

A Sufficient Stochastic Maximum Principle for Optimal Control of Jump Diffusions and Applications to Finance¹

N.C. FRAMSTAD² B. ØKSENDAL³ A. SULEM⁴

Revised in May 2005

Abstract

We give a verification theorem by employing Arrow's generalization of the Mangasarian sufficient condition to a general jump diffusion setting, and show the adjoint processes' connections to dynamic programming. The result is applied to financial optimization problems.

Key words and phrases: Jump diffusions, optimal control, sufficient maximum principle, mean-variance portfolio selection.

MSC (2000): 93E20, 60G51, 60H10, 60H30, 60J75, 91B28.

JEL classification: C61, D81, G11.

¹ The first author gratefully acknowledges financial support from the Research Council of Norway and the hospitality of the University of Kansas. This paper has benefited from comments from Jan Ube.

² Department of Mathematics, University of Oslo, Norway

³ Professor, Department of Mathematics, University of Oslo, Blindern, Oslo, Norway and Norwegian School of Economics and Business Administration, Bergen, Norway.

⁴ Member of the research institute INRIA, Domaine de Voluceau, Rocquencourt, Le Chesnay, France.

0 Introduction

The purpose of this paper is to prove a sufficient maximum principle for the optimal control of jump diffusions, show its connections to dynamic programming and give applications to financial optimization problems in a market described by such processes.

For *diffusions* (without jumps) the maximum principle has been studied by many researchers. To the best of our knowledge the earliest versions of a maximum principle for such process were given by Kushner [11] and Bismut [5]. Further progress on the subject were subsequently given by Bensoussan [3], Haussmann [7], Peng [12] and Cadenillas and Haussmann [6]; see also Yong and Zhou [18] and the references therein. For diffusions with jumps, a necessary maximum principle was given by Tang and Li [17]; see also Kabanov [8] and Kohlmann [9]. We consider general jump diffusions and prove a *sufficient* maximum principle for such processes. In other words, we prove a verification theorem saying that if a stochastic control satisfies the maximum principle conditions, then the control is indeed optimal for the stochastic control problem (Theorem 1.1). We believe that such a result constitutes a useful alternative to a verification theorem based on dynamic programming, since the latter involves a complicated integro-differential equation (the Hamilton-Jacobi-Bellman equation) in the jump diffusion case. We also show the connections to dynamic programming as interpretations of the adjoint processes (Section 2).

As an illustration of our sufficient maximum principle we use it to solve a mean-variance portfolio selection problem and a consumption-portfolio optimisation problem in a jump diffusion market.

1 The sufficient maximum principle

Suppose the state $X(t) = X^{(u)}(t)$ of a controlled jump diffusion in \mathbf{R}^n is given by

$$\begin{aligned} dX(t) &= b(t, X(t), u(t)) dt + \sigma(t, X(t), u(t)) dB(t) \\ &+ \int_{\mathbf{R}^n} \eta(t, X(t^-), u(t^-), z) \bar{N}(dt, dz); \quad t \in [0, T] \quad (T \text{ constant.}) \end{aligned} \quad (1)$$

Here $b : [0, T] \times \mathbf{R}^n \times U \mapsto \mathbf{R}^n$, $\sigma : [0, T] \times \mathbf{R}^n \times U \mapsto \mathbf{R}^{n \times m}$ and $\eta : [0, T] \times \mathbf{R}^n \times U \times \mathbf{R}^n \mapsto \mathbf{R}^{n \times \ell}$ are given continuous functions, and $B(t) = B(t, \omega)$; $t \geq 0$, $\omega \in \Omega$ is an m -dimensional standard Brownian motion (Wiener process.) $\bar{N}(dt, dz) = (\bar{N}_1(dt, dz), \dots, \bar{N}_\ell(dt, dz))^\top = (N_1(dt, dz) - \chi_1 d\lambda_1(z) dt, \dots, N_\ell(dt, dz) - \chi_\ell d\lambda_\ell(z) dt)^\top$ where the $\{N_j\}$

are independent $\mathbf{R}^{\ell \times 1}$ -valued Poisson random measures with Lévy measures λ_j respectively, on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$ satisfying the usual conditions. The truncation functions $\chi_j(z)$ serve the purpose of “compensating for small jumps” and are taken to be 1 for “small” jumps and 0 for “large”; the condition will be given implicitly in terms of existence of solution, but it may be convenient to put $\chi_j = 0$ (i.e. $d\bar{N}_j = dN_j$) or $\chi_j = 1$ (i.e. $d\bar{N}_j = d\tilde{N}_j = dN_j - d\lambda_j dt$, a *compensated* Poisson measure) whenever the appropriate integrability conditions hold. In this formulation, we will need the $\ell \times \ell$ matrix $I - \text{Diag}(\chi)$ where $\text{Diag}(\chi)$ is diagonal with χ as entries and I is the identity.

The process $u(t) = u(t, \omega)$; $t \in [0, T]$, $\omega \in \Omega$ is our *control process*. We assume that $u(t, \omega)$ has values in a given closed set $U \subset \mathbf{R}^k$ and that $u(t, \omega)$ is predictable, left-continuous and with right limits. In addition we require that $u(t, \omega)$ gives rise to a unique, strong solution $X(t) = X^{(u)}(t)$ of (1) for $t \in [0, T]$. Such controls are called *admissible* if also (3) below holds. The set of all admissible controls is denoted by \mathcal{A} . If $u \in \mathcal{A}$ and $X(t) = X^{(u)}(t)$ is the corresponding solution of (1), we call (u, X) an *admissible pair*.

We consider a *performance criterion* $J(u)$ of the form

$$J(u) = \mathbb{E} \left[\int_0^T f(t, X(t), u(t)) dt + g(X(T)) \right]; \quad u \in \mathcal{A} \quad (2)$$

where $f : [0, T] \times \mathbf{R}^n \times U \mapsto \mathbf{R}$ is continuous and $g : \mathbf{R}^n \mapsto \mathbf{R}$ is *concave*. As mentioned above, we require

$$\mathbb{E} \left[\int_0^T |f(t, X(t), u(t))| dt + \max\{0, g^-(X(T))\} \right] < \infty, \quad (3)$$

for u to be $\in \mathcal{A}$. The problem is to maximize $J(u)$ over all $u \in \mathcal{A}$, i.e. we seek $u^* \in \mathcal{A}$ such that

$$J(u^*) = \sup_{u \in \mathcal{A}} J(u). \quad (4)$$

Such controls u^* are called *optimal* controls. If $X^* = X^{(u^*)}$ is the corresponding solution of (1) then (X^*, u^*) is called an *optimal pair*.

Define the *Hamiltonian* $H : [0, T] \times \mathbf{R}^n \times U \times \mathbf{R}^n \times \mathbf{R}^{n \times m} \times \mathcal{R} \mapsto \mathbf{R}$ by

$$\begin{aligned} H(t, x, u, p, q, r) &= f(t, x, u) + b^\top(t, x, u)p + \text{tr}(\sigma^\top(t, x, u)q) \\ &+ \int_{\mathbf{R}^n} \left[\sum_{j=1}^{\ell} \left(\sum_{i=1}^n \eta_{ij}(t, x, u, z) r_{ij}(t, z) \right) \right. \\ &\quad \left. + (\eta^\top(t, x, u, z)p + x^\top r(t, z))(I - \text{Diag}(\chi)) \right] d\lambda(z) \end{aligned} \quad (5)$$

where \mathcal{R} is the set of functions $r : \mathbf{R}^{n+1} \mapsto \mathbf{R}^{n \times \ell}$ such that the integral in (5) converges. The *adjoint equation* (corresponding to the admissible pair (u, X)) in the unknown adapted processes $p(t) \in \mathbf{R}^n$, $q(t) \in \mathbf{R}^{n \times m}$ and $r(t, z) \in \mathbf{R}^{n \times \ell}$ is the backward stochastic differential equation (BSDE)

$$dp(t) = -\nabla_x H(t, X(t), u(t), p(t), q(t), r(t, \cdot)) dt + q(t) dB(t) + \int_{\mathbf{R}^n} r(t^-, z) \bar{N}(dt, dz) \quad (6)$$

with *terminal condition*

$$p(T) = \nabla g(X(T)), \quad (7)$$

where we for simplicity have assumed H differentiable w.r.t. x . As this paper concerns sufficient conditions, we can assume ad hoc that a solution exists; the reader is referred to [14] and [2] for BSDEs driven by both Brownian noise and Poisson jumps, while [1] treats the semimartingale case. We then have the following verification theorem for optimality:

