Malliavin Calculus and Optimal Control of Stochastic Volterra Equations

Nacira Agram [*] and Bernt Øksendal [†‡§]

11 February 2015

Abstract

Solutions of stochastic Volterra (integral) equations are not Markov processes, and therefore classical methods, like dynamic programming, cannot be used to study optimal control problems for such equations. However, we show 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.

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]; and for an example stemming from Newtonian motion in a random environment, see Exercise 5.12 in [11]. Stochastic Volterra equations can also be derived from stochastic *delay* equations. See [16] and the references 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, 14] and [15], 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 [20, 19, 16] and [18], the authors have obtained different types of maximum principles

---

[*]Faculty of economic sciences and management, University Med Khider, Po. Box 145, Biskra (07000) Algeria. Email: agramnacira@yahoo.fr

[†]Corresponding author

[‡]Dept. of Mathematics, University of Oslo, Box 1053 Blindern, N-0316 Oslo, Norway. Email: oksendal@math.uio.no

[§]Norwegian School of Economics (NHH), Helleveien 30, N-5045 Bergen, Norway.

for stochastic Volterra equations. In [20] 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. (A special class of backward stochastic Volterra equations without control had earlier been studied in [9].) An extension of the result of [20] to mean-field equations is obtained in [18]. In [19] 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 [16] 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 (in some way) a standard BSDE, not a Volterra type BSVIE as in [20, 19] and [18]. 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 SDEs with *noisy memory*. See [3].

Our sufficient maximum principle is new, even in the case without jumps. In our general setting also our necessary maximum principle is new. However, 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 [16]. For more general systems our maximum principle is simpler than the one in [16]. Moreover, there is no sufficient maximum principle in [16].

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 $(\Omega, \mathcal{F}, \mathbb{F} := \{\mathcal{F}_t\}_{0 \leq t \leq T}, P)$ satisfying the usual conditions, P is a reference probability measure and ν is the Lévy measure of N. We refer to [12] 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 $\mathbb{G} = \{\mathcal{G}_t\}_{t \geq 0}$ 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

$$\mathcal{G}_t = \mathcal{F}_{(t-\delta)^+} \qquad \text{(delayed information flow)}.$$

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

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

(1.1)

where $b(t, s, x, v) = b(t, s, x, v, \omega) : [0, T] \times [0, T] \times \mathbb{R} \times \mathbb{U} \times \Omega \mapsto \mathbb{R}, \sigma(t, s, x, v) = \sigma(t, s, x, v, \omega) : [0, T] \times [0, T] \times \mathbb{R} \times \mathbb{U} \times \Omega \mapsto \mathbb{R}$ and $\gamma(t, s, x, v, \zeta) = \gamma(t, s, x, v, \zeta, \omega) : [0, T] \times [0, T] \times \mathbb{R} \times \mathbb{U} \times \Omega \times \mathbb{R}_0 \mapsto \mathbb{R}$ are given functions, assumed to be $\mathbb{F}$-predictable with respect to the second variable s for all t, x, v, ζ and continuously

2

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 $\mathbb{R}_0 = \mathbb{R} - \{0\}$, and $\mathbb{U}$ denotes a given open set containing all possible admissible control values $u(t, \omega)$ for $(t, \omega) \in [0, T] \times \Omega, u \in \mathcal{A}$.

The *performance functional* is given by

$$J(u) = \mathbb{E}\left[\int_0^T f(s, X(s), u(s)) \ ds + g(X(T))\right]; u \in \mathcal{A}. \tag{1.2}$$

where $f(s, x, v) = f(s, x, v, \omega) : [0, T] \times \mathbb{R} \times \mathbb{U} \times \Omega \mapsto \mathbb{R}$ and $g(x) = g(x, \omega) : \mathbb{R} \times \Omega \mapsto \mathbb{R}$ are given random functions, f is adapted and g is $\mathcal{F}_T$ - measurable and C^1 with respect to x, and we assume that $J(u)$ exists for all $u \in \mathcal{A}$. The problem we study is the following:

Problem: Find $u^* \in \mathcal{A}$ such that

$$J(u^*) = \sup_{u \in \mathcal{A}} J(u). \tag{1.3}$$

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], [8], [10] and [17].

In view of the Lévy–Itô decomposition Theorem, which states that any Lévy process $Y(t)$ with

$$\mathbb{E}[Y^2(t)] < \infty \quad \text{for all} \quad t$$

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

$$\eta(\cdot) := \int_0 \int_{\mathbb{R}} \zeta \tilde{N}(ds, d\zeta)$$

separately.

2.1 Malliavin Calculus for $B(\cdot)$

A natural starting point is the Wiener-Itô chaos expansion Theorem, which states that any $F \in L^2(\mathcal{F}_T, P)$ can be written

$$F = \sum_{n=0}^{\infty} I_n(f_n) \tag{2.1}$$

for a unique sequence of symmetric deterministic functions $f_n \in L^2(\lambda^n)$, where λ is Lebesgue measure on $[0, T]$ and

$$I_n(f_n) = n! \int_0^T \int_0^{t_n} \cdots \int_0^{t_2} f_n(t_1, \cdots, t_n) dB(t_1) dB(t_2) \cdots dB(t_n) \tag{2.2}$$

(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

$$\mathbb{E}[F^2] = ||F||^2_{L^2(p)} = \sum_{n=0}^{\infty} n! ||f_n||^2_{L^2(\lambda^n)}. \tag{2.3}$$

DEFINITION 2.1 (MALLIAVIN DERIVATIVE D_t WITH RESPECT TO $B(\cdot)$)
Let $\mathbb{D}_{1,2}^{(B)}$ be the space of all $F \in L^2(\mathcal{F}_T, P)$ such that its chaos expansion (2.1) satisfies

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

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

$$D_t F = \sum_{n=1}^{\infty} n I_{n-1}(f_n(\cdot, t)), \tag{2.5}$$

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

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

so $(t, \omega) \to D_t F(\omega)$ 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

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

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

$$D_t \Big(\int_0^T u(s) \delta B(s) \Big) = \int_0^T D_t u(s) \delta B(s) + u(t) \text{ for a.a. } (t, \omega), \tag{2.7}$$

where $\int_0^T \psi(s) \delta B(s)$ denotes the Skorohod integral of a process ψ 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 [10], page 29)
   Suppose $F_1, \ldots, F_m \in \mathbb{D}_{1,2}^{(B)}$ and that $\psi : \mathbb{R}^m \to \mathbb{R}$ is C^1 with bounded partial derivatives. Then, $\psi(F_1, \cdots, F_m) \in \mathbb{D}_{1,2}$ and

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

4

(ii) **Duality formula**

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

$$\mathbb{E}[F \int_0^T u(t)dB(t)] = \mathbb{E}[\int_0^T u(t)D_tFdt]. \tag{2.9}$$

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

If φ is an $\mathbb{F}$-adapted process, then

$$D_s\varphi(t) = 0 \text{ for } s > t.$$

Remark 2.1 *We put $D_t\varphi(t) = \lim_{s \to t-} D_s\varphi(t)$ (if the limit exists).*

Remark 2.2 *It was proved in [1] that one can extend the Malliavin derivative operator D_t from $\mathbb{D}_{1,2}$ to all of $L^2(\mathcal{F}_T, P)$ in such a way that, also denoting the extended operator by D_t, for all $F \in L^2(\mathcal{F}_T, P)$ we have*

$$D_tF \in (\mathcal{S})^* \text{ and } (t, \omega) \mapsto \mathbb{E}[D_tF \mid \mathcal{F}_t] \text{ belongs to } L^2(\lambda \times P) \tag{2.10}$$

Here $(\mathcal{S})^*$ is the Hida space of stochastic distributions.

