Probab. Theory Relat. Fields (2008) 142:21–77
DOI 10.1007/s00440-007-0098-6

Well-posedness and regularity of backward stochastic Volterra integral equations

Jiongmin Yong

Received: 28 May 2006 / Revised: 8 July 2007 / Published online: 8 August 2007

Abstract Backward stochastic Volterra integral equations (BSVIEs, for short) are studied. Notion of adapted M-solution is introduced. Well-posedness of BSVIEs is established and some regularity results are proved for the adapted M-solutions via Malliavin calculus. A Pontryagin type maximum principle is presented for optimal controls of stochastic Volterra integral equations.

Keywords Backward stochastic Volterra integral equation · Adapted M-solution · Malliavin calculus · Optimal control · Pontryagin maximum principle

Mathematics Subject Classification (2000) 60H20 · 60H07 · 93E20 · 49K22

1 Introduction

Throughout this paper, we let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a complete filtered probability space on which a d-dimensional Brownian motion $W(\cdot)$ is defined with $\mathbb{F} \equiv \{\mathcal{F}_t\}_{t\geq 0}$ being its natural filtration augmented by all the $\mathbb{P}$-null sets in $\mathcal{F}$. Consider the following stochastic integral equation:

$$Y(t) = \psi(t) + \int_t^T g(t, s, Y(s), Z(t, s), Z(s, t))ds - \int_t^T Z(t, s)dW(s), \quad t \in [0, T], \tag{1.1}$$

This work is supported in part by NSF Grant DMS-0604309.

J. Yong (✉)
Department of Mathematics, University of Central Florida, Orlando, FL 32816, USA
e-mail: jyong@mail.ucf.edu

where $g : \Delta^c \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \times \mathbb{R}^{m\times d} \times \Omega \to \mathbb{R}^m$ and $\psi : [0, T] \times \Omega \to \mathbb{R}^m$ are given maps with $\Delta^c = \{(t, s) \in [0, T]^2 \bigm| t < s\}$. Such an equation is referred to as a *backward stochastic Volterra integral equation* (BSVIE, for short). The unknown processes, called an *adapted solution* of (1.1), that we are looking for is the pair $(Y(\cdot), Z(\cdot\,,\cdot))$ valued in $\mathbb{R}^m \times \mathbb{R}^{m\times d}$, with $Y(\cdot)$ being $\mathbb{F}$-adapted, and $Z(t,\cdot)$ being $\mathbb{F}$-adapted for almost all $t \in [0, T]$. Map g is referred to as the *generator* of BSVIE (1.1), and process $\psi(\cdot)$ is referred to as the *free term*.

An important special case of the above equation is the following:

$$Y(t) = \xi + \int_t^T f(s, Y(s), Z(s))ds - \int_t^T Z(s)dW(s), \quad t \in [0, T], \tag{1.2}$$

where $f : [0, T] \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \times \Omega \to \mathbb{R}^m$ and ξ is an $\mathcal{F}_T$-measurable L^p-integrable random variable ($p > 1$). This is the integral form of a so-called backward stochastic differential equation (BSDE, for short); see [5,12,15,18,22] for systematic discussions. A simple glance tells us that BSVIE (1.1) is a natural generalization of BSDE (1.2).

The main features of BSVIE (1.1) are the following: (i) the generator g depends on both t and s, which implies that the equation cannot be reduced to a BSDE in general; (ii) the generator g depends not only on $Z(t, s)$, but also on $Z(s, t)$ (see some comments about this at the end of this section); (iii) The process $\psi(\cdot)$ is allowed to be just $\mathcal{B}[0, T] \otimes \mathcal{F}_T$-measurable (not necessarily $\mathbb{F}$-adapted), where $\mathcal{B}[0, T]$ is the Borel σ-field of $[0, T]$.

There are several interesting problems motivating the study of BSVIEs, besides it being a natural generalization of BSDEs. Let us briefly mention them.

(i) Classical stochastic optimal control problems mainly focus on (forward) stochastic differential equations (FSDEs, for short) [34]. In reality, the state equation might contain memories. One way of describing such situations is to use (forward stochastic) Volterra integral equations (FSVIEs, for short); see [10] with some financial motivations (see also [27] and [23] for general results concerning FSVIEs). Hence, optimal control for FSVIEs becomes a very natural problem. In fact, the deterministic versions were actually discussed by many authors. The readers are referred to [31] (see [20] for a correction), [2,3,6,7,13,19,26,35], and so on. Now, for optimal control of FSVIEs, if one wants to look at the first order necessary conditions for optimal controls (the so-called Pontryagin maximum principle), one has to write the adjoint equation for the variational state equation (which is a linear FSVIE). In FSDE case, the adjoint equation is nothing but a BSDE [4,24,34]. Now, for FSVIE case, the adjoint equation will be a (linear) BSVIE. We will discuss this in more details in Sect. 5, which extends some relevant results of [33].

(ii) Stochastic recursive utility (also called stochastic differential utility) was introduced in [9] (see also [28,29] for some relevant works), which can be identified as the backward component of the adapted solution to a BSDE [12]. Such an identification is mainly due to the time-consistency of the differential utility. On the other hand, time-inconsistent preferences exist in real world [11,16,30]. When the preferences are time-inconsistent, to describe the corresponding recursive utility, one could use BSVIEs. Some further investigation along this line is still under our investigations and relevant results will be reported in a forthcoming paper.

(iii) Coherent risk measurement was introduced in [1]. See also [8,25,32]. These are mainly static or time-consistent. When one wants to study the dynamic version of coherent risk, and allowing possible time-inconsistent preferences, BSVIEs will and should play important roles. Detailed discussions along this line will appear elsewhere.

The above motivations show that study of BSVIEs not only has mathematical interest itself, but also has some interesting and important applications in stochastic optimal control, mathematical finance and risk management.

BSVIE of form (1.1) was firstly studied in [33] with an L^2-framework (A special case in which g is independent of $Z(s,t)$ and $\psi(t) \equiv X$ with X being a fixed $\mathcal{F}_T$-measurable random variable was discussed in [17]). Before going further, let us make a couple of observations on BSVIE (1.1), which will lead to the purposes of this paper. We first make an observation for BSDE (1.2). Suppose (1.2) is solvable on $[T-\delta, T]$ for some $\delta \in (0,T)$. Then the values $(Y(t), Z(t))$ of $(Y(\cdot), Z(\cdot))$ are determined for all $t \in [T-\delta, T]$. Thus, for $t \in [0, T-\delta]$, we may write (1.2) as follows:

$$Y(t) = Y(T-\delta) + \int_t^{T-\delta} g(s, Y(s), Z(s))ds - \int_t^{T-\delta} Z(s)dW(s), \tag{1.3}$$

where

$$Y(T-\delta) = \xi + \int_{T-\delta}^{T} g(s, Y(s), Z(s))ds - \int_{T-\delta}^{T} Z(s)dW(s), \tag{1.4}$$

which is $\mathcal{F}_{T-\delta}$-measurable (and completely determined by $(Y(s), Z(s))$, $s \in [T-\delta, T]$). Hence, (1.3) is a BSDE on $[0, T-\delta]$. Because of this feature, one can establish the well-posedness of BSDE (1.2) first on $[T-\delta, T]$, by martingale representation together with contraction mapping theorem, making use of the uniform Lipschitz continuity of the map $(y,z) \mapsto g(t,y,z)$. Then use the "similar argument" to prove the solvability of (1.3) on $[T-2\delta, T-\delta]$, and so on, eventually obtain the solvability of (1.2) over $[0,T]$.

Now, let us look at BSVIE (1.1). Suppose (1.1) admits an adapted solution $(Y(\cdot), Z(\cdot\,,\cdot))$ on $[T-\delta, T]$ for some $\delta \in (0,T)$. From the equation, we see that at this moment, the following are determined:

$$(Y(t), Z(t,s)), \quad (t,s) \in [T-\delta, T]^2. \tag{1.5}$$

Mimicking BSDEs, for $t \in [0, T-\delta]$, we write equation (1.1) as follows:

$$Y(t) = \psi^{T-\delta}(t) + \int_t^{T-\delta} g(t,s,Y(s),Z(t,s),Z(s,t))ds - \int_t^{T-\delta} Z(t,s)dW(s), \tag{1.6}$$

where [compare with (1.4)]

$$\psi^{T-\delta}(t) = \psi(t) + \int_{T-\delta}^{T} g(t, s, Y(s), Z(t, s), Z(s, t))ds - \int_{T-\delta}^{T} Z(t, s)dW(s), \quad t \in [0, T-\delta]. \tag{1.7}$$

In order (1.6) to be a BSVIE over $[0, T-\delta]$, we need $\psi^{T-\delta}(t)$ to be $\mathcal{F}_{T-\delta}$-measurable for almost all $t \in [0, T-\delta]$. Note that we only have the values $Y(t)$ and $Z(t, s)$ given in (1.5), whereas, in (1.7), one needs the values $Z(t, s)$ not only for $(t, s) \in [0, T-\delta] \times [T-\delta, T]$ (which are needed for defining the last term on the right hand side), but also for $(t, s) \in [T-\delta, T] \times [0, T-\delta]$ [which are needed due to the dependence of g on $Z(s, t)$]. However, these values $Z(t, s)$ of $Z(\cdot\,, \cdot)$ are not available at the moment. Thus, when (1.1) is solvable on $[T-\delta, T]$, one could not simply use the "similar arguments" (as BSDE case) to obtain the solvability of (1.1) over $[T-2\delta, T-\delta]$, etc. Hence, the arguments used in both [17] and [33] to establish the well-posedness of BSVIEs contained some gaps. The above observation also provides an interesting difference between BSDEs and BSVIEs. Note that we may regard (1.7) as a *stochastic Fredholm integral equation* with $(\psi^{T-\delta}(\cdot), Z(\cdot\,, \cdot))$ being the unknown process. We will present the solvability of such an equation in the process of establishing the well-posedness of BSVIEs in this paper. A more general theory for stochastic Fredholm integral equations will be presented in a forthcoming paper.

Next, we look at the following example which will bring us another interesting issue on the adapted solutions to BSVIEs.

Example 1.1 Let $d = 1$. Consider the following BSVIE:

$$Y(t) = \int_{t}^{T} Z(s, t)ds - \int_{t}^{T} Z(t, s)dW(s), \quad t \in [0, T]. \tag{1.8}$$

We can check that

$$\begin{cases} Y(t) = (T-t)\zeta(t), & t \in [0, T], \\ Z(t, s) = I_{[0,t]}(s)\zeta(s), & (t, s) \in [0, T] \times [0, T], \end{cases} \tag{1.9}$$

is an adapted solution of BSVIE (1.8) for any $\zeta(\cdot) \in L^2_{\mathbb{F}}(0, T; \mathbb{R})$, the set of all $\mathbb{F}$-adapted processes $\zeta\colon [0, T] \times \Omega \to \mathbb{R}$ such that $E\int_0^T |\zeta(t)|^2 dt < \infty$. Hence, the adapted solutions defined in an obvious or a "natural" way (as in [33], or [17]) are not unique.

From the above observations, we see that the theory of BSVIEs needs to be re-established. In this paper, we will introduce the notion of adapted M-solution, which refine the definition of adapted solution to BSVIEs introduced in [33]. Then, in

Sect. 3, we will establish the well-posedness of BSVIEs, which involves the solvability of a special type of stochastic Fredholm integral equation (1.7). This will fill the gaps left in [17] and [33]. We will also prove some regularity results for the adapted solutions to BSVIEs, by means of Malliavin calculus in Sect. 4. Further, as an important application, an optimal control problem for FSVIEs, which will extend some relevant results presented in [33].

We would like to emphasize that the dependence of the generator $g(t, s, y, z, \zeta)$ on ζ (i.e., the presence of $Z(s, t)$ on the right hand side of (1.1)) is very important in applications (see [33] and Sect. 5 below). On the other hand, such a dependence makes the BSVIEs nontrivial and that brings main technical difficulties to the solvability of the equation as well as the regularity of adapted M-solutions.

2 Preliminaries

In this section, we present some preliminaries.

2.1 Spaces

This subsection collects some definitions and notations for the spaces that we will use in the later sections. Readers can skip this subsection and come back when it is necessary.

Let $\mathbb{R}^m$ be m-dimensional Euclidean space with the usual L^2-norm, denoted by $|\cdot|$. Let $\mathbb{R}^{m\times d}$ be the Hilbert space of all $(m \times d)$ matrices with the inner product defined by the following:

$$\langle A, B \rangle = \mathrm{tr}\,[AB^T], \quad \forall A, B \in \mathbb{R}^{m\times d}. \tag{2.1}$$

Let $|\cdot|$ be the norm induced by the above inner product. Then one has

$$|A|^2 = \sum_{j=1}^{d} |a_j|^2 \equiv \sum_{i=1}^{m}\sum_{j=1}^{d} a_{ij}^2, \quad \forall A \equiv (a_1, \dots, a_d) \equiv (a_{ij}) \in \mathbb{R}^{m\times d}. \tag{2.2}$$

Next, let $\mathcal{B}(G)$ be the Borel σ-field of metric space G. For any $p, q \in [1, \infty)$, $H = \mathbb{R}^m, \mathbb{R}^{m\times d}$, and $S \in [0, T]$, we define

$$L^p_{\mathcal{F}_S}(\Omega) = \left\{\xi : \Omega \to H \;\middle|\; \xi \text{ is } \mathcal{F}_S\text{-measurable, } E|\xi|^p < \infty\right\},$$

$$L^p_{\mathcal{F}_S}(\Omega; L^q(0,T)) = \Bigg\{\varphi : (0,T)\times\Omega \to H \Big| \varphi(\cdot) \text{ is } \mathcal{B}([0,T])\otimes\mathcal{F}_S\text{-measurable,}$$

$$\times\, E\left(\int_0^T |\varphi(t)|^q dt\right)^{\frac{p}{q}} < \infty\Bigg\},$$

$$L^q(0,T;L^p_{\mathcal{F}_S}(\Omega))=\left\{\varphi:(0,T)\times\Omega\to H\,\Big|\,\varphi(\cdot)\text{ is }\mathcal{B}([0,T])\otimes\mathcal{F}_S\text{-measurable},\right.$$
$$\left.\times\int_0^T\left(E|\varphi(t)|^p\right)^{\frac{q}{p}}dt<\infty\right\}.$$

The spaces $L^\infty_{\mathcal{F}_S}(\Omega;L^q(0,T))$, $L^p_{\mathcal{F}_S}(\Omega;L^\infty(0,T))$, $L^\infty_{\mathcal{F}_S}(\Omega;L^\infty(0,T))$, $L^\infty(0,T;L^p_{\mathcal{F}_S}(\Omega))$, $L^q(0,T;L^\infty_{\mathcal{F}_S}(\Omega))$, and $L^\infty(0,T;L^\infty_{\mathcal{F}_S}(\Omega))$ can be defined in an obvious way. We identify

$$L^p_{\mathcal{F}_S}(\Omega;L^p(0,T))=L^p(0,T;L^p_{\mathcal{F}_S}(\Omega))\equiv L^p_{\mathcal{F}_S}(0,T),\quad p\in[1,\infty].$$

Next, we define

$$\begin{aligned}
&C([0,T];L^p_{\mathcal{F}_S}(\Omega))\\
&=\Big\{\varphi(\cdot)\in L^\infty(0,T;L^p_{\mathcal{F}_S}(\Omega))\;\Big|\;\varphi(t)\text{ is }\mathcal{F}_S\text{-measurable},\ \forall t\in[0,T],\\
&\qquad\varphi(\cdot)\text{ is continuous from }[0,T]\text{ to }L^p_{\mathcal{F}_S}(\Omega),\ \sup_{t\in[0,T]}E|\varphi(t)|^p<\infty\Big\},\\
&C^\#([0,T];L^p_{\mathcal{F}_S}(\Omega))\\
&=\Big\{\varphi(\cdot)\in C([0,T];L^p_{\mathcal{F}_S}(\Omega))\;\Big|\;\varphi(\cdot)\text{ has continuous paths a.s.}\Big\},\\
&L^p_{\mathcal{F}_S}(\Omega;C([0,T]))\\
&=\left\{\varphi(\cdot)\in C^\#([0,T];L^p_{\mathcal{F}_S}(\Omega))\;\Big|\;E\left[\sup_{t\in[0,T]}|\varphi(t)|^p\right]<\infty\right\}.
\end{aligned}$$

For notational simplicity, in the case $S=T$, we will omit the subscript $\mathcal{F}_T$ (unless it should be emphasized). Note that in the above definitions, $\mathbb{F}$-adaptiveness is not involved. We need to further define the following subspaces: For $p,q\in[1,\infty]$,

$$\begin{aligned}
&L^p_{\mathbb{F}}(\Omega;L^q(0,T))=\left\{\varphi(\cdot)\in L^p(\Omega;L^q(0,T))\;\big|\;\varphi(\cdot)\text{ is }\mathbb{F}\text{-adapted}\right\},\\
&L^q_{\mathbb{F}}(0,T;L^p(\Omega))=\left\{\varphi(\cdot)\in L^q(0,T;L^p(\Omega))\;\big|\;\varphi(\cdot)\text{ is }\mathbb{F}\text{-adapted}\right\},\\
&C_{\mathbb{F}}([0,T];L^p(\Omega))=\left\{\varphi(\cdot)\in C([0,T];L^p(\Omega))\;\big|\;\varphi(\cdot)\text{ is }\mathbb{F}\text{-adapted}\right\},\\
&C^\#_{\mathbb{F}}([0,T];L^p(\Omega))=\left\{\varphi(\cdot)\in C^\#([0,T];L^p(\Omega))\;\big|\;\varphi(\cdot)\text{ is }\mathbb{F}\text{-adapted}\right\},\\
&L^p_{\mathbb{F}}(\Omega;C([0,T]))=\left\{\varphi(\cdot)\in L^p(\Omega;C([0,T]))\;\big|\;\varphi(\cdot)\text{ is }\mathbb{F}\text{-adapted}\right\}.
\end{aligned}$$

Also, we identify

$$L^p_{\mathbb{F}}(\Omega;L^p(0,T))=L^p_{\mathbb{F}}(0,T;L^p(\Omega))\equiv L^p_{\mathbb{F}}(0,T),\quad p\in[1,\infty].$$

In the above definitions, we have suppressed the range space H in the notations, for simplicity. When the range space H needs to be emphasized, we will use the notations $L^p_{\mathcal{F}_S}(\Omega; H)$, $L^p_{\mathcal{F}_S}((0,T)\times\Omega; H)$, and so on. We now point out several facts about the spaces defined above. First of all, from the definition, the following chains of inclusions hold:

$$
\begin{aligned}
&L^p_{\mathcal{F}_S}(\Omega; C([0,T]))\subseteq C^{\#}([0,T]; L^p_{\mathcal{F}_S}(\Omega))\subseteq C([0,T]; L^p_{\mathcal{F}_S}(\Omega))\subseteq L^\infty(0,T; L^p_{\mathcal{F}_S}(\Omega))\\
&\quad\subseteq L^q(0,T; L^p_{\mathcal{F}_S}(\Omega))\subseteq L^p_{\mathcal{F}_S}(\Omega; L^q(0,T))\subseteq L^q_{\mathcal{F}_S}(0,T),\quad 1\le q\le p\le\infty,
\end{aligned}
$$

and

$$
\begin{aligned}
&L^\infty_{\mathcal{F}_S}(\Omega; C([0,T])) = C^{\#}([0,T]; L^\infty_{\mathcal{F}_S}(\Omega)) = C([0,T]; L^\infty_{\mathcal{F}_S}(\Omega))\\
&\quad\subseteq L^\infty(0,T; L^\infty_{\mathcal{F}_S}(\Omega))\equiv L^\infty_{\mathcal{F}_S}(0,T)\equiv L^\infty_{\mathcal{F}_S}(\Omega; L^\infty(0,T))\subseteq L^\infty_{\mathcal{F}_S}(\Omega; L^q(0,T))\\
&\quad\subseteq L^p_{\mathcal{F}_S}(\Omega; L^q(0,T))\subseteq L^q(0,T; L^p_{\mathcal{F}_S}(\Omega))\subseteq L^p_{\mathcal{F}_S}(0,T),\quad 1\le p\le q\le\infty.
\end{aligned}
$$

Similar inclusions hold for the corresponding spaces of $\mathbb{F}$-adapted processes.

Second, for $p\in[1,\infty)$, any $\varphi(\cdot)\in C([0,T]; L^p_{\mathcal{F}_S}(\Omega))$ only has the continuity of $t\mapsto\varphi(t)$ as a map from $[0,T]$ to $L^p_{\mathcal{F}_S}(\Omega)$, and it does not necessarily have continuous paths. One should note that space $C^{\#}([0,T]; L^p_{\mathcal{F}_S}(\Omega))$ is not complete under the norm

$$
\|\varphi(\cdot)\|_{C([0,T]; L^p_{\mathcal{F}_S}(\Omega))}\overset{\Delta}{=}\left\{\sup_{t\in[0,T]} E|\varphi(t)|^p\right\}^{\frac{1}{p}}, \tag{2.3}
$$

and $C([0,T]; L^p_{\mathcal{F}_S}(\Omega))$ is the completion of $C^{\#}([0,T]; L^p_{\mathcal{F}_S}(\Omega))$ under norm (2.3).

