Lifting of Volterra processes: optimal control and HJB equations

Giulia di Nunno *† Michele Giordano‡

29th October 2020

Keywords: Backward stochastic integral equation; Dynamic programming principle; Hamilton Jacobi Bellman; Optimal control; UMD Banach space; Markovian Lift;
MSC 2020: 60H10; 60H20; 93E20; 35R15; 49L20; 91B70;

Abstract

We study an optimization problem for a Volterra-type controlled forward equation with dependence from the past obtained via a convolution Kernel. In order to solve this problem via dynamic programming principle, we consider a lift of the whole problem (forward process and performance functional) in a Banach space and find optimality conditions via backward stochastic differential equations and HJB techniques. Examples will be given.

1 Introduction

Our goal is to minimize a performance functional of the form

$$J(t,x,u) = \mathbb{E}\left[\int_t^T F(t, X_t^u, u_t)dt + G(X_T^u)\right] \tag{1.1}$$

where F, G are functions with some hypothesis which are going to be stated later, and where the forward process follows controlled Volterra type dynamics given by

$$X_\tau^u = x + \int_t^\tau K(\tau - s)\Big[\beta(s, X_s^u)ds + \sigma(s, X_s^u)R(s, X_s^u, u_s)ds + \sigma(s, X_s^u)dW_s)\Big], \tag{1.2}$$

with initial value $X_t = x$, and where $\beta : [0,T] \times \mathbb{R} \longrightarrow \mathbb{R}$, $\sigma : [0,T] \times \mathbb{R} \longrightarrow \mathbb{R}$, and $R : [0,T] \times \mathbb{R} \times \mathcal{U} \longrightarrow \mathbb{R}$. K is a convolution kernel $K : [0,T]^2 \longrightarrow \mathbb{R}^+$ on which we assume

*Department of Mathematics, University of Oslo, P:O: Box 1053 Blindern, N-0316 Oslo, Email:giulian@math.uio.no.

†Department of Business and Management Science, NHH Norwegian School of Economics, Helleveien 30, N-5045 Bergen.

‡Department of Mathematics, University of Oslo, P:O: Box 1053 Blindern, N-0316 Oslo, Email:michelgi@math.uio.no

some hypothesis that will be discussed later on.
Usually, when dealing with such problems, one cannot directly apply a dynamic programming principle, due to the non markovianity of the framework. Many authors (see e.g. [2, 3, 17, 18]) tackled this problem by means of a maximum principle approach, which allows them to derive sufficient and necessary conditions for the optimality. In this paper though, similarly to what has been done in [1] for the linear-quadratic case, we aim to reformulate our optimization problem (1.2)-(1.1) in an infinite dimensional setting and solve the newly formulated problem by means of a dynamic programming principle.

The lift to the infinite dimensional setting that we are going to present, firstly introduced in [7], and then developed in [6] for the multi-dimensional case, and in [5] for a general Lévy driver, is a generalization of the case presented in [1]. This lift allows us to recover some Markov properties for the forward process (1.2) which, in turn, will let us derive a dynamic programming principle in terms of the HJB equations. The main difference between the work done in [1] and the one presented here is the more general class of kernel considered, together with a broader class of performance functionals. Here we propose a generalization of the work done in [1], where the authors consider the case of a convolution kernel K that can be expressed as the Laplace transform of a measure, and where the performance functional (1.1) is of linear-quadratic type. In fact, due to the recent development in [5, 6, 7], we can present a more general setting for both the convolution kernel and the performance functional.

Since we focus on markovian lifts, as introduced in [5, 6, 7], we assume that the kernel K can be represented as $K(t) = \langle g, \mathcal{S}_t^* \nu \rangle$, for $\mathcal{S}_t^*$ a strongly continuous semigroup acting on a Banach space Y^*, $\nu \in Y^*$, $g \in Y$ with Y the pre dual of Y^* and pairing $\langle \cdot, \cdot \rangle$. Even though this hypothesis seems quite restrictive, one can find several kernels that satisfy this condition, making the present framework quite general.

As anticipated, our goal is to find

$$J(t, x, \hat{u}) = \inf_{u \in \mathcal{A}} J(t, x, u)$$

for some admissible control set $\mathcal{A}$ defined as

$$\mathcal{A} = \{u : [0, T] \times \Omega \longrightarrow \mathcal{U},\ s.t.\ u\ is\ \mathcal{F}_\tau - predictable\}$$

with $\mathcal{U}$ a closed convex subset of $\mathbb{R}$. In order to do so, we are going to reformulate the optimization problem (1.2)-(1.1) in the Banach space Y^*. Our idea is in fact to solve an optimization problem equivalent to the original one, which does not involve dependence of the forward process from its past.

Once the original optimization problem is reformulated as a Banach valued problem, following the approach presented in [9, 12], and exploiting the Malliavin calculus in *UMD Banach spaces* as in, for example, [13] and the references therein, we are able to solve the lifted problem (2.5), with forward dynamics given by (2.6).

In order to maintain the highest generality for the volatility term in the forward equation, we will need to use some form of Malliavin calculus, which is not available in general Banach spaces. This means that we are going to restrict ourselves to the UMD Banach space setting, where the Malliavin calculus is fully developed (see e.g. [13]). Thanks to the structure of the lift, we also show that the UMD hypothesis is not as restrictive as it might seem at a first glance. The UMD hypothesis will also be lifted when σ does not depend on X, as one can follow the optimization approach presented in [12] in order to obtain the optimal control.

The present work introduces an element of novelty also in the infinite dimensional optimization part by considering a setting which is halfway between the one presented in [9] and the one in [12]. In [9] the authors considered a Hilbert valued forward controlled process, whereas in [12], a more general Banach space is considered, but the downside is that the drift term σ does not depend on X. Our work aims to present a new setting, taking the general volatility dynamics for the forward process presented in [9] and a more general setting than the Hilbert space setting in that paper. Unfortunately, in doing so, we are not able to consider a general Banach space as done in [12] but, as we will show in the following part, this is beyond our goal.

This paper is structured as follows: in section 2 we are going to present some preliminary results both on the Gâteaux differentiability in general Banach spaces and on the lift for Volterra processes. In section 3 we give an existence and continuity result for the forward equation, and in section 4 we introduce the backward equation and the Hamiltonian function associated with the lifted optimization problem and we present a solution method via HJB equations. Lastly, in section 5 we present a concrete example and we obtain an implicit characterization of the optimal control u.

2 Some preliminary results

We present now some results that we will use throughout the paper. First we are going to recall the definition of Gâteaux derivative and state some properties, then we present some results on Banach spaces and Malliavin differentiability in UMD Banach spaces, and lastly we recall how the Markovian lift is obtained.

For the proofs of the results that we are going to state we refer to [9] for the Gâteaux derivative and the results on Banach spaces, [8, 13] and [11] for the Malliavin calculus in UMD Banach spaces, and to [5, 7] for the ones concerning Markovian lifts.

2.1 The class of Gâteaux differentiable functions

We recall that, for a mapping $F : X \longrightarrow V$, with X, V two Banach spaces, the directional derivative at point $x \in X$ in the direction $h \in X$ is defined as

$$\nabla F(x; h) = \lim_{s \to 0} \frac{F(x + sh) - F(x)}{s}$$

whenever the limit exists in the topology of V. F is said to be Gâteaux differentiable at point x if it has directional derivative in every direction at x and there exists an element of $L(X, V)$ denoted with $\nabla F(x)$, called Gâteaux derivative, such that $\nabla F(x; h) = \nabla F(x)h$ for every $h \in X$.

Definition 2.1. We say that a mapping $F : X \longrightarrow V$ belongs to $\mathcal{G}^1(X; V)$ if it is continuous, Gâteaux differentiable on X and $\nabla F : X \longleftarrow L(X; V)$ is strongly continuous, i.e. the map $\nabla F(\cdot)h : X \longrightarrow V$ is continuous for every $h \in X$.

Remark 2.2. Let X, V, Z be three Banach spaces and $F \in \mathcal{G}^1(X, V)$. If $G \in \mathcal{G}^1(V, Z)$, then $G(F) \in \mathcal{G}^1(X, Z)$ and $\nabla(G(F))(x) = \nabla G(F)(x)\nabla F(x)$.

Definition 2.3. Given X, V, Z three Banach spaces, we say that the mapping $F : X \times Y \longrightarrow V$ belongs to the class $\mathcal{G}^{1,0}(X \times Y; V)$ if it is continuous, Gâteaux differentiable with respect to x on $X \times Y$ and $\nabla_x F : X \times Y \longrightarrow L(X, V)$ is strongly continuous.

Obviously when F depends on additional arguments the previous definition can be generalized.