Moreover, the following generalized Clark-Haussmann-Ocone formula *was proved:*

$$F = \mathbb{E}[F] + \int_0^T \mathbb{E}[D_tF \mid \mathcal{F}_t]dB(t) \tag{2.11}$$

for all $F \in L^2(\mathcal{F}_T, P)$. See Theorem 3.11 in [1] and also Theorem 6.35 in [5].

As also noted in [3] we can use this to get the following extension of the duality formula (2.9):

Proposition 2.3 The generalized duality formula

Let $F \in L^2(\mathcal{F}_T, P)$ and let $\varphi(t, \omega) \in L^2(\lambda \times P)$ be adapted. Then

$$\mathbb{E}[F \int_0^T \varphi(t)dB(t)] = \mathbb{E}[\int_0^T \mathbb{E}[D_tF \mid \mathcal{F}_t]\varphi(t)dt] \tag{2.12}$$

Proof. By (2.10) and (2.11) and the Itô isometry we get

$$\mathbb{E}[F \int_0^T \varphi(t)dB(t)] = \mathbb{E}[(\mathbb{E}[F] + \int_0^T \mathbb{E}[D_tF \mid \mathcal{F}_t]dB(t))(\int_0^T \varphi(t)dB(t))]$$

$$= \mathbb{E}[\int_0^T \mathbb{E}[D_tF \mid \mathcal{F}_t]\varphi(t)dt]. \tag{2.13}$$

$\square$

It is this extension of the Malliavin derivative we will use from now on.

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 $\eta(s) := \int_0^s \int_{\mathbb{R}_0} \zeta \tilde{N}(dr, d\zeta); \ 0 \le s \le t$) can be written as

$$F = \sum_{n=0}^{\infty} I_n(f_n); \ f_n \in \hat{L}^2((\lambda \times \nu)^n), \tag{2.14}$$

where $\hat{L}^2((\lambda \times \nu)^n)$ 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

$$I_n(f_n) := n! \int_0^T \int_{\mathbb{R}_0} \int_0^{t_n} \int_{\mathbb{R}_0} \cdots \int_0^{t_2} \int_{\mathbb{R}_0} f_n(t_1, \zeta_1, \cdots, t_n, \zeta_n) \tilde{N}(dt_1, d\zeta_1) \cdots \tilde{N}(dt_n, d\zeta_n) \qquad (2.15)$$

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

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

$$||F||^2_{L^2(P)} = \sum_{n=0}^{\infty} n! ||f_n||^2_{L^2((\lambda \times \nu)^n)}. \qquad (2.16)$$

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, ζ, where t represents time and $\zeta \neq 0$ represents a generic jump size.

DEFINITION 2.3 (MALLIAVIN DERIVATIVE $D_{t,\zeta}$ WITH RESPECT TO $\tilde{N}(\cdot, \cdot)$) [5] Let $\mathbb{D}_{1,2}^{(\tilde{N})}$ be the space of all $F \in L^2(\mathcal{F}_T, P)$ such that its chaos expansion (2.14) satisfies

$$||F||^2_{\mathbb{D}_{1,2}^{(\tilde{N})}} := \sum_{n=1}^{\infty} nn! ||f_n||^2_{L^2((\lambda \times \nu)^2)} < \infty. \qquad (2.17)$$

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

$$D_{t,\zeta}F := \sum_{n=1}^{\infty} n I_{n-1}(f_n(\cdot, t, \zeta)), \qquad (2.18)$$

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 $(t_1, \zeta_1), \cdots, (t_n, \zeta_n)$, keeping $(t_n, \zeta_n) = (t, \zeta)$ as a parameter.

In this case, we get the isometry.

$$\mathbb{E}[\int_0^T \int_{\mathbb{R}_0} (D_{t,\zeta}F)^2 \nu(d\zeta) dt] = \sum_{n=0}^{\infty} nn! ||f_n||^2_{L^2((\lambda \times \nu)^n)} = ||F||^2_{\mathbb{D}_{1,2}^{(\tilde{N})}}. \qquad (2.19)$$

(Compare with (2.6)).

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

$$D_{t,\zeta}F = f(t, \zeta) \text{ for a.a. } (t, \zeta).$$

More generally, if $\psi(s, \zeta)$ is Skorohod integrable with respect to $\tilde{N}(\delta s, d\zeta)$, $\psi(s, \zeta) \in \mathbb{D}_{1,2}^{(\tilde{N})}$ for a.a. s, ζ and $D_{t,z}\psi(s, \zeta)$ is Skorohod integrable for a.a. (t, z), then

$$D_{t,z}\left(\int_0^T \int_{\mathbb{R}} \psi(s, \zeta) \tilde{N}(\delta s, d\zeta)\right) = \int_0^T \int_{\mathbb{R}} D_{t,z}\psi(s, \zeta) \tilde{N}(\delta s, d\zeta) + u(t, z) \text{ for a.a. } t, z, \qquad (2.20)$$

where $\int_0^T \int_{\mathbb{R}} \psi(s, \zeta) \tilde{N}(\delta s, d\zeta)$ denotes the *Skorohod integral* of ψ with respect to $\tilde{N}(\cdot, \cdot)$. (See [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 [5]**

Suppose $F_1, \cdots, F_m \in \mathbb{D}_{1,2}^{(\tilde{N})}$ and that $\phi : \mathbb{R}^m \to \mathbb{R}$ is continuous and bounded. Then, $\phi(F_1, \cdots, F_m) \in \mathbb{D}_{1,2}^{(\tilde{N})}$ and

$$D_{t,\zeta}\phi(F_1, \cdots, F_m) = \phi(F_1 + D_{t,\zeta}F_1, \ldots, F_m + D_{t,\zeta}F_m) - \phi(F_1, \ldots, F_m). \tag{2.21}$$