The above are the spaces for the free term $\psi(\cdot)$ (for which the $\mathbb{F}$-adaptiveness is not required) and $Y(\cdot)$ (for which $\mathbb{F}$-adaptiveness is required). For the process $Z(\cdot\,,\cdot)$, we need to introduce the following spaces. First, for any $p,q\ge 1$, let $L^q(0,T; L^p_{\mathbb{F}}(\Omega; L^2(0,T)))$ be the set of all processes $Z:[0,T]^2\times\Omega\to\mathbb{R}^{m\times d}$ such that for almost all $t\in[0,T]$, $Z(t,\cdot)\in L^p_{\mathbb{F}}(\Omega; L^2(0,T))$ satisfying

$$
\int_0^T\left\{E\left(\int_0^T|Z(t,s)|^2ds\right)^{\frac{p}{2}}\right\}^{\frac{q}{p}}dt<\infty. \tag{2.4}
$$

The spaces $L^\infty(0,T; L^p_{\mathbb{F}}(\Omega; L^2(0,T)))$, $C([0,T]; L^p_{\mathbb{F}}(\Omega; L^2(0,T)))$, and $L^q(0,T; L^\infty_{\mathbb{F}}(\Omega; L^2(0,T)))$ can be defined similarly. Another space that we are going to use is denoted by $L^p_{\mathbb{F}}(\Omega; \widehat{C}([0,T]; L^2(0,T)))$ which consists of all $Z(\cdot\,,\cdot)$ with $Z(t,\cdot)\in L^2_{\mathbb{F}}(0,T)$, for almost all $t\in[0,T]$, such that

$$
E\left[\sup_{t\in[0,T]}\left|\int_0^t Z(t,s)dW(s)\right|^p\right]+E\left[\sup_{t\in[0,T]}\left|\int_t^T Z(t,s)dW(s)\right|^p\right]<\infty. \tag{2.5}
$$

Note that by Martingale Moment Inequalities [14, p. 163], one has

$$\begin{aligned} & E\left[\sup_{t\in[0,T]}\left|\int_0^t Z(t,s)dW(s)\right|^p\right]+E\left[\sup_{t\in[0,T]}\left|\int_t^T Z(t,s)dW(s)\right|^p\right] \\ & \geq E\left[\sup_{t\in[0,T]}\left(\left|\int_0^t Z(t,s)dW(s)\right|^p+\left|\int_t^T Z(t,s)dW(s)\right|^p\right)\right] \\ & \geq 2^{1-p}E\left[\sup_{t\in[0,T]}\left|\int_0^T Z(t,s)dW(s)\right|^p\right] \\ & \geq 2^{1-p}\sup_{t\in[0,T]}E\left|\int_0^T Z(t,s)dW(s)\right|^p \geq C\sup_{t\in[0,T]}E\left(\int_0^T |Z(t,s)|^2ds\right)^{\frac{p}{2}}, \end{aligned} \tag{2.6}$$

hereafter, $C > 0$ always represents a generic constant which can be different at different places. The above shows that

$$L^p_{\mathbb{F}}(\Omega;\widehat{C}([0,T];L^2(0,T)))\subseteq L^\infty(0,T;L^p_{\mathbb{F}}(\Omega;L^2(0,T))). \tag{2.7}$$

On the other hand, it is clear that

$$L^p_{\mathbb{F}}(\Omega;C([0,T];L^2(0,T)))\subseteq L^\infty(0,T;L^p_{\mathbb{F}}(\Omega;L^2(0,T))), \tag{2.8}$$

where the left hand side of the above is defined to be the set of all $Z(\cdot\,,\cdot)\in L^\infty(0,T;$ $L^p_{\mathbb{F}}(\Omega;L^2(0,T)))$ such that $t\mapsto Z(t,\cdot)$ is continuous from $[0,T]$ to $L^2(0,T)$, and

$$E\left(\sup_{t\in[0,T]}\int_0^T |Z(t,s)|^2ds\right)^{\frac{p}{2}}<\infty. \tag{2.9}$$

It is not clear to us if spaces $L^p_{\mathbb{F}}(\Omega;\widehat{C}([0,T];L^2(0,T)))$ and $L^p_{\mathbb{F}}(\Omega;C([0,T];$ $L^2(0,T)))$ have any interesting relations.

Finally, we point out that in all the definitions of the relevant spaces above, $[0,T]$ can be replaced by any $[R,S]$ with $0\leq R<S\leq T$.

Now, we briefly recall some relevant notations and results about Malliavin calculus, which will be used below. Let Ξ be the set of all (scalar) $\mathcal{F}_T$-measurable random variables ξ of form

$$\xi=f\left(\int_0^T h(s)dW(s)\right), \tag{2.10}$$

where $f \in C_b^1(\mathbb{R}^n)$ (the set of all bounded continuously differentiable functions with bounded first order partial derivatives), $h(\cdot) \equiv (h_1(\cdot), \dots, h_d(\cdot))$ with $h_i(\cdot) \in L^\infty_{\mathbb{F}}(0, T; \mathbb{R}^n)$, $1 \le i \le d$, and $n \ge 1$ is an arbitrary natural number. For any $\xi \in \Xi$, define

$$D_r^i \xi = \left\langle f_x \left(\int_0^T h(s) dW(s) \right), h_i(r) \right\rangle, \qquad 0 \le r \le T, \ 1 \le i \le d. \tag{2.11}$$

We call $D_r^i \xi$ the *Malliavin derivative* of ξ with respect to $W_i(\cdot)$. Note that, in general, for each $r \in [0, T]$ and $1 \le i \le d$, $D_r^i \xi$ is still (just) $\mathcal{F}_T$-measurable. Next, for any $\xi \in \Xi$ (of form (2.10)) and $1 \le q < \infty$, we have

$$\begin{aligned} \|\xi\|_{\mathbb{D}_{1,q}} &\stackrel{\Delta}{=} \left\{ E\left[|\xi|^q + \left(\int_0^T \sum_{i=1}^d |D_r^i \xi|^2 dr \right)^{\frac{q}{2}} \right] \right\}^{\frac{1}{q}} \\ &\le C \left\{ 1 + \left[E\left(\int_0^T |h(r)|^2 dr \right)^{\frac{q}{2}} \right]^{\frac{1}{q}} \right\} < \infty, \end{aligned} \tag{2.12}$$

and

$$\begin{aligned} \|\xi\|_{\mathbb{D}_{1,\infty}} &\stackrel{\Delta}{=} \operatorname*{esssup}_{\omega \in \Omega} \left\{ |\xi| + \left(\int_0^T \sum_{i=1}^d |D_r^i \xi|^2 dr \right)^{\frac{1}{2}} \right\} \\ &\le C \left\{ 1 + \operatorname*{esssup}_{\omega \in \Omega} \left(\int_0^T |h(r, \omega)|^2 dr \right)^{\frac{1}{2}} \right\} < \infty. \end{aligned} \tag{2.13}$$

Clearly, $\| \cdot \|_{\mathbb{D}_{1,q}}$ ($q \in [1, \infty]$) is a norm. Let $\mathbb{D}_{1,q}$ be the completion of Ξ under the norm $\| \cdot \|_{\mathbb{D}_{1,q}}$. It is known [21] that operator $D = (D^1, D^2, \dots, D^d)$ admits a closed extension on $\mathbb{D}_{1,q}$, and

$$\xi \text{ is } \mathcal{F}_t\text{-measurable} \quad \Rightarrow \quad D_r^i \xi = 0, \quad \forall r \in (t, T], \ 1 \le i \le d. \tag{2.14}$$

Further, for any $\xi \in \Xi$, we define (comparing with (2.12) and (2.13), respectively)

$$\|\xi\|_{\widetilde{\mathbb{D}}_{1,q}} \overset{\Delta}{=} \left\{ E\left[|\xi|^q + \int_0^T \sum_{i=1}^d |D_r^i \xi|^q dr \right] \right\}^{\frac{1}{q}} \\ \leq C \left\{ 1 + \left[E \int_0^T |h(r)|^q dr \right]^{\frac{1}{q}} \right\} < \infty, \quad q \in [1, \infty), \qquad (2.15)$$

and

$$\|\xi\|_{\widetilde{\mathbb{D}}_{1,\infty}} \overset{\Delta}{=} \operatorname*{esssup}_{\omega \in \Omega} \left\{ |\xi| + \sum_{i=1}^d |D_r^i \xi| \right\} \leq C \left\{ 1 + \operatorname*{esssup}_{r \in [0,T],\, \omega \in \Omega} |h(r, \omega)| \right\} < \infty. \qquad (2.16)$$

Let $\widetilde{\mathbb{D}}_{1,q}$ be the completion of Ξ under norm $\|\cdot\|_{\widetilde{\mathbb{D}}_{1,q}}$ ($q \in [1, \infty]$). By Hölder's inequality, one has the following inclusions:

$$\begin{cases} \mathbb{D}_{1,q} \subseteq \widetilde{\mathbb{D}}_{1,q}, & q \in [1, 2], \\ \widetilde{\mathbb{D}}_{1,q} \subseteq \mathbb{D}_{1,q}, & q \in [2, \infty]. \end{cases} \qquad (2.17)$$

We let $\mathbb{D}_{1,q}^n$ and $\mathbb{D}_{1,q}^{m \times n}$ be the set of all $\mathbb{R}^n$ and $\mathbb{R}^{m \times n}$-valued random variables with each component belonging to $\mathbb{D}_{1,q}$, respectively. The meaning of $\widetilde{\mathbb{D}}_{1,q}^n$ and $\widetilde{\mathbb{D}}_{1,q}^{m \times n}$ are similar. For any k-dimensional random vector η and any random field f defined on $[0, T] \times \mathbb{R}^k \times \Omega$, we will distinguish $D_r^i [f(t, \eta(\omega), \omega)]$ from $D_r^i f(t, x, \omega)\big|_{x=\eta(\omega)} \overset{\Delta}{=} [D_r^i f](t, \eta(\omega), \omega)$. Next, for each $1 \leq q < \infty$ and $0 \leq R < T$, we let $\mathbb{L}^{1,q}(R, T; \mathbb{R}^n)$ (resp. $\widetilde{\mathbb{L}}^{1,q}(R, T; \mathbb{R}^n)$) be the set of all progressively measurable processes $u : [R, T] \times \Omega \to \mathbb{R}^n$ such that for almost all $t \in [R, T]$, $u(t) \in \mathbb{D}_{1,q}^n$ (resp. $\widetilde{\mathbb{D}}_{1,q}^n$), for almost all $r \in [R, T]$ and each $1 \leq i \leq d$, $(t, \omega) \mapsto D_r^i u(t, \omega)$ admits a progressively measurable version, and

$$\|u(\cdot)\|_{\mathbb{L}^{1,q}(R,T;\mathbb{R}^n)}^q \overset{\Delta}{=} E\left\{ \left(\int_R^T |u(t)|^2 dt \right)^{\frac{q}{2}} + \int_R^T \left[\int_r^T \left(\sum_{i=1}^d |D_r^i u(t)|^2 \right) dt \right]^{\frac{q}{2}} dr \right\} < \infty. \qquad (2.18)$$

Similar to [21], for any $u(\cdot) \in \mathbb{L}^{1,q}(0, T; \mathbb{R}^d)$, with $q \geq 2$, we have

$$D_r^i \left[\int_0^t \langle u(s), dW(s) \rangle \right] = \left[u_i(r) + \int_r^t \langle D_r^i u(s), dW(s) \rangle \right] I_{[r \leq t]}, \\ (t, r) \in [0, T]^2, \ 1 \leq i \leq d. \qquad (2.19)$$

In general, the right hand side of the above is (merely) $\mathcal{F}_t$-measurable. From (2.19), it follows that

$$D_r^i\left[\int_t^T \langle u(s), dW(s)\rangle\right] = u_i(r)I_{[r>t]} + \int_t^T \langle D_r^i u(s), dW(s)\rangle, \quad (t,r)\in[0,T]^2,\ 1\le i\le d. \tag{2.20}$$

Further, for any $R, S \in [0, T]$, a direct computation shows that

$$\begin{aligned}\left\| D_\cdot^i\left[\int_S^T \langle u(s), dW(s)\rangle\right]\right\|^q_{L^q(\Omega;L^2(R,T))} &\triangleq E\left(\int_R^T \left|D_r^i\left[\int_S^T \langle u(s), dW(s)\rangle\right]\right|^2 dr\right)^{\frac{q}{2}} \\ &\le 2^{q-1}\left[1 \vee (T-R)^{\frac{q-2}{2}}\right]\|u(\cdot)\|^q_{\mathbb{L}^{1,q}(R,T;\mathbb{R}^d)}. \end{aligned}\tag{2.21}$$

In particular,

$$\begin{aligned}\left\| D_\cdot^i\left[\int_0^t \langle u(s), dW(s)\rangle\right]\right\|^q_{L^q(\Omega;L^2(0,t))} &\triangleq E\left(\int_0^t \left|D_r^i\left[\int_0^t \langle u(s), dW(s)\rangle\right]\right|^2 dr\right)^{\frac{q}{2}} \\ &\le 2^{q-1}(1 \vee t^{\frac{q-2}{2}})\|u(\cdot)\|^q_{\mathbb{L}^{1,q}(0,t;\mathbb{R}^d)}. \end{aligned}\tag{2.22}$$

2.2 BSDEs

In this subsection, we discuss the following BSDE:

$$Y(t) = \xi + \int_t^T f(s, Y(s), Z(s))ds - \int_t^T Z(s)dW(s), \quad t \in [0, T]. \tag{2.23}$$

Let us introduce the following assumption concerning the *generator* f of BSDE (2.23).

(H0) The map $f : [0, T] \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \times \Omega \to \mathbb{R}^m$ is $\mathcal{B}([0, T] \times \mathbb{R}^m \times \mathbb{R}^{m\times d}) \otimes \mathcal{F}_T$-measurable. For each $(y, z) \in \mathbb{R}^m \times \mathbb{R}^{m\times d}$, $t \mapsto f(t, y, z)$ is $\mathbb{F}$-progressive measurable, and $f(\cdot, 0, 0) \in L^p_{\mathbb{F}}(\Omega; L^1(0, T))$, i.e.,

$$E\left(\int_0^T |f(t,0,0)|dt\right)^p < \infty, \tag{2.24}$$

for some $p \in (1,\infty)$. There exist functions $L_y(\cdot) \in L^1(0,T)$ and $L_z(\cdot) \in L^2(0,T)$ such that

$$|f(t,y,z)-f(t,\bar{y},\bar{z})| \leq L_y(t)|y-\bar{y}|+L_z(t)|z-\bar{z}|, \quad t \in [0,T],\ y,\bar{y} \in \mathbb{R}^m,\ z,\bar{z} \in \mathbb{R}^{m\times d}. \tag{2.25}$$

In what follows, for any $p \in (1,\infty)$ and $0 \leq R < S \leq T$, we denote

$$\mathbb{H}^p[R,S] = L^p_{\mathbb{F}}(\Omega; C([R,S])) \times L^p_{\mathbb{F}}(\Omega; L^2(R,S)), \tag{2.26}$$

which is a Banach space under the norm:

$$\|(y(\cdot),z(\cdot))\|_{\mathbb{H}^p[R,S]} = \left\{E\left[\sup_{t\in[R,S]} |y(t)|^p + \left(\int_R^S |z(t)|^2 dt\right)^{\frac{p}{2}}\right]\right\}^{\frac{1}{p}}. \tag{2.27}$$

The following result will be useful later.

Proposition 2.1 *Let* (H0) *hold. Then for any $\xi \in L^p_{\mathcal{F}_T}(\Omega)$, with $p > 1$, BSDE* (2.23) *admits a unique adapted solution $(Y(\cdot), Z(\cdot\,,\cdot)) \in \mathbb{H}^p[0,T]$, and the following estimate holds*:

$$\|(Y(\cdot),Z(\cdot))\|^p_{\mathbb{H}^p[S,T]} \leq CE\left\{|\xi|^p + \left(\int_S^T |f(t,0,0)|dt\right)^p\right\}, \quad S \in [0,T). \tag{2.28}$$

Further, if $\bar{f}$ also satisfies (H0), *$\bar{\xi} \in L^p_{\mathcal{F}_T}(\Omega)$, and $(\bar{Y}(\cdot),\bar{Z}(\cdot)) \in \mathbb{H}^p[0,T]$ is the adapted solution to BSDE* (2.23) *with ξ and f replaced by $\bar{\xi}$ and $\bar{f}$, then*

$$\begin{aligned}&\|(Y(\cdot),Z(\cdot))-(\bar{Y}(\cdot),\bar{Z}(\cdot))\|^p_{\mathbb{H}^p[S,T]}\\ &\leq CE\left\{|\xi-\bar{\xi}|^p + \left(\int_S^T |f(t,Y(t),Z(t))-\bar{f}(t,Y(t),Z(t))|dt\right)^p\right\}, \quad S \in [0,T).\end{aligned} \tag{2.29}$$

Note that in (2.25), $L_y(\cdot)$ and $L_z(\cdot)$ are not constants. Thus, the classical results found in [12,18,22], and [5] do not directly apply, and some modifications are necessary. For reader's convenience, we sketch a proof here.

Proof Take any $(y(\cdot), z(\cdot)) \in \mathbb{H}^p[0, T]$. Consider the following BSDE:

$$Y(t) = \xi + \int_t^T f(s, y(s), z(s))ds - \int_t^T Z(s)dW(s), \quad t \in [0, T]. \tag{2.30}$$

Since

$$\begin{aligned} &E\left|\int_0^T f(s, y(s), z(s))ds\right|^p \\ &\le E\left\{\int_0^T \big(|f(s,0,0)| + L_y(s)|y(s)| + L_z(s)|z(s)|\big)\,ds\right\}^p \\ &\le CE\left\{\left(\int_0^T |f(s,0,0)|ds\right)^p + \left(\int_0^T L_y(s)ds\right)^p \left(\sup_{s\in[0,T]} |y(s)|^p\right)\right. \\ &\quad \left. + \left(\int_0^T L_z(s)^2 ds\right)^{\frac{p}{2}} \left(\int_0^T |z(s)|^2 ds\right)^{\frac{p}{2}}\right\} < \infty. \end{aligned} \tag{2.31}$$

By [12], BSDE (2.30) admits a unique adapted solution $(Y(\cdot), Z(\cdot)) \in \mathbb{H}^p[0, T]$. Thus, for any $S \in [0, T)$, we can define a map $\Phi : \mathbb{H}^p[S, T] \to \mathbb{H}^p[S, T]$ by

$$\Phi(y(\cdot), z(\cdot)) = (Y(\cdot), Z(\cdot)), \quad \forall (y(\cdot), z(\cdot)) \in \mathbb{H}^p[S, T]. \tag{2.32}$$

Next, we will show that Φ is a contraction. To this end, take another $(\bar{y}(\cdot), \bar{z}(\cdot)) \in \mathbb{H}^p[S, T]$, and let $(\bar{Y}(\cdot), \bar{Z}(\cdot)) \in \mathbb{H}^p[S, T]$ be the corresponding adapted solution of BSDE (2.30). Then we have the following:

$$\begin{aligned} &\|(Y(\cdot), Z(\cdot)) - (\bar{Y}(\cdot), \bar{Z}(\cdot))\|^p_{\mathbb{H}^p[S,T]} \\ &\le CE\left\{\int_S^T |f(s, y(s), z(s)) - f(s, \bar{y}(s), \bar{z}(s))|ds\right\}^p \\ &\le CE\left\{\int_S^T \big(L_y(s)|y(s) - \bar{y}(s)| + L_z(s)|z(s) - \bar{z}(s)|\big)\,ds\right\}^p \\ &\le C\left\{\left(\int_S^T L_y(s)ds\right)^p + \left(\int_S^T L_z(s)^2 ds\right)^{\frac{p}{2}}\right\} \|(y(\cdot), z(\cdot)) - (\bar{y}(\cdot), \bar{z}(\cdot))\|^p_{\mathbb{H}^p[S,T]}. \end{aligned} \tag{2.33}$$

By letting $T - S > 0$ small enough, we see that Φ is a contraction on $\mathbb{H}^p[S, T]$. Hence, Φ admits a unique fixed point $(Y(\cdot), Z(\cdot)) \in \mathbb{H}^p[S, T]$, which is the unique adapted solution to BSDE (2.30) on $[S, T]$. By induction, we obtain a unique adapted solution $(Y(\cdot), Z(\cdot)) \in \mathbb{H}^p[0, T]$. The rest of the conclusions are clear. □

3 Well-posedness of BSVIEs

In this section, we will refine the definition of adapted solutions to BSVIE (1.1) introduced in [33], and will establish the well-posedness of BSVIEs. This will fill the gaps that we left in [33] (as well as a similar gap found in [17]). In what follows, for any $0 \le R < S \le T$, we denote

$$\begin{cases} \Delta[R, S] = \Big\{(t, s) \in [R, S]^2 \bigm| R \le s \le t \le S\Big\}, \\ \Delta^c[R, S] = \Big\{(t, s) \in [R, S]^2 \bigm| R \le t < s \le S\Big\} \equiv [R, S]^2 \setminus \Delta[S, T]. \end{cases} \tag{3.1}$$

We simply denote $\Delta[0, T] = \Delta$, $\Delta^c[0, T] = \Delta^c$. Next, for any $0 \le R < S \le T$, we denote

$$\mathcal{H}^p[R, S] = L^p_{\mathbb{F}}(R, S) \times L^p(R, S; L^2_{\mathbb{F}}(R, S)), \quad p \in [1, \infty], \tag{3.2}$$

with the naturally induced norm. Let us first recall the following definition introduced in [33].