Lemma 2.4. *Given X, Y, V three Banach spaces, a continuous map $F : X \times Y \longrightarrow V$ belongs to $\mathcal{G}^{1,0}(X \times Y, V)$ provided the following conditions hold:*

1. *The directional derivatives $\nabla_x F(x, y; h)$ exist at every point $(x, y) \in X \times Y$ and in every directoin $h \in X$.*
2. *For every h, the mapping $\nabla F(\cdot, \cdot; h) : X \times Y \longrightarrow V$ is continuous.*
3. *For every (x, y) the mapping $h \mapsto \nabla_x F(x, y; h)$ is continuous from X to V.*

We are going to use the following parameter depending contraction principle in order to study regular dependence of solution of stochastic equations on their initial data:

Proposition 2.5. *Let X, Y, Z be Banach spaces and let $F : X \times Y \times Z \longrightarrow X$ a continuous mapping satisfying*

$$|F(x_1, y, z) - F(x_2, y, z)| \leq \alpha |x_1 - x_2|,$$

for some $\alpha \in [0, 1)$ and every x_1, x_2 in X, y in Y, z in Z. Let $\phi(y, z)$ denote the unique fixed point of the mapping $F(\cdot, y, z) : X \longrightarrow X$. Then $\phi : Y \times Z \longrightarrow X$ is continuous. If, in addition $F \in \mathcal{G}^{1,1,0}(X \times Y \times Z, X)$ then $\phi \in \mathcal{G}^{1,0}(Y \times Z, X)$ and

$$\nabla_y \phi(y, z) = \nabla_x F(\phi(y, z), y, z)\nabla_y \phi(y, z) + \nabla_y F(\phi(y, z), y, z)$$

In this framework we will need to use a Clark-Okone formula for UMD Banach spaces (see [11]) and the following chain rule (see [13] Proposition 3.8) that links the Malliavin derivative and the Gâteaux derivative:

Proposition 2.6. *Let E be a UMD Banach space and let $p \in (1, \infty)$. Suppose $\varphi : E \longrightarrow E$ is Gâteaux differentiable and has a continuous and bounded derivative. If $F \in \mathbb{D}^{1,p}(E)$, then $\varphi(F) \in \mathbb{D}^{1,p}(E)$ with*

$$D(\varphi(F)) = \nabla\varphi(F)DF$$

2.2 The Markovian Lift

As stated in the introduction we are interested in considering a forward equation with dynamics as in (1.2). In this paper we focus on Markovian lifts as presented in [5, 7]. Suppose to have a Volterra process of the form

$$X_t = f(t) + \int_0^t K(t-s)dV(s)$$

where $f(t)$ is a deterministic function, $K \in L^2_{loc}(\mathbb{R}_+; \mathbb{R})$ and V is a semimartingale depending on X. We suppose that the kernel K can be represented as $K(t) = \langle g, \mathcal{S}^*_t \nu \rangle$ with $(\mathcal{S}^*_t)_{t \geq 0}$ a strongly continuous semigroup acting on a banach space Y^*, $\nu \in Y^*$, $g \in Y$ and pairing $\langle \cdot, \cdot \rangle$ between Y and its dual space Y^*.

When the above hypothesis are satisfied, if we define

$$d\mathcal{Z}_t = \mathcal{A}^* \mathcal{Z}_t dt + \nu dV_t \tag{2.1}$$

we are able to rewrite $X_t = \langle g, \mathcal{Z}_t \rangle$. Recall now the definition of generalized Feller process:

Definition 2.7. A family of bounded linear operators P_t from $B^\rho(X)$ onto itself s.t.

(i) It has the semigroup property, and $P_0 f = Id$

(ii) $\lim_{t \to 0} T_t f - f = 0 \ \forall f \in C^\rho(X)$

(iii) There exist a constant $C \in \mathbb{R}$ and $\epsilon > 0$ s.t. $\forall t \in [0, \epsilon]$, $\|P_t\|_{L(B^\rho(X))} < C$

(iv) $P_t \geq 0$ for all $t \geq 0$

where $B^\rho(X) := \left\{ f : X \to \mathbb{R} : \|f\|_\rho := \sup_{\lambda \in X} \rho(\lambda)^{-1} \|f(\lambda)\| < \infty \right\}$, is a Generalized Feller process.

Then from [5] we have that

Theorem 2.8. *(i) The SPDE $\mathcal{Z}_t$ in* (2.2) *admits a unique Markovian solution in the sense that it generates a generalized Feller semigroup* *(ii) The associated generalized Feller process $(\mathcal{Z}_t)_t$ is also a probabilistycally weak and analytically mild càg solution on an appropriate probabilistic basis.*

(iii) For all $\mathcal{Z}_0 \in Y^*$, the corresponding jump diffusion stochastic Volterra equation $X_t = \langle g, \mathcal{Z}_t \rangle$ given by

$$X_t = \langle g, \mathcal{Z}_t \rangle = \langle g, \mathcal{S}^*_\tau \mathcal{Z}_0 \rangle + \int_t^\tau \langle g, \mathcal{S}^*_{\tau-s} \nu \rangle dV_t$$

with $f(t) = \langle g, \mathcal{S}^*_\tau \mathcal{Z}_0 \rangle$ admits a probabilistically weak solution with càg trajectories.

This allows us to recover some Markov properties for the forward equation (1.2) by lifting the original dynamics in an infinite dimensional Banach space. This theorem also guarantees that the forward Volterra equation (1.2) admits a solution.

Remark 2.9. Intuitively we have that, if our process $\mathcal{Z}$ follows the dynamics given by (2.1), then its solution is given by

$$\mathcal{Z}_t = \mathcal{S}^*_t \mathcal{Z}_0 + \mathcal{S}^*_{t-s} \nu V_s ds$$

We thus have that, by evaluating in g

$$\langle g, \mathcal{Z}_t \rangle = \langle g, \mathcal{S}^*_t \mathcal{Z}_0 \rangle + \int_0^t \langle g, \mathcal{S}^*_{t-s} \nu \rangle dV_s = X_t$$

Notice also that we considered in (1.2) a single kernel $K(t)$ that multiplies both the drift term and the noise term. This framework could be easily generalized to two different kernels $K_1(t)$ and $K_2(t)$ by simply increasing the dimension of our space and following the work done in [6]. More in general, if we had that the dynamics for X_t were given by

$$X_t = x + \int_0^t \mu(t, s, X_s, u_s) ds + \int_0^t \sigma(t, s, X_s, u_s) dW_s$$

by assuming that the terms μ and σ could be represented in terms of $\mathcal{S}^*_t$ as

$$\langle g, \mathcal{S}^*_{t-s} a(s, \mathcal{Z}) \rangle = \mu(t, s, \langle g, \mathcal{Z} \rangle),$$
$$\langle g, \mathcal{S}^*_{t-s} b(s, \mathcal{Z}) \rangle = \sigma(t, s, \langle g, \mathcal{Z} \rangle),$$

we would still be able to lift our forward equation X_t and represent it as $X_t = \langle g, \mathcal{Z}_t \rangle$, with

$$d\mathcal{Z}_t = \mathcal{A}^* \mathcal{Z}_t dt + a(t, \mathcal{Z}_t, u_t) dt + b(t, \mathcal{Z}_t, u_t) dW_t$$

Remark 2.10. In general, we could consider a broader classes of kernels by considering a subspace $Z \subset Y$ with their relative duals $Y^* \subset Z^*$. In this case we should have that

- Z and Y are Banach spaces $Z \subset Y$ and Z embeds continuously into Y.

- The semigroup $\mathcal{S}^*$ with generator $\mathcal{A}^*$ acts in a strongly continuous way on Y^* and Z^* with respect to the respective norm topologies.

- The map $\mathcal{Z} \mapsto \mathcal{S}_t^* \mathcal{Z}$ is weak-* continuous on Y^* and on Z^* for every $t \geq 0$

- The pre adjoing operator of $\mathcal{A}^*$, generates a strongly continuous semigroup on Z with respect to the respective norm topology (but not necessarily on Y).

In this case we can have that $K(t) = \langle g, \mathcal{S}_t^* \nu \rangle$ with $\nu \in Z^*$ if $S_t^* \nu$ is in Y^*. We would then need to assume that $\mathcal{S}_t^* \nu \in Y^*$ for all $t > 0$ and that $\int_0^t \|\mathcal{S}_s^* \nu\|_{Y^*}^2 ds < \infty$ for all $t > 0$ so that the previous terms are well defined.

Unfortunately though, while this setting allows us to lift a slightly bigger class of kernels, we cannot use this framework for the optimization purposes. We are thus going to be considering $\nu \in Y^*$, as originally stated.

