Optimal control of stochastic delay equations and time-advanced backward stochastic differential equations

Bernt Oksendal, Agnès Sulem, Tusheng Zhang

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Optimal control of stochastic delay equations and time-advanced backward stochastic differential equations

Bernt Øksendal — Agnès Sulem — Tusheng Zhang

N° 7518

January 2011

Stochastic Methods and Models

Rapport de recherche

ISRN INRIA/RR--7518--FR+ENG

SN 0249-6399

Optimal control of stochastic delay equations and time-advanced backward stochastic differential equations

Bernt Øksendal* † , Agnès Sulem ‡ , Tusheng Zhang§ *

Theme : Stochastic Methods and Models
Équipe-Projet Mathfi

Rapport de recherche n° 7518 — January 2011 — 29 pages

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.

Key-words: Optimal control, stochastic delay equations, Lévy processes, maximum principles, Hamiltonian, adjoint processes, time-advanced BSDEs

MSC (2010): 93EXX, 93E20, 60H10, 60H15, 60H20, 60J75, 49J55, 35R60

* Center of Mathematics for Applications (CMA), University of Oslo, Box 1053 Blindern, N-0316 Oslo, Norway. Email: oksendal@math.uio.no - The research leading to these results has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no [228087].
† Norwegian School of Economics and Business Administration (NHH), Helleveien 30, N-5045 Bergen, Norway.
‡ INRIA Paris-Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, Le Chesnay Cedex, 78153, France. Email: agnes.sulem@inria.fr
§ School of Mathematics, University of Manchester, Oxford Road, Manchester M139PL, United Kingdom. Email: tusheng.zhang@manchester.ac.uk

Contrôle optimal d'équations différentielles stochastiques avec retard et équations différentielles stochastiques rétrogrades anticipantes

Résumé : On étudie des problèmes de contrôle d'équations différentielles stochastiques à retard avec sauts. On établit des principes de maximum stochastiques nécessaires et suffisants pour le contrôle optimal de tels systèmes. On montre que les processus adjoints associés satisfont une équations différentielle stochastique rétrograde anticipante (en temps) (EDSRA). On prouve plusieurs résultats d'existence et d'unicité pour de telles EDSRA. Les résultats sont illustrés par un problème de consommation optimale d'un système à retard.

Mots-clés : Contrôle optimal, équations différentielles stochastiques avec retard, processus de Lévy, principes du maximum, Hamiltonien, processus adjoints, équations différentielles stochastiques rétrogrades anticipantes

1 Introduction and problem formulation

Let $B(t) = B(t,\omega)$ be a Brownian motion and $\tilde{N}(dt,dz) := N(dt,dz) - \nu(dz)dt$, 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 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\le t\le T}, P)$.

We consider a controlled stochastic delay equation of the form

$$\begin{aligned} dX(t) &= b(t,X(t),Y(t),A(t),u(t),\omega)dt + \sigma(t,X(t),Y(t),A(t),u(t),\omega)dB(t) \\ &\quad + \int_{\mathbb{R}} \theta(t,X(t),Y(t),A(t),u(t),z,\omega)\tilde{N}(dt,dz)\ ;\ t\in[0,T] \qquad (1.1) \\ X(t) &= x_0(t)\ ;\ t\in[-\delta,0], \qquad (1.2) \end{aligned}$$

where

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

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

$$\begin{aligned} b &: [0,T]\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}\times\mathcal{U}\times\Omega \to \mathbb{R} \\ \sigma &: [0,T]\times\mathbb{R}\times\mathbb{R}\times\mathbb{R}\times\mathcal{U}\times\Omega \to \mathbb{R} \end{aligned}$$

and

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

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 $z\in\mathbb{R}_0 := \mathbb{R}\backslash\{0\}$. The function $x_0(t)$ is assumed to be continuous, deterministic.

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 $X(\cdot)\in L^2(\lambda\times P)$ where λ denotes the Lebesgue measure on $[0,T]$.

The performance functional is assumed to have the form

$$J(u) = E\left[\int_0^T f(t,X(t),Y(t),A(t),u(t),\omega)dt + g(X(T),\omega)\right]\ ;\ u\in\mathcal{A}_{\mathcal{E}} \qquad (1.4)$$

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

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

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

$$\Phi(x_0) := \sup_{u \in \mathcal{A}_\mathcal{E}} J(u) = J(u^*). \tag{1.5}$$

Any control $u^* \in \mathcal{A}_\mathcal{E}$ satisfying (1.5) 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. optimal harvesting problem of biological systems, one is led to 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. 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 to [2, 8, 9, 10] in the stochastic Brownian case. To the best of our knowledge, despite the statement of a result in [20], 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 to [11, 7, 13, 14, 16]. The paper [4] represents an extension of [14] to the case when the equation is driven by a Lévy noise.

We also mention the paper [6], 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.

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 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 is Section 6.

2 Hamiltonian and time-advanced BSDEs for adjoint equations

We define the *Hamiltonian*

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

by

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

where $\mathcal{R}$ is the set of functions $r : \mathbb{R}_0 \to \mathbb{R}$ 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

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

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 following backward stochastic differential equation (BSDE):

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

where

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

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

In the case when there are no jumps and no integral term in (2.4), anticipative BSDEs (ABSDEs for short) have been studied by [19], who prove existence and uniqueness of such equations under certain conditions. They also relate 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 we throughout the study 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 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

$$Y_i(t) := X(t-\delta_i); \quad i = 1, ..., N, \tag{2.5}$$

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

$$A_j(t) := \int_{t-\delta}^t \phi_j(t,s)X(s)ds; \quad j = 1, ..., M, \tag{2.6}$$

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

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

$$-\int_t^{t+\delta} \frac{\partial H}{\partial a_j}(t, X(s), Y_1(s), ..., Y_N(s), A_1(s), ..., A_M(s), u(s), p(s), q(s), r(s,\cdot))\phi_j(t,s)\chi_{[0,T]}(s)ds; \quad 1 \leq j \leq M. \tag{2.7}$$

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

$$A^\lambda(t) := \int_{t-\delta}^t \phi(t,s)X(s)d\lambda(s), \tag{2.8}$$

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

$$\begin{aligned} dp(t) &= -\{E[\frac{\partial H}{\partial x}(t)|\mathcal{F}_t] + \sum_{i=1}^N E[\frac{\partial H}{\partial y}(t+\delta_i)|\mathcal{F}_t]\chi_{[0,T-\delta_i]}(t)\}dt \\ &\quad -E[(\int_t^{t+\delta} \frac{\partial H}{\partial a}(s)\phi(t,s)\chi_{[0,T]}(s)ds)|\mathcal{F}_t]d\lambda(t) \\ &\quad +q(t)dB(t) + \int_{\mathbb{R}_0} r(t,z)\tilde{N}(dt,dz)\ ;\ t \in [0,T] & (2.9) \\ p(T) &= g'(X(T)), & (2.10) \end{aligned}$$

where we have used the simplified notation $\frac{\partial H}{\partial x}(t) = \frac{\partial H}{\partial x}(t, X(t), Y_1(t), ..., Y_N(t), ...)$ etc.

However, for simplicity of presentation 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 principe 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 $\hat{X}(t), \hat{Y}(t), \hat{A}(t)$ and adjoint processes $\hat{p}(t), \hat{q}(t), \hat{r}(t,z)$, assumed to satisfy the ABSDE* (2.3)-(2.4)*. Suppose the following hold:*

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

$$(x,y,a,u) \to H(t,x,y,a,u,\hat{p}(t),\hat{q}(t),\hat{r}(t,\cdot)) \tag{3.1}$$

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

(ii)

$$E\left[\int_0^T \left\{\hat{p}(t)^2\left(\sigma^2(t) + \int_{\mathbb{R}_0}\theta^2(t,z)\nu(dz)\right) + X^2(t)\left(\hat{q}^2(t) + \int_{\mathbb{R}_0}\hat{r}^2(t,z)\nu(dz)\right)\right\}dt\right] < \infty \tag{3.2}$$

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

(iii)

$$
\begin{aligned}
&\max_{v\in\mathcal{U}} E\left[H(t,\hat{X}(t),\hat{X}(t-\delta),\hat{A}(t),v,\hat{p}(t),\hat{q}(t),\hat{r}(t,\cdot)) \mid \mathcal{E}_t\right] \\
&\qquad = E\left[H(t,\hat{X}(t),\hat{X}(t-\delta),\hat{A}(t),\hat{u}(t),\hat{p}(t),\hat{q}(t),\hat{r}(t,\cdot)) \mid \mathcal{E}_t\right]
\end{aligned}
\tag{3.3}
$$

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

Then $\hat{u}(t)$ is an optimal control for the problem (1.5).

