Forward backward stochastic differential equation games with delay and noisy memory

K. R. Dahl†

March 12, 2022

Abstract

The goal of this paper is to study a stochastic game connected to a system of forward backward stochastic differential equations (FBSDEs) involving delay and noisy memory. We derive sufficient and necessary maximum principles for a set of controls for the players to be a Nash equilibrium in the game. Furthermore, we study a corresponding FB-SDE involving Malliavin derivatives. This kind of equation has not been studied before. The maximum principles give conditions for determining the Nash equilibrium of the game. We use this to derive a closed form Nash equilibrium for an economic model where the players maximize their consumption with respect to recursive utility.

Key words: Forward backward stochastic differential equations. Stochastic game. Delay. Noisy memory.

AMS subject classification: 91A05. 91A15. 60H20. 60H10. 60J75. 34K50.

†Department of Mathematics, University of Oslo, Pb. 1053 Blindern, 0316 Oslo, Norway. kristrd@math.uio.no
1 Introduction

The aim of this paper is to study a stochastic game between two players. The game is based on a forward stochastic differential equation (SDE) for the process $X$. In applications to economy, this process can be thought of as the market situation, e.g., the financial market, the housing market or the oil market. This SDE includes two kinds of memory of the past; regular memory and noisy memory. Regular memory (also called delay, see e.g., the survey paper by Ivanov et al. [7]) means that the SDE can depend on previous values of the process $X$. That is, for some given $\delta > 0$, $X(t)$ depends on $X(t - \delta)$. For more on stochastic delay differential equations and optimal control with delay, see Øksendal et al. [18] and Agram and Øksendal [3]. In constrast, noisy memory means that the SDE may involve an Itô integral over previous values of the process, so for $\delta > 0$, $X(t)$ depends on $\int_{t-\delta}^{t} X(s) dB(s)$ where $\{B(s)\}_{s \in [0,T]}$ is a Brownian motion. For more on noisy memory, see Dahl et al. [6].

Connected to this SDE are two backward stochastic differential equations (BSDEs). These BSDEs are connected to the SDE in the sense that they depend on $\{X(t)\}_{t \in [0,T]}$, as well as the delay and noisy memory of this process. Hence, this forms an FBSDE system. Each of these BSDEs corresponds to one of the players in the stochastic game; corresponding to player $i = 1, 2$ is a BSDE in the process $\{W_i(t)\}_{t \in [0,T]}$. The length of memory can be different for the two players, so for $i = 1, 2$, player $i$ has memory span $\delta_i$. The players may also have different levels of information, which is included in the model by having (potentially) different filtrations $\{\mathcal{F}^{(i)}_t\}_{t \in [0,T]}$, $i = 1, 2$.

Each of the players aim to find an optimal control $u_i$ which maximizes their personal performance (objective) function, $J_i$. Seminal work in stochastic optimal control has been done by Krylov and his students, see e.g., Krylov [8] and [9]. The performance function of each of the agents will be defined in such a way that it depends on the player’s profit rate, the market process $X$ and the process $W_i$ coming from the player's BSDE (more on this in Section 2).
FBSDE games with delay and noisy memory

equation (3)). This kind of problem, where both players maximize their performance which depends on an FBSDE, is called an FBSDE stochastic game, and has been studied by e.g. Øksendal and Sulem [14]. However, they do not include memory in their model. We study conditions for a pair of controls \((u_1, u_2)\) to be a Nash equilibrium for such a stochastic game. That is, we would like to determine controls such that the players cannot benefit by changing their actions. In order to do so, we derive sufficient and necessary maximum principles giving conditions for a control to be Nash optimal. This is done in Sections 3 and 4. Maximum principles for forward backward stochastic differential equations (FBSDEs) have been studied by Wang and Wu [23] as well as Øksendal and Sulem [14], but these papers do not consider a stochastic game.

In connection with these maximum principles, there are adjoint equations (see e.g. Øksendal [12] for an introduction to stochastic maximum principles and adjoint equations, or Øksendal and Sulem [15] for maximum principles and adjoint equations where delay is involved). In our case, these adjoint equations are a system of coupled forward backward stochastic differential equations involving Malliavin derivatives (see Di Nunno et al. [5] for more on Malliavin derivatives). To the best of our knowledge, such equations have not been studied before. In Section 5 we study a slightly simplified version of these adjoint FBSDEs, and establish a connection between these equations and a system of FBSDEs without Malliavin derivatives. Finally, in Section 6 we apply our results to a specific example in order to determine the optimal consumption with respect to recursive utility.

2 The problem

Let \((\Omega, \mathcal{F}, P)\) be a probability space, and let \(B(t), t \in [0, T]\) be a Brownian motion in this space. Also, let \(\tilde{N}(t, \cdot)\) be an independent compensated Poisson random measure. Let \((\mathcal{F}_t)_{t \in [0, T]}\) be the \(P\)-augmented filtration generated by \(B(t)\) and \(\tilde{N}(t, \cdot)\).
We will consider a game between two players: player 1 and player 2. Let $u_i(t)$ be the control process chosen by player $i = 1, 2$, and denote $u(t) = (u_1(t), u_2(t))$. Let $A_i$, $i = 1, 2$, denote the set of admissible controls for player $i$ and $A = A_1 \times A_2$.

We consider a controlled forward stochastic differential equation for a process $X(t) = X_u(t, \omega)$, $\omega \in \Omega$, $t \in [0, T]$ determining the market situation (in the following, we omit the $\omega$ for notational ease unless it is important to highlight its dependence):

$$
\begin{align*}
    dX(t) &= b(t, X(t), Y(t), \Lambda(t), u(t), \omega)dt \\
    &\quad + \sigma(t, X(t), Y(t), \Lambda(t), u(t), \omega)dB(t) \\
    &\quad + \int_{\mathbb{R}} \gamma(t^-, X(t^-), Y(t^-), \Lambda(t^-), u(t^-), \zeta, \omega) \tilde{N}(dt, d\zeta)
\end{align*}
$$

(1)

$X(0) = x$

where $Y(t) = (Y_1(t), Y_2(t))$, $\Lambda(t) = (\Lambda_1(t), \Lambda_2(t))$, and $Y_i(t) := X(t - \delta_i)$, $\Lambda_i(t) := \int_{t-\delta_i}^t X(s)dB(s)$, and $\delta_i \geq 0$ for $i = 1, 2$. The superscript $t^-$ means that we are taking the left limit of the process is question (that is, the value before a potential jump at time $t$), see Øksendal and Sulem [13] for more on this.

Here, the delay processes $Y_i$, and the noisy memory processes $\Lambda_i$ correspond to player $i = 1, 2$ respectively. Hence, the two players may have memories for different time intervals, depending on the values of $\delta_i$. Also,

$$
\begin{align*}
    b : [0, T] \times \mathbb{R} \times \mathbb{R}^2 \times \mathbb{R}^2 \times A \times \Omega &\to \mathbb{R}, \\
    \sigma : [0, T] \times \mathbb{R} \times \mathbb{R}^2 \times \mathbb{R}^2 \times A \times \Omega &\to \mathbb{R}, \\
    \gamma : [0, T] \times \mathbb{R} \times \mathbb{R}^2 \times \mathbb{R}^2 \times A \times \Omega &\to \mathbb{R}
\end{align*}
$$

are predictable functions such that for each $u \in A$ the SDE (1) has a unique solution.
Remark 2.1. Existence and uniqueness of solution for the SDE (1) is guaranteed under certain, fairly unrestrictive, assumptions on the coefficient functions, see Dahl et al. [6], Assumption 1, for conditions ensuring existence and uniqueness of solution to (1). This can be seen by viewing equation (1) as a stochastic functional differential equation.

In addition to this, the players (potentially) have different levels of information, represented by different subfiltrations $\mathcal{E}_t^{(i)} \subseteq \mathcal{F}_t$ for all $t \in [0,T], i = 1,2$.

For $i = 1,2$, let $g_i(\cdot, x, y, \Lambda, w_i, z_i, k_i(\cdot), u, \omega)$ be a given predictable process, and let $h_i(x,\omega)$ be an $\mathcal{F}_T$-measurable function. Associated to the FSDE (1), we have a pair of backward stochastic differential equations (BSDEs) in the unknown stochastic processes $(W_i, Z_i, K_i), i = 1,2$:

$$dW_i(t) = -g_i(t, X(t), Y(t), A(t), W_i(t), Z_i(t), K_i(t, \cdot), u(t), \omega)dt + Z_i(t)dB(t) + \int_{\mathbb{R}} K_i(t, \zeta) \tilde{N}(dt, d\zeta)$$

$$W_i(T) = h_i(X(T), \omega).$$

Note that these BSDEs are coupled to the SDE (1) due to the dependency on $X$. Also, the BSDEs depend on the memory of the market process $X$, due to the dependency on the processes $Y$ and $A$. However, equation (2) is a standard BSDE, hence the conditions for existence and uniqueness of solution is well known, see e.g. Pardoux and Peng [20].

