A unified approach to well-posedness of type-I backward stochastic Volterra integral equations*

Camilo Hernández[†]   Dylan Possamaï[‡]

Abstract

We study a novel general class of multidimensional type-I backward stochastic Volterra integral equations. Toward this goal, we introduce an infinite family of standard backward SDEs and establish its well-posedness, and we show that it is equivalent to that of a type-I backward stochastic Volterra integral equation. We also establish a representation formula in terms of non-linear semi-linear partial differential equation of Hamilton–Jacobi–Bellman type. As an application, we consider the study of time-inconsistent stochastic control from a game-theoretic point of view. We show the equivalence of two current approaches to this problem from both a probabilistic and an analytic point of view.

1 Introduction

This paper is concerned with introducing a unified method to address the well-posedness of backward stochastic Volterra integral equations, BSVIEs for short. BSVIEs are regarded as natural extensions of backward stochastic differential equations, BS-DEs for short. On a complete filtered probability space $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$, supporting an n-dimensional Brownian motion B, and denoting by $\mathbb{G}$ the $\mathbb{P}$-augmented natural filtration generated by B, one is given data, that is to say a $\mathcal{G}_T$-measurable random variable ξ, and

†Columbia University, IEOR department, USA. E-mail: camilo.hernandez@columbia.edu.
‡ETH Zürich, Department of Mathematics, Switzerland. E-mail: dylan.possamai@math.ethz.ch

a mapping g, referred to respectively as the *terminal condition* and the *generator*. A solution to a BSDE is a pair of $\mathbb{G}$-adapted processes $(Y_\cdot, Z_\cdot)$ such that

$$Y_t = \xi + \int_t^T g_r(Y_r, Z_r)\mathrm{d}r - \int_t^T Z_r\mathrm{d}B_r, \ t \in [0,T], \ \mathbb{P}-\text{a.s.} \tag{1.1}$$

BSDEs of linear type were first introduced by Bismut [10, 11] as an adjoint equation in the Pontryagin stochastic maximum principle. Actually, the contemporary work of Davis and Varaiya [20][1] studied a precursor of a linear BSDE for characterising the value function and the optimal controls of stochastic control problems with drift control only. In the same context of the stochastic maximum principle, BSDEs of linear type are present in Arkin and Saksonov [7], Bensoussan [9] and Kabanov [33]. Remarkably, the extension to the non-linear case is due to Bismut [12], as a type of Riccati equation, as well as Chitashvili [16], and Chitashvili and Mania [17, 18]. Later, the seminal work of Pardoux and Peng [40] presented the first systematic treatment of BSDEs in the general nonlinear case, while the celebrated survey paper of El Karoui, Peng, and Quenez [25] collected a wide range of properties and applications of BSDEs to finance. Among such properties we recall the so-called *flow property*, that is to say, for any $0 \leq r \leq T$,

$$Y_t(T,\xi) = Y_t(r, Y_r(T,\xi)), \ t \in [0,r], \ \mathbb{P}-\text{a.s., and,}$$
$$Z_t(T,\xi) = Z_t(r, Y_r(T,\xi)), \ \mathrm{d}t \otimes \mathrm{d}\mathbb{P}-\text{a.e. on } [0,r] \times \Omega,$$

where $(Y(T,\xi), Z(T,\xi))$ denotes the solution to the BSDE with terminal condition ξ and final time horizon T.

A natural extension of (1.1) arises by considering a collection of $\mathcal{G}_T$-measurable random variables $(\xi(t))_{t \in [0,T]}$, referred in the literature of BSVIEs as the *free term*, as well as a generator g. In such a setting, a solution to a BSVIE is a pair $(Y_\cdot, Z_\cdot^\cdot)$ of processes such that

$$Y_t = \xi(t) + \int_t^T g_r(t, Y_r, Z_r^t, Z_t^r)\mathrm{d}r - \int_t^T Z_r^t\mathrm{d}B_r, \ \mathbb{P}-\text{a.s.}, \ t \in [0,T]. \tag{1.2}$$

Of noticeable interest is the case in which the term Z_t^r is absent in the generator, i.e.

$$Y_t = \xi(t) + \int_t^T g_r(t, Y_r, Z_r^t)\mathrm{d}r - \int_t^T Z_r^t\mathrm{d}B_r, \ \mathbb{P}-\text{a.s.}, \ t \in [0,T]. \tag{1.3}$$

Nowadays (1.3) and (1.2) are referred in the literature as type-I and type-II BSVIEs, respectively. The first mention of such equations is, to the best of our knowledge, due to Hu and Peng [31]. Indeed, in the context of well-posedness of BSDEs valued in a Hilbert space, a prototype of type-I BSVIEs (1.3) is considered, see the comments following [31, Remark 1.1]. Two decades passed before a direct consideration of BSVIEs of the form given by (1.3) was made by Lin [37], where the author studied the case $\xi(t) = \xi$, $t \in [0,T]$, for a $\mathcal{G}_T$-measurable ξ. The general form of (1.2) was first addressed in Yong [69, 71] in the context of optimal control of (forward) stochastic Volterra integral equations (FSVIEs, for short).

There are significant distinctions between BSDEs and BSVIEs. Nevertheless, a satisfactory concept of solution for such equations can be defined by extrapolating from the theory of BSDEs. In broad terms, a pair $(Y_\cdot, Z_\cdot^\cdot)$ is said to be a solution to a BSVIE, see [71], if for each $s \in [0,T]$, the mapping $t \longmapsto (Y_t, Z_t^s)$ is $\mathbb{G}$-adapted on $[s,T]$, (Y,Z) is appropriately integrable and satisfies (1.2). It is also worth pointing out the distinctions between type-I and type-II BSVIEs. As a consequence of the presence of

---

[1]Indeed, [20] was received for publication on October 27, 1971 and it is part of the bibliography of [11].

Z^s_t in the generator, to obtain a solution to a type-II BSVIE one has to determine Z^s_t for $(t, s) \in [0, T]^2$, and (1.2) alone does not give enough restrictions. Indeed, [71] showed that an adapted solution to the type-II BSVIE (1.2) may, in general, not be unique. This is in contrasts with type-I BSVIEs, where it suffices to determine Z^s_t for $(t, s) \in [0, T]^2$, $0 \le s \le t \le T$. Moreover, without additional assumptions, a solution to a general type-II BSVIE does not satisfy the flow property.

Since their introduction, BSVIEs have been extended to much more general frameworks than the one presented above. Hence, Wang [57] studies the case of random Lipschitz data; Wang and Zhang [66] and Shi and Wang [47] deal with general non-Lipschitz data; Coulibaly and Aman [19] study time-delayed generators; *mean-field* BSVIEs are considered in Shi, Wang, and Yong [48]; Lu [38], Hu and Øksendal [30] and Popier [42] studied extensions to general filtrations and the case where B is replaced by more general processes; Djordjević and Janković [22, 21] were interested in perturbed BSVIEs, i.e. when the coefficients depend additively on small perturbations; BSVIEs in Hilbert spaces have been investigated in Anh and Yong [5], Anh, Grecksch, and Yong [6], and Ren [45]; and an analysis of numerical schemes for BSVIEs has been proposed in Bender and Pokalyuk [8]. There is also a wide spectrum of applications of BSVIEs. Hence, dynamic risk measures have been considered in Yong [70], Wang and Shi [61, 62], Wang, Sun, and Yong [55] and Agram [1]. Kromer and Overbeck [35] also studied the question of dynamic capital allocations via BSVIEs. Wang and Shi [60] dealt with a risk minimisation problem by means of the maximum principle for FBSVIEs, while the optimal control of SVIEs and BSVIEs via the maximum principle has been studied in Chen and Yong [15], Wang [59], Agram, Øksendal, and Yakhlef [3, 4], Shi, Wang, and Yong [49], Shi, Wen, and Xiong [50], see also Wei and Xiao [67] for the case with state constraints.

Since their first appearance, a natural and non-trivial question for BSVIEs has been that of the regularity in time of their solutions. The best known probabilistic results for general type-II BSVIEs guarantee the regularity of the solutions in an $\mathbb{L}^p$ sense only, see Wang [58] and Li, Wu, and Wang [36]. Nevertheless, analytic results via a representation formula, guarantee the pathwise regularity of a solution to type-I BSVIEs, see Wang and Yong [64] and Wang, Yong, and Zhang [56] for results regarding the representation of BSVIEs in the Markovian and non-Markovian framework, respectively. Regarding BSVIEs driven by discontinuous processes, it is known that type-I BSVIEs are much more amenable to the analysis, for example [42; 66] are able to study the regularity of type-I BSVIEs with jumps by probabilistic methods. Extending these ideas to type-II BSVIEs is a more challenging task, see for instance Overbeck and Röder [39].[2] General discontinuous BSVIEs are out of the scope of our paper, and we leave a possible extension for future research.

Out of the class of processes described by BSVIEs, a broader family than that of standard type-I BSVIEs (1.3) is known to arise in the study of time-inconsistent control problems. Recently, Agram and Djehiche [2] studied reflected backward stochastic Volterra integral equations and their relations to a time-inconsistent optimal stopping problem. Earlier connections were suggested in the concluding remarks of Wang and Yong [64]. Indeed, BSVIEs provide a probabilistic representations of the system of partial differential equation (PDE, for short) appearing in the study of time-inconsistent optimal control problems, e.g. see Yong [72] and Wei, Yong, and Yu [68] for PDEs obtained via Pontryagin's and Bellman's principle, respectively. A natural link was then made rigorous independently by Wang and Yong [54, Section 5] and Hernández and Possamaï

---

[2]We highlight here a potential problematic argument in [39, Corollary 1], which was pointed out to us by a referee. Indeed, it is claimed there that if $t \longmapsto Y^s_t$ and $s \longmapsto Y^s_t$ are càdlàg, then $t \longmapsto Y^t_t$ is also càdlàg, which in general is not true. In our framework this will not be an issue.

[29, Lemma A.2.3]. Although following different approaches, their analyses lead to introduce type-I BSVIEs of the form

$$Y_t = \xi(t) + \int_t^T g_r(t, Y_r, Z_r^t, Z_r^r) \mathrm{d}r - \int_t^T Z_r^t \mathrm{d}B_r, \ \mathbb{P}-\text{a.s.}, \ t \in [0, T]. \qquad (1.4)$$

These are BSVIEs in which the diagonal of Z appears in the generator. We highlight that, until the present work, the only well-posedness results in the literature for type-I BSVIEs (1.4) are available in [54] and [29]. Both results hold for the particular case in which the driver g is linear in Z_r^t. Indeed, the argument in [54] follows as a consequence of the representation formula, i.e. an analytic argument via PDEs, and holds in a Markovian setting. On the other hand, the probabilistic argument in [29] holds in the non-Markovian case.

Likewise, Hamaguchi [26, 27] studied a time-inconsistent control problem where the cost functional is defined by the Y component of the solution of a type-I BSVIE (1.3), in which g depends on a control. Via Pontryagin's optimal principle, the author noticed that the adjoint equations correspond to an *extended* type-I BSVIE, as first introduced in Wang [53] in the context of generalising the celebrated Feynman–Kac formula. An extended type-I BSVIE consists of a pair $(Y^{\cdot}, Z^{\cdot})$, with appropriate integrability, such that $s \longmapsto Y^s$ is continuous in an appropriate sense for $s \in [0, T]$, $Y^s_{\cdot}$ is pathwise continuous, $Z^s_{\cdot}$ is predictable, and

$$Y_t^s = \xi(s) + \int_t^T g_r(s, Y_r^s, Z_r^s, Y_r^r) \mathrm{d}r - \int_t^T Z_r^s \mathrm{d}B_r, \ t \in [0, T], \ \mathbb{P}-\text{a.s.}, \ s \in [0, T]. \qquad (1.5)$$

We highlight that the noticeable feature of (1.4) and (1.5) is the appearance of the 'diagonal' processes $(Y_t^t)_{t \in [0,T]}$ and $(Z_t^t)_{t \in [0,T]}$, respectively. A prerequisite for rigorously introducing these processes is some regularity of the solution. Indeed, the regularity of $s \longmapsto (Y^s, Z^s)$ in combination with the pathwise continuity of Y and the introduction of a derivative of Z^s with respect to s, as first discussed in [29], make the analysis possible, see Remark 3.4 for details.

Put succinctly, type-I BSVIEs, understood in a broader sense than that of (1.3), provide a rich framework to address new classes of problems in mathematical finance and control. In the case of time-inconsistent control problems, (1.4) and (1.5) appear as a consequence of the study of such problems via Bellman's and Pontryagin's principles, respectively. Consequently, in this paper we want to build upon the strategy devised in [29] and address the well-posedness of a general and novel class of type-I BSVIEs. We let X be the solution to a drift-less stochastic differential equation (SDE, for short) under a probability measure $\mathbb{P}$, and $\mathbb{F}$ be the $\mathbb{P}$-augmentation of the filtration generated by X, see Section 2.1 for details, and consider a tuple $(Y^{\cdot}, Z^{\cdot}, N^{\cdot})$, of appropriately $\mathbb{F}$-adapted processes, which for any $s \in [0, T]$ satisfy, $\mathbb{P}-$a.s. for any $t \in [0, T]$, the equation

$$Y_t^s = \xi(s) + \int_t^T g_r(s, X, Y_r^s, Z_r^s, Y_r^r, Z_r^r) \mathrm{d}r - \int_t^T Z_r^s \mathrm{d}X_r - \int_t^T \mathrm{d}N_r^s, \qquad (1.6)$$

We remark that the additional process N corresponds to a martingale process which is $\mathbb{P}$-orthogonal to X. This is a consequence of the fact that we work with a general filtration $\mathbb{F}$. To the best of our knowledge, a theory for type-I BSVIEs, as general as the ones introduced above, remains absent in the literature. Moreover, such class of type-I BSVIEs has only been mentioned in [27, Remark 3.8] as an interesting generalisation of (1.5).

Our approach is based on the following class of infinite families of BSDEs, given for

any $s \in [0, T]$, by

$$\mathcal{Y}_t = \xi(T) + \int_t^T h_r(X, \mathcal{Y}_r, \mathcal{Z}_r, Y_r^r, Z_r^r) \mathrm{d}r - \int_t^T \mathcal{Z}_r \mathrm{d}X_r - \int_t^T \mathrm{d}\mathcal{N}_r, \ t \in [0, T], \ \mathbb{P}-\text{a.s.},$$

$$Y_t^s = \eta(s) + \int_t^T g_r(s, X, Y_r^s, Z_r^s, \mathcal{Y}_r, \mathcal{Z}_r) \mathrm{d}r - \int_t^T Z_r^s \mathrm{d}X_r - \int_t^T \mathrm{d}N_r^s, \ t \in [0, T], \ \mathbb{P}-\text{a.s.},$$

where $(\mathcal{Y}, \mathcal{Z}, \mathcal{N}, Y, Z, N)$ are unknown, and required to have appropriate integrability, see Section 3 and Equation $(\mathcal{S})$.

We first establish the well-posedness of $(\mathcal{S})$, see Theorem 3.6. For this it is important to be able to identify the proper spaces to carry out the analyses, see Remark 3.4. Moreover, we show that, for an appropriate choice of data for $(\mathcal{S})$, its well-posedness is equivalent to that of the type-I BSVIE (1.6), see Theorem 4.4. Noticeably, our approach can naturally be specialised to obtain the well-posedness of (1.3), (1.4) and (1.5) in the classic spaces, see Remark 4.5. Moreover, as our results provide an alternative approach to BSVIEs, it may allow for the future design of new numerical schemes to solve type-I BSVIEs, which to the best of our knowledge, remain limited to [8]. In addition, we recover classical results for this general class of multidimensional type-I BSVIEs. We provide *a priori* estimates, show the stability of solutions as well as a representation formula in terms of a semilinear PDEs, see Proposition 5.1. Given our multidimensional setting, we refrained from considering comparison results, see Wang and Yong [63] for the one-dimensional case.

As an application of our results, we consider the game-theoretic approach to time-inconsistent stochastic control problems. We recall this approach studies the problem faced by the, so-called, *sophisticated agent* who aware of the inconsistency of its preferences seeks for *consistent plans*, i.e. equilibria. We show that as a consequence of Theorem 4.4, one can reconcile two recent probabilistic approaches to this problem. Moreover, we provide, see Proposition 5.3, an equivalent result for two earlier analytic approaches, based on semi-linear PDEs. We believe this helps to elucidate connections between the different takes on the problem available in the literature.

The rest of the paper is structured as follows. Section 2 introduces the stochastic basis on a canonical space as well as the integrability spaces necessary to our analysis. Section 3 precisely formulates the class of infinite families of BSDEs $(\mathcal{S})$, which is the crux of our approach, and provides the statement of its well-posedness, while the proof is deferred to Section 6. Section 4 introduces the class of type-I BSVIEs which are the main object of this paper, and establishes the equivalence of its well-posedness with that of $(\mathcal{S})$ for a particular choice of data. Section 5 deals with the representation formula for the class of type-I BSVIEs considered, and presents the application of our results in the context of time-inconsistent stochastic control. Finally, Section 6 includes the analysis of $(\mathcal{S})$.

2 Preliminaries

Notations: we fix a time horizon $T > 0$. Given $(E, |\cdot|)$ a finite-dimensional Euclidean space, a positive integer d, and a non-negative integer q, $\mathcal{C}_q^d(E)$ (resp. $\mathcal{C}_{q,b}^d(E)$) will denote the space of functions from E to $\mathbb{R}^d$ which are q times continuously differentiable (resp. and bounded with bounded derivatives). When $d = 1$ we write $\mathcal{C}_q(E)$ and $\mathcal{C}_{q,b}(E)$. For $\phi \in \mathcal{C}_{0,q}([0, T] \times E)$ with $q \geq 2$, if $s \longmapsto \phi(s, \alpha)$ is uniformly continuous uniformly in α, we denote by $\rho_\phi : [0, T] \longrightarrow \mathbb{R}$ its modulus of continuity. $\partial_\alpha \phi$ and $\partial_{\alpha\alpha}^2 \phi$ denote the gradient and Hessian with respect to α, respectively. For $(u, v) \in (\mathbb{R}^p)^2$, $u \cdot v$ will denote their usual inner product, and $|u|$ the corresponding norm. For positive integers m and n, we denote by $M_{m,n}(\mathbb{R})$ the space of $m \times n$ matrices with real entries, and set $M_n(\mathbb{R}) := M_{n,n}(\mathbb{R})$.

For $M \in M_{m,n}(\mathbb{R})$, $M_{:i}$ and $M_{i:}$ denote the i-th column and row. $\mathbb{S}_n^+(\mathbb{R})$ denotes the set of $n \times n$ symmetric positive semi-definite matrices, while $\mathrm{Tr}[M]$ denotes the trace of $M \in M_m(\mathbb{R})$, and $|M| := \sqrt{\mathrm{Tr}[M^\top M]}$ for $M \in M_{m,n}(\mathbb{R})$.

For $(\Omega, \mathcal{F})$ a measurable space, $\mathrm{Prob}(\Omega)$ denotes the collection of probability measures on $(\Omega, \mathcal{F})$. For a filtration $\mathbb{F} := (\mathcal{F}_t)_{t \in [0,T]}$ on $(\Omega, \mathcal{F})$, $\mathcal{P}_{\mathrm{pred}}(E, \mathbb{F})$ (resp. $\mathcal{P}_{\mathrm{prog}}(E, \mathbb{F})$, $\mathcal{P}_{\mathrm{opt}}(E, \mathbb{F})$, $\mathcal{P}_{\mathrm{meas}}(E, \mathbb{F})$) denotes the set of E-valued, $\mathbb{F}$-predictable processes (resp. $\mathbb{F}$–progressively measurable processes, $\mathbb{F}$-optional processes, $\mathbb{F}$-adapted and measurable). For $\mathbb{P} \in \mathrm{Prob}(\Omega)$, $\mathbb{F}^{\mathbb{P}} := (\mathcal{F}_t^{\mathbb{P}})_{t \in [0,T]}$, denotes the $\mathbb{P}$-augmentation of $\mathbb{F}$, where for $t \in [0,T]$, $\mathcal{F}_t^{\mathbb{P}} := \mathcal{F}_t \vee \sigma(N^{\mathbb{P}})$, where $N^{\mathbb{P}} := \{N \subseteq \Omega : \exists B \in \mathcal{F}, N \subseteq B \text{ and } \mathbb{P}[B] = 0\}$. With this, $\mathbb{P} \in \mathrm{Prob}(\Omega)$ can be extended so that $(\Omega, \mathcal{F}, \mathbb{F}^{\mathbb{P}}, \mathbb{P})$ becomes a complete probability space, see Karatzas and Shreve [34, Chapter II.7]. $\mathbb{F}_+^{\mathbb{P}}$ denotes the right limit of $\mathbb{F}^{\mathbb{P}}$, i.e. $\mathcal{F}_{t+}^{\mathbb{P}} := \bigcap_{\varepsilon > 0} \mathcal{F}_{t+\varepsilon}^{\mathbb{P}}$, $t \in [0,T)$, and $\mathcal{F}_{T+}^{\mathbb{P}} := \mathcal{F}_T^{\mathbb{P}}$, so that $\mathbb{F}_+^{\mathbb{P}}$ is the minimal filtration that contains $\mathbb{F}$ and satisfies the usual conditions.

2.1 The stochastic basis on the canonical space

We fix two positive integers n and m, which represent respectively the dimension of the martingale which will drive our equations, and the dimension of the Brownian motion appearing in the dynamics of the former. We consider the canonical space $\mathcal{X} := \mathcal{C}([0,T], \mathbb{R}^n)$, with canonical process X. We let $\mathcal{F}$ be the Borel σ-algebra on $\mathcal{X}$ (for the topology of uniform convergence), and we denote by $\mathbb{F}^o := (\mathcal{F}_t^o)_{t \in [0,T]}$ the natural filtration of X. We fix a bounded Borel measurable map $\sigma : [0,T] \times \mathcal{X} \longrightarrow \mathbb{R}^{n \times m}$, $\sigma_{\cdot}(X) \in \mathcal{P}_{\mathrm{meas}}(\mathbb{R}^{n \times m}, \mathbb{F}^o)$, and an initial condition $x_0 \in \mathbb{R}^n$. We assume there is $\mathbb{P} \in \mathrm{Prob}(\mathcal{X})$ such that $\mathbb{P}[X_0 = x_0] = 1$ and X is martingale, whose quadratic variation, $\langle X \rangle = (\langle X \rangle_t)_{t \in [0,T]}$, is absolutely continuous with respect to Lebesgue measure, with density given by $\sigma \sigma^\top$. Enlarging the original probability space, see Stroock and Varadhan [51, Theorem 4.5.2], there is an $\mathbb{R}^m$-valued Brownian motion B with

$$X_t = x_0 + \int_0^t \sigma_r(X_{\cdot \wedge r}) \mathrm{d}B_r, \ t \in [0,T], \ \mathbb{P}-\text{a.s.}$$

We now let $\mathbb{F} := (\mathcal{F}_t)_{t \in [0,T]}$ be the (right-limit) of the $\mathbb{P}$-augmentation of $\mathbb{F}^o$. We stress that we will not assume $\mathbb{P}$ is unique. In particular, the predictable martingale representation property for $(\mathbb{F}, \mathbb{P})$-martingales in terms of stochastic integrals with respect to X might not hold.

Remark 2.1. We remark that the previous formulation on the canonical is by no means necessary. Indeed, any probability space supporting a Brownian motion B and a process X satisfying the previous SDE will do, and this can be found whenever that equation has a weak solution.

