Singular control of stochastic Volterra integral equations

Nacira Agram[1], Saloua Labed[2], Bernt Øksendal[3] and Samia Yakhlef[2]

14 April 2021

Abstract

This paper deals with optimal combined singular and regular controls for stochastic Volterra integral equations, where the solution $X^{u,\xi}(t) = X(t)$ is given by

$$\begin{aligned} X(t) &= \phi(t) + \int_0^t b\left(t,s,X(s),u(s)\right) ds + \int_0^t \sigma\left(t,s,X(s),u(s)\right) dB(s) \\ &\quad + \int_0^t h\left(t,s\right) d\xi(s). \end{aligned}$$

Here $dB(s)$ denotes the Brownian motion Itô type differential and ξ denotes the singular control (singular in time t with respect to Lebesgue measure) and u denotes the regular control (absolutely continuous with respect to Lebesgue measure).
Such systems may for example be used to model harvesting of populations with memory, where $X(t)$ represents the population density at time t, and the singular control process ξ represents the harvesting effort rate. The total income from the harvesting is represented by

$$J(u,\xi) = \mathbb{E}\Big[\int_0^T f_0(t,X(t),u(t))dt + \int_0^T f_1(t,X(t))d\xi(t) + g(X(T))\Big],$$

for given functions f_0, f_1 and g, where $T > 0$ is a constant denoting the terminal time of the harvesting. Note that it is important to allow the controls to be singular, because in some cases the optimal controls are of this type.

Using Hida-Malliavin calculus, we prove sufficient conditions and necessary conditions of optimality of controls. As a consequence, we obtain a new type of backward stochastic Volterra integral equations with singular drift.

Finally, to illustrate our results, we apply them to discuss optimal harvesting problems with possibly density dependent prices.

[1]Department of Mathematics, Linnaeus University (LNU), Växjö, Sweden. Email: nacira.agram@lnu.se.

[2]University Mohamed Khider of Biskra, Algeria. Email: s.labed@univ-biskra.dz, s.yakhlef@univ-biskra.dz

[3]Department of Mathematics, University of Oslo, P.O. Box 1053 Blindern, N–0316 Oslo, Norway. Email: oksendal@math.uio.no.

MSC(2010): 60H05, 60H20, 60J75, 93E20, 91G80,91B70.

Keywords: Stochastic maximum principle; stochastic Volterra integral equation; singular control; backward stochastic Volterra integral equation; Hida-Malliavin calculus.

1 Introduction

As a motivating example, consider the population of a certain type of fish in a lake, where the density $X(t)$ at time t can be modelled as the solution of the following stochastic Volterra integral equation (SVIE):

$$X(t) = x_0 + \int_0^t b_0(t,s)X(s)ds + \int_0^t \sigma_0(s)X(s)dB(s) - \int_0^t \gamma_0(t,s)d\xi(s),$$

where the coefficients b_0, σ_0 and γ_0 are bounded deterministic functions, and $B(t) = \{B(t,\omega)\}_{t\geq 0, \omega\in\Omega}$ is a Brownian motion defined on a complete probability space $(\Omega, \mathcal{F}, P)$. We associate to this space a natural filtration $\mathbb{F} = \{\mathcal{F}_t\}_{t\geq 0}$ generated by $B(t)$, assumed to satisfy the usual conditions. The process $\xi(t)$ is our control process. It is an $\mathbb{F}$- adapted, nondecreasing left-continuous process representing the harvesting effort. It is called singular, because as a function of time t it may be singular with respect to Lebesgue measure. The constant $\gamma_0 > 0$ is the harvesting efficiency coefficient. It turns out that in some cases the optimal process $\xi(t)$ can be represented as the local time of the solution $X(t)$ at some threshold curve. In this case $\xi(t)$ is increasing only on a set of Lebesgue measure 0.
Volterra equations are commonly used in population growth models, especially when age dependence plays a role. See e.g. Gripenberg *et al* [7]. Moreover, they are important examples of equations with memory.
We assume that the total expected utility from the harvesting is represented by

$$J(\xi) = \mathbb{E}[\theta X(T) + \int_0^T \log(X(t))d\xi(t)],$$

where $\mathbb{E}$ denotes the expectation with respect to P. The problem is then to maximise $J(\xi)$ over all admissible singular controls ξ. We will return to this example in Section 4.

Control problems for singular Volterra integral equations have been studied by Lin and Yong [12] in the deterministic case. In this paper we study singular control of SVIEs and we present a different approach based on a stochastic version of the Pontryagin maximum principle.
Stochastic control for Volterra integral equations has been studied by Yong [14] and subsequently by Agram *el al* [3], [5] who used the white noise calculus to obtain both sufficient and necessary conditions of optimality. In the latter, smoothness of coefficients is required.
The adjoint processes of our maximum principle satisfy a backward stochastic integral equation of Volterra type and with a singular term coming from the control. In our example one may consider the optimal singular term as the local time of the state process that is keeping it above/below a certain threshold curve. Hence in some cases we can associate this type of equations with reflected backward stochastic Volterra integral equations.

Partial result for existence and uniqueness of backward stochastic Volterra integral equation (BSVIE) in a continuous case can be found in Yong [14], [15], and for a discontinuous case, we refer for example to Agram *el al* [4], [2] where there are also some applications.

The paper is organised as follows: In the next section we give some preliminaries about the generalised Malliavin calculus, called Hida-Malliavin calculus, in the white noise space of Hida of stochastic distributions. Section 3 is addressed to the study of the stochastic maximum principle where both sufficient and necessary conditions of optimality are proved. Finally, in Section 4 we apply the results obtained in section 3 to discuss optimal harvesting problems with possibly density dependent prices.

2 Hida - Malliavin calculus

Let $\mathbb{G} = \{\mathcal{G}_t\}_{t\geq 0}$ be a subfiltration of $\mathbb{F}$, in the sense that $\mathcal{G}_t \subseteq \mathcal{F}_t$, for all $t \geq 0$. The given set $U \subset \mathbb{R}$ is assumed to be convex. The set of admissible controls, i.e. the strategies available to the controller, is given by a subset $\mathcal{A}$ of the càdlàg, U-valued and $\mathbb{G}$-adapted processes. Let $\mathcal{K}$ be the set of all $\mathbb{G}$-adapted processes $\xi(t)$ that are nondecreasing and left continuous with respect to t.
Next we present some preliminaries about the extension of the Malliavin calculus into the stochastic distribution space of Hida, for more details, we refer the reader to Aase *et al* [1], Di Nunno *et al* [11].
The classical Malliavin derivative is only defined on a subspace $\mathbb{D}_{1,2}$ of $\mathbb{L}^2(P)$. However, there are many important random variables in $\mathbb{L}^2(P)$ that do not belong to $\mathbb{D}_{1,2}$. For example, this is the case for the solutions of a backward stochastic differential equations or more generally the BSVIE. This is why the Malliavin derivative was extended to an operator defined on the whole of $\mathbb{L}^2(P)$ and with values in the Hida space $(\mathcal{S})^*$ of stochastic distributions. It was proved by Aase *et al* [1] that one can extend the Malliavin derivative operator D_t from $\mathbb{D}_{1,2}$ to all of $\mathbb{L}^2(\mathcal{F}_T, P)$ in such a way that, also denoting the extended operator by D_t, for all random variable $F \in \mathbb{L}^2(\mathcal{F}_T, P)$, we have

$$D_t F \in (\mathcal{S})^* \text{ and } (t,\omega) \mapsto \mathbb{E}[D_t F|\mathcal{F}_t] \text{ belongs to } \mathbb{L}^2(\lambda \times P), \tag{2.1}$$

where λ is Lebesgue measure on $[0, T]$. We now give a short introduction to Malliavin calculus and its extension to Hida-Malliavin calculus in the white noise setting:

Definition 2.1 **(i)** *Let $F \in \mathbb{L}^2(P)$ and let $\gamma \in \mathbb{L}^2(\mathbb{R})$ be deterministic. Then the directional derivative of F in $(\mathcal{S})^*$ (respectively, in $\mathbb{L}^2(P)$) in the direction γ is defined by*

$$D_\gamma F(\omega) = \lim_{\varepsilon\to 0} \frac{1}{\varepsilon}\big[F(\omega + \varepsilon\gamma) - F(\omega)\big] \tag{2.2}$$

whenever the limit exists in $(\mathcal{S})^*$ (respectively, in $\mathbb{L}^2(P)$).

