Malliavin calculus and optimal control of stochastic Volterra equations

Nacira AGRAM* and Bernt ØKSENDAL†‡§

2 June 2014

Abstract

Solutions of stochastic Volterra equations are not Markov processes, and therefore classical methods, like dynamic programming, cannot be used to study such control problems. However, we shall see that by using Malliavin calculus it is possible to formulate a modified functional type of maximum principle suitable for such systems. This principle also applies to situations where the controller has only partial information available to base her decisions upon. We present both a sufficient and a necessary maximum principle of this type, and then we use the results to study some specific examples. In particular, we solve an optimal portfolio problem in a financial market model with memory.

Keywords: Stochastic Volterra equations; Partial information; Malliavin calculus; Maximum principle.

MSC 2010 Mathematics Subject Classification: 60H07, 60H20, 93E20

1 Introduction

Stochastic Volterra equations appear naturally in many areas of mathematics such as integral transforms, transport equations, functional differential equations and so forth, and they also appear in applications in biology, physics and finance. For an example in economics (which also applies to population dynamics) see Example 3.4.1 in [7], for an example stemming from Newtonian motion in a random environment see Exercise 5.12 in [10]. Stochastic Volterra equations can also be derived from stochastic delay equations. See [15] and the reference therein. More generally, they represent interesting models for stochastic dynamic systems with memory. For more information on applications of Volterra integral equations, we refer to [6], [2], [13] and [14], the first two dealing with deterministic equations only.

In view of this it is important to find good methods to solve optimal control problems for such equations. In earlier papers [19], [18], [15] and [17], the authors have obtained different types of maximum principles

*Faculty of Science and Technology, University Med Khider, P. O. Box 145, Biskra (07000) Algeria. Email: agram-nacira@yahoo.fr
†Dept. of Mathematics, University of Oslo, Box 1053 Blindern, N-0316 Oslo, Norway. Email: oksendal@math.uio.no
‡The research leading to these results has received funding from the European Research Council under the European Community’s Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no [228087].
§Norwegian School of Economics (NHH), Helleveien 30, N-5045 Bergen, Norway.
for stochastic Volterra equations. In [19] a new type of backward stochastic Volterra integral equations (BSVIEs), driven by Brownian motion only, is studied, and it is proved that if a given control is optimal, then an associated BSVIE has a unique solution. An extension of this result to mean-field equations is obtained in [17]. In [18] the same type of BSVIEs are used and - still in the Brownian motion driven case - a necessary maximum principle is obtained for partial information and when the control domain is not necessarily convex. In [15] a Malliavin calculus approach is used, together with a perturbation argument, to get a necessary maximum principle with partial information.

In our paper, we use Malliavin calculus to obtain both a sufficient and a necessary maximum principle for optimal control of stochastic Volterra equations with jumps and partial information. We define a Hamiltonian which involves also the Malliavin derivatives of one of the adjoint processes. This has the advantage that the corresponding adjoint equation becomes a standard BSDE, not a Volterra type BSVIE as in [19], [18] and [17]. On the other hand, our BSDE involves the Malliavin derivative of the adjoint process. It is interesting to note that BSDEs involving Malliavin derivatives also appear in connection with optimal control of stochastic Volterra equations with noisy memory. See [3].

In the special case when the coefficients of the state equation do not depend on the state, we show that the necessary maximum principle we obtain, is equivalent to the one in [15]. However, for more general systems our maximum principle is simpler than the one in [15]. Moreover, there is no sufficient maximum principle we obtain, is equivalent to the one in [15]. However, for more general systems our maximum principle is simpler than the one in [15].

In the last part of the paper we illustrate our results by solving an optimal portfolio problem in a financial market modeled by a stochastic Volterra equation.

We now describe more precisely the general problem we consider:

From now on we let $B(t)$ and $\tilde{N}(dt,d\zeta):=N(dt,d\zeta)-\nu(d\zeta)dt$ denote a Brownian motion and an independent compensated Poisson random measure, respectively, on a filtered probability space $(\Omega,F,\mathbb{F}):=\{\mathcal{F}_t\}_{0\leq t\leq T},P)$ satisfying the usual conditions, $P$ is a reference probability measure and $\nu$ is the Lévy measure of $N$. We refer to [11] for an introduction to stochastic calculus for Lévy processes.

Let $\mathcal{A}$ be a given family of admissible controls, required to be $\mathcal{G}_t-$predictable, where $\mathcal{G}={\mathcal{G}_t}_{t\geq 0}$ is a given subfiltration of $\mathbb{F}={\mathcal{F}_t}_{t\geq 0}$, in the sense that $\mathcal{G}_t \subseteq \mathcal{F}_t$ for all $t$. For example, we could have

$$\mathcal{G}_t = \mathcal{F}_{(t-\delta)^+}$$

(delayed information flow).

Suppose the state dynamics is given by a controlled stochastic Volterra equation with jumps of the following form:

$$X(t) = X^{(\omega)}(t) = \xi(t) + \int_0^t b(t,s,X(s),u(s)) \, ds + \int_0^t \sigma(t,s,X(s),u(s)) \, dB(s) + \int_0^t \int_{\mathbb{R}} \gamma(t,s,X(s),u(s),\zeta) \tilde{N}(ds,d\zeta),$$

where $b(t,s,x,v,\omega):[0,T] \times [0,T] \times \mathbb{R} \times U \times \Omega) \mapsto \mathbb{R}, \sigma(t,s,x,v,\omega):[0,T] \times [0,T] \times \mathbb{R} \times U \times \Omega) \mapsto \mathbb{R}$ and $\gamma(t,s,x,v,\omega,\zeta):[0,T] \times [0,T] \times \mathbb{R} \times U \times \Omega \times \mathbb{R}_\nu) \mapsto \mathbb{R}$ are given functions, assumed to be $\mathbb{F}$-adapted with respect to the second variable $s$ for all $t,x,v,\zeta$ and continuously differentiable ($C^1$) with respect to the first variable $t$, with partial derivatives in $L^2([0,T] \times [0,T] \times \mathbb{R} \times \mathbb{R} \times \Omega)$ and in $L^2([0,T] \times [0,T] \times \mathbb{R} \times \mathbb{R} \times \Omega \times \nu)$, respectively. Here $\mathbb{R}_\nu = \mathbb{R} - \{0\}$ and $U$ denotes a given open set containing all possible admissible control values $u(t,\omega)$ for $(t,\omega) \in \Omega, u \in \mathcal{A}$.}

2
The performance functional is given by
\[ J(u) = \mathbb{E} \left[ \int_0^T f(s, X(s), u(s)) \, ds + g(X(T)) \right]; \quad u \in \mathcal{A}. \] (1.2)

The problem we study is the following:

**Problem:** Find \( u^* \in \mathcal{A} \) such that
\[ J(u^*) = \sup_{u \in \mathcal{A}} J(u) \] (1.3)

Such a control \( u^* \) is called an optimal control.

## 2 A brief review of Malliavin calculus for Lévy processes

In this section we recall the basic definition and properties of Malliavin calculus for Lévy processes related to this paper, for reader’s convenience. A general reference for this presentation is the book [5]. See also [4], [9] and [16].

In view of the Lévy–Itô decomposition theorem, which states that any Lévy process \( Y(t) \) with \( \mathbb{E}[Y^2(t)] < \infty \) for all \( t \) can be written
\[ Y(t) = at + bB(t) + \int_0^t \int_\mathbb{R} \zeta \tilde{N}(ds, d\zeta) \]
with constants \( a \) and \( b \), we see that it suffices to deal with Malliavin calculus for \( B(\cdot) \) and for
\[ \eta(\cdot) := \int_0^t \int_\mathbb{R} \zeta \tilde{N}(ds, d\zeta) \]
separately.