2.2 Functional spaces and norms

We now introduce our spaces. In the following, $(\Omega, \mathcal{F}_T, \mathbb{F}, \mathbb{P})$ is as in Section 2.1. We are given a finite-dimensional Euclidean space, i.e. $E = \mathbb{R}^k$ for some non-negative integer k and $|\cdot|$ denotes the Euclidean norm. For any $(p,q) \in (1,\infty)^2$, we introduce the spaces

- $\mathcal{L}^p(E)$ of $\mathcal{F}_T$-measurable, E-valued random variables ξ, $\|\xi\|_{\mathcal{L}^p}^p := \mathbb{E}^{\mathbb{P}}[|\xi|^p] < \infty$;

- $\mathbb{S}^p(E)$ of $Y \in \mathcal{P}_{\mathrm{opt}}(E, \mathbb{F})$, with $\mathbb{P}-$a.s. càdlàg paths, $\|Y\|_{\mathbb{S}^p}^p := \mathbb{E}^{\mathbb{P}}\left[\sup_{t \in [0,T]} |Y_t|^p\right] < \infty$;

- $\mathbb{L}^{q,p}(E)$ of $Y \in \mathcal{P}_{\mathrm{opt}}(E, \mathbb{F})$, $\|Y\|_{\mathbb{L}^{q,p}}^p := \mathbb{E}^{\mathbb{P}}\left[\left(\int_0^T |Y_r|^q \mathrm{d}r\right)^{\frac{p}{q}}\right] < \infty$;

- $\mathbb{H}^p(E)$ of $Z \in \mathcal{P}_{\mathrm{pred}}(E, \mathbb{F})$, defined $\sigma\sigma_t^\top \mathrm{d}t$−a.e., $\|Z\|_{\mathbb{H}^p}^p := \mathbb{E}^{\mathbb{P}}\left[\left(\int_0^T |\sigma_t Z_r|^2 \mathrm{d}r \right)^{\frac{p}{2}} \right] < \infty$;

- $\mathbb{M}^p(E)$ of martingales $M \in \mathcal{P}_{\mathrm{opt}}(E, \mathbb{F})$, $\mathbb{P}$-orthogonal to X (that is the product XM is an $(\mathbb{F}, \mathbb{P})$-martingale), with $\mathbb{P}$−a.s. càdlàg paths, $M_0 = 0$ and $\|M\|_{\mathbb{M}^p}^p := \mathbb{E}^{\mathbb{P}}\left[[M]_T^{\frac{p}{2}} \right] < \infty$;

- $\mathcal{L}^{p,2}(E)$ denotes the space of families $(\xi(s))_{s \in [0,T]}$ of $\mathcal{F}_T$-measurable E-valued random variables such that the mapping $([0,T] \times \Omega, \mathcal{B}([0,T]) \otimes \mathcal{F}) \longrightarrow (\mathcal{L}^p(E), \|\cdot\|_{\mathcal{L}^p}) : s \longmapsto \xi(s)$ is continuous, $\|\xi\|_{\mathcal{L}^{p,2}}^p := \sup_{s \in [0,T]} \|\xi\|_{\mathcal{L}^p}^p < \infty$;

- $\mathcal{P}_{\mathrm{meas}}^2(E, \mathcal{F}_T)$ of two parameter processes $(U_t^s)_{(s,t) \in [0,T]^2} : ([0,T]^2 \times \Omega, \mathcal{B}([0,T]^2) \otimes \mathcal{F}_T) \longrightarrow (E, \mathcal{B}(E))$ measurable.

  Finally, given an arbitrary integrability space $(\mathbb{I}^p(E), \|\cdot\|_{\mathbb{I}})$, we introduce the space

- $\mathbb{I}^{p,2}(E)$ of $(U_t^s)_{(s,t) \in [0,T]^2} \in \mathcal{P}_{\mathrm{meas}}^2(E, \mathcal{F}_T)$ such that the mapping $([0,T], \mathcal{B}([0,T])) \longrightarrow (\mathbb{I}^p(E), \|\cdot\|_{\mathbb{I}^p}) : s \longmapsto U^s$ is continuous and $\|U\|_{\mathbb{I}^{p,2}}^p := \sup_{s \in [0,T]} \|U^s\|_{\mathbb{I}^p}^p < \infty$.

  For example, $\mathbb{H}^{p,2}(E)$ denotes the space of $(Z_t^s)_{(s,t) \in [0,T]^2} \in \mathcal{P}_{\mathrm{meas}}^2(E, \mathcal{F}_T)$ such that the mapping $([0,T], \mathcal{B}([0,T])) \longrightarrow (\mathbb{H}^p(E), \|\cdot\|_{\mathbb{H}^p}) : s \longmapsto Z^s$ is continuous and $\|Z\|_{\mathbb{H}^{p,2}}^p := \sup_{s \in [0,T]} \|Z^s\|_{\mathbb{H}^p}^p < \infty$.

  Lastly we introduce the space, see Remark 2.2 for further details,

- $\overline{\mathbb{H}}^{p,2}(E)$ of $(Z_t^s)_{(s,t) \in [0,T]^2} \in \mathcal{P}_{\mathrm{meas}}^2(E, \mathcal{F}_T)$ such that $([0,T], \mathcal{B}([0,T])) \longrightarrow (\mathbb{H}^p(E), \|\cdot\|_{\mathbb{H}^p}) : s \longmapsto Z^s$ is absolutely continuous with respect to the Lebesgue measure[3], $\mathcal{Z} \in \mathbb{H}^2(E)$, where $\mathcal{Z} := (Z_t^t)_{t \in [0,T]}$ is given by

$$Z_t^t := Z_t^T - \int_t^T \eth Z_t^r \mathrm{d}r, \text{ and, } \|Z\|_{\overline{\mathbb{H}}^{p,2}}^2 := \|Z\|_{\mathbb{H}^{2,2}}^2 + \|\mathcal{Z}\|_{\mathbb{H}^2}^2 < \infty$$

Remark 2.2. When $p = q$, we will write $\mathbb{L}^p(E)$ (resp. $\mathbb{L}^{p,2}(E)$) for $\mathbb{L}^{q,p}(E)$ (resp. $\mathbb{L}^{q,p,2}(E)$). With this convention, $\mathbb{L}^2(E)$ (resp. $\mathbb{L}^{2,2}(E)$) will be $\mathbb{L}^{2,2}(E)$ (resp. $\mathbb{L}^{2,2,2}(E)$). Also, $\mathbb{S}^{p,2}(E)$, $\mathbb{L}^{q,p,2}(E)$ and $\mathbb{H}^{p,2}(E)$ are Banach spaces.

In addition, we remark that the space $\mathbb{H}^2(E)$ being closed implies $\overline{\mathbb{H}}^{p,2}(E)$ is a closed subspace of $\mathbb{H}^{p,2}(E)$ and thus a Banach space. The space $\overline{\mathbb{H}}^{p,2}(E)$ allows us to define a good candidate for $(Z_t^t)_{t \in [0,T]}$ as an element of $\mathbb{H}^2(E)$.[4] Let $\widetilde{\Omega} := [0,T] \times \mathcal{X}$, $\tilde{\omega} := (t,x) \in \widetilde{\Omega}$ and

$$\mathfrak{Z}_s(\tilde{\omega}) := Z_t^T(x) - \int_s^T \eth Z_t^r(x) \mathrm{d}r, \ \mathrm{d}t \otimes \mathrm{d}\mathbb{P}−\text{a.e. } \tilde{\omega} \in \widetilde{\Omega}, \ s \in [0,T],$$

so that the Radon–Nikodým property and Fubini's theorem imply $\mathfrak{Z}_s = Z^s$, $\mathrm{d}t \otimes \mathrm{d}\mathbb{P}$−a.e., $s \in [0,T]$. Lastly, as for $\tilde{\omega} \in \widetilde{\Omega}$, $s \longmapsto \mathfrak{Z}_s(\tilde{\omega})$ is continuous, we may define

$$Z_t^t := Z_t^T - \int_t^T \eth Z_t^r \mathrm{d}r, \text{ for } \mathrm{d}t \otimes \mathrm{d}\mathbb{P}−\text{a.e. } (t,x) \text{ in } [0,T] \times \mathcal{X}.$$

---

[3]We recall that $\mathbb{H}^2$, being a Hilbert space and in particular a reflexive Banach space, has the so-called Radon–Nikodým property, see [46, Corollary 5.45]. Thus the absolute continuity of the $\mathbb{H}^2$-valued mapping $s \longmapsto Z^s$ is equivalent to the existence of the density $(\eth Z^s)_{s \in [0,T]}$, which is unique for Lebesgue−a.e. $s \in [0,T]$. The fact $\mathcal{Z}$ is well-defined is argued in Remark 2.2

[4]This is based on [27, Section 2.1]. The mapping $s \longmapsto Z^s$ is assumed absolutely continuous, as opposed to continuously differentiable. We are indebted to Yushi Hamaguchi for pointing out inconsistencies in the definition of $\overline{\mathbb{H}}^{2,2}$ on an earlier version of this document.

3 An infinite family of BSDEs

We are given jointly measurable mappings h, g, ξ and η,

$$h : [0,T] \times \mathcal{X} \times \mathbb{R}^{d_1} \times \mathbb{R}^{n \times d_1} \times \mathbb{R}^{d_2} \times \mathbb{R}^{n \times d_2} \times \mathbb{R}^{d_2} \longrightarrow \mathbb{R}^{d_1}, \quad \xi : [0,T] \times \mathcal{X} \longrightarrow \mathbb{R}^{d_1},$$

$$g : [0,T]^2 \times \mathcal{X} \times \mathbb{R}^{d_2} \times \mathbb{R}^{n \times d_2} \times \mathbb{R}^{d_1} \times \mathbb{R}^{n \times d_1} \longrightarrow \mathbb{R}^{d_2}, \quad \eta : [0,T] \times \mathcal{X} \longrightarrow \mathbb{R}^{d_2},$$

such that for any $(y, z, u, v, \mathrm{u}) \in \mathbb{R}^{d_1} \times \mathbb{R}^{n \times d_1} \times \mathbb{R}^{d_2} \times \mathbb{R}^{n \times d_2} \times \mathbb{R}^{d_2}$, $h_\cdot(\cdot, y, z, u, v, \mathrm{u}) \in \mathcal{P}_{\mathrm{prog}}(\mathbb{R}^{d_1}, \mathbb{F})$, $g_\cdot(s, \cdot, u, v, y, z) \in \mathcal{P}_{\mathrm{prog}}(\mathbb{R}^{d_2}, \mathbb{F})$. Moreover, we work under the following set of assumptions.

Assumption 3.1. (i) $(s, u, v) \longmapsto g_t(s, x, u, v, y, z)$ (resp. $s \longmapsto \eta(s, x)$) is continuously differentiable. Moreover, the mapping $\nabla g : [0,T]^2 \times \mathcal{X} \times (\mathbb{R}^{d_2} \times \mathbb{R}^{n \times d_2})^2 \times \mathbb{R}^{d_1} \times \mathbb{R}^{n \times d_1} \longrightarrow \mathbb{R}^{d_2}$ defined by

$$\nabla g_t(s, x, \mathrm{u}, \mathrm{v}, u, v, y, z) := \partial_s g_t(s, x, u, v, y, z) + \partial_u g_t(s, x, u, v, y, z)\mathrm{u}$$
$$+ \sum_{i=1}^{n} \partial_{v_{:i}} g_t(s, x, u, v, y, z)\mathrm{v}_{i:},$$

satisfies $\nabla g_\cdot(s, \cdot, \mathrm{u}, \mathrm{v}, u, v, y, z) \in \mathcal{P}_{\mathrm{prog}}(\mathbb{R}^{d_2}, \mathbb{F})$;

(ii) $(y, z, u, v, \mathrm{u}) \longmapsto h_t(x, y, z, u, v, \mathrm{u})$ is uniformly Lipschitz continuous, i.e. $\exists L_h > 0$, such that for all $(t, x, y, \tilde{y}, z, \tilde{z}, u, \tilde{u}, v, \tilde{v}, \mathrm{u}, \tilde{\mathrm{u}})$

$$|h_t(x, y, z, u, v, \mathrm{u}) - h_t(x, \tilde{y}, \tilde{z}, \tilde{u}, \tilde{v}, \tilde{\mathrm{u}})|$$
$$\leq L_h \big(|y - \tilde{y}| + |\sigma_t(x)^\top (z - \tilde{z})| + |u - \tilde{u}| + |\sigma_t(x)^\top (v - \tilde{v})| + |\mathrm{u} - \tilde{\mathrm{u}}| \big);$$

(iii) for $\varphi \in \{g, \partial_s g\}$, $(u, v, y, z) \longmapsto \varphi_t(s, x, u, v, y, z)$ is uniformly Lipschitz continuous, i.e. $\exists L_\varphi > 0$, such that for all $(s, t, x, u, \tilde{u}, v, \tilde{v}, y, \tilde{y}, z, \tilde{z})$

$$|\varphi_t(s, x, u, v, y, z) - \varphi_t(s, x, \tilde{u}, \tilde{v}, \tilde{y}, \tilde{z})|$$
$$\leq L_\varphi \big(|u - \tilde{u}| + |\sigma_t(x)^\top (v - \tilde{v}')| + |y - \tilde{y}| + |\sigma_t(x)^\top (z - \tilde{z})| \big);$$

(iv) for $\mathbf{0} := (u, v, y, z)|_{(0,\dots,0)}$, $\big(\tilde{h}_\cdot, \tilde{g}_\cdot(s), \nabla \tilde{g}_\cdot(s)\big) := \big(h_\cdot(\cdot, \mathbf{0}, 0), g_\cdot(s, \cdot, \mathbf{0}), \partial_s g_\cdot(s, \cdot, \mathbf{0})\big) \in \mathbb{L}^{1,2}(\mathbb{R}^{d_1}) \times \big(\mathbb{L}^{1,2,2}(\mathbb{R}^{d_2})\big)^2$.

Remark 3.2. We comment on the set of requirements in Assumption 3.1. Of particular interest is Assumption 3.1.(i), the other being the standard Lipschitz assumptions on the generators as well as their integrability at zero. Anticipating the introduction of $(\mathcal{S})$ below and the discussion in Remark 3.4, Assumption 3.1.(i) will allow us to identify the second BSDE in the system as the antiderivative of the third one, see Remark 3.4.

Let us define the space $(\mathcal{H}, \|\cdot\|_{\mathcal{H}})$, of $\mathfrak{h} = (\mathcal{Y}, \mathcal{Z}, \mathcal{N}, U, V, M, \partial U, \partial V, \partial M)$, where

$$\mathcal{H} := \mathbb{S}^2(\mathbb{R}^{d_1}) \times \mathbb{H}^2(\mathbb{R}^{n \times d_1}) \times \mathbb{M}^2(\mathbb{R}^{d_1}) \times \mathbb{S}^{2,2}(\mathbb{R}^{d_2}) \times \overline{\mathbb{H}}^{2,2}(\mathbb{R}^{n \times d_2}) \times \mathbb{M}^{2,2}(\mathbb{R}^{d_2})$$
$$\times \mathbb{S}^{2,2}(\mathbb{R}^{d_2}) \times \mathbb{H}^{2,2}(\mathbb{R}^{n \times d_2}) \times \mathbb{M}^{2,2}(\mathbb{R}^{d_2}),$$

$$\|\mathfrak{h}\|_{\mathcal{H}}^2 := \|\mathcal{Y}\|_{\mathbb{S}^2}^2 + \|\mathcal{Z}\|_{\mathbb{H}^2}^2 + \|\mathcal{N}\|_{\mathbb{M}^2}^2 + \|U\|_{\mathbb{S}^{2,2}}^2 + \|V\|_{\overline{\mathbb{H}}^{2,2}}^2 + \|M\|_{\mathbb{M}^{2,2}}^2$$
$$+ \|\partial U\|_{\mathbb{S}^{2,2}}^2 + \|\partial V\|_{\mathbb{H}^{2,2}}^2 + \|\partial M\|_{\mathbb{M}^{2,2}}^2,$$

We are now ready to precise the class of systems subject to our study. Given $(\xi, \eta, \partial_s \eta) \in \mathcal{L}^2(\mathbb{R}^{d_1}) \times (\mathcal{L}^{2,2}(\mathbb{R}^{d_2}))^2$, $\partial_s \eta$ as in Assumption 3.1, we consider the system,

which for any $s \in [0, T]$ holds $\mathbb{P}-$a.s. for any $t \in [0, T]$,

$$
\mathcal{Y}_t = \xi(T, X) + \int_t^T h_r(X, \mathcal{Y}_r, \mathcal{Z}_r, U_r^r, V_r^r, \partial U_r^r) \mathrm{d}r - \int_t^T \mathcal{Z}_r^\top \mathrm{d}X_r - \int_t^T \mathrm{d}\mathcal{N}_r,
$$

$$
U_t^s = \eta(s, X) + \int_t^T g_r(s, X, U_r^s, V_r^s, \mathcal{Y}_r, \mathcal{Z}_r) \mathrm{d}r - \int_t^T V_r^{s\top} \mathrm{d}X_r - \int_t^T \mathrm{d}M_r^s, \qquad (\mathcal{S})
$$

$$
\partial U_t^s = \partial_s \eta(s, X) + \int_t^T \nabla g_r(s, X, \partial U_r^s, \partial V_r^s, U_r^s, V_r^s, \mathcal{Y}_r, \mathcal{Z}_r) \mathrm{d}r - \int_t^T \partial V_r^{s\top} \mathrm{d}X_r - \int_t^T \mathrm{d}\partial M_r^s.
$$

Definition 3.3. *We say $\mathfrak{h}$ is a solution to $(\mathcal{S})$ if $\mathfrak{h} \in \mathcal{H}$ and $(\mathcal{S})$ holds.*

Remark 3.4. We now expound on our choice for the set-up and the structure of $(\mathcal{S})$.

(i) We first highlight two aspects which are crucial to establish the connection between $(\mathcal{S})$ and type-I BSVIE (1.6). The first is the presence of ∂U in the generator of the first equation. This causes the system to be fully coupled but is nevertheless necessary in our methodology, this will be clear from the proof of Theorem 4.4 in Section 4. The second relates to our choice to write three equations instead of two. In fact, our approach is based on being able to identify ∂U as the derivative with respect to the s variable of U in an appropriate sense and, at least formally, it is clear that the third equation allows us to do so, see Lemma 6.1 for details. Alternatively, we could have chosen not to write the third equation and consider for any $s \in [0, T]$, the system, which holds $\mathbb{P}-$a.s. for any $t \in [0, T]$,

$$
\mathcal{Y}_t = \xi(T, X_{\cdot \wedge T}) + \int_t^T h_r(X, \mathcal{Y}_r, \mathcal{Z}_r, U_r^r, V_r^r, \partial U_r^r) \mathrm{d}r - \int_t^T \mathcal{Z}_r^\top \mathrm{d}X_r - \int_t^T \mathrm{d}\mathcal{N}_r,
$$

$$
U_t^s = \eta(s, X_{\cdot \wedge T}) + \int_t^T g_r(s, X, U_r^s, V_r^s, \mathcal{Y}_r, \mathcal{Z}_r) \mathrm{d}r - \int_t^T V_r^{s\top} \mathrm{d}X_r - \int_t^T \mathrm{d}M_r^s,
$$

$$
\partial U_t^s := \frac{\mathrm{d}}{\mathrm{d}s} U_{|(s,t)}^s,
$$

where $\frac{\mathrm{d}}{\mathrm{d}s} U^s$ corresponds to the density with respect to the Lebesgue measure of $s \longmapsto U^s$. Nevertheless, for the proof of well-posedness of $(\mathcal{S})$ that we present in Section 6, we have to derive appropriate estimates for $(\partial U_t^t)_{t \in [0,T]}$, and for this it is advantageous to do the identification by adding the third equation in $(\mathcal{S})$ and work on the space $(\mathcal{H}, \|\cdot\|_{\mathcal{H}})$.

(ii) We also emphasise that the presence of $(V_t^t)_{t \in [0,T]}$ in the generator of the first equation requires us to reduce the space of the solution from the classic $(\mathfrak{H}, \|\cdot\|_{\mathfrak{H}})$ to $(\mathcal{H}, \|\cdot\|_{\mathcal{H}})$ where

$$
\mathfrak{H} := \mathbb{S}^2(\mathbb{R}^{d_1}) \times \mathbb{H}^2(\mathbb{R}^{n \times d_1}) \times \mathbb{M}^2(\mathbb{R}^{d_1}) \times \mathbb{S}^{2,2}(\mathbb{R}^{d_2}) \times \mathbb{H}^{2,2}(\mathbb{R}^{n \times d_2}) \times \mathbb{M}^{2,2}(\mathbb{R}^{d_2})
$$
$$
\times \mathbb{S}^{2,2}(\mathbb{R}^{d_2}) \times \mathbb{H}^{2,2}(\mathbb{R}^{n \times d_2}) \times \mathbb{M}^{2,2}(\mathbb{R}^{d_2}),
$$

and $\|\cdot\|_{\mathfrak{H}}$ denotes the norm induced by $\mathfrak{H}$. Ultimately, this is due to the presence of $(Z_t^t)_{t \in [0,T]}$ in the type-I BSVIE (1.6). On this matter, we stress that to the best of our knowledge, our results constitute the first comprehensive study of type-I BSVIEs as general as (1.6). We remark that our identification of the appropriate space to carry out the analysis is based on [27, Section 2.1]. In the case where $(V_t^t)_{t \in [0,T]}$ (resp. $(Z_t^t)_{t \in [0,T]}$) does not appear in the generator of the first BSDE in $(\mathcal{S})$ (resp. type-I BSVIE (1.6)), Proposition 6.5 (resp. Remark 4.5) provide the arguments for how one can adapt our approach to yield a solution in the classical space. This shows that our methodology recovers existing results on type-I BSVIE (1.3) as well as the so-called extended type-I BSVIE (1.5).

Remark 3.5. In addition, we highlight two features of $(\mathcal{S})$ that will come into play in the setting of type-I BSVIE (1.6), and differ from the one in the classic literature. They are related to the fact we work under the general filtration $\mathbb{F}$. The first is the fact that the stochastic integrals in $(\mathcal{S})$ are with respect to the canonical process X. Recall that σ is not assumed to be invertible (it is not even a square matrix in general and can vanish), therefore the filtration generated by X is different from the one generated by B. This yields more general results and it allows for extra flexibility necessary in some applications, see [29] for an example. The second difference is the presence of the processes $(N, M, \partial M)$. As it was mentioned in Section 2.1, we work with a probability measure for which the martingale representation property for $\mathbb{F}$–local martingales in terms of stochastic integrals with respect to X does not necessarily hold. Therefore, we need to allow for orthogonal martingales in the representation. Certainly, there are known properties which are equivalent to the orthogonal martingales vanishing, i.e. $N = M = \partial M = 0$, for example when $\mathbb{P}$ is an extremal point of the convex hull of the probability measures that satisfy the properties in Section 2.1, see [32, Theorem 4.29].

Assumption 3.1 provides an appropriate framework to derive the well-posedness of $(\mathcal{S})$. The following is the main theorem of this section whose proof we postpone to Section 6.

