Backward Stochastic Volterra Integral Equations
— Representation of Adapted Solutions *

Tianxiao Wang† and Jiongmin Yong‡

February 13, 2018

Abstract

For backward stochastic Volterra integral equations (BSVIEs, for short), under some mild conditions, the so-called adapted solutions or adapted M-solutions uniquely exist. However, satisfactory regularity of the solutions is difficult to obtain in general. Inspired by the decoupling idea of forward-backward stochastic differential equations, in this paper, for a class of BSVIEs, a representation of adapted M-solutions is established by means of the so-called representation partial differential equations and (forward) stochastic differential equations. Well-posedness of the representation partial differential equations are also proved in certain sense.

Keywords. Backward stochastic Volterra integral equations, adapted solutions, representation partial differential equations, representation of adapted solutions.

AMS Mathematics subject classification. 60H20, 45D05, 35K15, 35K40.

1 Introduction.

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a complete filtered probability space on which a standard $d$-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. We consider the following stochastic integral equation in $\mathbb{R}^m$:

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

(1.1)

Such an equation is called a backward stochastic Volterra integral equation (BSVIE, for short). In the above, $\psi(\cdot)$, called a free term, is an $\mathcal{F}_T$-measurable stochastic process (not necessarily $\mathbb{F}$-adapted) and $g(\cdot)$, called the generator of the above BSVIE, is a given map, deterministic or random. The unknown that we are looking for is the pair $(Y(\cdot),Z(\cdot,\cdot))$. Let us look at a special case of the above BSVIE. Suppose

$$g(t,s,y,z,\zeta) = g(s,y,z), \quad \psi(t) = \xi, \quad \forall(t,s,y,z,\zeta),$$

with $\xi$ being an $\mathcal{F}_T$-measurable random variable, and $g(\cdot)$ being a proper map. Then the above BSVIE is reduced to the following form:

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

(1.2)

---

*This work is supported in part by NSF of China (Grant 11401404, 11471231) and NSF Grant DMS-1406776.
†School of Mathematics, Sichuan University, Chengdu 610065, China
‡Department of Mathematics, University of Central Florida, Orlando, FL 32816, USA
It is comparable with the integral form of \textit{backward stochastic differential equation} (BSDE, for short) which takes the following form:

\[
Y(t) = \xi + \int_t^T g(s, Y(s), Z(s))ds - \int_t^T Z(s)dW(s), \quad t \in [0, T].
\]

(1.3)

Now, if BSDE (1.3) admits a unique \textit{adapted solution} \((Y(\cdot), Z(\cdot))\), by which we mean that this pair is \(\mathbb{F}\)-adapted and satisfies (1.3) in the usual Itô sense, then, this solution must also be an \textit{adapted solution of BSVIE (1.2)} with \(Z(t, s) \equiv Z(s)\). From this point of view, BSVIE can be regarded as an extension of BSDE.

Linear BSDEs were firstly introduced by Bismut in 1973 ([4]) while he was studying stochastic linear-quadratic optimal control problems. In 1990, Pardoux–Peng generalized Bismut’s linear BSDEs to general nonlinear BSDEs ([26]). Shortly after, BSDE theory was found to have very interesting applications in mathematical finance (see for example, [9, 11]), and many other developments have been appearing thereafter, including nonlinear Feynman-Kac formula, nonlinear expectations, dynamic risk measures, path-dependent partial differential equations, etc., see for examples [28, 29, 30, 10], and references cited therein. Relevant theory of BSDEs can also be found in [5, 16, 6, 7]. On the other hand, an extension of BSDE to the so-called \textit{forward-backward stochastic differential equations} (FBSDEs, for short) was initiated by Antonelli in 1993 ([1]). A general form of FBSDE takes the following form:

\[
\begin{aligned}
&dX(t) = b(t, X(t), Y(t), Z(t))dt + \sigma(t, X(t), Y(t), Z(t))dW(t), \quad t \in [0, T], \\
&dY(t) = -g(t, X(t), Y(t), Z(t))dt + Z(t)dW(t), \quad t \in [0, T], \\
&X(0) = x, \quad Y(T) = h(X(T)),
\end{aligned}
\]

(1.4)

for some maps \(b, \sigma, g, h\). General theories on FBSDEs were developed in the past two and half decays, see [19, 21, 14, 36, 22], and references cited therein. A triple \((X(\cdot), Y(\cdot), Z(\cdot))\) of processes is called an \textit{adapted solution} to (1.4) if it is \(\mathbb{F}\)-adapted, and satisfies (1.4) in the usual Itô’s sense. It is known that under some very general conditions, FBSDE (1.4) admits a unique adapted solution \((X(\cdot), Y(\cdot), Z(\cdot))\), and the following estimate holds (for some \(p > 1\)): ([22, 36, 40])

\[
\mathbb{E}\left\{ \sup_{t \in [0,T]} |X(t)|^p + \sup_{t \in [0,T]} |Y(t)|^p + \left( \int_0^T |Z(t)|^2 dt \right)^{\frac{p}{2}} \right\} \leq K \left( 1 + |x|^p \right), \quad \forall x \in \mathbb{R}^n.
\]

(1.5)

Hereafter, \(K > 0\) will stand for a generic constant which can be different from line to line. We should point out that in general, \(Z(\cdot)\) only belongs to the following space:

\[
L^p_{\mathbb{F}}(\Omega; L^2(0, T; \mathbb{R}^{m \times d})) = \left\{ Z : [0, T] \times \Omega \to \mathbb{R}^{m \times d} \mid t \mapsto Z(t) \text{ is } \mathbb{F}\text{-adapted}, \mathbb{E}\left( \int_0^T |Z(t)|^2 dt \right)^{\frac{p}{2}} < \infty \right\}.
\]

Therefore, \(t \mapsto Z(t)\) is not necessarily continuous and for a given \(t \in [0, T]\), \(Z(t)\) might not be well-defined. This leads to some difficulties in numerical aspects for adapted solutions to BSDEs, and also restricts the usage of BSDEs/FBSDEs in applications. On the other hand, the so-called \textit{Four Step Scheme} ([19, 22], see also [13, 23, 20]) for FBSDEs gives a representation of the adapted solution \((Y(\cdot), Z(\cdot))\) of the BSDE in (1.4) via a solution to a relevant partial differential equation (PDE, for short), together with the solution \(X(\cdot)\) to the (forward) stochastic differential equation (FSDE, for short) in (1.4). Such kind of representation can help people to overcome the difficulties encountered in designing numerical algorithms for BSDEs/FBSDEs [43]. This also substantially broadens the applicability of BSDEs/FBSDEs in solving real problems.

Since the representation of adapted solutions to BSDEs/FBSDEs is very closely related to the main results in the current paper, and the techniques/ideas will be used below, we now elaborate the procedure here (See [19, 22] for more details).
Consider FBSDE (1.4). Suppose all the involved functions \(b, \sigma, g, h\) are deterministic. Inspired by the so-call \textit{invariant embedding} ([2, 3]), one could expect that there is a relation between the backward component \(Y(\cdot)\) and the forward component \(X(\cdot)\) as follows:

\[
Y(t) = \Theta(t, X(t)), \quad t \in [0, T],
\]

for some differentiable function \(\Theta : [0, T] \times \mathbb{R}^n \to \mathbb{R}^m\). If this is the case, then by Itô’s formula, one must have

\[
\begin{align*}
&-g(t, X(t), Y(t), Z(t))dt + Z(t)dW(t) = dY(t) = \left[\Theta_t(t, X(t)) + \Theta_x(t, X(t))b(t, X(t), Y(t), Z(t))
\right. \\
&\quad \left. + \frac{1}{2} \sigma(t, X(t), Y(t), Z(t))^\top \Theta_{xx}(t, X(t))\sigma(t, X(t), Y(t), Z(t))\right]dt + \Theta_x(t, X(t))\sigma(t, X(t), Y(t), Z(t))dW(t),
\end{align*}
\]

where

\[
\sigma(t, x, y, z)^\top \Theta_{xx}(t, x)\sigma(t, x, y, z) = \sum_{k=1}^{d} \left(\sigma_k(t, x, y, z)^\top \Theta_{1x}^k(t, x)\sigma_k(t, x, y, z)\right)
\]

with

\[
\sigma(t, x, y, z) = (\sigma_1(t, x, y, z), \sigma_2(t, x, y, z), \ldots, \sigma_d(t, x, y, z)), \quad \Theta(t, x) = \begin{pmatrix} \Theta^1(t, x) \\
\Theta^2(t, x) \\
\vdots \\
\Theta^m(t, x) \end{pmatrix}.
\]

Therefore, the following should hold:

\[
Z(t) = \Theta_x(t, X(t))\sigma(t, X(t), Y(t), Z(t)), \quad t \in [0, T].
\]

Suppose there exists a well-defined map \(\zeta : [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^{m \times d}\) satisfying

\[
\zeta(t, x, y) = \Theta_x(t, x)\sigma(t, x, y, \zeta(t, x, y)), \quad (t, x, y) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^m,
\]

which is the case, trivially, if \(\sigma(t, x, y, z) \equiv \sigma(t, x, y)\) (independent of \(z\)). Then \(\Theta(\cdot, \cdot)\) should solve the following system of quasi-linear parabolic PDE system:

\[
\begin{align*}
\left\{ \\
\Theta_t(t, x) + &\frac{1}{2} \sigma(t, x, \Theta(t, x), \zeta(t, x, \Theta(t, x)))^\top \Theta_{xx}(t, x)\sigma(t, x, \Theta(t, x), \zeta(t, x, \Theta(t, x))) \\
&+ \Theta_x(t, x)b(t, x, \Theta(t, x), \zeta(t, x, \Theta(t, x))) + g(t, x, \Theta(t, x), \zeta(t, x, \Theta(t, x))) = 0,
\end{align*}
\]

\[
\Theta(T, x) = h(x), \quad x \in \mathbb{R}^n.
\]

Suppose the above PDE has a classical solution \(\Theta(\cdot, \cdot)\). Then we solve the following FSDE:

\[
\begin{align*}
dX(t) &= b(t, X(t), \Theta(t, X(t)), \zeta(t, X(t), \Theta(t, X(t))))dt \\
&\quad + \sigma(t, X(t), \Theta(t, X(t)), \zeta(t, X(t), \Theta(t, X(t))))dW(t), \quad t \in [0, T], \\
X(0) &= x.
\end{align*}
\]

Now, if the above FSDE admits a solution \(X(\cdot)\), then we have the representation of the backward components \((Y(\cdot), Z(\cdot))\) in terms of the forward component \(X(\cdot)\):

\[
Y(t) = \Theta(t, X(t)), \quad Z(t) = \zeta(t, X(t), \Theta(t, X(t))), \quad t \in [0, T].
\]
In the above, \( \Theta(\cdot, \cdot) \) is called a \textit{decoupling field} of the FBSDE (1.4) ([20]), and (1.7) is called the \textit{representation PDE} since the solution \( \Theta(\cdot, \cdot) \) allows us to represent the backward component \( (Y(\cdot), Z(\cdot)) \) in terms of the forward component \( X(\cdot) \). From the above, we see that as long as all the involved functions are nice enough (in a suitable sense), the above representation (1.9) provides useful regularity information on \( (Y(\cdot), Z(\cdot)) \), especially for \( Z(\cdot) \). This actually has played some interesting roles in numerical aspects of BSDEs/FBSDEs ([8, 43]).

Note that in the case that both \( b \) and \( \sigma \) are independent of \( (Y(\cdot), Z(\cdot)) \), for which the FBSDE is decoupled, the representation PDE becomes

\[
\begin{aligned}
& \begin{cases}
\Theta_t(t,x) + \frac{1}{2} \sigma(t,x) \Theta_{xx}(t,x) \sigma(t,x) + \Theta_x(t,x)b(t,x) + g(t,x,\Theta(t,x), \Theta_x(t,x)\sigma(t,x)) = 0, \\
\Theta(T,x) = h(x),
\end{cases} \\
& (t,x) \in [0,T] \times \mathbb{R}^n,
\end{aligned}
\]  

(1.10)

whose solvability conditions are much simpler than those for (1.7). In this case, (1.9) becomes

\[
Y(t) = \Theta(t, X(t)), \quad Z(t) = \Theta_x(t, X(t))\sigma(t, X(t)), \quad t \in [0,T],
\]

(1.11)

with \( X(\cdot) \) being the solution of FSDE:

\[
\begin{aligned}
& \begin{cases}
\, dX(t) = b(t, X(t))dt + \sigma(t, X(t))dW(t), \\
\, X(0) = x.
\end{cases} \\
& t \in [0,T],
\end{aligned}
\]

(1.12)

Let us return to FBSDE (1.4). For any \((s,x) \in [0,T) \times \mathbb{R}^n\), let \((X(\cdot, s,x), Y(\cdot, s,x), Z(\cdot, s,x))\) be the (unique) adapted solution to (1.4) on \([s,T]\) with \(X(0) = x\) replaced by \(X(s) = x\). Then

\[
\Theta(s,x) = Y(s; s,x), \quad (s,x) \in [0,T) \times \mathbb{R}^n.
\]

(1.13)

Thus, the solution \( \Theta(\cdot, \cdot) \) to the PDE (1.7) admits a representation \( Y(\cdot, \cdot, \cdot), \) a part of the adapted solution to FBSDE (1.4). This is called a \textit{nonlinear Feynman-Kac formula} (see [27, 24]).

We now consider BSVIEs. In 2002, Lin firstly introduced a BSVIE ([18]) as an extension of BSDEs, in which the term \( Z(s,t) \) did not appear and \( \psi(t) \equiv \xi \) is a fixed \( \mathcal{F}_T \)-measurable random variable. The form (1.1), including the term \( Z(s,t) \) with general \( \psi(\cdot) \), was firstly introduced by the second author of the current paper in 2006 ([37]), motivated by optimal control of (forward) stochastic Volterra integral equations (FSVIEs, for short). When \( Z(s,t) \) is absent, the BSVIE (1.1) becomes:

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

(1.14)

Hereafter, we call (1.14) and (1.1) Type-I and Type-II BSVIEs, respectively. Thus, Type-I BSVIE is a special case of Type-II BSVIE.

Mimicking the case of BSDEs, a pair \((Y(\cdot), Z(\cdot, \cdot))\) is called an \textit{adapted solution} to BSVIE (1.1) if for each \( t \in [0,T) \), the map \( s \mapsto (Y(s), Z(t,s)) \) is \( \mathcal{F}_t \)-adapted on \([t,T]\), and satisfies equation (1.1) in the usual Itô sense. For Type-I BSVIE (1.14), one needs only to determine \( Z(t,s) \) for \((t,s) \in \Delta[0,T]\), where

\[
\Delta[0,T] = \left\{(t,s) \in [0,T]^2 \mid 0 \leq t \leq s \leq T\right\}.
\]

(1.15)

Therefore, under proper conditions, a Type-I BSVIE admits a unique adapted solution. However, for a Type-II BSVIE, due to the appearance of \( Z(s,t) \) in the equation, we need to determine \( Z(t,s) \) for \((t,s) \in [0,T]^2\),
and (1.1) alone does not give enough restrictions on $Z(t, s)$. Consequently, as pointed out in [39], the adapted solution to Type-II BSVIE (1.1) is not unique. Inspired by the duality principle needed in the optimal control of FSVIEs, the so-called *adapted M-solution* was introduced in [39]: A pair $(Y(\cdot), Z(\cdot, \cdot))$ is called an adapted M-solution to (1.1) if it is an adapted solution and moreover, the following holds:

$$Y(t) = \mathbb{E}[Y(t)] + \int_0^t Z(t, s)dW(s), \quad t \in [0, T], \text{ a.s.}$$  \hspace{1cm} (1.16)

It was proved in [39] that under certain conditions, Type-II BSVIE (1.1) admits a unique adapted M-solution. Moreover, the following estimate holds:

$$\mathbb{E}\left(\int_0^T |Y(t)|^2 ds + \int_0^T \int_0^T |Z(t, s)|^2 ds dt\right) \leq K\left(\int_0^T \int_s^T |g(t, s, 0, 0, 0, 0)|^2 ds dt + \mathbb{E}\int_0^T |\psi(s)|^2 ds\right).$$  \hspace{1cm} (1.17)

For some relevant results on BSVIEs, the readers are further referred to [31, 32, 33].

From [39], we see that to get some further regularities beyond the above estimate (1.17) for the process $(Y(\cdot), Z(\cdot, \cdot))$, many technical conditions have to be imposed, the proofs are quite technical, and unfortunately, the regularity results were still not satisfactory, especially that of the process $(t, s) \mapsto Z(t, s)$.

Inspired by the decoupling FBSDEs presented above, we naturally ask: Is it possible to get representation of adapted solutions for Type-I BSVIEs and adapted M-solutions for Type-II BSVIEs similar to (1.11) for BSDEs? More precisely, we will consider the following BSVIEs:

$$Y(t) = \psi(t, X(t), X(T)) + \int_t^T g(t, s, X(t), X(s), Y(s), Z(t, s))ds - \int_t^T Z(t, s)dW(s), \quad t \in [0, T],$$  \hspace{1cm} (1.18)

and

$$Y(t) = \psi(t, X(t), X(T)) + \int_t^T g(t, s, X(t), X(s), Y(s), Z(t, s), Z(s, t))ds - \int_t^T Z(t, s)dW(s), \quad t \in [0, T],$$  \hspace{1cm} (1.19)

with $X(\cdot)$ being the solution to the FSDE (1.12), and $\psi, g$ being some deterministic maps. Note that (1.18) and (1.19) are respectively Type-I and Type-II BSVIEs with random coefficients, for which the randomness all comes from the solution $X(\cdot)$ of FSDE (1.12). Our goal is to establish a representation of $(Y(\cdot), Z(\cdot, \cdot))$ in terms of $X(\cdot)$, via the solution to a suitable representation PDE system. More precisely, we will establish the following result.