(ii) *Suppose there exists a function $\psi : \mathbb{R} \mapsto (\mathcal{S})^*$ (respectively, $\psi : \mathbb{R} \mapsto \mathbb{L}^2(P)$) such that*

$$\begin{aligned} &\textstyle\int_{\mathbb{R}} \psi(t)\gamma(t)dt \quad \textit{exists in } (\mathcal{S})^* \textit{ (respectively, in } \mathbb{L}^2(P)) \textit{ and} \\ &D_\gamma F = \textstyle\int_{\mathbb{R}} \psi(t)\gamma(t)dt, \quad \textit{for all } \gamma \in \mathbb{L}^2(\mathbb{R}). \end{aligned} \tag{2.3}$$

Then we say that F is Hida-Malliavin differentiable *in $(\mathcal{S})^*$ (respectively, in $\mathbb{L}^2(P)$) and we write*

$$\psi(t) = D_t F, \quad t \in \mathbb{R}.$$

We call $D_t F$ the Hida-Malliavin derivative at t in $(\mathcal{S})^*$ (respectively, in $\mathbb{L}^2(P)$) *or the* stochastic gradient *of F at t.*

Let $F_1, ..., F_m \in \mathbb{L}^2(P)$ be Hida-Malliavin differentiable in $\mathbb{L}^2(P)$. Suppose that $\varphi \in C^1(\mathbb{R}^m)$, $D_t F_i \in \mathbb{L}^2(P)$, for all $t \in \mathbb{R}$, and $\frac{\partial \varphi}{\partial x_i}(F) D.F_i \in \mathbb{L}^2(\lambda \times P)$ for $i = 1, ..., m$, where $F = (F_1, ..., F_m)$. Then $\varphi(F)$ is Hida-Malliavin differentiable and

$$D_t \varphi(F) = \sum_{i=1}^m \frac{\partial \varphi}{\partial x_i}(F) D_t F_i. \tag{2.4}$$

We have the following *generalized duality formula,* for the Brownian motion:

Proposition 2.2 *Fix $s \in [0, T]$. If $t \mapsto \varphi(t, s, \omega) \in \mathbb{L}^2(\lambda \times P)$ is $\mathbb{F}$-adapted with $\mathbb{E}[\int_0^T \varphi^2(t, s) dt] < \infty$ and $F \in \mathbb{L}^2(\mathcal{F}_T, P)$, then we have*

$$\mathbb{E}[F \int_0^T \varphi(t, s) dB(t)] = \mathbb{E}[\int_0^T \mathbb{E}[D_t F | \mathcal{F}_t] \varphi(t, s) dt]. \tag{2.5}$$

We will need the following:

Lemma 2.3 *Let $t, s, \omega \mapsto G(t, s, \omega) \in \mathbb{L}^2(\lambda \times \lambda \times P)$ and $t, \omega \mapsto p(t) \in \mathbb{L}^2(\lambda \times P)$, then the followings hold:*

1. *The Fubini theorem combined with a change of variables gives*

$$\int_0^T p(t) (\int_0^t G(t, s) ds) dt = \int_0^T (\int_t^T p(s) G(s, t) ds) dt, \tag{2.6}$$

and

$$\int_0^T p(t) (\int_0^t G(t, s) ds) d\xi(t) = \int_0^T (\int_t^T p(s) G(s, t) ds) d\xi(t). \tag{2.7}$$

2. *The generalized duality formula* (2.5) *together with the Fubini theorem, yields*

$$\mathbb{E}[\int_0^T p(t) (\int_0^t G(t, s) dB(s)) dt] = \mathbb{E}[\int_0^T \int_t^T \mathbb{E}[D_t p(s) | \mathcal{F}_t] G(s, t) ds dt]. \tag{2.8}$$

3 Stochastic maximum principles

In this section, we study stochastic maximum principles of stochastic Volterra integral systems under partial information, i.e., the information available to the controller is given by a sub-filtration $\mathbb{G}$. Suppose that the state of our system $X^{u,\xi}(t) = X(t)$ satisfies the following SVIE

$$\begin{aligned} X(t) &= \phi(t) + \int_0^t b(t, s, X(s), u(s)) ds + \int_0^t \sigma(t, s, X(s), u(s)) dB(s) \\ &\quad + \int_0^t h(t, s) d\xi(s), \quad t \in [0, T], \end{aligned} \tag{3.1}$$

where $b(t,s,x,u) = b(t,s,x,u,\omega) : [0,T]^2 \times \mathbb{R} \times U \times \Omega \to \mathbb{R}$, $\sigma(t,s,x,u) = \sigma(t,s,x,u,\omega) : [0,T]^2 \times \mathbb{R} \times U \times \Omega \to \mathbb{R}$.

The *performance functional* has the form

$$J(u,\xi) \quad = \mathbb{E}[\int_0^T f_0(t,X(t),u(t))dt + \int_0^T f_1(t,X(t))d\xi(t) + g(X(T))], \tag{3.2}$$

with given functions $f_0(t,x,u) = f_0(t,x,u,\omega) : [0,T] \times \mathbb{R} \times U \times \Omega \to \mathbb{R}$, $f_1(t,x) = f_1(t,x,\omega) : [0,T] \times \mathbb{R} \times \Omega \to \mathbb{R}$ and $g(x) = g(x,\omega) : \mathbb{R} \times \Omega \to \mathbb{R}$. Let $\mathcal{A}, \mathcal{K}$ denote the family of admissible controls u, ξ, respectively. We let $\mathcal{A}$ be the set of all adapted of the càdlàg, processes $u(t,\omega) \in L^2(dt \times dP)$ and $\mathcal{K}$ consists of all adapted nondecreasing processes $\xi(t)$ with $\xi(0) = 0$. We study the following problem:

Problem 3.1 *Find a control pair* $(\hat{u}, \hat{\xi}) \in \mathcal{A} \times \mathcal{K}$ *such that*

$$J(\hat{u}, \hat{\xi}) = \sup_{(u,\xi) \in \mathcal{A} \times \mathcal{K}} J(u,\xi). \tag{3.3}$$

We impose the following assumptions on the coefficients:
The processes $b(t,s,x,u), \sigma(t,s,x,u), f_0(s,x,u)$, $f_1(t,x,\xi)$ and $h(t,s)$ are $\mathbb{F}$-adapted with respect to s for all $s \leq t$, and twice continuously differentiable (C^2) with respect to t, x, and continuously differentiable (C^1) with respect to u for each s. The driver g is assumed to be $\mathcal{F}_T$-measurable and (C^1) in x. Moreover, all the partial derivatives are supposed to be bounded.
Note that the performance functional (3.2) is not of Volterra type.