(ii) **Duality formula [5]**

Suppose $\Psi(t,\zeta)$ is $\mathcal{F}_t$-adapted and $\mathbb{E}[\int_0^T \int_{\mathbb{R}_0} \psi^2(t,\zeta)\nu(d\zeta)dt] < \infty$ and let $F \in \mathbb{D}_{1,2}^{(\tilde{N})}$. Then,

$$\mathbb{E}\Big[F \int_0^T \int_{\mathbb{R}_0} \Psi(t,\zeta)\tilde{N}(dt,d\zeta)\Big] = \mathbb{E}\Big[\int_0^T \int_{\mathbb{R}_0} \Psi(t,\zeta)D_{t,\zeta}F\nu(d\zeta)dt\Big]. \tag{2.22}$$

(iii) **Malliavin derivative and adapted processes [5]**

If φ is an $\mathbb{F}$-adapted process, then,

$$D_{s,\zeta}\varphi(t) = 0 \text{ for all } s > t.$$

Remark 2.4 *We put $D_{t,\zeta}\varphi(t) = \lim_{s\to t-} D_{s,\zeta}\varphi(t)$ (if the limit exists).*

Remark 2.5 *As in Remark 2.2 we note that there is an extension of the Malliavin derivative $D_{t,\zeta}$ from $\mathbb{D}_{1,2}^{(\tilde{N})}$ to $L^2(\mathcal{F}_t \times P)$ such that the following extension of the duality theorem holds:*

***Proposition 2.6* Generalized duality formula**

Suppose $\Psi(t,\zeta)$ is $\mathcal{F}_t$-adapted and $\mathbb{E}[\int_0^T \int_{\mathbb{R}_0} \psi^2(t,\zeta)\nu(d\zeta)dt] < \infty$ and let $F \in L^2(\mathcal{F}_T \times P)$. Then,

$$\mathbb{E}\Big[F \int_0^T \int_{\mathbb{R}_0} \Psi(t,\zeta)\tilde{N}(dt,d\zeta)\Big] = \mathbb{E}\Big[\int_0^T \int_{\mathbb{R}_0} \Psi(t,\zeta)\mathbb{E}[D_{t,\zeta}F \mid \mathcal{F}_t]\nu(d\zeta)dt\Big]. \tag{2.23}$$

We refer to Theorem 13.26 in [5].

Accordingly, note that from now on we are working with this generalized version of the Malliavin derivative. We emphasize that this generalized Malliavin derivative DX exists for all $X \in L^2(P)$ as an element of the Hida stochastic distribution space $(\mathcal{S})*$, and it has the property that the conditional expectation $\mathbb{E}[DX|\mathcal{F}_t]$ belongs to $L^2(\lambda \times P)$, where λ is Lebesgue measure on $[0,T]$. Therefore, using this generalized Malliavin derivative, combined with conditional expectation, no assumptions on Malliavin differentiability in the classical sense are needed; we can work on the whole space of random variables in $L^2(P)$.

3   A Sufficient Maximum Principle

Let $\mathcal{L}$ and $\mathcal{L}_\zeta$ be the set of all stochastic processes with parameter space $[0,T]$ and $[0,T] \times \mathbb{R}_0$, respectively. Define the *Hamiltonian functionals*

$$H_0 : [0,T] \times \mathbb{R} \times \mathbb{U} \times \mathbb{R} \times \mathbb{R} \times \mathcal{R} \mapsto \mathbb{R}$$

and

$$H_1 : [0,T] \times \mathbb{R} \times \mathbb{U} \times \mathbb{R} \times \mathcal{L} \times \mathcal{L}_\zeta \mapsto \mathbb{R}$$

by

$$H_0(t, x, v, p, q, r) := f(t, x, v) + b(t, t, x, v)p + \sigma(t, t, x, v)q + \int_{\mathbb{R}} \gamma(t, t, x, v, \zeta)r(\zeta)\nu(d\zeta) \qquad (3.1)$$

and

$$H_1(t, x, v, p, D_t p(\cdot), D_{t,\zeta} p(\cdot)) := \int_t^T \frac{\partial b}{\partial s}(s, t, x, v)p(s)ds + \int_t^T \frac{\partial \sigma}{\partial s}(s, t, x, v)\mathbb{E}[D_t p(s) \mid \mathcal{F}_t]ds$$

$$+ \int_t^T \int_{\mathbb{R}} \frac{\partial \gamma}{\partial s}(s, t, x, v, \zeta)\mathbb{E}[D_{t,\zeta} p(s) \mid \mathcal{F}_t]\nu(d\zeta)ds. \qquad (3.2)$$

Here, $\mathcal{R}$ denotes the set of all functions $r(\cdot) : \mathbb{R} \to \mathbb{R}$ such that the last integral above converges. We may regard $x, p, q, r = r(\cdot)$ as generic values for the processes $X(t), p(t), q(t), r(t, \cdot)$, respectively (see below).

Define

$$\mathcal{H}(t, x, v, p(\cdot), q(\cdot), r(\cdot)) := H_0(t, x, v, p, q, r) + H_1(t, x, v, p(\cdot), D_t p(\cdot), D_{t,\zeta} p(\cdot)). \qquad (3.3)$$

The BSDE for the adjoint processes $p(t), q(t), r(t, \cdot)$ is defined by

$$\begin{cases} dp(t) := -\frac{\partial \mathcal{H}}{\partial x}(t)dt + q(t)dB(t) + \int_{\mathbb{R}} r(t, \zeta)\tilde{N}(dt, d\zeta); 0 \le t \le T \\ p(T) := g'(X(T)), \end{cases} \qquad (3.4)$$

where we have used the simplified notation

$$\frac{\partial \mathcal{H}}{\partial x}(t) = \frac{\partial \mathcal{H}}{\partial x}(t, X(t), u(t), p(\cdot), q(\cdot), r(\cdot)). \qquad (3.5)$$

Note that from (1.1) we get

$$dX(t) = \xi'(t)dt + b\left(t, t, X(t), u(t)\right)dt + \left(\int_0^t \frac{\partial b}{\partial t}\left(t, s, X(s), u(s)\right)ds\right)dt$$

$$+ \sigma\left(t, t, X(t), u(t)\right)dB(t) + \left(\int_0^t \frac{\partial \sigma}{\partial t}\left(t, s, X(s), u(s)\right)dB(s)\right)dt \qquad (3.6)$$

$$+ \int_{\mathbb{R}} \gamma\left(t, t, X(t), u(t), \zeta\right)\tilde{N}(dt, d\zeta) + \left(\int_0^t \int_{\mathbb{R}} \frac{\partial \gamma}{\partial t}\left(t, s, X(s), u(s), \zeta\right)\tilde{N}(ds, d\zeta)\right)dt.$$

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.*

For fixed $t, \hat{p}(\cdot), \hat{q}(\cdot), \hat{r}(\cdot)$ define

$$\kappa(x, v) = \mathcal{H}(t, x, v, \hat{p}(\cdot), \hat{q}(\cdot), \hat{r}(\cdot)) \qquad (3.7)$$

Assume the following:

- *(The Arrow condition) The function*

$$x \to \hat{\kappa}(x) := \sup_{v \in \mathbb{U}} \kappa(x, v) \qquad (3.8)$$

  exists and is concave.