### 2.1 Malliavin calculus for \( B(\cdot) \)

A natural starting point is the Wiener-Itô chaos expansion theorem, which states that any Lévy process \( Y(t) \) with
\[ \mathbb{E}[Y^2(t)] < \infty \] for all \( t \)
can be written
\[ F = \sum_{n=0}^{\infty} I_n(f_n) \] (2.1)
for a unique sequence of symmetric deterministic functions \( f_n \in L^2(\lambda^n) \), where \( \lambda \) is Lebesgue measure on \([0, T]\) and
\[ I_n(f_n) = n! \int_0^T \int_0^{t_n} \cdots \int_0^{t_2} f_n(t_1, \cdots, t_n) dB(t_1) dB(t_2) \cdots dB(t_n) \] (2.2)
(the \( n \)-times iterated integral of \( f_n \) with respect to \( B(\cdot) \)) for \( n = 1, 2, \ldots \) and \( I_0(f_0) = f_0 \) when \( f_0 \) is a constant.

Moreover, we have the isometry
\[ \mathbb{E}[F^2] = ||F||_{L^2(\mathfrak{M})}^2 = \sum_{n=0}^{\infty} n! ||f_n||_{L^2(\lambda^n)}^2. \] (2.3)
Definition 2.1 (Malliavin derivative $D_t$ with respect to $B(\cdot)$)

Let $\mathbb{D}_{1,2}^{(B)}$ be the space of all $F \in L^2(F_T, P)$ such that its chaos expansion (2.1) satisfies

$$||F||_{\mathbb{D}_{1,2}^{(B)}}^2 := \sum_{n=1}^{\infty} nn! ||f_n||_{L^2(\lambda^n)}^2 < \infty. \quad (2.4)$$

For $F \in \mathbb{D}_{1,2}^{(B)}$ and $t \in [0, T]$, we define the Malliavin derivative of $F$ at $t$ (with respect to $B(\cdot)$), $D_tF$, by

$$D_tF = \sum_{n=1}^{\infty} nI_{n-1}(f_n(\cdot,t)), \quad (2.5)$$

where the notation $I_{n-1}(f_n(\cdot,t))$ means that we apply the $(n-1)$-times iterated integral to the first $n-1$ variables $t_1, \cdots, t_{n-1}$ of $f_n(t_1, t_2, \cdots, t_n)$ and keep the last variable $t_n = t$ as a parameter.

One can easily check that

$$E\left[\int_0^T (D_tF)^2 dt\right] = \sum_{n=1}^{\infty} nn! ||f_n||_{L^2(\lambda^n)}^2 = ||F||_{\mathbb{D}_{1,2}^{(B)}}^2, \quad (2.6)$$

so $(t, \omega) \rightarrow D_tF(\omega)$ belongs to $L^2(\lambda \times P)$.

Example 2.2 If $F = \int_0^T f(t)dB(t)$ with $f \in L^2(\lambda)$ deterministic, then

$$D_tF = f(t) \text{ for a.a. } t \in [0, T].$$

More generally, if $u(s)$ is Skorohod integrable, $u(s) \in \mathbb{D}_{1,2}$ for a.a. $s$ and $D_tu(s)$ is Skorohod integrable for a.a. $t$, then

$$D_t\left(\int_0^T u(s)dB(s)\right) = \int_0^T D_tu(s)dB(s) + u(t) \text{ for a.a. } (t, \omega), \quad (2.7)$$

where $\int_0^T \psi(s)dB(s)$ denotes the Skorohod integral of $\psi$ with respect to $B(\cdot)$.

Some other basic properties of the Malliavin derivative $D_t$ are the following:

(i) **Chain rule** (For a more general version see [9], page 29)

Suppose $F_1, \ldots, F_m \in \mathbb{D}_{1,2}^{(B)}$ and that $\psi : \mathbb{R}^m \rightarrow \mathbb{R}$ is $C^1$ with bounded partial derivatives. Then $\psi(F_1, \cdots, F_m) \in \mathbb{D}_{1,2}$ and

$$D_t\psi(F_1, \cdots, F_m) = \sum_{i=1}^{m} \frac{\partial \psi}{\partial x_i}(F_1, \cdots, F_m)D_tF_i. \quad (2.8)$$

(ii) **Duality formula**

Suppose $u(t)$ is $\mathcal{F}_t$-adapted with $E[\int_0^T u^2(t)dt] < \infty$ and let $F \in \mathbb{D}_{1,2}^{(B)}$. Then

$$E[F \int_0^T u(t)dB(t)] = E[\int_0^T u(t)D_tF(t)dt]. \quad (2.9)$$

(iii) **Malliavin derivative and adapted processes**

If $\varphi$ is an $\mathcal{F}$-adapted process, then $D_s\varphi(t) = 0$ for $s > t$.

Remark 2.1 We put $D_t\varphi(t) = \lim_{s \rightarrow t^-} D_s\varphi(t)$ (if the limit exists).
2.2 Malliavin calculus for \( \tilde{N}(\cdot) \)

The construction of a stochastic derivative/Malliavin derivative in the pure jump martingale case follows the same lines as in the Brownian motion case. In this case the corresponding Wiener-Itô chaos expansion theorem states that any \( F \in L^2(\mathcal{F}_T, P) \) (where in this case \( \mathcal{F}_t = \mathcal{F}_{t(\tilde{N})} \) is the \( \sigma \)-algebra generated by \( \eta(s) := \int_0^s \int_{\mathbb{R}_0} \tilde{N}(dr,d\zeta); 0 \leq s \leq t \) can be written as

\[
F = \sum_{n=0}^{\infty} I_n(f_n); \quad f_n \in \dot{L}^2((\lambda \times \nu)^n), \tag{2.10}
\]

where \( \dot{L}^2((\lambda \times \nu)^n) \) is the space of functions \( f_n(t_1, \zeta_1, \ldots, t_n, \zeta_n); \; t_i \in [0, T], \; \zeta_i \in \mathbb{R}_0 \) such that \( f_n \in L^2((\lambda \times \nu)^n) \) and \( f_n \) is symmetric with respect to the pairs of variables \( (t_1, \zeta_1), \ldots, (t_n, \zeta_n) \).

It is important to note that in this case the \( n \)-times iterated integral \( I_n(f_n) \) is taken with respect to \( \tilde{N}(dt,d\zeta) \) and not with respect to \( d\eta(t) \). Thus, we define

\[
I_n(f_n) = n! \int_0^T \int_{\mathbb{R}_0} \cdots \int_{\mathbb{R}_0} f_n(t_1, \zeta_1, \ldots, t_n, \zeta_n) \tilde{N}(dt_1, d\zeta_1) \cdots \tilde{N}(dt_n, d\zeta_n) \tag{2.11}
\]

for \( f_n \in \dot{L}^2((\lambda \times \nu)^n) \).

The Itô isometry for stochastic integrals with respect to \( \tilde{N}(dt,d\zeta) \) then gives the following isometry for the chaos expansion:

\[
\|F\|_{L^2(\mathcal{F}_T)}^2 = \sum_{n=0}^{\infty} n! \|f_n\|_{L^2((\lambda \times \nu)^n)}^2. \tag{2.12}
\]

As in the Brownian motion case we use the chaos expansion to define the Malliavin derivative. Note that in this case there are two parameters \( t, \zeta \), where \( t \) represents time and \( \zeta \neq 0 \) represents a generic jump size.

