,\mathscr{F},P^{\omega}_{N-1})$, where P^{ω}_{N-1} is the regular conditional probability $P(\cdot|\mathscr{F}_{t_{N-1}})(\omega)$. Using the conclusion on $[t_{N-1},t_N]$ we know that for P-a.e. $\omega\in\Omega$,

$$Y^{t,x}_s(\mathbf{x}^{(N-2)}) = u^N(s,\mathbf{x}^{(N-2)},X^{t,x}_{t_{N-1}}(\omega),X_s^{t_{N-1},X^{t,x}_{t_{N-1}}(\omega)}) \tag{4.7}$$

for all $s\in[t_{N-1},t_N]$, P^{ω}_{N-1}-a.s. But then, denoting $\hat{X}^{t,x,\omega,N}=X^{t_{N-1},X^{t,x}_{t_{N-1}}(\omega)}$ we have

$$\begin{aligned}&P\{\omega' : Y^{t,x}_s(\mathbf{x}^{(N-2)})(\omega')=u^N(s,\mathbf{x}^{(N-2)},X^{t,x}_{t_{N-1}}(\omega'),X^{t,x}_s(\omega'))\}\\ &=\int_\Omega P^{\omega}_{N-1}\{\omega' : Y^{t,x}_s(\mathbf{x}^{(N-2)})(\omega')= u^N(s,\mathbf{x}^{(N-2)},X^{t,x}_{t_{N-1}}(\omega),\hat{X}^{t,x,\omega,N}_s(\omega'))\}P(\mathrm{d}\omega)\\ &=1\end{aligned} \tag{4.8}$$

for all $s\in[t_{N-1},t_N]$. Hence,

$$Y^{t,x}_s(\mathbf{x}^{(N-2)}) = u^N(s,\mathbf{x}^{(N-2)},X^{t,x}_{t_{N-1}},X^{t,x}_s)\quad \forall s\in[t_{N-1},t_N],\ P\text{-a.s.}$$

Letting $s=t_{N-1}$ we derive (i).

Now replacing $g(\mathbf{x}^{(N-1)},x)$ by $u^N(t_{N-1},\mathbf{x}^{(N-2)},x,x)$ and $f(t,\mathbf{x}^{(N-1)},x,y,z)$ by $f(t,\mathbf{x}^{(N-2)},x,x,y,z)$ in (4.3), we can apply the "classical" nonlinear Feynman–Kac formula of Pardoux and Peng (1992) on the interval $[t_{N-2},t_{N-1}]$ to obtain that, for $t\in[t_{N-2},t_{N-1}]$,

$$u^{N-1}(t,\mathbf{x}^{(N-2)},x)\triangleq Y^{t,x}_t(\mathbf{x}^{(N-2)})$$

is the unique viscosity solution to PDE (4.2) with $k=N-1$.

Repeating the same arguments to the intervals $[t_{k-1}, t_k]$ for $k = N-2, N-3, \ldots, 1$, we complete the proof. $\square$

5. A measurable selection theorem

In the rest of the paper, we turn our attention to the existence of adapted solutions to BSDEs with continuous coefficients. Therefore, from now on we shall assume that the coefficients b, σ, f and g satisfy only (A2) and (A3). As an important tool in the proof, as well as an interesting problem in its own right, we first study a measurable selection problem which we now describe. Consider, for example, the following BSDE over the interval $[t_{N-1}, T]$:

$$Y_t = g(\mathbf{x}^{(N-1)}, X_T^{0,x}) + \int_t^T f(r, \mathbf{x}^{(N-1)}, X_r^{0,x}, Y_r, Z_r)\,\mathrm{d}r - \int_t^T Z_r\,\mathrm{d}W_r, \tag{5.1}$$

where $t_{N-1} \leqslant t \leqslant T$, and $\mathbf{x}^{(N-1)} \in \mathbb{R}^{(N-1)n}$. Suppose that there exists a pair of functions $(u^N, v^N): [t_{N-1}, T] \times \mathbb{R}^{(N-1)n} \times \mathbb{R}^n \mapsto \mathbb{R}^m \times \mathbb{R}^{m\times n}$ satisfying

(C-1) $\forall \mathbf{x}^{(N-1)} \in \mathbb{R}^{(N-1)n}$, *the mapping* $(t, y) \mapsto (u^N(t, \mathbf{x}^{(N-1)}, y), v^N(t, \mathbf{x}^{(N-1)}, y))$ *is* $\mathscr{B}([t_{N-1}, T] \times \mathbb{R}^n)$*-measurable.*

(C-2) *there exists a (Borelian) null-set* $A_{N-1} \subseteq \mathbb{R}^{(N-1)n}$, *such that for all* $\mathbf{x}^{(N-1)} \notin A_{N-1}$, *the processes*

$$\begin{aligned} Y_t^{x,N}(\mathbf{x}^{(N-1)}) &\triangleq u^N(t, \mathbf{x}^{(N-1)}, X_t^{0,x}) \\ Z_t^{x,N}(\mathbf{x}^{(N-1)}) &\triangleq v^N(t, \mathbf{x}^{(N-1)}, X_t^{0,x})\sigma(t, X_t^{0,x}), \end{aligned} \qquad t \in [t_{N-1}, T], \tag{5.2}$$

is an adapted solution to BSDE (5.1).

Our question is whether such a pair of functions can be chosen so that they are both jointly measurable in $(t, \mathbf{x}^{(N-1)}, y)$. The existence of such a "version" is essential for us to construct the adapted solution on the subsequent intervals $[t_{N-2}, t_{N-1}]$, $[t_{N-3}, t_{N-2}], \ldots$, whence a "global" adapted solution.

Note that our proof of the measurable selection applies to all intervals $[t_{k-1}, t_k]$ for $k \geqslant 2$. (In fact, on $[0, t_1]$ some of the arguments may fail because the density function $p(0, x; t, y)$ no longer has a positive lower bound!) However, note that the measurable selections are only needed for $[t_1, t_2], \ldots, [t_{N-1}, T]$, where the parameters $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(N-1)}$ are present. Namely the interval $[0, t_1]$ is not our concern. Therefore, without loss of generality we shall study only the case when $k = N$. Let us begin by taking a closer look at the functions (u^N, v^N) satisfying (C-1) and (C-2).

Let us denote $C_N = \|g\|_\infty + \|f\|_\infty (T - t_{N-1})$, and introduce the following class of functions

$$\mathscr{H}_{\mathrm{loc}}^{N,n,m\times n} = \{v \in L^2_{\mathrm{loc}}([t_{N-1}, T] \times \mathbb{R}^n; \mathbb{R}^{m\times n}) \mid |||v|||_N \leqslant C_N\}, \tag{5.3}$$

where

$$|||v|||_N^2 \triangleq \sup_{x\in\mathbb{R}^n} \int_{t_{N-1}}^T \int_{\mathbb{R}^n} |v(s,z)\sigma(s,z)|^2 p(0, x; s, z)\,\mathrm{d}s\,\mathrm{d}z. \tag{5.4}$$

Then, it is clear that $\mathscr{H}^{N,n,m\times n}_{\rm loc}$ is a closed subspace of $L^2_{\rm loc}([t_{N-1},T]\times\mathbb{R}^n;\mathbb{R}^{m\times n})$, and $(\mathscr{H}^{N,n,m\times n}_{\rm loc},\ |||\cdot|||_N)$ is a Banach space. In what follows, we shall always use $\mathscr{H}^{N,n,m\times n}_{\rm loc}$ in the latter sense. We have the following lemma.

Lemma 5.1. *Assume* (A2) *and* (A3). *Suppose that there exists a pair of functions* $(u^N,v^N):[t_{N-1},T]\times\mathbb{R}^{(N-1)n}\times\mathbb{R}^n\mapsto\mathbb{R}^m\times\mathbb{R}^{m\times n}$ *satisfying* (C-1) *and* (C-2). *Then for all* $\mathbf{x}^{(N-1)}\notin A_{N-1}$, *the null set in* (C-2), *the mapping*

$$(t,y)\mapsto(u^N(t,\mathbf{x}^{(N-1)},y),v^N(t,\mathbf{x}^{(N-1)},y))$$

must belong to the space $C_b([t_{N-1},T)\times\mathbb{R}^n;\mathbb{R}^m)\otimes\mathscr{H}^{N,n,m\times n}_{\rm loc}$.

Proof. Recall that we denoted (see Lemma 2.1) $\mu(t,x;s,\mathrm{d}y)=P\{X^{t,x}_s\in\mathrm{d}y\}$ to be the transition probability of $X^{t,x}$ and $p(t,x;s,y)$ $(s>t)$ to be its density.