Proof. Choose $u \in \mathcal{A}_{\mathcal{E}}$ and consider

$$
J(u) - J(\hat{u}) = I_1 + I_2 \tag{3.4}
$$

where

$$
I_1 = E\left[\int_0^T \{f(t,X(t),Y(t),A(t),u(t)) - f(t,\hat{X}(t),\hat{Y}(t),\hat{A}(t),\hat{u}(t))\}dt\right] \tag{3.5}
$$

$$
I_2 = E[g(X(T)) - g(\hat{X}(T))]. \tag{3.6}
$$

By the definition of H and concavity of H we have

$$
\begin{aligned}
I_1 = E\Bigg[\int_0^T &\{H(t,X(t),Y(t),A(t),u(t),\hat{p}(t),\hat{q}(t),\hat{r}(t,\cdot)) \\
&- H(t,\hat{X}(t),\hat{Y}(t),\hat{A}(t),\hat{u}(t),\hat{p}(t),\hat{q}(t),\hat{r}(t,\cdot)) \\
&- (b(t,X(t),Y(t),A(t),u(t)) - b(t,\hat{X}(t),\hat{Y}(t),\hat{A}(t),\hat{u}(t)))\hat{p}(t) \\
&- (\sigma(t,X(t),Y(t),A(t),u(t)) - \sigma(t,\hat{X}(t),\hat{Y}(t),\hat{A}(t),\hat{u}(t)))\hat{q}(t) \\
&- \int_{\mathbb{R}} (\theta(t,X(t),Y(t),A(t),u(t),z) - \theta(t,\hat{X}(t),\hat{Y}(t),\hat{A}(t),\hat{u}(t),z))\hat{r}(t,z)\nu(dz)\}dt\Bigg] \\
\le E\Bigg[\int_0^T &\{\frac{\partial \hat{H}}{\partial x}(t)(X(t)-\hat{X}(t)) + \frac{\partial \hat{H}}{\partial y}(t)(Y(t)-\hat{Y}(t)) + \frac{\partial \hat{H}}{\partial a}(t)(A(t)-\hat{A}(t)) \\
&+ \frac{\partial H}{\partial u}(t)(u(t)-\hat{u}(t)) - (b(t)-\hat{b}(t))\hat{p}(t) - (\sigma(t)-\hat{\sigma}(t))\hat{q}(t) \\
&- \int_{\mathbb{R}} (\theta(t,z)-\hat{\theta}(t,z))\hat{r}(t,z)\nu(dz)\}dt\Bigg],
\end{aligned}
\tag{3.7}
$$

where we have used the abbreviated notation

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

Since g is concave we have, by (3.2),

$$\begin{aligned}
I_2 &\le E[g'(\hat{X}(T))(X(T)-\hat{X}(T))] = E[\hat{p}(T)(X(T)-\hat{X}(T))] \\
&= E\left[\int_0^T \hat{p}(t)(dX(t)-d\hat{X}(t)) + \int_0^T (X(t)-\hat{X}(t))d\hat{p}(t)\right. \\
&\quad \left. + \int_0^T (\sigma(t)-\hat{\sigma}(t))\hat{q}(t)dt + \int_0^T \int_{\mathbb{R}} (\theta(t,z)-\hat{\theta}(t,z))\hat{r}(t,z)\nu(dz)dt\right] \\
&= E\left[\int_0^T (b(t)-\hat{b}(t))\hat{p}(t)dt + \int_0^T (X(t)-\hat{X}(t))E[\mu(t)|\mathcal{F}_t]dt\right. \\
&\quad \left. + \int_0^T (\sigma(t)-\hat{\sigma}(t))\hat{q}(t)dt + \int_0^T \int_{\mathbb{R}} (\theta(t,z)-\hat{\theta}(t,z))\hat{r}(t,z)\nu(dz)dt\right]. \qquad (3.8)
\end{aligned}$$

Combining (3.4)-(3.8) we get, using that $X(t)=\hat{X}(t)=x_0(t)$ for all $t\in[-\delta,0]$,

$$\begin{aligned}
J(u)-J(\hat{u}) &\le E\left[\int_0^T \left\{\frac{\partial H}{\partial x}(t)(X(t)-\hat{X}(t)) + \frac{\partial H}{\partial y}(t)(Y(t)-\hat{Y}(t))\right.\right. \\
&\quad \left.\left. + \frac{\partial \hat{H}}{\partial a}(t)(A(t)-\hat{A}(t)) + \frac{\partial \hat{H}}{\partial u}(t)(u(t)-\hat{u}(t)) + \mu(t)(X(t)-\hat{X}(t))\right\} dt\right] \\
&= E\left[\int_\delta^{T+\delta} \left\{\frac{\partial \hat{H}}{\partial x}(t-\delta) + \frac{\partial \hat{H}}{\partial y}(t)\chi_{[0,T]}(t) + \mu(t-\delta)\right\}(Y(t)-\hat{Y}(t))dt\right. \\
&\quad \left. + \int_0^T \frac{\partial \hat{H}}{\partial a}(t)(A(t)-\hat{A}(t))dt + \int_0^T \frac{\partial \hat{H}}{\partial u}(t)(u(t)-\hat{u}(t))dt\right]. \qquad (3.9)
\end{aligned}$$

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

$$\begin{aligned}
\int_0^T \frac{\partial \hat{H}}{\partial a}(s)(A(s)-\hat{A}(s))ds &= \int_0^T \frac{\partial \hat{H}}{\partial a}(s)\int_{s-\delta}^s e^{-\rho(s-r)}(X(r)-\hat{X}(r))drds \\
&= \int_0^T \left(\int_r^{r+\delta} \frac{\partial \hat{H}}{\partial a}(s)e^{-\rho s}\chi_{[0,T]}(s)ds\right)e^{\rho r}(X(r)-\hat{X}(r))dr \\
&= \int_\delta^{T+\delta} \left(\int_{t-\delta}^t \frac{\partial \hat{H}}{\partial a}(s)e^{-\rho s}\chi_{[0,T]}(s)ds\right)e^{\rho(t-\delta)}(X(t-\delta)-\hat{X}(t-\delta))dt. \qquad (3.10)
\end{aligned}$$

Combining this with (3.9) and using (2.4) we obtain

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

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. $\square$

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:

A 1 *For all $u \in \mathcal{A}_{\mathcal{E}}$ and all bounded $\beta \in \mathcal{A}_{\mathcal{E}}$ there exists $\varepsilon > 0$ such that*

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

A 2 *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*

$$\beta(t) = \alpha \chi_{[t_0,T]}(t) \; ; \; t \in [0,T] \tag{4.1}$$

belongs to $\mathcal{A}_{\mathcal{E}}$.

A 3 *For all bounded $\beta \in \mathcal{A}_{\mathcal{E}}$ the derivative process*

$$\xi(t) := \frac{d}{ds} X^{u+s\beta}(t) \mid_{s=0} \tag{4.2}$$

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 we for simplicity of notation have put

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

and we have used that

$$\frac{d}{ds}Y^{u+s\beta}(t) \mid_{s=0} = \frac{d}{ds}X^{u+s\beta}(t-\delta) \mid_{s=0} = \xi(t-\delta) \tag{4.4}$$

and

$$\begin{aligned} \frac{d}{ds}A^{u+s\beta}(t) \mid_{s=0} &= \frac{d}{ds}\left(\int_{t-\delta}^{t} e^{-\rho(t-r)}X^{u+s\beta}(r)dr\right) \mid_{s=0} \\ &= \int_{t-\delta}^{t} e^{-\rho(t-r)}\frac{d}{ds}X^{u+s\beta}(r) \mid_{s=0} dt = \int_{t-\delta}^{t} e^{-\rho(t-r)}\xi(r)dr. \end{aligned} \tag{4.5}$$

