The Annals of Probability
2011, Vol. 39, No. 4, 1422–1448
DOI: 10.1214/10-AOP588

BACKWARD STOCHASTIC DYNAMICS ON A FILTERED PROBABILITY SPACE[1]

BY GECHUN LIANG, TERRY LYONS AND ZHONGMIN QIAN

University of Oxford

We demonstrate that backward stochastic differential equations (BSDE) may be reformulated as ordinary functional differential equations on certain path spaces. In this framework, neither Itô's integrals nor martingale representation formulate are needed. This approach provides new tools for the study of BSDE, and is particularly useful for the study of BSDE with partial information. The approach allows us to study the following type of backward stochastic differential equations:

$$dY_t^j = -f_0^j(t, Y_t, L(M)_t)\,dt - \sum_{i=1}^d f_i^j(t, Y_t)\,dB_t^i + dM_t^j$$

with $Y_T = \xi$, on a general filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$, where B is a d-dimensional Brownian motion, L is a prescribed (nonlinear) mapping which sends a square-integrable M to an adapted process $L(M)$ and M, a correction term, is a square-integrable martingale to be determined. Under certain technical conditions, we prove that the system admits a unique solution (Y, M). In general, the associated partial differential equations are not only nonlinear, but also may be nonlocal and involve integral operators.

1. Introduction. Stochastic differential equations (SDE) may be considered as dynamical systems perturbed by random signals which are often modeled by Brownian motion. The important class of stochastic differential equations considered in the literature are Itô-type equations such as

$$dX_t^j = f_0^j(t, X_t)\,dt + \sum_{i=1}^d f_i^j(t, X_t)\,dB_t^i, \tag{1.1}$$

where $B = (B^1, \ldots, B^d)$ is Brownian motion in $\mathbf{R}^d$ on a completed probability space $(\Omega, \mathcal{F}, \mathbf{P})$, $f_i = \sum_{j=1}^{d'} f_i^j \frac{\partial}{\partial x^j}$ are bounded, smooth vector fields in $\mathbf{R}^{d'}$, $j = 1, \ldots, d'$, where d, d' are two positive integers. Itô gave the meaning of solutions to (1.1) by developing a theory of stochastic integration against Brownian motion (called Itô's calculus), and obtained strong solutions by specifying an initial data at a starting time T.

Received July 2009; revised March 2010.
[1]Supported in part by EPSRC Grant EP/F029578/1 and by the Oxford-Man Institute.
MSC2010 subject classifications. Primary 60H10, 60H30; secondary 60J45.
Key words and phrases. Brownian motion, BSDE, SDE, semimartingale.

SDE (1.1) has to be interpreted as an integral equation

$$X_t^j - X_0^j = \int_0^t f_0^j(s, X_s)\, ds + \sum_{i=1}^d \int_0^t f_i^j(s, X_s)\, dB_s^i,$$

which can be solved forward (i.e., for $t > 0$). Itô's calculus requires that a solution $X = (X_t)$ has to be adapted to Brownian motion $B = (B^1, \dots, B^d)$; it is thus not necessarily possible to solve (1.1) backward from a certain time T to $t < T$.

There are interesting applications on the other hand to be able to solve (1.1) backward. Suppose u is a smooth solution to the Cauchy problem of the quasi-linear parabolic equation

$$\left(\frac{1}{2}\Delta - \frac{\partial}{\partial t}\right)u + f(u, \nabla u) = 0 \qquad \text{on } [0, \infty) \times \mathbf{R}^d,$$

with the initial data $u(x, 0) = u_0(x)$. Let $T > 0$ and $h(t, x) = u(T - t, x)$ for $t \in [0, T]$. Then h solves the backward parabolic equation

$$\left(\frac{1}{2}\Delta + \frac{\partial}{\partial t}\right)h + f(h, \nabla h) = 0 \qquad \text{on } [0, T] \times \mathbf{R}^d,$$

and $h(x, T) = u_0(x)$. Let $Y_t = h(t, B_t)$ where B is Brownian motion in $\mathbf{R}^d$. According to Itô's formula

$$Y_T - Y_t = \int_t^T \left(\frac{\partial}{\partial s} + \frac{1}{2}\Delta\right)h(s, B_s)\, ds + M_T - M_t \tag{1.2}$$

for $t \le T$, where $M_t = \int_0^t \nabla h(s, B_s)\, dB_s$ is a martingale. Substituting $(\frac{\partial}{\partial s} + \frac{1}{2}\Delta)h$ by $-f_0(h, \nabla h)$ in (1.2) obtains

$$Y_T - Y_t = -\int_t^T f(Y_s, \nabla h(s, B_s))\, ds + M_T - M_t. \tag{1.3}$$

According to Itô's martingale representation theorem, the *density process* $Z_t = \nabla h(t, B_t)$ of M with respect to Brownian motion is uniquely determined as the unique predictable process Z_t such that

$$M_T - M_0 = \sum_{j=1}^d \int_0^T Z_t^j\, dB_t^j.$$

In terms of the pair (Y, Z) (1.3) may be written as

$$Y_T - Y_t = -\int_t^T f(Y_s, Z_s)\, ds + \sum_{j=1}^d \int_t^T Z_s^j\, dB_s^j$$

with the terminal data $Y_T = u_0(B_T)$, which is the integral form of the following backward stochastic differential equation:

$$dY_t = -f(Y_t, Z_t)\, dt + Z_t\, dB_t, \qquad Y_T = \xi, \tag{1.4}$$

introduced and studied by Pardoux and Peng [32].

In the past twenty years, there has been a large number of articles devoted to the theory of BSDE and its applications in various research areas. Our references listed at the end of the paper are by no means complete, and the reader should refer to excellent surveys such as articles in [18] edited by El Karoui and Mazliak, the recent paper by El Karoui, Hamadene and Matoussi [16], the book by Yong and Zhou [40] and the references therein for a guide to the BSDE literature.

To the knowledge of the present authors, it was Bismut [5] (see [6, 7]) who first formulated terminal problems for a class of stochastic differential equations in order to study stochastic optimal control problems by means of Pontryagin's maximum principal. His equations, called backward stochastic differential equations, have been extended and developed to a nonlinear case in the seminal paper [32] by Pardoux and Peng. A lot of efforts have been made to generalize the class of BSDE considered in [32]. For example, Lepeltier and San Martin [26] relaxed the Lipschitz continuous condition on the driver and studied BSDEs with coefficients of linear growth. Yong [39] employed the continuity method to prove the existence of solution with arbitrary time horizon. In [8] Briand et al. considered L^p-solutions for BSDE. It is also natural to consider BSDE coupled with a forward stochastic differential equation, called a forward–backward stochastic differential equations. Antonelli [1] first studied such FBSDE; his equation does not involve a density process Z in the driver. A definite account about FBSDE may be found in Ma, Protter and Young [27], Hu and Peng [23], Peng and Wu [33], the recent book [28] and the literature therein. Most authors consider BSDE on a probability space with Brownian filtration, and there are a few papers dealing with BSDE with jumps or with reflecting boundary conditions. Tang and Li [38] have studied BSDE with random jumps, and Barles, Buckdahn and Pardoux [4] have explored the connection between BSDE with random jumps and some parabolic integro-differential equations. Rong [36] proved the existence and uniqueness under non-Lipschitz continuous coefficients for this class of BSDE. Analogous to free-boundary PDE problems, El Karoui et al. [17] introduced an obstacle to BSDE such that the solution always stays above such obstacle. This so-called reflected BSDE is further developed to double reflected barriers by Cvitanić and Karatzas [14] and Hamadene, Lepeltier and Matoussi [21]. Furthermore, Bally, Pardoux and Stoica [3] have considered BSDE on the probability space associated with Dirichlet processes.

If the driver of BSDE is with quadratic growth of Z, the nature of equations is completely changed. This problem is first solved by Kobylanski [24] by using the Cole–Hopf transformation adopted from the PDE theory. Her results have been substantially developed and generalized by Briand and Hu [9, 10], where they extend to equations with convex drivers subject to unbounded terminal values. Most of the existing literature concentrates on solutions of BSDEs in a strong sense, that is, the underlying filtered probability space is given. One of the first attempts to introduce weak solutions for BSDEs was presented in Buckdahn, Engelbert and Răşcanu [12], and Buckdahn and Engelbert [11] who further proved the uniqueness of their weak solutions, while the coefficients of their BSDEs do not evolve a density process Z. On the other hand, the notion of weak solution for FBSDEs was introduced by Antonelli and Ma [2] and further developed by Ma, Zhang and Zheng [29] by employing the martingale problem approach.

The backward stochastic differential equations have found many connections with other research areas: stochastic control, PDE, mathematical finance, etc. To derive a maximum principle as necessary conditions for optimal control problems, one can observe that the adjoint equations to the optimal control problems satisfy certain backward equations. For stochastic control problems, the corresponding adjoint equations are stochastic rather than deterministic. Indeed Peng [34] established a general stochastic maximum principle by considering both first-order and second-order adjoint equations, and, on the other hand, Kohlmann and Zhou [25] interpreted BSDE as equivalent to stochastic control problems. Peng [35] and Pardoux and Pend [31] derived a probabilistic representation (a *Feynman–Kac representation*) for solutions of some quasi-linear PDEs, which was extended to other cases by Ma, Protter and Yong [27]. The later has been summarized as a four-step scheme of solving forward–backward stochastic differential equations (FBSDE) (see [28] by Ma and Yong for details). Cheridito, Soner, Touzi and Victoir [13] connected a class of second order BSDEs to fully nonlinear PDEs. In [15] Duffie and Epstein discovered a class of nonlinear BSDE in their study of recursive utility in economics. Later El Karoui, Peng and Quenez [19] applied BSDE to option pricing problems and provided a general framework for the application of BSDE in finance. In order to deal with utility maximization problems in incomplete markets, Rouge and El Karoui [37] introduced a class of BSDE with quadratic growth. Hu, Imkeller and Müller [22] further studied this class of BSDE in a more general setting.

In this article, we put forward a simple approach to deal with the kind of BSDE such as (1.4) which does not depend on any martingale representation, and thus allows us to study a wide class of backward stochastic dynamics. Our main idea and contribution in this article is to establish an *ordinary functional differential equation* which is equivalent to (1.4), which allows us to obtain alternative representations for solutions of BSDE and to consider a new interesting class of stochastic dynamical systems.

