

OPTIMAL CONTROL OF STOCHASTIC DELAY EQUATIONS AND TIME-ADVANCED BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS

BERNT ØKSENDAL,* *University of Oslo and Norwegian School of Economics and Business Administration*

AGNÈS SULEM,** *INRIA Paris-Rocquencourt*

TUSHENG ZHANG,*** *University of Manchester*

Abstract

We study optimal control problems for (time-)delayed stochastic differential equations with jumps. We establish sufficient and necessary stochastic maximum principles for an optimal control of such systems. The associated adjoint processes are shown to satisfy a (time-)advanced backward stochastic differential equation (ABSDE). Several results on existence and uniqueness of such ABSDEs are shown. The results are illustrated by an application to optimal consumption from a cash flow with delay.

Keywords: Optimal control; stochastic delay equation; Lévy process; maximum principle; Hamiltonian; adjoint process; time-advanced BSDE

2010 Mathematics Subject Classification: Primary 93EXX; 93E20; 60J75
 Secondary 60H10; 60H15; 60H20; 49J55;
 35R60

1. Introduction and problem formulation

Let $B(t) = B(t, \omega)$ be a Brownian motion, and let $\tilde{N}(dt, dz) := N(dt, dz) - \nu(dz) dt$, where ν is the Lévy measure of the jump measure $N(\cdot, \cdot)$, be an independent compensated Poisson random measure on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0 \leq t \leq T}, \mathbb{P})$.

We consider a controlled stochastic delay equation of the form

$$dX(t) = b(t, X(t), Y(t), A(t), u(t), \omega) dt + \sigma(t, X(t), Y(t), A(t), u(t), \omega) dB(t) + \int_{\mathbb{R}} \theta(t, X(t), Y(t), A(t), u(t), z, \omega) \tilde{N}(dt, dz), \quad t \in [0, T], \quad (1.1)$$

$$X(t) = x_0(t), \quad t \in [-\delta, 0], \quad (1.2)$$

where

$$Y(t) = X(t - \delta), \quad A(t) = \int_{t-\delta}^t e^{-\rho(t-r)} X(r) dr, \quad (1.3)$$

Received 14 September 2010; revision received 22 January 2011.

* Postal address: Center of Mathematics for Applications (CMA), University of Oslo, Box 1053 Blindern, N-0316 Oslo, Norway. Email address: oksendal@math.uio.no

** Postal address: INRIA Paris-Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 Le Chesnay Cedex, France. Email address: agnes.sulem@inria.fr

*** Postal address: School of Mathematics, University of Manchester, Oxford Road, Manchester M13 9PL, UK. Email address: tusheng.zhang@manchester.ac.uk

and $\delta > 0$, $\rho \geq 0$, and $T > 0$ are given constants. Here

$$b: [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathcal{U} \times \Omega \rightarrow \mathbb{R},$$

$$\sigma: [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathcal{U} \times \Omega \rightarrow \mathbb{R},$$

and

$$\theta: [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathcal{U} \times \mathbb{R}_0 \times \Omega \rightarrow \mathbb{R}$$

are given functions such that, for all t , $b(t, x, y, a, u, \cdot)$, $\sigma(t, x, y, a, u, \cdot)$, and $\theta(t, x, y, a, u, z, \cdot)$ are \mathcal{F}_t -measurable for all $x \in \mathbb{R}$, $y \in \mathbb{R}$, $a \in \mathbb{R}$, $u \in \mathcal{U}$, and $z \in \mathbb{R}_0 := \mathbb{R} \setminus \{0\}$. The function $x_0(t)$ is assumed to be continuous and deterministic. We refer the reader to [13] for more information about the optimal control of jump diffusions.

Let $\mathcal{E}_t \subseteq \mathcal{F}_t$, $t \in [0, T]$, be a given subfiltration of $\{\mathcal{F}_t\}_{t \in [0, T]}$, representing the information available to the controller who decides the value of $u(t)$ at time t . For example, we could have $\mathcal{E}_t = \mathcal{F}_{(t-c)^+}$ for some given $c > 0$. Let $\mathcal{U} \subset \mathbb{R}$ be a given set of admissible control values $u(t)$, $t \in [0, T]$, and let $\mathcal{A}_{\mathcal{E}}$ be a given family of admissible control processes $u(\cdot)$, included in the set of càdlàg, \mathcal{E} -adapted, and \mathcal{U} -valued processes $u(t)$, $t \in [0, T]$, such that (1.1)–(1.2) has a unique solution $X(\cdot) \in L^2(\lambda \times \mathbb{P})$, where λ denotes the Lebesgue measure on $[0, T]$.

The performance functional is assumed to have the form

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

where $f = f(t, x, y, a, u, \omega): [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathcal{U} \times \Omega \rightarrow \mathbb{R}$ and $g = g(x, \omega): \mathbb{R} \times \Omega \rightarrow \mathbb{R}$ are given \mathcal{C}^1 functions with respect to (x, y, a, u) such that

$$\mathbb{E} \left[\int_0^T \left\{ |f(t, X(t), A(t), u(t))| + \left| \frac{\partial f}{\partial x_i}(t, X(t), Y(t), A(t), u(t)) \right|^2 \right\} dt + |g(X(T))| + |g'(X(T))|^2 \right] < \infty \quad \text{for } x_i = x, y, a, \text{ and } u.$$

Here, and in the following, we suppress the ω for notational simplicity. The problem we consider in this paper is the following.

Find $\Phi(x_0)$ and $u^* \in \mathcal{A}_{\mathcal{E}}$ such that

$$\Phi(x_0) := \sup_{u \in \mathcal{A}_{\mathcal{E}}} J(u) = J(u^*). \tag{1.4}$$

Any control $u^* \in \mathcal{A}_{\mathcal{E}}$ satisfying (1.4) is called an *optimal control*.

This is an example of a stochastic control problem for a system with delay. Such problems appear in many applications. For example, for biological reasons, delays occur naturally in population dynamics models. Therefore, when dealing with, e.g. the optimal harvesting problem of biological systems, one is led to the optimal control of systems with delay.

Another area of applications is mathematical finance, where delays in the dynamics can represent memory or inertia in the financial system.

Variants of this problem have been studied in several papers. The stochastic control of delay systems is a challenging research area, because delay systems have, in general, an infinite-dimensional nature. Hence, the natural general approach to them is infinite-dimensional. For this kind of approach in the context of control problems, we refer the reader to [2], [6], [7],

and [8] in the stochastic Brownian case. To the best of the authors' knowledge, despite the statement of a result in [17], this kind of approach was not developed for delay systems driven by a Lévy noise.

Nonetheless, in some cases still very interesting for the applications, it happens that systems with delay can be reduced to finite-dimensional systems, since the information we need from their dynamics can be represented by a finite-dimensional variable evolving in terms of itself. In such a context, the crucial point is to understand when this finite-dimensional reduction of the problem is possible and/or to find conditions ensuring that. There are some papers dealing with this subject in the stochastic Brownian case: we refer the reader to [5], [9], [10], [11], and [12]. The paper [3] represents an extension of [11] to the case when the equation is driven by a Lévy noise.

We also mention the paper [4], where certain control problems of stochastic functional differential equations are studied by means of the Girsanov transformation. This approach, however, does not work if there is a delay in the noise components.

In [14] the problem is transformed into a stochastic control problem for Volterra equations.

Our approach in the current paper is different from all the above. Note that the presence of the terms $Y(t)$ and $A(t)$ in (1.1) makes the problem non-Markovian and we cannot use a (finite-dimensional) dynamic programming approach. However, we will show that it is possible to obtain a (Pontryagin–Bismut–Bensoussan-type) stochastic maximum principle for the problem.

The paper is organised as follows: we introduce the Hamiltonian and the time-advanced backward stochastic differential equation (BSDE) for the adjoint processes in Section 2, then we prove sufficient and necessary stochastic maximum principles in Sections 3 and 4. Section 5 is devoted to existence and uniqueness theorems for various time-advanced BSDEs with jumps. Finally, an example of application to an optimal consumption problem with delay is given in Section 6.

2. Hamiltonian and time-advanced BSDEs for adjoint equations

We define the *Hamiltonian*

$$H : [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathcal{U} \times \mathbb{R} \times \mathbb{R} \times \mathcal{R} \times \Omega \rightarrow \mathbb{R}$$

by

$$\begin{aligned} H(t, x, y, a, u, p, q, r(\cdot), \omega) &= H(t, x, y, a, u, p, q, r(\cdot)) \\ &= f(t, x, y, a, u) + b(t, x, y, a, u)p + \sigma(t, x, y, a, u)q \\ &\quad + \int_{\mathbb{R}_0} \theta(t, x, y, a, u, z)r(z)v(dz), \end{aligned} \tag{2.1}$$

where \mathcal{R} is the set of functions $r : \mathbb{R}_0 \rightarrow \mathbb{R}$ such that the last term in (2.1) converges.

We assume that b, σ , and θ are \mathcal{C}^1 functions with respect to (x, y, a, u) and that

$$\begin{aligned} \mathbb{E} \left[\int_0^T \left\{ \left| \frac{\partial b}{\partial x_i}(t, X(t), Y(t), A(t), u(t)) \right|^2 + \left| \frac{\partial \sigma}{\partial x_i}(t, X(t), Y(t), A(t), u(t)) \right|^2 \right. \right. \\ \left. \left. + \int_{\mathbb{R}_0} \left| \frac{\partial \theta}{\partial x_i}(t, X(t), Y(t), A(t), u(t), z) \right|^2 v(dz) \right\} dt \right] < \infty \end{aligned} \tag{2.2}$$

for $x_i = x, y, a$, and u .

Associated to H we define the adjoint processes $p(t), q(t), r(t, z), t \in [0, T], z \in \mathbb{R}_0$, by the BSDE

$$\begin{aligned} dp(t) &= E[\mu(t) \mid \mathcal{F}_t] dt + q(t) dB(t) + \int_{\mathbb{R}_0} r(t, z) \tilde{N}(dt, dz), \quad t \in [0, T], \\ p(T) &= g'(X(T)), \end{aligned} \tag{2.3}$$

where

$$\begin{aligned} \mu(t) &= -\frac{\partial H}{\partial x}(t, X(t), Y(t), A(t), u(t), p(t), q(t), r(t, \cdot)) \\ &\quad - \frac{\partial H}{\partial y}(t + \delta, X(t + \delta), Y(t + \delta), A(t + \delta), u(t + \delta), p(t + \delta), q(t + \delta), r(t + \delta, \cdot)) \\ &\quad \times \mathbf{1}_{[0, T-\delta]}(t) \\ &\quad - e^{\rho t} \left(\int_t^{t+\delta} \frac{\partial H}{\partial a}(s, X(s), Y(s), A(s), u(s), p(s), q(s), r(s, \cdot)) e^{-\rho s} \mathbf{1}_{[0, T]}(s) ds \right). \end{aligned} \tag{2.4}$$

Note that this BSDE is *anticipative*, or *time advanced* in the sense that the process $\mu(t)$ contains future values of $X(s), u(s), p(s), q(s), r(s, \cdot), s \leq t + \delta$.

In the case when there are no jumps and no integral term in (2.4), anticipative BSDEs (ABSDEs for short) have been studied by Peng and Yang [16], who proved existence and uniqueness of such equations under certain conditions. They also related a class of linear ABSDEs to a class of linear stochastic delay control problems where there is no delay in the noise coefficients. Thus, in our paper we extend this relation to general nonlinear control problems and general nonlinear ABSDEs by means of the maximum principle, where throughout the study we include the possibility of delays also in the noise coefficients, as well as the possibility of jumps. After this paper was written, we became aware of a recent paper [1] on a similar maximum principle for the stochastic control problem of delayed systems. However, the authors do not consider delays of moving average type and they do not allow jumps. On the other hand, they allow delay in the control.