Note that

$$\xi(t) = 0 \text{ for } t \in [-\delta, 0]. \tag{4.6}$$

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

Assume that

$$\begin{aligned} E\Bigg[\int_0^T \hat{p}^2(t) \Bigg\{ &\left(\frac{\partial \sigma}{\partial x}\right)^2(t)\hat{\xi}^2(t) + \left(\frac{\partial \sigma}{\partial y}\right)^2(t)\xi^2(t-\delta) \\ &+ \left(\frac{\partial \sigma}{\partial a}\right)^2(t)\left(\int_{t-\delta}^{t} e^{-\rho(t-r)}\hat{\xi}(r)dr\right)^2 + \left(\frac{\partial \sigma}{\partial u}\right)^2(t) \\ &+ \int_{\mathbb{R}_0}\Bigg\{\left(\frac{\partial \theta}{\partial x}\right)^2(t,z)\hat{\xi}^2(t) + \left(\frac{\partial \theta}{\partial y}\right)^2(t,z)\hat{\xi}^2(t-\delta) \\ &+ \left(\frac{\partial \theta}{\partial a}\right)^2(t,z)\left(\int_{t-\delta}^{t} e^{-\rho(t-r)}\hat{\xi}(r)dr\right)^2 + \left(\frac{\partial \theta}{\partial u}\right)^2(t,z)\Bigg\}\nu(dz)\Bigg\} dt \\ &+ \int_0^T \hat{\xi}^2(t)\left\{\hat{q}^2(t) + \int_{\mathbb{R}_0}\hat{r}^2(t,z)\nu(dz)\right\}dt\Bigg] < \infty. \end{aligned} \tag{4.7}$$

Then the following are equivalent:

(i) $\frac{d}{ds}J(\hat{u}+s\beta)\mid_{s=0}=0$ *for all bounded* $\beta \in \mathcal{A}_{\mathcal{E}}$.

(ii) $E\left[\frac{\partial H}{\partial u}(t,\hat{X}(t),\hat{Y}(t),\hat{A}(t),u,\hat{p}(t),\hat{q}(t),\hat{r}(t,\cdot)) \mid \mathcal{E}_t\right]_{u=\hat{u}(t)}=0$ *a.s. for all* $t\in[0,T]$.

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 **(i)** holds. Then

$$
\begin{aligned}
0 &= \frac{d}{ds}J(u+s\beta)\mid_{s=0}\\
&= \frac{d}{ds}E\left[\int_0^T f(t,X^{u+s\beta}(t),Y^{u+s\beta}(t),A^{u+s\beta}(t),u(t)+s\beta(t))dt+g(X^{u+s\beta}(T))\right]\mid_{s=0}\\
&= E\left[\int_0^T\left\{\frac{\partial f}{\partial x}(t)\xi(t)+\frac{\partial f}{\partial y}(t)\xi(t-\delta)+\frac{\partial f}{\partial a}(t)\int_{t-\delta}^t e^{-\rho(t-r)}\xi(r)dt+\frac{\partial f}{\partial u}(t)\beta(t)\right\}dt+g'(X(T))\xi(T)\right]\\
&= E\left[\int_0^T\left\{\frac{\partial H}{\partial x}(t)-\frac{\partial b}{\partial x}(t)p(t)-\frac{\partial \sigma}{\partial x}(t)q(t)-\int_{\mathbb{R}}\frac{\partial \theta}{\partial x}(t,z)r(t,z)\nu(dz)\right\}\xi(t)dt\right.\\
&\quad+\int_0^T\left\{\frac{\partial H}{\partial y}(t)-\frac{\partial b}{\partial y}(t)p(t)-\frac{\partial \sigma}{\partial y}(t)q(t)-\int_{\mathbb{R}}\frac{\partial \theta}{\partial y}(t,z)r(t,z)\nu(dz)\right\}\xi(t-\delta)dt\\
&\quad+\int_0^T\left\{\frac{\partial H}{\partial a}(t)-\frac{\partial b}{\partial a}(t)p(t)-\frac{\partial \sigma}{\partial a}(t)q(t)-\int_{\mathbb{R}}\frac{\partial \theta}{\partial a}(t,z)r(t,z)\nu(dz)\right\}\left(\int_{t-\delta}^t e^{-\rho(t-r)}\xi(r)dr\right)dt\\
&\quad\left.+\int_0^T\frac{\partial f}{\partial u}(t)\beta(t)dt+g'(X(T))\xi(T)\right].
\end{aligned}
\tag{4.8}
$$

By (4.3)

$$
\begin{aligned}
E[g'(X(T))\xi(T)] &= E[p(T)\xi(T)] = E\left[\int_0^T p(t)d\xi(t) + \int_0^T \xi(t)dp(t)\right.\\
&+ \int_0^T q(t)\left\{\frac{\partial\sigma}{\partial x}(t)\xi(t) + \frac{\partial\sigma}{\partial y}(t)\xi(t-\delta) + \frac{\partial\sigma}{\partial a}(t)\int_{t-\delta}^t e^{-\rho(t-r)}\xi(r)dr + \frac{\partial\sigma}{\partial u}(t)\beta(t)\right\}dt\\
&+ \int_0^T \int_{\mathbb{R}} r(t,z)\left\{\frac{\partial\theta}{\partial x}(t,z)\xi(t) + \frac{\partial\theta}{\partial y}(t,z)\xi(t-\delta) + \frac{\partial\theta}{\partial a}(t,z)\int_{t-\delta}^t e^{-\rho(t-r)}\xi(r)dr\right.\\
&\left.\left. + \frac{\partial\theta}{\partial u}(t)\beta(t)\right\}\nu(dz)dt\right]\\
= E&\left[\int_0^T p(t)\left\{\frac{\partial b}{\partial x}(t)\xi(t) + \frac{\partial b}{\partial y}(t)\xi(t-\delta) + \frac{\partial b}{\partial a}(t)\int_{t-\delta}^t e^{-\rho(t-r)}\xi(r)dr + \frac{\partial b}{\partial u}(t)\beta(t)\right\}dt\right.\\
&+ \int_0^T \xi(t)E[\mu(t)|\mathcal{F}_t]dt\\
&+ \int_0^T q(t)\left\{\frac{\partial\sigma}{\partial x}(t)\xi(t) + \frac{\partial\sigma}{\partial y}(t)\xi(t-\delta) + \frac{\partial\sigma}{\partial a}(t)\int_{t-\delta}^t e^{-\rho(t-r)}\xi(r)dr + \frac{\partial\sigma}{\partial u}(t)\beta(t)\right\}dt\\
&+ \int_0^T \int_{\mathbb{R}} r(t,z)\left\{\frac{\partial\theta}{\partial x}(t,z)\xi(t) + \frac{\partial\theta}{\partial y}(t,z)\xi(t-\delta) + \frac{\partial\theta}{\partial a}(t,z)\int_{t-\delta}^t e^{-\rho(t-r)}\xi(r)dr\right.\\
&\left.\left. + \frac{\partial\theta}{\partial u}(t,z)\beta(t)\right\}\nu(dz)dt\right]
\end{aligned}
\tag{4.9}
$$

Combining (4.8) and (4.9) we get

$$
\begin{aligned}
0 &= E\Bigg[\int_0^T \xi(t)\left\{\frac{\partial H}{\partial x}(t)+\mu(t)\right\}dt+\int_0^T \xi(t-\delta)\frac{\partial H}{\partial y}(t)dt \\
&\quad +\int_0^T\left(\int_{t-\delta}^t e^{-\rho(t-r)}\xi(r)dr\right)\frac{\partial H}{\partial a}(t)dt+\int_0^T\frac{\partial H}{\partial u}(t)\beta(t)dt\Bigg] \\
&= E\Bigg[\int_0^T \xi(t)\Bigg\{\frac{\partial H}{\partial x}(t)-\frac{\partial H}{\partial x}(t)-\frac{\partial H}{\partial y}(t+\delta)\chi_{[0,T-\delta]}(t) \\
&\quad -e^{\rho t}\left(\int_t^{t+\delta}\frac{\partial H}{\partial a}(s)e^{-\rho s}\chi_{[0,T]}(s)ds\right)\Bigg\}dt+\int_0^T\xi(t-\delta)\frac{\partial H}{\partial y}(t)dt \\
&\quad +\int_0^T\left(\int_{s-\delta}^s e^{-\rho(s-t)}\xi(t)dt\right)\frac{\partial H}{\partial a}(s)ds+\int_0^T\frac{\partial H}{\partial u}(t)\beta(t)dt\Bigg] \\
&= E\Bigg[\int_0^T \xi(t)\left\{-\frac{\partial H}{\partial y}(t+\delta)\chi_{[0,T-\delta]}(t)-e^{\rho t}\left(\int_t^{t+\delta}\frac{\partial H}{\partial a}(s)e^{-\rho s}\chi_{[0,T]}(s)ds\right)\right\}dt \\
&\quad +\int_0^T\xi(t-\delta)\frac{\partial H}{\partial y}(t)dt \\
&\quad +e^{\rho t}\int_0^T\left(\int_t^{t+\delta}\frac{\partial H}{\partial a}(s)e^{-\rho s}\chi_{[0,T]}(s)ds\right)\xi(t)dt+\int_0^T\frac{\partial H}{\partial u}(t)\beta(t)dt\Bigg] \\
&= E\left[\int_0^T\frac{\partial H}{\partial u}(t)\beta(t)dt\right],
\end{aligned}
\tag{4.10}
$$

where we again have used integration by parts.