Theorem 1.1. *Let (\hat{u}, \hat{X}) be an admissible pair and suppose there exists an adapted solution $(\hat{p}(t), \hat{q}(t), \hat{r}(t, z))$ of the corresponding adjoint equation (6) such that for all $u \in \mathcal{A}$, we have*

$$\begin{aligned} \mathbb{E} \left[\int_0^T (\hat{X}(t) - X^{(u)}(t))^T \{ \hat{q}(t) \hat{q}(t)^\top \right. \\ \left. + \int_{\mathbf{R}^n} [\hat{r}(t, z) \text{Diag}(\chi(z)) \hat{r}(t, z)^\top \lambda(dz)] \} (\hat{X}(t) - X^{(u)}(t)) dt \right] < \infty. \end{aligned} \quad (8)$$

and

$$\begin{aligned} \mathbb{E} \left[\int_0^T \hat{p}(t)^\top \{ \int [\eta(t, X(t^-), u(t), z) \text{Diag}(\chi(z)) \eta^\top(t, X(t^-), u(t), z) \lambda(dz)] \right. \\ \left. + (\sigma \sigma^\top)(t, X(t), u(t)) \} \hat{p}(t) dt \right] < \infty, \end{aligned} \quad (9)$$

ensuring that the integrals with respect to the local martingales (i.e. B and the “compensated small jumps” part) indeed have zero mean. Moreover, suppose that

$$H(t, \hat{X}(t), \hat{u}(t), \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot)) = \sup_{u \in U} H(t, \hat{X}(t), u, \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot)) \quad (10)$$

for all $t \in [0, T]$ and that

$$\hat{H}(x) := \max_{u \in U} H(t, x, u, \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot)) \quad (11)$$

exists and is a concave function of x for all $t \in [0, T]$. Then (\hat{u}, \hat{X}) is an optimal pair.

Remark 1.2. The concavity of \hat{H} defined by (11) is Arrow’s generalization of the frequently assumed a.e. concavity of $(x, v) \mapsto H(t, x, v, \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot))$. To the best of our knowledge, this was first proved rigorously in [15] in the deterministic setting.

To prove Theorem 1.1 we need the following auxiliary result, which is a special case of the Itô formula (see e.g. [4, Theorem 3.5.2, p. 265]):

Lemma 1.3 (Integration by parts formula for jump processes). *Suppose the processes $Y^{(1)}(t)$ and $Y^{(2)}(t)$ are given by*

$$\begin{cases} dY^{(j)}(t) = b^{(j)}(t, \omega) dt + \sigma^{(j)}(t, \omega) dB(t) + \int \eta^{(j)}(t, z, \omega) \bar{N}(dt, dz) \\ Y^{(j)}(0) = y^{(j)} \in \mathbf{R}^n ; \quad j = 1, 2 \end{cases} \quad (12)$$

where $b^{(j)} \in \mathbf{R}^n$, $\sigma^{(j)} \in \mathbf{R}^{n \times m}$ and $[\eta_{ik}^{(j)}] \in \mathbf{R}^{n \times \ell}$ are predictable processes such that the integrals in (12) exist. Then

$$\begin{aligned} & \mathbf{E}[Y^{(1)}(T) \cdot Y^{(2)}(T)] \\ &= y_1 \cdot y_2 + \mathbf{E} \left[\int_0^T Y^{(1)}(t^-) \cdot dY^{(2)}(t) + \int_0^T Y^{(2)}(t^-) \cdot dY^{(1)}(t) \right. \\ & \quad \left. + \int_0^T \text{tr}[\sigma^{(1)\top} \sigma^{(2)}](t) dt + \int_0^T \left(\int_{\mathbf{R}^n} \sum_{j=1}^{\ell} \left(\sum_{i=1}^n \eta_{ij}^{(1)}(t, z) \eta_{ij}^{(2)}(t, z) \right) d\lambda_j(z) \right) dt \right], \quad (13) \end{aligned}$$

provided the integrals and the mean exist. Here, as usual, $x \cdot y = x^\top y$ denotes the inner product of $x, y \in \mathbf{R}^n = \mathbf{R}^{n \times 1}$.

Proof of Theorem 1.1. Let (u, X) be an arbitrary admissible pair and consider

$$J(\hat{u}) - J(u) = \mathbb{E} \left[\int_0^T \{f(t, \hat{X}(t), \hat{u}(t)) - f(t, X(t), u(t))\} dt + g(\hat{X}(T)) - g(X(T)) \right] \quad (14)$$

By concavity of g and Lemma 1.3 we get the following, where the L^2 conditions (9) and (8) ensure that the stochastic integrals with respect to the local martingales have zero expectation:

$$\begin{aligned} & \mathbb{E}[g(\hat{X}(T)) - g(X(T))] \\ & \geq \mathbb{E}[(\hat{X}(T) - X(T))^\top \nabla g(\hat{X}(T))] \\ & = \mathbb{E}[(\hat{X}(T) - X(T))^\top \hat{p}(T)] \\ & = \mathbb{E} \left[\int_0^T (\hat{X}(t^-) - X(t^-))^\top d\hat{p}(t) + \int_0^T \hat{p}^\top(t^-) (d\hat{X}(t) - dX(t)) \right. \\ & \quad + \int_0^T \text{tr} \left[\{\sigma(t, \hat{X}(t), \hat{u}(t)) - \sigma(t, X(t), u(t))\}^\top \hat{q}(t) \right] dt \\ & \quad \left. + \int_0^T \int_{\mathbf{R}^n} \sum_{j=1}^{\ell} \sum_{i=1}^n \{\eta_{ij}(t, \hat{X}(t), \hat{u}(t), z) - \eta_{ij}(t, X(t), u(t), z)\} \hat{r}_{ij}(t, z) d\lambda_j(z) dt \right] \\ & = \mathbb{E} \left[\int_0^T \left\{ (\hat{X}(t) - X(t))^\top (-\nabla_x H(t, \hat{X}(t), \hat{u}(t), \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot))) \right. \right. \\ & \quad + \int_{\mathbf{R}^n} (\hat{X}(t) - X(t))^\top \hat{r}(t, z) (I - \text{Diag}(\chi)) d\lambda(z) \\ & \quad + \hat{p}^\top(t) \{b(t, \hat{X}(t), \hat{u}(t)) - b(t, X(t), u(t))\} \\ & \quad + \int_{\mathbf{R}^n} \hat{p}^\top(t) \{\eta(t, \hat{X}(t), \hat{u}(t), z) - \eta(t, X(t), u(t), z)\} (I - \text{Diag}(\chi)) d\lambda(z) \\ & \quad + \text{tr} \left[\{\sigma(t, \hat{X}(t), \hat{u}(t)) - \sigma(t, X(t), u(t))\}^\top \hat{q}(t) \right] \\ & \quad \left. \left. + \int_{\mathbf{R}^n} \sum_{j=1}^{\ell} \sum_{i=1}^n \{\eta_{ij}(t, \hat{X}(t), \hat{u}(t), z) - \eta_{ij}(t, X(t), u(t), z)\} \hat{r}_{ij}(t, z) d\lambda_j(z) \right\} dt \right]. \quad (15) \end{aligned}$$