For $i = 1,2$, let $f_i : [0,T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathcal{A} \times \Omega \to \mathbb{R}, \varphi_i : \mathbb{R} \to \mathbb{R}, \psi_i : \mathbb{R} \to \mathbb{R}$ be functions representing a profit rate, bequest function and risk evaluation. Then, the performance function of each player $i = 1,2$ is defined by:

$$J_i(u) = E\left[\int_0^T f_i(t, X^u(t), Y_i^u(t), A_i^u(t), u_i(t))dt + \varphi_i(X^u(T)) + \psi_i(W_i^u(0))\right]$$

where we must assume all conditions necessary for the integrals and the expec-
Also, note that the performance $J_i$ of player $i$ is a function of the control $u(t) = (u_1(t), u_2(t))$, which is determined by both players. Therefore, this problem setting specifies a stochastic game.

A pair of controls $(\hat{u}_1, \hat{u}_2)$ is called a Nash equilibrium for this stochastic game if the following holds:

$$J_1(u_1, \hat{u}_2) \leq J_1(\hat{u}_1, \hat{u}_2) \text{ for all } u_1 \in A_1$$

$$J_2(\hat{u}_1, u_2) \leq J_2(\hat{u}_1, \hat{u}_2) \text{ for all } u_2 \in A_2. \quad (4)$$

In words, this means that in the Nash equilibrium, neither player would like to change their control.

Assume there exists a Nash equilibrium for this forward-backward stochastic differential (FBSDE) game with delay and noisy memory. We would like to find this Nash equilibrium, and we will do so by proving sufficient and necessary maximum principles for this problem. Therefore, we define a Hamiltonian function for each player $i = 1, 2$ as follows:

$$H_i(t, x, y, A, w_i, z_i, k_i, u_1, u_2, \lambda_i, p_i, q_i, r_i) = f_i(t, y, A, u_i)$$

$$+ \lambda_i g_i(t, x, y, A, w_i, z_i, k_i, u_1, u_2) + p_i b(t, x, y, A, u_1, u_2)$$

$$+ q_i \sigma(t, x, y, A, u_1, u_2) + \int_{\mathbb{R}} r_i(\zeta) \gamma(t, x, y, A, u_1, u_2, \zeta) \nu(d\zeta). \quad (5)$$

Assume $H_i$ is $C^1$ in $x, y_1, y_2, A_1, A_2, w_i, z_i, k_i, u_1, u_2$ for $i = 1, 2$. In the following, for ease of notation, we will use the abbreviation

$$H_i(t) = H_i(t, x, y, A, w_i, z_i, k_i, u_1, u_2, \lambda_i, p_i, q_i, r_i)$$

For $i = 1, 2$, we define a system of FBSDEs associated to these Hamiltonians in the unknown adjoint processes $(\lambda_i, p_i, q_i, r_i)$:
FSDE in $\lambda_i$ (which depends on $p_i, q_i, r_i$):

$$
\begin{align*}
\frac{d\lambda_i(t)}{dt} &= \frac{\partial H_i}{\partial \omega_i}(t)dt + \frac{\partial H_i}{\partial y_i}(t)dB(t) + \int_{\mathbb{R}} \nabla_{k_i}(H_i(t, \zeta)) \tilde{N}(dt, d\zeta) \\
\lambda_i(0) &= \psi_i'(W_i(0)).
\end{align*}
$$

(6)

where $\nabla_{k_i}(H_i(t, \zeta))$ is the Fréchet derivative of $H_i$ at $k_i$, see the appendix in Øksendal and Sulem [14] for a closer explanation of this gradient.

We also define a BSDE in $p_i, q_i, r_i$, which depends on $\lambda_i$:

$$
\begin{align*}
\frac{dp_i(t)}{dt} &= E[\mu_i(t) | \mathcal{F}_t]dt + q_i(t)dB(t) + \int_{\mathbb{R}} r_i(t, \zeta) \tilde{N}(dt, d\zeta) \\
p_i(T) &= \varphi_i'(X(T)) + h_i'(X(T))\lambda_i(T)
\end{align*}
$$

(7)

where

$$
\mu_i(t) = -\frac{\partial H_i}{\partial x_i}(t) - \frac{\partial H_i}{\partial y_i}(t + \delta_i)1_{[0,T-\delta_i]}(t) - \int_t^{t+\delta_i} D_t\frac{\partial H_i}{\partial \Lambda_i}(s)1_{[0,T]}(s)ds
$$

and $D_t[\cdot]$ denotes the Malliavin derivative (see Remark 2.2). Note that the conditional expectation in (7) is well defined by the extension of the Malliavin derivative introduced by Aase et al. [1], see Remark 2.2. Equations (6)-(7) form an FBSDE-system involving Malliavin derivatives. To the best of our knowledge, such systems have not been studied before.

**Remark 2.2** We refer to Nualart [11], Sanz-Solè [22] and Di Nunno et al. [5] for information about the Malliavin derivative $D_t$ for Brownian motion $B(t)$ and, more generally, Lévy processes. In Aase et al. [1], $D_t$ was extended from the space $\mathbb{D}_{1,2}$ to $L^2(P)$, where $\mathbb{D}_{1,2}$ denotes the classical space of Malliavin differentiable $\mathcal{F}_T$-measurable random variables. The extension is such that for all $F \in L^2(\mathcal{F}_T, P)$, the following holds:

(i) $D_tF \in (S)^*$, where $(S)^* \subseteq L^2(P)$ denotes the Hida space of stochastic distributions,

(ii) the map $(t, \omega) \rightarrow E[D_tF | \mathcal{F}_t]$ belongs to $L^2(\mathcal{F}_T, \lambda \times P)$, where $\lambda$ denotes
the Lebesgue measure on \([0, T]\).

Moreover, the following generalized Clark-Ocone theorem holds:

\[(iii)\]
\[F = E[F] + \int_0^T E[D_t F | \mathcal{F}_t] dB(t). \quad (8)\]

See [1], Theorem 3.11, and also [5], Theorem 6.35.

Notice that by combining Itô’s isometry with the Clark-Ocone theorem, we obtain

\[
E\left[ \int_0^T E[D_t F | \mathcal{F}_t]^2 dt \right] = E\left[ \left( \int_0^T E[D_t F | \mathcal{F}_t] dB(t) \right)^2 \right] = E[F^2 - E[F]^2] \quad (9)
\]

(iv) As observed in Agram et al. [2], we can also apply the Clark-Ocone theorem to show the following generalized duality formula:

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

\[
E\left[ F \int_0^T \varphi(t) dB(t) \right] = E\left[ \int_0^T E[D_t F | \mathcal{F}_t] \varphi(t) dt \right] \quad (10)
\]

**Remark 2.3** Note that equation (6) is linear in \(\lambda_i\), and hence, if \(p_i, q_i, r_i\) were given, it could be solved by using the Itô formula. However, this solution will depend on the processes \(X, Y_i, \Lambda_i\) and \(W_i\), so in order to find an explicit solution for \(\lambda_i\), we must also solve the coupled FBSDE system (1)-(2).

The BSDE(7) is linear in \(p_i\), and hence, if \(\lambda_i\) was given, it would be possible to find a unique solution to this equation by using e.g. Proposition 6.2.1 in Pham [17] or Theorem 1.7 in Øksendal and Sulem [17]. However, as for the adjoint SDE (6), this solution will depend on the coupled FBSDE system (1)-(2).

In the remaining part of the paper, we will prove a sufficient (Section 3) and a necessary maximum principle (Section 4) for this kind of FBSDE game with delay and noisy memory. Then, we will study existence and uniqueness of solutions of the FBSDE system (6)-(7) (Section 5). Finally, we will present an
example which illustrates our results: optimal consumption rate with respect to recursive utility (see Section 3).

3 Sufficient maximum principle for FBSDE games with delay and noisy memory

We prove a sufficient maximum principle which roughly states that under concavity conditions, a control \((\hat{u}_1, \hat{u}_2)\) satisfying a conditional maximum principle and an \(L^2\)-condition is a Nash equilibrium for the stochastic game.