**Definition 2.3 (Malliavin derivative \( D_{t,\zeta} \) with respect to \( \tilde{N}(\cdot, \cdot) \))** \[23\] Let \( \mathcal{D}^{(\tilde{N})}_{1,2} \) be the space of all \( F \in L^2(\mathcal{F}_T, P) \) such that its chaos expansion (2.10) satisfies

\[
\|F\|_{\mathcal{D}^{(\tilde{N})}_{1,2}}^2 := \sum_{n=1}^{\infty} nn! \|f_n\|_{L^2((\lambda \times \nu)^n)}^2 < \infty. \tag{2.13}
\]

For \( F \in \mathcal{D}^{(\tilde{N})}_{1,2} \), we define the Malliavin derivative of \( F \) at \( (t, \zeta) \) (with respect to \( \tilde{N}(\cdot) \)), \( D_{t,\zeta} F \), by

\[
D_{t,\zeta} F = \sum_{n=1}^{\infty} nI_{n-1}(f_n(\cdot, t, \zeta)), \tag{2.14}
\]

where \( I_{n-1}(f_n(\cdot, t, \zeta)) \) means that we perform the \( (n - 1) \)-times iterated integral with respect to \( \tilde{N} \) to the first \( n - 1 \) variable pairs \( (t_1, \zeta_1), \ldots, (t_{n-1}, \zeta_{n-1}) \), keeping \( (t_n, \zeta_n) = (t, \zeta) \) as a parameter.

In this case we get the isometry.

\[
\mathbb{E}\left[ \int_0^T \int_{\mathbb{R}_0} (D_{t,\zeta} F)^2 \nu(d\zeta)dt \right] = \sum_{n=0}^{\infty} nn! \|f_n\|_{L^2((\lambda \times \nu)^n)}^2 = \|F\|_{\mathcal{D}^{(\tilde{N})}_{1,2}}^2. \tag{2.15}
\]

(Compare with (2.10).)
Example 2.4 If \( F = \int_0^T \int_{\mathbb{R}_0} f(t, \zeta) \tilde{N}(dt, d\zeta) \) for some deterministic \( f(t, \zeta) \in L^2(\lambda \times \nu) \), then
\[
D_{t, \zeta} F = f(t, \zeta) \quad \text{for a.a. } (t, \zeta).
\]

More generally, if \( \psi(s, \zeta) \) is Skorohod integrable with respect to \( \tilde{N}(\delta s, d\zeta) \), \( \psi(s, \zeta) \in \mathbb{D}_1^{(\tilde{N})} \) for a.a. \( s, \zeta \) and \( D_{t, z} \psi(s, \zeta) \) is Skorohod integrable for a.a. \( t, z \), then
\[
D_{t, z} (\int_0^T \int_{\mathbb{R}} \psi(s, \zeta) \tilde{N}(\delta s, d\zeta)) = \int_0^T \int_{\mathbb{R}} D_{t, z} \psi(s, \zeta) \tilde{N}(\delta s, d\zeta) + u(t, z) \quad \text{for a.a. } t, z,
\]
where \( \int_0^T \int_{\mathbb{R}} \psi(s, \zeta) \tilde{N}(\delta s, d\zeta) \) denotes the Skorohod integral of \( \psi \) with respect to \( \tilde{N}(. , \cdot) \). (See \cite{5} for a definition of such Skorohod integrals and for more details.)

The properties of \( D_{t, \zeta} \) corresponding to the properties (2.8) and (2.9) of \( D_t \) are the following:

(i) **Chain rule** \cite{5}
Suppose \( F_1, \cdots, F_m \in \mathbb{D}_1^{(\tilde{N})} \) and that \( \phi : \mathbb{R}^m \to \mathbb{R} \) is continuous and bounded. Then \( \phi(F_1, \cdots, F_m) \in \mathbb{D}_1^{(\tilde{N})} \) and
\[
D_{t, \zeta} \phi(F_1, \cdots, F_m) = \phi(F_1 + D_{t, \zeta} F_1, \cdots, F_m + D_{t, \zeta} F_m) - \phi(F_1, \cdots, F_m).
\]

(ii) **Duality formula** \cite{5}
Suppose \( \Psi(t, \zeta) \) is \( \mathcal{F}_t \)-adapted and \( \mathbb{E}\left[ \int_0^T \int_{\mathbb{R}_0} \psi^2(t, \zeta) \nu(d\zeta) dt \right] < \infty \) and let \( F \in \mathbb{D}_1^{(\tilde{N})} \). Then
\[
\mathbb{E}\left[ F \int_0^T \int_{\mathbb{R}_0} \Psi(t, \zeta) \tilde{N}(dt, d\zeta) \right] = \mathbb{E}\left[ \int_0^T \int_{\mathbb{R}_0} \Psi(t, \zeta) D_{t, \zeta} F \nu(d\zeta) dt \right].
\]

(iii) **Malliavin derivative and adapted processes** \cite{5}
If \( \phi \) is an \( \mathcal{F}_t \)-adapted process, then
\[
D_{s, \zeta} \phi(t) = 0 \quad \text{for all } s > t.
\]

**Remark 2.2** We put \( D_{t, \zeta} \phi(t) = \lim_{s\to t^-} D_{s, \zeta} \phi(t) \) (if the limit exists).

### 3 A sufficient maximum principle

Let \( \mathcal{L} \) and \( \mathcal{L}_\zeta \) be the set of all stochastic processes on \([0, T]\) and \([0, T] \times \mathbb{R}_0\), respectively. Define the Hamiltonian functionals

\[
H_0 : [0, T] \times \mathbb{R} \times \mathcal{U} \times \mathbb{R} \times \mathbb{R} \times \mathcal{R} \to \mathbb{R}
\]
and

\[
H_1 : [0, T] \times \mathbb{R} \times \mathcal{U} \times \mathbb{R} \times \mathcal{L} \times \mathcal{L}_\zeta \to \mathbb{R}
\]
by
\[
H_0(t, x, v, p, q, r) = f(t, x, v) + b(t, t, x, v)p + \sigma(t, t, x, v)q + \int_{\mathbb{R}} \gamma(t, t, x, v, \zeta) r(\zeta) \nu(d\zeta)
\]
and
\[
H_1(t, x, v, p, D_t p(\cdot), D_{t, \zeta} p(\cdot)) = \int_t^T \frac{\partial b}{\partial s}(s, t, x, v)p(s)ds + \int_t^T \frac{\partial \sigma}{\partial s}(s, t, x, v)D_t p(s)ds.
\]
Theorem 3.1
Let
\[ \mathcal{H}(t, x, v, p(\cdot), q(\cdot), r(\cdot)) = H_0(t, x, v, p, q, r) + H_1(t, x, v, p(\cdot), D_t p(\cdot), D_{t, \zeta} p(\cdot)) \]

The adjoint BSDE for \( p(t), q(t), r(t, \cdot) \) is defined by
\[
\begin{cases}
    dp(t) = -\mathbb{E}\left[ \frac{\partial \mathcal{H}}{\partial x}(t) \mid \mathcal{F}_t \right] dt + q(t) dB(t) + \int_{\mathbb{R}} r(t, \zeta) \tilde{N}(dt, d\zeta); 0 \leq t \leq T \\
    p(T) = g'(X(T))
\end{cases}
\]

where we have used the simplified notation
\[ \frac{\partial \mathcal{H}}{\partial x}(t, X(t), u(t), p(\cdot), q(\cdot), r(\cdot)) \]

Note that from (1.1) we get
\[
dX(t) = \xi'(t) dt + b(t, t, X(t), u(t)) dt + \left( \int_0^t \frac{\partial b}{\partial t} (s, X(s), u(s)) ds \right) dt \\
+ \sigma(t, t, X(t), u(t)) dB(t) + \left( \int_0^t \frac{\partial \sigma}{\partial t} (s, X(s), u(s)) dB(s) \right) dt \\
+ \int_{\mathbb{R}} \gamma(t, t, X(t), u(t), \zeta) \tilde{N}(dt, d\zeta) + \left( \int_0^t \frac{\partial \gamma}{\partial t} (s, X(s), u(s), \zeta) \tilde{N}(ds, d\zeta) \right) dt.
\]

Theorem 3.1 Let \( \hat{u} \in \mathcal{A} \), with corresponding solutions \( \hat{X}(t), (\hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot)) \) of (1.1) and (3.4), respectively.

