

Path dependent Feynman–Kac formula for forward backward stochastic Volterra integral equations

Hanxiao Wang^{1,a}, Jiongmin Yong^{2,b} and Jianfeng Zhang^{3,c}

¹*Department of Mathematics, National University of Singapore, Singapore 119076, Singapore. ^amathxw@nus.edu.sg*

²*Department of Mathematics, University of Central Florida, Orlando, FL 32816, USA. ^bjiongmin.yong@ucf.edu*

³*Department of Mathematics, University of Southern California, Los Angeles, CA 90089, USA. ^cjianfenz@usc.edu*

Received 13 April 2020; revised 23 January 2021; accepted 2 February 2021

Abstract. This paper is concerned with the relationship between forward–backward stochastic Volterra integral equations (FBSVIEs, for short) and a system of (nonlocal in time) path dependent partial differential equations (PPDEs, for short). Due to the nature of Volterra type equations, the usual flow property (or semigroup property) does not hold. Inspired by Viens–Zhang (*Ann. Appl. Probab.* **29** (2019) 3489–3540) and Wang–Yong (*Stochastic Process. Appl.* **129** (2019) 4926–4964), auxiliary processes are introduced so that the flow property of adapted solutions to the FBSVIEs is recovered in a suitable sense, and thus the functional Itô formula is applicable. Having achieved this stage, a natural PPDE is found so that the adapted solution of the backward SVIEs admits a representation in terms of the solution to the forward SVIE via the solution to a PPDE. On the other hand, the solution of the PPDE admits a representation in terms of adapted solution to the (path dependent) FBSVIE, which is referred to as a Feynman–Kac formula. This leads to the existence and uniqueness of a classical solution to the PPDE, under smoothness conditions on the coefficients of the FBSVIEs. Further, when the smoothness conditions are relaxed with the backward component of FBSVIE being one-dimensional, a new (and suitable) notion of viscosity solution is introduced for the PPDE, for which a comparison principle of the viscosity solutions is established, leading to the uniqueness of the viscosity solution. Finally, some results have been extended to coupled FBSVIEs and type-II BSVIEs, and a representation formula for the path derivatives of PPDE solution is obtained by a closer investigation of linear FBSVIEs.

Résumé. Cet article étudie les relations entre les équations intégrales de Volterra forward-backward stochastiques (FBSVIE) et un système d'équations aux dérivées partielles, non locales en temps, dépendant des trajectoires (PPDE). Du fait de la nature des équations de type Volterra, la propriété habituelle de flot, ou de semigroupe, n'est pas vérifiée. Influencés par les travaux de Viens–Zhang (*Ann. Appl. Probab.* **29** (2019) 3489–3540) et Wang–Yong (*Stochastic Process. Appl.* **129** (2019) 4926–4964), nous introduisons des processus auxiliaires tels que la propriété de flot de solutions adaptées aux FBSVIE peut être retrouvée en un sens approprié, de sorte que la formule d'Itô fonctionnelle peut être appliquée. Puis, nous exhibons une PPDE naturelle telle que la solution adaptée de la SVIE backward admet une représentation en termes de la solution de la SVIE forward via la solution de cette PPDE. Par ailleurs, la solution de la PPDE admet une représentation en termes de la solution à la FBSVIE (dépendant de la trajectoire), ce que nous interprétons comme une formule de Feynman–Kac. Ceci conduit à l'existence et l'unicité d'une solution classique à la PPDE, sous des conditions de régularité pour les coefficients de la FBSVIE. De plus, sous l'hypothèse que la composante backward de la FBSVIE est de dimension 1, on peut affaiblir ces conditions de régularité en introduisant une nouvelle notion de solution de viscosité pour la PPDE, et établir un principe de comparaison de ces solutions qui implique leur unicité. Enfin, certains résultats sont étendus à des FBSVIE couplées et des BSVIE de type II, et une formule de représentation des dérivées des trajectoires des solutions de la PPDE est obtenue par un examen approfondi des FBSVIE linéaires.

MSC2020 subject classifications: 60H20; 45D05; 35K10; 35D40; 60G22

Keywords: Forward backward stochastic Volterra integral equation; Path dependent partial differential equation; Feynman–Kac formula; Viscosity solution; Comparison principle

1. Introduction

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a complete filtered probability space, W a d -dimensional standard Brownian motion, $\mathbb{F} \equiv \{\mathcal{F}_t\}_{t \geq 0}$ the natural filtration generated by W augmented by all the $\mathbb{P}$ -null sets in $\mathcal{F}$ and $T > 0$ a fixed time horizon. Consider the

following (decoupled) *forward-backward stochastic differential equation* (FBSDE, for short): given initial data $(t, x) \in [0, T] \times \mathbb{R}^n$,

$$\begin{cases} X_s^{t,x} = x + \int_t^s b(r, X_r^{t,x}) dr + \int_t^s \sigma(r, X_r^{t,x}) dW_r, \\ Y_s^{t,x} = g(X_T^{t,x}) + \int_s^T f(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}) dr - \int_s^T Z_r^{t,x} dW_r, \end{cases} \quad s \in [t, T], \quad (1.1)$$

where the coefficients b, σ, f, g are deterministic functions. Such an FBSDE is associated with the following terminal value problem of a *partial differential equation* (PDE, for short):

$$\begin{cases} \partial_t u(t, x) + \frac{1}{2} \text{tr}[\partial_{xx}^2 u(t, x) \sigma(t, x) \sigma^\top(t, x)] + \partial_x u(t, x) b(t, x) \\ \quad + f(t, x, u(t, x), \partial_x u(t, x) \sigma(t, x)) = 0, \quad (t, x) \in [0, T] \times \mathbb{R}^n, \\ u(T, x) = g(x), \quad x \in \mathbb{R}^n. \end{cases} \quad (1.2)$$

By the seminal works Peng [48] and Pardoux–Peng [47], we have the nonlinear *Feynman–Kac formula*, representing the viscosity solution to PDE (1.2) by the adapted solution to FBSDE (1.1):

$$u(t, x) = Y_t^{t,x}, \quad (t, x) \in [0, T] \times \mathbb{R}^n, \quad (1.3)$$

and on the other hand the adapted solution $(Y^{t,x}, Z^{t,x})$ to the *backward stochastic differential equation* (BSDE, for short), namely the second equation in (1.1), has the following *representation formula* via the solution to PDE (1.2):

$$Y_s^{t,x} = u(s, X_s^{t,x}), \quad Z_s^{t,x} = \partial_x u(s, X_s^{t,x}) \sigma(s, X_s^{t,x}), \quad s \in [t, T], \quad (1.4)$$

provided u is smooth. The key for this PDE approach is the *flow property*, also called *semigroup property* and can be viewed as a type of *time consistency*, of the FBSDE. That is,

$$X_r^{t,x} = X_r^{s, X_s^{t,x}}, \quad Y_s^{t,x} = Y_s^{r, u(r, \cdot); t, x}, \quad Z_s^{t,x} = Z_s^{r, u(r, \cdot); t, x}, \quad t \leq s \leq r \leq T, \quad (1.5)$$

where $(Y^{r, u(r, \cdot); t, x}, Z^{r, u(r, \cdot); t, x})$ is the solution to the BSDE on $[t, r]$ with terminal condition $Y_r = u(r, X_r^{t,x})$. We remark that this approach remains effective for coupled FBSDEs (namely b, σ may depend on (Y, Z)); see Ma–Protter–Yong [44], and even for more general situations, where u plays the role of the *decoupling field* for the forward–backward equations.

In this paper, our objective is to consider the following decoupled *forward-backward stochastic Volterra integral equation* (FBSVIE, for short) with solution triple $(X_t, Y_t, Z_t^t), 0 \leq t \leq r \leq T$:

$$\begin{aligned} X_t &= \mathbf{x}_t + \int_0^t b(t, r, X_r) dr + \int_0^t \sigma(t, r, X_r) dW_r, \\ Y_t &= g(t, X_T) + \int_t^T f(t, r, X_r, Y_r, Z_r^t) dr - \int_t^T Z_r^t dW_r, \end{aligned} \quad t \in [0, T]. \quad (1.6)$$

Here, the coefficients b, σ, f involve two time variables; the initial condition is a continuous path $\mathbf{x} \in C([0, T]; \mathbb{R}^n)$; and the terminal condition g depends on t as well. A special case of the forward SVIE is the fractional Brownian motion, where $\mathbf{x} = 0, b = 0, \sigma = K(t, r)$ for some deterministic kernel K . FSVIE has received very strong attention in recent years due to its applications in *rough volatility* models; see, for example, Comte–Renault [10], Gatheral–Jaisson–Rosenbaum [30], El Euch–Rosenbaum [26,27] and Viens–Zhang [60]. On the other hand, BSVIE has become a popular tool for studying many problems in mathematical finance. For examples, Di Persio [18] on stochastic differential utility, Yong [69], Wang–Yong [64] and Agram [2] on dynamic risk measures, Kromer–Overbeck [38] on dynamic capital allocations, Wang–Sun–Yong [62] on equilibrium recursive utility and equilibrium dynamic risk measures, to mention a few. More interestingly, in recent years, time-inconsistent problems have attracted many researchers’ attention. Among others, the time-inconsistency could be caused by the time-preferences of the decision-makers, which can be described by nonexponential discounting. See the seminal paper by Strotz [59], and early follow-up works of Pollak [53] and Laibson [39]. For the recent works of time-inconsistent problems relevant to the nonexponential discounting, we mention Karp [37], Ekeland–Lazrak [22], Yong [71], Wei–Yong–Yu [67] and Hernandez–Possamai [34]. It is worthy of pointing out that the most suitable dynamic recursive cost functional allowing nonexponential discounting should be described by a BSVIE, as indicated in Wang–Yong [63]. We remark that the BSVIE in (1.6) is also called type-I BSVIE in the literature. A more general type-II BSVIE, where f depends not only on Z_r^t , but also on Z_t^t , appears naturally as an adjoint equation when one studies stochastic maximum principle for controlled FSVIE; see Yong [68,70].

Our goal of this paper is to extend the PDE approach to FBSVIEs. This on one hand will help us to understand the structure of FBSVIEs, and on the other hand is helpful for numerical computation of these equations. As mentioned, the PDE approach is based on the flow property of the equations. Unfortunately, due to the two time variable structure, neither FSVIE nor BSVIE satisfies the flow property in the standard sense: for $0 \leq t < s \leq T$,

$$X_s \neq X_t + \int_t^s b(s, r, X_r) dr + \int_t^s \sigma(s, r, X_r) dW_r, \quad Y_t \neq Y_s + \int_t^s f(t, r, X_r, Y_r, Z_r^t) dr - \int_t^s Z_r^t dW_r.$$

Our work is built on Viens–Zhang [60], Yong [72] and Wang–Yong [65]. By introducing auxiliary two-time variable processes $\tilde{X}_t^s, \tilde{Y}_t^s$ (see (2.6) and (2.11) below), [60] recovers the flow property of the FSVIE in certain sense, and [65, 72] recover the flow property of the BSVIE. We remark that in [60] the backward equation is a standard BSDE, while in [65, 72] the forward equation is a standard SDE. Putting together allows us to adopt the PDE approach to FBSVIE (1.6). We note that the associated PDE will intrinsically depend on the paths of $\tilde{X}_t^{[t, T]}$, and thus it becomes a *path dependent PDE* (PPDE, for short). Then, with a little extra effort, we can actually handle path dependent FBSVIEs, namely b, σ, f, g depend on the paths of X , as we will do in the paper. We shall emphasize though, even for the state dependent case (1.6), our results in the paper are new.

To be precise, we shall introduce a two-time variable function $U(t, s, \mathbf{x})$; $0 \leq t \leq s \leq T$, $\mathbf{x} \in C([0, T]; \mathbb{R}^n)$, which satisfies the PPDE with terminal condition $U(t, T, \mathbf{x}) = g(t, \mathbf{x})$:

$$\begin{aligned} \partial_s U(t, s, \mathbf{x}) + \frac{1}{2} \langle \partial_{\mathbf{x}\mathbf{x}}^2 U(t, s, \mathbf{x}), (\sigma_{[s, T]}^{s, \mathbf{x}}, \sigma_{[s, T]}^{s, \mathbf{x}}) \rangle + \langle \partial_{\mathbf{x}} U(t, s, \mathbf{x}), b_{[s, T]}^{s, \mathbf{x}} \rangle \\ + f(t, s, \mathbf{x}, U(s, s, \mathbf{x}), \langle \partial_{\mathbf{x}} U(t, s, \mathbf{x}), \sigma_{[s, T]}^{s, \mathbf{x}} \rangle) = 0. \end{aligned} \quad (1.7)$$

Here, $\partial_{\mathbf{x}} U, \partial_{\mathbf{x}\mathbf{x}}^2 U$ are the first- order and second-order Fréchet derivatives with respect to the perturbation of $\mathbf{x}_{[s, T]}$, and for $\varphi = b, \sigma, \varphi_{[s, T]}^{s, \mathbf{x}}$ refers to the path $\{\varphi(r, s, \mathbf{x})\}_{r \in [s, T]}$. Then we have the following relationship, which extends (1.4): denoting $\hat{X}_r^t := X_r 1_{[0, t)}(r) + \tilde{X}_r^t 1_{[t, T]}(r)$,

$$Y_t = U(t, t, \hat{X}^t), \quad Z_s^t = \langle \partial_{\mathbf{x}} U(t, s, \hat{X}^s), \sigma_{[s, T]}^{s, \hat{X}^s} \rangle, \quad \text{and} \quad \tilde{Y}_s^t = U(t, s, \hat{X}^s), \quad (1.8)$$

and similarly we can extend (1.3) to this case (see (3.6) below), and thus establish the Feynman–Kac formula for (1.7). Besides the key flow property, a crucial tool in this analysis is the functional Itô formula, initiated by Dupire [21] in standard SDE setting and extended to the SVIE setting by [60]. The PPDE (1.7) has several important features:

- The state variable $\mathbf{x}$ has a continuous path on $[0, T]$, and thus is infinite dimensional.
- U depends on two time variables (t, s) . In particular, the equation at $(t, s, \mathbf{x})$ involves the value $U(s, s, \mathbf{x})$, and thus is nonlocal in the first time variable t .
- Alternatively, noting that (1.7) does not involve derivatives with respect to the first time variable t , then we may view t as a parameter instead of a variable. That is, we may view (1.7) as a system of PPDEs with parameter t and solution $\{U(t, \cdot, \cdot)\}_{t \in [0, T]}$. Then this is an (uncountably) infinite dimensional system of PPDEs, which are self-interacted through the diagonal term $U(s, s, \mathbf{x})$.

We next prove the existence of classical solutions to PPDE (1.7), provided the coefficients are smooth enough in an appropriate sense, and thus establish the above connection between PPDE (1.7) and FBSVIE (1.6) rigorously. We remark that Peng–Wang [51] obtained the classical solution in the form $u(t, \mathbf{x}_{[0, t]})$ for a PPDE corresponding to PDE (1.2), associated with the path dependent version of the FBSDE (1.1). Our result generalizes [51] in several aspects. First, in [51] $u(t, \mathbf{x}_{[0, t]})$ depends on the path only up to t , in particular the path derivative $\partial_{\mathbf{x}} u$ there involves only the perturbation of $\mathbf{x}_t$, and thus is actually a finite dimensional derivative, while our path derivative is indeed a Fréchet derivative. Next, u is finite dimensional, while as mentioned (1.7) can be viewed as an infinite dimensional system. Moreover, when restricted to the state dependent case, the PPDE in [51] reduces back to the standard PDE (1.2), but (1.7) has the same features that both the state $\mathbf{x}$ and the solution U are infinite dimensional. We also obtain a representation formula for the path derivative $\partial_{\mathbf{x}} U(t, s, \mathbf{x})$, which is interesting in its own right and is new in the literature.

The more challenging part is the viscosity solution theory for PPDE (1.7), in the case that Y is scalar but the coefficients are less smooth. Note that the state space $C([0, T]; \mathbb{R}^n)$ is not locally compact, so the standard viscosity solution theory of Crandall–Ishii–Lions [15] does not work here. Moreover, we have some intrinsic adaptiveness requirement on the dependence of the path, which prevents us from applying the viscosity solution theory in infinite dimensional space; see, for example, Crandall–Lions [16], Li–Yong [41] and Fabbri–Gozzi–Swiech [28]. One exception in this direction is Ren–Rosestolato [55], which however requires some stronger type of regularity and is overall still not satisfactory for our

purpose. We shall follow the approach proposed by Ekren–Keller–Touzi–Zhang [23], where the pointwise optimization in [15] is replaced with an optimal stopping problem under certain nonlinear expectation, and thus the comparison principle can be obtained without requiring the local compactness of the state space. Our PPDE (1.7) has two major differences from [23]. First, the nonlinear expectation used in [23] relies on a family of semi-martingale measures, while our state process X is not a semimartingale. Second, the PPDE in [23] is one-dimensional and the comparison principle for classical solutions (if they exist) is quite straightforward, but as mentioned PPDE (1.7) is nonlocal (or viewed as infinite dimensional), and in fact the comparison principle fails in general even for classical solutions. Nevertheless, we shall propose a new notion of viscosity solution to PPDE (1.7) and establish its well-posedness, including the comparison principle, under an additional assumption that f is nondecreasing in y . We note that this monotonicity condition is essentially the *proper* condition required in [15] for elliptic equations. For a standard parabolic equation like (1.2), this condition is redundant because it is implied from the Lipschitz condition by a standard change variable argument. However, the change variable argument fails for (1.7) because of its nonlocal structure. We also note that Wang–Yong [64] proved the comparison principle for BSVIEs under the same monotonicity condition. As in the standard literature, since viscosity solution is a local notion (even with some nonlocal feature here), its comparison principle is much more challenging.

Finally, we investigate briefly two more general FBSVIEs, the coupled FBSVIE (with b, σ depending on Y) and the type-II BSVIEs, and extend the representation formula in these cases. The more detailed studies on these equations are left for interested readers. We note particularly that our new representation formula for $\partial_x U$ relies on a linear type-II BSVIE. For this purpose, we establish a duality result for linear path dependent FSVIE, which covers the corresponding results in Yong [68,70] and Peng–Yang [52], and provide an explicit solution for linear BSVIEs, which generalizes the result of Hu–Øksendal [36].

The rest of this paper is organized as follows. In Section 1.1, we provide a literature review on the closely related topics. Section 2 collects some preliminary results, which will be used in the paper. In Section 3, we establish the connection between FBSVIEs and PPDEs, and prove the existence of classical solutions under appropriate conditions. Section 4 is devoted to the viscosity solutions of the PPDE. We extend some results to coupled FBSVIEs and type-II BSVIEs in Section 5. Finally in Section 6, we obtain a representation formula for the path derivative $\partial_x U(t, s, \mathbf{x})$.

1.1. Literature review on some related topics

For FSVIEs, we first refer to Nualart [45] for a comprehensive exposition of fractional Brownian motion, which is a very special case of FSVIEs. In the state dependent case, the well-posedness of FSVIEs can be found in Berger–Mizel [8]. Since one cannot apply the Burkholder–Davis–Gundy inequalities for stochastic Volterra integral equations, the well-posedness of path dependent FSVIEs is actually more involved, and we refer to the recent work Ruan–Zhang [57]. There has been a growing number of publications on rough volatility models, for which FSVIE is a convenient tool. Besides [10,26,27,30,60], a partial list also includes Abi Jaber–Larsson–Pulido [1], Alos–Leon–Vives [3], Bayer–Friz–Gatheral [6], Bennedsen–Lunde–Pakkanen [7], Chronopoulou–Viens [9], Cuchiero–Teichmann [17], Fouque–Hu [29], Gatheral–Keller–Ressel [31] and Gulisashvili–Viens–Zhang [32].

BSVIE was first introduced by Lin [42] in a special form. The general form, including type-II BSVIEs, has been studied systematically by Yong [68,70], followed by Djordjevic–Jankovic [19,20], Shi–Wang–Yong [58], Wang–Yong [64], Wang–Zhang [66], Overbeck–Roder [46], Hu–Øksendal [36], Wang–Yong [65], Popier [54], Hernandez–Possamai [35], to mention a few. In particular, we note that Hamaguchi [33] proved the well-posedness of coupled FBSVIEs over small time horizon. The well-posedness of coupled FBSVIEs over arbitrary time horizon is still open, to our best knowledge. We also refer to [2,18,34,38,61–63,69] again for some applications of BSVIEs.

The notion of PPDE was first proposed by Peng [49]. A crucial tool is the functional Itô formula, initiated by Dupire [21] and further developed by Cont–Fournié [11,12]. In the semilinear case, Peng–Wang [51] obtained the classical solution and Ekren–Keller–Touzi–Zhang [23] established the viscosity solution theory. The viscosity solution approach of [23] has been successfully extended to the fully nonlinear case by Ekren–Touzi–Zhang [24,25], Ren–Touzi–Zhang [56] and Ren–Rosolatos [55]. We also refer to Barrasso–Russo [4], Cosso–Russo [13], Leao–Ohashi–Simas [40], Lukyanov [43], Peng–Song [50] for some related works, in particular to Cosso–Russo [14], Zhou [75] for some recent interesting developments and to the book Zhang [74] for more references. We shall remark though that the PPDEs in all the above works are in the semimartingale setting. Our PPDE is associated with SVIEs, and the corresponding functional Itô formula was proved by [60]. Another closely related work also beyond semi-martingale setting is Barrasso–Russo [5], which studies the so-called decoupled mild solution for a PPDE associated with Gaussian processes.

2. Preliminaries

Let $T > 0$ be the time horizon, $\Omega := C([0, T]; \mathbb{R}^d)$ the canonical space, W the canonical process (namely $W(\omega) = \omega$), $\mathbb{P}$ the Wiener measure (namely W is a standard d -dimensional Brownian motion under $\mathbb{P}$), and $\mathbb{F} := \mathbb{F}^W$ augmented with

the $\mathbb{P}$ -null sets. Denote

$$\begin{aligned}\mathbb{T} &= [0, T], & \mathbb{T}^2 &= [0, T] \times [0, T], \\ \mathbb{T}_-^2 &= \{(t, s) \mid 0 \leq s \leq t \leq T\}, & \mathbb{T}_+^2 &= \{(t, s) \mid 0 \leq t \leq s \leq T\}.\end{aligned}$$

Here, “ $-$ ” indicates the left neighborhood of t , and “ $+$ ” indicates the right neighborhood of t . For any Euclidean space $\mathbb{H}$ (say, $\mathbb{R}^n$, $\mathbb{R}^{m \times d}$, etc.), let

$$\mathbb{L}_{\mathbb{F}}^p(0, T; \mathbb{H}) = \left\{ \varphi : [0, T] \times \Omega \rightarrow \mathbb{H} \mid \varphi \text{ is } \mathbb{F}\text{-progressively measurable, } \mathbb{E} \int_0^T |\varphi(s)|^p ds < \infty \right\}.$$

Our state space is $\mathbb{X} := C([0, T]; \mathbb{R}^n)$, equipped with the uniform norm:

$$\|\mathbf{x}\| = \sup_{t \in [0, T]} |\mathbf{x}_t|, \quad \forall \mathbf{x} \in \mathbb{X}. \quad (2.1)$$

In this section, we review and present some basic results concerning forward and backward SVIEs, including a continuous-norm estimate for the adapted solution to a class of BSVIEs. Moreover, among other things, we shall introduce two auxiliary processes $\tilde{X}$ and $\tilde{Y}$ so that the flow property of the adapted solutions can be established in an extended sense. It turns out that such a property will play an essential role in proving the relation between FBSVIEs and PPDEs.

Before going further, we make a convention which will be used in the rest of the paper. For any map $\varphi : \mathbb{T}^2 \times \mathbb{X} \times \mathbb{H} \times \Omega \rightarrow \tilde{\mathbb{H}}$, where $\mathbb{H}$ and $\tilde{\mathbb{H}}$ are any Euclidean spaces (could be $\mathbb{R}^m$, $\mathbb{R}^m \times \mathbb{R}^{m \times d}$, etc.), we simply say that φ is *progressively measurable* if

$$\varphi(t, r, \mathbf{x}, h, \omega) = \varphi(t, r, \mathbf{x}_{r \wedge \cdot}, h, \omega_{r \wedge \cdot}), \quad \forall (t, r, \mathbf{x}, h, \omega) \in \mathbb{T}^2 \times \mathbb{X} \times \mathbb{H} \times \Omega, \quad (2.2)$$

and the above map is measurable. In the above, $\mathbb{T}^2$ can be replaced by $\mathbb{T}_{\pm}^2$; also some independent variables can be absent.

2.1. The well-posedness and flow property of FSVIEs

Given $\mathbf{x} \in \mathbb{X}$, consider an FSVIE:

$$X_t = \mathbf{x}_t + \int_0^t b(t, r, X_\cdot) dr + \int_0^t \sigma(t, r, X_\cdot) dW_r, \quad t \in \mathbb{T}. \quad (2.3)$$

We shall assume the following.

Assumption 2.1. The map $(b, \sigma) : \mathbb{T}_-^2 \times \mathbb{X} \rightarrow \mathbb{R}^n \times \mathbb{R}^{n \times d}$ is progressively measurable satisfying:

- (i) For some constant $C_0 > 0$, $|b(t, r, \mathbf{0})| + |\sigma(t, r, \mathbf{0})| \leq C_0$ for all $(t, r) \in \mathbb{T}_-^2$.
- (ii) The map $\mathbf{x} \mapsto (b(t, r, \mathbf{x}), \sigma(t, r, \mathbf{x}))$ is uniformly Lipschitz continuous under the norm $\|\cdot\|$.
- (iii) The map $t \mapsto (b(t, r, \mathbf{x}), \sigma(t, r, \mathbf{x}))$ is differentiable, with $\partial_t b$ and $\partial_t \sigma$ also satisfying the conditions as in (i) and (ii).

The following result follows from Ruan–Zhang [57].

Proposition 2.2. Under Assumption 2.1, FSVIE (2.3) admits a unique strong solution X such that X is continuous in t and the following estimate holds true: for any $p > 1$,

$$\mathbb{E}[\|X\|^p] \leq C_p[1 + \|\mathbf{x}\|^p]. \quad (2.4)$$

We remark that due to the first time variable t in σ , one cannot directly apply the Burkholder–Davis–Gundy inequalities in the Volterra setting. Assumption 2.1(iii) helps us to get around that. In the state dependent case: $\sigma = \sigma(t, r, X_r)$ (b can be path dependent, although in the literature, typically, it is also state dependent), the well-posedness of (2.3) follows from standard arguments; see, for example, Berger–Mizel [8]. The pathwise continuity of X as well as the norm estimate (2.4) hold true for b, σ satisfying weaker continuity in the spirit of (2.9) below. The arguments are similar to those of Proposition 2.4 below and we skip the details. It will be interesting to see if it is possible to weaken Assumption 2.1(iii) in the path dependent case.