Consider the following example of backward stochastic differential equations:

$$dY_t = -f(t, Y_t, Z_t)\,dt + \sum_{i=1}^{d} Z_t^i\,dB_t^i, \qquad Y_T = \xi, \tag{1.5}$$

where $B = (B_t)_{t\geq 0}$ is Brownian motion in $\mathbf{R}^d$, $\xi \in L^2(\Omega, \mathcal{F}_T, \mathbf{P})$ and $(\mathcal{F}_t)_{t\geq 0}$ is the Brownian filtration. The differential equation has to be interpreted as the integral equation

$$\xi - Y_t = -\int_t^T f(s, Y_s, Z_s)\,ds + \sum_{i=1}^{d} \int_t^T Z_s^i\,dB_s^i. \tag{1.6}$$

By applying the Picard iteration to (Y, Z), one shows that if f is globally Lipschitz continuous, then there is a unique pair (Y, Z) which satisfies (1.6) for all $t \le T$. This method relies on the martingale representation for Brownian motion and thus restricts the class of BSDE.

Our main idea is based on the following simple observation. Suppose that $Y = (Y_t)_{t\in[\tau,T]}$ is a solution of (1.6) back to time $\tau < T$, then Y must be a special semimartingale whose variation part is continuous. Let $Y_t = M_t - V_t$ be the Doob–Meyer decomposition into its martingale part M and its finite variation part $-V$. The decomposition over $[\tau, T]$ is unique up to a random variable measurable with respect to $\mathcal{F}_\tau$. Let us assume that the local martingale part M is indeed a martingale up to T. Then, since the terminal value $Y_T = \xi$ is given, $\xi = M_T - V_T$, so that $M_t = \mathbf{E}(\xi + V_T|\mathcal{F}_t)$ and $Y_t = \mathbf{E}(\xi + V_T|\mathcal{F}_t) - V_t$ for $t \in [\tau, T]$. The integral equation (1.6) thus can be written as

$$\xi - M_t + V_t = -\int_t^T f(s, Y_s, Z_s)\,ds + \sum_{i=1}^d \int_t^T Z_s^i\,dB_s^i$$

for every $t \in [\tau, T]$. Taking expectations, with both sides conditional on $\mathcal{F}_t$, one obtains

$$\mathbf{E}(\xi|\mathcal{F}_t) - M_t + V_t = -\mathbf{E}\left[\int_\tau^T f(s, Y_s, Z_s)\,ds\middle|\mathcal{F}_t\right] + \int_\tau^t f(s, Y_s, Z_s)\,ds.$$

By identifying the martingale parts and variational parts, we must have

$$V_t - V_\tau = \int_\tau^t f(s, Y_s, Z_s)\,ds, \tag{1.7}$$

where Y and Z are considered as functionals of V, namely

$$Y_t = \mathbf{E}(\xi + V_T|\mathcal{F}_t) - V_t, \qquad M_t = \mathbf{E}(\xi + V_T|\mathcal{F}_t), \tag{1.8}$$

and Z is determined uniquely by the martingale representation through

$$M_T - M_\tau = \sum_{i=1}^d \int_\tau^T Z_s^i\,dB_s^i.$$

Hence Y and Z are written as $Y(V)$ and $Z(V)$, respectively, if we wish to emphasize the fact that Y and Z are defined entirely through V. Observe that (1.7) is clearly the integral form of the functional differential equation

$$\frac{dV}{dt} = f(t, Y(V)_t, Z(V)_t),$$

which can be solved by Picard iteration applying to V alone, rather than the pair (Y, Z).

The approach may be made independent of the use of a martingale representation theorem, provided that one is willing to replace the density process Z by a functional of V, thus freeing us from the requirement of Brownian filtration. This kind of generalization of BSDE theory is a bit surprising and even overly rewarded, which is, however, not the only point we would like to emphasize. More precisely, we may consider the correction martingale part appearing in (1.5) as part of the solution rather than its density process Z. That is, by setting $M_t - M_\tau = \sum_{i=1}^d \int_\tau^t Z_s^i \, dB_s^i$, and regarding Z as a function of M, so denoted by $L(M)$, then (1.5) can be reformulated as

$$dY_t = -f(t, Y_t, L(M)_t)\, dt + dM_t, \qquad Y_T = \xi, \tag{1.9}$$

which is in turn equivalent to the functional integral equation

$$V_t - V_\tau = \int_\tau^t f(s, Y(V)_s, L(M(V))_s)\, ds, \tag{1.10}$$

where $Y(V)$ and $M(V)$ are given by (1.8). For (1.10), there is no need to insist that L sends a martingale M to its density process (if there is any), though the density process mapping L remains the most interesting case.

The approach might be applied to a more general setting of solving dynamical systems backward under other constraints, not necessarily the adaptedness to a filtration; even a probability setting is not necessary. One possible example can be the following. One may study the functional differential equation (1.7), where $Y : V \to Y(V)$ and $M : V \to M(V)$ are defined in terms of some kind of "projections" instead of conditional expectations. We, however, in this paper, make no attempt for such an extension.

To our knowledge, most of BSDE which currently exist in the literature may be studied in the framework of ordinary functional differential equations. Since our approach does not rely on the martingale representation theorem, we are able to study a class of BSDE on an arbitrary filtered probability space. We, however, would like to point out that this paper is not so much about generalizing the theory of BSDE to a general filtered probability space; our main contribution is the equivalence of BSDE and a class of ordinary functional integral equations. We allow a sufficient wide class of functionals $L(M)$ which, even in the classical setting, extends the associated PDE to some nonlocal integro-differential equations.

If $(\mathcal{F}_t)_{t\geq 0}$ is Brownian filtration, any martingale $M = (M_t)_{t\geq 0}$ has an Itô integral representation $M_t - M_0 = \sum_{j=1}^d \int_0^t Z_s^j \, dW_s^j$ which determines the density $Z = (Z^1, \dots, Z^d)$. Consider the functional over martingales

$$L(M)_t = \mathbf{E}\left\{ \int_t^T Z_s \mu(t, ds) \middle| \mathcal{F}_t \right\},$$

where $\mu(t, ds)$ is a transition kernel (not random for simplicity), and consider the corresponding BSDE

$$dY_t^j = -f_0^j\big(T - t, Y_t, L(M)_t\big)\, dt + dM_t^j, \qquad Y_T = \xi.$$

Our approach demonstrates the existence and uniqueness for this kind of BSDE, whose associated PDE is a system of integro-differential equations,

$$\frac{\partial}{\partial t}u - \frac{1}{2}\Delta u + f_0(t, u, H(u)) = 0,$$

where the nonlinear operator H involves space–time integration, and indeed

$$H(u)(t,x) = \int_t^T \int_{\mathbf{R}^d} \frac{\nabla u(T-s, z)}{(2\pi(s-t))^{d/2}} e^{-|x-z|^2/(2(s-t))} \, dz \, \mu(t, ds).$$

If $\mu(t, ds) = \delta_t(ds)$ then we recover the case considered in the current literature. By choosing different functionals $L(M)$ we may obtain even more general integro-differential equations. This kind of integro-differential equations often appears in the study of particle limiting models for PDE; one class of equations which has a similar nature is already in the literature, for example, in Majda [30].

In this paper we constrain ourselves to the study of the following type of backward stochastic differential equations:

$$dY_t^j = -f_0^j(t, Y_t, L(M)_t)\, dt - \sum_{i=1}^d f_i^j(t, Y_t)\, dB_t^i + dM_t^j, \tag{1.11}$$

subject to $Y_T = \xi$, on a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$, where B is a d-dimensional Brownian motion as given, $j = 1, \ldots, d'$, L is a given (nonlinear) functional on square-integrable martingales, while $(\mathcal{F}_t)_{t\geq 0}$ is not necessary to be Brownian filtration. A solution to (1.11) is a pair (Y, M), where $Y = (Y^j)$ are semimartingales and $M = (M^j)$ are square-integrable martingales, which satisfies the corresponding integral equations:

$$\begin{aligned} Y_T^j - Y_t^j = {} & -\int_t^T f_0^j(t, Y_t, L(M)_t)\, dt - \sum_{i=1}^d \int_t^T f_i^j(t, Y_t)\, dB_t^i \\ & + M_T^j - M_t^j. \end{aligned} \tag{1.12}$$

The term $L(M)$ appearing in the drift term f_0 on the right-hand side of (1.11) suggests that L is a mapping which sends a vector of square-integrable martingales $M = (M^j)$ to a progressively measurable process $L(M)$. The backward stochastic equation (1.11) is thus described by the driver f_0, the diffusion coefficients f_i together with the prescribed mapping L.

Finally, let us point out that similar ideas have been known in the PDE theory. Recall that, for any reasonable function u, u has the following decomposition:

$$u = H(u) + G(u),$$

where $H(u)$ is a harmonic function determined by a boundary integral against a Green function, and $G(u)$ is a potential. Thus the boundary condition (which corresponds to our case the terminal value) determines the harmonic function part $H(u)$.

The regularity theory for nonlinear PDE such as $\Delta u = f(u, \nabla u)$ may be developed via the previous decomposition, by studying the Newtonian potential $G(u)$, (Gilbarg and Trudinger [20]). In this way, backward stochastic dynamics, as a class of Markov processes, can be regarded as a generic extension of some nonlinear PDE problems of finite dimension to infinite-dimensional problems in path spaces. On the other hand, some nonlinear PDE can be considered as a pathwise version of backward stochastic dynamics. We will explore these ideas further in coming papers.

The paper is organized as follows. In Section 2 we present some elementary facts and basic assumptions. The main results of the existence of local and global solutions, and the uniqueness of the backward stochastic dynamics are presented and proved in Sections 3 and 4.

2. Preliminaries. Let $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$ (where $t \in [0, \infty)$) be a filtered probability space which satisfies the *usual conditions*: $(\Omega, \mathcal{F}, \mathbf{P})$ is a completed probability space, $(\mathcal{F}_t)_{t \geq 0}$ is a right-continuous filtration, each $\mathcal{F}_t$ contains all events in $\mathcal{F}$ with probability zero and $\mathcal{F} = \sigma\{\mathcal{F}_t : t \geq 0\}$. Under these technical assumptions, any martingale on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$ has a modification whose sample paths are right continuous with left-hand limits. Henceforth, by a martingale we always mean a martingale which is right continuous with left-hand limits.