- *The function*

$$x \to g(x) \text{ is concave.}$$

8

- *(The maximum condition)*

$$\sup_{v \in \mathbb{U}} \mathbb{E}\left[\mathcal{H}(t, \hat{X}(t), v, \hat{p}(\cdot), \hat{q}(\cdot), \hat{r}(\cdot)) \mid \mathcal{G}_t\right]$$

$$= \mathbb{E}\left[\mathcal{H}(t, \hat{X}(t), \hat{u}(t), \hat{p}(\cdot), \hat{q}(\cdot), \hat{r}(\cdot)) \mid \mathcal{G}_t\right] \forall t. \tag{3.9}$$

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 [13] 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

$$I_1 = \mathbb{E}\left[\int_0^T \left\{f(t) - \hat{f}(t)\right\} dt\right], \quad I_2 = \mathbb{E}\left[g(X(T)) - g(\hat{X}(T))\right], \tag{3.10}$$

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

$$b(t) = b(t, t, X(t), u(t)), \hat{b}(t) = b(t, t, \hat{X}(t), \hat{u}(t)) \text{ etc.,}$$

we get

$$I_1 = \mathbb{E}[\int_0^T \{H_0(t, X(t), u(t), \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot))$$

$$- H_0(t, \hat{X}(t), \hat{u}(t), \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot))$$

$$- [b(t) - \hat{b}(t)]\hat{p}(t) - [\sigma(t) - \hat{\sigma}(t)]\hat{q}(t) - \int_{\mathbb{R}} [\gamma(t, \zeta) - \hat{\gamma}(t, \zeta)]\hat{r}(t, \zeta)\nu(d\zeta)\}dt]. \tag{3.11}$$

Using concavity and the Itô Formula, we obtain

$$I_2 \leq \mathbb{E}\left[g'(\hat{X}(T))\left(X(T) - \hat{X}(T)\right)\right]$$

$$= \mathbb{E}\left[\hat{p}(T)\left(X(T) - \hat{X}(T)\right)\right]$$

$$= \mathbb{E}\left[\int_0^T \left\{\hat{p}(t)\left(b(t) - \hat{b}(t) + \int_0^t \left(\frac{\partial b}{\partial t}(t, s) - \frac{\partial \hat{b}}{\partial t}(t, s)\right) ds\right.\right.$$

$$+ \int_0^t \left(\frac{\partial \sigma}{\partial t}(t, s) - \frac{\partial \hat{\sigma}}{\partial t}(t, s)\right) dB(s)$$

$$+ \int_0^t \int_{\mathbb{R}} \left(\frac{\partial \gamma}{\partial t}(t, s, \zeta) - \frac{\partial \hat{\gamma}}{\partial t}(t, s, \zeta)\right) \tilde{N}(ds, d\zeta)\right)$$

$$- \frac{\partial \hat{\mathcal{H}}}{\partial x}(t)\left(X(t) - \hat{X}(t)\right)$$

$$+ \hat{q}(t)\left[\sigma(t) - \hat{\sigma}(t)\right] + \int_{\mathbb{R}} \hat{r}(t, \zeta)\left[\gamma(t, \zeta) - \hat{\gamma}(t, \zeta)\right]\nu(d\zeta)\right\} dt\right]. \tag{3.12}$$

9

By the Fubini Theorem, we get

$$\int_0^T (\int_0^t \frac{\partial b}{\partial t}(t,s)ds)\hat{p}(t)dt = \int_0^T (\int_s^T \frac{\partial b}{\partial t}(t,s)\hat{p}(t)dt)ds = \int_0^T (\int_t^T \frac{\partial b}{\partial s}(s,t)\hat{p}(s)ds)dt, \qquad (3.13)$$

and similarly, by the duality theorems,

$$\mathbb{E}[\int_0^T (\int_0^t \frac{\partial \sigma}{\partial t}(t,s)dB(s))\hat{p}(t)dt] = \int_0^T \mathbb{E}[\int_0^t \frac{\partial \sigma}{\partial t}(t,s)dB(s)\hat{p}(t)]dt = \int_0^T \mathbb{E}[\int_0^t \frac{\partial \sigma}{\partial t}(t,s)\mathbb{E}[D_s\hat{p}(t) \mid \mathcal{F}_s]ds]dt$$

$$= \int_0^T \mathbb{E}[\int_s^T \frac{\partial \sigma}{\partial t}(t,s)\mathbb{E}[D_s\hat{p}(t) \mid \mathcal{F}_s]dt]ds = \mathbb{E}[\int_0^T \int_t^T \frac{\partial \sigma}{\partial s}(s,t)\mathbb{E}[D_t\hat{p}(s) \mid \mathcal{F}_t]dsdt]$$

$$(3.14)$$

and

$$\mathbb{E}[\int_0^T (\int_0^t \int_\mathbb{R} \frac{\partial \gamma}{\partial t}(t,s,\zeta)\tilde{N}(ds,d\zeta))\hat{p}(t)dt] = \int_0^T \mathbb{E}[\int_0^t \int_\mathbb{R} \frac{\partial \gamma}{\partial t}(t,s,\zeta)\tilde{N}(ds,d\zeta))\hat{p}(t)]dt$$

$$= \int_0^T \mathbb{E}[\int_0^t \int_\mathbb{R} \frac{\partial \gamma}{\partial t}(t,s,\zeta)\mathbb{E}[D_{s,\zeta}\hat{p}(t) \mid \mathcal{F}_s]\nu(d\zeta)ds]dt = \int_0^T \mathbb{E}[\int_s^T \int_\mathbb{R} \frac{\partial \gamma}{\partial t}(t,s,\zeta)\mathbb{E}[D_{s,\zeta}\hat{p}(t) \mid \mathcal{F}_s]\nu(d\zeta)dt]ds$$

$$= \mathbb{E}[\int_0^T \int_t^T \int_\mathbb{R} \frac{\partial \gamma}{\partial s}(s,t,\zeta)\mathbb{E}[D_{t,\zeta}\hat{p}(s) \mid \mathcal{F}_t]\nu(d\zeta)dsdt].$$

$$(3.15)$$

Substituting (3.14) and (3.15) into (3.12), we get

$$I_2 \leq \mathbb{E}\left[\int_0^T \left\{ \hat{p}(t)[b(t) - \hat{b}(t)] + \hat{p}(t)\int_0^t (\frac{\partial b}{\partial t}(t,s) - \frac{\partial \hat{b}}{\partial t}(t,s))ds \right.\right.$$

$$+ \int_t^T (\frac{\partial \sigma}{\partial s}(s,t) - \frac{\partial \hat{\sigma}}{\partial s}(s,t))\mathbb{E}[D_t\hat{p}(s) \mid \mathcal{F}_t]ds$$

$$+ \int_t^T \int_\mathbb{R} (\frac{\partial \gamma}{\partial s}(s,t,\zeta) - \frac{\partial \hat{\gamma}}{\partial s}(s,t,\zeta))\mathbb{E}[D_{t,\zeta}\hat{p}(s) \mid \mathcal{F}_t]ds\nu(d\zeta)$$

$$- \frac{\partial \hat{\mathcal{H}}}{\partial x}(t)\left(X(t) - \hat{X}(t)\right)$$

$$\left.\left. + \hat{q}(t)[\sigma(t) - \hat{\sigma}(t)] + \int_\mathbb{R} \hat{r}(t,\zeta)[\gamma(t,\zeta) - \hat{\gamma}(t,\zeta)]\nu(d\zeta) \right\} dt \right]. \qquad (3.16)$$