3.1 The Hamiltonian and the adjoint equations

Define the *Hamiltonian functional* associated to our control problem (3.1) and (3.2), as

$$\begin{aligned} &\mathbb{H}(t,x,u,p,q)(dt, d\xi(t)) \\ &:= \left[H_0(t,x,u,p,q) + H_1(t,x,u,p)\right]dt + \left[\overline{H_0}(t,x,p) + \overline{H_1}(t,p)\right]d\xi(t), \end{aligned} \tag{3.4}$$

where

$$\begin{aligned} &H_0 : [0,T] \times \mathbb{R} \times U \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \\ &H_1 : [0,T] \times \mathbb{R} \times U \times \mathbb{R}^{[0,T]} \to \mathbb{R}, \\ &\overline{H_0} : [0,T] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \\ &\overline{H_1} : [0,T] \times \mathbb{R}^{[0,T]} \to \mathbb{R}, \end{aligned}$$

are defined as follows

$$\begin{aligned} H_0(t,x,u,p,q) &:= f_0(t,x,u) + p(t)b(t,t,x,u) + q(t,t)\sigma(t,t,x,u), \\ H_1(t,x,u,p) &:= \int_t^T p(s)\tfrac{\partial b}{\partial s}(s,t,x,u)ds + \int_t^T \mathbb{E}[D_t p(s)|\mathcal{F}_t]\tfrac{\partial \sigma}{\partial s}(s,t,x,u)ds, \\ \overline{H_0}(t,x,p) &:= f_1(t,x) + p(t)h(t,t), \\ \overline{H_1}(t,p) &:= \int_t^T p(s)\tfrac{\partial h}{\partial s}(s,t)ds. \end{aligned}$$

For convenience, we will use the following notation from now on:

$$\mathcal{H}(t,x,u,p,q) = H_0(t,x,u,p,q) + H_1(t,x,u,p), \tag{3.5}$$

$$\overline{\mathcal{H}}(t,x,p) = \overline{H_0}(t,x,p) + \overline{H_1}(t,p). \tag{3.6}$$

The BSVIE for the adjoint processes $p(t), q(t,s)$ is defined by

$$p(t) = \frac{\partial g}{\partial x}(X(T)) + \int_t^T \frac{\partial \mathcal{H}}{\partial x}(s)ds + \int_t^T \frac{\partial \overline{\mathcal{H}}}{\partial x}(s)d\xi(s) - \int_t^T q(t,s)dB(s), \tag{3.7}$$

where we have used the simplified notation

$$\frac{\partial \mathcal{H}}{\partial x}(t) = \frac{\partial \mathcal{H}}{\partial x}(t, X(t), u(t), p(t), q(t,t)),$$
$$\frac{\partial \overline{\mathcal{H}}}{\partial x}(t) = \frac{\partial \overline{\mathcal{H}}}{\partial x}(t, X(t), p(t)).$$

Note that from equation (3.1), we get the following equivalent formulation, for each $(t,s) \in [0,T]^2$,

$$\begin{aligned} dX(t) \quad &= \phi'(t)dt + b\left(t,t,X(t),u(t)\right)dt + \left(\int_0^t \frac{\partial b}{\partial t}\left(t,s,X(s),u(s)\right)ds\right)dt + \sigma\left(t,t,X(t),u(t)\right)dB(t) \\ &+ \left(\int_0^t \frac{\partial \sigma}{\partial t}\left(t,s,X(s),u(s)\right)dB(s)\right)dt + h\left(t,t\right)d\xi(t) + \left(\int_0^t \frac{\partial h}{\partial t}(t,s)d\xi(s)\right)dt. \end{aligned} \tag{3.8}$$

We assume that the map $t \mapsto q(t,s)$ is (C^1) for all s, ω and moreover,

$$\mathbb{E}\left[\int_0^T \int_0^T \left(\frac{\partial q(t,s)}{\partial t}\right)^2 dsdt\right] < \infty,$$

under which we can write the following differential form of equation (3.7):

$$\begin{cases} dp(t) = -\left[\frac{\partial \mathcal{H}}{\partial x}(t)dt + \frac{\partial \overline{\mathcal{H}}}{\partial x}(t)d\xi(t) + \int_t^T \frac{\partial q}{\partial t}(t,s)dB(s)dt\right] + q(t,t)dB(t), \\ p(T) = \frac{\partial g}{\partial x}(X(T)). \end{cases} \tag{3.9}$$

3.2 A sufficient maximum principle

We will see under which conditions the couple (u,ξ) is optimal, i.e. we will prove a sufficient version of the maximum principle approach (a verification theorem).

Theorem 3.2 (Sufficient maximum principle) *Let $(\widehat{u}, \widehat{\xi}) \in \mathcal{A} \times \mathcal{K}$, with corresponding solutions $\widehat{X}(t)$, $(\widehat{p}(t), \widehat{q}(t,s))$ of (3.1) and (3.7) respectively. Assume that the functions $x \mapsto g(x)$ and $(x,u,\xi) \mapsto \mathbb{H}(t,x,u,\widehat{p},\widehat{q})(dt, d\xi(t))$ are concave. Moreover, we impose the following optimal conditions for each control:*

- *(Maximum condition for u)*

$$\sup_{u \in U} \mathbb{E}[\mathbb{H}(t, \widehat{X}(t), u, \widehat{p}(t), \widehat{q}(t,t))|\mathcal{G}_t] = \mathbb{E}[\mathbb{H}(t, \widehat{X}(t), \widehat{u}(t), \widehat{p}(t), \widehat{q}(t,t))|\mathcal{G}_t], \text{ for a.a. } t, \ P\text{-a.s.} \tag{3.10}$$

  where we are using the notation

$$\begin{aligned} \mathbb{E}[\mathbb{H}(t, \widehat{X}(t), u, \widehat{p}(t), \widehat{q}(t,t))|\mathcal{G}_t] :&= \mathbb{E}[\mathcal{H}(t, \widehat{X}(t), u, \widehat{p}(t), \widehat{q}(t,t))|\mathcal{G}_t]dt \\ &+ \mathbb{E}[\overline{\mathcal{H}}(t, \widehat{X}(t), \widehat{p}(t))|\mathcal{G}_t]d\xi(t), \text{ for a.a. } t, \ P\text{-a.s.} \end{aligned}$$

- *(Maximum condition for ξ)*
  For all $\xi \in \mathcal{K}$ we have, in the sense of inequality between random measures,

$$\begin{aligned}&\mathbb{E}[\mathbb{H}(t,\widehat{X}(t),u,\widehat{p}(t),\widehat{q}(t,t))(dt,d\xi(t))|\mathcal{G}_t]\\ &\leq \mathbb{E}[\mathbb{H}(t,\widehat{X}(t),u,\widehat{p}(t),\widehat{q}(t,t))(dt,d\widehat{\xi}(t))|\mathcal{G}_t], \text{ for each } t, \ P\text{-a.s.}\end{aligned} \tag{3.11}$$

Then $(\widehat{u},\widehat{\xi})$ is an optimal pair.