Suppose that the functions
\[ x \to g(x) \text{ and } (x, v) \to \mathcal{H}(t, x, v, \hat{p}, \hat{q}, \hat{r}) \]
are concave.

and that
\[
\sup_{v \in U} \mathbb{E} \left[ \mathcal{H}(t, \hat{X}(t), v, \hat{p}(\cdot), \hat{q}(\cdot), \hat{r}(\cdot)) \mid \mathcal{G}_t \right] = \mathbb{E} \left[ \mathcal{H}(t, \hat{X}(t), \hat{u}(t), \hat{p}(\cdot), \hat{q}(\cdot), \hat{r}(\cdot)) \mid \mathcal{G}_t \right] \text{ for all } t.
\]

Then \( \hat{u} \) is an optimal control.

Proof. By considering a suitable increasing family of stopping times converging to \( T \) we may assume that all the local martingales appearing in the proof below are martingales. See the proof of Theorem 2.1 in [12] for details.

Choose an arbitrary \( u \in \mathcal{A} \) with corresponding \( X(t) \) and consider \( J(u) - J(\hat{u}) = I_1 + I_2 \), where
\[
I_1 = \mathbb{E} \left[ \int_0^T \left\{ f(t) - \hat{f}(t) \right\} dt \right], I_2 = \mathbb{E} \left[ g(X(T)) - \hat{g}(X(T)) \right],
\]

7
where \( f(t) = f(t, X(t), u(t)), \hat{f}(t) = f(t, \hat{X}(t), \hat{u}(t)) \).

Using a similar notation for

\[
\begin{align*}
\mathbb{E} \left[ \int_0^T \{ &H_0(t, X(t), u(t), \dot{p}(t), \dot{q}(t), \dot{r}(t, \cdot)) \\
&- H_0(t, \hat{X}(t), \hat{u}(t), \dot{p}(t), \dot{q}(t), \dot{r}(t, \cdot)) \\
&- [b(t) - \hat{b}(t)] \dot{p}(t) - [\sigma(t) - \hat{\sigma}(t)] \dot{q}(t) - \int_\mathbb{R} \gamma(t, \zeta) - \hat{\gamma}(t, \zeta) \nu(d\zeta) \right] dt.
\end{align*}
\]

By concavity we have

\[
\begin{align*}
I_2 &\leq \mathbb{E} \left[ g'(\hat{X}(T)) \left( X(T) - \hat{X}(T) \right) \right] \\
&= \mathbb{E} \left[ \hat{p}(T) \left( X(T) - \hat{X}(T) \right) \right] \\
&= \mathbb{E} \left[ \int_0^T \left\{ \dot{p}(t) \left( b(t) - \hat{b}(t) + \int_0^t \left( \frac{\partial b}{\partial t}(t, s) - \frac{\partial \hat{b}}{\partial t}(t, s) \right) ds \right) \\
+ \int_0^t \left( \frac{\partial \sigma}{\partial t}(t, s) - \frac{\partial \hat{\sigma}}{\partial t}(t, s) \right) dB(s) \\
+ \int_0^t \int_\mathbb{R} \left( \frac{\partial \gamma}{\partial t}(t, s, \zeta) - \frac{\partial \hat{\gamma}}{\partial t}(t, s, \zeta) \right) N(ds, d\zeta) \right] \\
&= \mathbb{E} \left[ \frac{\partial \mathcal{H}}{\partial x}(t) | \mathcal{F}_t \right] \left( X(t) - \hat{X}(t) \right) \\
+ \dot{q}(t) [\sigma(t) - \hat{\sigma}(t)] + \int_\mathbb{R} \dot{r}(t, \zeta) [\gamma(t, \zeta) - \hat{\gamma}(t, \zeta)] \nu(d\zeta) \right] dt.
\end{align*}
\]

By the Fubini theorem we get

\[
\int_0^T \left( \int_0^t \frac{\partial b}{\partial t}(t, s) ds \right) \dot{p}(t) dt = \int_0^T \left( \int_s^T \frac{\partial b}{\partial s}(t, s) \dot{p}(t) dt \right) ds = \int_0^T \left( \int_t^T \frac{\partial b}{\partial s}(s, t) \dot{p}(s) ds \right) dt
\]

and similarly, by the duality theorems,

\[
\begin{align*}
\mathbb{E} \int_0^T \left( \int_0^t \frac{\partial \sigma}{\partial t}(t, s) dB(s) \right) \dot{p}(t) dt &= \int_0^T \mathbb{E} \left[ \int_t^T \frac{\partial \sigma}{\partial s}(t, s) dB(s) \dot{p}(t) \right] dt = \int_0^T \mathbb{E} \left[ \int_t^T \frac{\partial \sigma}{\partial s}(t, s) D_s \dot{p}(t) ds \right] dt \\
&= \int_0^T \mathbb{E} \left[ \int_s^T \frac{\partial \sigma}{\partial s}(t, s) D_s \dot{p}(t) dt \right] ds = \mathbb{E} \left[ \int_t^T \frac{\partial \sigma}{\partial s}(s, t) D_t \dot{p}(s) ds dt \right]
\end{align*}
\]

and

\[
\begin{align*}
\mathbb{E} \int_0^T \left( \int_0^t \int_\mathbb{R} \frac{\partial \gamma}{\partial t}(t, s, \zeta) N(ds, d\zeta) \dot{p}(t) \right) dt &= \int_0^T \mathbb{E} \left[ \int_t^T \int_\mathbb{R} \frac{\partial \gamma}{\partial s}(t, s, \zeta) N(ds, d\zeta) \dot{p}(t) \right] dt \\
&= \int_0^T \mathbb{E} \left[ \int_s^T \int_\mathbb{R} \frac{\partial \gamma}{\partial s}(t, s, \zeta) D_s \dot{p}(t) \nu(d\zeta) dt \right] ds \\
&= \mathbb{E} \int_0^T \int_t^T \frac{\partial \gamma}{\partial s}(s, t, \zeta) D_t \dot{p}(s) \nu(d\zeta) ds dt
\end{align*}
\]

\( (3.13) \)
Substituting (3.12) and (3.13) into (3.10) we get
\begin{align*}
I_2 &\leq \mathbb{E} \left[ \int_0^T \left\{ \hat{p}(t)[b(t) - \hat{b}(t)] + \hat{p}(t) \int_0^t \left( \frac{\partial b}{\partial t}(t,s) - \frac{\partial \hat{b}}{\partial t}(t,s) \right) ds 
\right. \\
&\quad + \int_0^T \left( \frac{\partial \sigma}{\partial s}(s,t) - \frac{\partial \hat{\sigma}}{\partial s}(s,t) \right) D_t\hat{p}(s) ds \\
&\quad \left. + \int_0^T \int_\mathbb{R} \left( \frac{\partial \gamma}{\partial s}(s,t,\zeta) - \frac{\partial \hat{\gamma}}{\partial s}(s,t,\zeta) \right) D_{t,\zeta}\hat{p}(s) \nu(d\zeta) 
\right] ds \\
&\quad + \hat{q}(t) [\sigma(t) - \hat{\sigma}(t)] + \int_\mathbb{R} \hat{r}(t,\zeta) [\gamma(t,\zeta) - \hat{\gamma}(t,\zeta)] \nu(d\zeta) dt \right]. 
\end{align*}
(3.14)