**Theorem 1.1.** Let some suitable conditions hold.

(i) For Type-I BSVIE (1.18), the following representation holds:

$$\left\{ \begin{array}{l}
Y(s) = \Theta(s, s, X(s), X(s)), \quad s \in [0, T], \\
Z(t, s) = \Theta_2(t, s, X(t), X(s))\sigma(s, X(s)), \quad (t, s) \in \Delta[0, T],
\end{array} \right.$$  \hspace{1cm} (1.20)

with $\Theta(\cdot, \cdot, \cdot, \cdot)$ being the solution to the following PDE system:

$$\left\{ \begin{array}{l}
\Theta_4(t, s, \xi, x) + \frac{1}{2} \sigma(s, x)^T \Theta_{xx}(t, s, \xi, x)\sigma(s, x) + \Theta_4(t, s, \xi, x)b(s, x) \\
+ g(t, s, \xi, x, \Theta(s, s, x, x), \Theta_2(t, s, \xi, x)\sigma(s, x)) = 0, \quad (t, s, \xi, x) \in \Delta[0, T] \times \mathbb{R}^n \times \mathbb{R}^n,
\end{array} \right.$$  \hspace{1cm} (1.21)

$$\Theta(t, T, \xi, x) = \psi(t, \xi, x), \quad (t, \xi, x) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^n.$$
(ii) For Type-II BSVIE (1.19), the following representation holds:

\[
\begin{cases}
Y(s) = \Theta(s, s, X(s), X(s)), & s \in [0, T], \\
Z(t, s) = \Theta_x(t, s, X(t), X(s))\sigma(s, X(s)), & 0 \leq t \leq s \leq T, \\
Z(t, s) = \Gamma_x(t, s, X(s))\sigma(s, X(s)), & 0 \leq s \leq t \leq T,
\end{cases}
\]

with \((\Gamma, \Theta)\) being the solution to the following PDE system:

\[
\begin{cases}
\Gamma_x(t, s, x) + \frac{1}{2} \sigma(s, x)^T\Gamma_{xx}(t, s, x)\sigma(s, x) + \Gamma_x(t, s, x)b(s, x) = 0, & 0 \leq s \leq t \leq T, \ x \in \mathbb{R}^n, \\
\Theta_x(t, s, \xi, x) + \frac{1}{2} \sigma(s, x)^T\Theta_{xx}(t, s, \xi, x)\sigma(s, x) + \Theta_x(t, s, \xi, x)b(s, x) \\
\hspace{1cm}+ g(t, s, \xi, x, \Theta(s, s, x, x), \Theta_x(t, s, \xi, x)\sigma(s, x), \Gamma_x(t, s, \xi)\sigma(s, x)) = 0, & 0 \leq t \leq s \leq T, \ s, \xi \in \mathbb{R}^n, \\
\Gamma(t, t, x) = \Theta(t, t, x), & (t, x) \in [0, T] \times \mathbb{R}^n, \\
\Theta(t, T, \xi, x) = \psi(t, \xi, x), & (t, \xi, x) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^n.
\end{cases}
\]

The idea of obtaining the representation is to find a proper approximation of the BSVIE by BSDEs and then derive the correct form of the representation PDE system by the invariant embedding/decoupling technique. Once the correct form of PDE system is obtained, a standard application of Itô’s formula will lead to our representation. Partial results for Type-II BSVIEs of this paper was announced in [42] without detailed proofs.

The rest of the paper is organized as follows. In Section 2, approximation of Type-I BSVIEs by means of BSDEs is established. Section 3 is devoted to the derivation of representation for the adapted solutions of Type-I BSVIEs. In Section 4, we establish a representation for the adapted M-solutions of Type-II BSVIEs. The well-posedness of representation PDEs is established in Section 5. Some concluding remarks are collected in Section 6.

2 Approximation of Type-I BSVIEs

This section is devoted to an approximation of Type-I BSVIEs by a sequence of BSDEs. On one hand, such an approximation will be helpful for us to derive the representation of the adapted solutions to the Type-I BSVIEs. On the other hand, this will also be helpful for designing numerical scheme for such kind of BSVIEs ([34]). Before going further, let us first introduce the following assumption concerning the FSDE (1.12).

(H1) The maps \(b, \sigma : [0, T] \times \mathbb{R}^n \to \mathbb{R}^n\) are continuous such that for some constant \(L > 0\), it holds

\[|b(t, x) - b(t, x')| + |\sigma(t, x) - \sigma(t, x')| \leq L|x - x'|, \quad \forall t \in [0, T], \ x, x' \in \mathbb{R}^n.\]

It is standard that for any fixed \(x \in \mathbb{R}^n\), FSDE (1.12) admits a unique strong solution \(X(\cdot) \equiv X(\cdot; t, x)\), and the following holds:

\[\mathbb{E} |X(s) - X(t)|^2 \leq K_0|s - t|, \quad \forall s, t \in [0, T], \]

(2.1)

for some constant \(K_0 > 0\) independent of \(s\) and \(t\). Now, for such a given \(X(\cdot)\), we consider the following Type-I BSVIE:

\[Y(t) = \psi(t, X(t), X(T)) + \int_t^T g(t, s, X(t), X(s), Y(s), Z(t, s))ds - \int_t^T Z(t, s)dW(s), \quad t \in [0, T].\]

(2.2)
We introduce the following assumption, recalling the definition of $\Delta[0,T]$ (see (1.15)).

(H2) Functions $\psi: [0, T] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ and $g: \Delta[0,T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \times d \to \mathbb{R}^m$ are continuous. There exists a constant $L > 0$ such that

$$|\psi(t, \xi, x) - \psi(t', \xi', x')| \leq L(|t - t'| + |\xi - \xi'| + |x - x'|), \quad \forall (t, \xi, x), (t', \xi', x') \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^n,$$

$$|g(t, s, \xi, x, y) - g(t', s, \xi', x', y')| \leq L(|t - t'| + |\xi - \xi'| + |x - x'| + |y - y'| + |z - z'|),$$

$$\forall (t, s, \xi, x, y), (t', s, \xi', x', y', z') \in \Delta[0,T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m \times d.$$  

Under (H2), for the given $X(\cdot) = X(\cdot; 0, x)$, BSVIE (2.2) admits a unique adapted solution $(Y(\cdot), Z(\cdot, \cdot))$ on $[0, T]$ (see [39], for example). Let $\mathcal{P}[0, T]$ be the set of all partitions $\Pi$ of $[0, T]$ having the following form:

$$\Pi : \quad 0 = t_0 < t_1 < t_2 < \cdots < t_{N-1} < t_N = T,$$  

with some natural number $N > 1$. We define the mesh size $\|\Pi\|$ of $\Pi$ by the following:

$$\|\Pi\| = \max_{1 \leq i \leq N} (t_i - t_{i-1}).$$

For a partition $\Pi$ as above, let us make an observation. Keep in mind that when we discuss Type-I BSVIE (2.2), the process $X(\cdot)$ is given. Suppose $Y(s)$ and $Z(t, s)$ have been determined for $t_{k+1} \leq t \leq s \leq T$ (see the region marked ① in the figure below). Then for $t \in [t_k, t_{k+1})$, one has

$$Y(t) = \psi(t, X(t), X(T)) + \int_t^T g(t, s, X(t), X(s), Y(s), Z(t, s))ds - \int_t^T Z(t, s)dW(s)$$

$$= \psi(t, X(t), X(T)) + \int_{t_{k+1}}^T g(t, s, X(t), X(s), Y(s), Z(t, s))ds$$

$$+ \int_t^{t_{k+1}} g(t, s, X(t), X(s), Y(s), Z(t, s))ds - \int_t^T Z(t, s)dW(s).$$

In the above, we see that $Y(s)$, $s \in [t_{k+1}, T]$, is known. But, at the moment, $Z(t, s)$ has been determined only for $t_{k+1} \leq t \leq s \leq T$, and it is still unknown for $(t, s) \in [t_k, t_{k+1}) \times [t_{k+1}, T]$ (see the region marked ② in the figure below). Also, both $Y(s)$ and $Z(t, s)$ are unknown for $t_k \leq t \leq s \leq t_{k+1}$ (the region marked ③ in the figure below). Therefore, we want to find $(Y(s), Z(t, s))$ for $t_k \leq t \leq t_{k+1}$ and $t \leq s \leq T$ (the regions marked ② and ③).

(Figure 1)
We now construct an approximation of BSVIE (2.2). On \([t_{N-1}, T]\), we introduce the following BSDE:

\[
\begin{aligned}
&dY_{N-1}(s) = -g(t_{N-1}, s, X(t_{N-1}), X(s), Y_{N-1}(s), Z_{N-1}(s))ds + Z_{N-1}(s)dW(s), \\
&Y_{N-1}(T) = \psi(t_{N-1}, X(t_{N-1}), X(T)).
\end{aligned}
\] (2.4)

Under (H2), the above BSDE admits a unique adapted solution \((Y_{N-1}(\cdot), Z_{N-1}(\cdot))\). Next, on \([t_{N-2}, T]\), (not just on \([t_{N-2}, t_{N-1}]\)), we introduce the following BSDE:

\[
\begin{aligned}
&dY_{N-2}(s) = -g(t_{N-2}, s, X(t_{N-2}), X(s), Y_{N-1}(s), Z_{N-2}(s))ds + Z_{N-2}(s)dW(s), \\
&dY_{N-2}(s) = -g(t_{N-2}, s, X(t_{N-2}), X(s), Y_{N-2}(s), Z_{N-2}(s))ds + Z_{N-2}(s)dW(s), \\
&Y_{N-2}(T) = \psi(t_{N-2}, X(t_{N-2}), X(T)), \quad Y_{N-2}(t_{N-1}) = Y_{N-2}(t_{N-1} + 0).
\end{aligned}
\] (2.5)

Note that in the above, \(Y_{N-1}(s)\) is already determined by (2.4) on \([t_{N-1}, T]\), which has to stay unchanged. But, since \(t_{N-2}\) appears in \(g\) and \(\psi\), \((Y_{N-2}(\cdot), Z_{N-2}(\cdot))\) and \((Y_{N-1}(\cdot), Z_{N-1}(\cdot))\) satisfy different BSDEs on \([t_{N-1}, T]\). Consequently, \(Z_{N-2}(\cdot)\) is possibly different from \(Z_{N-1}(\cdot)\) on \([t_{N-1}, T]\), and therefore, \(Y_{N-2}(\cdot)\) could be different from \(Y_{N-1}(\cdot)\) on \([t_{N-1}, T]\). Under our conditions, the above BSDE admits a unique adapted solution on \([t_{N-2}, T]\).

In general, on \([t_k, T]\), we have a unique adapted solution \((Y^k(\cdot), Z^k(\cdot))\) to the following BSDE:

\[
\begin{aligned}
&dY^k(s) = -g(t_k, s, X(t_k), X(s), Y^\ell(s), Z^k(s))ds + Z^k(s)dW(s), \\
&dY^k(s) = -g(t_k, s, X(t_k), X(s), Y^k(s), Z^k(s))ds + Z^k(s)dW(s), \\
&Y^k(T) = \psi(t_k, X(t_k), X(T)), \quad Y^k(t_{\ell}) = Y^k(t_{\ell} + 0), \quad k + 1 \leq \ell \leq N - 1.
\end{aligned}
\] (2.6)

Define

\[
\begin{aligned}
&Y^\Pi(s) = \sum_{k=0}^{N-2} Y^k(s)I_{[t_k,t_{k+1}]}(s) + Y^N(s)I_{[t_{N-1}, T]}(s), \quad s \in [0, T], \\
&Z^\Pi(t, s) = \sum_{k=0}^{N-2} Z^k(s)I_{[t_k,t_{k+1}]}(t) + Z^\Pi(s)I_{[t_{N-1}, T]}(t), \quad 0 \leq t \leq s \leq T,
\end{aligned}
\]

The above defined \(Y^\Pi(\cdot)\) possibly has discontinuity at \(t_{N-1}, t_{N-2}, \ldots, t_1\). With the above definition, we may rewrite (2.6) as

\[
\begin{aligned}
&dY^k(s) = -g(t_k, s, X(t_k), X(s), Y^\Pi(s), Z^k(s))ds + Z^k(s)dW(s), \quad s \in [t_k, T], \\
&Y^k(T) = \psi(t_k, X(t_k), X(T)), \quad Y^k(t_{\ell}) = Y^k(t_{\ell} + 0), \quad k + 1 \leq \ell \leq N - 1.
\end{aligned}
\] (2.7)

Keep in mind that for each \(k = N, N - 1, N - 2, \ldots, 2, 1\), process \(Y^k(\cdot)\) is continuous, although \(Y^\Pi(\cdot)\) is not necessarily continuous. Now, we introduce

\[
\tau^\Pi(t) = \sum_{k=0}^{N-2} t_kI_{[t_k,t_{k+1}]}(t) + t_{N-1}I_{[t_{N-1}, T]}(t), \quad t \in [0, T].
\] (2.8)

Then

\[
0 \leq t - \tau^\Pi(t) \leq ||\Pi||, \quad t \in [0, T].
\] (2.9)
Hence, for any \( t \in [0, T) \), let \( t \in [t_k, t_{k+1}) \), we have the following:

\[
Y^{\Pi}(t) = Y^k(t) = \psi(t_k, X(t_k), X(T)) + \int_t^{t_{k+1}} g(t_k, s, X(t_k), X(s), Y^k(s), Z^k(s)) ds \\
+ \sum_{t=t_k+1}^{t_{k+1}} g(t_k, s, X(t_k), X(s), Y^k(s), Z^k(s)) ds - \int_t^T Z^k(s) dW(s)
\]

\[
= \psi(\tau^{\Pi}(t), X(\tau^{\Pi}(t)), X(T)) + \int_t^T g(\tau^{\Pi}(t), s, X(\tau^{\Pi}(t)), X(s), Y^{\Pi}(s), Z^{\Pi}(t, s)) ds - \int_t^T Z^{\Pi}(t, s) dW(s).
\]

Therefore, \((Y^{\Pi}(\cdot), Z^{\Pi}(\cdot, \cdot))\) satisfies a BSDE on each \((t_k, t_{k+1})\), and satisfies a BSVIE on \([0, T]\). Note that since \( t \mapsto g(\tau^{\Pi}(t), s, X(t), x, y, z)\) has possible jumps at \( t_k \), the resulting \( t \mapsto Y^{\Pi}(t)\) may also have jumps at \( t_k \), regardless of its integral form. For the above constructed \((Y^{\Pi}(\cdot), Z^{\Pi}(\cdot, \cdot))\), we have the following proposition.

**Proposition 2.1.** Suppose \((H1)–(H2)\) hold, \( X(\cdot) \) is the solution to the FSDE (1.12), and \((Y^{\Pi}(\cdot), Z^{\Pi}(\cdot, \cdot))\) is constructed as above. Then

\[
\mathbb{E} \int_0^T |Y^{\Pi}(t) - Y(t)|^2 dt + \mathbb{E} \int_0^T \int_t^T |Z^{\Pi}(t, s) - Z(t, s)|^2 ds dt \leq K \|\Pi\|^2, \tag{2.10}
\]

and

\[
\mathbb{E} \left[ \sup_{s \in [t_{k+1}, T]} |Y^{k+1}(s) - Y^k(s)|^2 \right] + \mathbb{E} \left( \int_{t_{k+1}}^T |Z^{k+1}(s) - Z^k(s)|^2 ds \right) \leq K \|\Pi\|, \quad 0 \leq k \leq N - 2. \tag{2.11}
\]

In particular,

\[
\mathbb{E} |Y^{\Pi}(t_k) - Y^{\Pi}(t_k - 0)| \leq K \|\Pi\|, \quad 1 \leq k \leq N - 1. \tag{2.12}
\]

**Proof.** By the stability of adapted solutions to BSVIEs ([39]), we have (note (2.9))

\[
\mathbb{E} \int_0^T |Y^{\Pi}(t) - Y(t)|^2 dt + \mathbb{E} \int_0^T \int_t^T |Z^{\Pi}(t, s) - Z(t, s)|^2 ds dt \\
\leq K \mathbb{E} \int_0^T |\psi(\tau^{\Pi}(t), X(t), X(T)) - \psi(t, X(t), X(T))|^2 dt \\
+ \mathbb{E} \int_0^T \int_t^T |g(\tau^{\Pi}(t), s, X(t), X(s), Y(s), Z(t, s)) - g(t, s, X(t), X(s), Y(s), Z(t, s))|^2 ds dt \\
\leq K \mathbb{E} \left\{ \int_0^T \|\Pi\|^2 dt + \int_0^T \|\Pi\|^2 ds \right\} \leq K \|\Pi\|^2.
\]

This proves (2.10). Next, one has

\[
\begin{cases}
    d \left[ Y^{k+1}(s) - Y^k(s) \right] = - \left[ g(t_{k+1}, s, X(t_{k+1}), X(s), Y^{k+1}(s), Z^{k+1}(s)) \\
    - g(t_k, s, X(t_k), X(s), Y^k(s), Z^k(s)) \right] ds + \left[ Z^{k+1}(s) - Z^k(s) \right] dW(s), \\
    s \in [t_{k+1}, T],
\end{cases}
\]

\[
\begin{cases}
    Y^{k+1}(T) - Y^k(T) = \psi(t_{k+1}, X(t_{k+1}), X(T)) - \psi(t_k, X(t_k), X(T)).
\end{cases}
\]

9
Hence,

\[
\mathbb{E}\left[ \sup_{s \in [t_{k+1}, T]} |Y^{k+1}(s) - Y^k(s)|^2 \right] + \mathbb{E}\left( \int_{t_{k+1}}^{T} |Z^{k+1}(s) - Z^k(s)|^2 ds \right) \\
\leq K \mathbb{E}\left\{ |\psi(t_{k+1}, X(t_{k+1})), X(T)) - \psi(t_k, X(t_k), X(T))|^2 \\
+ \left( \int_{t_{k+1}}^{T} |g(t_{k+1}, s, X(t_{k+1}), X(s), Y^k(s), Z^k(s)) - g(t_k, s, X(t_k), X(s), Y^k(s), Z^k(s))|ds \right)^2 \right\} \\
\leq K \mathbb{E}\left\{ (t_{k+1} - t_k)^2 + |X(t_{k+1}) - X(t_k)|^2 + \int_{t_{k+1}}^{T} \left( |t_{k+1} - t_k| + |X(t_{k+1}) - X(t_k)| \right) ds \right)^2 \leq K \|\Pi\|.
\]

This leads to our conclusion. \(\square\)

3 Representation of Adapted Solutions for Type-I BSVIEs

In this section, we will represent the adapted solution \((Y(\cdot), Z(\cdot, \cdot))\) of Type-I BSVIE (2.2) in terms of \(X(\cdot)\), the solution to FSDE (1.12), and the solution to the corresponding representation PDE.

Let \(X(\cdot)\) be the solution of (1.12) and \(\Pi \in \mathcal{P}[0, T] \) be of form (2.3). Let \((Y^k(\cdot), Z^k(\cdot))\) \((0 \leq k \leq N - 1)\) and \((Y^{\Pi}(\cdot), Z^{\Pi}(\cdot, \cdot))\) be constructed as in Section 2.