Let $0 \leq \tau < T$ be any but fixed numbers. $[\tau, T]$ serves as the region of the time parameter, although we are working with a fixed filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$. Let $\mathcal{C}([\tau, T]; \mathbf{R}^d)$ denote the space of all continuous, adapted processes $(V_t)_{t \in [\tau, T]}$ valued in $\mathbf{R}^d$ such that $\max_j \sup_{t \in [\tau, T]} |V_t^j|$ belongs to $L^2(\Omega, \mathcal{F}_T, \mathbf{P})$, equipped with the norm

$$\|V\|_{\mathcal{C}[\tau, T]} = \sqrt{\sum_{j=1}^{d} \mathbf{E} \sup_{t \in [\tau, T]} |V_t^j|^2}.$$

$\mathcal{C}([\tau, T]; \mathbf{R}^d)$ is a Banach space under $\|\cdot\|_{\mathcal{C}[\tau, T]}$, $\mathcal{M}^2([\tau, T]; \mathbf{R}^d)$ denotes the space of $\mathbf{R}^d$-valued square-integrable martingales on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$ from time τ up to time T (which, of course, can be uniquely extended to a martingale in $\mathcal{M}^2([0, T], \mathbf{R}^d)$), together with the norm $\|M\|_{\mathcal{C}[\tau, T]}$. We also need the direct sum space of $\mathcal{M}^2([\tau, T]; \mathbf{R}^d)$ and $\mathcal{C}([\tau, T]; \mathbf{R}^d)$, denoted by $\mathcal{S}([\tau, T]; \mathbf{R}^d)$. If $Y \in \mathcal{S}([\tau, T]; \mathbf{R}^d)$, its decomposition into an element in $\mathcal{M}^2([\tau, T]; \mathbf{R}^d)$ and the other in $\mathcal{C}([\tau, T]; \mathbf{R}^d)$ may not be unique, and there are various norms one can define on $\mathcal{S}([\tau, T]; \mathbf{R}^d)$. For our purposes, we choose the norm $\|Y\|_{\mathcal{C}[\tau, T]}$, although $\mathcal{S}([\tau, T]; \mathbf{R}^d)$ is not complete under $\|\cdot\|_{\mathcal{C}[\tau, T]}$. Finally $\mathcal{H}^2([\tau, T]; \mathbf{R}^{d' \times d})$ denotes the space of all *predictable* processes $Z = (Z_t^{j,i})_{t \in [\tau, T]}$ on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$ with running time $[\tau, T]$, endowed with the usual L^2-norm

$$\|Z\|_{\mathcal{H}^2_{[\tau, T]}} = \sqrt{\sum_{j=1}^{d'} \sum_{i=1}^{d} \mathbf{E} \int_\tau^T |Z_s^{i,j}|^2 \, ds}.$$

If Y is a semimartingale on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$ over time interval $[\tau, T]$ with its Doob–Meyer decomposition $Y_t = M_t - V_t$, such that M is an $\mathcal{F}_t$-martingale during $[\tau, T]$, V is a *continuous*, adapted process with finite variation on $[\tau, T]$ and V_T, Y_T are integrable, then $M_t = \mathbf{E}(Y_T + V_T|\mathcal{F}_t)$ and $Y_t = \mathbf{E}(Y_T + V_T|\mathcal{F}_t) - V_t$ for $t \in [\tau, T]$. Since we are interested in terminal value problems, in which $Y_T = \xi$ are given, therefore, for given $\xi = (\xi^i)$ where $\xi^i \in L^2(\Omega, \mathcal{F}_T, \mathbf{P})$, we consider two functionals on $\mathcal{C}([\tau, T]; \mathbf{R}^d) : V \to Y(V)$ and $V \to M(V)$ defined by

$$Y(V)_t = \mathbf{E}(\xi + V_T|\mathcal{F}_t) - V_t \qquad \text{for } t \in [\tau, T] \tag{2.1}$$

and

$$M(V)_t = \mathbf{E}(\xi + V_T|\mathcal{F}_t) \qquad \text{for } t \in [\tau, T] \tag{2.2}$$

for any $V \in \mathcal{C}([\tau, T]; \mathbf{R}^d)$. If we wish to indicate the dependence on the terminal value ξ as well, then we will use $Y(\xi, V)$ and $M(\xi, V)$ in places of $Y(V)$ and $M(V)$, respectively.

Note that $(Y(V)_t)_{t\in[\tau,T]}$ does not depend on the initial value V_τ, an important fact we will use in our construction of global solutions for the terminal value problem (1.11).

We consider the following type of backward stochastic differential equations:

$$dY_t^j = -f_0^j(t, Y_t, L(M)_t)\,dt - \sum_{i=1}^d f_i^j(t, Y_t)\,dB_t^i + dM_t^j, \qquad Y_T^j = \xi^j, \tag{2.3}$$

on the filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$ $(j = 1, \dots, d')$, where B is a d-dimensional Brownian motion on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$ as given, $T > 0$ is the terminal time, $\xi^j \in L^2(\Omega, \mathcal{F}_T, \mathbf{P})$ (for $j = 1, \dots, d'$) are terminal values, f_i^j $(i = 0, \dots, d$ and $j = 1, \dots, d')$ are locally bounded and Borel measurable, and L is a prescribed mapping on $\mathcal{M}^2([\tau, T]; \mathbf{R}^{d'})$ valued in $\mathcal{H}^2([\tau, T]; \mathbf{R}^{d'\times d})$ or in $\mathcal{C}([\tau, T]; \mathbf{R}^{d'})$.

A solution of (2.3) backward to time τ is a pair of adapted processes $(Y_t, M_t)_{t\in[\tau,T]}$ where $M^j = (M_t^j)_{t\in[\tau,T]}$ are square-integrable martingales and $Y_t^j = (Y_t^j)_{t\in[\tau,T]}$ are special semimartingales with continuous variation parts, which satisfies the integral equations

$$\begin{aligned} Y_t^j - \xi^j = {}& \int_t^T f_0^j(s, Y_s, L(M)_s)\,ds + \sum_{i=1}^d \int_t^T f_i^j(s, Y_s)\,dB_s^i \\ & + M_t^j - M_T^j \end{aligned} \tag{2.4}$$

for $t \in [\tau, T]$, $j = 1, \dots, d'$.

As we have seen in the Introduction, by writing $Y_t = M_t - V_t$, a solution (Y, M) to (2.4) is equivalent to a solution V of the functional integral equation

$$V_t - V_\tau = \int_\tau^t f_0(s, Y(V)_s, L(M(V))_s)\,ds + \sum_{i=1}^d \int_\tau^t f_i(s, Y(V)_s)\,dB_s^i, \tag{2.5}$$

where $M(V)_t = \mathbf{E}(\xi + V_T|\mathcal{F}_t)$ and $Y(V)_t = M(V)_t - V_t$ for $t \in [\tau, T]$. It is the integral equation (2.5) we are going to study.

The following standard assumptions are always imposed on our backward SDE (2.3). Additional conditions on L will be introduced later on to ensure local and global existence.

(1) $f_0 = (f_0^j)_{j \le d'}$ are Lipschitz continuous on $[0, \infty) \times \mathbf{R}^{d'} \times \mathbf{R}^m$ and $f_i = (f_i^j)_{j \le d'}$ $(i = 1, \dots, d)$ Lipschitz continuous on $[0, \infty) \times \mathbf{R}^{d'}$: there is a constant C_2 such that

$$|f_0(t, y, z)| \le C_2(1 + t + |y| + |z|),$$
$$|f_0(t, y, z) - f_0(t, y', z')| \le C_2(|y - y'| + |z - z'|),$$
$$|f_i(t, y)| \le C_2(1 + t + |y|)$$

and

$$|f_i(t, y) - f_i(t, y')| \le C_2|y - y'|$$

for $t \ge 0$, all $y, y' \in \mathbf{R}^{d'}$ and $z, z' \in \mathbf{R}^m$.

(2) The terminal value $\xi = (\xi^i)_{i=1,\dots,d'}$, $\xi^i \in L^2(\Omega, \mathcal{F}_T, \mathbf{P})$.

3. Local solutions and uniqueness. In this section, we prove two results: the uniqueness and the existence of a local solution to (2.3) under the assumption that L is Lipschitz continuous

(3) $L : \mathcal{M}^2([\tau, T]; \mathbf{R}^{d'}) \to \mathcal{H}^2([\tau, T]; \mathbf{R}^m)$ (resp., $\mathcal{C}([\tau, T]; \mathbf{R}^m)$):

$$\|L(M) - L(\tilde{M})\|_{\mathcal{H}^2} \le C_1 \|M - \tilde{M}\|_{\mathcal{C}}$$

[resp.,

$$\|L(M) - L(\tilde{M})\|_{\mathcal{C}} \le C_1 \|M - \tilde{M}\|_{\mathcal{C}}]$$

for any $M, \tilde{M} \in \mathcal{M}^2([\tau, T]; \mathbf{R}^{d'})$, where $\|M\|_{\mathcal{C}}$ means $\|M\|_{\mathcal{C}([\tau,T];\mathbf{R}^{d'})}$ etc. for simplicity.

By "local solution" we mean a solution from T back to τ, where $T - \tau$ is smaller than a certain constant depending on L and f_i^j.

In order to prove the uniqueness, we need to consider BSDE in a more general form than (2.3). More precisely, we are given another Brownian motion $W = (W^1, \dots, W^{m'})$ on $(\Omega, \mathcal{F}_t, \mathcal{F}, \mathbf{P})$ and $g_k : \mathbf{R}_+ \times \mathbf{R}^{d'} \to \mathbf{R}^{d'}$ which are Lipschitz continuous

$$|g_k(t, y)| \le C_2(1 + t + |y|)$$

and

$$|g_k(t, y) - g_k(t, y')| \le C_2|y - y'|$$

for all $t \geq 0$, $y, y' \in \mathbf{R}^{d'}$, $k = 1, \ldots, m'$. Define

$$\begin{aligned}\mathbf{L} : \mathcal{M}^2([\tau, T]; \mathbf{R}^{d'}) \times \mathcal{S}([\tau, T]; \mathbf{R}^{d'}) \\ \to \mathcal{H}^2([\tau, T]; \mathbf{R}^m) \qquad (\text{resp., } \mathcal{C}([\tau, T]; \mathbf{R}^m))\end{aligned}$$

by

$$\mathbf{L}(M, Y) = L\left(M - \sum_{k=1}^{m'} \int_\tau^\cdot g_k(s, Y_s)\, dW_s^k\right). \tag{3.1}$$