Theorem 3.6. *Let* Assumption 3.1 *hold. Then* $(\mathcal{S})$ *admits a unique solution in* $(\mathcal{H}, \|\cdot\|_{\mathcal{H}})$. *For any* $\mathfrak{h} \in \mathcal{H}$ *solution to* $(\mathcal{S})$ *there exists* $C > 0$ *such that*

$$\|\mathfrak{h}\|_{\mathcal{H}}^2 \leq C\Big(\|\xi\|_{\mathcal{L}^2}^2 + \|\eta\|_{\mathcal{L}^{2,2}}^2 + \|\partial_s \eta\|_{\mathcal{L}^{2,2}}^2 + \|\tilde{h}\|_{\mathbb{L}^{1,2}}^2 + \|\tilde{g}\|_{\mathbb{L}^{1,2,2}}^2 + \|\nabla\tilde{g}\|_{\mathbb{L}^{1,2,2}}^2\Big).$$

Moreover, if $\mathfrak{h}^i \in \mathcal{H}$ *denotes the solution to* $(\mathcal{S})$ *with coefficients* $(\xi^i, h^i, \eta^i, g^i, \partial_s \eta^i, \nabla g^i)$ *for* $i \in \{1, 2\}$, *then*

$$\|\delta\mathfrak{h}\|_{\mathcal{H}}^2 \leq C\Big(\|\delta\xi\|_{\mathcal{L}^2}^2 + \|\delta\eta\|_{\mathcal{L}^{2,2}}^2 + \|\delta\partial_s\eta\|_{\mathcal{L}^{2,2}}^2 + \|\delta_1 h\|_{\mathbb{L}^{1,2}}^2 + \|\delta_1 g\|_{\mathbb{L}^{1,2,2}}^2 + \|\delta_1\nabla g\|_{\mathbb{L}^{1,2,2}}\Big),$$

where for $\varphi \in \{\mathcal{Y}, \mathcal{Z}, \mathcal{N}, U, V, M, \partial U, \partial V, \partial M, \xi, \eta, \partial_s\eta\}$ *and* $\Phi \in \{h, g, \nabla g\}$

$$\delta\varphi := \varphi^1 - \varphi^2, \text{ and } \delta_1\Phi_t := \Phi_t^1(\mathcal{Y}_r^1, \mathcal{Z}_t^1, U_t^{1t}, V_t^{1t}) - \Phi_t^2(Y_r^1, Z_t^1, U_t^{1t}, V_t^{1t}), \ \mathrm{d}t \otimes \mathrm{d}\mathbb{P}-\text{a.e.}$$

Remark 3.7. The reader may wonder about our choice to leave out the diagonal of ∂V in the generator of the first equation in $(\mathcal{S})$. As we will argue below, this would require us to consider an auxiliary infinite family of quadratic BSDEs. Since the main purpose of this paper is to relate the well-posedness of $(\mathcal{S})$ to that of the type-I BSVIE (1.6), and inasmuch as we do not need to consider this case to establish Theorem 4.4, we have refrained from pursuing it in this document. Nevertheless, this case is covered as part of the study of the extension of $(\mathcal{S})$ to the quadratic case in Hernández [28]. If we were to study the system, which for any $s \in [0, T]$ satisfies

$$\mathcal{Y}_t = \xi(T, X) + \int_t^T h_r(X, \mathcal{Y}_r, \mathcal{Z}_r, U_r^r, V_r^r, \partial U_r^r, \partial V_r^r)\mathrm{d}r - \int_t^T \mathcal{Z}_r^\top \mathrm{d}X_r - \int_t^T \mathrm{d}N_r$$

$$U_t^s = \eta(s, X) + \int_t^T g_r(s, X, U_r^s, V_r^s, \mathcal{Y}_r, \mathcal{Z}_r)\mathrm{d}r - \int_t^T V_r^{s\top}\mathrm{d}X_r - \int_t^T \mathrm{d}M_r^s,$$

$$\partial U_t^s = \partial_s\eta(s, X) + \int_t^T \nabla g_r(s, X, \partial U_r^s, \partial V_r^s, U_r^s, V_r^s, \mathcal{Y}_r, \mathcal{Z}_r)\mathrm{d}r - \int_t^T \partial V_r^{s\top}\mathrm{d}X_r - \int_t^T \mathrm{d}\partial M_r^s,$$

and as it is clear from our analysis in Section 6, its well-posedness requires both having a rigorous method to define the mapping $t \longmapsto \partial V_t^t$, as well as deriving a priori estimates for the norm of ∂V_t^t. In analogy with Lemma 6.1 and Remark 3.4, both tasks require us

to make sense of the family of BSDEs with terminal condition $\partial_{ss}\eta$ and generator

$$\nabla^2 g_t(s, x, \tilde{u}, \tilde{v}, u, v, u, v, y, z) := \nabla g_t(s, x, \tilde{u}, \tilde{v}, u, v, y, z)$$
$$+ \sum_{\substack{(\tilde{\pi}_i, \tilde{\pi}_j) \in \tilde{\Pi}^2 \\ (\pi_i, \pi_j) \in \Pi^2}} \tilde{\pi}_i^\top \partial^2_{\pi_i \pi_j} g_t(s, x, u, v, y, z) \tilde{\pi}_j,$$

where $\Pi := (s, u, v_{1:}, ..., v_{n:})$, $\tilde{\Pi} := (1, u, v_{1:}, ..., v_{n:})$ and $\partial^2_{\pi_i \pi_j} g_t(s, x, u, v, y, z)$ denote the second order derivatives of g. Even though we could add assumptions ensuring that the second order derivatives are bounded, it is clear from the second term in the generator that we would necessarily need to consider a quadratic framework.

4 Well-posedness of type-I BSVIEs

We now address the well-posedness of type-I BSVIEs. Let d be a non-negative integer, and f and ξ be jointly measurable functionals such that for any $(s, y, z, u, v) \in [0, T] \times (\mathbb{R}^d \times \mathbb{R}^{n \times d})^2$

$$f : [0, T]^2 \times \mathcal{X} \times (\mathbb{R}^d \times \mathbb{R}^{n \times d})^2 \longrightarrow \mathbb{R}^d, \ f_\cdot(s, \cdot, y, z, u, v) \in \mathcal{P}_{\text{prog}}(\mathbb{R}^d, \mathbb{F}),$$
$$\xi : [0, T] \times \mathcal{X} \longrightarrow \mathbb{R}^d, \ \xi(s, \cdot) \text{ is } \mathcal{F}\text{-measurable}.$$

To derive the main result in this section, we will exploit the well-posedness of $(\mathcal{S})$. Therefore, we work under the following set of assumptions.

Assumption 4.1. (i) $(s, y, z) \longmapsto f_t(s, x, y, z, u, v)$ (resp. $s \longmapsto \xi(s, x)$) is continuously differentiable, uniformly in (t, x, u, v) (resp. in x). Moreover, the mapping $\nabla f : [0, T]^2 \times \mathcal{X} \times (\mathbb{R}^d \times \mathbb{R}^{n \times d})^3 \longrightarrow \mathbb{R}^d$ defined by

$$\nabla f_t(s, x, u, v, y, z, u, v) := \partial_s f_t(s, x, y, z, u, v) + \partial_y f_t(s, x, y, z, u, v)u$$
$$+ \sum_{i=1}^n \partial_{z_{:i}} f_t(s, x, y, z, u, v) v_{i:},$$

satisfies $\nabla f_\cdot(s, \cdot, y, z, u, v, u, v) \in \mathcal{P}_{\text{prog}}(\mathbb{R}^d, \mathbb{F})$ for all $s \in [0, T]$;

(ii) for $\varphi \in \{f, \partial_s f\}$, $(u, v, y, z) \longmapsto \varphi_t(s, x, y, z, u, v)$ is uniformly Lipschitz continuous, i.e. $\exists L_\varphi > 0$, such that for all $(s, t, x, y, \tilde{y}, z, \tilde{z}, u, \tilde{u}, v, \tilde{v})$

$$|\varphi_t(s, x, y, z, u, v) - \varphi_t(s, x, \tilde{y}, \tilde{z}, \tilde{u}, \tilde{v})|$$
$$\leq L_\varphi (|y - \tilde{y}| + |\sigma_t(x)^\top (z - \tilde{z})| + |u - \tilde{u}| + |\sigma_t(x)^\top (v - \tilde{v})|).$$

(iii) $(\tilde{f}, \tilde{f}_\cdot(s), \nabla \tilde{f}_\cdot(s)) := (f_\cdot(\cdot, \cdot, \mathbf{0}), f_\cdot(s, \cdot, \mathbf{0}), \partial_s f_\cdot(s, \cdot, \mathbf{0})) \in \mathbb{L}^{1,2}(\mathbb{R}^d) \times (\mathbb{L}^{1,2,2}(\mathbb{R}^d))^2$.

Let $(\mathcal{H}^\star, \|\cdot\|_{\mathcal{H}^\star})$ denote the space of $(Y, Z, N) \in \mathcal{H}^\star$ such that $\|(Y, Z, N)\|_{\mathcal{H}^\star} < \infty$ where

$$\mathcal{H}^\star := \mathbb{S}^{2,2}(\mathbb{R}^d) \times \overline{\mathbb{H}}^{2,2}(\mathbb{R}^{n \times d}) \times \mathbb{M}^{2,2}(\mathbb{R}^d), \ \|\cdot\|_{\mathcal{H}^\star} := \|Y\|^2_{\mathbb{S}^{2,2}} + \|Z\|^2_{\overline{\mathbb{H}}^{2,2}} + \|N\|^2_{\mathbb{M}^{2,2}}.$$

We consider the n-dimensional type-I BSVIE on $(\mathcal{H}^\star, \|\cdot\|_{\mathcal{H}^\star})$, which for any $s \in [0, T]$ holds $\mathbb{P}-$a.s. for any $t \in [0, T]$,

$$Y_t^s = \xi(s, X) + \int_t^T f_r(s, X, Y_r^s, Z_r^s, Y_r^r, Z_r^r) dr - \int_t^T Z_r^{s\top} dX_r - \int_t^T dN_r^s \qquad (4.1)$$

We work under the following notion of solution.

Definition 4.2. *We say (Y, Z, N) is a solution to the type-I BSVIE (4.1) if $(Y, Z, N) \in \mathcal{H}^\star$ verifies (4.1).*

Defining $h_t(x, y, z, u, v, \mathrm{u}) := f_t(t, x, y, z, u, v) - \mathrm{u}$, we may consider the system, given for any $s \in [0, T]$ by

$$\mathcal{Y}_t = \xi(T, X) + \int_t^T h_r(X, \mathcal{Y}_r, \mathcal{Z}_r, Y_r^r, Z_r^r, \partial Y_r^r) \mathrm{d}r - \int_t^T \mathcal{Z}_r^\top \mathrm{d}X_r - \int_t^T \mathrm{d}\mathcal{N}_r,$$

$$Y_t^s = \xi(s, X) + \int_t^T f_r(s, X, Y_r^s, Z_r^s, \mathcal{Y}_r, \mathcal{Z}_r) \mathrm{d}r - \int_t^T Z_r^{s\top} \mathrm{d}X_r - \int_t^T \mathrm{d}N_r^s, \qquad (\mathcal{S}_f)$$

$$\partial Y_t^s = \partial_s \xi(s, X) + \int_t^T \nabla f_r(s, X, \partial Y_r^s, \partial Z_r^s, Y_r^s, Z_r^s, \mathcal{Y}_r, \mathcal{Z}_r) \mathrm{d}r - \int_t^T \partial Z_r^{s\top} \mathrm{d}X_r - \int_t^T \mathrm{d}\partial N_r^s,$$

Remark 4.3. Let us comment on our set-up for the study type-I BSVIE (4.1).

(i) The necessity of the set of assumptions in Assumption 4.1 to our approach, based on the systems introduced in Section 3, is clear. Compared to the set of assumption made by recent works on BSVIEs in the literature we notice the main difference is the regularity with respect to the s variable we imposed on the data of the problem, i.e. Assumption 4.1.(i). In particular, we highlight that type-I BSVIE (1.5), in which the diagonal of Y, but not of Z is allowed in the generator, had been considered in [27; 53]. In such a scenario, the authors assumed $(\xi, f) \in \mathcal{L}^{2,2}(\mathbb{R}^d) \times \mathbb{L}^{1,2,2}(\mathbb{R}^d)$, and no additional condition is required to obtain the well-posedness of (1.5). As it will be clear from Proposition 6.5 and Remark 4.5 our procedure can be adapted to work under such set of assumptions provided the diagonal of Z is not considered in the generator.

(ii) Moreover, the spaces of the solution considered in [27; 53] also differ, echoing the absence of the diagonal of Z in the generator. The authors work with the notion of C-solution, that is, Y is assumed to be a jointly measurable process, such that $s \longmapsto Y^s$ is continuous in $\mathbb{L}^{1,p}(\mathbb{R}^d)$, $p \geq 2$, and for every $s \in [0, T]$, Y^s is $\mathbb{F}$-adapted with $\mathbb{P}-$a.s. continuous paths. This coincides with our definition of the space $\mathbb{L}^{1,p,2}(\mathbb{R}^d)$. Similarly, Z belongs to the space $\mathbb{H}^{2,2}(\mathbb{R}^{n \times d})$. On the other hand, [54] provides a representation formula for type-I BSVIEs for which the driver allows for the diagonal of Z, but not of Y. More precisely, they introduce a PDE, similar to the one we will introduce in Section 5, prove its well-posedness, and then a Feynman–Kac formula. Naturally, in this case (Y, Z) inherits the regularity of the underlying PDE.

(iii) The main contribution of our methodology to the field of BSVIEs is to be able to accommodate type-I BSVIEs for which the diagonal of Z appears in the generator. For this, the definition of the space $(\mathcal{H}^\star, \| \cdot \|_{\mathcal{H}^\star})$, notably the space $\overline{\mathbb{H}}^{2,2}(\mathbb{R}^{n \times d})$, and Assumption 4.1.($i$) play a central role. As first noticed in [29], see Lemma 6.1, under this assumption one can identify the density, with respect to the Lebesgue measure, of the $\mathbb{H}^2$-valued mapping $s \longmapsto Z^s$, namely the $\mathbb{H}^2$-valued mapping $s \longmapsto \partial Z^s$. This allows us to define the diagonal of Z following [27, Section 2.1], and consequently, introduce the space $\overline{\mathbb{H}}^{2,2}(\mathbb{R}^{n \times d})$.

(iv) We highlight that Theorem 4.4 below guarantees a unique solution to (4.1) in the space $\mathcal{H}^\star$, for which the mapping $s \longmapsto Z^s$ is assumed to be absolutely continuous, cf. [27, Section 2.1]. This allows for extra generality compare to assuming such mapping is, a priori, continuously differentiable. Nevertheless, as a by-product of Assumption 4.1.(i), the unique solution ensured by Theorem 4.4 is automatically $\mathcal{C}^1$ with derivative $s \longmapsto \partial Z^s$. Moreover, not only does our approach identify the dynamics of $(\partial Y, \partial Z, \partial N)$, but also, in the terminology of [27, Section 2.1], it does guarantee $(Z_t^t)_{t \in [0, T]}$ is the unique process that satisfies the (D)-property with respect to Z.

(v) Let us remark that Assumption 4.1.(i), being an assumption on the data of the BSVIE, is easier to verify in practice compare to the regularity required in [27]. Certainly, our results would still hold true if we require the differentiability of data (ξ, f) with respect to the parameter s in the $\mathcal{L}^2$ (resp. $\mathbb{L}^{1,2}$) sense, or, even better, absolute continuity,.

(vi) Lastly, we stress that the above type-I BSVIE is defined for $(s,t) \in [0,T]^2$, as opposed to $0 \leq s \leq t \leq T$. However, anticipating the result of Theorem 4.4, this could be handled by first solving on $(s,t) \in [0,T]^2$ and then consider the restriction to $0 \leq s \leq t \leq T$.

We are now in position to prove the main result of this paper. The next result shows that under the previous choice of data for $(\mathcal{S}_f)$, its solution solves the type-I BSVIE with data (ξ, f) and *vice versa*.

Theorem 4.4. *Let* Assumption 4.1 *hold. Then*

(i) *the well-posedness of $(\mathcal{S}_f)$ is equivalent to that of the* type-I BSVIE (4.1);

(ii) *the* type-I BSVIE (4.1) *is well-posed, and for any* $(Y, Z, N) \in \mathcal{H}^\star$ *solution to* type-I BSVIE (4.1) *there exists* $C > 0$ *such that*

$$\|(Y,Z,N)\|_{\mathcal{H}^\star} \leq C\Big(\|\xi\|_{\mathcal{L}^{2,2}}^2 + \|\partial_s\xi\|_{\mathcal{L}^{2,2}}^2 + \|\tilde{f}\|_{\mathbb{L}^{1,2,2}}^2 + \|\nabla\tilde{f}\|_{\mathbb{L}^{1,2,2}}^2 \Big). \qquad (4.2)$$

Moreover, if $(Y^i, Z^i, N^i) \in \mathcal{H}^\star$ *denotes the solution to* type-I BSVIE (4.1) *with data* (ξ^i, f^i) *for* $i \in \{1,2\}$, *we have*

$$\|(\delta Y, \delta Z, \delta N)\|_{\mathcal{H}^\star}^2 \leq C\Big(\|\delta\xi\|_{\mathcal{L}^2}^2 + \|\delta\partial_s\xi\|_{\mathcal{L}^{2,2}}^2 + \|\delta_1 f\|_{\mathbb{L}^{1,2}}^2 + \|\delta_1\nabla f\|_{\mathbb{L}^{1,2,2}}^2 \Big).$$

Proof. (ii) is a consequence of (i). Indeed, (4.2) follows from Proposition 6.3, and the well-posedness of type-I BSVIE (4.1) from that of $(\mathcal{S}_f)$, which holds by Assumption 4.1 and Theorem 3.6. We now argue (i).

Let $(\mathcal{Y}, \mathcal{Z}, \mathcal{N}, Y, Z, N, \partial Y, \partial Z, \partial N) \in \mathcal{H}$ be a solution to $(\mathcal{S}_f)$. It then follows from Lemma 6.2 that, $\mathbb{P}-$a.s. for any $t \in [0,T]$,

$$Y_t^t = \xi(T,X) + \int_t^T h_r(X, Y_r^r, Z_r^r, \mathcal{Y}_r, \mathcal{Z}_r, \partial Y_r^r)\mathrm{d}r - \int_t^T Z_r^{r\top}\mathrm{d}X_r - \int_t^T \mathrm{d}\widetilde{N}_r, \qquad (4.3)$$

where $\widetilde{N}_t := N_t^t - \int_0^t \partial N_r^r\mathrm{d}r$, $t \in [0,T]$, and $\widetilde{N} \in \mathbb{M}^2(\mathbb{R}^d)$. This shows that

$$\big((Y_t^t)_{t\in[0,T]}, (Z_t^t)_{t\in[0,T]}, \mathcal{Y}_\cdot, \mathcal{Z}_\cdot, (\widetilde{N}_t)_{t\in[0,T]}\big),$$

solves the first BSDE in $(\mathcal{S}_f)$. It then follows from the well-posedness of $(\mathcal{S}_f)$, which holds by Assumption 4.1 and Theorem 3.6, that $\big((Y_t^t)_{t\in[0,T]}, (Z_t^t)_{t\in[0,T]}, (\widetilde{N}_t)_{t\in[0,T]}\big) = (\mathcal{Y}_\cdot, \mathcal{Z}_\cdot, \mathcal{N}_\cdot)$ in $\mathbb{S}^2(\mathbb{R}^d) \times \mathbb{H}^2(\mathbb{R}^{n\times d}) \times \mathbb{M}^2(\mathbb{R}^d)$. Consequently, for any $s \in [0,T]$, it holds, $\mathbb{P}-$a.s. for any $t \in [0,T]$, that

$$Y_t^s = \xi(s,X) + \int_t^T f_r(s, X, Y_r^s, Z_r^s, Y_r^r, Z_r^r)\mathrm{d}r - \int_t^T Z_r^{s\top}\mathrm{d}X_r - \int_t^T \mathrm{d}N_r^s.$$

We are left to show the converse result. Let $(Y, Z, N) \in \mathcal{H}^\star$ be a solution to type-I BSVIE (4.1). We begin by noticing that the processes $\mathcal{Y} := (Y_t^t)_{t\in[0,T]}$, $\mathcal{Z} := (Z_t^t)_{t\in[0,T]}$, $\mathcal{N} := (N_t^t)_{t\in[0,T]}$ are well-defined. In particular, $\mathcal{Z} \in \mathbb{H}^2(\mathbb{R}^{n\times d})$ is well-defined as $Z \in \overline{\mathbb{H}}^{2,2}(\mathbb{R}^{n\times d})$. Moreover, $\mathcal{Y} \in \mathbb{L}^{2,2}(\mathbb{R}^d)$ follows from

$$\|\mathcal{Y}\|_{\mathbb{L}^2}^2 = \mathbb{E}\bigg[\int_0^T |Y_r^r|^2\mathrm{d}r\bigg] \leq \mathbb{E}\bigg[\int_0^T \sup_{t\in[0,T]} |Y_t^r|^2\mathrm{d}r\bigg] = \int_0^T \|Y^r\|_{\mathbb{S}^2}\mathrm{d}r < \infty.$$

Then, since $(\mathcal{Y}, \mathcal{Z}, Y, Z, N) \in \mathbb{L}^2(\mathbb{R}^d) \times \mathbb{H}^2(\mathbb{R}^{n \times d}) \times \mathbb{S}^{2,2}(\mathbb{R}^d) \times \mathbb{H}^{2,2}(\mathbb{R}^{n \times d}) \times \mathbb{M}^{2,2}(\mathbb{R}^d)$ and Assumption 4.1 holds, we can apply Lemma 6.1 and obtain the existence of $(\partial Y, \partial Z, \partial N) \in \mathbb{S}^{2,2}(\mathbb{R}^d) \times \mathbb{H}^{2,2}(\mathbb{R}^{n \times d}) \times \mathbb{M}^{2,2}(\mathbb{R}^d)$ such that for $s \in [0, T]$, $\mathbb{P}$−a.s. for $t \in [0, T]$,

$$\partial Y_t^s = \partial_s \xi(s, X) + \int_t^T \nabla f_r(s, X, \partial Y_r^s, \partial Z_r^s, Y_r^s, Z_r^s, Y_r^r, Z_r^r) \mathrm{d}r - \int_t^T \partial Z_r^{s\top} \mathrm{d}X_r - \int_t^T \mathrm{d}\partial N_r^s,$$

We claim that $\mathfrak{h} := (\mathcal{Y}, \mathcal{Z}, \widetilde{N}, Y, Z, N, \partial Y, \partial Z, \partial N)$ is a solution to $(\mathcal{S}_f)$, where $\widetilde{N}_t := N_t^t - \int_0^t \partial N_r^r \mathrm{d}r$, $t \in [0, T]$. For this, we first note that in light of Lemmata 6.1 and 6.2 we have that

$$\mathcal{Y}_t = \xi(T, X) + \int_t^T h_r(X, \mathcal{Y}_r, \mathcal{Z}_r, Y_r^r, Z_r^r, \partial Y_r^r) \mathrm{d}r - \int_t^T \mathcal{Z}_r^\top \mathrm{d}X_r - \int_t^T \mathrm{d}\widetilde{N}_r, \qquad (4.4)$$

and $\widetilde{\mathcal{N}} \in \mathbb{M}^{2,2}(\mathbb{R}^d)$. We are only left to argue $\mathcal{Y} \in \mathbb{S}^2(\mathbb{R}^d)$. Note that by Assumption 4.1 and Equation (6.1) there exists $C > 0$ such that

$$|\mathcal{Y}_t|^2 \leq C \bigg(|\xi|^2 + \bigg(\int_0^T |\tilde{f}_r| \mathrm{d}r \bigg)^2 + \int_0^T \big(|\mathcal{Y}_r|^2 + |\sigma_r^\top \mathcal{Z}_r|^2 + |Y_r^r|^2 + |\sigma_r^\top Z_r^r|^2 + |\partial Y_r^r|^2 \big) \mathrm{d}r$$
$$+ \bigg| \int_t^T \mathcal{Z}_r^\top \mathrm{d}X_r + \int_t^T \mathrm{d}\widetilde{N}_r \bigg|^2 \bigg).$$

Moreover, by Doob's inequality we have $\mathbb{E}\bigg[\sup_{t \in [0,T]} \bigg| \int_0^t \mathcal{Z}^\top \mathrm{d}X_r \bigg|^2 \bigg] \leq 4\|\mathcal{Z}\|_{\mathbb{H}^2}^2$ and thus Equation (6.16) yields

$$\|\mathcal{Y}\|_{\mathbb{S}^2}^2 \leq C \Big(\|\xi\|_{\mathcal{L}^{2,2}}^2 + \|\partial_s \xi\|_{\mathcal{L}^{2,2}}^2 + \|\tilde{f}\|_{\mathbb{L}^{1,2,2}}^2 + \|\nabla \tilde{f}\|_{\mathbb{L}^{1,2,2}}^2$$
$$+ \|\mathcal{Y}\|_{\mathbb{L}^2}^2 + \|Y\|_{\mathbb{L}^{2,2}}^2 + \|Z\|_{\overline{\mathbb{H}}^{2,2}}^2 + \|\partial Z\|_{\mathbb{H}^{2,2}}^2 \Big) < \infty.$$

We conclude $\|\mathfrak{h}\|_{\mathcal{H}} < \infty$, $\mathfrak{h} \in \mathcal{H}$ and thus $\mathfrak{h}$ solves $(\mathcal{S}_f)$. $\qquad \square$

Remark 4.5. There are two noticeable differences between Theorem 4.4 and the results in the literature on type-I BSVIEs (1.3), (1.4) and (1.5), as previously studied in [71], [54; 29] and [27; 53], respectively. The first is the additional terms, involving the derivative with respect to the parameter s of the data, appearing in the *a priori* estimates and the stability. The second one, is the space where the solution lives. Both differences are due to the fact that we are handling the diagonal term for Z in the generator.