Note that X is neither a Markov process nor a semimartingale. Even worse, in general the flow property fails in the following sense: for fixed t ,

$$X_s \neq X_t + \int_t^s b(s, r, X_\cdot) dr + \int_t^s \sigma(s, r, X_\cdot) dW_r, \quad s \in (t, T]. \quad (2.5)$$

One may refer to this as the *time inconsistency*. To overcome this deficiency, Viens–Zhang [60] introduced an auxiliary process with two-time variables:

$$\tilde{X}_t^s = \mathbf{x}_s + \int_0^t b(s, r, X_\cdot) dr + \int_0^t \sigma(s, r, X_\cdot) dW_r, \quad (t, s) \in \mathbb{T}_+^2. \quad (2.6)$$

This process enjoys the following nice properties:

- For fixed s , the process $[0, s] \ni t \mapsto \tilde{X}_t^s$ is an $\mathbb{F}$ -semimartingale with $\tilde{X}_t^t = X_t$;
- For fixed t , the process $[t, T] \ni s \mapsto \tilde{X}_t^s$ is $\mathcal{F}_t$ -measurable and continuous;
- The *flow property* holds in the following sense: for any $\mathbb{F}$ -stopping time τ ,

$$X_s = \tilde{X}_\tau^s + \int_\tau^s b(s, r, X_\cdot) dr + \int_\tau^s \sigma(s, r, X_\cdot) dW_r, \quad s \in [\tau, T]. \quad (2.7)$$

We remark that, in the state dependent case as in (1.6), (2.7) implies

$$X_s = \tilde{X}_t^s + \int_t^s b(s, r, X_r) dr + \int_t^s \sigma(s, r, X_r) dW_r, \quad (t, s) \in \mathbb{T}_+^2.$$

One can easily see that, conditional on $\tilde{X}_t^{[t, T]}$, $X_{[0, t]}$ and $X_{(t, T]}$ are conditionally independent. So this can be viewed as a generalized Markov property.

2.2. The well-posedness and flow property of BSVIEs

Consider the following path dependent BSVIE:

$$Y_t = g(t, X_\cdot) + \int_t^T f(t, r, X_\cdot, Y_r, Z_r^t) dr - \int_t^T Z_r^t dW_r, \quad t \in \mathbb{T}, \quad (2.8)$$

where Y is m -dimensional, and hence Z is $(m \times d)$ -dimensional. We shall assume

Assumption 2.3. The map $f : \mathbb{T}_+^2 \times \mathbb{X} \times \mathbb{R}^m \times \mathbb{R}^{m \times d} \rightarrow \mathbb{R}^m$ is progressively measurable and the map $g : \mathbb{T} \times \mathbb{X} \rightarrow \mathbb{R}^m$ is $\mathcal{F}_T$ -measurable satisfying:

- (i) For some constant $C_0 > 0$, it holds

$$|f(t, r, \mathbf{x}, 0, 0)| + |g(t, \mathbf{x})| \leq C_0[1 + \|\mathbf{x}\|], \quad \forall (t, r, \mathbf{x}) \in \mathbb{T}_+^2 \times \mathbb{X};$$

- (ii) The map $(y, z) \mapsto f(t, r, \mathbf{x}, y, z)$ is uniformly Lipschitz continuous;

- (iii) The map $t \mapsto (f(t, r, \mathbf{x}, y, z), g(t, \mathbf{x}))$ is locally uniformly continuous in the following sense: for some modulus of continuity function ρ ,

$$\begin{aligned} & |f(t - \delta, r, \mathbf{x}, y, z) - f(t, r, \mathbf{x}, y, z)| + |g(t - \delta, \mathbf{x}) - g(t, \mathbf{x})| \\ & \leq C[1 + \|\mathbf{x}\| + |y| + |z|]\rho(\delta), \quad \forall (t, r, \mathbf{x}, y, z) \in \mathbb{T}_+^2 \times \mathbb{X} \times \mathbb{R}^m \times \mathbb{R}^{m \times d}, \delta \in [0, t]. \end{aligned} \quad (2.9)$$

Proposition 2.4. Under Assumptions 2.1 and 2.3, BSVIE (2.8) admits a unique strong solution (Y, Z) such that Y is continuous in t and the following estimate holds true: for each $p > 1$,

$$\mathbb{E}\left[\sup_{0 \leq t \leq T} |Y_t|^p\right] + \sup_{0 \leq t \leq T} \mathbb{E}\left[\left(\int_t^T |Z_s^t|^2 ds\right)^{\frac{p}{2}}\right] \leq C_p[1 + \|\mathbf{x}\|^p]. \quad (2.10)$$

The proof of the well-posedness of BSVIE (2.8) is standard and could be found in Yong [70], where the pathwise continuity of Y was proved for a more general BSVIE, but under much stronger technical conditions. Our arguments for the pathwise continuity and the above estimate (2.10) seems to be new in the literature. We note that (2.9) is much weaker than Assumption 2.1(iii), because f is state dependent on (Y, Z) . To facilitate the proof, we introduce the following standard BSDE parameterized by $t \in \mathbb{T}$ with adapted solution $(\tilde{Y}^t, \tilde{Z}^t)$:

$$\tilde{Y}_s^t = g(t, X_\cdot) + \int_s^T f(t \wedge r, r, X_\cdot, Y_r, \tilde{Z}_r^t) dr - \int_s^T \tilde{Z}_r^t dW_r, \quad s \in [0, T]. \quad (2.11)$$

Proof. First, from [70] we know (2.8) admits a unique solution (Y, Z) such that $Y \in \mathbb{L}_{\mathbb{F}}^2(0, T; \mathbb{R}^m)$ and $Z^t \in \mathbb{L}_{\mathbb{F}}^2(t, T; \mathbb{R}^{m \times d})$. Compare (2.8) with the following linear BSDE on $[t, T]$:

$$\hat{Y}_s^t = g(t, X_\cdot) + \int_s^T f(t \wedge r, r, X_\cdot, Y_r, Z_r^t) dr - \int_s^T \hat{Z}_r^t dW_r, \quad s \in [t, T].$$

It is obvious that $\hat{Y}_t^t = Y_t$ and $\hat{Z}_r^t = Z_r^t$. This implies that $(\hat{Y}^t, \hat{Z}^t)$ satisfies (2.11). Then we have

$$\tilde{Y}_t^t = Y_t, \quad \tilde{Z}_r^t = Z_r^t, \quad (t, r) \in \mathbb{T}_+^2. \quad (2.12)$$

Next, by (2.4), (2.9) and the standard BSDE arguments, we have

$$\begin{aligned} \sup_{0 \leq t \leq T} \mathbb{E} \left[\sup_{0 \leq s \leq T} |\tilde{Y}_s^t|^p + \left(\int_0^T |Z_s^t|^2 ds \right)^{\frac{p}{2}} \right] &\leq C_p [1 + \|\mathbf{x}\|^p]; \\ |\tilde{Y}_s^t - \tilde{Y}_s^{t'}| &\leq C_p [1 + (\mathbb{E}_s[\|X\|^p])^{\frac{1}{p}}] \rho(|t - t'|), \quad \text{a.s., } \forall t, t', s \in \mathbb{T}. \end{aligned} \quad (2.13)$$

Note that $\tilde{Y}_s^t$, $\tilde{Y}_s^{t'}$, and $\mathbb{E}_s[\|X\|^p]$ are pathwise continuous in s , then we have

$$\sup_{s \in \mathbb{T}} |\tilde{Y}_s^t - \tilde{Y}_s^{t'}| \leq C_p \left[1 + \sup_{s \in \mathbb{T}} (\mathbb{E}_s[\|X\|^p])^{\frac{1}{p}} \right] \rho(|t - t'|), \quad \text{a.s., } \forall t, t' \in \mathbb{T}. \quad (2.14)$$

Note that, by (standard) Doob's maximum inequality,

$$\mathbb{E} \left[\sup_{s \in \mathbb{T}} \mathbb{E}_s[\|X\|^p] \right] \leq C_p (\mathbb{E}[\|X\|^{2p}])^{\frac{1}{2}} \leq C_p [1 + \|\mathbf{x}\|^p] < \infty.$$

Let $\{t_i\}_{i \geq 1}$ be the rationals in $[0, T]$. There exists an $\Omega_1 \subset \Omega$ such that $\mathbb{P}(\Omega_1) = 1$, $\tilde{Y}_s^{t_i}(\omega)$ is continuous in s , and

$$\begin{aligned} \sup_{s \in \mathbb{T}} |\tilde{Y}_s^{t_i} - \tilde{Y}_s^{t_j}|(\omega) &\leq C_p(\omega) \rho(|t_i - t_j|), \quad \forall (i, j), \forall \omega \in \Omega_1, \\ \text{where } C_p(\omega) &:= \left[1 + \sup_{s \in \mathbb{T}} (\mathbb{E}_s[\|X\|^p])^{\frac{1}{p}} \right](\omega) < \infty, \quad \forall \omega \in \Omega_1. \end{aligned} \quad (2.15)$$

For any $t \in \mathbb{T}$, by (2.14), there exists an $\Omega^t \subset \Omega$ such that $\mathbb{P}(\Omega^t) = 1$ and

$$\sup_{s \in \mathbb{T}} |\tilde{Y}_s^t - \tilde{Y}_s^{t_j}|(\omega) \leq C_p(\omega) \rho(|t - t_j|), \quad \forall j, \forall \omega \in \Omega^t \cap \Omega_1. \quad (2.16)$$

For any $(t, s) \in \mathbb{T}^2$, we define

$$\bar{Y}_s^t(\omega) := \limsup_{t_j \rightarrow t} \tilde{Y}_s^{t_j}(\omega), \quad \omega \in \Omega. \quad (2.17)$$

By (2.15), we see that the above $\limsup$ is actually a limit for $\omega \in \Omega_1$. Then, for any $\omega \in \Omega_1$,

$$\bar{Y}_s^t(\omega) \text{ is continuous in } s, \quad \sup_{s \in \mathbb{T}} |\bar{Y}_s^t - \bar{Y}_s^{t'}|(\omega) \leq C_p(\omega) \rho(|t - t'|), \quad \forall t, t' \in \mathbb{T}. \quad (2.18)$$

So $\bar{Y}$ is (uniformly) continuous in $(t, s) \in \mathbb{T}^2$ for all $\omega \in \Omega_1$. Moreover, by (2.16) we have

$$\tilde{Y}_s^t(\omega) = \bar{Y}_s^t(\omega), \quad \forall s \in \mathbb{T}, \forall \omega \in \Omega^t \cap \Omega_1. \quad (2.19)$$

Since $\mathbb{P}(\Omega^t \cap \Omega_1) = 1$, so $\bar{Y}$ is a desired version of $\tilde{Y}_s^t$, and thus, by always considering this version, $\tilde{Y}_s^t$ is jointly continuous in (t, s) , a.s. In particular, this implies that $Y_t = \tilde{Y}_t^t$ is continuous in t , a.s.

Finally, applying the standard BSDE estimates on (2.11) we have

$$|Y_t|^p = |\tilde{Y}_t^t|^p \leq C_p \mathbb{E}_t \left[1 + \|X\|^p + \int_t^T |Y_r|^p dr \right], \quad \text{a.s.}$$

Since $t \mapsto Y_t$ is continuous almost surely, we obtain from the Doob's maximum inequality that

$$\mathbb{E} \left[\sup_{0 \leq t \leq T} |Y_t|^p \right] \leq C_p \left(\mathbb{E} \left[1 + \|X\|^{2p} + \int_0^T |Y_r|^{2p} dr \right] \right)^{\frac{1}{2}} \leq C_p [1 + \|\mathbf{x}\|^p],$$

where the second inequality thanks to (2.4) and the first line of (2.13). This, together with the first line of (2.13) again, implies (2.10). $\square$

Similar to the forward case, in general the flow property fails in the following sense: for fixed s ,

$$Y_t \neq Y_s + \int_t^s f(t, r, X., Y_r, Z_r^t) dr - \int_t^s Z_r^t dW_r, \quad t \in [0, s). \quad (2.20)$$

However, we may recover the flow property by utilizing the auxiliary process $\tilde{Y}$:

$$Y_t = \tilde{Y}_s^t + \int_t^s f(t, r, X., Y_r, Z_r^t) dr - \int_t^s Z_r^t dW_r, \quad t \in [0, s]. \quad (2.21)$$

2.3. The functional Itô formula

The materials in this subsection follow from Viens–Zhang [60]. Recall $\mathbb{X} := C(\mathbb{T}; \mathbb{R}^n)$ and define

$$\begin{aligned} \mathbb{X}_t &:= C([t, T]; \mathbb{R}^n); \\ \widehat{\mathbb{X}} &:= D(\mathbb{T}; \mathbb{R}^n) = \{\mathbf{x} : \mathbb{T} \rightarrow \mathbb{R}^n \mid \mathbf{x} \text{ is right-continuous with left-limits}\}. \end{aligned} \quad (2.22)$$

Clearly, $\mathbb{X}$ is a subset of $\widehat{\mathbb{X}}$. Also, hereafter, for any $\eta \in \mathbb{X}_t$, we automatically extend it to be zero on $[0, t)$, still denote it by η . Then $\mathbb{X}_t \subseteq \widehat{\mathbb{X}}$. Next, we define

$$\Lambda := \mathbb{T} \times \mathbb{X}, \quad \widehat{\Lambda} := \{(t, \mathbf{x}) \in \mathbb{T} \times \widehat{\mathbb{X}} : \mathbf{x}|_{[t, T]} \in \mathbb{X}_t\}, \quad \mathbf{d}((t, \mathbf{x}), (t', \mathbf{x}')) := |t - t'| + \|\mathbf{x} - \mathbf{x}'\|,$$

with $\|\mathbf{x}\| = \sup_{t \in \mathbb{T}} |\mathbf{x}_t|$ for $\mathbf{x} \in \widehat{\mathbb{X}}$. It can be shown that $\mathbf{d}$ is a metric under which $\widehat{\Lambda}$ is a complete metric space. Now, let $C^0(\widehat{\Lambda})$ denote the set of all functions $u : \widehat{\Lambda} \rightarrow \mathbb{R}$, which are continuous under $\mathbf{d}$. For any $u \in C^0(\widehat{\Lambda})$ and given $(t, \mathbf{x}) \in \widehat{\Lambda}$, define

$$\partial_t u(t, \mathbf{x}) = \lim_{\delta \downarrow 0} \frac{u(t + \delta, \mathbf{x}) - u(t, \mathbf{x})}{\delta}, \quad (2.23)$$

provided the limit exists, and define $\partial_{\mathbf{x}} u(t, \mathbf{x})$ as the Fréchet derivative with respect to $\mathbf{x}|_{[t, T]}$, namely $\partial_{\mathbf{x}} u(t, \mathbf{x}) : \mathbb{X}_t \rightarrow \mathbb{R}$ is the linear functional satisfying the following:

$$u(t, \mathbf{x} + \eta) - u(t, \mathbf{x}) = \langle \partial_{\mathbf{x}} u(t, \mathbf{x}), \eta \rangle + o(\|\eta\|), \quad \forall \eta \in \mathbb{X}_t. \quad (2.24)$$

It is clear that this is equal to the Gâteaux derivative:

$$\langle \partial_{\mathbf{x}} u(t, \mathbf{x}), \eta \rangle = \lim_{\varepsilon \rightarrow 0} \frac{u(t, \mathbf{x} + \varepsilon \eta) - u(t, \mathbf{x})}{\varepsilon}, \quad \forall \eta \in \mathbb{X}_t. \quad (2.25)$$

Similarly, we define the second-order derivative $\partial_{\mathbf{xx}}^2 u(t, \mathbf{x})$ as a bilinear functional on $\mathbb{X}_t \times \mathbb{X}_t$:

$$\langle \partial_{\mathbf{x}} u(t, \mathbf{x} + \eta), \eta' \rangle - \langle \partial_{\mathbf{x}} u(t, \mathbf{x}), \eta' \rangle = \langle \partial_{\mathbf{xx}}^2 u(t, \mathbf{x}), (\eta, \eta') \rangle + o(\|\eta\|), \quad \forall \eta, \eta' \in \mathbb{X}_t. \quad (2.26)$$

Definition 2.5. Let $C_+^{1,2}(\widehat{\Lambda})$ denote the set of $u \in C^0(\widehat{\Lambda})$ such that $\partial_t u$, $\partial_{\mathbf{x}} u$, $\partial_{\mathbf{xx}}^2 u$ exist on $\mathbb{T} \times \widehat{\mathbb{X}}$ and satisfy:

(i) There exist constants $\kappa > 0$ and $C > 0$ such that, for any $(t, \mathbf{x})$,

$$|\partial_t u(t, \mathbf{x})| + \sup_{\eta \in \mathbb{X}_t, \|\eta\| \leq 1} |\langle \partial_{\mathbf{x}} u(t, \mathbf{x}), \eta \rangle| + \sup_{\eta, \eta' \in \mathbb{X}_t, \|\eta\|, \|\eta'\| \leq 1} |\langle \partial_{\mathbf{xx}}^2 u(t, \mathbf{x}), (\eta, \eta') \rangle| \leq C[1 + \|\mathbf{x}\|^\kappa].$$

(ii) For any $\eta, \eta' \in \mathbb{X}$, $\langle \partial_{\mathbf{x}} u(t, \mathbf{x}), \eta|_{[t, T]} \rangle$, $\langle \partial_{\mathbf{xx}}^2 u(t, \mathbf{x}), (\eta|_{[t, T]}, \eta'|_{[t, T]}) \rangle$ are continuous in $(t, \mathbf{x})$, where the continuity in t always means right continuity.

(iii) There exist $\kappa > 0$ and a modulus of continuity function ρ such that

$$|\langle \partial_{\mathbf{xx}}^2 u(t, \mathbf{x}) - \partial_{\mathbf{xx}}^2 u(t, \mathbf{x}'), (\eta, \eta) \rangle| \leq [1 + \|\mathbf{x}\|^\kappa + \|\mathbf{x}'\|^\kappa] \|\eta\|^2 \rho(\|\mathbf{x} - \mathbf{x}'\|).$$

We remark that the function u will be involved in some backward equations, so both in (2.23) and in Definition 2.5(ii) the time regularity is only required to be from right. We note that [60] assumes $\langle \partial_{\mathbf{x}} u(t, \mathbf{x}), \eta|_{[t, T]} \rangle$, etc. is continuous in t , but actually it can only be right continuous because of the indicator function in $\eta|_{[t, T]}$ and in all the arguments in [60] only right continuity is used. For any $u_1, u_2 \in C^{1,2}(\widehat{\Lambda})$, if $u_1 = u_2$ on Λ , by [60] we have, for any $(t, \mathbf{x}) \in \Lambda$ and $\eta \in \mathbb{X}_t$,

$$\begin{aligned} \partial_t u_1(t, \mathbf{x}) &= \partial_t u_2(t, \mathbf{x}), & \langle \partial_{\mathbf{x}} u_1(t, \mathbf{x}), \eta \rangle &= \langle \partial_{\mathbf{x}} u_2(t, \mathbf{x}), \eta \rangle, \\ \langle \partial_{\mathbf{xx}}^2 u_1(t, \mathbf{x}), (\eta, \eta) \rangle &= \langle \partial_{\mathbf{xx}}^2 u_2(t, \mathbf{x}), (\eta, \eta) \rangle. \end{aligned} \quad (2.27)$$

Definition 2.6. Let $C_+^{1,2}(\Lambda)$ be the set of functions $u : \Lambda \rightarrow \mathbb{R}$ such that there exists a $\widehat{u} \in C_b^{1,2}(\widehat{\Lambda})$ satisfying $u = \widehat{u}$ on Λ . For such a case, define

$$\partial_t u(t, \mathbf{x}) = \partial_t \widehat{u}(t, \mathbf{x}), \quad \partial_{\mathbf{x}} u(t, \mathbf{x}) = \partial_{\mathbf{x}} \widehat{u}(t, \mathbf{x}), \quad \partial_{\mathbf{xx}}^2 u(t, \mathbf{x}) = \partial_{\mathbf{xx}}^2 \widehat{u}(t, \mathbf{x}), \quad \forall (t, \mathbf{x}) \in \Lambda.$$

We emphasize that, by (2.27), $\partial_{\mathbf{xx}}^2 u(t, \mathbf{x})$ is well-defined (or say independent of the choice of $\widehat{u}$) only on (η, η) , rather than on general (η, η') . However, this is sufficient for our purpose.

Define, for $\mathbf{x} \in \mathbb{X}$, $\eta \in \mathbb{X}_t$, $\varphi : \mathbb{T}_-^2 \times \mathbb{X} \rightarrow \mathbb{R}^k$ with appropriate dimension k ,

$$(\mathbf{x} \oplus_t \eta)(s) := \mathbf{x}_s \mathbf{1}_{[0, t)}(s) + \eta_s \mathbf{1}_{[t, T]}(s), \quad s \in \mathbb{T}; \quad \varphi_s^{t, \mathbf{x}} := \varphi(s, t, \mathbf{x}), \quad s \in [t, T]. \quad (2.28)$$

The main result of [60] is the following functional Itô formula.

Proposition 2.7. Suppose Assumption 2.1 holds. Let X be the solution to FSVIE (2.3), $\widetilde{X}$ be the auxiliary process defined by (2.6) and $u \in C_+^{1,2}(\Lambda)$. Then, viewing $\widetilde{X}_t(\omega) \in \mathbb{X}_t$,

$$\begin{aligned} du(t, \widehat{X}^t) &= \left[\partial_t u(t, \widehat{X}^t) + \frac{1}{2} \langle \partial_{\mathbf{xx}}^2 u(t, \widehat{X}^t), (\sigma^{t, X}, \sigma^{t, X}) \rangle + \langle \partial_{\mathbf{x}} u(t, \widehat{X}^t), b^{t, X} \rangle \right] dt \\ &\quad + \langle \partial_{\mathbf{x}} u(t, \widehat{X}^t), \sigma^{t, X} \rangle dW_t, \quad \text{where } \widehat{X}^t := X \oplus_t \widetilde{X}_t. \end{aligned} \quad (2.29)$$

2.4. FBSVIEs with random coefficients

For later purpose, we shall consider a more general FSVIE with random coefficients:

$$\check{X}_t = \mathbf{x}_t + \int_0^t \check{b}(t, r, \omega, \check{X}_\cdot) dr + \int_0^t \check{\sigma}(t, r, \omega, \check{X}_\cdot) dW_r, \quad t \in \mathbb{T}. \quad (2.30)$$

Assumption 2.8. Let $(\check{b}, \check{\sigma}) : \mathbb{T}_-^2 \times \Omega \times \mathbb{X} \rightarrow \mathbb{R}^n \times \mathbb{R}^{n \times d}$ be progressively measurable satisfying:

- (i) The map $\mathbf{x} \mapsto (\check{b}(t, r, \omega, \mathbf{x}), \check{\sigma}(t, r, \omega, \mathbf{x}))$ is uniformly Lipschitz continuous under the norm $\|\cdot\|$.
- (ii) The map $t \mapsto (\check{b}(t, r, \omega, \mathbf{x}), \check{\sigma}(t, r, \omega, \mathbf{x}))$ is differentiable with $(\partial_t \check{b}, \partial_t \check{\sigma})$ also satisfying (i), and

$$I_p^p := \sup_{(t, r) \in \mathbb{T}_-^2} \mathbb{E}[|\check{b}(t, r, \cdot, \mathbf{0})|^p + |\check{\sigma}(t, r, \cdot, \mathbf{0})|^p + |\partial_t \check{b}(t, r, \cdot, \mathbf{0})|^p + |\partial_t \check{\sigma}(t, r, \cdot, \mathbf{0})|^p] < \infty. \quad (2.31)$$

We have the following result, also due to [57].

Proposition 2.9. Under Assumption 2.8, FSVIE (2.30) admits a unique strong solution $\check{X}$ such that $\check{X}$ is continuous in t and the following estimate holds true:

$$\mathbb{E}[\|\check{X}\|^p] \leq C_p[\|\eta\|^p + I_p^p], \quad \text{where } I_p^p \text{ is defined in (2.31).} \quad (2.32)$$

Similarly, we consider a more general BSVIE with random coefficients:

$$\check{Y}_t = \check{g}(t, \omega) + \int_t^T \check{f}(t, r, \omega, \check{Y}_r, \check{Z}_r^t) dr - \int_t^T \check{Z}_r^t dW_r, \quad t \in \mathbb{T}. \quad (2.33)$$

Assumption 2.10. The map $\check{f} : \mathbb{T}_+^2 \times \Omega \times \mathbb{R}^m \times \mathbb{R}^{m \times d} \rightarrow \mathbb{R}^m$ is progressively measurable and the map $\check{g} : \mathbb{T} \times \Omega \rightarrow \mathbb{R}^m$ is $\mathcal{F}_T$ -measurable satisfying:

$$\check{I}_p^p := \sup_{t \in \mathbb{T}} \mathbb{E} \left[\left(\int_t^T |\check{f}(t, r, \cdot, 0, 0)| dr \right)^p + |\check{g}(t, \cdot)|^p \right] < \infty, \quad (2.34)$$

and the map $(y, z) \mapsto \check{f}(t, r, \omega, y, z)$ is uniformly Lipschitz continuous.

From Shi–Wang–Yong [58], we have the following standard result.

Proposition 2.11. Under Assumption 2.10, BSVIE (2.33) admits a unique strong solution $(\check{Y}, \check{Z})$ such that the following estimate holds true:

$$\sup_{0 \leq t \leq T} \mathbb{E} \left[|\check{Y}_t|^p + \left(\int_t^T |\check{Z}_s^t|^2 ds \right)^{\frac{p}{2}} \right] \leq C_p \check{I}_p^p, \quad \text{where } \check{I}_p^p \text{ is defined in (2.34).} \quad (2.35)$$

Another important property of BSVIEs is the following comparison principle, due to Wang–Yong [64]. For $y, \tilde{y} \in \mathbb{R}^m$, we say $y \leq \tilde{y}$ if $y_i \leq \tilde{y}_i, i = 1, \dots, m$.

Proposition 2.12. For $k = 1, 2$, let $\check{f}^k, \check{g}^k$ satisfy Assumption 2.10 and $(\check{Y}^k, \check{Z}^k)$ be the solutions to the corresponding BSVIE (2.33). Assume $\check{f}^1 \leq \check{f}^2$ and $\check{g}^1 \leq \check{g}^2$. Assume further that, either for $k = 1$ or $k = 2$, $\check{f}^k$ is nondecreasing in y (in the componentwise sense), and $\check{f}_i^k$ does not depend on z_j , for $i \neq j$, where $z_j \in \mathbb{R}^d$ is the j th row of $z \in \mathbb{R}^{m \times d}$. Then we have $\check{Y}_t^1 \leq \check{Y}_t^2, 0 \leq t \leq T$, a.s.

3. The path dependent Feynman–Kac formula

In this section, we are going to establish the relations between PPDEs and FBSVIEs.