Consider the following mapping $\mathbb{L}$ defined on $\mathcal{C}_0([\tau, T]; \mathbf{R}^{d'})$, the space of all processes in $\mathcal{C}([\tau, T]; \mathbf{R}^{d'})$ with initial data $V_\tau = 0$, by

$$\begin{aligned}\mathbb{L}(V)_t = \int_\tau^t f_0(s, Y(V)_s, \mathbf{L}(M(V), Y(V))_s)\, ds \\ + \sum_{i=1}^d \int_\tau^t f_i(s, Y(V)_s)\, dB_s^i,\end{aligned} \tag{3.2}$$

where $M(V)_t = \mathbf{E}(\xi + V_T | \mathcal{F}_t)$ and $Y(V)_t = M(V)_t - V_t$ for $t \in [\tau, T]$, so that $Y(V)_T = \xi$. As we have seen, the functional integral equation: $V = \mathbb{L}(V)$, is equivalent to the following BSDE:

$$dY_t^j = -f_0^j(t, Y_t, \mathbf{L}(M, Y)_t)\, dt - \sum_{i=1}^d f_i^j(t, Y_t)\, dB_t^i + dM_t^j, \qquad Y_T = \xi. \tag{3.3}$$

LEMMA 3.1. $\mathbf{L}$ *defined by* (3.1) *is Lipschitz continuous*

$$\begin{aligned}&\|\mathbf{L}(M, Y) - \mathbf{L}(\tilde{M}, \tilde{Y})\|_{\mathcal{H}^2[\tau, T]} \\ &\quad \leq C_1 \|M - \tilde{M}\|_{\mathcal{C}[\tau, T]} + \frac{m' C_1 C_2}{\sqrt{2}} (T - \tau) \|Y - \tilde{Y}\|_{\mathcal{C}[\tau, T]}\end{aligned} \tag{3.4}$$

and

$$\begin{aligned}&\|\mathbf{L}(M, Y) - \mathbf{L}(\tilde{M}, \tilde{Y})\|_{\mathcal{C}[\tau, T]} \\ &\quad \leq C_1 \|M - \tilde{M}\|_{\mathcal{C}[\tau, T]} + 2m' C_1 C_2 \sqrt{T - \tau} \|Y - \tilde{Y}\|_{\mathcal{C}[\tau, T]}\end{aligned} \tag{3.5}$$

for any $M, \tilde{M} \in \mathcal{M}^2([\tau, T]; \mathbf{R}^{d'})$ *and* $Y, \tilde{Y} \in \mathcal{C}([\tau, T]; \mathbf{R}^{d'})$.

PROOF. Let us omit the subscript $[\tau, T]$ for simplicity. Then

$$\begin{aligned}&\|\mathbf{L}(M, Y) - \mathbf{L}(\tilde{M}, \tilde{Y})\|_{\mathcal{H}^2} \\ &\quad \leq C_1 \|M - \tilde{M}\|_{\mathcal{C}} + C_1 \sum_{k=1}^{m'} \left\| \int_\tau^\cdot (g_k(s, Y_s) - g_k(s, \tilde{Y}_s))\, dW_s^k \right\|_{\mathcal{H}^2}\end{aligned}$$

$$= C_1\|M - \tilde{M}\|_{\mathcal{C}} + C_1 \sum_{k=1}^{m} \sqrt{\mathbf{E}\int_\tau^T \left|\int_\tau^t \big(g_k(s, Y_s) - g_k(s, \tilde{Y}_s)\big)\, dW_s^k\right|^2 dt}$$

$$= C_1\|M - \tilde{M}\|_{\mathcal{C}} + C_1 \sum_{k=1}^{m'} \sqrt{\mathbf{E}\int_\tau^T \int_\tau^t \big|\big(g_k(s, Y_s) - g_k(s, \tilde{Y}_s)\big)\big|^2\, ds\, dt}$$

$$\le C_1\|M - \tilde{M}\|_{\mathcal{C}} + m'C_1C_2\sqrt{\mathbf{E}\int_\tau^T \int_\tau^t |Y_s - \tilde{Y}_s|^2\, ds\, dt}$$

$$\le C_1\|M - \tilde{M}\|_{\mathcal{C}} + \frac{m'C_1C_2}{\sqrt{2}}(T - \tau)\|Y - \tilde{Y}\|_{\mathcal{C}}.$$

The proof of the second inequality is similar. $\square$

The following is our basic local existence theorem.

THEOREM 3.2. *Under the assumptions on L, f_j^i and g_j^i described above. Let*

$$l_1 = \frac{1}{C_2^2[4C_1 + 6(1 + 2\sqrt{d}) + 3\sqrt{2}m'C_1C_2]^2} \wedge 1, \tag{3.6}$$

which depends on the Lipschitz constants C_1, C_2 and the dimensions, but is independent of the terminal data ξ. Suppose that $T - \tau \le l_1$, then $\mathbb{L}$ admits a unique fixed point on $\mathcal{C}_0([\tau, T]; \mathbf{R}^{d'})$.

PROOF. The proof is a standard use of the fixed point theorem applying to $\mathbb{L}$. To this end, we need to show that $\mathbb{L}$ is a contraction on $\mathcal{C}_0([\tau, T]; \mathbf{R}^{d'})$ as long as $T - \tau \le l_1$. This can be done by devising a priori estimates for $\mathbb{L}$. Let us prove the case that $L : \mathcal{M}^2([\tau, T]; \mathbf{R}^{d'}) \to \mathcal{H}^2([\tau, T]; \mathbf{R}^m)$ is Lipschitz; the other case can be treated similarly. To simplify our notation, let $\delta \equiv T - l_1$ be the life duration. Since

$$\|\mathbb{L}(V)\|_{\mathcal{C}} \le \sqrt{\delta}\sqrt{\mathbf{E}\int_\tau^T |f_0(s, Y(V)_s, \mathbf{L}(M(V), Y(V))_s)|^2\, ds}$$
$$+ 2\sqrt{\sum_{i=1}^{d} \mathbf{E}\int_\tau^T |f_i(s, Y(V)_s)|^2\, ds}.$$

Since f_0 and f_i are Lipschitz continuous, so that

$$\|\mathbb{L}(V)\|_{\mathcal{C}} \le 2C_2\big(\sqrt{\delta} + \sqrt{d}\big)\sqrt{\int_\tau^T (1 + s)^2\, ds}$$
$$+ 2C_2\big(\sqrt{\delta} + \sqrt{d}\big)\sqrt{\int_\tau^T \mathbf{E}|Y(V)_s|^2\, ds} \tag{3.7}$$
$$+ 2C_2\sqrt{\delta}\|\mathbf{L}(M(V), Y(V))\|_{\mathcal{H}^2}.$$

Together with the elementary estimates

$$\|Y(V)\|_{\mathcal{C}} \le 2\sqrt{\mathbf{E}|\xi|^2} + 3\|V\|_{\mathcal{C}}$$

and

$$\|M(V)\|_{\mathcal{C}} \le 2\sqrt{\mathbf{E}|\xi|^2} + 2\|V\|_{\mathcal{C}},$$

one deduces that

$$\begin{aligned}
\|\mathbb{L}(V)\|_{\mathcal{C}} \le{}& \frac{2}{\sqrt{3}} C_2(\sqrt{\delta} + \sqrt{3d})\delta\sqrt{\delta} \\
&+ 2\big[\sqrt{2}m' C_1 C_2^2 \delta + 2C_2\sqrt{\delta} + 2C_2\sqrt{d} + 2C_2C_1\big]\sqrt{\delta}\sqrt{\mathbf{E}|\xi|^2} \\
&+ \big[3\sqrt{2}m' C_1 C_2^2\delta + 6C_2\sqrt{\delta} + 4C_2C_1 + 6C_2\sqrt{d}\big]\sqrt{\delta}\|V\|_{\mathcal{C}}.
\end{aligned} \tag{3.8}$$

Similarly, for $V, \tilde{V} \in \mathcal{C}[\tau, T]$ such that $V_\tau = \tilde{V}_\tau = 0$ one has

$$\begin{aligned}
&\|\mathbb{L}(V) - \mathbb{L}(\tilde{V})\|_{\mathcal{C}} \\
&\quad \le \sqrt{\mathbf{E}\bigg(\int_\tau^T |f_0(s, Y_s, \mathbf{L}(M, Y)_s) - f_0(s, \tilde{Y}_s, \mathbf{L}(\tilde{M}, \tilde{Y})_s)|\, ds\bigg)^2} \\
&\qquad + \sqrt{\mathbf{E} \sup_{t\in[\tau,T]} \bigg|\sum_{i=1}^d \int_\tau^t [f_i(s, Y_s) - f_i(s, \tilde{Y}_s)]\, dB_s^i\bigg|^2},
\end{aligned}$$

where $M_t = \mathbf{E}(\xi + V_T|\mathcal{F}_t)$, $\tilde{M}_t = \mathbf{E}(\xi + \tilde{V}_T|\mathcal{F}_t)$, $Y_t = M_t - V_t$ and $\tilde{Y}_t = M_t - V_t$. Since f_i are Lipschitz continuous, so that

$$\begin{aligned}
&\sqrt{\mathbf{E}\bigg(\int_\tau^T |f_0(s, Y_s, \mathbf{L}(M, Y)_s) - f_0(s, \tilde{Y}_s, \mathbf{L}(\tilde{M}, \tilde{Y})_s)|\, ds\bigg)^2} \\
&\quad \le C_2 \sqrt{\mathbf{E}\bigg[\int_\tau^T \big(|Y_s - \tilde{Y}_s| + |\mathbf{L}(M, Y)_s - \mathbf{L}(\tilde{M}, \tilde{Y})_s|\big)\, ds\bigg]^2} \\
&\quad \le C_2\sqrt{\delta}\sqrt{\mathbf{E}\int_\tau^T \big(|Y_s - \tilde{Y}_s| + |\mathbf{L}(M, Y)_s - \mathbf{L}(\tilde{M}, \tilde{Y})_s|\big)^2\, ds} \\
&\quad \le C_2\delta\|Y - \tilde{Y}\|_{\mathcal{C}} + C_2\sqrt{\delta}\|\mathbf{L}(M, Y) - \mathbf{L}(\tilde{M}, \tilde{Y})\|_{\mathcal{H}^2} \\
&\quad \le C_2\bigg[1 + \sqrt{\delta}\frac{m' C_1 C_2}{\sqrt{2}}\bigg]\delta\|Y - \tilde{Y}\|_{\mathcal{C}} \\
&\qquad + C_2 C_1\sqrt{\delta}\|M - \tilde{M}\|_{\mathcal{C}},
\end{aligned}$$

where the last inequality follows from (3.5). Itô's integration term can be treated similarly. Applying Doob's inequality, one has

$$\begin{aligned}
&\sqrt{\mathbf{E}\sup_{t\in[\tau,T]}\left|\sum_{i=1}^{d}\int_{\tau}^{t}[f_i(s,Y_s)-f_i(s,\tilde{Y}_s)]\,dB_s^i\right|^2}\\
&\quad\le 2\sqrt{\mathbf{E}\left|\sum_{i=1}^{d}\int_{\tau}^{T}[f_i(s,Y_s)-f_i(s,\tilde{Y}_s)]\,dB_s^i\right|^2}\\
&\quad\le 2C_2\sqrt{d}\sqrt{\mathbf{E}\int_{\tau}^{T}|Y_s-\tilde{Y}_s|^2\,ds}\\
&\quad\le 2C_2\sqrt{d}\sqrt{\delta}\|Y-\tilde{Y}\|_{\mathcal{C}}.
\end{aligned}$$