Adding (3.11) and (3.16), we get

$$J(u) - J(\hat{u}) = I_1 + I_2 \qquad (3.17)$$

$$\leq \mathbb{E}\left[\int_0^T \left\{ \mathcal{H}(t) - \hat{\mathcal{H}}(t) - \frac{\partial \hat{\mathcal{H}}}{\partial x}(t)\left(X(t) - \hat{X}(t)\right) \right\} dt \right].$$

Finally, to prove that

$$\mathcal{H}(t) - \hat{\mathcal{H}}(t) - \frac{\partial \hat{\mathcal{H}}}{\partial x}(t)\left(X(t) - \hat{X}(t)\right) \leq 0, \qquad (3.18)$$

we use the Arrow condition (3.8) and the maximum condition (3.9), proceeding as in the proof of Corollary 11.3 in [12]. We omit the details. $\qquad \square$

4   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 a 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 controls 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, let $h \in [T - t, T]$ and define

$$\beta(s) := \alpha 1_{[t,t+h]}(s); s \in [0, T]. \tag{4.1}$$

Assume that

$$u + \lambda\beta \in \mathcal{A} \tag{4.2}$$

for all such α and all $u \in \mathcal{A}$, and all non-zero λ sufficiently small. Assume that the *derivative process* $Y(t)$, defined by

$$Y(t) = \frac{d}{d\lambda} X^{(u+\lambda\beta)}(t)\Big|_{\lambda=0}, \tag{4.3}$$

exists.

Then we see that

$$Y(t) = \int_0^t \left(\frac{\partial b}{\partial x}(t, s)Y(s) + \frac{\partial b}{\partial u}(t, s)\beta(s) \right) ds$$

$$+ \int_0^t \left(\frac{\partial \sigma}{\partial x}(t, s)Y(s) + \frac{\partial \sigma}{\partial u}(t, s)\beta(s) \right) dB(s)$$

$$+ \int_0^t \int_{\mathbb{R}} \left(\frac{\partial \gamma}{\partial x}(t, s, \zeta)Y(s) + \frac{\partial \gamma}{\partial u}(t, s, \zeta)\beta(s) \right) \tilde{N}(ds, d\zeta), \tag{4.4}$$

and hence

$$dY(t) = \left[\frac{\partial b}{\partial x}(t, t)Y(t) + \frac{\partial b}{\partial u}(t, t)\beta(t) + \int_0^t \left(\frac{\partial^2 b}{\partial t \partial x}(t, s)Y(s) + \frac{\partial^2 b}{\partial t \partial u}(t, s)\beta(s) \right) ds \right.$$

$$+ \int_0^t \left(\frac{\partial^2 \sigma}{\partial t \partial x}(t, s)Y(s) + \frac{\partial^2 \sigma}{\partial t \partial u}(t, s)\beta(s) \right) dB(s)$$

$$\left. + \int_0^t \int_{\mathbb{R}} \left(\frac{\partial^2 \gamma}{\partial t \partial x}(t, s, \zeta)Y(s) + \frac{\partial^2 \gamma}{\partial t \partial u}(t, s, \zeta)\beta(s) \right) \tilde{N}(ds, d\zeta) \right] dt$$

$$+ \left(\frac{\partial \sigma}{\partial x}(t, t)Y(t) + \frac{\partial \sigma}{\partial u}(t, t)\beta(t) \right) dB(t)$$

$$+ \int_{\mathbb{R}} \left(\frac{\partial \gamma}{\partial x}(t, t, \zeta)Y(t) + \frac{\partial \gamma}{\partial u}(t, t, \zeta)\beta(t) \right) \tilde{N}(dt, d\zeta). \tag{4.5}$$

We are now ready to formulate the result

Theorem 4.1 (Necessary maximum principle) *Suppose that $\hat{u} \in \mathcal{A}$ is such that, for all β as in (4.1),*

$$\frac{d}{d\lambda} J(\hat{u} + \lambda\beta)\Big|_{\lambda=0} = 0 \tag{4.6}$$

and the corresponding solution $\hat{X}(t), (\hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot))$ of (1.1) and (3.4) exists.

Then,

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

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

Proof.

For simplicity of notation we drop the "hat" everywhere and write u in stead of $\hat{u}$, X in stead of $\hat{X}$ etc in the following.

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 [13] for details.

Now consider

$$\frac{d}{d\lambda} J(u + \lambda\beta)\Big|_{\lambda=0}$$
$$= \mathbb{E}\left[\int_0^T \left\{\frac{\partial f}{\partial x}(t)Y(t) + \frac{\partial f}{\partial u}(t)\beta(t)\right\} dt + g'(X(T))Y(T)\right]. \tag{4.8}$$

Applying the Itô Formula, we get

$$\mathbb{E}\left[g'(X(T))Y(T)\right] = \mathbb{E}\left[p(T)Y(T)\right]$$
$$= \mathbb{E}\left[\int_0^T p(t)\left(\frac{\partial b}{\partial x}(t,t)Y(t) + \frac{\partial b}{\partial u}(t,t)\beta(t)\right) dt\right.$$
$$+ \int_0^T p(t)\left\{\int_0^t \left(\frac{\partial^2 b}{\partial t \partial x}(t,s)Y(s) + \frac{\partial^2 b}{\partial t \partial u}(t,s)\beta(s)\right) ds\right\} dt$$
$$+ \int_0^T p(t)\left\{\int_0^t \left(\frac{\partial^2 \sigma}{\partial t \partial x}(t,s)Y(s) + \frac{\partial^2 \sigma}{\partial t \partial u}(t,s)\beta(s)\right) dB(s)\right\} dt$$
$$+ \int_0^T p(t)\left\{\int_0^t \int_{\mathbb{R}} \left(\frac{\partial^2 \gamma}{\partial t \partial x}(t,s,\zeta)Y(s) + \frac{\partial^2 \gamma}{\partial t \partial u}(t,s,\zeta)\beta(s)\right) \tilde{N}(ds, d\zeta)\right\} dt$$
$$- \int_0^T Y(t)\frac{\partial \mathcal{H}}{\partial x}(t)dt + \int_0^T q(t)\left(\frac{\partial \sigma}{\partial x}(t,t)Y(t) + \frac{\partial \sigma}{\partial u}(t,t)\beta(t)\right) dt$$
$$+ \int_0^T \int_{\mathbb{R}} r(t,\zeta)\left(\frac{\partial \gamma}{\partial x}(t,t,\zeta)Y(t) + \frac{\partial \gamma}{\partial u}(t,t,\zeta)\beta(t)\right) \nu(d\zeta)dt\left.\right].$$