If \((Y^{\Pi}(\cdot), Z^{\Pi}(\cdot, \cdot))\) were represented by \(X(\cdot)\), together with the solution to certain PDE, then by sending \(\|\Pi\| \to 0\), we would get what we want. However, there are some difficulties in doing that directly (see below for some explanations). Therefore, instead, we construct a sequence of processes \((Y^{\Pi}(\cdot), Z^{\Pi}(\cdot, \cdot))\) which is close to \((Y^{\Pi}(\cdot), Z^{\Pi}(\cdot, \cdot))\) and which can be represented by \(X(\cdot)\), together with the solution of a certain PDE. Then by sending \(\|\Pi\| \to 0\), we will obtain the desired representation of \((Y(\cdot), Z(\cdot, \cdot))\).

Now, we carefully make this precise. First, let \((\bar{Y}^{N-1}(\cdot), \bar{Z}^{N-1}(\cdot))\) be the adapted solution to the following BSDE:

\[
\begin{cases}
    d\bar{Y}^{N-1}(s) = -g(t_{N-1}, s, X(t_{N-1}), X(s), \bar{Y}^{N-1}(s), \bar{Z}^{N-1}(s))ds + \bar{Z}^{N-1}(s)dW(s), & s \in [t_{N-1}, T], \\
    \bar{Y}^{N-1}(T) = \psi(t_{N-1}, X(t_{N-1}), X(T)),
\end{cases}
\]

which coincides with BSDE (2.4). Thus, one has

\[
(\bar{Y}^{N-1}(s), \bar{Z}^{N-1}(s)) = (Y^{N-1}(s), Z^{N-1}(s)), \quad t_{N-1} \leq t \leq s \leq T.
\]

Although in (3.1), the map \((s, x, y, z) \mapsto (g(t_{N-1}, s, X(t_{N-1}), x, y, z), \psi(t_{N-1}, X(t_{N-1}), x))\) is merely \(\mathcal{F}_{t_{N-1}}\)-measurable, not necessarily deterministic, since we are considering the BSDE on \([t_{N-1}, T]\), the (decoupling) technique introduced in [19, 22] will still work. In fact, we have the following representation:

\[
\begin{cases}
    \bar{Y}^{N-1}(s) = Y^{N-1}(s) = \Theta^{N-1}(s, X(t_{N-1}), X(s)), & s \in [t_{N-1}, T], \\
    \bar{Z}^{N-1}(s) = Z^{N-1}(s) = \Theta^{N-1}_x(s, X(t_{N-1}), X(s))\sigma(s, X(s)),
\end{cases}
\]

with \((s, x) \mapsto \Theta^{N-1}(s, x, x)\) being the solution to the following representation PDE:

\[
\begin{cases}
    \Theta^{N-1}(s, x) + \frac{1}{2} \sigma(s, x)^T \Theta^{N-1}_x(s, x)\sigma(s, x) + \Theta^{N-1}_x(s, x)b(s, x) \\
    + g(t_{N-1}, s, x, \Theta^{N-1}(s, x, x), \Theta^{N-1}_x(s, x)\sigma(s, x)) = 0, & (s, x) \in [t_{N-1}, T] \times \mathbb{R}^n, \\
    \Theta^{N-1}(T, x) = \psi(t_{N-1}, x), & x \in \mathbb{R}^n.
\end{cases}
\]

10
In the above, $\xi \in \mathbb{R}^n$ is treated as a parameter. With the representation (3.3), we can rewrite (3.1) as follows:

$$
\begin{align*}
\text{(3.5)} & \quad d\bar{Y}^{N-1}(s) = -g(t_{N-1}, s, X(t_{N-1}), X(s), \Theta^{N-1}(s, X(t_{N-1}), X(s)), \bar{Z}^{N-1}(s))ds + \bar{Z}^{N-1}(s)dW(s), \quad s \in [t_{N-1}, T], \\
\bar{Y}^{N-1}(T) & = \psi(t_{N-1}, X(t_{N-1}), X(T)).
\end{align*}
$$

Next, we construct $(\bar{Y}^{N-2}(\cdot), \bar{Z}^{N-2}(\cdot))$ on $[t_{N-2}, T]$. On $[t_{N-1}, T]$, we let $(\bar{Y}^{N-2}(\cdot), \bar{Z}^{N-2}(\cdot))$ be the adapted solution to the following BSDE:

$$
\begin{align*}
\text{(3.6)} & \quad d\bar{Y}^{N-2}(s) = -g(t_{N-2}, s, X(t_{N-2}), X(s), \Theta^{N-1}(s, X(t_{N-1}), X(s)), \bar{Z}^{N-2}(s))ds + \bar{Z}^{N-2}(s)dW(s), \quad s \in [t_{N-1}, T], \\
\bar{Y}^{N-2}(T) & = \psi(t_{N-2}, X(t_{N-2}), X(T)).
\end{align*}
$$

Note that on $[t_{N-1}, T]$, we have (3.2) and representation (3.3). Hence, by (2.5), we see that $(\bar{Y}^{N-2}(\cdot), \bar{Z}^{N-2}(\cdot))$ satisfies the following BSDE:

$$
\begin{align*}
\text{(3.7)} & \quad d\bar{Y}^{N-2}(s) = -g(t_{N-2}, s, X(t_{N-2}), X(s), \Theta^{N-1}(s, X(t_{N-1}), X(s)), \bar{Z}^{N-2}(s))ds + \bar{Z}^{N-2}(s)dW(s), \quad s \in [t_{N-1}, T], \\
\bar{Y}^{N-2}(T) & = \psi(t_{N-2}, X(t_{N-2}), X(T)).
\end{align*}
$$

Let us make two comparisons. First, (3.5) and (3.7) are different: $(t_{N-1}, X(t_{N-1}))$ in the former is replaced by $(t_{N-2}, X(t_{N-2}))$ in the latter at two places. Second, (3.7) and (3.6) are different: $\Theta^{N-1}(s, X(t_{N-1}), X(s))$ in the former is replaced by $\Theta^{N-1}(s, X(s), X(s))$ in the latter. Note that in (3.7), both $X(t_{N-1})$ and $X(t_{N-2})$ appear. This will cause some difficulties in passing to the limit as $\|\Pi\| \to 0$ later on. This is exactly the difficulty that we will encounter if we use $(Y^{\Pi}(\cdot), Z^{\Pi}(\cdot))$ directly trying to get our representation. On the other hand, since $\|\Pi\|$ will be small, $X(t_{N-1})$ and $X(s)$ will be close (in some sense), for $s \in [t_{N-1}, T]$, it should be harmless to replace $\Theta^{N-1}(s, X(t_{N-1}), X(s))$ by $\Theta^{N-1}(s, X(s), X(s))$ in the drift of the equation on $[t_{N-1}, T]$. Thus, if $\|\Theta^{N-1}_\xi\|^2_{\infty} < \infty$, by the stability of adapted solutions to BSDEs, we have

$$
\begin{align*}
\mathbb{E}\left[\sup_{s \in [t_{N-1}, T]} |\bar{Y}^{N-2}(s) - Y^{N-2}(s)|^2\right] & \leq \mathbb{E}\left[\int_{t_{N-1}}^{T} |g(t_{N-2}, s, X(t_{N-2}), X(s), \Theta^{N-1}(s, X(t_{N-1}), X(s)), \bar{Z}^{N-2}(s)) - g(t_{N-2}, s, X(t_{N-2}), X(s), \Theta^{N-1}(s, X(s), X(s)), \bar{Z}^{N-2}(s))|ds\right] \\
& \leq K_1 \mathbb{E}\left[\int_{t_{N-1}}^{T} |\Theta^{N-1}(s, X(s), X(s)) - \Theta^{N-1}(s, X(t_{N-1}), X(s))|^2ds\right] \\
& \leq K_1 L^2 \mathbb{E}\left[\int_{t_{N-1}}^{T} |X(s) - X(t_{N-1})|^2ds\right] \\
& \leq K_1 L^2 \|\Theta^{N-1}_\xi\|^2_{\infty} \mathbb{E}\left[\int_{t_{N-1}}^{T} |X(s) - X(t_{N-1})|^2ds\right] \\
& \leq K_1 L^2 \|\Theta^{N-1}_\xi\|^2_{\infty} (T - t_{N-1}) \int_{t_{N-1}}^{T} |X(s) - X(t_{N-1})|^2ds \\
& \leq K_1 L^2 \|\Theta^{N-1}_\xi\|^2_{\infty} K_0 (T - t_{N-1}) \int_{t_{N-1}}^{T} (s - t_{N-1})ds \\
& \leq \frac{K_0 K_1 L^2 \|\Theta^{N-1}_\xi\|^2_{\infty}}{2} \|\Pi\|^2 (T - t_{N-1})^2.
\end{align*}
$$

In the above, $K_0$ is the constant appears in (2.1), and $K_1$ is a constant appears in the stability estimate for the adapted solution of BSDEs, which can be made independent of the partition $\Pi$, under (H2).
Similar to the previous step, for (3.6), we have the following representation:

\[
\begin{aligned}
\dot{Y}^{N-2}(s) &= \Theta^{N-2}(s, X(t_{N-2}), X(s)), \\
\dot{Z}^{N-2}(s) &= \Theta_x^{N-2}(s, X(t_{N-2}), X(s))\sigma(s, X(s)), \\
\end{aligned}
\]  
\(s \in [t_{N-1}, T],\)  
(3.9)

with \((s, x) \mapsto \Theta^{N-2}(s, \xi, x)\) being the solution to the following representation PDE:

\[
\begin{aligned}
\Theta^{N-2}(s, \xi, x) + \frac{1}{2} \sigma(s, x)^T \Theta^{N-2}_{xx}(s, \xi, x)\sigma(s, x) + \Theta^{N-2}_x(s, \xi, x)b(s, x) \\
+ g(t_{N-2}, s, \xi, x, \Theta^{N-1}_x(s, \xi, x), \Theta^{N-2}_x(s, \xi, x)\sigma(s, x)) = 0, \\
(\Theta^{N-2}(T, \xi, x) = \psi(t_{N-2}, \xi, x), \\
x \in \mathbb{R}^n. \\
\end{aligned}
\]  
(3.10)

Note that equation (3.10) is different from (3.4), not just because \(t_{N-1}\) is replaced by \(t_{N-2}\) in \(g\), but also because \(\Theta^{N-1}(s, \xi, x)\) is replaced by \(\Theta^{N-1}(s, x, x)\) in \(g\). We expect that \(\Theta^{N-2}(s, \xi, x)|_{\xi=x}\) is close to \(\Theta^{N-1}(s, \xi, x)|_{\xi=x}\), for \(s \in [t_{N-1}, T]\), when \(t_{N-1} - t_{N-2} > 0\) is small.

So far, we have constructed \((\dot{Y}^{N-2}(\cdot), \dot{Z}^{N-2}(\cdot))\) on \([t_{N-1}, T]\). To construct \((\dot{Y}^{N-2}(\cdot), \dot{Z}^{N-2}(\cdot))\) on \([t_{N-2}, t_{N-1}]\), we introduce the following BSDE on \([t_{N-2}, t_{N-1}]\):

\[
\begin{aligned}
d\dot{Y}^{N-2}(s) &= -g(t_{N-2}, s, X(t_{N-2}), X(s), \dot{Y}^{N-2}(s), \dot{Z}^{N-2}(s))ds + \dot{Z}^{N-2}(s)dW(s), \\
\dot{Y}^{N-2}(t_{N-1} - 0) &= \Theta^{N-2}(t_{N-1}, X(t_{N-2}), X(t_{N-1})). \\
\end{aligned}
\]  
(3.11)

Now, on \([t_{N-2}, t_{N-1}]\), we have the following representation:

\[
\begin{aligned}
\dot{Y}^{N-2}(s) &= \Theta^{N-2}(s, X(t_{N-2}), X(s)), \\
\dot{Z}^{N-2}(s) &= \Theta_x^{N-2}(s, X(t_{N-2}), X(s))\sigma(s, X(s)), \\
\end{aligned}
\]  
\(s \in [t_{N-2}, t_{N-1}]\),  
(3.12)

with \((s, x) \mapsto \Theta^{N-2}(s, \xi, x)\) being the solution to the following representation PDE:

\[
\begin{aligned}
\Theta^{N-2}(s, \xi, x) + \frac{1}{2} \sigma(s, x)^T \Theta^{N-2}_{xx}(s, \xi, x)\sigma(s, x) + \Theta^{N-2}_x(s, \xi, x)b(s, x) \\
+ g(t_{N-2}, s, \xi, x, \Theta^{N-1}_x(s, \xi, x), \Theta^{N-2}_x(s, \xi, x)\sigma(s, x)) = 0, \\
(\Theta^{N-2}(t_{N-1}, \xi, x) = \Theta^{N-2}(t_{N-1} + 0, \xi, x), \\
x \in \mathbb{R}^n. \\
\end{aligned}
\]  
(3.13)

Note that unlike (3.10), in the above, \(\Theta^{N-2}(s, \xi, x)\) appears instead of \(\Theta^{N-2}(s, x, x)\) in \(g\). Next, since \((\dot{Y}^{N-2}(\cdot), \dot{Z}^{N-2}(\cdot))\) and \((Y^{N-2}(\cdot), Z^{N-2}(\cdot))\) satisfy the same equation on \([t_{N-2}, t_{N-1}]\) with different terminal conditions \(Y^{N-2}(t_{N-1})\) and \(Z^{N-2}(t_{N-1})\), we must have, making use of (3.8),

\[
\begin{aligned}
\mathbb{E}\left[ \sup_{s \in [t_{N-2}, t_{N-1}]} |\dot{Y}^{N-2}(s) - Y^{N-2}(s)|^2 + \int_{t_{N-2}}^{t_{N-1}} |\dot{Z}^{N-2}(s) - Z^{N-2}(s)|^2 ds \right] \\
\leq K_0 \mathbb{E} |\dot{Y}^{N-2}(t_{N-1}) - Y^{N-2}(t_{N-1})|^2 \leq \frac{K_0^2 K_1 L^2 ||\Theta_x^{N-1}||_{\infty}^2 \Pi}{2} (T - t_{N-1})^2. \\
\end{aligned}
\]  
(3.14)

To summarize the above, we have

\[
\begin{aligned}
\dot{Y}^{N-2}(s) &= \Theta^{N-2}(s, X(t_{N-2}), X(s)), \\
\dot{Z}^{N-2}(s) &= \Theta_x^{N-2}(s, X(t_{N-2}), X(s))\sigma(s, X(s)), \\
\end{aligned}
\]  
\(s \in [t_{N-2}, T],\)  
(3.15)
with \((s, x) \mapsto \Theta^{N-2}(s, \xi, x)\) being the solution to the following:

\[
\begin{aligned}
\Theta^{N-2}_s(s, \xi, x) + \frac{1}{2} \sigma(s, x)^T \Theta^{N-2}_{xx}(s, \xi, x) \sigma(s, x) + \Theta^{N-2}_x(s, \xi, x)b(s, x) \\
+ g(t_{N-2}, s, \xi, x, \Theta^{N-1}(s, x, x), \Theta^{N-2}(s, \xi, x) \sigma(s, x)) = 0, \quad (s, x) \in [t_{N-1}, T] \times \mathbb{R}^n,
\end{aligned}
\]

(3.16)

Note that although \(s \mapsto \bar{Y}^{N-2}(s)\) could be discontinuous at \(s = t_{N-1}\), the function \(s \mapsto \Theta^{N-2}(s, \xi, x)\) is continuous. Also, we point out that in the above system (3.16), the equations on \([t_{N-1}, T]\) and \([t_{N-2}, t_{N-1})\) are different: \(\Theta^{N-1}(s, x, x)\) appears in \(g\) for the former and \(\Theta^{N-2}(s, \xi, x)\) appears in \(g\) for the latter.

The above discussion seems not enough to obtain an inductive statement. In particular, we need to make sure that the estimate on the error between \(\bar{Y}^k(\cdot), \bar{Z}^k(\cdot)\) and \((Y^k(\cdot), Z^k(\cdot))\) will not be unboundedly accumulated. Thus, let us construct \((\bar{Y}^{N-3}(\cdot), \bar{Z}^{N-3}(\cdot))\) on \([t_{N-3}, T]\). To this end, we consider the following BSDE on \([t_{N-2}, T]\):

\[
\begin{aligned}
d\bar{Y}^{N-3}(s) &= -g(t_{N-3}, s, X(t_{N-3}), X(s), \Theta^{N-1}(s, X(s), X(s)), \bar{Z}^{N-3}(s))ds \\
&\quad + \bar{Z}^{N-3}(s)dW(s), \quad s \in [t_{N-1}, T],
\end{aligned}
\]

(3.17)

\[
\begin{aligned}
d\bar{Y}^{N-3}(s) &= -g(t_{N-3}, s, X(t_{N-3}), X(s), \Theta^{N-2}(s, X(s), X(s)), \bar{Z}^{N-3}(s))ds \\
&\quad + \bar{Z}^{N-3}(s)dW(s), \quad s \in [t_{N-2}, t_{N-1}],
\end{aligned}
\]

\[
\bar{Y}^{N-3}(t_{N-1} - 0) = \Theta^{N-3}(t_{N-1}, X(t_{N-3}), X(t_{N-1})),
\]

where \(\Theta^{N-3}(\cdot, \cdot, \cdot)\) is the solution to the following PDE:

\[
\begin{aligned}
\Theta^{N-3}_s(s, \xi, x) + \frac{1}{2} \sigma(s, x)^T \Theta^{N-3}_{xx}(s, \xi, x) \sigma(s, x) + \Theta^{N-3}_x(s, \xi, x)b(s, x) \\
+ g(t_{N-3}, s, \xi, x, \Theta^{N-1}(s, x, x), \Theta^{N-3}(s, \xi, x) \sigma(s, x)) = 0, \quad (s, x) \in [t_{N-1}, T] \times \mathbb{R}^n,
\end{aligned}
\]

(3.18)

Then we have the following representation:

\[
\begin{aligned}
\bar{Y}^{N-3}(s) &= \Theta^{N-3}(s, X(t_{N-3}), X(s)), \quad s \in [t_{N-2}, T],
\end{aligned}
\]

(3.19)