2.3 The lift approach

Having introduced the lift presented in [5, 6, 7], we exploit it in order to reformulate our optimization problem. As shown above, we rewrite (1.2) as $X_t^u = \langle g, \mathcal{Z}_t^u \rangle$, where the dynamics for $\mathcal{Z}$ are given by

$$d\mathcal{Z}_\tau^u = \mathcal{A}^* \mathcal{Z}_\tau^u d\tau + \nu \left(\beta(\tau, \mathcal{Z}_\tau^{u,g}) + \sigma(\tau, \mathcal{Z}_\tau^{u,g})[R(\tau, \mathcal{Z}_\tau^{u,g}, u_\tau) d\tau + dW_\tau]\right) \tag{2.2}$$

and $\mathcal{Z}_\tau^{u,g}$ is defined as $\mathcal{Z}_\tau^{u,g} := \langle g, \mathcal{Z}_\tau^u \rangle$. Here the initial condition at time t is $\mathcal{Z}_t = \zeta$ where $x = \langle g, \mathcal{S}_t^* \zeta \rangle$

One can in fact see that

$$\begin{aligned} X_\tau^u &= \langle g, \mathcal{Z}_\tau^u \rangle \qquad (2.3) \\ &= \langle g, \zeta + \int_t^\tau \mathcal{A}^* \mathcal{Z}_s^u + \nu \left(\beta(s, \mathcal{Z}_s^{u,g}) + \sigma(s, \mathcal{Z}_s^{u,g})[R(s, \mathcal{Z}_s^{u,g}, u_s) ds + dW_s]\right) \rangle. \end{aligned}$$

In a similar fashion, we rewrite the performance functional (1.1) in order to explicit its dependence from the lifted process $\mathcal{Z}_t$ as

$$\begin{aligned} J(t, x, u) &= \mathbb{E}\left[\int_t^T F(t, \mathcal{Z}_t^{u,g}, u_t) dt + G(\mathcal{Z}_T^{u,g})\right] \\ &= \mathbb{E}\left[\int_t^T F^g(t, \mathcal{Z}_t^u, u_t) dt + G^g(\mathcal{Z}_T^u)\right] := J^g(t, x, u) \qquad (2.4) \end{aligned}$$

Here we have that the function $F : [0, T] \times \mathbb{R} \times \mathcal{U} \longrightarrow \mathbb{R}$ is now considered as a function taking values from the Banach space Y^* instead. In fact, we have that $F^g : [0, T] \times Y^* \times \mathcal{U} \longrightarrow \mathbb{R}$ and it is defined as $F^g(\cdot, \mathcal{Z}_t^u, \cdot) := F(\cdot, \mathcal{Z}_t^{u,g}, \cdot)$. Analogously we can make the same lift for the function G and reformulate our performance functional as above. Our goal becomes to find:

$$J^g(t, \mathcal{Z}, \hat{u}) = \inf_{u \in \mathcal{A}} J^g(t, \mathcal{Z}, u) = \inf_{u \in \mathcal{A}} \mathbb{E}\left[\int_t^T F^g(t, \mathcal{Z}_t^u, u)dt + G^g(\mathcal{Z}_T^u)\right] \tag{2.5}$$

where the process $\mathcal{Z}^u$ takes value in the Banach space Y^*, and where the dynamics for the controlled process are given by (2.2). In order to simplify the notation in the following part we also rewrite (2.2) as

$$d\mathcal{Z}_\tau^u = \mathcal{A}^* \mathcal{Z}_\tau^u d\tau + \beta_\nu^g(\tau, \mathcal{Z}_\tau^u)ds + \sigma_\nu^g(\tau, \mathcal{Z}_\tau^u)[R^g(\tau, \mathcal{Z}_\tau^u, u_\tau)ds + dW_\tau] \tag{2.6}$$

where $\beta_\nu^g(\tau, \mathcal{Z}_\tau^u) := \nu\beta(\tau, \mathcal{Z}_s^{u,\tau})$ and similarly for σ and R. Notice that this change of notation embodies a crucial change of framework from a finite dimensional to an infinite dimensional setting, allowing us to move from functions β and σ taking values from $\mathbb{R}$ to new functions β^g and σ^g that now take values from Y^*.

This lift allows us to solve a new optimization problem, written on a probabilistic basis which is not the original one. Nonetheless, we have that $J(u) = J^g(u)$ for all $u \in \mathcal{U}$ and that finding the optimal pair $(\hat{u}, \mathcal{Z}^{\hat{u}})$ is equivalent to finding the optimal pair $(\hat{u}, X^{\hat{u}})$.

3 Background results on the lifted process

In order to be able to apply the Malliavin calculus in a Banach space we have to make an additional Hypothesis on Y^*, namely

Hypothesis 3.1. *UMD The space Y^* is a UMD (unconditional martingale differences) Banach Space.*

Definition 3.2. Let $(M_n)_{n=1}^N$ be a Banach-space valued martingale, the sequence $d_n = M_{n+1} - M_n$ is called the martingale difference sequence associated with $(M_n)_{n=1}^N$. A Banach space E is said to be a UMD_p, $(1 < p < \infty)$, space if there exists a constant β such that for all E-valued L^p-martingale difference sequences $(d_n)_{n=1}^N$ we have

$$\mathbb{E}\left\|\sum_{n=1}^N \epsilon_n d_n\right\|^p \leq \beta^p \mathbb{E}\left\|\sum_{n=1}^N d_n\right\|^p.$$

Thanks to [10] we also know that if a Banach space E is a UMD_p Banach space for some $1 < p < \infty$ then E is a UMD_p Banach space for all $p \in (1, \infty)$ and we simply call it a *UMD Banach space.*

In the context of stochastic analysis in Banach spaces, martingale difference sequences provide somehow a substitute for orthogonal sequences. As already anticipated, we will see that this hypothesis is not very restrictive, as the UMD Banach spaces include all Hilbert spaces, L^q spaces for $q \in (1, \infty)$, reflective Sobolev spaces and many others. In our framework $\mathcal{Z}_t$ is going to be in a UMD Banach space whenever we consider, for example, a Shift operator, a quasi-exponential kernel. For the Laplace transform case we are going to need some additional work.

For the stochastic calculus in UMD Banach spaces, and for all the results in UMD Banach spaces that we are going to use, we refer to [16] for the Burkholder–Davis–Gundy inequality, [15] for the Fubini Theorem, [11] for the Clark-Ocone formula and in general to [13] for the Malliavin calculus results.

3.1 The Forward equation

First of all we study the lifted forward equation. From [5] we already know that it admits a unique mild Markovian solution in Y^*, but we are interested in studying is continuous dependence from the initial parameter ζ. We start by considering the case where R in (1.2) is equal to 0 as the general case can be obtain via a Girsanov change of measure. Our lifted forward thus becomes

$$\begin{cases} \mathcal{Z}_\tau &= \mathcal{A}^* \mathcal{Z}_\tau d\tau + \beta^g_\nu(\tau, \mathcal{Z}_\tau) d\tau + \sigma^g_\nu(\tau, \mathcal{Z}_\tau) dW_\tau, \tau \in [t, T] \\ \mathcal{Z}_t &= \zeta \end{cases} \tag{3.1}$$

where W_t is a Brownian motion defined on a Hilbert space H.

We assume the following:

Hypothesis 3.3. *Suppose that*

i) The operator $\mathcal{A}^*$ is the generator of a strongly continuous semigroup in the Banach space Y^*.

ii) $\beta^g_\nu : [0,T] \times Y^* \longrightarrow \mathbb{R}$ is continuous and, for all $t \in [0,T]$ $\zeta_1, \zeta_2 \in Y^*$,

$$|\beta^g_\nu(t, \zeta) - \beta^g_\nu(t, \zeta_2)| \leq K_{lip} |\zeta_1 - \zeta_2|_{Y^*}$$

the map $\beta^g_\nu : [0,T] \times Y^* \longrightarrow Y^*$ is measurable.

iii) $\sigma^g_\nu : [0,T] \times Y^* \longrightarrow L(H)$ is such that, for every $v \in Y^{**}$ the map $\sigma^g_\nu v : [0,T] \times Y^* \longrightarrow H$ is measurable, $e^{s\mathcal{A}^*} \sigma^g_\nu(t, \zeta) \in L^2(H, Y^*)$ for every $s > 0$, $t \in [0,T]$ and $\zeta \in Y^*$, and

$$|e^{s\mathcal{A}^*} \sigma^g_\nu(t, \zeta)|_{L^2(H, Y^*)} \leq L s^{-\gamma} (1 + |\zeta|_{Y^*}).$$

$$|\sigma^g_\nu(t, \zeta)|_{L(H, Y^*)} \leq L(1 + |\zeta|)$$

for some constants $L > 0$ and $\gamma \in [0, 1/2)$.