Definition 3.1 A pair of processes $(Y(\cdot), Z(\cdot\,,\cdot)) \in \mathcal{H}^2[0, T]$ is called an *adapted solution* of BSVIE (1.1) if (1.1) is satisfied in the usual Itô sense for almost all $t \in [0, T]$.

From Example 1.1, we know that the adapted solution of BSVIE (1.1) defined by the above is not unique. Let us now make a further observation on BSVIE (1.1) to find the precise reason for that. Suppose $(Y(\cdot), Z(\cdot\,,\cdot)) \in \mathcal{H}^2[0, T]$ is an adapted solution of (1.1) (in the sense of Definition 3.1). Let $\zeta(\cdot\,,\cdot) \in L^2(0, T; L^2_{\mathbb{F}}(0, T))$ such that the following fails:

$$Z(t, s) = \zeta(t, s), \quad \text{a.e. } (s, t) \in \Delta, \quad \text{a.s.} \tag{3.3}$$

Consider BSVIE:

$$\widehat{Y}(t) = \psi(t) + \int_t^T g(t, s, \widehat{Y}(t), \widehat{Z}(t, s), \zeta(s, t))ds - \int_t^T \widehat{Z}(t, s)dW(s), \quad t \in [0, T]. \tag{3.4}$$

This can be regarded as a BSVIE with the generator independent of $Z(s, t)$. Under proper conditions (see below), we have an adapted solution $(\widehat{Y}(\cdot), \widehat{Z}(\cdot\,,\cdot)) \in \mathcal{H}^2[0, T]$ to the above. Now, if we redefine $\widehat{Z}(t, s)$ to be $\zeta(t, s)$ for $(t, s) \in \Delta$, then $(\widehat{Y}(\cdot), \widehat{Z}(\cdot\,,\cdot))$ is another adapted solution of (1.1), which is different from $(Y(\cdot), Z(\cdot\,,\cdot))$ since (3.3) does not hold. The point here is that in (3.4), only the values $\widehat{Z}(t, s)$ of $\widehat{Z}(\cdot\,,\cdot)$ for $(t, s) \in \Delta^c$ and only the values $\zeta(t, s)$ of $\zeta(\cdot\,,\cdot)$ for $(t, s) \in \Delta$ are used. Hence, roughly speaking, the values of $Z(\cdot\,,\cdot)$ have some freedom on Δ (as long as $Y(\cdot)$ is allowed to be correspondingly changed). This additional freedom leads to the failure of the uniqueness. Therefore, in order to have the uniqueness, some additional constraints should be imposed on $Z(t,s)$ for $(t,s)\in\Delta$.

Based on the above observation, we are now at the position of introducing the following definition which is a refinement of Definition 3.1.

Definition 3.2 Let $S\in[0,T)$. A pair $(Y(\cdot),Z(\cdot\,,\cdot))\in\mathcal{H}^1[S,T]$ is called an *adapted M-solution* of BSVIE (1.1) on $[S,T]$ if (1.1) holds in the usual Itô's sense for almost all $t\in[S,T]$ and, in addition, the following holds:

$$Y(t)=E[Y(t)\,\big|\,\mathcal{F}_S]+\int_S^t Z(t,s)dW(s),\quad \text{a.e. } t\in[S,T]. \tag{3.5}$$

In the above, "M" in "M-solution" stands for "a martingale representation" (for $Y(t)$ to determine $Z(\cdot\,,\cdot)$ on $\Delta[S,T]$). We will see that additional constraint (3.5) guarantees the uniqueness of corresponding adapted solutions.

It is natural to require that if $(Y(\cdot),Z(\cdot\,,\cdot))$ is an adapted M-solution of (1.1) on $[S,T]$, it should be an adapted M-solution of (1.1) on any $[\bar S,T]$ with $\bar S\in(S,T)$. This is actually the case. In fact, if (3.5) holds, then for any $\bar S\in(S,T)$ and almost all $t\in[\bar S,T]$, one has

$$E[Y(t)\,\big|\,\mathcal{F}_{\bar S}]=E[Y(t)\,\big|\,\mathcal{F}_S]+\int_S^{\bar S} Z(t,s)dW(s). \tag{3.6}$$

Consequently, for almost all $t\in[\bar S,T]$,

$$Y(t)=E[Y(t)\,\big|\,\mathcal{F}_S]+\int_S^t Z(t,s)dW(s)=E[Y(t)\,\big|\,\mathcal{F}_{\bar S}]+\int_{\bar S}^t Z(t,s)dW(s). \tag{3.7}$$

Before going further, let us look at BSDE (2.23) (which is a special case of BSVIE (1.1)). We know that under (H0), BSDE (2.23) admits a unique adapted solution $(Y(\cdot),Z(\cdot))\in\mathbb{H}^p[0,T]$. For each $t\in[0,T]$, we can find a unique $\zeta(t,\cdot)\in L^p_{\mathbb{F}}(\Omega;L^2(0,T))$ such that

$$Y(t)=EY(t)+\int_0^t \zeta(t,s)dW(s). \tag{3.8}$$

Then by defining

$$Z(t,s)=\begin{cases}\zeta(t,s), & (t,s)\in\Delta,\\ Z(s), & (t,s)\in\Delta^c,\end{cases} \tag{3.9}$$

we see that $(Y(\cdot),Z(\cdot\,,\cdot))\in\mathcal{H}^p[0,T]$ is an adapted M-solution of BSVIE (2.23) on $[0,T]$, in the sense of Definition 3.2, and it is actually the only adapted M-solution to such a BSVIE (see Theorem 3.7 below).

Next, for any $R, S \in [0, T)$, we consider the following stochastic integral equation:

$$\lambda(t, r) = \psi(t) + \int_r^T h(t, s, \mu(t, s))ds - \int_r^T \mu(t, s)dW(s), \quad r \in [R, T], \ t \in [S, T], \tag{3.10}$$

where $h : [S, T] \times [R, T] \times \mathbb{R}^{m\times d} \times \Omega \to \mathbb{R}^m$ is given. The unknown process is $(\lambda(\cdot\,, \cdot), \mu(\cdot\,, \cdot))$, for which $(\lambda(t, \cdot), \mu(t, \cdot))$ is $\mathbb{F}$-adapted for all $t \in [R, T]$. We may regard the above as a family of BSDEs on $[R, T]$, parameterized by $t \in [S, T]$, or a family of *stochastic Fredholm integral equations* (SFIEs, for short) on $[S, T]$, parameterized by $r \in [R, T]$. We introduce the following assumption concerning the *generator* h of Eq. (3.10).

(H0)′ Let $R, S \in [0, T)$, and $h : [S, T] \times [R, T] \times \mathbb{R}^{m\times d} \times \Omega \to \mathbb{R}^m$ be $\mathcal{B}([S, T] \times [R, T] \times \mathbb{R}^{m\times d}) \otimes \mathcal{F}_T$-measurable such that $s \mapsto h(t, s, z)$ is $\mathbb{F}$-progressively measurable for all $(t, z) \in [S, T] \times \mathbb{R}^{m\times d}$ and

$$\int_S^T E\left(\int_R^T |h(t, s, 0)|ds\right)^p dt < \infty, \tag{3.11}$$

for some $p > 1$. Moreover, the following holds:

$$|h(t, s, z) - h(t, s, \bar{z})| \leq L(t, s)|z - \bar{z}|, \quad (t, s) \in [S, T] \times [R, T], \ z, \bar{z} \in \mathbb{R}^{m\times d}, \text{ a.s. }, \tag{3.12}$$

where $L : [S, T] \times [R, T] \to [0, \infty)$ is a deterministic function such that for some $\varepsilon > 0$,

$$\sup_{t\in[S,T]} \int_R^T L(t, s)^{2+\varepsilon} ds < \infty. \tag{3.13}$$

We point out that condition (3.13) can be relaxed to that the integrability of $L_z(t, \cdot)^2$ is uniform in $t \in [S, T]$. The following result is a direct consequence of Proposition 2.1.

Lemma 3.3 *Let* (H0)′ *hold. Then for any* $\psi(\cdot) \in L^p_{\mathcal{F}_T}(S, T)$, *Eq.* (3.10), *regarded as a BSDE on* $[R, T]$, *admits a unique adapted solution* $(\lambda(t, \cdot), \mu(t, \cdot)) \in \mathbb{H}^p[R, T]$ *for almost all* $t \in [S, T]$, *and the following estimate holds*:

$$\begin{aligned} &\|(\lambda(t, \cdot), \mu(t, \cdot))\|^p_{\mathbb{H}^p[R,T]} \\ &\equiv E\left\{\sup_{r\in[R,T]} |\lambda(t, r)|^p + \left(\int_R^T |\mu(t, s)|^2 ds\right)^{\frac{p}{2}}\right\} \\ &\leq CE\left\{|\psi(t)|^p + \left(\int_R^T |h(t, s, 0)|ds\right)^p\right\}, \quad \text{a.e. } t \in [S, T]. \end{aligned} \tag{3.14}$$

If $\bar{h} : [S, T] \times [R, T] \times \mathbb{R}^{m\times d} \times \Omega \to \mathbb{R}^m$ *also satisfies* (H0)′, $\bar{\psi}(\cdot) \in L^p_{\mathcal{F}_T}(S, T)$, *and* $(\bar{\lambda}(t, \cdot), \bar{\mu}(t, \cdot)) \in \mathbb{H}^p[R, T]$ *is the unique adapted solution to* (3.10) *on* $[R, T]$ *with* (h, ψ) *replaced by* $(\bar{h}, \bar{\psi})$*, then*

$$
\begin{aligned}
&E\left\{\sup_{r\in[R,T]} |\lambda(t,r) - \bar{\lambda}(t,r)|^p + \left(\int_R^T |\mu(t,s) - \bar{\mu}(t,s)|^2 ds\right)^{\frac{p}{2}}\right\} \\
&\le CE\left\{|\psi(t) - \bar{\psi}(t)|^p + \left(\int_R^T |h(t,s,\mu(t,s)) - \bar{h}(t,s,\mu(t,s))| ds\right)^p\right\}, \\
&\qquad \text{a.e. } t \in [S,T]. \qquad (3.15)
\end{aligned}
$$

In particular,

$$
\begin{aligned}
&E\left\{\sup_{r\in[R,T]} |\lambda(t,r) - \lambda(\bar{t},r)|^p + \left(\int_R^T |\mu(t,s) - \mu(\bar{t},s)|^2 ds\right)^{\frac{p}{2}}\right\} \\
&\le CE\left\{|\psi(t) - \psi(\bar{t})|^p + \left(\int_R^T |h(t,s,\mu(t,s)) - h(\bar{t},s,\mu(t,s))| ds\right)^p\right\}, \\
&\qquad \text{a.e. } t, \bar{t} \in [S,T]. \qquad (3.16)
\end{aligned}
$$

Consequently, if $\psi(\cdot) \in C([0, T]; L^p_{\mathcal{F}_T}(\Omega))$ *and* $t \mapsto h(t, s, \mu)$ *is continuous in the following sense*

$$
\begin{aligned}
|h(t,s,\mu) - h(\bar{t},s,\mu)| \le C(1+|\mu|)\rho(|t-\bar{t}|), \quad & t, \bar{t} \in [S,T], \\
& s \in [R,T],\ \mu \in \mathbb{R}^{m\times d}, \text{a.s.}, \qquad (3.17)
\end{aligned}
$$

for some modulus of continuity $\rho(\cdot)$*, then*

$$
(\lambda(\cdot\,,\cdot), \mu(\cdot\,,\cdot)) \in C([S,T]; \mathbb{H}^p[R,T]). \qquad (3.18)
$$

Remark 3.4 Note that (3.18) implies that $\lambda(\cdot\,,\cdot) \in C([S, T] \times [R, T]; L^p_{\mathcal{F}_T}(\Omega))$, which means that, by definition, the map $(t, r) \mapsto \lambda(t, r)$ is continuous from $[S, T] \times [R, T]$ to $L^p_{\mathcal{F}_T}(\Omega)$. As a matter of fact, for any $t, \bar{t} \in [S, T]$ and $r, \bar{r} \in [R, T]$, from (3.16), we have

$$
\begin{aligned}
E|\lambda(t,r)-\lambda(\bar t,\bar r)|^p &\le 2^{p-1}E\left\{|\lambda(t,r)-\lambda(t,\bar r)|^p+|\lambda(t,\bar r)-\lambda(\bar t,\bar r)|^p\right\}\\
&\le CE\left\{\left(\int_{r\wedge\bar r}^{r\vee\bar r}|h(t,s,\mu(t,s))|ds\right)^p\right.\\
&\quad+\left(\int_{r\wedge\bar r}^{r\vee\bar r}|\mu(t,s)|^2ds\right)^{\frac{p}{2}}+|\psi(t)-\psi(\bar t)|^p\\
&\quad\left.+\left(\int_R^T|h(t,s,\mu(t,s))-h(\bar t,s,\mu(t,s))|ds\right)^p\right\}\to 0,\\
&\qquad\text{as } (\bar t,\bar r)\to(t,r). \qquad (3.19)
\end{aligned}
$$

In particular, for the case $R = S$, the map $t \mapsto \lambda(t,t)$ is continuous from $[S,T]$ to $L^p_{\mathcal{F}_T}(\Omega)$. This will be very useful below.

Now, let us look at two special cases of the above result. First, let $r = S \in [R,T)$ be fixed. Define

$$
\psi^S(t)=\lambda(t,S),\quad Z(t,s)=\mu(t,s),\qquad t\in[R,S],\quad s\in[S,T]. \qquad (3.20)
$$

Then (3.10) reads:

$$
\psi^S(t)=\psi(t)+\int_S^T h(t,s,Z(t,s))ds-\int_S^T Z(t,s)dW(s),\quad t\in[R,S]. \qquad (3.21)
$$

This is an SFIE. Any pair $(\psi^S(\cdot),Z(\cdot\,,\cdot))\in L^p_{\mathcal{F}_S}(R,S)\times L^p(R,S;L^2_{\mathbb{F}}(S,T))$ satisfying (3.21) in the usual Itô sense is called an adapted solution of (3.21). We note that $\psi^S(t)$ is (only) required to be $\mathcal{F}_S$-measurable for almost all $t\in[R,S]$, instead of $\mathbb{F}$-adaptiveness. In general, this is the best possibility since the Itô's integral on the right hand side is merely taken over $[S,T]$. According to Lemma 3.3, we have the following result whose proof is obvious.

Corollary 3.5 *Let* (H0)$'$ *hold. Then for any* $\psi(\cdot)\in L^p_{\mathcal{F}_T}(R,S)$, *SFIE* (3.21) *admits a unique adapted solution* $(\psi^S(\cdot),Z(\cdot\,,\cdot))\in L^p_{\mathcal{F}_S}(R,S)\times L^p(R,S;L^2_{\mathbb{F}}(S,T))$, *and the following estimate holds*:

$$
\begin{aligned}
&E\left\{|\psi^S(t)|^p+\left(\int_S^T|Z(t,s)|^2ds\right)^{\frac{p}{2}}\right\}\\
&\quad\le CE\left\{|\psi(t)|^p+\left(\int_S^T|h(t,s,0)|ds\right)^p\right\},\quad t\in[R,S]. \qquad (3.22)
\end{aligned}
$$

If $\bar h : [R,S]\times[S,T]\times\mathbb{R}^{m\times d}\times\Omega\to\mathbb{R}^m$ *also satisfies* (H0)′, $\bar\psi(\cdot)\in L^p_{\mathcal{F}_T}(R,S)$, *and* $(\bar\psi^S(\cdot),\bar Z(\cdot\,,\cdot))\in L^p_{\mathcal{F}_S}(R,S)\times L^p(R,S;L^2_{\mathbb{F}}(S,T))$ *is the unique adapted solution of SFIE* (3.21) *with* (h,ψ) *replaced by* $(\bar h,\bar\psi)$, *then*

$$E\left\{|\psi^S(t)-\bar\psi^S(t)|^p+\left(\int_S^T|Z(t,s)-\bar Z(t,s)|^2ds\right)^{\frac{p}{2}}\right\}\le CE\left\{|\psi(t)-\bar\psi(t)|^p+\left(\int_S^T|h(t,s,Z(t,s))-\bar h(t,s,Z(t,s))|ds\right)^p\right\},\quad t\in[R,S]. \tag{3.23}$$

Further, for any $t,\bar t\in[R,S]$,

$$E\left\{|\psi^S(t)-\psi^S(\bar t)|^p+\left(\int_S^T|Z(t,s)-Z(\bar t,s)|^2ds\right)^{\frac{p}{2}}\right\}\le CE\left\{|\psi(t)-\psi(\bar t)|^p+\left(\int_S^T|h(t,s,Z(t,s))-h(\bar t,s,Z(t,s))|ds\right)^p\right\}. \tag{3.24}$$

Consequently, if $\psi(\cdot)\in C([S,T];L^p(\Omega))$ *and* $t\mapsto h(t,s,z)$ *is continuous in the sense of* (3.17), *then*

$$(\psi^S(\cdot),Z(\cdot\,,\cdot))\in C([S,T];L^p_{\mathcal{F}_S}(\Omega))\times C([S,T];L^p_{\mathbb{F}}(\Omega;L^2(S,T))). \tag{3.25}$$

The second special case of (3.10) is the following: Let $R=S$, and define

$$\begin{cases} Y(t)=\lambda(t,t), & t\in[S,T],\\ Z(t,s)=\mu(t,s), & (t,s)\in\Delta^c[S,T].\end{cases} \tag{3.26}$$

Then (3.10) reads:

$$Y(t)=\psi(t)+\int_t^T h(t,s,Z(t,s))ds-\int_t^T Z(t,s)dW(s),\quad t\in[S,T]. \tag{3.27}$$

This is a special case of BSVIE (1.1) in which the generator h is independent of (y,ζ). Similar to (3.8), we define $Z(t,s)$ for $(t,s)\in\Delta[S,T]$ by the following relation (making use of the Martingale Representation Theorem):

$$Y(t)=E[Y(t)|\mathcal{F}_S]+\int_S^t Z(t,s)dW(s),\quad t\in[S,T]. \tag{3.28}$$

Note that $Z(t,s)$ and $\mu(t,s)$ might be different for $(t,s)\in\Delta[S,T]$. We have the following result.