Adding (3.9) and (3.14) and using concavity, we get
\begin{align*}
J(u) - J(\hat{u}) &= I_1 + I_2 \\
&\leq \mathbb{E} \left[ \int_0^T \left\{ \mathcal{H}(t) - \hat{\mathcal{H}}(t) - \frac{\partial \hat{\mathcal{H}}}{\partial x}(t) \left( X(t) - \hat{X}(t) \right) \right\} dt 
\right] \\
&\leq \mathbb{E} \left[ \int_0^T \frac{\partial \hat{\mathcal{H}}}{\partial v}(t) (u(t) - \hat{u}(t)) dt \right] \\
&= \mathbb{E} \left[ \int_0^T \mathbb{E} \left[ \frac{\partial \hat{\mathcal{H}}}{\partial v}(t) (u(t) - \hat{u}(t)) \bigg| \mathcal{G}_t \right] dt \right] \\
&= \mathbb{E} \left[ \int_0^T \mathbb{E} \left[ \frac{\partial \hat{\mathcal{H}}}{\partial v}(t) \mathcal{G}_t \right] (u(t) - \hat{u}(t)) dt \right] \leq 0 \text{ by (3.7)}. 
\end{align*}

\section{A necessary maximum principle}

The sufficient maximum principle proved in the previous section has the drawback that the required concavity conditions are not always satisfied. It is therefore important also to have maximum principle which does not need this assumption. In the following result, a necessary maximum principle, the concavity conditions are replaced by conditions related to the space of admissible control and the existence of the derivative process. The details are as follows:

For each given \( t \in [0, T] \) let \( \alpha = \alpha_t \) be a bounded \( \mathcal{G}_t \)-measurable random variable and define
\begin{equation}
\beta(s) = \alpha \mathbb{1}_{[t,t+h]}(s); s \in [0, T] 
\end{equation}
(4.1)

Assume that
\begin{equation}
u + \beta \in \mathcal{A}
\end{equation}
(4.2)
for all such \( \nu \) and all \( u \in \mathcal{A} \), and that the derivative process \( Y(t) \) defined by
\begin{equation}
Y(t) = \frac{d}{d\lambda} X^{(u+\lambda\hat{\beta})}(t) \bigg|_{\lambda=0}
\end{equation}
(4.3)
exists.
Then we see that
\[ Y(t) = \int_0^t \left( \frac{\partial b}{\partial x}(t,s)Y(s) + \frac{\partial b}{\partial u}(t,s)\beta(s) \right) ds \]
\[ + \int_0^t \left( \frac{\partial \sigma}{\partial x}(t,s)Y(s) + \frac{\partial \sigma}{\partial u}(t,s)\beta(s) \right) dB(s) \]
\[ + \int_0^t \int_\mathbb{R} \left( \frac{\partial \gamma}{\partial x}(t,s,\zeta)Y(s) + \frac{\partial \gamma}{\partial u}(t,s,\zeta)\beta(s) \right) \tilde{N}(ds,d\zeta) \]
\[ (4.4) \]
and hence
\[ dY(t) = \left[ \frac{\partial b}{\partial x}(t,t)Y(t) + \frac{\partial b}{\partial u}(t,t)\beta(t) + \int_0^t \left( \frac{\partial^2 b}{\partial t\partial x}(t,s)Y(s) + \frac{\partial^2 b}{\partial t\partial u}(t,s)\beta(s) \right) ds \right. \]
\[ + \int_0^t \left( \frac{\partial^2 \sigma}{\partial t\partial x}(t,s)Y(s) + \frac{\partial^2 \sigma}{\partial t\partial u}(t,s)\beta(s) \right) dB(s) \]
\[ + \int_0^t \int_\mathbb{R} \left( \frac{\partial^2 \gamma}{\partial t\partial x}(t,s,\zeta)Y(s) + \frac{\partial^2 \gamma}{\partial t\partial u}(t,s,\zeta)\beta(s) \right) \tilde{N}(ds,d\zeta) \]
\[ \left. + \left( \frac{\partial \sigma}{\partial x}(t,t)Y(t) + \frac{\partial \sigma}{\partial u}(t,t)\beta(t) \right) dB(t) \right] \]
\[ + \int_\mathbb{R} \left( \frac{\partial \gamma}{\partial x}(t,t,\zeta)Y(t) + \frac{\partial \gamma}{\partial u}(t,t,\zeta)\beta(t) \right) \tilde{N}(dt,d\zeta). \]
\[ (4.5) \]

We are now ready to formulate the result

**Theorem 4.1 (Necessary maximum principle)**

Suppose that \( u \in \mathcal{A} \) is such that, for all \( \beta \) above,
\[ \frac{d}{d\lambda} J(u + \lambda \beta) \bigg|_{\lambda=0} = 0 \]
and the corresponding solution \( X(t), (p(t), q(t), r(t, \cdot)) \) of (1.1) and (2.1) exist.

Then
\[ \mathbb{E} \left[ \frac{\partial \mathcal{H}}{\partial u}(t) \bigg| \mathcal{G}_t \right]_{u=\hat{u}(t)} = 0. \]
\[ (4.7) \]

Conversely, if (4.7) holds, then (4.6) holds.

**Proof.**

By considering a suitable increasing family of stopping times converging to \( T \) we may assume that all the local martingales appearing in the proof below are martingales. See the proof of Theorem 2.1 in [12] for details. Now consider
\[ \frac{d}{d\lambda} J(u + \lambda \beta) \bigg|_{\lambda=0} \]
\[ = \mathbb{E} \left[ \int_0^T \left\{ \frac{\partial f}{\partial x}(t)Y(t) + \frac{\partial f}{\partial u}(t)\beta(t) \right\} dt + g'(X(T))Y(T) \right] \]
\[ (4.8) \]
Applying Itô formula to \( g'(X(T))Y(T) \), we get

\[
\mathbb{E}[g'(X(T))Y(T)] = \mathbb{E}[p(T)Y(T)]
\]

\[
= \mathbb{E} \left[ \int_0^T p(t) \left( \frac{\partial b}{\partial x}(t,t)Y(t) + \frac{\partial b}{\partial u}(t,t)\beta(t) \right) dt \right]
+ \int_0^T p(t) \left\{ \int_0^t \left( \frac{\partial^2 b}{\partial s^2}(t,s)Y(s) + \frac{\partial^2 b}{\partial s\partial u}(t,s)\beta(s) \right) ds \right\} dt
+ \int_0^T p(t) \left\{ \int_0^t \left( \frac{\partial^2 \sigma}{\partial s^2}(t,s,Y(s)) + \frac{\partial^2 \sigma}{\partial s\partial u}(t,s,Y(s)) \right) dB(s) \right\} dt
+ \int_0^T p(t) \left\{ \int_0^t \left( \frac{\partial^2 \sigma}{\partial s^2}(t,s,Y(s)) + \frac{\partial^2 \sigma}{\partial s\partial u}(t,s,Y(s)) \right) \tilde{N}(ds,d\zeta) \right\} dt
+ \int_0^T Y(t)\mathbb{E} \left[ \frac{\partial^2 H}{\partial t^2}(t) \big| \mathcal{F}_t \right] dt + \int_0^T q(t) \left( \frac{\partial b}{\partial x}(t,t,Y(t) + \frac{\partial b}{\partial u}(t,t)\beta(t) \right) dt
+ \int_0^T \int_\mathbb{R} r(t,\zeta) \left( \frac{\partial^\gamma}{\partial x}(t,t,\zeta)Y(t) + \frac{\partial^\gamma}{\partial u}(t,t,\zeta)\beta(t) \right) \nu(d\zeta) dt \right]
\]