Proof. Choose $u\in\mathcal{A}$ and $\xi\in\mathcal{K}$, we want to prove that $J(u,\xi)-J(\widehat{u},\widehat{\xi})\leq 0$. We set

$$J(u,\xi)-J(\widehat{u},\widehat{\xi})=J(u,\xi)-J(u,\widehat{\xi})+J(u,\widehat{\xi})-J(\widehat{u},\widehat{\xi}).$$

Since we have one regular control and one singular, we will solve the problem by separating them, as follows:
First, we prove that ξ is optimal i.e., for all fixed $u\in U$, $J(u,\xi)-J(u,\widehat{\xi})\leq 0$. Then, we plug the optimal $\widehat{\xi}$ into the second part and we prove it for u, i.e., $J(u,\widehat{\xi})-J(\widehat{u},\widehat{\xi})\leq 0$. However, the case of regular controls u has been proved in Theorem 4.3 by Agram *et al* [4] . It rests to prove only the inequality for singular controls ξ.
From definition (3.2), we have

$$J(u,\xi)-J(u,\widehat{\xi})=A_1+A_2+A_3, \tag{3.12}$$

where we have used hereafter the shorthand notations

$$A_1=\mathbb{E}[\textstyle\int_0^T\widetilde{f}_0\,(t)\,dt],\quad A_2=\mathbb{E}[\int_0^T f_1\,(t)\,d\xi(t)-\int_0^T\widehat{f}_1\,(t)\,d\widehat{\xi}(t)],\quad A_3=\mathbb{E}[\widetilde{g}(T)],$$

with $\widetilde{f}_0(t)=f_0(t)-\widehat{f}_0(t)$, $\widetilde{g}(T)=g(X(T))-g(\widehat{X}(T))$, and similarly for $b(t,t)=b\,(t,t,X(t),u(t))$, and the other coefficients. By definition (3.5), we get

$$A_1=\mathbb{E}[\textstyle\int_0^T\{\widetilde{H_0}(t)-\widehat{p}(t)\widetilde{b}(t,t)-\widehat{q}(t,t)\widetilde{\sigma}(t,t)\}dt]. \tag{3.13}$$

Concavity of g together with the terminal value of the BSVIE (3.7), we obtain

$$A_3\quad\leq\mathbb{E}[\tfrac{\partial\widehat{g}}{\partial x}(T)\widetilde{X}(T)]\quad=\mathbb{E}[\widehat{p}(T)\widetilde{X}(T)].$$

Applying the integration by parts formula to the product $\widehat{p}(t)\widetilde{X}(t)$, we get

$$\begin{aligned}A_3&\leq\mathbb{E}[\widehat{p}(T)\widetilde{X}(T)]\\ &=\mathbb{E}[\textstyle\int_0^T\widehat{p}(t)\{\widetilde{b}(t,t)+\int_0^t\frac{\partial\widetilde{b}}{\partial t}(t,s)ds+\int_0^t\frac{\partial\widetilde{\sigma}}{\partial t}\,(t,s)\,dB(s)+\int_0^t\frac{\partial h}{\partial t}(t,s)d\widetilde{\xi}(s)\}dt\\ &+\textstyle\int_0^T\widehat{p}(t)\widetilde{\sigma}(t,t)dB(t)+\int_0^T\widehat{p}(t)h(t,t)d\tilde{\xi}(t)-\int_0^T\widetilde{X}(t)\frac{\partial\widehat{\mathcal{H}}}{\partial x}(t)dt-\int_0^T\widetilde{X}(t)\frac{\partial\widehat{\overline{\mathcal{H}}}}{\partial x}(t)d\widehat{\xi}(t)\\ &-\textstyle\int_0^T\widetilde{X}(t)(\int_t^T\frac{\partial\widehat{q}}{\partial t}(t,s)dB(s))dt+\int_0^T\widetilde{X}(t)\widehat{q}(t,t)dB(t)+\int_0^T\widehat{q}(t,t)\widetilde{\sigma}(t,t)dt].\end{aligned} \tag{3.14}$$

It follows from formulas (2.6)-(2.8), that

$$\begin{aligned}\mathbb{E}[\textstyle\int_0^T\widehat{p}(t)(\int_0^t\frac{\partial\widetilde{b}}{\partial t}(t,s)ds)dt]&=\mathbb{E}[\textstyle\int_0^T(\int_t^T\widehat{p}(s)\frac{\partial\widetilde{b}}{\partial s}(s,t)ds)dt],\\ \mathbb{E}[\textstyle\int_0^T\widehat{p}(t)(\int_0^t\frac{\partial h}{\partial t}(t,s)d\widetilde{\xi}(s))dt]&=\mathbb{E}[\textstyle\int_0^T(\int_t^T\widehat{p}(s)\frac{\partial h}{\partial s}(s,t)ds)d\widetilde{\xi}(t)],\\ \mathbb{E}[\textstyle\int_0^T\widehat{p}(t)(\int_0^t\frac{\partial\widetilde{\sigma}}{\partial t}(t,s)dB(s))dt]&=\mathbb{E}[\textstyle\int_0^T\int_t^T\mathbb{E}(D_t\widehat{p}(s)|\mathcal{F}_t)\frac{\partial\widetilde{\sigma}}{\partial s}(s,t,x,u)dsdt].\end{aligned}$$

Substituting the above into (3.12) , we obtain

$$J(\widehat{u},\xi)-J(\widehat{u},\widehat{\xi})\le\mathbb{E}[\int_0^T(\widetilde{H}_0(t)+\widetilde{H}_1(t))dt+\int_0^T f_1(t)d\xi(t)-\int_0^T\widehat{f}_1(t)d\widehat{\xi}(t)+\int_0^T\widehat{p}(t)h(t,t)d\widetilde{\xi}(t)$$
$$+\int_0^T(\int_0^T\widehat{p}(s)\tfrac{\partial h}{\partial s}(s,t)ds)d\widetilde{\xi}(t)-\int_0^T\widetilde{X}(t)\tfrac{\partial\widehat{\mathcal{H}}}{\partial x}(t)dt-\int_0^T\widetilde{X}(t)\tfrac{\partial\widehat{\overline{\mathcal{H}}}}{\partial x}(t)d\widehat{\xi}(t)]$$
$$=\mathbb{E}[\int_0^T(\mathcal{H}(t)-\widehat{\mathcal{H}}(t))dt+(\overline{\mathcal{H}}(t)d\xi(t)-\widehat{\overline{\mathcal{H}}}(t)d\widehat{\xi}(t))-\int_0^T\widetilde{X}(t)\tfrac{\partial\widehat{\mathcal{H}}}{\partial x}(t)dt$$
$$-\int_0^T\widetilde{X}(t)\tfrac{\partial\widehat{\overline{\mathcal{H}}}}{\partial x}(t)d\widehat{\xi}(t)].$$

Using the concavity of $\mathcal{H}$ and $\overline{\mathcal{H}}$ with respect to x and ξ, we have

$$J(\widehat{u},\xi)-J(\widehat{u},\widehat{\xi})$$
$$\le\mathbb{E}[\int_0^T\{\widetilde{X}(t)\tfrac{\partial\widehat{\mathcal{H}}}{\partial x}(t)-\widetilde{X}(t)\tfrac{\partial\widehat{\mathcal{H}}}{\partial x}(t)\}dt+\int_0^T\widetilde{X}(t)\tfrac{\partial\widehat{\overline{\mathcal{H}}}}{\partial x}(t)d\widehat{\xi}(t)$$
$$-\int_0^T\widetilde{X}(t)\tfrac{\partial\widehat{\overline{\mathcal{H}}}}{\partial x}(t)d\widehat{\xi}(t)+\int_0^T\widehat{\overline{\mathcal{H}}}(t)\{d\xi(t)-d\widehat{\xi}(t)\}]$$
$$=\mathbb{E}[\int_0^T\widehat{\overline{\mathcal{H}}}(t)\{d\xi(t)-d\widehat{\xi}(t)\}]$$
$$=\mathbb{E}[\int_0^T\mathbb{E}[\widehat{\overline{\mathcal{H}}}(t)|\mathcal{G}_t]\{d\xi(t)-d\widehat{\xi}(t)\}]$$
$$\le 0,$$

where the last inequality holds because of the maximum condition (3.11). We conclude that