If we apply (4.10) to

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

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

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

Differentiating with respect to s we obtain

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

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

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

This shows that **(i)** $\Rightarrow$ **(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 **(ii)** $\Rightarrow$ **(i)** by reversing the above argument. □

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

We now study time-advanced backward stochastic differential equations 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 BSDE 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 $X(\cdot) : \mathbb{R}_+ \to \mathbb{R}$, X_t will denote the function defined by $X_t(s) = X(t+s)$ for $s \in [0,\delta]$. Put $\mathcal{H} = L^2(\nu)$. Consider the L^2 spaces $V_1 := L^2([0,\delta], ds)$ and $V_2 := L^2([0,\delta] \to \mathcal{H}, ds)$. Let

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

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

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

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

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

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

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

$$\begin{aligned} dp(t) &= F\left(t, p(t), 0, 0, q(t), 0, 0, r(t,z), 0, 0\right) dt \\ &\quad + q(t)dB(t) + \int_{\mathbb{R}} r(t,z)\tilde{N}(dt,dz)\ ;\ t \in [T-\delta, T] \qquad (5.5) \\ p(T) &= G. \qquad (5.6) \end{aligned}$$

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

$$\begin{aligned} dp(t) &= E[F\left(t, p(t), p(t+\delta), p_t, q(t), q(t+\delta), q_t, r(t), r(t+\delta), r_t\right) |\mathcal{F}_t]dt \\ &\quad + q(t)dB(t) + \int_{\mathbb{R}} r(t,z)\tilde{N}(dt,dz)\ ;\ t \in [T-(k+1)\delta, T-k\delta] \qquad (5.7) \\ p(T-k\delta) &= \text{ the value found in Step } k-1 \qquad (5.8) \end{aligned}$$

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. 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 consider now the following backward stochastic differential equation in the unknown $\mathcal{F}_t$-adapted processes $(p(t), q(t), r(t,x))$:

$$\begin{aligned} dp(t) &= E[F(t, p(t), p(t+\delta), p_t, q(t), q(t+\delta), q_t, r(t), r(t+\delta), r_t)|\mathcal{F}_t]dt \\ &\quad + q(t)dB_t + \int_{\mathbb{R}} r(t,z)\tilde{N}(dt,dz),\ ;\ t \in [0,T] \qquad (5.9) \\ p(t) &= G(t), \qquad t \in [T, T+\delta], \qquad (5.10) \end{aligned}$$

where G is a given continuous $\mathcal{F}_t$-adapted stochastic process. We shall present 3 theorems with various conditions on F and G.

Theorem 5.2 *Assume $E[\sup_{T\leq t\leq T+\delta} |G(t)|^2] < \infty$ and that the condition (5.1) is satisfied. Then the backward stochastic differential equation (5.9)- (5.10) admits a unique solution $(p(t), q(t), r(t,z))$ such that*

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

Proof. *Step 1: Assume F is independent of p_1, p_2 and p.*

Set $q^0(t) := 0, r^0(t,x) = 0$. For $n \geq 1$, define $(p^n(t), q^n(t), r^n(t,x))$ to be the unique solution to the following backward stochastic differential equation equation:

$$\begin{aligned} dp^n(t) &= E[F(t, q^{n-1}(t), q^{n-1}(t+\delta), q_t^{n-1}, r^{n-1}(t,\cdot), r^{n-1}(t+\delta,\cdot), r_t^{n-1}(\cdot))|\mathcal{F}_t]dt \\ &\quad + q^n(t)dB_t + r^n(t,z)\tilde{N}(dt,dz), \quad t \in [0,T] \end{aligned} \tag{5.11}$$

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

It is a consequence of the martingale representation theorem that the above equation admits a unique solution, see, e.g. [23], [18]. We extend q^n, r^n to $[0, T+\delta]$ by setting $q^n(s) = 0$, $r^n(s,z) = 0$ 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

$$\begin{aligned} 0 &= |p^{n+1}(T) - p^n(T)|^2 = |p^{n+1}(t) - p^n(t)|^2 \\ &+ 2\int_t^T (p^{n+1}(s) - p^n(s))(E[F(s, q^n(s), q^n(s+\delta), q_s^n, r^n(s,\cdot), r^n(s+\delta,\cdot), r_s^n(\cdot))|\mathcal{F}_s] \\ &- E[F(s, q^{n-1}(s), q^{n-1}(s+\delta), q_s^{n-1}, r^{n-1}(s,\cdot), r^{n-1}(s+\delta,\cdot), r_s^{n-1}(\cdot)))|\mathcal{F}_s]ds \\ &+ \int_t^T \int_{\mathbb{R}} |r^{n+1}(s,z) - r^n(s,z)|^2 ds\, \nu(dz) + \int_t^T |q^{n+1}(s) - q^n(s)|^2 ds \\ &+ 2\int_t^T (p^{n+1}(s) - p^n(s))(q^{n+1}(s) - q^n(s))dB_s \\ &+ \int_t^T \int_{\mathbb{R}} \{|r^{n+1}(s,z) - r^n(s,z)|^2 + 2(p^{n+1}(s-) - p^n(s-))(r^{n+1}(s,z) - r^n(s,z))\}\tilde{N}(ds,dz) \end{aligned} \tag{5.12}$$

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

$$\begin{aligned} &E[|p^{n+1}(t) - p^n(t)|^2] \\ &\quad + E\left[\int_t^T \int_{\mathbb{R}} |r^{n+1}(s,z) - r^n(s,z)|^2 ds\nu(dz)\right] + E\left[\int_t^T |q^{n+1}(s) - q^n(s)|^2 ds\right] \\ &\leq 2E\left[\int_t^T |(p^{n+1}(s) - p^n(s))(E[F(s, q^n(s), q^n(s+\delta), r^n(s,\cdot), r^n(s+\delta,\cdot))\right. \\ &\quad \left. -F(s, q^{n-1}(s), q^{n-1}(s+\delta), r^{n-1}(s,\cdot), r^{n-1}(s+\delta,\cdot))|\mathcal{F}_s])|ds\right] \\ &\leq C_\varepsilon E\left[\int_t^T |p^{n+1}(s) - p^n(s)|^2 ds\right] + \varepsilon E\left[\int_t^T |q^n(s) - q^{n-1}(s)|^2 ds\right] \\ &\quad + \varepsilon E\left[\int_t^T |q^n(s+\delta) - q^{n-1}(s+\delta)|^2 ds\right] + \varepsilon E\left[\int_t^T (\int_s^{s+\delta} |q^n(u) - q^{n-1}(u)|^2 du) ds\right] \\ &\quad + \varepsilon E\left[\int_t^T |r^n(s) - r^{n-1}(s)|_{\mathcal{H}}^2 ds\right] \\ &\quad + \varepsilon E\left[\int_t^T |r^n(s+\delta) - r^{n-1}(s+\delta)|_{\mathcal{H}}^2 ds\right] + \varepsilon E\left[\int_t^T \left(\int_s^{s+\delta} |r^n(u) - r^{n-1}(u)|_{\mathcal{H}}^2 du\right) ds\right] \end{aligned} \tag{5.13}$$

Note that

$$E\left[\int_t^T |q^n(s+\delta)-q^{n-1}(s+\delta)|^2 ds\right] \le E\left[\int_t^T |q^n(s)-q^{n-1}(s)|^2 ds\right]. \tag{5.14}$$