From (3.11), (3.12) and (3.13), we have

\[
\mathbb{E}[p(T)Y(T)]
= \mathbb{E} \left[ \int_0^T \left( \frac{\partial b}{\partial x}(t,t)p(t) + \int_0^T \left( \frac{\partial^2 b}{\partial s^2}(s,t)p(s) + \frac{\partial^2 b}{\partial s\partial u}(s,t)D_t p(s) \right) ds \right) \right] Y(t) dt
+ \int_0^T \left\{ \frac{\partial b}{\partial u}(t,t)p(t) + \int_0^T \left( \frac{\partial^2 b}{\partial s^2}(s,t)p(s) + \frac{\partial^2 b}{\partial s\partial u}(s,t)D_t p(s) \right) ds \right\} \beta(t) dt
+ \int_0^T \int_\mathbb{R} r(t,\zeta) \left( \frac{\partial^\gamma}{\partial x}(t,t,\zeta)Y(t) + \frac{\partial^\gamma}{\partial u}(t,t,\zeta)\beta(t) \right) \nu(d\zeta) dt \right]
+ \int_0^T \int_\mathbb{R} \left( \frac{\partial^\gamma}{\partial x}(t,t,\zeta)Y(t) + \frac{\partial^\gamma}{\partial u}(t,t,\zeta)\beta(t) \right) r(t,\zeta) \nu(d\zeta) dt \right]
\]

Using the definition of \( \mathcal{H} \) in (3.3) and the definition of \( \beta \), we obtain

\[
\frac{d}{d\lambda} J(u + \lambda \beta) \bigg|_{\lambda=0} = \mathbb{E} \left[ \int_0^T \frac{\partial \mathcal{H}}{\partial u}(t)\beta(t) dt \right] = \mathbb{E} \left[ \int_0^{t+h} \frac{\partial \mathcal{H}}{\partial u}(t) dt \right].
\]  

(4.9)

Now suppose that

\[
\frac{d}{d\lambda} J(u + \lambda \beta) \bigg|_{\lambda=0} = 0.
\]  

(4.10)
Differentiating the right hand side of (4.9) at \( h = 0 \), we get

\[
\mathbb{E} \left[ \frac{\partial \mathcal{H}}{\partial u}(t) \alpha \right] = 0
\]

Since this holds for all bounded \( \alpha \), \( \mathcal{G}_t \)-measurable, we have

\[
\mathbb{E} \left[ \frac{\partial \mathcal{H}}{\partial u}(t) \bigg| \mathcal{G}_t \right] = 0. \tag{4.11}
\]

Conversely, if we assume that (4.11) holds, then we obtain (4.10) by using (4.9).

5 Applications

5.1 The case when the coefficients do not depend on \( x \)

Consider the case when the coefficients do not depend on \( x \), i.e., the system has the form:

\[
X(t) = \xi(t) + \int_0^t b(t, s, u(s))ds + \int_0^t \sigma(t, s, u(s))dB(s)
\]

\[
+ \int_0^t \int_\mathbb{R} \gamma(t, s, u(s), \zeta)\tilde{N}(ds, d\zeta)
\]

with performance functional

\[
J(u) = \mathbb{E} \left[ \int_0^T f(t, u(t))dt + g(X(T)) \right] \tag{5.2}
\]

In this case the Hamiltonian \( \mathcal{H} \) given in (2.3) takes the form

\[
\mathcal{H}(t, v, p(\cdot), q(\cdot), r(\cdot)) = f(t, v) + b(t, v)p(t) + \sigma(t, v)q(t) + \int_{\mathbb{R}} \gamma(t, v, \zeta)r(t, \zeta)\nu(d\zeta)
\]

\[
+ \int_t^T \frac{\partial b}{\partial s}(s, t, v)p(s)ds + \int_t^T \frac{\partial \sigma}{\partial s}(s, t, v)D_t p(s)ds
\]

\[
+ \int_t^T \int_{\mathbb{R}} \frac{\partial \gamma}{\partial s}(s, t, v, \zeta)D_{t, \zeta} p(s)\nu(d\zeta)ds \tag{5.3}
\]

The BSDE (2.4) for the adjoint variables \( p, q, r \) gets the form

\[
\begin{cases}
 dp(t) = q(t)dB(t) + \int_{\mathbb{R}} r(t, \zeta)\tilde{N}(dt, d\zeta) ; 0 \leq t \leq T \\
p(t) = g'(X(T))
\end{cases} \tag{5.4}
\]

which has the solution

\[
p(t) = \mathbb{E}[g'(X(T)) \mid \mathcal{F}_t] \tag{5.5}
\]

\[
q(t) = D_t p(t) = \mathbb{E}[D_t g'(X(T)) \mid \mathcal{F}_t] \tag{5.6}
\]

\[
r(t, \zeta) = D_{t, \zeta} p(t) = \mathbb{E}[D_{t, \zeta} g'(X(T)) \mid \mathcal{F}_t] \tag{5.7}
\]
Substituting (5.5)-(5.7) into (5.3) we get

\[ E[H(t, v(p(\cdot), q(\cdot)), r(\cdot)) \mid \mathcal{F}_t] = E[H_0(t, v, p, q, r) \mid \mathcal{F}_t] \]

where

\[ H_0(t, v, p, q, r) = f(t, v) + b(t, t, v)g'(X(T)) + \sigma(t, t, v)D_t g'(X(T)) \]
\[ + \int_{\mathbb{R}} \gamma(t, t, v, \zeta)D_{t, \zeta} g'(X(T)) \nu(d\zeta) \]
\[ + \int_t^T \frac{\partial b}{\partial s}(s, t, v)g'(X(T))ds + \int_t^T \frac{\partial \sigma}{\partial s}(s, t, v)D_t g'(X(T))ds \]
\[ + \int_t^T \int_{\mathbb{R}} \frac{\partial \gamma}{\partial s}(s, t, v, \zeta)D_{t, \zeta} g'(X(T)) \nu(d\zeta)ds \]  \hspace{1cm} (5.8)

Performing the ds-integrals we see that \( H_0(t, v, p, q, r) \) reduces to

\[ H_0(t, v, X(T)) = f(t, v) + b(T, t, v)g'(X(T)) + \sigma(T, t, v)D_t g'(X(T)) \]
\[ + \int_{\mathbb{R}} \gamma(T, t, v, \zeta)D_{t, \zeta} g'(X(T)) \nu(d\zeta) \]  \hspace{1cm} (5.9)

We conclude that in this case we have the following maximum principles:

**Theorem 5.1 (Sufficient maximum principle II)** Suppose that the coefficients \( f(t, v), b(t, s, v), \sigma(t, s, v) \) and \( \gamma(t, s, v, \zeta) \) of the stochastic control system (5.1)-(5.2) do not depend on \( x \).

Let \( \hat{u} \in \mathcal{A} \) with associated solution \( \hat{X} \) of (5.1). Suppose that the functions \( x \to g(x) \) and \( v \to H_0(t, v, X(T)) \) are concave and that, for all \( t \),

\[ \max_{v \in U} E[H_0(t, v, \hat{X}(T)) \mid \mathcal{G}_t] = E[H_0(t, \hat{u}(t), \hat{X}(T)) \mid \mathcal{G}_t]. \]  \hspace{1cm} (5.10)

Then \( \hat{u} \) is an optimal control, i.e.

\[ \sup_{u \in \mathcal{A}} J(u) = J(\hat{u}) \]  \hspace{1cm} (5.11)

**Theorem 5.2 (Necessary maximum principle II)** Let \( X(t) \) and \( J(u) \) be as in Theorem 4.1.

Let \( \hat{u} \in \mathcal{A} \) with associated solution \( \hat{X} \) of (5.1).

Then the following, (i) and (ii), are equivalent:

(i) \( \hat{u} \) is a critical point for \( J(u) \), i.e.

\[ \frac{d}{dy} J(\hat{u} + yw) \bigg|_{y=0} = 0 \]

for all \( w \in \mathcal{A} \) such that \( \hat{u} + yw \in \mathcal{A} \) for all \( y \) small enough.