$$J(\widehat{u},\xi)-J(\widehat{u},\widehat{\xi})\le 0.$$

The proof is complete. □

3.3 A necessary maximum principle

Since the concavity condition is not always satisfied, it is useful to have a necessary condition of optimality where this condition is not required. Suppose that a control $(\widehat{u},\widehat{\xi})\in\mathcal{A}\times\mathcal{K}$ is an optimal pair and that $(v,\zeta)\in\mathcal{A}\times\mathcal{K}$. Define $u^\lambda=u+\lambda v$ and $\xi^\lambda=\xi+\lambda\zeta$, for a non-zero sufficiently small λ. Assume that $(u^\lambda,\xi^\lambda)\in\mathcal{A}\times\mathcal{K}$. For each given $t\in[0,T]$, let $\eta=\eta(t)$ be a bounded $\mathcal{G}_t$-measurable random variable, let $h\in[T-t,T]$ and define

$$v(s):=\eta\mathbf{1}_{[t,t+h]}(s);s\in[0,T]\,. \tag{3.15}$$

Assume that the *derivative process* $Y(t)$, defined by $Y(t):=\frac{d}{d\lambda}X^{u^\lambda,\xi}(t)|_{\lambda=0}$ exists. Then we see that

$$Y(t)=\int_0^t\{\tfrac{\partial b}{\partial x}(t,s)Y(s)+\tfrac{\partial b}{\partial u}(t,s)v(s)\}ds+\int_0^t\{\tfrac{\partial\sigma}{\partial x}(t,s)Y(s)+\tfrac{\partial\sigma}{\partial u}(t,s)v(s)\}dB(s),$$

and hence

$$dY(t)=[\tfrac{\partial b}{\partial x}(t,t)Y(t)+\tfrac{\partial b}{\partial u}(t,t)v(t)+\int_0^t(\tfrac{\partial^2 b}{\partial t\partial x}(t,s)Y(s)+\tfrac{\partial^2 b}{\partial t\partial u}(t,s)v(s))ds+\int_0^t(\tfrac{\partial^2\sigma}{\partial t\partial x}(t,s)Y(s)$$
$$+\tfrac{\partial^2\sigma}{\partial t\partial u}(t,s)v(s))dB(s)]dt+(\tfrac{\partial\sigma}{\partial x}(t,t)Y(t)+\tfrac{\partial\sigma}{\partial u}(t,t)v(t))dB(t). \tag{3.16}$$

Similarly, we define the derivative process $Z(t) := \frac{d}{d\lambda} X^{u,\xi^\lambda}(t)|_{\lambda=0}$, as follows

$$Z(t) = \int_0^t \frac{\partial b}{\partial x}(t,s)Z(s)ds + \int_0^t \frac{\partial \sigma}{\partial x}(t,s)Z(s)dB(s) + \int_0^t h(t,s)d\zeta(s),$$

which is equivalent to

$$\begin{aligned} dZ(t) &= [\tfrac{\partial b}{\partial x}(t,t)Z(t) + \int_0^t \tfrac{\partial^2 b}{\partial t \partial x}(t,s)Z(s)ds]dt + \tfrac{\partial \sigma}{\partial x}(t,t)Z(t)dB(t) \\ &+ \int_0^t \tfrac{\partial^2 \sigma}{\partial t \partial x}(t,s)Z(s)dB(s)dt + h(t,t)d\zeta(t) + \int_0^t \tfrac{\partial h}{\partial t}(t,s)d\zeta(s)dt. \end{aligned} \tag{3.17}$$

We shall prove the following theorem:

Theorem 3.3 (Necessary maximum principle) *1. For fixed $\xi \in \mathcal{K}$, suppose that $\widehat{u} \in \mathcal{A}$ is such that, for all β as in (3.15),*

$$\tfrac{d}{d\lambda} J(\widehat{u} + \lambda\beta, \xi)|_{\lambda=0} = 0 \tag{3.18}$$

and the corresponding solution $\widehat{X}(t), (\widehat{p}(t), \widehat{q}(t,t))$ of (3.1) and (3.7) exists. Then,

$$\mathbb{E}[\tfrac{\partial \mathbb{H}}{\partial u}(t)|\mathcal{G}_t]_{u=\widehat{u}(t)} = 0. \tag{3.19}$$

2. Conversely, if (3.19) holds, then (3.18) holds.

3. Similarly, for fixed $\widehat{u} \in \mathcal{A}$, suppose that $\widehat{\xi} \in \mathcal{K}$ is optimal. Then the following variational inequalities hold:

$$\mathbb{E}[\widehat{f}_1(t) + \widehat{p}(t)h(t,t) + \int_t^T \widehat{p}(s)\tfrac{\partial h}{\partial s}(s,t)ds|\mathcal{G}_t] \leq 0, \tag{3.20}$$

and

$$\mathbb{E}[\widehat{f}_1(t) + \widehat{p}(t)h(t,t) + \int_t^T \widehat{p}(s)\tfrac{\partial h}{\partial s}(s,t)ds|\mathcal{G}_t]d\widehat{\xi}(t) = 0. \tag{3.21}$$

Proof. For simplicity of notation we drop the "hat" notation in the following.
Points 1-2 are direct consequence of Theorem 4.4 in Agram *et al* [4]. We proceed to prove point 3. Since $\widehat{u}$ is fixed we drop the hat from the notation. Set

$$\tfrac{d}{d\lambda} J(\xi^\lambda)|_{\lambda=0} = \mathbb{E}[\int_0^T \{\tfrac{\partial f_0}{\partial x}(t)Z(t)dt + \int_0^T \tfrac{\partial f_1}{\partial x}(t)Z(t)d\xi(t) + \int_0^T f_1(t)d\zeta(t) + \tfrac{\partial g}{\partial x}(T)Z(T)]. \tag{3.22}$$

Applying the Itô formula, we get

$$\begin{aligned} &\mathbb{E}[\tfrac{\partial g}{\partial x}(T)Z(T)] = \mathbb{E}[p(T)Z(T)] \\ &= \mathbb{E}[\int_0^T p(t)\{\tfrac{\partial b}{\partial x}(t,t)Z(t) + \int_0^t \tfrac{\partial^2 b}{\partial t \partial x}(t,s)Z(s)ds\}dt \\ &+ \int_0^T p(t)\tfrac{\partial \sigma}{\partial x}(t,t)Z(t)dB(t) + \int_0^T p(t)(\int_0^t \tfrac{\partial^2 \sigma}{\partial t \partial x}(t,s)Z(s)dB(s))dt \\ &+ \int_0^T p(t)h(t,t)d\zeta(t) + \int_0^T p(t)(\int_0^t \tfrac{\partial h}{\partial t}(t,s)d\zeta(s))dt \\ &- \int_0^T Z(t)\tfrac{\partial \mathcal{H}}{\partial x}(t)dt - \int_0^T Z(t)\tfrac{\partial \overline{\mathcal{H}}}{\partial x}(t)d\xi(t) - \int_0^T Z(t)(\int_t^T \tfrac{\partial q}{\partial t}(t,s)dB(s))dt \\ &+ \int_0^T Z(t)q(t,t)dB(t) + \int_0^T Z(t)\tfrac{\partial \sigma}{\partial x}(t,t)q(t,t)dt]. \end{aligned}$$

Therefore, from (2.6)-(2.8), we obtain

$$
\begin{aligned}
&\mathbb{E}\left[p(T)Z(T)\right] \\
&= \mathbb{E}[\int_0^T Z(t)\{\tfrac{\partial b}{\partial x}(t,t)p(t) + \int_t^T \left(\tfrac{\partial^2 b}{\partial s \partial x}(s,t)p(s) + \mathbb{E}(D_t p(s)|\mathcal{F}_t)\tfrac{\partial^2 \sigma}{\partial s \partial x}(s,t)\right) ds\}dt \\
&+ \int_0^T p(t)h(t,t)d\zeta(t) + \int_0^T (\int_t^T p(s)\tfrac{\partial h}{\partial s}(s,t)ds)d\zeta(t) - \int_0^T Z(t)\tfrac{\partial \mathcal{H}}{\partial x}(t)dt \\
&- \int_0^T Z(t)\tfrac{\partial \overline{\mathcal{H}}}{\partial x}(t)d\xi(t) + \int_0^T Z(t)\tfrac{\partial \sigma}{\partial x}(t,t)q(t,t)dt].
\end{aligned}
$$