From (3.13), (3.14) and (3.15), we have

$$
\begin{aligned}
&\mathbb{E}\left[p(T)Y(T)\right] \\
&= \mathbb{E}\left[\int_0^T \left\{ \frac{\partial b}{\partial x}(t,t)p(t) + \int_t^T \left(\frac{\partial^2 b}{\partial s \partial x}(s,t)p(s) + \frac{\partial^2 \sigma}{\partial s \partial x}(s,t)\mathbb{E}[D_t p(s) \mid \mathcal{F}_t] \right.\right.\right.\\
&\quad \left.\left. + \int_{\mathbb{R}} \frac{\partial^2 \gamma}{\partial s \partial x}(s,t,\zeta)\mathbb{E}[D_{t,\varsigma}p(s) \mid \mathcal{F}_t]\nu(d\zeta) \right) ds \right\} Y(t)dt \\
&\quad + \int_0^T \left\{ \frac{\partial b}{\partial u}(t,t)p(t) + \int_t^T \left(\frac{\partial^2 b}{\partial s \partial u}(s,t)p(s) + \frac{\partial^2 \sigma}{\partial s \partial u}(s,t)\mathbb{E}[D_t p(s) \mid \mathcal{F}_t] \right.\right.\\
&\quad \left.\left. + \int_{\mathbb{R}} \frac{\partial^2 \gamma}{\partial s \partial u}(s,t,\zeta)\mathbb{E}[D_{t,\varsigma}p(s) \mid \mathcal{F}_t]\nu(d\zeta) \right) ds \right\} \beta(t)dt \\
&\quad - \int_0^T \frac{\partial \mathcal{H}}{\partial x}(t)Y(t)dt + \int_0^T \left(\frac{\partial \sigma}{\partial x}(t,t)Y(t) + \frac{\partial \sigma}{\partial u}(t,t)\beta(t) \right) q(t)dt \\
&\quad + \int_0^T \int_{\mathbb{R}} \left(\frac{\partial \gamma}{\partial x}(t,t,\zeta)Y(t) + \frac{\partial \gamma}{\partial u}(t,t,\zeta)\beta(t) \right) r(t,\zeta)\nu(d\zeta)dt \Bigg].
\end{aligned}
$$

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

$$
\frac{d}{d\lambda}J(u+\lambda\beta)\Big|_{\lambda=0} = \mathbb{E}\left[\int_0^T \frac{\partial \mathcal{H}}{\partial u}(s)\beta(s)ds\right] = \mathbb{E}\left[\int_t^{t+h} \frac{\partial \mathcal{H}}{\partial u}(s)ds\,\alpha\right]. \tag{4.9}
$$

Now suppose that

$$
\frac{d}{d\lambda}J(u+\lambda\beta)\Big|_{\lambda=0} = 0. \tag{4.10}
$$

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 $\mathcal{G}_t$-measurable α, we have

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

Conversely, if we assume that (4.11) holds, then we obtain (4.10) by reversing the argument we used to obtain (4.9).

$\square$

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:

$$
X(t) = \xi(t) + \int_0^t b(t,s,u(s))ds + \int_0^t \sigma(t,s,u(s))dB(s) \tag{5.1}
$$

13

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

with performance functional

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

This special case is of interest because any linear SDE with delay can be written on this form. See [16] and the references therein.

In this case the Hamiltonian $\mathcal{H}$ given in (3.3) takes the form

$$
\begin{aligned}
\mathcal{H}&(t, v, p(\cdot), q(\cdot), r(\cdot)) \\
=\ & f(t, v) + b(t, t, v)p(t) + \sigma(t, t, v)q(t) + \int_{\mathbb{R}} \gamma(t, t, v, \zeta)r(t, \zeta)\nu(d\zeta) \\
& + \int_t^T \frac{\partial b}{\partial s}(s, t, v)p(s)ds + \int_t^T \frac{\partial \sigma}{\partial s}(s, t, v)\mathbb{E}[D_t p(s) \mid \mathcal{F}_t]ds \\
& + \int_t^T \int_{\mathbb{R}} \frac{\partial \gamma}{\partial s}(s, t, v, \zeta)\mathbb{E}[D_{t,\zeta} p(s) \mid \mathcal{F}_t]\nu(d\zeta)ds. \tag{5.3}
\end{aligned}
$$

The BSDE (3.4) for the adjoint variables p, q, r gets the form

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

which has the solution

$$
\begin{aligned}
p(t) &= \mathbb{E}[g'(X(T)) \mid \mathcal{F}_t] & (5.5) \\
q(t) &= D_t p(t) = \mathbb{E}[D_t g'(X(T)) \mid \mathcal{F}_t] & (5.6) \\
r(t, \zeta) &= D_{t,\zeta} p(t) = \mathbb{E}[D_{t,\zeta} g'(X(T)) \mid \mathcal{F}_t]. & (5.7)
\end{aligned}
$$

Substituting (5.5)-(5.7) into (5.3) we get

$$\mathbb{E}[\mathcal{H}(t, v, p(\cdot), q(\cdot), r(\cdot)) \mid \mathcal{F}_t] = \mathbb{E}[\mathcal{H}_0(t, v, p, q, r) \mid \mathcal{F}_t],$$

where

$$
\begin{aligned}
\mathcal{H}_0(t, v, p, q, r) =\ & f(t, v) + b(t, t, v)g'(X(T)) + \sigma(t, t, v)\mathbb{E}[D_t g'(X(T)) \mid \mathcal{F}_t] \\
& + \int_{\mathbb{R}} \gamma(t, t, v, \zeta)\mathbb{E}[D_{t,\zeta} g'(X(T)) \mid \mathcal{F}_t]\nu(d\zeta) \\
& + \int_t^T \frac{\partial b}{\partial s}(s, t, v)g'(X(T))ds + \int_t^T \frac{\partial \sigma}{\partial s}(s, t, v)\mathbb{E}[D_t g'(X(T)) \mid \mathcal{F}_t]ds \\
& + \int_t^T \int_{\mathbb{R}} \frac{\partial \gamma}{\partial s}(s, t, v, \zeta)\mathbb{E}[D_{t,\zeta} g'(X(T)) \mid \mathcal{F}_t]\nu(d\zeta)ds. \tag{5.8}
\end{aligned}
$$