Corollary 3.6 *Let* (H0)′ *hold. Then for any* $\psi(\cdot)\in L^p_{\mathcal{F}_T}(S,T)$*, BSVIE* (3.27) *admits a unique adapted M-solution* $(Y(\cdot),Z(\cdot\,,\cdot))\in\mathcal{H}^p[S,T]$*, and the following estimate holds*:

$$
\begin{aligned}
&E\left\{|Y(t)|^p+\left(\int_S^T|Z(t,s)|^2ds\right)^{\frac{p}{2}}\right\}\\
&\quad\le CE\left\{|\psi(t)|^p+\left(\int_t^T|h(t,s,0)|ds\right)^p\right\},\quad t\in[S,T].
\end{aligned}\tag{3.29}
$$

If $\bar h$ *also satisfies* (H0)′, $\bar\psi(\cdot)\in L^p_{\mathcal{F}_T}(S,T)$*, and* $(\bar Y(\cdot),\bar Z(\cdot\,,\cdot))\in\mathcal{H}^p[S,T]$ *is the unique adapted M-solution of BSVIE* (3.27) *with* (h,ψ) *replaced by* $(\bar h,\bar\psi)$*, then*

$$
\begin{aligned}
&E\left\{|Y(t)-\bar Y(t)|^p+\left(\int_S^T|Z(t,s)-\bar Z(t,s)|^2ds\right)^{\frac{p}{2}}\right\}\\
&\le CE\left\{|\psi(t)-\bar\psi(t)|^p+\left(\int_t^T|h(t,s,Z(t,s))-\bar h(t,s,Z(t,s))|ds\right)^p\right\},\\
&\quad t\in[S,T].
\end{aligned}\tag{3.30}
$$

Further, for any $t,\bar t\in[S,T]$*,*

$$
\begin{aligned}
&E\left\{|Y(t)-Y(\bar t)|^p+\left(\int_{t\vee\bar t}^T|Z(t,s)-Z(\bar t,s)|^2ds\right)^{\frac{p}{2}}\right.\\
&\qquad\left.+\left(\int_S^{t\wedge\bar t}|Z(t,s)-Z(\bar t,s)|^2ds\right)^{\frac{p}{2}}\right\}\\
&\le CE\left\{|\psi(t)-\psi(\bar t)|^p+\left(\int_{t\wedge\bar t}^{t\vee\bar t}|h(t,s,Z(t,s))|ds\right)^p+\left(\int_{t\wedge\bar t}^{t\vee\bar t}|Z(t,s)|^2ds\right)^{\frac{p}{2}}\right.\\
&\qquad\left.+\left(\int_{t\vee\bar t}^T|h(t,s,Z(t,s))-h(\bar t,s,Z(t,s))|ds\right)^p\right\}.
\end{aligned}\tag{3.31}
$$

Consequently, if $\psi(\cdot) \in C([S,T]; L^p_{\mathcal{F}_T}(\Omega))$ *and* $t \mapsto h(t,s,z)$ *is continuous in the sense of* (3.17)*, then*

$$\begin{cases} Y(\cdot) \in C_{\mathbb{F}}([S,T]; L^p(\Omega)), \\ Z(\cdot\,,\cdot) \in C([R,T]; L^p_{\mathbb{F}}(\Omega; L^2(S,R)) \bigcap C([S,R]; L^p_{\mathbb{F}}(\Omega; L^2(R,T)), \\ \quad \forall R \in (S,T). \end{cases} \tag{3.32}$$

Proof The existence and uniqueness follows from Lemma 3.3 and relation (3.28). Note that by (3.22) (with $t = R = S$), we have estimate

$$\begin{aligned} &E\left[|Y(t)|^p + \left(\int_t^T |Z(t,s)|^2 ds \right)^{\frac{p}{2}} \right] \\ &\quad \le CE\left\{ |\psi(t)|^p + \left(\int_t^T |h(t,s,0)| ds \right)^p \right\}, \quad t \in [S,T]. \end{aligned} \tag{3.33}$$

On the other hand, by (3.28), using Martingale Moment Inequality [14], we obtain

$$\begin{aligned} E\left(\int_S^t |Z(t,s)|^2 ds \right)^{\frac{p}{2}} &\le CE\left| \int_S^t Z(t,s) dW(s) \right|^p \\ &= CE\Big| Y(t) - E[Y(t)|\mathcal{F}_S]|^p \le CE|Y(t)|^p. \end{aligned} \tag{3.34}$$

Combining the above two, we obtain (3.29). Similarly, we obtain (3.30).

To obtain (3.31), we take $t, \bar{t} \in [S,T]$. Without loss of generality, let $S \le t < \bar{t} \le T$. Then by (3.16), with $R = t \vee \bar{t} = \bar{t}$,

$$\begin{aligned} &E\left\{ \sup_{r \in [\bar{t},T]} |\lambda(t,r) - \lambda(\bar{t},r)|^p + \left(\int_{\bar{t}}^T |\mu(t,s) - \mu(\bar{t},s)|^2 ds \right)^{\frac{p}{2}} \right] \\ &\quad \le CE\left\{ |\psi(t) - \psi(\bar{t})|^p + \left(\int_{\bar{t}}^T |h(t,s,\mu(t,s)) - h(\bar{t},s,\mu(t,s))| ds \right)^p \right\}, \end{aligned} \tag{3.35}$$

which implies

$$
\begin{aligned}
&E\left\{|Y(t)-Y(\bar t)|^p+\left(\int_{\bar t}^T|Z(t,s)-Z(\bar t,s)|^2ds\right)^{\frac p2}\right\}\\
&\le CE\left\{|\lambda(t,t)-\lambda(t,\bar t)|^p+\sup_{r\in[\bar t,T]}|\lambda(t,r)-\lambda(\bar t,r)|^p\right.\\
&\quad\left.+\left(\int_{\bar t}^T|\mu(t,s)-\mu(\bar t,s)|^2ds\right)^{\frac p2}\right]\\
&\le CE\left\{\left(\int_t^{\bar t}|h(t,s,Z(t,s))|ds\right)^p+\left(\int_t^{\bar t}|Z(t,s)|^2ds\right)^{\frac p2}\right.\\
&\quad\left.+|\psi(t)-\psi(\bar t)|^p+\left(\int_{\bar t}^T|h(t,s,Z(t,s))-h(\bar t,s,Z(t,s))|ds\right)^p\right\},
\end{aligned}
\tag{3.36}
$$

Now, from (3.28) with $R=S$, we obtain

$$
\begin{aligned}
&E\left(\int_S^t|Z(t,s)-Z(\bar t,s)|^2ds\right)^{\frac p2}+E\left(\int_t^{\bar t}|Z(\bar t,s)|^2ds\right)^{\frac p2}\\
&\le CE\left|\int_S^tZ(t,s)dW(s)-\int_S^{\bar t}Z(\bar t,s)dW(s)\right|^p\\
&=CE\left|Y(t)-Y(\bar t)-E[Y(t)-Y(\bar t)\mid\mathcal F_S]\right|^p\le CE|Y(t)-Y(\bar t)|^p.
\end{aligned}
\tag{3.37}
$$

Then (3.31) follows.

Finally, from (3.31), it is easy to see that in the case $\psi(\cdot)\in C([S,T];L^p_{\mathcal F_T}(\Omega))$ and $t\mapsto h(t,s,z)$ is continuous in the sense of (3.17), the first inclusion in (3.32) follows easily, by fixing $t\in[S,T]$ and letting $\bar t\to t$ (note that $\bar t$ is allowed to approach to t from both sides!). For the continuity of $Z(\cdot\,,\cdot)$, take any $R\in(S,T)$, from (3.31), one has

$$
\begin{aligned}
&E\left(\int_R^T |Z(t,s)-Z(\bar t,s)|^2 ds\right)^{\frac{p}{2}} \le E\left(\int_{t\vee\bar t}^T |Z(t,s)-Z(\bar t,s)|^2 ds\right)^{\frac{p}{2}}\\
&\le CE\left\{|\psi(t)-\psi(\bar t)|^p + \left(\int_{t\wedge\bar t}^{t\vee\bar t} |h(t,s,Z(t,s))|ds\right)^p + \left(\int_{t\wedge\bar t}^{t\vee\bar t} |Z(t,s)|^2 ds\right)^{\frac{p}{2}}\right.\\
&\quad\left.+\left(\int_{t\vee\bar t}^T |h(t,s,Z(t,s))-h(\bar t,s,Z(t,s))|ds\right)^p\right\}, \quad t,\bar t\in[S,R], \qquad (3.38)
\end{aligned}
$$

and

$$
\begin{aligned}
&E\left(\int_S^R |Z(t,s)-Z(\bar t,s)|^2 ds\right)^{\frac{p}{2}} \le E\left(\int_S^{t\wedge\bar t} |Z(t,s)-Z(\bar t,s)|^2 ds\right)^{\frac{p}{2}}\\
&\le CE\left\{|\psi(t)-\psi(\bar t)|^p + \left(\int_{t\wedge\bar t}^{t\vee\bar t} |h(t,s,Z(t,s))|ds\right)^p + \left(\int_{t\wedge\bar t}^{t\vee\bar t} |Z(t,s)|^2 ds\right)^{\frac{p}{2}}\right.\\
&\quad\left.+\left(\int_{t\vee\bar t}^T |h(t,s,Z(t,s))-h(\bar t,s,Z(t,s))|ds\right)^p\right\}, \quad t,\bar t\in[R,T]. \qquad (3.39)
\end{aligned}
$$

Thus, the second inclusion in (3.32) follows. □

Note that the second inclusion in (3.32) does not mean that $Z(\cdot\,,\cdot)\in C([S,T];L^p_{\mathbb{F}}(\Omega;L^2(S,T))$, which we do not expect. In fact, (3.31) does not give an estimate for $E\left(\int_{t\wedge\bar t}^{t\vee\bar t}|Z(t,s)-Z(\bar t,s)|^2ds\right)^{\frac{p}{2}}$. Recall that the different definitions of $Z(t,s)$ for (t,s) crossing the "diagonal line" $t=s$ might cause some kind of discontinuity for the process $t\mapsto Z(t,\cdot)$ as a map from $[S,T]$ to $L^p_{\mathbb{F}}(\Omega;L^2(S,T))$, in some sense.

The above two corollaries will have some interesting applications below. We now introduce the following standing assumption which will be used below. In what follows, we denote

$$
g_0(t,s)=g(t,s,0,0,0). \qquad (3.40)
$$

(H1) Let $g:\Delta^c\times\mathbb{R}^m\times\mathbb{R}^{m\times d}\times\mathbb{R}^{m\times d}\times\Omega\to\mathbb{R}^m$ be $\mathcal{B}(\Delta^c\times\mathbb{R}^m\times\mathbb{R}^{m\times d}\times\mathbb{R}^{m\times d})\otimes\mathcal{F}_T$-measurable such that $s\mapsto g(t,s,y,z,\zeta)$ is $\mathbb{F}$-progressively measurable for all $(t,y,z,\zeta)\in[0,T]\times\mathbb{R}^m\times\mathbb{R}^{m\times d}\times\mathbb{R}^{m\times d}$ and

$$
E\int_0^T\left(\int_t^T|g_0(t,s)|ds\right)^2dt<\infty. \qquad (3.41)
$$

Moreover, it holds

$$
\begin{aligned}
&|g(t,s,y,z,\zeta)-g(t,s,\bar y,\bar z,\bar\zeta)|\le L(t,s)\Big(|y-\bar y|+|z-\bar z|+|\zeta-\bar\zeta|\Big),\\
&\quad\forall (t,s)\in\Delta^c,\ \ y,\bar y\in\mathbb{R}^m,\ \ z,\bar z,\zeta,\bar\zeta\in\mathbb{R}^{m\times d},\ \text{a.s.}\,, \qquad (3.42)
\end{aligned}
$$

where $L:\Delta^c\to\mathbb{R}$ is a deterministic function such that the following holds:

$$
\sup_{t\in[0,T]}\int_t^T L(t,s)^{2+\varepsilon}ds<\infty, \qquad (3.43)
$$

for some $\varepsilon>0$.

Note that if $L(\cdot\,,\cdot)$ is uniformly bounded, condition (3.43) automatically holds. Similar to (H0)$'$, condition (3.43) can be relaxed to the integrability of $L(t,\cdot)^2$ being uniform in $t\in[0,T]$.

Our main result of this section is the following well-posedness result for BSVIE (1.1).

Theorem 3.7 *Let* (H1) *hold. Then for any* $\psi(\cdot)\in L^2_{\mathcal{F}_T}(0,T)$, *BSVIE* (1.1) *admits a unique adapted M-solution* $(Y(\cdot),Z(\cdot\,,\cdot))\in\mathcal{H}^2[0,T]$ *on* $[0,T]$. *Moreover, the following estimate holds*:

$$
\begin{aligned}
&\|(Y(\cdot),Z(\cdot\,,\cdot))\|^2_{\mathcal{H}^2[S,T]}\equiv E\left\{\int_S^T|Y(t)|^2dt+\int_S^T\int_S^T|Z(t,s)|^2dsdt\right\}\\
&\quad\le CE\left\{\int_S^T|\psi(t)|^2dt+\int_S^T\left(\int_t^T|g_0(t,s)|ds\right)^2dt\right\},\quad\forall S\in[0,T]. \qquad (3.44)
\end{aligned}
$$

Let $\bar g:[0,T]\times[0,T]\times\mathbb{R}^m\times\mathbb{R}^{m\times d}\times\mathbb{R}^{m\times d}\times\Omega\to\mathbb{R}^m$ *also satisfy* (H1). *Let* $\bar\psi(\cdot)\in L^2_{\mathcal{F}_T}(0,T)$ *and* $(\bar Y(\cdot),\bar Z(\cdot\,,\cdot))\in\mathcal{H}^2[0,T]$ *be the adapted M-solution of* (1.1) *with* g *and* $\psi(\cdot)$ *replaced by* $\bar g$ *and* $\bar\psi(\cdot)$, *respectively, then*

$$
\begin{aligned}
&E\left\{\int_S^T|Y(t)-\bar Y(t)|^2dt+\int_S^T\int_S^T|Z(t,s)-\bar Z(t,s)|^2dsdt\right\}\\
&\le CE\left\{\int_S^T|\psi(t)-\bar\psi(t)|^2dt\right.\\
&\quad\left.+\int_S^T\left(\int_t^T|g(t,s,Y(s),Z(t,s),Z(s,t))-\bar g(t,s,Y(s),Z(t,s),Z(s,t))|ds\right)^2dt\right\},\\
&\quad\forall S\in[0,T]. \qquad (3.45)
\end{aligned}
$$

Proof We split the proof into several steps.

Step 1 Existence and uniqueness of the adapted M-solution to BSVIE (1.1) on $[S, T]$ for some $S \in [0, T)$.

For any $S \in [0, T)$, let $\mathcal{M}^2[S, T]$ be the space of all $(y(\cdot), z(\cdot\,,\cdot)) \in \mathcal{H}^2[S, T]$ such that

$$y(t) = E[y(t) \,\big|\, \mathcal{F}_S] + \int_S^t z(t, s)dW(s), \quad \text{a.e. } t \in [S, T], \text{ a.s.} \tag{3.46}$$

Clearly, $\mathcal{M}^2[S, T]$ is a nontrivial closed subspace of $\mathcal{H}^2[S, T]$. Furthermore, for any $(y(\cdot), z(\cdot\,,\cdot)) \in \mathcal{M}^2[S, T]$, from (3.46), we have

$$E\int_S^t |z(t, s)|^2 ds = E|y(t)|^2 - E|E[y(t)\big|\mathcal{F}_S]|^2 \le E|y(t)|^2, \quad t \in [S, T], \text{ a.s.} \tag{3.47}$$

Thus, for any $(y(\cdot), z(\cdot\,,\cdot)) \in \mathcal{M}^2[S, T]$,

$$\begin{aligned} &E\left\{\int_S^T |y(t)|^2 dt + \int_S^T\int_t^T |z(t, s)|^2 ds dt\right\} \\ &\le E\left\{\int_S^T |y(t)|^2 dt + E\int_S^T\int_S^T |z(t, s)|^2 ds dt\right\} \\ &\le E\left\{\int_S^T |y(t)|^2 dt + \int_S^T\int_S^t |z(t, s)|^2 ds dt + \int_S^T\int_t^T |z(t, s)|^2 ds dt\right\} \\ &\le 2E\left\{\int_S^T |y(t)|^2 dt + \int_S^T\int_t^T |z(t, s)|^2 ds dt\right\}. \end{aligned} \tag{3.48}$$

Hence, we may take

$$\|(y(\cdot), z(\cdot\,,\cdot))\|_{\mathcal{M}^2[S,T]} \triangleq \left\{E\left[\int_S^T |y(t)|^2 dt + \int_S^T\int_t^T |z(t, s)|^2 ds dt\right]\right\}^{\frac{1}{2}} \tag{3.49}$$

as an equivalent norm for $\mathcal{M}^2[S, T]$.

Next, for any $\psi(\cdot) \in L^2_{\mathcal{F}_T}(S,T)$, and $(y(\cdot), z(\cdot\,,\cdot)) \in \mathcal{M}^2[S,T]$, we consider the following BSVIE:

$$Y(t) = \psi(t) + \int_t^T g(t,s,y(s),Z(t,s),z(s,t))ds - \int_t^T Z(t,s)dW(s), \quad t \in [S,T]. \tag{3.50}$$

By Corollary 3.6 (with $p = 2$), the above BSVIE admits a unique adapted M-solution $(Y(\cdot), Z(\cdot\,,\cdot)) \in \mathcal{H}^2[S,T]$ and

$$\begin{aligned}
&E\left\{|Y(t)|^2 + \int_t^T |Z(t,s)|^2 ds\right\}\\
&\le CE\left\{|\psi(t)|^2 + \left(\int_t^T |g(t,s,y(s),0,z(s,t))|ds\right)^2\right\}\\
&\le CE\left\{|\psi(t)|^2 + \left(\int_t^T |g_0(t,s)|ds\right)^2 + \left(\int_t^T L(t,s)^2 ds\right)\left(\int_t^T |y(s)|^2 ds\right)\right.\\
&\qquad\left. + \left(\int_t^T L(t,s)^2 ds\right)\left(\int_t^T |z(s,t)|^2 ds\right)\right\}\\
&\le CE\left\{|\psi(t)|^2 + \left(\int_t^T |g_0(t,s)|ds\right)^2 + \int_t^T |y(s)|^2 ds + \int_t^T |z(s,t)|^2 ds\right\}.
\end{aligned} \tag{3.51}$$

Consequently, [note (3.48)]

$$\begin{aligned}
\|(Y(\cdot), Z(\cdot\,,\cdot))\|^2_{\mathcal{M}^2[S,T]} &\equiv E\left\{\int_S^T |Y(t)|^2 dt + \int_S^T \int_t^T |Z(t,s)|^2 ds dt\right\}\\
&\le CE\left\{\int_S^T |\psi(t)|^2 dt + \int_S^T \left(\int_t^T |g_0(t,s)|ds\right)^2 dt + \|(y(\cdot), z(\cdot\,,\cdot))\|^2_{\mathcal{M}^2[S,T]}\right\}.
\end{aligned} \tag{3.52}$$

Hence, such a $(Y(\cdot), Z(\cdot\,,\cdot)) \in \mathcal{M}^2[S,T]$. Now, let us define a map $\Theta : \mathcal{M}^2[S,T] \to \mathcal{M}^2[S,T]$ by

$$\Theta(y(\cdot), z(\cdot\,,\cdot)) = (Y(\cdot), Z(\cdot\,,\cdot)), \qquad \forall (y(\cdot), z(\cdot\,,\cdot)) \in \mathcal{M}^2[S,T]. \tag{3.53}$$

We claim that this map is a contraction when $T - S > 0$ is small. To prove this, let $(\bar{y}(\cdot), \bar{z}(\cdot\,,\cdot)) \in \mathcal{M}^2[S,T]$, and $(\bar{Y}(\cdot), \bar{Z}(\cdot\,,\cdot)) = \Theta(\bar{y}(\cdot), \bar{z}(\cdot\,,\cdot))$. Then, by Corollary 3.6 again, for $t \in [S,T]$,

$$
\begin{aligned}
&E\left\{|Y(t)-\bar{Y}(t)|^2 + \int_t^T |Z(t,s)-\bar{Z}(t,s)|^2 ds\right\} \\
&\quad \le CE\left(\int_t^T |g(t,s,y(s),0,z(s,t)) - g(t,s,\bar{y}(s),0,\bar{z}(s,t))|ds\right)^2 \\
&\quad \le CE\left\{\int_t^T L(t,s)\Big(|y(s)-\bar{y}(s)| + |z(s,t)-\bar{z}(s,t)|\Big)ds\right\}^2 \\
&\quad \le C(T-t)^{\frac{\varepsilon}{2+\varepsilon}}\left(\int_t^T L(t,s)^{2+\varepsilon}ds\right)^{\frac{2}{2+\varepsilon}} \\
&\qquad \cdot E\left\{\int_t^T |y(t)-\bar{y}(t)|^2 dt + \int_t^T |z(t,s)-\bar{z}(t,s)|^2 ds dt\right\} \\
&\quad \le C(T-S)^{\frac{\varepsilon}{2+\varepsilon}} E\left\{\int_t^T |y(t)-\bar{y}(t)|^2 dt + \int_t^T |z(t,s)-\bar{z}(t,s)|^2 ds dt\right\}. \qquad (3.54)
\end{aligned}
$$

Consequently, similar to (3.52),

$$
\begin{aligned}
&\|\Theta(y(\cdot),z(\cdot\,,\cdot)) - \Theta(\bar{y}(\cdot),\bar{z}(\cdot\,,\cdot))\|_{\mathcal{M}^2[S,T]} \\
&\quad \equiv \|(Y(\cdot),Z(\cdot\,,\cdot)) - (\bar{Y}(\cdot),\bar{Z}(\cdot\,,\cdot))\|^2_{\mathcal{M}^2[S,T]} \\
&\quad \equiv E\int_S^T |Y(t)-\bar{Y}(t)|^2 dt + E\int_S^T\int_t^T |Z(t,s)-\bar{Z}(t,s)|^2 ds dt \\
&\quad \le C(T-S)^{\frac{\varepsilon}{2+\varepsilon}} E\left[\int_S^T |y(t)-\bar{y}(t)|^2 dt + \int_S^T\int_t^T |z(t,s)-\bar{z}(t,s)|^2 ds dt\right] \\
&\quad \le C(T-S)^{\frac{\varepsilon}{2+\varepsilon}} \|(y(\cdot),z(\cdot\,,\cdot)) - (\bar{y}(\cdot),\bar{z}(\cdot\,,\cdot))\|^2_{\mathcal{M}^2[S,T]}. \qquad (3.55)
\end{aligned}
$$

Thus, when $T-S>0$ is small, the map $\Theta : \mathcal{M}^2[S,T] \to \mathcal{M}^2[S,T]$ is a contraction. Hence, it admits a unique fixed point $(Y(\cdot), Z(\cdot\,,\cdot)) \in \mathcal{M}^2[S,T]$ which is the unique adapted M-solution of (1.1) over $[S,T]$. Also, estimate (3.44) holds for this S. We emphasize that the constant on the right hand side of the last inequality in (3.55) is an absolute one.