By the definition (5) of H we have

$$\begin{aligned}
& \mathbb{E} \left[\int_0^T \{f(t, \hat{X}(t), \hat{u}(t)) - f(t, X(t), u(t))\} dt \right] \\
&= \mathbb{E} \left[\int_0^T \left\{ H(t, \hat{X}(t), \hat{u}(t), \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot)) - H(t, X(t), u(t), \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot)) \right. \right. \\
&\quad - \{b(t, \hat{X}(t), \hat{u}(t)) - b(t, X(t), u(t))\}^\top \hat{p}(t) \\
&\quad - \text{tr} \{ (\sigma(t, \hat{X}(t), \hat{u}(t)) - \sigma(t, X(t), u(t)))^\top \hat{q}(t) \} \\
&\quad - \int_{\mathbf{R}^n} \sum_{j=1}^{\ell} \sum_{i=1}^n \{ \eta_{ij}(t, \hat{X}(t), \hat{u}(t)) - \eta_{ij}(t, X(t), u(t)) \} \hat{r}_{ij}(t, z) d\lambda_j(z) \\
&\quad - \int_{\mathbf{R}^n} \hat{p}^\top(t) \{ \eta(t, \hat{X}(t), \hat{u}(t)) - \eta(t, X(t), u(t)) \} (I - \text{Diag}(\chi)) d\lambda(z) \\
&\quad \left. - \int_{\mathbf{R}^n} (\hat{X}(t) - X(t))^\top \hat{r}(t, z) (I - \text{Diag}(\chi)) d\lambda(z) \right\} dt \right]. \tag{16}
\end{aligned}$$

Adding (15) and (16) we get

$$\begin{aligned}
& J(\hat{u}) - J(u) \\
&\geq \mathbb{E}^x \left[\int_0^T \left\{ H(t, \hat{X}(t), \hat{u}(t), \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot)) - H(t, X(t), u(t), \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot)) \right. \right. \\
&\quad \left. \left. - (\hat{X}(t) - X(t))^\top \nabla_x H(t, \hat{X}(t), \hat{u}(t), \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot)) \right\} dt \right]. \tag{17}
\end{aligned}$$

As in the deterministic case the maximality of $\hat{u}(t)$ and the concavity of \hat{H} yield that the integrand in the r.h.s. of (17) is nonnegative for all t , a.s. See e.g. [16, p. 108]. For completeness we give the details: To simplify the notation put

$$h(t, x, u) = H(t, x, u, \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot)) \tag{18}$$

and

$$\hat{h}(t, x) = \max_{u \in U} h(t, x, u). \tag{19}$$

Then by (10) we have

$$h(t, \hat{X}(t), \hat{u}(t)) = \hat{h}(t, \hat{X}(t)) \tag{20}$$

and by (11)

$$h(t, x, u(t)) \leq \hat{h}(t, x) \quad \text{for all } t, x . \quad (21)$$

Therefore, subtracting (20) from (21) we get

$$h(t, x, u) - h(t, \hat{X}(t), \hat{u}(t)) \leq \hat{h}(t, x) - \hat{h}(t, \hat{X}(t)) \quad \text{for all } (t, x, u) . \quad (22)$$

Hence, to prove that the integrand in (17) is nonnegative it suffices to prove that on $[0, T]$ we almost surely have

$$\hat{h}(t, X(t)) - \hat{h}(t, \hat{X}(t)) - \nabla_x h(t, \hat{X}(t), \hat{u}(t)) \cdot (X(t) - \hat{X}(t)) \leq 0. \quad (23)$$

Fix $t \in [0, T]$. Since $x \mapsto \hat{h}(t, x)$ is concave, it follows by a standard separating hyperplane argument (see e.g. [13, Chapter 5, Section 23]) that there exists a *supergradient* $a \in \mathbf{R}^n$ for $\hat{h}(t, x)$ at $x = \hat{X}(t)$, i.e.

$$\hat{h}(t, x) - \hat{h}(t, \hat{X}(t)) - a \cdot (x - \hat{X}(t)) \leq 0 \quad \text{for all } x . \quad (24)$$

Define $\phi(x) = h(t, x, \hat{u}(t)) - h(t, \hat{X}(t), \hat{u}(t)) - a \cdot (x - \hat{X}(t))$. Then by (22) and (24) we have $\phi(x) \leq 0$ for all x . Moreover, we clearly have $\phi(\hat{X}(t)) = 0$, and therefore $\nabla \phi(\hat{X}(t)) = 0$, i.e. $\nabla_x h(t, \hat{X}(t), \hat{u}(t)) = a$. Substituting this into (24) we get (23). We therefore conclude that $J(\hat{u}) \geq J(u)$ for all $u \in \mathcal{A}$, which proves that \hat{u} is optimal. \square

2 Relation to dynamic programming

In the diffusion case it is well-known that there is a relation between the maximum principle and dynamic programming. More precisely, the adjoint processes $p(t)$, $q(t)$ can be expressed in terms of the derivatives of the value function $V(t, x)$. See e.g. [18]. In this section we prove that a similar relation holds in the jump diffusion case. We show that – under some conditions – the three adjoint processes $p(t)$, $q(t)$, $r(t, \cdot)$ for the jump diffusion case are given by

$$p_i(t) = \frac{\partial V}{\partial x_i}(t, X^*(t)) \quad (25a)$$

$$q_{ik}(t) = \sum_{j=1}^n \sigma_{jk}(t, X^*(t), u^*(t)) \frac{\partial^2 V}{\partial x_i \partial x_j}(t, X^*(t)) \quad (25b)$$

$$r_{ij}(t, z) = \frac{\partial V}{\partial x_i}(t, X^*(t) + \eta^{(j)}(t, X^*(t), u^*(t), z)) - \frac{\partial V}{\partial x_i}(t, X^*(t)) \quad (25c)$$

for all $i = 1, \dots, n$, $j = 1, \dots, \ell$, $k = 1, \dots, m$, where $X^*(t)$ is the solution of (1) corresponding to an optimal control $u^*(t)$ (see Theorem 2.1 below). In particular, note that the adjoint process $r(t, \cdot)$ represents the jumps of the x -gradient of $V(t, x)$.

We now proceed to prove these formulas. To put our problem (4) into a Markovian framework suitable for dynamic programming we let $X(t) = X^{s,x}(t)$ be the solution of (1) for $t \geq s$ with initial value $X(s) = x$ and we put

$$J_u(s, x) = \mathbb{E} \left[\int_s^T f(t, X^{s,x}(t), u(t)) dt + g(X^{s,x}(T)) \right]; \quad u \in \mathcal{A} \quad (26)$$

where f, g are as in (2). Then we define the *value function* $V(s, x)$ of the problem (4) by

$$V(s, x) = \sup_{u \in \mathcal{A}} J_u(s, x). \quad (27)$$

Assume that an optimal Markovian (feedback) control $u^*(t, x) = u^*$ exists for this problem and let $X^*(t)$ be the corresponding optimal state process, i.e. $X^*(t)$ is the solution of (1) when $u = u^*(t, X(t))$. Then – under some conditions – the following *Hamilton-Jacobi-Bellman* (HJB) equation of dynamic programming holds:

$$\sup_{u \in U} F(t, x, u) = F(t, x, u^*(t, x)) = 0 \quad (28)$$

where $F(t, x, u)$ is given by

$$\begin{aligned} F(t, x, u) := & f(t, x, u) + \frac{\partial V}{\partial t}(t, x) + \sum_{i=1}^n b_i(t, x, u) \frac{\partial V}{\partial x_i}(t, x) \\ & + \frac{1}{2} \sum_{i,j=1}^n (\sigma \sigma^\top)_{ij}(t, x, u) \frac{\partial^2 V}{\partial x_i \partial x_j}(t, x) + \sum_{k=1}^{\ell} \int_{\mathbf{R}^n} \left\{ V(t, x + \eta^{(k)}(t, x, u, z)) - V(t, x) \right. \\ & \left. - \chi_k(z) \sum_{i=1}^n \eta_{ik}(t, x, u, z) \cdot \frac{\partial V}{\partial x_i}(t, x) \right\} d\lambda_k(z), \end{aligned} \quad (29)$$