In fact, in light of Proposition 6.5 for the case of type-I BSVIEs (1.5), i.e. where only the diagonal of Y is allowed in the generator, one can work in the more standard (compared to the existing literature) space $(\mathfrak{H}^\star, \|\cdot\|_{\mathfrak{H}^\star})$ given by

$$\mathfrak{H}^\star := \mathbb{S}^{2,2}(\mathbb{R}^d) \times \mathbb{H}^{2,2}(\mathbb{R}^{n \times d}) \times \mathbb{M}^{2,2}(\mathbb{R}^d), \quad \|(Y, Z, N)\|_{\mathfrak{H}^\star}^2 := \|Y\|_{\mathbb{S}^{2,2}}^2 + \|Z\|_{\mathbb{H}^{2,2}}^2 + \|N\|_{\mathbb{M}^{2,2}}^2.$$

Then, the *a priori* estimate (4.2) simplifies to

$$\|(Y, Z, N)\|_{\mathfrak{H}^\star} \leq C \big(\|\xi\|_{\mathcal{L}^{2,2}}^2 + \|\tilde{f}\|_{\mathbb{L}^{1,2,2}}^2 \big),$$

and for (Y^i, Z^i, N^i) the solution to type-I BSVIE (1.5) with data (ξ^i, f^i) for $i \in \{1, 2\}$, we obtain

$$\|(\delta Y, \delta Z, \delta N)\|_{\mathfrak{H}^\star}^2 \leq C \Big(\|\delta \xi\|_{\mathcal{L}^2}^2 + \|\delta_1 f\|_{\mathbb{L}^{1,2}}^2 \Big).$$

5 Type-I BSVIEs, parabolic PDEs and time-inconsistent control

This section is devoted to the application of our results in Section 4 to the problem of time-inconsistent control for sophisticated agents. Moreover, we also reconcile seemingly different approaches to the study of this problem.

5.1 Representation formula for adapted solutions of type-I BSVIEs

Building upon the fact that second-order, parabolic, semilinear PDEs of HJB type admit a non–linear Feynman–Kac representation formula, we can identify the family of PDEs associated to Type-I BSVIEs. This is similar to the representation of forward backward stochastic differential equations, FBSDEs for short, see [64].

For $(s,t,x,u,y,v,z,\gamma,\Sigma) \in [0,T]^2 \times \mathcal{X} \times (\mathbb{R}^d)^2 \times (\mathbb{R}^{n \times d})^2 \times (\mathbb{R}^{n \times n})^d \times \mathbb{R}^{n \times m}$, define $\mathrm{Tr}\big[\Sigma\Sigma^\top\gamma\big] \in \mathbb{R}^d$ by $\big(\mathrm{Tr}\big[\Sigma\Sigma^\top\gamma\big]\big)_i := \mathrm{Tr}\big[\Sigma\Sigma^\top\gamma_i\big]$, $i \in \{1,\ldots,d\}$. Let f and σ be as in the preceding section, and $b : [0,T] \times \mathcal{X} \longrightarrow \mathbb{R}^m$ be bounded, $b_\cdot(X) \in \mathcal{P}_{\mathrm{meas}}(\mathbb{R}^m,\mathbb{F})$ and

$$G\big(s,t,x,u,v,y,z,\gamma\big) := v^\top\sigma(t,x)b(t,x) + \frac{1}{2}\mathrm{Tr}[\sigma\sigma^\top(t,x)\gamma] + f(s,t,x,u,v,y,z).$$

Proposition 5.1. Suppose $\varphi_t(X,\cdot) = \varphi_t(X_t,\cdot)$ for $\varphi \in \{b,\sigma,f,\partial_s f\}$, and $\varphi(s,X) = \varphi(s,X_T)$ for $\varphi \in \{\xi,\partial_s\xi\}$. For $\mathcal{O} := [0,T) \times [0,T] \times \mathbb{R}^n$, let $\nu \in \mathcal{C}^d_{1,1,2}([0,T]^2 \times \mathbb{R}^n)$ be a classical solution to

$$\begin{cases} \partial_t\nu(s,t,x) + G\big(s,t,x,\nu(s,t,x),\partial_x\nu(s,t,x),\nu(t,t,x),\partial_x\nu(t,t,x),\partial^2_{xx}\nu(s,t,x)\big) = 0, \text{ on } \mathcal{O}, \\ \nu(s,T,x) = \xi(s,x), \ (s,x) \in [0,T] \times \mathbb{R}^n, \end{cases}$$

for which $\nu(s,t,x)$ and $\partial_x\nu(s,t,x)$ have uniform exponential growth in x[5], i.e. $\exists C > 0$, $\forall(s,t,x) \in [0,T]^2 \times \mathcal{X}$

$$|\nu(t,x)| + |\partial_x\nu(t,x)| \leq C\exp(C|x|_1).$$

Then, $Y^s_t := \nu(s,t,X_t)$, and $Z^s_t := \partial_x\nu(s,t,X_t)$ define a solution to the type-I BSVIE given for every $s \in [0,T]$ by

$$Y^s_t = Y^s_T + \int_t^T \big(f_r(s,X_r,Y^s_r,Z^s_r,Y^r_r,Z^r_r) + Z^{s\top}_r\sigma_r(X_r)b_r(X_r)\big)\mathrm{d}r - \int_t^T Z^{s\top}_r\mathrm{d}X_r, \quad (5.1)$$

Proof. Let $s \in [0,T]$ and $\mathbb{P}$ as in Section 2. Applying Itō's formula to the process Y^s_t we find that $\mathbb{P}-$a.s.

$$Y^s_t = Y^s_T - \int_t^T \Big(\partial_t\nu(s,r,X_r) + \frac{1}{2}\mathrm{Tr}[\sigma\sigma^\top_r(X_r)\partial^2_{xx}\nu(s,r,X_r)]\Big)\mathrm{d}r - \int_t^T Z^{s\top}_r\mathrm{d}X_r$$

$$= Y^s_T + \int_t^T \big(f_r(s,X_r,Y^s_r,Z^s_r,Y^r_r,Z^r_r) + Z^{s\top}_r\sigma_r(X_r)b_r(X_r)\big)\mathrm{d}r - \int_t^T Z^{s\top}_r\mathrm{d}X_r.$$

We verify the integrability of (Y,Z). As σ is bounded, X_t has exponential moments of any order which are bounded on $[0,T]$, i.e. $\exists C > 0$, such that $\sup_{t \in [0,T]} \mathbb{E}^\mathbb{P}[\exp(c|X_t|_1)] \leq C < \infty$, for any $c > 0$, where C depends on T and the bound on σ. This together with the growth condition on $\nu(s,t,x)$ and $\partial_x\nu(s,t,x)$ yield the integrability. $\qquad\square$

Remark 5.2. In the previous result the type-I BSVIEs has an additional term linear in z. This is a consequence of the dynamics of X under $\mathbb{P}$, see Section 2.1. Nevertheless, as b is bounded, we can define $\mathbb{P}^b \in \mathrm{Prob}(\mathcal{X})$, equivalent to $\mathbb{P}$ so that by Girsanov's theorem $B^b := B - \int_0^\cdot b_r(X_r)\mathrm{d}r$ is a $\mathbb{P}^b-$Brownian motion and

$$X_t = x_0 + \int_0^t \sigma_r(X_r)b_r(X_r)\mathrm{d}r + \int_0^t \sigma_r(X_r)\mathrm{d}B^b_r, \ t \in [0,T], \ \mathbb{P}^b-\text{a.s.},$$

[5]Here, $|\cdot|_1$ denotes the ℓ_1 norm in $\mathbb{R}^n$, i.e. for $x \in \mathbb{R}^n$, $|x|_1 := \sum_{i=1}^d |x_i|$

and consequently

$$Y_t^s = Y_T^s + \int_t^T f_r(s, X_r, Y_r^s, Z_r^s, Y_r^r, Z_r^r) \mathrm{d}r - \int_t^T Z_r^\top \mathrm{d}X_r, \ t \in [0, T], \ \mathbb{P}-\text{a.s.}, \ s \in [0, T].$$

5.2 On equilibria and their value function in time-inconsistent control problems

The game theoretic approach to time-inconsistent control problems in continuous-time started with the Markovian setting, and is grounded in the notion of equilibrium first proposed in Ekeland and Pirvu [24], Ekeland and Lazrak [23], and the infinite family of PDEs, or Hamilton–Jacobi–Bellman equation, provided by Björk, Khapko, and Murgoci [13], see Equation (5.2) below. Soon after, Wei, Yong, and Yu [68] presented a verification argument via a one dimensional PDE, but over an extended domain, see Equation (5.3) below. Both approaches have generated independent lines of research in the community, including both analytic and probabilistic methods, but no compelling connections have been established, as far as we know.

BSDEs and BSVIEs appear naturally as part of the probabilistic study of these problems. This approach allows extensions to a non-Markovian framework, and to reward functionals given by recursive utilities. Indeed, the approaches in [29] and [54] address these directions, and can be regarded as extensions of [13] and [68], respectively. As such, it is not surprising that in order to characterise an equilibrium and its associated value function, both [29] and [54] lay down an infinite family of BSDEs, and a type-I BSVIEs, respectively. In fact, [29, Theorem 3.7] and [54, Theorem 5.1] establish representation formulae for the analytic, i.e. PDEs, counterparts. Moreover, [29] noticed that their approach through BSDEs led to the well-posedness of a BSVIE. This is nothing but a manifestation of Theorem 4.4 which reconciles, at the probabilistic level, the findings of [29] and [54]. Moreover, we also include Proposition 5.3 which does the same at the PDE level. To sum up, we can visualise this in the next picture.
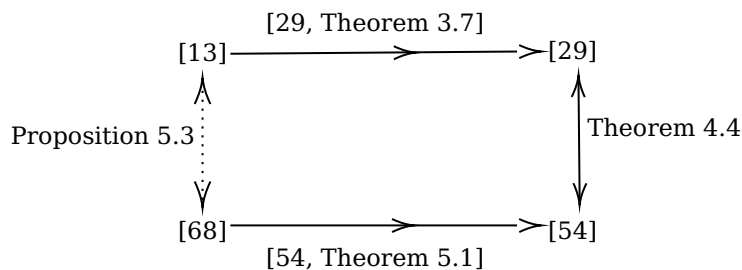
Let $A \subseteq \mathbb{R}^p$ be a compact set, $\bar{f} : [0, T]^2 \times \mathbb{R}^n \times A \longrightarrow \mathbb{R}$, $b : [0, T] \times \mathbb{R}^n \times A \longrightarrow \mathbb{R}^m$, bounded, with $f.(s, \cdot, a) \in \mathcal{P}_{\text{prog}}(\mathbb{R}, \mathbb{F})$ and $b.(\cdot, a) \in \mathcal{P}_{\text{meas}}(\mathbb{R}^m, \mathbb{F})$ for $(s, a) \in [0, T] \times A$. With this we may define

$$\bar{g}(s, t, x, a, v) := \bar{f}(s, t, x, a) + v \cdot \sigma(t, x)b(t, x, a), \ H(s, t, x, v) := \sup_{a \in A} \bar{g}(s, t, x, a, v),$$

$$\nabla \bar{g}(s, t, x, a, \mathrm{v}) := \partial_s \bar{f}(s, t, x, a) + \mathrm{v} \cdot \sigma(t, x)b(t, x, a),$$

and assume there exists $a^\star : [0, T]^2 \times \mathbb{R}^n \times \mathbb{R}^n \longrightarrow A$, measurable, such that, $H(s, t, x, v) = \bar{g}(s, t, x, a^\star(s, t, x, v), v)$.

Following the approach of [54], let us assume that given an admissible A-valued strategy ν over the interval $[s, T]$, the reward at $s \in [0, T]$ is given by the value at s of the Y coordinate of the solution to the type-I BSVIE given by

$$Y_t^\nu = \xi(t, X_T) + \int_t^T \bar{g}_r(t, X_r, \nu_r, Z_r^t) \mathrm{d}r - \int_t^T Z_r^t \cdot \mathrm{d}X_r, \ t \in [s, T], \ \mathbb{P}-\text{a.s.}$$

[54] finds that the value along the *equilibrium policy* coincides with the Y coordinate of the following type-I BSVIE

$$Y_t = \xi(t, X_T) + \int_t^T \bar{g}_r\big(t, X_r, a^\star(r, r, X_r, Z_r^r), Z_r^t\big)\mathrm{d}r - \int_t^T Z_r^t \cdot \mathrm{d}X_r, \ t \in [0, T], \ \mathbb{P}-\text{a.s.},$$

where the diagonal of Z appears in the generator. However, decoupling the dependence between the time variable and the variable source of time-inconsistency, we can define, for $s \in [0, T]$,

$$Y_t^s := \xi(s, X_T) + \int_t^T \bar{g}_r\big(s, X_r, a^\star(r, r, X_r, Z_r^r), Z_r^s\big)\mathrm{d}r - \int_t^T Z_r^s \cdot \mathrm{d}X_r, \ t \in [0, T], \ \mathbb{P}-\text{a.s.}$$

It then follows from Theorem 4.4 that this approach is equivalent to that of [29] based on the system, which for any $s \in [0, T]$ holds $\mathbb{P}-$a.s. for $t \in [0, T]$,

$$Y_t = \xi(T, X_T) + \int_t^T \big(H_r(r, X_r, Z_r) - \partial Y_r^r\big)\mathrm{d}r - \int_t^T Z_r \cdot \mathrm{d}X_r,$$

$$\partial Y_t^s = \partial_s \xi(s, X_T) + \int_t^T \nabla \bar{g}_r\big(s, X_r, a^\star(r, r, X_r, Z_r^r), \partial Z_r^s\big)\mathrm{d}r - \int_t^T \partial Z_r^s \cdot \mathrm{d}X_r.$$

We now move on to establish the connection of the analyses at the PDE level. The original result of [13] is based on the semi-linear PDE system of HJB type given for $\big(V(t, x), \mathcal{J}(s, t, x)\big) \in \mathcal{C}_{1,2}([0, T] \times \mathbb{R}^n) \times \mathcal{C}_{1,1,2}([0, T]^2 \times \mathbb{R}^n)$ by

$$\begin{cases} \partial_t V(t, x) + \dfrac{1}{2}\mathrm{Tr}[\sigma\sigma^\top(t, x)\partial_{xx}V(t, x)] \\ \qquad + H(t, t, x, \partial_x V(t, x)) - \partial_s \mathcal{J}(t, t, x) = 0, \ \text{on } \mathcal{O}, \\ \partial_t \mathcal{J}(s, t, x) + \dfrac{1}{2}\mathrm{Tr}[\sigma\sigma^\top(t, x)\partial_{xx}\mathcal{J}(s, t, x)] \\ \qquad + \bar{g}\big(s, t, x, \partial_x \mathcal{J}(s, t, x), a^\star(t, t, x, \partial_x V(t, x))\big) = 0, \ \text{on } \mathcal{O}, \\ V(T, x) = \xi(T, x), \ \mathcal{J}(s, T, x) = \xi(s, x), \ (s, x) \in [0, T] \times \mathbb{R}^d. \end{cases} \quad (5.2)$$

On the other hand, [68] considers the equilibrium HJB equation for $\mathcal{J}(s, t, x) \in \mathcal{C}_{1,1,2}([0, T]^2 \times \mathbb{R}^n)$ given by

$$\begin{cases} \partial_t \mathcal{V}(s, t, x) + \dfrac{1}{2}\mathrm{Tr}[\sigma\sigma^\top(t, x)\partial_{xx}\mathcal{V}(s, t, x)] \\ \qquad + \bar{g}\big(s, t, x, \partial_x \mathcal{V}(t, t, x), a^\star(t, t, x, \partial_x \mathcal{V}(t, t, x))\big) = 0, \ \text{on } \mathcal{O}, \\ \mathcal{V}(s, T, x) = \xi(s, x), \ (s, x) \in [0, T] \times \mathbb{R}^d. \end{cases} \quad (5.3)$$

By setting $\mathcal{V}(s, t, x) = \mathcal{J}(s, t, x)$, it is immediate that a solution to (5.2) defines a solution to (5.3). The next proposition establishes the converse result.

Proposition 5.3. Suppose (5.2) and (5.3) are well-posed.

(*i*) Let $\mathcal{J}(s, t, x) \in \mathcal{C}_{1,1,2}([0, T]^2 \times \mathbb{R}^n)$ solve (5.2). Then $\mathcal{V}(s, t, x) := \mathcal{J}(s, t, x)$ solves (5.3).

(*ii*) Let $\mathcal{V}(s, t, x) \in \mathcal{C}_{1,1,2}([0, T]^2 \times \mathbb{R}^n)$ solve (5.3). Then $V(t, x) := \mathcal{V}(t, t, x)$ and $\mathcal{J}(s, t, x) := \mathcal{V}(s, t, x)$ solve (5.2).

Proof. We are only left to argue (*ii*). It is clear $\big(V(t, x), \mathcal{J}(s, t, x)\big) \in \mathcal{C}_{1,2}([0, T] \times \mathbb{R}^n) \times \mathcal{C}_{1,1,2}([0, T]^2 \times \mathbb{R}^n)$, the results follows as $-\partial_t V(t, x) + \partial_s \mathcal{J}(t, t, x) = -\partial_t \mathcal{V}(t, t, x)$. $\qquad \square$

6 Analysis of the BSDE system

Let us first recall the elementary inequalities

$$\left(\sum_{i=1}^{n} a_i\right)^2 \le n \sum_{i=1}^{n} a_i^2, \tag{6.1}$$

valid for any positive integer n and any collection $(a_i)_{1 \le i \le n}$ of non-negative numbers, as well as, Young's inequality which guarantees that for any $\varepsilon > 0$, $2ab \le \varepsilon a^2 + \varepsilon^{-1} b^2$.

In order to alleviate notations, and as it is standard in the literature, we suppress the dependence on ω, i.e. on X, in the functionals, and, write $\mathbb{E}$ instead of $\mathbb{E}^{\mathbb{P}}$ as the underlying probability measure $\mathbb{P}$ is fixed. Moreover, we will write $\mathbb{I}^2$ instead of $\mathbb{I}^2(E)$ for any of the integrability spaces involved, the specific space E is fixed and understood without ambiguity.

6.1 Regularity of the system and the diagonal processes

In preparation to the proof of Theorem 3.6, we present next a couple of lemmata from which we will benefit in the following. As a historical remark, we mention the following is in the spirit of the analysis in Protter [43, Section 3] and Pardoux and Protter [41] of forward Volterra integral equations.

A technical detail is to identify appropriate spaces so that given $(\partial U, U, V, M)$ one can rigorously define the processes $\big((U_t^t)_{t \in [0,T]}, (V_t^t)_{t \in [0,T]}, (M_t^t)_{t \in [0,T]}, (\partial U_t^t)_{t \in [0,T]}\big)$. It is known that for $U \in \mathbb{S}^{2,2}$, the diagonal process $(U_t^t)_{t \in [0,T]}$ is well-defined. Indeed, this follows from the pathwise regularity of U^s for $s \in [0,T]$ and has been noticed since [27; 29; 65]. The same argument works for $(\partial U, M) \in \mathbb{S}^{2,2} \times \mathbb{M}^{2,2}$. Unfortunately, the same reasoning cannot be applied for arbitrary $V \in \mathbb{H}^{2,2}$ and motives the introduction of the space $\overline{\mathbb{H}}^{2,2}$, see Remark 3.4.

Lemma 6.1. Let Assumption 3.1 hold and $(\mathcal{Y}, \mathcal{Z}) \in \mathbb{L}^2 \times \mathbb{H}^2$. Let $(U, V, M) \in \mathbb{L}^{2,2} \times \mathbb{H}^{2,2} \times \mathbb{M}^{2,2}$ be a solution to for $s \in [0, T]$

$$U_t^s = \eta(s) + \int_t^T g_r(s, U_r^s, V_r^s, \mathcal{Y}_r, \mathcal{Z}_r)\mathrm{d}r - \int_t^T V_r^{s\top}\mathrm{d}X_r - \int_t^T \mathrm{d}M_r^s, \ t \in [0,T], \ \mathbb{P}-\text{a.s.}$$

(i) There exists $(\partial U, \partial V, \partial M) \in \mathbb{S}^{2,2} \times \mathbb{H}^{2,2} \times \mathbb{M}^{2,2}$ unique solution to the equation,

$$\partial U_t^s = \partial_s \eta(s) + \int_t^T \nabla g_r(s, \partial U_r^s, \partial V_r^s, U_r^s, V_r^s, \mathcal{Y}_r, \mathcal{Z}_r)\mathrm{d}r$$

$$- \int_t^T \partial V_r^{s\top}\mathrm{d}X_r - \int_t^T \mathrm{d}\partial M_r^s, \ t \in [0,T], \ \mathbb{P}-\text{a.s.}, s \in [0,T]; \tag{6.2}$$

(ii) there exist $C > 0$, such that for all $c > 4L_g$ and $(s, t) \in [0,T]^2$

$$\mathbb{E}\left[\int_t^T \mathrm{e}^{cr}|\sigma_r^\top \partial V_r^s|^2 \mathrm{d}r\right] \le C\mathbb{E}\left[\mathrm{e}^{cT}|\partial_s \eta(s)|^2 + \left(\int_t^T \mathrm{e}^{\frac{c}{2}r}|\nabla \tilde{g}_r(s)|\mathrm{d}r\right)^2\right.$$

$$\left. + \int_t^T \mathrm{e}^{cr}\big(|U_r^s|^2 + |\sigma_r^\top V_r^s|^2 + |\mathcal{Y}_r|^2 + |\sigma_r^\top \mathcal{Z}_r|^2\big)\mathrm{d}r\right]; \tag{6.3}$$

(iii) for any $s \in [0,T]$, as elements of $\mathbb{S}^2 \times \mathbb{H}^2 \times \mathbb{M}^2$

$$\left(\int_s^T \partial U^r \mathrm{d}r, \int_s^T \partial V^r \mathrm{d}r, \int_s^T \partial M^r \mathrm{d}r\right) = (U^T - U^s, V^T - V^s, M^T - M^s); \tag{6.4}$$

(iv) $V \in \overline{\mathbb{H}}^{2,2}$. Moreover, for $\mathcal{V} := (V_t^t)_{t \in [0,T]}$ and $\varepsilon > 0$, $\mathbb{P}-$a.s.

$$\int_t^T |\sigma_u^\top \mathcal{V}_u|^2 \mathrm{d}u \leq \int_t^T |\sigma_u^\top V_u^t|^2 \mathrm{d}u + \int_t^T \int_r^T \varepsilon |\sigma_u^\top V_u^r|^2 + \varepsilon^{-1} |\sigma_u^\top \partial V_u^r|^2 \mathrm{d}u \mathrm{d}r, \ t \in [0,T]. \quad (6.5)$$

Proof. Note that in light of Assumption 3.1.(i), for $(t,s,x,u,v,y,z) \in [0,T]^2 \times \mathcal{X} \times \mathbb{R}^{d_2} \times \mathbb{R}^{d_2 \times n} \times \mathbb{R}^{d_1} \times \mathbb{R}^{d_1 \times n}$, $(\mathrm{u},\mathrm{v}) \longmapsto \nabla g_t(s,x,\mathrm{u},\mathrm{v},u,v,y,z)$ is linear. Therefore, for any $s \in [0,T]$ the second equation defines a linear BSDE, in $(\partial U^s, \partial V^s)$, whose generator at zero, by Assumption 3.1.(iii), is in $\mathbb{L}^{1,2}$. Therefore, its solution $(\partial U^s, \partial V^s, \partial M^s) \in \mathbb{S}^2 \times \mathbb{H}^2 \times \mathbb{M}^2$ is well-defined from classic results, see for instance Zhang [73] or [25]. The continuity of the applications $s \longmapsto (\partial U^s, \partial V^s, \partial M^s)$, e.g. $([0,T], \mathcal{B}([0,T])) \longrightarrow (\mathbb{S}^2, \|\cdot\|_{\mathbb{S}^2}) : s \longmapsto \partial U^s$, follows from the classical stability results of BSDE, and that by assumption $(U,V,Y,Z) \in \mathbb{L}^{2,2} \times \mathbb{H}^{2,2} \times \mathbb{L}^2 \times \mathbb{H}^2$ and $s \longmapsto (\partial_s \eta(s), \nabla \tilde{g}(s))$ is continuous. This establishes $(\partial U, \partial V, \partial M) \in \mathbb{S}^{2,2} \times \mathbb{H}^{2,2} \times \mathbb{M}^{2,2}$.

