Maximum principles for optimal control of forward-backward stochastic differential equations with jumps

Bernt Øksendal[*] [†]   Agnès Sulem[‡]

Revised July 2009

Abstract

We present various versions of the maximum principle for optimal control of forward-backward SDEs with jumps. Our study is motivated by risk minimization via g-expectations. We first prove a general sufficient maximum principle for optimal control with partial information of a stochastic system consisting of a forward and a backward SDE driven by Lévy processes. We then present a Malliavin calculus approach which allows us to handle non-Markovian systems. Finally we give examples of applications.

1   Introduction and motivation example

The purpose of this paper is to discuss solution methods of the maximum principle type for the optimal control of systems of forward-backward stochastic differential equations driven by Lévy processes. One of the motivations of this study is the problem of finding risk minimizing portfolios in finance, where the risk is represented in terms of g-expectations. We now explain this in more detail.

Let $\eta(t) = \eta(t, \omega)$; $t \geq 0$, $\omega \in \Omega$, be a Lévy process on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$. Let $\mathbb{F}$ be a family of $\mathcal{F}_T$-measurable random variables $F : \Omega \to \mathbb{R}$, where $T > 0$ is a fixed constant. We interpret $F \in \mathbb{F}$ as a *financial position* of a trader in a financial market driven by the Lévy process $\eta(t)$.

For simplicity we assume that

$$E[\eta^2(t)] < \infty \qquad \text{for all } t \geq 0.$$

[*]Center of Mathematics for Applications (CMA), Dept. of Mathematics, University of Oslo, P.O. Box 1053 Blindern, N–0316 Oslo, Norway, email: `oksendal@math.uio.no`

[†]Norwegian School of Economics and Business Administration, Helleveien 30, N–5045 Bergen, Norway

[‡]INRIA Paris-Rocquencourt, Domaine de Voluceau, BP 105, Le Chesnay Cedex, 78153, France, email: `agnes.sulem@inria.fr`

This implies that $\eta(t)$ has the representation

$$\eta(t) = at + bB(t) + \int_0^t \int_{\mathbb{R}_0} z\tilde{N}(ds, dz); \qquad t \geq 0 \tag{1.1}$$

where a, b are constants, $B(t)$ is an $\mathcal{F}_t$-Brownian motion, $\mathbb{R}_0 = \mathbb{R} \setminus \{0\}$ and

$$\tilde{N}(dt, dz) = N(dt, dz) - \nu(dz)dt$$

is the *compensation* of the *jump measure* $N(\cdot, \cdot)$ of $\eta(\cdot)$, ν being the *Lévy measure* of $\eta(\cdot)$.

We refer e.g. to Øksendal and Sulem [ØS2] and Applebaum [A] for more information about the stochastic control and calculus of processes driven by Lévy processes.

Let $\mathcal{R}$ denote the set of all functions $k : \mathbb{R}_0 \to \mathbb{R}$. In the following we fix a *convex function*

$$g : \mathbb{R} \times \mathbb{R} \times \mathcal{R} \to \mathbb{R}$$

and we consider the backward stochastic differential equation (BSDE)

$$\begin{cases} dX(t) = -g(X(t), Y(t), K(t, \cdot))dt + Y(t)dB(t) + \displaystyle\int_{\mathbb{R}_0} K(t, z)\tilde{N}(dt, dz) \\ X(T) = F. \end{cases} \tag{1.2}$$

in the 3 unknown $\mathcal{F}_t$-predictable processes $X(t), Y(t), K(t, z)$. We assume that this equation has a unique solution, which we denote by $(X_g^F(t), Y_g^F(t), K_g^F(t, z))$. BSDE's of this form are related to the concept of *convex risk measures*, described as follows:

Risk measures. Let $\mathbb{F} = L^p(\mathcal{F}_T, P)$ $(p \in [1, \infty])$ represent the family of all financial standings at time T. A *convex risk measure* is a map

$$\rho : \mathbb{F} \to \mathbb{R}$$

satisfying the following conditions

(convexity) $\rho(\lambda F + (1 - \lambda)G) \leq \lambda\rho(F) + (1 - \lambda)\rho(G)$ for all, $F, G \in \mathbb{F}, \ \lambda \in (0, 1).$
$$\tag{1.3}$$

(monotonicity) $F \leq G \Rightarrow \rho(G) \leq \rho(F)$; $F, G \in \mathbb{F}.$ $\tag{1.4}$

(See e.g. Frittelli and Rosazza-Gianin [FR1], Föllmer and Schied [FS] and Biagini and Frittelli [BF]).

Intuitively, $\rho(F)$ may be interpreted as the the amount the agent has to hold to cancel the risk associated with his risky position F, that is $\rho(F + \rho(F)) = 0$ (see [BE]).

There are several useful representations of convex risk measures. One of them is the following:

Theorem 1.1 (Frittelli and Rosazza-Gianin [FR1], Föllmer and Schied [FS])

Every convex risk measure $\rho : \mathbb{F} \to \mathbb{R}$ which also satisfies the translation invariance property

$$\rho(F + a) = \rho(F) - a \; ; \; \forall F \in \mathbb{F}, \; a \in \mathbb{R} \tag{1.5}$$

is of the form

$$\rho(F) = \sup_{Q \in \mathcal{P}} \{ E_Q[-F] - \zeta(Q) \} \; ; \; F \in \mathbb{F} \tag{1.6}$$

for some family $\mathcal{P}$ of probability measures absolutely continuous with respect to P and some convex "penalty" function $\zeta : \mathcal{P} \to \mathbb{R}$.

For example, if $\mathcal{P}$ is the set of all probability measures $Q \ll P$ and ζ is *the relative entropy*

$$\zeta(Q) = E\left[\frac{dQ}{dP} \ln\left(\frac{dQ}{dP} \right) \right] = H(Q, P) \tag{1.7}$$

then

$$\rho(F) := \sup_{Q \ll P} \{ E_Q[-F] - H(Q, P) \} \tag{1.8}$$

is called the *entropic risk measure.*

If one uses the representation (1.6) for the convex risk measure and wants to find the portfolio that *minimizes the risk of the corresponding terminal wealth*; one arrives at a *stochastic differential game.* This approach has been studied recently by Mataramvura and Øksendal [MØ]. A corresponding *risk indifference pricing method* is studied in Øksendal and Sulem [ØS1].

In this project we study the risk minimization problem by using another representation of the convex risk measure, namely the representation in terms of *g-expectation* (See e.g. Peng [P], Rosazza-Gianin [R]).

Definition 1.2 *The risk $\rho(F) = \rho_g(F)$ (associated to the convex function g) of a financial position $F \in \mathbb{F}$ is defined by*

$$\rho(F) := \mathcal{E}_g[-F] := X_g^{-F}(0) \in \mathbb{R}, \tag{1.9}$$

where $X_g^{-F}(0)$ is the value at $t = 0$ of the solution $X(t)$ of the BSDE (1.2) with terminal value $-F$.

Remark 1.3 The number $X_g^{-F}(0) = \mathcal{E}_g(-F)$ is called *the g-expectation of $-F$.* In the *Brownian motion case* this concept has been studied extensively in the literature. See e.g. Peng [P] or Frittelli & Rosazza-Gianin [FR1] for more information about g-expectations and its relation to convex risk measures. In this paper we extend this concept to *the Lévy process case*, and we use it to study *risk minimization* in financial markets.

3

Remark 1.4 One can show that since g is convex the function $\rho_g : \mathbb{F} \to \mathbb{R}$ is indeed a convex risk measure. The function g may be regarded as a measure of risk aversion. For example, one can show that (see Coquet et al [CHMP])

$$g \leq \tilde{g} \Rightarrow \rho_g(F) \leq \rho_{\tilde{g}}(F) \; ; \; F \in \mathbb{F}.$$

If the function g does not depend on x, i.e. if

$$g(x, y, k) = g(y, k),$$

then ρ_g also satisfies the translation property (1.5) and then g is related to the penalty function ζ in the representation (1.6). For example, if we choose $\nu = 0$ and

$$g(x, y, k) = \frac{1}{2}y^2 \; ; \; x \in \mathbb{R}$$

then

$$\rho_g(F) = \mathcal{E}_g[-F] = \sup_{Q \ll P} \{E_Q[-F] - H(Q, P)\}$$

coincides with the *entropic risk measure*. See Barrieu and El Karoui [BE] for more information.

Risk minimizing portfolios   Suppose we have a financial market with the following 2 investment possibilities:

$$A \textit{ risk free asset}, \text{ with unit price } S_0(t) = 1 \quad \text{for all } t \in [0, T] \tag{1.10}$$

$$A \text{ risky asset, with unit price } S(t) \text{ at time } t \text{ given by} \tag{1.11}$$

$$dS(t) = S(t^-)\Big[\alpha(t)dt + \beta(t)dB(t) + \int_{\mathbb{R}_0} \theta(t, z)\tilde{N}(dt, dz)\Big]; \quad S(0) > 0$$

where $\alpha(t), \beta(t)$ and $\theta(t, z)$ are $\mathcal{F}_t$-predictable processes such that $\theta(t, z) \geq -1 + \epsilon$ for some $\epsilon > 0$ and

$$\int_0^T \Big\{|\alpha(t)| + \beta^2(t) + \int_{\mathbb{R}_0} \theta^2(t, z)\nu(dz)\Big\}dt < \infty \quad \text{a.s.}$$

Suppose we are given a subfiltration

$$\mathcal{E}_t \subseteq \mathcal{F}_t \qquad \text{for all } t \in [0, T]$$

representing the *information available* to the trader at time t. This means that the *portfolio* $\pi(t)$ of the trader, representing the fraction of the total wealth invested in the risky asset at t, is required to be $\mathcal{E}_t$-*predictable*. The wealth process $A(t) = A_\pi(t)$ corresponding to the portfolio $\pi(t)$ is given by

$$\begin{cases} dA(t) = A(t^-)\pi(t)\Big[\alpha(t)dt + \beta(t)dB(t) + \int_{\mathbb{R}_0} \theta(t, z)\tilde{N}(dt, dz)\Big] \\ A(0) = a > 0. \end{cases} \tag{1.12}$$

4

We say that the portfolio $\pi(t)$ is *admissible* and write $\pi \in \mathcal{A}_{\mathcal{E}}$ if π is $\mathcal{E}_t$-*predictable* and

$$\int_0^T \left\{ |\alpha(t)\pi(t)| + \pi^2(t)\beta^2(t) + \pi^2(t)\int_{\mathbb{R}_0} \theta^2(t,z)\nu(dz) \right\} dt < \infty \quad \text{a.s.}$$

The risk minimizing portfolio problem is to find the portfolio $\pi^* \in \mathcal{A}_{\mathcal{E}}$ which minimizes the risk of the terminal wealth, i.e.

$$\inf_{\pi \in \mathcal{A}_{\mathcal{E}}} \rho_g(A_\pi(T)) = \rho_g(A_{\pi^*}(T)). \tag{1.13}$$

In view of (1.9) this is equivalent to

$$\inf_{\pi \in \mathcal{A}_{\mathcal{E}}} X_g^{-A_\pi(T)}(0) = X_g^{-A_{\pi^*}(T)}(0), \tag{1.14}$$

where $X_g^{-A_\pi(T)}(t) = X(t)$ *is given by* (1.2) *with* $F = -A_\pi(T)$. This is an example of a *partial information stochastic control problem of a system of forward* (1.12)-*backward* (1.2) *stochastic differential equations (FBSDEs), driven by Lévy processes.*

In section 2 we study the general partial information optimal control problem for FBSDEs and we prove a partial information *sufficient* maximum principle for such problems. In Section 3 we prove a partial information *necessary* maximum principle. More precisely, we prove the equivalence between being a critical point for the performance functional of a partial information FBSDE problem and being a conditional critical point for the associated Hamiltonian. A drawback of the methods of the previous sections is the assumption of the existence of solutions of the BSDEs of the adjoint processes. In Section 4 we present a method based on Malliavin calculus, where this assumption is not needed. Moreover, this approach allows us to handle non-Markovian systems. Finally, in Section 5 we apply our general results to the risk minimization problem mentioned in the introduction, and we find explicit expressions for the minimal risk in special cases.

2 A maximum principle for optimal control of forward-backward SDE

In this section we present a general sufficient maximum principle for optimal control with *partial information* of a stochastic system consisting of a forward and a backward SDE driven by *Lévy processes*. Such a maximum principle is also studied by Bahlali [B] and Shi and Wu [SW] in the complete information case (and [B] with the Brownian motion case only), and their proofs differ from ours. Related earlier results are Framstad, Øksendal and Sulem [FØS] (maximum principle for (forward) jump diffusions), Baghery and Øksendal [BØ] (partial information case (2007)) and Meyer-Brandis, Øksendal and Zhou [MØZ] (partial information, non-Markovian system, Malliavin calculus for Lévy processes). See also [PW].