Theorem 3.1 Let \(\hat{u}_1 \in A_1\) and \(\hat{u}_2 \in A_2\) with corresponding solutions \(\hat{X}(t), \hat{Y}(t), \hat{A}(t), \hat{W}(t), \hat{Z}(t), \hat{K}(t, \cdot)\), \(\hat{\lambda}_1(t), \hat{\lambda}_2(t)\), \(\hat{\phi}_1(t), \hat{\phi}_2(t)\), \(\hat{r}_1(t, \cdot)\) and \(\hat{r}_2(t, \cdot)\) of the FSDE (1), the BSDE (2), and the FBSDE systems (3) - (7) for \(i = 1, 2\). Also, assume that:

- (Concavity I) The functions \(x \rightarrow h_i(x), x \rightarrow \varphi_i(x), x \rightarrow \psi_i(x)\) are concave for \(i = 1, 2\).
- (The conditional maximum principle)

\[
\text{ess sup}_{\nu \in \mathcal{A}_1} E[H_1(t, \hat{X}(t), \hat{Y}(t), \hat{A}(t), \hat{W}(t), \hat{Z}_1(t), \hat{K}_1(t, \cdot),
\hat{\lambda}_1(t), \hat{\phi}_1(t), \hat{r}_1(t, \cdot))|\mathcal{F}_t^{(1)}] \\
= E[H_1(t, \hat{X}(t), \hat{Y}(t), \hat{A}(t), \hat{W}_1(t), \hat{Z}_1(t), \hat{K}_1(t, \cdot),
\hat{u}_1(t), \hat{u}_2(t), \hat{\lambda}_1(t), \hat{\phi}_1(t), \hat{r}_1(t, \cdot))|\mathcal{F}_t^{(1)}] \\
\]

and similarly

\[
\text{ess sup}_{\nu \in \mathcal{A}_2} E[H_2(t, \hat{X}(t), \hat{Y}(t), \hat{A}(t), \hat{W}_2(t), \hat{Z}_2(t), \hat{K}_2(t, \cdot),
\hat{\lambda}_2(t), \hat{\phi}_2(t), \hat{r}_2(t, \cdot))|\mathcal{F}_t^{(2)}] \\
= E[H_2(t, \hat{X}(t), \hat{Y}(t), \hat{A}(t), \hat{W}_2(t), \hat{Z}_2(t), \hat{K}_2(t, \cdot),
\hat{u}_1(t), \hat{u}_2(t), \hat{\lambda}_2(t), \hat{\phi}_2(t), \hat{r}_2(t, \cdot))|\mathcal{F}_t^{(2)}].
\]
(Concavity II) The functions

\[
\hat{H}_1(t, x, y, \Lambda, w, z, k) := \text{ess sup } v \in A E[H_1(t, x, y, \hat{y}, \hat{\Lambda}, w, z, k, v, \hat{u}, \hat{\lambda}, \hat{p}, \hat{q}, \hat{r}) | E(t)]
\]

and

\[
\hat{H}_2(t, x, y, \Lambda, w, z, k) := \text{ess sup } v \in A E[H_2(t, x, y, \hat{y}, \hat{\Lambda}, w, z, k, v, \hat{u}, \hat{\lambda}, \hat{p}, \hat{q}, \hat{r}) | E(t)]
\]

are concave for all \( t \) a.s.

Finally, assume that the following \( L^2 \) conditions hold:

\[
E\left[ \int_0^T \left\{ \tilde{p}^2(t) \left[ (\sigma(t) - \sigma(t))^2 + \int_\mathbb{R} (\hat{r}(t, \zeta) - \hat{\hat{r}}(t, \zeta))^2 \nu(d\zeta) \right] \\
+ (\tilde{X}(t) - \hat{\tilde{X}}(t))^2 \tilde{q}^2(t) + \int_\mathbb{R} \tilde{r}^2(t, \zeta) \nu(d\zeta) \right] \\
+ (\tilde{Y}(t) - \hat{\tilde{Y}}(t))^2 \left( \frac{\partial \hat{H}_i}{\partial z} \right)^2(t) + \int_\mathbb{R} ||\nabla_k \hat{H}_i(t, \zeta)||^2 \nu(d\zeta) \\
+ \hat{\lambda}^2(t) \left[ (\Lambda_i(t) - \hat{\Lambda}_i(t))^2 + \int_\mathbb{R} (K_i(t, \zeta) - \hat{K}_i(t, \zeta))^2 \nu(d\zeta) \right] \right\} < \infty
\]

for \( i = 1, 2 \).

Then, \((\hat{u}_1, \hat{u}_2)\) is a Nash equilibrium.

Proof. We would like to show that \( J_1(u_1, \hat{u}_2) \leq J_1(\hat{u}_1, \hat{u}_2) \) for all \( u_1 \in A_1 \).

Choose \( u_1 \in A_1 \). By the definition of the performance function \( J_1 \),

\[
\delta := J_1(u_1, \hat{u}_2) - J_1(\hat{u}_1, \hat{u}_2) = I_1 + I_2 + I_3
\]

where

\[
I_1 = E\left[ \int_0^T \left\{ f_1(t, x, y, \Lambda, u) - f_1(t, \hat{x}, \hat{y}, \hat{\Lambda}, \hat{u}) \right\} dt \right],
\]

\[
I_2 = E[\varphi_1(X(T)) - \varphi_1(\hat{X}(T))],
\]

\[
I_3 = E\left[ \int_0^T \left\{ \tilde{p}^2(t) \left[ (\sigma(t) - \sigma(t))^2 + \int_\mathbb{R} (\hat{r}(t, \zeta) - \hat{\hat{r}}(t, \zeta))^2 \nu(d\zeta) \right] \\
+ (\tilde{X}(t) - \hat{\tilde{X}}(t))^2 \tilde{q}^2(t) + \int_\mathbb{R} \tilde{r}^2(t, \zeta) \nu(d\zeta) \right] \\
+ (\tilde{Y}(t) - \hat{\tilde{Y}}(t))^2 \left( \frac{\partial \hat{H}_i}{\partial z} \right)^2(t) + \int_\mathbb{R} ||\nabla_k \hat{H}_i(t, \zeta)||^2 \nu(d\zeta) \\
+ \hat{\lambda}^2(t) \left[ (\Lambda_i(t) - \hat{\Lambda}_i(t))^2 + \int_\mathbb{R} (K_i(t, \zeta) - \hat{K}_i(t, \zeta))^2 \nu(d\zeta) \right] \right\}
\]
\[ I_3 = E[\psi_1(W_1(0)) - \psi_1(\hat{W}_1(0))]. \]

Note that from the definition of the Hamiltonian,

\[
I_1 = E[\int_0^T (H_1(t) - \hat{H}_1(t) - \dot{\lambda}_1(t)(g_1(t) - \hat{g}_1(t)) - \dot{p}_1(t)(b(t) - \hat{b}(t)) - \dot{q}_1(t)(\sigma(t) - \hat{\sigma}(t)) - \int_\mathbb{R} \hat{r}_1(t, \zeta)(\gamma(t, \zeta) - \hat{\gamma}(t, \zeta)\nu(d\zeta))\}dt]
\]

where we have used the abbreviation

\[ \hat{H}_1(t) := H_1(t, \hat{X}(t), \hat{Y}(t), \hat{\Delta}(t), \hat{\hat{W}}_1(t), \hat{\hat{Z}}_1(t), \hat{\hat{K}}_1(t, \cdot), \hat{u}, \hat{\lambda}_1, \hat{p}_1, \hat{q}_1, \hat{r}_1, \omega) \]

and corresponding abbreviations for \( H_1(t), b(t), \hat{b}(t), \sigma, \hat{\sigma}(t), \gamma(t) \) and \( \hat{\gamma}(t) \).

Also,

\[
I_2 = E[\varphi_1(X(T)) - \varphi_1(\hat{X}(T))]
\]

\[ \leq E[\varphi_1'(\hat{X}(T))(X(T) - \hat{X}(T))] \]

\[ = E[\dot{p}_1(T)(X(T) - \hat{X}(T))]|_{t=0}^T \]

\[ = E[\dot{q}_1(T)(\sigma(t) - \hat{\sigma}(t)) dt + \int_0^T \hat{r}_1(t, \zeta)(\gamma(t, \zeta) - \hat{\gamma}(t, \zeta)\nu(d\zeta))dt] \]

\[ - E[\hat{\lambda}_1(T)\hat{h}_1'(\hat{X}(T))(X(T) - \hat{X}(T))] \]

\[ = E[\int_0^T \dot{p}_1(t)(b(t) - \hat{b}(t))dt + \int_0^T (X(t) - \hat{X}(t))\frac{\partial \hat{H}_1}{\partial x}(t)\]

\[ + \int_0^T \dot{q}_1(t)(\sigma(t) - \hat{\sigma}(t)) dt + \int_0^T \hat{r}_1(t, \zeta)(\gamma(t, \zeta) - \hat{\gamma}(t, \zeta)\nu(d\zeta)dt] \]

\[ - E[\hat{\lambda}_1(T)\hat{h}_1'(\hat{X}(T))(X(T) - \hat{X}(T))] \]

where the first inequality follows from the concavity of \( \varphi_1 \), the second equality follows from equation (7), the fourth equality from Itô’s product rule applied to
the fifth equality follows from equation (7), the double expectation rule and equation (11).