This step determines the values $(Y(t), Z(t,s))$ for $(t,s)\in[S,T]\times[S,T]$.

Step 2 Application of the Martingale Representation Theorem to determine the values $Z(t,s)$ of $Z(\cdot\,,\cdot)$ for $(t,s)\in[S,T]\times[R,S]$, with any given $R\in[0,S)$.

Since $E[Y(t)\,\big|\,\mathcal{F}_S]\in L^2(S,T;L^2_{\mathcal{F}_S}(\Omega))$, there is a unique $Z(\cdot\,,\cdot)\in L^2(S,T;L^2_{\mathcal{F}}(R,S))$ such that

$$E[Y(t)\mid\mathcal{F}_S]=E[Y(t)\mid\mathcal{F}_R]+\int_R^S Z(t,s)dW(s),\quad t\in[S,T],\tag{3.56}$$

which implies

$$E\int_R^S|Z(t,s)|^2ds=E|Y(t)|^2-|EY(t)|^2,\quad t\in[S,T].\tag{3.57}$$

Consequently,

$$\begin{aligned}E\int_S^T\int_R^S|Z(t,s)|^2dsdt&\le E\int_S^T|Y(t)|^2dt\\&\le CE\left\{\int_S^T|\psi(t)|^2dt+\int_S^T\left(\int_t^T|g_0(t,s)|ds\right)^2dt\right\}.\end{aligned}\tag{3.58}$$

Combining Steps 1–2, we have (uniquely) determined $(Y(t), Z(t,s))$ for $(t,s)\in[S,T]\times[R,T]$, and the following estimate holds [combining (3.44) and (3.58)]:

$$\begin{aligned}&E\left\{\int_S^T|Y(t)|^2dt+\int_S^T\int_R^T|Z(t,s)|^2dsdt\right\}\\&\le CE\left\{\int_S^T|\psi(t)|^2dt+\int_S^T\left(\int_t^T|g_0(t,s)|ds\right)^2dt\right\}.\end{aligned}\tag{3.59}$$

Step 3 A related SFIE is solvable on $[S,T]$.

For $(t,s)\in[R,S]\times[S,T]$, from Steps 1–2, we know that the values $Y(s)$ and $Z(s,t)$ are all already determined. Hence, the following can be defined:

$$g^S(t,s,z)=g(t,s,Y(s),z,Z(s,t)),\qquad (t,s,z)\in[R,S]\times[S,T]\times\mathbb{R}^{m\times d}.\tag{3.60}$$

Now, consider the following SFIE:

$$\psi^S(t) = \psi(t) + \int_S^T g^S(t,s,Z(t,s))ds - \int_S^T Z(t,s)dW(s), \quad t \in [R,S], \tag{3.61}$$

By Corollary 3.5 (with $p = 2$), the above SFIE admits a unique adapted solution $(\psi^S(\cdot), Z(\cdot\,,\cdot)) \in L^2_{\mathcal{F}_S}(R,S) \times L^2(R,S;L^2_{\mathbb{F}}(S,T))$, and the following holds:

$$\begin{aligned}
&E\left\{|\psi^S(t)|^2 + \int_S^T |Z(t,s)|^2 ds\right\} \\
&\leq CE\left[|\psi(t)|^2 + \left(\int_S^T |g(t,s,Y(s),0,Z(s,t))|ds\right)^2\right] \\
&\leq CE\left\{|\psi(t)|^2 + \left(\int_S^T |g_0(t,s)|ds\right)^2\right. \\
&\qquad \left. + \int_S^T |Y(s)|^2 ds + \int_S^T |Z(s,t)|^2 ds\right\}, \quad t \in [R,S].
\end{aligned} \tag{3.62}$$

Consequently, making use of (3.59), we have

$$\begin{aligned}
&E\left\{\int_R^S |\psi^S(t)|^2 dt + \int_R^S \int_S^T |Z(t,s)|^2 ds dt\right\} \\
&\leq CE\int_R^S \left\{|\psi(t)|^2 + \left(\int_S^T |g_0(t,s)|ds\right)^2 + \int_S^T |Y(s)|^2 ds + \int_S^T |Z(s,t)|^2 ds\right\} dt \\
&\leq CE\left\{\int_R^T |\psi(t)|^2 dt + \int_R^T \left(\int_t^T |g_0(t,s)|ds\right)^2 dt\right\}.
\end{aligned} \tag{3.63}$$

This step uniquely determines the values $Z(t,s)$ for $(t,s) \in [R,S] \times [S,T]$, and by the definition of $g^S(t,s,z)$, we see that $(\psi^S(\cdot), Z(\cdot\,,\cdot))$ satisfies

$$\begin{aligned}
\psi^S(t) &= \psi(t) + \int_S^T g(t,s,Y(s),Z(t,s),Z(s,t))ds \\
&\qquad - \int_S^T Z(t,s)dW(s), \quad t \in [R,S].
\end{aligned} \tag{3.64}$$

Step 4 Complete the proof by induction. Combining Steps 1–3, we have uniquely determined

$$\begin{cases} Y(t), & t \in [S,T], \\ Z(t,s), & (t,s) \in ([S,T]\times[R,T]) \bigcup ([R,S]\times[S,T]), \end{cases} \tag{3.65}$$

and the following estimate holds [see (3.59) and (3.63)]:

$$\begin{aligned} &E\left\{\int_S^T |Y(t)|^2 dt + \int_S^T\int_R^T |Z(t,s)|^2 ds dt + \int_R^S\int_S^T |Z(t,s)|^2 ds dt\right\} \\ &\quad \le CE\left\{\int_R^T |\psi(t)|^2 dt + \int_R^T \left(\int_t^T |g_0(t,s)| ds\right)^2 dt\right\}. \end{aligned} \tag{3.66}$$

Now, we consider

$$\begin{aligned} Y(t) = \psi^S(t) &+ \int_t^S g(t,s,Y(s),Z(t,s),Z(s,t)) ds \\ &- \int_t^S Z(t,s) dW(s), \quad t \in [R,S]. \end{aligned} \tag{3.67}$$

Since $\psi^S(\cdot) \in L^2_{\mathcal{F}_S}(R,S)$, (3.67) is a BSVIE over $[R,S]$. Hence, similar to Step 1, we are able to show that that (3.67) is solvable on $[R,S]$ when $S-R>0$ is small. Moreover, the following estimate holds:

$$\begin{aligned} &E\left\{\int_R^S |Y(t)|^2 dt + \int_R^S\int_R^S |Z(t,s)|^2 ds dt\right\} \\ &\quad \le CE\left\{\int_R^S |\psi^S(t)|^2 dt + \int_R^S \left(\int_t^S |g_0(t,s)| ds\right)^2 dt\right\} \\ &\quad \le CE\left\{\int_R^T |\psi(t)|^2 dt + \int_R^T \left(\int_t^T |g_0(t,s)| ds\right)^2 dt\right\}. \end{aligned} \tag{3.68}$$

This solvability determines $(Y(t), Z(t,s))$ for $(t,s) \in [R,S]\times[R,S]$. Note that for $t \in [R,S]$,

$$
\begin{aligned}
Y(t) &= \psi^S(t) + \int_t^S g(t,s,Y(s),Z(t,s),Z(s,t))ds - \int_t^S Z(t,s)dW(s) \\
&= \psi(t) + \int_S^T g(t,s,Y(s),Z(t,s),Z(s,t))ds - \int_S^T Z(t,s)dW(s) \\
&\quad + \int_t^S g(t,s,Y(s),Z(t,s),Z(s,t))ds - \int_t^S Z(t,s)dW(s) \\
&= \psi(t) + \int_t^T g(t,s,Y(s),Z(t,s),Z(s,t))ds - \int_t^T Z(t,s)dW(s). \qquad (3.69)
\end{aligned}
$$

Hence, we obtain the (unique) solvability of BSVIE (1.1) on $[R, T]$, with the following estimate holds [combining (3.66) and (3.68)]:

$$
E\left\{\int_R^T |Y(t)|^2 dt + \int_R^T\int_R^T |Z(t,s)|^2 ds dt\right\} \le CE\left\{\int_R^T |\psi(t)|^2 dt + \int_R^T \left(\int_t^T |g(t,s)|ds\right)^2 dt\right\}. \qquad (3.70)
$$

Then we can use induction to finish the proof of the existence and uniqueness of adapted M-solution to BSVIE (1.1).

Finally, we prove the stability estimate. To this end, let $(Y(\cdot), Z(\cdot\,,\cdot))$ and $(\bar{Y}(\cdot), \bar{Z}(\cdot\,,\cdot))$ be adapted M-solutions of (1.1) corresponding to (g, ψ) and $(\bar{g}, \bar{\psi})$, respectively. Let

$$
\begin{aligned}
&\widehat{Y}(t) = Y(t) - \bar{Y}(t), \qquad \widehat{Z}(t,s) = Z(t,s) - \bar{Z}(t,s), \qquad \hat{\psi}(t) = \psi(t) - \bar{\psi}(t), \\
&\hat{g}(t,s) = |g(t,s,Y(s),Z(t,s),Z(s,t)) - \bar{g}(t,s,Y(s),Z(t,s),Z(s,t))|, \\
&\bar{g}_y(t,s) = \frac{[\bar{g}(t,s,Y(s),Z(t,s),Z(s,t)) - \bar{g}(t,s,\bar{Y}(s),Z(t,s),Z(s,t))]\widehat{Y}(s)^T}{|\widehat{Y}(s)|^2} I_{[\widehat{Y}(s)\neq 0]},
\end{aligned}
$$

Similarly, we define $\bar{g}_{z_i}(t,s)$ and $\bar{g}_{\zeta_i}(t,s)$. Then one has

$$
\begin{aligned}
\widehat{Y}(t) = \hat{\psi}(t) + \int_t^T \Big\{ &\bar{g}_y(t,s)\widehat{Y}(s) + \sum_{i=1}^d \Big[\bar{g}_{z_i}(t,s)\widehat{Z}_i(t,s) \\
&+ \bar{g}_{\zeta_i}(t,s)\widehat{Z}_i(s,t)\Big] + \hat{g}(t,s)\Big\} ds \\
&- \int_t^T \widehat{Z}(t,s)dW(s). \qquad (3.71)
\end{aligned}
$$

Similar to the above Steps 1–4, we have

$$
\begin{aligned}
&E\int_S^T |\widehat{Y}(t)|^2dt + E\int_S^T\int_S^T |\widehat{Z}(t,s)|^2 dsdt \\
&\quad \le CE\left\{\int_S^T |\hat{\psi}(t)|^2dt + \int_S^T\left(\int_t^T \hat{g}(t,s)ds\right)^2 dt\right\}.
\end{aligned}
$$

Then our stability estimate follows. In particular, we have the uniqueness. □

The following result will be useful in the sequel.

Corollary 3.8 *Let* (H1) *hold and* $(Y(\cdot), Z(\cdot\,,\cdot)) \in \mathcal{M}^2[0,T]$ *be the unique adapted M-solution of* (1.1). *For any* $t \in [0,T]$, *let* $(\lambda^t(\cdot), \mu^t(\cdot)) \in \mathbb{H}^2[t,T]$ *be the adapted solution of the following BSDE*:

$$
\lambda^t(r) = \psi(t) + \int_r^T g(t,s,Y(s),\mu^t(s),Z(s,t))ds - \int_r^T \mu^t(s)dW(s), \quad r \in [t,T]. \tag{3.72}
$$

Let

$$
\begin{cases} \bar{Y}(t) = \lambda^t(t), & t \in [0,T], \\ \bar{Z}(t,s) = \mu^t(s), & (t,s) \in \Delta^c, \end{cases} \tag{3.73}
$$

and let the values $\bar{Z}(t,s)$ *of* $\bar{Z}(\cdot\,,\cdot)$ *for* $(t,s) \in \Delta$ *be defined through*

$$
\bar{Y}(t) = E\bar{Y}(t) + \int_0^t \bar{Z}(t,s)dW(s), \quad t \in [0,T]. \tag{3.74}
$$

Then

$$
\begin{cases} \bar{Y}(t) = Y(t), & t \in [0,T], \\ \bar{Z}(t,s) = Z(t,s), & (t,s) \in \Delta^c. \end{cases} \tag{3.75}
$$

Proof Take $r = t$ in (3.72), we obtain

$$
\bar{Y}(t) = \psi(t) + \int_t^T g(t,s,Y(s),\bar{Z}(t,s),Z(s,t))ds - \int_t^T \bar{Z}(t,s)dW(s), \quad t \in [0,T]. \tag{3.76}
$$

Thus, $(\bar{Y}(\cdot), \bar{Z}(\cdot\,,\cdot)) \in \mathcal{M}^2[0,T]$ is an adapted M-solution of the above BSVIE. Since $(Y(\cdot), Z(\cdot\,,\cdot)) \in \mathcal{M}^2[0,T]$ is another adapted solution of the above, we obtain (3.75) by uniqueness of adapted M-solutions. □

To conclude this section, we present an interesting result concerning the adapted M-solution of our BSVIE (1.1).

Proposition 3.9 *Let* (H1) *hold. Let* $\psi(\cdot) \in L^2_{\mathcal{F}_T}(0,T)$ *and* $(Y(\cdot), Z(\cdot\,,\cdot)) \in \mathcal{H}^2[0,T]$ *be the unique adapted M-solution of BSVIE* (1.1) *on* $[0,T]$. *Then for all* $S \in [0,T)$,

$$\psi^S(t) \stackrel{\Delta}{=} \psi(t) + \int_S^T g(t,s,Y(s),Z(t,s),Z(s,t))ds - \int_S^T Z(t,s)dW(s) \tag{3.77}$$

is $\mathcal{F}_S$*-measurable for almost all* $t \in [0,S]$.

Proof Note that for almost all $t \in [0,T]$, the involved functions $\psi(t)$, $g(t,\cdot\,,Y(\cdot),$ $Z(t,\cdot), Z(\cdot\,,t))$, etc. in (3.77) are well-defined. Take such a t. Suppose for some $S \in [0,T]$, $\psi^S(t)$ is not $\mathcal{F}_S$-measurable. According to the proof of Theorem 3.7 given above, there exists an $R \in [0,S)$ such that

$$\psi^R(t) \stackrel{\Delta}{=} \psi(t) + \int_R^T g(t,s,Y(s),Z(t,s),Z(s,t))ds - \int_R^T Z(t,s)dW(s) \tag{3.78}$$

is $\mathcal{F}_R$-measurable. On the other hand, we have

$$\psi^R(t) = \psi^S(t) + \int_R^S g(t,s,Y(s),Z(t,s),Z(s,t))ds - \int_R^S Z(t,s)dW(s), \tag{3.79}$$

and the last two terms on the right hand side of the above are $\mathcal{F}_S$-measurable. Hence, if $\psi^S(t)$ is not $\mathcal{F}_S$-measurable, neither the whole right hand side. This contradicts the $\mathcal{F}_S$-measurability of the left hand side since $R < S$. □

4 Regularity of adapted M-solutions

In this section, we are going to discuss some regularity for the adapted M-solutions to BSVIE (1.1) by means of Malliavin calculus. Among other things, we will establish an estimate stronger than (3.44), under, of course, some stronger conditions. Moreover, we will establish the continuity of $t \mapsto Y(t)$, allowing the dependence of g on ζ.

Let us first make an observation. By Corollary 3.6 (with $p = 2$), together with Gronwall's inequality, we see that if (H1) holds, and $(Y(\cdot), Z(\cdot\,,\cdot)) \in \mathcal{M}^2[0,T]$ is the unique adapted M-solution of BSVIE (1.1), then

$$E\left\{|Y(t)|^2 + \int_0^T |Z(t,s)|^2 ds\right\}$$
$$\leq CE\left\{|\psi(t)|^2 + \int_t^T |\psi(s)|^2 ds + \left(\int_t^T |g(t,s,0,0,Z(s,t))|ds\right)^2\right.$$

$$+\int_t^T\left(\int_s^T|g(s,\tau,0,0,Z(\tau,s))|d\tau\right)^2dt\Bigg\}$$

$$\leq CE\Bigg\{|\psi(t)|^2+\int_t^T|\psi(s)|^2ds+\left(\int_t^T|g_0(t,s)|ds\right)^2+\int_t^T\left(\int_s^T|g_0(s,\tau)|d\tau\right)^2dt$$

$$+\left(\int_t^T L(t,s)|Z(s,t)|ds\right)^2+\int_t^T\left(\int_s^T L(s,\tau)|Z(\tau,s)|d\tau\right)^2ds\Bigg\}$$

$$\leq CE\Bigg\{|\psi(t)|^2+\int_t^T|\psi(s)|^2ds+\left(\int_t^T|g_0(t,s)|ds\right)^2+\int_t^T\left(\int_s^T|g_0(s,\tau)|d\tau\right)^2dt$$

$$+\int_t^T|Z(s,t)|^2ds\Bigg\}. \tag{4.1}$$

Thus, in order to estimate the left hand side of the above, it is crucial to estimate the last term on the right hand side of the above. The difficulty here is that we hope to get some kind of pointwise behavior (instead of just integrability behavior) of $t\mapsto Z(s,t)$. To achieve this, we need to use Malliavin calculus. Roughly speaking, if $Y(\cdot)$ and $Z(\cdot,\cdot)$ are Malliavin differentiable, then from

$$Y(t)=EY(t)+\int_0^t Z(t,s)dW(s),\quad t\in[0,T], \tag{4.2}$$

we have

$$D_r^iY(t)=Z_i(t,r)+\int_r^t D_r^iZ(t,s)dW(s),\quad (t,r)\in\Delta,\ 1\leq i\leq d. \tag{4.3}$$

Thus, the last term on the right hand side of (4.1) can be estimated through the following:

$$E\int_t^T|Z_i(s,t)|^2ds=7E\int_t^T|D_t^iY(s)|^2ds-E\int_t^T\int_t^s|D_t^iZ(s,\tau)|^2d\tau ds$$

$$\leq E\int_t^T|D_t^iY(s)|^2ds. \tag{4.4}$$

In this section, among other things, we will mainly make some efforts to estimate the right hand side of (4.4), which will lead to an estimate for the left hand side of (4.1) and some other related estimates.

To begin with, we introduce some more spaces. For any $0 \le R < S \le T$, let $\Psi[R, S]$ be the space consists of all processes $\psi(\cdot) \in L^\infty(R, S; L^2_{\mathcal{F}_T}(\Omega))$ such that

$$\|\psi\|_{\Psi[R,S]} \triangleq \sup_{t\in[R,S]} E\left[|\psi(t)|^2 + \int_t^S \sum_{i=1}^d |D^i_t\psi(s)|^2 ds\right]^{\frac{1}{2}} < \infty. \tag{4.5}$$

We let $\mathcal{Y}[R, S]$ be the space of all processes $y(\cdot) \in \Psi[R, S]$ which are $\mathbb{F}$-adapted. Likewise, let $\mathcal{Z}[R, S]$ be the space consists of all processes $z : [R, S]^2 \times \Omega \to \mathbb{R}^{m\times d}$ such that $s \mapsto z(t, s)$ is $\mathbb{F}$-adapted for almost all $t \in [R, S]$, and the following holds:

$$\|z\|_{\mathcal{Z}[R,S]} \triangleq \left\{\sup_{t\in[R,S]} E\left[\int_R^S |z(t, s)|^2 ds + \int_t^S |z(s, t)|^2 ds + \int_t^S\int_t^S \sum_{i=1}^d |D^i_t z(\tau, s)|^2 ds d\tau\right]\right\}^{\frac{1}{2}} < \infty. \tag{4.6}$$

Clearly, $\|\cdot\|_{\Psi[R,S]}$ and $\|\cdot\|_{\mathcal{Z}[R,S]}$ are norms under which $\Psi[R, S]$ and $\mathcal{Z}[R, S]$ are Banach spaces, respectively; and $\mathcal{Y}[R, S]$ is a closed subspace of $\Psi[R, S]$. For notational better-looking, when $y(\cdot) \in \mathcal{Y}[R, S]$, we use $\|y(\cdot)\|_{\mathcal{Y}[R,S]}$ instead of $\|y(\cdot)\|_{\Psi[R,S]}$ for its norm.