For simplicity we present the 1-dimensional case only. Suppose the state $(A(t), X(t))$ of our system is described by the following coupled forward-backward system of SDE's:

(*Forward* system in the controlled process $A(t)$)

$$\begin{cases} dA(t) = b(t, A(t), u(t))dt + \sigma(t, A(t), u(t))dB(t) \\ \quad + \displaystyle\int_{\mathbb{R}_0} \gamma(t, A(t), u(t), z)\tilde{N}(dt, dz); \quad t \in [0, T] \\ A(0) = a \in \mathbb{R} \end{cases} \tag{2.1}$$

(*Backward* system in the unknown processes $X(t), Y(t), K(t, z)$)

$$\begin{cases} dX(t) = -g(t, A(t), X(t), Y(t), u(t))dt + Y(t)dB(t) \\ \quad + \displaystyle\int_{\mathbb{R}_0} K(t, z)\tilde{N}(dt, dz); \quad t \in [0, T] \\ X(T) = cA(T), \text{ where } c \in \mathbb{R}\backslash\{0\} \text{ is a given constant.} \end{cases} \tag{2.2}$$

Suppose in addition that we are given a subfiltration

$$\mathcal{E}_t \subseteq \mathcal{F}_t; \qquad t \in [0, T]$$

representing the information available to the controller at time t. For example, we could have

$$\mathcal{E}_t = \mathcal{F}_{(t-\delta)^+} \qquad (\delta > 0 \text{ constant})$$

meaning that the controller gets a *delayed* information flow compared to $\mathcal{F}_t$.

Let $\mathcal{A}_\mathcal{E}$ denote a given family of controls, contained in the set of $\mathcal{E}_t$-predictable controls $u(\cdot)$ such that the system (2.1)–(2.2) has a unique strong solution. If $u \in \mathcal{A}_\mathcal{E}$ we call u an *admissible* control. Let $U \in \mathbb{R}$ be a given convex set such that $u(t) \in U$ for all $t \in [0, T]$ a.s., for all $u \in \mathcal{A}_\mathcal{E}$.

Suppose we are given a *performance functional* of the form

$$\begin{aligned} J(u) = E\Big[&\int_0^T f(t, A(t), X(t), Y(t), K(t, \cdot), u(t))dt \\ &+ h_1(X(0)) + h_2(A(T))\Big]; \qquad u \in \mathcal{A}_\mathcal{E} \end{aligned} \tag{2.3}$$

where E denotes expectation with respect to P and f, h_1, h_2 are given functions such that

$$E\Big[\int_0^T |f(t, A(t), X(t), Y(t), K(t, \cdot), u(t))|dt + |h_1(X(0))| + |h_2(A(T))| \Big] < \infty.$$

The problem we consider is the following:

Problem 2.1 (Partial information optimal control of forward-backward SDEs) Find $\Phi_\mathcal{E} \in \mathbb{R}$ and $u^* \in \mathcal{A}_\mathcal{E}$ such that

$$\Phi_\mathcal{E} = \sup_{u \in \mathcal{A}_\mathcal{E}} J(u) = J(u^*). \tag{2.4}$$

The *Hamiltonian*

$$H : [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times L^2(\nu) \times U \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times L^2(\nu) \to \mathbb{R}$$

is defined by

$$
\begin{aligned}
H(t, & a, x, y, k, u, \lambda, p, q, r(.)) \\
& = f(t, a, x, y, k, u) + g(t, a, x, y, u)\lambda + b(t, a, u)p \\
& + \sigma(t, a, u)q + \int_{\mathbb{R}_0} \gamma(t, a, u, z)r(z)\nu(dz)
\end{aligned}
\tag{2.5}
$$

We assume that H is Frechet differentiable (C^1) in the variables a, x, y, k (See Remark 2.2 below). To this problem we associate a pair of forward-backward SDEs in the *adjoint processes* $\lambda(t), (p(t), q(t), r(t, \cdot))$ as follows

(Forward system in the unknown process $\lambda(t)$)

$$
\begin{cases}
d\lambda(t) = & \dfrac{\partial H}{\partial x}(t, A(t), X(t), Y(t), K(t, \cdot), u(t), \lambda(t), p(t), q(t), r(t, \cdot))dt \\
& + \dfrac{\partial H}{\partial y}(t, A(t), X(t), Y(t), K(t, \cdot), u(t), \lambda(t), p(t), q(t), r(t, \cdot))dB(t) \\
& + \displaystyle\int_{\mathbb{R}_0} \nabla_k H(t, A(t), X(t), Y(t), K(t, \cdot), u(t), \lambda(t), p(t), q(t), r(t, \cdot))\tilde{N}(dt, dz) \\
\lambda(0) = & h_1'(X(0)) \quad \left(= \dfrac{dh_1}{dx}(X(0))\right)
\end{cases}
\tag{2.6}
$$

(Backward system in the unknown processes $p(t), q(t), r(t, \cdot)$)

$$
\begin{cases}
dp(t) = & -\dfrac{\partial H}{\partial a}(t, A(t), X(t), Y(t), K(t, \cdot), \lambda(t), p(t), q(t), r(t, \cdot))dt \\
& + q(t)dB(t) + \displaystyle\int_{\mathbb{R}} r(t, z)\tilde{N}(dt, dz); \quad t \in [0, T] \\
p(T) = & c\lambda(T) + h_2'(A(T))
\end{cases}
\tag{2.7}
$$

Remark 2.2 Let V be an open subset of a Banach space χ and let $F : V \to \mathbb{R}$.

(i) We say that F has a directional derivative (or Gateaux derivative) at $x \in V$ in the direction $y \in \chi$ if

$$D_y F(x) := \lim_{\epsilon \to 0} \frac{1}{\epsilon}(F(x + \epsilon y) - F(x)) \quad \text{exists}$$

(ii) We say that F is Fréchet differentiable at $x \in V$ if there exists a linear map

$$L : \chi \to \mathbb{R}$$

such that

$$\lim_{\substack{h \to 0 \\ h \in \chi}} \frac{1}{\|h\|} |F(x + h) - F(x) - L(h)| = 0.$$

In this case we call L the Fréchet derivative of F at x and we write

$$L = \nabla_x F.$$

7

(iii) If F is Fréchet differentiable, then F has a directional derivative in all directions $y \in \chi$ and

$$D_y F(x) = \nabla_x F(y).$$

Theorem 2.3 (Sufficient conditional maximum principle) *Let $\hat{u} \in \mathcal{A}_\mathcal{E}$ with corresponding solutions $\hat{A}(t), \hat{X}(t), \hat{Y}(t), \hat{K}(t, z), \hat{\lambda}(t), \hat{p}(t), \hat{q}(t)$ and $\hat{r}(t, z)$ of equations (2.1), (2.2), (2.6) and (2.7). Suppose that*

(Concavity) The functions $x \to h_i(x)$; $i = 1, 2$ and

$(a, x, y, k, u) \to H(t, a, x, y, k, u, \hat{\lambda}(t), \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot))$

are concave, for all $t \in [0, T]$. $\qquad\qquad$ (2.8)

(The conditional maximum principle)

$$\max_{v \in U} E[H(t, \hat{A}(t), \hat{X}(t), \hat{Y}(t), \hat{K}(t, \cdot), v, \hat{\lambda}(t), \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot)) \mid \mathcal{E}_t]$$
$$= E[H(t, \hat{A}(t), \hat{X}(t), \hat{Y}(t), \hat{K}(t, \cdot), \hat{u}(t), \hat{\lambda}(t), \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot)) \mid \mathcal{E}_t]. \qquad (2.9)$$

Moreover, suppose that for all $u \in \mathcal{A}_\mathcal{E}$ the following growth conditions hold:

$$E\Big[\int_0^T X^2(t) \Big\{ \Big(\frac{\partial \hat{H}}{\partial y}(t) \Big)^2 + \int_{\mathbb{R}_0} \|\nabla_k \hat{H}(t, z)\|^2 \nu(dz) \Big\} dt \Big] < \infty \qquad (2.10)$$

$$E\Big[\int_0^T \hat{\lambda}(t)^2 \Big\{ Y^2(t) + \int_{\mathbb{R}_0} K^2(t, z) \nu(dz) \Big\} dt \Big] < \infty \qquad (2.11)$$

$$E\Big[\int_0^T A^2(t) \Big\{ \hat{q}^2(t) + \int_{\mathbb{R}_0} \hat{r}(t, z)^2 \nu(dz) \Big\} dt \Big] < \infty \qquad (2.12)$$

$$E\Big[\int_0^T \hat{p}(t)^2 \Big\{ \sigma^2(t) + \int_{\mathbb{R}_0} \gamma^2(t, z) \nu(dz) \Big\} dt \Big] < \infty, \qquad (2.13)$$

where $A(t), X(t), Y(t), K(t, z)$ are the solutions of (2.1), (2.2) corresponding to u, and we are using the notation

$$\frac{\partial \hat{H}}{\partial y}(t) = \frac{d}{dy} H\big(t, \hat{A}(t), \hat{X}(t), y, \hat{K}(t, \cdot), \hat{u}(t), \hat{\lambda}(t), \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot)\big)_{y = \hat{Y}(t)}$$

and similarly with $\nabla_k \hat{H}(t, z)$.

Then $\hat{u}(t)$ is an optimal control for Problem 2.1, i.e.

$$J(\hat{u}) = \sup_{u \in \mathcal{A}_\mathcal{E}} J(u).$$

Proof.   Choose $u \in \mathcal{A}$ with corresponding solutions $A(t), X(t), Y(t), K(t, \cdot), \lambda(t), p(t), q(t),$ $r(t, \cdot)$. In the following we write (see (2.5))

$$\hat{H}(t) = H(t, \hat{A}(t), \hat{X}(t), \hat{Y}(t), \hat{K}(t, \cdot), \hat{u}(t), \hat{\lambda}(t), \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot))$$
$$= f(t, \hat{A}(t), \hat{X}(t), \hat{Y}(t), \hat{K}(t, \cdot), \hat{u}(t)) + g(t, \hat{A}(t), \hat{X}(t), \hat{Y}(t), \hat{u}(t))\hat{\lambda}(t)$$
$$+ b(t, \hat{A}(t), \hat{u}(t))\hat{p}(t) + \sigma(t, \hat{A}(t), \hat{u}(t))\hat{q}(t) + \int_{\mathbb{R}_0} \gamma(t, \hat{A}(t), \hat{u}(t), z)\hat{r}(t, z)\nu(dz)$$

$$H(t) = H(t, A(t), X(t), Y(t), K(t, \cdot), u(t), \hat{\lambda}(t), \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot))$$
$$= f(t, A(t), X(t), Y(t), K(t, \cdot), u(t)) + g(t, A(t), X(t), Y(t), u(t))\hat{\lambda}(t)$$
$$+ b(t, A(t), u(t))\hat{p}(t) + \sigma(t, A(t), u(t))\hat{q}(t) + \int_{\mathbb{R}_0} \gamma(t, A(t), u(t), z)\hat{r}(t, z)\nu(dz)$$

and similarly
$$\hat{f}(t) = f(t, \hat{A}(t), \hat{X}(t), \hat{Y}(t), \hat{K}(t, \cdot), \hat{u}(t)); \quad f(t) = f(t, A(t), X(t), Y(t), K(t, \cdot), u(t)) \dots \text{etc.}$$
Then
$$J(\hat{u}) - J(u) = I_1 + I_2, \tag{2.14}$$

where
$$I_1 = E\Big[\int_0^T \{\hat{f}(t) - f(t)\}dt \Big] \tag{2.15}$$

and
$$I_2 = E[h_1(\hat{X}(0)) - h_1(X(0)) + h_2(\hat{A}(T)) - h_2(A(T))]. \tag{2.16}$$

By the definition of H we have

$$I_1 = E\Big[\int_0^T \{\hat{H}(t) - H(t) - (\hat{g}(t) - g(t))\hat{\lambda}(t)$$
$$- (\hat{b}(t) - b(t))\hat{p}(t) - (\hat{\sigma}(t) - \sigma(t))\hat{q}(t)$$
$$- \int_{\mathbb{R}_0} (\hat{\gamma}(t, z) - \gamma(t, z))\hat{r}(t, z)\nu(dz)\}dt \Big]. \tag{2.17}$$

Since h_1 is concave, we have
$$h_1(\hat{X}(0)) - h_1(X(0)) \geq (\hat{X}(0) - X(0))h_1'(\hat{X}(0)) = (\hat{X}(0) - X(0))\hat{\lambda}(0). \tag{2.18}$$

Since h_2 is concave, we have
$$h_2(\hat{A}(T)) - h_2(A(T)) \geq (\hat{A}(T) - A(T))h_2'(\hat{A}(T)). \tag{2.19}$$