Performing the ds-integrals we see that $\mathcal{H}_0(t, v, p, q, r)$ reduces to

$$\mathbb{H}_0(t, v, X(T)) \ :\ = f(t, v) + b(T, t, v)g'(X(T))$$

14

$$+\sigma(T, t, v)\mathbb{E}[D_t g'(X(T)) \mid \mathcal{F}_t]$$
$$+ \int_{\mathbb{R}} \gamma(T, t, v, \zeta)\mathbb{E}[D_{t,\zeta} g'(X(T)) \mid \mathcal{F}_t]\nu(d\zeta). \tag{5.9}$$

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

$$x \to g(x) \tag{5.10}$$

and

$$v \to \mathbb{H}_0(t, v, \hat{X}(T)) \tag{5.11}$$

are concave and that, for all t,

$$\max_{v \in \mathbb{U}} \mathbb{E}\left[\mathbb{H}_0(t, v, \hat{X}(T)) \mid \mathcal{G}_t\right] = \mathbb{E}\left[\mathbb{H}_0(t, \hat{u}(t), \hat{X}(T)) \mid \mathcal{G}_t\right]. \tag{5.12}$$

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

$$\sup_{u \in \mathcal{A}} J(u) = J(\hat{u}). \tag{5.13}$$

Theorem 5.2 (Necessary maximum principle II) *Let $X(t)$ and $J(u)$ be as in Theorem 5.1.*

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

(ii)

$$\mathbb{E}\left[\frac{\partial \mathbb{H}_0}{\partial v}(t, v, \hat{X}(T)) \mid \mathcal{G}_t\right]_{v=\hat{u}(t)} = 0.$$

Remark 5.3 *Theorem 5.2 is identical to Theorem 3.2 in [16]. However, the method in [16] is different, being based on perturbation techniques and complicated stochastic expansions. In the general case the necessary maximum principle of [16] is completely different from our Theorem 3.1. There is no corresponding* sufficient *maximum principle in [16].*

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

$$X(t) = x + \int_0^t b_0(t,s)\pi(s)X(s)ds + \int_0^t \sigma_0(t,s)\pi(s)X(s)dB(s); t \geq 0 \tag{5.14}$$

or, in differential form,

$$\begin{cases} dX(t) = b_0(t,t)\pi(t)X(t)dt + \sigma_0(t,t)\pi(t)X(t)dB(t) \\ + \left[\int_0^t \frac{\partial b_0}{\partial t}(t,s)\pi(s)X(s)ds + \int_0^t \frac{\partial \sigma_0}{\partial t}(t,s)\pi(s)X(s)dB(s) \right] dt; t \geq 0 \\ X(0) = x. \end{cases} \tag{5.15}$$

Thus we see that (5.15) 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 bounded derivatives with respect to t for all s, a.s.

We also assume that

$$\sigma_0(t,s) \geq c_0 \text{ a.s. for all } t, s \in [0,T] \text{ for some constant } c_0 > 0. \tag{5.16}$$

We choose $\mathbb{G} = \mathbb{F}$ in this example and we say that π is admissible and write $\pi \in \mathcal{A}$ if π is $\mathbb{F}$-adapted, $\pi \in L^2(d\lambda \times dP)$ and equation (5.14) has a unique solution with $\pi X \in L^2(d\lambda \times dP)$.

We assume that $x > 0$. If $\pi \in \mathcal{A}$, then it follows that $X_\pi(t) > 0$ for all $t \in [0,T]$. To see this, note that from (5.15) we get

$$\begin{aligned} X_\pi(t) = x \exp &\left(\int_0^t \sigma_0(s,s)\pi(s)dB(s) \right. \\ &\left. + \int_0^t \left\{ b_0(s,s)\pi(s) - \tfrac{1}{2}\sigma_0^2(s,s)\pi^2(s) + \alpha(s) \right\} ds \right) > 0, \end{aligned} \tag{5.17}$$

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

$$\sup_{\pi \in \mathcal{A}} \mathbb{E}\left[U(X_\pi(T)) \right] = \mathbb{E}\left[U(X_{\hat{\pi}}(T)) \right], \tag{5.18}$$

where $U : [0,\infty) \rightarrow [-\infty,\infty)$ is a given *utility* function, assumed to be strictly increasing, concave, and C^1 on $(0,\infty)$. This is a control problem of the type studied in Section 3 and 4, and we apply the results from there:

The Hamiltonian $\mathcal{H}$ given by (3.3) gets the form

$$\begin{aligned} \mathcal{H}(t,x,\pi,p,q) = &b_0(t,t)\pi x p + \sigma_0(t,t)\pi x q \\ &+ \int_t^T \frac{\partial b_0}{\partial s}(s,t)\pi x p(s)ds + \int_t^T \frac{\partial \sigma_0}{\partial s}(s,t)\pi x \mathbb{E}[D_t p(s) \mid \mathcal{F}_t]ds \end{aligned} \tag{5.19}$$

Suppose there exists an optimal control $\hat{\pi} \in \mathcal{A}$ for (5.18) with corresponding $\hat{X}, \hat{p}, \hat{q}$. Then,

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

i.e.,

$$\mathbb{E}\left[b_0(t, t)\hat{X}(t)\hat{p}(t) + \sigma_0(t, t)\hat{X}(t)\hat{q}(t)\right.$$
$$\left. + \int_t^T \frac{\partial b_0}{\partial s}(s, t)\hat{X}(t)\hat{p}(s)ds + \int_t^T \frac{\partial \sigma_0}{\partial s}(s, t)\hat{X}(t)\mathbb{E}[D_t\hat{p}(s) \mid \mathcal{F}_t]ds \Big| \mathcal{F}_t\right] = 0.$$

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

$$b_0(t, t)\hat{p}(t) + \sigma_0(t, t)\hat{q}(t)$$
$$+ \mathbb{E}\left[\int_t^T \left\{\frac{\partial b_0}{\partial s}(s, t)\hat{p}(s) + \int_t^T \frac{\partial \sigma_0}{\partial s}(s, t)\mathbb{E}[D_t\hat{p}(s) \mid \mathcal{F}_t]\right\} ds \Big| \mathcal{F}_t\right] = 0. \tag{5.20}$$

We deduce that the corresponding BSDE (3.4) reduces to

$$\begin{cases} d\hat{p}(t) = \hat{q}(t)dB(t); 0 \le t \le T \\ \hat{p}(T) = U'(\hat{X}(T)), \end{cases} \tag{5.21}$$

which has the unique solution

$$\hat{p}(t) = \mathbb{E}\left[U'(\hat{X}(T)) \Big| \mathcal{F}_t\right], \hat{q}(t) = D_t\hat{p}(t). \tag{5.22}$$