We first fix any $t\in[t_{N-1},T)$, and take conditional expectation $E\{\cdot|X^{0,x}_t=y\}$ on both sides of (5.1). Using the relations in assumption (C-2) and the Markovian property of process $X^{0,x}$ we see that

$$\begin{aligned}
&u^N(t,\mathbf{x}^{(N-1)},y)\\
&=E\{Y^{x,N}_t(\mathbf{x}^{(N-1)})|X^{0,x}_t=y\}\\
&=\int_{\mathbb{R}^n}g(\mathbf{x}^{(N-1)},z)p(t,y;T,z)\,\mathrm{d}z\\
&\quad+\int_t^T\int_{\mathbb{R}^n}f(s,\mathbf{x}^{(N-1)},z,u^N(s,\mathbf{x}^{(N-1)},z),v^N(s,\mathbf{x}^{(N-1)},z)\sigma(s,z))\\
&\quad\times p(t,y;s,z)\,\mathrm{d}s\,\mathrm{d}z.
\end{aligned}\tag{5.5}$$

Recall from Lemma 2.1 that the transition density $p(t,y;s,z)$ is continuous in (t,y), and note that f and g are bounded by (A2) and (A3), we see that $u^N(\cdot,\mathbf{x}^{(N-1)},\cdot)\in C_b([t_{N-1},T)\times\mathbb{R}^n)$. In fact, from (5.5) we deduce easily that

$$\begin{aligned}
\|u(\cdot,\mathbf{x}^{(N-1)},\cdot)\|_\infty&=\sup_{(t,y)\in[t_{N-1},T]\times\mathbb{R}^n}E\{|Y^{x,N}_t(\mathbf{x}^{(N-1)})||X^{0,x}_t=y\}\\
&\leqslant\|g\|_\infty+\|f\|_\infty(T-t_{N-1})=C_N.
\end{aligned}\tag{5.6}$$

To see that $v^N(\cdot,\mathbf{x}^{(N-1)},\cdot)\in\mathscr{H}^{N,n,m\times n}_{\rm loc}$, we first note that being an adapted solution to BSDE (5.1), $(Y^{x,N}(\mathbf{x}^{(N-1)}),Z^{x,N}(\mathbf{x}^{(N-1)}))$ must satisfy

$$\begin{aligned}
&E\left\{|Y^{x,N}_{t_{N-1}}(\mathbf{x}^{(N-1)})|^2+\int_{t_{N-1}}^T|Z^{x,N}_s(\mathbf{x}^{(N-1)})|^2\,\mathrm{d}s\right\}\\
&=E\left|Y^{x,N}_{t_{N-1}}(\mathbf{x}^{(N-1)})+\int_{t_{N-1}}^TZ^{x,N}_s(\mathbf{x}^{(N-1)})\,\mathrm{d}W_s\right|^2\\
&=E\left|g(\mathbf{x}^{(N-1)},X^{0,x}_T)+\int_{t_{N-1}}^Tf(r,\mathbf{x}^{(N-1)},X^{0,x}_r,Y^{x,N}_r(\mathbf{x}^{(N-1)}),Z^{x,N}_r(\mathbf{x}^{(N-1)}))\,\mathrm{d}r\right|^2\\
&\leqslant C_N^2.
\end{aligned}\tag{5.7}$$

In light of assumption (C-2) we then have

$$E\left\{\int_{t_{N-1}}^{T}|Z_s^{x,N}(\mathbf{x}^{(N-1)})|^2\,\mathrm{d}s\right\}$$

$$=\int_{t_{N-1}}^{T}\int_{\mathbb{R}^n}|v^N(s,\mathbf{x}^{(N-1)},y)\sigma(s,y)|^2\,p(0,x;s,y)\,\mathrm{d}s\,\mathrm{d}y\leqslant C_N^2,\qquad\forall x\in\mathbb{R}^n.$$

Thus $|||v^N(\cdot,\mathbf{x}^{(N-1)},\cdot)|||_N\leqslant C_N$. Furthermore, applying the density estimate (for p) in Lemma 2.1 to the above inequality we obtain that, for each $x\in\mathbb{R}^n$,

$$\int_{t_{N-1}}^{T}\int_{\mathbb{R}^n}|v^N(s,\mathbf{x}^{(N-1)},y)\sigma(s,y)|^2\frac{m}{s^{n/2}}\exp\left\{-\frac{\lambda|y-x|^2}{2s}\right\}\mathrm{d}s\,\mathrm{d}y\leqslant C_N^2. \tag{5.8}$$

Consequently, one must have, for every compact set $K\subset\mathbb{R}^n$,

$$\int_{t_{N-1}}^{T}\int_{K}|v^N(s,\mathbf{x}^{(N-1)},y)|^2\,\mathrm{d}s\,\mathrm{d}y<\infty,$$

proving that $v^N(\cdot,\mathbf{x}^{(N-1)},\cdot)\in L^2_{\mathrm{loc}}([t_{N-1},T]\times\mathbb{R}^n;\mathbb{R}^{m\times n})$, whence $\mathscr{H}_{\mathrm{loc}}^{N,n,m\times n}$. $\square$

Now let us introduce a device that is useful for our measurable selection. For notational simplicity in what follows we denote $\mathscr{C}_b^{N,n,m}=C_b([t_{N-1},T)\times\mathbb{R}^n;\mathbb{R}^m)$. Also, for any $(\mathbf{y},u,v)\in\mathbb{R}^{(N-1)n}\times\mathscr{C}_b^{N,n,m}\times\mathscr{H}_{\mathrm{loc}}^{N,n,m\times n}$, we denote

$$F[\mathbf{y},u,v](t,X_t^{0,x})\triangleq f(t,\mathbf{y},X_t^{0,x},u(t,X_t^{0,x}),v(t,X_t^{0,x})\sigma(t,X_t^{0,x})).$$

Now let A_{N-1} be the exceptional set in condition (C-2). Consider a mapping $G_N:\mathbb{R}^{(N-1)n}\times\mathscr{C}_b^{N,n,m}\times\mathscr{H}_{\mathrm{loc}}^{N,n,m\times n}\to L^0(\mathbb{R}^n;\mathbb{R}_+)$ defined by

$$G_N[\mathbf{y},u,v](x)\triangleq E\left\{\int_{t_{N-1}}^{T}\left|u(t,X_t^{0,x})-g(\mathbf{y},X_T^{0,x})-\int_t^T F[\mathbf{y},u,v](r,X_r^{0,x})\,\mathrm{d}r\right.\right.$$

$$\left.\left.+\int_t^T v(r,X_r^{0,x})\sigma(r,X_r^{0,x})\,\mathrm{d}W_r\right|\,\mathrm{d}t\right\}1_{A_{N-1}}(\mathbf{y}),\qquad x\in\mathbb{R}^n.$$

The following two lemmas give the main properties of the function G_N.

Lemma 5.2. *Assume* (A2) *and* (A3). *Then the mapping $x\mapsto G_N[\mathbf{y},u,v](x)$ is continuous for each fixed $(\mathbf{y},u,v)$, and $\|G_N[\mathbf{y},u,v](\cdot)\|_{C(\mathbb{R}^n,\mathbb{R}_+)}\leqslant 3C_N(T-t_{N-1})$.*

Proof. The bound for G_N is a direct consequence of estimates (5.6) and (5.7). We prove the continuity of $G_N[\mathbf{y},u,v](\cdot)$. To begin with, let us choose, for $\varepsilon>0$, continuous functions g^ε, F^ε and v^ε, such that

(a) g^ε and F^ε are uniformly bounded (uniformly in ε as well),
(b) $\|g^\varepsilon(\mathbf{y},\cdot)-g(\mathbf{y},\cdot)\|_{L^2_{\mathrm{loc}}(\mathbb{R}^n;\mathbb{R}^m)}\to 0$, $\|F^\varepsilon-F[\mathbf{y},u,v]\|_{L^2_{\mathrm{loc}}([t_{N-1},T]\times\mathbb{R}^n;\mathbb{R}^m)}\to 0$,
(c) $\|v^\varepsilon-v\|_{L^2_{\mathrm{loc}}([t_{N-1},T]\times\mathbb{R}^n;\mathbb{R}^{m\times n})}\to 0$.

Now assume that $x \to x_0$. Without loss of generality, we may assume that $x \in B(x_0; a)$ for some $a > 0$, where $B(x_0; a)$ denotes the closed ball centered at x_0, and with radius a. For each $M > 0$, we define

$$\tau_M = \inf\{s \geqslant t_{N-1} \mid |X_s^{0,x}| + |X_s^{0,x_0}| > M\} \wedge T.$$

Then it is clear that for some constant $C' > 0$,

$$P\{\tau_M < T\} = P\left\{\sup_{t_{N-1} \leqslant s \leqslant T} (|X_s^{0,x}| + |X_s^{0,x_0}|) > M\right\} \leqslant \frac{C'(1 + |x|^2 + |x_0|^2)}{M^2}.$$

Therefore $\lim_{M \to \infty} P\{\tau_M < T\} = 0$, and the limit is uniform for all $x \in B(x_0; a)$. Further, for notational simplicity, let us denote for any $\xi \in L^2(\mathbf{F}, \Omega; C([0,T]; \mathbb{R}^n))$ and $(\mathbf{y}, u, v) \in \mathbb{R}^{(N-1)n} \times \mathscr{C}_b^{N,n,m} \times \mathscr{H}_{\text{loc}}^{N,n,m\times n}$,

$$\Gamma[\mathbf{y}, u, v](t, \xi) \triangleq g(\mathbf{y}, \xi_T) + \int_t^T F[\mathbf{y}, u, v](r, \xi_r)\,\mathrm{d}r - \int_t^T v(r, \xi_r)\sigma(r, \xi_r)\,\mathrm{d}W_r,$$

Then we can write

$$G_N[\mathbf{y}, u, v](x) = E\left\{\int_{t_{N-1}}^T |u(t, X_t^{0,x}) - \Gamma[\mathbf{y}, u, v](t, X^{0,x})|\,\mathrm{d}t\right\} 1_{A_{N-1}}(\mathbf{y}) \tag{5.9}$$

and consequently,

$$\begin{aligned}
&|G_N[y, u, v](x) - G_N[y, u, v](x_0)| \\
&\quad \leqslant E\left\{\int_{t_{N-1}}^T \{|u(t, X_t^{0,x}) - u(t, X_t^{0,x_0})| \right. \\
&\qquad \left. + |\Gamma[\mathbf{y}, u, v](t, X^{0,x}) - \Gamma[\mathbf{y}, u, v](t, X^{0,x_0})|\}\,\mathrm{d}t\right\}.
\end{aligned} \tag{5.10}$$

A simple application of Lebesgue's Dominated Convergence Theorem shows that

$$\lim_{x \to x_0} E\left\{\int_{t_{N-1}}^T |u(t, X_t^{0,x}) - u(t, X_t^{0,x_0})|\,\mathrm{d}t\right\} = 0. \tag{5.11}$$