$\eta^{(k)}$ being column number k of the $n \times \ell$ matrix η . If we differentiate $F(t, x, u^*(t, x))$

with respect to x_h and evaluate the result at $x = X^*(t)$ we get

$$\begin{aligned}
0 &= \frac{\partial f}{\partial x_h}(t, X^*(t), u^*(t, X^*(t))) + \frac{\partial^2 V}{\partial t \partial x_h}(t, X^*(t)) \\
&+ \sum_{i=1}^n b_i(t, X^*(t), u^*(t, X^*(t))) \cdot \frac{\partial^2 V}{\partial x_i \partial x_h}(t, X^*(t)) \\
&+ \sum_{i=1}^n \frac{\partial b_i}{\partial x_h}(t, X^*(t), u^*(t, X^*(t))) \cdot \frac{\partial V}{\partial x_i}(t, X^*(t)) \\
&+ \frac{1}{2} \sum_{i,j=1}^n (\sigma \sigma^\top)_{ij}(t, X^*(t), u^*(t, X^*(t))) \cdot \frac{\partial^3 V}{\partial x_i \partial x_j \partial x_h}(t, X^*(t)) \\
&+ \frac{1}{2} \sum_{i,j=1}^n \frac{\partial}{\partial x_h} (\sigma \sigma^\top)_{ij}(t, X^*(t), u^*(t, X^*(t))) \cdot \frac{\partial^2 V}{\partial x_i \partial x_j}(t, X^*(t)) \\
&+ \sum_{k=1}^{\ell} \int_{\mathbf{R}^n} \left\{ \sum_{i=1}^n \frac{\partial V}{\partial x_i}(t, X^*(t) + \eta^{(k)}(t, X^*(t), u^*(t, X^*(t)), z)) \right. \\
&\quad \cdot \left(\delta_{ih} + \frac{\partial \eta_{ik}}{\partial x_h}(t, X^*(t), u^*(t, X^*(t)), z) \right) \\
&\quad - \chi_k(z) \sum_{i=1}^n \eta_{ik}(t, X^*(t), u^*(t, X^*(t)), z) \cdot \frac{\partial^2 V}{\partial x_i \partial x_h}(t, X^*(t)) \\
&\quad - \chi_k(z) \sum_{i=1}^n \frac{\partial \eta_{ik}}{\partial x_h}(t, X^*(t), u^*(t, X^*(t)), z) \cdot \frac{\partial V}{\partial x_i}(t, X^*(t)) \\
&\quad \left. - \frac{\partial V}{\partial x_h}(t, X^*(t)) \right\} d\lambda_k(z), \tag{30}
\end{aligned}$$

where $\delta_{ih} = 1$ if $i = h$, zero if not. Note that the terms containing the derivatives of $F(t, x, u)$ with respect to u vanish at $u = u^*$, because $F(t, x, u)$ is maximal at $u = u^*$. Define

$$Y_h(t) = \frac{\partial V}{\partial x_h}(t, X^*(t)); \quad h = 1, 2, \dots, n. \tag{31}$$

Then by the Itô formula (see [4, Theorem 3.5.2, p. 265])

$$\begin{aligned}
dY_h(t) &= \sum_{i=1}^n \frac{\partial^2 V}{\partial x_i \partial x_h}(t, X^*(t)) (b_i(t, X^*(t), u^*(t)) dt + \sigma_i(t, X^*(t), u^*(t)) dB(t)) \\
&+ \frac{\partial^2 V}{\partial t \partial x_h}(t, X^*(t)) dt + \frac{1}{2} \sum_{i,j=1}^n (\sigma \sigma^\top)_{ij}(t, X^*(t), u^*(t, X^*(t))) \cdot \frac{\partial^3 V}{\partial x_i \partial x_j \partial x_h}(t, X^*(t)) dt \\
&+ \sum_{k=1}^{\ell} \left[\int_{\mathbf{R}^n} \left\{ \frac{\partial V}{\partial x_h}(t, X^*(t^-) + \eta^{(k)}(t, X^*(t), u^*(t, X^*(t)), z)) - \frac{\partial V}{\partial x_h}(t, X^*(t^-)) \right. \right. \\
&\quad \left. \left. - \sum_{i=1}^n \frac{\partial^2 V}{\partial x_i \partial x_h}(t, X^*(t)) \eta_{ik}(t, X^*(t), u^*(t, X^*(t)), z) \right\} \chi_k(z) d\lambda_k(z) dt \right. \\
&\quad \left. + \int_{\mathbf{R}^n} \left\{ \frac{\partial V}{\partial x_h}(t, X^*(t^-) + \eta^{(k)}(t, X^*(t), u^*(t, X^*(t)), z)) - \frac{\partial V}{\partial x_h}(t, X^*(t^-)) \right\} \bar{N}_k(dt, dz) \right] \tag{32}
\end{aligned}$$

and substituting into (32) the value we get for $\frac{\partial^2 V}{\partial t \partial x_h}(t, X^*(t))$ from (30) we get for each $h = 1, \dots, n$:

$$\begin{aligned}
dY_h(t) &= - \left[\frac{\partial f}{\partial x_h} + \sum_{i=1}^n b_i \frac{\partial^2 V}{\partial x_i \partial x_h} + \sum_{i=1}^n \frac{\partial b_i}{\partial x_h} \cdot \frac{\partial V}{\partial x_i} \right. \\
&\quad + \frac{1}{2} \sum_{i,j=1}^n (\sigma \sigma^\top)_{ij} \frac{\partial^3 V}{\partial x_i \partial x_j \partial x_h} + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial}{\partial x_h} (\sigma \sigma^\top)_{ij} \frac{\partial^2 V}{\partial x_i \partial x_j} \\
&\quad + \sum_{k=1}^{\ell} \int_{\mathbf{R}^n} \left\{ \sum_{i=1}^n \left(\left(\frac{\partial V}{\partial x_i}(t, x + \eta^{(k)}) - \chi_k(z) \frac{\partial V}{\partial x_i} \right) \frac{\partial \eta_{ik}}{\partial x_h} - \eta_{ik} \frac{\partial^2 V}{\partial x_i \partial x_h} \chi_k(z) \right) \right. \\
&\quad \left. + \left(\frac{\partial V}{\partial x_h}(t, x + \eta^{(k)}) - \frac{\partial V}{\partial x_h} \right) \right\} d\lambda_k(z) \Big] dt \\
&+ \sum_{i=1}^n \frac{\partial^2 V}{\partial x_i \partial x_h} (b_i dt + \sigma_i dB(t)) + \frac{1}{2} \sum_{i,j=1}^n (\sigma \sigma^\top)_{ij} \cdot \frac{\partial^3 V}{\partial x_i \partial x_j \partial x_h} dt \\
&+ \sum_{k=1}^{\ell} \int_{\mathbf{R}^n} \left\{ \frac{\partial V}{\partial x_h}(t, x + \eta^{(k)}) - \frac{\partial V}{\partial x_h}(t, x) - \sum_{i=1}^n \frac{\partial^2 V}{\partial x_i \partial x_h} \cdot \eta_{ik} \right\} \chi_k(z) d\lambda_k(z) dt \\
&+ \sum_{k=1}^{\ell} \int_{\mathbf{R}^n} \left\{ \frac{\partial V}{\partial x_h}(t, x + \eta^{(k)}) - \frac{\partial V}{\partial x_h}(t, x) \right\} \bar{N}_k(dt, dz)
\end{aligned}$$