(ii) follows from classic a priori estimates, but when the norms are considered over $[t,T]$ instead of $[0,T]$. Indeed, following the argument in [25, Proposition 2.1], applying Itō's formula to $\mathrm{e}^{ct}|\partial U_t^s|^2$ we may find $C > 0$ such that for any $c > 4L_g$ and $(s,t) \in [0,T]^2$

$$\mathbb{E}\left[\sup_{r \in [t,T]} \left\{ \mathrm{e}^{cr} |\partial U_r^s|^2 \right\} + \int_t^T \mathrm{e}^{cr} |\sigma_r^\top \partial V_r^s|^2 \mathrm{d}r + \int_t^T \mathrm{e}^{cr-} \mathrm{d}[\partial M^s]_r \right]$$

$$\leq C \mathbb{E}\left[\mathrm{e}^{cT} |\partial_s \eta(s)|^2 + \left(\int_t^T \mathrm{e}^{\frac{c}{2}r} |\nabla g_t(s,0,0,U_r^s,V_r^s,\mathcal{Y}_r,\mathcal{Z}_r)| \mathrm{d}r \right)^2 \right]$$

$$= C \mathbb{E}\left[\mathrm{e}^{cT} |\partial_s \eta(s)|^2 + \left(\int_t^T \mathrm{e}^{\frac{c}{2}r} |\partial_s g_r(s,U_r^s,V_r^s,\mathcal{Y}_r,\mathcal{Z}_r)| \mathrm{d}r \right)^2 \right]$$

$$\leq C \mathbb{E}\left[\mathrm{e}^{cT} |\partial_s \eta(s)|^2 + \left| \int_t^T \mathrm{e}^{\frac{c}{2}r} |\nabla \tilde{g}_r(s)| \mathrm{d}r \right|^2 \right.$$

$$\left. + \int_t^T \mathrm{e}^{cr} \left(|U_t^s|^2 + |\sigma_t^\top V_t^s|^2 + |\mathcal{Y}_t|^2 + |\sigma_t^\top \mathcal{Z}_t|^2 \right) \mathrm{d}t \right],$$

where in the second inequality we exploited the fact $(u,v,y,z) \longmapsto \partial_s g(t,s,x,u,v,y,z)$ is Lipschitz, see Assumption 3.1.(iii), and $C > 0$ was appropriately updated.

Next, we assume (iii) and show (iv). In light of (i) and (iii), $s \longmapsto \partial V^s$ is the density of $s \longmapsto V^s$ with respect to the Lebesgue measure. Arguing as in Remark 2.2, we obtain that we can define

$$V_t^t := V_t^T - \int_t^T \partial V_t^r \mathrm{d}r, \ \mathrm{d}t \otimes \mathrm{d}\mathbb{P}-\text{a.e. in } [0,T] \times \mathcal{X}. \quad (6.6)$$

We now verify (6.5). By definition of $\mathcal{V}$, Fubini's theorem and Young's inequality we have that for $\varepsilon > 0$

$$\int_t^T |\sigma_u^\top V_u^u|^2 - |\sigma_u^\top V_u^t|^2 \mathrm{d}u = \int_t^T \int_t^u 2\mathrm{Tr}\left[V_u^{r\top} \sigma_u \sigma_u^\top \partial V_u^r \right] \mathrm{d}r \mathrm{d}u$$

$$= \int_t^T \int_r^T 2\mathrm{Tr}\left[V_u^{r\top} \sigma_u \sigma_u^\top \partial V_u^r \right] \mathrm{d}u \mathrm{d}r$$

$$\leq \int_t^T \int_r^T \varepsilon |\sigma_u^\top V_u^r|^2 + \varepsilon^{-1} |\sigma_u^\top \partial V_u^r|^2 \mathrm{d}u \mathrm{d}r.$$

Thus $\|\mathcal{V}\|_{\mathbb{H}^2} < \infty$ and consequently $V \in \overline{\mathbb{H}}^{2,2}$. This proves (iv).

We now argue (iii). We also remark that a similar argument to the one in (i) shows that under Assumption 3.1 $U \in \mathbb{S}^{2,2}$. We know the mapping $[0,T] \ni s \longmapsto (\partial Y^s, \partial Z^s, \partial M^s)$

is continuous, in particular integrable. A formal integration with respect to s to (6.2) leads to

$$\int_s^T \partial U_t^r \mathrm{d}r = \int_s^T \int_t^T \partial_s g_u(r, U_u^r, V_u^r, \mathcal{Y}_u, \mathcal{Z}_u) + \partial_y g_u(r, U_u^r, V_u^r, \mathcal{Y}_u, \mathcal{Z}_u) \partial U_u^r \mathrm{d}u \mathrm{d}r$$

$$+ \int_t^T \int_s^T \sum_{i=1}^n \partial_{v_i} g_u(r, U_u^r, V_u^r, \mathcal{Y}_u, \mathcal{Z}_u)(\partial V_u^r)_{:i} \mathrm{d}u \mathrm{d}r$$

$$- \int_s^T \int_t^T \partial Z_u^{r\top} \mathrm{d}X_u \mathrm{d}r - \int_t^T \int_s^T \mathrm{d}\partial M_u^r \mathrm{d}r + \int_s^T \partial_s \eta(r) \mathrm{d}r.$$

Fix $s \in [0, T]$ and let $(\Pi^\ell)_\ell$ be a properly chosen sequence of partitions of $[s, T]$, as in van Neerven [52, Theorem 1], $\Pi^\ell = (s_i)_{i=1,\dots,n_\ell}$ with $\|\Pi^\ell\| := \sup_i |s_{i+1} - s_i| \leq \ell$. To ease the notation, we set $\Delta s_i^\ell := s_i^\ell - s_{i-1}^\ell$, and, for a generic family process $(\phi^s)_{s \in [0,T]}$, and a mapping $s \longmapsto \partial_s \eta(s, x)$, we define $\delta\eta(x) := \eta(T, x) - \eta(s, x)$, and

$$I^\ell(\phi) := \sum_{i=0}^{n_\ell} \Delta s_i^\ell \phi^{s_i^\ell}, \ \delta\phi := \phi^T - \phi^s, \ I^\ell(\partial_s \eta(\cdot, x)) := \sum_{i=0}^n \Delta s_i^\ell \partial_s \eta(s_i^\ell, x).$$

We then notice that for any $t \in [0, T]$

$$I^\ell(\partial U)_t - (\delta U)_t$$

$$= I^\ell(\partial_s \eta(\cdot)) - \delta\eta + \int_t^T \left[I^\ell\big(\partial_s g_u(\cdot, U_u^{\cdot}, V_u^{\cdot}, \mathcal{Y}_u, \mathcal{Z}_u)\big) + I^\ell\big(\partial_y g_u(\cdot, U_u^{\cdot}, V_u^{\cdot}, \mathcal{Y}_u, \mathcal{Z}_u)\partial U_u^{\cdot}\big) \right] \mathrm{d}u$$

$$+ \int_t^T \left[I^\ell\big(\partial_z g_u(\cdot, U_u^{\cdot}, V_u^{\cdot}, \mathcal{Y}_u, \mathcal{Z}_u)\partial V_u^{\cdot}\big) - \delta g_u(\cdot, U_u^{\cdot}, V_u^{\cdot}, \mathcal{Y}_u, \mathcal{Z}_u) \right] \mathrm{d}u$$

$$- \int_t^T \big(I^\ell(\partial V)_r - (\delta V)_r\big)^\top \mathrm{d}X_u - I^\ell(\partial M_T^{\cdot} - \partial M_t^{\cdot}) - \delta(M_T - M_t).$$

We now note that the integrability of $(\partial U, \partial V)$ and (U, V) yields $I^\ell(\partial U) - (\delta U) \in \mathbb{S}^{2,2}$ and $I^\ell(\partial V) - (\delta V) \in \mathbb{H}^{2,2}$. Therefore, Bouchard, Possamaï, Tan, and Zhou [14, Theorem 2.2] yields

$$\|I^\ell(\partial U) - (\delta U)\|_{\mathbb{S}^{2,2}}^2 + \|I^\ell(\partial V) - (\delta V)\|_{\mathbb{H}^{2,2}}^2 + \|I^\ell(\partial M) - (\delta M)\|_{\mathbb{M}^{2,2}}^2$$

$$\leq C\mathbb{E}\left[\left| I^\ell(\partial_s \eta(\cdot)) - \delta\eta \right|^2 + \left| \int_t^T \left| I^\ell\big(\partial_s g_u(\cdot, U_u^{\cdot}, V_u^{\cdot}, \mathcal{Y}_u, \mathcal{Z}_u)\big) - \delta g_u(\cdot, U_u^{\cdot}, V_u^{\cdot}, \mathcal{Y}_u, \mathcal{Z}_u) \right| \mathrm{d}u \right|^2 \right]$$

To conclude, we first note that by our choice of $(\Pi^\ell)_\ell$, $I^\ell(\partial U)$ converges to the Lebesgue integral of ∂U^s. In addition, the uniform continuity of $s \longmapsto \partial_s g(s, x, u, v, y, z)$ and $s \longmapsto \partial_s \eta(s, x)$, see Assumption 3.1.$(i)$, justifies, via bounded convergence, the convergence in $\mathbb{S}^{2,2}$ (resp. $\mathbb{H}^{2,2}$) of $I^\ell(\partial U^s)$ to $U^T - U^s$ (resp. $I^\ell(\partial V^s)$ to $V^T - V^s$) as $\ell \longrightarrow 0$. The result follows in virtue of the uniqueness of (U, V, M). $\qquad \square$

The next lemma identifies the dynamics of $(U_t^t)_{t \in [0,T]}$.

Lemma 6.2. Let $(\mathcal{Y}, \mathcal{Z}) \in \mathbb{L}^2 \times \mathbb{H}^2$ and $(U, \partial U, V, \partial V, M, \partial M) \in (\mathbb{L}^{2,2})^2 \times \overline{\mathbb{H}}^{2,2} \times \mathbb{H}^{2,2} \times (\mathbb{M}^{2,2})^2$ satisfy, for any $s \in [0, T]$ $\mathbb{P}-$a.s. for any $t \in [0, T]$,

$$U_t^s = \eta(s, X) + \int_t^T g_r(s, X, U_r^s, V_r^s, \mathcal{Y}_r, \mathcal{Z}_r) \mathrm{d}r - \int_t^T V_r^{s\top} \mathrm{d}X_r - \int_t^T \mathrm{d}M_r^s,$$

$$\partial U_t^s = \partial_s \eta(s, X) + \int_t^T \nabla g_r(s, X, \partial U_r^s, \partial V_r^s, U_r^s, V_r^s, \mathcal{Y}_r, \mathcal{Z}_r) \mathrm{d}r - \int_t^T \partial V_r^{s\top} \mathrm{d}X_r - \int_t^T \mathrm{d}\partial M_r^s,$$

Then

$$U_t^t = U_T^T + \int_t^T \left(g_r(r, U_r^r, V_r^r, \mathcal{Y}_r, \mathcal{Z}_r) - \partial U_r^r \right) dr - \int_t^T V_r^r dX_r - \int_t^T d\widetilde{M}_r, \ t \in [0, T], \ \mathbb{P}-\text{a.s.}$$

where $\widetilde{M} := (\widetilde{M}_t)_{t \in [0,T]}$ is given by $\widetilde{M}_t := M_t^t - \int_0^t \partial M_r^r dr$. Moreover, $\widetilde{M} \in \mathbb{M}^2$.

Proof. We show that $\mathbb{P}-$a.s., for any $t \in [0, T]$

$$\int_t^T \partial U_r^r dr = U_T^T - U_t^t + \int_t^T g_r(r, U_r^r, V_r^r, \mathcal{Y}_r, \mathcal{Z}_r) dr - \int_t^T V_r^r dX_r - \int_t^T d\widetilde{M}_r.$$

Indeed, note that, $\mathbb{P}-$a.s., for any $t \in [0, T]$

$$\int_t^T \partial U_r^r dr = \int_t^T \partial_s \eta(r) dr + \int_t^T \left(\int_r^T \nabla g_u(r, \partial U_u^r, \partial V_u^r, U_u^r, V_u^r, \mathcal{Y}_u, \mathcal{Z}_u) du \right) dr$$
$$- \int_t^T \left(\int_r^T \partial V_u^{r\top} dX_u + \int_r^T d\partial M_u^r \right) dr.$$

By Fubini's theorem, the change of variables formula for the Lebesgue integral, [44, Theorem 54], and Lemma 6.1 we have that, $\mathbb{P}-$a.s., for any $t \in [0, T]$

$$\int_t^T \int_r^T \nabla g_u(r, \partial U_u^r, \partial V_u^r, U_u^r, V_u^r, \mathcal{Y}_u, \mathcal{Z}_u) du dr$$
$$= \int_t^T \int_t^u \nabla g_u(r, \partial U_u^r, \partial V_u^r, U_u^r, V_u^r, \mathcal{Y}_u, \mathcal{Z}_u) dr du$$
$$= \int_t^T g_u(u, U_u^u, V_u^u, \mathcal{Y}_u, \mathcal{Z}_u) - g_u(t, U_u^t, V_u^t, \mathcal{Y}_u, \mathcal{Z}_u) du.$$

Similarly, given $\partial V \in \mathbb{H}^{2,2}$, the version of Fubini's theorem for stochastic integration, see [44, Theorem 65], yields

$$\int_t^T \int_r^T \partial V_u^{r\top} dX_u dr = \int_t^T \int_t^u \partial V_u^{r\top} dr dX_u = \int_t^T \left(V_u^u - V_u^t \right)^\top dX_u.$$

Now, by Lemma 6.1 we have that, $\mathbb{P}-$a.s., for any $t \in [0, T]$

$$\int_t^T \int_r^T d\partial M_u^r dr = \int_t^T \left(\partial M_T^r - \partial M_r^r \right) dr = \widetilde{M}_T - \widetilde{M}_t + M_t^t - M_T^t.$$

We are left to verify $\widetilde{M} \in \mathbb{M}^2$. Indeed, note that $\widetilde{M}_0 = 0$ and

(i) $\widetilde{M}$ has càdlàg paths. This follows from the fact that, as $\mathbb{F}$ satisfies the usual conditions, there exist a càdlàg modification of $(M_t^t)_{t \in [0,T]}$, which, abusing notations, we still denote by $(M_t^t)_{t \in [0,T]}$. Indeed for $t \in [0, T]$

$$M_t^t = M_t^T - \int_t^T \partial M_t^r dr, \ \mathbb{P}-\text{a.s.}$$

(ii) $\widetilde{M}$ is a martingale. Indeed by Lemma 6.1, for $0 \le u \le t \le T$, $\mathbb{P}-$a.s.

$$\mathbb{E}\left[\widetilde{M}_t \big| \mathcal{F}_u \right] = \mathbb{E}\left[M_t^t \big| \mathcal{F}_u \right] - \int_0^u \partial M_r^r dr - \int_u^t \mathbb{E}\left[\partial M_r^r \big| \mathcal{F}_u \right] dr$$
$$= M_u^t - \int_0^u \partial M_r^r dr - \int_u^t \partial M_u^r dr = \widetilde{M}_u,$$

(iii) $\widetilde{M}$ is orthogonal to X. For $0 \leq u \leq t \leq T$, $\mathbb{P}-$a.s.

$$\mathbb{E}\big[X_t\widetilde{M}_t|\mathcal{F}_u\big] = \mathbb{E}\big[X_tM_t^t|\mathcal{F}_u\big] - \mathbb{E}\Big[X_t\int_0^u \partial M_r^r \mathrm{d}r\Big|\mathcal{F}_u\Big] - \int_u^t \mathbb{E}\big[\mathbb{E}[X_t|\mathcal{F}_r]\partial M_r^r|\mathcal{F}_u\big]\mathrm{d}r$$

$$= X_uM_u^t - X_u\int_0^u \partial M_r^r \mathrm{d}r - X_u\int_u^t \partial M_u^r \mathrm{d}r$$

$$= X_uM_u^t - X_u\int_0^u \partial M_r^r \mathrm{d}r - X_uM_u^t + X_uM_u^u$$

$$= X_u\widetilde{M}_u,$$

where in the second equality we used the fact $\int_0^u \partial M_r^r \mathrm{d}r$ is $\mathcal{F}_u$-measurable, the tower property and the orthogonality of M^s and X and of ∂M^s and X for $s \in [0,T]$. The third equality follows from Lemma 6.1.

(iv) $\|\widetilde{M}\|_{\mathbb{M}^2} < \infty$,

$$\|\widetilde{M}\|_{\mathbb{M}^2}^2 = \mathbb{E}\bigg[\bigg(M_T^T - \int_0^T \partial M_r^r \mathrm{d}r\bigg)^2\bigg] \leq 2\bigg(\mathbb{E}\big[[M^T]_T\big] + T\int_0^T \mathbb{E}\big[[\partial M^r]_r\big]\mathrm{d}r\bigg)$$

$$\leq 2\Big[\|M^T\|_{\mathbb{M}^2}^2 + T^2\|\partial M\|_{\mathbb{M}^{2,2}}^2\Big] < \infty. \qquad \square$$

6.2 *A priori* **estimates**

We now establish *a priori* estimates for $(\mathcal{S})$. To ease the readability, recall that for any $s \in [0,T]$ it holds $\mathbb{P}-$a.s. for any $t \in [0,T]$,

$$\begin{aligned}
\mathcal{Y}_t &= \xi(T,X) + \int_t^T h_r(X,\mathcal{Y}_r,\mathcal{Z}_r,U_r^r,V_r^r,\partial U_r^r)\mathrm{d}r - \int_t^T \mathcal{Z}_r^\top \mathrm{d}X_r - \int_t^T \mathrm{d}\mathcal{N}_r, \\
U_t^s &= \eta(s,X) + \int_t^T g_r(s,X,U_r^s,V_r^s,\mathcal{Y}_r,\mathcal{Z}_r)\mathrm{d}r - \int_t^T V_r^{s\top}\mathrm{d}X_r - \int_t^T \mathrm{d}M_r^s, \qquad (\mathcal{S}) \\
\partial U_t^s &= \partial_s\eta(s,X) + \int_t^T \nabla g_r(s,X,\partial U_r^s,\partial V_r^s,U_r^s,V_r^s,\mathcal{Y}_r,\mathcal{Z}_r)\mathrm{d}r - \int_t^T \partial V_r^{s\top}\mathrm{d}X_r - \int_t^T \mathrm{d}\partial M_r^s.
\end{aligned}$$

Let us introduce $(\mathcal{H}^o, \|\cdot\|_{\mathcal{H}^o})$ and $(\mathfrak{H}^o, \|\cdot\|_{\mathfrak{H}^o})$ where $\|\cdot\|_{\mathcal{H}^o}$ and $\|\cdot\|_{\mathfrak{H}^o}$ denote the norms induced by

$$\mathcal{H}^o := \mathbb{L}^2 \times \mathbb{H}^2 \times \mathbb{M}^2 \times \mathbb{L}^{2,2} \times \overline{\mathbb{H}}^{2,2} \times \mathbb{M}^{2,2} \times \mathbb{L}^{2,2} \times \mathbb{H}^{2,2} \times \mathbb{M}^{2,2},$$

$$\mathfrak{H}^o := \mathbb{L}^2 \times \mathbb{H}^2 \times \mathbb{M}^2 \times \big(\mathbb{L}^{2,2} \times \mathbb{H}^{2,2} \times \mathbb{M}^{2,2}\big)^2.$$

To obtain estimates between the difference of solutions, it is more convenient to work with norms defined by adding exponential weights. We recall, for instance, that for $c \in [0,\infty)$ the norm $\|\cdot\|_{\mathbb{H}^{2,c}}$ is given by

$$\|\mathcal{Z}\|_{\mathbb{H}^{2,c}}^2 = \mathbb{E}\bigg[\int_0^T \mathrm{e}^{ct}|\sigma_t^\top\mathcal{Z}_t|^2\mathrm{d}t\bigg],$$

and they are equivalent for any two values of c, since $[0,T]$ is compact. With this, we define the norm $\|\cdot\|_{\mathcal{H}^{o,c}}$. In the following, we take the customary approach of introducing arbitrary constants $C > 0$ to our analysis. These constants will typically depend on the data of the problem, e.g. the Lipschitz constants and T and on the value of c unless otherwise stated.

Proposition 6.3. Let Assumption 3.1 hold and $\mathfrak{h} \in \mathcal{H}^o$ satisfy $(\mathcal{S})$. Then $(\mathcal{Y}, U, \partial U) \in \mathbb{S}^2 \times \mathbb{S}^{2,2} \times \mathbb{S}^{2,2}$. Furthermore, there exists a constant $C > 0$ such that for $\|\cdot\|_{\mathcal{H}}$ as in Section 3

$$\|\mathfrak{h}\|_{\mathcal{H}}^2 \leq C \underbrace{\left(\|\xi\|_{\mathcal{L}^2}^2 + \|\eta\|_{\mathcal{L}^{2,2}}^2 + \|\partial_s\eta\|_{\mathcal{L}^{2,2}}^2 + \|\tilde{h}\|_{\mathbb{L}^{1,2}}^2 + \|\tilde{g}\|_{\mathbb{L}^{1,2,2}}^2 + \|\nabla\tilde{g}\|_{\mathbb{L}^{1,2,2}}^2 \right)}_{=:I_0^2} < \infty.$$

Proof. We proceed in several steps. We recall that in light of Assumption 3.1, $\mathrm{d}t \otimes \mathrm{d}\mathbb{P}-$a.e.

$$|h_r(\mathcal{Y}_r, \mathcal{Z}_r, U_r^r, V_r^r, \partial U_r^r))| \leq |\tilde{h}| + L_h\big(|\mathcal{Y}_r| + |\sigma_r^\top \mathcal{Z}_r| + |U_r^r| + |\sigma_r^\top V_r^r| + |\partial U_r^r|\big),$$

$$|g_r(s, U_r^s, V_r^s, \mathcal{Y}_r, \mathcal{Z}_r))| \leq |\tilde{g}(s)| + L_g\big(|U_r^s| + |\sigma_r^\top V_r^s| + |\mathcal{Y}_r| + |\sigma_r^\top \mathcal{Z}_r|\big), \quad (6.7)$$

$$|\nabla g_r(s, \partial U_r^s, \partial V_r^s, U_r^s, V_r^s, \mathcal{Y}_r, \mathcal{Z}_r))| \leq |\nabla\tilde{g}_r(s)| + L_{\partial_s g}\big(|U_r^s| + |\sigma_r^\top V_r^s| + |\mathcal{Y}_r| + |\sigma_r^\top \mathcal{Z}_r|\big)$$
$$+ L_g\big(|\partial U_r^s| + |\sigma_r^\top \partial V_r^s|\big).$$

Step 1: we start with auxiliary estimates. By Meyer–Itô's formula for $\mathrm{e}^{\frac{c}{2}t}|\partial U_t^s|$, see Protter [44, Theorem 70], for any $t \in [0, T]$,

$$\mathrm{e}^{\frac{c}{2}t}|\partial U_t^s| + L_T^0 - \int_t^T \mathrm{e}^{\frac{c}{2}r}\operatorname{sgn}(\partial U_r^s)\cdot\partial V_r^{s\top}\mathrm{d}X_r - \int_t^T \mathrm{e}^{\frac{c}{2}r-}\operatorname{sgn}(\partial U_{r-}^s)\cdot\mathrm{d}\partial M_r^s$$
$$= \mathrm{e}^{\frac{c}{2}T}|\partial_s\eta(s)| + \int_t^T \mathrm{e}^{\frac{c}{2}r}\bigg(\operatorname{sgn}(\partial U_r^s)\cdot\nabla g_r(s, \partial U_r^s, \partial V_r^s, U_r^s, V_r^s, \mathcal{Y}_r, \mathcal{Z}_r) - \frac{c}{2}|\partial U_r^s| \bigg)\mathrm{d}r, \tag{6.8}$$

where $L^0 := L^0(\partial U^s)$ denotes the non-decreasing and pathwise-continuous local time of the semi-martingale ∂U^s at 0, see [44, Chapter IV, pp. 216]. We also notice that for any $s \in [0, T]$ the last two terms on the left-hand side are martingales, recall that $\partial V^s \in \mathbb{H}^2$.