9

By the Itô formula and (2.2) and (2.7) we get

$$
\begin{aligned}
E[(\hat{X}(0 - X(0))\hat{\lambda}(0)] &= E[(\hat{X}(T) - X(T))\hat{\lambda}(T)] \\
&\quad - E\Big[\int_0^T (\hat{X}(t) - X(t))d\hat{\lambda}(t) + \int_0^T \hat{\lambda}(t)d(\hat{X}(t) - X(t)) \\
&\quad + \int_0^T \frac{\partial \hat{H}}{\partial y}(t)(\hat{Y}(t) - Y(t))dt + \int_0^T \int_{\mathbb{R}_0} \nabla_k \hat{H}(t, z)(\hat{K}(t, z) - K(t, z))\nu(dz)dt \Big] \\
&= E[(\hat{A}(T) - A(T))(\hat{p}(T) - h_2'(\hat{A}(T)))] \\
&\quad - E\Big[\int_0^T \Big\{ (\hat{X}(t) - X(t))\frac{\partial \hat{H}}{\partial x}(t) - \hat{\lambda}(t)(\hat{g}(t) - g(t)) \\
&\quad + \frac{\partial \hat{H}}{\partial y}(t)(\hat{Y}(t) - Y(t)) + \int_{\mathbb{R}_0} \nabla_k \hat{H}(t, z)(\hat{K}(t, z) - K(t, z))\nu(dz) \Big\}dt \\
&= E\Big[\int_0^T (\hat{A}(t) - A(t))d\hat{p}(t) + \int_0^T \hat{p}(t)d(\hat{A}(t) - A(t)) \\
&\quad + \int_0^T (\hat{\sigma}(t) - \sigma(t))\hat{q}(t)dt + \int_0^T \int_{\mathbb{R}_0} (\hat{\gamma}(t, z) - \gamma(t, z))\hat{r}(t, z)\nu(dz)dt \\
&\quad - \int_0^T \Big\{ (\hat{X}(t) - X(t))\frac{\partial \hat{H}}{\partial x}(t) - \hat{\lambda}(t)(\hat{g}(t) - g(t)) \\
&\quad + (\hat{Y}(t) - Y(t))\frac{\partial \hat{H}}{\partial y}(t) + \int_{\mathbb{R}_0} \nabla_k \hat{H}(t, z)(\hat{K}(t, z) - K(t, z))\nu(dz) \Big\}dt \Big] \\
&\quad - E[(\hat{A}(T) - A(T))h_2'(\hat{A}(T))]. \\
&= E\Big[\int_0^T \Big\{ (\hat{b}(t) - b(t))\hat{p}(t) + (\hat{\sigma}(t) - \sigma(t))\hat{q}(t) \\
&\quad + \int_{\mathbb{R}_0} (\hat{\gamma}(t, z) - \gamma(t, z))\hat{r}(t, z)\nu(dz) - (\hat{A}(t) - A(t))\frac{\partial \hat{H}}{\partial a}(t) \\
&\quad - (\hat{X}(t) - X(t))\frac{\partial \hat{H}}{\partial x}(t) + \hat{\lambda}(t)(\hat{g}(t) - g(t)) \\
&\quad - (\hat{Y}(t) - Y(t))\frac{\partial \hat{H}}{\partial y}(t) - \int_{\mathbb{R}_0} \nabla_k \hat{H}(t, z)(\hat{K}(t, z)) - K(t, z))\nu(dz) \Big\}dt \Big] \\
&\quad - E[(\hat{A}(T) - A(T))h_2'(\hat{A}(T))]. \qquad (2.20)
\end{aligned}
$$

Combining (2.14)–(2.20) we get

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

$$\geq E\Big[\int_0^T \Big\{ \hat{H}(t) - H(t) - (\hat{A}(t) - A(t))\frac{\partial\hat{H}}{\partial a}(t) - (\hat{X}(t) - X(t))\frac{\partial\hat{H}}{\partial x}(t)$$

$$- (\hat{Y}(t) - Y(t))\frac{\partial\hat{H}}{\partial y}(t) - \int_{\mathbb{R}_0} \nabla_k \hat{H}(t,z)(\hat{K}(t,z) - K(t,z))\nu(dz)\Big\}dt\Big]$$

$$= E\Big[\int_0^T E\Big[\Big\{ \hat{H}(t) - H(t) - (\hat{A}(t) - A(t))\frac{\partial\hat{H}}{\partial a}(t) - (\hat{X}(t) - X(t))\frac{\partial\hat{H}}{\partial x}(t)$$

$$- (\hat{Y}(t) - Y(t))\frac{\partial\hat{H}}{\partial y}(t) - \int_{\mathbb{R}_0} \nabla_k \hat{H}(t,z)(\hat{K}(t,z) - K(t,z))\nu(dz)\Big\} \mid \mathcal{E}_t\Big]dt\Big]. \qquad (2.21)$$

Since the function

$$(a,x,y,k,u) \rightarrow H(t,a,x,y,k,u,\hat{\lambda}(t),\hat{p}(t),\hat{q}(t),\hat{r}(t,\cdot))$$

is concave, we have

$$\hat{H}(t) - H(t) \geq \frac{\partial\hat{H}}{\partial a}(t)(\hat{A}(t) - A(t)) + \frac{\partial\hat{H}}{\partial x}(t)(\hat{X}(t) - X(t))$$

$$+ \frac{\partial\hat{H}}{\partial y}(t)(\hat{Y}(t) - Y(t)) + \int_{\mathbb{R}_0} \nabla_k \hat{H}(t,z)(\hat{K}(t,z) - K(t,z))\nu(dz)$$

$$+ \frac{\partial\hat{H}}{\partial u}(t)(\hat{u}(t) - u(t)). \qquad (2.22)$$

Since $u = \hat{u}(t)$ maximizes

$$u \rightarrow E[H(t,\hat{A}(t),\hat{X}(t),\hat{Y}(t),\hat{K}(t,\cdot),u,\hat{\lambda}(t),\hat{p}(t),\hat{q}(t),\hat{r}(t,\cdot)) \mid \mathcal{E}_t]$$

we deduce that

$$\frac{d}{du}E[H(t,\hat{A}(t),\hat{X}(t),\hat{Y}(t),\hat{K}(t,\cdot),u,$$

$$\hat{\lambda}(t),\hat{p}(t),\hat{q}(t),\hat{r}(t,\cdot)) \mid \mathcal{E}_t]_{u=\hat{u}(t)}(\hat{u}(t) - u(t)) \geq 0$$

i.e.

$$E\Big[\frac{\partial\hat{H}}{\partial u}(t)(\hat{u}(t) - u(t)) \mid \mathcal{E}_t\Big] \geq 0. \qquad (2.23)$$

Combining this with (2.21) and (2.22) we conclude that

$$J(\hat{u}) - J(u) \geq 0.$$

Since this holds for all $u \in \mathcal{A}_\mathcal{E}$, $\hat{u}$ is optimal. $\qquad\square$

3 A partial information equivalence principle for forward-backward SDE's

A drawback with the result in the previous section is that in many applications the concavity condition may not hold. In this section, we prove a version of the maximum principle which does not need this assumption. Instead, we assume the following:

(A1) For all $s \in [0, T)$ and all bounded $\mathcal{E}_s$-measurable random variables $\theta(\omega)$ the control β_s defined by

$$\beta_s(t) = \theta(\omega)\chi_{(s,T]}(t) \; ; \; t \in [0, T]$$

is in $\mathcal{A}_\mathcal{E}$.

(A2) For all $u, \beta \in \mathcal{A}_\mathcal{E}$ where β is bounded there exists $\delta > 0$ such that the control

$$u(t) + y\beta(t) \; ; \; t \in [0, T]$$

belongs to $\mathcal{A}_\mathcal{E}$ for all $y \in (-\delta, \delta)$.

Theorem 3.1 (Partial information equivalence principle) *Suppose $u \in \mathcal{A}_\mathcal{E}$ with corresponding solutions $A(t)$, $X(t)$, $Y(t)$, $K(t, z)$, $\lambda(t)$, $p(t)$, $q(t)$ and $r(t, z)$ of (2.1), (2.2), (2.6) and (2.7). Assume that (2.10)-(2.13) hold. Then the following are equivalent:*

(i) $\dfrac{d}{dy}J(u + y\beta) \mid_{y=0} = 0$ *for all bounded $\beta \in \mathcal{A}_\mathcal{E}$*

(ii) $E\left[\dfrac{\partial}{\partial u}H(t, A(t), X(t), Y(t), K(t, \cdot), u, \lambda(t), p(t), q(t), r(t, \cdot))_{u=u(t)} \mid \mathcal{E}_t\right] = 0$

Proof.   Define

$$\alpha(t) = \frac{d}{dy}A_{u+y\beta}(t) \mid_{y=0}$$

$$\xi(t) = \frac{d}{dy}X_{u+y\beta}(t) \mid_{y=0}$$

$$\eta(t) = \frac{d}{dy}Y_{u+y\beta}(t) \mid_{y=0}$$

$$\zeta(t, z) = \frac{d}{dy}K_{u+y\beta}(t, z) \mid_{y=0}$$

Note that

$$\alpha(0) = \frac{d}{dy}A_{u+y\beta}(0) \mid_{y=0} = 0$$

and

$$\alpha(T) = \frac{d}{dy}A_{u+y\beta}(T) \mid_{y=0} = \frac{1}{c}\frac{d}{dy}X_{u+y\beta}(T) \mid_{y=0} = \frac{1}{c}\xi(T).$$

12

With this notation we have, writing

$$\frac{\partial b}{\partial a}(t) = \frac{\partial b}{\partial a}(t, A(t), u(t)) \text{ etc.,}$$

$$d\alpha(t) = \left\{ \frac{\partial b}{\partial a}(t)\alpha(t) + \frac{\partial b}{\partial u}(t)\beta(t) \right\} dt$$

$$+ \left\{ \frac{\partial \sigma}{\partial a}(t)\alpha(t) + \frac{\partial \sigma}{\partial u}(t)\beta(t) \right\} dB(t)$$

$$+ \int_{\mathbb{R}_0} \left\{ \frac{\partial \gamma}{\partial a}(t, z)\alpha(t) + \frac{\partial \gamma}{\partial u}(t, z)\beta(t) \right\} \tilde{N}(dt, dz) \tag{3.1}$$

$$d\xi(t) = \left\{ -\frac{\partial g}{\partial a}(t)\alpha(t) - \frac{\partial g}{\partial x}(t)\xi(t) - \frac{\partial g}{\partial y}(t)\eta(t) - \frac{\partial g}{\partial u}(t)\beta(t) \right\} dt$$

$$+ \eta(t)dB(t) + \int_{\mathbb{R}_0} \zeta(t, z)\tilde{N}(dt, dz) \tag{3.2}$$

Assume that (i) holds. Then

$$0 = \frac{d}{dy} J(u + y\beta) \mid_{y=0}$$

$$= E \left[\int_0^T \left\{ \frac{\partial f}{\partial a}(t)\alpha(t) + \frac{\partial f}{\partial x}\xi(t) + \frac{\partial f}{\partial y}(t)\eta(t) + \nabla_k f(t, z)\zeta(t, z) + \frac{\partial f}{\partial u}(t)\beta(t) \right\} dt \right.$$

$$\left. + h_1'(X(0))\xi(0) + h_2'(A(T))\alpha(T) \right] \tag{3.3}$$

By the Itô formula,

$$E[h_1'(X(0))\xi(0)] = E[\lambda(0)\xi(0)]$$

$$= E \left[\lambda(T)\xi(T) - \int_0^T \lambda(t)d\xi(t) - \int_0^T \xi(t)d\lambda(t) \right.$$

$$\left. - \int_0^T \frac{\partial H}{\partial y}(t)\eta(t)dt - \int_0^T \int_{\mathbb{R}_0} \nabla_k H(t, z)\zeta(t, z)\nu(dz)dt \right]$$

$$= E \left[\lambda(T)\xi(T) - \int_0^T \lambda(t) \left\{ -\frac{\partial g}{\partial a}(t)\alpha(t) - \frac{\partial g}{\partial x}(t)\xi(t) \right. \right.$$

$$\left. - \frac{\partial g}{\partial y}(t)\eta(t) - \frac{\partial g}{\partial u}(t)\beta(t) \right\} dt - \int_0^T \xi(t)\frac{\partial H}{\partial x}(t)dt$$

$$\left. - \int_0^T \eta(t)\frac{\partial H}{\partial y}(t)dt - \int_0^T \int_{\mathbb{R}_0} \nabla_k H(t, z)\zeta(t, z)\nu(dz)dt \right]. \tag{3.4}$$

13

Similarly, by (2.7),

$$
\begin{aligned}
E[h_2'(A(T))\alpha(T)] &= E[(p(T) - c\lambda(T))\alpha(T)] \\
&= E[p(T)\alpha(T)] - E[\lambda(T)\xi(T)] \\
&= \left[\int_0^T p(t)d\alpha(t) + \int_0^T \alpha(t)dp(t) \right. \\
&\quad + \int_0^T q(t) \left\{ \frac{\partial \sigma}{\partial a}(t)\alpha(t) + \frac{\partial \sigma}{\partial u}(t)\beta(t) \right\} dt \\
&\quad + \left. \int_0^T \int_{\mathbb{R}_0} r(t,z) \left\{ \frac{\partial \gamma}{\partial a}(t,z)\alpha(t) + \frac{\partial \gamma}{\partial u}(t,z)\beta(t) \right\} \nu(dz)dt \right] \\
&\quad - E[\lambda(T)\xi(T)] \\
&= E\left[\int_0^T \left(p(t) \left\{ \frac{\partial b}{\partial a}(t)\alpha(t) + \frac{\partial b}{\partial u}(t)\beta(t) \right\} - \frac{\partial H}{\partial a}(t)\alpha(t) \right. \right. \\
&\quad + q(t) \left\{ \frac{\partial \sigma}{\partial a}(t,z)\alpha(t) + \frac{\partial \sigma}{\partial u}(t)\beta(t) \right\} \\
&\quad + \left. \left. \int_{\mathbb{R}} r(t,z) \left\{ \frac{\partial \gamma}{\partial a}(t,z)\alpha(t) + \frac{\partial \gamma}{\partial u}(t,z)\beta(t) \right\} \nu(dz) \right) dt \right] \\
&\quad - E[\lambda(T)\xi(T)].
\end{aligned}
\tag{3.5}
$$