Remark 2.1. The methods used in this paper extend easily to more general delay systems. For example, we can add finitely many delay terms of the form

$$Y_i(t) := X(t - \delta_i), \quad i = 1, \dots, N,$$

where the $\delta_i > 0$ are given, in all the coefficients. Moreover, we can include more general moving average terms of the form

$$A_j(t) := \int_{t-\delta}^t \phi_j(t, s) X(s) ds, \quad j = 1, \dots, M,$$

where the $\phi_j(t, s), 0 \leq s \leq t \leq T$, are given locally bounded \mathcal{F}_s -adapted processes.

In this case the process $\mu(t)$ in (2.4) must be modified accordingly. More precisely, the last term in (2.4) must be changed to the sum of

$$\begin{aligned} & - \int_t^{t+\delta} \frac{\partial H}{\partial a_j}(t, X(s), Y_1(s), \dots, Y_N(s), A_1(s), \dots, A_M(s), u(s), p(s), q(s), r(s, \cdot)) \\ & \quad \times \phi_j(t, s) \mathbf{1}_{[0, T]}(s) ds, \quad 1 \leq j \leq M. \end{aligned}$$

Even more generally, when modified appropriately, our method could also deal with delay terms of the form

$$A^\lambda(t) := \int_{t-\delta}^t \phi(t, s)X(s) \, d\lambda(s),$$

where λ is a given (positive) measure on $[0, T]$. In this case the BSDE (2.3) must be modified to

$$\begin{aligned} dp(t) = & - \left\{ \mathbb{E} \left[\frac{\partial H}{\partial x}(t) \mid \mathcal{F}_t \right] + \sum_{i=1}^N \mathbb{E} \left[\frac{\partial H}{\partial y}(t + \delta_i) \mid \mathcal{F}_t \right] \mathbf{1}_{[0, T-\delta_i]}(t) \right\} dt \\ & - \mathbb{E} \left[\left(\int_t^{t+\delta} \frac{\partial H}{\partial a}(s) \phi(t, s) \mathbf{1}_{[0, T]}(s) \, ds \right) \mid \mathcal{F}_t \right] d\lambda(t) + q(t) \, dB(t) \\ & + \int_{\mathbb{R}_0} r(t, z) \tilde{N}(dt, dz), \quad t \in [0, T], \\ p(T) = & g'(X(T)), \end{aligned}$$

where we have used the simplified notation

$$\frac{\partial}{\partial x} H(t) = \frac{\partial}{\partial x} H(t, X(t), Y_1(t), \dots, Y_N(t), \dots), \quad \text{etc.}$$

For simplicity of presentation, however, we have chosen to focus on just the two types of delay, $Y(t)$ and $A(t)$ given in (1.3).

3. A sufficient maximum principle

In this section we establish a maximum principle of sufficient type, i.e. we show that, under some assumptions, maximizing the Hamiltonian leads to an optimal control.

Theorem 3.1. (Sufficient maximum principle.) *Let $\hat{u} \in \mathcal{A}_\mathcal{E}$ with corresponding state processes $\hat{X}(t)$, $\hat{Y}(t)$, and $\hat{A}(t)$ and adjoint processes $\hat{p}(t)$, $\hat{q}(t)$, and $\hat{r}(t, z)$, assumed to satisfy the ABSDE (2.3)–(2.4). Suppose that the following assertions hold.*

(i) *The functions $x \rightarrow g(x)$ and*

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

are concave for each $t \in [0, T]$ almost surely (a.s.).

(ii)

$$\begin{aligned} \mathbb{E} \left[\int_0^T \left\{ \hat{p}(t)^2 \left(\sigma^2(t) + \int_{\mathbb{R}_0} \theta^2(t, z) \nu(dz) \right) \right. \right. \\ \left. \left. + X^2(t) \left(\hat{q}^2(t) + \int_{\mathbb{R}_0} \hat{r}^2(t, z) \nu(dz) \right) \right\} dt \right] < \infty \end{aligned} \tag{3.1}$$

for all $u \in \mathcal{A}_\mathcal{E}$.

(iii)

$$\begin{aligned} \max_{v \in \mathcal{U}} \mathbb{E}[H(t, \hat{X}(t), \hat{X}(t - \delta), \hat{A}(t), v, \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot)) \mid \mathcal{E}_t] \\ = \mathbb{E}[H(t, \hat{X}(t), \hat{X}(t - \delta), \hat{A}(t), \hat{u}(t), \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot)) \mid \mathcal{E}_t] \end{aligned}$$

for all $t \in [0, T]$ a.s.

Then $\hat{u}(t)$ is an optimal control for problem (1.4).

Proof. Choose $u \in \mathcal{A}_g$, and consider

$$J(u) - J(\hat{u}) = I_1 + I_2, \tag{3.2}$$

where

$$I_1 = \mathbb{E} \left[\int_0^T \{f(t, X(t), Y(t), A(t), u(t)) - f(t, \hat{X}(t), \hat{Y}(t), \hat{A}(t), \hat{u}(t))\} dt \right], \tag{3.3}$$

$$I_2 = \mathbb{E}[g(X(T)) - g(\hat{X}(T))]. \tag{3.4}$$

By the definition of H and the concavity of H , we have

$$\begin{aligned} I_1 &= \mathbb{E} \left[\int_0^T \{H(t, X(t), Y(t), A(t), u(t), \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot)) \right. \\ &\quad - H(t, \hat{X}(t), \hat{Y}(t), \hat{A}(t), \hat{u}(t), \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot)) \\ &\quad - (b(t, X(t), Y(t), A(t), u(t)) - b(t, \hat{X}(t), \hat{Y}(t), \hat{A}(t), \hat{u}(t)))\hat{p}(t) \\ &\quad - (\sigma(t, X(t), Y(t), A(t), u(t)) - \sigma(t, \hat{X}(t), \hat{Y}(t), \hat{A}(t), \hat{u}(t)))\hat{q}(t) \\ &\quad \left. - \int_{\mathbb{R}} (\theta(t, X(t), Y(t), A(t), u(t), z) - \theta(t, \hat{X}(t), \hat{Y}(t), \hat{A}(t), \hat{u}(t), z)) \right. \\ &\quad \left. \times \hat{r}(t, z) \nu(dz) \} dt \right] \\ &\leq \mathbb{E} \left[\int_0^T \left\{ \frac{\partial \hat{H}}{\partial x}(t)(X(t) - \hat{X}(t)) + \frac{\partial \hat{H}}{\partial y}(t)(Y(t) - \hat{Y}(t)) + \frac{\partial \hat{H}}{\partial a}(t)(A(t) - \hat{A}(t)) \right. \right. \\ &\quad + \frac{\partial H}{\partial u}(t)(u(t) - \hat{u}(t)) - (b(t) - \hat{b}(t))\hat{p}(t) - (\sigma(t) - \hat{\sigma}(t))\hat{q}(t) \\ &\quad \left. \left. - \int_{\mathbb{R}} (\theta(t, z) - \hat{\theta}(t, z))\hat{r}(t, z) \nu(dz) \right\} dt \right], \tag{3.5} \end{aligned}$$

where we have used the abbreviated notation

$$\begin{aligned} \frac{\partial}{\partial x} \hat{H}(t) &= \frac{\partial}{\partial x} H(t, \hat{X}(t), \hat{Y}(t), \hat{A}(t), \hat{u}(t), \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot)), \\ b(t) &= b(t, X(t), Y(t), A(t), u(t)), \\ \hat{b}(t) &= b(t, \hat{X}(t), \hat{Y}(t), \hat{A}(t), \hat{u}(t)), \end{aligned}$$

etc. Since g is concave, we have, by (3.1),

$$\begin{aligned} I_2 &\leq \mathbb{E}[g'(\hat{X}(T))(X(T) - \hat{X}(T))] \\ &= \mathbb{E}[\hat{p}(T)(X(T) - \hat{X}(T))] \\ &= \mathbb{E} \left[\int_0^T \hat{p}(t)(dX(t) - d\hat{X}(t)) + \int_0^T (X(t) - \hat{X}(t)) d\hat{p}(t) \right. \\ &\quad \left. + \int_0^T (\sigma(t) - \hat{\sigma}(t))\hat{q}(t) dt + \int_0^T \int_{\mathbb{R}} (\theta(t, z) - \hat{\theta}(t, z))\hat{r}(t, z) \nu(dz) dt \right] \end{aligned}$$

$$\begin{aligned}
 &= \mathbb{E} \left[\int_0^T (b(t) - \hat{b}(t)) \hat{p}(t) dt + \int_0^T (X(t) - \hat{X}(t)) \mathbb{E}[\mu(t) \mid \mathcal{F}_t] dt \right. \\
 &\quad \left. + \int_0^T (\sigma(t) - \hat{\sigma}(t)) \hat{q}(t) dt + \int_0^T \int_{\mathbb{R}} (\theta(t, z) - \hat{\theta}(t, z)) \hat{r}(t, z) \nu(dz) dt \right]. \tag{3.6}
 \end{aligned}$$

Combining (3.2)–(3.6) we obtain, using the fact that $X(t) = \hat{X}(t) = x_0(t)$ for all $t \in [-\delta, 0]$,

$$\begin{aligned}
 J(u) - J(\hat{u}) &\leq \mathbb{E} \left[\int_0^T \left\{ \frac{\partial H}{\partial x}(t)(X(t) - \hat{X}(t)) + \frac{\partial H}{\partial y}(t)(Y(t) - \hat{Y}(t)) \right. \right. \\
 &\quad \left. \left. + \frac{\partial \hat{H}}{\partial a}(t)(A(t) - \hat{A}(t)) + \frac{\partial \hat{H}}{\partial u}(t)(u(t) - \hat{u}(t)) \right. \right. \\
 &\quad \left. \left. + \mu(t)(X(t) - \hat{X}(t)) \right\} dt \right] \\
 &= \mathbb{E} \left[\int_{\delta}^{T+\delta} \left\{ \frac{\partial \hat{H}}{\partial x}(t - \delta) + \frac{\partial \hat{H}}{\partial y}(t) \mathbf{1}_{[0, T]}(t) + \mu(t - \delta) \right\} (Y(t) - \hat{Y}(t)) dt \right. \\
 &\quad \left. + \int_0^T \frac{\partial \hat{H}}{\partial a}(t)(A(t) - \hat{A}(t)) dt + \int_0^T \frac{\partial \hat{H}}{\partial u}(t)(u(t) - \hat{u}(t)) dt \right]. \tag{3.7}
 \end{aligned}$$

Using integration by parts and substituting $r = t - \delta$, we obtain

$$\begin{aligned}
 &\int_0^T \frac{\partial \hat{H}}{\partial a}(s)(A(s) - \hat{A}(s)) ds \\
 &= \int_0^T \frac{\partial \hat{H}}{\partial a}(s) \int_{s-\delta}^s e^{-\rho(s-r)} (X(r) - \hat{X}(r)) dr ds \\
 &= \int_0^T \left(\int_r^{r+\delta} \frac{\partial \hat{H}}{\partial a}(s) e^{-\rho s} \mathbf{1}_{[0, T]}(s) ds \right) e^{\rho r} (X(r) - \hat{X}(r)) dr \\
 &= \int_{\delta}^{T+\delta} \left(\int_{t-\delta}^t \frac{\partial \hat{H}}{\partial a}(s) e^{-\rho s} \mathbf{1}_{[0, T]}(s) ds \right) e^{\rho(t-\delta)} (X(t - \delta) - \hat{X}(t - \delta)) dt.
 \end{aligned}$$