3.1. From PPDE to FBSVIE

Recall the FSVIEs (2.3)–(2.6) and BSVIEs (2.8)–(2.11). Recall (1.7), let us introduce the following system of PPDEs:

$$\begin{cases} \mathcal{L}U(t, s, \mathbf{x}) := \partial_s U(t, s, \mathbf{x}) + \frac{1}{2} \langle \partial_{\mathbf{x}\mathbf{x}}^2 U(t, s, \mathbf{x}), (\sigma^{s, \mathbf{x}}, \sigma^{s, \mathbf{x}}) \rangle + \langle \partial_{\mathbf{x}} U(t, s, \mathbf{x}), b^{s, \mathbf{x}} \rangle \\ \quad + f(t, s, \mathbf{x}, U(t, s, \mathbf{x}), \langle \partial_{\mathbf{x}} U(t, s, \mathbf{x}), \sigma^{s, \mathbf{x}} \rangle) = 0, \quad (t, s, \mathbf{x}) \in \mathbb{T}_+^2 \times \mathbb{X}, \\ U(t, T, \mathbf{x}) = g(t, \mathbf{x}), \quad (t, \mathbf{x}) \in [0, T] \times \mathbb{X}, \end{cases} \quad (3.1)$$

where, for $\varphi = b, \sigma, \varphi^{s, \mathbf{x}}$ is defined by (2.28). As we see in (2.29), $\mathbf{x}$ will correspond to $X \oplus_s \tilde{X}_s$, rather than X . However, due to the required adaptiveness, one has

$$\varphi_r^{s, X \oplus_s \tilde{X}_s} = \varphi(r, s, X \oplus_s \tilde{X}_s) = \varphi(r, s, X) = \varphi_r^{s, X}, \quad r \in [s, T], \text{ for } \varphi = b, \sigma.$$

We emphasize that the derivatives in $\mathcal{L}U(t, s, \mathbf{x})$ are with respect to $(s, \mathbf{x})$ only. As mentioned in Introduction, since (3.1) involves the diagonal value $U(s, s, \mathbf{x})$, it is nonlocal in the first time variable t . Alternatively, if we view t as a parameter rather than an independent variable, then (3.1) is an (uncountably) infinite dimensional PPDE system self-interacted through $U(s, s, \mathbf{x})$.

We call $U \in C^0(\mathbb{T}_+^2 \times \mathbb{X})$ a *classical solution* to the PPDE (3.1) if $U(t, \cdot) \in C_+^{1,2}([t, T] \times \mathbb{X})$ for all $t \in \mathbb{T}$, where $C_+^{1,2}([t, T] \times \mathbb{X})$ is defined in the spirit of Definitions 2.5 and 2.6, but restrict to $s \in [t, T]$ only, and (3.1) is satisfied in the classical sense.

Theorem 3.1. *Under Assumptions 2.1 and 2.3, if the PPDE (3.1) has a classical solution U , then, for any $(t, s) \in \mathbb{T}_+^2$ and recalling the $\widehat{X}$ in (2.29),*

$$\widetilde{Y}_s^t = U(t, s, \widehat{X}^s), \quad Y_t = U(t, t, \widehat{X}^t), \quad Z_s^t = \widetilde{Z}_s^t = \langle \partial_{\mathbf{x}} U(t, s, \widehat{X}^s), \sigma^{s,X} \rangle. \quad (3.2)$$

Proof. Fix t . Applying the functional Itô formula (2.29) to $U(t, \cdot, \widehat{X}^\cdot)$, we have

$$\begin{aligned} dU(t, s, \widehat{X}^s) &= \partial_s U(t, s, \widehat{X}^s) ds + \langle \partial_{\mathbf{x}} U(t, s, \widehat{X}^s), b^{s,X} \rangle ds \\ &\quad + \frac{1}{2} \langle \partial_{\mathbf{xx}}^2 U(t, s, \widehat{X}^s), (\sigma^{s,X}, \sigma^{s,X}) \rangle ds + \langle \partial_{\mathbf{x}} U(t, s, \widehat{X}^s), \sigma^{s,X} \rangle dW_s. \end{aligned}$$

Since U satisfies the PPDE (3.1), the above implies that

$$dU(t, s, \widehat{X}^s) = -f(t, s, X, U(t, s, \widehat{X}^s), \langle \partial_{\mathbf{x}} U(t, s, \widehat{X}^s), \sigma^{s,X} \rangle) ds + \langle \partial_{\mathbf{x}} U(t, s, \widehat{X}^s), \sigma^{s,X} \rangle dW_s. \quad (3.3)$$

Note that $U(t, T, X) = g(t, T, X)$, integrating (3.3) over $[t, T]$ we have

$$\begin{aligned} U(t, t, \widehat{X}^t) &= g(t, T, X) + \int_t^T f(t, r, X, U(t, r, \widehat{X}^r), \langle \partial_{\mathbf{x}} U(t, r, \widehat{X}^r), \sigma^{r,X} \rangle) dr \\ &\quad - \int_t^T \langle \partial_{\mathbf{x}} U(t, r, \widehat{X}^r), \sigma^{r,X} \rangle dW_r. \end{aligned}$$

That is, $(\widehat{Y}_t, \widehat{Z}_t^t) := (U(t, t, \widehat{X}^t), \langle \partial_{\mathbf{x}} U(t, t, \widehat{X}^t), \sigma^{t,X} \rangle)$ satisfies BSVIE (2.8). Then, from the uniqueness of BSVIEs we obtain $\widehat{Y} = Y$ and $\widehat{Z} = Z$, hence the last two formulae in (3.2). Moreover, by substituting these into (3.3), we have

$$dU(t, s, \widehat{X}^s) = -f(t, s, X, Y_s, Z_s^t) ds + Z_s^t dW_s.$$

This clearly implies the first formula in (3.2). □

3.2. From FBSVIE to PPDE

In this subsection, we proceed with the opposite direction: constructing $U(t, s, \mathbf{x})$ by using FBSVIEs. We emphasize again that the $\mathbf{x}$ here corresponds to $X \oplus_s \widehat{X}_s$.

First, for any $(s, \mathbf{x}) \in \Lambda$, denote $X_l^{s,\mathbf{x}} := \mathbf{x}_l$, $l \in [0, s]$, and consider the following FBSVIE:

$$\begin{aligned} X_l^{s,\mathbf{x}} &= \mathbf{x}_l + \int_s^l b(l, r, X_r^{s,\mathbf{x}}) dr + \int_s^l \sigma(l, r, X_r^{s,\mathbf{x}}) dW_r; \\ Y_l^{s,\mathbf{x}} &= g(l, X_l^{s,\mathbf{x}}) + \int_l^T f(l, r, X_r^{s,\mathbf{x}}, Y_r^{s,\mathbf{x}}, Z_r^{l,s,\mathbf{x}}) dr - \int_l^T Z_r^{l,s,\mathbf{x}} dW_r, \end{aligned} \quad l \in [s, T]. \quad (3.4)$$

Next, given $(t, s, \mathbf{x}) \in \mathbb{T}_+^2 \times \mathbb{X}$, consider the following standard FBSDE:

$$\begin{aligned} \widetilde{X}_{r,l}^{t,s,\mathbf{x}} &:= \mathbf{x}_l + \int_s^r b(l, r', X_{r'}^{s,\mathbf{x}}) dr' + \int_s^r \sigma(l, r', X_{r'}^{s,\mathbf{x}}) dW_{r'}, \quad s \leq r \leq l \leq T; \\ \widetilde{Y}_l^{t,s,\mathbf{x}} &= g(t, X_l^{s,\mathbf{x}}) + \int_l^T f(t, r, X_r^{s,\mathbf{x}}, Y_r^{s,\mathbf{x}}, \widetilde{Z}_r^{t,s,\mathbf{x}}) dr - \int_l^T \widetilde{Z}_r^{t,s,\mathbf{x}} dW_r, \quad l \in [s, T]. \end{aligned} \quad (3.5)$$

We now define

$$U(t, s, \mathbf{x}) := \widetilde{Y}_s^{t,s,\mathbf{x}}, \quad (t, s, \mathbf{x}) \in \mathbb{T}_+^2 \times \mathbb{X}. \quad (3.6)$$

It is obvious that $\widetilde{Y}_l^{t,s,\mathbf{x}}$ is $\sigma(W_r - W_s, r \in [s, l])$ -measurable, so the above $U(t, s, \mathbf{x})$ is deterministic.

Theorem 3.2. *Under Assumptions 2.1 and 2.3, if the function U defined by (3.6) is continuous in all variables and, for any fixed t , $\partial_{\mathbf{x}}U(t, \cdot)$, $\partial_{\mathbf{xx}}^2U(t, \cdot)$ exist and satisfy the requirements in Definition 2.5 (in the sense of Definition 2.6), then U is a classical solution of the PPDE (3.1).*

Proof. First, note that $U(s, s, \mathbf{x}) = \tilde{Y}_s^{s,s,\mathbf{x}} = Y_s^{s,\mathbf{x}}$. For any $s \leq r \leq l \leq T$, by (2.7) we have

$$X_l^{s,\mathbf{x}} = X_l^{r,\hat{X}^{s,\mathbf{x},r}}, \quad \text{where } \hat{X}_{r'}^{s,\mathbf{x},r} := (X^{s,\mathbf{x}} \oplus \tilde{X}_r^{s,\mathbf{x}})_{r'} := X_{r'}^{s,\mathbf{x}} \mathbf{1}_{[0,r)}(r') + \tilde{X}_{r,r'}^{s,\mathbf{x}} \mathbf{1}_{[r,T]}(r'). \quad (3.7)$$

Then by the uniqueness of BSVIEs and BSDEs we have

$$Y_r^{s,\mathbf{x}} = U(r, r, \hat{X}^{s,\mathbf{x},r}), \quad \tilde{Y}_r^{t,s,\mathbf{x}} = U(t, r, \hat{X}^{s,\mathbf{x},r}). \quad (3.8)$$

We next establish the representation for $\tilde{Z}_r^{t,s,\mathbf{x}}$:

$$\tilde{Z}_r^{t,s,\mathbf{x}} = \langle \partial_{\mathbf{x}}U(t, r, \hat{X}^{s,\mathbf{x},r}), \sigma^{r,\hat{X}^{s,\mathbf{x},r}} \rangle, \quad \text{a.s., } s \leq r \leq T. \quad (3.9)$$

Fix $\delta > 0$ and let $s = s_0 < \dots < s_n = T$ be such that $\Delta s_i := s_i - s_{i-1} \leq \delta$. Denote

$$Y_r^n := U(t, s_{i+1}, \hat{X}^{s,\mathbf{x},r}), \quad Z_r^n := \langle \partial_{\mathbf{x}}U(t, s_{i+1}, \hat{X}^{s,\mathbf{x},r}), \sigma^{r,\hat{X}^{s,\mathbf{x},r}} \rangle, \quad r \in [s_i, s_{i+1}].$$

Note that (Y_r^n, Z_r^n) is $\mathcal{F}_r$ -measurable. Fix t and then apply the functional Itô formula (2.29) to $U(t, s_{i+1}, \hat{X}^{s,\mathbf{x},\cdot})$ (with time variable fixed), we get

$$dY_r^n = \left[\frac{1}{2} \langle \partial_{\mathbf{xx}}^2U(t, s_{i+1}, \hat{X}^{s,\mathbf{x},r}), (\sigma^{r,\hat{X}^{s,\mathbf{x},r}}, \sigma^{r,\hat{X}^{s,\mathbf{x},r}}) \rangle + \langle \partial_{\mathbf{x}}U(t, s_{i+1}, \hat{X}^{s,\mathbf{x},r}), b^{r,\hat{X}^{s,\mathbf{x},r}} \rangle \right] dr + Z_r^n dW_r.$$

Denote $\Delta Y_r^n := Y_r^n - \tilde{Y}_r^{t,s,\mathbf{x}}$, $\Delta Z_r^n := Z_r^n - \tilde{Z}_r^{t,s,\mathbf{x}}$. Note that $\Delta Y_{s_{i+1}}^n = 0$, and

$$\begin{aligned} d\Delta Y_r^n &= \hat{f}(t, r, \hat{X}^{s,\mathbf{x},r}, Y_r^{s,\mathbf{x}}, Z_r^n + \Delta Z_r^n) dr + \Delta Z_r^n dW_r, \quad \text{where} \\ \hat{f}(t, r, \hat{\mathbf{x}}, y, z) &:= \frac{1}{2} \langle \partial_{\mathbf{xx}}^2U(t, s_{i+1}, \hat{\mathbf{x}}), (\sigma^{r,\hat{\mathbf{x}}}, \sigma^{r,\hat{\mathbf{x}}}) \rangle + \langle \partial_{\mathbf{x}}U(t, s_{i+1}, \hat{\mathbf{x}}), b^{r,\hat{\mathbf{x}}} \rangle + f(t, r, \hat{\mathbf{x}}, y, z). \end{aligned} \quad (3.10)$$

By standard BSDE arguments we have

$$\begin{aligned} &\mathbb{E} \left[\sup_{s_i \leq r \leq s_{i+1}} |\Delta Y_r^n|^2 + \int_{s_i}^{s_{i+1}} |\Delta Z_r^n|^2 dr \right] \\ &\leq C \mathbb{E} \left[\left(\int_{s_i}^{s_{i+1}} |\hat{f}(t, r, \hat{X}^{s,\mathbf{x},r}, Y_r^{s,\mathbf{x}}, Z_r^n)|^2 dr \right)^2 \right] \leq C[1 + \|\mathbf{x}\|^{4+2\kappa}] \delta \Delta s_{i+1}, \end{aligned} \quad (3.11)$$

where κ is the generic order of polynomial growth in Definition 2.5. Thus

$$\mathbb{E} \left[\sum_{i=0}^{n-1} \int_{s_i}^{s_{i+1}} |\tilde{Z}_r^{t,s,\mathbf{x}} - \langle \partial_{\mathbf{x}}U(t, s_{i+1}, \hat{X}^{s,\mathbf{x},r}), \sigma^{r,\hat{X}^{s,\mathbf{x},r}} \rangle|^2 dr \right] \leq C[1 + \|\mathbf{x}\|^{4+2\kappa}] \delta.$$

Send $\delta \rightarrow 0$, by the (right) continuity of $\partial_{\mathbf{x}}U$ we obtain (3.9).

Moreover, set $s_1 := s + \delta$, by (3.10) again we have

$$U(t, s + \delta, \mathbf{x}) - U(t, s, \mathbf{x}) = \Delta Y_s^n = -\mathbb{E} \left[\int_s^{s+\delta} \bar{f}(t, r, X^{s,\mathbf{x}}, Y_r^{s,\mathbf{x}}, Z_r^n) dr \right] - R(\delta), \quad (3.12)$$

where

$$\begin{aligned} \bar{f}(t, r, \hat{\mathbf{x}}, y, z) &:= \frac{1}{2} \langle \partial_{\mathbf{xx}}^2U(t, r, \hat{\mathbf{x}}), (\sigma^{r,\hat{\mathbf{x}}}, \sigma^{r,\hat{\mathbf{x}}}) \rangle + \langle \partial_{\mathbf{x}}U(t, r, \hat{\mathbf{x}}), b^{r,\hat{\mathbf{x}}} \rangle + f(t, r, \hat{\mathbf{x}}, y, z), \\ R(\delta) &:= \mathbb{E} \left[\int_s^{s+\delta} [\hat{f}(t, r, X^{s,\mathbf{x}}, Y_r^{s,\mathbf{x}}, Z_r^n + \Delta Z_r^n) - \hat{f}(t, r, X^{s,\mathbf{x}}, Y_r^{s,\mathbf{x}}, Z_r^n)] dr \right. \\ &\quad \left. + \int_s^{s+\delta} [\hat{f}(t, r, X^{s,\mathbf{x}}, Y_r^{s,\mathbf{x}}, Z_r^n) - \bar{f}(t, r, X^{s,\mathbf{x}}, Y_r^{s,\mathbf{x}}, Z_r^n)] dr \right]. \end{aligned}$$

By (3.11) and the regularity of U , we have

$$|R(\delta)|^2 \leq C\delta\mathbb{E}\left[\int_s^{s+\delta} |\Delta Z_r^n|^2 dr\right] + o(\delta^2) \leq C[1 + \|\mathbf{x}\|^{4+2\kappa}]\delta^3 + o(\delta^2).$$

Divide the both sides of (3.12) by δ and send $\delta \rightarrow 0$, by the desired continuity we see that $\partial_s U(t, s, \mathbf{x})$ exists and

$$\begin{aligned} -\partial_s U(t, s, \mathbf{x}) &= \bar{f}(t, s, \mathbf{x}, U(s, s, \mathbf{x}), \langle \partial_{\mathbf{x}} U(t, s, \mathbf{x}), \sigma^{s, \mathbf{x}} \rangle) \\ &= \frac{1}{2} \langle \partial_{\mathbf{xx}}^2 U(t, s, \mathbf{x}), (\sigma^{s, \mathbf{x}}, \sigma^{s, \mathbf{x}}) \rangle + \langle \partial_{\mathbf{x}} U(t, s, \mathbf{x}), b^{s, \mathbf{x}} \rangle + f(t, s, \mathbf{x}, U(s, s, \mathbf{x}), \langle \partial_{\mathbf{x}} U(t, s, \mathbf{x}), \sigma^{s, \mathbf{x}} \rangle). \end{aligned}$$

This implies that $\partial_s U(t, s, \mathbf{x})$ has the desired regularity and $\mathcal{L}U(t, s, \mathbf{x}) = 0$.

Finally, clearly $U(t, T, \mathbf{x}) = \tilde{Y}_T^{t, T, \mathbf{x}} = g(t, T, \mathbf{x})$, thus U is a classical solution of PPDE (3.1). $\square$

3.3. Classical solutions of the PPDE

In this subsection, we provide some sufficient conditions so that the function U defined by (3.6) has the desired regularity, and thus is the unique classical solution of the PPDE (3.1).

We first note that, since the derivatives of U involve càdlàg paths, we shall assume the coefficients b, σ, f, g can be extended to $\widehat{\mathbb{X}}$, and we will use the same notations. The derivatives of f with respect to (y, z) and those of (b, σ, f, g) with respect to the first time variable t are in the standard sense, while those with respect to the second time variable will not be needed. Given an adapted function $\varphi : \widehat{\Lambda} \rightarrow \mathbb{R}$, the derivative with respect to $\mathbf{x}$ is the Fréchet derivative as a linear operator on $\widehat{\mathbb{X}}$:

$$\varphi(t, \mathbf{x} + \eta) - \varphi(t, \mathbf{x}) = \langle D\varphi(t, \mathbf{x}), \eta \rangle + o(\|\eta\|). \quad (3.13)$$

We emphasize that in (2.24) and (2.26) the perturbation path η is on $[t, T]$, while here η is on $[0, T]$ (actually on $[0, t]$ due to the adaptedness). Similarly, we define $D^2\varphi$ as a bilinear operator on $\widehat{\mathbb{X}} \times \widehat{\mathbb{X}}$.

We say $D\varphi$ is bounded if $|\langle D\varphi(t, \mathbf{x}), \eta \rangle| \leq C\|\eta\|$ for all $(t, \mathbf{x}, \eta) \in \widehat{\Lambda} \times \widehat{\mathbb{X}}$, and $D\varphi$ is continuous if, for any $\eta \in \widehat{\mathbb{X}}$, the mapping $(t, \mathbf{x}) \in \widehat{\Lambda} \mapsto \langle D\varphi(t, \mathbf{x}), \eta \rangle$ is continuous. Similarly, $D^2\varphi$ is bounded if $|\langle D^2\varphi(t, \mathbf{x}), (\eta', \eta) \rangle| \leq C\|\eta'\|\|\eta\|$ and continuous if $(t, \mathbf{x}) \in \widehat{\Lambda} \mapsto \langle D^2\varphi(t, \mathbf{x}), (\eta', \eta) \rangle$ is continuous. When the mapping $\mathbf{x} \in \widehat{\mathbb{X}} \mapsto D\varphi(t, \mathbf{x})$ is continuous, one can easily show that

$$\varphi(t, \mathbf{x} + \eta) - \varphi(t, \mathbf{x}) = \int_0^1 \langle D\varphi(t, \mathbf{x} + \theta\eta), \eta \rangle d\theta. \quad (3.14)$$

Moreover, we may switch the order of differentiation: $\partial_t D\varphi = D\partial_t \varphi$, if one of them is continuous.

Assumption 3.3. Assumptions 2.1 and 2.3 hold, and the dependence of b, σ, f, g on $\mathbf{x}$ can be extended to $\widehat{\Lambda}$, stilled using the same notations, such that

- (i) For $\varphi = b, \sigma, g$, φ is twice differentiable in $\mathbf{x}$ with bounded derivatives, and $D^2\varphi$ is uniformly Lipschitz continuous in $\mathbf{x}$;
- (ii) f is jointly twice differentiable in $(\mathbf{x}, y, z)$ with bounded derivatives, and the second-order derivatives are uniformly Lipschitz continuous in $(\mathbf{x}, y, z)$;
- (iii) b, σ are differentiable in the first time variable t , and $\partial_t b, \partial_t \sigma$ satisfy the requirements in (i).

Theorem 3.4. Under Assumption 3.3, the function U defined by (3.6) is the unique classical solution of PPDE (3.1).

Proof. By Theorem 3.2, it suffices to verify the required regularities of U . We shall repeatedly apply Propositions 2.9 and 2.11. In the proof, we may abuse the notations $\Delta X, \tilde{b}$, etc., and we may omit the variable ω . We proceed in three steps.

Step 1. In this step, we show that

$$\langle \partial_{\mathbf{x}} U(t, s, \mathbf{x}), \eta \rangle = \nabla_{\eta} \tilde{Y}_s^{t, s, \mathbf{x}}, \quad (s, \mathbf{x}) \in \widehat{\Lambda}, t \leq s, \eta \in \mathbb{X}_s, \quad (3.15)$$

where $(\nabla_\eta X^{s,\mathbf{x}}, \nabla_\eta Y^{s,\mathbf{x}}, \nabla_\eta \tilde{Y}^{t,s,\mathbf{x}})$ solve the following linear system with random coefficients on $[s, T]$: denoting $\nabla_\eta X_l^{s,\mathbf{x}} := 0$ for $l \in [0, s]$,

$$\begin{aligned} \nabla_\eta X_l^{s,\mathbf{x}} &= \eta_l + \int_s^l \langle Db(l, r, X^{s,\mathbf{x}}), \nabla_\eta X^{s,\mathbf{x}} \rangle dr + \int_s^l \langle D\sigma(l, r, X^{s,\mathbf{x}}), \nabla_\eta X^{s,\mathbf{x}} \rangle dW_r; \\ \nabla_\eta Y_l^{s,\mathbf{x}} &= \langle Dg(l, X^{s,\mathbf{x}}), \nabla_\eta X^{s,\mathbf{x}} \rangle - \int_l^T \nabla_\eta Z_r^{l,s,\mathbf{x}} dW_r \\ &\quad + \int_l^T [\langle Df(\cdot), \nabla_\eta X^{s,\mathbf{x}} \rangle + \partial_y f(\cdot) \nabla_\eta Y_r^{s,\mathbf{x}} + \partial_z f(\cdot) \nabla_\eta Z_r^{l,s,\mathbf{x}}](l, r, X^{s,\mathbf{x}}, Y_r^{s,\mathbf{x}}, Z_r^{l,s,\mathbf{x}}) dr; \\ \nabla_\eta \tilde{Y}_l^{t,s,\mathbf{x}} &= \langle Dg(t, X^{s,\mathbf{x}}), \nabla_\eta X^{s,\mathbf{x}} \rangle - \int_l^T \nabla_\eta \tilde{Z}_r^{t,s,\mathbf{x}} dW_r \\ &\quad + \int_l^T [\langle Df(\cdot), \nabla_\eta X^{s,\mathbf{x}} \rangle + \partial_y f(\cdot) \nabla_\eta Y_r^{s,\mathbf{x}} + \partial_z f(\cdot) \nabla_\eta \tilde{Z}_r^{t,s,\mathbf{x}}](t, r, X^{s,\mathbf{x}}, Y_r^{s,\mathbf{x}}, \tilde{Z}_r^{t,s,\mathbf{x}}) dr. \end{aligned} \quad (3.16)$$

Indeed, first by Propositions 2.9 and 2.11, and by standard BSDE arguments (see, e.g., [74, Chapter 4]) we see that the above system (3.16) is well-posed, and

$$\begin{aligned} \mathbb{E}[\|\nabla_\eta X^{s,\mathbf{x}}\|^p] &\leq C_p \|\eta\|^p; \\ \sup_{l \in [s, T]} \mathbb{E} \left[|\nabla_\eta Y_l^{s,\mathbf{x}}|^p + \left(\int_l^T |\nabla_\eta Z_r^{l,s,\mathbf{x}}|^2 dr \right)^{\frac{p}{2}} \right] &\leq C_p \mathbb{E}[\|\nabla_\eta X^{s,\mathbf{x}}\|^p] \leq C_p \|\eta\|^p; \\ \mathbb{E} \left[\sup_{l \in [s, T]} |\nabla_\eta \tilde{Y}_l^{t,s,\mathbf{x}}|^p + \left(\int_s^T |\nabla_\eta \tilde{Z}_r^{t,s,\mathbf{x}}|^2 dr \right)^{\frac{p}{2}} \right] &\leq C_p \mathbb{E}[\|\nabla_\eta X^{s,\mathbf{x}}\|^p] + C_p \sup_{l \in [s, T]} \mathbb{E}[|\nabla_\eta Y_l^{s,\mathbf{x}}|^p] \leq C_p \|\eta\|^p. \end{aligned} \quad (3.17)$$