Putting these estimates together we obtain

$$\begin{aligned}
\|\mathbb{L}(V)-\mathbb{L}(\tilde{V})\|_{\mathcal{C}} &\le C_2\left[1+\sqrt{\delta}\frac{m'C_1C_2}{\sqrt{2}}\right]\delta\|Y-\tilde{Y}\|_{\mathcal{C}}\\
&\quad + C_2\big(C_1+2\sqrt{d}\big)\sqrt{\delta}\|M-\tilde{M}\|_{\mathcal{C}}.
\end{aligned}\tag{3.9}$$

On the other hand it is easy to see that

$$\begin{aligned}
\|M-\tilde{M}\|_{\mathcal{C}} &= \sqrt{\mathbf{E}\sup_{t\in[\tau,T]}\mathbf{E}(V_T-\tilde{V}_T|\mathcal{F}_t)^2}\\
&\le 2\|V-\tilde{V}\|_{\mathcal{C}}
\end{aligned}$$

and

$$\|Y-\tilde{Y}\|_{\mathcal{C}} \le 3\|V-\tilde{V}\|_{\mathcal{C}}.$$

Inserting these estimates into (3.9) we finally obtain

$$\begin{aligned}
&\|\mathbb{L}(V)-\mathbb{L}(\tilde{V})\|_{\mathcal{C}}\\
&\quad\le C_2\left[2C_1+6\sqrt{d}+3\sqrt{\delta}+3\delta\frac{m'C_1C_2}{\sqrt{2}}\right]\sqrt{\delta}\|V-\tilde{V}\|_{\mathcal{C}}.
\end{aligned}\tag{3.10}$$

Since $\delta \le l_1$, the constant in front of the norm on the right-hand side is less than $\frac{1}{2}$, so that

$$\|\mathbb{L}(V)-\mathbb{L}(\tilde{V})\|_{\mathcal{C}} \le \tfrac{1}{2}\|V-\tilde{V}\|_{\mathcal{C}}.$$

Therefore $\mathbb{L}$ is a contraction on $\mathcal{C}_0([\tau,T];\mathbf{R}^d)$ as long as $T-\tau\le l_1$, so there is a unique fixed point in $\mathcal{C}_0[\tau,T]$. This completes the proof. □

We are now in a position to show the local existence and uniqueness of solutions to BSDE (2.3).

THEOREM 3.3. *Let L, f_j^i be Lipschitz continuous with Lipschitz constants C_1, C_2 and*

$$l_2 = \frac{1}{C_2^2[4C_1 + 6(1 + 2\sqrt{d}) + 3\sqrt{2}d' C_1 C_2]^2} \wedge 1,$$

which is independent of the terminal data $\xi \in L^2(\Omega, \mathcal{F}_T, \mathbf{P})$. Suppose that $T - \tau \leq l_2$ and $L(M) = L(M - M_\tau)$ for any $M \in \mathcal{M}^2([\tau, T]; \mathbf{R}^{d'})$. Then there is a pair (Y, M), where $Y = (Y_t)_{t \in [\tau, T]}$ is a special semimartingale, $M = (M_t)_{t \in [\tau, T]}$ is a square-integrable martingale, which solves the backward stochastic differential equation (2.3) to time τ. Moreover, such a pair of solution is unique in the sense that if (Y, M) and $(\tilde{Y}, \tilde{M})$ are two pairs of solutions, then $Y = \tilde{Y}$ and $M - M_\tau = \tilde{M} - \tilde{M}_\tau$ on $[\tau, T]$.

PROOF. By Theorem 3.2 (applying to the case that all $g_k = 0$), there is a unique $V \in \mathcal{C}_0[\tau, T]$ such that

$$V_t = \int_\tau^t f_0(s, Y_s, L(M)_s)\, ds + \sum_{i=1}^d \int_\tau^t f_i(s, Y_s)\, dB_s^i \qquad \forall t \in [\tau, T],$$

where $M_t = \mathbf{E}(\xi + V_T | \mathcal{F}_t)$ and $Y_t = M_t - V_t$. It is clear that $Y_T = \xi$ and

$$Y_t - \xi = \int_t^T f_0(s, Y_s, L(M)_s)\, ds + \sum_{i=1}^d \int_t^T f_i(s, Y_s)\, dB_s^i + M_t - M_T \tag{3.11}$$

for all $t \in [\tau, T]$. Therefore (Y, M) solves the backward stochastic differential equations (2.3).

Suppose that (Y, M) and $(\hat{Y}, \hat{M})$ are two solutions satisfying (3.11), where Y and $\hat{Y}$ are two special semimartingales. Let

$$Z_t = M_t + \sum_{i=1}^d \int_\tau^t f_i(s, Y_s)\, dB_s^i.$$

Then

$$Y_t - \xi = \int_t^T f_0(s, Y_s, \mathbf{L}(Z, Y)_s)\, ds + Z_t - Z_T \tag{3.12}$$

for $t \in [\tau, T]$, where

$$\mathbf{L}(Z, Y) = L\left(Z - \sum_{i=1}^d \int_\tau^\cdot f_i(s, Y_s)\, dB_s^i\right).$$

It follows that

$$Y_t = \mathbf{E}[\xi + A_T | \mathcal{F}_t] - A_t,$$

where

$$A_t = \int_\tau^t f_0(s, Y_s, \mathbf{L}(Z, Y)_s)\, ds \qquad \forall t \in [\tau, T].$$

Hence $Y_t = Y(A)_t$ and the integral equation (3.12) becomes

$$Y_t = A_T - Z_T + \xi - A_t + Z_t.$$

Since $A_\tau = 0$ so that

$$Y_\tau = A_T - Z_T + \xi + Z_\tau,$$

and thus we may rewrite the previous identity as

$$Y_t = Y_\tau + (Z_t - Z_\tau) - A_t.$$

By the uniqueness of the decompositions for special semimartingales we must have

$$Y_\tau + (Z_t - Z_\tau) = \mathbf{E}[\xi + A_T | \mathcal{F}_t] = M(A)_t.$$

Since $L(M) = L(M - M_\tau)$ for any $M \in \mathcal{M}^2([\tau, T]; \mathbf{R}^{d'})$, so that $\mathbf{L}(Z, Y) = \mathbf{L}(M(A), Y)$. Hence

$$A_t = \int_\tau^t f_0(s, Y(A)_s, \mathbf{L}(M(A), Y(A))_s)\, ds.$$

The same argument applies to $(\tilde{Y}, \tilde{M})$, so that we also have

$$\tilde{A}_t = \int_\tau^t f_0(s, Y(\tilde{A})_s, \mathbf{L}(M(\tilde{A}), Y(\tilde{A}))_s)\, ds.$$

By Theorem 3.2, $A = \tilde{A}$, which yields that $Y = \tilde{Y}$. It follows then

$$Z_t - Z_\tau = \tilde{Z}_t - \tilde{Z}_\tau \qquad \forall t \in [\tau, T]$$

thus $M - M_\tau = \tilde{M} - \tilde{M}_\tau$ which completes the proof. $\square$

One of course wonders whether the global existence can be established, by means of weighted norms, for example, as in the BSDE literature. The present authors were unable to achieve better results than the local existence even with different choices of norms or spaces to which we apply the fixed point theorem. In fact, under the Lipschitz condition only on the mapping L, the local existence is the best we can hope. This is because $L(M)_t$ may depend on the whole path from τ to T, and therefore the corresponding stochastic functional differential equation

$$dV_t = f_0(t, Y(V)_t, L(M(V))_t)\, ds + \sum_{i=1}^{d} f_i(t, Y(V)_t)\, dB_t^i, \qquad V_\tau = 0,$$

is neither local nor Markovian. This can be best demonstrated by its associated differential and integral equation. For example, it is not difficult to show that

$$L_c(M)_t = \sqrt{\mathbf{E}(\langle M^c, M^c\rangle_T - \langle M^c, M^c\rangle_t | \mathcal{F}_t)}$$

for $t \in [\tau, T]$ is Lipschitz continuous, where M^c is its continuous martingale part such that $M_0^c = 0$, and therefore we have

COROLLARY 3.4. *Suppose $T \le l_2$. Then there is a unique special semimartingale $Y = (Y_t)_{t\in[0,T]}$ such that $Y_T = \xi$ and*

$$Y_t - \xi = \int_t^T f_0(s, Y_s, L_c(M))\, ds + M_t - M_T. \tag{3.13}$$

Moreover M is unique up to a random variable measurable with respect to $\mathcal{F}_0$.

Let us apply Corollary 3.4 to the case that $(\mathcal{F}_t)_{t\ge 0}$ is the Brownian filtration of Brownian motion $B = (B^1, \ldots, B^d)$. Then, by Itô's martingale representation theorem,

$$L_c(M)_t = \sqrt{\int_t^T \sum_{i=1}^d \mathbf{E}(|Z_s^i|^2 | \mathcal{F}_t)\, ds},$$

where Z^i are predictable processes such that

$$M_T - M_\tau = \sum_{i=1}^d \int_\tau^T Z_t^i\, dB_t^i.$$

Suppose u is a bounded, smooth function which solves the backward parabolic nonlinear equation

$$\frac{\partial}{\partial t} u + \frac{1}{2}\Delta u + f_0(t, u, K(u)) = 0 \qquad \text{on } [\tau, T] \times R^d, \tag{3.14}$$

with $u(T, \cdot) = \varphi$, where

$$K(u)(t, x) = \sqrt{\int_t^T P_{s-t} |\nabla u|^2(s, x)\, ds},$$

where $(P_t)_{t\ge 0}$ is the heat semi-group in $\mathbf{R}^d$, that is, $P_t = e^{(t\Delta)/2}$. In particular, the differential and integral equation (3.14) is not local, and is a nonlinear equation involving space–time integration and partial derivatives.