which simplifies to

$$\begin{aligned}
dY_h(t) = & - \left[\frac{\partial f}{\partial x_h} + \sum_{i=1}^n \frac{\partial b_i}{\partial x_h} \cdot \frac{\partial V}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial}{\partial x_h} (\sigma \sigma^\top)_{ij} \frac{\partial^2 V}{\partial x_i \partial x_j} \right. \\
& + \sum_{k=1}^{\ell} \int_{\mathbf{R}^n} \left\{ \left(\frac{\partial V}{\partial x_h}(t, x + \eta^{(k)}) - \frac{\partial V}{\partial x_h} \right) (1 - \chi_k(z)) \right. \\
& \quad \left. + \sum_{i=1}^n \left(\frac{\partial V}{\partial x_i}(t, x + \eta^{(k)}) - \chi_k(z) \frac{\partial V}{\partial x_i} \right) \frac{\partial \eta_{ik}}{\partial x_h} \right\} d\lambda_k(z) \Big] dt \\
& + \sum_{k=1}^{\ell} \int_{\mathbf{R}^n} \left\{ \frac{\partial V}{\partial x_h}(t, x + \eta^{(k)}) - \frac{\partial V}{\partial x_h} \right\} \bar{N}_k(dt, dz) + \sum_{i=1}^n \sum_{j=1}^m \frac{\partial^2 V}{\partial x_i \partial x_h} \sigma_{ij} dB_j \quad (33)
\end{aligned}$$

where all functions are evaluated at $x = X^*(t^-)$, $u = u^*(t, X^*(t^-))$. From (5) we note that

$$\begin{aligned}
\frac{\partial H}{\partial x_h}(t, x, u, p, q, r) = & \frac{\partial f}{\partial x_h}(t, x, u) + \frac{\partial b^\top}{\partial x_h}(t, x, u)p + \text{tr} \left(\frac{\partial \sigma^\top}{\partial x_h}(t, x, u)q \right) \\
& + \sum_{k=1}^{\ell} \int_{\mathbf{R}^n} \left\{ \sum_{i=1}^n \frac{\partial \eta_{ik}}{\partial x_h}(t, x, u, z) r_{ik}(t, z) \right. \\
& \quad \left. + \left(\frac{\partial \eta_k^\top}{\partial x_h}(t, x, u, z)p + r_{hk}(t, z) \right) (1 - \chi_k(z)) \right\} d\lambda_k(z). \quad (34)
\end{aligned}$$

Note that

$$\begin{aligned}
\frac{1}{2} \sum_{i,j=1}^n \frac{\partial}{\partial x_h} (\sigma \sigma^\top)_{ij} \frac{\partial^2 V}{\partial x_i \partial x_j} &= \frac{1}{2} \sum_{i,j=1}^n \frac{\partial}{\partial x_h} \left(\sum_{k=1}^m \sigma_{ik} \sigma_{jk} \right) \frac{\partial^2 V}{\partial x_i \partial x_j} \\
&= \frac{1}{2} \sum_{i,j=1}^n \sum_{k=1}^m \left(\frac{\partial \sigma_{ik}}{\partial x_h} \sigma_{jk} + \sigma_{ik} \frac{\partial \sigma_{jk}}{\partial x_h} \right) \frac{\partial^2 V}{\partial x_i \partial x_j} = \sum_{j=1}^n \sum_{k=1}^m \left(\sum_{i=1}^n \sigma_{ik} \frac{\partial^2 V}{\partial x_i \partial x_j} \right) \frac{\partial \sigma_{jk}}{\partial x_h}. \quad (35)
\end{aligned}$$

On the other hand

$$\text{tr} \left(\frac{\partial \sigma^\top}{\partial x_h} q \right) = \sum_{k=1}^m \left[\frac{\partial \sigma^\top}{\partial x_h} q \right]_{kk} = \sum_{k=1}^m \sum_{j=1}^n q_{jk} \frac{\partial \sigma_{jk}}{\partial x_h}. \quad (36)$$

Therefore, if we write $x(t) = X^*(t)$, $u(t) = u^*(t, X^*(t))$ and define $p_i(t)$, $q_{jk}(t)$ and

$r_{ik}(t, z)$ by (25), we get the following:

$$\begin{aligned}
dY_h(t) &= -\frac{\partial H}{\partial x_h}(t, x(t), u(t), p(t), q(t), r(t, \cdot)) dt + q_h(t) dB(t) \\
&\quad + \sum_{k=1}^{\ell} \int_{\mathbf{R}^n} r_{hk}(t^-, z) \bar{N}_k(dt, dz) + \sum_{k=1}^{\ell} \int_{\mathbf{R}^n} \sum_{i=1}^n \frac{\partial \eta_{ik}}{\partial x_h}(t, x(t), u(t), z) \\
&\quad \cdot (r_{ik} + p_i(t)(1 - \chi_k(z)) - (\frac{\partial V}{\partial x_i}(t, x + \eta^{(k)}) - \chi_k(z) \frac{\partial V}{\partial x_i})) d\lambda_k(z) dt \\
&= -\frac{\partial H}{\partial x_h}(t, x(t), u(t), p(t), q(t), r(t, \cdot)) dt + q_h(t) dB(t) \\
&\quad + \sum_{k=1}^{\ell} \int_{\mathbf{R}^n} r_{hk}(t^-, z) \bar{N}_k(dt, dz) \tag{37}
\end{aligned}$$

and we see that $p(t), q(t)$ and $r(t, \cdot)$ solve the adjoint equation (6). We thus have proved:

Theorem 2.1. *Assume that the value function $V(s, x)$ defined in (27) belongs to $C^{1,3}(\mathbf{R} \times \mathbf{R}^n)$ and that there exists an optimal Markovian control $u^*(t, x)$ for problem (27) with corresponding optimal state process $X^*(t)$ solving (1). Then the processes $p(t), q(t), r(t, \cdot)$ given by (25) solve the adjoint equation (6).*

3 Applications to finance

In this section, we shall give two examples arising from financial optimization. Suppose we have a mathematical market consisting of two investment possibilities:

- (i) a *risk free* security (e.g. a bond), where the price $S_0(t)$ at time t is given by

$$dS_0(t) = \rho_t S_0(t) dt ; \quad S_0(0) > 0 \tag{38a}$$

where ρ_t is a locally bounded deterministic function,

- (ii) a *risky* security (e.g. a stock), where the price $S_1(t)$ at time t is given by

$$dS_1(t) = S_1(t^-) \left[\mu_t dt + \sigma_t dB(t) + \int_{\mathbf{R}} \eta(t, z) \tilde{N}(dt, dz) \right], \quad S_1(0) > 0, \tag{38b}$$

where $\mu_t, \sigma_t \neq 0$ and $\eta(t, z)$ are locally bounded deterministic functions, $\mu_t > \rho_t$ and, as above, \tilde{N} is a compensated random measure. To ensure that $S_1(t) > 0$ for all t we assume that

$$\eta(t, z) = 0 \quad \text{for } z \in (-\infty, -1] . \quad (39)$$

We also assume that $t \mapsto \int_{\mathbf{R}} \eta^2(t, z) d\lambda(z)$ is a locally bounded function, where λ is the Lévy measure of N as in the previous sections.

A *portfolio* is a predictable process $\theta(t) = (\theta_0(t), \theta_1(t)) \in \mathbf{R}^2$ giving the number of units held at time t of the risk-free and the risky security, respectively. The corresponding *wealth process* $X(t) = X(t)$ is then given by

$$X^\theta(t) = \theta_0(t)S_0(t) + \theta_1(t)S_1(t) ; \quad t \geq 0 . \quad (40)$$

The portfolio is called *self-financing* if

$$dX^\theta(t) = \theta_0(t) dS_0(t) + \theta_1(t^-) dS_1(t) \quad (41)$$

or

$$X^\theta(t) = X^\theta(0) + \int_0^t \theta_0(s) dS_0(s) + \int_0^t \theta_1(s^-) dS_1(s) ; \quad t \geq 0 .$$

Let $v(t) := \theta_1(t)S_1(t)$ denote the *amount* invested in the risky security. Combining (40) and (41), we get the wealth dynamics

$$dX(t) = \{\rho_t X(t) + (\mu_t - \rho_t)v(t)\} dt + \sigma_t v(t) dB(t) + v(t^-) \int_{\mathbf{R}} \eta(t, z) \tilde{N}(dt, dz) \quad (42a)$$

where $X(0) \in \mathbf{R}$ is given. In Example 3.3 below we shall also allow the agent to withdraw consumption from his wealth, thus giving

$$dX(t) = \{\rho_t X(t) + (\mu_t - \rho_t)v(t) - c(t)\} dt + \sigma_t v(t) dB(t) + v(t^-) \int_{\mathbf{R}} \eta(t, z) \tilde{N}(dt, dz) \quad (42b)$$

as wealth process. The control v is called *tame* if the corresponding wealth process (42a) is square integrable with respect to $dt \times dP$ over $[0, T] \times \Omega$. Such a requirement is necessary to exclude doubling strategies that would gain arbitrary profit at time T , but with the economically unrealistic consequence of unbounded intermediate losses.