Furthermore, note that (suppressing $(\mathbf{y}, u, v)$ in $\Gamma[\cdots]$)

$$\begin{aligned}
&E\left\{\int_{t_{N-1}}^T |\Gamma(t, X^{0,x}) - \Gamma(t, X^{0,x_0})|\,\mathrm{d}t\right\} \\
&\quad = E\left\{\int_{t_{N-1}}^T \{\mathbf{1}_{\{\tau_M < T\}} |\Gamma(t, X^{0,x}) - \Gamma(t, X^{0,x_0})| \right. \\
&\qquad \left. + \mathbf{1}_{\{\tau_M = T\}} |\Gamma(t, X^{0,x}) - \Gamma(t, X^{0,x_0})|\}\,\mathrm{d}t\right\} \\
&\quad \triangleq I_M^{N,1} + I_M^{N,2},
\end{aligned} \tag{5.12}$$

where $I_M^{N,1}$ and $I_M^{N,2}$ are defined in an obvious way. Since $\lim_{M\to\infty} P\{\tau_M < T\} = 0$, uniformly for $x \in B(x_0; a)$, we have

$$\begin{aligned}
I_M^{N,1} &\leqslant E\left\{\int_{t_{N-1}}^{T} \mathbf{1}_{\{\tau_M<T\}}\{|\Gamma(t,X^{0,x})| + |\Gamma(t,X^{0,x_0})|\}\,\mathrm{d}t\right\}\\
&\leqslant P\{\tau_M < T\}^{1/2}\int_{t_{N-1}}^{T} E\{(|\Gamma(t,X^{0,x})| + |\Gamma(t,X^{0,x_0})|)^2\}^{1/2}\,\mathrm{d}t \to 0,\\
&\qquad \text{as } M\to\infty,
\end{aligned}$$

uniformly for $x \in B(x_0; a)$. Thus for any $\delta > 0$, we can choose an $M = M(\delta) > 0$ such that $I_M^{N,1} < \delta$. Fixing this M we now estimate $I_M^{N,2}$. Bearing the indicator function $\mathbf{1}_{\{\tau_M=T\}}$ in mind, we have

$$\begin{aligned}
I_M^{N,2} &\leqslant E\{|g(\mathbf{y},X^{0,x}_{\tau_M}) - g^{\varepsilon}(\mathbf{y},X^{0,x}_{\tau_M})| + |g(\mathbf{y},X^{0,x_0}_{\tau_M}) - g^{\varepsilon}(\mathbf{y},X^{0,x_0}_{\tau_M})|\}(T-t_{N-1})\\
&\quad + E\{|g^{\varepsilon}(\mathbf{y},X^{0,x}_{T}) - g^{\varepsilon}(\mathbf{y},X^{0,x_0}_{T})|\}(T-t_{N-1})\\
&\quad + \int_{t_{N-1}}^{T} E\left\{\left|\int_{t\wedge\tau_M}^{\tau_M}(F[\mathbf{y},u,v]-F^{\varepsilon})(r,X_r^{0,x})\,\mathrm{d}r\right|\right\}\,\mathrm{d}t\\
&\quad + \int_{t_{N-1}}^{T} E\left\{\left|\int_{t\wedge\tau_M}^{\tau_M}(F[\mathbf{y},u,v]-F^{\varepsilon})(r,X_r^{0,x_0})\,\mathrm{d}r\right|\right\}\,\mathrm{d}t\\
&\quad + \int_{t_{N-1}}^{T} E\left\{\int_{t_{N-1}}^{T}|F^{\varepsilon}(r,X_r^{0,x}) - F^{\varepsilon}(r,X_r^{0,x_0})|\,\mathrm{d}r\right\}\,\mathrm{d}t\\
&\quad + C_{N,M,T}\int_{t_{N-1}}^{T} E\left\{\int_{t\wedge\tau_M}^{\tau_M}|(v-v^{\varepsilon})\sigma(r,X_r^{0,x})|^2\,\mathrm{d}r\right\}^{1/2}\,\mathrm{d}t\\
&\quad + C_{N,M,T}\int_{t_{N-1}}^{T} E\left\{\int_{t\wedge\tau_M}^{\tau_M}|(v-v^{\varepsilon})\sigma(r,X_r^{0,x_0})|^2\,\mathrm{d}r\right\}^{1/2}\,\mathrm{d}t\\
&\quad + C_{N,M,T}\int_{t_{N-1}}^{T} E\left\{\int_{t_{N-1}}^{T}|(v^{\varepsilon}\sigma)(r,X_r^{0,x}) - (v^{\varepsilon}\sigma)(r,X_r^{0,x_0})|^2\,\mathrm{d}r\right\}^{1/2}\,\mathrm{d}t.
\end{aligned}$$

Here we used the Burkholder's inequality to treat the stochastic integrals, and $C_{T,N,M}$ denotes a generic constant depending only on T, N, M, and $B(x_0; a)$, which is allowed to vary from line to line. Clearly, for fixed M we can choose ε small enough so that

$$E|g(\mathbf{y},X^{0,x}_{\tau_M}) - g^{\varepsilon}(\mathbf{y},X^{0,x}_{\tau_M})| \leqslant \int_{|y|\leqslant M}|g(\mathbf{y},y) - g^{\varepsilon}(\mathbf{y},y)|^2\,p(0,x;T,y)\,\mathrm{d}y$$

$$\leqslant C_{T,N,M}\int_{|y|\leqslant M}|g(\mathbf{y},y)-g^{\varepsilon}(\mathbf{y},y)|^2\,\mathrm{d}y$$
$$<\frac{\delta}{T-t_{N-1}}.$$

In fact, we can choose ε small enough so that all the following hold:

$$E|g(\mathbf{y},X_{\tau_M}^{0,x_0})-g^{\varepsilon}(\mathbf{y},X_{\tau_M}^{0,x_0})|<\frac{\delta}{T-t_{N-1}},$$

$$\int_{t_{N-1}}^{T}E\left\{\left|\int_{t\wedge\tau_M}^{\tau_M}(F[\mathbf{y},u,v]-F^{\varepsilon})(r,X_r^{0,x})\,\mathrm{d}r\right|\right\}\,\mathrm{d}t<\delta,$$

$$\int_{t_{N-1}}^{T}E\left\{\left|\int_{t\wedge\tau_M}^{\tau_M}(F[\mathbf{y},u,v]-F^{\varepsilon})(r,X_r^{0,x_0})\,\mathrm{d}r\right|\right\}\,\mathrm{d}t<\delta,$$

$$\int_{t_{N-1}}^{T}E\left\{\left|\int_{t\wedge\tau_M}^{\tau_M}|(v-v^{\varepsilon})\sigma(r,X_r^{0,x})|^2\,\mathrm{d}r\right\}^{1/2}\,\mathrm{d}t<\frac{\delta}{C_{N,M,T}},$$

$$\int_{t_{N-1}}^{T}E\left\{\left|\int_{t\wedge\tau_M}^{\tau_M}|(v-v^{\varepsilon})\sigma(r,X_r^{0,x_0})|^2\,\mathrm{d}r\right\}^{1/2}\,\mathrm{d}t<\frac{\delta}{C_{N,M,T}}.$$

Consequently, we have

$$\begin{aligned}I_M^{N,2}&\leqslant 6\delta+E\{|g^{\varepsilon}(\mathbf{y},X_T^{0,x})-g^{\varepsilon}(\mathbf{y},X_T^{0,x_0})|\}(T-t_{N-1})\\&\quad+\int_{t_{N-1}}^{T}E\left\{\int_{t_{N-1}}^{T}|F^{\varepsilon}(r,X_r^{0,x})-F^{\varepsilon}(r,X_r^{0,x_0})|\,\mathrm{d}r\right\}\,\mathrm{d}t\\&\quad+C_{N,M,T}\int_{t_{N-1}}^{T}E\left\{\int_{t_{N-1}}^{T}|(v^{\varepsilon}\sigma)(r,X_r^{0,x})-(v^{\varepsilon}\sigma)(r,X_r^{0,x_0})|^2\,\mathrm{d}r\right\}^{1/2}\,\mathrm{d}t.\end{aligned}$$

Since for any $t\in[t_{N-1},T]$, $X_t^{0,x}$ converges in probability to X_t^{0,x_0}, one can apply the Dominated Convergence Theorem again to show that the last three terms on the right-hand side above all converge to 0 as $x\to x_0$ (while M and ε being fixed!). In other words, we have $\overline{\lim}_{x\to x_0}I_M^{N,2}\leqslant 6\delta$. Plugging this, as well as $I_M^{N,1}<\delta$, into (5.12) we have

$$\overline{\lim_{x\to x_0}}E\left\{\int_{t_{N-1}}^{T}|\Gamma(t,X^{0,x})-\Gamma(t,X^{0,x_0})|\,\mathrm{d}t\right\}<7\delta.$$

Since δ is arbitrary, combining this with (5.10) and (5.11) we obtain that

$$\lim_{x\to x_0}|G_N[y,u,v](x)-G_N[y,u,v](x_0)|=0.$$

That is, the mapping G_N can be considered as a functional from $\mathscr{C}_b^{N,n,m} \times \mathscr{H}_{\mathrm{loc}}^{N,n,m\times d}$ to $C(\mathbb{R}^n;\mathbb{R}_+)$, proving the lemma. □

The next lemma concerns the measurability of the mapping G_N. Let us endow the space $C(\mathbb{R}^n,\mathbb{R}_+)$ with a metric, $\rho(\cdot,\cdot)$, that is equivalent to the uniform convergence on compacts (e.g., $\rho(\varphi_1,\varphi_2) \triangleq \sum_{k=1}^{\infty}(\|\varphi_1-\varphi_2\|_k \wedge 1)/2^k$, where $\|\varphi\|_k = \sup_{|x|\leqslant k}|\varphi(x)|$). In what follows all the measurability involving $C(\mathbb{R}^n;\mathbb{R}_+)$ will be in terms of the topological Borel field of $C(\mathbb{R}^n;\mathbb{R}_+)$ under this metric. We denote $\|\varphi\|_C = \rho(\varphi,0)$.