Combining (3.3), (3.4) and (3.5) we get

$$
\begin{aligned}
0 = E\left[\int_0^T \right. & \left\{ \frac{\partial f}{\partial a}(t)\alpha(t) + \frac{\partial f}{\partial x}(t)\xi(t) + \frac{\partial f}{\partial y}(t)\eta(t) \right. \\
& + \int_{\mathbb{R}_0} \nabla_k f(t,z)\zeta(t,z)\nu(dz) + \frac{\partial f}{\partial u}(t)\beta(t) \\
& + \lambda(t) \left(\frac{\partial g}{\partial a}(t)\alpha(t) + \frac{\partial g}{\partial x}(t)\xi(t) + \frac{\partial g}{\partial y}(t)\eta(t) + \frac{\partial g}{\partial u}(t)\beta(t) \right) \\
& - \frac{\partial H}{\partial x}(t)\xi(t) - \frac{\partial H}{\partial y}(t)\eta(t) - \int_{\mathbb{R}_0} \nabla_k H(t,z)\zeta(t,z)\nu(dz) \\
& + p(t) \left(\frac{\partial b}{\partial a}(t)\alpha(t) + \frac{\partial b}{\partial u}(t)\beta(t) \right) - \frac{\partial H}{\partial a}(t)\alpha(t) \\
& + q(t) \left(\frac{\partial \sigma}{\partial a}(t)\alpha(t) + \frac{\partial \sigma}{\partial u}(t)\beta(t) \right) \\
& \left. \left. + \int_{\mathbb{R}_0} r(t,z) \left(\frac{\partial \gamma}{\partial a}(t,z)\alpha(t) + \frac{\partial \gamma}{\partial u}(t,z)\beta(t) \right) \nu(dz) \right\} dt \right].
\end{aligned}
$$

Collecting the terms with $\alpha(t)$, $\xi(t)$, $\eta(t)$, $\zeta(t,z)$ and $\beta(t)$, the above can be written

$$
\begin{aligned}
E\Bigg[\int_0^T \Bigg\{ &\left(\frac{\partial f}{\partial a}(t) + \frac{\partial g}{\partial a}(t)\lambda(t) + \frac{\partial b}{\partial a}(t)p(t) + \frac{\partial \sigma}{\partial a}(t)q(t)\right. \\
&\left. + \int_{\mathbb{R}_0} \frac{\partial \gamma}{\partial a}(t,z)r(t,z)\nu(dz) - \frac{\partial H}{\partial a}(t)\right)\alpha(t) \\
&+ \left(\frac{\partial f}{\partial x}(t) + \frac{\partial g}{\partial x}(t)\lambda(t) - \frac{\partial H}{\partial x}(t)\right)\xi(t) \\
&+ \left(\frac{\partial f}{\partial y}(t) + \frac{\partial g}{\partial y}(t)\lambda(t) - \frac{\partial H}{\partial y}(t)\right)\eta(t) \\
&+ \int_{\mathbb{R}_0} (\nabla_k f(t,z) - \nabla_k H(t,z))\zeta(t,z)\nu(dz) \\
&+ \left(\frac{\partial f}{\partial u}(t) + \frac{\partial g}{\partial u}(t)\lambda(t) + \frac{\partial b}{\partial u}(t)p(t) + \frac{\partial \sigma}{\partial u}(t)q(t)\right. \\
&\left. + \int_{\mathbb{R}_0} \frac{\partial \gamma}{\partial u}(t,z)r(t,z)\nu(dz)\right)\beta(t)\Bigg\} dt\Bigg] = 0
\end{aligned}
\tag{3.6}
$$

By the definition of H the coefficients of $\alpha(t)$, $\xi(t)$, $\eta(t)$ and $\zeta(t)$ are all 0, and we conclude that

$$
E\left[\int_0^T \frac{\partial H}{\partial u}(t)\beta(t)dt\right] = 0 \; ; \; \beta \in \mathcal{A}_{\mathcal{E}} \text{ bounded.}
\tag{3.7}
$$

In particular, this holds for all $\beta \in \mathcal{A}_{\mathcal{E}}$ of the form

$$
\beta(t) = \beta_s(t,\omega) = \theta(\omega)\chi_{[s,T]}(t) \; ; \; t \in [0,T]
\tag{3.8}
$$

for a fixed $s \in [0,T)$ where $\theta(\omega)$ is a bounded $\mathcal{E}_s$-measurable random variable.

This gives

$$
E\left[\int_s^T \frac{\partial H}{\partial u}(t)\theta dt\right] = 0.
$$

Differentiating with respect to s we arrive at

$$
E\left[\frac{\partial H}{\partial u}(s)\theta\right] = 0.
$$

Since this holds for all bounded $\mathcal{E}_s$-measurable random variables θ, we conclude that

$$
E\left[\frac{\partial H}{\partial u}(s) \mid \mathcal{E}_s\right] = 0.
\tag{3.9}
$$

This proves that (i) $\Rightarrow$ (ii).

Conversely, since every bounded $\beta \in \mathcal{A}_{\mathcal{E}}$ can be approximated by linear combinations of controls β_s of the form (3.8), we can prove that (ii) $\Rightarrow$ (i) by reversing the above argument.
$\square$

4 A Malliavin calculus approach

A major difficulty with the maximum principle formulations in Sections 2 and 3 is that they involve the adjoint processes $p(t), q(t)$ and $r(t, z)$. They are defined in terms of backward stochastic differential equations (BSDEs), which are usually hard to solve. In this section we propose a new approach based on Malliavin calculus. With this approach the adjoint processes above are replaced by processes $\tilde{p}(t)$, $\tilde{q}(t)$ and $\tilde{r}(t, z)$ which are given directly in terms of the parameters and state of the system, not by BSDEs. Moreover, this approach allows us to handle non-Markovian systems. Our method is an adaptation of the method in [MØZ] to our setting of stochastic control of forward-backward SDEs.

We keep the setup and notation in (2.1)-(2.4) above, except that now we allow the coefficients to be non-Markovian, i.e. we have

$$b(t, A(t), u(t)) = b(t, A(t), u(t), \omega),$$
$$\sigma(t, A(t), u(t)) = \sigma(t, A(t), u(t), \omega),$$
$$\gamma(t, A(t), u(t), z) = \gamma(t, A(t), u(t), z, \omega),$$
$$g(t, A(t), X(t), Y(t), u(t)) = g(t, A(t), X(t), Y(t), u(t), \omega),$$
$$f(t, A(t), X(t), Y(t), K(t, \cdot), u(t)) = f(t, A(t), X(t), Y(t), K(t, \cdot), u(t), \omega)$$

and

$$h_2(A(T)) = h_2(A(T), \omega),$$

where $\omega \to b(t, a, u, \omega)$ is $\mathcal{F}_t$-measurable for each constant a, u, and similarly with σ, γ, g and f.

In the following, $D_t F$ denotes the *Malliavin derivative with respect to $B(\cdot)$ (at t)* of a given (Malliavin differentiable) random variable $F = F(\omega)$; $\omega \in \Omega$. Similarly, $D_{t,z} F$ denotes the *Malliavin derivative with respect to $\tilde{N}(\cdot, \cdot)$ (at t, z)* of F. We let $\mathbb{D}_{1,2}$ denote the set of all random variables which are Malliavin differentiable with respect to both $B(\cdot)$ and $N(\cdot, \cdot)$. We will use the following *duality formulae* for Malliavin derivatives

$$E\left[F \int_0^T \varphi(s) dB(s)\right] = E\left[\int_0^T \varphi(s) D_s F ds\right] \tag{4.1}$$

$$E\left[F \int_0^T \int_{\mathbb{R}_0} \psi(s, z) \tilde{N}(ds, dz)\right] = E\left[\int_0^T \int_{\mathbb{R}_0} \psi(s, z) D_{s,z} F \nu(dz) ds\right], \tag{4.2}$$

valid for all Malliavin differentiable F and $\mathcal{F}_t$-predictable processes φ and ψ such that the integrals on the right converge absolutely. We also need the following basic properties of Malliavin derivatives:

If $F \in \mathbb{D}_{1,2}$ is $\mathcal{F}_s$- measurable, then

$$D_t F = D_{t,z} F = 0 \text{ for all } t > s \tag{4.3}$$

(Fundamental theorems)

16

$$D_t \left(\int_0^T \varphi(s)dB(s) \right) = \int_0^T D_t\varphi(s)dB(s) + \varphi(t) \tag{4.4}$$

$$D_{t,z} \left(\int_0^T \int_{\mathbb{R}_0} \psi(s,y)\tilde{N}(ds,dy) \right) = \int_0^T \int_{\mathbb{R}_0} D_{t,z}\psi(s,y)\tilde{N}(ds,dy) + \psi(t,z), \tag{4.5}$$

provided that all terms involved are well-defined. We refer to [DØP] for more information about the Malliavin calculus for Lévy processes and its applications.

We now define *modified adjoint processes* $\tilde{p}(t)$, $\tilde{q}(t)$, $\tilde{r}(t,z)$, $\tilde{\lambda}(t)$ as follows:

$$\tilde{p}(t) = \kappa(t) + \int_t^T \frac{\partial H_0}{\partial a}(s)G(t,s)ds \tag{4.6}$$

$$\tilde{q}(t) = D_t\tilde{p}(t) \tag{4.7}$$

$$\tilde{r}(t,z) = D_{t,z}\tilde{p}(t), \tag{4.8}$$

with

$$\kappa(t) = h_2'(A(T)) + c\tilde{\lambda}(T) + \int_t^T \frac{\partial f}{\partial a}(s)ds \tag{4.9}$$

$$H_0(s,a,x,u) = \kappa(s)b(s,a,u) + D_s\kappa(s)\sigma(s,a,u)$$
$$+ \int_{\mathbb{R}_0} D_{s,z}\kappa(s)\gamma(s,a,u,z)\nu(dz) + g(s,a,x,u)\tilde{\lambda}(s) \tag{4.10}$$

Above and in the following, We use the shorthand notation $H_0(s) = H_0(s,A(s),X(s),u(s))$.

The process $\tilde{\lambda}(t)$ is given by the similar (forward) equation as in (2.6) but with p,q,r replaced by $\tilde{p},\tilde{q},\tilde{r}$:

$$\begin{cases} d\tilde{\lambda}(t) &= \dfrac{\partial H}{\partial x}(t,A(t),X(t),Y(t),K(t,\cdot),u(t),\tilde{\lambda}(t),\tilde{p}(t),\tilde{q}(t),\tilde{r}(t,z))dt \\ &+ \dfrac{\partial H}{\partial y}(t,A(t),X(t),Y(t),K(t,\cdot),u(t),\tilde{\lambda}(t),\tilde{p}(t),\tilde{q}(t),\tilde{r}(t,z))dB(t) \\ &+ \displaystyle\int_{\mathbb{R}_0} \nabla_k H(t,A(t),X(t),Y(t),K(t,\cdot),u(t),\tilde{\lambda}(t),\tilde{p}(t),\tilde{q}(t),\tilde{r}(t,z),z)\tilde{N}(dt,dz) \; ; \; t \in [0,T] \\ \tilde{\lambda}(0) &= h_1'(X(0)), \end{cases} \tag{4.11}$$

with the Hamiltonian H defined in (2.5).

We can now state and prove the main result of this section. We first introduce the notation:

$$G(t,s) = \exp\left(\int_t^s \left\{ \frac{\partial b}{\partial a}(r) - \frac{1}{2}\left(\frac{\partial \sigma}{\partial a}(r)\right)^2 \right\}dr + \int_t^s \frac{\partial \sigma}{\partial a}(r)dB(r) \right.$$
$$+ \int_t^s \int_{\mathbb{R}_0} \ln\left(1 + \frac{\partial \gamma}{\partial a}(r,z)\right)\tilde{N}(dr,dz)$$
$$\left. + \int_t^s \int_{\mathbb{R}_0} \left[\ln\left(1 + \frac{\partial \gamma}{\partial a}(r,z)\right) - \frac{\partial \gamma}{\partial a}(r,z)\right]\nu(dz)dr \right) \; ; \; s > t. \tag{4.12}$$

$$F(T) := h_2'(A(T)) + c\tilde{\lambda}(T) \tag{4.13}$$

$$\Phi(t,s) := \frac{\partial H_0}{\partial a}(s)G(t,s) \tag{4.14}$$

Theorem 4.1 *Let $u \in \mathcal{A}_\varepsilon$ with corresponding solutions $A(t), X(t), Y(t), K(t,z)$ and $\tilde{\lambda}(t)$ of (2.1), (2.2) and (4.11). Assume that the random variables $F(T)$, $\frac{\partial f}{\partial a}(t)$ and $\Phi(t,s)$ belong to $\mathbb{D}_{1,2}$ for all $0 \le t \le s \le T$ and that*

$$E\left[\int_0^T \left\{ (\frac{\partial \sigma}{\partial a}(s))^2 \alpha^2(s) + (\frac{\partial \sigma}{\partial u}(s))^2 + \int_{\mathbb{R}_0} \left\{ (\frac{\partial \gamma}{\partial a}(s,z))^2 \alpha^2(s) + (\frac{\partial \gamma}{\partial u}(s,z))^2 \right\} \nu(dz) \right\} ds \right] < \infty$$

$$E\left[\int_0^T \int_0^T \left\{ (D_s(\frac{\partial f}{\partial a}(t)))^2 + \int_{\mathbb{R}_0} (D_{s,z}(\frac{\partial f}{\partial a}(t)))^2 \nu(dz) \right\} ds\, dt \right] < \infty$$

$$E\left[\int_0^T \int_0^T \left\{ (D_r \Phi(t,s))^2 + \int_{\mathbb{R}_0} (D_{r,z}\Phi(t,s))^2 \nu(dz) \right\} dr\, ds \right] < \infty. \tag{4.15}$$

Then the following are equivalent:

(i) $\dfrac{d}{dy} J(u + y\beta)\,|_{y=0} = 0$ *for all bounded $\beta \in \mathcal{A}_\mathcal{E}$*

(ii) $E\left[\dfrac{\partial}{\partial u} H(t, A(t), X(t), Y(t), K(t, \cdot), u, \tilde{\lambda}(t), \tilde{p}(t), \tilde{q}(t), \tilde{r}(t,z))_{u=u(t)} \mid \mathcal{E}_t \right] = 0$
for a.a. $(t, \omega) \in [0, T] \times \Omega$.