(3.3) and (3.15), we see that

\[
\begin{aligned}
dY^{N-3}(s) &= -g(t_{N-3}, s, X(t_{N-3}), X(s), \Theta^{N-1}(s, X(t_{N-1}), X(s)), Z^{N-3}(s))ds \\
&\quad + Z^{N-3}(s)dW(s), \quad s \in [t_{N-1}, T],
\end{aligned}
\]

\[
\begin{aligned}
dY^{N-3}(s) &= -g(t_{N-3}, s, X(t_{N-3}), X(s), \Theta^{N-2}(s, X(t_{N-2}), X(s)), Z^{N-3}(s))ds \\
&\quad + Z^{N-3}(s)dW(s), \quad s \in [t_{N-2}, t_{N-1}],
\end{aligned}
\]

\[
Y^{N-3}(T) = \psi(t_{N-3}, X(t_{N-3}), X(T)).
\]
Thus, by the stability of adapted solutions to BSDEs, one has

\[
\mathbb{E} \left[ \sup_{s \in [t_{N-2}, T]} |\bar{Y}^{N-3}(s) - Y^{N-3}(s)|^2 + \int_{t_{N-2}}^{T} |\bar{Z}^{N-3}(s) - Z^{N-3}(s)|^2 ds \right] \\
\leq K_1 \mathbb{E} \left( \int_{t_{N-1}}^{T} |g(t_{N-3}, s, X(t_{N-3}), \Theta^{N-1}(s, X(s), X(s)), Z^{N-3}(s)) - g(t_{N-3}, s, X(t_{N-3}), \Theta^{N-1}(s, X(t_{N-1}), X(s)), Z^{N-3}(s))| ds \\
+ \int_{t_{N-2}}^{t_{N-1}} |g(t_{N-3}, s, X(t_{N-3}), \Theta^{N-2}(s, X(s), X(s)), Z^{N-3}(s)) - g(t_{N-3}, s, X(t_{N-3}), \Theta^{N-2}(s, X(t_{N-2}), X(s)), Z^{N-3}(s))| ds \right)^2 \\
\leq K_1 L^2 \mathbb{E} \left( \int_{t_{N-1}}^{T} |\Theta^{N-1}(s, X(s), X(s)) - \Theta^{N-1}(s, X(t_{N-1}), X(s))| ds \\
+ \int_{t_{N-2}}^{t_{N-1}} |\Theta^{N-2}(s, X(s), X(s)) - \Theta^{N-2}(s, X(t_{N-2}), X(s))| ds \right)^2 \\
\leq K_1 L^2 \|\Theta\|_{\infty}^2 \mathbb{E} \left( \int_{t_{N-1}}^{T} |X(s) - X(t_{N-1})| ds + \int_{t_{N-2}}^{t_{N-1}} |X(s) - X(t_{N-2})| ds \right)^2 \\
\leq K_1 L^2 \|\Theta\|_{\infty}^2 \mathbb{E} \left( \int_{t_{N-1}}^{T} \mathbb{E} |X(s) - X(t_{N-1})|^2 ds + \int_{t_{N-2}}^{t_{N-1}} \mathbb{E} |X(s) - X(t_{N-2})|^2 ds \right) \\
\leq K_1 L^2 \|\Theta\|_{\infty}^2 \mathbb{E} \left( \int_{t_{N-1}}^{T} (s - t_{N-1}) ds + \int_{t_{N-2}}^{t_{N-1}} (s - t_{N-2}) ds \right) \\
\leq \frac{K_0 K_1 L^2 \sqrt{\|\Theta\|_{\infty}^2}}{2} \left( T - t_{N-2} \right)^2 \left( T - t_{N-1} \right)^2 + (t_{N-1} - t_{N-2})^2 \\
\leq \frac{K_0 K_1 L^2 \sqrt{\|\Theta\|_{\infty}^2}}{2} \left( T - t_{N-2} \right)^2 \|\Pi\|,
\]

where

\[ \|\Theta\|_{\infty} = \|\Theta^{N-1}\|_{\infty} \vee \|\Theta^{N-2}\|_{\infty}. \]

Next, on \([t_{N-3}, t_{N-2})\), \((\bar{Y}^{N-3}(\cdot), \bar{Z}^{N-3}(\cdot))\) and \((Y^{N-3}(\cdot), Z^{N-3}(\cdot))\) satisfy the same equation with possibly different terminal conditions at \(t = t_{N-2}\). Hence, we have

\[
\mathbb{E} \left[ \sup_{s \in [t_{N-3}, t_{N-2})} |\bar{Y}^{N-3}(s) - Y^{N-3}(s)|^2 + \int_{t_{N-3}}^{t_{N-2}} |\bar{Z}^{N-3}(s) - Z^{N-3}(s)|^2 ds \right] \leq K_1 \mathbb{E} |\bar{Y}^{N-3}(t_{N-2}) - Y^{N-3}(t_{N-2})|^2 \leq \frac{K_0 K_1 L^2 \sqrt{\|\Theta\|_{\infty}^2}}{2} \left( T - t_{N-2} \right)^2 \|\Pi\|. 
\]

Also, we have

\[
\begin{align*}
\bar{Y}^{N-3}(s) &= \Theta^{N-3}(s, X(t_{N-3}), X(s), X(s)) , \\
\bar{Z}^{N-3}(s) &= \Theta_x^{N-3}(s, X(t_{N-3}), X(s)) \sigma(s, X(s)), \\
&\quad \quad \quad s \in [t_{N-3}, t_{N-2}),
\end{align*}
\]

with \(\Theta^{N-3}(\cdot, \cdot, \cdot)\) satisfying

\[
\begin{align*}
\Theta^{N-3}_s(s, \xi, x) &= \frac{1}{2} \sigma(s, x)^T \Theta_x^{N-3}(s, \xi, x) \sigma(s, x) + \Theta^{N-3}(s, \xi, x) b(s, x), \\
&+ g(t_{N-3}, s, \xi, x, \Theta^{N-3}(s, \xi, x), \Theta_x^{N-3}(s, \xi, x) \sigma(s, x)) = 0, \quad (s, x) \in [t_{N-3}, t_{N-2}) \times \mathbb{R}^n, \\
\Theta^{N-3}(t_{N-2}, \xi, x) &= \Theta^{N-3}(t_{N-2} + 0, \xi, x), \quad x \in \mathbb{R}^n.
\end{align*}
\]

14
Now, we look at the general case. For each $k = 0, 1, \cdots, N - 1$, on $[t_k, T]$, we consider the following BSDE:

$$
\begin{cases}
    d\bar{Y}^k(s) = -g(t_k, s, X(t_k), X(s), \Theta^f(s, X(s), X(s)), \bar{Z}^k(s))ds + \bar{Z}^k(s)dW(s), \\
    s \in [t_k, t_{k+1}), \quad k + 1 \leq \ell \leq N - 1, \\
    dY^k(s) = -g(t_k, s, X(t_k), X(s), Y^k(s), \bar{Z}^k(s))ds + \bar{Z}^k(s)dW(s), \\
    s \in [t_k, t_{k+1}), \\
    Y^k(T) = \psi(t_k, X(t_k), X(T)).
\end{cases}
$$

(3.22)

Then the following representation holds:

$$
\begin{cases}
    \bar{Y}^k(s) = \Theta^k(s, X(t_k), X(s)), \\
    \bar{Z}^k(s) = \Theta^s_k(s, X(t_k), X(s))\sigma(s, X(s)), \\
    s \in [t_k, T],
\end{cases}
$$

(3.23)

where $(s, x) \mapsto \Theta^k(s, \xi, x)$ is the solution to the following PDE:

$$
\begin{cases}
    \Theta^k(s, \xi, x) + \frac{1}{2} \sigma(s, x)^\top \Theta_{xx}^k(s, \xi, x)\sigma(s, x) + \Theta^s_k(s, \xi, x)b(s, x) \\
    + g(t_k, s, \xi, x, \Theta^f(s, x, x), \Theta^s_k(s, \xi, x)\sigma(s, x)) = 0, \\
    (s, x) \in [t_\ell, t_{k+1}) \times \mathbb{R}^n, \quad k + 1 \leq \ell \leq N - 1, \\
    \Theta^k(s, \xi, x) + \frac{1}{2} \sigma(s, x)^\top \Theta_{xx}^k(s, \xi, x)\sigma(s, x) + \Theta^s_k(s, \xi, x)b(s, x) \\
    + g(t_k, s, \xi, x, \Theta^k(s, \xi, x), \Theta^s_k(s, \xi, x)\sigma(s, x)) = 0, \quad (s, x) \in [t_k, t_{k+1}] \times \mathbb{R}^n, \\
    \Theta^k(T, \xi, x) = \psi(t_k, \xi, x), \quad \Theta^k(t_j, \xi, x) = \Theta^k(t_j + 0, \xi, x), \\
    j = N - 1, \cdots, k + 1, \quad x \in \mathbb{R}^n.
\end{cases}
$$

(3.24)

We recall the definition of $\tau^{[\Pi]}(t)$ (see (2.8)), and define

$$
\bar{\tau}^{[\Pi]}(t) = \sum_{k=0}^{N-2} t_{k+1}I_{[t_k, t_{k+1})}(t) + t_NI_{[t_{N-1}, T]}(t), \quad t \in [0, T].
$$

(3.25)

Then

$$
0 \leq \bar{\tau}^{[\Pi]}(t) - t \leq \|\Pi\|, \quad \forall t \in [0, T],
$$

and

$$
[\tau^{[\Pi]}(t), \bar{\tau}^{[\Pi]}(t)] = [t_k, t_{k+1}), \quad \forall t \in [t_k, t_{k+1}), \quad 0 \leq k \leq N - 1.
$$

Let

$$
\Theta^{[\Pi]}(t, s, \xi, x) = \sum_{k=0}^{N-1} \Theta^k(s, \xi, x)I_{[t_k, t_{k+1})}(t), \quad 0 \leq t \leq s \leq T, \quad x, \xi \in \mathbb{R}^n.
$$

For $t \in [t_k, t_{k+1})$, with $k = 0, 1, \cdots, N - 1$, $s \in [\tau^{[\Pi]}(t), T] = [t_k, T]$,

$$
\Theta^k(s, \xi, x) = \Theta^{[\Pi]}(t, s, \xi, x), \quad \Theta^s_k(s, \xi, x) = \Theta^{[\Pi]}(t, s, \xi, x), \\
\Theta^s_k(s, \xi, x) = \Theta^{[\Pi]}(t, s, \xi, x), \quad \Theta^s_{xx}(s, \xi, x) = \Theta^{[\Pi]}_{xx}(t, s, \xi, x),
$$

and

$$
\sum_{\ell=k+1}^{N-1} \Theta^f(s, x, x)I_{[t_\ell, t_{\ell+1})}(s) = \Theta^{[\Pi]}(s, s, x, x), \quad s \in [t_{k+1}, T].
$$
Then the above PDE (3.24) can be written as

\[
\begin{cases}
\Theta^H_s(t, s, \xi, x) + \frac{1}{2} \sigma(s, x) \sigma(s, x) + \Theta^H_s(t, s, \xi, x) b(s, x) \\
+ g(\tau^H(t), s, \xi, x) \Theta^H_s(t, s, \xi, x) \sigma(s, x) + \Theta^H_x(t, s, \xi, x) b(s, x) \\
= 0, \quad (s, \xi, x) \in [\tau^H(t), T] \times \mathbb{R}^n \times \mathbb{R}^n,
\end{cases}
\]

(3.26)

Also,

\[
\mathbb{E} \left[ \sup_{s \in [t_{k+1}, T]} |\bar{Y}^k(s) - Y^k(s)|^2 + \int_{t_{k+1}}^T |\bar{Z}^k(s) - Z^k(s)|^2 ds \right] \\
\leq K_1 \mathbb{E} \left( \sum_{\ell=k+1}^{N-1} \int_{t_\ell}^{t_{\ell+1}} |g(t_k, s, X(t_k), \Theta^\ell(s, X(s), s), Z^\ell(s)) - g(t_k, s, X(t_k), \Theta^\ell(s, X(t_\ell), s), Z^\ell(s))| ds \right) \\
\leq K_1 L^2 \mathbb{E} \left( \sum_{\ell=k+1}^{N-1} \int_{t_\ell}^{t_{\ell+1}} \left| \Theta^\ell(s, X(s), s) - \Theta^\ell(s, X(t_\ell), s) \right| ds \right)^2 \\
\leq K_1 L^2 \|\Theta_\ell\|_\infty^2 \mathbb{E} \left( \sum_{\ell=k+1}^{N-1} \int_{t_\ell}^{t_{\ell+1}} |X(s) - X(t_\ell)| ds \right)^2 \\
\leq K_1 L^2 \|\Theta_\ell\|_\infty^2 (T - t_{k+1}) \left( \sum_{\ell=k+1}^{N-1} \int_{t_\ell}^{t_{\ell+1}} \mathbb{E} |X(s) - X(t_\ell)|^2 ds \right) \\
\leq K_1 L^2 \|\Theta_\ell\|_\infty^2 (T - t_{k+1}) K_1 \left( \sum_{\ell=k+1}^{N-1} \int_{t_\ell}^{t_{\ell+1}} (s - t_\ell) ds \right) \\
\leq \frac{K_0 K_1 L^2 \|\Theta_\ell\|_\infty^2}{2} (T - t_{k+1}) \sum_{\ell=k+1}^{N-1} (t_{\ell+1} - t_\ell)^2 \leq \frac{K_0 K_1 L^2 \|\Theta_\ell\|_\infty^2}{2} (T - t_{k+1})^2 \|\Pi\|,
\]

where

\[
\|\Theta_\ell\|_\infty \leq \max_{k+1 \leq \ell \leq N} \|\Theta^\ell\|_\infty,
\]

and

\[
\mathbb{E} \left[ \sup_{s \in [t_{k+1}, T]} |\bar{Y}^k(s) - Y^k(s)|^2 + \int_{t_k}^{t_{k+1}} |\bar{Z}^k(s) - Z^k(s)|^2 ds \right] \\
\leq K_1 \mathbb{E} |\bar{Y}^k(t_{k+1}) - Y^k(t_{k+1})|^2 \leq \frac{K_0 K_1^2 L^2 \|\Theta_\ell\|_\infty^2}{2} (T - t_{k+1})^2 \|\Pi\|.
\]

(3.27)

Now, let

\[
\begin{align*}
\bar{Y}^H(s) &= \sum_{k=0}^{N-1} \bar{Y}^k(s) I_{[t_k, t_{k+1})}(s), \quad s \in [0, T), \\
\bar{Z}^H(t, s) &= \sum_{k=0}^{N-1} \bar{Z}^k(s) I_{[t_k, t_{k+1})}(t), \quad 0 \leq t \leq s \leq T.
\end{align*}
\]
Then
\[
\begin{aligned}
\bar{Y}^\Pi(t) &= \sum_{k=0}^{N-1} \Theta^k(t,\xi,x)I_{[t_k,t_{k+1})}(s) = \Theta^\Pi(t,\xi,x), \\
\bar{Z}^\Pi(t) &= \sum_{k=0}^{N-1} \Theta^k_x(t,\xi,x)\sigma(s,\xi,x)I_{[t_k,t_{k+1})}(t) = \Theta^\Pi(t,\xi,x)\sigma(s,\xi,x).
\end{aligned}
\]

Consequently, for any \(s \in [0,T]\), let \(s \in [t_k, t_{k+1})\).
\[
\mathbb{E}|Y^\Pi(s) - \Theta^\Pi(s,\xi,x)|^2 \leq 2\mathbb{E}|Y^\Pi(s) - \bar{Y}^\Pi(s)|^2 + 2\mathbb{E}|\Theta^\Pi(s,\xi,x)|^2
\leq K_0K_1^2 L^2\|\Theta\|_\infty^2 (T-t_{k+1})^2 \|\Pi\| + 2\|\Theta\|_\infty^2 \mathbb{E}|X(t_k) - \bar{X}(s)|^2
\leq K_0K_1^2 L^2\|\Theta\|_\infty^2 T^2 \|\Pi\| + 2\|\Theta\|_\infty^2 K_0 \|\Pi\| \leq K \|\Pi\|.
\]

Also,
\[
\begin{aligned}
\mathbb{E}\int_0^T \int_t^T |Z^\Pi(t) - \Theta^\Pi_x(t,\xi,x)\sigma(s,\xi,x)|^2 ds dt &\leq \sum_{k=0}^{N-1} \mathbb{E}\int_{t_k}^{t_{k+1}} \int_{t_k}^{t_{k+1}} |Z^k(s) - \Theta^k_x(s,\xi,x)\sigma(s,\xi,x)|^2 ds dt \\
&\leq \sum_{k=0}^{N-1} \mathbb{E}\int_{t_k}^{t_{k+1}} \int_{t_k}^{t_{k+1}} \left(|Z^k(s) - \bar{Z}^k(s)|^2 + |\Theta^k_x(s,\xi,x)\sigma(s,\xi,x)|^2\right) ds dt \\
&\leq \sum_{k=0}^{N-1} 2\mathbb{E}\int_{t_k}^{t_{k+1}} K \|\Pi\| ds + \sum_{k=0}^{N-1} 2\mathbb{E}\int_{t_k}^{t_{k+1}} K \|\Pi\| dt \leq K \|\Pi\|.
\end{aligned}
\]

Hence, at the limit (as \(\|\Pi\| \to 0\)), we have the following representation:
\[
\begin{aligned}
Y(s) &= \Theta(s,\xi,x), \quad s \in [0,T], \\
Z(t) &= \Theta_x(t,\xi,x)\sigma(s,\xi,x), \quad (t,s) \in \Delta[0,T],
\end{aligned}
\tag{3.28}
\]

if \(\Theta(\cdot,\cdot,\cdot,\cdot)\) satisfies (1.21) which is rewritten here:
\[
\begin{aligned}
\Theta_x(t,\xi,x) + \frac{1}{2} \sigma(s,\xi,x)^T \Theta_{xx}(t,\xi,x)\sigma(s,\xi,x) + \Theta_x(t,\xi,x) b(s,\xi,x) \\
+ g(t,\xi,x,\Theta(s,\xi,x),\Theta_x(t,\xi,x)\sigma(s,\xi,x)) = 0, \quad (t,\xi,x) \in \Delta[0,T] \times \mathbb{R}^n \times \mathbb{R}^n,
\end{aligned}
\tag{3.29}
\]

The above derivation tells us that if everything is fine, (3.28)–(3.29) should give us the right representation. This can actually be proved directly.

**Theorem 3.1.** Let (H1)–(H2) hold. Let \(\Theta : \Delta[0,T] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}\) be the classical solution of the representation PDE (3.29). Let \((Y(\cdot),Z(\cdot,\cdot,\cdot))\) be the adapted solution to the Type-I BSVIE (1.18) with \(X(\cdot)\) being the solution to SDE (1.12). Then representation (3.28) holds.