Also, note that

\[ I_3 = E[\psi_1(W_1(0)) - \psi_1(\hat{W}_1(0))] \]
\[ \leq E[\psi_1'(\hat{W}_1(0))(W_1(0) - \hat{W}_1(0))] \]
\[ = E[\lambda_1(T)(W_1(T) - \hat{W}_1(T))] - \{ E[\int_0^T (W_1(t) - \hat{W}_1(t))d\lambda_1(t)\] 
\[ + \int_0^T \lambda_1(t)(dW_1(t) - d\hat{W}_1(t)) + \int_0^T \frac{\partial H_1}{\partial z_1}(t)(Z_1(t) - \hat{Z}_1(t))dt \] 
\[ + \int_0^T \int_{\mathbb{R}} \nabla_k H_1(t)(K_1(t) - \hat{K}_1(t))\nu(d\zeta)dt]\} \]
\[ = E[\lambda_1(T)(h_1(X(T)) - h_1(\hat{X}(T))) - \{ E[\int_0^T \frac{\partial H_1}{\partial w_1}(t)(W_1(t) - \hat{W}_1(t))dt \] 
\[ + \int_0^T \lambda_1(t)(-g_1(t) + \hat{g}_1(t))dt + \int_0^T \frac{\partial H_1}{\partial z_1}(t)(Z_1(t) - \hat{Z}_1(t))dt \] 
\[ + \int_0^T \int_{\mathbb{R}} \nabla_k \hat{H}_1(t)(K_1(t) - \hat{K}_1(t))\nu(d\zeta)dt]\} \]
\[ \leq E[\lambda_1(T)k_1(\hat{X}(T))(X(T) - \hat{X}(T))] - \{ E[\int_0^T \frac{\partial H_1}{\partial w_1}(t)(W_1(t) - \hat{W}_1(t))dt \] 
\[ + \int_0^T \lambda_1(t)(-g_1(t) + \hat{g}_1(t))dt + \int_0^T \frac{\partial H_1}{\partial z_1}(t)(Z_1(t) - \hat{Z}_1(t))dt \] 
\[ + \int_0^T \int_{\mathbb{R}} \nabla_k \hat{H}_1(t)(K_1(t) - \hat{K}_1(t))\nu(d\zeta)dt]\} \]
\[ \leq E[\lambda_1(T)k_1(\hat{X}(T))(X(T) - \hat{X}(T))] - \{ E[\int_0^T \frac{\partial H_1}{\partial w_1}(t)(W_1(t) - \hat{W}_1(t))dt \] 
\[ + \int_0^T \lambda_1(t)(-g_1(t) + \hat{g}_1(t))dt + \int_0^T \frac{\partial H_1}{\partial z_1}(t)(Z_1(t) - \hat{Z}_1(t))dt \] 
\[ + \int_0^T \int_{\mathbb{R}} \nabla_k \hat{H}_1(t)(K_1(t) - \hat{K}_1(t))\nu(d\zeta)dt]\} \]
(13)

where the first inequality follows from the concavity of \( \psi_1 \), the second equality follows from equation (5), the third equality follows from Itô’s product rule applied to \( \lambda_1 Y_1 \) and \( \hat{\lambda}_1 Y_1 \), the fourth equality follows from equation (4) as well as equation (11). The final inequality follows from the concavity of \( h_1 \) and that \( \hat{\lambda}_1(T) \geq 0 \).

Hence,
\[
\Delta = I_1 + I_2 + I_3 \\
\leq E\left[\int_0^T \{H_1(t) - \hat{H}_1(t) - \left(\frac{\partial H_1}{\partial w}\right)(t) + \frac{\partial \hat{H}_1}{\partial \nu}(t + \delta_1)1_{[0,T-\delta_1]}(t)\right] \\
+ \int_t^{t+\delta_1} D_t\left(\frac{\partial H_1}{\partial \nu}(s)\right)1_{[0,T]}(s)ds \right] (X(t) - \hat{X}(t))dt \right) \\
- \int_0^T \left(\frac{\partial \hat{H}_1}{\partial \nu}(t)(W_1(t) - \hat{W}_1(t)) + \frac{\partial \hat{H}_1}{\partial \nu}(t)(Z_1(t) - \hat{Z}_1(t))\right) \\
+ \int_\mathcal{F}_t \nabla_{\delta_1} \hat{H}_1(t)(K_1(t, \zeta) - \hat{K}_1(t, \zeta)) \nu(d\zeta) \right] dt].
\]

Note that by changing the order of integration and using the duality formula
for Malliavin derivatives (see Di Nunno et al., [5]), we get:

\[
E\left[\int_0^T \frac{\partial H_1}{\partial X}(s)(\Lambda_1(s) - \hat{\Lambda}_1(s))ds\right] \\
= E\left[\int_0^T \frac{\partial H_1}{\partial X}(s) \int_{s-\delta_1}^s (X(t) - \hat{X}(t))dB(t)ds\right] \\
= \int_0^T E\left[\frac{\partial H_1}{\partial X}(s) \int_{s-\delta_1}^s (X(t) - \hat{X}(t))dB(t)\right]ds \\
= \int_0^T E\left[\int_{s-\delta_1}^s E[D_t\left(\frac{\partial H_1}{\partial X}(s)\right)|\mathcal{F}_s](X(t) - \hat{X}(t))dt\right]ds \\
= E\left[\int_0^T \int_{s-\delta_1}^{s+\delta_1} E[D_t\left(\frac{\partial H_1}{\partial X}(s)\right)|\mathcal{F}_s]1_{[0,T]}(s)ds(X(t) - \hat{X}(t))dt\right] \\
= E\left[\int_0^T \int_{s-\delta_1}^{s+\delta_1} D_t\left(\frac{\partial H_1}{\partial X}(s)\right)1_{[0,T]}(s)ds(X(t) - \hat{X}(t))dt\right].
\]

Also, note that

\[
E\left[\int_0^T \frac{\partial H_1}{\partial \nu}(s)(Y_1(s) - \hat{Y}_1(s))ds\right] \\
= E\left[\int_0^T \frac{\partial H_1}{\partial \nu}(t)(X(t) - \delta) - \hat{X}(t - \delta_1))dt\right] \\
= E\left[\int_0^T \frac{\partial H_1}{\partial \nu}(t + \delta_1)1_{[0,T-\delta_1]}(t)(X(t) - \hat{X}(t)) dt\right].
\]

Hence, by the inequality (14) combined with equations (15) and (16),

13
FBSDE games with delay and noisy memory

$$\Delta \leq E\left[\int_0^T \left( H_1(t) - \hat{H}_1(t) - \frac{\partial H_1}{\partial x}(t)(X(t) - \hat{X}(t)) - \frac{\partial H_1}{\partial y_1}(t)(Y_1(t) - \hat{Y}_1(t)) \right. \right.$$ 
$$-\frac{\partial H_1}{\partial \Lambda_1}(t)(\Lambda_1(t) - \hat{\Lambda}_1(t)) dt - \frac{\partial H_1}{\partial w_1}(t)(W_1(t) - \hat{W}_1(t)) - \frac{\partial H_1}{\partial z_1}(t)(Z_1(t) - \hat{Z}_1(t))$$ 
$$+ \int_{\mathbb{R}} \nabla_k \hat{H}_1(t)(K_1(t, \zeta) - \hat{K}_1(t, \zeta)) \nu(d\zeta) dt \left. \right] \right]. \tag{17}$$

By assumption, $\hat{H}_1$ is concave, so it is superdifferentiable (see Rockafellar [21]) at the point $\vec{x} := (\hat{X}, \hat{Y}_1, \hat{\Lambda}_1, \hat{W}_1, \hat{Z}_1, \hat{K}_1)$. Thus, there exists a super-gradient $\vec{a} := (a_0, a_1, a_2, a_3, a_4, a_5)$ such that for all $\vec{y} := (x, y, \Lambda, w, z, k)$, the following holds:

$$\hat{H}_1(\vec{x}) + \vec{a} \cdot (\vec{y} - \vec{x}) \geq \hat{H}_1(\vec{y}). \tag{18}$$

Define

$$\phi_1(x, y, \Lambda, w, z, k) := \hat{H}_1(x, y, \Lambda, w, z, k) - \hat{H}_1(\hat{X}, \hat{Y}_1, \hat{\Lambda}_1, \hat{W}_1, \hat{Z}_1, \hat{K}_1)$$
$$- \left\{ a_0(x - \hat{X}) + a_1(y - \hat{Y}_1) + a_2(\Lambda - \hat{\Lambda}_1) + a_3(w - \hat{W}_1) + a_4(z - \hat{Z}_1) \right.$$ 
$$+ \int_{\mathbb{R}} a_5(\zeta)(k - \hat{K}_1) \nu(d\zeta) \right\}. \tag{19}$$

Then, by equation (18)

$$\phi_1(x, y, \Lambda, w, z, k) \leq 0 \text{ for all } x, y, \Lambda, w, z, k, \text{ and } \phi_1(\hat{X}, \hat{Y}_1, \hat{\Lambda}_1, \hat{W}_1, \hat{Z}_1, \hat{K}_1) = 0 \text{ (by definition).} \tag{20}$$

Therefore, by differentiating equation (19) and using equation (20), we find that

*Defined similarly as subdifferentiability for convex functions.*
Therefore, it follows from this, equation (17) and equation (20) that

\[ \Delta = \phi(X(t), Y_1(t), \Lambda_1(t), W_1(t), Z_1(t), K_1(t, \cdot)) \leq 0 \]

where the final inequality follows since \( \hat{H}_1 \) is concave.