Example 3.1 (Mean-variance portfolio selection). In this application, we shall consider the process (42a) (i.e. without consumption.) Our objective is to find an admissible portfolio $v(t)$ which minimizes the variance

$$\text{Var}[X(T)] = \mathbb{E}[(X(T) - \mathbb{E}[X(T)])^2]$$

under the condition that

$$\mathbb{E}[X(T)] = A ,$$

where A is given real number. Using the Lagrange multiplier method we see that the problem can be reduced to the problem of minimizing, for a given $a \in \mathbf{R}$,

$$\mathbb{E}[(X(T) - a)^2]$$

without any constraints. To see this, note that

$$\mathbb{E}[(X(T) - A)^2 - \lambda(X(T) - A)] = \mathbb{E}[(X(T) - (A + \frac{\lambda}{2}))^2] + \frac{\lambda^2}{4},$$

if $\lambda \in \mathbf{R}$ is a constant.

In the following we study the following equivalent problem

$$\sup_{v \in \mathcal{A}} \mathbb{E} \left[-\frac{1}{2} (X^{(v)}(T) - a)^2 \right] , \quad (43)$$

where $X(t) = X^{(v)}(t)$ is given by (42a) and the set \mathcal{A} of admissible strategies consists of the predictable tame portfolios $v(t)$ such that (42a) has a strong solution in the interval $[0, T]$.

In the no jump case ($\eta = 0$) the solution of this problem is well-known. We refer to [10] for more information. We now illustrate our maximum principle (Theorem 1.1) to solve the problem in the jump diffusion case. In this case the Hamiltonian (5) gets the form

$$H(t, x, v, p, q, r) = \{\rho_t x + (\mu_t - \rho_t)v\}p + \sigma_t v q + v \int_{\mathbf{R}} \eta(t, z) r(t, z) d\lambda(z). \quad (44)$$

Therefore the adjoint equation (6) is

$$\begin{cases} dp(t) = -\rho_t p(t) dt + q(t) dB(t) + \int_{\mathbf{R}} r(t, z) \tilde{N}(dt, dz) \\ p(T) = -X(T) + a \end{cases} \quad (45)$$

To solve this we try a process $p(t)$ of the form

$$p(t) = \phi_t X(t) + \psi_t, \quad (46)$$

where ϕ_t, ψ_t are deterministic differentiable functions. Using (42a) this leads to

$$\begin{aligned} dp(t) &= \phi_t [\{\rho_t X(t) + (\mu_t - \rho_t)v(t)\} dt + \sigma_t v(t) dB(t) \\ &\quad + v(t^-) \int_{\mathbf{R}} \eta(t, z) \tilde{N}(dt, dz)] + X(t) \phi'_t dt + \psi'_t dt \\ &= [\phi_t \rho_t X(t) + \phi_t (\mu_t - \rho_t)v(t) + X(t) \phi'_t + \psi'_t] dt \\ &\quad + \phi_t \sigma_t v(t) dB(t) + \phi_t v(t^-) \int_{\mathbf{R}} \eta(t, z) \tilde{N}(dt, dz). \end{aligned} \quad (47)$$

Comparing (45) and (47) we get

$$\phi_t \rho_t X(t) + \phi_t (\mu_t - \rho_t)v(t) + X(t) \phi'_t + \psi'_t = -\rho_t (\phi_t X(t) + \psi_t) \quad (48a)$$

$$q(t) = \phi_t \sigma_t v(t) \quad (48b)$$

$$r(t, z) = \phi_t v(t) \eta(t, z). \quad (48c)$$

Let $\hat{v} \in \mathcal{A}$ be a candidate for an optimal control and let $\hat{X}(t)$ be the corresponding wealth process with corresponding solution $(\hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot))$ of the adjoint equation. Then

$$\begin{aligned} H(t, \hat{X}(t), v, \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot)) \\ = \rho_t \hat{X}(t) \hat{p}(t) + v \left[(\mu_t - \rho_t) \hat{p}(t) + \sigma_t \hat{q}(t) + \int_{\mathbf{R}} \eta(t, z) \hat{r}(t, z) d\lambda(z) \right]. \end{aligned}$$

Since this is a linear expression in v , we guess that the coefficient of v vanishes, i.e. that

$$(\mu_t - \rho_t) \hat{p}(t) + \sigma_t \hat{q}(t) + \int_{\mathbf{R}} \eta(t, z) \hat{r}(t, z) d\lambda(z) = 0. \quad (49)$$

Substituting (from (48b) and (48c))

$$\hat{q}(t) = \phi_t \sigma_t \hat{v}(t); \quad \hat{r}(t, z) = \phi_t \hat{v}(t) \eta(t, z) \quad (50)$$

into (49) and writing

$$\Lambda_t := \sigma_t^2 + \int_{\mathbf{R}} \eta^2(t, z) d\lambda(z), \quad (51)$$

we get

$$\hat{v}(t) = \frac{(\rho_t - \mu_t)\hat{p}(t)}{\phi_t\Lambda_t} = \frac{(\rho_t - \mu_t)(\phi_t\hat{X}(t) + \psi_t)}{\phi_t\Lambda_t}. \quad (52a)$$

On the other hand, (48a) gives

$$\hat{v}(t) = \frac{(\phi_t\rho_t + \phi'_t)\hat{X}(t) + \rho_t(\phi_t\hat{X}(t) + \psi_t) + \psi'_t}{\phi_t(\rho_t - \mu_t)}. \quad (52b)$$

Combining (52a) and (52b) we get

$$(\rho_t - \mu_t)^2(\phi_t\hat{X}(t) + \psi_t) = [(\phi_t\rho_t + \phi'_t)\hat{X}(t) + \rho_t(\phi_t\hat{X}(t) + \psi_t) + \psi'_t]\Lambda_t. \quad (53)$$

Comparing the terms containing $\hat{X}(t)$ we obtain the two equations

$$(\rho_t - \mu_t)^2\phi_t - [2\rho_t\phi_t + \phi'_t]\Lambda_t = 0 \quad (54a)$$

$$(\rho_t - \mu_t)^2\psi_t - [\rho_t\psi_t + \psi'_t]\Lambda_t = 0 \quad (54b)$$

which are equivalent to

$$\phi'_t = \left(\frac{(\rho_t - \mu_t)^2}{\Lambda_t} - 2\rho_t \right) \phi_t, \quad \phi(T) = -1 \quad (55a)$$

and

$$\psi'_t = \left(\frac{(\rho_t - \mu_t)^2}{\Lambda_t} - \rho_t \right) \psi_t, \quad \psi(T) = a. \quad (55b)$$

The solutions of these equations are

$$\phi_t = -\exp\left(\int_t^T \left\{ \frac{(\rho_s - \mu_s)^2}{\Lambda_s} - 2\rho_s \right\} ds\right); \quad 0 \leq t \leq T \quad (56a)$$

$$\psi_t = a \exp\left(\int_t^T \left\{ \frac{(\rho_s - \mu_s)^2}{\Lambda_s} - \rho_s \right\} ds\right); \quad 0 \leq t \leq T. \quad (56b)$$

With this choice of ϕ_t and ψ_t the processes

$$\hat{p}(t) = \phi_t\hat{X}(t) + \psi_t, \quad \hat{q}(t) = \phi_t\sigma_t\hat{v}(t), \quad \hat{r}(t, z) = \phi_t\hat{v}(t)\eta(t, z)$$

solve the adjoint equation (45) with $\hat{v}(t)$ given by (52a). Moreover, with this choice of $\hat{v}(t)$ conditions (10) and (11) of Theorem 1.1 is satisfied. Therefore we have the following:

Theorem 3.2. *The solution $v^* \in \mathcal{A}$ of the mean-variance portfolio selection problem (43) when X obeys (42a) is given (in feedback form) by*

$$v^*(t, X^*) = \frac{(\rho_t - \mu_t)(\phi_t X^* + \psi_t)}{\phi_t \Lambda_t} \quad (57)$$

with ϕ_t, ψ_t given by (56) and Λ_t given by (51).