Moreover, for $s > 0$, $t \in [0,T]$, $\zeta_1, \zeta_2 \in Y^*$

$$|e^{s\mathcal{A}^*} \sigma(t, \zeta_1) - e^{s\mathcal{A}^*} \sigma(t, \zeta_2)|_{L^2(H, Y^*)} \leq L s^{-\gamma} |\zeta_1 - \zeta_2|_{Y^*}$$

iv) For every $s > 0$, $t \in [0,T]$, $v \in H$

$$\beta_\nu^g(t,\cdot) \in \mathcal{G}^1(Y^*,Y^*), \quad e^{s\mathcal{A}^*}\sigma_\nu^g(t,\cdot) \in \mathcal{G}^1(Y^*,L^2(H,Y^*))$$

Proposition 3.4. *Assume Hypothesis 3.3, then for every $p \in [2,\infty)$, the following hold:*

i) The map $x \mapsto \mathcal{Z}(\cdot,t,x)$ belongs to $\mathcal{G}^{0,1}([0,T]\times Y^*; L^p(\Omega; C([0,T];Y^*)))$

ii) Denoting by $\nabla_x \mathcal{Z}$ the partial Gâteaux derivative, for every $h \in Y^*$ the directional derivative process $\nabla_x\mathcal{Z}(\tau,t,x)h$, $\tau \in [0,T]$ solves, $\mathbb{P}$-a.s., the equation

$$\nabla_x\mathcal{Z}(\tau,t,x)h = e^{(\tau-t)\mathcal{A}^*}h + \int_t^\tau e^{(\tau-s)\mathcal{A}^*}\nabla_x\beta_\nu^g(s,\mathcal{Z}(s,t,x))\nabla_x\mathcal{Z}(s,t,x)hds$$
$$+ \int_t^\tau \nabla_x\left(e^{(\tau-s)\mathcal{A}^*}\sigma_\nu^g(s,\mathcal{Z}(s,t,x))\right)\nabla_x\mathcal{Z}(s,t,x)hdW_s, \ \tau \in [t,T]$$
$$\nabla_x\mathcal{Z}(\tau,t,x,)h = h, \ \tau \in [0,t)$$

iii) $|\nabla_x\mathcal{Z}(\tau,t,x)h|_{L^p(\omega;C([0,T];Y^*))} \le c|h|$ for some constant c.

We also find once again, as a result of this theorem that

iv) (3.1) admits a unique predictable solution $\mathcal{Z} \in L^p(\Omega, C([t,T]);H)$ and

$$\mathbb{E}\left[\sup_{\tau\in[t,T]}|\mathcal{Z}_\tau|^p\right] \le C(1+|\zeta|^p) \tag{3.2}$$

for some constant C depending only on p,γ,T,L and $M = \sup_{\tau\in[t,T]}|e^{\tau\mathcal{A}^*}|$

Proof. The proof follows the lines of [9] Proposition 3.2 - 3.3. The main difference between our work paper and the cited one, is that we are considering a UMD Banach space instead of a Hilbert one. We are thus going to give the main ideas of the proof and point out the differences between our work and the one in [9].

We start by defining a map from $L^p(\Omega;C([t,T];Y^*))$ to itself by the formula

$$\Phi(\mathcal{Z})_\tau = e^{(\tau-t)\mathcal{A}^*}\mathcal{Z} + \int_t^\tau e^{(\tau-s)\mathcal{A}^*}\beta_\nu^g(s,\mathcal{Z}_s)ds$$
$$+ \int_t^\tau e^{(\tau-s)\mathcal{A}^*}\sigma_\nu^g(s,\mathcal{Z}_s)dW_s \ \tau \in [t,T]$$

and we show that it is a contraction under an equivalent norm. For simplicity we set $t = 0$ and $\beta = 0$, as the general case can be handled similarly. We define the norm

$$\|\mathcal{Z}\|^p = \mathbb{E}\left[\sup_{\tau\in[0,T]} e^{-\vartheta\tau p}|\mathcal{Z}_\tau|^p\right]$$

In the space $L^p(\Omega; C([t,T]; Y^*))$ this norm is equivalent to the original one. We take $p > 2$ and $\alpha \in (0,1)$ such that

$$\frac{1}{p} < \alpha < \frac{1}{2} \text{ and let } c_\alpha^{-1} = \int_\eta^\tau (\tau - s)^{\alpha-1}(s-\eta)^{-\alpha} ds$$

Then, by the Fubini Theorem, as in [15], we get that

$$\begin{aligned}\Phi(\mathcal{Z})_\tau &= e^{\tau\mathcal{A}^*}\mathcal{Z} + c_\alpha \int_0^\tau \int_\eta^\tau (\tau - s)^{\alpha-1}(s-\eta)^{-\alpha} e^{(\tau-s)\mathcal{A}^*} e^{(s-\eta)\mathcal{A}^*} ds \sigma_\nu^g(\eta, \mathcal{Z}_\eta) dW_\eta \\ &= e^{\tau\mathcal{A}^*}\mathcal{Z} + c_\alpha \int_0^\tau (\tau - s)^{\alpha-1} e^{(\tau-s)\mathcal{A}^*} Y_s ds\end{aligned}$$

where we define

$$Y_s := \int_0^s (s-\eta)^{-\alpha} e^{(s-\eta)\mathcal{A}^*} \sigma^g(\eta, \mathcal{Z}_\eta) dW_\eta$$

By Holder inequality we now get that

$$\begin{aligned}&e^{-\vartheta\tau} \left| \int_0^\tau (\tau - s)^{\alpha-1} e^{(\tau-s)\mathcal{A}^*} Y_s ds \right| \\ &\quad \leq M \left(\int_0^T e^{-p'\vartheta s} s^{(\alpha-1)p'} ds \right)^{1/p'} \left(\int_0^T e^{-p\vartheta s} |Y_s|^p ds \right)^{1/p}\end{aligned}$$

where we set $M := \sup_{\tau \in [0,T]} |e^{\tau\mathcal{A}^*}|$ and $p' = p/(p-1)$. We want now to estimate $\mathbb{E}[|Y_s|^p]$. In order to do so we use the BDG inequalities, as in [16], and we get that, for some constant c_p depending only on p,

$$\mathbb{E}\left[|Y_s|^p\right] \leq K^p c_p \mathbb{E}\left[\sup_{\eta \in [0,s]} \left[(1 + |\mathcal{Z}_\eta|)^p e^{-p\vartheta\eta} \right] \left(\int_0^s (s-\eta)^{-2\alpha-2\gamma} e^{2\vartheta\eta} d\eta \right)^{p/2} \right]$$

from which we have that

$$e^{-p\vartheta s}\mathbb{E}[|Y_s|^p] \leq L^p c_p (1 + \|\mathcal{Z}\|^p) \left(\int_0^T \eta^{-2\alpha-2\gamma} e^{-2\vartheta\eta} d\eta \right)^{p/2}$$

From which have that Φ is a well defined mapping. If now we take ζ^1, ζ^2 in $L^p(\Omega; C([0,T]; Y^*))$, similarly to what has been done above, we can prove that

$$\begin{aligned}\|\Phi(\zeta^1) - \Phi(\zeta^2)\| &\leq M L c_\alpha (T c_p)^{1/p} \|\zeta^1 - \zeta^2\| \\ &\quad \cdot \left(e^{-p'\vartheta s} s^{(\alpha-1)p'} ds \right)^{(1/p')} \left(\int_0^T \eta^{-2\alpha-2\gamma} 2^{-2\vartheta\eta} d\eta \right)^{1/2}\end{aligned}$$

and thus, for ϑ large enough, Φ is a contraction. Thanks to the fixed point Theorem we find once again that (3.1) admits a unique solution. Moreover noticing that $\|\mathcal{Z}\| \leq C(1 + |\zeta|)$, we also have (3.2).

We are now interested in considering Φ as a mapping

$$\Phi(\zeta, t, x,)_\tau : L^p(\Omega; C([0,T]; Y^*)) \times [0,T] \times Y^* \longrightarrow L^p(\Omega; C([0,T]; Y^*))$$

defined as

$$\begin{aligned}\Phi(\mathcal{Z}, t, x)_\tau &:= \mathcal{S}^*(\tau - t)x + \int_0^\tau \mathbb{1}_{[t,T]}(\eta)\mathcal{S}^*(\tau-\eta)\beta_\nu^g(\eta, \mathcal{Z}_\eta)d\eta \\ &+ \int_0^\tau \mathbb{1}_{[t,T]}(\eta)\mathcal{S}^*(\tau-\eta)\sigma_\nu^g(\eta, \mathcal{Z}_\eta)dW_\eta\end{aligned}$$

for $\tau \in [0,T]$. Since $\Phi(\cdot, t, x)$ is a contraction uniformly with respect to t, x, by Proposition 2.5 we just need to show that

$$\Phi \in \mathcal{G}^{1,0,1}\left(L^p(\Omega; C([0,T]; Y^*)) \times [0,T] \times Y^*, L^p(\Omega; C([0,T]; Y^*))\right)$$

and this follows by an extension of Lemma 2.4 if we prove the following parts.
1. Φ is continuous: as we noticed that $\Phi(\cdot, t, x)$ is a contraction uniformly with respect to $t \in [0,T]$ and $x \in Y^*$. Moreover, for fixed $\mathcal{Z}$ also $\Phi(\mathcal{Z}, \cdot, \cdot)$ is continuous from $[0,T] \times Y^*$ to $L^p(\Omega; C([0,T]; Y^*))$.
2. The directional derivative $\nabla_{\mathcal{Z}}\Phi(\mathcal{Z}, t, x; N)$ in the direction $N \in L^p(\Omega; C([0,T]; Y^*))$ is the process