Interchanging the order of integration,

$$\begin{aligned}
&E\left[\int_t^T \left(\int_s^{s+\delta} |q^n(u)-q^{n-1}(u)|^2 du\right) ds\right] = E\left[\int_t^{T+\delta} |q^n(u)-q^{n-1}(u)|^2 du(\int_{u-\delta}^u ds\right] \\
&\le \delta E\left[\int_t^T |q^n(s)-q^{n-1}(s)|^2 ds\right].
\end{aligned} \tag{5.15}$$

Similar inequalities hold also for $r^n - r^{n-1}$. It follows from (5.13) that

$$\begin{aligned}
&E[|p^{n+1}(t)-p^n(t)|^2] \\
&\quad + E\left[\int_t^T \int_{\mathbb{R}} |r^{n+1}(s,z)-r^n(s,z)|^2 ds\nu(dz)\right] + E\left[\int_t^T |q^{n+1}(s)-q^n(s)|^2 ds\right] \\
&\le C_\varepsilon E\left[\int_t^T |p^{n+1}(s)-p^n(s)|^2 ds\right] + (2+M)\varepsilon E\left[\int_t^T |q^n(s)-q^{n-1}(s)|^2 ds\right] \\
&+ 3\varepsilon E\left[\int_t^T |r^n(s)-r^{n-1}(s)|^2_{\mathcal{H}} ds\right].
\end{aligned} \tag{5.16}$$

Choose $\varepsilon > 0$ sufficiently small so that

$$\begin{aligned}
&E[|p^{n+1}(t)-p^n(t)|^2] \\
&\quad + E\left[\int_t^T \int_{\mathbb{R}} |r^{n+1}(s,z)-r^n(s,z)|^2 ds\ \nu(dz)\right] + E\left[\int_t^T |q^{n+1}(s)-q^n(s)|^2 ds\right] \\
&\le C_\varepsilon E\left[\int_t^T |p^{n+1}(s)-p^n(s)|^2 ds\right] + \frac{1}{2} E\left[\int_t^T |q^n(s)-q^{n-1}(s)|^2 ds\right] \\
&+ \frac{1}{2} E\left[\int_t^T |r^n(s)-r^{n-1}(s)|^2_{\mathcal{H}} ds\right].
\end{aligned} \tag{5.17}$$

This implies that

$$\begin{aligned}
&-\frac{d}{dt}\left(e^{C_\varepsilon t} E\left[\int_t^T |p^{n+1}(s)-p^n(s)|^2 ds\right]\right) \\
&\quad + e^{C_\varepsilon t} E\left[\int_t^T \int_{\mathbb{R}} |r^{n+1}(s,z)-r^n(s,z)|^2 ds\ \nu(dz)\right] + e^{C_\varepsilon t} E\left[\int_t^T |q^{n+1}(s)-q^n(s)|^2 ds\right] \\
&\le \frac{1}{2} e^{C_\varepsilon t} E\left[\int_t^T |q^n(s)-q^{n-1}(s)|^2 ds\right] + \frac{1}{2} e^{C_\varepsilon t} E\left[\int_t^T |r^n(s)-r^{n-1}(s)|^2_{\mathcal{H}} ds\right].
\end{aligned} \tag{5.18}$$

Integrating the last inequality we get

$$
\begin{aligned}
&E\left[\int_0^T |p^{n+1}(s)-p^n(s)|^2 ds\right] + \int_0^T dt\; e^{C_\varepsilon t} E\left[\int_t^T |q^{n+1}(s)-q^n(s)|^2 ds\right] \\
&+ \int_0^T dt\; e^{C_\varepsilon t} E\left[\int_t^T \int_{\mathbb{R}} |r^{n+1}(s,z)-r^n(s,z)|^2 ds\nu(dz)\right] \\
&\le \frac{1}{2}\int_0^T dt\; e^{C_\varepsilon t} E\left[\int_t^T |q^n(s)-q^{n-1}(s)|^2 ds\right] + \frac{1}{2}\int_0^T dt\; e^{C_\varepsilon t} E\left[\int_t^T |r^n(s)-r^{n-1}(s)|^2_{\mathcal{H}} ds\right]
\end{aligned} \tag{5.19}
$$

In particular,

$$
\begin{aligned}
&\int_0^T dt\; e^{C_\varepsilon t} E\left[\int_t^T \int_{\mathbb{R}} |r^{n+1}(s,z)-r^n(s,z)|^2 ds\nu(dz)\right] + \int_0^T dt\; e^{C_\varepsilon t} E\left[\int_t^T |q^{n+1}(s)-q^n(s)|^2 ds\right] \\
&\quad \le \frac{1}{2}\int_0^T dt\; e^{C_\varepsilon t} E\left[\int_t^T |q^n(s)-q^{n-1}(s)|^2 ds\right] + \frac{1}{2}\int_0^T dt\; e^{C_\varepsilon t} E\left[\int_t^T |r^n(s)-r^{n-1}(s)|^2_{\mathcal{H}} ds\right]
\end{aligned} \tag{5.20}
$$

This yields

$$
\begin{aligned}
&\int_0^T dt\; e^{C_\varepsilon t} E\left[\int_t^T \int_{\mathbb{R}} |r^{n+1}(s,z)-r^n(s,z)|^2 ds\nu(dz)\right] + \int_0^T dt\; e^{C_\varepsilon t} E\left[\int_t^T |q^{n+1}(s)-q^n(s)|^2 ds\right] \\
&\quad \le \left(\frac{1}{2}\right)^n C
\end{aligned} \tag{5.21}
$$

for some constant C. It follows from (5.19) that

$$
E\left[\int_0^T |p^{n+1}(s)-p^n(s)|^2 ds\right] \le \left(\frac{1}{2}\right)^n C. \tag{5.22}
$$

(5.16) and ((5.19) further gives

$$
E\left[\int_0^T \int_{\mathbb{R}} |r^{n+1}(s,z)-r^n(s,z)|^2 ds\nu(dz)\right] + E\left[\int_0^T |q^{n+1}(s)-q^n(s)|^2 ds\right] \le \left(\frac{1}{2}\right)^n CnC_\varepsilon. \tag{5.23}
$$

In view of (5.16), (5.19) and (5.20), we conclude that there exist progressively measurable processes $(p(t), q(t), r(t,z))$ such that

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

Letting $n \to \infty$ in (5.11) we see that $(p(t), q(t), r(t,z))$ satisfies

$$\begin{aligned} p(t) &+ \int_t^T E[F(s, q(s), q(s+\delta), q_s, r(s,\cdot), r(s+\delta,\cdot), r_s(\cdot))|\mathcal{F}_s]ds \\ &+ \int_t^T q(s)dB_s + \int_t^T \int_{\mathbb{R}} r(s,z)\tilde{N}(ds,dz) = g(T) \end{aligned} \tag{5.24}$$

i.e., $(p(t), q(t), r(t,z))$ is a solution. Uniqueness follows easily from the Ito's formula, a similar calculation of deducing (5.12) and (5.13), 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:

$$\begin{aligned} dp^n(t) &= E[F(t, p^{n-1}(t), p^{n-1}(t+\delta), p_t^{n-1}, q^n(t), q^n(t+\delta), q_t^n, r^n(t,\cdot), r^n(t+\delta,\cdot), r_t^n(\cdot))|\mathcal{F}_t]dt \\ &\quad + q^n(t)dB_t + r^n(t,z)\tilde{N}(dt,dz), \quad t \in [0,T] && (5.25) \\ p^n(t) &= G(t); \qquad t \in [T, T+\delta] && (5.26) \end{aligned}$$

The existence of $(p^n(t), q^n(t), r^n(t,z))$ is proved in Step 1. By the same arguments leading to (5.16), we deduce that

$$\begin{aligned} E[|p^{n+1}(t) - p^n(t)|^2] &+ \frac{1}{2}E\left[\int_t^T \int_{\mathbb{R}} |r^{n+1}(s,z) - r^n(s,z)|^2 ds\nu(dz)\right] \\ &+ \frac{1}{2}E\left[\int_t^T |q^{n+1}(s) - q^n(s)|^2 ds\right] \\ &\leq CE\left[\int_t^T |p^{n+1}(s) - p^n(s)|^2 ds\right] + \frac{1}{2}E\left[\int_t^T |p^n(s) - p^{n-1}(s)|^2 ds\right] \end{aligned} \tag{5.27}$$