**Proof.** For fixed \(t \in [0,T]\), applying Itô’s formula to \(s \mapsto \Theta(t,s,t,t),X(t))\) on \([t,T]\), we have
\[
d\Theta(t,s,t,t),X(t),X(s)) = \left( \Theta_x(t,s,t,t),X(t)) + \Theta_x(t,s,t,t) b(s,\xi,x) \right) ds + \Theta_x(t,s,t,t)\sigma(s,\xi,x) dW(s).
\tag{3.30}
\]

17
Since $\Theta(\cdot, \cdot, \cdot)$ satisfies PDE (3.29), one has

$$d\Theta(t, s, X(t), X(s)) = -g(t, s, X(t), X(s), \Theta(t, s, X(s), X(s)), \Theta_x(t, s, X(t), X(s))\sigma(s, X(s)))ds + \Theta_x(t, s, X(t), X(s))\sigma(s, X(s))dW(s),$$

and

$$\Theta(t, T, X(t), X(T)) = \psi(t, X(t), X(T)).$$

Now, we define

$$\lambda(t, s) := \Theta(t, s, X(t), X(s)), \quad Z(t, s) := \Theta_x(t, s, X(t), X(s))\sigma(s, X(s)), \quad s \geq t.$$ (3.32)

Then

$$\lambda(t, s) = \psi(t, X(t), X(T)) + \int_s^T g(t, r, X(t), X(r), \lambda(r, r), Z(t, r))dr - \int_s^T Z(t, r)dW(r).$$ (3.33)

Let $t = s$ and $Y(s) := \lambda(s, s)$, we then see that $(Y(\cdot), Z(\cdot, \cdot))$ satisfies BSVIE (1.18) and desired representation is obtained.

4 Representation of Adapted M-solutions for Type-II BSVIEs

In this section, we are going to establish a representation of adapted M-solutions for Type-II BSVIE (1.19), where both $Z(t, s)$ and $Z(s, t)$ appear in the drift. We still let $X(\cdot)$ be the solution to FSDE (1.12). Let us first present the following result which is interesting itself.

**Proposition 4.1.** Let $\Lambda : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^m$ be continuous. Let (H1) hold and the following PDE system admit a classical solution $\Gamma(\cdot, \cdot, \cdot)$:

$$\begin{cases}
\Gamma_x(t, s, x) + \frac{1}{2} \sigma(s, x)^\top \Gamma_{xx}(t, s, x)\sigma(s, x) + \Gamma_x(t, s, x)b(s, x) = 0, & 0 \leq s \leq t \leq T, \quad x \in \mathbb{R}^n, \\
\Gamma(t, t, x) = \Lambda(t, x), & (t, x) \in [0, T] \times \mathbb{R}^n,
\end{cases}$$ (4.1)

where the meaning of $\sigma(s, x)^\top \Gamma_{xx}(t, s, x)\sigma(s, x)$ is similar to (1.6). Then

$$\Lambda(t, X(t)) = E\Lambda(t, X(t)) + \int_0^t \Gamma_x(t, s, X(s))\sigma(s, X(s))dW(s), \quad t \in [0, T].$$ (4.2)

**Proof.** We consider the following (decoupled) FBSDE on $[0, t]$:

$$\begin{cases}
dX(s) = b(s, X(s))ds + \sigma(s, X(s))dW(s), & s \in [0, t], \\
d\eta(t, s) = \zeta(t, s)dW(s), & s \in [0, t], \\
\eta(t, t) = \Lambda(t, X(t)),
\end{cases}$$ (4.3)

where $t \in [0, T]$ is a parameter. Then the following representation holds:

$$\begin{cases}
\eta(t, s) = \Gamma(t, s, X(s)), & s \in [0, t], \\
\zeta(t, s) = \Gamma_x(t, s, X(s))\sigma(s, X(s)),
\end{cases}$$ (4.4)

where $\Gamma(\cdot, \cdot, \cdot)$ is the solution to (4.1). Consequently,

$$\Lambda(t, X(t)) = \eta(t, t) = \eta(t, 0) + \int_0^t \zeta(t, s)dW(s).$$
Taking expectation, we have\[\mathbb{E} \Lambda(t, X(t)) = \eta(t, 0).\]

Therefore, (4.2) follows. \[\square\]

From the above, we see that when \((t, s) \mapsto \Gamma_x(t, s, x)\) and \(s \mapsto \sigma(s, x)\) are continuous, the map \((t, s) \mapsto \zeta(t, s)\) is continuous (see (4.4)).

Now, we consider Type-II BSVIE (1.19). Let \((Y(\cdot), Z(\cdot, \cdot))\) be the adapted S-solution. Then we have (1.16). Suppose \[Y(t) = \Lambda(t, X(t)), \quad t \in [0, T],\]

for some undetermined continuous function \(\Lambda(\cdot, \cdot)\). By Proposition 4.1, we have \[Z(t, s) = \Gamma_x(t, s, X(s))\sigma(s, X(s)), \quad 0 \leq s \leq t \leq T.\]

Thus, switching \(s\) and \(t\), one has \[Z(s, t) = \Gamma_x(s, t, X(t))\sigma(t, X(t)), \quad 0 \leq t \leq s \leq T.\]

We consider the following Type-I BSVIE:

\begin{equation}
Y(t) = \psi(t, X(t), X(T)) + \int_t^T g(t, s, X(t), X(s), Y(s), Z(t, s), \Gamma_x(s, t, X(t))\sigma(t, X(t))) ds
- \int_t^T Z(t, s) dW(s), \quad t \in [0, T].
\end{equation}

If we let \[\bar{g}(t, s, \xi, x, y, z) = g(t, s, \xi, x, y, z, \Gamma_x(s, t, \xi)\sigma(t, \xi)),\]

then (4.5) becomes

\begin{equation}
Y(t) = \psi(t, X(t), X(T)) + \int_t^T \bar{g}(t, s, X(t), X(s), Y(s), Z(t, s)) ds - \int_t^T Z(t, s) dW(s), \quad t \in [0, T].
\end{equation}

Now, from the result of the previous section, we have the following representation:

\[
\begin{cases}
Y(s) = \Theta(s, s, X(s), X(s)), & s \in [0, T], \\
Z(t, s) = \Theta_x(t, s, X(t), X(s))\sigma(s, X(s)), & 0 \leq t \leq s \leq T, \\
Z(t, s) = \Gamma_x(t, s, X(s))\sigma(s, X(s)), & 0 \leq s \leq t \leq T,
\end{cases}
\]

with \((\Gamma, \Theta)\) being the solution to (1.23) which is rewritten here

\[
\begin{cases}
\Gamma_x(t, s, x) + \frac{1}{2} \sigma(s, x)^\top \Gamma_{xx}(t, s, x)\sigma(s, x) + \Gamma_x(t, s, x)b(s, x) = 0, & 0 \leq s \leq t \leq T, \ x \in \mathbb{R}^n, \\
\Theta_x(t, s, \xi, x) + \frac{1}{2} \sigma(s, x)^\top \Theta_{xx}(t, s, \xi, x)\sigma(s, x) + \Theta_x(t, s, \xi, x)b(s, x) + g(t, s, \xi, x, \Theta(s, s, x, x), \Theta_x(s, s, x, x), \Gamma_x(s, s, x)\sigma(s, x)) = 0, & (t, s, \xi, x) \in \Delta[0, T] \times \mathbb{R}^n \times \mathbb{R}^n, \\
\Gamma(t, t, x) = \Theta(t, t, x, x), & (t, t) \in [0, T] \times \mathbb{R}^n, \\
\Theta(t, s, \xi, x) = \psi(t, \xi, x), & (t, s, \xi, x) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^n.
\end{cases}
\]

We now state the representation result as follows.
Theorem 4.2. Let \( \Theta, \Gamma \) be the solution of representation PDE (4.8). Let \((Y(\cdot), Z(\cdot, \cdot))\) be the adapted M-solution to Type-II BSVIE (1.19) with \(X(\cdot)\) being the solution to SDE (1.12). Then representation (4.7) holds.

Proof. For given \( t \in [0, T] \), applying Itô's formula to \( s \mapsto \Theta(t, s, X(t), X(s)) \) on \([t, T]\), one has

\[
d\Theta(t, s, X(t), X(s)) = \left( \Theta_x(t, s, X(t), X(s)) + \Theta_x(t, s, X(t), X(s))b(s, X(s)) \right. \\
+ \left. \frac{1}{2} \sigma(s, X(s))^{\top} \Theta_{xx}(t, s, X(t), X(s)) \sigma(s, X(s)) \right) ds + \Theta_x(t, s, X(t), X(s)) \sigma(s, X(s)) dW(s) \\
+ \frac{1}{2} \sigma(s, X(s))^{\top} \Theta_{xx}(t, s, X(t), X(s)) \sigma(s, X(s)) ds.
\]

Since \( \Theta \) satisfies the second PDE of (4.8), one has

\[
d\Theta(t, s, X(t), X(s)) = \Theta_x(t, s, X(t), X(s)) \sigma(s, X(s)) dW(s) - g(t, s, X(t), X(s), \Theta(t, s, X(t), X(s)), \Theta_x(t, s, X(t), X(s)) \sigma(s, X(s)), \Gamma(t, s, X(t)) \sigma(s, X(s))) ds,
\]

and

\[
\Theta(t, T, X(t), X(T)) = \psi(t, X(t), X(T)).
\]

Set

\[
\lambda(t, s) := \Theta(t, s, X(t), X(s)), \quad Z(t, s) := \Theta_x(t, s, X(t), X(s)) \sigma(s, X(s)), \quad s \geq t.
\]

Then

\[
\lambda(t, s) = \psi(t, X(t), X(T)) + \int_s^T g(t, r, X(t), X(r), \lambda(r, r), Z(t, r), \Gamma(r, t, X(t)) \sigma(r, X(r))) dr \\
- \int_s^T Z(t, r) dW(r).
\]

Let \( t = s \) and \( Y(s) := \lambda(s, s) \), we then obtain

\[
Y(t) = \psi(t, X(t), X(T)) + \int_t^T g(t, r, X(t), X(r), \lambda(r, r), Z(t, r), \Gamma(r, t, X(t)) \sigma(r, X(r))) dr \\
- \int_t^T Z(t, r) dW(r).
\]

Note that

\[
Y(t) = \Theta(t, t, X(t), X(t)) = \Gamma(t, t, X(t))
\]

where \( \Gamma \) satisfies the first PDE in (4.8). By Proposition 4.1, we know that

\[
Y(t) = \mathbb{E}Y(t) + \int_0^t \Gamma_x(t, s, X(s)) \sigma(s, X(s)) dW(s), \quad t \geq s.
\]

Consequently, by defining \( Z(t, s) := \Gamma_x(t, s, X(s)) \sigma(s, X(s)) \) with \( t \geq s \), we can rewrite above BSVIE as

\[
Y(t) = \psi(t, X(t), X(T)) + \int_t^T g(t, r, X(t), X(r), \lambda(r, r), Z(t, r), Z(t, r)) dr - \int_t^T Z(t, r) dW(r).
\]

The conclusion then follows easily. \( \square \)
5 Well-posedness of the Representation PDEs

In this section, we will establish the well-posedness of the representation PDEs (3.29) (which is a copy of (1.21)) and (4.8) (which is a copy of (1.23)), in certain sense. Let us first look at the representation PDE (3.29) for Type-I BSVIEs, which is recalled here, for convenience:

\[
\begin{align*}
\Theta_s(t, s, \xi, x) + \frac{1}{2} \sigma(s, x) \cdot \Theta_{xx}(t, s, \xi, x) + \Theta_x(t, s, \xi, x) b(s, x) \\
+ g(t, s, \xi, x, \Theta(s, s, x, x), \Theta_x(t, s, \xi, x) \sigma(s, x)) = 0, \\
(\Theta(t, T, \xi, x) = \psi(t, \xi, x), \quad (t, \xi, x) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^n.
\end{align*}
\] (5.1)

If we denote \( \Theta = (\Theta^1, \cdots, \Theta^m) \) and

\[
\frac{1}{2} \sigma(s, x) \sigma(s, x)^\top = a(s, x) = (a_{ij}(s, x)), \quad b(s, x) = (b_1(s, x), \cdots, b_n(s, x))^\top,
\]

then (5.1) can be rewritten as the following system (parameterized by \((t, \xi) \in [0, T] \times \mathbb{R}^n\)):

\[
\begin{align*}
\Theta^k_s(t, s, \xi, x) + \sum_{i,j=1}^n a_{ij}(s, x) \Theta^k_{x_i x_j}(t, s, \xi, x) + \sum_{i=1}^n b_i(s, x) \Theta^k_{x_i}(t, s, \xi, x) \\
+ g^k(t, s, \xi, x, \Theta(s, s, x, x), \Theta_x(t, s, \xi, x) \sigma(s, x))) = 0, \\
(\Theta^k(t, T, \xi, x) = \psi^k(t, \xi, x), \quad x \in \mathbb{R}^n, \quad 1 \leq k \leq m,
\end{align*}
\] (5.2)

which is a quasilinear parabolic system for unknown functions \( \Theta^1, \cdots, \Theta^m \), with the same leading part for each equation.

5.1 Linear parabolic PDEs

To study parabolic system (5.1) or its equivalent form (5.2), let us first adopt some notations from [17] (Chapter 1, pp.7–8). For any suitable function \( \varphi: [S, T] \times \mathbb{R}^n \rightarrow \mathbb{R} \), with \( \alpha \in (0, 1) \) and \( S \in [0, T] \), let

\[
\begin{align*}
|\varphi|^{(0)} &= \|\varphi\|_{L^\infty([S, T] \times \mathbb{R}^n)}, \quad |\varphi|^{(1)} = |\varphi|^{(0)} + |\varphi_{x_1}|^{(0)}, \quad |\varphi|^{(2)} = |\varphi|^{(1)} + |\varphi_s|^{(0)} + |\varphi_{xx}|^{(0)}, \\
\langle \varphi \rangle_s^{(\frac{1}{2})} &= \sup_{x \in [S, T], x \in \mathbb{R}^n} \frac{|\varphi(s, x) - \varphi(s', x)|}{|s - s'|^{\frac{1}{2}}}, \quad \langle \varphi \rangle_s^{(\alpha)} = \sup_{x \in [S, T], x \in \mathbb{R}^n} \frac{|\varphi(s, x) - \varphi(s, x')|}{|s - s'|^{\alpha}}, \\
|\varphi|^{(1+\alpha)} &= |\varphi|^{(1)} + \langle \varphi \rangle_s^{(\alpha)}, \\
|\varphi|^{(2+\alpha)} &= |\varphi|^{(2)} + \langle \varphi \rangle_s^{(\alpha)} + \langle \varphi_{xx} \rangle_s^{(\alpha)} + \langle \varphi_x \rangle_s^{(\alpha+\frac{1}{2})}, \\
\langle \varphi \rangle_s^{(\alpha)} &= \langle \varphi \rangle^{(\frac{1}{2})} + \langle \varphi_{xx} \rangle^{(\alpha)} + \langle \varphi_{s} \rangle^{(\alpha+\frac{1}{2})}.
\end{align*}
\] (5.3)

When \([S, T] \times \mathbb{R}^n\) needs to be emphasized, we use, say, \(|\varphi|^{(1)}_{[S, T] \times \mathbb{R}^n}\), etc. We denote

\[
C^{\frac{1}{2}, \alpha}([S, T] \times \mathbb{R}^n) = \left\{ \varphi: [S, T] \times \mathbb{R}^n \rightarrow \mathbb{R} \mid |\varphi|^{(\alpha)}_{[S, T] \times \mathbb{R}^n} < \infty \right\}.
\]

Clearly, \( \varphi(\cdot, \cdot) \in C^{\frac{1}{2}, \alpha}([S, T] \times \mathbb{R}^n) \) if and only if \( \varphi(\cdot, \cdot) \in L^\infty([S, T] \times \mathbb{R}^n) \) and

\[
|\varphi(s, x) - \varphi(s', x')| \leq K \left( |s - s'|^{\frac{1}{2}} + |x - x'|^{\alpha} \right), \quad \forall s, s' \in [S, T], \quad x, x' \in \mathbb{R}^n, \quad |x - x'| \leq 1.
\]

Also, we denote

\[
\begin{align*}
C^{\frac{1}{2}, 1+\alpha}([S, T] \times \mathbb{R}^n) &= \left\{ \varphi: [S, T] \rightarrow \mathbb{R}^n \mid |\varphi|^{(1+\alpha)}_{[S, T] \times \mathbb{R}^n} < \infty \right\}, \\
C^{1+\frac{1}{2}, 2+\alpha}([S, T] \times \mathbb{R}^n) &= \left\{ \varphi: [S, T] \rightarrow \mathbb{R}^n \mid |\varphi|^{(2+\alpha)}_{[S, T] \times \mathbb{R}^n} < \infty \right\}.
\end{align*}
\]
Let us consider the following Cauchy problem for linear equation:

\[ \begin{align*}
  v_s(t,s,x) + \sum_{i,j=1}^{n} a_{ij}(s,x)v_{x_jx_i}(s,x) + \sum_{i=1}^{n} b_i(s,x)v_{x_i}(s,x) + f(s,x) &= 0, \quad (s,x) \in [t,T] \times \mathbb{R}^n, \\
  v(T,x) &= h(x), \quad x \in \mathbb{R}^n.
\end{align*} \tag{5.4} \]

We introduce the following hypotheses:

**P1** Operator \( \mathcal{L} \) is uniformly parabolic, i.e., there exist constants \( \lambda_0 > 0 \) such that

\[ \sum_{i,j=1}^{n} a_{ij}(s,x)\xi_i\xi_j = \langle a(s,x)\xi,\xi \rangle \geq \lambda_0 |\xi|^2, \quad \forall \xi \in \mathbb{R}^n, \quad (s,x) \in [0,T] \times \mathbb{R}^n. \]

**P2** Functions \( a_{ij}(s,x), b_i(s,x) \) are continuous and bounded, and for some \( \alpha \in (0,1) \),

\[ \begin{align*}
  |a_{ij}(s,x) - a_{ij}(s',x')| &\leq K (|s-s'|^{\frac{2}{\alpha}} + |x-x'|^\alpha), \\
  |b_i(s,x) - b_i(s',x')| &\leq K |x-x'|^\alpha,
\end{align*} \]

\( (s,x), (s',x') \in [0,T] \times \mathbb{R}^n. \)

By [17] (Chapter IV, Sections 13–14) (see also [12], Chapter 1, Section 7) we have the following result.