We now take conditional expectation with respect to $\mathcal{F}_t$ in Equation (6.8). We may use (6.7) and the fact L^0 is non-decreasing to derive that for $t \in [0, T]$ and $c > 2L_g$

$$\mathrm{e}^{\frac{c}{2}t}|\partial U_t^s|$$
$$\leq \mathbb{E}_t\bigg[\mathrm{e}^{\frac{c}{2}T}|\partial_s\eta(s)| + \int_t^T \mathrm{e}^{\frac{c}{2}r}\Big(|\partial U_r^s|(L_g - c/2) + |\nabla\tilde{g}_r(s)| + L_g|\sigma_r^\top\partial V_r^s|\Big)\mathrm{d}r$$
$$+ \int_t^T \mathrm{e}^{\frac{c}{2}r}L_{\partial_s g}\big(|U_r^s| + |\sigma_r^\top V_r^s| + |\mathcal{Y}_r| + |\sigma_r^\top \mathcal{Z}_r|\big)\mathrm{d}r\bigg]$$
$$\leq \mathbb{E}_t\bigg[\mathrm{e}^{\frac{c}{2}T}|\partial_s\eta(s)|$$
$$+ \int_t^T \mathrm{e}^{\frac{c}{2}r}\Big(|\nabla\tilde{g}_r(s)| + \bar{L}\big(|\sigma_r^\top\partial V_r^s| + |U_r^s| + |\sigma_r^\top V_r^s| + |\mathcal{Y}_r| + |\sigma_r^\top \mathcal{Z}_r|\big)\Big)\mathrm{d}r\bigg]. \tag{6.9}$$

where $\bar{L} := \max\{L_g, L_{\partial_s g}\}$. Squaring in (6.9), we may use (6.1) and Jensen's inequality to derive for $t \in [0, T]$

$$\frac{\mathrm{e}^{ct}}{7}|\partial U_t^t|^2 \leq \mathbb{E}_t\bigg[\mathrm{e}^{cT}|\partial_s\eta(t)|^2 + \bigg(\int_t^T \mathrm{e}^{\frac{c}{2}r}|\nabla\tilde{g}_r(t)|\mathrm{d}r\bigg)^2\bigg]$$
$$+ \mathbb{E}_t\bigg[T\bar{L}^2\int_t^T \mathrm{e}^{cr}\big(|U_r^t|^2 + |\mathcal{Y}_r|^2 + |\sigma_r^\top\partial V_r^t|^2 + |\sigma_r^\top V_r^t|^2 + |\sigma_r^\top \mathcal{Z}_r|^2\big)\mathrm{d}r\bigg].$$

Integrating the previous expression and taking expectation, it follows from the tower

property that for any $t \in [0, T]$

$$\frac{1}{7} \mathbb{E}\bigg[\int_t^T e^{cr} |\partial U_r^r|^2 \mathrm{d}r \bigg]$$

$$\leq \mathbb{E}\bigg[\int_t^T e^{cT} |\partial_s \eta(r)|^2 \mathrm{d}r \bigg] + \mathbb{E}\bigg[\int_t^T \bigg(\int_r^T e^{\frac{c}{2}u} |\nabla \tilde{g}_u(r)| \mathrm{d}u \bigg)^2 \mathrm{d}r \bigg]$$

$$+ T \bar{L}^2 \mathbb{E}\bigg[\int_t^T \int_r^T e^{cu} \big(|U_u^r|^2 + |\mathcal{Y}_u|^2 + |\sigma_u^\top \partial V_u^r|^2 + |\sigma_u^\top V_u^r|^2 + |\sigma_u^\top \mathcal{Z}_u|^2 \big) \mathrm{d}u \mathrm{d}r \bigg]$$

$$\leq T \sup_{r \in [0,T]} \bigg\{ \|e^{cT} |\partial_s \eta(r)|^2]\|_{\mathcal{L}^2} + \mathbb{E}\bigg[\bigg(\int_t^T e^{\frac{c}{2}u} |\nabla \tilde{g}_u(r)| \mathrm{d}u \bigg)^2 \bigg] \bigg\}$$

$$+ T^2 \bar{L}^2 \sup_{r \in [0,T]} \bigg\{ \mathbb{E}\bigg[\int_t^T e^{cu} \big(|U_u^r|^2 + |\sigma_u^\top V_u^r|^2 + |\sigma_u^\top \partial V_u^r|^2 \big) \mathrm{d}u \bigg] \bigg\}$$

$$+ T^2 \bar{L}^2 \mathbb{E}\bigg[\int_t^T e^{cu} \big(|\mathcal{Y}_u|^2 + |\sigma_u^\top \mathcal{Z}_u|^2 \big) \mathrm{d}u \bigg].$$

Thus, we obtain there is $C_{\partial u} > 0$ such that for any $c > 2L_g$ and $t \in [0, T]$

$$\frac{1}{C_{\partial u}} \mathbb{E}\bigg[\int_t^T e^{cr} |\partial U_r^r|^2 \mathrm{d}r \bigg]$$

$$\leq e^{cT} \big(\|\partial_s \eta\|_{\mathcal{L}^{2,2}}^2 + \|\nabla \tilde{g}\|_{\mathbb{L}^{1,2,2}}^2 \big) + \mathbb{E}\bigg[\int_t^T e^{cr} \big(|\mathcal{Y}_r|^2 + |\sigma_r^\top \mathcal{Z}_r|^2 \big) \mathrm{d}r \bigg] \qquad (6.10)$$

$$+ \sup_{s \in [0,T]} \mathbb{E}\bigg[\int_t^T e^{cr} \big(|U_r^s|^2 + |\sigma_r^\top V_r^s|^2 + |\sigma_r^\top \partial V_r^s|^2 \big) \mathrm{d}r \bigg].$$

Similarly, we may find $C_u > 0$ such that for any $c > 2L_g$ and $t \in [0, T]$

$$\frac{1}{C_u} \mathbb{E}\bigg[\int_t^T e^{cr} |U_r^r|^2 \mathrm{d}r \bigg] \leq e^{cT} \big(\|\eta\|_{\mathcal{L}^{2,2}}^2 + \|\tilde{g}\|_{\mathbb{L}^{1,2,2}}^2 \big) + \mathbb{E}\bigg[\int_t^T e^{cr} \big(|\mathcal{Y}_r|^2 + |\sigma_r^\top \mathcal{Z}_r|^2 \big) \mathrm{d}r \bigg]$$

$$+ \sup_{s \in [0,T]} \mathbb{E}\bigg[\int_t^T e^{cr} |\sigma_r^\top V_r^s|^2 \mathrm{d}r \bigg]. \qquad (6.11)$$

We now estimate the term V_t^t. In light of (6.3) and (6.5), with $e^{\frac{c}{2}t} \sigma_t^\top$ instead of $\sigma_t^\top$, there exists $C > 0$ such that for any $\varepsilon > 0$, $c > 4L_g$ and $t \in [0, T]$

$$\mathbb{E}\bigg[\int_t^T e^{cu} |\sigma_u^\top V_u^u|^2 \mathrm{d}u \bigg]$$

$$\leq \mathbb{E}\bigg[\int_t^T e^{cu} |\sigma_u^\top V_u^t|^2 \mathrm{d}u \bigg] + \int_t^T \mathbb{E}\bigg[\int_r^T e^{cu} \big(\varepsilon |\sigma_u^\top V_u^r|^2 + \varepsilon^{-1} |\sigma_u^\top \partial V_u^r|^2 \big) \mathrm{d}u \bigg] \mathrm{d}r$$

$$\leq \sup_{r \in [0,T]} \mathbb{E}\bigg[\int_t^T e^{cu} |\sigma_u^\top V_u^r|^2 \mathrm{d}u \bigg] + T \sup_{r \in [0,T]} \mathbb{E}\bigg[\int_t^T e^{cu} \big(\varepsilon |\sigma_u^\top V_u^r|^2 + \varepsilon^{-1} |\sigma_u^\top \partial V_u^r|^2 \big) \mathrm{d}u \bigg]$$

$$\leq (1 + \varepsilon T) \sup_{r \in [0,T]} \mathbb{E}\bigg[\int_t^T e^{cu} |\sigma_u^\top V_u^r|^2 \mathrm{d}u \bigg] + TC\varepsilon^{-1} \sup_{r \in [0,T]} \mathbb{E}\bigg[\int_t^T e^{cu} \big(|U_u^r|^2 + |\sigma_u^\top V_u^r|^2 \big) \mathrm{d}u \bigg]$$

$$+ TC\varepsilon^{-1} \mathbb{E}\bigg[\int_t^T e^{cu} \big(|\mathcal{Y}_u|^2 + |\sigma_u^\top \mathcal{Z}_u|^2 \big) \mathrm{d}u \bigg]$$

$$+ TC\varepsilon^{-1} e^{cT} \sup_{r \in [0,T]} \mathbb{E}\bigg[|\partial_s \eta(r)|^2 + \bigg(\int_t^T |\nabla \tilde{g}_u(r)| \mathrm{d}u \bigg)^2 \bigg].$$

Thus, taking $\varepsilon = TC$ we may find $C_v > 0$ such that for any $c > 4L_g$ and $t \in [0,T]$

$$\frac{1}{C_v}\mathbb{E}\bigg[\int_t^T e^{cu}|\sigma_r^\top V_r^r|^2 \mathrm{d}r\bigg]$$

$$\leq e^{cT}\big(\|\partial_s\eta\|_{\mathcal{L}^{2,2}}^2 + \|\nabla\tilde{g}\|_{\mathbb{L}^{1,2,2}}^2\big) + \mathbb{E}\bigg[\int_t^T e^{cr}\big(|\mathcal{Y}_r|^2 + |\sigma_r^\top \mathcal{Z}_r|^2\big)\mathrm{d}r\bigg] \qquad (6.12)$$

$$+ \sup_{s\in[0,T]}\mathbb{E}\bigg[\int_t^T e^{cr}\big(|U_r^s|^2 + |\sigma_r^\top V_r^s|^2\big)\mathrm{d}r\bigg].$$

We emphasise that the constants $(C_{\partial u}, C_u, C_v) \in (0,\infty)^3$ depend only of the data of the problem and are universal for any value of $c > 4L_g$.

Step 2: Let $s \in [0,T]$, we show that $(\mathcal{Y}, U, \partial U) \in \mathbb{S}^2 \times \mathbb{S}^{2,2} \times \mathbb{S}^{2,2}$. To alleviate the notation we introduce

$$\mathfrak{Y} := (\mathcal{Y}, U^s, \partial U^s), \ \mathfrak{Z} := (\mathcal{Z}, V^s, \partial V^s), \ \mathfrak{N} := (\mathcal{N}, M^s, \partial M^s),$$

whose elements we may denote with superscripts, e.g. $\mathcal{Y}^1, \mathcal{Y}^2, \mathcal{Y}^3$ correspond to $\mathcal{Y}$, U^s, ∂U^s.

By (6.1) and (6.7), we obtain that there exists $C > 0$, which may change value from line to line, such that

$$|U_t^s|^2 \leq C\bigg(|\eta(s)|^2 + \bigg(\int_0^T |\tilde{g}_r(s)|\mathrm{d}r\bigg)^2 + \int_0^T \big(|U_r^s|^2 + |\sigma_r^\top V_r^s|^2 + |\mathcal{Y}_r|^2 + |\sigma_r^\top \mathcal{Z}_r|^2\big)\mathrm{d}r$$

$$+ \bigg|\int_t^T V_r^{s\top}\mathrm{d}X_r + \int_t^T \mathrm{d}M_r^s\bigg|^2\bigg).$$

We note that by Doob's inequality

$$\mathbb{E}\bigg[\sup_{t\in[0,T]}\bigg|\int_0^t V_r^{s\top}\mathrm{d}X_r\bigg|^2\bigg] \leq 4\|V^s\|_{\mathbb{H}^2}^2. \qquad (6.13)$$

Taking supremum over $t \in [0,T]$ and expectation we obtain for $s \in [0,T]$

$$\|U^s\|_{\mathbb{S}^2}^2 \leq C\big(\|\eta(s)\|_{\mathcal{L}^2}^2 + \|\tilde{g}(s)\|_{\mathbb{L}^{1,2}}^2 \qquad (6.14)$$
$$+ \|U^s\|_{\mathbb{L}^2}^2 + \|V^s\|_{\mathbb{H}^2}^2 + \|\mathcal{Y}\|_{\mathbb{L}^2}^2 + \|\mathcal{Z}\|_{\mathbb{H}^2}^2 + \|M^s\|_{\mathbb{M}^2}^2\big) < \infty.$$

Similarly, we obtain that there exists $C > 0$ such that for $s \in [0,T]$

$$\|\partial U^s\|_{\mathbb{S}^2}^2 \leq C\bigg(\|\partial_s\eta(s)\|_{\mathcal{L}^2}^2 + \|\nabla\tilde{g}(s)\|_{\mathbb{L}^{1,2}}^2 \qquad (6.15)$$
$$+ \sum_{i=1}^3 \|\mathfrak{Y}^i\|_{\mathbb{L}^2}^2 + \|\mathfrak{Z}^i\|_{\mathbb{H}^2}^2 + \|\partial M^s\|_{\mathbb{M}^2}^2\bigg) < \infty.$$

Given $(\eta, \partial_s\eta, \tilde{g}, \partial\tilde{g}) \in (\mathcal{L}^{2,2})^2 \times (\mathbb{L}^{1,2,2})^2$, $(U, \partial U, V, \partial V, M, \partial M) \in (\mathbb{L}^{2,2})^2 \times (\mathbb{H}^{2,2})^2 \times (\mathbb{M}^{2,2})^2$, (6.14) and (6.15), the mapping

$$([0,T], \mathcal{B}([0,T])) \longrightarrow (\mathbb{S}^2, \|\cdot\|_{\mathbb{S}^2}) : s \longmapsto \mathfrak{Y}^{is} \text{ is continuous for } i \in \{2,3\},$$

and $\|\mathfrak{Y}^i\|_{\mathbb{S}^{2,2}} < \infty$. Consequently, $\mathfrak{Y}^i \in \mathbb{S}^{2,2}$ for $i \in \{2,3\}$. Arguing as above we may also derive,

$$|\mathcal{Y}_t|^2 \leq C\bigg(|\xi|^2 + \bigg(\int_0^T |\tilde{h}_r|\mathrm{d}r\bigg)^2 + \int_0^T \big(|\mathcal{Y}_r|^2 + |\sigma_r^\top \mathcal{Z}_r|^2 + |U_r^r|^2 + |\sigma_r^\top V_r^r|^2 + |\partial U_r^r|^2\big)\mathrm{d}r$$

$$+ \bigg|\int_t^T \mathcal{Z}_r^\top\mathrm{d}X_r\bigg|^2 + \bigg|\int_t^T \mathrm{d}N_r\bigg|^2\bigg),$$

which in turn yields, in combination with (6.10), (6.11) and (6.12),

$$\|\mathcal{Y}\|_{\mathbb{S}^2}^2 \leq C\big(I_0^2 + \|\mathcal{Y}\|_{\mathbb{L}^2}^2 + \|\mathcal{Z}\|_{\mathbb{H}^2}^2 + \|U\|_{\mathbb{L}^{2,2}}^2 + \|V\|_{\mathbb{H}^{2,2}}^2 + \|\partial V\|_{\mathbb{H}^{2,2}}^2 + \|N\|_{\mathbb{M}^2}^2\big) < \infty. \quad (6.16)$$

Finally, taking sup over $s \in [0, T]$ and adding together (6.16) (6.14) and (6.15) we obtain

$$\|\mathcal{Y}\|_{\mathbb{S}^2}^2 + \|U^s\|_{\mathbb{S}^{2,2}}^2 + \|\partial U\|_{\mathbb{S}^{2,2}}^2 \leq C\big(I_0^2 + \|(\mathcal{Y}, \mathcal{Z}, \mathcal{N}, U, V, M, \partial U, \partial V, \partial M)\|_{\mathcal{H}^o}\big). \quad (6.17)$$

Step 3: We obtain the estimate of the norm. To ease the notation we introduce

$$h_r := h_r(\mathcal{Y}_r, \mathcal{Z}_r, U_r^r, V_r^r, \partial U_r^r), \; g_r(s) := g_r(s, U_r^s, V_r^s, \mathcal{Y}_r, \mathcal{Z}_r),$$
$$\nabla g_r(s) := \nabla g_r(s, \partial U_r^s, \partial V_r^s, U_r^s, V_r^s, \mathcal{Y}_r, \mathcal{Z}_r).$$

By applying Itô's formula to $\mathrm{e}^{ct}\big(|\mathcal{Y}_t|^2 + |U_t^s|^2 + |\partial U_t^s|^2\big)$ we obtain, $\mathbb{P}-$a.s.

$$\sum_{i=1}^3 \mathrm{e}^{ct}|\mathfrak{Y}_t^i|^2 + \int_t^T \mathrm{e}^{cr}|\sigma_r^\top \mathfrak{Z}_r^i|^2 \mathrm{d}r + \int_t^T \mathrm{e}^{cr-}\mathrm{d}\big[\mathfrak{N}^i\big]_r + \mathfrak{M}_t - \mathfrak{M}_T$$

$$= \mathrm{e}^{cT}(|\xi(T)|^2 + |\eta(s)|^2 + |\partial_s \eta(s)|^2)$$

$$+ \int_t^T \mathrm{e}^{cr}\big(2\mathcal{Y}_r \cdot h_r + 2U_r^s \cdot g_r(s) + 2\partial U_r^s \cdot \nabla g_r(s) - c(|\mathcal{Y}_r|^2 + |U_r^s|^2 + |\partial U_r^s|^2)\big)\mathrm{d}r,$$

where we used the orthogonality of X and both M^s and N, and we introduced the martingale

$$\mathfrak{M}_t := 2\sum_{i=1}^3 \int_0^t \mathrm{e}^{cr}\mathfrak{Y}_r^i \cdot \mathfrak{Z}_r^{i\top} \mathrm{d}X_r + \int_0^t \mathrm{e}^{cr-}\mathfrak{Y}_{r-}^i \cdot \mathrm{d}\,\mathfrak{N}_r^i.$$

Indeed, the Burkholder–Davis–Gundy inequality in combination with the fact that $(\mathcal{Y}, U^s, \partial U^s) \in (\mathbb{S}^2)^3$ shows that $\mathfrak{M}$ is uniformly integrable, and consequently a true martingale. Moreover, we insist on the fact that the integrals with respect to N, M^s and ∂M^s account for possible jumps, see [32, Lemma 4.24].

Moreover, as $(\mathcal{Y}, U, \partial U) \in \mathbb{S}^2 \times (\mathbb{S}^{2,2})^2$, from (6.7) and with Young's inequality we obtain that for any $(\varepsilon, \tilde{\varepsilon}) \in (0, \infty)^2$, there is $C(\tilde{\varepsilon}) \in (0, \infty)$ such that the left-hand side above is smaller than

$$\leq \mathrm{e}^{cT}\big(|\xi|^2 + |\eta(s)|^2 + |\partial_s \eta(s)|^2\big) + \int_t^T \mathrm{e}^{cr}(C(\tilde{\varepsilon}) - c)\big(|\mathcal{Y}_r|^2 + |U_r^s|^2 + |\partial U_r^s|^2\big)\mathrm{d}r$$

$$+ \int_t^T \tilde{\varepsilon}\mathrm{e}^{cr}\left(|\sigma_r^\top \mathcal{Z}_r|^2 + |\sigma_r^\top V_r^s|^2 + |\sigma_r^\top \partial V_r^s|^2 + \frac{1}{C_u}|U_r^r|^2 + \frac{1}{C_v}|V_r^r|^2 + \frac{1}{C_{\partial u}}|\partial U_r^r|^2\right)\mathrm{d}r$$

$$+ \varepsilon\sum_{i=1}^3 \sup_{r \in [0,T]} \mathrm{e}^{cr}|\mathfrak{Y}_t^i|^2 + \frac{1}{\varepsilon}\left|\int_0^T \mathrm{e}^{\frac{c}{2}r}|\tilde{h}_r|\mathrm{d}r\right|^2 + \frac{1}{\varepsilon}\left|\int_0^T \mathrm{e}^{\frac{c}{2}r}|\tilde{g}_r(s)|\mathrm{d}r\right|^2 + \frac{1}{\varepsilon}\left|\int_0^T \mathrm{e}^{\frac{c}{2}r}|\nabla \tilde{g}_r(s)|\mathrm{d}r\right|^2$$

with $(C_{\partial u}, C_u, C_v)$ as in (6.10)–(6.12). Taking expectation and letting $c > 4L_g$, we find there is $C > 0$ such that

$$\mathbb{E}\left[\sum_{i=1}^3 \mathrm{e}^{ct}|\mathfrak{Y}_t^i|^2 + \int_t^T \mathrm{e}^{cr}|\sigma_r^\top \mathfrak{Z}_r^i|^2 \mathrm{d}r + \int_t^T \mathrm{e}^{cr-}\mathrm{d}\big[\mathfrak{N}^i\big]_r\right]$$

$$\leq \sup_{s \in [0,T]} \mathbb{E}\left[\int_t^T \mathrm{e}^{cr}(|\mathcal{Y}_r|^2 + |U_r^s|^2)(C(\tilde{\varepsilon}) - c)\mathrm{d}r\right] + \sup_{s \in [0,T]} \mathbb{E}\left[\int_t^T \mathrm{e}^{cr}|\partial U_r^s|^2(C(\tilde{\varepsilon}) - c)\mathrm{d}r\right]$$

$$+ (1 + \varepsilon^{-1} + \tilde{\varepsilon})CI_0^2 + \tilde{\varepsilon}C\big(\|\mathcal{Z}\|_{\mathbb{H}^{2,c}}^2 + \|V\|_{\mathbb{H}^{2,2,c}}^2 + \|\partial V\|_{\mathbb{H}^{2,2,c}}^2\big)$$

$$+ \varepsilon C\big(\|\mathcal{Y}\|_{\mathbb{S}^{2,c}}^2 + \|U\|_{\mathbb{S}^{2,2,c}}^2 + \|\partial U\|_{\mathbb{S}^{2,2,c}}^2\big).$$

We then let $\tilde{\varepsilon} = 1/(2^4 C)$, $c \geq \max\{4L_g, C(\tilde{\varepsilon})\}$, and take sup over $t \in [0, T]$ (resp. $(s, t) \in [0, T]^2$) to each term on the left side separately. Adding these terms up we find there is $C > 0$, such that for any $\varepsilon > 0$

$$
\begin{aligned}
&\frac{1}{T}\|(\mathcal{Y}, \mathcal{Z}, \mathcal{N}, U, V, M, \partial U, \partial V, \partial M)\|_{\mathfrak{H}^\circ} \\
&\leq \sup_{t \in [0, T]} \mathbb{E}\big[e^{ct}|\mathcal{Y}_t|^2\big] + \sup_{(s,t) \in [0,T]^2} \mathbb{E}\big[e^{ct}|U_t^s|^2\big] + \sup_{(s,t) \in [0,T]^2} \mathbb{E}\big[e^{ct}|\partial U_t^s|^2\big] \\
&\quad + \|\mathcal{Z}\|_{\mathbb{H}^2}^2 + \|V\|_{\mathbb{H}^{2,2}}^2 + \|\partial V\|_{\mathbb{H}^{2,2}}^2 + \|\mathcal{N}\|_{\mathbb{M}^2}^2 + \|M\|_{\mathbb{M}^{2,2}}^2 + \|\partial M\|_{\mathbb{M}^{2,2}}^2 \\
&\leq (1 + \varepsilon^{-1}) C I_0^2 + \varepsilon C\big(\|\mathcal{Y}\|_{\mathbb{S}^2}^2 + \|U\|_{\mathbb{S}^{2,2}}^2 + \|\partial U\|_{\mathbb{S}^{2,2}}^2\big).
\end{aligned}
\tag{6.18}
$$

We can the use (6.18) back in (6.17) to find $\varepsilon \in (0, \infty)$ small enough so that[6]

$$
\|\mathfrak{h}\|_{\mathfrak{H}}^2 \leq C I_0^2.
$$

The result in terms of the norm $\|\cdot\|_{\mathcal{H}}$ follows from (6.12). $\qquad \square$