Lemma 5.3. *Assume* (A2) *and* (A3). *The mapping* $(\mathbf{y},u,v)\mapsto G_N[\mathbf{y},u,v](\cdot)$ *is jointly measurable, as a function from* $\mathbb{R}^{(N-1)n}\times\mathscr{C}_b^{N,n,m}\times\mathscr{H}_{\mathrm{loc}}^{N,n,m\times n}$ *to* $C(\mathbb{R}^n;\mathbb{R}_+)$.

Proof. We first note that for fixed (u,v), the mapping $\mathbf{y}\mapsto G_N(\mathbf{y},u,v)(\cdot)$ is obviously (Borel) measurable in $\mathbf{y}\in\mathbb{R}^{(N-1)n}$. We need only check that for each fixed $\mathbf{y}\in\mathbb{R}^{(N-1)n}$, the mapping $(u,v)\mapsto G_N[\mathbf{y},u,v]$ is continuous.

To this end, let $(u^\ell,v^\ell)\to(u,v)$ in $\mathscr{C}_b^{N,n,m}\times\mathscr{H}_{\mathrm{loc}}^{N,n,m\times n}$, and fix $\mathbf{y}\notin A_{N-1}$. Recall (5.9) and the definitions of $F[\mathbf{y},u,v]$ and $\Gamma[\mathbf{y},u,v]$ we see that for any $x\in\mathbb{R}^n$ it holds that

$$
\begin{aligned}
&|G_N(\mathbf{y},u^\ell,v^\ell)(x)-G_N(\mathbf{y},u,v)(x)|\\
&\quad\leqslant E\left\{\int_{t_{N-1}}^T |u^l(t,X_t^{0,x})-u(t,X_t^{0,x})|\right.\\
&\qquad\left.+|\Gamma[\mathbf{y},u^\ell,v^\ell](t,X^{0,x})-\Gamma[\mathbf{y},u,v](t,X^{0,x})|\,\mathrm{d}t\right\}\\
&\quad\leqslant (\|u^\ell-u\|_\infty+|||v^\ell-v|||_N)(T-t_{N-1})\\
&\qquad+E\left\{\int_{t_{N-1}}^T\int_t^T |F[\mathbf{y},u^\ell,v^\ell](r,X_r^{0,x})-F[\mathbf{y},u,v](r,X_r^{0,x})|\,\mathrm{d}r\,\mathrm{d}t\right\}.
\end{aligned}
$$

Clearly, by the nature of the metric $\|\cdot\|_C$ we need only show that for any $R>0$,

$$
\lim_{\ell\to\infty}\sup_{|x|\leqslant R} E\left\{\int_{t_{N-1}}^T |F[\mathbf{y},u^\ell,v^\ell](t,X_t^{0,x})-F[\mathbf{y},u,v](t,X_t^{0,x})|\,\mathrm{d}t\right\}=0. \tag{5.13}
$$

To see this, denote $\Delta F^\ell(\mathbf{y};t,y)=F[\mathbf{y},u^\ell,v^\ell](t,y)-F[\mathbf{y},u,v](t,y)$, and note that

$$
E\left\{\int_{t_{N-1}}^T |\Delta F^\ell(\mathbf{y};t,X_t^{0,x})|\,\mathrm{d}t\right\}=\int_{t_{N-1}}^T\int_{\mathbb{R}^n}|\Delta F^\ell(\mathbf{y};t,y)|\,p(0,x;t,y)\,\mathrm{d}t\,\mathrm{d}y.
$$

Since $|\Delta F^\ell(\mathbf{y};t,y)|\leqslant 2\|f\|_\infty$, and for fixed $R>0$ it holds that

$$
\lim_{K\to\infty}\sup_{|x|\leqslant R}\int_{t_{N-1}}^T\int_{|y|>K} p(0,x;t,y)\,\mathrm{d}t\,\mathrm{d}y=0,
$$

we have, for any $\delta > 0$, there exists $K = K(\delta, R) > 0$, such that

$$\sup_{|x|\leqslant R} \int_{t_{N-1}}^{T} \int_{|y|>K} |\Delta F^{\ell}(\mathbf{y}; t, y)|\, p(0,x;t,y)\,\mathrm{d}t\,\mathrm{d}y < \delta.$$

Therefore

$$\sup_{|x|\leqslant R} \int_{t_{N-1}}^{T} \int_{\mathbb{R}^n} |\Delta F^{\ell}(\mathbf{y}; t, y)|\, p(0,x;t,y)\,\mathrm{d}t\,\mathrm{d}y$$

$$\leqslant \delta + \sup_{|x|\leqslant R} \int_{t_{N-1}}^{T} \int_{|y|\leqslant K} |\Delta F^{\ell}(\mathbf{y}; t, y)|\, p(0,x;t,y)\,\mathrm{d}t\,\mathrm{d}y. \tag{5.14}$$

Now denote, for any $\gamma > 0$, $A_{\ell}^{K,\gamma} \triangleq \{(t,y) \,|\, |y| \leqslant K, |v^{\ell}(t,y) - v(t,y)| > \gamma\}$, and let $|A_{\ell}^{K,\gamma}|$ denote the Lebesgue measure of $A_{\ell}^{K,\gamma}$ in $[t_{N-1}, T] \times \mathbb{R}^n$. Using the uniform ellipticity of $\sigma\sigma^*$, thanks to (A3), we have

$$\begin{aligned} |||v^{\ell} - v|||_N^2 &\geqslant \int_{t_{N-1}}^{T} \int_{|y|\leqslant K} c|v^{\ell}(t,y) - v(t,y)|^2 p(0,x;t,y)\,\mathrm{d}t\,\mathrm{d}y \\ &\geqslant c\gamma^2 |A_{\ell}^{K,\gamma}| c_{R,K,N} \quad \forall x \in B(0;R), \end{aligned}$$

where $c_{R,K,N} = \inf_{|x|\leqslant R, |y|\leqslant K, t\in[t_{N-1},T]} p(0,x,t,y) > 0$, and c is the constant in (2.5) of (A3). In other words, v^{ℓ} converges to v in measure on $[t_{N-1}, T] \times B(0;K)$, uniformly for $|x| \leqslant R$. The continuity of f then implies that $\lim_{\ell\to\infty} \Delta F^{\ell}(\mathbf{y};\cdot,\cdot) = 0$ in measure on $[t_{N-1}, T] \times B(0;K)$, uniformly for $|x| \leqslant R$, as well. Consequently, if we denote $C_{R,K,N} = \sup_{|x|\leqslant R, |y|\leqslant K, t\in[t_{N-1},T]} p(0,x;t,y)$, then

$$\sup_{|x|\leqslant R} \int_{t_{N-1}}^{T} \int_{|y|\leqslant K} |\Delta F^{\ell}(\mathbf{y}; t, y)|\, p(0,x;t,y)\,\mathrm{d}t\,\mathrm{d}y$$

$$\leqslant C_{R,K,N} \int_{t_{N-1}}^{T} \int_{|y|\leqslant K} |\Delta F^{\ell}(\mathbf{y}; t, y)|\,\mathrm{d}t\,\mathrm{d}y \to 0, \quad \text{as } \ell \to \infty,$$

thanks to the Dominated Convergence Theorem. This, together with (5.14), leads to (5.13). The proof is complete. □

We are now ready to prove the main result of this section.

Theorem 5.4. *Assume* (A2) *and* (A3), *and suppose that there exists a pair of functions* (u^N, v^N) *satisfying conditions* (C-1) *and* (C-2). *Then there exists a pair of functions* $(\bar{u}^N, \bar{v}^N): [0,T] \times \mathbb{R}^{(N-1)n} \times \mathbb{R}^n \mapsto \mathbb{R}^m \times \mathbb{R}^{m\times n}$, *such that*

(i) *for each* $\mathbf{x}^{(N-1)}$, $\bar{u}^N(\cdot, \mathbf{x}^{(N-1)}, \cdot) \in C_b([t_{N-1}, T) \times \mathbb{R}^n; \mathbb{R}^m)$ *and* $\bar{v}^N(\cdot, \mathbf{x}^{(N-1)}, \cdot) \in \mathscr{H}_{\mathrm{loc}}^{N,n,m\times n}$;

(ii) *there exists a* (*Borelian*) *null set* $A'_{N-1} \subset \mathbb{R}^{(N-1)n}$ *such that for each* $\mathbf{x}^{(N-1)} \notin A'_{N-1}$, *the pair of processes* $(\bar{Y}^{x,N}(\mathbf{x}^{(N-1)}), \bar{Z}^{x,N}(\mathbf{x}^{(N-1)}))$ *defined by*

$$\bar{Y}^{x,N}_t(\mathbf{x}^{(N-1)}) = \bar{u}^N(t, \mathbf{x}^{(N-1)}, X^{0,x}_t), \quad \bar{Z}^{x,N}_t(\mathbf{x}^{(N-1)}) = \bar{v}^N(t, \mathbf{x}^{(N-1)}, X^{0,x}_t)\sigma(t, X^{0,x}_t)$$

is an adapted solution to BSDE (5.1); *and*

(iii) $(\bar{u}^N(\cdot,\cdot,\cdot), \bar{v}^N(\cdot,\cdot,\cdot))$ *is* $\mathscr{B}([t_{N-1},T] \times \mathbb{R}^{(N-1)n} \times \mathbb{R}^n)/\mathscr{B}(\mathbb{R}^m) \otimes \mathscr{B}(\mathbb{R}^{m\times n})$*-jointly measurable.*