Next, we let $\Psi_c[R, S]$ be the space consists of all $\psi(\cdot) \in \Psi[R, S]$ such that

$$\lim_{\substack{t,\bar t\in[R,S]\\ |t-\bar t|\to 0}} E\left\{|\psi(t) - \psi(\bar t)|^2 + \int_{t\vee\bar t}^S \sum_{i=1}^d |D^i_t\psi(s) - D^i_{\bar t}\psi(s)|^2 ds\right\} = 0, \tag{4.7}$$

and let $\mathcal{Y}_c[R, S]$ be the space of all process $y(\cdot) \in \Psi_c[R, S]$ that are $\mathbb{F}$-adapted. Likewise, let $\mathcal{Z}_c[R, S]$ be the space of all $z(\cdot\, , \cdot) \in \mathcal{Z}[R, S]$ such that

$$\lim_{\substack{t,\bar t\in[R,S]\\ |t-\bar t|\to 0}} E\left\{\int_R^{t\wedge\bar t} |z(t, s) - z(\bar t, s)|^2 ds + \int_{t\vee\bar t}^S |z(t, s) - z(\bar t, s)|^2 ds + \int_{t\vee\bar t}^S |z(s, t) - z(s, \bar t)|^2 ds + \int_{t\vee\bar t}^S\int_{t\vee\bar t}^S \sum_{i=1}^d |D^i_t z(s, \tau) - D^i_{\bar t} z(s, \tau)|^2 d\tau ds\right\} = 0. \tag{4.8}$$

It is easy to see that $\Psi_c[R, S]$, $\mathcal{Y}_c[R, S]$ and $\mathcal{Z}_c[R, S]$ are closed subspaces of $\Psi[R, S]$, $\mathcal{Y}[R, S]$ and $\mathcal{Z}[R, S]$, respectively; and $\mathcal{Y}_c[R, S]$ is a closed subspace of $\Psi_c[R, S]$.

Also, we note that

$$\mathcal{Y}_c[R,S] \subseteq C_{\mathbb{F}}([R,S];L^2(\Omega)). \tag{4.9}$$

Thus, any process in $\mathcal{Y}_c[R,S]$ is continuous from $[R,S]$ to $L^2(\Omega)$.

Now, we introduce the following assumption [comparing with (H1)].

(H2) Let $g:\Delta^c\times\mathbb{R}^m\times\mathbb{R}^{m\times d}\times\mathbb{R}^{m\times d}\times\Omega\to\mathbb{R}^m$ be $\mathcal{B}(\Delta^c\times\mathbb{R}^m\times\mathbb{R}^{m\times d}\times\mathbb{R}^{m\times d})\otimes\mathcal{F}_T$-measurable, with $s\mapsto g(t,s,y,z,\zeta)$ being $\mathbb{F}$-adapted for all $(t,y,z,\zeta)\in[0,T]\times\mathbb{R}^m\times\mathbb{R}^{m\times d}\times\mathbb{R}^{m\times d}$ satisfying (3.41) such that $(y,z,\zeta)\mapsto g(t,s,y,z,\zeta)$ is continuously differentiable, and $(y,z,\zeta)\mapsto[D_r^i g](t,s,y,z,\zeta)$ is also continuous. Moreover, there are deterministic functions $L,L_y,L_z:\Delta^c\to[0,\infty)$, and a process $L_0:\Delta^c\times\Omega\to[0,\infty)$ satisfying (3.43) for some $\varepsilon>0$, and

$$E\int_0^T\left(\int_t^T L_0(t,s)ds\right)^2dt+\int_0^T\int_t^T L_y(t,s)^2dsdt+\sup_{t\in[0,T]}\int_t^T L_z(t,s)^2ds<\infty, \tag{4.10}$$

such that

$$\begin{cases}\sum_{i=1}^d\left|[D_r^i g](t,s,y,z,\zeta)\right|\le L_0(t,s)+L_y(t,s)|y|+L_z(t,s)(|z|+|\zeta|),\\ |g_y(t,s,y,z,\zeta)|\le L(t,s),\\ \left|\sum_{j=1}^d g_{z_j}(t,s,y,z,\zeta)\eta_j\right|\le L(t,s)|\eta|,\qquad\left|\sum_{j=1}^d g_{\zeta_j}(t,s,y,z,\zeta)\eta_j\right|\le L(t,s)|\eta|,\\ \forall(t,s)\in\Delta^c,\ (y,z,\zeta)\in\mathbb{R}^m\times\mathbb{R}^{m\times d}\times\mathbb{R}^{m\times d},\eta\equiv(\eta_1,\cdots,\eta_d)\in\mathbb{R}^{m\times d},\ r\in[0,T],\ \text{a.s.}\end{cases} \tag{4.11}$$

Note that the last three conditions in (4.11) imply the Lipschitz condition for the map $(y,z,\zeta)\mapsto g(t,s,y,z,\zeta)$ with the Lipschitz constants depending on (t,s). Hence, (H2) is stronger than (H1). Also, it should be pointed out that we only assume the continuity and linear growth of the map $(y,z,\zeta)\mapsto[D_r^i g](t,s,y,z,\zeta)$, instead of the Lipschitz continuity. We introduce the following a little stronger assumption.

(H2)$'$ Let (H2) hold and moreover,

$$\sup_{t\in[0,T]}E\left(\int_t^T g_0(t,s)ds\right)^2<\infty. \tag{4.12}$$

The first main result of this section is the following result.

Theorem 4.1 *Let* (H2) *hold,* $\psi(\cdot)\in\Psi[0,T]$*, and* $(Y(\cdot),Z(\cdot\,,\cdot))\in\mathcal{M}^2[0,T]$ *be the adapted M-solution of* (1.1)*. Then for any* $r\in[0,T)$*, and* $S\in[0,T)$*,*

$$E\sum_{i=1}^{d}\left\{\int_S^T |D_r^iY(t)|^2dt+\int_S^T\int_S^T |D_r^iZ(t,s)|^2dsdt\right\}$$
$$\leq CE\left\{\int_S^T |\psi(t)|^2dt+\sum_{i=1}^{d}\int_S^T| D_r^i\psi(t)|^2dt+\int_S^T\left(\int_t^T g_0(t,s)ds\right)^2 dt\right.$$
$$\left.+\int_S^T\left(\int_t^T L_0(t,s)ds\right)^2 dt\right\}, \tag{4.13}$$

and $(D_r^iY(\cdot),\, D_r^iZ(\cdot\,,\cdot))$ *is the adapted M-solution of the following BSVIE*:

$$\begin{aligned}D_r^iY(t)=D_r^i\psi(t)+\int_t^T\Big\{&[D_r^ig](t,s,Y(s),Z(t,s),Z(s,t))\\&+g_y(t,s,Y(s),Z(t,s),Z(s,t))D_r^iY(s)\\&+\sum_{i=1}^{d}\Big[g_{z_j}(t,s,Y(s),Z(t,s),Z(s,t))D_r^iZ_j(t,s)\\&+g_{\zeta_j}(t,s,Y(s),Z(t,s),Z(s,t))D_r^iZ_j(s,t)\Big]\Big\}ds\\&-\int_t^T D_r^iZ(t,s)dW(s),\quad t\in[r,T].\end{aligned} \tag{4.14}$$

Moreover,

$$D_r^iY(t)=Z_i(t,r)+\int_r^t D_r^iZ(t,s)dW(s),\quad (t,r)\in\Delta, \tag{4.15}$$

which implies

$$Z_i(t,r)=E\left[D_r^iY(t)\;\middle|\;\mathcal{F}_r\right],\quad (t,r)\in\Delta,\ \text{a.s.},\quad 1\leq i\leq d. \tag{4.16}$$

In addition,

$$\begin{aligned}Z_i(t,r)=D_r^i\psi(t)+\int_r^T\Big\{&[D_r^ig](t,s,Y(s),Z(t,s),Z(s,t))\\&+g_y(t,s,Y(s),Z(t,s),Z(s,t))D_r^iY(s)\\&+\sum_{j=1}^{d}[g_{z_j}(t,s,Y(s),Z(t,s),Z(s,t))D_r^iZ_j(t,s)\Big\}ds\end{aligned}$$

$$-\int_r^T D_r^i Z(t,s)dW(s),$$
$$(t,r)\in\Delta^c,\quad 1\le i\le d. \tag{4.17}$$

Further, if (H2)$'$ *holds, then*

$$\|Y(\cdot)\|^2_{\mathcal{Y}[0,T]}+\|Z(\cdot\,,\cdot)\|^2_{\mathcal{Z}[0,T]}$$
$$\le C\left\{\|\psi(\cdot)\|^2_{\Psi[0,T]}+\sup_{t\in[0,T]}E\left(\int_t^T g_0(t,s)ds\right)^2+E\int_0^T\left(\int_t^T L_0(t,s)ds\right)^2dt\right\}. \tag{4.18}$$

Proof Since (H2) is stronger than (H1), by Theorem 3.7, for any $\psi(\cdot)\in\Psi[0,T]$, BSVIE (1.1) admits a unique adapted solution $(Y(\cdot),Z(\cdot))\in\mathcal{M}^2[0,T]$ and estimate (3.44) holds. Next, we introduce the following BSVIE [which is a formal Malliavin differentiation of (1.1)]:

$$\begin{aligned}\widehat{Y}^{r,i}(t)=D_r^i\psi(t)+\int_t^T&\Big\{[D_r^i g](t,s,Y(s),Z(t,s),Z(s,t))\\ &+g_y(t,s,Y(s),Z(t,s),Z(s,t))\widehat{Y}^{r,i}(s)\\ &+\sum_{j=1}^d\Big[g_{z_j}(t,s,Y(s),Z(t,s),Z(s,t))\widehat{Z}_j^{r,i}(t,s)\\ &+g_{\zeta_j}(t,s,Y(s),Z(t,s),Z(s,t))\widehat{Z}^{r,i}(s,t)\Big]\Big\}ds\\ &-\int_t^T\widehat{Z}^{r,i}(t,s)dW(s),\quad t\in[0,T].\end{aligned} \tag{4.19}$$

By Theorem 3.7, the above admits a unique adapted solution $(\widehat{Y}^{r,i}(\cdot),\widehat{Z}^{r,i}(\cdot\,,\cdot))\in\mathcal{H}^2[0,T]$, and for each $S\in[0,T]$, the following estimate holds: [note (3.44)]

$$\begin{aligned}&\|(\widehat{Y}^{r,i}(\cdot),\widehat{Z}^{r,i}(\cdot\,,\cdot))\|^2_{\mathcal{H}^2[S,T]}\\ &\equiv E\sum_{i=1}^d\left\{\int_S^T|\widehat{Y}^{r,i}(t)|^2ds+\int_S^T\int_S^T|\widehat{Z}^{r,i}(t,s)|^2dsdt\right\}\\ &\le CE\sum_{i=1}^d\left\{\int_S^T|D_r^i\psi(t)|^2dt+\int_S^T\left(\int_t^T|[D_r^ig](t,s,Y(s),Z(t,s),Z(s,t))|ds\right)^2dt\right\}\end{aligned}$$

$$\leq CE\left\{\sum_{i=1}^{d}\int_S^T |D_r^i\psi(t)|^2dt + \int_S^T\left(\int_t^T L_0(t,s)ds\right)^2 dt\right.$$
$$+\left(\int_S^T\int_t^T L_y(t,s)^2dsdt\right)\int_S^T |Y(s)|^2ds$$
$$\left.+\left(\sup_{t\in[S,T]}\int_t^T L_z(t,s)^2ds\right)\int_S^T\int_S^T |Z(t,s)|^2dsdt\right\}$$
$$\leq CE\left\{\int_S^T |\psi(t)|^2dt + \sum_{i=1}^{d}\int_S^T |D_r^i\psi(t)|^2dt + \int_S^T\left(\int_t^T g_0(t,s)ds\right)^2\right.$$
$$\left.+\int_S^T\left(\int_t^T L_0(t,s)ds\right)^2 dt\right\}. \tag{4.20}$$

Further, we see from the proof of Theorem 3.7 that when $T - S > 0$ small, the map Θ defined by (3.53) is a contraction on $\mathcal{H}^2[S, T]$. Therefore, a Picard iteration sequence converges to the unique adapted M-solution. Namely, if we define:

$$\begin{cases}(Y^0(\cdot), Z^0(\cdot)) = 0,\\ (Y^{k+1}(\cdot), Z^{k+1}(\cdot)) = \Theta(Y^k(\cdot), Z^k(\cdot)), \quad k \geq 0,\end{cases} \tag{4.21}$$

then

$$\lim_{k\to\infty} ||(Y^k(\cdot), Z^k(\cdot)) - (Y(\cdot), Z(\cdot))||_{\mathcal{H}^2[S,T]} = 0. \tag{4.22}$$

Next, from

$$Y^{k+1}(t) = \psi(t) + \int_t^T g(t, s, Y^k(s), Z^{k+1}(t, s), Z^k(s, t))ds - \int_t^T Z^{k+1}(t, s)dW(s), \tag{4.23}$$

similar to [12], we can recursively show that

$$(D_r^i Y^k(\cdot), D_r^i Z^k(\cdot\,,\cdot)) \in \mathcal{H}^2[S, T], \quad k \geq 0, \tag{4.24}$$

and

$$
\begin{aligned}
D_r^i Y^{k+1}(t) = D_r^i \psi(t) + \int_t^T \Big\{ & [D_r^i g](t,s,Y^k(s),Z^k(t,s),Z^k(s,t)) \\
& + g_y(t,s,Y^k(s),Z^k(t,s),Z^k(s,t)) D_r^i Y^k(s) \\
& + \sum_{j=1}^d \Big[g_{z_j}(t,s,Y^k(s),Z^k(t,s),Z^k(s,t)) D_r^i Z_j^k(t,s) \\
& + g_{\zeta_j}(t,s,Y^k(s),Z^k(t,s),Z^k(s,t)) D_r^i Z_j^k(s,t) \Big] \Big\} ds \\
& - \int_t^T D_r^i Z^{k+1}(t,s) dW(s), \quad t \in [S,T].
\end{aligned} \tag{4.25}
$$

Then (we suppress (t,s) in $[D_r^i g](t,s,\cdot,\cdot,\cdot)$, etc.)

$$
\begin{aligned}
\theta_{k+1} & \stackrel{\Delta}{=} \|(D_r^i Y^{k+1}(\cdot), D_r^i Z^{k+1}(\cdot\,,\cdot)) - (\widehat{Y}^{r,i}(\cdot), \widehat{Z}^{r,i}(\cdot))\|^2_{\mathcal{H}^2[S,T]} \\
& \le CE \Bigg\{ \int_S^T \Bigg(\int_t^T |[D_r^i g](Y^k(s), Z^k(t,s), Z^k(s,t)) \\
& \qquad - [D_r^i g](Y(s), Z(t,s), Z(s,t))| ds \Bigg)^2 dt \\
& \quad + \int_S^T \Bigg(\int_t^T |g_y(Y^k(s), Z^k(t,s), Z^k(s,t)) \\
& \qquad - g_y(Y(s), Z(t,s), Z(s,t))| \, |\widehat{Y}^{r,i}(s)| ds \Bigg)^2 dt \\
& \quad + \sum_{j=1}^d \int_S^T \Bigg(\int_t^T |g_{z_j}(Y^k(s), Z^k(t,s), Z^k(s,t)) \\
& \qquad - g_{z_j}(Y(s), Z(t,s), Z(s,t))| \, |\widehat{Z}_j^{r,i}(t,s)| ds \Bigg)^2
\end{aligned}
$$

$$+\sum_{j=1}^{d}\int_S^T\left(\int_t^T |g_{\zeta_j}(Y^k(s), Z^k(t,s), Z^k(s,t))\right.$$

$$\left.- g_{z_j}(Y(s), Z(t,s), Z(s,t))|\,|\widehat{Z}_j^{r,i}(s,t)|ds\right)^2 dt$$

$$+\left[\int_S^T\int_t^T L(t,s)^2 dsdt + \sup_{t\in[S,T]}\int_t^T L(t,s)^2 ds\right]$$

$$\left.\cdot\,\|(D_r^i Y^k(\cdot), D_r^i Z^k(\cdot\,,\cdot)) - (\widehat{Y}^{r,i}(\cdot), \widehat{Z}^{r,i}(\cdot\,,\cdot))\|^2_{\mathcal{H}^2[S,T]}\right\} \le \eta_k + \alpha\theta_k, \tag{4.26}$$

where

$$\eta_k \triangleq CE\left\{\int_S^T\left(\int_t^T |\,[D_r^i g](Y^k(s), Z^k(t,s), Z^k(s,t))\right.\right.$$

$$\left.- [D_r^i g](Y(s), Z(t,s), Z(s,t))|ds\right)^2 dt$$

$$+\int_S^T\left(\int_t^T |g_y(Y^k(s), Z^k(t,s), Z^k(s,t))\right.$$

$$\left.-g_y(Y(s), Z(t,s), Z(s,t))|\,|\widehat{Y}^{r,i}(s)|ds\right)^2 dt$$

$$+\sum_{j=1}^{d}\int_S^T\left(\int_t^T |g_{z_j}(Y^k(s), Z^k(t,s), Z^k(s,t))\right.$$

$$\left.- g_{z_j}(Y(s), Z(t,s), Z(s,t))|\,|\widehat{Z}_j^{r,i}(t,s)|ds\right)^2 dt$$

$$+\sum_{j=1}^{d}\int_S^T\left(\int_t^T |g_{\zeta_j}(Y^k(s), Z^k(t,s), Z^k(s,t))\right.$$

$$\left.- g_{z_j}(Y(s), Z(t,s), Z(s,t))|\,|\widehat{Z}_j^{r,i}(s,t)|ds\right)^2 dt, \tag{4.27}$$

and if necessary, we shrink $T-S$ further so that

$$CE\left[\int_S^T\int_t^T L(t,s)^2ds+\sup_{t\in[S,T]}\int_t^T L(t,s)^2ds\right]\le\alpha<1. \tag{4.28}$$

By the convergence (4.22) and the dominated convergence theorem, we see that

$$\lim_{k\to\infty}\eta_k=0. \tag{4.29}$$

Then (4.26) implies

$$\lim_{k\to\infty}\theta_k=0. \tag{4.30}$$

Since operator D_r^i is closed, we must have

$$\widehat{Y}^{r,i}(t)=D_r^iY(t),\quad \widehat{Z}^{r,i}(t,s)=D_r^iZ(t,s),\qquad t,s\in[S,T],\ \text{a.s.} \tag{4.31}$$

This proves (4.13)–(4.17) for $T-S>0$ small.

Next, by (4.3), we have

$$\begin{aligned}&E\int_S^T\int_r^S|D_r^iZ(t,s)|^2dsdt\le E\int_S^T\int_r^t|D_r^iZ(t,s)|^2dsdt\le E\int_S^T|D_r^iY(t)|^2dt\\&\le CE\left\{\int_S^T|\psi(t)|^2dt+\sum_{i=1}^d\int_S^T|D_r^i\psi(t)|^2dt+\int_S^T\left(\int_t^T g_0(t,s)ds\right)^2\right.\\&\qquad\left.+\int_S^T\left(\int_t^T L_0(t,s)ds\right)^2dt\right\}.\end{aligned} \tag{4.32}$$

Combining what we have proved, one has

$$\begin{aligned}&E\left\{\int_S^T|D_r^iY(t)|^2dt+\int_S^T\int_r^T|D_r^iZ(t,s)|^2dsdt\right\}\\&\le CE\left\{\int_S^T|\psi(t)|^2dt+\sum_{i=1}^d\int_S^T|D_r^i\psi(t)|^2dt+\int_S^T\left(\int_t^T g_0(t,s)ds\right)^2\right.\\&\qquad\left.+\int_S^T\left(\int_t^T L_0(t,s)ds\right)^2dt\right\}.\end{aligned} \tag{4.33}$$

Next, for any fixed $t \in [0, S]$, applying Proposition 2.2 to the following BSDE:

$$\lambda^t(\tau) = \psi(t) + \int_\tau^T g(t, s, Y(s), \mu^t(s), Z(s,t))ds - \int_\tau^T \mu^t(s)dW(s), \quad \tau \in [S, T], \tag{4.34}$$

we obtain (with $r \in [0, S]$)

$$\begin{aligned}
&E\left\{\sup_{\tau\in[S,T]}\sum_{i=1}^d |D_r^i\lambda^t(\tau)|^2 + \sum_{i=1}^d\left(\int_S^T |D_r^i\mu^t(s)|^2 ds\right)\right\} \\
&\le CE\left\{|\psi(t)|^2 + \sum_{i=1}^d |D_r^i\psi(t)|^2 + \left(\int_S^T |g(t,s,Y(s),0,Z(s,t))|ds\right)^2\right. \\
&\quad + \sum_{i=1}^d\left[\int_S^T \Big(|\,[D_r^i g](t,s,Y(s),0,Z(s,t))| \right. \\
&\qquad\qquad \left.\left. + L(t,s)[\,|D_r^iY(s)| + |D_r^iZ(s,t)|\,]\Big)ds\right]^2\right\} \\
&\le CE\left\{|\psi(t)|^2 + \sum_{i=1}^d |D_r^i\psi(t)|^2 + \left(\int_S^T g_0(t,s)ds\right)^2\right. \\
&\quad + \left(\int_S^T L(t,s)^2ds\right)\int_S^T \Big(|Y(s)|^2 + |Z(s,t)|^2\Big)ds + \left(\int_S^T L_0(t,s)ds\right)^2 \\
&\quad + \left(\int_S^T L_y(t,s)^2ds\right)\int_S^T |Y(s)|^2ds + \left(\int_S^T L_\zeta(t,s)^2ds\right)\int_S^T |Z(s,t)|^2ds \\
&\quad \left. + \left(\int_S^T L(t,s)^2ds\right)\int_S^T \Big(|D_r^iY(s)|^2 + |D_r^iZ(s,t)|^2\Big)ds\right\}.
\end{aligned} \tag{4.35}$$

Hence, noting (3.18) and (4.33), for $R \in [0, S]$,

$$\begin{aligned}
&E\left\{\int_R^S \sum_{i=1}^d |D_r^i\psi^S(t)|^2dt + \sum_{i=1}^d\left(\int_R^S\int_S^T |D_r^iZ(t,s)|^2dsdt\right)\right\} \\
&\le CE\left\{\int_R^S |\psi(t)|^2dt + \sum_{i=1}^d\int_R^S |D_r^i\psi(t)|^2dt\right.
\end{aligned}$$

$$
\begin{aligned}
&+\int_R^S\left(\int_S^T g_0(t,s)ds\right)^2 dt+\int_S^T |Y(s)|^2 ds\\
&+\int_R^S\int_S^T |Z(s,t)|^2 ds dt+\int_R^S\left(\int_S^T L_0(t,s)ds\right)^2 dt+\int_S^T |D_r^i Y(s)|^2 ds\\
&+\int_R^S\int_S^T |D_r^i Z(s,t)|^2 ds dt\Bigg\}\\
\le CE&\left\{\int_R^T |\psi(t)|^2 dt+\sum_{i=1}^d\int_R^T |D_r^i\psi(t)|^2 dt+\int_R^T\left(\int_t^T g_0(t,s)ds\right)^2 dt\right.\\
&\left.+\int_R^T\left(\int_t^T L_0(t,s)ds\right)^2 dt\right\}. \qquad (4.36)
\end{aligned}
$$

Now, we can consider the following BSVIE:

$$
Y(t)=\psi^S(t)+\int_t^S g(t,s,Y(s),Z(t,s),Z(s,t))ds-\int_t^S Z(t,s)dW(s),\quad t\in[0,S]. \qquad (4.37)
$$

By induction, we can obtain (4.13)–(4.17). Conclusions (4.15)–(4.16) are obvious.