Next, denote

$$\Delta_\eta X^{s,\mathbf{x}} := X^{s,\mathbf{x}+\eta} - X^{s,\mathbf{x}} - \nabla_\eta X^{s,\mathbf{x}}, \quad \Delta_\eta \tilde{Y}^{t,s,\mathbf{x}} := \tilde{Y}^{t,s,\mathbf{x}+\eta} - \tilde{Y}^{t,s,\mathbf{x}} - \nabla_\eta \tilde{Y}^{t,s,\mathbf{x}},$$

and similarly for $\Delta_\eta Y^{s,\mathbf{x}}$, $\Delta_\eta Z^{l,s,\mathbf{x}}$, $\Delta_\eta \tilde{Z}^{t,s,\mathbf{x}}$. Then $\Delta_\eta X_l^{s,\mathbf{x}} = 0$ for $l \in [0, s]$, and for $l \in [s, T]$,

$$\begin{aligned} \Delta_\eta X_l^{s,\mathbf{x}} &= \int_s^l \check{b}(l, r, \Delta_\eta X^{s,\mathbf{x}}) dr + \int_s^l \check{\sigma}(l, r, \Delta_\eta X^{s,\mathbf{x}}) dW_r; \\ \Delta_\eta Y_l^{s,\mathbf{x}} &= \check{g}(l, \Delta_\eta X^{s,\mathbf{x}}) + \int_l^T \check{f}(l, r, \Delta_\eta X^{s,\mathbf{x}}, \Delta_\eta Y_r^{s,\mathbf{x}}, \Delta_\eta Z_r^{l,s,\mathbf{x}}) dr - \int_l^T \Delta_\eta Z_r^{l,s,\mathbf{x}} dW_r; \\ \Delta_\eta \tilde{Y}_l^{t,s,\mathbf{x}} &= \check{g}(t, \Delta_\eta X^{s,\mathbf{x}}) + \int_l^T \check{f}(t, r, \Delta_\eta X^{s,\mathbf{x}}, \Delta_\eta Y_r^{s,\mathbf{x}}, \Delta_\eta \tilde{Z}_r^{t,s,\mathbf{x}}) dr - \int_l^T \Delta_\eta \tilde{Z}_r^{t,s,\mathbf{x}} dW_r, \end{aligned} \quad (3.18)$$

where, for $\varphi = b, \sigma, g$,

$$\begin{aligned} \check{\varphi}(l, r, \mathbf{x}') &:= \varphi(l, r, X^{s,\mathbf{x}} + \nabla_\eta X^{s,\mathbf{x}} + \mathbf{x}') - \varphi(l, r, X^{s,\mathbf{x}}) - \langle D\varphi(l, r, X^{s,\mathbf{x}}), \nabla_\eta X^{s,\mathbf{x}} \rangle; \\ \check{f}(l, r, \mathbf{x}', y, z) &:= f(l, r, X^{s,\mathbf{x}} + \nabla_\eta X^{s,\mathbf{x}} + \mathbf{x}', Y_r^{s,\mathbf{x}} + \nabla_\eta Y_r^{s,\mathbf{x}} + y, Z_r^{l,s,\mathbf{x}} + \nabla_\eta Z_r^{l,s,\mathbf{x}} + z) \\ &\quad - [f(\cdot) + \langle Df(\cdot), \nabla_\eta X^{s,\mathbf{x}} \rangle + \partial_y f(\cdot) \nabla_\eta Y_r^{s,\mathbf{x}} + \partial_z f(\cdot) \nabla_\eta Z_r^{l,s,\mathbf{x}}](l, r, X^{s,\mathbf{x}}, Y_r^{s,\mathbf{x}}, Z_r^{l,s,\mathbf{x}}). \end{aligned}$$

Again by Propositions 2.9 and 2.11, recalling (3.14) we have

$$\begin{aligned}
\mathbb{E}[\|\Delta_\eta X_l^{s,\mathbf{x}}\|^p] &\leq C_p \sup_{(l,r)} \sum_{\varphi=b,\sigma} \mathbb{E}[|\check{\varphi}(l,r,\mathbf{0})|^p + |\partial_t \check{\varphi}(l,r,\mathbf{0})|^p] \\
&\leq C_p \sup_{(l,r)} \sum_{\varphi=b,\sigma} \mathbb{E}\left[\left|\int_0^1 \langle D\varphi(l,r,X^{s,\mathbf{x}} + \theta \nabla_\eta X^{s,\mathbf{x}}) - D\varphi(l,r,X^{s,\mathbf{x}}), \nabla_\eta X^{s,\mathbf{x}} \rangle d\theta \right|^p \right. \\
&\quad \left. + \left|\int_0^1 \partial_t [D\varphi(l,r,X^{s,\mathbf{x}} + \theta \nabla_\eta X^{s,\mathbf{x}}) - D\varphi(l,r,X^{s,\mathbf{x}}), \nabla_\eta X^{s,\mathbf{x}}] d\theta \right|^p \right] \\
&\leq C_p \mathbb{E}[\|\nabla_\eta X^{s,\mathbf{x}}\|^{2p}] \leq C_p \|\eta\|^{2p},
\end{aligned}$$

and

$$\begin{aligned}
&\sup_{l \in [s,T]} \mathbb{E}\left[|\Delta_\eta Y_l^{s,\mathbf{x}}|^p + \left(\int_l^T |\Delta_\eta Z_r^{l,s,\mathbf{x}}|^2 dr\right)^{\frac{p}{2}}\right] \\
&\leq C_p \sup_{l \in [s,T]} \mathbb{E}\left[|\check{g}(l, \Delta_\eta X^{s,\mathbf{x}})|^p + \left(\int_l^T |\check{f}(l,r, \Delta_\eta X^{s,\mathbf{x}}, 0, 0)| dr\right)^p\right] \\
&\leq C_p \sup_{l \in [s,T]} \mathbb{E}\left[|\check{g}(l, \mathbf{0})|^p + \left(\int_l^T |\check{f}(l,r, \mathbf{0}, 0, 0)| dr\right)^p + \|\Delta_\eta X^{s,\mathbf{x}}\|^p\right] \\
&\leq C_p \sup_{l \in [s,T]} \mathbb{E}\left[\left(\int_l^T (\|\nabla_\eta X^{s,\mathbf{x}}\| + |\nabla_\eta Y_r^{s,\mathbf{x}}| + |\nabla_\eta Z_r^{l,s,\mathbf{x}}|)^2 dr\right)^p + \|\Delta_\eta X^{s,\mathbf{x}}\|^p\right] \\
&\leq C_p \|\eta\|^{2p}.
\end{aligned}$$

Then it follows from standard BSDE arguments that

$$\begin{aligned}
&\mathbb{E}\left[\sup_{l \in [s,T]} |\Delta_\eta \tilde{Y}_l^{t,s,\mathbf{x}}|^p + \left(\int_s^T |\Delta_\eta \tilde{Z}_r^{t,s,\mathbf{x}}|^2 dr\right)^{\frac{p}{2}}\right] \\
&\leq C_p \mathbb{E}\left[|\check{g}(t, \Delta_\eta X^{s,\mathbf{x}})|^p + \left(\int_s^T |\check{f}(t,r, \Delta_\eta X^{s,\mathbf{x}}, \Delta_\eta Y_r^{s,\mathbf{x}}, 0)| dr\right)^p\right] \\
&\leq C_p \mathbb{E}\left[|\check{g}(t, \Delta_\eta X^{s,\mathbf{x}})|^p + \left(\int_s^T (|\check{f}(t,r, \Delta_\eta X^{s,\mathbf{x}}, 0, 0)| + |\Delta_\eta Y_r^{s,\mathbf{x}}|) dr\right)^p\right] \\
&\leq C_p \mathbb{E}[\|\Delta_\eta X_l^{s,\mathbf{x}}\|^p] + C_p \sup_{l \in [s,T]} \mathbb{E}[|\Delta_\eta Y_l^{s,\mathbf{x}}|^p] \leq C_p \|\eta\|^{2p}.
\end{aligned}$$

In particular, taking $l = s$, this implies

$$|U(t, s, \mathbf{x} + \eta) - U(t, s, \mathbf{x}) - \nabla_\eta \tilde{Y}_s^{t,s,\mathbf{x}}| = |\Delta_\eta \tilde{Y}_s^{t,s,\mathbf{x}}| \leq C \|\eta\|^2,$$

which exactly means (3.15).

Step 2. Denote

$$\begin{aligned}
G(l, \mathbf{x}') &:= \langle Dg(l, X^{s,\mathbf{x}}), \mathbf{x}' \rangle + \langle D^2g(l, X^{s,\mathbf{x}}), (\nabla_{\eta'} X^{s,\mathbf{x}}, \nabla_\eta X^{s,\mathbf{x}}) \rangle; \\
F(l, r, \mathbf{x}', y, z) &:= [Df(\cdot), \mathbf{x}'] + \partial_y f(\cdot) y + \partial_z f(\cdot) z \\
&\quad + \langle D^2f(\cdot), (\nabla_{\eta'} X^{s,\mathbf{x}}, \nabla_\eta X^{s,\mathbf{x}}) \rangle + \partial_y \langle Df(\cdot), \nabla_\eta X^{s,\mathbf{x}} \rangle \nabla_{\eta'} Y_r^{s,\mathbf{x}} \\
&\quad + \partial_z \langle Df(\cdot), \nabla_\eta X^{s,\mathbf{x}} \rangle \nabla_{\eta'} Z_r^{l,s,\mathbf{x}} + \langle D\partial_y f(\cdot), \nabla_{\eta'} X^{s,\mathbf{x}} \rangle \nabla_\eta Y_r^{s,\mathbf{x}} + \partial_y^2 f(\cdot) \nabla_\eta Y_r^{s,\mathbf{x}} \nabla_{\eta'} Y_r^{s,\mathbf{x}} \\
&\quad + \partial_{yz}^2 f(\cdot) \nabla_\eta Y_r^{s,\mathbf{x}} \nabla_{\eta'} Z_r^{l,s,\mathbf{x}} + \langle D\partial_z f(\cdot), \nabla_{\eta'} X^{s,\mathbf{x}} \rangle \nabla_\eta Z_r^{l,s,\mathbf{x}} + \partial_{yz}^2 f(\cdot) \nabla_{\eta'} Y_r^{s,\mathbf{x}} \nabla_\eta Z_r^{l,s,\mathbf{x}} \\
&\quad + \partial_{zz}^2 f(\cdot) \nabla_{\eta'} Z_r^{l,s,\mathbf{x}} \nabla_\eta Z_r^{l,s,\mathbf{x}}](l, r, X^{s,\mathbf{x}}, Y_r^{s,\mathbf{x}}, Z_r^{l,s,\mathbf{x}}).
\end{aligned}$$

Following similar arguments as in Step 1, we can show that

$$\langle \partial_{\mathbf{x}\mathbf{x}}^2 U(t, s, \mathbf{x}), (\eta', \eta) \rangle = \nabla_{\eta', \eta} \tilde{Y}_s^{t, s, \mathbf{x}}, \quad (s, \mathbf{x}) \in \widehat{\Lambda}, t \leq s, \eta', \eta \in \mathbb{X}_s, \quad (3.19)$$

with $\nabla_{\eta', \eta} \tilde{Y}^{t, s, \mathbf{x}}$ solving the following linear system on $[s, T]$: denoting $\nabla_{\eta', \eta} X_l^{s, \mathbf{x}} := 0$ for $l \in [0, s]$,

$$\begin{aligned} \nabla_{\eta', \eta} X_l^{s, \mathbf{x}} &= \int_s^l [\langle Db(l, r, X^{s, \mathbf{x}}), \nabla_{\eta', \eta} X^{s, \mathbf{x}} \rangle + \langle D^2 b(l, r, X^{s, \mathbf{x}}), (\nabla_{\eta'} X^{s, \mathbf{x}}, \nabla_{\eta} X^{s, \mathbf{x}}) \rangle] dr \\ &\quad + \int_s^l [\langle D\sigma(l, r, X^{s, \mathbf{x}}), \nabla_{\eta', \eta} X^{s, \mathbf{x}} \rangle + \langle D^2 \sigma(l, r, X^{s, \mathbf{x}}), (\nabla_{\eta'} X^{s, \mathbf{x}}, \nabla_{\eta} X^{s, \mathbf{x}}) \rangle] dW_r; \\ \nabla_{\eta', \eta} Y_l^{s, \mathbf{x}} &= G(l, \nabla_{\eta', \eta} X^{s, \mathbf{x}}) - \int_l^T \nabla_{\eta', \eta} Z_r^{l, s, \mathbf{x}} dW_r \\ &\quad + \int_l^T F(l, r, \nabla_{\eta', \eta} X^{s, \mathbf{x}}, \nabla_{\eta', \eta} Y_r^{s, \mathbf{x}}, \nabla_{\eta', \eta} Z_r^{l, s, \mathbf{x}}) dr; \\ \nabla_{\eta', \eta} \tilde{Y}_l^{t, s, \mathbf{x}} &= G(l, \nabla_{\eta', \eta} X^{s, \mathbf{x}}) - \int_l^T \nabla_{\eta', \eta} \tilde{Z}_r^{t, s, \mathbf{x}} dW_r \\ &\quad + \int_l^T F(t, r, \nabla_{\eta', \eta} X^{s, \mathbf{x}}, \nabla_{\eta', \eta} Y_r^{s, \mathbf{x}}, \nabla_{\eta', \eta} \tilde{Z}_r^{t, s, \mathbf{x}}) dr. \end{aligned}$$

Step 3. It remains to show that $U, \partial_{\mathbf{x}} U(t, \cdot)$ and $\partial_{\mathbf{x}\mathbf{x}}^2 U(t, \cdot)$ have the desired regularity required in Theorem 3.2. We emphasize that these functions here are already defined in $\widehat{\mathbb{X}}$.

Step 3.1. We first show the continuity in $\mathbf{x}$. Fix $(s, \mathbf{x}) \in \widehat{\Lambda}, t \leq s$, and $\mathbf{x}' \in \widehat{\mathbb{X}}$. By abusing the notation, denote $\Delta_{\mathbf{x}'} X^{s, \mathbf{x}} := X^{s, \mathbf{x} + \mathbf{x}'} - X^{s, \mathbf{x}}$ and similarly for the other terms, and

$$\begin{aligned} \check{\varphi}(l, r, \mathbf{x}') &:= \varphi(l, r, X^{s, \mathbf{x}} + \mathbf{x}') - \varphi(l, r, X^{s, \mathbf{x}}), \quad \text{for } \varphi = b, \sigma, g; \\ \check{f}(l, r, \mathbf{x}', y, z) &:= f(l, r, X^{s, \mathbf{x} + \mathbf{x}'}, Y_r^{s, \mathbf{x}} + y, Z_r^{l, s, \mathbf{x}} + z) - f(l, r, X^{s, \mathbf{x}}, Y_r^{s, \mathbf{x}}, Z_r^{l, s, \mathbf{x}}). \end{aligned}$$

We can see that

$$\begin{aligned} \Delta_{\mathbf{x}'} X_l^{s, \mathbf{x}} &= \mathbf{x}'_l, \quad l \in [0, s]; \\ \Delta_{\mathbf{x}'} X_l^{s, \mathbf{x}} &= \mathbf{x}'_l + \int_s^l \check{b}(l, r, \Delta_{\mathbf{x}'} X^{s, \mathbf{x}}) dr + \int_s^l \check{\sigma}(l, r, \Delta_{\mathbf{x}'} X^{s, \mathbf{x}}) dW_r, \quad l \in [s, T]; \end{aligned}$$

and $\Delta_{\mathbf{x}'} Y^{s, \mathbf{x}}, \Delta_{\mathbf{x}'} \tilde{Y}^{t, s, \mathbf{x}}$ satisfy equations similar to the last two equations in (3.18). Following the same arguments as in Step 1, we can easily show that

$$|U(t, s, \mathbf{x} + \mathbf{x}') - U(t, s, \mathbf{x})| \leq C \|\mathbf{x}'\|, \quad \forall \mathbf{x}' \in \widehat{\mathbb{X}}. \quad (3.20)$$

Similarly, for any fixed $\eta, \eta' \in \mathbb{X}_s$ with $\|\eta\|, \|\eta'\| \leq 1$, one can show that $\nabla_{\eta} \tilde{Y}_s^{t, s, \mathbf{x}}$ and $\nabla_{\eta', \eta} \tilde{Y}_s^{t, s, \mathbf{x}}$ are uniformly Lipschitz continuous in $\mathbf{x}$; that is, for any $\mathbf{x}' \in \widehat{\mathbb{X}}$,

$$|\langle \partial_{\mathbf{x}} U(t, s, \mathbf{x} + \mathbf{x}') - \partial_{\mathbf{x}} U(t, s, \mathbf{x}), \eta \rangle| + |\langle \partial_{\mathbf{x}\mathbf{x}}^2 U(t, s, \mathbf{x} + \mathbf{x}') - \partial_{\mathbf{x}\mathbf{x}}^2 U(t, s, \mathbf{x}), (\eta', \eta) \rangle| \leq C \|\mathbf{x}'\|. \quad (3.21)$$

Step 3.2. We next show the right continuity in s . Recall (3.7) and (3.8), we have

$$\tilde{Y}_l^{t, s, \mathbf{x}} = U(t, s + \delta, \widehat{X}^{s, \mathbf{x}, s + \delta}) + \int_l^{s + \delta} f(t, r, X^{s, \mathbf{x}}, Y_r^{s, \mathbf{x}}, \tilde{Z}_r^{t, s, \mathbf{x}}) dr - \int_l^{s + \delta} \tilde{Z}_r^{t, s, \mathbf{x}} dW_r,$$

for $l \in [s, s + \delta]$. Then by Propositions 2.9 and 2.11, we get

$$\begin{aligned}
& |U(t, s, \mathbf{x}) - U(t, s + \delta, \mathbf{x})|^2 \\
&= \mathbb{E}[\|\mathbb{E}_s[\tilde{Y}_s^{t,s,\mathbf{x}} - U(t, s + \delta, \mathbf{x})]\|^2] \\
&\leq C\mathbb{E}\left[|U(t, s + \delta, \widehat{X}^{s,\mathbf{x},s+\delta}) - U(t, s + \delta, \mathbf{x})|^2 + \left(\int_s^{s+\delta} |f(t, r, X^{s,\mathbf{x}}, Y_r^{s,\mathbf{x}}, \tilde{Z}_r^{t,s,\mathbf{x}})| dr\right)^2\right] \\
&\leq C\mathbb{E}\left[\|\widehat{X}^{s,\mathbf{x},s+\delta} - \mathbf{x}\|^2 + \delta \int_s^{s+\delta} (1 + \|X^{s,\mathbf{x}}\|^2 + |Y_r^{s,\mathbf{x}}|^2 + |\tilde{Z}_r^{t,s,\mathbf{x}}|^2) dr\right] \\
&\leq C(1 + \|\mathbf{x}\|^2)\delta.
\end{aligned} \tag{3.22}$$

Thus

$$|U(t, s + \delta, \mathbf{x}) - U(t, s, \mathbf{x})| \leq C(1 + \|\mathbf{x}\|)\sqrt{\delta}. \tag{3.23}$$

Similarly, fix $\eta \in \mathbb{X}_s$, by (3.5), (3.7) and (3.8) again, we have

$$\begin{aligned}
\nabla_\eta \tilde{Y}_{s+\delta}^{t,s,\mathbf{x}} &= \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} [\tilde{Y}_{s+\delta}^{t,s,\mathbf{x}+\varepsilon\eta} - \tilde{Y}_{s+\delta}^{t,s,\mathbf{x}}] = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} [U(t, s + \delta, \widehat{X}^{s,\mathbf{x}+\varepsilon\eta,s+\delta}) - U(t, s + \delta, \widehat{X}^{s,\mathbf{x},s+\delta})] \\
&= \langle \partial_{\mathbf{x}} U(t, s + \delta, \widehat{X}^{s,\mathbf{x},s+\delta}), \nabla_\eta \widehat{X}^{s,\mathbf{x},s+\delta} \rangle,
\end{aligned}$$

where

$$\begin{aligned}
\nabla_\eta \widehat{X}_l^{s,\mathbf{x},s+\delta} &:= \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} [\widehat{X}_l^{s,\mathbf{x}+\varepsilon\eta,s+\delta} - \widehat{X}_l^{s,\mathbf{x},s+\delta}] = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} [\tilde{X}_{s+\delta,l}^{s,\mathbf{x}+\varepsilon\eta} - \tilde{X}_{s+\delta,l}^{s,\mathbf{x}}] \\
&= \eta_l + \int_s^{s+\delta} \langle \partial_{\mathbf{x}} b(l, r, X^{s,\mathbf{x}}), \nabla_\eta X^{s,\mathbf{x}} \rangle dr + \int_s^{s+\delta} \langle \partial_{\mathbf{x}} \sigma(l, r, X^{s,\mathbf{x}}), \nabla_\eta X^{s,\mathbf{x}} \rangle dW_r,
\end{aligned}$$

with $\nabla_\eta X^{s,\mathbf{x}}$ determined by (3.16). From the above and (3.16), note that, for $l \in [s, s + \delta]$,

$$\begin{aligned}
\nabla_\eta \tilde{Y}_l^{t,s,\mathbf{x}} &= \langle \partial_{\mathbf{x}} U(t, s + \delta, \widehat{X}^{s,\mathbf{x},s+\delta}), \nabla_\eta \widehat{X}^{s,\mathbf{x},s+\delta} \rangle - \int_l^{s+\delta} \nabla_\eta \tilde{Z}_r^{t,s,\mathbf{x}} dW_r \\
&\quad + \int_\tau^{s+\delta} [Df(\cdot), \nabla_\eta X^{s,\mathbf{x}} + \partial_y f(\cdot) \nabla_\eta Y_r^{s,\mathbf{x}} + \partial_z f(\cdot) \nabla_\eta \tilde{Z}_r^{t,s,\mathbf{x}}](t, r, X^{s,\mathbf{x}}, Y_r^{s,\mathbf{x}}, \tilde{Z}_r^{t,s,\mathbf{x}}) dr.
\end{aligned}$$

Then similar to (3.22) we have

$$\begin{aligned}
& |\langle \partial_{\mathbf{x}} U(t, s, \mathbf{x}), \eta \rangle - \langle \partial_{\mathbf{x}} U(t, s + \delta, \mathbf{x}), \eta \rangle|^2 \\
&= |\nabla_\eta \tilde{Y}_s^{t,s,\mathbf{x}} - \langle \partial_{\mathbf{x}} U(t, s + \delta, \mathbf{x}), \eta \rangle|^2 \\
&\leq C\mathbb{E}\left[|\langle \partial_{\mathbf{x}} U(t, s + \delta, \widehat{X}^{s,\mathbf{x},s+\delta}), \nabla_\eta \widehat{X}^{s,\mathbf{x},s+\delta} \rangle - \langle \partial_{\mathbf{x}} U(t, s + \delta, \mathbf{x}), \eta \rangle|^2\right. \\
&\quad \left.+ \left(\int_s^{s+\delta} |Df(\cdot), \nabla_\eta X^{s,\mathbf{x}} + \partial_y f(\cdot) \nabla_\eta Y_r^{s,\mathbf{x}} + \partial_z f(\cdot) \nabla_\eta \tilde{Z}_r^{t,s,\mathbf{x}}| dr\right)^2\right] \\
&\leq C\mathbb{E}\left[\|\widehat{X}^{s,\mathbf{x},s+\delta} - \mathbf{x}\|^2 \|\nabla_\eta \widehat{X}^{s,\mathbf{x},s+\delta}\|^2 + \|\nabla_\eta \widehat{X}^{s,\mathbf{x},s+\delta} - \eta\|^2\right. \\
&\quad \left.\times \delta \int_s^{s+\delta} (\|\nabla_\eta X^{s,\mathbf{x}}\|^2 + |\nabla_\eta Y_r^{s,\mathbf{x}}|^2 + |\nabla_\eta \tilde{Z}_r^{t,s,\mathbf{x}}|^2) dr\right] \\
&\leq C(1 + \|\mathbf{x}\|^2)\|\eta\|^2\delta.
\end{aligned}$$

Thus

$$|\langle \partial_{\mathbf{x}} U(t, s + \delta, \mathbf{x}), \eta \rangle - \langle \partial_{\mathbf{x}} U(t, s, \mathbf{x}), \eta \rangle| \leq C(1 + \|\mathbf{x}\|)\|\eta\|\sqrt{\delta}.$$

Similarly, by using (3.19) we can show that

$$\left| \langle \partial_{\mathbf{x}\mathbf{x}}^2 U(t, s + \delta, \mathbf{x}), (\eta', \eta) \rangle - \langle \partial_{\mathbf{x}\mathbf{x}}^2 U(t, s, \mathbf{x}), (\eta', \eta) \rangle \right| \leq C(1 + \|\mathbf{x}\|) \|\eta'\| \|\eta\| \sqrt{\delta}.$$

Step 3.3. Finally, by (2.9) and standard BSDE arguments we have

$$|U(t - \delta, s, \mathbf{x}) - U(t, s, \mathbf{x})| = |\tilde{Y}_s^{t-\delta, s, \mathbf{x}} - \tilde{Y}_s^{t, s, \mathbf{x}}| \leq C(1 + \|\mathbf{x}\|) \rho(\delta).$$

This, together with (3.20) and (3.23), implies that U is continuous in all variables $(t, s, \mathbf{x})$. $\square$

Combining Theorems 3.1 and 3.4, under Assumption 3.3, the *path dependent Feynman–Kac formula* of FBSVIEs is established in the contexts of classical solutions. In Section 6.3 below, we shall obtain a more explicit representation formula for $\partial_{\mathbf{x}} U(t, s, \mathbf{x})$.

4. Viscosity solution of the PPDE

Inspired by Proposition 2.12, in this section, we investigate viscosity solutions for the PPDE system (3.1) in the case $m = 1$. Since the state space $\mathbb{X}$ is not locally compact here, we shall take the approach of Ekren–Keller–Touzi–Zhang [23], rather than the standard approach of Crandall–Ishii–Lions [15]. However, we shall emphasize that the paths here are on the whole interval $[0, T]$, due to the Volterra nature of the state process, which is different from the setting in [23]. In particular, our work covers the PPDE in Viens–Zhang [60] (under some stronger technical conditions though).

Throughout this section, we shall assume the following.

Assumption 4.1. Let Assumptions 2.1 and 2.3 hold and $m = 1$. Moreover,

- (i) f, g are bounded and uniformly Lipschitz continuous in $\mathbf{x}$.
- (ii) f is nondecreasing in y .

We remark that the monotonicity condition in Assumption 4.1(ii) is essentially the *proper* condition in [15] for elliptic PDEs. For standard parabolic PDE like (1.2), such a condition is redundant because, for any Lipschitz continuous function $f, \tilde{u}(t, x) := e^{-\lambda t} u(t, x)$ will satisfy a PDE whose corresponding $\tilde{f}$ is nondecreasing in y whenever λ is large enough. However, due to the two-time variable structure, this change variable technique does not work for PPDE (3.1). Indeed, if we remove the monotonicity condition, the comparison principle may fail even for classical solutions.

Let $C_b^0(\mathbb{T}_+^2 \times \mathbb{X})$ denote the set of functions $U : \mathbb{T}_+^2 \times \mathbb{X} \rightarrow \mathbb{R}$ such that U is bounded, uniformly continuous in all variables, and progressively measurable. Following the arguments in the proof of Theorem 3.4, Step 3, we have the following.