Proof.

(i) $\Rightarrow$ (ii): Assume that (i) holds. Then, as in (3.3),

$$0 = \frac{d}{dy} J(u + y\beta)\,|_{y=0}$$

$$= E\left[\int_0^T \left\{ \frac{\partial f}{\partial a}(t)\alpha(t) + \frac{\partial f}{\partial x}(t)\xi(t) + \frac{\partial f}{\partial y}(t)\eta(t) + \int_{\mathbb{R}_0} \nabla_k f(t,z)\zeta(t,z)\nu(dz) \right. \right.$$

$$\left. \left. + \frac{\partial f}{\partial u}(t)\beta(t) \right\} dt + h_1'(X(0))\xi(0) + (h_2'(A(T)) + c\tilde{\lambda}(T) - c\tilde{\lambda}(T))\alpha(T) \right]. \tag{4.16}$$

By (3.1) and the duality formulae (4.1), (4.2) we have, with $F(T)$ defined in (4.13),

$$E[F(T)\alpha(T)] = E\left[F(T)\left(\int_0^T \left\{ \frac{\partial b}{\partial a}(t)\alpha(t) + \frac{\partial b}{\partial u}(t)\beta(t) \right\} dt \right. \right.$$

$$\left. + \int_0^T \left\{ \frac{\partial \sigma}{\partial a}(t)\alpha(t) + \frac{\partial \sigma}{\partial u}(t)\beta(t) \right\} dB(t) + \int_{\mathbb{R}_0} \left\{ \frac{\partial \gamma}{\partial a}(t,z)\alpha(t) + \frac{\partial \gamma}{\partial u}(t,z)\beta(t) \right\} \tilde{N}(dt,dz) \right)\right]$$

$$= E\left[\int_0^T \left\{ F(T)\left[\frac{\partial b}{\partial a}(t) + \frac{\partial b}{\partial u}(t)\beta(t) \right] + D_t F(T)\left[\frac{\partial \sigma}{\partial a}(t)\alpha(t) + \frac{\partial \sigma}{\partial u}(t)\beta(t) \right] \right. \right.$$

$$\left. \left. + \int_{\mathbb{R}_0} D_{t,z}F(T)\left[\frac{\partial \gamma}{\partial a}(t,z)\alpha(t) + \frac{\partial \gamma}{\partial u}(t,z)\beta(t) \right] \nu(dz) \right\} dt \right]. \tag{4.17}$$

Similarly we have, using the Fubini theorem,

$$E\left[\int_0^T \frac{\partial f}{\partial a}(t)\alpha(t)dt\right] =$$

$$= E\left[\int_0^T \frac{\partial f}{\partial a}(t)\left(\int_0^t \left\{\frac{\partial b}{\partial a}(s)\alpha(s) + \frac{\partial b}{\partial u}(s)\beta(s)\right\}ds\right.\right.$$

$$+ \int_0^t \left\{\frac{\partial \sigma}{\partial a}(s)\alpha(s) + \frac{\partial \sigma}{\partial u}(s)\beta(s)\right\}dB(s)$$

$$\left.\left.+ \int_0^t \int_{\mathbb{R}_0}\left\{\frac{\partial \gamma}{\partial a}(s,t)\alpha(s) + \frac{\partial \gamma}{\partial u}(s,t)\beta(s)\right\}\tilde{N}(ds,dz)\right)\right]$$

$$= E\left[\int_0^T \left(\int_0^t \left\{\frac{\partial f}{\partial a}(t)\left[\frac{\partial b}{\partial a}(s)\alpha's) + \frac{\partial b}{\partial u}(s)\beta(s)\right]\right.\right.$$

$$+ D_s\left(\frac{\partial f}{\partial a}(t)\right)\left[\frac{\partial \sigma}{\partial a}(s)\alpha(s) + \frac{\partial \sigma}{\partial u}(s)\beta(s)\right]$$

$$\left.\left.+ \int_{\mathbb{R}_0} D_{s,z}\left(\frac{\partial f}{\partial a}(t)\right)\left[\frac{\partial \gamma}{\partial a}(s,z)\alpha(s) + \frac{\partial \gamma}{\partial u}(s,z)\beta(s)\right]\nu(dz)\right\}ds\right)dt\right]$$

$$= E\left[\int_0^T \left\{\left(\int_s^T \frac{\partial f}{\partial a}(t)dt\right)\left[\frac{\partial b}{\partial a}(s)\alpha(s) + \frac{\partial b}{\partial u}(s)\beta(s)\right]\right.\right.$$

$$+ \left(\int_s^T D_s\left(\frac{\partial f}{\partial a}(t)\right)dt\right)\left[\frac{\partial \sigma}{\partial a}(s)\alpha(s) + \frac{\partial \sigma}{\partial u}(s)\beta(s)\right]$$

$$\left.\left.+ \int_{\mathbb{R}_0}\left(\int_s^T D_{s,z}\left(\frac{\partial f}{\partial a}(t)\right)dt\right)\left[\frac{\partial \gamma}{\partial a}(s,z)\alpha(s) + \frac{\partial \gamma}{\partial u}(s,z)\beta(s)\right]\nu(dz)\right\}ds\right].$$

Changing the notation $s \leftrightarrow t$ this becomes

$$= E\left[\int_0^T \left\{\left(\int_t^T \frac{\partial f}{\partial a}(s)ds\right)\left[\frac{\partial b}{\partial a}(t)\alpha(t) + \frac{\partial b}{\partial u}(t)\beta(t)\right]\right.\right.$$

$$+ \left(\int_t^T D_t\left(\frac{\partial f}{\partial a}(s)\right)ds\right)\left[\frac{\partial \sigma}{\partial a}(t)\alpha(t) + \frac{\partial \sigma}{\partial u}(t)\beta(t)\right]$$

$$\left.\left.+ \int_{\mathbb{R}_0}\left(\int_t^T \left(\frac{\partial f}{\partial a}(s)\right)ds\right)\left[\frac{\partial \gamma}{\partial a}(t,z)\alpha(t) + \frac{\partial \gamma}{\partial u}(t,z)\beta(t)\right]\nu(dz)\right\}dt\right]. \qquad (4.18)$$

Combining (4.17) and (4.18), and using (4.9) we get

$$E\left[\int_0^T \left\{\frac{\partial f}{\partial a}(t)\alpha(t) + \frac{\partial f}{\partial u}(t)\beta(t)\right\} dt + h_2'(A(T))\alpha(T)\right]$$

$$= E\left[\int_0^T \left\{\kappa(t)\left[\frac{\partial b}{\partial a}(t)\alpha(t) + \frac{\partial b}{\partial u}(t)\beta(t)\right]\right.\right.$$

$$+ D_t\kappa(t)\left[\frac{\partial \sigma}{\partial a}(t)\alpha(t) + \frac{\partial \sigma}{\partial u}(t)\beta(t)\right]$$

$$+ \int_{\mathbb{R}_0} D_{t,z}\kappa(t)\left[\frac{\partial \gamma}{\partial a}(t,z)\alpha(t) + \frac{\partial \gamma}{\partial u}(t,z)\beta(t)\right] + \frac{\partial f}{\partial u}(t)\beta(t)\right\}\right]$$

$$- E[\tilde{\lambda}(T)\xi(T)], \text{ using that } c\alpha(T) = \xi(T). \tag{4.19}$$

Then by the Itô formula and (4.11),

$$E[h_1'(X(0))\xi(0)] = E[\tilde{\lambda}(0)\xi(0)]$$

$$= E\left[\tilde{\lambda}(T)\xi(T) - \int_0^T \tilde{\lambda}(t)d\xi(t) - \int_0^T \xi(t)d\tilde{\lambda}(t)\right.$$

$$- \int_0^T \frac{\partial H}{\partial y}(t)\eta(t)dt - \int_0^T \int_{\mathbb{R}_0} \nabla_k H(t,z)\zeta(t,z)\nu(dz)dt\right]$$

$$= E\left[\tilde{\lambda}(T)\xi(T) - \int_0^T \tilde{\lambda}(t)\left\{-\frac{\partial g}{\partial a}(t)\alpha(t) - \frac{\partial g}{\partial x}(t)\xi(t)\right.\right.$$

$$\left.-\frac{\partial g}{\partial y}(t)\eta(t) - \frac{\partial g}{\partial u}(t)\beta(t)\right\} dt - \int_0^T \xi(t)\frac{\partial H}{\partial x}(t)dt$$

$$\left.- \int_0^T \eta(t)\frac{\partial H}{\partial y}(t)dt - \int_0^T \int_{\mathbb{R}_0} \nabla_k H(t,z)\zeta(t,z)\nu(dz)dt\right].$$

Now

$$\frac{\partial H}{\partial x}(t) = \frac{\partial f}{\partial x}(t) + \frac{\partial g}{\partial x}(t)\tilde{\lambda}$$

$$\frac{\partial H}{\partial y}(t) = \frac{\partial f}{\partial y}(t) + \frac{\partial g}{\partial y}(t)\tilde{\lambda}$$

$$\nabla_k H(t,z) = \nabla_k f(t,z).$$

Hence the above simplifies to

$$E[h_1'(X(0))\xi(0)] = E[\tilde{\lambda}(T)\xi(T)$$

$$+ \int_0^T \left\{\tilde{\lambda}(t)\left[\frac{\partial g}{\partial a}(t)\alpha(t) + \frac{\partial g}{\partial u}(t)\beta(t)\right] - \frac{\partial f}{\partial x}(t)\xi(t) - \frac{\partial f}{\partial y}(t)\eta(t)\right.$$

$$\left.- \int_{\mathbb{R}_0} \nabla_k f(t,z)\zeta(t,z)\nu(dz)\right\} dt\right]. \tag{4.20}$$

20

Combining (4.16), (4.19) and (4.20) we get

$$0 = \frac{d}{dy} J(u + y\beta) \mid_{y=0}$$

$$= E\left[\int_0^T \left\{ \kappa(t) \left[\frac{\partial f}{\partial a}(t)\alpha(t) + \frac{\partial f}{\partial u}(t)\beta(t) \right] + D_t \kappa(t) \left[\frac{\partial \sigma}{\partial a}(t)\alpha(t) + \frac{\partial \sigma}{\partial u}(t)\beta(t) \right] \right.\right.$$

$$+ \int_{\mathbb{R}_0} D_{t,z}\kappa(t) \left[\frac{\partial \gamma}{\partial a}(t,z)\alpha(t) + \frac{\partial \gamma}{\partial u}(t,z)\beta(t) \right] \nu(dz) + \frac{\partial f}{\partial u}\beta(t)$$

$$\left. + \tilde{\lambda}(t) \left[\frac{\partial g}{\partial a}(t)\alpha(t) + \frac{\partial g}{\partial u}(t)\beta(t) \right] \right\} dt \right]$$

$$= E\left[\int_0^T \left\{ \left[\kappa(t)\frac{\partial b}{\partial a}(t) + D_t \kappa(t)\frac{\partial \sigma}{\partial a}(t) + \int_{\mathbb{R}_0} D_{t,z}\kappa(t)\frac{\partial \gamma}{\partial a}(t,z)\nu(dz) + \tilde{\lambda}(t)\frac{\partial g}{\partial a}(t) \right] \alpha(t) \right.\right.$$

$$+ \left[\kappa(t)\frac{\partial b}{\partial u}(t) + D_t \kappa(t)\frac{\partial \sigma}{\partial u}(t) + \int_{\mathbb{R}_0} D_{t,z}\kappa(t)\frac{\partial \gamma}{\partial u}(t,z)\nu(dz) \right.$$

$$\left.\left.\left. + \frac{\partial f}{\partial u}(t) + \tilde{\lambda}(t)\frac{\partial g}{\partial u}(t) \right] \beta(t) \right\} dt \right. \tag{4.21}$$

This holds for all $\beta \in \mathcal{A}_\mathcal{E}$. In particular, if we apply this to

$$\beta_\theta = \beta_\theta(s) = \theta(\omega)\chi_{(t,t+h]}(s)$$

where $\theta(\omega)$ is $\mathcal{E}_t$-measurable and $0 \le t \le t + h \le T$ we get, by (3.1),

$$\alpha = \alpha^{(\beta_\theta)}(s) = 0 \text{ for } 0 \le s \le t$$

and (4.21) can be written

$$L_1(h) + L_2(h) = 0, \tag{4.22}$$

where

$$L_1(h) = E\left[\int_t^T \left\{ \kappa(s)\frac{\partial b}{\partial a}(s) + D_s\kappa(s)\frac{\partial \sigma}{\partial a}(s) + \int_{\mathbb{R}_0} D_{s,z}\kappa(s)\frac{\partial \gamma}{\partial a}(s,z)\nu(dz) \right.\right.$$

$$\left.\left. + \tilde{\lambda}(s)\frac{\partial g}{\partial a}(s) \right\} \alpha(s)ds \right] \tag{4.23}$$

and

$$L_2(h) = E\left[\theta \int_t^{t+h} \left\{ \kappa(s)\frac{\partial b}{\partial u}(s) + D_s\kappa(s)\frac{\partial \sigma}{\partial u}(s) + \int_{\mathbb{R}_0} D_{s,z}\kappa(s)\frac{\partial \gamma}{\partial u}(s,z)\nu(dz) \right.\right.$$

$$\left.\left. + \frac{\partial f}{\partial u}(s) + \tilde{\lambda}(s)\frac{\partial g}{\partial u}(s) \right\} ds \right] \tag{4.24}$$