Applying Itô's formula to the process $Y_t = u(t, B_t)$ one has

$$\begin{aligned} Y_T - Y_t &= \int_t^T \left(\frac{\partial}{\partial t} + \frac{1}{2}\Delta\right) u(s, B_s)\, ds + M_T - M_t \\ &= -\int_t^T f_0(s, Y_s, K(u)(s, B_s))\, ds + M_T - M_t, \end{aligned}$$

where $M_t = \int_0^t \nabla u(s, B_s)\, dB_s$ is a square-integrable martingale, and one recognizes that

$$
\begin{aligned}
L_c(M)_t &= \sqrt{\mathbf{E}(\langle M, M\rangle_T - \langle M, M\rangle_t | \mathcal{F}_t)} \\
&= \sqrt{\mathbf{E}\biggl(\int_t^T |\nabla u|^2(s, B_s)\, ds | \mathcal{F}_t\biggr)} \\
&= \sqrt{\int_t^T P_{s-t} |\nabla u|^2(s, B_t)\, ds} \\
&= K(u)(t, B_t).
\end{aligned}
$$

Therefore (Y, M) is the unique solution to (3.13), and we have a probability representation

$$
u(t, x) = \mathbf{E}\{Y_t | B_t = x\}.
$$

Since the nonlinear equation (3.14) depends on the "future" of the solution from time T, it is not always possible that a solution exists back to any time τ. In turn, we thus cannot expect that the general BSDE (2.3) have a solution that is global in time without further restrictions on L.

4. Global solutions. In the previous section, under only the Lipschitz conditions on L we are able to construct a solution to the backward stochastic differential equation (2.3) back to time τ such that $T - \tau \le l_2$.

In this section we construct the unique global solution to (2.3) if L satisfies further regularity conditions.

We assume that the mapping $L : \mathcal{M}^2([0, T]; \mathbf{R}^{d'}) \to \mathcal{H}^2([0, T]; \mathbf{R}^m)$ (resp., $\mathcal{C}([0, T]; \mathbf{R}^m)$) satisfies three technical conditions (a), (b) and (c) below: the local-in-time property, the differential property and the Lipschitz condition. The last one is standard, but the first two properties are motivated by the example of density processes in Itô's martingale representations.

For any $[T_2, T_1] \subset [0, T]$, define the restriction

$$
L_{[T_2,T_1]} : \mathcal{M}^2([T_2, T_1]; \mathbf{R}^{d'}) \to \mathcal{H}^2([T_2, T_1]; \mathbf{R}^m) \qquad (\text{resp., } \mathcal{C}([T_2, T_1]; \mathbf{R}^{d'}))
$$

by $L_{[T_2,T_1]}(N)_t = L(\hat{N})_t$ for any $N \in \mathcal{M}^2([T_2, T_1]; \mathbf{R}^{d'})$ and $t \in [T_2, T_1]$, where $\hat{N} \in \mathcal{M}^2([0, T]; \mathbf{R}^{d'})$ defined by $\hat{N}_t = \mathbf{E}(N_{T_1} | \mathcal{F}_t)$ for $t \le T_1$ and $\hat{N}_t = N_{T_1}$ for $t \ge T_1$.

(a) (*Local-in-time property.*) For every pair of nonnegative rational numbers $T_2 < T_1 \le T$, and for any $M \in \mathcal{M}^2([0, T]; \mathbf{R}^{d'})$, $L(M) = L_{[T_2,T_1]}(\tilde{M})$ on (T_2, T_1), where $\tilde{M} = (M_t)_{t\in[T_2,T_1]}$ is restriction of M on $[T_2, T_1]$. The local-in-time property requires that $L(M)_t$ is locally defined, that is, $L(M)_t$ depends only on $(M_s)_{s\in[t,t+\varepsilon)}$ for however small the $\varepsilon > 0$.

(b) (*Differential property*.) For every pair of nonnegative rational numbers $T_1 < T_2 \le T$ and $M \in \mathcal{M}^2([T_2, T_1]; \mathbf{R}^{d'})$, one has $L_{[T_2,T_1]}(M - M_{T_2}) = L_{[T_2,T_1]}(M)$ on (T_2, T_1). The differential property requires that $L_{[T_2,T_1]}(M)_t$ depends only on the increments $\{M_s - M_{T_2} : s \ge t\}$ for $t \in [T_2, T_1]$.

(c) (*Lipschitz continuity*.) $L : \mathcal{M}^2([0, T]; \mathbf{R}^{d'}) \to \mathcal{H}^2([0, T]; \mathbf{R}^m)$ (resp., $\mathcal{C}([0, T]; \mathbf{R}^m)$) is bounded and Lipschitz continuous: there is a constant C_1 such that

$$\|L(M)\|_{\mathcal{H}^2_{[T_2,T_1]}} \le C_1 \|M\|_{\mathcal{C}[T_2,T_1]} \tag{4.1}$$

and

$$\|L(M) - L(\tilde{M})\|_{\mathcal{H}^2_{[T_2,T_1]}} \le C_1 \|M - \tilde{M}\|_{\mathcal{C}[T_2,T_1]} \tag{4.2}$$

[resp.,

$$\|L(M)\|_{\mathcal{C}[T_2,T_1]} \le C_1 \|M\|_{\mathcal{C}[T_2,T_1]} \tag{4.3}$$

and

$$\|L(M) - L(\tilde{M})\|_{\mathcal{C}[T_2,T_1]} \le C_1 \|M - \tilde{M}\|_{\mathcal{C}[T_2,T_1]}] \tag{4.4}$$

for any $M, \tilde{M} \in \mathcal{M}^2([0, T]; \mathbf{R}^{d'})$ and for any rationales T_1 and T_2 such that $0 \le T_2 < T_1 \le T$. That is to say $L_{[T_2,T_1]}$ are Lipschitz continuous with Lipschitz constant independent of $[T_2, T_1] \subset [0, T]$.

The first example below provides the most interesting examples of L in applications, which are, however, variations of the classical example considered in the literature.

EXAMPLE 1. Suppose that $(\mathcal{F}_t)_{t\ge 0}$ is the Brownian filtration generated by a $d + N$-dimensional Brownian motion $B = (B^1, \dots, B^d, W^1, \dots, W^N)$ on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$. If $M \in \mathcal{M}^2([0, T]; \mathbf{R}^{d'})$, then, according to Itô's martingale representation theorem, M is continuous, and there are unique predictable processes $(Z_t^{j,i})_{t\in[0,T]}$ such that

$$M_t^j - M_0^j = \sum_{i=1}^d \int_0^t Z_s^{j,i}\, dB_s^i + \sum_{k=1}^N \int_0^t Z_s^{j,k+d}\, dW_s^k, \qquad j = 1, \dots, d', \tag{4.5}$$

for all $t \in [0, T]$. Assign $M \in \mathcal{M}^2([0, T]; \mathbf{R}^{d'})$ with $L(M) = (Z^{j,i})_{j\le d', i\le d}$. For $0 \le T_2 < T_1 \le T$, the restriction of M on $[T_2, T_1]$, denoted again by M, belongs to $\mathcal{M}^2([T_2, T_1]; \mathbf{R}^{d'})$. By the uniqueness of Itô's representation we can see that L satisfies the local-in-time and differential properties. It is also easy to show that $L : \mathcal{M}^2([0, T]; \mathbf{R}^{d'}) \to \mathcal{H}^2([0, T]; \mathbf{R}^{d'\times d})$ satisfies the Lipschitz condition.

Another class of interesting examples of L is presented in the following example.

EXAMPLE 2. Let $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$ be a filtered probability space which satisfies the technical conditions described at the beginning of Section 2, but not necessary to be a Brownian filtration. Let $B = (B_t)_{t\geq 0}$ be a Brownian motion in $\mathbf{R}^m$ adapted to $(\mathcal{F}_t)_{t\geq 0}$, therefore $(\mathcal{F}_t)_{t\geq 0}$ is in general bigger than the Brownian filtration generated by B. Let $\mathcal{M}_B$ denote the closed stable sub-space of $\mathcal{M}_2$ determined by B, that is,

$$\mathcal{M}_B = \left\{ \sum_{j=1}^m \int_0^{\cdot} Z_s^j \, dB_s^j : Z^j \text{ are predictable and } \mathbf{E} \int_0^T |H_s^j|^2 \, ds < \infty \right\}.$$

Then any martingale M has a unique decomposition

$$M_t - M_0 = \sum_{j=1}^m \int_0^t Z_s^j \, dB_s^j + M_t',$$

where $M' \in \mathcal{M}^2([0, T]; \mathbf{R})$ orthogonal to $\mathcal{M}_B$. Then $L(M) = (Z^j)$ satisfies the local-in-time and differential properties, as well as the Lipschitz condition.

In the following theorems we retain the basic assumptions on the coefficients f_j^i and the terminal values ξ^i.

THEOREM 4.1. *Assume that L satisfies conditions* (a), (b) *and* (c) *listed above. Then there exists a pair of processes (Y, M), where $Y = (Y_t)_{t\in[0,T]}$ is a special semimartingale, and $M = (M_t)_{t\in[0,T]}$ is a square integrable martingale, which solves the backward equation*

$$dY_t = -f_0(t, Y_t, L(M)_t)\, dt - \sum_{i=1}^d f_i(t, Y_t)\, dB_t^i + dM_t, \qquad Y_T = \xi. \tag{4.6}$$

The solution Y is unique, and its martingale correction term M is unique up to a random variable measurable with respect to $\mathcal{F}_0$.

The remainder of this section is devoted to the proof of Theorem 4.1.

PROOF OF THEOREM 4.1. Recall that

$$l_2 = \frac{1}{C_2^2[4C_1 + 6(1 + 2\sqrt{d}) + 3\sqrt{2}dC_1C_2]^2} \wedge 1,$$

which is positive and independent of ξ.

By Theorem 3.2, if the terminal time $T \leq l_2$, the nonlinear mapping $\mathbb{L}$ on $\mathcal{C}_0([0, T]; \mathbf{R}^{d'})$ admits a unique fixed point, where

$$\mathbb{L}(V)_t = \int_0^t f_0(s, Y(V)_s, L(M(V))_s)\, ds + \sum_{i=1}^d \int_0^t f_i(s, Y(V)_s)\, dB_s^i.$$

Next we consider the case $T > l_2$. In this case we divide the interval $[0, T]$ into subintervals with length not exceeding l_2. More precisely, let

$$T = T_0 > T_1 > \cdots > T_k = 0$$

so that $0 < T_{i-1} - T_i \le l_2$ where T_i are rationales except $T_0 = T$.