Combining this with (3.7) and using (2.4), we obtain

$$\begin{aligned}
 &J(u) - J(\hat{u}) \\
 &\leq \left[\int_{\delta}^{T+\delta} \left\{ \frac{\partial \hat{H}}{\partial x}(t - \delta) + \frac{\partial \hat{H}}{\partial y}(t) \mathbf{1}_{[0, T]}(t) + \left(\int_{t-\delta}^t \frac{\partial \hat{H}}{\partial a}(s) e^{-\rho s} \mathbf{1}_{[0, T]}(s) ds \right) e^{\rho(t-\delta)} \right. \right. \\
 &\quad \left. \left. + \mu(t - \delta) \right\} (Y(t) - \hat{Y}(t)) dt + \int_0^T \frac{\partial \hat{H}}{\partial u}(t)(u(t) - \hat{u}(t)) dt \right] \\
 &= \mathbb{E} \left[\int_0^T \frac{\partial \hat{H}}{\partial u}(t)(u(t) - \hat{u}(t)) dt \right] \\
 &= \mathbb{E} \left[\int_0^T \mathbb{E} \left[\frac{\partial \hat{H}}{\partial u}(t)(u(t) - \hat{u}(t)) \mid \mathcal{E}_t \right] dt \right] \\
 &= \mathbb{E} \left[\int_0^T \mathbb{E} \left[\frac{\partial \hat{H}}{\partial u}(t) \mid \mathcal{E}_t \right] (u(t) - \hat{u}(t)) dt \right] \\
 &\leq 0.
 \end{aligned}$$

The last inequality holds because $v = \hat{u}(t)$ maximizes $E[H(t, \hat{X}(t), \hat{Y}(t), \hat{A}(t), v, \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot) \mid \mathcal{E}_t]$ for each $t \in [0, T]$. This proves that \hat{u} is an optimal control.

4. A necessary maximum principle

A drawback with the sufficient maximum principle in Section 3 is the condition of concavity, which does not always hold in the applications. In this section we will prove a result going in the other direction. More precisely, we will prove the equivalence between being a directional critical point for $J(u)$ and a critical point for the conditional Hamiltonian. To this end, we need to make the following assumptions.

(A1) For all $u \in \mathcal{A}_\varepsilon$ and all bounded $\beta \in \mathcal{A}_\varepsilon$, there exists $\varepsilon > 0$ such that

$$u + s\beta \in \mathcal{A}_\varepsilon \quad \text{for all } s \in (-\varepsilon, \varepsilon).$$

(A2) For all $t_0 \in [0, T]$ and all bounded \mathcal{E}_{t_0} -measurable random variables α , the control process $\beta(t)$ defined by

$$\beta(t) = \alpha \mathbf{1}_{[t_0, T]}(t), \quad t \in [0, T], \tag{4.1}$$

belongs to \mathcal{A}_ε .

(A3) For all bounded $\beta \in \mathcal{A}_\varepsilon$, the derivative process

$$\xi(t) := \left. \frac{d}{ds} X^{u+s\beta}(t) \right|_{s=0} \tag{4.2}$$

exists and belongs to $L^2(\lambda \times P)$.

It follows from (1.1) that

$$\begin{aligned} d\xi(t) = & \left\{ \frac{\partial b}{\partial x}(t)\xi(t) + \frac{\partial b}{\partial y}(t)\xi(t - \delta) + \frac{\partial b}{\partial a}(t) \int_{t-\delta}^t e^{-\rho(t-r)} \xi(r) \, dr + \frac{\partial b}{\partial u}(t)\beta(t) \right\} dt \\ & + \left\{ \frac{\partial \sigma}{\partial x}(t)\xi(t) + \frac{\partial \sigma}{\partial y}(t)\xi(t - \delta) + \frac{\partial \sigma}{\partial a}(t) \int_{t-\delta}^t e^{-\rho(t-r)} \xi(r) \, dr + \frac{\partial \sigma}{\partial u}(t)\beta(t) \right\} dB(t) \\ & + \int_{\mathbb{R}_0} \left\{ \frac{\partial \theta}{\partial x}(t, z)\xi(t) + \frac{\partial \theta}{\partial y}(t, z)\xi(t - \delta) \right. \\ & \left. + \frac{\partial \theta}{\partial a}(t) \int_{t-\delta}^t e^{-\rho(t-r)} \xi(r) \, dr + \frac{\partial \theta}{\partial u}(t)\beta(t) \right\} \tilde{N}(dt, dz), \end{aligned} \tag{4.3}$$

where, for simplicity of notation, we have set

$$\frac{\partial}{\partial x} b(t) = \frac{\partial}{\partial x} b(t, X(t), X(t - \delta), A(t), u(t)),$$

etc., and we have used the facts that

$$\left. \frac{d}{ds} Y^{u+s\beta}(t) \right|_{s=0} = \left. \frac{d}{ds} X^{u+s\beta}(t - \delta) \right|_{s=0} = \xi(t - \delta)$$

and

$$\begin{aligned} \frac{d}{ds} A^{u+s\beta}(t) \Big|_{s=0} &= \frac{d}{ds} \left(\int_{t-\delta}^t e^{-\rho(t-r)} X^{u+s\beta}(r) dr \right) \Big|_{s=0} \\ &= \int_{t-\delta}^t e^{-\rho(t-r)} \frac{d}{ds} X^{u+s\beta}(r) \Big|_{s=0} dt \\ &= \int_{t-\delta}^t e^{-\rho(t-r)} \xi(r) dr. \end{aligned}$$

Note that

$$\xi(t) = 0 \quad \text{for } t \in [-\delta, 0].$$

Theorem 4.1. (Necessary maximum principle.) *Suppose that $\hat{u} \in \mathcal{A}_\xi$ with corresponding solutions $\hat{X}(t)$ of (1.1)–(1.2) and $\hat{p}(t)$, $\hat{q}(t)$, and $\hat{r}(t, z)$ of (2.2)–(2.3), and corresponding derivative process $\hat{\xi}(t)$ given by (4.2).*

Assume that

$$\begin{aligned} E \left[\int_0^T \hat{p}^2(t) \left\{ \left(\frac{\partial \sigma}{\partial x} \right)^2(t) \hat{\xi}^2(t) + \left(\frac{\partial \sigma}{\partial y} \right)^2(t) \xi^2(t - \delta) \right. \right. \\ + \left(\frac{\partial \sigma}{\partial a} \right)^2(t) \left(\int_{t-\delta}^t e^{-\rho(t-r)} \hat{\xi}(r) dr \right)^2 + \left(\frac{\partial \sigma}{\partial u} \right)^2(t) \\ + \int_{\mathbb{R}_0} \left\{ \left(\frac{\partial \theta}{\partial x} \right)^2(t, z) \hat{\xi}^2(t) + \left(\frac{\partial \theta}{\partial y} \right)^2(t, z) \xi^2(t - \delta) \right. \\ \left. \left. + \left(\frac{\partial \theta}{\partial a} \right)^2(t, z) \left(\int_{t-\delta}^t e^{-\rho(t-r)} \hat{\xi}(r) dr \right)^2 + \left(\frac{\partial \theta}{\partial u} \right)^2(t, z) \right\} \nu(dz) \right\} dt \\ + \int_0^T \hat{\xi}^2(t) \left\{ \hat{q}^2(t) + \int_{\mathbb{R}_0} \hat{r}^2(t, z) \nu(dz) \right\} dt \Big] < \infty. \end{aligned}$$

Then the following assertions are equivalent.

(i) For all bounded $\beta \in \mathcal{A}_\xi$,

$$\frac{d}{ds} J(\hat{u} + s\beta) \Big|_{s=0} = 0.$$

(ii) For all $t \in [0, T]$,

$$E \left[\frac{\partial}{\partial u} H(t, \hat{X}(t), \hat{Y}(t), \hat{A}(t), u, \hat{p}(t), \hat{q}(t), \hat{r}(t, \cdot)) \Big| \mathcal{E}_t \right]_{u=\hat{u}(t)} = 0 \quad \text{a.s.}$$

Proof. For simplicity of notation, we write $\hat{u} = u$, $\hat{X} = X$, $\hat{p} = p$, $\hat{q} = q$, and $\hat{r} = r$ in the following. Suppose that assertion (i) holds. Then

$$\begin{aligned} 0 &= \frac{d}{ds} J(u + s\beta) \Big|_{s=0} \\ &= \frac{d}{ds} E \left[\int_0^T f(t, X^{u+s\beta}(t), Y^{u+s\beta}(t), A^{u+s\beta}(t), u(t) + s\beta(t)) dt + g(X^{u+s\beta}(T)) \right] \Big|_{s=0} \\ &= E \left[\int_0^T \left\{ \frac{\partial f}{\partial x}(t) \xi(t) + \frac{\partial f}{\partial y}(t) \xi(t - \delta) + \frac{\partial f}{\partial a}(t) \int_{t-\delta}^t e^{-\rho(t-r)} \xi(r) dt + \frac{\partial f}{\partial u}(t) \beta(t) \right\} dt \right. \\ &\quad \left. + g'(X(T)) \xi(T) \right] \end{aligned}$$

$$\begin{aligned}
 &= \mathbb{E} \left[\int_0^T \left\{ \frac{\partial H}{\partial x}(t) - \frac{\partial b}{\partial x}(t)p(t) - \frac{\partial \sigma}{\partial x}(t)q(t) - \int_{\mathbb{R}} \frac{\partial \theta}{\partial x}(t, z)r(t, z)v(dz) \right\} \xi(t) dt \right. \\
 &\quad + \int_0^T \left\{ \frac{\partial H}{\partial y}(t) - \frac{\partial b}{\partial y}(t)p(t) - \frac{\partial \sigma}{\partial y}(t)q(t) - \int_{\mathbb{R}} \frac{\partial \theta}{\partial y}(t, z)r(t, z)v(dz) \right\} \xi(t - \delta) dt \\
 &\quad + \int_0^T \left\{ \frac{\partial H}{\partial a}(t) - \frac{\partial b}{\partial a}(t)p(t) - \frac{\partial \sigma}{\partial a}(t)q(t) - \int_{\mathbb{R}} \frac{\partial \theta}{\partial a}(t, z)r(t, z)v(dz) \right\} \\
 &\quad \quad \times \left(\int_{t-\delta}^t e^{-\rho(t-r)} \xi(r) dr \right) dt \\
 &\quad \left. + \int_0^T \frac{\partial f}{\partial u}(t)\beta(t) dt + g'(X(T))\xi(T) \right]. \tag{4.4}
 \end{aligned}$$