Proposition 6.4. Let $(\xi^i, \eta^i, \partial_s \eta^i) \in \mathcal{L}^2 \times (\mathcal{L}^{2,2})^2$ and $(h^i, g^i, \partial_s g^i)$ for $i \in \{1, 2\}$ satisfy Assumption 3.1 and suppose in addition that $\mathfrak{H}^i \in \mathcal{H}^\circ$ is a solution to $(\mathcal{S})$ with coefficients $(\xi^i, h^i, \eta^i, g^i, \partial_s \eta^i, \nabla g^i)$, $i \in \{1, 2\}$. Then

$$
\|\delta \mathfrak{H}\|_{\mathcal{H}}^2 \leq C\Big(\|\delta \xi\|_{\mathcal{L}^2}^2 + \|\delta \eta\|_{\mathcal{L}^{2,2}}^2 + \|\delta \partial \eta\|_{\mathcal{L}^{2,2}}^2 + \|\delta_1 h\|_{\mathbb{L}^{1,2}}^2 + \|\delta_1 g\|_{\mathbb{L}^{1,2,2}}^2 + \|\delta_1 \nabla g\|_{\mathbb{L}^{1,2,2}}\Big),
$$

where for $\varphi \in \{\mathcal{Y}, \mathcal{Z}, \mathcal{N}, U, V, M, \partial U, \partial V, \partial M, \xi, \eta, \partial_s \eta\}$ and $\Phi \in \{h, g, \nabla g\}$

$$
\delta \varphi := \varphi^1 - \varphi^2, \text{ and, } \delta_1 \Phi_t := \Phi_t^1(\mathcal{Y}_r^1, \mathcal{Z}_t^1, U_t^{1t}, V_t^{1t}) - \Phi_t^2(\mathcal{Y}_r^1, \mathcal{Z}_t^1, U_t^{1t}, V_t^{1t}), \text{ } dt \otimes d\mathbb{P}-\text{a.e.}
$$

Proof. Note that by the Lipschitz assumption on h and g there exist bounded processes with appropriate dimensions $(\alpha^i, \beta^i, \gamma^i, \varepsilon^i)$, $i \in \{1, 2, 3\}$, ρ and ϱ such that for every $s \in [0, T]$, $\mathbb{P}-$a.s., $t \in [0, T]$

$$
\begin{aligned}
\delta \mathcal{Y}_t = &\int_t^T \big(\delta_1 h_r + \gamma_r^1 \delta \mathcal{Y}_r + \alpha_r^{1\top} \sigma_r^\top \delta \mathcal{Z}_r + \beta_r^1 \delta U_r^r + \varepsilon_r^{1\top} \sigma_r^\top \delta V_r^r\big) dr \\
&+ \delta \xi(T) - \int_t^T \delta \mathcal{Z}_r^\top dX_r - \int_t^T d\delta N_r,
\end{aligned}
$$

$$
\begin{aligned}
\delta U_t^s = &\int_t^T \big(\delta_1 g_r(s) + \beta_r^2 \delta U_r^s + \varepsilon_r^{2\top} \sigma_r^\top \delta V_r^s + \gamma_r^2 \delta \mathcal{Y}_r + \alpha_r^{2\top} \sigma_r^\top \delta \mathcal{Z}_r\big) dr \\
&+ \delta \eta(s) - \int_t^T \delta V_r^{s\top} dX_r - \int_t^T d\delta M_r^s,
\end{aligned}
$$

$$
\begin{aligned}
\delta \partial U_t^s = &\int_t^T \big(\delta_1 \nabla g_r(s) + \rho_r \delta \partial U_r^s + \varrho_r \delta \partial V_r^s + \beta_r^3 \delta U_r^s + \varepsilon_r^{3\top} \sigma_r^\top \delta V_r^s + \gamma_r^3 \delta \mathcal{Y}_r + \alpha_r^{3\top} \sigma_r^\top \delta \mathcal{Z}_r\big) dr \\
&+ \delta \partial_s \eta(s) - \int_t^T \delta \partial V_r^{s\top} dX_r - \int_t^T d\delta \partial M_r^s.
\end{aligned}
$$

We can therefore apply Proposition 6.3 and the result follows. $\qquad \square$

When the data of the system is chosen so as to study the class of type-I BSVIEs considered in Section 4, our approach can be specialised so as to enlarge the initial space and simplify the *a priori* estimates obtained in Proposition 6.3. We let $(\mathcal{H}^{\star, \circ}, \|\cdot\|_{\mathcal{H}^{\star, \circ}})$ be

$$
\mathcal{H}^{\star, \circ} := \mathbb{L}^2 \times \mathbb{H}^2 \times \mathbb{M}^2 \times \mathbb{L}^{2,2} \times \mathbb{H}^{2,2} \times \mathbb{M}^{2,2},
$$

$$
\|(\mathcal{Y}, \mathcal{Z}, \mathcal{N}, Y, Z, N)\|_{\mathcal{H}^{\star, \circ}}^2 := \|\mathcal{Y}\|_{\mathbb{L}^2}^2 + \|\mathcal{Z}\|_{\mathbb{H}^2}^2 + \|\mathcal{N}\|_{\mathbb{M}^2}^2 + \|Y\|_{\mathbb{L}^{2,2}}^2 + \|Z\|_{\mathbb{H}^{2,2}}^2 + \|N\|_{\mathbb{M}^{2,2}}^2.
$$

---

[6]Recall the norm $\|\cdot\|_{\mathfrak{H}}$ is the norm induced by the space $\mathfrak{H}$ as defined in Remark 3.4.(ii)

Proposition 6.5. Let Assumption 4.1 hold and consider $(\mathcal{Y}, \mathcal{Z}, \mathcal{N}, Y, Z, N) \in \mathcal{H}^{\star,o}$ solution, which holds for any $s \in [0, T]$ $\mathbb{P}-$a.s. for any $t \in [0, T]$, to

$$
\begin{aligned}
\mathcal{Y}_t &= \xi(T, X) + \int_t^T \big(f_r(r, X, \mathcal{Y}_r, \mathcal{Z}_r, Y_r^r) - \partial Y_r^r \big) \mathrm{d}r - \int_t^T \mathcal{Z}_r^\top X_r - \int_t^T \mathcal{N}_r, \\
Y_t^s &= \xi(s, X) + \int_t^T f_r(s, X, Y_r^s, Z_r^s, \mathcal{Y}_r) \mathrm{d}r - \int_t^T Z_r^{s\top} \mathrm{d}X_r - \int_t^T \mathrm{d}N_r^s,
\end{aligned}
\tag{$\mathcal{S}_f^o$}
$$

with ∂Y given as in Lemma 6.1. Then $(\mathcal{Y}, Y) \in \mathbb{S}^2 \times \mathbb{S}^{2,2}$ and there exist $C > 0$ such that

$$
\|(\mathcal{Y}, \mathcal{Z}, \mathcal{N}, Y, Z, N)\|_{\mathcal{H}^\star}^2 \leq C\Big(\|\xi\|_{\mathcal{L}^{2,2}}^2 + \|\tilde{f}\|_{\mathbb{L}^{1,2,2}}^2 \Big) < \infty.
$$

Proof. We first note that $Y \in \mathbb{S}^{2,2}$ follows as in Proposition 6.3. Thus, in light of Lemma 6.1, there exists $(\partial Y, \partial Z, \partial N) \in \mathbb{S}^{2,2} \times \mathbb{H}^{2,2} \times \mathbb{M}^{2,2}$ solution to the BSDE with data $(\partial_s \xi, \partial f)$, and, $\big((Y_t^t)_{t \in [0,T]}, (Z_t^t)_{t \in [0,T]} \big) \in \mathbb{S}^2 \times \mathbb{H}^2$ are well-defined. Moreover, ∂Y_r^r is well-defined as an element of $\mathbb{L}^{1,2}$, i.e. $\mathrm{d}t \otimes \mathrm{d}\mathbb{P}-$a.e. on $[0, T] \times \mathcal{X}$, as a consequence of the path-wise continuity of ∂Y^s. With this, we conclude $\mathcal{Y} \in \mathbb{S}^2$.

Let us now note that given $(Y_t^t)_{t \in [0,T]}, (Z_t^t)_{t \in [0,T]}$ and $(\partial Y_t^t)_{t \in [0,T]}$, the first equation, being a Lipschitz BSDE, admits a unique solution $(\mathcal{Y}, \mathcal{Z}, \mathcal{N})$. In addition, for $\widetilde{N}$ as in Lemma 6.2 we obtain

$$
Y_t^t = Y_T^T + \int_t^T \big(f_r(r, Y_r^r, Z_r^r, \mathcal{Y}_r) - \partial Y_r^r \big) \mathrm{d}r - \int_t^T Z_r^{r\top} \mathrm{d}X_r - \int_t^T \mathrm{d}\widetilde{N}_r, \ t \in [0, T], \ \mathbb{P}-\text{a.s.}
\tag{6.19}
$$

Thus, $\big(\mathcal{Y}_\cdot, \mathcal{Z}_\cdot, \mathcal{N}_\cdot \big) = \big((Y_t^t)_{t \in [0,T]}, (Z_t^t)_{t \in [0,T]}, (\widetilde{N}_t)_{t \in [0,T]} \big)$ in $\mathbb{S}^2 \times \mathbb{H}^2 \times \mathbb{M}^2$ and we obtain

$$
\mathcal{Y}_t = \xi(t) + \int_t^T f_r(t, Y_r^t, Z_r^t, \mathcal{Y}_r) \mathrm{d}r - \int_t^T Z_r^{t\top} \mathrm{d}X_r - \int_t^T \mathrm{d}N_r^t, \ t \in [0, T], \ \mathbb{P}-\text{a.s.}
\tag{6.20}
$$

With this equation we can simplify our estimates. Let us introduce the system

$$
\begin{aligned}
\mathcal{Y}_t &= \xi(t) + \int_t^T f_r(t, Y_r^t, Z_r^t, \mathcal{Y}_r) \mathrm{d}r - \int_t^T Z_r^{t\top} \mathrm{d}X_r - \int_t^T \mathrm{d}N_r^t, \\
Y_t^s &= \xi(s) + \int_t^T f_r(s, Y_r^s, Z_r^s, \mathcal{Y}_r) \mathrm{d}r - \int_t^T Z_r^{s\top} \mathrm{d}X_r - \int_t^T \mathrm{d}N_r^s, \ t \in [0, T].
\end{aligned}
\tag{$\mathcal{A}$}
$$

Then, following the same reasoning of Proposition 6.3, i.e. applying Itô's formula to $\mathrm{e}^{ct} \big(|\mathcal{Y}_t| + |Y_t^s| \big)$ in combination with Young's inequality, we obtain there is $C > 0$ such that $\|(\mathcal{Y}, \mathcal{Z}, \mathcal{N}, Y, Z, N)\|_{\mathcal{H}^\star}^2 \leq C\big(\|\xi\|_{\mathcal{L}^{2,2}}^2 + \|\tilde{f}\|_{\mathbb{L}^{1,2,2}}^2 \big) < \infty.$ $\qquad \square$

6.3 Well-posedness

Before we present the proof of Theorem 3.6 we recall that in light of Proposition 6.3 and Proposition 6.4 once the result is obtained for $\mathcal{H}^o$ the existence of a unique solution in $\mathcal{H}$ follows immediately.

Proof of Theorem 3.6. Note that uniqueness follows from Proposition 6.4. To show existence, let us define the map

$$
\mathfrak{T} : \mathcal{H}^o \longrightarrow \mathcal{H}^o
$$
$$
(y, z, n, u, v, m, \mathrm{u}, \mathrm{v}, \mathrm{m}) \longmapsto (Y, Z, N, U, V, M, \partial U, \partial V, \partial M),
$$

with $(\mathcal{Y}, \mathcal{Z}, \mathcal{N}, U, V, M, \partial U, \partial V, \partial M)$ given for any $s \in [0, T]$, $\mathbb{P}-$a.s. for any $t \in [0, T]$ by

$$\mathcal{Y}_t = \xi(T, X) + \int_t^T h_r(X, y_r, z_r, U_r^r, V_r^r, \partial U_r^r)\mathrm{d}r - \int_t^T \mathcal{Z}_r^\top dX_r - \int_t^T \mathrm{d}N_r,$$

$$U_t^s = \eta(s, X) + \int_t^T g_r(s, X, U_r^s, V_r^s, y_r, z_r)\mathrm{d}r - \int_t^T V_r^{s\top}\mathrm{d}X_r - \int_t^T \mathrm{d}M_r^s,$$

$$\partial U_t^s = \partial_s \eta(s, X) + \int_t^T \nabla g_r(s, X, \partial U_r^s, \partial V_r^s, U_r^s, V_r^s, y_r, z_r)\mathrm{d}r - \int_t^T \partial V_r^{s\top}\mathrm{d}X_r - \int_t^T \mathrm{d}\partial M_r^s.$$

Step 1: We first show $\mathfrak{T}$ is well defined. Let $(y, z, n, u, v, m, \mathrm{u}, \mathrm{v}, \mathrm{m}) \in \mathcal{H}^o$.

(i) Let us first consider the tuples (U, V, M) and $(\partial U, \partial V, \partial M)$. Let us first consider the second equation. Given $(y, z) \in \mathbb{S}^2 \times \mathbb{H}^2$ and Assumption 3.1, this equation is a standard Lipschitz BSDE whose well-posedness follows by classical arguments, see [73; 25]. This yields $(U^s, V^s, M^s) \in \mathbb{S}^2 \times \mathbb{H}^2 \times \mathbb{M}^2$ for all $s \in [0, T]$.

Let us argue the continuity of $([0, T], \mathcal{B}([0, T])) \longrightarrow (\mathbb{L}^2, \| \cdot \|_{\mathbb{L}^2}) : s \longmapsto U^s$. Let $(s_n)_n \subseteq [0, T], s_n \xrightarrow{n \to \infty} s_0 \in [0, T]$ and define for $\varphi \in \{U, V, \eta\}$, $\Delta \varphi^n := \varphi^{s_n} - \varphi^{s_0}$. From the classic stability result for BSDEs we obtain that there is $C > 0$ such that

$$\mathbb{E}\left[\int_0^T |\Delta U_t^n|^2 \mathrm{d}r\right] \leq 2T\left(\|\Delta \eta^n\|_{\mathcal{L}^2}^2 + TL_g^2\left(\rho_g^2(|s_n - s_0|)\right)\right).$$

We conclude $\|U\|_{\mathbb{L}^{2,2}} < \infty$ and $U \in \mathbb{L}^{2,2}$. Given $(U^s, V^s) \in \mathbb{L}^2 \times \mathbb{H}^2$ together with $(y, z) \in \mathbb{S}^2 \times \mathbb{H}^2$, the argument for $(\partial U^s, \partial V^s, \partial M^s)$ for fixed $s \in [0, T]$ is identical.

(ii) We now show that $(V, \partial V, M, \partial M) \in (\mathbb{H}^{2,2})^2 \times (\mathbb{M}^{2,2})^2$. Again, the argument for $(\partial V, \partial M)$ is completely analogous. Applying Itô's formula to $|U_r^s|^2$ we obtain

$$|U_t^s|^2 + \int_t^T |\sigma_r^\top V_r^s|^2 \mathrm{d}r + \int_t^T \mathrm{d}[M^s]_r = |\eta(s)|^2 + 2\int_t^T U_r^s \cdot g_r(s, U_r^s, V_r^s, y_r, z_r)\mathrm{d}r$$

$$- 2\int_t^T U_r^s \cdot V_r^{s\top}\mathrm{d}X_r - 2\int_t^T U_{r-}^s \cdot \mathrm{d}M_r^s.$$

First note $U^s \in \mathbb{S}^2$ guarantees that the last two terms are true martingale for any $s \in [0, T]$. To show that $([0, T], \mathcal{B}([0, T])) \longrightarrow (\mathbb{H}^2, \| \cdot \|_{\mathbb{H}^2})$ (resp. $(\mathbb{M}^2, \| \cdot \|_{\mathbb{M}^2})$) $: s \longmapsto V^s$ (resp. M^s) is continuous, let $(s_n)_n \subseteq [0, T], s_n \xrightarrow{n \to \infty} s_0 \in [0, T]$. We then deduce there is $C > 0$ such that

$$\mathbb{E}\left[\int_0^T |\sigma_r^\top \Delta V_r^n|^2 \mathrm{d}r + [\Delta M^s]_T\right] \leq C\left(\|\Delta \eta\|_{\mathcal{L}^2}^2 + \rho_g^2(|s_n - s_0|)\right),$$

where ρ_g denotes the modulus of continuity of g as defined in the notations section above. Likewise, we obtain

$$\sup_{s \in [0, T]} \mathbb{E}\left[\int_0^T |\sigma_r^\top V_r^s|^2 \mathrm{d}r + [M^s]_T\right] \leq C\left(\|\eta\|_{\mathcal{L}^{2,2}}^2 + \|\tilde{g}\|_{\mathbb{L}^{1,2}}^2\right) < \infty.$$

Since the first term on the right-hand side is finite from Assumption 3.1, we obtain $\|V\|_{\mathbb{H}^{2,2}} + \|M\|_{\mathbb{M}^{2,2}} < \infty$.

We are left to argue $V \in \overline{\mathbb{H}}^{2,2}$. Applying Lemma 6.1 to the system, which holds for any $s \in [0, T]$ $\mathbb{P}-$a.s. for any $t \in [0, T]$,

$$U_t^s = \eta(s, X + \int_t^T g_r(s, X, U_r^s, V_r^s, y_r, z_r)\mathrm{d}r - \int_t^T V_r^{s\top}\mathrm{d}X_r - \int_t^T \mathrm{d}M_r^s,$$

$$\partial U_t^s = \partial_s \eta(s, X) + \int_t^T \nabla g_r(s, X, \partial U_r^s, \partial V_r^s, U_r^s, V_r^s, y_r, z_r)\mathrm{d}r - \int_t^T \partial V_r^{s\top}\mathrm{d}X_r - \int_t^T \mathrm{d}\partial M_r^s,$$

we obtain $s \longmapsto V^s$ is absolutely continuous with density $s \longmapsto \partial V^s$. Consequently, we may define the diagonal of V, denoted by $\mathcal{V} := (V_t^t)_{t \in [0,T]}$, as in Equation (6.6), with

$$V_t^t := V_t^T - \int_t^T \partial V_t^r \mathrm{d}r, \ \mathrm{d}t \otimes \mathrm{d}\mathbb{P}-\text{a.e. in } [0,T] \times \mathcal{X}.$$

The fact $\|\mathcal{V}\|_{\mathbb{H}^2} < \infty$ follows as in Lemma 6.1.(iv). This yields $\mathcal{V} \in \mathbb{H}^2$ and we conclude $V \in \overline{\mathbb{H}}^{2,2}$.

(iii) We derive an auxiliary estimate. Recall $(U, V, M, \partial U, \partial V, \partial M)$ satisfy (6.4). Now, in light of Assumption 3.1 and (ii), we may find, as in **Step** 1 in Proposition 6.3, a universal constant $C > 0$ such that for any $c > 4L_g$ and $t \in [0,T]$

$$\frac{1}{C}\mathbb{E}\bigg[\int_t^T \mathrm{e}^{cr}\big(|U_r^r|^2 + |\partial U_r^r|^2 + |\sigma_r^\top V_r^r|^2\big)\mathrm{d}r\bigg]$$

$$\leq \mathrm{e}^{cT}\big(\|\eta\|_{\mathcal{L}^{2,2}}^2 + \|\tilde{g}\|_{\mathbb{L}^{1,2}}^2 + \|\partial_s\eta\|_{\mathcal{L}^{2,2}}^2 + \|\nabla\tilde{g}\|_{\mathbb{L}^{1,2,2}}^2\big) + \mathbb{E}\bigg[\int_t^T \mathrm{e}^{cr}\big(|y_r|^2 + |\sigma_r^\top z_r|^2\big)\mathrm{d}r\bigg].$$

(iv) We now argue for the tuple $(\mathcal{Y}, \mathcal{Z}, \mathcal{N})$. Notice that

$$\widetilde{\mathcal{Y}}_t := \mathbb{E}\bigg[\xi(T) + \int_0^T h_r(y_r, z_r, U_r^r, V_r^r, \partial U_r^r)\mathrm{d}r \bigg| \mathcal{F}_t\bigg], \text{ is a square integrable } \mathbb{F}\text{-martingale.}$$

Indeed, under Assumption 3.1, h is uniformly Lipschitz in (y, z, u), so (iii) yields

$$\mathbb{E}\big[|\widetilde{\mathcal{Y}}_t|^2\big]$$

$$\leq 6\bigg(\|\xi\|_{\mathcal{L}^2}^2 + \|\tilde{h}\|_{\mathbb{L}^{1,2}}^2 + TL_h^2\bigg(\|y\|_{\mathbb{H}^2}^2 + \|z\|_{\mathbb{H}^2}^2 + \mathbb{E}\bigg[\int_0^T \big(|U_r^r|^2 + |V_r^r|^2 + |\partial U_r^r|^2\big)\mathrm{d}r\bigg]\bigg)\bigg) < \infty$$

Integrating the above expression, Fubini's theorem implies that $\widetilde{\mathcal{Y}} \in \mathbb{L}^2$, thus the predictable martingale representation property for local martingales guarantees the existence of a unique $(\mathcal{Z}, \mathcal{N}) \in \mathbb{H}^2 \times \mathbb{M}^2$ such that $(\mathcal{Y}, \mathcal{Z}, \mathcal{N})$ satisfies the correct dynamics and Doob's inequality implies $\mathcal{Y} \in \mathbb{S}^2$, where

$$\mathcal{Y} := \widetilde{\mathcal{Y}} - \mathbb{E}\bigg[\int_0^\cdot h_r(y_r, z_r, U_r^r, V_r^r, \partial U_r^r)\mathrm{d}r\bigg].$$

All together, we have shown that $\mathfrak{T}(y, z, n, u, v, m, \partial u, \partial v, \partial m) \in \mathcal{H}^o$.

Step 2: We show $\mathfrak{T}$ is a contraction under the equivalent norm $\|\cdot\|_{\mathcal{H}^{o,c}}$, for some $c > 0$ large enough. Let $(y^i, z^i, n^i, u^i, v^i, m^i) \in \mathcal{H}^o$, $\mathfrak{h}^i = \mathfrak{T}(y^i, z^i, n^i, u^i, v^i, m^i, \partial u^i, \partial v^i, \partial m^i)$ for $i \in \{1, 2\}$. We first note that by Lemma 6.2

$$\mathcal{U}_t^i = \eta(T) + \int_t^T \big(g_r(r, \mathcal{U}_r^i, \mathcal{V}_r^i, y_r^i, z_r^i) - \partial\mathcal{U}_r^i\big)\mathrm{d}r - \int_t^T \mathcal{V}_r^i \mathrm{d}X_r - \int_t^T \mathrm{d}\widetilde{\mathcal{M}}_r^i, \ t \in [0,T], \ \mathbb{P}-\text{a.s.,}$$

where $(\mathcal{U}^i, \mathcal{V}^i, \widetilde{\mathcal{M}}^i, \partial\mathcal{U}^i) := \big((U_t^{it})_{t \in [0,T]}, (V_t^{it})_{t \in [0,T]}, (\widetilde{M}_t^i)_{t \in [0,T]}, (\partial U_t^{it})_{t \in [0,T]}\big)$ and $\widetilde{M}^i$ as in Lemma 6.2 for $i \in \{1, 2\}$.