$$\begin{aligned}\nabla_{\mathcal{Z}}\Phi(\mathcal{Z}, t, x; N)_\tau &= \int_t^\tau e^{(\tau-\eta)\mathcal{A}^*}\nabla_x\beta_\nu^g(\eta, \mathcal{Z}_\eta)N_\eta d\eta \\ &+ \int_t^\tau \nabla_x\left(e^{(\tau-\eta)\mathcal{A}^*}\sigma_\nu^g(\eta, \mathcal{Z}_\eta)\right)N_\eta dW_\eta \ \tau \in [t,T] \\ \nabla_{\mathcal{Z}}\Phi(\mathcal{Z}, t, x; N)_\tau &= 0 \ \tau \in [0,t)\end{aligned}$$

moreover the mappings $(\mathcal{Z}, t, x) \mapsto \nabla_{\mathcal{Z}}\Phi(\mathcal{Z}, t, x; N)$ and $N \mapsto \nabla_{\mathcal{Z}}\Phi(\mathcal{Z}, t, x; N)$ are continuous.
3. The directional derivative $\nabla_x\Phi(\mathcal{Z}, t, x; h)$ in the direction $h \in Y^*$ is the process given by

$$\nabla_x\Phi(\mathcal{Z}, t, x; h)_\tau = \begin{cases} e^{(\tau-t)\mathcal{A}^*}h, & \tau \in [t,T] \\ h, & \tau \in [0,t), \end{cases}$$

and the mappings $(\mathcal{Z}, t, x) \mapsto \nabla_x\Phi(\mathcal{Z}, t, x; h)$ and $h \mapsto \nabla_x\Phi(\mathcal{Z}, t, x; h)$ are continuous.

to complete the proof we notice that (ii) is a consequence of (i) and the fact that $|\nabla_{\mathcal{Z}}\Phi|$ is uniformly bounded by a constant < 1 by the contraction property of Φ.

□

4 The forward-backward system

In this section we are interested in studying an equation like (4.4) with $\mathcal{H}$ depending on the generalized Feller process $\mathcal{Z}_t$ wich is the solution of (2.2) with initial condition ζ at the initial time t.

We can now consider the forward-backward system given by

$$\begin{cases} d\mathcal{Z}_s &= \mathcal{A}^*\mathcal{Z}_s ds + \beta_\nu^g(s,\mathcal{Z}_s)ds + \sigma_\nu^g(s,\mathcal{Z}_s)dW_s, \quad s \in [t,T] \\ \mathcal{Z}_t &= \zeta \\ dp_s &= -\mathcal{H}(s,\mathcal{Z}_s,q_s)ds + q_s dW_s, \quad s \in [t,T] \\ p_T &= G(\mathcal{Z}_T^u) \end{cases} \tag{4.1}$$

where $\mathcal{H} : [0,T] \times Y^* \times Y^{**} \longrightarrow \mathbb{R}$ is the Hamiltonian function defined as

$$\mathcal{H}(t,\mathcal{Z},\xi) = \inf_{u\in\mathcal{U}} \left[F^g(t,\mathcal{Z},u) + \xi R^g(t,\mathcal{Z},u)\right]. \tag{4.2}$$

We are interested in studying the existence and uniqueness of a solution of this forward-backward system and its regular dependence on the initial value ζ. The solution will be denoted by $(\mathcal{Z}_t, p_t, q_t)$. We will sometime write $p_s = p(s,t,\zeta)$ when we want to emphasize the dependence of p from the initial value ζ at time t.

If we define

$$v(t,\zeta) := p(t,t,\zeta) \tag{4.3}$$

where p is the solution to the previous backward SDE, we can prove that $J^g(t,\mathcal{Z}^{\hat{u}},\hat{u}) = v(t,\mathcal{Z}_t)$. Our first step in this direction to prove that, when $v(t,\zeta)$ is differentiable with respect to ζ, the identification $q_t = \nabla v(t,\mathcal{Z}_t)\sigma_\nu^g(t,\mathcal{Z}_t)$ holds.

In the following part, we will thus make use of the additional Hypothesis **??** on the Banach space Y^* in order to be able to use Malliavin calculus techniques that will allow us to follow what has been done in [9], even though we are in a slightly more general setting.

In order to prove the identification $q_t = \nabla v(t,\mathcal{Z}_t)\sigma^g(t,\mathcal{Z}_t)$, $p(t,\zeta)$ has to be differentiable with respect to ζ, and this can be proved by requiring that the Hamiltonian $\mathcal{H}$ is differentiable with respect to ζ and by using the fact that the process $\mathcal{Z}_t^u$ is differentiable with respect to the initial condition ζ.

We are thus going to study the dependence between of the pair (p_t,q_t) on $\zeta \in Y^*$, where (p_t,q_t) is a pair of processes wich solves the BSDE of the form

$$p_t = G^g(X_T) + \int_t^T \mathcal{H}(s,\mathcal{Z}_s^u,q_s)ds - \int_t^T q_t dW_t, \ t \in [0,T]. \tag{4.4}$$

We consider the forward-backward system (4.1). In the previous section we already studied existence and uniqueness reesults for the forward equation (3.1) and the differentiable dependence of this solution on the initial condition ζ. We are thus going to study the dependence of the pair (p_t, q_t) on $\zeta \in Y^*$.

We are now ready prove a crucial representation Theorem. This result, in the case where σ^g does not depend on X, can be proved without having to recur to the Malliavin derivative as done in [12]. In the particular case of a constant σ^g, we would not need Hypothesis **??** and thus we could have a higher generality.

Proposition 4.1. *Assume that Hypothesis 3.3 hold. Then for almost all s, τ such that $t \le s \le \tau \le T$ we have that*

$$D_s \mathcal{Z}(\tau, t, \zeta) = \nabla_\zeta \mathcal{Z}(\tau, s, \mathcal{Z}(s, t, \zeta)) \sigma^g_\nu(s, \mathcal{Z}(s, t, \zeta)),\ \mathbb{P} - a.s. \tag{4.5}$$

moreover $D_s \mathcal{Z}(T, t, \zeta) = \nabla_\zeta \mathcal{Z}(T, s, \mathcal{Z}(s, t, \zeta)) \sigma^g_\nu(s, \mathcal{Z}(s, t, \zeta))$, $\mathbb{P} - a.s.$ for almost any s.

Proof. Thanks to Proposition 3.4, for every $s \in [0, T]$ and every direction $h \in Y^*$, the directional derivative process $\nabla_\zeta \mathcal{Z}(\tau, s, \zeta)h$, $\tau \in [s, T]$ solves $\mathbb{P}$-a.s. the equation

$$\begin{aligned}
\nabla_\zeta \mathcal{Z}(\tau, t, \zeta)h &= e^{(\tau - t)\mathcal{A}^*} h + \int_t^\tau e^{(\tau - s)\mathcal{A}^*} \nabla_\zeta \beta^g_\nu(s, \mathcal{Z}(s, t, \zeta)) \nabla_\zeta \mathcal{Z}(s, t, \zeta) h ds \\
&\quad + \int_t^\tau \nabla_\zeta \left(e^{(\tau - s)\mathcal{A}^*} \sigma^g_\nu(s, \mathcal{Z}(s, t, \zeta)) \right) \nabla_\zeta \mathcal{Z}(s, t, \zeta) h dW_s,\ \tau \in [t, T] \\
\nabla_\zeta \mathcal{Z}(\tau, t, \zeta)h &= h,\ \tau \in [0, t)
\end{aligned}$$

Given $v \in Y^*$ and $t \in [0, s]$ we can replace ζ by $\mathcal{Z}(s, t, \zeta)$ and h by $\sigma^g_\nu(s, \mathcal{Z}(s, t, \zeta))v$ in the previous equation, since $\mathcal{Z}(s, t, \zeta)$ is $\mathcal{F}_s$ measurable. Note now that,

$$\mathcal{Z}(\eta, s, \mathcal{Z}(s, t, \zeta)) = \mathcal{Z}(\eta, t, \zeta),\ \mathbb{P} - a.s.$$

for $\eta \in [s, T]$ as a consequence of the uniqueness of the solution of (3.1), and we obtain that, $\mathbb{P}$-a.s.,

$$\begin{aligned}
&\nabla_\zeta \mathcal{Z}(\tau, s\mathcal{Z}(s, t, \zeta)) \sigma^g_\nu(s, \mathcal{Z}(s, t, \zeta))v = e^{(\tau - s)\mathcal{A}^*} \sigma^g_\nu(s, \mathcal{Z}(s, t, \zeta))v \\
&\quad + \int_s^\tau e^{(\tau - \eta)\mathcal{A}^*} \nabla_\zeta \beta^g_\nu(\eta, \mathcal{Z}(\eta, t, \zeta)) \nabla_\zeta \mathcal{Z}(\eta, s, \mathcal{Z}(s, t, \zeta)) \sigma^g(s, \mathcal{Z}(s, t, \zeta)) v d\eta \\
&\quad + \int_s^\tau \nabla_\zeta (e^{(\tau - \eta)\mathcal{A}^*} \sigma^g_\nu(\eta, \mathcal{Z}(\eta, t, \zeta))) \nabla_\zeta \mathcal{Z}(\eta, s, \mathcal{Z}(s, t, \zeta)) \sigma^g(s, \mathcal{Z}(s, t, \zeta)) v dW_\eta
\end{aligned}$$