Lemma 4.2. *Under Assumption 4.1, the function U defined by (3.6) is in $C_b^0(\mathbb{T}_+^2 \times \mathbb{X})$.*

To introduce our notion of viscosity solutions, for any $U \in C_b^0(\mathbb{T}_+^2 \times \mathbb{X})$ and $\phi \in C_+^{1,2}(\Lambda)$, define

$$\begin{aligned} \mathcal{L}_U \phi(t, s, \mathbf{x}) &:= [\mathcal{L}_U \phi](t, s, \mathbf{x}) \\ &:= \partial_s \phi(s, \mathbf{x}) + \frac{1}{2} \langle \partial_{\mathbf{x}\mathbf{x}}^2 \phi(s, \mathbf{x}), (\sigma^{s, \mathbf{x}}, \sigma^{s, \mathbf{x}}) \rangle + \langle \partial_{\mathbf{x}} \phi(s, \mathbf{x}), b^{s, \mathbf{x}} \rangle \\ &\quad + f(t, s, \mathbf{x}, U(s, s, \mathbf{x}), \langle \partial_{\mathbf{x}} \phi(s, \mathbf{x}), \sigma^{s, \mathbf{x}} \rangle). \end{aligned} \quad (4.1)$$

We emphasize that we use $U(s, s, \mathbf{x})$ instead of $\phi(s, \mathbf{x})$ inside f . It is clear that, for any fixed t ,

$$\mathcal{L}_U U_t(t, s, \mathbf{x}) = \mathcal{L}U(t, s, \mathbf{x}), \quad \text{where } U_t := U(t, \cdot). \quad (4.2)$$

For any $s \in \mathbb{T}$ and $L > 0$, denote $\mathbb{F}^s := \{\mathcal{F}_r^s\}_{r \in [s, T]}$ with $\mathcal{F}_r^s := \sigma(W_l - W_s, l \in [s, r])$. Let $\mathcal{T}_s$ be the set of $\mathbb{F}^s$ -stopping times, $\mathcal{T}_s^+$ the subset of $\tau \in \mathcal{T}_s$ such that $\tau > s$, a.s., $\mathcal{U}_s^L$ the set of $\mathbb{F}^s$ -progressively measurable processes on $[s, T]$ bounded by L , and

$$M_r^\theta := \exp\left(\int_s^r \theta_l dW_l - \frac{1}{2} \int_s^r |\theta_l|^2 dl\right), \quad r \in [s, T], \theta \in \mathcal{U}_s^L. \quad (4.3)$$

Given $U \in C_b^0(\mathbb{T}_+^2 \times \mathbb{X})$ and $(t, s, \mathbf{x}) \in \mathbb{T}_+^2 \times \mathbb{X}$, denote

$$\begin{aligned} \underline{\mathcal{A}}^L U(t, s, \mathbf{x}) &:= \left\{ \phi \in C_+^{1,2}([s, T] \times \mathbb{X}; \mathbb{R}) \mid \exists H \in \mathcal{T}_s^+ \text{ such that} \right. \\ &\quad \left. \phi(s, \mathbf{x}) - U(t, s, \mathbf{x}) = 0 = \inf_{\theta \in \mathcal{U}_s^L} \inf_{H \geq \tau \in \mathcal{T}_s} \mathbb{E}[M_\tau^\theta[\phi - U_t](\tau, \widehat{X}^{s, \mathbf{x}, \tau})] \right\}; \\ \overline{\mathcal{A}}^L U(t, s, \mathbf{x}) &:= \left\{ \phi \in C_+^{1,2}([s, T] \times \mathbb{X}; \mathbb{R}) \mid \exists H \in \mathcal{T}_s^+ \text{ such that} \right. \\ &\quad \left. \phi(s, \mathbf{x}) - U(t, s, \mathbf{x}) = 0 = \sup_{\theta \in \mathcal{U}_s^L} \sup_{H \geq \tau \in \mathcal{T}_s} \mathbb{E}[M_\tau^\theta[\phi - U_t](\tau, \widehat{X}^{s, \mathbf{x}, \tau})] \right\}. \end{aligned} \quad (4.4)$$

We note that, if $\phi \in \underline{\mathcal{A}}^L U(t, s, \mathbf{x})$ with the corresponding $H \in \mathcal{T}_s^+$, then for any $\theta \in \mathcal{U}_s^L$ and $H \geq \tau \in \mathcal{T}_s$, we have

$$M_s^\theta[\phi - U_t](s, \mathbf{x}) = 0 \leq \mathbb{E}[M_\tau^\theta[\phi - U_t](\tau, \widehat{X}^{s, \mathbf{x}, \tau})]. \quad (4.5)$$

Definition 4.3. Let $U \in C_b^0(\mathbb{T}_+^2 \times \mathbb{X})$.

(i) We say U is an L -viscosity subsolution of PPDE (3.1) if

$$\mathcal{L}_U \phi(t, s, \mathbf{x}) \geq 0 \quad \text{for any } (t, s, \mathbf{x}) \in \mathbb{T}_+^2 \times \mathbb{X} \text{ and any } \phi \in \underline{\mathcal{A}}^L U(t, s, \mathbf{x}). \quad (4.6)$$

(ii) We say U is an L -viscosity supersolution of PPDE (3.1) if

$$\mathcal{L}_U \phi(t, s, \mathbf{x}) \leq 0 \quad \text{for any } (t, s, \mathbf{x}) \in \mathbb{T}_+^2 \times \mathbb{X} \text{ and any } \phi \in \overline{\mathcal{A}}^L U(t, s, \mathbf{x}). \quad (4.7)$$

(iii) We say U is an L -viscosity solution of PPDE (3.1) if it is both an L -viscosity subsolution and an L -viscosity supersolution. Moreover, we say U is a viscosity solution of PPDE (3.1) if it is an L -viscosity solution for some $L > 0$.

For consistency, we say U is a classical subsolution (resp., classical supersolution) of PPDE (3.1) if $U_t(\cdot) \in C_+^{1,2}([s, T] \times \mathbb{X}; \mathbb{R})$ and satisfies

$$\mathcal{L}_U U_t(t, s, \mathbf{x}) = \mathcal{L}U(t, s, \mathbf{x}) \geq (\text{resp. } \leq) 0.$$

From now on, we let

$$L_0 \text{ denote the Lipschitz constant of } f \text{ with respect to } z. \quad (4.8)$$

We first have the consistency result.

Proposition 4.4. Assume $U \in C_b^0(\mathbb{T}_+^2 \times \mathbb{X})$ and $U(t, \cdot, \cdot) \in C_+^{1,2}([t, T] \times \mathbb{X})$. Then U is a classical subsolution of PPDE (3.1) if and only if it is a viscosity subsolution of PPDE (3.1).

Proof. We first assume U is an L -viscosity subsolution for some L . Clearly $U_t \in \underline{\mathcal{A}}^L U(t, s, \mathbf{x})$. Then $\mathcal{L}U(t, s, \mathbf{x}) = \mathcal{L}_U U_t(t, s, \mathbf{x}) \geq 0$, which implies that U is a classical subsolution.

On the other hand, assume U is a classical subsolution. For any $\phi \in \underline{\mathcal{A}}^{L_0} U(t, s, \mathbf{x})$ with the corresponding $H \in \mathcal{T}_s^+$, applying the functional Itô formula we have

$$\begin{aligned} d\phi(r, \widehat{X}^{s, \mathbf{x}, r}) &= \left[\partial_r \phi + \frac{1}{2} \langle \partial_{\mathbf{x}\mathbf{x}}^2 \phi, (\sigma^\cdot, \sigma^\cdot) \rangle + \langle \partial_{\mathbf{x}} \phi, b^\cdot \rangle \right] (r, \widehat{X}^{s, \mathbf{x}, r}) dr + \langle \partial_{\mathbf{x}} \phi, \sigma^\cdot \rangle (r, \widehat{X}^{s, \mathbf{x}, r}) dW_r \\ &= [\mathcal{L}_U \phi(t, \cdot) - f(t, \cdot, U(r, \cdot), \langle \partial_{\mathbf{x}} \phi, \sigma^\cdot \rangle)](r, \widehat{X}^{s, \mathbf{x}, r}) dr + \langle \partial_{\mathbf{x}} \phi, \sigma^\cdot \rangle (r, \widehat{X}^{s, \mathbf{x}, r}) dW_r. \end{aligned} \quad (4.9)$$

Similarly, we have

$$dU_t(r, \widehat{X}^{s, \mathbf{x}, r}) = [\mathcal{L}U(t, \cdot) - f(t, \cdot, U(r, \cdot), \langle \partial_{\mathbf{x}} U_t, \sigma^\cdot \rangle)](r, \widehat{X}^{s, \mathbf{x}, r}) dr + \langle \partial_{\mathbf{x}} U_t, \sigma^\cdot \rangle (r, \widehat{X}^{s, \mathbf{x}, r}) dW_r.$$

Denote

$$\Delta Y_r := [\phi - U_t](r, \widehat{X}^{s, \mathbf{x}, r}), \quad \Delta Z_r := \langle \partial_{\mathbf{x}} [\phi - U_t], \sigma^\cdot \rangle (r, \widehat{X}^{s, \mathbf{x}, r}).$$

Then

$$d[\Delta Y_r] = [\mathcal{L}_U \phi - \mathcal{L}U](t, r, \widehat{X}^{s, \mathbf{x}, r}) dr - \theta_r \Delta Z_r dr + \Delta Z_r dW_r,$$

for some $|\theta| \leq L_0$. This implies that

$$d(M_r^\theta \Delta Y_r) = M_r^\theta [\mathcal{L}_U \phi - \mathcal{L}U](t, r, \widehat{X}^{s, \mathbf{x}, r}) dr + M_r^\theta \Delta Z_r dW_r,$$

where M_r^θ is defined by (4.3). Then, for any $\tau \leq H$, by (4.4) (or (4.5)) we have

$$0 \leq \mathbb{E}[M_\tau^\theta \Delta Y_\tau - M_s^\theta \Delta Y_s] = \mathbb{E}\left[\int_s^\tau M_r^\theta [\mathcal{L}_U \phi - \mathcal{L}U](t, r, \widehat{X}^{s, \mathbf{x}, r}) dr\right].$$

Since $\tau \geq t$ is arbitrary and $\mathcal{L}_U \phi - \mathcal{L}U$ is continuous, using the fact that $\mathcal{L}U(t, s, \mathbf{x}) \geq 0$, we have

$$0 \leq [\mathcal{L}_U \phi - \mathcal{L}U](t, s, \mathbf{x}) \leq \mathcal{L}_U \phi(t, s, \mathbf{x}),$$

implying the viscosity subsolution property. □

Next, we have the following existence of the viscosity solutions to PPDE (3.1).

Theorem 4.5. *Under Assumption 4.1, the function U defined by (3.6) is an L_0 -viscosity solution of PPDE (3.1).*

Proof. Without loss of generality, we shall only verify the viscosity subsolution property. Fix $(t, s, \mathbf{x}) \in \mathbb{T}_+^2 \times \mathbb{X}$. Recall (4.9) and (3.5), and denote

$$\Delta Y_r := \phi(r, \widehat{X}^{s, \mathbf{x}, r}) - \widetilde{Y}_r^{t, s, \mathbf{x}}, \quad \Delta Z_r := \langle \partial_{\mathbf{x}} \phi, \sigma \cdot \rangle(r, \widehat{X}^{s, \mathbf{x}, r}) - \widetilde{Z}_r^{t, s, \mathbf{x}}.$$

Then

$$\begin{aligned} d(\Delta Y_r) = & \left[[\mathcal{L}_U \phi(t, \cdot) - f(t, \cdot, U(r, \cdot), \langle \partial_{\mathbf{x}} \phi, \sigma \cdot \rangle)](r, \widehat{X}^{s, \mathbf{x}, r}) \right. \\ & \left. + f(t, r, X_r^{s, \mathbf{x}}, Y_r^{s, \mathbf{x}}, \widetilde{Z}_r^{t, s, \mathbf{x}}) \right] dr + \Delta Z_r dW_r. \end{aligned}$$

Recall (3.7) and (3.8), the above implies

$$d(\Delta Y_r) = \mathcal{L}_U \phi(t, r, \widehat{X}^{s, \mathbf{x}, r}) dr - \theta_r \Delta Z_r dr + \Delta Z_r dW_r,$$

for some $|\theta| \leq L_0$. Then, for the M_r^θ defined by (4.3),

$$d(M_r^\theta \Delta Y_r) = M_r^\theta \mathcal{L}_U \phi(t, r, \widehat{X}^{s, \mathbf{x}, r}) dr + M_r^\theta \Delta Z_r dW_r.$$

Recall (3.8) again that $\Delta Y_r = [\phi - U_t](r, \widehat{X}^{s, \mathbf{x}, r})$. Then, for any $\tau \leq H$, by (4.4) (or (4.5)) we have

$$0 \leq \mathbb{E}[M_\tau^\theta \Delta Y_\tau - M_s^\theta \Delta Y_s] = \mathbb{E}\left[\int_s^\tau M_r^\theta \mathcal{L}_U \phi(t, r, \widehat{X}^{s, \mathbf{x}, r}) dr\right].$$

Since $\tau \geq t$ is arbitrary and $\mathcal{L}_U \phi$ is continuous, we obtain $\mathcal{L}_U \phi(t, s, \mathbf{x}) \geq 0$. □

The key for the viscosity solution theory is the following partial comparison principle.

Theorem 4.6. *Let Assumption 4.1 hold and U_1 (resp., U_2) be a viscosity subsolution (resp., supersolution) of PPDE (3.1). Assume $U_1(t, T, \mathbf{x}) \leq U_2(t, T, \mathbf{x})$ for all $(t, \mathbf{x}) \in \Lambda$. If one of U_1, U_2 is smooth, then $U_1 \leq U_2$.*

Proof. Without loss of generality, we assume that U_2 is a classical supersolution. Fix $\delta > 0$, which will be specified later. We shall first prove $U_1(t, s, \mathbf{x}) \leq U_2(t, s, \mathbf{x})$ whenever $s \in [T - \delta, T]$. Assume by contradiction that

$$c := \sup_{s \in [T - \delta, T], t \in [0, s], \mathbf{x} \in \mathbb{X}} [U_1 - U_2](t, s, \mathbf{x}) > 0. \quad (4.10)$$

Then there exists desired $(t_0, s_0, \mathbf{x}^0)$ such that $[U_1 - U_2](t_0, s_0, \mathbf{x}^0) \geq \frac{c}{2} > 0$. Fix t_0 and denote

$$V(s, \mathbf{x}) := [U_1 - U_2](t_0, s, \mathbf{x}) - \frac{c}{4(T - s_0)}[T - s], \quad \psi(s, \mathbf{x}) := \sup_{\tau \in \mathcal{T}_s} \sup_{\theta \in \mathcal{U}_s^{L_0}} \mathbb{E}[M_\tau^\theta V(\tau, \widehat{X}^{s, \mathbf{x}, \tau})], \quad (4.11)$$

where $\widehat{X}^{s, \mathbf{x}, \tau}$ and M_τ^θ are defined by (3.7) and (4.3), respectively. Similar to Lemma 4.2, ψ is bounded and uniformly continuous in $(s, \mathbf{x})$. Moreover, by standard BSDE results (see [74], e.g.), $\mathcal{Y}_s := \psi(s, \widehat{X}^{s_0, \mathbf{x}^0, s})$ is the solution to the following reflected BSDE:

$$\begin{aligned} \mathcal{Y}_s &= V(T, \widehat{X}^{s_0, \mathbf{x}^0, T}) + L_0 \int_s^T |\mathcal{Z}_r| dr - \int_s^T \mathcal{Z}_r dW_r + K_T - K_s; \\ \mathcal{Y}_s &\geq V(s, \widehat{X}^{s_0, \mathbf{x}^0, s}), \quad [\mathcal{Y}_s - V(s, \widehat{X}^{s_0, \mathbf{x}^0, s})] dK_s = 0. \end{aligned} \quad (4.12)$$

Denote

$$\tau^* := \inf\{s \geq s_0 : \mathcal{Y}_s = V(s, \widehat{X}^{s_0, \mathbf{x}^0, s})\}. \quad (4.13)$$

Then $dK_s = 0$ for $s \in [s_0, \tau^*]$. From (4.11), we note

$$\mathcal{Y}_{s_0} \geq V(s_0, \widehat{X}^{s_0, \mathbf{x}^0, s_0}) = V(s_0, \mathbf{x}^0) = [U_1 - U_2](t_0, s_0, \mathbf{x}^0) - \frac{c}{4} \geq \frac{c}{4} > 0, \quad (4.14)$$

$$\mathcal{Y}_T = [U_1 - U_2](t_0, T, \widehat{X}^{s_0, \mathbf{x}^0, T}) \leq 0. \quad (4.15)$$

Then it is clear that $\mathbb{P}(\tau^* < T) > 0$. Indeed, if $\mathbb{P}(\tau^* < T) = 0$, we have $dK_s \equiv 0$ and the reflected BSDE (4.12) becomes a standard BSDE. Then the terminal condition (4.15) implies $\mathcal{Y}_{s_0} \leq 0$, which contradicts (4.14). Therefore, there exists $\omega^* \in \Omega$ such that

$$\tau^*(\omega^*) < T \quad \text{and} \quad \psi(s^*, \mathbf{x}^*) = V(s^*, \mathbf{x}^*), \quad \text{where } s^* := \tau^*(\omega^*), \mathbf{x}^* := \widehat{X}^{s_0, \mathbf{x}^0, s^*}(\omega^*).$$

Now define

$$\phi(s, \mathbf{x}) := U_2(t_0, s, \mathbf{x}) + \frac{c}{4(T - s_0)}(T - s) + \psi(s^*, \mathbf{x}^*).$$

Then $\phi \in C_+^{1,2}([s_0, T] \times \mathbb{X})$, $\phi(s^*, \mathbf{x}^*) = U_1(t_0, s^*, \mathbf{x}^*)$, and, for any $\theta \in \mathcal{U}_{s^*}^{L_0}$ and any $\tau \in \mathcal{T}_{s^*}$,

$$\mathbb{E}[M_\tau^\theta [\phi - U_1(t_0, \cdot)](s, \widehat{X}^{s^*, \mathbf{x}^*, s})] = \psi(s^*, \mathbf{x}^*) - \mathbb{E}[M_\tau^\theta V(s, \widehat{X}^{s^*, \mathbf{x}^*, s})] \geq 0. \quad (4.16)$$

That is, $\phi \in \underline{\mathcal{A}}^{L_0} U_1(t_0, s^*, \mathbf{x}^*)$, and thus by the viscosity subsolution property of U_1 we have

$$\begin{aligned} 0 &\leq \mathcal{L}_{U_1} \phi(t_0, s^*, \mathbf{x}^*) \\ &= \left[\partial_s \phi + \frac{1}{2} \langle \partial_{\mathbf{xx}}^2 \phi, (\sigma \cdot, \sigma \cdot) \rangle + \langle \partial_{\mathbf{x}} \phi, b \cdot \rangle + f(t_0, \cdot, U_1(s^*, \cdot), \langle \partial_{\mathbf{x}} \phi, \sigma \cdot \rangle) \right] (s^*, \mathbf{x}^*) \\ &= -\frac{c}{4(T - s_0)} + \left[\partial_s U_2(t_0, \cdot) + \frac{1}{2} \langle \partial_{\mathbf{xx}}^2 U_2(t_0, \cdot), (\sigma \cdot, \sigma \cdot) \rangle + \langle \partial_{\mathbf{x}} U_2(t_0, \cdot), b \cdot \rangle \right. \\ &\quad \left. + f(t_0, \cdot, U_1(s^*, \cdot), \langle \partial_{\mathbf{x}} U_2(t_0, \cdot), \sigma \cdot \rangle) \right] (s^*, \mathbf{x}^*). \end{aligned} \quad (4.17)$$

Recall (4.10) we have

$$U_1(s^*, s^*, \mathbf{x}^*) \leq U_2(s^*, s^*, \mathbf{x}^*) + c.$$

Let L denote the Lipschitz constant of f with respect to y . Then, by Assumption 4.1(ii) we have

$$0 \leq -\frac{c}{4\delta} + \left[\partial_s U_2(t_0, \cdot) + \frac{1}{2} \langle \partial_{\mathbf{xx}}^2 U_2(t_0, \cdot), (\sigma \cdot, \sigma \cdot) \rangle + \langle \partial_{\mathbf{x}} U_2(t_0, \cdot), b \cdot \rangle \right]$$

$$\begin{aligned}
& + f(t_0, \cdot, U_2(s^*, \cdot), \langle \partial_{\mathbf{x}} U_2(t_0, \cdot), \sigma \cdot \rangle) \Big] (s^*, \mathbf{x}^*) + Lc \\
& = \mathcal{L}U_2(t_0, s^*, \mathbf{x}^*) - \frac{c}{4\delta} + Lc \leq Lc - \frac{c}{4\delta},
\end{aligned}$$

thanks to the supersolution property of U_2 . Set $\delta := \frac{1}{8L}$, we obtain the desired contradiction, and hence $U_1(t, s, \mathbf{x}) \leq U_2(t, s, \mathbf{x})$ whenever $s \in [T - \delta, T]$.

Now consider the PPDE (3.1) on $[0, T - \delta]$. Since $U_1(t, T - \delta, \mathbf{x}) \leq U_2(t, T - \delta, \mathbf{x})$ for all $(t, \mathbf{x}) \in [0, T - \delta] \times \mathbb{X}$, by the same arguments as above we can show that $U_1(t, s, \mathbf{x}) \leq U_2(t, s, \mathbf{x})$ whenever $s \in [T - 2\delta, T - \delta]$. Repeat the arguments backwardly in time, we show that $U_1 \leq U_2$ over the whole space. $\square$

Our final result in this section is the following comparison principle.

Theorem 4.7. *Let b, σ satisfy the requirements in Assumption 3.3 and f, g satisfy the requirements in Assumption 4.1. Let U_1 (resp., U_2) be a viscosity subsolution (resp., supersolution) of PPDE (3.1). Assume $U_1(t, T, \mathbf{x}) \leq g(t, \mathbf{x}) \leq U_2(t, T, \mathbf{x})$ for all $(t, \mathbf{x}) \in \Lambda$, then $U_1 \leq U_2$ on $\mathbb{T}_+^2 \times \mathbb{X}$.*

Proof. Without loss of generality, we shall assume $U_1(t, T, \mathbf{x}) \leq g(t, \mathbf{x})$ and prove only $U_1 \leq U$, where U is defined by (3.6). We shall approximate (f, g) by (f_n, g_n) , which satisfy Assumption 3.3. However, since $\mathbf{x}$ is a path, the standard mollification does not work and the approximations may not be uniform in terms of $\mathbf{x}$. In particular, we may not have $U_1(t, T, \mathbf{x}) \leq g_n(t, \mathbf{x})$ for all $(t, \mathbf{x}) \in \Lambda$. Therefore, instead of directly applying the partial comparison principle, we will follow its arguments. As in Theorem 4.6, it suffices to prove $U_1(t, s, \mathbf{x}) \leq U(t, s, \mathbf{x})$ for $s \in [T - \delta, T]$, where $\delta := \frac{1}{8L}$. Assume by contradiction that, for some $s_0 \in [T - \delta, T]$, $t_0 \leq s_0$ and $\mathbf{x}^0 \in \mathbb{X}$,

$$c := \sup_{s \in [T - \delta, T], t \in [0, s], \mathbf{x} \in \mathbb{X}} [U_1 - U](t, s, \mathbf{x}) > 0 \quad \text{and} \quad [U_1 - U](t_0, s_0, \mathbf{x}^0) \geq \frac{c}{2}. \quad (4.18)$$

We now mollify (f, g) . By first discretizing $\mathbf{x} \in \mathbb{X}$, one can easily construct f_n, g_n such that, for each n , f_n, g_n satisfy Assumption 3.3, and

$$\sup_{(t, s) \in \mathbb{T}_+^2} \sup_{y, z} |[f_n - f](t, s, \mathbf{x}, y, z)| + \sup_{t \in \mathbb{T}} |[g_n - g](t, T, \mathbf{x})| \leq C \left[\frac{1}{n} + \text{OSC}_{\frac{1}{n}}(\mathbf{x}) \right], \quad (4.19)$$

$$\text{where} \quad \text{OSC}_{\delta}(\mathbf{x}) := \sup_{|t-s| \leq \delta} |\mathbf{x}_t - \mathbf{x}_s|.$$

By Theorem 3.4, the PPDE (3.1) with coefficients (b, σ, f_n, g_n) has a classical solution U_n . As in (4.11), fix t_0 and denote

$$V_n(s, \mathbf{x}) := [U_1 - U_n](t_0, s, \mathbf{x}) - \frac{c}{4(T - s_0)}[T - s], \quad \psi_n(s, \mathbf{x}) := \sup_{\tau \in \mathcal{T}_s} \sup_{\theta \in \mathcal{U}_s^{L_0}} \mathbb{E}[M_{\tau}^{\theta} V_n(\tau, \widehat{X}^{s, \mathbf{x}, \tau})]. \quad (4.20)$$

Denote further $\mathcal{Y}_s^n := \psi_n(s, \widehat{X}^{s_0, \mathbf{x}^0, s})$. Then (4.12) and (4.13) become:

$$\begin{aligned}
\mathcal{Y}_s^n &= V_n(T, \widehat{X}^{s_0, \mathbf{x}^0, T}) + L_0 \int_s^T |\mathcal{Z}_r^n| dr - \int_s^T \mathcal{Z}_r^n dW_r + K_T^n - K_s^n; \\
\mathcal{Y}_s^n &\geq V_n(s, \widehat{X}^{s_0, \mathbf{x}^0, s}), \quad [\mathcal{Y}_s^n - V_n(s, \widehat{X}^{s_0, \mathbf{x}^0, s})] dK_s^n = 0; \\
\tau_n^* &:= \inf\{s \geq s_0 : \mathcal{Y}_s^n = V_n(s, \widehat{X}^{s_0, \mathbf{x}^0, s})\}.
\end{aligned} \quad (4.21)$$

Note that

$$\begin{aligned}
\mathcal{Y}_{s_0}^n &\geq V_n(s_0, \widehat{X}^{s_0, \mathbf{x}^0, s_0}) = V_n(s_0, \mathbf{x}^0) = [U_1 - U_n](s_0, \mathbf{x}^0) - \frac{c}{4} \geq \frac{c}{4} - [U - U_n](s_0, \mathbf{x}^0), \\
\mathcal{Y}_T^n &= [U_1 - U_n](t_0, T, \widehat{X}^{s_0, \mathbf{x}^0, T}) \leq [U - U_n](t_0, T, \widehat{X}^{s_0, \mathbf{x}^0, T}).
\end{aligned}$$

By (4.19) and noting that $\mathbf{x} \in \mathbb{X}$, one can easily show that

$$\lim_{n \rightarrow \infty} \mathbb{E}[|\text{OSC}_{\frac{1}{n}}(\widehat{X}^{s_0, \mathbf{x}^0, \tau})|^2 + |U - U_n|^2(t_0, \tau, \widehat{X}^{s_0, \mathbf{x}^0, \tau})] = 0, \quad \text{for any } \tau \in \mathcal{T}_{s_0}. \quad (4.22)$$

In particular, for any $\varepsilon > 0$ small, this implies that, for n large enough,

$$\mathcal{Y}_{s_0}^n \geq \frac{c}{8} > 0 \quad \text{and} \quad \mathbb{E}[|\mathcal{Y}_T^n|^2 \mathbf{1}_{\{\mathcal{Y}_T^n \geq 0\}}] \leq \varepsilon.$$