This means that \( J_1(u_1, \hat{u}_2) \leq J_1(\hat{u}_1, \hat{u}_2) \) for all \( u_1 \in A_1 \).

In a similar way, one can prove that \( J_2(\hat{u}_1, u_2) \leq J_2(\hat{u}_1, \hat{u}_2) \) for all \( u_2 \in A_2 \).

This completes the proof that \( (\hat{u}_1, \hat{u}_2) \) is a Nash-equilibrium.

\[ \square \]

4 Necessary maximum principle for FBSDE games with delay and noisy memory

In the following, we need some additional assumptions and notation:

- For all \( t_0 \in [0, T] \) and all bounded \( \mathcal{E}(t) \)-measurable random variables \( \alpha_i(\omega) \), the control

\[ \beta_i(t) := 1_{(t_0, T)}(t)\alpha_i(\omega) \text{ is in } A_i \text{ for } i = 1, 2. \]

- For all \( u_i, \beta_i \in A_i \) with \( \beta_i \) bounded, there exists \( \kappa_i > 0 \) such that the control
belongs to $A_i$ for all $s \in (-\kappa_i, \kappa_i)$, $i = 1, 2$.

• Also, assume that the following derivative processes exist and belong to $L^2([0, T] \times \Omega)$:

\[
\begin{align*}
x_1(t) &= \frac{d}{ds} X^{(u_1 + s\beta_1, u_2)}(t)|_{s=0}, \\
y_1(t) &= \frac{d}{ds} Y_{1}^{(u_1 + s\beta_1, u_2)}(t)|_{s=0}, \\
\Lambda_1(t) &= \frac{d}{ds} \Lambda_{1}^{(u_1 + s\beta_1, u_2)}(t)|_{s=0}, \\
w_1(t) &= \frac{d}{ds} W_{1}^{(u_1 + s\beta_1, u_2)}(t)|_{s=0}, \\
z_1(t) &= \frac{d}{ds} Z_{1}^{(u_1 + s\beta_1, u_2)}(t)|_{s=0}, \\
k_1(t) &= \frac{d}{ds} K_{1}^{(u_1 + s\beta_1, u_2)}(t)|_{s=0}, \\
\end{align*}
\]

and similarly for $x_2(t) = \frac{d}{ds} X^{(u_1, u_2 + s\beta_2)}(t)|_{s=0}$ etc. Notice that $x_i(0) = 0$ for $i = 1, 2$ since $X(0) = x$.

If these assumptions hold, we can prove a necessary maximum principle for our noisy memory FBSDE game. The proof of the following theorem is based on the same idea as the proof of Theorem 2.2 in Øksendal and Sulem [14], however the presence of noisy memory in our problem requires some extra care.

**Theorem 4.1** Suppose that $u \in A$ with corresponding solutions $X(t), Y_i(t), \Lambda_i(t), W_i(t), Z_i(t), K_i(t, \zeta), \lambda_i(t), p_i(t), q_i(t), r_i(t, \zeta), i = 1, 2$, of equations (1), (2), (6) and (7). Also, assume that conditions (21) - (23) hold. Then, the following are equivalent:

(i) $\frac{\partial}{\partial s} J_1(u_1 + s\beta_1, u_2)|_{s=0} = \frac{\partial}{\partial s} J_2(u_1, u_2 + s\beta_2)|_{s=0} = 0$ for all bounded $\beta_1 \in A_1, \beta_2 \in A_2$.

(ii) $E[\frac{\partial H_1(t, X(t), Y(t), \Lambda(t), W(t), Z(t), K(t, \zeta), \lambda(t), p(t), q(t), r(t, \zeta))}{\partial v_1}]|_{v_1 = u_1(t)}$
FBSDE games with delay and noisy memory

\[
\begin{align*}
E\left[ \frac{\partial H_2}{\partial v_2} (t, X(t), Y(t), \Lambda(t), \omega_2(t), \xi_2(t), z_2(t), k_2(t, \cdot), u_2(t)) \right]_{v_2=u_2(t)} \\
= 0.
\end{align*}
\]

Proof. We only prove that \( \frac{\partial}{\partial s} J_1 (u_1 + s \beta_1, u_2) \big|_{s=0} = 0 \) for all bounded \( \beta_1 \in A_1 \) is equivalent to

\[
E\left[ \frac{\partial H_1}{\partial v_1} (t, X(t), Y(t), \Lambda(t), \omega_1(t), \xi_1(t), z_1(t), k_1(t, \cdot), u_1(t), \lambda_1(t), p_1(t), q_1(t), r_1(t)) \right]_{v_1=u_1(t)} = 0.
\]

The remaining part of the theorem (i.e., the same statement for \( J_2 \) and \( H_2 \)) is proved in a similar way.

Note that, by the definition of \( J_1 \) and by interchanging differentiation and integration,

\[
D_1 := \frac{\partial}{\partial s} J_1 (u_1 + s \beta_1, u_2) \big|_{s=0} = E\left[ \int_0^T \left\{ \frac{\partial f_1}{\partial x_1} (t) x_1(t) + \frac{\partial f_1}{\partial y_1} (t) y_1(t) + \frac{\partial f_1}{\partial \Lambda_1} (t) \tilde{\Lambda}_1(t) \frac{\partial f_1}{\partial u_1} (t) \beta_1(t) \right\} dt \\
+ \varphi'_1(X(T)) x_1(T) + \varphi'_1(W_1(0)) x_1(0) \right].
\]

We study the different parts of \( D_1 \) separately. First, by the Itô product rule, the adjoint BSDE \( \Box \) and the definition of \( x_1(t) \),
\[ I_1 := E[\varphi_1'(X(T))x_1(T)] \]
\[ = E[p_1(T)x_1(T)] - E[b_1'(X(T))\lambda_1(T)x_1(T)] \]
\[ = E[p_1(0)x_1(0)] + E}\int_0^T p_1(t)dx_1(t) + \int_0^T x_1(t)dp_1(t) + \int_0^T d[p_1, x_1](t) - E[b_1'(X(T))\lambda_1(T)x_1(T)] \]
\[ = E}\int_0^T p_1(t)\left(\frac{\partial}{\partial x_1}(t)x_1(t) + \frac{\partial}{\partial y_1}(t)y_1(t) + \frac{\partial}{\partial \lambda_1}(t)\tilde{\Lambda}_1(t) + \frac{\partial}{\partial \beta_1}(t)\beta_1(t)\right)dt \]
\[ + E}\int_0^T x_1(t)E[p_1(t)|F_t]dt \]
\[ + E}\int_0^T q_1(t)\left(\frac{\partial}{\partial x_1}(t)x_1(t) + \frac{\partial}{\partial y_1}(t)y_1(t) + \frac{\partial}{\partial \lambda_1}(t)\tilde{\Lambda}_1(t) + \frac{\partial}{\partial \beta_1}(t)\beta_1(t)\right)dt \]
\[ + E}\int_0^T \int_\mathbb{R} r_1(t, \zeta)\left(\frac{\partial}{\partial x_1}(t)x_1(t) + \frac{\partial}{\partial y_1}(t)y_1(t) + \frac{\partial}{\partial \lambda_1}(t)\tilde{\Lambda}_1(t) + \frac{\partial}{\partial \beta_1}(t)\beta_1(t)\right)d\nu(\zeta)dt \]
\[ - E[b_1'(X(T))\lambda_1(T)x_1(T)]. \]
\[ (24) \]