Next, in the case that (H2)′ holds, from (4.4) and (4.13), we have

$$
\begin{aligned}
E\int_t^T |Z(s,t)|^2 ds &\le \sum_{i=1}^d E\int_t^T |D_t^i Y(s)|^2 ds\\
&\le CE\left\{\int_t^T |\psi(s)|^2 ds+\sum_{i=1}^d\int_t^T |D_t^i\psi(s)|^2 ds+\int_t^T\left(\int_s^T g_0(s,\tau)d\tau\right)^2 ds\right.\\
&\qquad\left.+\int_0^T\left(\int_s^T L_0(s,\tau)d\tau\right)^2 ds\right\}\\
&\le C\left\{\|\psi(\cdot)\|^2_{\Psi[0,T]}+\sup_{t\in[0,T]} E\left(\int_t^T g_0(t,s)ds\right)^2+E\int_0^T\left(\int_t^T L_0(t,s)ds\right)^2 dt\right\}. \qquad (4.38)
\end{aligned}
$$

Combining the above with (5.1), we obtain

$$E\left\{|Y(t)|^2+\int_0^T|Z(t,s)|^2ds+\int_t^T|Z(s,t)|^2ds\right\}$$
$$\leq C\left\{\|\psi(\cdot)\|^2_{\Psi[0,T]}+\sup_{t\in[0,T]}E\left(\int_t^T g_0(t,s)ds\right)^2+E\int_0^T\left(\int_t^T L_0(t,s)ds\right)^2dt\right\}. \tag{4.39}$$

Then (4.18) follows from the above and (4.13). □

Note that the above theorem does not give the continuity of $t\mapsto Y(t)$ in any sense. As a matter of fact, due to the dependence of the generator $g(t,s,y,z,\zeta)$ on ζ (namely, the term $Z(s,t)$ appears in the drift of the BSVIE), the continuity of $Y(\cdot)$ becomes very subtle. The next result is about the continuity of $Y(\cdot)$.

Theorem 4.2 *Let* (H2)$'$ *hold. Then*

$$E\left\{|Y(t)-Y(\bar t)|^2+\int_{t\vee\bar t}^T|D_t^iY(s)-D_{\bar t}^iY(s)|^2ds+\int_{t\vee\bar t}^T|Z(t,s)-Z(\bar t,s)|^2ds\right.$$
$$+\int_S^{t\wedge\bar t}|Z(t,s)-Z(\bar t,s)|^2ds+\int_{t\vee\bar t}^T|Z_i(s,t)-Z_i(s,\bar t)|^2ds$$
$$\left.+\int_{t\vee\bar t}^T\int_{t\vee\bar t}^T|D_t^iZ(s,\tau)-D_{\bar t}^iZ(s,\tau)|^2d\tau ds\right\}$$
$$\leq CE\left\{|\psi(t)-\psi(\bar t)|^2+\int_{t\vee\bar t}^T|D_t^i\psi(s)-D_{\bar t}^i\psi(s)|^2ds\right.$$
$$+\left(\int_{t\vee\bar t}^T|g(t,s,Y(s),Z(t,s),Z(s,t))-g(\bar t,s,Y(s),Z(t,s),Z(s,t))|ds\right)^2$$
$$+\int_{t\vee\bar t}^T\left(\int_{t\vee\bar t}^T|[D_t^ig](s,\tau,Y(\tau),Z(s,\tau),Z(\tau,s))\right.$$
$$\left.-[D_{\bar t}^ig](s,\tau,Y(\tau),Z(s,\tau),Z(\tau,s))|d\tau\right)^2ds$$
$$+\left(\int_{t\wedge\bar t}^{t\vee\bar t}|g(t,s,Y(s),Z(t,s),Z(s,t))|ds\right)^2+\int_{t\wedge\bar t}^{t\vee\bar t}|Z(t,s)|^2ds$$

$$+\int_{t\vee\bar t}^{T}\Big|E[D_t^iY(s)\mid\mathcal{F}_t]-E[D_t^iY(s)\mid\mathcal{F}_{\bar t}]\Big|^2ds\Bigg\},\qquad\forall t,\bar t\in[0,T].\tag{4.40}$$

Consequently, in the case that $\psi(\cdot)\in\Psi_c[0,T]$ *and* $t\mapsto g(t,s,y,z,\zeta)$ *and* $t\mapsto[D_t^ig](s,\tau,s,y,z,\zeta)$ $(1\le i\le d)$ *are all continuous, we have*

$$(Y(\cdot),Z(\cdot\,,\cdot))\in\mathcal{Y}_c[0,T]\times\mathcal{Z}_c[0,T].\tag{4.41}$$

Proof By Corollary 3.6 with $p=2$ [see (3.31) with $h(t,s,z)=g(t,s,Y(s),z,Z(s,t))$],

$$\begin{aligned}
&E\Bigg\{|Y(t)-Y(\bar t)|^2+\int_{t\vee\bar t}^{T}|Z(t,s)-Z(\bar t,s)|^2ds+\int_{S}^{t\wedge\bar t}|Z(t,s)-Z(\bar t,s)|^2ds\Bigg\}\\
&\le CE\Bigg\{|\psi(t)-\psi(\bar t)|^2+\Bigg(\int_{t\wedge\bar t}^{t\vee\bar t}|g(t,s,Y(s),Z(t,s),Z(s,t))|ds\Bigg)^2\\
&\qquad+\int_{t\wedge\bar t}^{t\vee\bar t}|Z(t,s)|^2ds\\
&\quad+\Bigg(\int_{t\vee\bar t}^{T}|g(t,s,Y(s),Z(t,s),Z(s,t))-g(\bar t,s,Y(s),Z(t,s),Z(s,\bar t))|ds\Bigg)^2\Bigg\}\\
&\le CE\Bigg\{|\psi(t)-\psi(\bar t)|^2+\Bigg(\int_{t\wedge\bar t}^{t\vee\bar t}|g(t,s,Y(s),Z(t,s),Z(s,t))|ds\Bigg)^2\\
&\qquad+\int_{t\wedge\bar t}^{t\vee\bar t}|Z(t,s)|^2ds\\
&\quad+\Bigg(\int_{t\vee\bar t}^{T}|g(t,s,Y(s),Z(t,s),Z(s,t))-g(\bar t,s,Y(s),Z(t,s),Z(s,t))|ds\Bigg)^2\\
&\quad+\int_{t\vee\bar t}^{T}|Z(s,t)-Z(s,\bar t)|^2ds\Bigg\}.
\end{aligned}\tag{4.42}$$

Thus, we need to estimate the last term on the right hand side of the above. To this end, let $t,\bar t\in[0,T]$. By (4.17), we have

$$\begin{aligned} D_t^i Y(s) - D_{\bar t}^i Y(s) &= D_t^i \psi(s) - D_{\bar t}^i \psi(s) \\ &+ \int_s^T \Big\{ [D_t^i g](s,\tau,Y(\tau),Z(s,\tau),Z(\tau,s)) - [D_{\bar t}^i g](s,\tau,Y(\tau),Z(s,\tau),Z(\tau,s)) \\ &\quad + g_y(s,\tau,Y(\tau),Z(s,\tau),Z(\tau,s))[D_t^i Y(\tau) - D_{\bar t}^i Y(\tau)] \\ &\quad + \sum_{i=1}^d \Big[g_{z_j}(s,\tau,Y(\tau),Z(s,\tau),Z(\tau,s))[D_t^i Z_j(s,\tau) - D_{\bar t}^i Z_j(s,\tau)] \\ &\quad + g_{\zeta_j}(s,\tau,Y(\tau),Z(s,\tau),Z(\tau,s))[D_t^i Z_j(\tau,s) - D_{\bar t}^i Z_j(\tau,s)] \Big\} d\tau \\ &- \int_s^T [D_t^i Z(s,\tau) - D_{\bar t}^i Z(s,\tau)] dW(\tau), \quad s \in [t \vee \bar t, T]. \end{aligned} \tag{4.43}$$

Consequently, by Theorem 3.7,

$$\begin{aligned} &E\left\{ \int_{t\vee\bar t}^T |D_t^i Y(s) - D_{\bar t}^i Y(s)|^2 ds + \int_{t\vee\bar t}^T \int_{t\vee\bar t}^T |D_t^i Z(s,\tau) - D_{\bar t}^i Z(s,\tau)|^2 d\tau ds \right\} \\ &\le CE\left\{ \int_{t\vee\bar t}^T |D_t^i \psi(s) - D_{\bar t}^i \psi(s)|^2 ds \right. \\ &\qquad + \int_{t\vee\bar t}^T \Bigg(\int_{t\vee\bar t}^T |[D_t^i g](s,\tau,Y(\tau),Z(s,\tau),Z(\tau,s)) \\ &\qquad \left. - [D_{\bar t}^i g](s,\tau,Y(\tau),Z(s,\tau),Z(\tau,s))| d\tau \Bigg)^2 ds \right\}. \end{aligned} \tag{4.44}$$

Next, by (4.16), we have

$$\begin{aligned} E\int_{t\vee\bar t}^T |Z_i(s,t) - Z_i(s,\bar t)|^2 ds &= E\int_{t\vee\bar t}^T |E[D_t^i Y(s) \mid \mathcal{F}_t] - E[D_{\bar t}^i Y(s) \mid \mathcal{F}_{\bar t}]|^2 ds \\ &\le 2E\int_{t\vee\bar t}^T \Big\{ |E[D_t^i Y(s) \mid \mathcal{F}_t] - E[D_t^i Y(s) \mid \mathcal{F}_{\bar t}]|^2 + |E[D_t^i Y(s) - D_{\bar t}^i Y(s) \mid \mathcal{F}_{\bar t}]|^2 \Big\} ds \\ &\le 2E\int_{t\vee\bar t}^T \Big\{ |E[D_t^i Y(s) \mid \mathcal{F}_t] - E[D_t^i Y(s) \mid \mathcal{F}_{\bar t}]|^2 + |D_t^i Y(s) - D_{\bar t}^i Y(s)|^2 \Big\} ds. \end{aligned} \tag{4.45}$$

Then (4.40) follows, which further yields (4.41). □

5 An optimal control problem

In this section, we discuss an optimal control problem for stochastic Volterra integral equations with Bolza type cost functionals. More precisely, we consider the following state equation:

$$X(t)=\varphi(t)+\int_0^t b(t,s,X(s),u(s))ds+\int_0^t \sigma(t,s,X(s),u(s))dW(s),\quad t\in[0,T], \tag{5.1}$$

where $X(\cdot)$ and $u(\cdot)$ are *state* and *control processes*, respectively; $b:\Delta\times\mathbb{R}^n\times U\to\mathbb{R}^n$, $\sigma:\Delta\times\mathbb{R}^n\times U\to\mathbb{R}^{n\times d}$ are given maps and $\varphi(\cdot)\in L^p_{\mathbb{F}}(\Omega;C([0,T];\mathbb{R}^n))$; $U\subseteq\mathbb{R}^n$ is convex. The cost functional is defined to be the following Bolza form:

$$J(u(\cdot))=E\left[\int_0^T g(t,X(t),u(t))dt+h(X(T))\right], \tag{5.2}$$

with $g:[0,T]\times\mathbb{R}^n\times U\to\mathbb{R}$ and $h:\mathbb{R}^n\to\mathbb{R}$ being given maps as well. In the above, all the functions can be random. Note that in [33], only the case $h=0$ (and one-dimensional situation) was considered.

We now introduce the following assumption. The conditions assumed are more than sufficient. One can relax many of them. But we prefer these strong conditions to make the presentation simple.

(H3) Let b,σ, and h be continuous in (t,s,X,u), and differentiable in the variables X and u, with bounded derivatives. Also,

$$|b(t,s,0,u)|+|\sigma(t,s,0,u)|\le C,\quad\forall(t,s)\in\Delta,\ u\in U. \tag{5.3}$$

We let

$$\mathcal{U}\stackrel{\Delta}{=}\left\{u:[0,T]\times\Omega\to U\ \Big|\ u(\cdot)\text{ is }\mathbb{F}\text{-progressively measurable}\right\}. \tag{5.4}$$

It is not hard to show that under (H3), for any $\varphi(\cdot)\in L^p_{\mathbb{F}}(\Omega;C([0,T];\mathbb{R}^n))$ and $u(\cdot)\in\mathcal{U}$, (5.1) admits a unique solution $X(\cdot)\in L^p_{\mathbb{F}}(\Omega;C([0,T];\mathbb{R}^n))$. Thus the cost functional $J(u(\cdot))$ is well-defined. Our optimal control problem can be stated as follows.

Problem (C). Find a $\bar u(\cdot)\in\mathcal{U}$ such that

$$J(\bar u(\cdot))=\inf_{u(\cdot)\in\mathcal{U}}J(u(\cdot)). \tag{5.5}$$

Any $\bar u(\cdot)$ satisfying (5.5) is called an *optimal control* of Problem (C), the corresponding state process $\bar X(\cdot)$ is called an *optimal state process* and $(\bar X(\cdot),\bar u(\cdot))$ is called an *optimal pair*.

Next result is called the *duality principle* of linear stochastic integral equations, which will play an important role below.

Theorem 5.1 *Let* $A_i(\cdot\,,\cdot) \in L^\infty([0,T]; L^\infty_{\mathbb{F}}(0,T;\mathbb{R}^{n\times n}))$ $(i = 0,1,\ldots,d)$, $\bar\varphi(\cdot) \in L^2_{\mathbb{F}}(0,T;\mathbb{R}^n)$, *and* $\psi(\cdot) \in L^2((0,T)\times\Omega;\mathbb{R}^n)$. *Let* $\xi(\cdot) \in L^2_{\mathbb{F}}(0,T;\mathbb{R}^n)$ *be the solution of FSVIE*:

$$\xi(t) = \varphi(t) + \int_0^t A_0(t,s)\xi(s)ds + \int_0^t \sum_{i=1}^d A_i(t,s)\xi(s)dW_i(s), \quad t \in [0,T], \tag{5.6}$$

and $(Y(\cdot), Z(\cdot\,,\cdot)) \in \mathcal{H}^2[0,T]$ *be the adapted M-solution to the following BSVIE:*

$$\begin{aligned} Y(t) = \psi(t) + \int_t^T \Big[A_0(s,t)^T Y(s) + \sum_{i=1}^d A_i(s,t)^T Z_i(s,t) \Big] ds \\ - \int_t^T Z(t,s)dW(s), \quad t \in [0,T]. \end{aligned} \tag{5.7}$$

Then the following relation holds:

$$E\int_0^T \langle\, \xi(t), \psi(t)\,\rangle\, dt = E\int_0^T \langle\, \varphi(t), Y(t)\,\rangle\, dt. \tag{5.8}$$

Further, suppose for each $i = 0,1,\cdots,d$, $t \mapsto A_i(t,s)$ is continuous at $t = T$, $\varphi(\cdot) \in C([0,T];\mathbb{R}^n)$, and $\eta \in L^2_{\mathcal{F}_T}(\Omega;\mathbb{R}^n)$. Let $(Y(\cdot), Z(\cdot\,,\cdot)) \in \mathcal{H}^2[0,T]$ be the adapted solution to the following BSVIE:

$$\begin{aligned} Y(t) &= \psi(t) + A_0(T,t)^T\eta + \sum_{i=1}^d A_i(T,t)^T \zeta_i(t) \\ &\quad + \int_t^T \Big[A_0(s,t)^T Y(s) + \sum_{i=1}^d A_i(s,t)^T Z_i(s,t) \Big] ds \\ &\quad - \int_t^T Z(t,s)dW(s), \quad t \in [0,T], \end{aligned} \tag{5.9}$$

where $\zeta(\cdot) \equiv (\zeta_1(\cdot),\ldots,\zeta_d(\cdot)) \in L^2_{\mathbb{F}}(0,T;\mathbb{R}^{n\times d})$ is the unique process that satisfies

$$\eta = E\eta + \int_0^T \zeta(t)dW(t). \tag{5.10}$$

Then the following holds:

$$E\left\{\langle\,\xi(T),\eta\,\rangle+\int_0^T\langle\,\xi(t),\psi(t)\,\rangle\,dt\right\}=E\left\{\langle\,\bar\varphi(T),\eta\,\rangle+\int_0^T\langle\,\bar\varphi(t),Y(t)\,\rangle\,dt\right\}. \tag{5.11}$$

Proof We first look at (5.8). Observe the following:

$$\begin{aligned}
&E\int_0^T\langle\,\bar\varphi(t),Y(t)\,\rangle\,dt\\
&=E\int_0^T\left\langle\xi(t)-\int_0^t A_0(t,s)\xi(s)ds-\sum_{i=1}^d\int_0^t A_i(t,s)\xi(s)dW_i(s),Y(t)\right\rangle dt\\
&=E\int_0^T\langle\,\xi(t),Y(t)\,\rangle\,dt-E\int_0^T\int_s^T\langle\,\xi(s),A_0(t,s)^TY(t)\,\rangle\,dtds\\
&\quad-\sum_{i=1}^d E\int_0^T\left\langle\int_0^t A_i(t,s)\xi(s)dW_i(s),EY(t)+\int_0^t Z(t,s)dW(s)\right\rangle dt\\
&=E\int_0^T\langle\,\xi(t),Y(t)\,\rangle\,dt-E\int_0^T\int_t^T\langle\,\xi(t),A_0(s,t)^TY(s)\,\rangle\,dsdt\\
&\quad-\sum_{i=1}^d E\int_0^T\int_0^t\langle\,A_i(t,s)\xi(s),Z_i(t,s)\,\rangle\,dsdt\\
&=E\int_0^T\left\langle\xi(t),Y(t)-\int_t^T\left[A_0(s,t)^TY(s)+\sum_{i=1}^d A_i(s,t)^TZ_i(s,t)\right]ds\right\rangle dt\\
&=E\int_0^T\left\langle\xi(t),\psi(t)-\int_t^T Z(t,s)dW(s)\right\rangle dt=E\int_0^T\langle\,\xi(t),\psi(t)\,\rangle\,dt.
\end{aligned} \tag{5.12}$$

Thus, (5.8) holds. Now, for (5.11), we denote

$$\hat\psi(t)=\psi(t)+A_0(T,t)^T\eta+\sum_{i=1}^d A_i(T,t)^T\zeta_i(t). \tag{5.13}$$

Then by what we have proved, one has

$$
\begin{aligned}
E\int_0^T \langle\, \bar\varphi(t), Y(t)\,\rangle\, dt &= E\int_0^T \langle\, \xi(t), \hat\psi(t)\,\rangle\, dt \\
&= E\int_0^T \langle\, \xi(t), \psi(t) + A_0(T,t)^T\eta + \sum_{i=1}^d A_i(T,t)^T\zeta_i(t)\,\rangle\, dt \\
&= E\int_0^T \Big(\langle\, \xi(t), \psi(t)\,\rangle + \langle\, A_0(T,t)\xi(t), \eta\,\rangle + \sum_{i=1}^d \langle\, A_i(T,t)\xi(t), \zeta_i(t)\,\rangle \Big) dt \\
&= E\int_0^T \langle\, \xi(t), \psi(t)\,\rangle\, dt + E\,\langle\, \xi(T) - \bar\varphi(T), \eta\,\rangle\, .
\end{aligned}
\tag{5.14}
$$

This proves (5.11). □

We now prove the following theorem called Pontryagin's maximum principle.