Proof. It only remains to check the L^2 conditions (8) and (9) and tameness of v^* . The control v^* leads to a linear stochastic differential equation with bounded coefficients for X^* , and is therefore tame. Hence v^* itself is square integrable, and (48b) and (48c) now imply (8) and (9). \square

Our solution (Theorem 3.2) is a generalization to jump diffusions of the (known) solution in the diffusion case.

Example 3.3 (Consumption-portfolio optimization with almost sure terminal condition). Consider now the wealth process (42b), i.e. with consumption. Our objective is to solve the following consumption-portfolio optimization problem

$$\sup_{(c,v) \in \mathcal{A}} \mathbb{E} \left[\int_0^T e^{-\int_0^t \delta_s ds} \frac{c(t)^\gamma}{\gamma} dt \right] \quad (58)$$

subject to an almost sure terminal wealth constraint:

$$X(T) \geq 0 \text{ a.s.} \quad (59)$$

The set \mathcal{A} is defined to be the class of predictable consumption-portfolio pairs $u = (c, v)$ with v tame, c nonnegative, such that (42b) has a strong solution in $[0, T]$ and (59) holds. Here, δ_s is a given bounded deterministic function and γ is a given nonzero constant, $\gamma < 1$.

The dynamic programming approach is not directly applicable to this type of constrained stochastic control problems with the a.s. terminal condition (59). However, it is easy to see that our sufficient stochastic maximum principle (Theorem 1.1) still applies, if we replace the terminal condition (7)

$$p(T) = \nabla g(X(T))$$

by the *transversality condition*

$$\mathbb{E}[(\hat{X}(T) - X(T))^T \nabla g(X(T))] \geq \mathbb{E}[(\hat{X}(T) - X(T))\hat{p}(T)] \quad (60)$$

(see the step leading to (15) in the proof of Theorem 1.1).

Note the easy generalization to nonzero minimum terminal wealth, say, \bar{x} by considering $X - \bar{x}$ instead of X .

Here, the Hamiltonian (5) gets the form

$$H(t, x, c, v, p, q, r) = e^{-\int_0^t \delta_s ds} \frac{c^\gamma}{\gamma} - pc + p\rho x + v \cdot \left\{ p(\mu - \rho) + q\sigma + \int_{\mathbf{R}} r(t^-, z) \eta(t, z) d\lambda(z) \right\} \quad (61)$$

and the (modified) adjoint equation (6) becomes

$$dp(t) = -\rho(t)p(t) dt + q(t) dB(t) + \int_{\mathbf{R}} r(t^-, z) d\tilde{N}(dt, dz) \quad (62)$$

$$\mathbb{E}[(\hat{X}(T) - X(T))\hat{p}(T)] \leq 0 \quad (\text{since } g = 0). \quad (63)$$

Now, let $(\hat{c}, \hat{v}) \in \mathcal{A}$ with corresponding solution \hat{X} , $(\hat{p}, \hat{q}, \hat{r})$ of (42b) and (62) respectively.

The value of c which maximizes $H(t, \hat{X}(t), \hat{v}, \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot))$ is

$$c = \hat{c}(t) = \left(e^{\int_0^t \delta_s ds} \hat{p}(t) \right)^{\frac{1}{\gamma-1}}. \quad (64)$$

Since the expression involving v in the Hamiltonian is linear, the maximum principle (10) suggests that the v -coefficient should vanish, i.e.

$$\hat{p}(t)(\mu_t - \rho_t) + \sigma_t \hat{q}(t) + \int_{\mathbf{R}} \eta(t, z) \hat{r}(t, z) d\lambda(z) = 0. \quad (65)$$

We now guess that it is optimal to consume at a rate proportional to the current wealth $\hat{X}(t)$. By (64), this suggests that:

$$\hat{p}(t) = f(t) \hat{X}(t)^{\gamma-1} \quad (66)$$

for some deterministic differentiable function f (to be determined). Differentiating (66), we get:

$$\begin{aligned} d\hat{p}(t) &= f'(t) \hat{X}(t)^{\gamma-1} dt + (\gamma - 1) f(t) \hat{X}(t)^{\gamma-2} \left[d\hat{X}(t) - v(t) \int_{\mathbf{R}} \eta(t, z) \tilde{N}(dt, dz) \right] \\ &+ \frac{1}{2} (\gamma - 1)(\gamma - 2) f(t) \hat{X}(t)^{\gamma-3} \sigma_t^2 \hat{v}(t)^2 dt \\ &+ \int_{\mathbf{R}} f(t) \{ (\hat{X}(t) + \eta(t, z) \hat{v}(t))^{\gamma-1} - \hat{X}(t)^{\gamma-1} - (\gamma - 1) \hat{X}(t)^{\gamma-2} \hat{v}(t) \eta(t, z) \} d\lambda(z) dt \\ &+ \int_{\mathbf{R}} f(t) \{ (\hat{X}(t) + \eta(t, z) \hat{v}(t))^{\gamma-1} - \hat{X}(t)^{\gamma-1} \} \tilde{N}(dt, dz). \end{aligned}$$

Using now (42b) and comparing with (62), we get, by equating the $d\tilde{N}$, dB and dt - coefficients respectively:

$$\hat{r}(t, z) = f(t)\hat{X}(t)^{\gamma-1}\{(1 + \hat{v}(t)\eta(t, z)\hat{X}^{-1}(t))^{\gamma-1} - 1\} \quad (67)$$

$$\hat{q}(t) = (\gamma - 1)f(t)\sigma_t\hat{v}(t)\hat{X}(t)^{\gamma-2} \quad (68)$$

and

$$f'(t) + \alpha_t f(t) + (1 - \gamma)e^{\int_0^t \frac{\delta_s}{\gamma-1} ds} f(t)^{\frac{\gamma}{\gamma-1}} = 0 \quad (69)$$

where

$$\begin{aligned} \alpha_t = & \gamma\rho_t + (\gamma - 1)(\mu_t - \rho_t)\hat{v}(t)\hat{X}^{-1}(t) + \frac{1}{2}(\gamma - 1)(\gamma - 2)\sigma_t^2\hat{v}^2(t)\hat{X}^{-2}(t) \\ & + \int_{\mathbf{R}} \{(1 + \eta(t, z)\hat{v}(t)\hat{X}^{-1}(t))^{\gamma-1} - 1 - (\gamma - 1)\hat{v}(t)\hat{X}^{-1}(t)\eta(t, z)\} d\lambda(z). \end{aligned} \quad (70)$$