Denote $\theta_s^n := L_0 \text{sign}(\mathcal{Z}_s^n)$. Note that $dK_s^n = 0$ for $s \in [s_0, \tau_n^*]$, then $\mathcal{Y}_{s_0}^n = \mathbb{E}[M_{\tau_n^*}^{\theta^n} \mathcal{Y}_{\tau_n^*}^n]$. Thus

$$\begin{aligned} \frac{c}{8} &\leq \mathcal{Y}_{s_0}^n = \mathbb{E}[M_{\tau_n^*}^{\theta^n} \mathcal{Y}_{\tau_n^*}^n \mathbf{1}_{\{\tau_n^* < T\}} + M_T^{\theta^n} \mathcal{Y}_T^n \mathbf{1}_{\{\tau_n^* = T\}}] \\ &\leq \mathbb{E}[M_{\tau_n^*}^{\theta^n} \mathcal{Y}_{\tau_n^*}^n \mathbf{1}_{\{\tau_n^* < T\}} + M_T^{\theta^n} \mathcal{Y}_T^n \mathbf{1}_{\{\tau_n^* = T\}} \mathbf{1}_{\{\mathcal{Y}_T^n \geq 0\}}] \\ &\leq C \left[\sqrt{\mathbb{P}(\tau_n^* < T)} + \sqrt{\varepsilon} \right]. \end{aligned}$$

Then, for $\varepsilon > 0$ small, we have

$$\mathbb{P}(\tau_n^* < T) \geq \frac{c^2}{C}, \quad \text{for all } n \text{ large enough.} \quad (4.23)$$

Moreover, by (4.19) and (4.22), we have $\mathbb{E}[|\Delta_n|^2] \leq \varepsilon^3$ for n large enough, where

$$\Delta_n := \sup_{s \in [s_0, T], t \in [0, s]} \sup_{y, z} |f_n - f|(t, s, \widehat{X}^{s_0, \mathbf{x}^0, s}, y, z) + |U - U_n|(t_0, \tau_n^*, \widehat{X}^{s_0, \mathbf{x}^0, \tau_n^*}), \quad (4.24)$$

which implies that

$$\mathbb{P}(\Delta_n > \varepsilon) \leq \frac{1}{\varepsilon^2} \mathbb{E}[|\Delta_n|^2] \leq \varepsilon.$$

Together with (4.23), for $\varepsilon < \frac{c^2}{C}$, we have $\mathbb{P}(\{\tau_n^* < T\} \cap \{\Delta_n \leq \varepsilon\}) > 0$, for all n large enough. Therefore, there exists ω_n^* such that

$$\tau^*(\omega_n^*) < T, \quad \Delta_n(\omega_n^*) \leq \varepsilon, \quad \text{and} \quad \psi_n(s_n^*, \mathbf{x}_n^*) = V_n(s_n^*, \mathbf{x}_n^*),$$

$$\text{where } s_n^* := \tau_n^*(\omega_n^*), \quad \mathbf{x}_n^* := \widehat{X}^{s_0, \mathbf{x}^0, s_n^*}(\omega_n^*).$$

Now define

$$\phi_n(s, \mathbf{x}) := U_n(t_0, s, \mathbf{x}) + \frac{c}{4(T - s_0)}(T - s) + \psi_n(s_n^*, \mathbf{x}_n^*).$$

Similar to (4.16)–(4.17), we have $\phi_n \in \mathcal{A}^{L_0} U_1(t_0, s_n^*, \mathbf{x}_n^*)$ and then, recalling that $T - s_0 \leq \delta = \frac{1}{8L}$,

$$\begin{aligned} 0 &\leq -2Lc + \left[\partial_s U_n(t_0, \cdot) + \frac{1}{2} \langle \partial_{\mathbf{x}\mathbf{x}}^2 U_n(t_0, \cdot), (\sigma \cdot, \sigma \cdot) \rangle + \langle \partial_{\mathbf{x}} U_n(t_0, \cdot), b \cdot \rangle \right. \\ &\quad \left. + f(t_0, \cdot, U_1(s_n^*, \cdot), \langle \partial_{\mathbf{x}} U_n(t_0, \cdot), \sigma \cdot \rangle) \right] (s_n^*, \mathbf{x}_n^*). \end{aligned}$$

Since U_n is a classical solution of the corresponding PPDE, this implies

$$2Lc \leq [f(t_0, \cdot, U_1(s_n^*, \cdot), \langle \partial_{\mathbf{x}} U_n(t_0, \cdot), \sigma \cdot \rangle) - f_n(t_0, \cdot, U_n(s_n^*, \cdot), \langle \partial_{\mathbf{x}} U_n(t_0, \cdot), \sigma \cdot \rangle)](s_n^*, \mathbf{x}_n^*).$$

Then, by (4.24) and recalling $\Delta_n(\omega_n^*) \leq \varepsilon$, we have

$$2Lc \leq [f(t_0, \cdot, U_1(s_n^*, \cdot), \langle \partial_{\mathbf{x}} U_n(t_0, \cdot), \sigma \cdot \rangle) - f(t_0, \cdot, U(s_n^*, \cdot), \langle \partial_{\mathbf{x}} U_n(t_0, \cdot), \sigma \cdot \rangle)](s_n^*, \mathbf{x}_n^*) + C\varepsilon.$$

Note further that (4.18) leads to $U_1(s_n^*, s_n^*, \mathbf{x}_n^*) \leq U(s_n^*, s_n^*, \mathbf{x}_n^*) + c$, and since f is Lipschitz continuous and nondecreasing in y , we have $2Lc \leq Lc + C\varepsilon$, and thus $Lc \leq C\varepsilon$. This is a desired contradiction since ε can be arbitrarily small, thus $U_1(t, s, \mathbf{x}) \leq U(t, s, \mathbf{x})$ whenever $s \in [T - \delta, T]$.

Now similar to the end of Theorem 4.6, we can show that $U_1 \leq U$ over the whole space. $\square$

Remark 4.8. In Theorem 4.7, the assumptions imposed on b, σ are somewhat strong.

(i) In [23] and the subsequent works [24,25], general semimartingale measures are used to define the corresponding set of test functions $\underline{A}^L U$. In this paper, X is not a semimartingale and inside U we need to use $\widehat{X}$, so in (4.4) we are using the exact process $\widehat{X}$. Consequently, we are not allowed to mollify (b, σ) , which will change the process $\widehat{X}$. Therefore, we assume b, σ are smooth so that, together with mollified (f_n, g_n) , the corresponding PPDE has a classical solution U_n . It will be desirable to allow the $\widehat{X}$ in (4.4) to have more general distributions, in the spirit of [24,25]. Then it may become possible to mollify (b, σ) , and even to allow b, σ to depend on some controls. We shall leave this to future research.

(ii) If $X \equiv B^H$ is a fractional Brownian motion (with the Hurst parameter $H \neq \frac{1}{2}$), namely $b \equiv 0$ and $\sigma(t, r, \mathbf{x}) \equiv \sigma(t, r)$, following our arguments we may prove our results without Assumption 2.1(iii) and Assumption 3.3(iii). Thus, in the setting that the randomness of f, g comes from some fractional Brownian motions, the viscosity theory of the corresponding PPDEs still holds true.

5. Coupled forward backward SVIEs and type-II BSVIEs

In this section, we investigate briefly two more general BSVIEs.

5.1. Coupled FBSVIEs

We now consider the following coupled FBSVIEs:

$$\begin{cases} X_t = \mathbf{x}_t + \int_0^t b(t, r, X_r, Y_r) dr + \int_0^t \sigma(t, r, X_r, Y_r) dW_r, & t \in \mathbb{T}, \\ Y_t = g(t, X_t) + \int_t^T f(t, r, X_r, Y_r, Z_r^t) dr - \int_t^T Z_r^t dW_r, & t \in \mathbb{T}, \end{cases} \quad (5.1)$$

and the associated PPDE:

$$\begin{cases} \partial_s U(t, s, \mathbf{x}) + \frac{1}{2} \langle \partial_{\mathbf{x}\mathbf{x}}^2 U(t, s, \mathbf{x}), (\widehat{\sigma}^{s, \mathbf{x}}, \widehat{\sigma}^{s, \mathbf{x}}) \rangle + \langle \partial_{\mathbf{x}} U(t, s, \mathbf{x}), \widehat{b}^{s, \mathbf{x}} \rangle \\ \quad + f(t, s, \mathbf{x}, U(t, s, \mathbf{x}), \langle \partial_{\mathbf{x}} U(t, s, \mathbf{x}), \widehat{\sigma}^{s, \mathbf{x}} \rangle) = 0, & (t, s) \in \mathbb{T}_+^2, \mathbf{x} \in \mathbb{X}, \\ U(t, T, \mathbf{x}) = g(t, T, \mathbf{x}), & t \in \mathbb{T}, \mathbf{x} \in \mathbb{X}, \end{cases} \quad (5.2)$$

where, for $\varphi = b, \sigma$, $\widehat{\varphi}_r^{s, \mathbf{x}} := \varphi(r, s, \mathbf{x}, U(s, s, \mathbf{x}))$, $r \in [s, T]$.

When T is small, Hamaguchi [33] proved the well-posedness of FBSVIE (5.1) with the forward being a SDE. Following Ma–Protter–Yong [44], in this subsection we prove the well-posedness of (5.1) for arbitrary T , provided PPDE (5.2) has a classical solution. The existence of such classical solution, as well as the well-posedness of (5.1) in general, remains a challenging problem and we shall leave it for future research. For simplicity, in the following result we do not specify the precise technical conditions.

Theorem 5.1. Assume b, σ, f, g are sufficiently smooth with all the related derivatives being bounded. If PPDE (5.2) has a classical solution U with bounded $\partial_{\mathbf{x}} U$, then the coupled FBSVIE (5.1) admits a unique strong solution (X, Y, Z) and the following representation holds:

$$\begin{aligned} Y_t &= U(t, t, \widehat{X}^t), & Z_s^t &= \langle \partial_{\mathbf{x}} U(t, s, \widehat{X}^s), \widehat{\sigma}^{s, \widehat{X}^s} \rangle, & (t, s) &\in \mathbb{T}_+^2, \\ \text{where } \widehat{X}^t &:= X \oplus_t \widetilde{X}_t, & \widetilde{X}_t^s &:= \mathbf{x}_s + \int_0^t b(s, r, X_r, Y_r) dr + \int_0^t \sigma(s, r, X_r, Y_r) dW_r. \end{aligned} \quad (5.3)$$

Proof. We proceed in two steps. Fix an arbitrary T .

Step 1. We first show the existence. Let $\delta > 0$ be a small number which will be specified later. Introduce a mapping Φ on $\mathbb{L}_{\mathbb{F}}^2([0, \delta]; \mathbb{R}^m)$ by $\Phi(\mathbf{y}) := Y^{\mathbf{y}}$, where, for any $\mathbf{y} \in \mathbb{L}_{\mathbb{F}}^2([0, \delta]; \mathbb{R}^m)$,

$$\begin{aligned} X_t^{\mathbf{y}} &= \mathbf{x}_t + \int_0^t b(t, r, X_r^{\mathbf{y}}, \mathbf{y}_r) dr + \int_0^t \sigma(t, r, X_r^{\mathbf{y}}, \mathbf{y}_r) dW_r, & t &\in [0, \delta]; \\ \widetilde{X}_t^{\mathbf{y}, s} &= \mathbf{x}_s + \int_0^t b(s, r, X_r^{\mathbf{y}}, \mathbf{y}_r) dr + \int_0^t \sigma(s, r, X_r^{\mathbf{y}}, \mathbf{y}_r) dW_r, & s &\in [t, T]; \\ \widehat{X}^{\mathbf{y}, t} &:= X^{\mathbf{y}} \oplus_t \widetilde{X}_t^{\mathbf{y}}, & Y_t^{\mathbf{y}} &:= U(t, t, \widehat{X}^{\mathbf{y}, t}), & t &\in [0, \delta]. \end{aligned} \quad (5.4)$$

We emphasize that here we do not need to assume $T \leq \delta$. We shall show that Φ is a contraction mapping when δ is small enough.

Indeed, let $\mathbf{y}, \mathbf{y}' \in \mathbb{L}_{\mathbb{F}}^2([0, \delta]; \mathbb{R}^m)$. Denote $\Delta \mathbf{y} := \mathbf{y} - \mathbf{y}'$, $\Delta X := X^{\mathbf{y}} - X^{\mathbf{y}'}$, and similarly for the other notations. First, applying Proposition 2.9 one can easily have

$$\mathbb{E} \left[\sup_{0 \leq t \leq \delta} |\Delta X_t|^2 \right] \leq C \mathbb{E} \left[\int_0^\delta |\Delta \mathbf{y}_r|^2 dr \right]. \quad (5.5)$$

Then by standard SDE estimates we have

$$\mathbb{E} [|\Delta \tilde{X}_t^s|^2] \leq C \mathbb{E} \left[\int_0^\delta |\Delta \mathbf{y}_r|^2 dr \right], \quad t \in [0, \delta], s \in [t, T].$$

Moreover, since $\partial_t b$, $\partial_t \sigma$ satisfy the desired regularity, following the arguments in [57] we have

$$\mathbb{E} \left[\sup_{s \in [t, T]} |\Delta \tilde{X}_t^s|^2 \right] \leq C \mathbb{E} \left[\int_0^\delta |\Delta \mathbf{y}_r|^2 dr \right], \quad t \in [0, \delta].$$

This, together with (5.5), implies that

$$\mathbb{E} \left[\sup_{s \in [0, T]} |\Delta \hat{X}_s^t|^2 \right] \leq C \mathbb{E} \left[\int_0^\delta |\Delta \mathbf{y}_r|^2 dr \right], \quad t \in [0, \delta].$$

Therefore, since $\partial_{\mathbf{x}} U$ is bounded,

$$\mathbb{E} [|\Delta Y_t|^2] \leq C \mathbb{E} [\|\Delta \hat{X}^t\|^2] \leq C \mathbb{E} \left[\int_0^\delta |\Delta \mathbf{y}_r|^2 dr \right], \quad t \in [0, \delta],$$

and thus

$$\mathbb{E} \left[\int_0^\delta |\Delta Y_t|^2 dt \right] \leq C \delta \mathbb{E} \left[\int_0^\delta |\Delta \mathbf{y}_t|^2 dt \right].$$

Choose $\delta := \frac{1}{2C}$, we see that Φ is a contraction mapping. Consequently, Φ has a unique fixed-point $\mathbf{y}^*$. Denote $X_t^* := X_t^{\mathbf{y}^*}$, $\tilde{X}_t^{*,s} := \tilde{X}_t^{\mathbf{y}^*,s}$, $t \in [0, \delta]$, $s \in [t, T]$.

Next, we introduce another mapping Φ_2 on $\mathbb{L}_{\mathbb{F}}^2([\delta, 2\delta]; \mathbb{R}^m)$ by $\Phi_2(\mathbf{y}) := Y^{\mathbf{y}}$, where, by abusing the notation, for any $\mathbf{y} \in \mathbb{L}_{\mathbb{F}}^2([\delta, 2\delta]; \mathbb{R}^m)$,

$$\begin{aligned} X_t^{\mathbf{y}} &= \tilde{X}_\delta^{*,t} + \int_\delta^t b(t, r, X^* \oplus_\delta X^{\mathbf{y}}, \mathbf{y}_r) dr + \int_\delta^t \sigma(t, r, X^* \oplus_\delta X^{\mathbf{y}}, \mathbf{y}_r) dW_r, \quad t \in [\delta, 2\delta]; \\ \tilde{X}_t^{\mathbf{y},s} &= \tilde{X}_\delta^{*,s} + \int_\delta^t b(s, r, X^* \oplus_\delta X^{\mathbf{y}}, \mathbf{y}_r) dr + \int_\delta^t \sigma(s, r, X^* \oplus_\delta X^{\mathbf{y}}, \mathbf{y}_r) dW_r, \quad s \in [t, T]; \\ \hat{X}^{\mathbf{y},t} &:= X^* \oplus_\delta X^{\mathbf{y}} \oplus_t \tilde{X}_t^{\mathbf{y}}, \quad Y_t^{\mathbf{y}} := U(t, t, \hat{X}^{\mathbf{y},t}), \quad t \in [\delta, 2\delta]. \end{aligned} \quad (5.6)$$

Following the same arguments, we can show that Φ_2 is also a contraction mapping, and thus we may extend the unique fixed-point $\mathbf{y}^*$ to $[0, 2\delta]$. Repeat the arguments and we will obtain a fixed-point $\mathbf{y}^* \in \mathbb{L}_{\mathbb{F}}^2([0, T]; \mathbb{R}^m)$ such that

$$\begin{aligned} X_t^* &= \mathbf{x}_t + \int_0^t b(t, r, X^*, \mathbf{y}_r^*) dr + \int_0^t \sigma(t, r, X^*, \mathbf{y}_r^*) dW_r, \quad t \in [0, T]; \\ \tilde{X}_t^{*,s} &= \mathbf{x}_s + \int_0^t b(s, r, X^*, \mathbf{y}_r^*) dr + \int_0^t \sigma(s, r, X^*, \mathbf{y}_r^*) dW_r, \quad s \in [t, T]; \\ \hat{X}^{*,t} &:= X^* \oplus_t \tilde{X}_t^*, \quad \mathbf{y}_t^* := U(t, t, \hat{X}^{*,t}), \quad t \in [0, T]. \end{aligned} \quad (5.7)$$

Now applying the functional Itô formula (2.29) on $U(t, s, \hat{X}^{*,t})$ and utilizing the PPDE (5.2), one can easily see that $(X^*, \mathbf{y}^*)$ satisfy FBSVIE (5.1) and the representation (5.3) holds true.

Step 2. We next show the uniqueness. Let (X, Y, Z) be an arbitrary solution, and $\tilde{X}, \tilde{Y}, \hat{X}$ be defined in an obvious way: $\tilde{X}^t := X \oplus_t \hat{X}_t$, and

$$\begin{aligned}\tilde{X}_t^s &= \mathbf{x}_s + \int_0^t b(s, r, X., Y_r) dr + \int_0^t \sigma(s, r, X., Y_r) dW_r, \\ \tilde{Y}_s^t &= g(t, X.) + \int_s^T f(t, r, X., Y_r, Z_r^t) dr - \int_s^T Z_r^t dW_r,\end{aligned}\quad (t, s) \in \mathbb{T}_+^2.$$

Now denote $\varphi_r^{s, (\mathbf{x}, y)} := \varphi(r, s, \mathbf{x}, y)$, $r \in [s, T]$, for $\varphi = b, \sigma$, and

$$\begin{aligned}\mathcal{Y}_t &:= U(t, t, \hat{X}^t), \quad \tilde{\mathcal{Y}}_s^t := U(t, s, \hat{X}^s), \quad \mathcal{Z}_s^t := \langle \partial_{\mathbf{x}} U(t, s, \hat{X}^s), \sigma^{s, (X, Y_s)} \rangle; \\ \Delta \mathcal{Y} &:= \mathcal{Y} - Y, \quad \Delta \tilde{\mathcal{Y}} := \tilde{\mathcal{Y}} - \tilde{Y}, \quad \Delta \mathcal{Z} := \mathcal{Z} - Z.\end{aligned}$$

Applying the functional Itô formula (2.29) and then utilizing the PPDE (5.2), we have

$$\begin{aligned}d\tilde{\mathcal{Y}}_s^t &= \mathcal{Z}_s^t dW_s + \left[\partial_s U + \frac{1}{2} \langle \partial_{\mathbf{xx}}^2 U, (\sigma^{s, X, Y_s}, \sigma^{s, X, Y_s}) \rangle + \langle \partial_{\mathbf{x}} U, b^{s, X, Y_s} \rangle \right] (t, s, \hat{X}^s) ds \\ &= \mathcal{Z}_s^t dW_s + \left[\frac{1}{2} \langle \partial_{\mathbf{xx}}^2 U, (\sigma^{s, X, Y_s}, \sigma^{s, X, Y_s}) \rangle + \langle \partial_{\mathbf{x}} U, b^{s, X, Y_s} \rangle \right] (t, s, \hat{X}^s) ds \\ &\quad - \left[\frac{1}{2} \langle \partial_{\mathbf{xx}}^2 U, (\hat{\sigma}^{s, \hat{X}^s}, \hat{\sigma}^{s, \hat{X}^s}) \rangle + \langle \partial_{\mathbf{x}} U, \hat{b}^{s, \hat{X}^s} \rangle + f(\cdot, \mathcal{Y}_s, \langle \partial_{\mathbf{x}} U, \hat{\sigma}^{s, \hat{X}^s} \rangle) \right] (t, s, \hat{X}^s) ds.\end{aligned}$$

Then

$$\begin{aligned}d\Delta \tilde{\mathcal{Y}}_s^t &= \Delta \mathcal{Z}_s^t dW_s + \frac{1}{2} [\langle \partial_{\mathbf{xx}}^2 U, (\sigma^{s, X, Y_s}, \sigma^{s, X, Y_s}) \rangle - \langle \partial_{\mathbf{xx}}^2 U, (\hat{\sigma}^{s, \hat{X}^s}, \hat{\sigma}^{s, \hat{X}^s}) \rangle] (t, s, \hat{X}^s) ds \\ &\quad + \langle \partial_{\mathbf{x}} U, b^{s, X, Y_s} - \hat{b}^{s, \hat{X}^s} \rangle (t, s, \hat{X}^s) ds + [f(\cdot, Y_s, \mathcal{Z}_s^t) - f(\cdot, \mathcal{Y}_s, \langle \partial_{\mathbf{x}} U, \hat{\sigma}^{s, \hat{X}^s} \rangle)] (t, s, \hat{X}^s) ds.\end{aligned}$$

Noting that $\Delta \tilde{\mathcal{Y}}_T^t = 0$, by standard BSDE arguments, this implies

$$\begin{aligned}\mathbb{E} \left[|\Delta \tilde{\mathcal{Y}}_s^t|^2 + \int_s^T |\Delta \mathcal{Z}_r^t|^2 dr \right] &\leq C \mathbb{E} \left[\int_s^T (|\Delta \tilde{\mathcal{Y}}_r^t|^2 + \|\sigma^{r, X, Y_r} - \hat{\sigma}^{r, \hat{X}^r}\|^2 + \|b^{r, X, Y_r} - \hat{b}^{r, \hat{X}^r}\|^2 \right. \\ &\quad \left. + |\Delta \mathcal{Y}_r|^2 + |\Delta \tilde{\mathcal{Y}}_r^t| |Z_r^t - \langle \partial_{\mathbf{x}} U(t, r, \hat{X}^r), \hat{\sigma}^{r, \hat{X}^r} \rangle|) dr \right].\end{aligned}$$

Note that, for $\varphi = b, \sigma$,

$$\begin{aligned}\|\varphi^{r, X, Y_r} - \hat{\varphi}^{r, \hat{X}^r}\| &= \sup_{l \in [r, T]} |\varphi(l, r, X, Y_r) - \varphi(l, r, X, U(r, r, \hat{X}^r))| \leq C |\Delta \mathcal{Y}_r|; \\ |Z_r^t - \langle \partial_{\mathbf{x}} U(t, r, \hat{X}^r), \hat{\sigma}^{r, \hat{X}^r} \rangle| &\leq |\Delta \mathcal{Z}_r^t| + |\langle \partial_{\mathbf{x}} U(t, r, \hat{X}^r), \sigma^{r, X, Y_r} - \hat{\sigma}^{r, \hat{X}^r} \rangle| \\ &\leq |\Delta \mathcal{Z}_r^t| + C \|\sigma^{r, X, Y_r} - \hat{\sigma}^{r, \hat{X}^r}\| \leq |\Delta \mathcal{Z}_r^t| + C |\Delta \mathcal{Y}_r|.\end{aligned}$$

Then

$$\mathbb{E} \left[|\Delta \tilde{\mathcal{Y}}_s^t|^2 + \int_s^T |\Delta \mathcal{Z}_r^t|^2 dr \right] \leq \mathbb{E} \left[\int_s^T [C |\Delta \tilde{\mathcal{Y}}_r^t|^2 + C |\Delta \mathcal{Y}_r|^2 + C |\Delta \tilde{\mathcal{Y}}_r^t| |Z_r^t|] dr \right],$$

which implies

$$\mathbb{E} [|\Delta \tilde{\mathcal{Y}}_s^t|^2] \leq \mathbb{E} \left[|\Delta \tilde{\mathcal{Y}}_s^t|^2 + \frac{1}{2} \int_s^T |\Delta \mathcal{Z}_r^t|^2 dr \right] \leq C \mathbb{E} \left[\int_s^T [|\Delta \tilde{\mathcal{Y}}_r^t|^2 + |\Delta \mathcal{Y}_r|^2] dr \right]. \quad (5.8)$$

Now applying the Grönwall inequality we obtain

$$\mathbb{E} [|\Delta \tilde{\mathcal{Y}}_s^t|^2] \leq C \mathbb{E} \left[\int_s^T |\Delta \mathcal{Y}_r|^2 dr \right], \quad s \in [t, T]. \quad (5.9)$$

Set $s = t$ at above, we have

$$\mathbb{E}[|\Delta \mathcal{Y}_t|^2] = \mathbb{E}[|\Delta \tilde{\mathcal{Y}}_t|^2] \leq C \mathbb{E}\left[\int_t^T |\Delta \mathcal{Y}_r|^2 dr\right], \quad t \in [0, T].$$

Apply the Grönwall inequality again, we have $\Delta \mathcal{Y} = 0$. Plug this into (5.9) and then to (5.8), we see that $\Delta \tilde{\mathcal{Y}} = \Delta \mathcal{Z} = 0$. In particular, this implies that $\mathbf{y} := Y$ is a fixed point in Step 1. By the uniqueness of the fixed point, we see that Y is unique, which implies immediately that X and Z are also unique. $\square$

5.2. Type-II BSVIEs

Let X be the solution to FSVIE (2.3). In this subsection, we consider the following type-II BSVIE:

$$\begin{aligned} Y_t &= g(t, X_t) + \int_t^T f(t, r, X_r, Y_r, Z_r^t, Z_r^t) dr - \int_t^T Z_r^t dW_r, \\ Y_t &= \mathbb{E}[Y_t] + \int_0^t Z_r^t dW_r, \end{aligned} \quad t \in \mathbb{T}. \quad (5.10)$$

We note that here f depends on both Z_r^t and Z_r^t , where Z_r^t for $t \leq r$ is determined by the martingale representation of Y_r , as in the second line of (5.10). The $\mathbb{F}$ -adapted solution to (5.10) is called an M -solution, with M referring to martingale. By Yong [70], under suitable conditions BSVIE (5.10) admits a unique M -solution. Inspired by Wang–Yong [65], we introduce the following PPDE:

$$\begin{cases} \partial_s V(t, s, \mathbf{x}) + \frac{1}{2} \langle \partial_{\mathbf{x}\mathbf{x}}^2 V(t, s, \mathbf{x}), (\sigma^{s,\mathbf{x}}, \sigma^{s,\mathbf{x}}) \rangle + \langle \partial_{\mathbf{x}} V(t, s, \mathbf{x}), b^{s,\mathbf{x}} \rangle = 0, & (t, s) \in \mathbb{T}_-^2, \\ \partial_s U(t, s, \mathbf{x}', \mathbf{x}) + \frac{1}{2} \langle \partial_{\mathbf{x}\mathbf{x}}^2 U(t, s, \mathbf{x}', \mathbf{x}), (\sigma^{s,\mathbf{x}}, \sigma^{s,\mathbf{x}}) \rangle + \langle \partial_{\mathbf{x}} U(t, s, \mathbf{x}', \mathbf{x}), b^{s,\mathbf{x}} \rangle \\ \quad + f(t, s, \mathbf{x}, U(t, s, \mathbf{x}, \mathbf{x}), \langle \partial_{\mathbf{x}} U(t, s, \mathbf{x}', \mathbf{x}), \sigma^{s,\mathbf{x}} \rangle, \langle \partial_{\mathbf{x}} V(t, s, \mathbf{x}'), \sigma^{t,\mathbf{x}'} \rangle) = 0, & (t, s) \in \mathbb{T}_+^2, \\ V(t, t, \mathbf{x}) = U(t, t, \mathbf{x}, \mathbf{x}), \quad U(t, T, \mathbf{x}', \mathbf{x}) = g(t, \mathbf{x}), \quad t \in \mathbb{T}, \mathbf{x}', \mathbf{x} \in \mathbb{X}. \end{cases} \quad (5.11)$$

Note that here $V : \mathbb{T}_-^2 \times \mathbb{X} \rightarrow \mathbb{R}^m$ and $U : \mathbb{T}_+^2 \times \mathbb{X} \times \mathbb{X} \rightarrow \mathbb{R}^m$.