**Proposition 5.1.** Let (P1)–(P2) hold. Assume that

\[ f(\cdot, \cdot) \in C^{\frac{2}{\alpha},\alpha}([0,T] \times \mathbb{R}^n), \quad h(\cdot) \in C^{2+\alpha}(\mathbb{R}^n), \]

for some \( \alpha \in (0,1) \). Then Cauchy problem (5.4) admits a unique classical solution \( v(\cdot, \cdot) \in C^{1+\frac{2}{\alpha},2+\alpha}([0,T] \times \mathbb{R}^n) \). Moreover, \( v(\cdot, \cdot) \) is represented as follows:

\[ v(s,x) = \int_{\mathbb{R}^n} G(s,x;\tau,\eta)h(\eta)d\eta + \int_{s}^{T} \int_{\mathbb{R}^n} G(s,x;\tau,\eta)f(\tau,\eta)d\eta d\tau, \quad (s,x) \in [0,T] \times \mathbb{R}^n. \]

Here \( G(s,x;\tau,\eta) \) is called the fundamental solution of the parabolic operator \( \mathcal{L} \), which satisfies the following: There exists a \( \lambda > 0 \) such that for any \( x, \eta \in \mathbb{R}^n \),

\[ \begin{align*}
  |G(s,x;\tau,\eta)| &\leq \frac{K}{(\tau-s)^{\frac{2}{\alpha}}} e^{-\frac{\lambda |x-\eta|^2}{\tau-s}}, \\
  |G_x(s,x;\tau,\eta)| &\leq \frac{K}{(\tau-s)^{\frac{\alpha+2}{\alpha}}} e^{-\frac{\lambda |x-\eta|^2}{\tau-s}}, \quad 0 \leq s < \tau \leq T. \tag{5.5}
\end{align*} \]

Moreover,

\[ |v|^{(2+\alpha)}_{[0,T] \times \mathbb{R}^n} \leq K (|f|^{(\alpha)}_{[0,T] \times \mathbb{R}^n} + |h|^{(2+\alpha)}_{\mathbb{R}^n}), \tag{5.6} \]

and for any \( S \in [0,T), \)

\[ |v|^{(1+\alpha)}_{[S,T] \times \mathbb{R}^n} \leq |h|^{(1+\alpha)}_{\mathbb{R}^n} + K(T-S)^{\frac{2}{\alpha}} (|f|^{(\alpha)}_{[S,T] \times \mathbb{R}^n} + |h|^{(2+\alpha)}_{\mathbb{R}^n}). \tag{5.7} \]

**Proof.** The proof up to (5.6) is standard (see [17, 12]). Let us look at (5.7) which will play a crucial role below.

Note that by defining

\[ \begin{align*}
  \bar{v}(s,x) &= v(s,x) - h(x), \\
  \bar{f}(s,x) &= f(s,x) + \frac{1}{2} \sigma(s,x)^T h_{xx}(x) + h_x(x)b(s,x), \quad (s,x) \in [0,T] \times \mathbb{R}^n,
\end{align*} \]

...
we see that $\tilde{v}(\cdot, \cdot)$ is the solution to the following:

$$\begin{align*}
\begin{cases}
\tilde{v}_s(s, x) + \frac{1}{2} \sigma(s, x)\sigma(s, x) + \tilde{v}_x(s, x)b(s, x) + \tilde{f}(s, x) = 0, \\
\tilde{v}(T, x) = 0,
\end{cases} 
\quad (s, x) \in [0, T] \times \mathbb{R}^n,
\end{align*}
(5.8)$$

Thus, the following representation holds:

$$\tilde{v}(s, x) = \int_s^T \int_{\mathbb{R}^n} G(s, x; \tau, \eta) \tilde{f}(\tau, \eta) d\eta d\tau, 
\quad (s, x) \in [0, T] \times \mathbb{R}^n.
(5.9)$$

Following the steps of proving the inequality (5.6) in [17] (Chapter IV, Section 14), we have the following useful estimates:

$$\begin{align*}
\begin{cases}
|\tilde{v}_s|_{[S,T] \times \mathbb{R}^n} \leq K(T - S) |\tilde{f}|_{[S,T] \times \mathbb{R}^n}, \\
|\tilde{v}_x|_{[S,T] \times \mathbb{R}^n} \leq K(T - S)^{\frac{1+\alpha}{2\alpha}} |\tilde{f}|_{[S,T] \times \mathbb{R}^n}, \\
|\tilde{v}|_{[S,T] \times \mathbb{R}^n} \leq K(T - S)^{\frac{\alpha}{2}} |\tilde{f}|_{[S,T] \times \mathbb{R}^n}.
\end{cases}
(5.10)
\end{align*}$$

Also, for any $s, s' \in [S, T], x, x' \in \mathbb{R}^n$,

$$\begin{align*}
\begin{cases}
|\tilde{v}(s, x) - \tilde{v}(s', x)| \leq K |s - s'|^{\frac{1+\alpha}{\alpha}} |\tilde{f}|_{[S,T] \times \mathbb{R}^n}, \\
|\tilde{v}_x(s, x) - \tilde{v}_x(s', x)| \leq K |s - s'|^{\frac{\alpha}{2}} |\tilde{f}|_{[S,T] \times \mathbb{R}^n}.
\end{cases}
(5.11)
\end{align*}$$

Now, from (5.10), for any $s, s' \in [S, T], x, x' \in \mathbb{R}^n, |x - x'| \leq 1$, we further have

$$\begin{align*}
|\tilde{v}(s, x) - \tilde{v}(s', x)| & \leq \int_0^1 |\tilde{v}_s(s + \mu(s - s'), x)| d\mu |s - s'| \leq K(T - S)^{\frac{\alpha}{2}} |\tilde{f}|_{[S,T] \times \mathbb{R}^n} |s - s'| \\
& \leq K |\tilde{f}|_{[S,T] \times \mathbb{R}^n} |(T - S)^{\frac{\alpha}{2}}|s - s'|^2 + (T - S)^{\frac{\alpha}{6}} |s - s'|^\frac{1+\alpha}{\alpha},
\end{align*}$$

$$\begin{align*}
|\tilde{v}(s, x) - \tilde{v}(s, x')| & \leq \int_0^1 |\tilde{v}_x(s, x' + \mu(x - x'))| d\mu |x - x'| \leq K(T - S)^{\frac{1+\alpha}{2}} |\tilde{f}|_{[S,T] \times \mathbb{R}^n} |x - x'| \\
& \leq K(T - S)^{\frac{1+\alpha}{2}} |\tilde{f}|_{[S,T] \times \mathbb{R}^n} |x - x'|^\alpha,
\end{align*}$$

which leads to

$$\langle \tilde{v} \rangle_{[S,T] \times \mathbb{R}^n} = \langle \tilde{v} \rangle_{s, [S,T] \times \mathbb{R}^n} + \langle \tilde{v}_x \rangle_{x, [S,T] \times \mathbb{R}^n} \leq K(T - S)^{\frac{1+\alpha}{2}} |\tilde{f}|_{[S,T] \times \mathbb{R}^n}.
(5.12)$$

Next, the first inequality in (5.11) implies that

$$|\tilde{v}_x(s, x) - \tilde{v}_x(s', x)| \leq K |\tilde{f}|_{[S,T] \times \mathbb{R}^n} |s - s'|^{\frac{1+\alpha}{\alpha}} \leq K(T - S)^{\frac{\alpha}{2}} |\tilde{f}|_{[S,T] \times \mathbb{R}^n} |s - s'|^2.$$

Similar to the above, for any $s \in [S, T], x, x' \in \mathbb{R}^n, |x - x'| \leq 1$, making use of the third inequality in (5.10), we have

$$|\tilde{v}_x(s, x) - \tilde{v}_x(s, x')| \leq \int_0^1 |\tilde{v}_xx(s, x' + \mu(x - x'))| d\mu |x - x'| \leq K(T - S)^{\frac{\alpha}{2}} |\tilde{f}|_{[S,T] \times \mathbb{R}^n} |x - x'|,$$

which leads to

$$\langle \tilde{v}_x \rangle_{[S,T] \times \mathbb{R}^n} \leq K(T - S)^{\frac{\alpha}{2}} |\tilde{f}|_{[S,T] \times \mathbb{R}^n}.
(5.13)$$

Hence, combining the above, we end up with

$$|\tilde{v}|_{[S,T] \times \mathbb{R}^n} \leq |\tilde{v}_s|_{[S,T] \times \mathbb{R}^n} + |\tilde{v}_x|_{[S,T] \times \mathbb{R}^n} + \langle \tilde{v} \rangle_{s, [S,T] \times \mathbb{R}^n} + \langle \tilde{v}_x \rangle_{x, [S,T] \times \mathbb{R}^n} \leq K(T - S)^{\frac{\alpha}{2}} |\tilde{f}|_{[S,T] \times \mathbb{R}^n}.
(5.14)$$

This implies (5.7).
5.2 The first representation PDE

Now, let us return to the first representation PDE (5.1). We impose the following further assumption.

(H3) The maps \( b(s, x), \sigma(s, x), \psi(t, \xi, x), \) and \( g(t, s, \xi, x, y, z) \) are bounded, have all required differentiability with bounded derivatives. Moreover, there exists a constant \( \sigma > 0 \) such that

\[
|\sigma(s, x)|^2, \quad \forall (s, x, \xi) \in [0, T] \times \mathbb{R}^n \times \mathbb{R}^n. \tag{5.15}
\]

The above assumption is much more than enough. However, in this paper, we prefer not to get into the most generality in this aspect, to reduce the complexity of presentation. Also, we will extend \( (t, s) \mapsto g(t, s, \xi, x, y, z) \) from \( \Delta[0, T] \) to \( [0, T] \) by letting

\[
g(t, s, \xi, x, y, z) = g(s, t, \xi, x, y, z), \quad \forall t, s \in [0, T].
\]

We now state the following theorem.

**Theorem 5.2.** Let (H3) hold. Then for any \((t, \xi) \in [0, T) \times \mathbb{R}^n\), system (5.1) admits a unique classical solution \( \Theta(t, \cdot, \xi, \cdot) \in C^{1 + \frac{\alpha}{2}, 2+\alpha}([t, T] \times \mathbb{R}^n) \), and the following holds

\[
\sup_{(t, \xi) \in [0, T] \times \mathbb{R}^n} |\Theta(t, \cdot, \xi, \cdot)|^{(2+\alpha)}_{[t, T] \times \mathbb{R}^n} \leq K \left( 1 + \sup_{(t, \xi) \in [0, T] \times \mathbb{R}^n} |\psi(t, \xi, \cdot)|^{(2+\alpha)}_{\mathbb{R}^n} \right). \tag{5.16}
\]

Moreover, if \( \hat{\Theta}(t, \cdot, \xi, \cdot) \) is the solution to the system (5.1) with the pair \((\psi, g)\) replaced by \((\hat{\psi}, \hat{g})\) that also satisfies (H3), then the following stability estimate holds:

\[
\sup_{(t, \xi) \in [0, T] \times \mathbb{R}^n} |\Theta(t, \cdot, \xi, \cdot) - \hat{\Theta}(t, \cdot, \xi, \cdot)|^{(2+\alpha)}_{[t, T] \times \mathbb{R}^n} \leq K \sup_{(t, \xi) \in [0, T] \times \mathbb{R}^n} \left( |\psi(t, \xi, \cdot) - \hat{\psi}(t, \xi, \cdot)|^{(2+\alpha)}_{\mathbb{R}^n} + |g(t, t, \cdot, \xi, \cdot) - \hat{g}(t, \cdot, \xi, \cdot)|^{(2+\alpha)}_{[t, T] \times \mathbb{R}^n} \right),
\]

\[
\forall (t, \xi) \in [0, T) \times \mathbb{R}^n,
\]

where

\[
g(t, s, \xi, x) = g(s, t, \xi, x, s, x), \quad (t, s, \xi, x, s, x) \in \Delta[0, T] \times \mathbb{R}^n \times \mathbb{R}^n.
\]

**Proof.** We split the proof into several steps.

**Step 1.** A reduction. First, we let

\[
\Theta(t, s, \xi, x) = \Theta(t, s, \xi, x) - \psi(t, \xi, x), \quad (t, s) \in \Delta[0, T], \quad x, \xi \in \mathbb{R}^n.
\]

Then \( \Theta(\cdot, \cdot, \cdot, \cdot) \) is a solution of (5.1) if and only if \( \Theta(\cdot, \cdot, \cdot, \cdot) \) is a solution to the following:

\[
\begin{align*}
\Theta_s(t, s, \xi, x) + \frac{1}{2} \sigma(s, x)^\top \Theta_{xx}(t, s, \xi, x) \sigma(s, x) + \Theta_x(t, s, \xi, x) b(s, x) \\
+ \hat{g}(t, s, \xi, x, \Theta(s, s, x, x), \Theta_x(t, s, \xi, x) \sigma(s, x)) = 0, \quad (s, x) \in [t, T] \times \mathbb{R}^n,
\end{align*}
\]

\[
\Theta(t, T, \xi, x) = 0, \quad x \in \mathbb{R}^n,
\]

where

\[
\hat{g}(t, s, \xi, x, y, z) = g(t, s, \xi, x, y + \psi(t, \xi, x), \psi(t, \xi, x) \sigma(s, x) + z \sigma(s, x))
\]

\[
+ \frac{1}{2} \sigma(s, x)^\top \psi_{xx}(t, \xi, x) \sigma(s, x) + \psi_x(t, \xi, x) b(s, x) = 0, \quad (s, x) \in [t, T] \times \mathbb{R}^n,
\]
Hence, without loss of generality, we may consider (5.1) with $\psi(t, \xi, x) \equiv 0$.

**Step 2. The solution map of a parabolic PDE.** Let

$$\widehat{\Delta}[S, T] = \left( [0, S] \times [S, T] \right) \cup \Delta[S, T] = \left\{ (t, s) \in [0, T]^2 \mid 0 \leq t \vee S \leq s \leq T \right\}.$$  

(5.20)

See the figure below for illustration.

Let $\mathcal{X}[S, T]$ be the set of all measurable functions $\theta : \widehat{\Delta}[S, T] \times \mathbb{R}^n \to \mathbb{R}^m$ such that

\[
\|\theta\|_{\mathcal{X}[S, T]} \equiv \sup_{(t, \xi) \in [0, T] \times \mathbb{R}^n} \left( |\theta(t, \cdot, \xi, \cdot)|_{[\mathcal{V}_S, T] \times \mathbb{R}^n}^{(1+\alpha)} \right) \left[ |\theta(x, \cdot, \xi, \cdot)|_{[\mathcal{V}_S, T] \times \mathbb{R}^n} + |\theta_t(t, \cdot, \xi, \cdot)|_{[\mathcal{V}_S, T] \times \mathbb{R}^n}^{(0)} + |\theta_{\xi}(t, \cdot, \xi, \cdot)|_{[\mathcal{V}_S, T] \times \mathbb{R}^n}^{(0)} \right) \\
+ \left( |\theta_{x\xi}(t, \cdot, \xi, \cdot)|_{[\mathcal{V}_S, T] \times \mathbb{R}^n}^{(1+\alpha)} \right) + \left( |\theta(t, \cdot, \xi, \cdot)|_{[\mathcal{V}_S, T] \times \mathbb{R}^n}^{(0)} + |\theta_t(t, \cdot, \xi, \cdot)|_{[\mathcal{V}_S, T] \times \mathbb{R}^n}^{(0)} + |\theta_{\xi}(t, \cdot, \xi, \cdot)|_{[\mathcal{V}_S, T] \times \mathbb{R}^n}^{(0)} \right) \\
+ |\theta_{x\xi}(t, \cdot, \xi, \cdot)|_{[\mathcal{V}_S, T] \times \mathbb{R}^n}^{(0)} < \infty.
\]

(5.21)

Clearly, $\| \cdot \|_{\mathcal{X}[S, T]}$ is a norm under which $\mathcal{X}[S, T]$ is a Banach space.

Let $S \in [0, T)$ be fixed. For any $\theta(\cdot, \cdot, \cdot, \cdot) \in \mathcal{X}[S, T]$, denote

$$g(t, \tau, \xi, \eta) = g(t, \tau, \xi, \eta, \theta(\tau, \eta, \xi, \cdot), \theta(\tau, \eta, \xi, \cdot)) \sigma(\tau, \eta).$$

(5.22)

We claim that

$$|g(t, \cdot, \xi, \cdot)|_{[\mathcal{V}_S, T] \times \mathbb{R}^n}^{(0)} \leq K \left( 1 + \|\theta\|_{\mathcal{X}[S, T]} \right), \quad \forall (t, \xi) \in [S, T] \times \mathbb{R}^n.$$

(5.23)

Let us prove this. By boundedness of $g$, one has

$$|g(t, \tau, \xi, \eta, \theta(\tau, \eta, \xi, \cdot), \theta(\tau, \eta, \xi, \cdot))| \leq K.$$

(5.24)
Next, for $\tau, \tau' \in [t \lor S, T]$ and $\xi, \eta \in \mathbb{R}^n$, 
\[ |g(t, \tau, \xi, \eta, \theta(\tau, \xi, \eta), \theta_0(t, \tau, \xi, \eta) \sigma(\tau, \eta)) - g(t, \tau', \xi, \eta, \theta(\tau', \xi, \eta), \theta_0(t, \tau', \xi, \eta) \sigma(\tau', \eta))| \]
\[ \leq K \left( |\tau - \tau'| + |\theta(\tau, \xi, \eta) - \theta(\tau', \xi, \eta)| \right) \]
\[ \leq K \left( |\tau - \tau'| + \sup_{(\bar{\xi}, \bar{\eta}) \in [t \lor S, T] \times \mathbb{R}^n} |\theta_0(\bar{\xi}, \bar{\eta})||_{[t \lor S, T] \times \mathbb{R}^n} |\tau - \tau'| \right) \]
\[ \leq K \left( 1 + \|\theta\|_{[t \lor S, T]} \right) |\tau - \tau'|. \]  
(5.25)

Likewise, for $\tau \in [t \lor S, T]$ and $\xi, \eta, \eta' \in \mathbb{R}^n$, 
\[ |g(t, \tau, \xi, \eta, \theta(\tau, \xi, \eta), \theta_0(t, \tau, \xi, \eta) \sigma(\tau, \eta))| \]
\[ \leq K \left( \|\eta - \eta'\| + |\theta(\tau, \xi, \eta) - \theta(\tau, \xi, \eta')| \right) \]
\[ \leq K \left( \|\eta - \eta'\| + \sup_{(\bar{\xi}, \bar{\eta}) \in [t \lor S, T] \times \mathbb{R}^n} |\theta_0(\bar{\xi}, \bar{\eta})||_{[t \lor S, T] \times \mathbb{R}^n} \|\eta - \eta'\| \right) \]
\[ \leq K \left( 1 + \|\theta\|_{[t \lor S, T]} \right) \|\eta - \eta'\|. \]
(5.26)