Proof. Consider the function $G_N[\mathbf{y}, u, v]$ defined by (5.9). Define the set

$$\mathscr{A}_N \triangleq \{(\mathbf{y}, u, v) : G_N[\mathbf{y}, u, v](x) = 0, \forall x \in \mathbb{R}^n\},$$

then $\mathscr{A}_N$ is a Borel set in $\mathscr{B}(\mathbb{R}^{(N-1)n} \times \mathscr{C}_b^{N,n,m} \times \mathscr{H}^{N,n,m\times n}_{\mathrm{loc}})$. Also, the existence of solution to (5.1) and the relation (5.2) imply that $\mathrm{Proj}_{\mathbb{R}^{(N-1)n}}(\mathscr{A}_N) = \mathbb{R}^{(N-1)n}$. Therefore, applying the measurable selection theorem (see, e.g., Dellacherie and Meyer, 1978, Appendix to III-81 or Bertsekas and Shreve, 1978, Proposition 7.49), we can find a pair of functions

$$(\bar{u}^N, \bar{v}^N) : \mathbb{R}^{(N-1)n} \mapsto \mathscr{C}_b^{N,n,m} \times \mathscr{H}^{N,n,m\times n}_{\mathrm{loc}}$$

and a Lebesgue null set A''_{N-1} such that for any $\mathbf{y} \notin A''_{N-1}$, $G_N(\mathbf{y}, \bar{u}^N(\mathbf{y}), \bar{v}^N(\mathbf{y}))(x) = 0$, for all $x \in \mathbb{R}^n$ and that the mapping $\mathbf{y} \mapsto (\bar{u}^N(\mathbf{y}), \bar{v}^N(\mathbf{y}))$ is $\mathscr{B}(\mathbb{R}^{(N-1)n})/\mathscr{B}(\mathscr{C}_b^{N,n,m} \times \mathscr{H}^{N,n,m\times n}_{\mathrm{loc}})$-measurable. This, together with the definition of the set $\mathscr{A}_N$, proves (i) and (ii) (replacing $\mathbf{y}$ by $\mathbf{x}^{(N-1)}$!).

To see (iii), we note that for each $\mathbf{y}$, the mapping $(t, y) \mapsto \bar{u}^N(t, \mathbf{y}, y)$ is continuous, and for (t, y), the mapping $\mathbf{y} \mapsto \bar{u}^N(t, \mathbf{y}, y)$ is measurable. Thus, $\bar{u}^N(\cdot,\cdot,\cdot)$ is jointly measurable. It remains to show that $\bar{v}^N(\cdot,\cdot,\cdot)$ is also jointly measurable.

To this end, noting that $\lim_{k\to+\infty} \bar{v}^N(\mathbf{y})\mathbf{1}_{\{|y|\leqslant k\}} = \bar{v}^N(\mathbf{y})$, it suffices to prove the joint measurability of $\bar{v}^N$ assuming that $\bar{v}^N(\mathbf{y}) \in L^2([t_{N-1},T] \times \mathbb{R}^n; \mathbb{R}^{m\times n})$. We note that the space $L^2([t_{N-1},T] \times \mathbb{R}^n; \mathbb{R}^{m\times n})$ is a separable Hilbert space. Let $\{\xi^i\}_{i\geqslant 1}$ be a standard orthogonal basis of $L^2([t_{N-1},T] \times \mathbb{R}^n; \mathbb{R}^{m\times n})$, so that the functions $\bar{v}^N$ can be written as $\bar{v}^N(\mathbf{y}) = \sum_{i=1}^\infty b_i(\mathbf{y})\xi^i$, in $L^2([t_{N-1},T] \times \mathbb{R}^n; \mathbb{R}^{m\times n})$ and a.e., where

$$b_i(\mathbf{y}) \triangleq \int_{t_{N-1}}^T \int_{\mathbb{R}^n} \bar{v}^N(\mathbf{y})\xi^i(t, y)\,\mathrm{d}y\,\mathrm{d}t.$$

Clearly, for each i, $b_i(\cdot)$ is $\mathscr{B}(\mathbb{R}^{(N-1)n})/\mathscr{B}(\mathbb{R})$-measurable, thus all the mappings

$$(t, \mathbf{y}, x) \mapsto b_i(\mathbf{y})\xi^i(t, x)$$

are $\mathscr{B}([t_{N-1},T] \times \mathbb{R}^{(N-1)n} \times \mathbb{R}^n)/\mathscr{B}(\mathbb{R}^{m\times n})$-jointly measurable. Thus (iii) follows. □

6. Existence of solutions for BSDEs with continuous coefficients

We now prove the existence of adapted solutions to BSDEs with continuous coefficients in the discrete functional form. To be more precise, we shall consider the following (F)BSDE:

$$\begin{aligned} X_t &= x + \int_0^t b(r,X_r)\,\mathrm{d}r + \int_0^t \sigma(r,X_r)\,\mathrm{d}W_r, \\ Y_t &= g(\mathbf{X}^{(N)}) + \int_t^T f(r,\mathbf{X}_r^{(N)},Y_r,Z_r)\,\mathrm{d}r - \int_t^T Z_r\,\mathrm{d}W_r, \end{aligned} \qquad t \in [0,T]. \tag{6.1}$$

We assume that the coefficients g and f satisfy only (A2), that is, g is only bounded measurable, and f is only bounded and continuous. We assume that b and σ satisfy (A3).

Our plan of attack is the following, we first mollify the coefficients g and f so that the representation results of the previous sections can be applied. We then pass to the limit, in the spirit of the method proposed in Hamadene et al. (1997), to obtain a candidate of the solution. Then, by using the measurable selection theorem established in Section 5 we verify that the candidate solution is indeed what we are looking for. Since the discussion is quite lengthy, we shall split it into several lemmas.

To begin with let us choose a sequence of mollifiers $\{(g^\varepsilon, f^\varepsilon)\}_{\varepsilon>0}$, such that

(i) for all $\varepsilon > 0$, g^ε and f^ε are uniformly bounded (uniformly in ε as well!);
(ii) $g^\varepsilon \in C^\infty(\mathbb{R}^{Nn};\mathbb{R}^m)$ and $f^\varepsilon \in C^\infty([0,T]\times\mathbb{R}^{Nn}\times\mathbb{R}^m\times\mathbb{R}^{m\times d};\mathbb{R}^m)$, with bounded derivatives of all orders; and
(iii) for fixed $(t,\mathbf{x}^{(N)},y,z)\in[0,T]\times\mathbb{R}^{Nn}\times\mathbb{R}^m\times\mathbb{R}^{m\times d}$, it holds that

$$f^\varepsilon(t,\mathbf{x}^{(N)},y,z) \to f(t,\mathbf{x}^{(N)},y,z), \quad \text{as } \varepsilon \to 0 \tag{6.2}$$

and for any $p \geqslant 1$, $g^\varepsilon(\cdot) \to g(\cdot)$, in $L^p_{\mathrm{loc}}(\mathbb{R}^{Nn})$, as $\varepsilon \to 0$.

By a diagonalisation procedure, we can choose a subsequence (still denote by itself), such that g^ε converges to g a.e.

Now for each $\varepsilon > 0$ consider the FBSDE (6.1) with g and f being replaced by g^ε and f^ε. Clearly, the adapted solution exists and is unique, we denote it by $(Y^\varepsilon, Z^\varepsilon)$. Furthermore, since b and σ satisfy (A3), and g^ε and f^ε are both bounded and smooth with derivatives of all orders, it is known that all PDEs in the system (3.3) (with g^ε and f^ε) will have classical solutions (see the remarks at the beginning of Section 3). Therefore, applying the result of Section 3 we can construct a pair of functions $u^\varepsilon : [0,T]\times\mathbb{R}^{Nn} \mapsto \mathbb{R}^m$ and $v^\varepsilon : [0,T]\times\mathbb{R}^{Nn}\mapsto\mathbb{R}^{m\times n}$ such that

$$Y_s^\varepsilon = u^\varepsilon(s,\mathbf{X}_s^{(N)}), \qquad Z_s^\varepsilon = v^\varepsilon(s,\mathbf{X}_s^{(N)})\sigma(s,X_s).$$

The same relation holds when we consider the FSDE (6.1) starting from (t,x) with $t\in[t_{k-1},t_k)$. Denote the corresponding solution by $(Y^{\varepsilon,t,x}(\mathbf{x}^{(k-1)}), Z^{\varepsilon,t,x}(\mathbf{x}^{(k-1)}))$.

Our next step is to look at the limit of the family $(Y^{\varepsilon,t,x}(\mathbf{x}^{(k-1)}), Z^{\varepsilon,t,x}(\mathbf{x}^{(k-1)}))$, as $\varepsilon\to 0$, which is one of the building blocks of the desired adapted solution. The main technicality in this step can be roughly described as follows. Since in general one only knows that the family $\{\bar{Y}^{\varepsilon,\cdot,\cdot}(\mathbf{x}^{(k-1)}), \{\bar{Z}^{\varepsilon,\cdot,\cdot}(\mathbf{x}^{(k-1)})\}$ is precompact for each fixed $\mathbf{x}^{(k-1)}$, a limit point of this family will depend on the choice of the subsequence, whence on $\mathbf{x}^{(k-1)}$, the measurability of the limit points of this family on $\mathbf{x}^{(k-1)}$ thus requires more careful consideration.

We first state a simple lemma to facilitate our discussion.