for $\tau \in [s, T]$. This shows that the process

$$\{\nabla_\zeta \mathcal{Z}(\tau, t, \mathcal{Z}(s, t, \zeta,)) \sigma^g(s, \mathcal{Z}(s, t, \zeta))v\}_{t \le s \le \tau \le T}$$

is a solution of the equation

$$Q_{s,\tau} = e^{(\tau - s)\mathcal{A}^*} \sigma^g_\nu(s, \mathcal{Z}_s)v + \int_s^\tau e^{(\tau - \eta)\mathcal{A}^*} \nabla_\zeta \beta^g_\nu(\eta, \mathcal{Z}_\eta) Q_{s,\eta} d\eta$$

$$+\int_s^\tau \nabla_\zeta (e^{(\tau-\eta)\mathcal{A}^*}\sigma_\nu^g(\eta,\mathcal{Z}_\eta))Q_{s,\eta}dW_\eta$$

where $Q_{s,\tau} := D_s\mathcal{Z}_\tau v$. Thus the thesis follows from the uniqueness property, as proved in [9] Proposition 3.5.

To prove the last assertion we take a sequence $\tau_n \uparrow T$ such that (4.5) holds for every τ_n, and we let $n \to \infty$. The result follows from the regularity properties of $D\mathcal{Z}$ and $\nabla_\zeta \mathcal{Z}$, as well as the closedness of the operator D. □

In this framework, exploiting what has been done in [11] and [13] we can have the same kind of result as in Proposition 5.6 of [9], and have that

Proposition 4.2. *Assume Hypothesis 3.3 - 4.3 Then for a.a.* s, τ *such that* $t \le s \le \tau \le T$ *we have that*

$$D_s p(\tau,t,\zeta) = \nabla_\zeta p(\tau,s,\mathcal{Z}(s,t,\zeta))\sigma_\nu^g(s,\mathcal{Z}(s,t,\zeta)) \ \mathbb{P}-a.s. \tag{4.6}$$

$$D_s q(\tau,t,\zeta) = \nabla_\zeta q(\tau,s,\mathcal{Z}(s,t,\zeta))\sigma_\nu^g(s,\mathcal{Z}(s,t,\zeta)) \ \mathbb{P}-a.s. \tag{4.7}$$

Moreover, for a.a. $s \in [t,T]$,

$$q(s,t,\zeta) = \nabla_\zeta p(s,s,\mathcal{Z}(s,t,\zeta))\sigma_\nu^g(s,\mathcal{Z}(s,t,\zeta)) \ \mathbb{P}-a.s. \tag{4.8}$$

Proof. The proof of this Theorem follows the proof of Proposition 5.6 of [9]. The main difference between our work and [9] is the spaces at play. Having proved proposition (4.1), we have that the extension of Proposition 5.6 of [9] to our setting is trivial. □

Having this result, we can state our optimal value Theorem. The proof of such theorem is going to follow the same approach of the ones proved in [8, 9] and [12], but in our case we are considering a UMD Banach space with a non constant volatility, whereas in [9] a Hilbert space is considered and in [12] the volatility is constant. Notice also that, if the volatility σ^g was a linear functional not depending on ζ, we could have dropped the UMD hypothesis and obtained the optimality by simply following the work done in [12].

We also have to make some additional assumptions on $\mathcal{H}$ that will guarantee that

$$\mathbb{E}\left[\int_t^T |\mathcal{H}(s,0,0)|^s \, ds\right] < \infty.$$

Notice also that for the pair process (p_t, q_t) the following a priori estimate holds (see [12] and [9] Proposition 4.3):

$$\mathbb{E}\left[\sup_{\tau\in[t,T]} |Y_\tau|^2\right] + \mathbb{E}\left[\int_t^T \|q_\eta\|^2_{Y^*{}^*} d\eta\right] \le c\mathbb{E}\left[\int_t^T |\mathcal{H}(\eta,0,0)|^2 d\eta\right] + c\mathbb{E}\left[|G^g(\mathcal{Z}_T)|^2\right]$$

Hypothesis 4.3. *Let us assume that*

1) There exists $L > 0$ such that

$$|\mathcal{H}(t, \mathcal{Z}, \xi_1) - \mathcal{H}(t, \mathcal{Z}, \xi_2)| \leq L\|\xi_1 - \xi_2\|_{Y^{**}}$$

for every $t \in [0, T]$, $\zeta \in Y^*$ and $\xi_1, \xi_2 \in Y^{**}$

2) For every $t \in [0, T]$ we have $\mathcal{H}(t, \cdot, \cdot) \in \mathcal{G}^{1,1}(Y^* \times Y^{**})$

3) There exist $L > 0$ and $m \geq 0$ such that

$$|\nabla_{\mathcal{Z}} \mathcal{H}(t, \mathcal{Z}, \xi)h| \leq L\|h\|_{Y^*}(1 + \|\mathcal{Z}\|_{Y^*})^m(1 + \|\xi\|_{Y^{**}})$$

for every $t \in [0, T]$, $\zeta, h \in Y^*$ and $\xi \in Y^{**}$

4) $G^g \in \mathcal{G}^1(Y^*)$ and there exists $L > 0$ such that, for every $x, y \in Y^*$

$$|G^g(x) - G^g(y)| \leq L\|x - y\|_{Y^*}$$

Proposition 4.4. *Assume that Hypothesis 3.3 and 4.3 hold true. Then* (4.4) *admits a unique solution (p, q) such that the map $\zeta \mapsto (p(\cdot, \zeta), q(\cdot, \zeta))$ belongs to $\mathcal{G}^1(L^\rho(\Omega; C([0, T]; Y^*)), \mathcal{K}_{cont}([0, T]))$ for $\rho = p(m+1)(m+2)$, where $\mathcal{K}_{cont}([0, T])$ is the space of predictable processes (p, q) taking values in $\mathbb{R} \times H^*$ such that p has continuous paths and*

$$\mathbb{E}\left[\sup_{\tau \in [0,T]} |Y_\tau^2|\right] + \mathbb{E}\left[\int_t^T \|q_s\|_{H^*}^2 ds\right] < \infty.$$

Moreover, for every $p \geq 2$.

$$\left(\mathbb{E}\left[\sup_{t \in [0,T]} |\nabla_{\mathcal{Z}} p(t, \zeta)h|^p\right]\right)^{1/p} \leq C\|h\|_{Y^*}\left(1 + \|\zeta\|_{Y^*}^{(m+1)^2}\right)$$

Proof. For the proof check either [12] Proposition 4.2 or [9]. □

Corollary 4.5. *Assume that Hypothesis 3.3 and 4.3 hold true. Then the function $v(t, \zeta) := p(t, \zeta)$ is continuous and for every $t \in [0, T]$, $v(t, \cdot)$ belongs to $\mathcal{G}^1(Y^*, \mathbb{R})$ and there exists $C > 0$ such that*

$$|\nabla_{\mathcal{Z}} v(t, x)h| \leq C\|h\|_{Y^*}(1 + \|\zeta\|_{Y^*}^{(m+1)^2})$$

4.1 The optimal control problem

We are now interested in solving the lifted optimal control problem, where the process $\mathcal{Z}$ follows the controlled dynamics given by

$$\begin{cases} d\mathcal{Z}^u_\tau = & \mathcal{A}^*\mathcal{Z}^u_\tau d\tau + \beta^g_\nu(\tau,\mathcal{Z}^u_\tau)d\tau + \sigma^g_\nu(\tau,\mathcal{Z}^u_\tau)R^g(\tau,\mathcal{Z}^u_\tau,u_\tau)d\tau \\ & +\sigma^g(\tau,\mathcal{Z}^u_\tau)dW_\tau \\ \mathcal{Z}_0 = & \zeta \end{cases} \tag{4.9}$$

and our performance functional is (2.5) In order for our results to hold we also have to add some Hypothesis on R^g:

Hypothesis 4.6. *$R : [0,T]\times\mathbb{R}\times\mathcal{U} \longrightarrow \mathbb{R}$ is measurable and $\|R(\tau,x,u)\|_{\mathbb{R}} \le K_R$ for a suitable positive constant $K_R > 0$ and every $\tau \in [0,T]$, $x \in \mathbb{R}$, $u \in \mathcal{U}$.*

We now define the, possibly empty, set

$$\Gamma(t,\mathcal{Z},\xi) = \{u \in \mathcal{U} : F^g(t,\mathcal{Z},u) + \xi R^g(t,\mathcal{Z},u) = \mathcal{H}(t,\mathcal{Z},\xi)\} \tag{4.10}$$

Thanks to the Filippov theorem (see [4]) if Γ is non empty, there exists a Borel measurable map $\Gamma_0 : [0,T]\times Y^*\times H^* \longrightarrow \mathcal{U}$ such that, for $t\in[0,T]$, $\zeta \in Y^*$ and $\xi \in Y^{**}$, $\Gamma_0(t,\mathcal{Z},\xi) \in \Gamma(t,\mathcal{Z},\xi)$.