Using the definition of $\mathcal{H}$ and $\overline{\mathcal{H}}$ in (3.5) – (3.6),

$$
\begin{aligned}
&\tfrac{d}{d\lambda}J(u,\xi^\lambda)|_{\lambda=0} \\
&= \mathbb{E}[\int_0^T \left\{p(t)h(t,t) + f_1(t) + \int_0^T\int_t^T p(s)\tfrac{\partial h}{\partial s}(s,t)ds\right\}d\zeta(t)].
\end{aligned}
\tag{3.23}
$$

Thus,

$$
\begin{aligned}
0 &\geq \tfrac{d}{d\lambda}J(u,\xi^\lambda)|_{\lambda=0} \\
&= \mathbb{E}[\int_0^T \left\{p(t)h(t,t) + f_1(t) + \int_t^T p(s)\tfrac{\partial h}{\partial s}(s,t)ds\right\}d\zeta(t)],
\end{aligned}
$$

for all $\zeta \in \mathcal{K}(\widehat{\xi})$.
If we choose ζ to be a pure jump process of the form $\zeta(t) = \sum_{0\leq t_i \leq T} \alpha(t_i)$ where $\alpha(t_i) > 0$ is $\mathcal{G}_{t_i}$-measurable for all t_i, then $\zeta \in \mathcal{K}(\widehat{\xi})$ and (3.23) gives

$$
\mathbb{E}[(f_1(t) + p(t)h(t,t) + \int_t^T p(s)\tfrac{\partial h}{\partial s}(s,t)ds)\alpha(t_i)] \leq 0 \text{ for each } t_i \text{ a.s.}
$$

Since this holds for all such ζ with arbitrary t_i, we conclude that

$$
\mathbb{E}[(f_1(t) + p(t)h(t,t) + \int_t^T p(s)\tfrac{\partial h}{\partial s}(s,t)ds)|\mathcal{G}_t] \leq 0 \text{ for each } t \in [0,T] \text{ a.s.}
$$

Finally, applying (3.23) to $\zeta_1 = \widehat{\xi} \in \mathcal{K}(\widehat{\xi})$ and to $\zeta_2 = -\widehat{\xi} \in \mathcal{K}(\widehat{\xi})$, we get for all $t \in [0,T]$

$$
\mathbb{E}[(f_1(t) + p(t)h(t,t) + \int_t^T p(s)\tfrac{\partial h}{\partial s}(s,t)ds)|\mathcal{G}_t]d\widehat{\xi}(t) = 0 \text{ for each } t \in [0,T] \text{ a.s.}
$$

□

4 Application to optimal harvesting with memory

4.1 Optimal harvesting with density-dependent prices

Let $X^\xi(t) = X(t)$ be a given population density (or cash flow) process, modelled by the following stochastic Volterra equation:

$$
X(t) = x_0 + \int_0^t b_0(t,s)X(s)ds + \int_0^t \sigma_0(s)X(s)dB(s) - \int_0^t h(t,s)d\xi(s),
\tag{4.1}
$$

or, in differential form,

$$\begin{cases} dX(t) = b_0(t,t)X(t)dt + \sigma_0(t)X(t)dB(t) - h(t,t)\xi(t) \\ \qquad +[\int_0^t \frac{\partial b_0}{\partial t}(t,s)X(s)ds - \int_0^t \frac{\partial h}{\partial t}(t,s)d\xi(s)]dt, \quad t \geq 0. \\ X(0) = x_0. \end{cases} \tag{4.2}$$

We see that the dynamics of $X(t)$ contains a history or memory term represented by the ds-integral. We assume that $b_0(t,s)$ and $\sigma_0(s)$ are given deterministic functions of t, s, with values in $\mathbb{R}$, and that $b_0(t,s), h(t,s)$ are continuously differentiable with respect to t for each s and $h(t,s) > 0$. For simplicity we assume that these functions are bounded, and the initial value $x_0 \in \mathbb{R}$.

We want to solve the following maximisation problem:

Problem 4.1 *Find* $\widehat{\xi} \in \mathcal{K}$, *such that*

$$\sup_{\xi} J(\xi) = J(\widehat{\xi}), \tag{4.3}$$

where

$$J(\xi) = \mathbb{E}[\theta X(T) + \int_0^T X(t)d\xi(t)]. \tag{4.4}$$

Here $\theta = \theta(\omega)$ is a given $\mathcal{F}_T$-measurable square integrable random variable.
In this case the Hamiltonian $\mathbb{H}$ takes the form

$$\begin{aligned} \mathbb{H}(t,x,p,q) = [b_0(t,t)xp + \sigma_0(t)xq + \int_t^T \tfrac{\partial b_0}{\partial s}(s,t)xp(s)ds - \int_t^T \tfrac{\partial h}{\partial s}(s,t)p(s)d\xi(s)]dt \\ +[x - h(t,t)p]d\xi(t). \end{aligned} \tag{4.5}$$

Note that $\mathbb{H}$ is not concave with respect to x, so the sufficient maximum principle does not apply. However, we can use the necessary maximum principle as follows: The adjoint equation takes the form

$$\begin{cases} dp(t) = -\Big[p(t)b_0(t,t) + \sigma_0(t)q(t,t) + \int_t^T \frac{\partial b_0}{\partial s}(s,t)p(s)ds\Big]dt + d\xi(t) + q(t,t)dB(t) \\ p(T) = \theta, \end{cases}$$

equivalently

$$p(t) = \theta + \int_t^T \{b_0(t,s)p(s) + \sigma_0(s)q(t,s)\}ds + \int_t^T d\xi(s) - \int_t^T q(t,s)dB(s). \tag{4.6}$$

The variational inequalities for an optimal control $\widehat{\xi}$ and the corresponding $\widehat{p}$ are:

$$\widehat{X}(t) - h(t,t)\widehat{p}(t) - \int_t^T \tfrac{\partial h}{\partial s}(s,t)\widehat{p}(s)ds \leq 0, \tag{4.7}$$

and

$$\big\{\widehat{X}(t) - h(t,t)\widehat{p}(t) - \int_t^T \tfrac{\partial h}{\partial s}(s,t)\widehat{p}(s)ds\big\}d\widehat{\xi}(t) = 0. \tag{4.8}$$

We have proved:

Theorem 4.2 *Suppose* $\hat{\xi}$ *is an optimal control for Problem 4.1, with corresponding solution* $\hat{X}$ *of* (4.1). *Then* (4.7) *and* (4.8) *hold, i.e.*

$$\gamma_0(t,t)\widehat{p}(t) + \int_t^T \tfrac{\partial h}{\partial s}(s,t)\widehat{p}(s)ds \geq \widehat{X}(t) \quad a.s., \quad t \in [0,T] \tag{4.9}$$

and

$$\big\{\gamma_0(t,t)\widehat{p}(t) + \int_t^T \tfrac{\partial h}{\partial s}(s,t)\widehat{p}(s)ds - \widehat{X}(t)\big\}d\widehat{\xi}(t) = 0. \tag{4.10}$$

Remark 4.3 The above result states that $\widehat{\xi}(t)$ increases only when

$$\gamma_0(t,t)\widehat{p}(t) + \int_t^T \frac{\partial h}{\partial s}(s,t)\widehat{p}(s)ds - \log(\widehat{X}(t)) = 0. \tag{4.11}$$

Combining this with (4.7) we can conclude that the optimal control can be associated to the solution of a system of reflected forward-backward SVIEs with barrier given by (4.9).