Lemma 6.1. *Suppose that $g:\mathbb{R}^{Nn} \mapsto \mathbb{R}^m$ is a bounded measurable function, and $\{g^\varepsilon(\cdot)\}$ is a family of smooth mollifiers of g converging to g a.e. Then there exists a (Borelian) null measurable set $A_{N-1} \subset \mathbb{R}^{(N-1)n}$, independent of the initial state x, such that*

$$\lim_{\varepsilon\to 0} E|g^\varepsilon(\mathbf{x}^{(N-1)}, X_T^{0,x}) - g(\mathbf{x}^{(N-1)}, X_T^{0,x})| = 0 \quad \forall \mathbf{x}^{(N-1)} \notin A_{N-1}.$$

Proof. Since g^ε converges to g a.e., i.e.,

$$\int_{\mathbb{R}^{(N-1)n}} \int_{\mathbb{R}^n} \mathbf{1}_{\{\overline{\lim}_{\varepsilon\to 0}|g^\varepsilon(\mathbf{x}^{(N-1)},y_N)-g(\mathbf{x}^{(N-1)},y_N)|>0\}}(\mathbf{x}^{(N-1)}, y_N)\,\mathrm{d}\mathbf{x}^{(N-1)}\,\mathrm{d}y_N = 0.$$

Thus, there exists a (Borelian) null set A_{N-1}, such that $\forall \mathbf{x}^{(N-1)} \notin A_{N-1}$,

$$\int_{\mathbb{R}^n} \mathbf{1}_{\{\overline{\lim}_{\varepsilon\to 0}|g^\varepsilon(\mathbf{x}^{(N-1)},y_N)-g(\mathbf{x}^{(N-1)},y_N)|>0\}}(\mathbf{x}^{(N-1)}, y_N)\,\mathrm{d}y_N = 0.$$

That is, $\overline{\lim}_{\varepsilon\to 0} |g^\varepsilon(\mathbf{x}^{(N-1)}, y_N) - g(\mathbf{x}^{(N-1)}, y_N)| = 0$, $\mathrm{d}y_N$-a.e., for all $\forall \mathbf{x}^{(N-1)} \notin A_{N-1}$. Consequently,

$$\begin{aligned}
&\lim_{\varepsilon\to 0} E|g^\varepsilon(\mathbf{x}^{(N-1)}, X_T^{0,x}) - g(\mathbf{x}^{(N-1)}, X_T^{0,x})| \\
&\quad = \lim_{\varepsilon\to 0} \int_{\mathbb{R}^n} |g^\varepsilon(\mathbf{x}^{(N-1)}, y) - g(\mathbf{x}^{(N-1)}, y)|\,p(0,x;T,y)\,\mathrm{d}y = 0 \quad \forall \mathbf{x}^{(N-1)} \notin A_{N-1},
\end{aligned}$$

thanks to the Dominated Convergence Theorem, proving the lemma. □

We now state and prove our main result.

Theorem 6.2. *Assume* (A2) *and* (A3). *Then, there exists a pair of functionals (u,v): $[0,T]\times\mathbb{R}^{Nn} \mapsto \mathbb{R}^m \times \mathbb{R}^{m\times n}$, such that*

(i) *$u(\cdot,\cdot)$ (resp. $v(\cdot,\cdot)$) is $\mathscr{B}([0,T]\times\mathbb{R}^{Nn})/\mathscr{B}(\mathbb{R}^m)$ (resp. $\mathscr{B}(\mathbb{R}^{m\times n})$)-measurable;*
(ii) *if we define*

$$Y_t^x = u(t, \mathbf{X}_t^{(N)}), \quad Z_t^x = v(t, \mathbf{X}_t^{(N)})\sigma(t, X_t^{0,x}), \quad t \in [0,T], \tag{6.3}$$

then (Y^x, Z^x) is an adapted solution to FBSDE (6.1) over $[0,T]$.

Proof. We shall construct the desired functional interval by interval, starting from $[t_{N-1}, T]$. More precisely, in each interval we proceed with two-steps: the first step is to prove the existence of the pair functions (u^k, v^k) satisfying conditions (C-1) and (C-2) in Section 5; and the second step is to modify these functions, with the help of Theorem 5.4, so that they have required measurability so that a recursive argument can be carried out. Consider now the interval $[t_{N-1}, t_N] = [t_{N-1}, T]$.

Step 1: Let $\{\varepsilon_\ell^N\}$ be a sequence such that

$$\lim_{l\to\infty} g^{\varepsilon_l^N}(\mathbf{x}^{(N-1)},x)=g(\mathbf{x}^{(N-1)},x),\ \ \mathrm{d}\mathbf{x}^{(N-1)}\,\mathrm{d}x\text{-a.e.}$$

We denote it by $\{\varepsilon_\ell\}$ for notational simplicity. Also, we denote the exceptional set in Lemma 6.1 by $A_{N-1}\subset\mathbb{R}^{(N-1)n}$, and let $x\in\mathbb{R}^n$ be fixed.

For each $\mathbf{x}^{(N-1)}\in\mathbb{R}^{(N-1)n}$, let us denote the solution of the following (decoupled) FBSDE, $(\tau,\xi)\in[t_{N-1},T)\times\mathbb{R}^n$, $s\in[\tau,T]$,

$$X_s=\xi+\int_\tau^s b(r,X_r)\,\mathrm{d}r+\int_\tau^s\sigma(r,X_r)\,\mathrm{d}W_r, \tag{6.4}$$

$$Y_s=g^{\varepsilon_\ell}(\mathbf{x}^{(N-1)},X_T)+\int_s^T f^{\varepsilon_\ell}(r,\mathbf{x}^{(N-1)},Y_r,Z_r)\,\mathrm{d}r-\int_s^T Z_r\,\mathrm{d}W_r \tag{6.5}$$

by $(X^{\tau,\xi},Y^{\ell,\tau,\xi}(\mathbf{x}^{(N-1)}),Z^{\ell,\tau,\xi}(\mathbf{x}^{(N-1)}))$. Then, thanks to Theorem 3.2, we have the following representation: for $(\tau,\xi)\in[t_{N-1},T)\times\mathbb{R}^n$ and $s\in[\tau,T]$,

$$\begin{aligned}Y_s^{\ell,\tau,\xi}(\mathbf{x}^{(N-1)})&=u^{\ell,N}(s,\mathbf{x}^{(N-1)},X_s^{\tau,\xi}),\\ Z_s^{\ell,\tau,\xi}(\mathbf{x}^{(N-1)})&=v^{\ell,N}(s,\mathbf{x}^{(N-1)},X_s^{\tau,\xi})\sigma(s,X_s^{\tau,\xi}),\end{aligned} \tag{6.6}$$

where $u^{\ell,N}(\tau,\mathbf{x}^{(N-1)},\xi)=Y_\tau^{\ell,\tau,\xi}(\mathbf{x}^{(N-1)})$ is the classical solution to the PDE (3.3) for $(\tau,\xi)\in[t_{N-1},T]\times\mathbb{R}^n$, and the following relation holds

$$v^{\ell,N}(\tau,\mathbf{x}^{(N-1)},\xi)=\partial_x u^{\ell,N}(\tau,\mathbf{x}^{(N-1)},\xi),\quad (\tau,\xi)\in[t_{N-1},T]\times\mathbb{R}^n.$$

We shall prove that for fixed $\mathbf{x}^{(N-1)}\notin A_{N-1}$, there exists a subsequence $\{\varepsilon_{\ell'}\}$ (this time it may depend on $\mathbf{x}^{(N-1)}$!), such that for each fixed $(\tau,\xi)\in[t_{N-1},T]\times\mathbb{R}^n$,

$$\lim_{\varepsilon_{\ell'}\to 0} u^{\ell',N}(\tau,\mathbf{x}^{(N-1)},\xi)=u^N(\tau,\mathbf{x}^{(N-1)},\xi),$$

where $u^N(\cdot,\mathbf{x}^{(N-1)},\cdot)$ is some measurable function.

To this end we fix $\mathbf{x}^{(N-1)}\notin A_{N-1}$, and denote, for each $\ell>0$,

$$\begin{aligned}&F^{\ell,N}(r,\mathbf{x}^{(N-1)},y)\\ &\quad\triangleq f^{\varepsilon_\ell}(r,\mathbf{x}^{(N-1)},y,u^{\ell,N}(r,\mathbf{x}^{(N-1)},y),v^{\ell,N}(r,\mathbf{x}^{(N-1)},y)\sigma(r,y)).\end{aligned} \tag{6.7}$$

Then clearly $F^{\ell,N}$ is a bounded measurable function on $(r,y)\in[t_{N-1},T]\times\mathbb{R}^n$. Now consider the family $\{F^{\ell,N}(\cdot,\mathbf{x}^{(N-1)},\cdot)\}_{\ell>0}$. Since f^{ε_ℓ} is uniformly bounded, this family is a bounded set in $L^\infty([t_{N-1},T]\times\mathbb{R}^n)$, whence weak*-precompact. In other words, there exists a subsequence $\{\ell'\}$ (may depend on $\mathbf{x}^{(N-1)}$!), and a function $F^{0,N}(\cdot,\mathbf{x}^{(N-1)},\cdot)\in L^\infty([t_{N-1},t_N]\times\mathbb{R}^n)$ such that for any $\varphi\in L^1([t_{N-1},t_N]\times\mathbb{R}^n)$, we have

$$\lim_{\ell'\to\infty}\int_{t_{N-1}}^T\int_{\mathbb{R}^n}[F^{\ell',N}(r,\mathbf{x}^{(N-1)},y)-F^{0,N}(r,\mathbf{x}^{(N-1)},y)]\varphi(r,y)\,\mathrm{d}y\,\mathrm{d}r=0.$$

Consequently, for all $(\tau,\xi)\in[t_{N-1},T)\times\mathbb{R}^n$ and $\delta>0$ sufficiently small, we have

$$\lim_{\ell'\to\infty}\int_{\tau+\delta}^{T}\int_{\mathbb{R}^n}[F^{\ell',N}(r,\mathbf{x}^{(N-1)},y)-F^{0,N}(r,\mathbf{x}^{(N-1)},y)]\,p(\tau,\xi;r,y)\,\mathrm{d}y\,\mathrm{d}r$$
$$=0. \tag{6.8}$$