Begin with the top interval $[T_1, T_0]$, together with the terminal value $Y_{T_0} = \xi$ and the filtration starting from $\mathcal{F}_{T_1}$. Applying Lemma 3.2 to the interval $[T_1, T_0]$ and $\mathbb{L}_1$, where

$$\begin{aligned}(\mathbb{L}_1 V)_t = {} & \int_{T_1}^t f_0\big(s, Y_1(V)_s, L_{[T_1,T_0]}(M_1(V))_s\big)\, ds \\ & + \sum_{i=1}^d \int_{T_1}^t f_i(s, Y_1(V)_s)\, dB_s^i,\end{aligned}$$

where

$$M_1(V)_t = \mathbf{E}(\xi + V_{T_0} | \mathcal{F}_t), \qquad Y_1(V)_t = M_1(V)_t - V_t$$

for any $V \in \mathcal{C}([T_1, T_0]; \mathbf{R}^{d'})$ and $t \in [T_1, T_0]$. Then, there exists a unique $V(1) \in \mathcal{C}_0([T_1, T_0]; \mathbf{R}^{d'})$ such that $\mathbb{L}_1 V(1) = V(1)$.

Repeat the same argument to each interval $[T_j, T_{j-1}]$ (for $2 \le j \le k$) with the terminal value $Y_{j-1}(V(j-1))_{T_{j-1}}$, the filtration starting from $\mathcal{F}_{T_j}$, and the nonlinear mapping $\mathbb{L}_j$ defined on $\mathcal{C}_0([T_j, T_{j-1}]; \mathbf{R}^{d'})$ by

$$\begin{aligned}(\mathbb{L}_j V)_t = {} & \int_{T_j}^t f_0\big(s, Y_j(V)_s, L_{[T_j,T_{j-1}]}(M(V_j))_s\big)\, ds \\ & + \sum_{i=1}^N \int_{T_j}^t f_i(s, Y_j(V)_s)\, dB_s^i,\end{aligned}$$

where $V \in \mathcal{C}([T_j, T_{j-1}]; \mathbf{R}^{d'})$ and

$$\begin{aligned}M_j(V)_t &= \mathbf{E}\big(Y_{j-1}\big(V(j-1)\big)_{T_{j-1}} + V_{T_{j-1}} | \mathcal{F}_t\big), \\ Y_j(V)_t &= M_j(V)_t - V_t\end{aligned}$$

for $t \in [T_j, T_{j-1}]$.

Therefore, for $1 \le j \le k$, there exists a unique $V(j) \in C([T_j, T_{j-1}]; \mathbf{R}^{d'})$ such that

$$\begin{aligned}V(j)_t = {} & \int_{T_j}^t f_0\big(s, Y(j)_s, L_{[T_j,T_{j-1}]}(M(j))_s\big)\, ds \\ & + \sum_{i=1}^N \int_{T_j}^t f_i(s, Y(j)_s)\, dB_s^i\end{aligned}$$

for $t \in [T_j, T_{j-1}]$, where $Y(0)_{T_0} = \xi$, $Y(j-1)_{T_{j-1}} = Y(j)_{T_{j-1}}$ for $2 \le j \le k$, and

$$M(j)_t = \mathbf{E}\big(Y(j-1)_{T_{j-1}} + V(j)_{T_{j-1}} | \mathcal{F}_t\big),$$
$$Y(j)_t = M(j)_t - V(j)_t$$

for $t \in [T_j, T_{j-1}]$.

Since $Y(j-1)_{T_{j-1}} = Y(j)_{T_{j-1}}$ for $2 \le j \le k$, $Y = (Y_t)_{t \in [0,T]}$ given by

$$Y_t = Y(j)_t \qquad \text{if } t \in [T_j, T_{j-1}]$$

for $1 \le j \le k$, is well defined. Define V by shifting it at the partition points,

$$V_t = \begin{cases} V(k)_t, & \text{if } t \in [0, T_{k-1}], \\ V(k-1)_t + V(k)_{T_{k-1}}, & \text{if } t \in [T_{k-1}, T_{k-2}], \\ \cdots & \\ V(1)_t + \displaystyle\sum_{l=2}^{k} V(l)_{T_{l-1}}, & \text{if } t \in [T_1, T]. \end{cases}$$

Then $V \in \mathcal{C}([0, T]; \mathbf{R}^{d'})$. Finally we define

$$M_t = Y_t + V_t \qquad \text{for } t \in [0, T].$$

It remains to show that M is a martingale.

LEMMA 4.2. *M defined above has the expression*

$$M_t = M(j)_t + \sum_{l=j+1}^{k} V(l)_{T_{l-1}} \qquad \text{if } t \in [T_j, T_{j-1}] \tag{4.7}$$

for $1 \le j \le k$, and moreover, M is an $(\mathcal{F}_t)$-martingale up to time T, so that

$$M_t = \mathbf{E}(\xi + V_T | \mathcal{F}_t).$$

PROOF. We first prove the expression (4.7). Since for $1 \le j \le k$,

$$Y(j)_t = M(j)_t - V(j)_t \qquad \text{if } t \in [T_j, T_{j-1}]$$

so that

$$Y_t = M(j)_t + \sum_{l=j+1}^{k} V(l)_{T_{l-1}} - V_t \qquad \text{if } t \in [T_j, T_{j-1}],$$

one may conclude that

$$M_t = M(j)_t + \sum_{l=j+1}^{k} V(l)_{T_{l-1}} \qquad \text{if } t \in [T_j, T_{j-1}].$$

It is clear that M is adapted to $(\mathcal{F}_t)$, so we only need to show $\mathbf{E}(M_t|\mathcal{F}_s) = M_s$ for any $0 \le s \le t \le T$. If $s, t \in [T_j, T_{j-1}]$ for some j, then

$$M_t - M_s = M(j)_t - M(j)_s$$

so that

$$\mathbf{E}(M_t - M_s|\mathcal{F}_s) = \mathbf{E}\big(M(j)_t - M(j)_s|\mathcal{F}_s\big) = 0.$$

If $s \in [T_i, T_{i-1}]$ and $t \in [T_j, T_{j-1}]$ for some $i > j$, then according to (4.7),

$$M_s = M(i)_s + \sum_{l=i+1}^{k} V(l)_{T_{l-1}}$$

and

$$M_t = M(j)_t + \sum_{l=j+1}^{k} V(l)_{T_{l-1}}.$$

Since $M(j)$ is a martingale on $[T_j, T_{j-1}]$ so that

$$\mathbf{E}(M_t|\mathcal{F}_{T_j}) = M(j)_{T_j} + \sum_{l=j+1}^{k} V(l)_{T_{l-1}},$$

conditional on $\mathcal{F}_{T_{j+1}} \subset \mathcal{F}_{T_j}$ we obtain

$$\mathbf{E}(M_t|\mathcal{F}_{T_{j+1}}) = \mathbf{E}\big(M(j)_{T_j} + V(j+1)_{T_j}|\mathcal{F}_{T_{j+1}}\big) + \sum_{l=j+2}^{k} V(l)_{T_{l-1}}. \tag{4.8}$$

On the other hand, $M(j)_{T_j} = Y_{T_j} + V(j)_{T_j} = Y_{T_j}$ so that

$$\begin{aligned}\mathbf{E}\big(M(j)_{T_j} + V(j+1)_{T_j}|\mathcal{F}_{T_{j+1}}\big) &= \mathbf{E}\big(Y_{T_j} + V(j+1)_{T_j}|\mathcal{F}_{T_{j+1}}\big)\\ &= M(j+1)_{T_{j+1}}.\end{aligned}$$

Substituting it into (4.8) we obtain

$$\mathbf{E}(M_t|\mathcal{F}_{T_{j+1}}) = M(j+1)_{T_{j+1}} + \sum_{l=j+2}^{k} V(l)_{T_{l-1}}. \tag{4.9}$$

By repeating the same argument we may establish

$$\mathbf{E}(M_t|\mathcal{F}_{T_{i-1}}) = M(i-1)_{T_{i-1}} + \sum_{l=i}^{k} V(l)_{T_{l-1}}. \tag{4.10}$$

Since $s \in [T_i, T_{i-1}]$, conditional on $\mathcal{F}_s$,

$$\begin{aligned}
\mathbf{E}(M_t|\mathcal{F}_s) &= \mathbf{E}\big(M(i-1)_{T_{i-1}} + V(i)_{T_{i-1}}|\mathcal{F}_s\big) + \sum_{l=i+1}^{k} V(l)_{T_{l-1}} \\
&= \mathbf{E}\big(Y_{T_{i-1}} + V(i)_{T_{i-1}}|\mathcal{F}_s\big) + \sum_{l=i+1}^{k} V(l)_{T_{l-1}} \\
&= M(i)_s + \sum_{l=i+1}^{k} V(l)_{T_{l-1}} \\
&= M_s,
\end{aligned}$$

which proves M is an $\mathcal{F}_t$-adapted martingale up to T. $\square$

Since L satisfies the local-in-time property and the differential property, so that

$$L_{[T_j,T_{j-1}]}(M(V_j))_s = L(M)_s \qquad \text{for } s \in [T_j, T_{j-1}],$$

hence

$$V(j)_t = \int_{T_j}^{t} f_0(s, Y_s, L(M)_s)\,ds + \sum_{i=1}^{d} \int_{T_j}^{t} f_i(s, Y_s)\,dB_s^i$$

for any $t \in [T_j, T_{j-1}]$ and $j = 2, \dots, k$. Therefore

$$V_t = \int_0^t f_0(s, Y_s, L(M)_s)\,ds + \sum_{i=1}^{d} \int_0^t f_i(s, Y_s)\,dB_s^i \qquad \forall t \in [0, T]$$

and $Y = M - V$, $Y_T = \xi$, which together imply that

$$M_t - Y_t = \int_0^t f_0(s, Y_s, L(M)_s)\,ds + \sum_{i=1}^{d} \int_0^t f_i(s, Y_s)\,dB_s^i \qquad \forall t \in [0, T].$$

Thus (Y, M) solves the backward equation (2.3). Uniqueness follows from the fact the solution $(Y(j), M(j) - M(j)_{T_j})$ is unique for any j.

The proof of Theorem 4.1 is complete. $\square$

We end this article with several comments about the main results.