Theorem 5.2 *Let* (H3) *hold. Let* $(\bar X(\cdot), \bar u(\cdot))$ *be an optimal pair of Problem* (C)*. Then there exists a unique adapted M-solution* $(Y(\cdot), Y_0(\cdot), \eta(\cdot); Z(\cdot\,,\cdot), Z_0(\cdot\,,\cdot), \zeta(\cdot))$ *of the following BSVIE*:

$$
\begin{cases}
Y(t) = g_x(t,\bar X(t),\bar u(t)) + b_x(T,t,\bar X(t),\bar u(t))^T h_x(\bar X(T)) + \sum_{i=1}^d \sigma_x^i(T,t,\bar X(t),\bar u(t))^T\zeta_i(t) \\
\qquad + \int_t^T \Big[b_x(s,t,\bar X(t),\bar u(t))^T Y(s) + \sum_{i=1}^d \sigma_x^i(s,t,\bar X(t),\bar u(t)) Z_i(s,t) \Big] ds - \int_t^T Z(t,s)dW(s), \\
Y_0(t) = b_u(T,t,\bar X(t),\bar u(t))^T h_x(\bar X(T)) + \sum_{i=1}^d \sigma_u^i(T,t,\bar X(t),\bar u(t))^T\zeta_i(t) \\
\qquad + \int_t^T \Big[b_u(s,t,\bar X(t),\bar u(t))^T Y(s) + \sum_{i=1}^d \sigma_u^i(s,t,\bar X(t),\bar u(t))^T Z_i(s,t) \Big] ds - \int_t^T Z_0(t,s)dW(s), \\
\eta(t) = h_x(\bar X(T)) - \int_t^T \zeta(s)dW(s),
\end{cases}
\tag{5.15}
$$

such that

$$
\langle\, Y_0(t) + g_u(t,\bar X(t),\bar u(t)), u - \bar u(t)\,\rangle \ge 0, \quad \forall u\in U,\ t\in[0,T],\ \text{a.s.} \tag{5.16}
$$

Note that the third equation in (5.15) is actually a BSDE, and the second equation in (5.15) is a BSVIE in which the drift and the free term do not depend on $(Y_0(\cdot), Z_0(\cdot\,,\cdot))$.

Proof Let $(\bar{X}(\cdot), \bar{u}(\cdot))$ be an optimal pair of Problem (C). Take any $u(\cdot) \in \mathcal{U}$. Since $\mathcal{U}$ is convex, for any $\varepsilon \in (0, 1)$,

$$u_\varepsilon(\cdot) \stackrel{\Delta}{=} \bar{u}(\cdot) + \varepsilon[u(\cdot) - \bar{u}(\cdot)] \in \mathcal{U}. \tag{5.17}$$

Let $X_\varepsilon(\cdot)$ be the solution of (5.1) corresponding to $u_\varepsilon(\cdot)$. Define

$$\xi_\varepsilon(t) = \frac{X_\varepsilon(t) - \bar{X}(t)}{\varepsilon}, \quad t \in [0, T]. \tag{5.18}$$

Then $\xi_\varepsilon(\cdot) \to \xi(\cdot)$ in $L^2_{\mathbb{F}}(0, T; \mathbb{R}^n)$ with $\xi(\cdot)$ satisfying the following:

$$\begin{aligned}\xi(t) &= \int_0^t \big\{b_x(t, s, \bar{X}(s), \bar{u}(s))\xi(s) + b_u(t, s, \bar{X}(s), \bar{u}(s))[u(s) - \bar{u}(s)]\big\}\, ds\\ &\quad + \int_0^t \sum_{i=1}^d \Big\{\sigma_x^i(t, s, \bar{X}(s), \bar{u}(s))\xi(s) + \sigma_u^i(t, s, \bar{X}(s), \bar{u}(s))[u(s) - \bar{u}(s)]\Big\}\, dW_i(s)\\ &\equiv \bar{\varphi}(t) + \int_0^t b_x(t, s, \bar{X}(s), \bar{u}(s))\xi(s)ds + \int_0^t \sum_{i=1}^d \sigma_x^i(t, s, \bar{X}(s), \bar{u}(s))\xi(s)dW_i(s),\end{aligned} \tag{5.19}$$

where

$$\begin{aligned}\bar{\varphi}(t) &= \int_0^t b_u(t, s, \bar{X}(s), \bar{u}(s))[u(s) - \bar{u}(s)]ds\\ &\quad + \int_0^t \sum_{i=1}^d \sigma_u^i(t, s, \bar{X}(s), \bar{u}(s))[u(s) - \bar{u}(s)]dW_i(s), \qquad t \in [0, T].\end{aligned} \tag{5.20}$$

Now, let $(Y(\cdot), Y_0(\cdot), Z(\cdot\,,\cdot), Z_0(\cdot\,,\cdot))$ be the unique adapted M-solution to BSVIE (5.15). By the optimality of $(\bar{X}(\cdot), \bar{u}(\cdot))$, and the duality principle (Theorem 5.1), we have

$$\begin{aligned}0 &\le \frac{J(u_\varepsilon(\cdot)) - J(\bar{u}(\cdot))}{\varepsilon}\\ &\to E\Bigg\{\langle h_x(\bar{X}(T)), \xi(T)\rangle + \int_0^T \big[\langle g_x(t, \bar{X}(t), \bar{u}(t)), \xi(t)\rangle\\ &\qquad + \langle g_u(t, \bar{X}(t), \bar{u}(t)), u(t) - \bar{u}(t)\rangle\big]\, dt\Bigg\}\end{aligned}$$

$$
\begin{aligned}
&= E\left\{\langle h_x(\bar{X}(T)), \bar{\varphi}(T)\rangle + \int_0^T [\langle Y(t), \bar{\varphi}(t)\rangle \right.\\
&\qquad \left. + \langle g_u(t, \bar{X}(t), \bar{u}(t)), u(t) - \bar{u}(t)\rangle\big] dt\right\}\\
&= E\int_0^T \langle b_u(T, t, \bar{X}(t), \bar{u}(t))^T h_x(\bar{X}(T)) + \sum_{i=1}^d \sigma_u^i(T, t, \bar{X}(t), \bar{u}(t))^T \zeta_i(t)\\
&\quad + \int_t^T \left[b_u(s, t, \bar{X}(t), \bar{u}(t))^T Y(s) + \sum_{i=1}^d \sigma_u^i(s, t, \bar{X}(t), \bar{u}(t))^T Z_i(s, t)\right] ds\\
&\quad + g_u(t, \bar{X}(t), \bar{u}(t)), u(t) - \bar{u}(t)\rangle dt\\
&= E\int_0^T \langle Y_0(t) + g_u(t, \bar{X}(t), \bar{u}(t)), u(t) - \bar{u}(t)\rangle dt.
\end{aligned}
$$

Since the above holds for all $u(\cdot) \in \mathcal{U}[0, T]$, we obtain (5.16). □

Now, we define

$$
\begin{aligned}
&H(t, \bar{X}(t), \bar{u}(t), Y(\cdot), Z(\cdot\,, t), u) \triangleq -\left[Y_0(t) + g_u(t, \bar{X}(t), \bar{u}(t))\right]^T u\\
&= -E\left\{ g_u(t, \bar{X}(t), \bar{u}(t)) + b_u(T, t, \bar{X}(t), \bar{u}(t))^T h_x(\bar{X}(T)) \right.\\
&\qquad + \sum_{i=1}^d \sigma_u^i(T, t, \bar{X}(t), \bar{u}(t))^T \zeta_i(t)\\
&\quad \left. + \int_t^T \left[b_u(s, t, \bar{X}(t), \bar{u}(t))^T Y(s) + \sum_{i=1}^d \sigma_u^i(s, t, \bar{X}(t), \bar{u}(t))^T Z_i(s, t)\right] ds \,\Big|\, \mathcal{F}_t \right\}^T u,
\end{aligned}
\tag{5.21}
$$

then (5.16) can be written as

$$
H(t, \bar{X}(t), \bar{u}(t), Y(\cdot), Z(\cdot\,, t), \bar{u}(t)) = \max_{u \in U} H(t, \bar{X}(t), \bar{u}(t), Y(\cdot), Z(\cdot\,, t), u).
\tag{5.22}
$$

We call $H(\cdot)$ defined by (5.21) the *Hamiltonian* of our optimal control problem, call (5.16) (and (5.22)) the *maximum condition*, and call the first BSVIE in (4.1) the *adjoint equation* of (4.5), along the optimal pair $(\bar{X}(\cdot), \bar{u}(\cdot))$.

By putting (1.1), and (5.15)–(5.16) together (dropping the bars in $(\bar{X}(\cdot), \bar{u}(\cdot))$), we obtain the following system:

$$
\begin{cases}
X(t)=\varphi(t)+\int_0^t b(t,s,X(s),u(s))ds+\int_0^t \sigma(t,s,X(s),u(s))dW(s),\\
Y(t)=g_x(t,X(t),u(t))+b_x(T,t,X(t),u(t))^T h_x(X(T))+\sum_{i=1}^d \sigma_x^i(T,t,X(t),u(t))^T\zeta_i(t)\\
\qquad +\int_t^T\Big[b_x(s,t,X(t),u(t))^T Y(s)+\sum_{i=1}^d \sigma_x^i(s,t,X(t),u(t))^T Z_i(s,t)\Big]ds-\int_t^T Z(t,s)dW(s),\\
Y_0(t)=b_u(T,t,X(t),u(t))^T h_x(X(T))+\sum_{i=1}^d \sigma_u^i(T,t,X(t),u(t))^T\zeta_i(t)\\
\qquad +\int_t^T\Big[b_u(s,t,X(t),u(t))^T Y(s)+\sum_{i=1}^d \sigma_u^i(s,t,X(t),u(t))^T Z_i(s,t)\Big]ds-\int_t^T Z_0(t,s)dW(s),\\
\eta(t)=h_x(X(T))-\int_t^T \zeta(s)dW(s),\\
[Y_0(t)+g_u(t,X(t),u(t))]^T\,[v-u(t)]\ge 0,\quad \forall v\in U,\ t\in[0,T],\ \text{a.s.}
\end{cases}
\tag{5.23}
$$

This is a couple systems of FSVIE and BSVIE. The coupling is through the maximum condition (via $u(\cdot)$). We call (5.23) a forward–backward stochastic Volterra integral equation (FBSVIE, for short). To get some feeling about the above results, let us look at an important special case—the linear-quadratic (LQ, for short) problem. Thus, we consider the following linear state equation:

$$
\begin{aligned}
X(t)=\varphi(t)+&\int_0^t [A_0(t,s)X(s)+B_0(t,s)u(s)]\,dt\\
&+\sum_{i=1}^d\int_0^t [A_i(t,s)X(s)+B_i(t,s)u(s)]\,dW_i(s),
\end{aligned}
\tag{5.24}
$$

with the cost functional:

$$
\begin{aligned}
J(\varphi(\cdot),u(\cdot))=\frac{1}{2}E\Bigg\{&\int_0^T [\langle\, Q(t)X(t),X(t)\,\rangle+\langle\, R(t)u(t),u(t)\,\rangle]\,dt\\
&+\langle\, GX(T),X(T)\,\rangle\Bigg\}.
\end{aligned}
\tag{5.25}
$$

For the simplicity of presentation, we make the following assumption.

(H4) Let $T>0$. Let maps $A_0,\dots,A_d:\Delta\to\mathbb{R}^{n\times n}$, $B_0,\dots,B_d:\Delta\to\mathbb{R}^{n\times m}$, $Q:[0,T]\to\mathcal{S}^n$, $R:[0,T]\to\mathcal{S}^m$ be all continuous, and $G\in\mathcal{S}^n$, where $\mathcal{S}^n$ is the set of all $(n\times n)$ symmetric matrices. Assume that Q and G are positive semi-definite, R is uniform positive definite. Let $\varphi(\cdot)\in L^2_{\mathbb{F}}(\Omega;C([0,T];\mathbb{R}^n))$.

We refer to the problem of minimizing (5.25) subject to (5.24) over the control set $\mathcal{U} = L^2_{\mathbb{F}}(0, T; \mathbb{R}^m)$ as Problem (LQ). Note that all the coefficients are assumed to be deterministic. In the case that some or all coefficients are allowed to be random, the discussion can still be carried out, with a little more complicated looking. It is not hard to see that under (H4), for any $\varphi(\cdot) \in L^2_{\mathbb{F}}(\Omega; C([0, T]; \mathbb{R}^n))$, the functional $u(\cdot) \mapsto J(\varphi(\cdot), u(\cdot))$ is quadratic and uniformly positive definite. Thus, there exists a unique optimal control $u(\cdot)$. Further, by a result similar to Theorem 5.2, and regarding our LQ problem as an abstract minimization problem for a quadratic functional with linear constraints, we see that the following FBSVIE admits a unique adapted M-solution:

$$
\begin{cases}
X(t) = \varphi(t) + \displaystyle\int_0^t \Big[A_0(t,s)X(s) - B_0(t,s)R(s)^{-1}Y_0(s)\Big]dt \\
\qquad + \displaystyle\sum_{i=1}^d \int_0^t \Big[A_i(t,s)X(s) - B_i(t,s)R(s)^{-1}Y_0(s)\Big]dW_i(s), \\
Y(t) = Q(t)X(t) + A_0(T,t)^T GX(T) + \displaystyle\sum_{i=1}^d A_i(T,t)^T \zeta_i(t) \\
\qquad + \displaystyle\int_t^T \Big[A_0(s,t)^T Y(s) + \sum_{i=1}^d A_i(s,t)^T Z_i(s,t)\Big] ds - \int_t^T Z(t,s)dW(s), \\
Y_0(t) = B_0(T,t)^T GX(T) + \displaystyle\sum_{i=1}^d B_i(T,t)^T \zeta_i(t) \\
\qquad + \displaystyle\int_t^T \Big[B_0(s,t)^T Y(s) + \sum_{i=1}^d B_i(s,t)^T Z_i(s,t)\Big] ds - \int_t^T Z_0(t,s)dW(s), \\
\eta(t) = GX(T) - \displaystyle\int_t^T \zeta(s)dW(s),
\end{cases}
\tag{5.26}
$$

by which we mean that $X(\cdot)$ satisfies the first FSVIE in the usual sense and $(Y(\cdot), Y_0(\cdot), \eta(\cdot), Z(\cdot\,,\cdot), Z_0(\cdot\,,\cdot), \zeta(\cdot))$ is the adapted M-solution of the corresponding BSVIE. The optimal control is given by

$$
u(t) = -R(t)^{-1}Y_0(t), \qquad t \in [0, T]. \tag{5.27}
$$

Since all the coefficients are deterministic, in the case that $\varphi(\cdot) \in \Psi_c[0, T]$, for any $1 \le i \le d$ and $r \in [0, T]$, $(D^i_r X(\cdot), D^i_r Y(\cdot), D^i_r Y_0(\cdot), D^i_r \eta(\cdot), D^i_r Z(\cdot\,,\cdot), D^i_r Z_0(\cdot\,,\cdot), D^i_r \zeta(\cdot))$ will satisfy the same FBSVIE as (5.26), with $\varphi(\cdot)$ replaced by $D^i_r\varphi(\cdot)$. This will lead to $X(\cdot), \zeta_i(\cdot) \in \Psi_c[0, T]$. Then by the regularity results from Section 4, we have the continuity of $Y(\cdot)$ and $Y_0(\cdot)$. Hence the optimal control $u(\cdot)$ given by (5.27) is also continuous. Along this line, there are some other related interesting problems, including the so-called causality of the optimal control ([26,35]), corresponding

Riccati type equations, etc. Relevant investigations are still undergoing, and results will be reported in our forthcoming papers.

References

1. Artzner, Ph., Delbaen, F., Jean-Marc, E., Heath, D.: Coherent measures of risk. Math. Finance **9**, 203–228 (1999)
2. Belbas, S.A.: Iterative schemes for optimal control of Volterra integral equations. Nonlinear Anal. **37**, 57–79 (1999)
3. Belbas, S.A., Schmidt, W.H.: Optimal control of Volterra equations with impulses. Appl. Math. Comput. **166**, 696–723 (2005)
4. Bismut, J.B.: Théorie Probabiliste du Contrôle des Diffusions. Mem. Am. Math. Soc. 176, Providence, Rhode Island (1973)
5. Briand, P., Delyon, B., Hu, Y., Pardoux, E., Stoica, L.: L^p solutions of backward stochastic differential equations. Stoch. Proc. Appl. **108**, 109–129 (2003)
6. Burnap, C., Kazemi, M.: Optimal control of a system governed by nonlinear Volterra integral equations with delay. IMA J. Math. Control Inform. **16**, 73–89 (1999)
7. Carlson, D.A.: An elementary proof of the maximum principle for optimal control problems governed by a Volterra integral equation. J. Optim. Theory Appl. **54**, 43–61 (1987)
8. Cheridito, P., Delbaen, F., Kupper, M.: Coherent and convex monetary risk measures for bounded càdlàg processes. Stoch. Proc. Appl. **112**, 1–22 (2004)
9. Duffie, D., Epstein, L.: Stochastic differential utility. Econometrica **60**, 353–394 (1992)
10. Duffie, D., Huang, C.F.: Stochastic production-exchange equilibria. Research paper No. 974, Graduate School of Business, Stanford University, Stanford (1986)
11. Ekeland, I., Lazrak, A.: Non-commitment in continuous time. Preprint (2006)
12. El Karoui, N., Peng, S., Quenez, M.C.: Backward stochastic differential equations in finance. Math. Finance **7**, 1–71 (1997)
13. Kamien, M.I., Muller, E.: Optimal control with integral state equations. Rev. Econ. Stud. **43**, 469–473 (1976)
14. Karatzas, I., Shreve, S.E.: Brownian Motion and Stochastic Calculus. Springer, Heidelberg (1988)
15. Kobylanski, M.: Backward stochastic differential equations and partial differential equations with quadratic growth. Ann. Probab. **28**, 558–602 (2000)
16. Laibson, D.: Golden eggs and hyperbolic discounting. Q. J. Econ. **42**, 443–477 (1997)
17. Lin, J.: Adapted solution of a backward stochastic nonlinear Volterra integral equation. Stoch. Anal. Appl. **20**, 165–183 (2002)
18. Ma, J., Yong, J.: Forward-Backward Stochastic Differential Equations and Their Applications, Springer-Verlag, Berlin. Springer, Berlin (1999)
19. Medhin, N.G.: Optimal processes governed by integral equations. J. Math. Anal. Appl. **120**, 1–12 (1986)
20. Neustadt, L.W., Warga, J.: Comments on the paper "Optimal control of processes described by integral equations. I" by V.R. Vinokurov. SIAM J. Control **8**, 572 (1970)
21. Nualart, D.: The Malliavin Calculus and Related Topics. Springer, Heidelberg (1995)
22. Pardoux, E., Peng, S.: Adapted solutions of backward stochastic equations. Syst. Control Lett. **14**, 55–61 (1990)
23. Pardoux, E., Protter, P.: Stochastic Volterra equations with anticipating coefficients. Ann. Probab. **18**, 1635–1655 (1990)
24. Peng, S.: A general stochastic maximum principle for optimal control problems. SIAM J. Control Optim. **28**, 966–979 (1990)
25. Peng, S.: Nonlinear expectations, nonlinear evaluations and risk measures. Stochastic Methods in Finance, Lecture Notes in Math., vol. 1856, pp. 165–253. Springer, Heidelberg (2004)
26. Pritchard, A.J., You, Y.: Causal feedback optimal control for Volterra integral equations. SIAM J. Control Optim. **34**, 1874–1890 (1996)
27. Protter, P.: Volterra equations driven by semimartingales. Ann. Probab. **13**, 519–530 (1985)
28. Schroder, M., Skiadas, C.: Optimal consumption and portfolio selection with stochastic utility. J. Econ. Theory **89**, 68–126 (1999)
29. Skiadas, C.: Robust control and recursive utility. Finance Stoch. **7**, 475–489 (2003)

30. Strotz, R.H.: Myopia and inconsistency in dynamic utility maximization. Rev. Econ. Stud. **23**, 165–180 (1956)
31. Vinokurov, V.R.: Optimal control of processes described by integral equations, I, II, III, Izv. Vysš. Učebn. Zaved. Matematika **7**(62), 21–33; **8**(63) 16–23; **9**(64), 16–25; (in Russian) English transl. in SIAM J. Control **7** 324–336, 337–345, 346–355 (1967)
32. Wang, T.: A class of dynamic risk measures. Preprint, University of British Columbia, BC (2002)
33. Yong, J.: Backward stochastic Volterra integral equations and some related problems. Stoch. Proc. Appl. **116**, 779–795 (2006)
34. Yong, J., Zhou, X.Y.: Stochastic Controls: Hamiltonian Systems and HJB Equations. Springer, New York (1999)
35. You, Y.: Quadratic integral games and causal synthesis. Trans. Am. Math. Soc. **352**, 2737–2764 (2000)