We now borrow some idea from Hamadene et al. (1997) to prove that the sequence $\{u^{\ell',N}(\tau,\mathbf{x}^{(N-1)},\xi)\}$ converges pointwisely for every $(\tau,\xi)\in[t_{N-1},t_N]\times\mathbb{R}^n$ and $\mathbf{x}^{(N-1)}\notin A_{N-1}$. First, we claim that the sequence is Cauchy in ℓ'. Indeed, from BSDE (6.5) and representation (6.6), we see that for $(\tau,\xi)\in[t_{N-1},T)\times\mathbb{R}^n$ and $\delta>0$ is sufficiently small, then for any $\ell'_1,\ell'_2\in\{\ell'\}$,

$$\begin{aligned}
&|u^{\ell'_1,N}(\tau,\mathbf{x}^{(N-1)},\xi)-u^{\ell'_2,N}(\tau,\mathbf{x}^{(N-1)},\xi)|\\
&\quad=\left|E\int_\tau^T[F^{\ell'_1,N}(r,\mathbf{x}^{(N-1)},X_r^{\tau,\xi})-F^{\ell'_2,N}(r,\mathbf{x}^{(N-1)},X_r^{\tau,\xi})]\,\mathrm{d}r\right|\\
&\qquad+E|g^{\ell'_1}(\mathbf{x}^{(N-1)},X_T^{\tau,\xi})-g^{\ell'_2}(\mathbf{x}^{(N-1)},X_T^{\tau,\xi})|\\
&\quad=\left|\int_\tau^T\int_{\mathbb{R}^n}[F^{\ell'_1,N}(r,\mathbf{x}^{(N-1)},y)-F^{\ell'_2,N}(r,\mathbf{x}^{(N-1)},y)]\,p(\tau,\xi;r,y)\,\mathrm{d}r\,\mathrm{d}y\right|\\
&\qquad+E|g^{\ell'_1}(\mathbf{x}^{(N-1)},X_T^{\tau,\xi})-g^{\ell'_2}(\mathbf{x}^{(N-1)},X_T^{\tau,\xi})|\\
&\quad\leqslant\left|\int_\tau^{\tau+\delta}\int_{\mathbb{R}^n}[\Delta F^{\ell'_1,\ell'_2,N}(r,\mathbf{x}^{(N-1)},y)]\,p(\tau,\xi;r,y)\,\mathrm{d}r\,\mathrm{d}y\right|\\
&\qquad+\left|\int_{\tau+\delta}^T\int_{\mathbb{R}^n}[\Delta F^{\ell'_1,\ell'_2,N}(r,\mathbf{x}^{(N-1)},y)]\,p(\tau,\xi;r,y)\,\mathrm{d}r\,\mathrm{d}y\right|\\
&\qquad+E|g^{\ell'_1}(\mathbf{x}^{(N-1)},X_T^{\tau,\xi})-g^{\ell'_2}(\mathbf{x}^{(N-1)},X_T^{\tau,\xi})|\\
&\quad=I_1(\delta,\ell'_1,\ell'_2)+I_2(\delta,\ell'_1,\ell'_2)+I_3(\ell'_1,\ell'_2),
\end{aligned}\tag{6.9}$$

where

$$\Delta F^{\ell'_1,\ell'_2,N}(r,\mathbf{x}^{(N-1)},y)\triangleq F^{\ell'_1,N}(r,\mathbf{x}^{(N-1)},y)-F^{\ell'_2,N}(r,\mathbf{x}^{(N-1)},y)$$

and I_1, I_2 and I_3 are defined in an obvious way. We now analyze the convergence of I_1,I_2,I_3 one by one. First, by Lemma 6.1 we see that $I_3(\ell'_1,\ell'_2)\to0$, as $\ell'_1,\ell'_2\to\infty$. Next, applying Lemma 2.1 and using the boundedness of the function f, one shows that $I_1(\delta,\ell'_1,\ell'_2)\leqslant C_1\delta$, where $C_1>0$ is a constant independent of ℓ'. Finally, (6.8) implies that $I_2(\delta,\ell'_1,\ell'_2)\to0$ as $\ell'_1,\ell'_2\to\infty$, thanks to the weak*-precompactness of the sequence $\{F^{\ell',N}(\cdot,\mathbf{x}^{(N-1)},\cdot)\}$ in $L^\infty([t_{N-1},T]\times\mathbb{R}^n)$. Consequently, first letting $\ell'_1,\ell'_2\to\infty$ on both sides of (6.9) and then letting $\delta\to0$ one shows that the sequence $\{u^{\ell',N}(\tau,\mathbf{x}^{(N-1)},\xi)\}$ is Cauchy. Thus it converges, for fixed $\mathbf{x}^{(N-1)}\notin A_{N-1}$, to a (measurable) function $u^N(\tau,\mathbf{x}^{(N-1)},\xi)$, $(\tau,\xi)\in[t_{N-1},T]\times\mathbb{R}^n$. (Bearing in mind that the subsequence $\{\ell'\}$ depends on $\mathbf{x}^{(N-1)}$, and we do not claim any measurability of u^N in $\mathbf{x}^{(N-1)}$!).

Now let us consider the process

$$Y_t^{\ell',N}(\mathbf{x}^{(N-1)}) = u^{\ell',N}(t,\mathbf{x}^{(N-1)},X_t^{0,x}), \qquad (t,x)\in[t_{N-1},T]\times\mathbb{R}^n.$$

The pointwise convergence of the sequence $\{u^{\ell',N}(\cdot,\mathbf{x}^{(N-1)},\cdot)\}$ implies that for fixed $\mathbf{x}^{(N-1)}$, $Y^{\ell',N}(\mathbf{x}^{(N-1)})$ converges to a process $Y_t^N(\mathbf{x}^{(N-1)}) \triangleq u^N(t,\mathbf{x}^{(N-1)},X_t^{0,x})$, $t\in[t_{N-1},T]$, almost surely. Furthermore, applying Itô's formula to $|Y_t^{l_1',N} - Y_t^{l_2',N}|^2$ and following the same argument as that in Hamadene et al. (1997) one shows easily that $Z^{\ell',N}$ converges also to a process $Z^N(\mathbf{x}^{(N-1)})\in L^2([t_{N-1},T]\times\Omega;\mathbb{R}^{m\times d})$. Note that for each $\mathbf{x}^{(N-1)}\notin A_{N-1}$, and for each $\varepsilon>0$, $(Y^{\varepsilon,N}(\mathbf{x}^{(N-1)}),Z^{\varepsilon,N}(\mathbf{x}^{(N-1)}))$ solves the BSDE on $[t_{N-1},T]$:

$$Y_t^\varepsilon = g^\varepsilon(\mathbf{x}^{(N-1)},X_T^{0,x}) + \int_t^T f^\varepsilon(r,\mathbf{x}^{(N-1)},X_r^{0,x},Y_r^\varepsilon,Z_r^\varepsilon)\,\mathrm{d}r - \int_t^T Z_r^\varepsilon\,\mathrm{d}W_r. \qquad (6.10)$$

Applying Lemma 6.1 and using the continuity of the function f and the fact that f^ε converges to f uniformly on compact sets, as $\varepsilon'(=\varepsilon_{\ell'})\to 0$, it is standard to show that for each $\mathbf{x}^{(N-1)}\notin A_{N-1}$, the limiting process $(Y^N(\mathbf{x}^{(N-1)}),Z^N(\mathbf{x}^{(N-1)}))$ satisfies the limiting BSDE (suppressing "$\mathbf{x}^{(N-1)}$" in the notation) on $[t_{N-1},T]$:

$$Y_t^N = g(\mathbf{x}^{(N-1)},X_T^{0,x}) + \int_t^T f(r,\mathbf{x}^{(N-1)},X_r^{0,x},Y_r^N,Z_r^N)\,\mathrm{d}r - \int_t^T Z_r^N\,\mathrm{d}W_r. \qquad (6.11)$$

Defining $u^N(t,\mathbf{x}^{(N-1)},x)$ arbitrarily for $\mathbf{x}^{(N-1)}\in A_{N-1}$, we complete the construction of u^N.

We now show that there exists a function $v^N(t,\mathbf{x}^{(N-1)},\xi)$ such that, whenever $\mathbf{x}^{(N-1)}\notin A_{N-1}$ it holds that

$$Z_t^N(\mathbf{x}^{(N-1)}) = v^N(t,\mathbf{x}^{(N-1)},X_t^{0,x})\sigma(t,X_t^{0,x}), \qquad \mathrm{d}P\times\,\mathrm{d}t\text{-a.s.}$$

Indeed, denote

$$v^N(t,\mathbf{x}^{(N-1)},\xi) = \overline{\lim_{\varepsilon'\to 0}}\, v^{\varepsilon',N}(t,\mathbf{x}^{(N-1)},\xi).$$

Then, since $Z^{\varepsilon',N}$ converges to $Z^N(\mathbf{x}^{(N-1)})\in L^2([t_{N-1},T]\times\Omega;\mathbb{R}^{m\times d})$, along a subsequence, may denote $\{\varepsilon'\}$ itself, the convergence is $\mathrm{d}P\times\mathrm{d}t$-almost sure. Therefore, one has

$$\begin{aligned} v^N(t,\mathbf{x}^{(N-1)},X_t^{0,x})\sigma(t,X_t^{0,x}) &= \overline{\lim_{\varepsilon'\to 0}}\, v^{\varepsilon',N}(t,\mathbf{x}^{(N-1)},X_t^{0,x})\sigma(t,X_t^{0,x}) \\ &= \overline{\lim_{\varepsilon'\to 0}}\, Z_t^{\varepsilon',N}(\mathbf{x}^{(N-1)}) = Z_t^N(\mathbf{x}^{(N-1)}). \end{aligned}$$

This completes Step 1 of the construction.