This implies that

$$-\frac{d}{dt}\left(e^{Ct}E\left[\int_t^T |p^{n+1}(s) - p^n(s)|^2 ds\right]\right) \leq \frac{1}{2}e^{Ct}E\left[\int_t^T |p^n(s) - p^{n-1}(s)|^2 ds\right] \tag{5.28}$$

Integrating (5.28) from u to T we get

$$\begin{aligned} E\left[\int_u^T |p^{n+1}(s) - p^n(s)|^2 ds\right] &\leq \frac{1}{2}\int_u^T dt e^{C(t-u)}E\left[\int_t^T |p^n(s) - p^{n-1}(s)|^2 ds\right] \\ &\leq e^{CT}\int_u^T dt E[\int_t^T |p^n(s) - p^{n-1}(s)|^2 ds]. \end{aligned} \tag{5.29}$$

Iterating the above inequality we obtain that

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

Using 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 equation (5.9). □

Theorem 5.3 *Assume* $E\left[\sup_{T\le t\le T+\delta}|G(t)|^{2\alpha}\right] < \infty$ *for some* $\alpha > 1$ *and that the following condition hold:*

$$\begin{aligned}|F(t,p_1,p_2,p,q_1,q_2,q,r_1,r_2,r) - F(t,\bar{p}_1,\bar{p}_2,\bar{p},\bar{q}_1,\bar{q}_2,\bar{q},\bar{r}_1,\bar{r}_2,\bar{r})| \\ \le C(|p_1-\bar{p}_1| + |p_2-\bar{p}_2| + \sup_{0\le s\le\delta}|p(s)-\bar{p}(s)| + |q_1-\bar{q}_1| + |q_2-\bar{q}_2| + |q-\bar{q}|_{V_1} \\ + |r_1-\bar{r}_1|_{\mathcal{H}} + |r_2-\bar{r}_2|_{\mathcal{H}} + |r-\bar{r}|_{V_2}).\end{aligned} \tag{5.30}$$

Then the BSDE (5.9) *admits a unique solution* $(p(t), q(t), r(t,z))$ *such that*

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

Proof.

Step 1: Assume F is independent of p_1, p_2 *and* p.

In this case condition (5.30) reduces to assumption (5.1). By the Step 1 in the proof of Theorem 5.2, there is a unique solution $(p(t), q(t), r(t,z))$ to equation (5.9).

Step 2: General case.

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

$$\begin{aligned}dp^n(t) =& E[F(t,p^{n-1}(t),p^{n-1}(t+\delta),p^{n-1}_t,q^n(t),q^n(t+\delta),q^n_t,r^n(t,\cdot),r^n(t+\delta,\cdot),r^n_t(\cdot))|\mathcal{F}_t]dt \\ &+ q^n(t)dB_t + r^n(t,z)\tilde{N}(dt,dz),\end{aligned} \tag{5.31}$$

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

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

$$\begin{aligned}&|p^{n+1}(t)-p^n(t)|^2 + \int_t^T\int_{\mathbb{R}}|r^{n+1}(s,z)-r^n(s,z)|^2 ds\nu(dz) + \int_t^T|q^{n+1}(s)-q^n(s)|^2 ds \\ &\quad = -2\int_t^T(p^{n+1}(s)-p^n(s)) \\ &\times [E[F(s,p^n(s),p^n(s+\delta),p^n_s,q^{n+1}(s),q^{n+1}(s+\delta),q^{n+1}_s,r^{n+1}(s,\cdot),r^{n+1}(s+\delta,\cdot),r^{n+1}_s(\cdot)) \\ &- F(s,p^{n-1}(s),p^{n-1}(s+\delta),p^{n-1}_s,q^n(s),q^n(s+\delta),q^n_s,r^n(s,\cdot),r^n(s+\delta,\cdot),r^n_s(\cdot))|\mathcal{F}_s]]ds \\ &- 2\int_t^T(p^{n+1}(s)-p^n(s))(q^{n+1}(s)-q^n(s))dB_s \\ &- \int_t^T\int_{\mathbb{R}}[|r^{n+1}(s,z)-r^n(s,z)|^2 + 2(p^{n+1}(s-)-p^n(s-))(r^{n+1}(s,z)-r^n(s,z))]\tilde{N}(ds,dz)\end{aligned} \tag{5.32}$$

Take conditional expectation with respect to $\mathcal{F}_t$, take the supremum over the interval $[u,T]$ and use the condition (5.30) to get

$$\begin{aligned}
&\sup_{u\le t\le T}|p^{n+1}(t)-p^n(t)|^2+\sup_{u\le t\le T}E\left[\int_t^T|q^{n+1}(s)-q^n(s)|^2ds|\mathcal{F}_t\right]\\
&\quad+\sup_{u\le t\le T}E\left[\int_t^T\int_{\mathbb{R}}|r^{n+1}(s,z)-r^n(s,z)|^2ds\nu(dz)|\mathcal{F}_t\right]\\
&\le C_\varepsilon\sup_{u\le t\le T}E\left[\int_u^T|p^{n+1}(s)-p^n(s)|^2ds|\mathcal{F}_t\right]\\
&\quad+C_1\varepsilon\sup_{u\le t\le T}E\left[\int_u^T|p^n(s)-p^{n-1}(s)|^2ds|\mathcal{F}_t\right]\\
&\quad+C_2\varepsilon\sup_{u\le t\le T}E\left[\int_u^T E[\sup_{s\le v\le T}|p^n(v)-p^{n-1}(v)|^2|\mathcal{F}_s]ds|\mathcal{F}_t\right]\\
&\quad+C_3\varepsilon\sup_{u\le t\le T}E\left[\int_t^T|q^{n+1}(s)-q^n(s)|^2ds|\mathcal{F}_t\right]\\
&\quad+C_4\varepsilon\sup_{u\le t\le T}E\left[\int_t^T\int_{\mathbb{R}}|r^{n+1}(s,z)-r^n(s,z)|^2ds\;\nu(dz)|\mathcal{F}_t\right]
\end{aligned}\tag{5.33}$$

Choosing $\varepsilon>0$ such that $C_3\varepsilon<1$ and $C_4\varepsilon<1$ it follows from (5.33) that

$$\begin{aligned}
\sup_{u\le t\le T}|p^{n+1}(t)-p^n(t)|^2&\le C_\varepsilon\sup_{u\le t\le T}E\left[\int_u^T|p^{n+1}(s)-p^n(s)|^2ds|\mathcal{F}_t\right]\\
&\quad+(C_1+C_2)\varepsilon\sup_{u\le t\le T}E\left[\int_u^T E[\sup_{s\le v\le T}|p^n(v)-p^{n-1}(v)|^2|\mathcal{F}_s]ds|\mathcal{F}_t\right]
\end{aligned}\tag{5.34}$$

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

$$\begin{aligned}
&E\left[\left(\sup_{u\le t\le T}E\left[\int_u^T|p^{n+1}(s)-p^n(s)|^2ds|\mathcal{F}_t\right]\right)^\alpha\right]\le c_\alpha E\left[\left(\int_u^T|p^{n+1}(s)-p^n(s)|^2ds\right)^\alpha\right]\\
&\le c_{T,\alpha}E\left[\int_u^T\sup_{s\le v\le T}|p^{n+1}(v)-p^n(v)|^{2\alpha}ds\right],
\end{aligned}\tag{5.35}$$

and

$$E\left[\left(\sup_{u\le t\le T} E\left[\int_u^T E[\sup_{s\le v\le T}|p^n(v)-p^{n-1}(v)|^2|\mathcal{F}_s]ds|\mathcal{F}_t\right]\right)^\alpha\right]$$
$$\le c_{T,\alpha}E\left[\int_u^T E[\sup_{s\le v\le T}|p^n(v)-p^{n-1}(v)|^{2\alpha}|\mathcal{F}_s]ds\right]$$
$$\le c_{T,\alpha}E\left[\int_u^T \sup_{s\le v\le T}|p^n(v)-p^{n-1}(v)|^{2\alpha}ds\right], \tag{5.36}$$