Also, by the FSDE (3), the BSDE (2), the definition of \( x_1(t) \) and the Itô product rule,

\[ I_2 := E[\varphi_1'(W_1(0))w_1(0)] \]
\[ = E[\lambda_1(0)x_1(0)] \]
\[ = E[\lambda_1(T)x_1(T)] - E}\int_0^T \lambda_1(t)dw_1(t) + \int_0^T w_1(t)d\lambda_1(t) \]
\[ + \int_0^T z_1(t)\left(\frac{\partial}{\partial x_1}(t)x_1(t) + \frac{\partial}{\partial y_1}(t)y_1(t) + \frac{\partial}{\partial \lambda_1}(t)\tilde{\Lambda}_1(t) + \frac{\partial}{\partial \beta_1}(t)\beta_1(t)\right)dt \]
\[ + E}\int_0^T \int_\mathbb{R} \nabla_k H_1(t, \zeta)k_1(t, \zeta)\nu(d\zeta)dt \]
\[ = E[\lambda_1(T)H_1'(X(T))x_1(T)] + E}\int_0^T \lambda_1(t)\left(\frac{\partial}{\partial x_1}(t)x_1(t) + \frac{\partial}{\partial y_1}(t)y_1(t) + \frac{\partial}{\partial \lambda_1}(t)\tilde{\Lambda}_1(t) + \frac{\partial}{\partial \beta_1}(t)\beta_1(t)\right)dt \]
\[ + \frac{\partial}{\partial \lambda_1}(t)\tilde{\Lambda}_1(t) + \frac{\partial}{\partial \beta_1}(t)\beta_1(t)\right)dt - E}\int_0^T \nabla_k H_1(t, \zeta)k_1(t, \zeta)\nu(d\zeta)dt \]
\[ - E}\int_0^T z_1(t)\left(\frac{\partial}{\partial x_1}(t)x_1(t) + \frac{\partial}{\partial y_1}(t)y_1(t) + \frac{\partial}{\partial \lambda_1}(t)\tilde{\Lambda}_1(t) + \frac{\partial}{\partial \beta_1}(t)\beta_1(t)\right)dt \]
\[ + \int_0^T \int_\mathbb{R} \nabla_k H_1(t, \zeta)k_1(t, \zeta)\nu(d\zeta)dt. \]
\[ (25) \]

By the definition of \( D_1 \) as well as equations (24) and (25),
\[ D_1 = A + E[\int_0^T \beta_1(t) \left( \frac{\partial H_1}{\partial u_1}(t) + \frac{\partial H_1}{\partial b_1}(t) \right) p_1(t) + \frac{\partial H_1}{\partial y_1}(t) q_1(t) + \frac{\partial H_1}{\partial \xi_1}(t) r_1(t) \right) dt + E[\int_0^T w_1(t) \left( -\frac{\partial H_1}{\partial w_1}(t) + \frac{\partial H_1}{\partial \lambda_1}(t) \right) \right) dt \\
+ \int_0^T z_1(t) \left( -\frac{\partial H_1}{\partial z_1}(t) + \frac{\partial H_1}{\partial \lambda_1}(t) \right) dt + \int_0^T k_1(t) \left( -\nabla k_1 H_1(t) + \nabla k g_1(t) \lambda_1(t) \right) \right] \] 

(26)

where

\[ A := E[\int_0^T x_1(t) \left( \frac{\partial H_1}{\partial x_1}(t) + \frac{\partial H_1}{\partial y_1}(t) p_1(t) + E[\mu_1(t)] \right) \int_0^T y_1(t) \left( \frac{\partial H_1}{\partial y_1}(t) + \frac{\partial H_1}{\partial \lambda_1}(t) \right) \right) dt + \int_0^T \lambda_1(t) \left( \frac{\partial H_1}{\partial \lambda_1}(t) \right) \right) dt \\
+ \int_0^T \frac{\partial H_1}{\partial \lambda_1}(t) \int_0^T \lambda_1(t) \left( \frac{\partial H_1}{\partial \lambda_1}(t) \right) \right) dt + \int_0^T k_1(t) \left( -\nabla k_1 H_1(t) + \nabla k g_1(t) \lambda_1(t) \right) \right] \] 

(27)

Then, by using the definition of the Hamiltonian \( H_1 \), see equation (5), we see that everything inside the curly brackets in equation (26) is equal to zero. Hence,

\[ D_1 = A + E[\int_0^T \beta_1(t) \frac{\partial H_1}{\partial u_1}(t) dt]. \]

Recall that from the definitions of \( y_1 \) and \( \lambda_1 \),

\[ y_1(t) = x_1(t - \delta_1) \text{ and } \lambda_1(t) = \int_{t-\delta_1}^t x_1(u) dB(u). \]

This implies, by change of variables
FBSDE games with delay and noisy memory

\[
E\left[\int_0^T y_1(t) \frac{\partial H_1}{\partial y_1}(t) dt\right] = E\left[\int_0^T x_1(t - \delta_1) \frac{\partial H_1}{\partial y_1}(t) dt\right] \\
= \int_{\delta_1}^T x_1(u) \frac{\partial H_1}{\partial y_1}(u + \delta_1) du \\
= E\left[\int_0^T x_1(u)[1_{[0,T-\delta_1]}(u)] \frac{\partial H_1}{\partial y_1}(u + \delta_1) du\right].
\]

Also, by the duality formula for Malliavin derivatives (see Di Nunno et al. [5]) and changing the order of integration

\[
E\left[\int_0^T \tilde{\Lambda}_1(t) \frac{\partial H_1}{\partial \Lambda_1}(t) dt\right] = E\left[\int_0^T \int_t^{t+\delta_1} x_1(u) dB(u) \frac{\partial H_1}{\partial \Lambda_1}(t) dt \right] \\
= E\left[\int_0^T \int_t^{t+\delta_1} E[D_u(\frac{\partial H_1}{\partial \Lambda_1}(t))|F_u] x_1(u) du dt\right].
\]

But, from the definition of \(\mu_1\),

\[
E\left[\int_0^T x_1(t) E[\mu_1(t)|F_t] dt\right] = E\left[\int_0^T E[x_1(t)\mu_1(t)|F_t] dt\right] \\
= E\left[\int_0^T E[x_1(t)\{-\frac{\partial H_1}{\partial x_1}(t) - \frac{\partial H_1}{\partial y_1}(t + \delta_1) 1_{[0,T-\delta_1]}\} \\
- \int_t^{t+\delta_1} D_u(\frac{\partial H_1}{\partial \Lambda_1}(s))|1_{[0,T]}(s) ds|F_t]\right] dt.
\]

So, by the rule of double expectation and the calculations above, \(A = 0\).

This implies that \(D_1 = E[\int_0^T \beta_1(t) \frac{\partial H_1}{\partial u_1}(t) dt]\), so

\[
\frac{\partial}{\partial s} J_1(u_1 + s\beta_1, u_2)|_{s=0} = E\left[\int_0^T \beta_1(t) \frac{\partial H_1}{\partial u_1}(t) dt\right]
\]

which was what we wanted to prove.

\(\square\)
5 Solution of the noisy memory FBSDE

In this section, we consider a slightly simplified version of the system of noisy memory FBSDEs in equations (6) and (7). Instead, consider the following noisy memory FBSDE:

**FSDE in $\lambda$,**

\[
d\lambda(t) = \frac{\partial H}{\partial w}(t)dt + \frac{\partial H}{\partial z}(t)dB(t) + \int_{\mathbb{R}} \nabla_k H(t, \zeta) \tilde{N}(dt, d\zeta) \\
\lambda(0) = \phi'(W(0)).
\] (28)

**BSDE in $p, q$ and $r$,**

\[
dp(t) = -E[\mu(t)|\mathcal{F}_t]dt + q(t)dB(t) + \int_{\mathbb{R}} r(t, \zeta) \tilde{N}(dt, d\zeta) \\
p(T) = \varphi'(X(T)) + h'(X(T))\lambda(T)
\] (29)

where

\[
H(t, x, y_1, y_2, \Lambda_1, \Lambda_2, w, z, k, u_1, u_2, \lambda, p, q, r) \\
= f(t, x, y, \Lambda, u_1, u_2) + \lambda g(t, x, y_1, y_2, \Lambda_1, \Lambda_2, w, z, k, u_1, u_2) \\
+ pb(t, x, y_1, y_2, \Lambda_1, \Lambda_2, u_1, u_2) + q\sigma(t, x, y_1, y_2, \Lambda_1, \Lambda_2, u_1, u_2) \\
+ \int_{\mathbb{R}} r(\zeta)\gamma(t, x, y_1, y_2, \Lambda_1, \Lambda_2, u_1, u_2, \zeta)\nu(d\zeta)
\]

and

\[
\mu(t) = \frac{\partial H}{\partial x}(t) + \frac{\partial H}{\partial y}(t + \delta)1_{[0, T-\delta]}(t) + \int_{t}^{t+\delta} E[D_{t}^{\Lambda}[\frac{\partial H}{\partial \Lambda}(s)]|\mathcal{F}_t]1_{[0, T]}(s)ds.
\]

Note that the set of equations (6) and (7) are two such systems such as (28)-(29) involving the same $X$ process as well as the same controls $u_1, u_2$.