(ii)

\[ E \left[ \frac{\partial H_0}{\partial v}(t, v, \hat{X}(T)) \mid \mathcal{G}_t \right]_{v=\hat{u}(t)} = 0 \]

**Remark 5.3** Theorem 5.2 is identical to Theorem 3.2 in [15]. However, the method in [15] is different, being based on perturbation techniques and complicated stochastic expansions. In the general case the necessary maximum principle of [15] is completely different from our Theorem 3.1. There is no corresponding sufficient maximum principle in [15].
5.2 Optimal investment in a financial market modeled by a Volterra equation

Consider a financial market with the following two investment possibilities:

(i) A risk free asset with unit price $S_0(t) = 1; t \geq 0$

(ii) A risky asset, in which investments have long term (memory) effects, in the following sense:

If we at time $s \geq 0$ decide to invest the fraction $\pi(s)$ of the current total wealth $X(s)$ in this asset, then we assume that the wealth $X(t) = X_\pi(t)$ at time $t$ is described by the linear stochastic Volterra equation

$$X(t) = x + \int_0^t b_0(t,s)\pi(s)X(s)ds + \int_0^t \sigma_0(t,s)\pi(s)X(s)dB(s); t \geq 0$$

(5.12)

or, in differential form,

$$\begin{aligned}
&dX(t) = b_0(t,t)\pi(t)X(t)dt + \sigma_0(t,t)\pi(t)X(t)dB(t) \\
&+ \left[ \int_0^t \frac{\partial b_0}{\partial t}(t,s)\pi(s)X(s)ds + \int_0^t \frac{\partial \sigma_0}{\partial t}(t,s)\pi(s)X(s)dB(s) \right] dt; t \geq 0 \\
&X(0) = x.
\end{aligned}$$

(5.13)

Thus we see that (5.13) differs from the classical Black-Scholes type of wealth equation by the last two integral terms on the right hand side. These terms represent long term (memory) effects of the investment strategy $\pi(\cdot)$.

We assume that $b_0(t,s) = b_0(t,s,\omega)$ and $\sigma_0(t,s) = \sigma_0(t,s,\omega)$ are given bounded processes, and that $b_0(t,s)$ and $\sigma_0(t,s)$ are $\mathcal{F}_s$-measurable for all $s,t$ and differentiable with respect to $t$ for all $s$, a.s with derivatives in $L^2(d\lambda \times d\lambda \times dP)$, where $\lambda$ denotes Lebesgue measure on $[0,T]$.

We choose $\mathcal{G} = \mathcal{F}$ in this example and we say that $\pi$ is admissible and write $\pi \in \mathcal{A}$ if $\pi$ is $\mathcal{F}$-adapted, $\pi \in L^2(d\lambda \times dP)$ and equation (5.12) has a unique solution with $\pi X \in L^2(d\lambda \times dP)$.

We assume that $x > 0$. If $\pi \in \mathcal{A}$ it follows that $X_\pi(t) > 0$ for all $t \in [0,T]$. To see this, note that from (5.13) we get

$$X_\pi(t) = x \exp \left( \int_0^t \sigma_0(s,s)\pi(s)dB(s) \\
+ \int_0^t \left\{ b_0(s,s)\pi(s) - \frac{1}{2} \sigma_0^2(s,s)\pi^2(s) + \alpha(s) \right\} ds \right) > 0,$$

(5.14)

where

$$\alpha(s) := \int_0^s \frac{\partial b_0}{\partial s}(s,r)\pi(r)X(r)dr + \int_0^s \frac{\partial \sigma_0}{\partial s}(s,r)\pi(r)X(r)dB(r).$$

We now study the following optimal investment problem:

Find $\hat{\pi} \in \mathcal{A}$ such that

$$\sup_{\pi \in \mathcal{A}} \mathbb{E} \left[ U(X_\pi(T)) \right] = \mathbb{E} \left[ U(X_{\hat{\pi}}(T)) \right]$$

(5.15)

where $U : (0,\infty) \rightarrow (-\infty, \infty)$ is a given utility function, assumed to be strictly increasing, $C^1$ and concave.

This is a control problem of the type studied in Section 2 and 3, and we apply the results from there: The Hamiltonian $\mathcal{H}$ given by (3.3) gets the form

$$\mathcal{H}(t,x,v,p,q) = b_0(t,t)xv + \sigma_0(t,t)xq$$

$$+ \int_t^T \frac{\partial b_0}{\partial s}(s,t)vx ds + \int_t^T \frac{\partial \sigma_0}{\partial s}(s,t)vxdP(s)ds$$

(5.16)
Suppose there exists an optimal control $\hat{\pi} \in A$ for (5.15) with corresponding $\hat{X}, \hat{p}, \hat{q}$. Then

$$E \left[ \frac{\partial}{\partial \pi} \mathcal{H}(t, \hat{X}(t), \pi, \hat{p}, \hat{q}) \right]_{\pi = \hat{\pi}(t)} = 0$$

i.e.

$$E \left[ b_0(t, t)\dot{X}(t)\dot{p}(t) + \sigma_0(t, t)\dot{X}(t)\dot{q}(t) + \int_t^T \frac{\partial b_0}{\partial x}(s, t)\dot{X}(t)\dot{p}(s) ds + \int_t^T \frac{\partial \sigma_0}{\partial x}(s, t)\dot{X}(t)D_t\dot{p}(s) ds \right] = 0.$$

Since $\dot{X}(t) > 0$ this is equivalent to

$$b_0(t, t)\dot{p}(t) + \sigma_0(t, t)\dot{q}(t) + \int_t^T \frac{\partial b_0}{\partial x}(s, t)\dot{p}(s) + \int_t^T \frac{\partial \sigma_0}{\partial x}(s, t)D_t\dot{p}(s) ds = 0. \quad (5.17)$$

We deduce that the corresponding BSDE (3.4) reduces to

$$\begin{cases}
  d\dot{p}(t) = \dot{q}(t)dB(t); 0 \leq t \leq T \\
  \dot{p}(T) = U'(\hat{X}(T)),
\end{cases} \quad (5.18)$$

which has the unique solution

$$\dot{p}(t) = E \left[ U'(\hat{X}(T)) \right]_{\mathcal{F}_t}, \dot{q}(t) = D_t\dot{p}(t). \quad (5.19)$$

Substituted into (5.19) this gives the equation

$$E \left[ b_0(t, t)U'(\hat{X}(T)) + \sigma_0(t, t)D_tU'(\hat{X}(T)) + \int_t^T \frac{\partial b_0}{\partial x}(s, t)U'(\hat{X}(T)) ds \right]_{\mathcal{F}_t} = 0 \quad (5.20)$$

where we have used that

$$D_tE \left[ U'(\hat{X}(T)) \right]_{\mathcal{F}_t} = E \left[ D_tU'(\hat{X}(T)) \right]_{\mathcal{F}_t}, \quad (5.21)$$

which is an identity that follows easily from the definition (2.14) of the Malliavin derivative. Equation (5.20) can be simplified to

$$b_0(t, t)E \left[ U'(\hat{X}(T)) \right]_{\mathcal{F}_t} + \sigma_0(t, t)E \left[ D_tU'(\hat{X}(T)) \right]_{\mathcal{F}_t}$$

$$+ E \left[ \int_t^T \frac{\partial b_0}{\partial x}(s, t)U'(\hat{X}(T)) ds \right]_{\mathcal{F}_t}$$

$$+ E \left[ \int_t^T \frac{\partial \sigma_0}{\partial x}(s, t)D_tU'(\hat{X}(T)) ds \right]_{\mathcal{F}_t} = 0. \quad (5.22)$$

or

$$\sigma_0(T, t)D_tE \left[ U'(\hat{X}(T)) \right]_{\mathcal{F}_t} + b_0(T, t)E \left[ U'(\hat{X}(T)) \right]_{\mathcal{F}_t} = 0. \quad (5.23)$$

Now assume that
\[ \sigma_0(t,s) > 0 \text{ for all } s,t \in [0,T] \]  