Note that with $\alpha(s) = \alpha^{(\beta_\theta)}(s)$ we have, for $s \ge t + h$,

$$d\alpha(s) = \alpha(s^-) \left\{ \frac{\partial b}{\partial a}(s)ds + \frac{\partial \sigma}{\partial a}(s)dB(s) + \int_{\mathbb{R}_0} \frac{\partial \gamma}{\partial a}(s,z)\tilde{N}(ds,dz) \right\} \tag{4.25}$$

21

Hence, by the Itô formula

$$\alpha(s) = \alpha(t+h)G(t+h,s) \; ; \; s \geq t+h \tag{4.26}$$

where G is defined in (4.12). Note that $G(t,s)$ does not depend on h. Then

$$L_1 h = E\left[\int_t^T \frac{\partial H_0}{\partial a}(s)\alpha(s)ds\right]. \tag{4.27}$$

where H_0 is defined in (4.10). Differentiating with respect to h at $h=0$ gives

$$L_1'(0) = \frac{d}{dh} E\left[\int_t^{t+h} \frac{\partial H_0}{\partial a}(s)\alpha(s)ds\right]_{h=0} + \frac{d}{dh} E\left[\int_{t+h}^T \frac{\partial H_0}{\partial a}(s)\alpha(s)ds\right]_{h=0}. \tag{4.28}$$

Since $\alpha(t) = 0$ we see that

$$\frac{d}{dh} E\left[\int_t^{t+h} \frac{\partial H_0}{\partial a}(s)\alpha(s)ds\right]_{h=0} = 0. \tag{4.29}$$

Therefore, by (4.26)

$$\begin{aligned}
L_1'(0) &= \frac{d}{dh} E\left[\int_{t+h}^T \frac{\partial H_0}{\partial a}(s)\alpha(t+h)G(t+h,s)ds\right]_{h=0} \\
&= \int_t^T \frac{d}{dh} E\left[\frac{\partial H_0}{\partial a}(s)\alpha(t+h)G(t+h,s)ds\right]_{h=0} ds \\
&= \int_t^T \frac{d}{dh} E\left[\frac{\partial H_0}{\partial a}(s)G(t,s)\alpha(t+h)\right]_{h=0} ds.
\end{aligned} \tag{4.30}$$

By (3.1) we have

$$\begin{aligned}
\alpha(t+h) &= \theta \int_t^{t+h} \left\{\frac{\partial b}{\partial u}(r)dr + \frac{\partial \sigma}{\partial u}(r)dB(r) + \int_{\mathbb{R}_0} \frac{\partial \gamma}{\partial u}(r,z)\tilde{N}(dr,dz)\right\} \\
&\quad + \int_t^{t+h} \alpha(r^-) \left\{\frac{\partial b}{\partial a}(r)dr + \frac{\partial \sigma}{\partial a}(r)dB(r) + \int_{\mathbb{R}_0} \frac{\partial \gamma}{\partial a}(r,z)\tilde{N}(dr,dz)\right\}. \tag{4.31}
\end{aligned}$$

Therefore, by (4.30) and (4.31)

$$L_1'(0) = \Gamma_1 + \Gamma_2, \tag{4.32}$$

where

$$\begin{aligned}
\Gamma_1 &= \int_t^T \frac{d}{dh} E\left[\frac{\partial H_0}{\partial a}(s)G(t,s)\theta \int_t^{t+h} \left\{\frac{\partial b}{\partial u}(r)dr + \frac{\partial \sigma}{\partial u}(r)dB(r)\right.\right. \\
&\quad \left.\left. + \int_{\mathbb{R}_0} \frac{\partial \gamma}{\partial u}(r,z)\tilde{N}(dr,dz)\right\}\right]_{h=0} ds \tag{4.33}
\end{aligned}$$

22

and

$$\Gamma_2 = \int_t^{t+h} \frac{d}{dh} E\left[\frac{\partial H_0}{\partial a}(s)G(t,s)\int_t^{t+h}\alpha(r^-)\left\{\frac{\partial b}{\partial a}(r)dr + \frac{\partial \sigma}{\partial a}(r)dB(r)\right.\right.$$
$$\left.\left. + \int_{\mathbb{R}_0}\frac{\partial \gamma}{\partial a}(r,z)\tilde{N}(dr,dz)\right\}\right]_{h=0} ds. \tag{4.34}$$

By the duality formulae (4.1), (4.2) we have, using (4.14)

$$\Gamma_1 = \int_t^T \frac{d}{dh}E\left[\theta\int_t^{t+h}\left\{\frac{\partial b}{\partial u}(r)\Phi(t,s) + \frac{\partial \sigma}{\partial u}(r)D_r\Phi(t,s)\right.\right.$$
$$\left.\left. + \int_{\mathbb{R}_0}\frac{\partial \gamma}{\partial u}(r,z)D_{r,z}\Phi(t,s)\nu(dz)\right\}dr\right]_{h=0} ds$$
$$= \int_t^T E\left[\theta\left\{\frac{\partial b}{\partial u}(t)\Phi(t,s) + \frac{\partial \sigma}{\partial u}(t)D_t\Phi(t,s)\right.\right.$$
$$\left.\left. + \int_{\mathbb{R}_0}\frac{\partial \gamma}{\partial u}(t,z)D_{t,z}\Phi(t,s)\nu(dz)\right\}\right] ds, \tag{4.35}$$

Since $\alpha(t) = 0$ we see that

$$\Gamma_2 = 0. \tag{4.36}$$

We conclude from (4.32)-(4.36) that

$$L_1'(0) = \Gamma_1. \tag{4.37}$$

Moreover, we see directly that

$$L_2'(0) = E\left[\theta\left\{\kappa(t)\frac{\partial b}{\partial u}(t) + D_t\kappa(t)\frac{\partial \sigma}{\partial u}(t)\right.\right.$$
$$\left.\left. + \int_{\mathbb{R}_0}D_{t,z}\kappa(t)\frac{\partial \gamma}{\partial u}(t,z)\nu(dz) + \frac{\partial f}{\partial u}(t) + \tilde{\lambda}(t)\frac{\partial g}{\partial u}(t)\right\}\right]. \tag{4.38}$$

By differentiating (4.22) with respect to h at $h = 0$ we thus obtain the equation

$$E\left[\theta\left\{\left(\kappa(t) + \int_t^T \Phi(t,s)ds\right)\frac{\partial b}{\partial u}(t) + D_t\left(\kappa(t) + \int_t^T \Phi(t,s)ds\right)\frac{\partial \sigma}{\partial u}(t)\right.\right.$$
$$\left.\left. + \int_{\mathbb{R}_0}D_{t,z}\left(\kappa(t) + \int_t^T \Phi(t,s)ds\right)\frac{\partial \gamma}{\partial u}(t,z)\nu(dz) + \frac{\partial f}{\partial u}(t) + \tilde{\lambda}(t)\frac{\partial g}{\partial u}(t)\right\}\right] = 0. \tag{4.39}$$

Using (4.6), equation (4.39) can be written

$$E\left[\theta\frac{\partial}{\partial u}\left\{f(t,A(t),X(t),Y(t),K(t,\cdot),u)\right.\right.$$
$$+ \tilde{p}(t)b(t,A(t),u) + \tilde{\lambda}(t)g(t,A(t),X(t),Y(t),u)$$
$$\left.\left. + D_t\tilde{p}(t)\sigma(t,A(t),u) + \int_{\mathbb{R}_0}D_{t,z}\tilde{p}(t)\gamma(t,A(t),u,z)\nu(dz)\right\}_{u=u(t)}\right] = 0. \tag{4.40}$$

Since this holds for all $\mathcal{E}_t$-measurable θ we conclude that

$$E\left[\frac{\partial}{\partial u}H(t, A(t), X(t), Y(t), K(t, \cdot), u, \tilde{p}(t), \tilde{q}(t), \tilde{r}(t, z), \tilde{\lambda}(t))_{u=u(t)} \mid \mathcal{E}_t\right] = 0. \qquad (4.41)$$

(ii) $\Rightarrow$ **(i)**: Conversely, suppose (4.41) holds for some $u \in \mathcal{A}_{\mathcal{E}}$. Then we can reverse the argument to get that (4.22) holds for all $\beta = \beta_\theta$. Then (4.22) holds for all linear combinations of such β_θ. Since all bounded $\beta \in \mathcal{A}_{\mathcal{E}}$ can be approximated by such linear combinations, it follows that (4.22) holds for all bounded $\beta \in \mathcal{A}_{\mathcal{E}}$. Hence, by reversing the remaning part of the argument above, we conclude that (ii) $\Rightarrow$ (i). $\qquad \square$

5 Applications in Finance

5.1 Risk minimizing portfolios

To illustrate our results, we now apply them to study the risk minimizing portfolio problem stated in the introduction. Note that the Malliavin calculus approach in Section 4 does not require that the system is Markovian. So we assume that the wealth process $A(t) = A_u(t)$ corresponding to the portfolio u is as in (1.12), i.e.

$$\begin{cases} dA(t) = u(t^-)\left[\alpha(t)dt + \beta(t)dB(t) + \displaystyle\int_{\mathbb{R}_0} \theta(t, z)\tilde{N}(dt, dz)\right] \\ A(0) = a > 0 \end{cases} \qquad (5.1)$$

where α, β and θ are given predictable processes. Here $u(t) = \pi(t)A(t^-)$ is the *amount* invested in the risky asset at time t.

The corresponding BSDE for $X(t) = X_u(t)$, $Y(t) = Y_u(t)$ and $K(t, z) = K_u(t, z)$ in (1.2) becomes

$$\begin{cases} dX(t) = -g(t, X(t), \omega)dt + Y(t)dB(t) + \displaystyle\int_{\mathbb{R}_0} K(t, z)\tilde{N}(dt, dz) \\ X(T) = -A_u(T) \end{cases} \qquad (5.2)$$

where $g : [0, T] \times \mathbb{R} \times \Omega \to \mathbb{R}$ is a given function, such that $(t, \omega) \to g(t, x, \omega)$ is $\mathcal{F}_t$-predictable for each given x.

The performance functional $J(u)$ in (2.3) simplifies to

$$J(u) = X_u(0),$$

i.e. we have

$$f = 0, \quad h_1(x) = x \text{ and } h_2 = 0. \qquad (5.3)$$

Moreover, here we have

$$c = -1$$
$$b(t, a, u, \omega) = u\alpha(t, \omega) = u\alpha(t)$$
$$\sigma(t, a, u, \omega) = u\beta(t, \omega) = u\beta(t)$$
$$\gamma(t, a, u, z, \omega) = u\theta(t, z, \omega) = u\theta(t, z), \text{ for short.}$$

24

Therefore, the Hamiltonian becomes (dropping the ω in the notation for simplicity)

$$H(t, a, x, y, k, u, p, q, r(.), \lambda, \omega)$$
$$= \lambda g(t, x) + u\alpha(t)p + u\beta(t)q + \int_{\mathbb{R}_0} u\theta(t, z)r(z)\nu(dz), \qquad (5.4)$$

The modified adjoint processes are given by:

$$\tilde{p}(t) = \kappa(t) + \int_t^T \frac{\partial H_0}{\partial a}(s)G(t, s)ds = -\tilde{\lambda}(T) \qquad (5.5)$$

$$\tilde{q}(t) = -D_t\tilde{\lambda}(T) \qquad (5.6)$$

$$\tilde{r}(t, z) = -D_{t,z}\tilde{\lambda}(T) \; ; \; 0 \le t \le T. \qquad (5.7)$$

$\tilde{\lambda}(t)$ is given by (4.11), i.e.

$$\begin{cases} d\tilde{\lambda}(t) &= \tilde{\lambda}(t)g'(t, X(t))dt \\ \tilde{\lambda}(0) &= 1, \end{cases} \qquad (5.8)$$

where $g'(t, x) = \dfrac{\partial}{\partial x}g(t, x)$. The equation (5.8) has the solution

$$\tilde{\lambda}(t) = \exp\left(\int_0^t g'(s, X(s))ds\right) \; ; \; 0 \le t \le T. \qquad (5.9)$$

The condition (ii) in Theorem 4.1 for an optimal control $\hat{u}(t)$ is

$$E\left[\alpha(t)\tilde{\lambda}(T) + \beta(t)D_t\tilde{\lambda}(T) + \int_{\mathbb{R}_0} \theta(t, z)D_{t,z}\tilde{\lambda}(T)\nu(dz) \mid \mathcal{E}_t\right] = 0. \qquad (5.10)$$

Equation (5.10) is a (linear, homogeneous) *partial information, Malliavin-differential type equation* in the unknown random variable $\tilde{\lambda}(T)$. In the Appendix, we solve this equation and then extend the solution to the in-homogeneous equation. An equation of this type was also encountered and solved in [MØZ] in the case when $\mathcal{E}_t = \mathcal{F}_t$ and $\nu = 0$.