To ease the readability we define $\mathrm{d}t \otimes \mathrm{d}\mathbb{P}-$a.e.

$$\delta h_r := h_r(y_r^1, z_r^1, \mathcal{U}_r^1, \mathcal{V}_r^1, \partial\mathcal{U}_r^1) - h_r(y_r^2, z_r^2, \mathcal{U}_r^2, \mathcal{V}_r^2, \partial\mathcal{U}_r^2),$$

$$\delta\hat{g}_r := g_r(r, \mathcal{U}_r^1, \mathcal{V}_r^1, y_r^1, z_r^1) - \partial\mathcal{U}_r^1 - g_r(r, \mathcal{U}_r^2, \mathcal{V}_r^2, y_r^2, z_r^2) + \partial\mathcal{U}_r^2,$$

$$\delta g_r(s) := g_r(s, U_r^{1s}, V_r^{1s}, y_r^1, z_r^1) - g_r(s, U_r^{2s}, V_r^{2s}, y_r^2, z_r^2),$$

$$\delta\nabla g_r(s) := \nabla g_r(s, \partial U_r^{s1}, \partial V_r^{s1}, U_r^{s1}, V_r^{s1}, y_r^{s1}, z_r^1) - \nabla g_r(s, \partial U_r^{s2}, \partial V_r^{s2}, U_r^{s2}, V_r^{s2}, y_r^{s2}, z_r^2),$$

and
$$\delta \mathfrak{Y} := (\delta \mathcal{Y}, \delta \mathcal{U}, \delta U^s, \delta \partial U^s), \ \delta \mathfrak{Z} := (\delta \mathcal{Z}, \delta \mathcal{V}, \delta V^s, \delta \partial V^s), \ \delta \mathfrak{N} := (\delta \mathcal{N}, \delta \widetilde{\mathcal{M}}, \delta M^s, \delta \partial M^s),$$

whose elements we may denote with superscripts, e.g. $\delta \mathfrak{Y}^1, \delta \mathfrak{Y}^2, \delta \mathfrak{Y}^3, \delta \mathfrak{Y}^4$ correspond to $\delta \mathcal{Y}, \delta \mathcal{U}, \delta U^s, \delta \partial U^s$.

(i) In light of (6.4), as in **Step** 1 in Proposition 6.3, we may find that for $c > 4L_g$ there exists a universal constant $C \in (0, \infty)$ such that for $t \in [0, T]$

$$\mathbb{E}\left[\int_t^T \mathrm{e}^{cr} \big(|\delta \mathcal{U}_r|^2 + |\delta \partial \mathcal{U}_r|^2 + |\sigma_r^\top \delta \mathcal{V}_r|^2 \big) \mathrm{d}r \right] \leq C \mathbb{E}\left[\int_t^T \mathrm{e}^{cr} \big(|\delta y_r|^2 + |\sigma_r^\top \delta z_r|^2 \big) \mathrm{d}r \right]. \quad (6.21)$$

(ii) Applying Itô's formula to $\mathrm{e}^{cr}\big(|\delta \mathcal{Y}_r|^2 + |\delta \mathcal{U}_r|^2 + |\delta U_r^s|^2 + |\delta \partial U_r^s|^2 \big)$ and noticing that $(\delta \mathcal{Y}_T, \delta \mathcal{U}_T, \delta U_T, \delta \partial U_T) = (0,0,0,0)$ we obtain

$$\sum_{i=1}^4 \mathrm{e}^{ct} \big| \delta \mathfrak{Y}_t^i \big|^2 + \int_t^T \mathrm{e}^{cr} \big| \sigma_r^\top \delta \mathfrak{Z}_r^i \big|^2 \mathrm{d}r + \int_t^T \mathrm{e}^{cr-} \mathrm{d}\big[\delta \mathfrak{N}^i \big]_r + \widetilde{\mathfrak{M}}_T^s - \widetilde{\mathfrak{M}}_t^s$$

$$= \int_t^T \mathrm{e}^{cr} \Big(2\delta Y_r \cdot \delta h_r + 2\delta U_r^s \cdot \delta g_r(s) + 2\delta \partial U_r^s \cdot \delta \nabla g_r(s) - c\big(|\delta Y_r|^2 + |\delta U_r^s|^2 + |\delta \partial U_r^s|^2 \big) \Big) \mathrm{d}r,$$

where $\widetilde{\mathfrak{M}}_t^s = 2 \sum_{i=1}^4 \int_0^t \mathrm{e}^{cr} \delta \mathfrak{Y}_r^i \cdot \delta \mathfrak{Z}_r^{i\top} \mathrm{d}X_r + \int_0^t \mathrm{e}^{cr-} \delta \mathfrak{Y}_{r-}^i \cdot \mathrm{d}\delta \mathfrak{N}_r^i$. Again, the fact that $(\delta \mathcal{Y}, \delta \mathcal{U}, \delta U, \delta \partial U) \in (\mathbb{S}^2)^2 \times (\mathbb{S}^{2,2})^2$ guarantees, via the Burkholder–Davis–Gundy inequality, that $\widetilde{\mathfrak{M}}^s$ is a uniformly integrable martingale, and thus a true martingale for all $s \in [0,T]$.

Additionally, under Assumption 3.1.(ii) and 3.1.(iii), $\mathrm{d}t \otimes \mathrm{d}\mathbb{P}$−a.e.

$$|\delta h_r| \leq L_h \big(|\delta y_r| + |\sigma_t^\top \delta z_r| + |\delta \mathcal{U}_r| + |\sigma_r^\top \delta \mathcal{V}_r| + |\delta \partial \mathcal{U}_r^r| \big),$$
$$|\delta \hat{g}_r| \leq L_g \big(|\delta y_r| + |\sigma_t^\top \delta z_r| + |\delta \mathcal{U}_r| + |\sigma_r^\top \delta \mathcal{V}_r| \big) + |\delta \partial \mathcal{U}_r^r|,$$
$$|\delta g_r(s)| \leq L_g \big(|\delta U_r^s| + |\sigma_r^\top \delta V_r^s| + |\delta y_r| + |\sigma_r^\top \delta z_r| \big),$$
$$|\delta \nabla g_r(s)| \leq L_{\partial_s g} \big(|\delta U_r^s| + |\sigma_r^\top \delta V_r^s| + |\delta y_r| + |\sigma_r^\top \delta z_r| \big) + L_g \big(|\delta \partial U_r^s| + |\sigma_r^\top \delta \partial V_r^s| \big).$$

In turn, this implies together with Young's inequality and (6.21), that for any $c > 4L_g$ there exists a universal constant $C > 0$ such that for any $\varepsilon > 0$

$$\mathbb{E}\left[\sum_{i=1}^4 \mathrm{e}^{ct} \big| \delta \mathfrak{Y}_t^i \big|^2 + \int_t^T \mathrm{e}^{cr} \big| \sigma_r^\top \delta \mathfrak{Z}_r^i \big|^2 \mathrm{d}r + \int_t^T \mathrm{e}^{cr-} \mathrm{d}\big[\delta \mathfrak{N}^i \big]_r \right]$$

$$\leq \mathbb{E}\left[\int_t^T \mathrm{e}^{cr} \Big(\big(|\delta \mathcal{Y}_r|^2 + |\delta \mathcal{U}_r|^2 + |\delta U_r^s|^2 + |\delta \partial U_r^s|^2 \big)(C\varepsilon^{-1} - c) \Big) \mathrm{d}r \right]$$

$$+ \varepsilon \mathbb{E}\left[\int_t^T \mathrm{e}^{cr} \big(|\delta y_r|^2 + |\sigma_r^\top \delta z_r|^2 + |\delta \mathcal{U}_r| + |\sigma_r^\top \delta \mathcal{V}_r| + |\delta \partial \mathcal{U}_r| \big) \mathrm{d}r \right]$$

$$\leq \mathbb{E}\left[\int_t^T \mathrm{e}^{cr} \big(|\delta \mathcal{Y}_r|^2 + |\delta \mathcal{U}_r|^2 \big)(C\varepsilon^{-1} - c) \mathrm{d}r \right] + \sup_{s \in [0,T]} \mathbb{E}\left[\int_t^T \mathrm{e}^{cr} |\delta U_r^s|^2 (C\varepsilon^{-1} - c) \mathrm{d}r \right]$$

$$+ \sup_{s \in [0,T]} \mathbb{E}\left[\int_t^T \mathrm{e}^{cr} |\delta \partial U_r^s|^2 (C\varepsilon^{-1} - c) \mathrm{d}r \right] + \varepsilon C \mathbb{E}\left[\int_0^T \mathrm{e}^{cr} \big(|\delta y_r|^2 + |\sigma_r^\top \delta z_r|^2 \big) \mathrm{d}r \right],$$

where in the second inequality C is appropriately updated. Choosing $\varepsilon = Cc^{-1}$ we obtain

$$\mathbb{E}\left[\sum_{i=1}^4 \mathrm{e}^{ct} \big| \delta \mathfrak{Y}_t^i \big|^2 + \int_t^T \mathrm{e}^{cr} \big| \sigma_r^\top \delta \mathfrak{Z}_r^i \big|^2 \mathrm{d}r + \int_t^T \mathrm{e}^{cr-} \mathrm{d}\big[\delta \mathfrak{N}^i \big]_r \right] \leq \frac{C}{c} \Big(\|\delta y\|_{\mathbb{L}^{2,c}} + \|\delta z\|_{\mathbb{H}^{2,c}} \Big),$$

which yields

$$\|\delta \mathfrak{h}\|_{\mathcal{H}^{o,c}}^2 \leq \frac{C}{c} \Big(\|\delta y\|_{\mathbb{L}^{2,c}} + \|\delta z\|_{\mathbb{H}^{2,c}} \Big).$$

We conclude $\mathfrak{T}$ has a fixed point as it is a contraction for c large enough. $\qquad \square$

References

[1] N. Agram. Dynamic risk measure for BSVIE with jumps and semimartingale issues. *Stochastic Analysis and Applications*, 37(3):361–376, 2019. MR-3946615

[2] N. Agram and B. Djehiche. Reflected backward stochastic Volterra integral equations and related time-inconsistent optimal stopping problems. *ArXiv preprint arXiv:2004.11654*, 2020.

[3] N. Agram, B. Øksendal, and S. Yakhlef. Optimal control of forward–backward stochastic Volterra equations. In F. Gesztesy, H. Hanche-Olsen, E.R. Jakobsen, Y.I. Lyubarskii, N.H. Risebro, and K. Seip, editors, *Non–linear partial differential equations, mathematical physics, and stochastic analysis: the Helge Holden anniversary volume*, number 250768 in EMS Series of Congress Reports, pages 3–35. European Mathematical Society Publishing House, Zürich, Switzerland, 2018. MR-3823837

[4] N. Agram, B. Øksendal, and S. Yakhlef. New approach to optimal control of stochastic Volterra integral equations. *Stochastics: An International Journal of Probability and Stochastic Processes*, 91(6):873–894, 2019. MR-3985801

[5] V.V. Anh and J. Yong. Backward stochastic Volterra integral equations in Hilbert spaces. In P.A. Ravi and P Kanishka, editors, *Differential & difference equations and applications*, pages 57–66. Hindawi, New York, 2006. MR-2307328

[6] V.V. Anh, W. Grecksch, and J. Yong. Regularity of backward stochastic Volterra integral equations in Hilbert spaces. *Stochastic Analysis and Applications*, 29(1):146–168, 2010. MR-2763554

[7] V.I. Arkin and M.T. Saksonov. Necessary optimality conditions in control problems for stochastic differential equations. *Doklady Akademii Nauk SSSR*, 244(1):11–15, 1979. MR-0517478

[8] C. Bender and S. Pokalyuk. Discretization of backward stochastic Volterra integral equations. In T. Gerstner and P. Kloeden, editors, *Recent revelopments in computational finance – foundations, algorithms and applications*, volume 14 of *Interdisciplinary mathematical sciences*, pages 245–278. World Scientific Publishing Company, 2013. MR-3288637

[9] A. Bensoussan. Maximum principle and dynamic programming approaches of the optimal control of partially observed diffusions. *Stochastics: An International Journal of Probability and Stochastic Processes*, 9(3):169–222, 1983. MR-0705471

[10] J.-M. Bismut. *Analyse convexe et probabilités.* PhD thesis, Faculté des sciences de Paris, 1973. MR-0324734

[11] J.-M. Bismut. Conjugate convex functions in optimal stochastic control. *Journal of Mathematical Analysis and Applications*, 44(2):384–404, 1973. MR-0329726

[12] J.-M. Bismut. Contrôle des systèmes linéaires quadratiques: applications de l'intégrale stochastique. *Séminaire de probabilités de Strasbourg*, XII:180–264, 1978. MR-0520007

[13] T. Björk, M. Khapko, and A. Murgoci. On time-inconsistent stochastic control in continuous time. *Finance and Stochastics*, 21(2):331–360, 2017. MR-3626618

[14] B. Bouchard, D. Possamaï, X. Tan, and C. Zhou. A unified approach to a priori estimates for supersolutions of BSDEs in general filtrations. *Annales de l'institut Henri Poincaré (B), Probabilités et Statistiques*, 54(1):154–172, 2018. MR-3765884

[15] S. Chen and J. Yong. A linear quadratic optimal control problem for stochastic Volterra integral equations. In S. Tang and J. Yong, editors, *Control theory and related topics – in memory of professor Xunjing Li, Fudan university, China, 3–5 June 2005*, pages 44–66, 2007. MR-2391817

[16] R.J. Chitashvili. Martingale ideology in the theory of controlled stochastic processes. In J.V. Prokhorov and K. Itō, editors, *Probability theory and mathematical statistics. Proceedings of the fourth USSR–Japan symposium, held at Tbilisi, USSR, August 23–29, 1982*, volume 1021 of *Lecture notes in mathematics*, pages 73–92. Springer, 1983. MR-0735975

[17] R.J. Chitashvili and M.G. Mania. Optimal locally absolutely continuous change of measure. Finite set of decisions. Part I. *Stochastics: An International Journal of Probability and Stochastic Processes*, 21(2):131–185, 1987. MR-0897099

[18] R.J. Chitashvili and M.G. Mania. Optimal locally absolutely continuous change of measure. Finite set of decisions. Part II: optimization problems. *Stochastics: An International Journal of Probability and Stochastic Processes*, 21(3):187–229, 1987. MR-0900113

[19] H. Coulibaly and A. Aman. Backward stochastic Volterra integral equations with time delayed generators. *Journal of Mathematics Research*, 11(3):94, 2019. MR-4192905

[20] M.H.A. Davis and P. Varaiya. Dynamic programming conditions for partially observable stochastic systems. *SIAM Journal on Control and Optimization*, 11(2):226–261, 1973. MR-0319642

[21] J. Djordjević and S. Janković. On a class of backward stochastic Volterra integral equations. *Applied Mathematics Letters*, 26(12):1192–1197, 2013. MR-3250433

[22] J. Djordjević and S. Janković. Backward stochastic Volterra integral equations with additive perturbations. *Applied Mathematics and Computation*, 265(174007):903–910, 2015. MR-3373532

[23] I. Ekeland and A. Lazrak. The golden rule when preferences are time inconsistent. *Mathematics and Financial Economics*, 4(1):29–55, 2010. MR-2746576

[24] I. Ekeland and T.A. Pirvu. Investment and consumption without commitment. *Mathematics and Financial Economics*, 2(1):57–86, 2008. MR-2461340

[25] N. El Karoui, S. Peng, and M.-C. Quenez. Backward stochastic differential equations in finance. *Mathematical Finance*, 7(1):1–71, 1997. MR-1434407

[26] Y. Hamaguchi. Small-time solvability of a flow of forward–backward stochastic differential equations. *Applied Mathematics and Optimization,* to appear, 2020. MR-3200726

[27] Y. Hamaguchi. Extended backward stochastic Volterra integral equations and their applications to time–inconsistent stochastic recursive control problems. *Mathematical Control and Related Fields*, 11(2):433–478, 2021. MR-4218119

[28] C. Hernández. On quadratic type-I BSVIEs appearing in time-inconsistent control problems. *Working paper*, 2020.

[29] C. Hernández and D. Possamaï. Me, myself and I: a general theory of non-Markovian time-inconsistent stochastic control for sophisticated agents. *ArXiv preprint arXiv:2002.12572*, 2020.

[30] Y. Hu and B. Øksendal. Linear Volterra backward stochastic integral equations. *Stochastic Processes and their Applications*, 129(2):626–633, 2019. MR-3907012

[31] Y. Hu and S. Peng. Adapted solution of a backward semilinear stochastic evolution equation. *Stochastic Analysis and Applications*, 9(4):445–459, 1991. MR-1133858

[32] J. Jacod and A.N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der mathematischen Wissenschaften*. Springer–Verlag Berlin Heidelberg, 2nd edition, 2003. MR-1943877

[33] Yu.M. Kabanov. On the Pontryiagin maximum principle for the linear stochastic differential equations. In *Probabilistic models and control of economic processes*. TsEMI akademii nauk SSSR, Moscow, 1978.

[34] I. Karatzas and S.E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate texts in mathematics*. Springer, 1991. MR-1121940

[35] E. Kromer and L. Overbeck. Differentiability of BSVIEs and dynamic capital allocations. *International Journal of Theoretical and Applied Finance*, 20(7):1–26, 2017. MR-3724287

[36] W. Li, R. Wu, and K. Wang. Existence and uniqueness of M-solutions for backward stochastic Volterra integral equations. *Electronic Journal of Differential Equations*, 2014(178):1–16, 2014. MR-3262049

[37] J. Lin. Adapted solution of a backward stochastic nonlinear Volterra integral equation. *Stochastic Analysis and Applications*, 20(1):165–183, 2002. MR-1891483

[38] W. Lu. Backward stochastic Volterra integral equations associated with a Lévy process and applications. *ArXiv preprint arXiv:1106.6129*, 2016. MR-4192905

[39] L. Overbeck and J.A.L. Röder. Path-dependent backward stochastic Volterra integral equations with jumps, differentiability and duality principle. *Probability, Uncertainty and Quantitative Risk*, 3(1), 2018. MR-3848765

[40] É. Pardoux and S. Peng. Adapted solution of a backward stochastic differential equation. *System and Control Letters*, 14(1):55–61, 1990. MR-1037747

[41] É. Pardoux and P.E. Protter. Stochastic Volterra equations with anticipating coefficients. *The Annals of Probability*, 18(4):1635–1655, 1990. MR-1071815

[42] A. Popier. Backward stochastic Volterra integral equations with jumps in a general filtration. *ESAIM: PS*, 25:133–203, 2021. MR-4234131

[43] P.E. Protter. Volterra equations driven by semimartingales. *The Annals of Probability*, 13(2):519–530, 1985. MR-0781420

[44] P.E. Protter. *Stochastic integration and differential equations*, volume 21 of *Stochastic modelling and applied probability*. Springer–Verlag Berlin Heidelberg, 2nd edition, 2005. MR-2273672

[45] Y. Ren. On solutions of backward stochastic Volterra integral equations with jumps in Hilbert spaces. *Journal of Optimization Theory and Applications*, 144(2):319–333, 2010. MR-2581109

[46] R.A. Ryan. *Introduction to tensor products of Banach spaces*. Springer monographs in mathematics. Springer London, 2002. MR-1888309

[47] Y. Shi and T. Wang. Solvability of general backward stochastic Volterra integral equations. *Journal of the Korean Mathematical Society*, 49(6):1301–1321, 2012. MR-3025531

[48] Y. Shi, T. Wang, and J. Yong. Mean-field backward stochastic Volterra integral equation. *Discrete and Continuous Dynamical Systems–Series B*, 18(7):1929–1967, 2013. MR-3066329

[49] Y. Shi, T. Wang, and J. Yong. Optimal control problems of forward–backward stochastic Volterra integral equations. *Mathematical Control and Related Fields*, 5(3):613–649, 2015. MR-3371959

[50] Y. Shi, J. Wen, and J. Xiong. Backward doubly stochastic Volterra integral equations and applications to optimal control problems. *Journal of Differential Equations*, 269(9):6492–6528, 2020. MR-4107060

[51] D.W. Stroock and S.R.S. Varadhan. *Multidimensional diffusion processes*, volume 233 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag Berlin Heidelberg, 1997. MR-0532498

[52] J.M.A.M. van Neerven. Approximating Bochner integrals by Riemann sums. *Indagationes Mathematicae*, 13(2):197–208, 2002. MR-2016338

[53] H. Wang. Extended backward stochastic Volterra integral equations, quasilinear parabolic equations, and Feynman–Kac formula. *Stochastics and Dynamics*, to appear, 2020. MR-4192905

[54] H. Wang and J. Yong. Time–inconsistent stochastic optimal control problems and backward stochastic Volterra integral equations. *ESAIM: COCV*, 27(22), 2021. MR-4238770

[55] H. Wang, J. Sun, and J. Yong. Recursive utility processes, dynamic risk measures and quadratic backward stochastic Volterra integral equations. *Applied Mathematics and Optimization*, to appear, 2019. MR-4192905

[56] H. Wang, J. Yong, and J. Zhang. Path dependent Feynman–Kac formula for forward backward stochastic Volterra integral equations. *ArXiv preprint arXiv:2004.05825*, 2020. MR-4192905

[57] T. Wang. Backward stochastic volterra integral equation under local lipschitz condition. *Journal of Shandong University (Natural Science)*, 46(7):112–115, 2011. MR-2932889

[58] T. Wang. L^p solutions of backward stochastic Volterra integral equations. *Acta Mathematica Sinica, English Series*, 28(9):1875–1882, 2012. MR-2956187

[59] T. Wang. Linear quadratic control problems of stochastic Volterra integral equation. *Control, Optimisation and Calculus of Variations*, 24:1849–1879, 2018. MR-3922428

[60] T. Wang and Y. Shi. A maximum principle for forward–backward stochastic Volterra integral equations and applications in finance. *ArXiv preprint arXiv:1004.2206*, 2010.

[61] T. Wang and Y. Shi. Symmetrical solutions of backward stochastic Volterra integral equations and their applications. *Discrete and Continuous Dynamical Systems–Series B*, 14(1):251–274, 2010. MR-2644263

[62] T. Wang and Y. Shi. A class of time inconsistent risk measures and backward stochastic Volterra integral equations. *Risk and Decision Analysis*, 4(1):17–24, 2013. MR-2956187

[63] T. Wang and J. Yong. Comparison theorems for some backward stochastic Volterra integral equations. *Stochastic Processes and their Applications*, 125(5):1756–1798, 2015. MR-3315612

[64] T. Wang and J. Yong. Backward stochastic Volterra integral equation—representation of adapted solutions. *Stochastic Processes and their Applications*, 129(12):4926–4964, 2019. MR-4025696

[65] T. Wang and H. Zhang. Optimal control problems of forward–backward stochastic Volterra integral equations with closed control regions. *SIAM Journal on Control and Optimization*, 55(4):2574–2602, 2017. MR-3686787

[66] Z. Wang and X. Zhang. Non-Lipschitz backward stochastic Volterra type equations with jumps. *Stochastics and Dynamics*, 7(4):479–496, 2007. MR-2378580

[67] Q. Wei and X. Xiao. An optimal control problem of forward–backward stochastic Volterra integral equations with state constraints. *Abstract and Applied Analysis*, 432718(2014), 2014. MR-3176745

[68] Q. Wei, J. Yong, and Z. Yu. Time-inconsistent recursive stochastic optimal control problems. *SIAM Journal on Control and Optimization*, 55(6):4156–4201, 2017. MR-3738841

[69] J. Yong. Backward stochastic Volterra integral equations and some related problems. *Stochastic Processes and their Applications*, 116(5):779–795, 2006. MR-2218335

[70] J. Yong. Continuous-time dynamic risk measures by backward stochastic Volterra integral equations. *Applicable Analysis*, 86(11):1429–1442, 2007. MR-2373675

[71] J. Yong. Well-posedness and regularity of backward stochastic Volterra integral equations. *Probability Theory and Related Fields*, 142(1-2):21–77, 2008. MR-2413266

[72] J. Yong. Time-inconsistent optimal control problems and the equilibrium HJB equation. *Mathematical Control and Related Fields*, 2(3):271–329, 2012. MR-2991570

[73] J. Zhang. *Backward stochastic differential equations: from linear to fully nonlinear theory*, volume 86 of *Probability theory and stochastic modelling*. Springer, 2017. MR-3699487

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards

- Free for authors, free for readers

- Quick publication (no backlog)

- Secure publication (LOCKSS[1])

- Easy interface (EJMS[2])

Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4] , ProjectEuclid[5]

- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)

- Submit your best articles to EJP-ECP

- Choose EJP-ECP over for-profit journals

---

[1]LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
[2]EJMS: Electronic Journal Management System http://www.vtex.lt/en/ejms.html
[3]IMS: Institute of Mathematical Statistics http://www.imstat.org/
[4]BS: Bernoulli Society http://www.bernoulli-society.org/
[5]Project Euclid: https://projecteuclid.org/
[6]IMS Open Access Fund: http://www.imstat.org/publications/open.htm