Following now what has been done in [9, 12] we can thus present the following Theorem

Theorem 4.7. *Assume that Hypothesis 3.3 - 4.3 and 4.6 hold true. For all admissible controls we have that*

$$J^g(t,x,\mathcal{Z}^u) \ge v(t,x)$$

and the equality holds true if and only if

$$u_\tau \in \Gamma(\tau,\mathcal{Z}^u_\tau,\nabla v(\tau,\mathcal{Z}^u_\tau)\sigma^g_\nu(\tau,\mathcal{Z}^u_\tau))\ \mathbb{P}-a.s.\ \textit{for a.a.}\ \tau\in[t,T] \tag{4.11}$$

Moreover, let us denote by Γ_0 a measurable selection of Γ, then a control satisfying the feedback law

$$u_\tau = \Gamma_0(\tau,\mathcal{Z}^u_\tau,\nabla v(\tau,\mathcal{Z}^u_\tau)\sigma^g_\nu(\tau,\mathcal{Z}^u_\tau))\ \mathbb{P}-a.s.\ \textit{for a.a.}\ \ \tau\in[t,T] \tag{4.12}$$

is optimal. The closed loop equation

$$\begin{cases} \widetilde{\mathcal{Z}}_\tau = & [\mathcal{A}^*\widetilde{\mathcal{Z}}_\tau + \beta^g_\nu(\tau,\widetilde{\mathcal{Z}}_\tau) + \sigma^g_\nu(\tau,\widetilde{\mathcal{Z}}_\tau)R^g(\tau,\widetilde{\mathcal{Z}}_\tau,\Gamma_0(\nabla v(\tau,\widetilde{\mathcal{Z}}_\tau)\sigma^g_\nu(\tau,\widetilde{\mathcal{Z}}_\tau)))]d\tau \\ & +\sigma^g_\nu(\tau,\widetilde{\mathcal{Z}}_\tau)dW_\tau,\quad \tau\in[t,T] \\ \widetilde{\mathcal{Z}}_t = & \zeta \end{cases} \tag{4.13}$$

admits a weak solution which is unique in law, and the corresponding pair $(u, \widetilde{\mathcal{Z}}^u)$ is optimal.

Proof. The proof follow what is done in [12] Theorem 5.7. For completeness we are going to present the main ideas.

Using the definition of $v(t,x)$ and Proposition 4.2 we can rewrite $v(t,x)$ as

$$\begin{aligned} v(t,x) =& J(t,x,u) + \int_t^T [\mathcal{H}(\tau, \mathcal{Z}^u_\tau, \nabla v(\tau, \mathcal{Z}^u_\tau)\sigma^g_\nu(\tau, \mathcal{Z}^u_\tau)) \\ & -\nabla v(\tau, \mathcal{Z}^u_\tau)\sigma^g_\nu(\tau, \mathcal{Z}^u_\tau))R^g(\tau, \mathcal{Z}^u_t, u_\tau) - F^g(\tau, \mathcal{Z}^u_\tau, u_\tau)]\, d\tau \end{aligned}$$

The closed loop equation can be solved in the weak sense via a Girsanov change of measure. Namely we fix an arbitrary cylindrical Wiener process $(W_\tau)_{\tau\geq 0}$ on a probability space $(\Omega, \mathcal{F}, P)$ with values in H and we define a $(\hat{W}_\tau)_{\tau\geq 0}$ that is a Wiener process under a new probability $\hat{\mathbb{P}}$ on Ω.

Then we can find a unique mild solution to this new process relative to the probability $\hat{\mathbb{P}}$ and the wiener process $\hat{W}_\tau$, which implies that also the closed loop equation (4.13) always admits a solution in the weak sense.

So if Γ is nonempty, by the Filippov theorem, a measurable selection Γ_0 of Γ exists and it is possible to obtain the optimal control. □

4.2 The HJB equation

If we now formally define

$$\begin{aligned} \mathcal{L}_t[f](x) :=& \frac{1}{2}\text{Trace}(G^g(t,x)G^g(t,x)^*\nabla^2 f(x)) \\ & + \langle \mathcal{A}^* x, \nabla f(x)\rangle_{Y^*\times Y^{**}} + \langle \beta^g_\nu(t,x), \nabla f(x)\rangle_{Y^*\times Y^{**}} \end{aligned}$$

we can consider the Hamilton-Jacobi-Bellman equation associated with out control problem, which is given by

$$\begin{cases} \frac{\partial v}{\partial t}(t,z) & = -\mathcal{L}_t v(t,z) - \mathcal{H}(t,z,\nabla v(t,z)\sigma^g_\nu(t,z)) \\ v(T,z) & = G^g(z) \end{cases} \tag{4.14}$$

We are interested in finding mild solutions to the previous equation, which we are going to define soon. This problem has been solved both in [9] (for a Hilbert space) in the case of a general σ^g_ν, and in [12] (for a Banach space) in the case of a constant σ^g_ν.

Let now $\mathcal{Z}(\tau, t, \zeta)$ be a solution to (3.1), with $\mathcal{A}^*$, β^g_ν and σ^g_ν that satisfies Hypothesis 3.3 - 4.3. We recall that this solution is a Y^*-valued Markov proces. We thus can define the transition semigroup on ontinuous and bounded functions $\varphi : Y^* \longrightarrow \mathbb{R}$ as

$$P_{t,\tau}[\varphi](z) = \mathbb{E}[\varphi(\mathcal{Z}(\tau, t, z))]$$

Moreover we have that this semigroup is also well defined on continuous functions $\varphi : Y^* \longrightarrow \mathbb{R}$ with polynomial growth with respect to z.

We look thus for mild solutions of (4.14), that is a function $v(t,z)$ such that

$$v(t,z) = P_{t,T}[G^g](z) + \int_t^T P_{t,\tau}[\mathcal{H}(\tau,\cdot,\nabla v(\tau,\cdot)\sigma_{\mathcal{V}}^g(t,\cdot))](z)d\tau, \tag{4.15}$$

for $t \in [0,T], z \in Y^*$.

Definition 4.8. A function $v : [0,T] \times Y^* \longrightarrow \mathbb{R}$ is a mild solution of the Hamilton-Jacobi-Bellman equation (4.14) if:

- v is continuous, for every $t \in [0,T]$ $v(t,\cdot) \in \mathcal{G}^1(Y^*)$, and the map $(t,z) \mapsto v(t,z)$ is measurable from $[0,T] \times Y^*$ with values in Y^**
- There exists $C > 0$ such that $|v(t,z)| \leq C(1+\|z\|_{Y^*}^j)$ and $|\nabla_z v(t,z)h| \leq C\|h\|_{Y^*}(1+\|z\|_{Y^*}^k)$ for every $t \in [0,T]$, for $z,h \in Y^*$ and for some positive integers j and k.
- equality (4.15) holds

In order to prove that there exists a unique solution of (4.14) we are also going to need the Forward-Backward system

$$\begin{cases} d\mathcal{Z}_\tau &= \mathcal{A}^*\mathcal{Z}_\tau d\tau + \beta_{\mathcal{V}}^g(\tau,\mathcal{Z}_\tau) + \sigma_{\mathcal{V}}^g(\tau,\mathcal{Z}_\tau)dW_\tau, \qquad \tau \in [t,T] \\ dp_t &= -\mathcal{H}(\tau,\mathcal{Z}_\tau,q_\tau)d\tau + q_\tau dW_\tau \qquad \tau \in [t,T] \\ \mathcal{Z}_t &= \zeta \\ p_T &= G^g(\mathcal{Z}_T) \end{cases} \tag{4.16}$$

Theorem 4.9. *Assume that Hypothesis 3.3 hold true and let G^g and $\mathcal{H}$ satisfy Hypothesis 4.3. Then there exists a unique mild solution of the Hamilton-Jacobi-Bellman equation* (4.14) *given by*

$$v(t,z) = p(t,t,z), \tag{4.17}$$

where $(\mathcal{Z},p,q)$ is the solution of the forward-backward system (4.16).

Proof. The proof is based on [12] Theorem 6.2 and [9] Theorem 6.2, and adapted to this context. □

5 A rough cashflow dynamics example

We consider a cash flow dynamics for our forward process, exposed to a compsuntion rate c_t given by

$$X_t^c = x + \int_0^t K(t-\tau)\mu(\tau,X_\tau^c)d\tau + \int_0^t K(t-\tau)\sigma(\tau,X_\tau^c)(-c_\tau d\tau + dW_\tau) \tag{5.1}$$

Where we supposed that $X_0 = x \in \mathbb{R}_+$, and μ, σ satisfy Hypothesis 3.3. Here $K(t) = \int_0^\infty e^{-tx}\nu(dx)$ is the Laplace transform of a finite signed Borel measure.