Step 2: We note here that the functions $u^N(t,\mathbf{x}^{(N-1)},\xi)$ and $v^N(t,\mathbf{x}^{(N-1)},\xi)$ constructed in Step 1 are only measurable in (t,ξ), for each fixed $\mathbf{x}^{(N-1)}$ and $x\in\mathbb{R}^d$. Also, only for $\mathbf{x}^{(N-1)}\notin A_{N-1}$, the processes

$$Y_t^N \triangleq u^N(t,\mathbf{x}^{(N-1)},X_t^{0,x}), \quad Z_t^N \triangleq v^N(t,\mathbf{x}^{(N-1)},X_t^{0,x})\sigma(t,X_t^{0,x})$$

satisfy BSDE (6.11). However, applying the measurable selection Theorem 5.4 we can choose a pair of $(t,\mathbf{x}^{(N-1)},\xi)$-jointly measurable functions $(\bar{u}^N,\bar{v}^N)$, and define, for $t\in[t_{N-1},T]$,

$$\bar{Y}_t^{x,N}=\bar{u}^N(t,\mathbf{X}^{(N-1)},X_t^{0,x}), \quad \bar{Z}_t^{x,N}=\bar{v}^N(t,\mathbf{X}^{(N-1)},X_t^{0,x})\sigma(t,X_t^{0,x}).$$

Then, noting that $\mathbf{X}^{(N-1)}$ has a positive density, it can be checked, using the by now standard argument of regular conditional probabilities, that $(\bar{Y}^N,\bar{Z}^N)$ solves BSDE (6.1) on $[t_{N-1},T]$. This completes the second step, whence the construction of the adapted solution on $[t_{N-1},T]$.

To proceed further, we need to make sure that the same arguments can be applied to the subsequent interval $[t_{N-2},t_{N-1}]$. Let us drop the sign "–" and superscript x from the aforementioned solutions $(\bar{Y}^N,\bar{Z}^N)$.

Similar to the estimate (5.6), we now have

$$\begin{aligned}|Y^N_{t_{N-1}}| &= |u^N(t_{N-1},\mathbf{X}^{(N-1)},X_{t_{N-1}})| = |u^N(t_{N-1},\mathbf{X}^{(N-2)},X_{t_{N-1}},X_{t_{N-1}})|\\ &= \left|E\left\{g(\mathbf{X}^{(N-1)},X_T^{0,x})+\int_{t_{N-1}}^T f(r,\mathbf{X}^{(N-1)},X_r^{0,x},Y_r^N,Z_r^N)\,\mathrm{d}r\right\}\right|\\ &\leqslant \|g\|_\infty+\|f\|_\infty(T-t_{N-1})=C_N. \end{aligned}\tag{6.12}$$

Let us denote

$$g^{(N-1)}(\mathbf{x}^{(N-2)},x)=u^N(t_{N-1},\mathbf{x}^{(N-2)},x,x)\mathbf{1}_{\{|u^N(t_{N-1},\mathbf{x}^{(N-2)},x,x)|\leqslant C_N\}},$$

$$f^{(N-1)}(t,\mathbf{x}^{(N-2)},x,y,z)=f(t,\mathbf{x}^{(N-2)},x,x,y,z).$$

Then $g^{(N-1)}$ and $f^{(N-1)}$ satisfy (A2) with N being replaced by $N-1$. Moreover, recalling (5.6) we see that BSDE (6.1) can be written, for $t\in[t_{N-2},t_{N-1}]$ as

$$\begin{aligned}Y_t &= Y^N_{t_{N-1}}+\int_t^{t_{N-1}} f(r,\mathbf{X}^{(N-1)},X_r,X_r,Y_r,Z_r)\,\mathrm{d}r-\int_t^{t_{N-1}} Z_r\,\mathrm{d}W_r\\ &= u^N(t_{N-1},\mathbf{X}^{(N-1)},X_{t_{N-1}})\\ &\quad+\int_t^{t_{N-1}} f^{(N-1)}(r,\mathbf{X}^{(N-2)},X_r,Y_r,Z_r)\,\mathrm{d}r-\int_t^{t_{N-1}} Z_r\,\mathrm{d}W_r\\ &= g^{(N-1)}(\mathbf{X}^{(N-2)},X_{t_{N-1}})\\ &\quad+\int_t^{t_{N-1}} f^{(N-1)}(r,\mathbf{X}^{(N-2)},X_r,Y_r,Z_r)\,\mathrm{d}r-\int_t^{t_{N-1}} Z_r\,\mathrm{d}W_r. \end{aligned}\tag{6.13}$$

Thus, we can repeat the same argument as before to obtain a pair of measurable functions (u^{N-1}, v^{N-1}) so that the process

$$Y_t^{x,N-1} = u^{N-1}(t, \mathbf{X}^{(N-2)}, X_t), \quad Z_t^{x,N-1} = v^{N-1}(t, \mathbf{X}^{(N-2)}, X_t)$$

is the adapted solution of BSDE (6.13). Continuing for N steps, we obtain N pairs of processes $(Y^{x,k}, Z^{x,k})$ and N pairs of functions (u^k, v^k) with

$$Y_t^{x,k} = u^k(t, \mathbf{X}^{(k-1)}, X_t), \quad Z_t^{x,k} = v^k(t, \mathbf{X}^{(k-1)}, X_t)\sigma(t, X_t), \quad t \in [t_{k-1}, t_k]$$

solves BSDE (6.1) on $[t_{k-1}, t_k]$. Finally, define the functionals u and v as in (3.6) and then the processes Y and Z as in (3.7) we obtain an adapted solution of BSDE (6.1) on $[0, T]$ as desired. □

Remark 6.3. Let $t \in [t_{k-1}, t_k)$ for some $1 \leqslant k \leqslant N$. We define $(Y^{t,x}, Z^{t,x})$ by (3.13), where (u, v) is what we have obtained in the above theorem. Then $(Y^{t,x}, Z^{t,x})$ solves the BSDE (3.12).

Acknowledgements

We are deeply indebted to the anonymous referee for his/her extremely enthusiastic reading of our original manuscript and many insightful comments and suggestions. Our paper has improved greatly during the revision process because of that.

References

Aronson, D.G., 1967. Bounds for the fundamental solution of a parabolic equation. Bull. Amer. Math. Soc. 73, 890–896.

Aronson, D.G., 1968. Non-negative solutions of linear parabolic equations. Ann. Scuola Norm. Sup. Pisa 22, 607–694.

Bertsekas, D.P., Shreve, S.E., 1978. Stochastic Optimal Control. The Discrete Time Case. Academic Press, New York.

Bismut, J.M., 1976. Théorie Probabiliste du Contrôle des Diffusions. Mem. Amer. Math. Soc., Vol. 4 (167). Amer. Math. Soc., Providence, RI.

Çinlar, E., Jacod, J., Protter, P., Sharpe, M.J., 1980. Semimartingales and Markov processes. Z. Wahrsch. Verw. Gebiete 54 (2), 161–219.

Dellacherie, C., Meyer, P.A., 1978. Probabilities and Potential. North-Holland Publishing Co., Amsterdam.

El Karoui, N., Mazliak, L. (Eds.), 1997. Backward Stochastic Differential Equations. Pitman Research Notes in Mathematics Series, Vol. 364, Longman, Harlow.

El Karoui, N., Peng, S., Quenez, M.C., 1997. Backward Stochastic Differential Equations in finance. Math. Finance 7 (1), 1–71.

Hamadene, S., 1996. Équations différentielles stochastiques rétrogrades: les cas localement Lipschitzien. Ann. Inst. H. Poincaré Probab. Statist. 32 (5), 645–659.

Hamadene, S., Lepeltier, J.P., Peng, S., 1997. BSDEs with continuous coefficients and stochastic differential games. In: El Karoui, N., Mazliak, L. (Eds.), Backward Stochastic Differential Equations (Paris, 1995–1996), Pitman Research Notes in Mathematics Series, Vol. 364. Longman, Harlow, pp. 115–128.

Kobylanski, M., 2000. Backward stochastic differential equations and partial differential equations with quadratic growth. Ann. Probab. 28 (2), 558–602.

Ladyzenskaja, O.A., Solonnikov, V.A., Ural'ceva, N.N., 1968. Linear and Quasilinear Equations of Parabolic Type. American Mathematical Society, Providence, RI.

Lepeltier, J.P., San Martin, J., 1997. Backward stochastic differential equations with continuous coefficient. Statist. Probab. Lett. 32 (4), 425–430.

Lepeltier, J.P., San Martin, J., 1998. Existence for BSDE with superlinear-quadratic coefficient. Stochastics Stochastics Rep. 63 (3–4), 227–240.

Ma, J., Yong, J., 1999. Forward–backward Stochastic Differential Equations and Their Applications. Lecture Notes in Mathematics, Vol. 1702. Springer, Berlin.

Ma, J., Zhang, J., 2002. Representation theorems for backward stochastic differential equations. Ann. Appl. Probab. 12 (4), 1390–1418.

Ma, J., Protter, P., Yong, J., 1994. Solving forward–backward stochastic differential equations explicitly—a four step scheme. Probab. Theory Relat. Fields 98 (3), 339–359.

Pardoux, E., Peng, S., 1990. Adapted solution of a backward stochastic differential equation. Systems Control Lett. 14 (1), 55–61.

Pardoux, E., Peng, S., 1992. Backward Stochastic Differential Equations and Quasilinear Parabolic Partial Differential Equations. Stochastic Partial Differential Equations and Their Applications (Charlotte, NC, 1991). Lecture Notes in Control and Information Science, Vol. 176. Springer, Berlin, pp. 200–217.

Zhang, Y., Zheng, W., 2002. Discretizing a backward stochastic differential equation. Internat. J. Math. Math. Sci. 32 (2), 103–116.