Then (5.23) can be written

\[ \frac{D_t Y(t)}{Y(t)} = -\frac{b_0(T,t)}{\sigma_0(T,t)} \]  

where

\[ Y(t) = \mathbb{E} \left[ U'(\hat{X}(T)) \right] \bigg| F_t \]  

By the chain rule for Malliavin derivatives, we deduce from (5.25) that

\[ D_t (\ln Y(t)) = -\frac{b_0(T,t)}{\sigma_0(T,t)} \]  

On the other hand, since \( Y(t) \) is a positive martingale there exists an adapted process \( \theta_0(t) \) such that

\[ dY(t) = \theta_0(t) Y(t) dB(t) \]

i.e.

\[ Y(t) = Y(0) \exp \left( \int_0^t \theta_0(s) dB(s) - \frac{1}{2} \int_0^t \theta_0^2(s) ds \right) \]  

From (5.28) we get

\[ D_t (\ln Y(t)) = D_t \left( \int_0^t \theta_0(s) dB(s) - \frac{1}{2} \int_0^t \theta_0^2(s) ds \right) = \theta_0(t), \]  

since

\[ D_t \theta_0(s) = D_t (\theta_0^2(s)) = 0 \text{ for all } s < t \]  

(because \( \theta_0 \) is adapted).

Comparing (5.27) and (5.29) we conclude that

\[ \theta_0(t) = -\frac{b_0(T,t)}{\sigma_0(T,t)} \]  

and hence, by (5.28),

\[ \mathbb{E} \left[ U'(\hat{X}(T)) \right] \bigg| F_t = Y(t) = \mathbb{E} \left[ U'(\hat{X}(T)) \right] \exp \left( \int_0^t \theta_0(s) dB(s) - \frac{1}{2} \int_0^t \theta_0^2(s) ds \right). \]  

It remains to find the constant

\[ c := \mathbb{E} \left[ U'(\hat{X}(T)) \right]. \]  

From (5.31) we get

\[ \hat{X}(T) = (U')^{-1}(c \exp \left( \int_0^T \theta_0(s) dB(s) - \frac{1}{2} \int_0^T \theta_0^2(s) ds \right)) =: F(c). \]  

On the other hand, if we define

\[ \hat{Z}_c(t,s) = \sigma_0(t,s)\hat{\pi}(s)\hat{X}(T) \]  

then by (5.12) the pair \( (\hat{X}, \hat{Z}_c) \) solves the backward stochastic Volterra equation (BSVE)
\begin{equation}
\hat{X}_c(t) = F(c) - \int_t^T \frac{b_0(t,s)\hat{Z}_c(t,s)}{\sigma_0(t,s)} ds - \int_t^T \hat{Z}_c(t,s) dB(s); \quad 0 \leq s \leq T. \tag{5.35}
\end{equation}

By theorem 3.2 in [19] the solution of this equation is unique. Putting \( t = 0 \) and taking expectation in (4.35) we get

\begin{equation}
x = F(c) - \int_0^T \mathbb{E} \left[ \frac{b_0(t,s)}{\sigma_0(t,s)} \hat{Z}_c(t,s) \right] ds. \tag{5.36}
\end{equation}

This equation determines implicitly the value of \( c \).

Hence by (5.35) we have found the optimal terminal wealth \( \hat{X}(T) \). Then finally we obtain the optimal portfolio \( \hat{\pi} \) by (5.34), assuming that \( \sigma_0(t,s) > 0 \) is bounded away from 0.

Conversely, since the functions \( x \to U(X) \) and \( (x,\pi) \to \mathcal{H}(t,x,\pi,\hat{\mu},\hat{\nu}) \) are concave, we see that \( \hat{\pi} \) found above satisfies the conditions of Theorem 2.1 and hence \( \hat{\pi} \) is indeed optimal.

We summarize what we have proved as follows:

**Theorem 5.4** Assume that \( \sigma_0(t,s) > 0 \) is bounded away from 0, for \( s, t \in [0,T] \).

Then the optimal portfolio \( \hat{\pi} \) for problem (5.15) is

\[ \hat{\pi}(s) = \frac{\hat{Z}_c(t,s)}{\sigma_0(t,s)\hat{X}(s)}; \quad s \in [0,T] \]

where \((\hat{X}, \hat{Z}_c)\) is the unique solution of the BSVIE (5.35) with \( F \) defined by (5.33), and the constant \( c \) is the solution of (5.36).

**References**

[1] F. Baghery and B. Øksendal: A maximum principle for stochastic control with partial information. Stoch. Anal. Appl. 2007, 25:706-717.

[2] S. A. Belbas: A new method for optimal control of Volterra integral equations. Appl. Math. Comput. 2007, 189: 1902-1915.

[3] K.R. Dahl, S.-E. A. Mohammed, B. Øksendal and E. R. Rose; Optimal control with noisy memory and BSDe with Malliavin derivatives. arXiv 1403.4034 (2014).

[4] G. Di Nunno, T. Meyer-Brandis, B. Øksendal and F. Proske: Malliavin calculus and anticipative Itô formulae for Lévy processes, Infinite Dimens. Anal. Quantum Probab. Rel. Topics, 2005, 8: 235-258.

[5] G. Di Nunno, B. Øksendal and F. Proske: Malliavin Calculus for Lévy processes with Applications to Finance. Springer 2009.

[6] G. Gripenberg, S.-O. Londen and O. Staffans: Volterra Integral and Functional Equations. Cambridge University Press 1990.

[7] H. Holden, B. Øksendal, J. Ubøe and T. Zhang: Stochastic Partial Differential Equations. Second Edition, Springer 2010.
[8] J. Lin: Adapted solution of backward stochastic nonlinear Volterra integral equation. Stoch. Anal. Appl. 2002, 20: 165-183.

[9] D. Nualart: The Malliavin Calculus and Related Topics. Second Edition, Springer 2006.

[10] B. Øksendal: Stochastic Differential Equations. Sixth Edition, Springer 2013.

[11] B. Øksendal and A. Sulem: Applied Stochastic Control of Jump Diffusions. Second Edition, Springer 2007.

[12] B. Øksendal and A. Sulem: Risk minimization in financial markets modeled by Itō-Lévy processes. arXiv 1402.3131, Afrika Matematika (to appear), DOI 10.1007/s13370-014-0248-9.

[13] B. Øksendal and T. Zhang: The stochastic Volterra equation. In D. Nualart, M.S. Solè (eds): Barcelona Seminar on Stochastic Analysis. Birkhauser 1993, pp.168-202.

[14] B. Øksendal and T. Zhang: The general linear stochastic Volterra equations with anticipating coefficients. In Davis, Truman and Elworthy (eds): Stochastic Analysis and Applications. World Scientific 1996, pp. 343-366.

[15] B. Øksendal and T. Zhang: Optimal control with partial information for stochastic Volterra equations. Intern. J. Stoch. Anal. 2010, doi:10.1115/2010/329185.

[16] M. Santz-Solé: Malliavin Calculus, with Applications to Stochastic Partial Differential Equations. EPFL Press/CRC Press 2005.

[17] Y. Shi, T. Wang and J. Yong: Mean-field backward stochastic Volterra integral equations. arXiv:1104.4725v2 (2011). Discrete and Continuous Dynamical Systems Series B 18 (2013), 1929-1967.

[18] W. Tianxiao, Z. Qingfeng and S. Yufeng: Necessary and sufficient conditions of optimality for stochastic integral systems with partial information. Proceedings of the 30th Chinese Control Conference. July 22-24, 2011, Yantai, China.

[19] J. Yong: Backward stochastic Volterra integral equations and some related problems. Stochastic Proc. Appl. 2006, 116: 779–795.

[20] J. Yong: Well-posedness and regularity of backward stochastic Volterra integral equation. Probab. Theory Relat. Fields 2008, 142: 21-77.