(5.34), (5.35) and (5.36) yield that for $\alpha > 1$,

$$E\left[\sup_{u\le t\le T}|p^{n+1}(t)-p^n(t)|^{2\alpha}\right] \le C_{1,\alpha}E\left[\int_u^T \sup_{s\le v\le T}|p^{n+1}(v)-p^n(v)|^{2\alpha}ds\right]$$
$$+C_{2,\alpha}E\left[\int_u^T \sup_{s\le v\le T}|p^n(v)-p^{n-1}(v)|^{2\alpha}ds\right] \tag{5.37}$$

Put

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

(5.37) implies that

$$-\frac{d}{dt}(e^{C_{1,\alpha}u}g_{n+1}(u)) \le e^{C_{1,\alpha}u}C_{2,\alpha}g_n(u) \tag{5.38}$$

Integrating (5.38) from t to T we get

$$g_{n+1}(t) \le c_{2,\alpha}\int_t^T e^{C_{1,\alpha}(s-t)}g_n(s)ds \le C_{2,\alpha}e^{C_{1,\alpha}T}\int_t^T g_n(s)ds. \tag{5.39}$$

Iterating the above inequality we obtain that

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

Using 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 equation (5.9). □

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* $E\left[\sup_{T\le t\le T+\delta}|G(t)|^2\right] < \infty$ *and* F *satisfies* (5.30), *i.e.*

$$|F(t,p_1,p_2,p)-F(t,\bar{p}_1,\bar{p}_2,\bar{p})| \le C(|p_1-\bar{p}_1|+|p_2-\bar{p}_2|+\sup_{0\le s\le\delta}|p(s)-\bar{p}(s)|). \tag{5.40}$$

Then the backward stochastic differential equation (5.9) *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:

$$dp^n(t)=E[F(t,p^{n-1}(t),p^{n-1}(t+\delta),p_t^{n-1})|\mathcal{F}_t]dt+q^n(t)dB_t+r^n(t,z)\tilde{N}(dt,dz), \tag{5.41}$$
$$p^n(t)=G(t) \qquad t\in[T,T+\delta].$$

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 conditional expectation with respect to $\mathcal{F}_t$ we get

$$\begin{aligned}&p^{n+1}(t)-p^n(t)\\ =\ &-E[\int_t^T (E[F(s,p^n(s),p^n(s+\delta),p_s^n)|\mathcal{F}_s]\\ &\quad -E[F(s,p^{n-1}(s),p^{n-1}(s+\delta),p_s^{n-1})|\mathcal{F}_s])ds|\mathcal{F}_t]\end{aligned} \tag{5.42}$$

Take the supremum over the interval $[u,T]$ and use the assumption (5.40) to get

$$\begin{aligned}\sup_{u\leq t\leq T}|p^{n+1}(t)-p^n(t)|^2 &\leq C\sup_{u\leq t\leq T}\left(E\left[\int_u^T |p^n(s)-p^{n-1}(s)|ds|\mathcal{F}_t\right]\right)^2\\ &+C\sup_{u\leq t\leq T}\left(E\left[\int_u^T E[\sup_{s\leq v\leq T}|p^n(v)-p^{n-1}(v)||\mathcal{F}_s]ds|\mathcal{F}_t\right]\right)^2\end{aligned} \tag{5.43}$$

By the Martingale Inequality, we have

$$\begin{aligned}&E\left[\left(\sup_{u\leq t\leq T}E\left[\int_u^T |p^n(s)-p^{n-1}(s)|ds|\mathcal{F}_t\right]\right)^2\right]\leq cE\left[\left(\int_u^T |p^n(s)-p^{n-1}(s)|ds\right)^2\right]\\ &\leq c_T E\left[\int_u^T \sup_{s\leq v\leq T}|p^n(v)-p^{n-1}(v)|^2ds\right],\end{aligned} \tag{5.44}$$

and

$$\begin{aligned}&E\left[\left(\sup_{u\leq t\leq T}E\left[\int_u^T E[\sup_{s\leq v\leq T}|p^n(v)-p^{n-1}(v)||\mathcal{F}_s]ds|\mathcal{F}_t\right]\right)^2\right]\\ \leq\ & c_T E\left[\int_u^T E[\sup_{s\leq v\leq T}|p^n(v)-p^{n-1}(v)|^2|\mathcal{F}_s]ds\right],\end{aligned} \tag{5.45}$$

Taking expectation on both sides of (5.43) gives

$$E\left[\sup_{u\leq t\leq T}|p^{n+1}(t)-p^n(t)|^2\right]\leq C\int_u^T E\left[\sup_{s\leq v\leq T}|p^n(v)-p^{n-1}(v)|^2\right]ds \tag{5.46}$$

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 equation (5.9). □

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, $\alpha(t)$ deterministic. Assume that $\int_{\mathbb{R}} \gamma^2(t,z)\nu(dz) < \infty$. Consider a cash flow $X^0(t)$ with dynamics

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

$$X^0(t) = x_0(t) > 0 \; ; \; t \in [-\delta, 0], \tag{6.2}$$

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 $X(t) = X^c(t)$ is

$$dX(t) = [X(t-\delta)\alpha(t) - c(t)]dt + X(t-\delta)\beta(t)dB(t) + X(t-\delta)\int_{\mathbb{R}} \gamma(t,z)\tilde{N}(dt,dz) \; ; \; t \in [0,T] \tag{6.3}$$

$$X(t) = x_0(t) \; ; \; t \in [-\delta, 0]. \tag{6.4}$$

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

$$\begin{aligned} t \to U_1(t,c,\omega) &\text{ is } \mathcal{F}_t\text{-adapted for each } c \geq 0, \\ c \to U_1(t,c,\omega) &\text{ is } \mathcal{C}^1,\ \frac{\partial U_1}{\partial c}(t,c,\omega) > 0, \\ c \to \frac{\partial U_1}{\partial c}&(t,c,\omega) \text{ is strictly decreasing} \\ \lim_{c\to\infty} \frac{\partial U_1}{\partial c}(t,c,\omega) &= 0 \text{ for all } t,\omega \in [0,T]\times\Omega. \end{aligned} \tag{6.5}$$

Put $v_0(t,\omega) = \dfrac{\partial U_1}{\partial c}(t,0,\omega)$ and define

$$I(t,v,\omega) = \begin{cases} 0 & \text{if } v \geq v_0(t,\omega) \\ \left(\dfrac{\partial U_1}{\partial c}(t,\cdot,\omega)\right)^{-1}(v) & \text{if } 0 \leq v < v_0(t,\omega) \end{cases} \tag{6.6}$$

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

$$J(\hat{c}) = \sup\{J(c) \; ; \; c \in \mathcal{A}\} \tag{6.7}$$

where

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

Here $g : \mathbb{R} \to \mathbb{R}$ 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 $E[|g(X(T))|] < \infty$.

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

$$H(t,x,y,a,u,p,q,r(\cdot),\omega) = U_1(t,c,\omega) + (\alpha(t)y - c)p + y\beta(t)q + y\int_{\mathbb{R}} \gamma(t,z)r(z)\nu(dz). \tag{6.8}$$

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

$$\frac{\partial U_1}{\partial c}(t,\hat{c}(t),\omega) = p(t). \tag{6.9}$$

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

$$dp(t) = -E[\Big\{\alpha(t)p(t+\delta) + \beta(t)q(t+\delta) + \int_{\mathbb{R}} \gamma(t,z)r(t+\delta,z)\nu(dz)\Big\}\chi_{[0,T-\delta]}(t)|\mathcal{F}_t]dt + q(t)dB(t) + \int_{\mathbb{R}} r(t,z)\tilde{N}(dt,dz)\ ;\ t\in[0,T] \tag{6.10}$$

$$p(T) = g'(X(T)). \tag{6.11}$$

We solve this BSDE (6.10)-(6.11) recursively by proceeding as in Theorem 5.1:

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

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

which has the solution

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

with corresponding $q(t),\ r(t,z)$ 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 get by Step 1 the BSDE:

$$dp(t) = -E[\alpha(t)p(t+\delta) + \beta(t)q(t+\delta) + \int_{\mathbb{R}} \gamma(t,z)r(t+\delta,z)\nu(dz) \mid \mathcal{F}_t]dt + q(t)dB(t) + \int_{\mathbb{R}} r(t,z)\tilde{N}(dt,dz)\ ;\ t\in[T-2\delta,T-\delta].$$

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 $p(t),\ q(t),\ r(t,\cdot)\ ; t\in[T-2\delta,T-\delta]$. We continue like this by induction up to and including step j, where j is such that $T-j\delta\le 0< T-(j-1)\delta$. With this procedure we end up with a solution $p(t)=p_{X(T)}(t)$ of (6.10)-(6.11) which depends on the (optimal) terminal value $X(T)$. If

$$0\le p(t)\le v_0(t,\omega)\ \text{ for all } t\in[0,T], \tag{6.12}$$

then the optimal consumption rate $\hat{c}(t)$ is by (6.9) given by

$$\hat{c}(t) = \hat{c}_{\hat{X}(T)}(t) = I(t, p(t), \omega);\ t \in [0, T]. \tag{6.13}$$

Substituting this expression for $\hat{c}(t)$ into (6.3) 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 $p(t), q(t), r(t,z)$ be the solution of the BSDE* (6.10)-(6.10)*, as described above. Suppose* (6.12) *holds. Then the optimal consumption rate $\hat{c}(t)$ and the corresponding optimal terminal wealth $X(t)$ are given implicitly by the coupled equations* (6.13) *and* (6.3)-(6.4).