Combining (5.24)–(5.26), we obtain (5.23). Now, for any $\theta(\cdot, \cdot, \cdot, \cdot) \in \mathcal{B}_r[S, T]$, consider the following linear parabolic system, parametrized by $(\xi, \eta) \in [0, T) \times \mathbb{R}^n$: 
\[
\begin{cases}
\Theta_s(s, \xi, x) + \frac{1}{2} \sigma(s, x)^T \Theta_x(x, s, \xi, x) \sigma(s, x) + \Theta_x(s, t, \xi, x) b(s, x) \\
\quad + g(t, s, \xi, x, \theta(s, s, x, x), \theta_x(s, t, \xi, x) \sigma(s, x)) = 0, \quad (s, x) \in [t \lor S, T] \times \mathbb{R}^n, \\
\Theta(t, t, \xi, x) = 0, \quad x \in \mathbb{R}^n.
\end{cases}
\]
(5.27)

Then the corresponding solution $\Theta(t, \cdot, \xi, \cdot)$ uniquely exists and the following holds: 
\[
\Theta(t, t, \xi, x) = \int_t^T \int_{\mathbb{R}^n} G(s, t; \tau, \eta) g(t, \tau, \eta, \theta(\tau, \xi, \eta), \theta_0(t, \tau, \xi, \eta) \sigma(\tau, \eta)) d\eta d\tau.
\]
(5.28)

Due to (5.23), we have $\Theta(t, \cdot, \xi, \cdot) \in C^{2+\alpha}([t \lor S, T] \times \mathbb{R}^n)$. On the other hand, by (5.7) and (5.23), we have 
\[
|\Theta(t, \cdot, \xi, \cdot)||_{[t \lor S, T] \times \mathbb{R}^n} \leq K(T - S)^{\frac{\alpha}{2}} |g(t, \cdot, \xi, \cdot)||_{[t \lor S, T] \times \mathbb{R}^n} \leq K(T - S)^{\frac{\alpha}{2}} (1 + \|\theta\|_{[t \lor S, T]}).
\]
(5.29)

Next, 
\[
\Theta_1(t, s, \xi, x) = \int_s^T \int_{\mathbb{R}^n} G(s, x; \tau, \eta) g_1(t, \tau, \xi, \eta, \theta(\tau, \xi, \eta), \theta_0(t, \tau, \xi, \eta) \sigma(\tau, \eta)) d\eta d\tau \\
+ \int_s^T \int_{\mathbb{R}^n} G(s, x; \tau, \eta) g_2(t, \tau, \xi, \eta, \theta(\tau, \xi, \eta), \theta_0(t, \tau, \xi, \eta) \sigma(\tau, \eta)) \theta_{\eta t}(t, \tau, \xi, \eta) \sigma(\tau, \eta) d\eta d\tau,
\]

26
Moreover, we shrink

Likewise, we have

Consequently, it holds that

where

Similarly,

Thus,

Using (5.5), one has

where

Similarly,

Consequently, it holds that

Likewise, we have

Then combining (5.29) with (5.30)–(5.31), one obtains

with \( K > 0 \) being an absolute constant. Hence, we have defined a map \( S : \mathcal{X}[S, T] \to \mathcal{X}[S, T] \) by

Moreover, we shrink \( T - S > 0 \) (if necessary) so that \( K(T - S)^{\frac{1}{2}} \leq \frac{4}{3} \). Then for any \( M \geq 1 \),

Thus, \( S \) maps a ball in \( \mathcal{X}[S, T] \), centered at 0 with radius \( M \) to itself.

**Step 3. Contraction of the solution map.** Let \( \theta, \tilde{\theta} \in \mathcal{X}[S, T] \) such that

\[ \|\theta\|_{\mathcal{X}[S, T]}, \|\tilde{\theta}\|_{\mathcal{X}[S, T]} \leq M, \]
with $S \in [0,T)$ and $M > 0$ obtained as above. Let

$$
\Theta(t,s,\xi,x) = S[\theta](t,s,\xi,x), \quad \tilde{\Theta}(t,s,\xi,x) = S[\tilde{\theta}](t,s,\xi,x),
$$

$$
v(t,s,\xi,x) = \Theta(t,s,\xi,x) - \tilde{\Theta}(t,s,\xi,x).
$$

Then $v(t,\cdot,\xi,\cdot)$ satisfies the following:

$$
\begin{cases}
\displaystyle v_s(t,s,\xi,x) + \frac{1}{2} \sigma(s,x) \sigma(s,x) + \Theta_x(t,s,\xi,x) b(s,x) + f(t,s,\xi,x) = 0, \\
\quad (s,x) \in [t \vee S,T] \times \mathbb{R}^n,
\end{cases}
$$

(5.33)

with

$$
f(t,s,\xi,x) = g(t,s,\xi,x,\theta(s,x,y,z),\tilde{\theta}(s,x,y,z)) - g(t,s,\xi,x,\tilde{\theta}(s,x,y,z),\tilde{\theta}(s,x,y,z)).
$$

Consequently,

$$
v(t,s,\xi,x) = \int_s^T \int_{\mathbb{R}^n} G(s,x;\tau,\eta) f(t,\tau,\xi,\eta) d\eta d\tau.
$$

(5.34)

Similar to the proof of (5.23), we can show that

$$
|f(t,\cdot,\xi,\cdot)(\cdot)|_{[t\vee S,T] \times \mathbb{R}^n} \leq K \|\theta - \tilde{\theta}\|_{X[S,T]}, \quad \forall (t,\xi) \in [S,T] \times \mathbb{R}^n.
$$

(5.35)

Then one has

$$
\|\Theta - \tilde{\Theta}\|_{X[S,T]} \leq K(T-S)^{\frac{2}{\alpha}} \|\theta - \tilde{\theta}\|_{X[S,T]},
$$

(5.36)

for some absolute constant $K > 0$. Hence, by choosing $\delta = T - S > 0$ small, we obtain a contraction mapping $S$ on $X[T - \delta,T]$. Consequently, $S$ has a unique fixed point which is the solution $\Theta(t,\cdot,\cdot,\cdot)$ on $X[T - \delta,T] \times \mathbb{R}^n \times \mathbb{R}^n$, which is rewritten here:

$$
\begin{cases}
\Theta_s(t,s,\xi,x) + \frac{1}{2} \sigma(s,x) \sigma(s,x) + \Theta_x(t,s,\xi,x) b(s,x) \\
\quad + g(t,s,\xi,x,\Theta(s,x,y,z),\Theta_x(t,s,\xi,x) \sigma(s,x)) = 0, \\
\Theta(t,T,\xi,x) = 0, \quad (t,\xi,\cdot) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^n.
\end{cases}
$$

(5.37)

Note that $(t,\xi) \mapsto g(t,s,\xi,x,y,z)$ is assumed to be continuously differentiable with bounded derivatives. Therefore, by a standard argument, we know that

$$
|\Theta_t(t,\cdot,\cdot,\cdot)|_{[t\vee(T-\delta),T] \times \mathbb{R}^n} + |\Theta_x(t,\cdot,\cdot,\cdot)|_{[t\vee(T-\delta),T] \times \mathbb{R}^n} < \infty.
$$

Now, we denote

$$
\tilde{\psi}(t,\xi,x) = \Theta(t,T-\delta,\xi,x), \quad (t,\xi,x) \in [0,T-\delta] \times \mathbb{R}^n \times \mathbb{R}^n,
$$

and consider the following equation:

$$
\begin{cases}
\Theta_s(t,s,\xi,x) + \frac{1}{2} \sigma(s,x) \sigma(s,x) + \Theta_x(t,s,\xi,x) b(s,x) \\
\quad + g(t,s,\xi,x,\Theta(s,x,y,z),\Theta_x(t,s,\xi,x) \sigma(s,x)) = 0, \\
\Theta(t,T-\delta,\xi,x) = \tilde{\psi}(t,\xi,x), \quad x \in \mathbb{R}^n.
\end{cases}
$$

(5.38)

Then, we may repeat the above procedure, to get a unique solution $\Theta(t,\cdot,\cdot,\cdot)$ on $X[T - 2\delta,T - \delta] \times \mathbb{R}^n \times \mathbb{R}^n$. By continuing such a procedure, we obtain the existence and unique solution $\Theta(t,\cdot,\cdot,\cdot)$ to the representation PDE (5.1), and (5.16) holds.
Step 4. Stability estimates. Let \((\hat{\psi}, \hat{g})\) be another pair of maps such that, together with \(b\) and \(\sigma\), satisfy (H3) as well. Let \(\hat{\Theta}\) be the corresponding solution. Let
\[
v(t, s, \xi, x) = \Theta(t, s, \xi, x) - \hat{\Theta}(t, s, \xi, x).
\]
Then \(v(\cdot, \cdot, \xi, \cdot)\) satisfies the following:
\[
\begin{cases}
  v_s(t, s, \xi, x) + \frac{1}{2} \sigma(s, x)\nabla v(t, s, \xi, x)\sigma(s, x) + v_x(t, s, \xi, x)b(s, x) + f(t, s, \xi, x) + \hat{f}(t, s, \xi, x) = 0, \\
  v(t, T, \xi, x) = \varphi(t, \xi, x), \quad x \in \mathbb{R}^n,
\end{cases}
\]
with
\[
f(t, s, \xi, x) = \hat{g}(t, s, \xi, x), \Theta_x(t, s, \xi, x) = \hat{g}(t, s, \xi, x), \hat{\Theta}_x(t, s, \xi, x) - \hat{g}(t, s, \xi, x), \\
\hat{f}(t, s, \xi, x) = \hat{g}(t, s, \xi, x), \Theta_x(t, s, \xi, x) = \hat{g}(t, s, \xi, x), \hat{\Theta}_x(t, s, \xi, x) \sigma(s, x),
\]
\[
\varphi(t, \xi, x) = \psi(t, \xi, x) - \hat{\psi}(t, \xi, x).
\]
Then
\[
v(t, s, \xi, x) = \int_{\mathbb{R}^n} G(s, x; T, \eta)\varphi(t, \xi, \eta)d\eta + \int_t^T \int_{\mathbb{R}^n} G(s, x; \tau, \eta)\hat{f}(t, \tau, \xi, \eta)d\eta d\tau + \int_t^T \int_{\mathbb{R}^n} G(s, x; \tau, \eta)f(t, \tau, \xi, \eta)d\eta d\tau.
\]
From (5.7), one has
\[
\left|\int_{\mathbb{R}^n} G(\cdot, \cdot; T, \eta)\varphi(t, \xi, \eta)d\eta + \int_t^T \int_{\mathbb{R}^n} G(\cdot, \cdot; \tau, \eta)\hat{f}(t, \tau, \xi, \eta)d\eta d\tau\right|^{2+\alpha}_{[t, T] \times \mathbb{R}^n} \leq K \left(\|\varphi(t, \xi, \cdot)\|_{L^2[\mathbb{R}^n]}^{2+\alpha} + \|\hat{f}(t, \cdot, \xi, \cdot)\|_{L^2[\mathbb{R}^n]}^{(\alpha)}\right).
\]
Also, by Step 3 above, we have
\[
\left\|\int_t^T \int_{\mathbb{R}^n} G(\cdot, \cdot; \tau, \eta)f(t, \cdot, \tau, \cdot, \eta)d\eta d\tau\right\|_{L^2[\mathbb{R}^n]} \leq K (1 + M)(T - S)^{\frac{\alpha}{2}} \|v\|_{L^2[\mathbb{R}^n]}.\]
Hence, for \(T - S > 0\) small, we obtain
\[
\|v\|_{L^2[\mathbb{R}^n]} \leq K \sup_{(t, \xi) \in [S, T] \times \mathbb{R}^n} \left(\|\varphi(t, \xi, \cdot)\|_{L^2[\mathbb{R}^n]}^{2+\alpha} + \|\hat{f}(t, \cdot, \xi, \cdot)\|_{L^2[\mathbb{R}^n]}^{(\alpha)}\right).
\]
Repeating the same argument, we obtain
\[
\|v\|_{L^2[\mathbb{R}^n]} \leq K \sup_{(t, \xi) \in [0, T] \times \mathbb{R}^n} \left(\|\varphi(t, \xi, \cdot)\|_{L^2[\mathbb{R}^n]}^{2+\alpha} + \|\hat{f}(t, \cdot, \xi, \cdot)\|_{L^2[\mathbb{R}^n]}^{(\alpha)}\right).
\]
Then (5.36) follows.

5.3 The second representation PDE

We now look at the second representation PDE. Similar to the previous subsection, without loss of generality, we again assume that \(\psi(t, \xi, x) = 0\). Thus, we consider the following family of parabolic systems
(parameterized by \((t, \xi) \in (S, T) \times \mathbb{R}^n\)):
\[
\begin{aligned}
\Gamma(t, s, x) + \frac{1}{2} \sigma(s, x) \Gamma_x(t, s, x) \sigma(s, x) + \Gamma_x(t, s, x) b(s, x) &= 0, \quad (s, x) \in [0, t] \times \mathbb{R}^n, \\
\Theta(t, s, x, t,\xi) + \frac{1}{2} \sigma(s, x) \Theta_x(t, s, x, t,\xi) \sigma(s, x) + \Theta_x(t, s, x, t,\xi) b(s, x) \\
&\quad + g(t, s, \xi, x, \Theta(t, s, x, t,\xi), \Theta_x(t, s, x, t,\xi) \sigma(s, x), \Gamma(t, s, \xi) \sigma(s, x)) = 0, \quad (s, x) \in [t, T] \times \mathbb{R}^n, \\
\Gamma(t, t, x) &= \Theta(t, t, x), \quad x \in \mathbb{R}^n, \\
\Theta(t, T, x) &= 0,
\end{aligned}
\tag{5.41}
\]

Note that for any given \((t, \xi) \in (0, T) \times \mathbb{R}^n\), the equation for \(\Gamma(t, t, \cdot)\) is to be solved on \([0, t]\) and the equation for \(\Theta(t, \cdot, \xi, \cdot)\) is to be solved on \([t, T]\). The coupling appears at two places: \(\Gamma(t, s, \xi)\) (with \(0 \leq t \leq s\)) appears in the equation for \(\Theta(t, t, \xi, \cdot)\) and \(\Theta(t, t, x)\) appears as the terminal value for \(\Gamma(t, s, x)\) at \(s = t\).

\[\text{See the above figure for the domains in which } \Theta(t, x, \xi, x) \text{ and } \Gamma(t, s, \xi) \text{ are defined. Let us make some observations. Suppose } (\Gamma(\cdot, \cdot, \cdot), \Theta(\cdot, \cdot, \cdot, \cdot)) \text{ is a solution to the above (5.41). Then}
\]

\[
\Gamma(t, s, x) = \int_{\mathbb{R}^n} G(s, x; t, \eta) \Theta(t, t, \eta, \eta) d\eta, \quad 0 \leq s \leq t \leq T, \quad x \in \mathbb{R}^n,
\tag{5.42}
\]

and thus,
\[
\Gamma(t, s, x) = \int_{\mathbb{R}^n} G(t, \xi; \tau, \eta) \Theta(\tau, \tau, \eta, \eta) d\eta, \quad 0 \leq t \leq \tau \leq T, \quad \xi \in \mathbb{R}^n.
\]

On the other hand,
\[
\Theta(t, s, x, t,\xi) = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} G(s, x; t, \eta) g(t, \tau, \xi, \eta, \Theta(\tau, t, \eta, \eta), \Theta_y(t, \tau, x, \xi, \eta) \sigma(\tau, \eta), \Gamma(\tau, \sigma(\tau, \eta)) d\eta d\tau
\]
\[
= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} G(s, x; t, \eta) g(t, \tau, \xi, \eta, \Theta(\tau, t, \eta, \eta), \Theta_y(t, \tau, x, \xi, \eta) \sigma(\tau, \eta), \Gamma(\tau, \sigma(\tau, \eta)) d\eta d\tau
\]
\[
= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} G(s, x; t, \eta) \Theta(\tau, \tau, \eta, \eta) d\eta d\tau, \quad (t, s, x, \xi) \in \Delta[0, T] \times \mathbb{R}^n \times \mathbb{R}^n.
\tag{5.43}
\]

The above tells us that if (5.41) admits a classical solution \((\Gamma, \Theta)\), then \(\Theta\) must be a solution to the above nonlinear integral equation (5.43). Conversely, if nonlinear integral equation (5.43) admits a smooth solution \(\Theta\), by defining \(\Gamma\) as (5.42), we have a solution \((\Gamma, \Theta)\) to the second representation PDE (5.41). Hence, we could introduce the following definition.
**Definition 5.3.** A pair of functions \((\Gamma, \Theta)\) is called a *mild solution* of (5.41) if \(\Theta\) is a solution to the integral equation (5.28) and \(\Gamma\) is defined by (5.42).

Now, we introduce the following hypothesis for system (5.41).

**(H4)** The maps \(b(s, x), \sigma(s, x)\), and \(g(t, s, \xi, x, y, z, \zeta)\) are bounded and have all needed orders of bounded derivatives. Moreover (5.15) holds for some \(\bar{\sigma} > 0.0\).

The main result of this subsection is the following theorem.

**Theorem 5.4.** Let (H4) hold. Then (5.41) admits a unique mild solution \((\Gamma, \Theta)\).

**Proof.** Let \(1 < p < 2, S \in [0, T)\) and recall \(\hat{\Delta}[S, T]\) defined by (5.20). Let \(\mathcal{Y}[S, T]\) be the set of all functions \(\theta : \hat{\Delta}[S, T] \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^m\) such that

\[
\|\theta\|_{\mathcal{Y}[S, T]} = \sup_{t \in [0, T]} \left( \int_0^T \frac{p}{s} \frac{d}{ds} \right)^{\frac{p}{2}} + \sup_{(t, s) \in \hat{\Delta}[S, T]} |\theta(t, s, \xi, x)| < \infty.
\]

Clearly, \(\| \cdot \|_{\mathcal{Y}[S, T]}\) is a norm under which \(\mathcal{Y}[S, T]\) is a Banach space.