Also, consider the following system consisting of an FSDE and two BSDEs:
FSDE in $\lambda$,

$$
\begin{align*}
    d\hat{\lambda}(t) &= \frac{\partial H}{\partial w}(t) dt + \frac{\partial H}{\partial z}(t) dB(t) + \int_{\mathbb{R}} \nabla_k H(t, \zeta) \tilde{N}(dt, d\zeta) \\
    \hat{\lambda}(0) &= \phi'(W(0)).
\end{align*}
$$

(30)

BSDE in $p_1, q_1$ and $r_1$,

$$
\begin{align*}
    dp_1(t) &= -E[\mu_1(t)|\mathcal{F}_t] dt + q_1(t) dB(t) + \int_{\mathbb{R}} r_1(t, \zeta) \tilde{N}(dt, d\zeta) \\
    p_1(T) &= \varphi'(X(T)) + h'(X(T))\hat{\lambda}(T).
\end{align*}
$$

(31)

BSDE in $p_2, q_2$ and $r_2$,

$$
\begin{align*}
    dp_2(t) &= -E[\mu_2(t)|\mathcal{F}_t] dt + q_2(t) dB(t) + \int_{\mathbb{R}} r_2(t, \zeta) \tilde{N}(dt, d\zeta) \\
    p_2(T) &= 0
\end{align*}
$$

(32)

where

$$
\begin{align*}
    H(t, x, y_1, y_2, \Lambda_1, \Lambda_2, w, z, k, u_1, u_2, \hat{\lambda}, p_1, p_2, q_1, q_2, r_1, r_2) &= q_2(t) x + H(t, x, y_1, y_2, \Lambda_1, \Lambda_2, w, z, k, u_1, u_2, \hat{\lambda}, p_1, p_2, q_1, q_2) \\
    \mu_1(t) &= q_2(t) + \frac{\partial H}{\partial x}(t) + \frac{\partial H}{\partial y}(t + \delta) 1_{[0,T-\delta]}(t)
\end{align*}
$$

(33)

and

$$
\begin{align*}
    \mu_2(t) &= \frac{\partial H}{\partial \Lambda}(t) - \frac{\partial H}{\partial \Lambda}(t + \delta) 1_{[0,T-\delta]}(t).
\end{align*}
$$

Note that $\frac{\partial H}{\partial \Lambda}(t) = \frac{\partial H}{\partial \Lambda}(t) = q_2(t) + \frac{\partial H}{\partial y}(t)$ and $\frac{\partial H}{\partial \Lambda}(t) = \frac{\partial H}{\partial y}(t)$. Hence, equations (28) and (30) are structurally equal.

Then, by similar techniques as in Dahl et al. [6], we can show the following theorem:

**Theorem 5.1** Assume that $(p_i, q_i, r_i)$ for $i = 1, 2$ and $\hat{\lambda}$ solve the FBSDE sys-
Define $\lambda = \lambda_1$, $p(t) = p_1(t)$, $q(t) = q_1(t)$ and $r(t, \cdot) = r_1(t, \cdot)$ and assume that $E[\int_0^T (\frac{\partial H(t)}{\partial z})^2]dt < \infty$. Then, $(p, q, r, \lambda)$ solves the noisy memory FBSDE (28)–(29) and

$$q_2(t) = \int_t^{t+\delta} E[D_t[\frac{\partial H(s)}{\partial \Lambda}]_1[0,T](s)] ds.$$  

**Proof.** The jump terms do not make a difference here, so assume for simplicity that $r = r_1 = r_2 = 0$ everywhere.

In general, we know that if $dp_2(t) = -\theta(t, p_2, q_2) dt + q_2(t) dB(t)$, $p_2(T) = F$, then

$$q_2(t) = D_t p_2(t).$$  

(34)

Now, note that the solution $p_2$ of the BSDE (32) can be written

$$p_2(t) = -E[\int_t^T E[\mu_2(s)|\mathcal{F}_s]|\mathcal{F}_t]$$

$$= -\int_t^T E[\mu_2(s)|\mathcal{F}_t] ds$$

$$= -\int_t^T \frac{\partial H(t)}{\partial \Lambda} - \frac{\partial H(t+\delta)}{\partial \Lambda} 1_{[0,T-\delta]}(t)|\mathcal{F}_t] ds$$

$$= -\int_t^{t+\delta} E[\frac{\partial H(s)}{\partial \Lambda}]_1[0,T](s)|\mathcal{F}_t] ds$$

where the equalities follow from Fubini’s theorem, the rule of double expectation, the definition of $\mu_2$ and a change of variables. Hence, by equation (34):

$$q_2(t) = D_t p_2(t)$$

$$= D_t \int_t^{t+\delta} E[\frac{\partial H(s)}{\partial \Lambda}]_1[0,T](s)|\mathcal{F}_t] ds$$

and

$$= \int_t^{t+\delta} E[D_t(\frac{\partial H(s)}{\partial \Lambda})]|\mathcal{F}_t]_1[0,T](s)|\mathcal{F}_t] ds$$

which is part of what we wanted to prove.

By inserting this expression for $q_2$ into the definition of $\mu_1$, we see that
\[
\mu_1(t) = \int_t^{t+\delta} E[D_t\left(\frac{\partial H(s)}{\partial \Lambda}\right)\mathcal{F}_t]1_{[0,T]}(s)ds + \frac{\partial H(t)}{\partial x} + \frac{\partial H(t+\delta)}{\partial y}1_{[0,T]}(t+\delta).
\]

Hence, we see that the BSDE (31) is the same as (29), so they have the same solution. This completes the proof of the theorem.

\[\square\]

We can also prove the following converse result.

**Theorem 5.2** If \(p, q, r, \lambda\) solve the FBSDE (28) – (29) and we define \(\tilde{\lambda} = \lambda, p_1 = p, q_1 = q, r_1 = r\) and

\[
\begin{align*}
p_2(t) &= \int_t^{t+\delta} E[D_t\left(\frac{\partial H(s)}{\partial \Lambda}\right)\mathcal{F}_t]1_{[0,T-\delta]}(s)ds \\
q_2(t) &= \int_t^{t+\delta} E[D_t\left(\frac{\partial H(s)}{\partial \Lambda}\right)\mathcal{F}_t]1_{[0,T-\delta]}(s)ds \\
r_2(t, \cdot) &= 0.
\end{align*}
\]

Then, \((p_i, q_i, r_i)\) for \(i = 1, 2\) and \(\tilde{\lambda}\) solve the system of equations (30) – (32).

**Proof.** Again, the jump parts make no crucial difference, so we consider the no-jump situation for simplicity.

It is clear that equation (30) holds from the assumptions above (from the definition of \(\mathcal{H}\), see (33)). Also, the BSDE (31) holds: Clearly, the terminal condition holds, and by the computations in the proof of Theorem 5.1, the remaining part of equation (31) also holds. Therefore, it only remains to prove that the BSDE (32) holds.

By the Itô isometry and the Clark-Ocone formula,

\[
E[\int_0^T E[D_s\left(\frac{\partial H(r)}{\partial \Lambda}\right)\mathcal{F}_s]d\mathcal{F}_s] = E[\left(\int_0^T E[D_s\left(\frac{\partial H(r)}{\partial \Lambda}\right)\mathcal{F}_s]dB_s\right)^2] = E[(\frac{\partial H(r)}{\partial \Lambda}(r))^2] = E[(\frac{\partial H(r)}{\partial \Lambda}(r))^2].
\]
Hence,

$$\int_0^T E[\int_0^T E[D_s(\frac{\partial H(r)}{\partial \Lambda})|\mathcal{F}_s]^2 ds] \, dr = \int_0^T (E[\frac{\partial H(r)}{\partial \Lambda}]^2 - E[\frac{\partial H(r)}{\partial \Lambda}]^2)^2 \, dt < \infty.$$  

Note that from the Clark-Ocone theorem,

$$\frac{\partial H(r)}{\partial \Lambda} = E[\frac{\partial H(r)}{\partial \Lambda} | \mathcal{F}_t] + \int_t^r E[D_s(\frac{\partial H(r)}{\partial \Lambda})|\mathcal{F}_s] dB(s).$$

Therefore, by the definition of $q_2$ in the theorem and the Fubini theorem

$$\int_t^T q_2(s) dB(s) = \int_t^T \int_t^T E[D_s(\frac{\partial H(r)}{\partial \Lambda})|\mathcal{F}_s] 1_{[s, s+\delta]}(r) dr dB(s) = \int_t^T \int_t^T E[D_s(\frac{\partial H(r)}{\partial \Lambda})|\mathcal{F}_s] 1_{[r-\delta, r]}(s) dB(s) dr.$$  

By some algebra and the Clark-Ocone theorem, we have

$$\int_t^T \int_t^T E[D_s(\frac{\partial H(r)}{\partial \Lambda})|\mathcal{F}_s] 1_{[r-\delta, r]}(s) dB(s) dr = \int_t^T \int_{r-\delta}^T E[D_s(\frac{\partial H(r)}{\partial \Lambda})|\mathcal{F}_s] dB(s) dr = \int_t^T \frac{\partial H(r)}{\partial \Lambda} - E[\frac{\partial H(r)}{\partial \Lambda} | \mathcal{F}_{r-\delta}] \, dr.$$

By splitting the integrals and using change of variables (twice) as well as some algebra,

$$= \int_t^T \frac{\partial H(s)}{\partial \Lambda} \, ds - \int_{T-\delta}^T E[\frac{\partial H(s+\delta)}{\partial \Lambda}|\mathcal{F}_s] \, ds$$

$$= \int_t^T \frac{\partial H(s)}{\partial \Lambda} \, ds - \int_t^T E[\frac{\partial H(s+\delta)}{\partial \Lambda}|\mathcal{F}_s] 1_{[0, T-\delta]}(s) \, ds$$

$$- \int_t^{T-\delta} E[\frac{\partial H(s)}{\partial \Lambda}|\mathcal{F}_s] 1_{[0, T-\delta]}(s) \, ds$$

$$= \int_t^T E[\frac{\partial H(s)}{\partial \Lambda} - \frac{\partial H(s+\delta)}{\partial \Lambda} 1_{[0, T-\delta]}(s)|\mathcal{F}_s] \, ds - p_2(t).$$

This proves that the BSDE holds as well.