Theorem 5.2. Assume b, σ, f, g are sufficiently smooth with all the related derivatives being bounded, and let $X, \tilde{X}, \hat{X}$ be determined by FSVIE (2.3) in the obvious sense. Assume PPDE (5.11) has a classical solution (V, U) with bounded derivatives. Then the unique M -solution of BSVIE (5.10) satisfies: for any $0 \leq r \leq t \leq s \leq T$,

$$Y_t = U(t, t, \hat{X}^t, \hat{X}^t), \quad Z_s^t = \langle \partial_{\mathbf{x}} U(t, s, \hat{X}^t, \hat{X}^s), \sigma^{s,X} \rangle, \quad Z_r^t = \langle \partial_{\mathbf{x}} V(t, r, \hat{X}^r), \sigma^{r,X} \rangle. \quad (5.12)$$

Proof. Define Y, Z as in (5.12) and $\tilde{Y}_s^t := U(t, s, \hat{X}^t, \hat{X}^s)$, $\tilde{Y}_r^t = V(t, r, \hat{X}^r)$ for $0 \leq r \leq t \leq s \leq T$. We shall verify that they satisfy BSVIE (5.10).

First, fix t , and apply functional Itô formula (2.29) on $V(t, r, \hat{X}^r)$; $r \in [0, t]$, we have

$$d\tilde{Y}_r^t = \left[\partial_r V + \frac{1}{2} \langle \partial_{\mathbf{x}\mathbf{x}}^2 V, (\sigma^{r,X}, \sigma^{r,X}) \rangle + \langle \partial_{\mathbf{x}} V, b^{r,X} \rangle \right] (t, r, \hat{X}^r) ds + Z_r^t dW_r = Z_r^t dW_r,$$

where the second equality is due to (5.11). Since $\tilde{Y}_t^t = Y_t$, this verifies the second line of (5.10).

Next, fix $(t, \hat{X}^t)$, and apply functional Itô formula (2.29) on $U(t, s, \hat{X}^t, \hat{X}^s)$; $s \in [t, T]$, we have

$$\begin{aligned} d\tilde{Y}_s^t &= \left[\partial_s U + \frac{1}{2} \langle \partial_{\mathbf{x}\mathbf{x}}^2 U, (\sigma^{s,X}, \sigma^{s,X}) \rangle + \langle \partial_{\mathbf{x}} U, b^{s,X} \rangle \right] (t, s, \hat{X}^t, \hat{X}^s) ds + Z_s^t dW_s \\ &= -f(t, s, Y_s, Z_s^t, Z_s^t) ds + Z_s^t dW_s, \end{aligned}$$

where the second equality is also due to (5.11). This verifies the first line of (5.10) immediately. $\square$

6. Probabilistic representation of $\partial_{\mathbf{x}} U(t, s, \mathbf{x})$

In this section, we shall investigate linear FBSVIEs more closely and then use it to obtain an explicit representation formula for $\partial_{\mathbf{x}} U(t, s, \mathbf{x})$.

6.1. A duality result for linear FSVIE

For the ease of presentation, in the rest of the paper we restrict to one-dimensional processes only. However, all our results hold true in the multiple dimensional situation, and we provide a multiple dimensional setting in Remark 6.4 below.

Note that the dual space of $C([0, T])$ consists of signed measures on $[0, T]$ (see [73], p. 119, e.g.). That is, for a continuous linear mapping $\Phi : C([0, T]) \rightarrow \mathbb{R}$, there exists a unique function $\check{F}$ on $[0, T]$ with finite variation such that $\langle \Phi, \eta \rangle = \int_0^T \eta_t \check{F}(dt)$. Then we may view the Db , $D\sigma$ in (3.16) as signed measures. For this purpose, in this subsection we consider the following FSVIE:

$$\mathcal{X}_t = \eta_t + \int_0^t \int_0^s \mathcal{X}_r \check{b}(t, s, dr) ds + \int_0^t \int_0^s \mathcal{X}_r \check{\sigma}(t, s, dr) dW_s. \quad (6.1)$$

Here, $\check{b}, \check{\sigma} : (t, s, r, \omega) \in \mathbb{T}_+^3 \times \Omega \rightarrow \mathbb{R}$ are progressively measurable, the adaptedness of ω is with respect to the second time variable s , and the dependence on the third time variable r is right continuous with finite variation. We are interested in the term

$$\mathbb{E} \left[\int_0^T \mathcal{X}_r \check{g}(dr) \right], \quad (6.2)$$

where $\check{g} : \mathbb{T} \times \Omega \rightarrow \mathbb{R}$ is $\mathcal{F}_T$ measurable in ω and right continuous and finite varied in r . Our goal is to find a finite varied function $r \in \mathbb{T} \mapsto \tilde{\mathcal{Y}}(r, 0)$ such that the following duality principle holds:

$$\mathbb{E} \left[\int_0^T \mathcal{X}_r \check{g}(dr) \right] = \int_0^T \eta_r \tilde{\mathcal{Y}}(dr, 0). \quad (6.3)$$

This will give us an explicit representation for the linear mapping $\check{F}$:

$$\Phi = \check{F}(\cdot) = \tilde{\mathcal{Y}}(\cdot, 0). \quad (6.4)$$

We shall approach the problem dynamically. Define, for $0 \leq t \leq r \leq T$,

$$\tilde{\mathcal{X}}_t^r := \eta_r + \int_0^r \int_0^{s \wedge t} \mathcal{X}_l \check{b}(r, s, dl) ds + \int_0^r \int_0^{s \wedge t} \mathcal{X}_l \check{\sigma}(r, s, dl) dW_s. \quad (6.5)$$

Note that $\eta_r = \tilde{\mathcal{X}}_0^r$. In light of (6.3), we want to find $\tilde{\mathcal{Y}}$ such that

$$\mathbb{E}_t \left[\int_t^T \mathcal{X}_r \check{g}(dr) \right] = \int_t^T \tilde{\mathcal{X}}_t^r \tilde{\mathcal{Y}}(dr, t), \quad t \in [0, T]. \quad (6.6)$$

For this purpose, we introduce the following type-II BSVIE:

$$\begin{aligned} \mathcal{Y}_t &= \check{g}(t) - \int_{t \leq l \leq s \leq r \leq T} [\check{b}(r, s, dl) \tilde{\mathcal{Y}}(dr, s) + \check{\sigma}(r, s, dl) \mathcal{Z}(dr, s)] ds - \int_t^T \mathcal{Z}(t, s) dW_s; \\ \tilde{\mathcal{Y}}(t, s) &= \mathcal{Y}_t - \int_s^t \mathcal{Z}(t, r) dW_r, \quad 0 \leq s \leq t. \end{aligned} \quad (6.7)$$

We emphasize that, for fixed t ,

- The mapping $s \in [0, t] \rightarrow \tilde{\mathcal{Y}}(t, s)$ is an $\mathbb{F}$ -martingale;
- The mappings $s \in [t, T] \rightarrow (\tilde{\mathcal{Y}}(s, t), \mathcal{Z}(s, t))$ are $\mathcal{F}_t$ -measurable and finite varied.

The second requirement above, of course, will add difficulty for the existence of solutions, which we shall leave for future research.

Theorem 6.1. *Let $\mathcal{X}$ and $\mathcal{Y}$, $\tilde{\mathcal{Y}}$, $\mathcal{Z}$ be the solution to (6.1) and (6.7), respectively. Then (6.6) holds, and in particular (6.3) holds.*

Proof. We shall only prove (6.3), the arguments for (6.6) are similar.

Since η is continuous, by taking time partitions we have

$$\int_0^T \eta_t \tilde{\mathcal{Y}}(dt, 0) = \lim_{N \rightarrow \infty} \sum_{i=0}^{N-1} [\tilde{\mathcal{Y}}(t_{i+1}, 0) - \tilde{\mathcal{Y}}(t_i, 0)] \eta_{t_i}.$$

Now fix a time partition with large N , by the second line of (6.7), we see that

$$\begin{aligned} \int_0^T \eta_t \tilde{\mathcal{Y}}(dt, 0) &\approx \sum_{i=0}^{N-1} [\tilde{\mathcal{Y}}(t_{i+1}, 0) - \tilde{\mathcal{Y}}(t_i, 0)] \eta_{t_i} = \sum_{i=0}^{N-1} \mathbb{E}[\mathcal{Y}_{t_{i+1}} - \mathcal{Y}_{t_i} | \eta_{t_i}] \\ &= \sum_{i=0}^{N-1} \mathbb{E} \left[[\mathcal{Y}_{t_{i+1}} - \mathcal{Y}_{t_i}] \left[\mathcal{X}_{t_i} - \int_0^{t_i} \int_0^s \mathcal{X}_r \check{b}(t_i, s, dr) ds - \int_0^{t_i} \int_0^s \mathcal{X}_r \check{\sigma}(t_i, s, dr) dW_s \right] \right] \\ &= \sum_{i=0}^{N-1} \mathbb{E} \left[[\mathcal{Y}_{t_{i+1}} - \mathcal{Y}_{t_i}] \mathcal{X}_{t_i} - [\mathcal{Y}(t_{i+1}, t_i) - \mathcal{Y}(t_i, t_i)] \right. \\ &\quad \times \left. \left[\int_0^{t_i} \int_0^s \mathcal{X}_r \check{b}(t_i, s, dr) ds + \int_0^{t_i} \int_0^s \mathcal{X}_r \check{\sigma}(t_i, s, dr) dW_s \right] \right] \\ &= \sum_{i=0}^{N-1} \mathbb{E} \left[[\mathcal{Y}_{t_{i+1}} - \mathcal{Y}_{t_i}] \mathcal{X}_{t_i} - \int_0^{t_i} [\tilde{\mathcal{Y}}(t_{i+1}, s) - \tilde{\mathcal{Y}}(t_i, s)] \left[\int_0^s \mathcal{X}_r \check{b}(t_i, s, dr) \right] ds \right. \\ &\quad \left. - \int_0^{t_i} [\mathcal{Z}(t_{i+1}, s) - \mathcal{Z}(t_i, s)] \left[\int_0^s \mathcal{X}_r \check{\sigma}(t_i, s, dr) \right] ds \right] \\ &\approx \sum_{k=0}^{N-1} \mathbb{E}[\mathcal{Y}_{t_{k+1}} - \mathcal{Y}_{t_k}] \mathcal{X}_{t_k} - \frac{T}{n} \sum_{0 \leq k < j < i \leq N-1} [[\tilde{\mathcal{Y}}(t_{i+1}, t_j) - \tilde{\mathcal{Y}}(t_i, t_j)] \\ &\quad \times [\check{b}(t_i, t_j, t_{k+1}) - \check{b}(t_i, t_j, t_k)] + [\mathcal{Z}(t_{i+1}, t_j) - \mathcal{Z}(t_i, t_j)][\check{\sigma}(t_i, t_j, t_{k+1}) - \check{\sigma}(t_i, t_j, t_k)]] \mathcal{X}_{t_k}. \end{aligned}$$

Here and in the sequel, we are using $\approx$ to denote a difference of $o(1)$ term when $N \rightarrow \infty$. Then

$$\begin{aligned} \int_0^T \eta_t \tilde{\mathcal{Y}}(dt, 0) - \mathbb{E} \left[\int_0^T \mathcal{X}_t \check{g}(dt) \right] \\ \approx \int_0^T \eta_t \tilde{\mathcal{Y}}(dt, 0) - \sum_{k=0}^{N-1} \mathbb{E}[[\check{g}(t_{k+1}) - \check{g}(t_k)] \mathcal{X}_{t_k}] \approx \sum_{k=0}^{N-1} \mathbb{E}[\mathcal{X}_{t_k} I_k], \end{aligned} \quad (6.8)$$

where, for each k ,

$$\begin{aligned} I_k &:= \int_{t_k \leq l \leq s \leq r \leq T} [\check{b}(r, s, dl) \tilde{\mathcal{Y}}(dr, s) + \check{\sigma}(r, s, dl) \mathcal{Z}(dr, s)] ds \\ &\quad - \int_{t_{k+1} \leq l \leq s \leq r \leq T} [\check{b}(r, s, dl) \tilde{\mathcal{Y}}(dr, s) + \check{\sigma}(r, s, dl) \mathcal{Z}(dr, s)] ds \\ &\quad - \frac{T}{n} \sum_{k < j < i \leq N-1} [[\check{b}(t_i, t_j, t_{k+1}) - \check{b}(t_i, t_j, t_k)][\tilde{\mathcal{Y}}(t_{i+1}, t_j) - \tilde{\mathcal{Y}}(t_i, t_j)] \\ &\quad + [\check{\sigma}(t_i, t_j, t_{k+1}) - \check{\sigma}(t_i, t_j, t_k)][\mathcal{Z}(t_{i+1}, t_j) - \mathcal{Z}(t_i, t_j)]]. \end{aligned}$$

One may easily check that

$$I_k \approx \int_{t_k \leq s \leq r \leq T} [[\check{b}(r, s, s) - \check{b}(r, s, t_k)] \tilde{\mathcal{Y}}(dr, s) + [\check{\sigma}(r, s, s) - \check{\sigma}(r, s, t_k)] \mathcal{Z}(dr, s)] ds$$

$$\begin{aligned}
& - \int_{t_{k+1} \leq s \leq r \leq T} [[\check{b}(r, s, s) - \check{b}(r, s, t_{k+1})] \tilde{\mathcal{Y}}(dr, s) + [\check{\sigma}(r, s, s) - \check{\sigma}(r, s, t_{k+1})] \mathcal{Z}(dr, s)] ds \\
& - \int_{t_{k+1} \leq s \leq r \leq T} [[\check{b}(r, s, t_{k+1}) - \check{b}(r, s, t_k)] \tilde{\mathcal{Y}}(dr, s) + [\check{\sigma}(r, s, t_{k+1}) - \check{\sigma}(r, s, t_k)] \mathcal{Z}(dr, s)] ds \\
& = \int_{t_k \leq s \leq r \leq T} [[\check{b}(r, s, s) - \check{b}(r, s, t_k)] \tilde{\mathcal{Y}}(dr, s) + [\check{\sigma}(r, s, s) - \check{\sigma}(r, s, t_k)] \mathcal{Z}(dr, s)] ds \\
& - \int_{t_{k+1} \leq s \leq r \leq T} [[\check{b}(r, s, s) - \check{b}(r, s, t_k)] \tilde{\mathcal{Y}}(dr, s) + [\check{\sigma}(r, s, s) - \check{\sigma}(r, s, t_k)] \mathcal{Z}(dr, s)] ds \\
& = \int_{t_k}^{t_{k+1}} \int_s^T [[\check{b}(r, s, s) - \check{b}(r, s, t_k)] \tilde{\mathcal{Y}}(dr, s) + [\check{\sigma}(r, s, s) - \check{\sigma}(r, s, t_k)] \mathcal{Z}(dr, s)] ds.
\end{aligned}$$

Substituting the above into (6.8) implies that

$$\begin{aligned}
\int_0^T \eta_t \tilde{\mathcal{Y}}(dt, 0) - \mathbb{E} \left[\int_0^T \mathcal{X}_t \check{g}(dt) \right] & \approx \sum_{k=0}^{N-1} \mathbb{E} \left[\int_{t_k}^{t_{k+1}} \int_s^T [[\check{b}(r, s, s) - \check{b}(r, s, t_k)] \tilde{\mathcal{Y}}(dr, s) \right. \\
& \left. + [\check{\sigma}(r, s, s) - \check{\sigma}(r, s, t_k)] \mathcal{Z}(dr, s)] ds \mathcal{X}_{t_k} \right].
\end{aligned} \tag{6.9}$$

Using the fact that the finite variated function is a.e continuous, we get

$$\lim_{t_k \uparrow s} [\check{b}(r, s, s) - \check{b}(r, s, t_k)] = \lim_{t_k \uparrow s} [\check{\sigma}(r, s, s) - \check{\sigma}(r, s, t_k)] = 0, \quad \text{for a.e. } s.$$

Then from (6.9) we see that (6.3) holds true by letting $N \rightarrow \infty$. □

Remark 6.2. In the state dependent case, the measures are degenerate:

$$\check{b}(t, s, dr) = b(t, s) \delta_s(r), \quad \check{\sigma}(t, s, dr) = \sigma(t, s) \delta_s(r), \quad \check{g}(dr) = g(r) dr,$$

with the Dirac measure δ being defined by

$$\delta_s \mathbf{x} = \mathbf{x}_s, \quad s \in [0, T], \mathbf{x} \in C([0, T]).$$

Then (6.1) and (6.7) become

$$\begin{aligned}
\mathcal{X}_t &= \eta_t + \int_0^t b(t, s) \mathcal{X}_s ds + \int_0^t \sigma(t, s) \mathcal{X}_s dW_s; \\
\mathcal{Y}_t &= \check{g}(t) - \int_t^T \int_s^T [b(r, s) \tilde{\mathcal{Y}}(dr, s) + \sigma(r, s) \mathcal{Z}(dr, s)] ds - \int_t^T \mathcal{Z}(t, s) dW_s; \\
\tilde{\mathcal{Y}}(t, s) &= \mathcal{Y}_t - \int_s^t \mathcal{Z}(t, r) dW_r, \quad 0 \leq s \leq t.
\end{aligned} \tag{6.10}$$

Let (Y, Z) denote the solution to the following type-II BSVIE:

$$\begin{aligned}
Y_t &= g(t) + \int_t^T [b(s, t) Y_s + \sigma(s, t) Z(s, t)] ds - \int_t^T Z(t, s) dW_s; \\
Y_t &= \mathbb{E}[Y_T] + \int_0^t Z(t, s) dW_s.
\end{aligned} \tag{6.11}$$

We can easily check that

$$\mathcal{Y}_t = \mathcal{Y}_0 + \int_0^t Y_r dr, \quad \tilde{\mathcal{Y}}(t, s) = \mathcal{Y}_s + \int_s^t \mathbb{E}_s[Y_r] dr, \quad \mathcal{Z}(t, s) = \int_s^t Z(r, s) dr, \quad 0 \leq s \leq t \leq T$$

satisfy the BSVIE in (6.10). Then (6.3) becomes

$$\mathbb{E} \left[\int_0^T g(t) \mathcal{X}_t dt \right] = \mathbb{E} \left[\int_0^T \mathcal{X}_t \check{g}(dt) \right] = \int_0^T \eta_t \tilde{\mathcal{Y}}(dt, 0) = \int_0^T \eta_t \mathbb{E}[Y_t] dt = \mathbb{E} \left[\int_0^T \eta_t Y_t dt \right].$$

This is exactly the duality in Yong [68,70]. So our result here is a generalization of these works.

Remark 6.3. Our result also generalizes the duality between delayed SDEs and anticipated BSDEs in Peng–Yang [52]. Let (X, Y, Z) denote the solution to the following equations:

$$\begin{cases} dX_s^\xi = (\mu_s X_s^\xi + \bar{\mu}_{s-\theta} X_{s-\theta}^\xi) ds + (\sigma_s X_s^\xi + \bar{\sigma}_{s-\theta} X_{s-\theta}^\xi) dW_s, & s \in [t, T + \theta], \\ X_t = \xi, & X_s = 0, & s \in [t - \theta, t]; \\ -dY_s = (\mu_s Y_s + \bar{\mu}_s \mathbb{E}_s[Y_{s+\theta}] + \sigma_s Z_s + \bar{\sigma}_s \mathbb{E}_s[Z_{s+\theta}] + l_s) ds - Z_s dW_s, & s \in [t, T], \\ Y_s = Q_s, & Z_s = P_s, & s \in [T, T + \theta], \end{cases} \quad (6.12)$$

where $\theta > 0$ is a fixed delay time. From [52, Theorem 2.1], we get the duality

$$\langle Y_t, \xi \rangle = \mathbb{E}_t \left[X_T^\xi Q_T + \int_t^T X_s^\xi l_s ds + \int_T^{T+\theta} (Q_s \bar{\mu}_{s-\theta} + P_s \bar{\sigma}_{s-\theta}) X_{s-\theta}^\xi ds \right] := \mathbb{E}_t \left[\int_t^T X_s^\xi \check{g}(ds) \right], \quad (6.13)$$

which shows that Y_t is an explicit representation of the linear functional $\xi \mapsto \mathbb{E}_t[\int_t^T X_s^\xi \check{g}(ds)]$. Since FSVIE (6.1) is more general than the delayed SDE in (6.12), we can also use Theorem 6.1 to give such an explicit representation for $\xi \mapsto \mathbb{E}_t[\int_t^T X_s^\xi \check{g}(ds)]$. Indeed, take

$$\begin{aligned} \check{b}(s, dr) &= \mu_s \delta_s(r) + \bar{\mu}_{s-\theta} \delta_{s-\theta}(r), & \check{\sigma}(s, dr) &= \sigma(s) \delta_s(r) + \bar{\sigma}_{s-\theta} \delta_{s-\theta}(r), & s \in [t + \theta, T], \\ \check{b}(s, dr) &= \mu_s \delta_s(r), & \check{\sigma}(s, dr) &= \sigma(s) \delta_s(r), & s \in [t, t + \theta], \\ \check{g}(dr) &= l(r) dr + \mathbb{E}_r[Q_{r+\theta} \bar{\mu}_r + P_{r+\theta} \bar{\sigma}_r] 1_{[T-\theta, T]}(r) dr + Q(T) \delta_T(r). \end{aligned}$$

Note that $\check{b}(\tau, s, dr), \check{\sigma}(\tau, s, dr)$ are independent of τ , the corresponding BSVIE (6.7) reads

$$\begin{aligned} \mathcal{Y}_\tau &= \check{g}(\tau) - \int_\tau^T \mu_s [\tilde{\mathcal{Y}}(T, s) - \tilde{\mathcal{Y}}(s, s)] ds - \int_{\tau+\theta}^T \bar{\mu}_{s-\theta} [\tilde{\mathcal{Y}}(T, s) - \tilde{\mathcal{Y}}(s, s)] ds \\ &\quad - \int_\tau^T \sigma_s [\mathcal{Z}(T, s) - \mathcal{Z}_s] ds - \int_{\tau+\theta}^T \bar{\sigma}_{s-\theta} [\mathcal{Z}(T, s) - \mathcal{Z}_s] ds - \int_\tau^T \mathcal{Z}_s dW_s; \\ \tilde{\mathcal{Y}}(\tau, s) &= \mathcal{Y}_\tau - \int_s^\tau \mathcal{Z}(\tau, r) dW_r, \quad 0 \leq s \leq \tau. \end{aligned}$$

Then it is easy to check

$$\begin{aligned} Y_t &= \mathcal{Y}(T, t) - \mathcal{Y}_t, & Z_t &= \mathcal{Z}(T, t) - \mathcal{Z}_t, \\ \langle \mathcal{Y}(T, t) - \mathcal{Y}_t, \xi \rangle &= \langle Y_t, \xi \rangle = \mathbb{E}_t \left[\int_t^T X_s^\xi \check{g}(ds) \right]. \end{aligned}$$

Thus Theorem 6.1 covers the duality in [52].

Remark 6.4. The duality (6.6) still holds true in the multidimensional case, where the FSVIE (6.1) and type-II BSVIE (6.7) become

$$\begin{aligned}\mathcal{X}_t &= \eta_t + \int_0^t \int_0^s \mathcal{X}_r \check{b}(t, s, dr) ds + \sum_{j=1}^d \int_0^t \int_0^s \mathcal{X}_r \check{\sigma}^j(t, s, dr) dW_s^j; \\ \mathcal{Y}_t &= \check{g}(t) - \int_{t \leq l \leq s \leq r \leq T} \left[\check{b}(r, s, dl) \tilde{\mathcal{Y}}(dr, s) + \sum_{j=1}^d \check{\sigma}^j(r, s, dl) \mathcal{Z}^j(dr, s) \right] ds - \sum_{j=1}^d \int_t^T \mathcal{Z}^j(t, s) dW_s^j; \\ \tilde{\mathcal{Y}}(t, s) &= \mathcal{Y}_t - \sum_{j=1}^d \int_s^t \mathcal{Z}^j(t, r) dW_r^j, \quad 0 \leq s \leq t,\end{aligned}$$

with $\check{b}, \check{\sigma}^j : (t, s, r, \omega) \in \mathbb{T}_-^3 \times \Omega \rightarrow \mathbb{R}^{n \times n}$ and $\check{g} : \mathbb{T} \times \Omega \rightarrow \mathbb{R}^{m \times n}$ being proper maps.

6.2. An explicit solution for linear BSVIEs

In this subsection, we investigate the following linear BSVIE:

$$\mathcal{Y}_t = \xi_t + \int_t^T [\alpha(t, r) \mathcal{Y}_r + \beta(t, r) \mathcal{Z}_r^t] dr - \int_t^T \mathcal{Z}_r^t dW_r, \quad (6.14)$$

where $\xi : \mathbb{T} \times \Omega \rightarrow \mathbb{R}$, $\alpha, \beta : \mathbb{T}_+^2 \times \Omega \rightarrow \mathbb{R}$ are progressively measurable (omitting the variable ω).