For any \(\theta \in \mathcal{Y}[S, T]\), define

\[
\mathcal{S}[\theta](t, s, \xi, x) = \int_s^t \int_{\mathbb{R}^n} G(s, x; \xi, \eta)g(t, s, \xi, \eta, \theta(t, s, \xi, \eta), \theta_\xi(t, s, \xi, \eta)) \sigma(t, \eta) d\eta d\tau,
\]

(5.44)

Note that

\[
\mathcal{S}[\theta](0, s, \xi, x) = \int_s^T \int_{\mathbb{R}^n} G(s, x; \xi, \eta)g(t, s, \xi, \eta, 0, 0, 0) d\eta d\tau, \quad (t, s, \xi, x) \in \hat{\Delta}[S, T] \times \mathbb{R}^n \times \mathbb{R}^n.
\]

Thus,

\[
\mathcal{S}[\theta](s, s, \xi, x) = \int_s^T \int_{\mathbb{R}^n} G(s, x; \xi, \eta)g(s, s, \xi, \eta, 0, 0, 0) d\eta d\tau, \quad (s, s, \xi, x) \in [S, T] \times \mathbb{R}^n,
\]

and

\[
\mathcal{S}[\theta](t, s, \xi, x) = \int_s^T \int_{\mathbb{R}^n} G(s, x; \xi, \eta)g(t, s, \xi, \eta, 0, 0, 0) d\eta d\tau, \quad (t, s, \xi, x) \in \hat{\Delta}[S, T] \times \mathbb{R}^n \times \mathbb{R}^n.
\]

Then

\[
|\mathcal{S}[\theta](t, s, \xi, x)| \leq K \int_s^T \int_{\mathbb{R}^n} \frac{1}{(t-s)^\frac{n-1}{2}} e^{-\lambda (t-s)^2} d\eta d\tau = K \int_s^T \int_{\mathbb{R}^n} e^{-\lambda \tau^2} dz d\tau \leq K.
\]

Also,

\[
|\mathcal{S}[\theta](t, s, \xi, x)| \leq K \int_s^T \int_{\mathbb{R}^n} \frac{1}{(t-s)^\frac{n-1}{2}} e^{-\lambda (t-s)^2} d\eta d\tau = K \int_s^T \int_{\mathbb{R}^n} \frac{1}{(t-s)^\frac{n-1}{2}} e^{-\lambda \tau^2} dz d\tau \leq K.
\]

Hence, \(\mathcal{S}[\theta] \in \mathcal{Y}[S, T]\).

Next, let \(\theta, \hat{\theta} \in \mathcal{Y}[S, T]\). We estimate the following:

\[
|\mathcal{S}[\theta](t, s, \xi, x) - \mathcal{S}[\hat{\theta}](t, s, \xi, x)| \leq K \int_s^T \int_{\mathbb{R}^n} |G(s, x; \xi, \eta)||\theta(t, s, \xi, \eta) - \hat{\theta}(t, s, \xi, \eta)|
\]

\[
+ |\theta_\xi(t, s, \xi, \eta) - \hat{\theta}_\xi(t, s, \xi, \eta)| + \int_{\mathbb{R}^n} |G_\xi(t, s, \xi, \eta)||\theta(t, s, \xi, \eta) - \hat{\theta}(t, s, \xi, \eta)| d\eta d\tau.
\]
\begin{align*}
= K \int_s^T \int_{\mathbb{R}^n} |G(s, x; \tau, \eta)| \left| \theta_\eta(t, \tau, \xi, \eta) - \hat{\theta}_\eta(t, \tau, \xi, \eta) \right| d\eta d\tau \\
+ K \int_s^T \int_{\mathbb{R}^n} \left( |G(s, x; \tau, \eta)| + |G_\xi(t, \xi; \tau, \eta)| \int_{\mathbb{R}^n} |G(s, x; \tau, \eta)| d\eta \right) |\theta(t, \tau, \eta, \eta) - \hat{\theta}(t, \tau, \eta, \eta)| d\eta d\tau \\
\leq K \int_s^T \int_{\mathbb{R}^n} \left( e^{-\lambda (s-t) \eta^2} \right) \sup_{\eta, \eta \in \mathbb{R}^n} |\theta_\eta(t, \tau, \xi, \eta) - \hat{\theta}_\eta(t, \tau, \xi, \eta)| d\tau \\
+ K \int_s^T \int_{\mathbb{R}^n} \left( e^{-\lambda (s-t) \eta^2} \right) \int_{\mathbb{R}^n} \left( e^{-\lambda (s-t) \eta^2} \right) d\eta d\tau \sup_{\eta, \eta \in [s, T] \times \mathbb{R}^n} |\theta(t, \tau, \eta, \eta) - \hat{\theta}(t, \tau, \eta, \eta)| \\
\leq K \int_s^T \int_{\mathbb{R}^n} \left( e^{-\lambda \eta^2} \right) \sup_{\xi, \eta \in \mathbb{R}^n} |\theta_\xi(t, \tau, \xi, x) - \hat{\theta}_\xi(t, \tau, \xi, x)| d\tau \\
+ K \int_s^T \left( \int_{\mathbb{R}^n} \left( e^{-\lambda \eta^2} \right) \left( \int_{\mathbb{R}^n} \left( e^{-\lambda \eta^2} \right) \right) d\eta d\tau \right) \sup_{\eta, \eta \in [s, T] \times \mathbb{R}^n} |\theta(t, \tau, \eta, \eta) - \hat{\theta}(t, \tau, \eta, \eta)| \\
\leq K \left( T - s \right)^{\frac{1}{2}} \left( \int_s^T \sup_{\xi, \eta \in \mathbb{R}^n} \left| \theta_\xi(t, \tau, \xi, x) - \hat{\theta}_\xi(t, \tau, \xi, x) \right|^p d\tau \right)^{\frac{1}{p}} \\
+ K \left( T - s \right)^{\frac{1}{2}} \sup_{\tau, \eta \in \mathbb{R}^n} \left| \theta(t, \tau, \xi, x) - \hat{\theta}(t, \tau, \xi, x) \right|. \\
\end{align*}

This leads to (note $e^{-1} < \frac{1}{2}$)

$$
\sup_{\tau, \eta \in \mathbb{R}^n} \left| \left[ S \theta \right](t, s, \xi, x) - \left[ \hat{\theta} \right](t, s, \xi, x) \right| \leq K \left( T - S \right)^{\frac{1}{p-1}} \left\| \theta - \hat{\theta} \right\|_{[S, T]}.
$$

(5.45)

Next,

$$
\left| \left[ S \theta \right]_z(t, s, \xi, x) - \left[ \hat{\theta} \right]_z(t, s, \xi, x) \right| \leq K \int_s^T \int_{\mathbb{R}^n} \left| G_x(s, x; \tau, \eta) \right| \left( |\theta_\eta(t, \tau, \xi, \eta) - \hat{\theta}_\eta(t, \tau, \xi, \eta)| + \left| G_\xi(t, \xi; \tau, \eta) \right| \left| \theta(t, \tau, \xi, \eta) - \hat{\theta}(t, \tau, \xi, \eta) \right| \right) d\eta d\tau
$$

$$
= K \int_s^T \int_{\mathbb{R}^n} \left| G_x(s, x; \tau, \eta) \right| \left| \theta_\eta(t, \tau, \xi, \eta) - \hat{\theta}_\eta(t, \tau, \xi, \eta) \right| d\eta d\tau \\
+ K \int_s^T \int_{\mathbb{R}^n} \left( |G_\xi(t, \xi; \tau, \eta)| \int_{\mathbb{R}^n} \left| G_x(s, x; \tau, \eta) \right| d\eta \right) |\theta(t, \tau, \xi, \eta) - \hat{\theta}(t, \tau, \xi, \eta)| d\eta d\tau \\
\leq K \int_s^T \int_{\mathbb{R}^n} \left( e^{-\lambda (s-t) \eta^2} \right) \sup_{\eta, \eta \in \mathbb{R}^n} |\theta_\eta(t, \tau, \xi, \eta) - \hat{\theta}_\eta(t, \tau, \xi, \eta)| d\tau \\
+ K \int_s^T \int_{\mathbb{R}^n} \left( e^{-\lambda (s-t) \eta^2} \right) \left( \int_{\mathbb{R}^n} \left( e^{-\lambda (s-t) \eta^2} \right) d\eta \right) d\tau \sup_{\eta, \eta \in [s, T] \times \mathbb{R}^n} |\theta(t, \tau, \eta, \eta) - \hat{\theta}(t, \tau, \eta, \eta)| \\
\leq K \int_s^T \int_{\mathbb{R}^n} \frac{1}{(s-t)^{\frac{1}{2}}} \sup_{\eta, \eta \in \mathbb{R}^n} \left| \theta_\eta(t, \tau, \xi, x) - \hat{\theta}_\eta(t, \tau, \xi, x) \right| d\tau \\
+ K \int_s^T \left( \int_{\mathbb{R}^n} \frac{1}{(s-t)^{\frac{1}{2}}} \left( \int_{\mathbb{R}^n} \frac{1}{(s-t)^{\frac{1}{2}}} \right) d\tau \right) \sup_{\eta, \eta \in [s, T] \times \mathbb{R}^n} \left| \theta(t, \tau, \xi, \eta) - \hat{\theta}(t, \tau, \xi, \eta) \right|.
$$

32
Then noting $1 < p < 2$, by Young’s inequality, we have

\[
\left( \int_{t \vee S}^{T} \sup_{\xi, x \in \mathbb{R}^n} |S[\theta]_x(t, s, \xi, x) - S[\tilde{\theta}]_x(t, s, \xi, x)|^p \, ds \right)^{\frac{1}{p}} \\
\leq K(T - S)^{\frac{1}{2}} \left( \left( \int_{t \vee S}^{T} \sup_{\xi, x \in \mathbb{R}^n} |\theta_x(t, \tau, \xi, x) - \tilde{\theta}_x(t, \tau, \xi, x)|^p \, d\tau \right)^{\frac{1}{p}} + \sup_{(t, s) \in \hat{\Delta}[S, T]} |\theta(t, s, \xi, x) - \tilde{\theta}(t, s, \xi, x)| \right) \\
\leq K(T - S)^{\frac{1}{2}} \|\theta - \tilde{\theta}\|_{\mathcal{Y}[S, T]}.
\]

Combining the above, we obtain

\[
\|S[\theta] - S[\tilde{\theta}]\|_{\mathcal{Y}[S, T]} \leq K(T - S)^{\frac{p-1}{p}} \|\theta - \tilde{\theta}\|_{\mathcal{Y}[S, T]},
\]

By taking $\tilde{\theta} = 0$, we see that

\[
\|S[\theta]\|_{\mathcal{Y}[S, T]} \leq \|S[0]\|_{\mathcal{Y}[S, T]} + K(T - S)^{\frac{p-1}{p}} \|\theta\|_{\mathcal{Y}[S, T]}, \quad \forall \theta \in \mathcal{Y}[S, T].
\]

Consequently, $S : \mathcal{Y}[S, T] \to \mathcal{Y}[S, T]$ and it is a contraction when $\delta = T - S > 0$ is small. Hence, it admits a unique fixed point on $[T - \delta, T]$, which gives a solution to (5.28) on $\hat{\Delta}[T - \delta, T] \times \mathbb{R}^{2n}$. By repeating the same argument, we will be able to get a unique solution of (5.28) on $\Delta[0, T] \times \mathbb{R}^{2n}$. Then we define $\Gamma$ by (5.42). This gives the existence of a mild solution $(\Gamma, \Theta)$ of (5.41).

The argument used to established the contractiveness of the solution map $S$ also gives the uniqueness of the mild solution.

\[\square\]

6 Concluding Remarks

In this paper, we have derived the representations of adapted solutions of Type-I BSVIEs and adapted M-solutions of Type-II BSVIEs in terms of the solution to forward SDEs via the solutions of representation PDEs. For Type-I BSVIEs, the well-posedness of representation PDE is established in the classical sense, and for Type-II BSVIEs, the well-posedness of representation PDE is established in the mild solution. It remains open at the moment whether the representation PDE for Type-II BSVIEs admits a unique classical solution, which we believe it to be true, under certain conditions.

On the other hand, our results could also be regarded as Feynman-Kac formula, from which the solutions to the PDE systems of forms (1.21) and (1.23) can be represented by the solutions to the corresponding BSVIEs.

It is worthy of pointing out that, to our best knowledge, representation PDEs of form (1.21) appeared for the first time in the study of time-inconsistent optimal control problems ([41], see also [35, 25]). This indicates that there should be some intrinsic relationship between BSVIEs and time-inconsistent optimal control problems. We hope to explore that in our future publications.

References

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

[2] R. Bellman, R. Kalaba and G. M. Wing, *Invariant imbedding and the reduction of two-point boundary value problems to initial value problems*, Proc. Nat. Acad. Sci. U.S.A., 46 (1960), 1646–1649.
[3] R. Bellman and G. M. Wing, An Introduction to Invariant Imbedding, Wiley-Interscience, New York, 1975.

[4] J.-M. Bismut, Théorie Probabiliste du Contrôle des Diffusions, Mem. Amer. Math. Soc. 176, Providence, Rhode Island, 1973.

[5] Ph. Priand, D. Delyon, Y. Hu, E. Pardoux and L. Stoica, $L^p$ solutions of backward stochastic differential equations, Stoch. Proc. Appl., 108 (2003), 109–129.

[6] Ph. Briand and Y. Hu, BSDE with quadratic growth and unbounded terminal value, Probab. Theory Rel. Fields, 136 (2006), 604–618.

[7] F. Delbaen, Y. Hu and X. Bao, Backward SDEs with superquadratic growth, Probab. Theory Rel. Fields, 150 (2011), 145–192.

[8] J. Douglas, Jr., J. Ma and P. Protter, Numerical methods for forward-backward stochastic differential equations, Ann. Appl. Probab., 6 (1996), 940–968.

[9] D. Duffie and L. Epstein, Stochastic differential utility, Econometrica, 60 (1992), 353–394.

[10] I. Ekren, C. Keller, N. Touzi and J. Zhang, On viscosity solutions of path dependent PDEs, Ann. Probab., 42 (2014), 204–236.

[11] N. El Karoui, S. Peng, and M. C. Quenez, Backward stochastic differential equation in finance, Math. Finance, 7 (1997), 1–71.

[12] A. Friedman, Partial Differential Equations of Parabolic Type, Prentice-Hall, Englewood Cliffs, N.J., 1964.

[13] Y. Hu and J. Ma, Nonlinear Feynman-Kac formula and discrete-functional-type BSDEs with continuous coefficients, Stoch. Proc. Appl., 112 (2004), 23–51.

[14] Y. Hu and S. Peng, Solution of forward-backward stochastic differential equations, Probab. Theory Rel. Fields, 103 (1995), 273–283.

[15] I. Karatzas and S. E. Shreve, Brownian Motion and Stochastic Calculus, Springer-Verlag, Berlin, 1988.

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

[17] O. A. Ladyzenskaja, V. A. Solonnikov, and N. N. Ural’ceva, Linear and Quasi-linear Equations of Parabolic Type, AMS, Providence, R.I., 1968.

[18] J. Lin, Adapted solution of a backward stochastic nonlinear Volterra integral equation, Stoch. Anal. Appl., 20 (2002), 165–183.

[19] J. Ma, P. Protter, and J. Yong, Solving forward-backward stochastic differential equations explicitly — a four step scheme, Prob. Theory Rel. Fields, 98 (1994), 339–359.

[20] J. Ma, Z. Wu, D. Zhang, J. Zhang, On well-posedness of forward-backward SDEs — a unified approach, Ann. Appl. Probab., 25 (2015), 2168–2214.

[21] J. Ma and J. Yong, Solvability of forward backward SDEs and the nodal set of Hamilton-Jaccobi-Bellman Equations, Chin. Ann. Math., 16B (1995), 279–298.
[22] J. Ma and J. Yong, Forward-backward stochastic differential equations and their applications, Lecture Notes in Mathematics, 1702. Springer-Verlag, Berlin, 1999.

[23] J. Ma, J. Yong and Y. Zhao, Four step scheme for general Markovian forward-backward SDEs, J. Syst. Sci. Complex., 23 (2010), 546–571.

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

[25] H. Mei and J. Yong, Equilibrium strategies for time-inconsistent stochastic switching systems, arXiv:1712.09505v1[math.OC]27 Dec 2017.

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

[27] S. Peng, A nonlinear Feynman-Kac formula and applications, Control Theory, Stochastic Analysis and Applications, S. Chen and J. Yong, eds, World Scientific, Singapore, 1991, 173–184.

[28] S. Peng, Nonlinear expectations, nonlinear evaluations and risk measures, Stochastic Methods in Finance, 165–253, Lecture Notes in Math., 1856, Springer, Berlin, 2004.

[29] S. Peng, Backward stochastic differential equation, nonlinear expectation and their applications, Proceedings of the International Congress of Mathematicians, Volume I, 393–432, Hindustan Book Agency, New Delhi, 2010.

[30] H. M. Soner, N. Touzi and J. Zhang, Wellposedness of second order backward SDEs, Probab. Theory Rel. Fields, 153 (2012), 149–190.

[31] Y. Shi, T. Wang and J. Yong, Mean-field backward stochastic Volterra integral equations, Discrete Continuous Dynamic Systems, 18 (2013), 1929–1967.

[32] Y. Shi, T. Wang and J. Yong, Optimal control problems of forward-backward stochastic Volterra integral equations, Math. Control Rel. Fields, 5 (2015), 613–649.

[33] T. Wang and J. Yong, Comparison theorems for backward stochastic integral equations, Stoch. Proc. Appl., 125 (2015), 1756–1798.

[34] Y. Wang, A numerical scheme for BSVIEs, arXiv:1605.04865v1 [math.NA] 16 May 2016.

[35] Q. Wei, J. Yong and Z. Yu, Time-inconsistent recursive stochastic optimal control problems, SIAM J. Control Optim., 55 (2017), 4156–4201.

[36] J. Yong, Finding adapted solution of forward-backward stochastic differential equations — method of continuation, Probab. Theory Rel. Fields, 107 (1997) 537–572.

[37] J. Yong, Backward stochastic Volterra integral equations and some related problems, Stoch. Proc. Appl, 116 (2006), 779–795.

[38] J. Yong, Continuous-time dynamic risk measures by backward stochastic Volterra integral equations, Appl. Anal, 86 (2007), 1429–1442.

[39] J. Yong, Well-posedness and regularity of backward stochastic Volterra integral equations, Probab. Theory Rel. Fields, 142 (2008), 21–77.
[40] J. Yong, Forward-backward stochastic differential equations with mixed initial-terminal conditions, Trans. Amer. Math. Soc., 362 (2010), 1047–1096.

[41] J. Yong, Time-inconsistent optimal control problems and the equilibrium HJB equation, Math. Control Rel. Fields, 2 (2012), 271–329.

[42] J. Yong, Representation of adapted solutions to backward stochastic Volterra integral equations, Scientia Sinica Mathematica, 47 (2017), 305–345 (in Chinese).

[43] J. Zhang, A numerical scheme for BSDEs, Ann Appl. Probab., 14 (2004), 459–488.