The local and global existence results remain valid even if the driver f_0^j and the diffusion coefficients f_i^j of the BSDE are random as long as the global Lipschitz conditions are maintained. For example, if $f_0^j : \mathbf{R}_+ \times \Omega \times \mathbf{R}^{d'} \times \mathbf{R}^m \to \mathbf{R}^{d'}$ and $f_i^j : \mathbf{R}_+ \times \Omega \times \mathbf{R}^{d'} \to \mathbf{R}^{d'}$ are jointly measurable such that for any special semi-martingale Y and $Z \in \mathcal{H}^2([0,T]; \mathbf{R}^m)$ (resp., $\mathcal{C}([0,T]; \mathbf{R}^m)$), $f_0^j(t, \cdot, Y_t, Z_t)$ and

$f_i^j(t,\cdot,Y_t)$ are progressively measurable and

$$\sqrt{\mathbf{E}\int_{T_2}^{T_1}|f_0^j(t,\cdot,Y_t,Z_t)-f_0^j(t,\cdot,\tilde{Y}_t,\tilde{Z}_t)|^2\,dt}$$
$$\leq C_3\|Y-Y\|_{\mathcal{C}[T_2,T_1]}+C_3\|Z-Z\|_{\mathcal{H}^2[T_2,T_1]}$$

and

$$\sqrt{\mathbf{E}\int_{T_2}^{T_1}|f_i^j(t,\cdot,Y_t)-f_i^j(t,\cdot,\tilde{Y}_t)|^2\,dt}\leq C_3\|Y-Y\|_{\mathcal{C}[T_2,T_1]}$$

for any $[T_2,T_1]\subset[0,T]$, $Y,\tilde{Y}\in\mathcal{S}([0,T];\mathbf{R}^d)$ and $Z,\tilde{Z}\in\mathcal{H}^2([0,T];\mathbf{R}^m)$ (and similarly for the case $Z,\tilde{Z}\in\mathcal{C}([0,T];\mathbf{R}^m)$ with norm $\mathcal{C}[T_2,T_1]$ instead of $\mathcal{H}^2[T_2,T_1]$), then all our local and global results remain true. We leave the details of the proofs for the reader who may be interested in such a generalization.

Acknowledgments. The authors wish to thank Professor Yves LeJan and the referee for their comments and suggestions on the presentation.

REFERENCES

[1] ANTONELLI, F. (1993). Backward–forward stochastic differential equations. *Ann. Appl. Probab.* **3** 777–793. MR1233625

[2] ANTONELLI, F. and MA, J. (2003). Weak solutions of forward–backward SDE's. *Stoch. Anal. Appl.* **21** 493–514. MR1978231

[3] BALLY, V., PARDOUX, E. and STOICA, L. (2005). Backward stochastic differential equations associated to a symmetric Markov process. *Potential Anal.* **22** 17–60. MR2127730

[4] BARLES, G., BUCKDAHN, R. and PARDOUX, E. (1997). Backward stochastic differential equations and integral-partial differential equations. *Stochastics Stochastics Rep.* **60** 57–83. MR1436432

[5] BISMUT, J. M. (1973). Analyse convexe et probabilitiés. These, faculté des sciences de Paris, Paris.

[6] BISMUT, J.-M. (1976). Théorie probabiliste du contrôle des diffusions. *Mem. Amer. Math. Soc.* **4** 1–130. MR0453161

[7] BISMUT, J.-M. (1978). An introductory approach to duality in optimal stochastic control. *SIAM Rev.* **20** 62–78. MR0469466

[8] BRIAND, P., DELYON, B., HU, Y., PARDOUX, E. and STOICA, L. (2003). L^p solutions of backward stochastic differential equations. *Stochastic Process. Appl.* **108** 109–129. MR2008603

[9] BRIAND, P. and HU, Y. (2006). BSDE with quadratic growth and unbounded terminal value. *Probab. Theory Related Fields* **136** 604–618. MR2257138

[10] BRIAND, P. and HU, Y. (2008). Quadratic BSDEs with convex generators and unbounded terminal conditions. *Probab. Theory Related Fields* **141** 543–567. MR2391164

[11] BUCKDAHN, R. and ENGELBERT, H. J. (2007). On the continuity of weak solutions of backward stochastic differential equations. *Teor. Veroyatn. Primen.* **52** 190–199. MR2354579

[12] BUCKDAHN, R., ENGELBERT, H. J. and RĂŞCANU, A. (2004). On weak solutions of backward stochastic differential equations. *Teor. Veroyatn. Primen.* **49** 70–108. MR2141331

[13] CHERIDITO, P., SONER, M., TOUZI, N. and VICTOIR, N. (2007). Second order backward stochastic differential equations and fully non-linear parabolic PDEs. *Comm. Pure Appl. Math.* **60** 1081–1110.

[14] CVITANIĆ, J. and KARATZAS, I. (1996). Backward stochastic differential equations with reflection and Dynkin games. *Ann. Probab.* **24** 2024–2056. MR1415239

[15] DUFFIE, D. and EPSTEIN, L. G. (1992). Stochastic differential utility. With an appendix by the authors and C. Skiadas. *Econometrica* **60** 353–394. MR1162620

[16] EL KAROUI, N., HAMADENE, S. and MATOUSSI, A. (2009). *BSDEs and Applications. Indifference Pricing: Theory and Applications* 267–320. Princeton Univ. Press, Princeton, NJ.

[17] EL KAROUI, N., KAPOUDJIAN, C., PARDOUX, E., PENG, S. and QUENEZ, M. C. (1997). Reflected solutions of backward SDE's, and related obstacle problems for PDE's. *Ann. Probab.* **25** 702–737. MR1434123

[18] EL KAROUI, N. and MAZLIAK, L., eds. (1997). *Backward Stochastic Differential Equations. Pitman Research Notes in Mathematics Series* **364**. Longman, Harlow. MR1752671

[19] EL KAROUI, N., PENG, S. and QUENEZ, M. C. (1997). Backward stochastic differential equations in finance. *Math. Finance* **7** 1–71. MR1434407

[20] GILBARG, D. and TRUDINGER, N. S. (2001). *Elliptic Partial Differential Equations of Second Order*. Springer, Berlin. MR1814364

[21] HAMADENE, S., LEPELTIER, J. P. and MATOUSSI, A. (1997). Double barrier backward SDEs with continuous coefficient. In *Backward Stochastic Differential Equations* (*Paris*, 1995–1996). *Pitman Research Notes in Mathematics Series* **364** 161–175. Longman, Harlow. MR1752681

[22] HU, Y., IMKELLER, P. and MÜLLER, M. (2005). Utility maximization in incomplete markets. *Ann. Appl. Probab.* **15** 1691–1712. MR2152241

[23] HU, Y. and PENG, S. (1995). Solution of forward–backward stochastic differential equations. *Probab. Theory Related Fields* **103** 273–283. MR1355060

[24] KOBYLANSKI, M. (2000). Backward stochastic differential equations and partial differential equations with quadratic growth. *Ann. Probab.* **28** 558–602. MR1782267

[25] KOHLMANN, M. and ZHOU, X. Y. (2000). Relationship between backward stochastic differential equations and stochastic controls: A linear-quadratic approach. *SIAM J. Control Optim.* **38** 1392–1407 (electronic). MR1766421

[26] LEPELTIER, J. P. and SAN MARTIN, J. (1997). Backward stochastic differential equations with continuous coefficient. *Statist. Probab. Lett.* **32** 425–430. MR1602231

[27] MA, J., PROTTER, P. and YONG, J. M. (1994). Solving forward–backward stochastic differential equations explicitly—a four step scheme. *Probab. Theory Related Fields* **98** 339–359. MR1262970

[28] MA, J. and YONG, J. (1999). *Forward–Backward Stochastic Differential Equations and Their Applications. Lecture Notes in Math.* **1702**. Springer, Berlin. MR1704232

[29] MA, J., ZHANG, J. and ZHENG, Z. (2008). Weak solutions for forward–backward SDEs—a martingale problem approach. *Ann. Probab.* **36** 2092–2125. MR2478677

[30] MAJDA, A. (1984). *Compressible Fluid Flow and Systems of Conservation Laws in Several Space Variables. Applied Mathematical Sciences* **53**. Springer, New York. MR0748308

[31] PARDOUX, É. and PENG, S. (1992). Backward stochastic differential equations and quasilinear parabolic partial differential equations. In *Stochastic Partial Differential Equations and Their Applications* (*Charlotte*, *NC*, 1991). *Lecture Notes in Control and Inform. Sci.* **176** 200–217. Springer, Berlin. MR1176785

[32] PARDOUX, É. and PENG, S. G. (1990). Adapted solution of a backward stochastic differential equation. *Systems Control Lett.* **14** 55–61. MR1037747

[33] PENG, S. and WU, Z. (1999). Fully coupled forward–backward stochastic differential equations and applications to optimal control. *SIAM J. Control Optim*. **37** 825–843. MR1675098

[34] PENG, S. G. (1990). A general stochastic maximum principle for optimal control problems. *SIAM J. Control Optim*. **28** 966–979. MR1051633

[35] PENG, S. G. (1991). Probabilistic interpretation for systems of quasilinear parabolic partial differential equations. *Stochastics Stochastics Rep*. **37** 61–74. MR1149116

[36] RONG, S. (1997). On solutions of backward stochastic differential equations with jumps and applications. *Stochastic Process*. *Appl*. **66** 209–236. MR1440399

[37] ROUGE, R. and EL KAROUI, N. (2000). Pricing via utility maximization and entropy. *Math*. *Finance* **10** 259–276. MR1802922

[38] TANG, S. J. and LI, X. J. (1994). Maximum principle for optimal control of distributed parameter stochastic systems with random jumps. In *Differential Equations*, *Dynamical Systems*, *and Control Science*. *Lecture Notes in Pure and Appl*. *Math*. **152** 867–890. Dekker, New York. MR1243240

[39] YONG, J. (1997). Finding adapted solutions of forward–backward stochastic differential equations: Method of continuation. *Probab. Theory Related Fields* **107** 537–572. MR1440146

[40] YONG, J. and ZHOU, X. Y. (1999). *Stochastic Controls*: *Hamiltonian Systems and HJB Equations*. *Applications of Mathematics* (*New York*) **43**. Springer, New York. MR1696772

MATHEMATICAL INSTITUTE
UNIVERSITY OF OXFORD
OXFORD OX1 3LB
UNITED KINGDOM
AND
OXFORD-MAN INSTITUTE
UNIVERSITY OF OXFORD
OXFORD OX2 6ED
UNITED KINGDOM
E-MAIL: liangg@maths.ox.ac.uk
tlyons@maths.ox.ac.uk
qianz@maths.ox.ac.uk