Notice that in this case we can express $K(t) = \langle g, S_t^*\nu\rangle$ where $g = 1$, $S_t^*\zeta = e^{t\cdot}\zeta$ and ν belongs to $\mathcal{M}([0,\infty])$, namely the space of the finite, signed Borel measures on $[0,\infty]$. In this case the pairing $\langle\cdot,\cdot\rangle$ is given by

$$\langle g, \nu\rangle = \int_0^\infty g(x)\nu(dx)$$

This Hypothesis guarantees that we are in the framework presented in Remark 2.10.

Our goal is to maximize the performance functional given by

$$J(t,x,c) = \mathbb{E}\left[\int_t^T F(\tau, X_\tau^c, c_\tau)d\tau + G(X_T^c)\right] \tag{5.2}$$

for some function G that satisfies Hypothesis 4.3. In this case we have that the problem can be reformulated in Y^* as having forward dynamics $\mathcal{Z}_\tau^c$ are given by

$$\mathcal{Z}_\tau^c = \zeta + \int_t^\tau \mathcal{A}^*\mathcal{Z}_s^c + \mu_\nu^g(s,\mathcal{Z}_s^c)ds + \sigma_\nu^g(s,\mathcal{Z}_s^c)(-c_s ds + dW_s) \tag{5.3}$$

with ζ such that $x \equiv \langle g, \mathcal{S}_t^*\zeta\rangle$ for all $t \in [0,T]$, and where our goal is to maximize

$$J^g(t,\zeta,c) = \mathbb{E}\left[\int_t^T F^g(\tau,\mathcal{Z}_\tau^c, c_\tau)d\tau + G^g(\mathcal{Z}_T^c)\right] \tag{5.4}$$

In this case our Hamiltonian function (4.2) is given by

$$\mathcal{H}(t,\mathcal{Z},\xi) = \inf_{c\in\mathcal{U}}[F^g(t,\mathcal{Z},c) - \xi c] \tag{5.5}$$

and the forward-backward system is given by

$$\begin{cases} d\mathcal{Z}_\tau &= \mathcal{A}^*\mathcal{Z}_\tau d\tau + \mu_\nu^g(s,\mathcal{Z}_\tau)d\tau + \sigma_\nu(\tau,\mathcal{Z}_\tau)dW_\tau, \quad \tau\in[t,T] \\ \mathcal{Z}_0 &= \zeta \\ dp_\tau &= -\mathcal{H}(\tau,\mathcal{Z}_\tau^c, q_\tau)d\tau + q_\tau dW_\tau, \quad \tau\in[t,T] \\ p_T &= G^g(\mathcal{Z}_T^c) \end{cases} \tag{5.6}$$

In particular, using (5.5), we have that

$$p_t = G^g(\mathcal{Z}_T^c) - \int_t^T \inf_{c\in\mathcal{U}}\left(F^g(s,\mathcal{Z}_s^c,c_s) - q_s c_s\right)ds + q_s dW_s \;\; s\in[t,T].$$

We thus get that the set Γ defined in (4.10) is

$$\Gamma(t,\mathcal{Z},\xi) = \{c\in\mathcal{U} : F^g(t,\mathcal{Z},c) - \xi u = \mathcal{H}(t,\mathcal{Z},c)\} \tag{5.7}$$

and thus the optimal u_τ can be characterized by Theorem 4.7 as

$$c_\tau = \Gamma_0(\tau, \mathcal{Z}^c_\tau, \nabla v(\tau, \mathcal{Z}^c_\tau)\sigma^g_\nu(\tau, \mathcal{Z}^c_\tau)) \;\; \mathbb{P} - a.s. \text{ for a.a. } \tau \in [0, T] \tag{5.8}$$

for a certain function Γ_0 such that $\Gamma_0(t, \mathcal{Z}, \xi) \in \Gamma(t, \mathcal{Z}, \xi)$. In this case the HJB equations (4.14) become

$$\begin{cases} \frac{\partial v}{\partial t}(t, \zeta) & = -\mathcal{L}_t v(t, \mathcal{Z}(t, t, \zeta)) \\ & \quad - \inf_{c \in \mathcal{U}}[F^g(t, \mathcal{Z}(t, t, \zeta), c_t) - \nabla v(t, \mathcal{Z}(t, t, \zeta))\sigma^g_\nu(t, \mathcal{Z}(t, t, \zeta))c_t] \\ v(T, \zeta) & = G^g(\zeta) \end{cases} \tag{5.9}$$

Were, we remind, $\mathcal{Z}(\tau, t, \zeta) = \mathcal{Z}(\tau)$, $\tau \in [t, T]$ and $\mathcal{Z}(t) = \zeta$. Now thanks to Theorem 4.9 we have that

Theorem 5.1. *Equation* (5.9) *has a unique mild solution v, and if the cost is defined as in* (5.4)*, then for all admissible couples $(c, \mathcal{Z}^c)$ we have that $J(t, \mathcal{Z}_t, c) \geq v(t, \mathcal{Z}_t)$, and the equality holds if and only if* (4.11) *holds. If* (5.8) *holds, there exists an optimal couple $(c, \mathcal{Z}^c)$.*

Remark 5.2. This give us a complete implicit characterization of the optimal control for our lifted problem (5.4) and thus also for (5.2): in fact in our case we have that $X_t = \langle 1, \mathcal{Z}_t \rangle$. Thanks to this we are able to find the optimal process $X^{\hat{u}}$. Moreover we notice that, the optimal control for the lifted problem and the optimal control for the original problem coincide, which allows us to retrieve the optimal pair $(\hat{u}, X^{\hat{u}})$ for the optimal control problem (5.2). Lastly we notice that the HJB equations (5.9), gives us the optimal value $J^g(t, x, \hat{u})$, which in turn gives us the optimal value of $J(t, x, \hat{u})$, thanks to (2.4).

Acknowledgment

The research leading to these results is within the project STORM: Stochastics for Time-Space Risk Models, receiving founding from the Research Council of Norway (RCN). Project number: 274410.

References

[1] E. Abi Jaber, E. Miller, and H. Pham. Linear-Quadratic control for a class of stochastic Volterra equations: solvability and approximation. 2019.

[2] N. Agram and B. Øksendal. Malliavin Calculus and Optimal Control of Stochastic Volterra Equations. *Journal of Optimization Theory and Applications*, 167:1070–1094, 2015.

[3] N. Agram, B. Øksendal, and S. Yakhlef. Optimal Control of Forward-Backward Stochastic Volterra Equations. *Non-linear Partial Differential Equations, Mathematical Physics, and Stochastic Analysis: The Helge Holden Anniversary Volume*, pages 3–36, 2009.

[4] J.-P. Aubin and H. Frankowska. *Set-valued analysis*. 1990.

[5] C. Cuchiero and G. Di Nunno. Notes - markovian lifts. 2019.

[6] C. Cuchiero and J. Teichman. Markovian lifts of positive semidefinite affine Volterra-type processes. *Decisions in Economics and Finance*, 42:407–448, 2019.

[7] C. Cuchiero and J. Teichman. Generalized Feller processes and Markovian lifts of stochastic Volterra processes: the affine case. *Journal of Evolution Equations*, pages 1–48, 2020.

[8] G. Fabbri, F. Gozzi, and A. Swiech. *Stochastic Optimal Control in Infinite Dimension Dynamic Programming and HJB Equations*. Springer, 2017.

[9] M. Fuhrman and G. Tessitore. Nonlinear Kolmogorov equations in infinite dimensional spaces: the backward stochastic differential equations approach and applications to optimal control. *The Annals of Probability*, 30:1397–1465, 2002.

[10] T. Hytönen, J. van Neerven, M. Veraar, and L. Weis. *Analysis in Banach Spaces*, volume I: Martingales and Littlewood-Paley Theory. Springer, Cham, 2016.

[11] J. Maas and J. Van Neerven. A Clark-Ocone formula in UMD Banach spaces. *Electronic communications in probability*, 2008.

[12] F. Masiero. Stochastic optimal control problems and parabolic equations in Banach spaces. *SIAM Journal on Control and Optimization*, 47:251–300, 2008.

[13] M. Pronk and M. Veraar. Tools for Malliavin calculus in UMD Banach spaces. *Potential Analysis*, 40:307–344, 2014.

[14] P. Protter. Volterra Equations Driven by Semimartingales. *The Annals of Probability*, 13:519–530, 1993.

[15] J. Van Neerven and M. Veraar. On the stochastic Fubini theorem in infinite dimensions. 2005.

[16] J. Van Neerven, M. Veraar, and L. Weis. Stochastic integration in UMD Banach spaces. *The Annals of Probability*, 35, 2007.

[17] J. Yong. Backward Stochastic Volterra Integral Equations and some Related Problems. *Stochastic Processes and their Applications*, 116:770–795, 2006.

[18] J. Yong. Well-Posedness and Regularity of Backward Stochastic Volterra Integral Equations. *Probability Theory and Related Fields*, 142:21–77, 2007.