By (4.3),

$$\begin{aligned}
 &\mathbb{E}[g'(X(T))\xi(T)] \\
 &= \mathbb{E}[p(T)\xi(T)] \\
 &= \mathbb{E} \left[\int_0^T p(t) d\xi(t) + \int_0^T \xi(t) dp(t) \right. \\
 &\quad + \int_0^T q(t) \left\{ \frac{\partial \sigma}{\partial x}(t)\xi(t) + \frac{\partial \sigma}{\partial y}(t)\xi(t - \delta) + \frac{\partial \sigma}{\partial a}(t) \int_{t-\delta}^t e^{-\rho(t-r)} \xi(r) dr \right. \\
 &\quad \quad \left. + \frac{\partial \sigma}{\partial u}(t)\beta(t) \right\} dt \\
 &\quad + \int_0^T \int_{\mathbb{R}} r(t, z) \left\{ \frac{\partial \theta}{\partial x}(t, z)\xi(t) + \frac{\partial \theta}{\partial y}(t, z)\xi(t - \delta) \right. \\
 &\quad \quad \left. + \frac{\partial \theta}{\partial a}(t, z) \int_{t-\delta}^t e^{-\rho(t-r)} \xi(r) dr + \frac{\partial \theta}{\partial u}(t)\beta(t) \right\} v(dz) dt \Big] \\
 &= \mathbb{E} \left[\int_0^T p(t) \left\{ \frac{\partial b}{\partial x}(t)\xi(t) + \frac{\partial b}{\partial y}(t)\xi(t - \delta) + \frac{\partial b}{\partial a}(t) \int_{t-\delta}^t e^{-\rho(t-r)} \xi(r) dr \right. \right. \\
 &\quad \quad \left. \left. + \frac{\partial b}{\partial u}(t)\beta(t) \right\} dt + \int_0^T \xi(t) \mathbb{E}[\mu(t) | \mathcal{F}_t] dt \right. \\
 &\quad + \int_0^T q(t) \left\{ \frac{\partial \sigma}{\partial x}(t)\xi(t) + \frac{\partial \sigma}{\partial y}(t)\xi(t - \delta) + \frac{\partial \sigma}{\partial a}(t) \int_{t-\delta}^t e^{-\rho(t-r)} \xi(r) dr \right. \\
 &\quad \quad \left. + \frac{\partial \sigma}{\partial u}(t)\beta(t) \right\} dt \\
 &\quad + \int_0^T \int_{\mathbb{R}} r(t, z) \left\{ \frac{\partial \theta}{\partial x}(t, z)\xi(t) + \frac{\partial \theta}{\partial y}(t, z)\xi(t - \delta) \right. \\
 &\quad \quad \left. + \frac{\partial \theta}{\partial a}(t, z) \int_{t-\delta}^t e^{-\rho(t-r)} \xi(r) dr + \frac{\partial \theta}{\partial u}(t, z)\beta(t) \right\} v(dz) dt \Big]. \tag{4.5}
 \end{aligned}$$

Combining (4.4) and (4.5) we obtain

$$\begin{aligned}
 0 &= \mathbb{E} \left[\int_0^T \xi(t) \left\{ \frac{\partial H}{\partial x}(t) + \mu(t) \right\} dt + \int_0^T \xi(t - \delta) \frac{\partial H}{\partial y}(t) dt \right. \\
 &\quad \left. + \int_0^T \left(\int_{t-\delta}^t e^{-\rho(t-r)} \xi(r) dr \right) \frac{\partial H}{\partial a}(t) dt + \int_0^T \frac{\partial H}{\partial u}(t)\beta(t) dt \right]
 \end{aligned}$$

$$\begin{aligned}
 &= \mathbb{E} \left[\int_0^T \xi(t) \left\{ \frac{\partial H}{\partial x}(t) - \frac{\partial H}{\partial x}(t) - \frac{\partial H}{\partial y}(t + \delta) \mathbf{1}_{[0, T-\delta]}(t) \right. \right. \\
 &\quad \left. \left. - e^{\rho t} \left(\int_t^{t+\delta} \frac{\partial H}{\partial a}(s) e^{-\rho s} \mathbf{1}_{[0, T]}(s) ds \right) \right\} dt + \int_0^T \xi(t - \delta) \frac{\partial H}{\partial y}(t) dt \right. \\
 &\quad \left. + \int_0^T \left(\int_{s-\delta}^s e^{-\rho(s-t)} \xi(t) dt \right) \frac{\partial H}{\partial a}(s) ds + \int_0^T \frac{\partial H}{\partial u}(t) \beta(t) dt \right] \\
 &= \mathbb{E} \left[\int_0^T \xi(t) \left\{ -\frac{\partial H}{\partial y}(t + \delta) \mathbf{1}_{[0, T-\delta]}(t) - e^{\rho t} \left(\int_t^{t+\delta} \frac{\partial H}{\partial a}(s) e^{-\rho s} \mathbf{1}_{[0, T]}(s) ds \right) \right\} dt \right. \\
 &\quad \left. + \int_0^T \xi(t - \delta) \frac{\partial H}{\partial y}(t) dt + e^{\rho t} \int_0^T \left(\int_t^{t+\delta} \frac{\partial H}{\partial a}(s) e^{-\rho s} \mathbf{1}_{[0, T]}(s) ds \right) \xi(t) dt \right. \\
 &\quad \left. + \int_0^T \frac{\partial H}{\partial u}(t) \beta(t) dt \right] \\
 &= \mathbb{E} \left[\int_0^T \frac{\partial H}{\partial u}(t) \beta(t) dt \right], \tag{4.6}
 \end{aligned}$$

where again we have used integration by parts.

If we apply (4.6) to

$$\beta(t) = \alpha(\omega) \mathbf{1}_{[s, T]}(t),$$

where $\alpha(\omega)$ is bounded and \mathcal{E}_{t_0} -measurable, $s \geq t_0$, we obtain

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

Differentiating with respect to s we obtain

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

Since this holds for all $s \geq t_0$ and all α , we conclude that

$$\mathbb{E} \left[\frac{\partial}{\partial u} H(t_0) \mid \mathcal{E}_{t_0} \right] = 0.$$

This shows that assertion (i) implies assertion (ii).

Conversely, since every bounded $\beta \in \mathcal{A}_\mathcal{E}$ can be approximated by linear combinations of controls β of the form (4.1), we can prove that assertion (ii) implies assertion (i) by reversing the above argument.

5. Existence and uniqueness theorems for time-advanced BSDEs with jumps

We now study time-advanced BSDEs driven both by Brownian motion $B(t)$ and compensated Poisson random measures $\tilde{N}(dt, dz)$. We first provide a constructive procedure to compute the solution of time-advanced BSDEs of the form (2.3)-(2.4), typically satisfied by the adjoint processes of the Hamiltonian. Then we turn to more general BSDEs and provide several theorems on the existence and uniqueness of the solutions under different sets of assumptions on the driver and the terminal condition, which require different treatments in the proof.

5.1. Time-advanced BSDEs for adjoint processes

We introduce the following framework.

Given a positive constant δ , denote by $D([0, \delta], \mathbb{R})$ the space of all càdlàg paths from $[0, \delta]$ into \mathbb{R} . For a path $X(\cdot): \mathbb{R}_+ \rightarrow \mathbb{R}$, X_t will denote the function defined by $X_t(s) = X(t + s)$ for $s \in [0, \delta]$. Set $\mathcal{H} = L^2(\nu)$. Consider the L^2 spaces $V_1 := L^2([0, \delta], ds)$ and $V_2 := L^2([0, \delta] \rightarrow \mathcal{H}, ds)$. Let

$$F: \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R} \times V_1 \times \mathbb{R} \times \mathbb{R} \times V_1 \times \mathcal{H} \times \mathcal{H} \times V_2 \times \Omega \rightarrow \mathbb{R}$$

be a predictable function. Introduce the following Lipschitz condition. There exists a constant C such that

$$\begin{aligned} &|F(t, p_1, p_2, p, q_1, q_2, q, r_1, r_2, r, \omega) - F(t, \bar{p}_1, \bar{p}_2, \bar{p}, \bar{q}_1, \bar{q}_2, \bar{q}, \bar{r}_1, \bar{r}_2, \bar{r}, \omega)| \\ &\leq C(|p_1 - \bar{p}_1| + |p_2 - \bar{p}_2| + |p - \bar{p}|_{V_1} + |q_1 - \bar{q}_1| + |q_2 - \bar{q}_2| + |q - \bar{q}|_{V_1} \\ &\quad + |r_1 - \bar{r}_1|_{\mathcal{H}} + |r_2 - \bar{r}_2|_{\mathcal{H}} + |r - \bar{r}|_{V_2}). \end{aligned} \tag{5.1}$$

Consider the following time-advanced BSDE in the unknown \mathcal{F}_t -adapted processes $(p(t), q(t), r(t, z))$:

$$\begin{aligned} dp(t) = &E[F(t, p(t), p(t + \delta) \mathbf{1}_{[0, T-\delta]}(t), p_t \mathbf{1}_{[0, T-\delta]}(t), q(t), q(t + \delta) \mathbf{1}_{[0, T-\delta]}(t), \\ &q_t \mathbf{1}_{[0, T-\delta]}(t), r(t), r(t + \delta) \mathbf{1}_{[0, T-\delta]}(t), r_t \mathbf{1}_{[0, T-\delta]}(t)) \mid \mathcal{F}_t] dt \\ &+ q(t) dB(t) + \int_{\mathbb{R}} r(t, z) \tilde{N}(dt, dz), \quad t \in [0, T], \end{aligned} \tag{5.2}$$

$$p(T) = G, \tag{5.3}$$

where G is a given \mathcal{F}_T -measurable random variable such that $E[G^2] < \infty$.

Note that the time-advanced BSDE (2.3)–(2.4) for the adjoint processes of the Hamiltonian is of this form. For this type of time-advanced BSDEs, we have the following result.

Theorem 5.1. *Assume that condition (5.1) is satisfied. Then the BSDE (5.2)–(5.3) has a unique solution $(p(t), q(t), r(t, z))$ such that*

$$E \left[\int_0^T \left\{ p^2(t) + q^2(t) + \int_{\mathbb{R}} r^2(t, z) \nu(dz) \right\} dt \right] < \infty. \tag{5.4}$$

Moreover, the solution can be found by inductively solving a sequence of BSDEs backwards as follows.

Step 0. In the interval $[T - \delta, T]$ we let $p(t), q(t)$, and $r(t, z)$ be defined as the solution of the classical BSDE

$$\begin{aligned} dp(t) = &F(t, p(t), 0, 0, q(t), 0, 0, r(t, z), 0, 0) dt + q(t) dB(t) \\ &+ \int_{\mathbb{R}} r(t, z) \tilde{N}(dt, dz), \quad t \in [T - \delta, T], \\ p(T) = &G. \end{aligned}$$

Step $k, k \geq 1$. If the values of $(p(t), q(t), r(t, z))$ have been found for $t \in [T - k\delta, T - (k - 1)\delta]$ then, if $t \in [T - (k + 1)\delta, T - k\delta]$, the values of $p(t + \delta), p_t, q(t + \delta), q_t, r(t + \delta, z)$,

and r_t are known and, hence, the BSDE

$$dp(t) = E[F(t, p(t), p(t + \delta), p_t, q(t), q(t + \delta), q_t, r(t), r(t + \delta), r_t) \mid \mathcal{F}_t] dt + q(t) dB(t) + \int_{\mathbb{R}} r(t, z) \tilde{N}(dt, dz); t \in [T - (k + 1)\delta, T - k\delta],$$

$p(T - k\delta) =$ the value found in step $k - 1$,

has a unique solution in $[T - (k + 1)\delta, T - k\delta]$.

We proceed like this until k is such that $T - (k + 1)\delta \leq 0 < T - k\delta$ and then we solve the corresponding BSDE on the interval $[0, T - k\delta]$.

Proof of Theorem 5.1. The proof follows directly from the above inductive procedure. The estimate (5.4) is a consequence of known estimates for classical BSDEs.