From now on, we assume that

$$\mathcal{E}_t = \mathcal{F}_t \quad \text{for all } t \in [0, T] \qquad (5.11)$$

By Theorem A.1 we get,

$$\tilde{\lambda}(T) = E[\tilde{\lambda}(T)] \exp\left(\int_0^T \sigma(s)dB(s) - \frac{1}{2}\int_0^T \sigma^2(s)ds\right.$$
$$+ \int_0^T \int_{\mathbb{R}_0} \ln(1 + \gamma(s, z))\tilde{N}(ds, dz)$$
$$\left. + \int_0^T \int_{\mathbb{R}_0} \{\ln(1 + \gamma(s, z)) - \gamma(s, z)\}\nu(dz)ds\right) \qquad (5.12)$$

25

for some $\mathcal{F}_t$-predictable processes $\sigma(t)$ and $\gamma(t, z)$ such that

$$\alpha(t) + \beta(t)\sigma(t) + \int_{\mathbb{R}_0} \theta(t, z)\gamma(t, z)\nu(dz) = 0 \text{ for a.a. } t, \omega. \tag{5.13}$$

Remark 5.1 Note that condition (5.13) *is saying that the measure Q defined by*

$$dQ(\omega) = \frac{\tilde{\lambda}(T)}{E[\tilde{\lambda}(T)]} dP(\omega) \text{ on } \mathcal{F}_T \tag{5.14}$$

is an equivalent local martingale measure (ELMM) for the process $A(t)$ given by (5.1). *(See Øksendal and A. Sulem [ØS2, Theorem 1.31]).*
In particular, if σ and γ satisfy the Novikov condition

$$E\left[\exp\left(\frac{1}{2}\int_0^t \left\{\sigma^2(s) + \int_{\mathbb{R}_0} [(1 - \gamma(s, z))\ln(1 - \gamma(s, z)) + \gamma(s, z)]\nu(dz)\right\} ds\right)\right] < \infty \tag{5.15}$$

then Q is an equivalent martingale measure (EMM) (see [KS, p. 408]).

We now study some special cases in more detail.

Lemma 5.2 *Suppose the driver $g(t, x, \omega)$ satisfies the equation*

$$mxg'(t, x) - g(t, x) + c_1(t)x + c_0(t) = 0 \tag{5.16}$$

for some constant $m > 0$ and some bounded predictable processes $c_1(t), c_0(t)$. Define

$$\lambda(t) = \lambda_u(t) = \exp\left(\int_0^t g'(s, X(s))ds\right), \tag{5.17}$$

where $X(s) = X_u(s)$ is as in (5.2). *Then*

$$X(0) = -E\left[\lambda(T)^m \exp\left(\int_0^T c_1(s)ds\right) A(T)\right.$$
$$\left. - \int_0^T \lambda(t)^m \exp\left(\int_0^t c_1(s)ds\right) c_0(t)dt\right]. \tag{5.18}$$

Proof.   Assume that $g(t, x)$ has the form

$$g(t, x) = h(t, x)x + c_0(t) \tag{5.19}$$

for some predictable processes $h(t, x), c_0(t)$. Then the equation for $X(t)$ becomes

$$dX(t) = -h(t, X(t))X(t)dt - c_0(t)dt + dM(t),$$

where
$$dM(t) = Y(t)dB(t) + \int_{\mathbb{R}_0} K(t,z)\tilde{N}(dt,dz).$$

This can be written
$$d(J(t)X(t)) = J(t)[-c_0(t)dt + dM(t)],$$

where
$$J(t) = \exp\left(\int_0^t h(s,X(s))ds\right). \tag{5.20}$$

Hence,
$$J(T)X(T) = X(0) + \int_0^T J(s)[-c_0(s)ds + dM(s)]. \tag{5.21}$$

In particular, if
$$h(t,x) = mg'(t,x) + c_1(t), \tag{5.22}$$

where $m > 0$ is a constant and $c_1(t)$ is a predictable process, then
$$J(t) = \exp\left(m\int_0^t g'(s,X(s))ds + \int_0^t c_1(s)ds\right)$$
$$= \lambda(t)^m \exp\left(\int_0^t c_1(s)ds\right). \tag{5.23}$$

Therefore, by (5.21),
$$\lambda(T)^m \exp\left(\int_0^T c_1(s)ds\right) X(T) =$$
$$= X(0) - \int_0^T \lambda(t)^m \exp\left(\int_0^t c_1(s)ds\right) c_0(s)ds + \int_0^T J(s)dM(s).$$

Taking expectation and noting that
$$X(T) = -A(T) = -A_u(T),$$

we obtain
$$X(0) = E\left[-\lambda(T)^m \exp\left(\int_0^T c_1(s)ds\right) A(T)\right.$$
$$\left. + \int_0^T \lambda(t)^m \exp\left(\int_0^t c_1(s)ds\right) c_0(s)ds\right], \tag{5.24}$$

as claimed. $\qquad\square$

Lemma 5.3 *The general predictable solution $g(t,x,\omega)$ of equation (5.16) is as follows:*

(i) *If $m = 1$, then*

$$g(t, x) = -c_1(t)x \ln x + c_0(t) + c(t)x, \tag{5.25}$$

where $c(t)$ is an arbitrary predictable process.

(ii) *If $m \neq 1$, then*

$$g(t, x) = -\frac{c_1(t)}{m - 1}x + c_0(t) + c(t)x^{\frac{1}{m}}, \tag{5.26}$$

where $c(t)$ is an arbitrary predictable process.

Proof.   The proof is straightforward and hence omitted.   □

In particular, if we choose $m = 1$ and $c_0(t) = 0$ in (5.16) we get the driver

$$g(t, x) = -c_1(t)x \ln x + c(t)x. \tag{5.27}$$

This gives the following result:

Theorem 5.4 *Suppose $g(t, x)$ is as in (5.27) and that $c_1(t)$ is deterministic and $\mathcal{E}_t = \mathcal{F}_t$. Let $\tilde{\lambda}(T)$ be the solution of (5.10) as given by (5.12) and suppose σ, γ satisfies (5.15). Then the minimal risk $\hat{X}(0)$ satisfies the equation*

$$\hat{X}(0) + a \exp\left(\int_0^T c_1(s)ds\right) E[\tilde{\lambda}(T)] = 0. \tag{5.28}$$

Proof.   By Lemma 5.2

$$\hat{X}(0) = -E\left[\tilde{\lambda}(T) \exp\left(\int_0^T c_1(s)ds\right) \hat{A}(T),\right]$$

$$= -\exp\left(\int_0^T c_1(s)ds\right) E[\tilde{\lambda}(T)\hat{A}(T)].$$

By Remark 5.1 it follows that the measure Q defined by

$$dQ(\omega) := \frac{\tilde{\lambda}(T)}{E[\tilde{\lambda}(T)]}dP(\omega) \text{ on } \mathcal{F}_T \tag{5.29}$$

is an EMM for $A(t)$. Hence $E_Q[\hat{A}(T)] = a$ and

$$\hat{X}(0) = -a \exp\left(\int_0^T c_1(s)ds\right) E[\tilde{\lambda}(T)].$$

□

Remark 5.5 The remaining problem is that we only know $\tilde{\lambda}(T)$ up to a multiplicative constant (see Theorem A.1).

As a second example, suppose we choose $m = 1$ and $c_1(t) = 0$ in (5.16). Then by (5.25) we get the driver

$$g(t, x) = c_0(t) + c(t)x, \tag{5.30}$$

where $c_0(t)$ and $c(t)$ are given predictable processes. In this case we are able to find explicitly the minimal risk $\hat{X}(0)$ as follows:

Theorem 5.6 *Suppose the driver $g(t, x) = g(t, x, \omega)$ is given by (5.30). Suppose $\mathcal{E}_t = \mathcal{F}_t$ and that*

$$E\left[\exp\left(\int_0^T c(s)ds\right)\left(1 + \int_0^T |c_0(s)| ds\right)\right] < \infty.$$

Suppose an optimal control $\hat{u}$ exists. Then the minimal risk is

$$\hat{X}(0) = -aE\left[\exp\left(\int_0^T c(s)ds\right)\right] + \int_0^T E\left[c_0(t)\exp\left(\int_0^t c(s)ds\right)\right] dt, \tag{5.31}$$

which is attained at $u = \hat{u} = 0$.

Proof.   In this case we get

$$\lambda_u(t) = \exp\left(\int_0^t g'(s, X_u(s))ds\right) = \exp\left(\int_0^t c(s)ds\right).$$

Hence by (5.18) we have, for $u \in \mathcal{A}$,

$$X_u(0) = -E\left[\lambda_u(T)A_u(T) - \int_0^T \lambda_u(t)c_0(t)dt\right].$$

In the optimal case $u = \hat{u}$ we have $\lambda_{\hat{u}}(T) = \tilde{\lambda}(T)$ and by Remark 5.1 and (5.29) we get

$$\begin{aligned}
\hat{X}(0) = X_{\hat{u}}(0) &= -E[\tilde{\lambda}(T)A_{\hat{u}}(T)] + E\left[\int_0^T \tilde{\lambda}(t)c_0(t)dt\right] \\
&= -E[\tilde{\lambda}(T)]E_Q[A_{\hat{u}}(T)] + E\left[\int_0^T \tilde{\lambda}(t)c_0(t)dt\right] \\
&= -aE[\tilde{\lambda}(T)] + E\left[\int_0^T \tilde{\lambda}(t)c_0(t)dt\right] \\
&= -aE\left[\exp\left(\int_0^T c(s)ds\right)\right] + \int_0^T E\left[c_0(t)\exp\left(\int_0^t c(s)ds\right)\right] dt.
\end{aligned}$$

By direct computation we see that this value is attained by using $u = 0$, which therefore is optimal. $\qquad\square$

Remark 5.7 If g is as in Theorem 5.6, then the BSDE for $X_u(t)$ is linear and we can solve for $X_u(t)$ as follows:

$$X_u(t) = E[X_u(T)\exp(\int_t^T c(s)ds) + \int_t^T \exp(\int_t^s c(r)dr)c_0(s)ds \mid \mathcal{F}_t]. \qquad (5.32)$$

In particular, if we put $t = 0$ and take expectation we get

$$X_u(0) = E[X_u(T)\exp(\int_0^T c(s)ds) + \int_0^T \exp(\int_0^s c(r)dr)c_0(s)ds]$$

$$= -E[A_u(T)\exp(\int_0^T c(s))ds] + R, \qquad (5.33)$$

where

$$R = E[\int_0^T \exp(\int_0^s c(r)dr)c_0(s)ds]. \qquad (5.34)$$

Therefore, to minimize $X_u(0)$ is the same as to maximize

$$J_0(u) := E[A_u(T)\exp(\int_0^T c(s)ds)]. \qquad (5.35)$$

If $c(s)$ is *deterministic* we see that

$$J_0(u) = \exp(\int_0^T c(s)ds)(a + E[\int_0^T \alpha(t)u(t)dt]) \qquad (5.36)$$

which has a finite maximum iff

$$\alpha(t) = 0 \text{ for a.a. } t, \omega \in [0, T] \times \Omega. \qquad (5.37)$$

And if (5.37) holds, an optimal control is

$$\hat{u} = 0 \text{ for a.a. } t, \omega \qquad (5.38)$$

and hence

$$\hat{X}(0) = X_0(0) = -a\exp(\int_0^T c(s)ds) + R, \qquad (5.39)$$

as obtained in Theorem 5.6.

However, we cannot get this conclusion from the above argument if $c(s)$ is *stochastic*. Theorem 5.6 states that it also holds in this case (provided an optimal control exists at all).

Remark 5.8 In (5.11) and the remaining part of Section 5 we have assumed that the controller has full information, i.e. that $\mathcal{E}_t = \mathcal{F}_t$ for all $t \in [0, T]$. However, it is possible to do a similar analysis also in some partial information cases, e.g. when $\mathcal{E}_t = \mathcal{F}_{t-\delta}$ for some positive constant δ. In this case one can solve the corresponding partial information Malliavin-differential type equation (5.10) by using that the process $\eta(t) := \eta(t - \delta)$ is a Lévy process with respect to $\mathcal{F}_{t-\delta}$.

5.2 Utility optimization under risk constraint

Our general forward-backward control formulation also covers the situation when one wants to maximize (over all admissible portfolios) the expected utility of the terminal wealth, subject to the constraint that the risk must not exceed a given threshold.

To explain this, consider again the forward-backward system (5.1)-(5.2).

Let $U : [0, \infty) \to [-\infty, \infty)$ be a given utility function and let ρ_0 be given real number. The risk constrained maximal utility problem is the following:

Maximize $E[U(A_u(T))]$ over all $u \in \mathcal{A}_{\mathcal{E}}$, subject to

$$\rho(A_u(T)) \leq \rho_0. \tag{5.40}$$

As before the risk ρ is interpreted as

$$\rho(A_u(T)) = X_u(0).$$

We can use the Lagrange multiplier method to study this problem:

Fix $y > 0$ and consider the *unconstrained* problem to maximize

$$E[U(A_u(T))] + y(\rho_0 - X_u(0)). \tag{5.41}$$

This is of the form (2.1)-(2.3), with $c = -1$, $f = 0$, $h_2(a) = U(a)$; $a > 0$ and $h_1(x) = -yx$; $x \in \mathbb{R}$.