Substituted into (5.20), this gives the equation

$$\mathbb{E}\left[b_0(t, t)U'(\hat{X}(T)) + \sigma_0(t, t)D_tU'(\hat{X}(T))\right.$$
$$+ \int_t^T \frac{\partial b_0}{\partial s}(s, t)\mathbb{E}\left[U'(\hat{X}(T)) \Big| \mathcal{F}_s\right] ds \tag{5.23}$$
$$\left. + \int_t^T \frac{\partial \sigma_0}{\partial s}(s, t)\mathbb{E}\left[D_tU'(\hat{X}(T)) \Big| \mathcal{F}_s\right] ds \Big| \mathcal{F}_t\right] = 0,$$

where we have used that

$$D_t\mathbb{E}\left[U'(\hat{X}(T)) \Big| \mathcal{F}_t\right] = \mathbb{E}\left[D_tU'(\hat{X}(T)) \Big| \mathcal{F}_t\right], \tag{5.24}$$

which is an identity that follows easily from the definition (2.5) of the Malliavin derivative. Equation (5.23) can be simplified to

$$b_0(t, t)\mathbb{E}\left[U'(\hat{X}(T)) \Big| \mathcal{F}_t\right] + \sigma_0(t, t)\mathbb{E}\left[D_tU'(\hat{X}(T)) \Big| \mathcal{F}_t\right]$$
$$+ \mathbb{E}\left[\int_t^T \frac{\partial b_0}{\partial s}(s, t)U'(\hat{X}(T))ds \Big| \mathcal{F}_t\right] \tag{5.25}$$
$$+ \mathbb{E}\left[\int_t^T \frac{\partial \sigma_0}{\partial s}(s, t)D_tU'(\hat{X}(T))ds \Big| \mathcal{F}_t\right] = 0,$$

or

$$\sigma_0(T, t)D_t\mathbb{E}\left[U'(\hat{X}(T)) \Big| \mathcal{F}_t\right] + b_0(T, t)\mathbb{E}\left[U'(\hat{X}(T)) \Big| \mathcal{F}_t\right] = 0. \tag{5.26}$$

By (5.16) we see that (5.26) can be written

$$\frac{D_t Y(t)}{Y(t)} = -\frac{b_0(T,t)}{\sigma_0(T,t)}, \tag{5.27}$$

where

$$Y(t) = \mathbb{E}\left[U'(\hat{X}(T)) \middle| \mathcal{F}_t\right]. \tag{5.28}$$

By the chain rule for Malliavin derivatives, we deduce from (5.28) that

$$D_t(\ln Y(t)) = -\frac{b_0(T,t)}{\sigma_0(T,t)}. \tag{5.29}$$

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.,

$$Y(t) = Y(0)\exp(\int_0^t \theta_0(s)dB(s) - \frac{1}{2}\int_0^t \theta_0^2(s)ds). \tag{5.30}$$

From (5.30) we get

$$D_t(\ln Y(t)) = D_t(\int_0^t \theta_0(s)dB(s) - \frac{1}{2}\int_0^t \theta_0^2(s)ds) = \theta_0(t), \tag{5.31}$$

since

$$D_t\theta_0(s) = D_t(\theta_0^2(s)) = 0 \text{ for all } s < t \text{ (because } \theta_0 \text{ is adapted)}.$$

Comparing (5.29) and (5.31) we conclude that

$$\theta_0(t) = -\frac{b_0(T,t)}{\sigma_0(T,t)} \tag{5.32}$$

and hence, by (5.28),

$$\mathbb{E}\left[U'(\hat{X}(T)) \middle| \mathcal{F}_t\right] = Y(t) = \mathbb{E}\left[U'(\hat{X}(T))\right]\exp(\int_0^t \theta_0(s)dB(s) - \int_0^t \frac{1}{2}\theta_0^2(s)ds). \tag{5.33}$$

It remains to find the constant

$$c := \mathbb{E}\left[U'(\hat{X}(T))\right]. \tag{5.34}$$

From (5.33) with $t = T$ we get

$$\hat{X}(T) = (U')^{-1}(c\exp(\int_0^T \theta_0(s)dB(s) - \frac{1}{2}\int_0^T \theta_0^2(s)ds)) =: F(c). \tag{5.35}$$

On the other hand, if we define

$$\hat{Z}_c(t,s) := \sigma_0(t,s)\hat{\pi}(s)\hat{X}(s), \tag{5.36}$$

then by (5.14), the pair $(\hat{X}, \hat{Z}_c)$ solves the following *(Yong type) backward stochastic Volterra integral equation (BSVIE)*

$$\hat{X}(t) = F(c) - \int_t^T \frac{b_0(t,s)\hat{Z}_c(t,s)}{\sigma_0(t,s)}ds - \int_t^T \hat{Z}_c(t,s)dB(s); 0 \leq s \leq T. \tag{5.37}$$

18

By Theorem 3.2 in [20] the solution of this equation is unique. Putting $t = 0$ and taking expectation in (5.37), we get

$$x = \mathbb{E}[F(c)] - \int_0^T \mathbb{E}\left[\frac{b_0(t,s)}{\sigma_0(t,s)}\hat{Z}_c(t,s)\right]ds. \tag{5.38}$$

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 (5.36).

Conversely, since the functions $x \to U(x)$ and $(x,\pi) \to \mathcal{H}(t,x,\pi,\hat{p},\hat{q})$ are concave, we see that $\hat{\pi}$ found above satisfies the conditions of Theorem 3.1, 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 the problem (5.18) *is*

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

where $(\hat{X}, \hat{Z}_c)$ is the unique solution of the BSVIE (5.37) *with F defined by* (5.35), *and the constant c is the solution of* (5.38).

6   Conclusions

In this paper we study the problem of optimal control of a system described by a stochastic Volterra equation with jumps. We use Malliavin calculus to obtain both a sufficient and a necessary maximum principle for optimal control of such systems, with 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 (in some way) a standard BSDE, not a Volterra type BSVIE as in [20, 19] and [18]. 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 SDEs with *noisy memory*. See [3].

Our sufficient maximum principle is new, even in the case without jumps. In our general setting also our necessary maximum principle is new. However, 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 [16]. For more general systems our maximum principle is simpler than the one in [16]. Moreover, there is no sufficient maximum principle in [16].

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

References

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

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

[3] K.R. Dahl, S.-E. A. Mohammed, B. Øksendal and E. R. Røse; Optimal control with noisy memory and BSDEs 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, 8 (2005), 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] Y. Ishikawa: Stochastic Calculus of Variations for Jump Processes. De Greuter 2013.

[9] J. Lin: Adapted solution of backward stochastic nonlinear Volterra integral equation. Stoch. Anal. Appl. 20 (2002), 165-183.

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

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

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

[13] B. Øksendal and A. Sulem: Risk minimization in financial markets modeled by Itô-Lévy processes. arXiv: 1402.3131, Afrika Matematika (2014), DOI 10.1007/s 13370-014-0248-9.

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

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

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

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

[18] 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.

[19] 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.

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

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