5.2. General time-advanced BSDEs

We now consider the following BSDE in the unknown \mathcal{F}_t -adapted processes $(p(t), q(t), r(t, x))$:

$$dp(t) = E[F(t, p(t), p(t + \delta), p_t, q(t), q(t + \delta), q_t, r(t), r(t + \delta), r_t) \mid \mathcal{F}_t] dt + q(t) dB_t + \int_{\mathbb{R}} r(t, z) \tilde{N}(dt, dz), \quad t \in [0, T], \tag{5.5}$$

$$p(t) = G(t), \quad t \in [T, T + \delta], \tag{5.6}$$

where G is a given continuous \mathcal{F}_t -adapted stochastic process. We shall present three theorems with various conditions on F and G .

Theorem 5.2. *Assume that $E[\sup_{T \leq t \leq T + \delta} |G(t)|^2] < \infty$ and that condition (5.1) is satisfied. Then the BSDE (5.5)–(5.6) admits a unique solution $(p(t), q(t), r(t, z))$ such that*

$$E \left[\int_0^T \left\{ p^2(t) + q^2(t) + \int_{\mathbb{R}} r^2(t, z) \nu(dz) \right\} dt \right] < \infty.$$

Proof. Step 1: assume that F is independent of $p_1, p_2,$ and p . Set $q^0(t) := 0$ and $r^0(t, x) = 0$. For $n \geq 1$, define $(p^n(t), q^n(t), r^n(t, x))$ to be the unique solution to the following BSDE:

$$dp^n(t) = E[F(t, q^{n-1}(t), q^{n-1}(t + \delta), q_t^{n-1}, r^{n-1}(t, \cdot), r^{n-1}(t + \delta, \cdot), r_t^{n-1}(\cdot)) \mid \mathcal{F}_t] dt + q^n(t) dB_t + r^n(t, z) \tilde{N}(dt, dz), \quad t \in [0, T], \tag{5.7}$$

$p^n(t) = G(t), \quad t \in [T, T + \delta].$

It is a consequence of the martingale representation theorem that the above equation admits a unique solution; see, e.g. [15] and [18]. We extend q^n and r^n to $[0, T + \delta]$ by setting $q^n(s) = 0$ and $r^n(s, z) = 0$ for $T \leq s \leq T + \delta$. We are going to show that $(p^n(t), q^n(t), r^n(t, x))$ forms

a Cauchy sequence. By Itô's formula we have

$$\begin{aligned}
 0 &= |p^{n+1}(T) - p^n(T)|^2 \\
 &= |p^{n+1}(t) - p^n(t)|^2 \\
 &\quad + 2 \int_t^T (p^{n+1}(s) - p^n(s)) \\
 &\quad \quad \times (\mathbb{E}[F(s, q^n(s), q^n(s + \delta), q_s^n, r^n(s, \cdot), r^n(s + \delta, \cdot), r_s^n(\cdot)) \mid \mathcal{F}_s] \\
 &\quad \quad - \mathbb{E}[F(s, q^{n-1}(s), q^{n-1}(s + \delta), q_s^{n-1}, r^{n-1}(s, \cdot), r^{n-1}(s + \delta, \cdot), \\
 &\quad \quad \quad r_s^{n-1}(\cdot)) \mid \mathcal{F}_s]) ds \\
 &\quad + \int_t^T \int_{\mathbb{R}} |r^{n+1}(s, z) - r^n(s, z)|^2 ds \nu(dz) + \int_t^T |q^{n+1}(s) - q^n(s)|^2 ds \\
 &\quad + 2 \int_t^T (p^{n+1}(s) - p^n(s))(q^{n+1}(s) - q^n(s)) dB_s \\
 &\quad + \int_t^T \int_{\mathbb{R}} \{|r^{n+1}(s, z) - r^n(s, z)|^2 \\
 &\quad \quad + 2(p^{n+1}(s-) - p^n(s-))(r^{n+1}(s, z) - r^n(s, z))\} \tilde{N}(ds, dz). \tag{5.8}
 \end{aligned}$$

Rearranging terms, in view of (5.1), we obtain

$$\begin{aligned}
 &\mathbb{E}[|p^{n+1}(t) - p^n(t)|^2] + \mathbb{E}\left[\int_t^T \int_{\mathbb{R}} |r^{n+1}(s, z) - r^n(s, z)|^2 ds \nu(dz)\right] \\
 &+ \mathbb{E}\left[\int_t^T |q^{n+1}(s) - q^n(s)|^2 ds\right] \\
 &\leq 2 \mathbb{E}\left[\int_t^T |(p^{n+1}(s) - p^n(s))\right. \\
 &\quad \times \mathbb{E}[F(s, q^n(s), q^n(s + \delta), r^n(s, \cdot), r^n(s + \delta, \cdot)) \\
 &\quad \quad \left. - F(s, q^{n-1}(s), q^{n-1}(s + \delta), r^{n-1}(s, \cdot), r^{n-1}(s + \delta, \cdot)) \mid \mathcal{F}_s] ds\right] \\
 &\leq C_\varepsilon \mathbb{E}\left[\int_t^T |p^{n+1}(s) - p^n(s)|^2 ds\right] + \varepsilon \mathbb{E}\left[\int_t^T |q^n(s) - q^{n-1}(s)|^2 ds\right] \\
 &\quad + \varepsilon \mathbb{E}\left[\int_t^T |q^n(s + \delta) - q^{n-1}(s + \delta)|^2 ds\right] \\
 &\quad + \varepsilon \mathbb{E}\left[\int_t^T \left(\int_s^{s+\delta} |q^n(u) - q^{n-1}(u)|^2 du\right) ds\right] \\
 &\quad + \varepsilon \mathbb{E}\left[\int_t^T |r^n(s) - r^{n-1}(s)|_{\mathcal{H}}^2 ds\right] + \varepsilon \mathbb{E}\left[\int_t^T |r^n(s + \delta) - r^{n-1}(s + \delta)|_{\mathcal{H}}^2 ds\right] \\
 &\quad + \varepsilon \mathbb{E}\left[\int_t^T \left(\int_s^{s+\delta} |r^n(u) - r^{n-1}(u)|_{\mathcal{H}}^2 du\right) ds\right]. \tag{5.9}
 \end{aligned}$$

Note that

$$\mathbb{E} \left[\int_t^T |q^n(s + \delta) - q^{n-1}(s + \delta)|^2 ds \right] \leq \mathbb{E} \left[\int_t^T |q^n(s) - q^{n-1}(s)|^2 ds \right].$$

Interchanging the order of integration,

$$\begin{aligned} \mathbb{E} \left[\int_t^T \left(\int_s^{s+\delta} |q^n(u) - q^{n-1}(u)|^2 du \right) ds \right] &= \mathbb{E} \left[\int_t^{T+\delta} |q^n(u) - q^{n-1}(u)|^2 du \int_{u-\delta}^u ds \right] \\ &\leq \delta \mathbb{E} \left[\int_t^T |q^n(s) - q^{n-1}(s)|^2 ds \right]. \end{aligned}$$

Similar inequalities hold also for $r^n - r^{n-1}$. It follows from (5.9) that

$$\begin{aligned} &\mathbb{E}[|p^{n+1}(t) - p^n(t)|^2] + \mathbb{E} \left[\int_t^T \int_{\mathbb{R}} |r^{n+1}(s, z) - r^n(s, z)|^2 ds \nu(dz) \right] \\ &+ \mathbb{E} \left[\int_t^T |q^{n+1}(s) - q^n(s)|^2 ds \right] \\ &\leq C_\varepsilon \mathbb{E} \left[\int_t^T |p^{n+1}(s) - p^n(s)|^2 ds \right] + (2 + M)\varepsilon \mathbb{E} \left[\int_t^T |q^n(s) - q^{n-1}(s)|^2 ds \right] \\ &\quad + 3\varepsilon \mathbb{E} \left[\int_t^T |r^n(s) - r^{n-1}(s)|_{\mathcal{H}}^2 ds \right]. \end{aligned} \tag{5.10}$$

Choose $\varepsilon > 0$ sufficiently small so that

$$\begin{aligned} &\mathbb{E}[|p^{n+1}(t) - p^n(t)|^2] + \mathbb{E} \left[\int_t^T \int_{\mathbb{R}} |r^{n+1}(s, z) - r^n(s, z)|^2 ds \nu(dz) \right] \\ &+ \mathbb{E} \left[\int_t^T |q^{n+1}(s) - q^n(s)|^2 ds \right] \\ &\leq C_\varepsilon \mathbb{E} \left[\int_t^T |p^{n+1}(s) - p^n(s)|^2 ds \right] + \frac{1}{2} \mathbb{E} \left[\int_t^T |q^n(s) - q^{n-1}(s)|^2 ds \right] \\ &\quad + \frac{1}{2} \mathbb{E} \left[\int_t^T |r^n(s) - r^{n-1}(s)|_{\mathcal{H}}^2 ds \right]. \end{aligned}$$

This implies that

$$\begin{aligned} &-\frac{d}{dt} \left(e^{C_\varepsilon t} \mathbb{E} \left[\int_t^T |p^{n+1}(s) - p^n(s)|^2 ds \right] \right) \\ &+ e^{C_\varepsilon t} \mathbb{E} \left[\int_t^T \int_{\mathbb{R}} |r^{n+1}(s, z) - r^n(s, z)|^2 ds \nu(dz) \right] + e^{C_\varepsilon t} \mathbb{E} \left[\int_t^T |q^{n+1}(s) - q^n(s)|^2 ds \right] \\ &\leq \frac{1}{2} e^{C_\varepsilon t} \mathbb{E} \left[\int_t^T |q^n(s) - q^{n-1}(s)|^2 ds \right] + \frac{1}{2} e^{C_\varepsilon t} \mathbb{E} \left[\int_t^T |r^n(s) - r^{n-1}(s)|_{\mathcal{H}}^2 ds \right]. \end{aligned}$$

Integrating the last inequality yields

$$\begin{aligned}
 & \mathbb{E} \left[\int_0^T |p^{n+1}(s) - p^n(s)|^2 ds \right] + \int_0^T dt e^{C_\varepsilon t} \mathbb{E} \left[\int_t^T |q^{n+1}(s) - q^n(s)|^2 ds \right] \\
 & + \int_0^T dt e^{C_\varepsilon t} \mathbb{E} \left[\int_t^T \int_{\mathbb{R}} |r^{n+1}(s, z) - r^n(s, z)|^2 ds \nu(dz) \right] \\
 & \leq \frac{1}{2} \int_0^T dt e^{C_\varepsilon t} \mathbb{E} \left[\int_t^T |q^n(s) - q^{n-1}(s)|^2 ds \right] \\
 & + \frac{1}{2} \int_0^T dt e^{C_\varepsilon t} \mathbb{E} \left[\int_t^T |r^n(s) - r^{n-1}(s)|_{\mathcal{H}}^2 ds \right].
 \end{aligned} \tag{5.11}$$

In particular,

$$\begin{aligned}
 & \int_0^T dt e^{C_\varepsilon t} \mathbb{E} \left[\int_t^T \int_{\mathbb{R}} |r^{n+1}(s, z) - r^n(s, z)|^2 ds \nu(dz) \right] \\
 & + \int_0^T dt e^{C_\varepsilon t} \mathbb{E} \left[\int_t^T |q^{n+1}(s) - q^n(s)|^2 ds \right] \\
 & \leq \frac{1}{2} \int_0^T dt e^{C_\varepsilon t} \mathbb{E} \left[\int_t^T |q^n(s) - q^{n-1}(s)|^2 ds \right] \\
 & + \frac{1}{2} \int_0^T dt e^{C_\varepsilon t} \mathbb{E} \left[\int_t^T |r^n(s) - r^{n-1}(s)|_{\mathcal{H}}^2 ds \right].
 \end{aligned} \tag{5.12}$$