Let the optimal control (portfolio) corresponding to y be denoted by $\hat{u}^{(y)}$ and let $\hat{X}^{(y)}(0)$ be the corresponding risk. Suppose we can find $y^* > 0$ such that the corresponding minimal risk $\hat{X}^{(y^*)}(0)$ satisfies (5.40) with *equality*, i.e.

$$\hat{X}^{(y^*)}(0) = \rho_0.$$

Then $u^* := \hat{u}^{(y^*)}$ is an optimal control for the constrained problem, because

$$
\begin{aligned}
\sup_{u:X_u(0)\leq\rho_0} E[U(A_u(T))] &\leq \sup_u E[U(A_u(T))] + y^*(\rho_0 - X_u(0))] \\
&= E[U(A_{u^*}(T))] + y^*(\rho - \hat{X}^{(y^*)}(0)) \\
&= E[U(A_{u^*}(T))] \\
&\leq \sup_{u:X_u(0)\leq\rho_0} E[U(A_u(Y))].
\end{aligned}
$$

We conclude that it suffices to study the problem to maximize (5.41) for each given $y > 0$ without constraints. For this we can proceed as in subsection 5.1.

Appendix

A   Solution of linear Malliavin differential equations

We first consider the *homogeneous* equation (5.10). We assume $\mathcal{E}_t = \mathcal{F}_t$. Let $\alpha(t)$, $\beta(t)$, and $\theta(t, z)$ be given $\mathcal{F}_t$-predictable processes. We want to find all $\mathcal{F}_T$-measurable Malliavin differentiable random variables $Z = Z(\omega) > 0$ such that

$$\alpha(t)E[Z \mid \mathcal{F}_t] + \beta(t)E[D_t Z \mid \mathcal{F}_t]$$
$$+ \int_{\mathbb{R}_0} \theta(t, z)E[D_{t,z}Z \mid \mathcal{F}_t]\nu(dz) = 0 \; ; \; t \in [0, T]. \tag{A.1}$$

Theorem A.1 *The general solution $Z > 0$ of the homogeneous equation* (A.1) *has the form*

$$Z = c \exp \left(\int_0^T \sigma(s)dB(s) - \frac{1}{2} \int_0^T \sigma^2(s)ds \right.$$
$$+ \int_0^T \int_{\mathbb{R}_0} \ln(1 + \gamma(s, z))\tilde{N}(ds, dz)$$
$$\left. + \int_0^T \int_{\mathbb{R}_0} \{\ln(1 + \gamma(s, z)) - \gamma(s, z)\} \, \nu(dz)ds \right) \tag{A.2}$$

for any constant $c > 0$, where $\sigma(s)$ and $\gamma(s, z)$ are any predictable processes such that

$$\alpha(t) + \beta(t)\sigma(t) + \int_{\mathbb{R}_0} \theta(t, z)\gamma(t, z)\nu(dz) = 0 \text{ for a.a. } t, \omega. \tag{A.3}$$

Proof.   Let $Z \in \mathbb{D}_{1,2}$ be $\mathcal{F}_T$-measurable. By the Itô representation theorem there exist $\mathcal{F}_t$-predictable processes $\sigma_0(t)$ and $\gamma_0(t, z)$ such that

$$Z = E[Z] + \int_0^T \sigma_0(s)dB(s) + \int_0^T \int_{\mathbb{R}_0} \gamma_0(s, z)\tilde{N}(ds, dz). \tag{A.4}$$

This gives

$$E[Z \mid \mathcal{F}_t] = E[Z] + \int_0^t \sigma_0(s)dB(s) + \int_0^t \int_{\mathbb{R}_0} \gamma_0(s, z)\tilde{N}(ds, dz), \tag{A.5}$$

$$E[D_t Z \mid \mathcal{F}_t] = E\left[\sigma_0(t) + \int_t^T D_t\sigma_0(s)dB(s) + \int_t^T \int_{\mathbb{R}_0} D_t\gamma_0(s, z)\tilde{N}(ds, dz) \mid \mathcal{F}_t\right]$$
$$= \sigma_0(t), \tag{A.6}$$

and

$$E[D_{t,z}Z \mid \mathcal{F}_t] = E\left[\gamma_0(t, z) + \int_t^T \int_{\mathbb{R}_0} D_{t,z}\gamma_0(s, \zeta)\tilde{N}(ds, d\zeta) \mid \mathcal{F}_t\right] = \gamma_0(t, z). \tag{A.7}$$

Substituting (A.5)-(A.7) into (A.11) we get

$$\alpha(t)E[Z \mid \mathcal{F}_t] + \beta(t)\sigma_0(t) + \int_{\mathbb{R}_0} \theta(t,z)\gamma_0(t,z)\nu(dz) = 0$$

or

$$\alpha(t) + \beta(t)\sigma(t) + \int_{\mathbb{R}_0} \theta(t,z)\gamma(t,z)\nu(dz) = 0, \tag{A.8}$$

where

$$\sigma(t) = \frac{\sigma_0(t)}{E[Z \mid \mathcal{F}_t]} \text{ and } \gamma(t,z) = \frac{\gamma_0(t,z)}{E[Z \mid \mathcal{F}_t]}. \tag{A.9}$$

Substituting (A.9) into (A.4) we get

$$Z = E[Z] + \int_0^T \sigma(s)E[Z \mid \mathcal{F}_s]dB(s) + \int_0^T \int_{\mathbb{R}_0} \gamma(s,z)E[Z \mid \mathcal{F}_{s^-}]\tilde{N}(ds,dz). \tag{A.10}$$

Therefore, if we define

$$Z(t) := E[Z \mid \mathcal{F}_t]. \tag{A.11}$$

We have

$$dZ(t) = Z(t^-)\left[\sigma(t)dB(t) + \int_{\mathbb{R}_0} \gamma(t,z)\tilde{N}(dt,dz)\right]. \tag{A.12}$$

This stochastic differential equation has the solution

$$Z(t) = E[Z] \exp\left(\int_0^t \sigma(s)dB(s) - \frac{1}{2}\int_0^t \sigma^2(s)ds \right.$$
$$+ \int_0^t \int_{\mathbb{R}_0} \ln(1+\gamma(s,z))\tilde{N}(ds,dz)$$
$$\left. + \int_0^t \int_{\mathbb{R}_0} \{\ln(1+\gamma(s,z)) - \gamma(s,z)\}\nu(dz)ds\right). \tag{A.13}$$

That completes the proof of Theorem A.1. □

We proceed to study the *inhomogeneous equation.*
Let $\alpha(t)$, $\beta(t)$, $\zeta(t)$ and $\theta(t,z)$ be given $\mathcal{F}_t$-predictable processes. We want to find all $\mathcal{F}_T$-measurable Malliavin differentiable random variables $Z = Z(\omega)$ such that

$$\alpha(t)E[Z \mid \mathcal{F}_t] + \beta(t)E[D_t Z \mid \mathcal{F}_t]$$
$$+ \int_{\mathbb{R}_0} \theta(t,z)E[D_{t,z}Z \mid \mathcal{F}_t]\nu(dz) = \zeta(t) \text{ for a.a. } (t,\omega) \in [0,T] \times \Omega. \tag{A.14}$$

Theorem A.2 *Suppose $Z = Z(T) \in \mathbb{D}_{1,2}$, where $Z(t)$ is a solution of the SDE*

$$dZ(t) = (\xi(t) + \sigma(t)Z(t))dB(t)$$
$$+ \int_{\mathbb{R}_0} (\eta(t,z) + \gamma(t,z)Z(t^-))\tilde{N}(dt,dz) \; ; \; Z(0) \in \mathbb{R}. \tag{A.15}$$

33

Then Z is a solution of (A.14) if the processes $\xi(t)$, $\sigma(t)$, $\eta(t, z)$ and $\gamma(t, z)$ are $\mathcal{F}_t$-predictable and satisfy the equations

$$\alpha(t) + \beta(t)\sigma(t) + \int_{\mathbb{R}_0} \theta(t, z)\gamma(t, z)\nu(dz) = 0, \ \ a.a. \ (t, \omega) \tag{A.16}$$

and

$$\beta(t)\xi(t) + \int_{\mathbb{R}_0} \theta(t, z)\eta(t, z)\nu(dz) = \zeta(t), \ \ a.a. \ (t, \omega). \tag{A.17}$$

Proof.   We proceed as in the proof of Theorem A.1. With Z as in (A.4) we get, by substituting (A.5)-(A.7) into (A.14),

$$\alpha(t)E[Z \mid \mathcal{F}_t] + \beta(t)\sigma_0(t) + \int_{\mathbb{R}_0} \theta(t, z)\gamma_0(t, z)\nu(dz) = \zeta(t). \tag{A.18}$$

Now choose processes $\sigma(t)$, $\xi(t)$, $\gamma(t, z)$ and $\eta(t, z)$ such that

$$\sigma_0(t) = \sigma(t)E[Z \mid \mathcal{F}_t] + \xi(t) \tag{A.19}$$

and

$$\gamma_0(t, z) = \gamma(t, z)E[Z \mid \mathcal{F}_t] + \eta(t, z). \tag{A.20}$$

Then (A.18) becomes

$$\left[\alpha(t) + \beta(t)\sigma(t) + \int_{\mathbb{R}_0} \theta(t, z)\gamma(t, z)\nu(dz)\right] E[Z \mid \mathcal{F}_t]$$
$$+ \beta(t)\xi(t) + \int_{\mathbb{R}_0} \theta(t, z)\eta(t, z)\nu(dz) = \zeta(t). \tag{A.21}$$

We conclude that the random variable $Z = Z(T)$ with $Z(t) = E[Z \mid \mathcal{F}_t]$ defined by (A.15) satisfies the equation (A.14) for any choice of σ, ξ, γ and η such that (A.16) and (A.17) hold. $\square$

References

[A] D. Applebaum: Lévy Processes and Stochastic Calculus. Cambridge University Press 2004.

[BØ] F. Baghery and B. Øksendal: A maximum principle for stochastic control with partial information. Stoch. Anal. Appl. 25 (2007), 705-717.

[B] S. Bahlali: Necessary and sufficient optimality conditions for relaxed and strict control problems of forward-backward systems. Manuscript 2008.

[BE] P. Barrieu and N. El Karoui: Optimal derivatives design under dynamic risk measures. Mathematics of Finance. Contemporary Mathematics (A.M.S. Proceedings), 13-26, 2004.

[BF] S. Biagini and M. Frittelli: On the extension of the Namioka-Klee theorem and on the Fatou property for risk measures. Manuscript 2007.

[CHMP] F. Coquet, Y. Hu, J. Mémin and S. Peng: Filtration consistent nonlinear expectations and related g-expectations. Probab. Th. Rel. Fields 123 (2002), 1-27.

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

[EPQ] N. El Karoui, S. Peng and M.-C. Quenez: Backward stochastic differential equations in finance. Mathematical Finance 7 (1997), 1-71.

[FØS] N.C. Framstad, B. Øksendal and A. Sulem: Sufficient stochastic maximum principle for the optimal control of jump diffusions and applications to finance. J. Optim. Theory Appl. 121 (2004), 77-98 [Errata 124(2005), 511-512].

[FR1] M. Frittelli and E. Rosazza-Gianin: Putting order in risk measures. J. Banking and Finance 26 (2002), 1473-1486.

[FR2] M. Frittelli and E. Rosazza-Gianin: Dynamic convex risk measures. In G. Szegö (editor): Risk Measures in the 21st Century, Wiley 2004, 227-248.

[FS] H. Föllmer and A. Schied: Convex measures of risk and trading constraints. Finance & Stoch. 6 (2002), 429-447.

[KS] J. Kallsen and A. Shiryaev: The cumulant process and Esscher's change of measure. Finance and Stochastics 6 (2002), 397-428.

[KO] Y. Karatzas and D. Ocone: A generalized Calrk representation formula, with application to optimal portfolios. Stochastics Stochastics Rep. 34 (1991), 187-220.

[MØ] S. Mataramvura and B. Øksendal: Risk minimizing portfolios and HJB equations for stochastic differential games. Stochastics 80 (2008), 317-337.

[MØZ] T. Meyer-Brandis,B. Øksendal and X.Y. Zhou: A stochastic maximum principle via Malliavin calculus . Eprint, University of Oslo, 10/2008.

[O] Y.Y. Okur: White noise generalization of the Clark-Ocone formula under change of measure. Eprint, University of Oslo 15/2007.

[ØS1] B. Øksendal and A. Sulem: Risk indifference pricing in jump diffusion markets. Math. Finance (to appear).

[ØS2] B. Øksendal and A. Sulem: Applied Stochastic Control of Jump Diffusions. Second Edition, Springer 2007.

[P] S. Peng: Backward SDE and related g-expectations. In N. El Karoui and L. Mazliak (editors): Backward Stochastic Differential Equations. Pitman Research Notes in Mathematics Series 364 (1997), 141-159.

[Pa] E. Pardoux: Backward stochastic differential equations and viscosity solutions of systems of semilinear parabolic and elliptic PDEs of second order. In L. Decreusefond et al. (editors): Stochastic Analysis and related Topics VI, The Geilo Workshop, 1996. Birkhäuser 1998, pp. 79-127.

[PW] S. Peng and Z. Wu: Fully coupled forward-backward stochastic differential equations and applications to optimal control. SIAM J. Control Optim. 37 (1999), 825-843.

[R] E. Rosazza Gianin: Risk measures via g-expectations. Insurance: Mathematics and Economics 39 (2006), 19-34.

[SW] J.-T. Shi and Z. Wu: Maximum principle for forward-backward stochastic control system with random jumps and applications to finance. Manuscript.