In particular, if we choose $h = 1$ the variational inequalities become

$$\widehat{p}(t) \geq \widehat{X}(t) \text{ for all } t \quad a.s., \quad t \in [0,T] \tag{4.12}$$

and

$$\{\widehat{p}(t) - \widehat{X}(t)\}d\widehat{\xi}(t) = 0. \tag{4.13}$$

Remark 4.4 This is a coupled system $(\widehat{X}(t), \widehat{p}(t))$ consisting of the solution $X(t)$ of the singularly controlled forward SDE

$$X(t) = x_0 + \int_0^t b_0(t,s)X(s)ds + \int_0^t \sigma_0(s)X(s)dB(s) - \int_0^t d\xi(s), \tag{4.14}$$

and the backward reflected SDE

$$p(t) = \theta + \int_t^T \{b_0(t,s)p(s) + \sigma_0(s)q(t,s)\}ds - \int_t^T q(t,s)dB(s) + \int_t^T d\xi(s). \tag{4.15}$$

with barrier $\widehat{X}(t)$ and solution $\widehat{p}(t)$, if we choose $\xi = \widehat{\xi}$. The optimal control is the process $\xi(t)$ which makes (4.12) - (4.15) satisfied. To the best of our knowledge such a forward-backward singularly controlled system has not been studied before. This is an interesting topic for future research.

4.2 Optimal harvesting with density-independent prices

Consider again equation (4.14) but now with performance functional

$$J(\xi) = \mathbb{E}[\theta X(T) + \int_0^T \rho(t)d\xi(t)],$$

for some positive deterministic function ρ. We want to find an optimal $\widehat{\xi} \in \mathcal{K}$, such that

$$\sup_{\xi} J(\xi) = J(\widehat{\xi}).$$

In this case the Hamiltonian $\mathbb{H}$ gets the form

$$\begin{aligned}\mathbb{H}(t,x,p,q) = [b_0(t,t)xp + \sigma_0(t)xq + \int_t^T \frac{\partial b_0}{\partial s}(s,t)xp(s)ds - \int_t^T \frac{\partial h}{\partial s}(s,t)p(s)d\xi(s)]dt \\ +[\rho(t) - h(t,t)p]d\xi(t).\end{aligned}$$

Note that $\mathbb{H}(x)$ is concave in this case. Therefore we can apply the sufficient maximum principle here. The adjoint equation gets the form

$$\begin{cases} dp(t) &= -\Big[p(t)b_0(t,t) + \sigma_0(t)q(t,t) + \int_t^T \frac{\partial b_0}{\partial s}(s,t)p(s)ds\Big]dt + q(t,t)dB(t), \\ p(T) &= \theta, \end{cases} \tag{4.16}$$

equivalently

$$p(t) = \theta + \int_t^T \{b_0(t,s)p(s) + \sigma_0(s)q(t,s)\}ds - \int_t^T q(t,s)dB(s).$$

A closed form expression for $p(t)$ is given in the Appendix (Theorem 5.1).

In this case the variational inequalities for an optimal control $\widehat{\xi}$ and the corresponding $\widehat{p}$ are:

$$\gamma_0(t,t)\widehat{p}(t) + \int_t^T \frac{\partial h}{\partial s}(s,t)\widehat{p}(s)ds \geq \rho(t) \tag{4.17}$$

and

$$\{\gamma_0(t,t)\widehat{p}(t) + \int_t^T \frac{\partial h}{\partial s}(s,t)\widehat{p}(s)ds - \rho(t)\}d\widehat{\xi}(t) = 0. \tag{4.18}$$

We have proved:

Theorem 4.5 *Suppose $\hat{\xi}$ with corresponding solution $\widehat{p}(t)$ of the BSVIE (4.16) satisfies the equations (4.17) - (4.18). Then $\widehat{\xi}$ is an optimal control for Problem 4.1.*

Remark 4.6 Note that (4.17) - (4.18) constitute a sufficient condition for optimality. We can for example get this equation satisfied by choosing $(\widehat{p}(t), \widehat{\xi}(t))$ as the solution of the BSVIE (4.16) reflected downwards at the barrier given by

$$\gamma_0(t,t)\widehat{p}(t) + \int_t^T \frac{\partial h}{\partial s}(s,t)\widehat{p}(s)ds - \rho(t) = 0. \tag{4.19}$$

5 Appendix

Theorem 5.1 *Consider the following linear BSVIE with singular drift*

$$p(t) = \theta + \int_t^T \{b_0(t,s)p(s) + \sigma_0(s)q(t,s)\}ds + \int_t^T \frac{1}{X(s)}d\xi(s) - \int_t^T q(t,s)dB(s). \tag{5.1}$$

The first component $p(t)$ of the solution $(p(t), q(t))$ can be written in closed formula as follows

$$p(t) = \mathbb{E}[\{\theta + \theta\int_t^T \Psi(t,s)ds + \int_t^T\int_t^T \Psi(t,s)\frac{1}{X(r)}d\xi(r)ds\}K(T)|\mathcal{F}_t],$$

where

$$\Psi(t,r) := \Sigma_{n=1}^{\infty} b_0^n(t,r) \tag{5.2}$$

and $K(T)$ is given by

$$K(T) = \exp(\int_0^T \sigma_0(s)dB(s) - \frac{1}{2}\int_0^T \sigma_0^2(s)ds).$$

Proof. The proof is an extension of Theorem 3.1 in Hu and Øksendal [9] to BSVIE with singular drift. Define the measure Q by

$$dQ = M(T)dP \text{ on } \mathcal{F}_T,$$

where $M(t)$ satisfies the equation

$$\begin{cases} dM(t) &= M(t)\sigma_0(t)dB(t), \quad t \in [0,T], \\ M(0) &= 1, \end{cases}$$

which has the solution

$$M(t) := \exp(\int_0^t \sigma_0(s)dB(s) - \frac{1}{2}\int_0^t \sigma_0^2(s)ds), \quad t \in [0,T].$$

Then under the measure Q the process

$$B_Q(t) := B(t) - \int_0^t \sigma_0(s)ds, \quad t \in [0,T] \tag{5.3}$$

is a Q-Brownian motion.
For all $0 \le t \le r \le T$, define

$$b_0^1(t,r) = b_0(t,r)\,, \quad b_0^2(t,r) = \int_t^r b_0(t,s)b_0(s,r)ds,$$

and inductively

$$b_0^n(t,r) = \int_t^r b_0^{n-1}(t,s)b_0(s,r)ds\,, \quad n = 3,4,\cdots.$$

Note that if $|b_0(t,r)| \le C$ (constant) for all t, r, then by induction on $n \in \mathbb{N} : |b_0^n(t,r)| \le \frac{C^n T^n}{n!}$, for all t, r, n. Hence,

$$\Psi(t,r) := \Sigma_{n=1}^{\infty} |b_0^n(t,r)| < \infty,$$

for all t, r. By changing of measure, we can rewrite equation (4.15) as

$$p(t) = \theta + \int_t^T b_0(t,s)p(s)ds + \int_t^T X^{-1}(s)d\xi(s) - \int_t^T q(t,s)dB_Q(s), \quad 0 \le t \le T, \tag{5.4}$$

where the process B_Q is defined by (5.3). Taking the conditional Q-expectation on $\mathcal{F}_t$, we get