This yields

$$\begin{aligned}
 & \int_0^T dt e^{C_\varepsilon t} \mathbb{E} \left[\int_t^T \int_{\mathbb{R}} |r^{n+1}(s, z) - r^n(s, z)|^2 ds \nu(dz) \right] \\
 & + \int_0^T dt e^{C_\varepsilon t} \mathbb{E} \left[\int_t^T |q^{n+1}(s) - q^n(s)|^2 ds \right] \\
 & \leq \left(\frac{1}{2}\right)^n C
 \end{aligned}$$

for some constant C . It follows from (5.11) that

$$\mathbb{E} \left[\int_0^T |p^{n+1}(s) - p^n(s)|^2 ds \right] \leq \left(\frac{1}{2}\right)^n C.$$

Equation (5.10) and (5.11) further give

$$\begin{aligned}
 & \mathbb{E} \left[\int_0^T \int_{\mathbb{R}} |r^{n+1}(s, z) - r^n(s, z)|^2 ds \nu(dz) \right] + \mathbb{E} \left[\int_0^T |q^{n+1}(s) - q^n(s)|^2 ds \right] \\
 & \leq \left(\frac{1}{2}\right)^n C n C_\varepsilon.
 \end{aligned}$$

In view of (5.10), (5.11), and (5.12), we conclude that there exist progressively measurable

processes $(p(t), q(t), r(t, z))$ such that

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{E}[|p^n(t) - p(t)|^2] &= 0, \\ \lim_{n \rightarrow \infty} \int_0^T \mathbb{E}[|p^n(t) - p(t)|^2] dt &= 0, \\ \lim_{n \rightarrow \infty} \int_0^T \mathbb{E}[|q^n(t) - q(t)|^2] dt &= 0, \\ \lim_{n \rightarrow \infty} \int_0^T \int_{\mathbb{R}} \mathbb{E}[|r^n(t, z) - r(t, z)|^2] \nu(dz) dt &= 0. \end{aligned}$$

Letting $n \rightarrow \infty$ in (5.7) we see that $(p(t), q(t), r(t, z))$ satisfies

$$\begin{aligned} p(t) + \int_t^T \mathbb{E}[F(s, q(s), q(s + \delta), q_s, r(s, \cdot), r(s + \delta, \cdot), r_s(\cdot)) \mid \mathcal{F}_s] ds \\ + \int_t^T q(s) dB_s + \int_t^T \int_{\mathbb{R}} r(s, z) \tilde{N}(ds, dz) = g(T), \end{aligned}$$

i.e. $(p(t), q(t), r(t, z))$ is a solution. Uniqueness follows easily from Itô’s formula, a similar calculation used to deduce (5.8) and (5.9), and Gronwall’s lemma.

Step 2: general case. Let $p^0(t) = 0$. For $n \geq 1$, define $(p^n(t), q^n(t), r^n(t, z))$ to be the unique solution to the following BSDE:

$$\begin{aligned} dp^n(t) &= \mathbb{E}[F(t, p^{n-1}(t), p^{n-1}(t + \delta), p_t^{n-1}, q^n(t), q^n(t + \delta), q_t^n, r^n(t, \cdot), \\ &\quad r^n(t + \delta, \cdot), r_t^n(\cdot)) \mid \mathcal{F}_t] dt + q^n(t) dB_t + r^n(t, z) \tilde{N}(dt, dz), \quad t \in [0, T], \\ p^n(t) &= G(t), \quad t \in [T, T + \delta]. \end{aligned}$$

The existence of $(p^n(t), q^n(t), r^n(t, z))$ is proved in step 1. By the same arguments leading to (5.10), we deduce that

$$\begin{aligned} \mathbb{E}[|p^{n+1}(t) - p^n(t)|^2] + \frac{1}{2} \mathbb{E} \left[\int_t^T \int_{\mathbb{R}} |r^{n+1}(s, z) - r^n(s, z)|^2 ds \nu(dz) \right] \\ + \frac{1}{2} \mathbb{E} \left[\int_t^T |q^{n+1}(s) - q^n(s)|^2 ds \right] \\ \leq C \mathbb{E} \left[\int_t^T |p^{n+1}(s) - p^n(s)|^2 ds \right] + \frac{1}{2} \mathbb{E} \left[\int_t^T |p^n(s) - p^{n-1}(s)|^2 ds \right]. \end{aligned}$$

This implies that

$$-\frac{d}{dt} \left(e^{Ct} \mathbb{E} \left[\int_t^T |p^{n+1}(s) - p^n(s)|^2 ds \right] \right) \leq \frac{1}{2} e^{Ct} \mathbb{E} \left[\int_t^T |p^n(s) - p^{n-1}(s)|^2 ds \right]. \tag{5.13}$$

Integrating (5.13) from u to T we obtain

$$\begin{aligned} \mathbb{E} \left[\int_u^T |p^{n+1}(s) - p^n(s)|^2 ds \right] &\leq \frac{1}{2} \int_u^T dt e^{C(t-u)} \mathbb{E} \left[\int_t^T |p^n(s) - p^{n-1}(s)|^2 ds \right] \\ &\leq e^{CT} \int_u^T dt \mathbb{E} \left[\int_t^T |p^n(s) - p^{n-1}(s)|^2 ds \right]. \end{aligned}$$

Iterating the above inequality we obtain

$$\mathbb{E} \left[\int_0^T |p^{n+1}(s) - p^n(s)|^2 ds \right] \leq \frac{e^{CnT} T^n}{n!}.$$

Using the above inequality and a similar argument as in step 1, it can be shown that $(p^n(t), q^n(t), r^n(t, z))$ converges to some limit $(p(t), q(t), r(t, z))$, which is the unique solution of (5.5).

Theorem 5.3. Assume that $\mathbb{E}[\sup_{T \leq t \leq T+\delta} |G(t)|^{2\alpha}] < \infty$ for some $\alpha > 1$ and that the following condition holds:

$$\begin{aligned} & |F(t, p_1, p_2, p, q_1, q_2, q, r_1, r_2, r) - F(t, \bar{p}_1, \bar{p}_2, \bar{p}, \bar{q}_1, \bar{q}_2, \bar{q}, \bar{r}_1, \bar{r}_2, \bar{r})| \\ & \leq C \left(|p_1 - \bar{p}_1| + |p_2 - \bar{p}_2| + \sup_{0 \leq s \leq \delta} |p(s) - \bar{p}(s)| + |q_1 - \bar{q}_1| + |q_2 - \bar{q}_2| \right. \\ & \quad \left. + |q - \bar{q}|_{V_1} + |r_1 - \bar{r}_1|_{\mathcal{H}} + |r_2 - \bar{r}_2|_{\mathcal{H}} + |r - \bar{r}|_{V_2} \right). \end{aligned} \tag{5.14}$$

Then the BSDE (5.5) admits a unique solution $(p(t), q(t), r(t, z))$ such that

$$\mathbb{E} \left[\sup_{0 \leq t \leq T} |p(t)|^{2\alpha} + \int_0^T \left\{ q^2(t) + \int_{\mathbb{R}} r^2(t, z) \nu(dz) \right\} dt \right] < \infty.$$

Proof. Step 1: assume that F is independent of p_1, p_2 , and p . In this case condition (5.14) reduces to assumption (5.1). By step 1 in the proof of Theorem 5.2, there is a unique solution $(p(t), q(t), r(t, z))$ to (5.5).

Step 2: general case. Let $p^0(t) = 0$. For $n \geq 1$, define $(p^n(t), q^n(t), r^n(t, z))$ to be the unique solution to the following BSDE:

$$\begin{aligned} dp^n(t) &= \mathbb{E}[F(t, p^{n-1}(t), p^{n-1}(t + \delta), p_t^{n-1}, q^n(t), q^n(t + \delta), q_t^n, r^n(t, \cdot), r^n(t + \delta, \cdot), \\ & \quad r_t^n(\cdot) \mid \mathcal{F}_t] dt + q^n(t) dB_t + r^n(t, z) \tilde{N}(dt, dz), \\ p^n(t) &= G(t), \quad t \in [T, T + \delta]. \end{aligned}$$

By step 1, $(p^n(t), q^n(t), r^n(t, z))$ exists. We are going to show that $(p^n(t), q^n(t), r^n(t, z))$ forms a Cauchy sequence. Using Itô's formula, we have

$$\begin{aligned} & |p^{n+1}(t) - p^n(t)|^2 + \int_t^T \int_{\mathbb{R}} |r^{n+1}(s, z) - r^n(s, z)|^2 ds \nu(dz) + \int_t^T |q^{n+1}(s) - q^n(s)|^2 ds \\ & = -2 \int_t^T (p^{n+1}(s) - p^n(s)) \\ & \quad \times [\mathbb{E}[F(s, p^n(s), p^n(s + \delta), p_s^n, q^{n+1}(s), q^{n+1}(s + \delta), q_s^{n+1}, r^{n+1}(s, \cdot), \\ & \quad r^{n+1}(s + \delta, \cdot), r_s^{n+1}(\cdot)) \\ & \quad - F(s, p^{n-1}(s), p^{n-1}(s + \delta), p_s^{n-1}, q^n(s), q^n(s + \delta), q_s^n, r^n(s, \cdot), \\ & \quad r^n(s + \delta, \cdot), r_s^n(\cdot) \mid \mathcal{F}_s]] ds \end{aligned}$$

$$\begin{aligned}
 & - 2 \int_t^T (p^{n+1}(s) - p^n(s))(q^{n+1}(s) - q^n(s)) dB_s \\
 & - \int_t^T \int_{\mathbb{R}} [|r^{n+1}(s, z) - r^n(s, z)|^2 \\
 & \quad + 2(p^{n+1}(s-) - p^n(s-))(r^{n+1}(s, z) - r^n(s, z))] \tilde{N}(ds, dz).
 \end{aligned}$$

Take the conditional expectation with respect to \mathcal{F}_t , take the supremum over the interval $[u, T]$, and use condition (5.14) to obtain

$$\begin{aligned}
 & \sup_{u \leq t \leq T} |p^{n+1}(t) - p^n(t)|^2 + \sup_{u \leq t \leq T} \mathbb{E} \left[\int_t^T |q^{n+1}(s) - q^n(s)|^2 ds \mid \mathcal{F}_t \right] \\
 & + \sup_{u \leq t \leq T} \mathbb{E} \left[\int_t^T \int_{\mathbb{R}} |r^{n+1}(s, z) - r^n(s, z)|^2 ds \nu(dz) \mid \mathcal{F}_t \right] \\
 & \leq C_\varepsilon \sup_{u \leq t \leq T} \mathbb{E} \left[\int_u^T |p^{n+1}(s) - p^n(s)|^2 ds \mid \mathcal{F}_t \right] \\
 & \quad + C_1 \varepsilon \sup_{u \leq t \leq T} \mathbb{E} \left[\int_u^T |p^n(s) - p^{n-1}(s)|^2 ds \mid \mathcal{F}_t \right] \\
 & \quad + C_2 \varepsilon \sup_{u \leq t \leq T} \mathbb{E} \left[\int_u^T \mathbb{E} \left[\sup_{s \leq v \leq T} |p^n(v) - p^{n-1}(v)|^2 \mid \mathcal{F}_s \right] ds \mid \mathcal{F}_t \right] \\
 & \quad + C_3 \varepsilon \sup_{u \leq t \leq T} \mathbb{E} \left[\int_t^T |q^{n+1}(s) - q^n(s)|^2 ds \mid \mathcal{F}_t \right] \\
 & \quad + C_4 \varepsilon \sup_{u \leq t \leq T} \mathbb{E} \left[\int_t^T \int_{\mathbb{R}} |r^{n+1}(s, z) - r^n(s, z)|^2 ds \nu(dz) \mid \mathcal{F}_t \right]. \tag{5.15}
 \end{aligned}$$