Proposition 6.5. Assume α, β are bounded and $\sup_{t \in \mathbb{T}} \mathbb{E}[|\xi_t|^2] < \infty$. Then

$$\mathcal{Y}_t = \mathbb{E}_t \left[M_T^t \xi_t + \int_t^T \Gamma(t, r) M_T^r \xi_r dr \right], \quad (6.15)$$

where M is the solution to the following SDE:

$$dM_r^t = M_r^t \beta(t, r) dW_r, \quad (t, r) \in \mathbb{T}_+^2; \quad M_t^t = I_m, \quad (6.16)$$

and

$$\Gamma(t, s) := \sum_{n=1}^{\infty} K_n(t, s), \quad K_1(t, s) := M_s^t \alpha(t, s), \quad K_{n+1}(t, s) := \int_t^s K_1(t, r) K_n(r, s) dr. \quad (6.17)$$

Proof. First, by Proposition 2.11 we see that (6.14) is well-posed. Next, since α, β are bounded, it is clear that $\mathbb{E}_t[|K_1(t, s)|^2] \leq C_0 < \infty$. Note that

$$\mathbb{E}_t[|K_{n+1}(t, s)|^2] \leq (s - t) \int_t^s \mathbb{E}_t[|K_1(t, r)|^2 \mathbb{E}_r[|K_n(r, s)|^2]] dr.$$

Then by induction one can easily show that

$$\mathbb{E}_t[|K_{n+1}(t, s)|^2] \leq \frac{C_0^{n+1} (s - t)^{2n}}{(2n - 1)!!}, \quad \text{and thus} \quad \mathbb{E}_t[|\Gamma(t, s)|^2] \leq C < \infty.$$

We now let $(\tilde{\mathcal{Y}}, \tilde{\mathcal{Z}})$ satisfy the following BSDE:

$$\tilde{\mathcal{Y}}_s^t = \xi_t + \int_s^T [\alpha(t, r) \mathcal{Y}_r + \beta(t, r) \tilde{\mathcal{Z}}_r^t] dr - \int_s^T \tilde{\mathcal{Z}}_r^t dW_r. \quad (6.18)$$

Then

$$\mathcal{Y}_t = \tilde{\mathcal{Y}}_t^t, \quad \mathcal{Z}_s^t = \tilde{\mathcal{Z}}_s^t.$$

Apply Itô formula to the mapping $s \mapsto M_s^t \tilde{\mathcal{Y}}_s^t$ on $[t, T]$, we get

$$\mathcal{Y}_t = \tilde{\mathcal{Y}}_t^t = \mathbb{E}_t \left[M_T^t \xi_t + \int_t^T M_s^t \alpha(t, s) \mathcal{Y}_s ds \right]. \quad (6.19)$$

Moreover, note that

$$\Gamma(t, s) = K_1(t, s) + \int_t^s K_1(t, r) \Gamma(r, s) dr.$$

Then

$$\begin{aligned} \mathcal{Y}_t &= \mathbb{E}_t \left[M_T^t \xi_t + \int_t^T K_1(t, s) \mathbb{E}_s \left[M_T^s \xi_s + \int_s^T \Gamma(s, r) M_T^r \xi_r dr \right] ds \right] \\ &= \mathbb{E}_t \left[M_T^t \xi_t + \int_t^T K_1(t, s) M_T^s \xi_s ds + \int_t^T K_1(t, s) \int_s^T \Gamma(s, r) M_T^r \xi_r dr ds \right] \\ &= \mathbb{E}_t \left[M_T^t \xi_t + \int_t^T K_1(t, s) M_T^s \xi_s ds + \int_t^T \int_t^r K_1(t, s) \Gamma(s, r) ds M_T^r \xi_r dr \right] \\ &= \mathbb{E}_t \left[M_T^t \xi_t + \int_t^T \left[K_1(t, s) + \int_t^s K_1(t, r) \Gamma(r, s) dr \right] M_T^s \xi_s ds \right] \\ &= \mathbb{E}_t \left[M_T^t \xi_t + \int_t^T \Gamma(t, s) M_T^s \xi_s ds \right] = \mathcal{Y}_t. \end{aligned}$$

This implies that $(\mathcal{Y}, \mathcal{Z})$ satisfy (6.14). The result then follows from the uniqueness of (6.14). $\square$

Remark 6.6. The representation (6.15) of $\mathcal{Y}$ is exactly the so-called *variation of constants formula* for linear BSVIEs. A similar result was first obtained by Hu–Øksendal [36] for the linear BSVIEs driven by a Brownian motion and a compensated Poisson random measure. However, in [36] the coefficients α, β are assumed to be deterministic functions and $\beta(t, r) \equiv \beta(r)$ is required to be independent of t . Thus our result is a generalized version of [36, Theorem 3.1].

6.3. Representation of $\partial_{\mathbf{x}} U$

In this subsection, we assume Assumption 3.3 holds true, and let U be the classical solution to PPDE (3.1), corresponding to the decoupled FBSVIE (2.3)–(2.8). We shall use type-II BSVIE to provide an explicit representation formula for $\partial_{\mathbf{x}} U(t, s, \mathbf{x})$, which is determined by (3.15)–(3.16).

We first apply Proposition 6.5 to the middle equation of (3.16) with

$$\begin{aligned} \xi_l &:= \langle Dg(l, X^{s, \mathbf{x}}), \nabla_{\eta} X^{s, \mathbf{x}} \rangle + \int_{l \vee s}^T \langle Df(l, r, X^{s, \mathbf{x}}, Y_r^{s, \mathbf{x}}, Z_r^{l, s, \mathbf{x}}), \nabla_{\eta} X^{s, \mathbf{x}} \rangle dr; \\ \alpha(l, r) &:= \partial_y f(l, r, X^{s, \mathbf{x}}, Y_r^{s, \mathbf{x}}, Z_r^{l, s, \mathbf{x}}), \quad \beta(l, r) := \partial_z f(l, r, X^{s, \mathbf{x}}, Y_r^{s, \mathbf{x}}, Z_r^{l, s, \mathbf{x}}), \quad l \vee s \leq r; \\ \alpha(t, r) &:= 0, \quad \beta(t, r) := 0, \quad t \leq r < s. \end{aligned} \quad (6.20)$$

Define $M_r^l, K_1(l, r)$ and $\Gamma(l, r)$ by (6.16)–(6.17), then

$$\nabla_{\eta} Y_l^{s, \mathbf{x}} = \mathbb{E}_l \left[M_T^l \xi_l + \int_l^T \Gamma(l, r) M_T^r \xi_r dr \right], \quad l \in [s, T].$$

Note that $M_s^t = 1$ and $K_1(t, r) = 0$ for $r \in [t, s]$, thanks to the third line of (6.20). Then, by (3.15) and the last equation of (3.16) we have

$$\begin{aligned} \partial_{\mathbf{x}} U(t, s, \mathbf{x}) &= \nabla_{\eta} \tilde{Y}_s^{t, s, \mathbf{x}} = \mathbb{E} \left[M_T^t \xi_t + \int_s^T M_l^t \alpha(t, l) \nabla_{\eta} Y_l^{s, \mathbf{x}} dl \right] \\ &= \mathbb{E} \left[M_T^t \xi_t + \int_s^T K_1(t, l) \left[M_T^l \xi_l + \int_l^T \Gamma(l, r) M_T^r \xi_r dr \right] dl \right] \end{aligned}$$

$$\begin{aligned}
&= \mathbb{E} \left[M_T^l \xi_t + \int_s^T \left[K_1(t, l) + \int_s^r K_1(t, l) \Gamma(l, r) dl \right] M_T^r \xi_r dr \right] \\
&= \mathbb{E} \left[M_T^l \xi_t + \int_s^T \Gamma(t, r) M_T^r \xi_r dr \right].
\end{aligned}$$

Plug the first line of (6.20) into this, we obtain

$$\begin{aligned}
\partial_{\mathbf{x}} U(t, s, \mathbf{x}) &= \mathbb{E} \left[\langle G^{s, \mathbf{x}}(t), \nabla_{\eta} X^{s, \mathbf{x}} \rangle \right], \\
\text{where } G^{s, \mathbf{x}}(t) &:= M_T^l Dg(t, X^{s, \mathbf{x}}) + \int_s^T \Gamma(t, l) M_T^l Dg(l, X^{s, \mathbf{x}}) dl + \int_s^T H^{s, \mathbf{x}}(t, r) dr; \\
H^{s, \mathbf{x}}(t, r) &:= M_r^l Df(t, r, X^{s, \mathbf{x}}, Y_r^{s, \mathbf{x}}, Z_r^{t, s, \mathbf{x}}) + \int_s^r \Gamma(t, l) M_r^l Df(l, r, X^{s, \mathbf{x}}, Y_r^{s, \mathbf{x}}, Z_r^{l, s, \mathbf{x}}) dl.
\end{aligned} \tag{6.21}$$

Next, recall the first equation of (3.16). We set

$$\begin{aligned}
\check{\varphi}(t, s, \mathbf{x}; t', s', dr') &:= D\varphi(t', s', X^{s, \mathbf{x}})(dr'), \quad \text{for } \varphi = b, \sigma; \\
\check{g}(t, s, \mathbf{x}; dt') &:= G^{s, \mathbf{x}}(t)(dt').
\end{aligned} \tag{6.22}$$

We now introduce the type-II BSVIE on $[s, T]$:

$$\begin{aligned}
\mathcal{Y}_{t'} &= \check{g}(t') - \int_{t' \leq l' \leq s' \leq r' \leq T} [\check{b}(t, s, \mathbf{x}; r', s', dl') \tilde{\mathcal{Y}}(dr', s') \\
&\quad + \check{\sigma}(t, s, \mathbf{x}; r', s', dl') \mathcal{Z}(dr', s')] ds' - \int_{t'}^T \mathcal{Z}(t', s') dW_{s'}; \\
\tilde{\mathcal{Y}}(t', s') &= \mathcal{Y}_{t'} - \int_{s'}^{t'} \mathcal{Z}(t', r') dW_{r'}, \quad 0 \leq s' \leq t'.
\end{aligned} \tag{6.23}$$

By Theorem 6.1, we obtain the following explicit representation formula for $\partial_{\mathbf{x}} U(t, s, \mathbf{x})$.

Theorem 6.7. *For any fixed $(t, s, \mathbf{x}) \in \mathbb{T}_+^2 \times \mathbb{X}$, let $\tilde{\mathcal{Y}}$ be determined by (6.23). Then the path derivative of the solution U to PPDE (3.1) can be represented explicitly as follows:*

$$\partial_{\mathbf{x}} U(t, s, \mathbf{x}) = \tilde{\mathcal{Y}}(\cdot, s). \tag{6.24}$$

Funding

The first author was supported by Singapore MOE AcRF Grants R-146-000-271-112. The second author was supported in part by NSF Grant DMS-1812921. The third author was supported in part by NSF Grant DMS-1908665.

References

- [1] E. Abi Jaber, M. Larsson and S. Pulido. Affine Volterra processes. *Ann. Appl. Probab.* **29** (2019) 3155–3200. MR4019885 <https://doi.org/10.1214/19-AAP1477>
- [2] N. Agram. Dynamic risk measure for BSVIE with jumps and semimartingale issues. *Stoch. Anal. Appl.* **37** (2019) 1–16. MR3946615 <https://doi.org/10.1080/07362994.2019.1569531>
- [3] E. Alos, J. Leon and J. Vives. On the short-time behavior of the implied volatility for jump-diffusion models with stochastic volatility. *Finance Stoch.* **11** (2007) 571–589. MR2335834 <https://doi.org/10.1007/s00780-007-0049-1>
- [4] A. Barrasso and F. Russo. Decoupled mild solutions of path-dependent PDEs and integro PDEs represented by BSDEs driven by cadlag martingales. *Potential Anal.* **53** (2020) 449–481. MR4125098 <https://doi.org/10.1007/s11118-019-09775-x>
- [5] A. Barrasso and F. Russo. Gâteaux type path-dependent PDEs and BSDEs with Gaussian forward processes. Preprint. Available at [arXiv:1907.13366](https://arxiv.org/abs/1907.13366).
- [6] C. Bayer, P. Friz and J. Gatheral. Pricing under rough volatility. *Quant. Finance* **16** (2016) 887–904. MR3494612 <https://doi.org/10.1080/14697688.2015.1099717>
- [7] M. Bennedsen, A. Lunde and M. Pakkanen. Decoupling the short- and long-term behavior of stochastic volatility. Preprint. Available at [arXiv:1610.00332](https://arxiv.org/abs/1610.00332).

- [8] M. Berger and V. Mizel. Volterra equation with Itô integrals, I,II. *J. Integral Equ.* **2** (1980) 187–245, 319–337. MR0596459
- [9] A. Chronopoulou and F. G. Alexandra. Stochastic volatility and option pricing with long-memory in discrete and continuous time. *Quant. Finance* **12** (2012) 635–649. MR2909603 <https://doi.org/10.1080/14697688.2012.664939>
- [10] F. Comte and E. Renault. Long memory in continuous-time stochastic volatility models. *Math. Finance* **8** (1998) 291–323. MR1645101 <https://doi.org/10.1111/1467-9965.00057>
- [11] R. Cont and D. A. Fournié. Change of variable formulas for non-anticipative functionals on path space. *J. Funct. Anal.* **259** (2010) 1043–1072. MR2652181 <https://doi.org/10.1016/j.jfa.2010.04.017>
- [12] R. Cont and D. A. Fournié. Functional Itô calculus and stochastic integral representation of martingales. *Ann. Probab.* **41** (2013) 109–133. MR3059194 <https://doi.org/10.1214/11-AOP721>
- [13] A. Cosso and F. Russo. Strong-viscosity solutions: Semilinear parabolic PDEs and path-dependent PDEs. *Osaka J. Math.* **56** (2019) 323–373. MR3934979
- [14] A. Cosso and F. Russo. Crandall–Lions viscosity solutions for path-dependent PDEs: The case of heat equation. Preprint. Available at arXiv:1911.13095.
- [15] M. G. Crandall, H. Ishii and P. L. Lions. User’s guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc.* **27** (1992) 1–67. MR1118699 <https://doi.org/10.1090/S0273-0979-1992-00266-5>
- [16] M. G. Crandall and P. L. Lions. Hamilton–Jacobi equations in infinite dimensions, V: Unbounded linear terms and B -continuous solutions. *J. Funct. Anal.* **97** (1991) 417–465. MR1111190 [https://doi.org/10.1016/0022-1236\(91\)90010-3](https://doi.org/10.1016/0022-1236(91)90010-3)
- [17] C. Cuchiero and J. Teichmann. Generalized Feller processes and Markovian lifts of stochastic Volterra processes: The affine case. *J. Evol. Equ.* (2020). MR4181950 <https://doi.org/10.1007/s00028-020-00557-2>
- [18] L. Di Persio. Backward stochastic Volterra integral equation approach to stochastic differential utility. *Int. Electron. J. Pure Appl. Math.* **8** (2014) 11–15. MR3258002 <https://doi.org/10.12732/iejpm.v8i1.2>
- [19] J. Djordjević and S. Janković. On a class of backward stochastic Volterra integral equations. *Appl. Math. Lett.* **26** (2013) 1192–1197. MR3250433 <https://doi.org/10.1016/j.aml.2013.07.006>
- [20] J. Djordjević and S. Janković. Backward stochastic Volterra integral equations with additive perturbations. *Appl. Math. Comput.* **265** (2015) 903–910. MR3373532 <https://doi.org/10.1016/j.amc.2015.05.077>
- [21] B. Dupire. Functional Itô calculus. *Quant. Finance* **19** (2019) 721–729. MR3939653 <https://doi.org/10.1080/14697688.2019.1575974>
- [22] I. Ekeland and A. Lazrak. The golden rule when preferences are time inconsistent. *Math. Financ. Econ.* **4** (2010) 29–55. MR2746576 <https://doi.org/10.1007/s11579-010-0034-x>
- [23] I. Ekren, C. Keller, N. Touzi and J. Zhang. On viscosity solutions of path dependent PDEs. *Ann. Probab.* **42** (2014) 204–236. MR3161485 <https://doi.org/10.1214/12-AOP788>
- [24] I. Ekren, N. Touzi and J. Zhang. Viscosity solutions of fully nonlinear parabolic path dependent PDEs: Part I. *Ann. Probab.* **44** (2016) 1212–1253. MR3474470 <https://doi.org/10.1214/14-AOP999>
- [25] I. Ekren, N. Touzi and J. Zhang. Viscosity solutions of fully nonlinear parabolic path dependent PDEs: Part II. *Ann. Probab.* **44** (2016) 2507–2553. MR3531674 <https://doi.org/10.1214/15-AOP1027>
- [26] O. El Euch and M. Rosenbaum. Perfect hedging in rough Heston models. *Ann. Appl. Probab.* **28** (2018) 3813–3856. MR3861827 <https://doi.org/10.1214/18-AAP1408>
- [27] O. El Euch and M. Rosenbaum. The characteristic function of rough Heston models. *Math. Finance* **29** (2019) 3–38. MR3905737 <https://doi.org/10.1111/mafi.12173>
- [28] G. Fabbri, F. Gozzi and A. Swiech. *Stochastic Optimal Control in Infinite Dimensions*. Springer, 2017. MR3674558 <https://doi.org/10.1007/978-3-319-53067-3>
- [29] J. Fouque and R. Hu. Optimal portfolio under fractional stochastic environment. *Math. Finance* **29** (2019) 697–734. MR3974280 <https://doi.org/10.1111/mafi.12195>
- [30] J. Gatheral, T. Jaisson and M. Rosenbaum. Volatility is rough. *Quant. Finance* **18** (2018) 933–949. MR3805308 <https://doi.org/10.1080/14697688.2017.1393551>
- [31] J. Gatheral and M. Keller-Ressel. Affine forward variance models. *Finance Stoch.* **23** (2019) 501–533. MR3968277 <https://doi.org/10.1007/s00780-019-00392-5>
- [32] A. Gulisashvili, F. Viens and X. Zhang. Small-time asymptotics for Gaussian self-similar stochastic volatility models. *Appl. Math. Optim.* **82** (2020) 183–223. MR4123286 <https://doi.org/10.1007/s00245-018-9497-6>
- [33] Y. Hamaguchi. Small-time solvability of a flow of forward-backward stochastic differential equations. *Appl. Math. Optim.* (2020). MR1654566 <https://doi.org/10.1007/s002459900100>
- [34] C. Hernández and D. Possamai. Me, myself and I: A general theory of non-Markovian time-inconsistent stochastic control for sophisticated agents. Preprint. Available at arXiv:2002.12572.
- [35] C. Hernández and D. Possamai. A unified approach to well-posedness of type-I backward stochastic Volterra integral equations. Preprint. Available at arXiv:2007.12258. MR2956187 <https://doi.org/10.1007/s10114-012-9738-6>
- [36] Y. Hu and B. Øksendal. Linear Volterra backward stochastic integral equations. *Stochastic Process. Appl.* **129** (2019) 626–633. MR3907012 <https://doi.org/10.1016/j.spa.2018.03.016>
- [37] L. Karp. Non-constant discounting in continuous time. *J. Econom. Theory* **132** (2007) 557–568. MR2285623 <https://doi.org/10.1016/j.jet.2005.07.006>
- [38] E. Kromer and L. Overbeck. Differentiability of BSVIEs and dynamical capital allocations. *Int. J. Theor. Appl. Finance* **20** (07) (2017), 1750047. MR3724287 <https://doi.org/10.1142/S0219024917500479>
- [39] D. Laibson. Golden eggs and hyperbolic discounting. *Q. J. Econ.* **112** (1997) 443–478.
- [40] D. Leao, A. Ohashi and A. B. Simas. A weak version of path-dependent functional Itô calculus. *Ann. Probab.* **46** (2018) 3399–3441. MR3857859 <https://doi.org/10.1214/17-AOP1250>
- [41] X. Li and J. Yong. *Optimal Control Theory for Infinite Dimensional Systems*. Birkhäuser, Boston, 1995. MR1312364 <https://doi.org/10.1007/978-1-4612-4260-4>
- [42] J. Lin. Adapted solution of a backward stochastic nonlinear Volterra integral equation. *Stoch. Anal. Appl.* **20** (2002) 165–183. MR1891483 <https://doi.org/10.1081/SAP-120002426>

- [43] N. Yu. Lukoyanov. On viscosity solution of functional Hamilton–Jacobi type equations for hereditary systems. *Proc. Steklov Inst. Math.* **259** (2007) 190–200. MR2479421 <https://doi.org/10.1134/S1064562408010134>
- [44] J. Ma, P. Protter and J. Yong. Solving forward-backward stochastic differential equations explicitly – a four step scheme. *Probab. Theory Related Fields* **98** (1994) 339–359. MR1262970 <https://doi.org/10.1007/BF01192258>
- [45] D. Nualart. Malliavin calculus and related topics. In *Stochastic Processes and Related Topics*. Springer, Berlin, 2006. MR2200233
- [46] L. Overbeck and J. A. L. Röder. Path-dependent backward stochastic Volterra integral equations with jumps, differentiability and duality principle. *Probab. Uncertain. Quant. Risk* (2018). MR3848765 <https://doi.org/10.1186/s41546-018-0030-2>
- [47] E. Pardoux and S. Peng. Backward stochastic differential equations and quasilinear parabolic partial differential equations. In *Stochastic Partial Differential Equations and Their Applications* 200–217. Springer, Berlin, Heidelberg, 1992. MR1176785 <https://doi.org/10.1007/BFb0007334>
- [48] S. Peng. A nonlinear Feynman–Kac formula and applications. In *Proc. Symposium of System Sciences and Control Theory* 173–184. Chen & Yong (Eds). World Scientific, Singapore, 1991. MR1171673
- [49] S. Peng. Backward stochastic differential equation, nonlinear expectation and their applications. In *Proc. International Congress of Mathematicians 2010 I* 393–432. MR2827899
- [50] S. Peng and Y. Song. G-expectation weighted Sobolev spaces, backward SDE and path dependent PDE. *J. Math. Soc. Japan* **67** (2015) 1725–1757. MR3417511 <https://doi.org/10.2969/jmsj/06741725>
- [51] S. Peng and F. Wang. BSDE, path-dependent PDE and nonlinear Feynman–Kac formula. *Sci. China Math.* **59** (2016) 19–36. MR3436993 <https://doi.org/10.1007/s11425-015-5086-1>
- [52] S. Peng and Z. Yang. Anticipated backward stochastic differential equations. *Ann. Probab.* **37** (2009) 877–902. MR2537524 <https://doi.org/10.1214/08-AOP423>
- [53] R. A. Pollak. Consistent planning. *Rev. Econ. Stud.* **35** (1968) 185–199.
- [54] A. Popier. Backward stochastic Volterra integral equations with jumps in a general filtration. Preprint. Available at arXiv:2002.06992. MR4234131 <https://doi.org/10.1051/ps/2021006>
- [55] Z. Ren and M. Rosetolato. Viscosity solutions of path-dependent PDEs with randomized time. *SIAM J. Math. Anal.* **52** (2020) 1943–1979. MR4088380 <https://doi.org/10.1137/18M122666X>
- [56] Z. Ren, N. Touzi and J. Zhang. Comparison of viscosity solutions of fully nonlinear degenerate parabolic path-dependent PDEs. *SIAM J. Math. Anal.* **49** (2017) 4093–4116. MR3715377 <https://doi.org/10.1137/16M1090338>
- [57] J. Ruan and J. Zhang. Numerical methods for high dimensional path dependent PDEs driven by stochastic Volterra integral equations. Working paper.
- [58] Y. Shi, T. Wang and J. Yong. Optimal control problems of forward–backward stochastic Volterra integral equations. *Math. Control Relat. Fields* **5** (2015) 613–649. MR3371959 <https://doi.org/10.3934/mcrf.2015.5.613>
- [59] R. H. Strotz. Myopia and inconsistency in dynamic utility maximization. *Rev. Econ. Stud.* **23** (1955–1956) 165–180.
- [60] F. Viens and J. Zhang. A martingale approach for fractional Brownian motions and related path dependent PDEs. *Ann. Appl. Probab.* **29** (2019) 3489–3540. MR4047986 <https://doi.org/10.1214/19-AAP1486>
- [61] H. Wang. Extended backward stochastic Volterra integral equations, quasilinear parabolic equations, and Feynman–Kac formula. *Stoch. Dyn.* **21** (2021), 2150004. MR4192905 <https://doi.org/10.1142/S0219493721500040>
- [62] H. Wang, J. Sun and J. Yong. Recursive utility processes, dynamic risk measures and quadratic backward stochastic Volterra integral equations. *Appl. Math. Optim.* (2019). <https://doi.org/10.1007/s00245-019-09641-7>
- [63] H. Wang and J. Yong. Time-inconsistent stochastic optimal control problems and backward stochastic Volterra integral equations. Preprint. Available at arXiv:1911.04995. MR4192905 <https://doi.org/10.1142/S0219493721500040>
- [64] T. Wang and J. Yong. Comparison theorems for some backward stochastic Volterra integral equations. *Stochastic Process. Appl.* **125** (2015) 1756–1798. MR3315612 <https://doi.org/10.1016/j.spa.2014.11.013>
- [65] T. Wang and J. Yong. Backward stochastic Volterra integral equations – representation of adapted solutions. *Stochastic Process. Appl.* **129** (2019) 4926–4964. MR4025696 <https://doi.org/10.1016/j.spa.2018.12.016>
- [66] T. Wang and H. Zhang. Optimal control problems of forward–backward stochastic Volterra integral equations with closed control regions. *SIAM J. Control Optim.* **55** (2017) 2574–2602. MR3686787 <https://doi.org/10.1137/16M1059801>
- [67] Q. Wei, J. Yong and Z. Yu. Time-inconsistent recursive stochastic optimal control problems. *SIAM J. Control Optim.* **55** (2017) 4156–4201. MR3738841 <https://doi.org/10.1137/16M1079415>
- [68] J. Yong. Backward stochastic Volterra integral equations and some related problems. *Stochastic Process. Appl.* **116** (2006) 779–795. MR2218335 <https://doi.org/10.1016/j.spa.2006.01.005>
- [69] J. Yong. Continuous-time dynamic risk measures by backward stochastic Volterra integral equations. *Appl. Anal.* **86** (2007) 1429–1442. MR2373675 <https://doi.org/10.1080/00036810701697328>
- [70] J. Yong. Well-posedness and regularity of backward stochastic Volterra integral equations. *Probab. Theory Related Fields* **142** (2008) 21–77. MR2413266 <https://doi.org/10.1007/s00440-007-0098-6>
- [71] J. Yong. Time-inconsistent optimal control problems and the equilibrium HJB equation. *Math. Control Relat. Fields* **2** (2012) 271–329. MR2991570 <https://doi.org/10.3934/mcrf.2012.2.271>
- [72] J. Yong. Representation of adapted solutions to backward stochastic Volterra integral equations. *Sci. Sin., Math.* **47** (2017) 1355–1366. (in Chinese).
- [73] K. Yosida. *Functional Analysis*, 6th edition. Springer-Verlag, Berlin, 1980. MR1336382 <https://doi.org/10.1007/978-3-642-61859-8>
- [74] J. Zhang. *Backward Stochastic Differential Equations: From Linear to Fully Nonlinear Theory*. Springer, New York, 2017. MR3699487 <https://doi.org/10.1007/978-1-4939-7256-2>
- [75] J. Zhou. Viscosity solutions to second order path-dependent Hamilton–Jacobi–Bellman equations and applications. Preprint. Available at arXiv:2005.05309.