$$\begin{aligned} p(t) &= \mathbb{E}_Q[\theta + \int_t^T b_0(t,s)p(s)ds + \int_t^T X^{-1}(s)d\xi(s)|\mathcal{F}_t] \\ &= \tilde{F}(t) + \int_t^T b_0(t,s)\mathbb{E}_Q[p(s)|\mathcal{F}_t]ds + \mathbb{E}_Q[\int_t^T X^{-1}(s)d\xi(s)|\mathcal{F}_t], \quad 0 \le t \le T, \end{aligned} \tag{5.5}$$

where

$$\tilde{F}(s) = \mathbb{E}_Q[\theta|\mathcal{F}_s\,].$$

Fix $r \in [0,t]$. Taking the conditional Q-expectation on $\mathcal{F}_r$ of (5.5), we get

$$\mathbb{E}_Q\left[\hat{p}(t)|\mathcal{F}_r\right] = \tilde{F}(r) + \int_t^T b_0(t,s)\mathbb{E}_Q[p(s)|\mathcal{F}_r]ds + \mathbb{E}_Q[\int_t^T X^{-1}(s)d\xi(s)|\mathcal{F}_r], \quad r \le t \le T\,.$$

Put

$$\tilde{p}(s) = \mathbb{E}_Q\left[p(s)|\mathcal{F}_r\right], \quad r \le s \le T\,.$$

Then the above equation can be written as

$$\tilde{p}(t) = \tilde{F}(r) + \int_t^T b_0(t,s)\tilde{p}(s)ds + \mathbb{E}_Q[\int_t^T X^{-1}(s)d\xi(s)|\mathcal{F}_r], \quad r \le t \le T\,.$$

Substituting $\tilde{p}(s) = \tilde{F}(r) + \int_s^T b_0(s,\alpha)\tilde{p}(\alpha)d\alpha + \mathbb{E}_Q[\int_s^T X^{-1}(\alpha)d\xi(\alpha)|\mathcal{F}_r]$ in the above equation, we obtain

$$\begin{aligned} \tilde{p}(t) &= \tilde{F}(r) + \int_t^T b_0(t,s)\{\tilde{F}(r) + \int_s^T b_0(s,\alpha)\tilde{p}(\alpha)d\alpha + \mathbb{E}_Q[\int_s^T X^{-1}(\alpha)d\xi(\alpha)|\mathcal{F}_r]\}ds \\ &= \tilde{F}(r) + \int_t^T b_0(t,s)\tilde{F}(r)ds + \int_t^T b_0(t,s)\mathbb{E}_Q[\int_s^T X^{-1}(\alpha)d\xi(\alpha)|\mathcal{F}_r]ds \\ &\quad + \int_t^T b_0^{(2)}(t,\alpha)\tilde{p}(\alpha)d\alpha, \quad r \le t \le T\,. \end{aligned}$$

Repeating this, we get by induction

$$\begin{aligned}\tilde{p}(t) &= \tilde{F}(r) + \sum_{n=1}^{\infty}\int_t^T b_0^n(t,\alpha)\tilde{F}(r)d\alpha + \sum_{n=1}^{\infty}\int_t^T b_0^n(t,\alpha)\mathbb{E}_Q[\int_s^T X^{-1}(\alpha)d\xi(\alpha)|\mathcal{F}_r]d\alpha \\ &= \tilde{F}(r) + \int_t^T \Psi(t,\alpha)\tilde{F}(r)d\alpha + \int_t^T \Psi(t,\alpha)\mathbb{E}_Q[\int_s^T X^{-1}(\alpha)d\xi(\alpha)|\mathcal{F}_r]d\alpha.\end{aligned}$$

where Ψ is defined by (5.2). Now substituting $\tilde{p}(s)$ in (5.5), for $r = t$, we obtain

$$\begin{aligned}p(t) &= \tilde{F}(t) + \int_t^T \Psi(t,s)\tilde{F}(t)ds + \int_t^T \Psi(t,s)\mathbb{E}_Q[\int_s^T X^{-1}(\alpha)d\xi(\alpha)|\mathcal{F}_t]ds \\ &= \mathbb{E}_Q[\theta + \theta\int_t^T \Psi(t,s)ds + \int_t^T \Psi(t,s)\int_s^T X^{-1}(\alpha)d\xi(\alpha)ds|\mathcal{F}_t] \\ &= \mathbb{E}_Q[\theta + \theta\int_t^T \Psi(t,s)ds + \int_t^T \Psi(t,s)ds\int_t^T X^{-1}(\alpha)d\xi(\alpha)|\mathcal{F}_t].\end{aligned}$$

Acknowledgements. Nacira Agram and Bernt Øksendal are gratefully acknowledge the financial support provided by the Swedish Research Council grant (2020-04697) and the Norwegian Research Council grant (250768/F20), respectively.

References

[1] Aase, K., Øksendal, B., Privault, N., & Ubøe, J. (2000). White noise generalizations of the Clark-Haussmann-Ocone theorem with application to mathematical finance. Finance and Stochastics, 4(4), 465-496.

[2] Agram, N. (2019). Dynamic risk measure for BSVIE with jumps and semimartingale issues. Stochastic Analysis and Applications, 37(3), 361-376.

[3] Agram, N. & Øksendal, B. (2015). Malliavin calculus and optimal control of stochastic Volterra equations. Journal of Optimization Theory and Applications, 167(3), 1070-1094.

[4] Agram, N., Øksendal, B., & Yakhlef, S. (2019). New approach to optimal control of stochastic Volterra integral equations. Stochastics, 91(6), 873-894.

[5] Agram, N.,Øksendal, B. & Yakhlef, S. (2018) Optimal control of forward-backward stochastic Volterra equations. In F. Gesztezy et al (editors): Non-linear Partial Differential equations, Mathematical Physics, and Stochastic Analysis. The Helge Holden Anniversary Volume. EMS Congress Reports, pp. 3-35. http://arxiv.org/abs/1606.03280v4.

[6] Belbas, S. A. (2007). A new method for optimal control of Volterra integral equations. Applied mathematics and computation, 189(2), 1902-1915.

[7] Gripenberg, G., Londen. S.-O. & Staffans, O. (1990). Volterra integral and functional equations (No. 34). Cambridge University Press.

[8] Hida, T., Kuo, H. H., Potthoff, J. & Streit, L. (2013). White noise: an infinite dimensional calculus (Vol. 253). Springer.

[9] Hu,Y. & Øksendal, B. (2019). Linear backward stochastic Volterra equations. Stochastic Processes and their Applications, 129(2), 626-633.

[10] Malliavin, P. (1978). (1978). Stochastic calculus of variations and hypoelliptic operators. In Proc. Internat. Symposium on Stochastic Differential Equations, Kyoto Univ., Kyoto, 1976. Wiley.

[11] Di Nunno, G., Øksendal, B. K., & Proske, F. (2009). Malliavin Calculus for Lévy Processes with Applications to Finance. Second Edition. Springer.

[12] Lin, P., & Yong, J. (2020). Controlled singular Volterra integral equations and Pontryagin maximum principle. SIAM Journal on Control and Optimization, 58(1), 136-164.

[13] Wang, T., Zhu, Q., & Shi, Y. (2011, July). Necessary and sufficient conditions of optimality for stochastic integral systems with partial information. In Proceedings of the 30th Chinese Control Conference (pp. 1950-1955). IEEE.

[14] Yong, J. (2006). Backward stochastic Volterra integral equations and some related problems. Stochastic Processes and their Applications, 116(5), 779-795.

[15] Yong, J. (2008). Well-posedness and regularity of backward stochastic Volterra integral equations. Probability Theory and Related Fields, 142(1), 21-77.