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

Since k is deterministic, we can choose $q = r = 0$ in (6.10)-(6.11) and the BSDE becomes

$$dp(t) = -\alpha(t)p(t+\delta)\chi_{[0,T-\delta]}(t)dt\ ;\ t < T \tag{6.14}$$

$$p(t) = k \text{ for } t \in [T-\delta, T]. \tag{6.15}$$

To solve this we introduce

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

Then

$$\begin{aligned} dh(t) &= -dp(T-t) = \alpha(T-t)p(T-t+\delta)dt \\ &= \alpha(T-t)h(t-\delta)dt \end{aligned} \tag{6.16}$$

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

$$h(t) = p(T-t) = k \text{ for } t \in [-\delta, 0]. \tag{6.17}$$

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

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

$$h(t) = h(j\delta) + \int_0^t h'(s)ds = h(j\delta) + \int_{j\delta}^t \alpha(T-s)h(s-\delta)ds\ ;\ j \in [j\delta, (j+1)\delta]. \tag{6.18}$$

We have proved:

Proposition 6.2 *Assume that $\alpha(t)$ is* deterministic *and $g(x) = kx,\ k > 0$. The optimal consumption rate $\hat{c}_\delta(t)$ for the problem* (6.3)-(6.4), (6.7) *is given by*

$$\hat{c}_\delta(t) = I(t, h_\delta(T-t), \omega), \tag{6.19}$$

where $h_\delta(\cdot) = h(\cdot)$ is determined by (6.17)-(6.18).

Remark 6.3 Assume that $\alpha(t) = \alpha(deterministic) > 0$ for all $t \in [0, T]$. Then we see by induction on (6.18) *that*

$$0 \le \delta_1 < \delta_2 \Rightarrow h_{\delta_1}(t) > h_{\delta_2(t)} \textit{ for all } t \in (0, T]$$

and hence, perhaps suprisingly,

$$0 \le \delta_1 < \delta_2 \Rightarrow \hat{c}_{\delta_1}(t) < \hat{c}_{\delta_2}(t) \textit{ for all } t \in [0, T).$$

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.

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

References

[1] Chen, L., Wu, Z., Maximum principle for the stochastic optimal control problem with delay and application. *Automatica* **46**, pp. 1074-1080, 2010.

[2] Chojnowska-Michalik A., *Representation theorem for general stochastic delay equations*, Bull. Acad. Polon. Sci. Ser. Sci. Math. Astronom. Phys., **26**, 7, pp. 635-642, 1978.

[3] Da Prato G., Zabczyk J., *Stochastic Equations in Infinite Dimensions*, Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge (UK), 1992.

[4] David D., *Optimal control of stochastic delayed systems with jumps*, preprint, 2008.

[5] Diekmann O., Van Gils S.A., Verduyn Lunel S.M., Walther H.O., *Delay Equations.* Functional-, Complex-, and Nonlinear Analysis, Springer-Verlag, New York, 1995.

[6] El Karoui, N., Hamadène, S., *BSDEs and risk-sensitive control, zero-sum and nonzero-sum game problems of stochastic functional differential equations*, Stochastic Processes and their Applications, **107**, pp 145–169, 2003.

[7] Elsanousi I., Øksendal B., Sulem A., *Some solvable stochastic control problems with delay*, Stochastics and Stochastics Reports, **71**, pp 69–89, 2000.

[8] Federico S., *A stochastic control problem with delay arising in a pension fund model*, to appear in Finance and Stochastics, 2009.

[9] Gozzi F., Marinelli C., *Stochastic optimal control of delay equations arising in advertising models*, Da Prato (ed.) et al., Stochastic partial differential equations and applications VII - Papers of the 7th meeting, Levico Terme, Italy, January 5–10, 2004. Boca Raton, FL: Chapman & Hall/CRC. Lecture Notes in Pure and Applied Mathematics, **245**, pp. 133–148, 2004.

[10] Gozzi F., Marinelli C., Savin S., *On controlled linear diffusions with delay in a model of optimal advertising under uncertainty with memory effects*, Journal of Optimization, Theory and Applications, to appear.

[11] Kolmanovski, V.B., Shaikhet, L.E. *Control of Systems with Aftereffect*, American Mathematical Society 1996.

[12] Ikeda, N., Watanabe, S., *Stochastic Differential Equations and Diffusion Processes, Second Edition*, North-Holland/Kodansha 1989.

[13] Larssen B., *Dynamic programming in stochastic control of systems with delay*, Stochastics and Stochastics Reports, **74**, 3-4, pp 651-673, 2002.

[14] Larssen B., Risebro N.H., *When are HJB-equations for control problems with stochastic delay equations finite-dimensional?*, Stochastic Analysis and Applications, **21**, 3, pp. 643-671, 2003.

[15] Øksendal B., Sulem A., *A maximum principle for optimal control of stochastic systems with delay with applications to finance*, Optimal Control and PDE, Essays in Honour of Alain Bensoussan, eds J.L. Menaldi, E. Rofman and A. Sulem, IOS Press, Amsterdam, pp. 64-79, 2001.

[16] Øksendal B., Sulem A., *Applied Stochastic Control of Jump Diffusions. Second Edition*, Springer, 2007.

[17] Øksendal B., Zhang T., *Optimal control with partial information for stochastic Volterra equations*, International Journal of Stochastic Analysis 2010, 25 pages, DOI: 10/1155/2010/329185.

[18] Pardoux E. and Peng S., *Adapted solution of a backward stochastic differential equation*, Systems Control Letter 14, 55-61, 1990.

[19] Peng, S., Yang, Z., *Anticipated backward stochastic differential equations*, The Annals of Probability, **37**,3, pp. 877–902, 2009.

[20] Peszat S., Zabczyk J., *Stochastic Partial Differential Equations with Lévy Noise*, Encyclopedia of Mathematics and its Applications, Vol. 113, Cambridge University Press, Cambridge (UK), 2008.

[21] Protter P.E., *Stochastic Integration and Differential Equations*, 2nd edition, Springer-Verlag, Berlin-Heidelberg-New York, 2003.

[22] Reiss M., Riedle M., Van Gaans O., *Delay differential equations driven by Lévy processes: stationarity and Feller's properties*, Stochastic Processes and their Applications, Vol.116, pp. 1409–1432, 2006.

[23] Situ R., *On solutions of backward stochastic differential equations with jumps and applications*, Stochastic Processes and their Applications 66, 209-236, 1997.

Contents

Centre de recherche INRIA Paris – Rocquencourt
Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Centre de recherche INRIA Bordeaux – Sud Ouest : Domaine Universitaire - 351, cours de la Libération - 33405 Talence Cedex
Centre de recherche INRIA Grenoble – Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier
Centre de recherche INRIA Lille – Nord Europe : Parc Scientifique de la Haute Borne - 40, avenue Halley - 59650 Villeneuve d'Ascq
Centre de recherche INRIA Nancy – Grand Est : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex
Centre de recherche INRIA Rennes – Bretagne Atlantique : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex
Centre de recherche INRIA Saclay – Île-de-France : Parc Orsay Université - ZAC des Vignes : 4, rue Jacques Monod - 91893 Orsay Cedex
Centre de recherche INRIA Sophia Antipolis – Méditerranée : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399