Choosing $\varepsilon > 0$ such that $C_3 \varepsilon < 1$ and $C_4 \varepsilon < 1$, it follows from (5.15) that

$$\begin{aligned}
 & \sup_{u \leq t \leq T} |p^{n+1}(t) - p^n(t)|^2 \\
 & \leq C_\varepsilon \sup_{u \leq t \leq T} \mathbb{E} \left[\int_u^T |p^{n+1}(s) - p^n(s)|^2 ds \mid \mathcal{F}_t \right] \\
 & \quad + (C_1 + C_2) \varepsilon \sup_{u \leq t \leq T} \mathbb{E} \left[\int_u^T \mathbb{E} \left[\sup_{s \leq v \leq T} |p^n(v) - p^{n-1}(v)|^2 \mid \mathcal{F}_s \right] ds \mid \mathcal{F}_t \right]. \tag{5.16}
 \end{aligned}$$

Note that

$$\mathbb{E} \left[\int_u^T |p^{n+1}(s) - p^n(s)|^2 ds \mid \mathcal{F}_t \right]$$

and

$$\mathbb{E} \left[\int_u^T \mathbb{E} \left[\sup_{s \leq v \leq T} |p^n(v) - p^{n-1}(v)|^2 \mid \mathcal{F}_s \right] ds \mid \mathcal{F}_t \right]$$

are right-continuous martingales on $[0, T]$ with terminal random variables $\int_u^T |p^{n+1}(s) - p^n(s)|^2 ds$ and $\int_u^T E[\sup_{s \leq v \leq T} |p^n(v) - p^{n-1}(v)|^2 | \mathcal{F}_s] ds$. Thus, for $\alpha > 1$, we have

$$\begin{aligned} & E \left[\left(\sup_{u \leq t \leq T} E \left[\int_u^T |p^{n+1}(s) - p^n(s)|^2 ds \mid \mathcal{F}_t \right] \right)^\alpha \right] \\ & \leq c_\alpha E \left[\left(\int_u^T |p^{n+1}(s) - p^n(s)|^2 ds \right)^\alpha \right] \\ & \leq c_{T,\alpha} E \left[\int_u^T \sup_{s \leq v \leq T} |p^{n+1}(v) - p^n(v)|^{2\alpha} ds \right], \end{aligned} \tag{5.17}$$

and

$$\begin{aligned} & E \left[\left(\sup_{u \leq t \leq T} E \left[\int_u^T E \left[\sup_{s \leq v \leq T} |p^n(v) - p^{n-1}(v)|^2 \mid \mathcal{F}_s \right] ds \mid \mathcal{F}_t \right] \right)^\alpha \right] \\ & \leq c_{T,\alpha} E \left[\int_u^T E \left[\sup_{s \leq v \leq T} |p^n(v) - p^{n-1}(v)|^{2\alpha} \mid \mathcal{F}_s \right] ds \right] \\ & \leq c_{T,\alpha} E \left[\int_u^T \sup_{s \leq v \leq T} |p^n(v) - p^{n-1}(v)|^{2\alpha} ds \right]. \end{aligned} \tag{5.18}$$

Equations (5.16), (5.17), and (5.18) yield, for $\alpha > 1$,

$$\begin{aligned} E \left[\sup_{u \leq t \leq T} |p^{n+1}(t) - p^n(t)|^{2\alpha} \right] & \leq C_{1,\alpha} E \left[\int_u^T \sup_{s \leq v \leq T} |p^{n+1}(v) - p^n(v)|^{2\alpha} ds \right] \\ & \quad + C_{2,\alpha} E \left[\int_u^T \sup_{s \leq v \leq T} |p^n(v) - p^{n-1}(v)|^{2\alpha} ds \right]. \end{aligned} \tag{5.19}$$

Set

$$g_n(u) = E \left[\int_u^T \sup_{t \leq s \leq T} |p^n(s) - p^{n-1}(s)|^{2\alpha} dt \right].$$

Equation (5.19) implies that

$$-\frac{d}{dt} (e^{C_{1,\alpha}u} g_{n+1}(u)) \leq e^{C_{1,\alpha}u} C_{2,\alpha} g_n(u). \tag{5.20}$$

Integrating (5.20) from t to T we obtain

$$g_{n+1}(t) \leq c_{2,\alpha} \int_t^T e^{C_{1,\alpha}(s-t)} g_n(s) ds \leq C_{2,\alpha} e^{C_{1,\alpha}T} \int_t^T g_n(s) ds.$$

Iterating the above inequality we obtain

$$E \left[\int_0^T \sup_{t \leq s \leq T} |p^{n+1}(s) - p^n(s)|^{2\alpha} dt \right] \leq \frac{e^{CnT} T^n}{n!}.$$

Using the above inequality and a similar argument as in step 1, we can show that $(p^n(t), q^n(t), r^n(t, z))$ converges to some limit $(p(t), q(t), r(t, z))$, which is the unique solution of (5.5).

Finally, we present a result when the function F is independent of q_1, q_2, q, r_1, r_2 , and r .

Theorem 5.4. Assume that $E[\sup_{T \leq t \leq T+\delta} |G(t)|^2] < \infty$ and F satisfies (5.14), i.e.

$$|F(t, p_1, p_2, p) - F(t, \bar{p}_1, \bar{p}_2, \bar{p})| \leq C \left(|p_1 - \bar{p}_1| + |p_2 - \bar{p}_2| + \sup_{0 \leq s \leq \delta} |p(s) - \bar{p}(s)| \right). \tag{5.21}$$

Then the BSDE (5.5) admits a unique solution such that (5.4) holds.

Proof. Let $p^0(t) = 0$. For $n \geq 1$, define $(p^n(t), q^n(t), r^n(t, z))$ to be the unique solution to the following BSDE:

$$\begin{aligned} dp^n(t) &= E[F(t, p^{n-1}(t), p^{n-1}(t + \delta), p_t^{n-1}) \mid \mathcal{F}_t] dt + q^n(t) dB_t + r^n(t, z) \tilde{N}(dt, dz), \\ p^n(t) &= G(t), \quad t \in [T, T + \delta]. \end{aligned}$$

We will show that $(p^n(t), q^n(t), r^n(t, z))$ forms a Cauchy sequence. Subtracting p^n from p^{n+1} and taking the conditional expectation with respect to \mathcal{F}_t , we obtain

$$\begin{aligned} p^{n+1}(t) - p^n(t) &= -E \left[\int_t^T (E[F(s, p^n(s), p^n(s + \delta), p_s^n) \mid \mathcal{F}_s] \right. \\ &\quad \left. - E[F(s, p^{n-1}(s), p^{n-1}(s + \delta), p_s^{n-1}) \mid \mathcal{F}_s]) ds \mid \mathcal{F}_t \right]. \end{aligned}$$

Take the supremum over the interval $[u, T]$ and use assumption (5.21) to obtain

$$\begin{aligned} &\sup_{u \leq t \leq T} |p^{n+1}(t) - p^n(t)|^2 \\ &\leq C \sup_{u \leq t \leq T} \left(E \left[\int_u^T |p^n(s) - p^{n-1}(s)| ds \mid \mathcal{F}_t \right] \right)^2 \\ &\quad + C \sup_{u \leq t \leq T} \left(E \left[\int_u^T E \left[\sup_{s \leq v \leq T} |p^n(v) - p^{n-1}(v)| \mid \mathcal{F}_s \right] ds \mid \mathcal{F}_t \right] \right)^2. \tag{5.22} \end{aligned}$$

By the martingale inequality we have

$$\begin{aligned} &E \left[\left(\sup_{u \leq t \leq T} E \left[\int_u^T |p^n(s) - p^{n-1}(s)| ds \mid \mathcal{F}_t \right] \right)^2 \right] \\ &\leq c E \left[\left(\int_u^T |p^n(s) - p^{n-1}(s)| ds \right)^2 \right] \\ &\leq c_T E \left[\int_u^T \sup_{s \leq v \leq T} |p^n(v) - p^{n-1}(v)|^2 ds \right], \end{aligned}$$

and

$$\begin{aligned} &E \left[\left(\sup_{u \leq t \leq T} E \left[\int_u^T E \left[\sup_{s \leq v \leq T} |p^n(v) - p^{n-1}(v)| \mid \mathcal{F}_s \right] ds \mid \mathcal{F}_t \right] \right)^2 \right] \\ &\leq c_T E \left[\int_u^T E \left[\sup_{s \leq v \leq T} |p^n(v) - p^{n-1}(v)|^2 \mid \mathcal{F}_s \right] ds \right]. \end{aligned}$$

Taking the expectation on both sides of (5.22) gives

$$E \left[\sup_{u \leq t \leq T} |p^{n+1}(t) - p^n(t)|^2 \right] \leq C \int_u^T E \left[\sup_{s \leq v \leq T} |p^n(v) - p^{n-1}(v)|^2 ds \right].$$

It follows easily from here that $(p^n(t), q^n(t), r^n(t, z))$ converges to some limit $(p(t), q(t), r(t, z))$, which is the unique solution of (5.5).

6. Application to optimal consumption from a cash flow with delay

Let $\alpha(t)$, $\beta(t)$, and $\gamma(t, z)$ be given bounded adapted processes, with $\alpha(t)$ assumed to be deterministic. Suppose that $\int_{\mathbb{R}} \gamma^2(t, z)v(dz) < \infty$. Consider a cash flow $X^0(t)$ with dynamics

$$dX^0(t) = X^0(t - \delta) \left[\alpha(t) dt + \beta(t) dB(t) + \int_{\mathbb{R}} \gamma(t, z) \tilde{N}(dt, dz) \right], \quad t \in [0, T],$$

$$X^0(t) = x_0(t) > 0, \quad t \in [-\delta, 0],$$

where $x_0(t)$ is a given bounded deterministic function.

Suppose that at time $t \in [0, T]$ we consume at the rate $c(t) \geq 0$, a càdlàg adapted process. Then the dynamics of the corresponding net cash flow $X(t) = X^c(t)$ are

$$dX(t) = [X(t - \delta)\alpha(t) - c(t)] dt + X(t - \delta)\beta(t) dB(t) + X(t - \delta) \int_{\mathbb{R}} \gamma(t, z) \tilde{N}(dt, dz), \quad t \in [0, T], \tag{6.1}$$

$$X(t) = x_0(t), \quad t \in [-\delta, 0]. \tag{6.2}$$

Let $U_1(t, c, \omega) : [0, T] \times \mathbb{R}^+ \times \Omega \rightarrow \mathbb{R}$ be a given stochastic utility function satisfying the following conditions:

$$t \rightarrow U_1(t, c, \omega) \text{ is } \mathcal{F}_t\text{-adapted for each } c \geq 0,$$

$$c \rightarrow U_1(t, c, \omega) \text{ is } \mathcal{C}^1, \quad \frac{\partial}{\partial c} U_1(t, c, \omega) > 0,$$

$$c \rightarrow \frac{\partial}{\partial c} U_1(t, c, \omega) \text{ is strictly decreasing,}$$

$$\lim_{c \rightarrow \infty} \frac{\partial}{\partial c} U_1(t, c, \omega) = 0 \quad \text{for all } t, \omega \in [0, T] \times \Omega.$$

Set $v_0(t, \omega) = \partial U_1(t, 0, \omega) / \partial c$, and define

$$I(t, v, \omega) = \begin{cases} 0 & \text{if } v \geq v_0(t, \omega), \\ \left(\frac{\partial}{\partial c} U_1(t, \cdot, \omega) \right)^{-1}(v) & \text{if } 0 \leq v < v_0(t, \omega). \end{cases}$$

Suppose that we want to find the consumption rate $\hat{c}(t)$ such that

$$J(\hat{c}) = \sup\{J(c), c \in \mathcal{A}\}, \tag{6.3}$$

where

$$J(c) = \mathbb{E} \left[\int_0^T U_1(t, c(t), \omega) dt + g(X(T)) \right].$$

Here $g : \mathbb{R} \rightarrow \mathbb{R}$ is a given concave \mathcal{C}^1 function and \mathcal{A} is the family of all càdlàg, \mathcal{F}_t -adapted processes $c(t) \geq 0$ such that $\mathbb{E}[|g(X(T))|] < \infty$.