Substituting (66), (67) and (68) into (65), we get

$$F(\hat{v}(t)\hat{X}^{-1}(t)) = 0,$$

where

$$F(\pi) = \mu_t - \rho_t + \sigma_t^2(\gamma - 1)\pi + \int_{\mathbf{R}} \eta(t, z)\{(1 + \eta(t, z)\pi)^{\gamma-1} - 1\} d\lambda(z),$$

which is easily seen to have a zero $\hat{\pi}(t) > 0$, i.e.:

$$F(\hat{\pi}(t)) = 0. \quad (71)$$

With this choice of

$$\hat{v}(t)\hat{X}^{-1}(t) = \hat{\pi}(t) \quad (72)$$

and α_t given by (70), we can proceed to solve equation (69) by performing the change of variable $h(t) = (e^{\int_0^t \delta_s ds} f(t))^{\frac{1}{1-\gamma}}$. We obtain that

$$f(t) = e^{-\int_0^t \delta_s ds} \left(f(T)^{\frac{1}{1-\gamma}} e^{\int_0^T \frac{\delta_s}{1-\gamma} ds} e^{\int_t^T \frac{\alpha_r - \delta_r}{1-\gamma} dr} + \int_t^T e^{-\int_s^t \frac{\alpha_r - \delta_r}{1-\gamma} dr} ds \right)^{1-\gamma} \quad (73)$$

solves equation (69). Using (72), (64) and (66) we get that

$$\hat{c}(t) = e^{\int_0^t \frac{\delta_s}{\gamma-1} ds} f(t)^{\frac{1}{\gamma-1}} \hat{X}(t). \quad (74)$$

The corresponding equation (42b) for $\hat{X}(t)$ becomes

$$\begin{aligned} d\hat{X}(t) = & \hat{X}(t^-) \left\{ \left(\rho_t + (\mu_t - \rho_t) \hat{\pi}(t) - e^{\int_0^t \frac{\delta_s}{\gamma-1} ds} f(t)^{\frac{1}{\gamma-1}} \right) dt \right. \\ & \left. + \sigma_t \hat{\pi}(t) dB(t) + \hat{\pi}(t^-) \int_{\mathbf{R}} \eta(t, z) \tilde{N}(dt, dz) \right\}. \end{aligned}$$

The solution of this equation is

$$\begin{aligned} \hat{X}(t) = & \hat{X}(0) \exp \left\{ \int_0^t \left(\rho_s + (\mu_s - \rho_s) \hat{\pi}(s) - e^{\int_0^s \frac{\delta_r}{\gamma-1} dr} f(s)^{\frac{1}{\gamma-1}} - \frac{1}{2} \hat{\pi}(s) \sigma_s^2 \right) ds \right. \\ & \left. + \int_0^t \sigma_s \hat{\pi}(s) dB(s) + \int_0^t \left(\int_{\mathbf{R}} \ln(1 + \eta(t, s)) N(ds, dz) - \int_{\mathbf{R}} \eta(s, z) d\lambda(z) \right) ds \right\}. \end{aligned} \quad (75)$$

It is reasonable to guess, by economic intuition, that the optimal wealth process will satisfy the terminal condition with equality, i.e. $\hat{X}(T) = 0$ a.s., as excess wealth is worthless. To achieve this, we must have, by (75),

$$f(T) = 0,$$

which gives, by (73)

$$f(t) = e^{-\int_0^t \delta_s ds} \left(\int_t^T e^{-\int_s^t \frac{\alpha_r - \delta_r}{1-\gamma} dr} ds \right)^{1-\gamma}. \quad (76)$$

Then $f(s) \sim (T-s)^{1-\gamma}$ as $s \rightarrow T^-$ and hence

$$\int^T f(s)^{\frac{1}{\gamma-1}} ds \sim \int^T (T-s)^{-1} ds = \infty$$

which by (75) gives $\hat{X}(T) = 0$, as required. With $\hat{\pi}(t), \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot)$ and $f(t)$ as in (71), (66), (68), (67) and (76) respectively, we see that all the conditions of the maximum principle are satisfied, including the transversality conditions (63). Therefore, we have the theorem:

Theorem 3.4. *An optimal control $u^* = (c^*, v^*)$ for problem (58) subject to the dynamics (42b) is given in feedback form by*

$$\begin{aligned} c^*(t, x) &= e^{\int_0^t \frac{\delta_s}{\gamma-1} ds} f(t)^{\frac{1}{\gamma-1}} x \\ v^*(t, x) &= \hat{\pi}(t) x \end{aligned} \quad (77)$$

with $f(t)$ given by (76) and $\hat{\pi}(t)$ given by (71).

Proof. Again, we only need to check the L^2 conditions (8) and (9) and tameness of u^* , which hold by arguments similar to the proof of Theorem 3.2. \square

Remark 3.5. Note that we in fact get $X^*(t) \geq 0$ a.s. for all $t \in [0, T]$, even though we only require that $X^{(u)}(T) \geq 0$ a.s. Therefore Problem (58) has the same solution if the terminal condition (59) is replaced by the stronger condition

$$X^{(u)}(t) \geq 0 \quad \text{a.s. for all } t \in [0, T].$$

In this case the problem could also have been solved by dynamic programming (the Hamilton-Jacobi-Bellman equation). However, it is difficult to see how dynamic programming could be used without knowing that the two constrained problems are equivalent.

References

- [1] F. ANTONELLI, *Backward-forward stochastic differential equations*, Ann. Appl. Probab., 3 (1993), pp. 777–793.
- [2] G. BARLES, R. BUCKDAHN, AND E. PARDOUX, *Backward stochastic differential equations and integral-partial differential equations*, Stochastics Stochastics Rep., 60 (1997), pp. 57–83.
- [3] A. BENSOUSSAN, *Maximum principle and dynamic programming approaches of the optimal control of partially observed diffusions*, Stochastics, 9 (1983), pp. 169–222.
- [4] A. BENSOUSSAN AND J.-L. LIONS, *Impulse control and quasivariational inequalities*, Gauthier-Villars, Montrouge, 1984.
- [5] J.-M. BISMUT, *Conjugate convex functions in optimal stochastic control*, J. Math. Anal. Appl., 44 (1973), pp. 384–404.
- [6] A. CADENILLAS AND U. G. HAUSSMANN, *The stochastic maximum principle for a singular control problem*, Stochastics Stochastics Rep., 49 (1994), pp. 211–237.
- [7] U. G. HAUSSMANN, *A stochastic maximum principle for optimal control of diffusions*, Longman Scientific & Technical, Harlow, 1986.

- [8] Y. M. KABANOV, *On the Pontryagin maximum principle for SDEs with a Poisson-type driving noise*, in *Statistics and control of stochastic processes* (Moscow, 1995/1996), Y. M. Kabanov, B. L. Rozovskii, and A. N. Shiryaev, eds., World Sci. Publishing, River Edge, NJ, 1997, pp. 173–190.
- [9] M. KOHLMANN, *Optimality conditions in optimal control of jump processes – extended abstract*. *Proceedings in Operations Research*, 7 (Sixth Annual Meeting, Deutsch. Gesellsch. Operations Res., Christian-Albrechts-Univ., Kiel, 1977), pp. 48–57, Physica, Wrzburg, 1978.
- [10] M. KOHLMANN AND X. Y. ZHOU, *Relationship between backward stochastic differential equations and stochastic controls: a linear-quadratic approach*, *SIAM J. Control Optim.*, 38 (2000), pp. 1392–1407 (electronic).
- [11] H. J. KUSHNER, *Necessary conditions for continuous parameter stochastic optimization problems*, *SIAM J. Control*, 10 (1972), pp. 550–565.
- [12] S. G. PENG, *A general stochastic maximum principle for optimal control problems*, *SIAM J. Control Optim.*, 28 (1990), pp. 966–979.
- [13] R. T. ROCKAFELLAR, *Convex analysis*, Princeton University Press, Princeton, N.J., 1970.
- [14] S. RONG, *On solutions of backward stochastic differential equations with jumps and applications*, *Stochastic Process. Appl.*, 66 (1997), pp. 209–236.
- [15] A. SEIERSTAD AND K. SYDSÆTER, *Sufficient conditions in optimal control theory*, *Internat. Econom. Rev.*, 18 (1977), pp. 367–391.
- [16] ———, *Optimal control theory with economic applications*, North-Holland Publishing Co., Amsterdam, 1987.
- [17] S. J. TANG AND X. J. LI, *Necessary conditions for optimal control of stochastic systems with random jumps*, *SIAM J. Control Optim.*, 32 (1994), pp. 1447–1475.
- [18] J. YONG AND X. Y. ZHOU, *Stochastic controls*, Springer-Verlag, New York, 1999. Hamiltonian systems and HJB equations.