In this case the Hamiltonian given by (2.1) takes the form

$$H(t, x, y, a, u, p, q, r(\cdot), \omega) = U_1(t, c, \omega) + (\alpha(t)y - c)p + y\beta(t)q + y \int_{\mathbb{R}} \gamma(t, z)r(z)v(dz).$$

Maximizing H with respect to c gives the following first-order condition for an optimal $\hat{c}(t)$:

$$\frac{\partial}{\partial c} U_1(t, \hat{c}(t), \omega) = p(t). \tag{6.4}$$

The time-advanced BSDE for $p(t)$, $q(t)$, and $r(t, z)$ is, by (2.3)–(2.4),

$$\begin{aligned} dp(t) = & -\mathbb{E} \left[\left\{ \alpha(t)p(t + \delta) + \beta(t)q(t + \delta) + \int_{\mathbb{R}} \gamma(t, z)r(t + \delta, z)v(dz) \right\} \right. \\ & \left. \times \mathbf{1}_{[0, T - \delta]}(t) \mid \mathcal{F}_t \right] dt, \\ & + q(t) dB(t) + \int_{\mathbb{R}} r(t, z)\tilde{N}(dt, dz), \quad t \in [0, T], \end{aligned} \tag{6.5}$$

$$p(T) = g'(X(T)). \tag{6.6}$$

We solve the BSDE (6.5)–(6.6) recursively by proceeding as in Theorem 5.1.

Step 1. If $t \in [T - \delta, T]$, the BSDE takes the form

$$\begin{aligned} dp(t) &= q(t) dB(t) + \int_{\mathbb{R}} r(t, z)\tilde{N}(dt, dz), \quad t \in [T - \delta, T], \\ p(T) &= g'(X(T)), \end{aligned}$$

which has the solution

$$p(t) = \mathbb{E}[g'(X(T)) \mid \mathcal{F}_t], \quad t \in [T - \delta, T],$$

with corresponding $q(t)$ and $r(t, z)$ given by the martingale representation theorem (found, e.g. by using the Clark–Ocone theorem).

Step 2. If $t \in [T - 2\delta, T - \delta]$ and $T - 2\delta > 0$, we obtain by step 1 the BSDE

$$\begin{aligned} dp(t) = & -\mathbb{E} \left[\left\{ \alpha(t)p(t + \delta) + \beta(t)q(t + \delta) + \int_{\mathbb{R}} \gamma(t, z)r(t + \delta, z)v(dz) \right\} \mid \mathcal{F}_t \right] dt \\ & + q(t) dB(t) + \int_{\mathbb{R}} r(t, z)\tilde{N}(dt, dz), \quad t \in [T - 2\delta, T - \delta], \end{aligned}$$

with $p(T - \delta)$ known from step 1. Note that $p(t + \delta)$, $q(t + \delta)$, and $r(t + \delta)$ are also known from step 1. Therefore, this is a simple BSDE which can be solved for $p(t)$, $q(t)$, and $r(t, \cdot)$, $t \in [T - 2\delta, T - \delta]$. We continue like this by induction up to and including step j , where j is such that $T - j\delta \leq 0 < T - (j - 1)\delta$. With this procedure we end up with a solution $p(t) = p_{X(T)}(t)$ of (6.5)–(6.6) which depends on the (optimal) terminal value $X(T)$. If

$$0 \leq p(t) \leq v_0(t, \omega) \quad \text{for all } t \in [0, T] \tag{6.7}$$

then the optimal consumption rate $\hat{c}(t)$ is, by (6.4), given by

$$\hat{c}(t) = \hat{c}_{\hat{X}(T)}(t) = I(t, p(t), \omega), \quad t \in [0, T]. \tag{6.8}$$

Substituting this expression for $\hat{c}(t)$ into (6.1) we end up with a stochastic differential equation for the optimal wealth process $X(t)$. Solving this we find $X(T)$ and, hence, $\hat{c}(t)$.

We summarize the above in the following.

Proposition 6.1. *Let $p(t)$, $q(t)$, and $r(t, z)$ be the solution of the BSDE (6.5)–(6.5), as described above. Suppose that (6.7) holds. Then the optimal consumption rate $\hat{c}(t)$ and the corresponding optimal terminal wealth $X(t)$ are given implicitly by the coupled equations (6.8) and (6.1)–(6.2).*

To obtain a more explicit solution presentation, let us now assume that $\alpha(t)$ is deterministic and that $g(x) = kx$, $k > 0$.

Since k is deterministic, we can choose $q = r = 0$ in (6.5)–(6.6) and the BSDE becomes

$$\begin{aligned} dp(t) &= -\alpha(t)p(t + \delta) \mathbf{1}_{[0, T-\delta]}(t) dt, & t < T, \\ p(t) &= k \quad \text{for } t \in [T - \delta, T]. \end{aligned}$$

To solve this, we introduce

$$h(t) := p(T - t), \quad t \in [-\delta, T].$$

Then

$$dh(t) = -dp(T - t) = \alpha(T - t)p(T - t + \delta) dt = \alpha(T - t)h(t - \delta) dt$$

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

$$h(t) = p(T - t) = k \quad \text{for } t \in [-\delta, 0]. \tag{6.9}$$

This determines $h(t)$ inductively on each interval $[j\delta, (j + 1)\delta]$, $j = 1, 2, \dots$, as follows.

If $h(s)$ is known on $[(j - 1)\delta, j\delta]$ then

$$h(t) = h(j\delta) + \int_0^t h'(s) ds = h(j\delta) + \int_{j\delta}^t \alpha(T - s)h(s - \delta) ds \tag{6.10}$$

for $j \in [j\delta, (j + 1)\delta]$.

We have thus proved the following result.

Proposition 6.2. *Assume that $\alpha(t)$ is deterministic and that $g(x) = kx$, $k > 0$. The optimal consumption rate $\hat{c}_\delta(t)$ for the problem (6.1)–(6.2), (6.3) is given by*

$$\hat{c}_\delta(t) = I(t, h_\delta(T - t), \omega), \tag{6.11}$$

where $h_\delta(\cdot) = h(\cdot)$ is determined by (6.9)–(6.10).

Remark 6.1. Assume that $\alpha(t) = \alpha(\text{deterministic}) > 0$ for all $t \in [0, T]$. Then we see by induction on (6.10) that

$$0 \leq \delta_1 < \delta_2 \implies h_{\delta_1}(t) > h_{\delta_2}(t) \quad \text{for all } t \in (0, T],$$

and, hence, perhaps surprisingly,

$$0 \leq \delta_1 < \delta_2 \implies \hat{c}_{\delta_1}(t) < \hat{c}_{\delta_2}(t) \quad \text{for all } t \in [0, T].$$

Thus, the optimal consumption rate *increases* if the delay increases. The explanation for this may be that the delay postpones the negative effect on the growth of the cash flow caused by the consumption.

Acknowledgements

We thank Joscha Diehl, Martin Schweizer, and an anonymous referee for helpful comments and suggestions.

The research leading to these results has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013), grant agreement number 228087.

References

- [1] CHEN, L. AND WU, Z. (2010). Maximum principle for the stochastic optimal control problem with delay and application. *Automatica* **46**, 1074–1080.
- [2] CHOJNOWSKA-MICHALIK, A. (1978). Representation theorem for general stochastic delay equations. *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.* **26**, 635–642.
- [3] DAVID D. (2008). Optimal control of stochastic delayed systems with jumps. Preprint.
- [4] EL-KAROUI, N. AND HAMADÈNE, S. (2003). BSDEs and risk-sensitive control, zero-sum and nonzero-sum game problems of stochastic functional differential equations. *Stoch. Process. Appl.* **107**, 145–169.
- [5] ELSANOUSI, I., ØKSENDAL, B. AND SULEM, A. (2000). Some solvable stochastic control problems with delay. *Stoch. Stoch. Reports* **71**, 69–89.
- [6] FEDERICO, S. (2009). A stochastic control problem with delay arising in a pension fund model. To appear in *Finance Stoch.*
- [7] GOZZI, F. AND MARINELLI, C. (2004). Stochastic optimal control of delay equations arising in advertising models. In *Stochastic Partial Differential Equations and Applications—VII* (Lecture Notes Pure Appl. Math. **245**), Chapman and Hall/CRC, Boca Raton, FL, pp. 133–148.
- [8] GOZZI, F., MARINELLI, C. AND SAVIN, S. (2009). On controlled linear diffusions with delay in a model of optimal advertising under uncertainty with memory effects. *J. Optimization Theory Appl.* **142**, 291–321.
- [9] KOLMANOVSKI, V. B. AND SHAIKHET, L. E. (1996). *Control of Systems with Aftereffect*. American Mathematical Society, Providence, RI.
- [10] LARSEN, B. (2002). Dynamic programming in stochastic control of systems with delay. *Stoch. Stoch. Reports* **74**, 651–673.
- [11] LARSEN, B. AND RISEBRO, N. H. (2003). *When are HJB-equations in stochastic control of delay systems finite dimensional?* *Stoch. Anal. Appl.* **21**, 643–671.
- [12] ØKSENDAL, B. AND SULEM, A. (2001). A maximum principle for optimal control of stochastic systems with delay with applications to finance. In *Optimal Control and Partial Differential Equations*, eds J. L. Menaldi, E. Rofman and A. Sulem, IOS Press, Amsterdam, pp. 64–79.
- [13] ØKSENDAL, B. AND SULEM, A. (2007). *Applied Stochastic Control of Jump Diffusions*, 2nd edn. Springer, Berlin.
- [14] ØKSENDAL, B. AND ZHANG, T. (2010). Optimal control with partial information for stochastic Volterra equations. *Internat. J. Stoch. Anal.* **2010**, 25 pp.
- [15] PARDOUX, É. AND PENG, S. G. (1990). Adapted solution of a backward stochastic differential equation. *Systems Control Lett.* **14**, 55–61.
- [16] PENG, S. AND YANG, Z. (2009). Anticipated backward stochastic differential equations. *Ann. Probab.* **37**, 877–902.
- [17] PESZAT, S. AND ZABCZYK, J. (2008). *Stochastic Partial Differential Equations with Lévy Noise* (Encyclopedia Math. Appl. **113**). Cambridge University Press.
- [18] SITU, R. (1997). On solutions of backward stochastic differential equations with jumps and applications. *Stoch. Process. Appl.* **66**, 209–236.