□

Now, we have expressed the solution of the Malliavin FBSDE via the solution of the “double” FBSDE system. What kind of system of equations is
FBSDE games with delay and noisy memory

this? The system consists of two connected BSDEs in \((p_1, q_1, r_1)\) and \((p_2, q_2, r_2)\) respectively, and these are again connected to a FBSDE in \(\lambda\). However, from equation (32) and the definition of \(\mu_2\), we see that the right hand side of (32) does not depend on \(p_2\). Hence, the BSDE (32) can be rewritten

\[
dp_2(t) = h(t, \lambda, p_1, q_1, r_1(\cdot))dt + q_2(t)dB(t) + \int_{\mathbb{R}} r_2(t, \zeta) \tilde{N}(dt, d\zeta)
\]

\[
p_2(T) = 0.
\]

This can be solved to express \(p_2\) using \(\lambda, p_1, q_1\) and \(r_1(\cdot)\) by letting \(q_2(t) = r_2(t, \cdot) = 0\) for all \(t\) and

\[
p_2(t) = E[\int_t^T h(t, \lambda, p_1, q_1, r_1(\cdot))dt | \mathcal{F}_t].
\]

Now, we can substitute this solution for \(p_2(t)\) into the FBSDE system (30)-(31). The resulting set of equations is a regular system of time advanced FBSDEs with jumps. There are to the best of our knowledge, no general results on existence and uniqueness of such systems of FBSDEs. However, if we simplify by removing the jumps and there was no time-advanced part (i.e., no delay process \(Y_i\) in the original FSDE (11)), there are some results by Ma et al. [10].

6 Optimal consumption rate with respect to recursive utility

In this section, we apply the previous results to the problem of determining an optimal consumption rate with respect to recursive utility (see also Øksendal and Sulem [16] and Dahl and Øksendal [4]). Let \(X(t) = X^c(t)\), where the consumption rate \(c(t)\) is our control, and assume that
\[dX(t) = X(t)[\mu(t)dt + \sigma(t)dB(t) + \int_{\mathbb{R}} \gamma(t, \zeta) \tilde{N}(dt, d\zeta)] - [c_1(t) + c_2(t)]X(t)dt, \]
\[X(0) = x > 0\] (35)

and \(W_i(t)\) is given by

\[dW_i(t) = -[\alpha_i(t)W_i(t) + \eta_i(t) \ln(Y_i(t)) + \kappa_i(t) \ln(\Lambda_i(t)) + \ln(c_i(t)X(t))] + Z_i(t)dB(t) + \int_{\mathbb{R}} K_i(t, \zeta) \tilde{N}(dt, d\zeta)\]
\[W_i(T) = 0.\]

Let the performance functional be defined by
\[J_i(c_1, c_2) = W_i(0),\] i.e., \(J_i\) is the recursive utility for player \(i\). Also, assume that both players have full information, so \(\mathbb{E}(i)_{t} = (\mathcal{F}_t)_{t}\) for \(i = 1, 2\).

We would like to find a Nash equilibrium for this FBSDE game with delay. To do so we will use the maximum principle Theorem 3.1. Note that \(f_i = \varphi_i = h_i = 0\) and that \(\psi_i(w) = w\) for \(i = 1, 2\). The Hamiltonians are:

\[H_i(t, x, y_1, y_2, A_1, A_2, w_i, z_i, c_1, c_2, \lambda_i, p_i, q_i, r_i(\zeta)) = \lambda_i(\alpha_i(t)w_i + \eta_i(t) \ln(y_i) + \ln(c_iX)) + p_i(x\mu(t) - (c_1 + c_2)x) + q_i\sigma(t)x + \int_{\mathbb{R}} x\gamma_i(t, \zeta)\gamma(t, \zeta)\nu(d\zeta)\] for \(i = 1, 2\).

The adjoint BSDEs are

\[dp_i(t) = E[\mu_i(t)|\mathcal{F}_t]dt + q_i(t)dB(t) + \int_{\mathbb{R}} r_i(t, \zeta)\tilde{N}(dt, d\zeta),\]
\[p_i(T) = 0\]

where

\[\mu_i(t) = -\frac{\lambda_i(t)}{X_i(t)} - \frac{\lambda_i(t+\delta)}{X_i(t+\delta)}\mathbb{1}_{[0,T-\delta]}(t) - p_i(t)(\mu(t) - (c_1(t) + c_2(t))) + q_i(t)\sigma(t) + \int_{\mathbb{R}} r_i(t, \zeta)\gamma(t, \zeta)\nu(d\zeta)\]
for $i = 1, 2$. Note that by the definition of $Y_i$, $Y_i(t + \delta_i) = X(\{t + \delta_i\} - \delta_i) = X(t)$.

The adjoint BSDEs are linear, and the solutions are given by (see Øksendal and Sulem [17])

$$
\Gamma_i(t)p_i(t) = E[\int_t^T \left( \lambda_i(s) + \frac{\lambda_i(s+\delta_i)\eta_i(s+\delta_i)}{X(s+\delta_i)} \right) 1_{[0,T-\delta_i]}(s) \Gamma_i(s) ds | \mathcal{F}_t] \quad (36)
$$

where

$$
\Gamma_i(0) = 1 \quad \text{for} \quad i = 1, 2.
$$

Note that by the SDE (35),

$$
x \Gamma_i(t) = X(t).
$$

Hence, by combining equations (36) and (37), we see that

$$
X(t)p_i(t) = E[\int_t^T (\lambda_i(s) + \lambda_i(s + \delta_i)\eta_i(s + \delta_i)1_{[0,T-\delta_i]}(s)) ds | \mathcal{F}_t]. \quad (38)
$$

The adjoint FSDEs are

$$
d\lambda_i(t) = \lambda_i(t)c_i(t) dt \\
\lambda_i(0) = 1, \text{ for } i = 1, 2.
$$

These are (non-stochastic) differential equation with solution $\lambda_i(t) = \exp(\int_0^t c_i(s) ds)$ for $i = 1, 2$.

We maximize $H_i$ with respect to $c_i$. For $i = 1, 2$, the first order condition is:

$$
\hat{c}_i(t) = \frac{\lambda_i(t)}{p_i(t)X(t)}.
$$
By substituting equation (38) into this, we find (by the sufficient maximum principle, Theorem 3.1) that the consumption rates leading to a Nash equilibrium for the recursive utility problem are given by:

\[ c^*_i(t) = \lambda_i(t) E\left[ \int_t^T \lambda_i(s) + \lambda_i(s + \delta_i)\eta_i(s + \delta_i)\mathbf{1}_{[0,T-\delta_i]}(t)ds|\mathcal{F}_t \right]. \]

where \( \lambda_i(t) = \exp\left(\int_0^t \alpha_i(s)ds\right) \) for \( i = 1, 2 \).

7 Conclusion

In this paper, we have analyzed a two-player stochastic game connected to a set of FBSDEs involving delay and noisy memory of the market process. We have derived sufficient and necessary maximum principles for a set of controls for the two players to be a Nash equilibrium in this game. We have also studied the associated FBSDE involving Malliavin derivatives, and connected this to a system of FBSDEs not involving Malliavin derivatives. Finally, we were able to derive a closed form Nash equilibrium solution to a game where the aim is to find the optimal consumption with respect to recursive utility.

References

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

[2] Agram, N. and Øksendal, B. 2015. Malliavin calculus and optimal control of stochastic Volterra equations, *Journal of Optimization Theory and Applications*, 3 (167): 1070-1094.

[3] Agram, N. and Øksendal, B. 2014. Infinite horizon optimal control of forward-backward stochastic differential equations with delay, *Journal of Computational and Applied Mathematics*, 259 (B): 336-349.
[4] Dahl, K. and Øksendal, B. 2017. Singular recursive utility, *Stochastics*, so far only published online.

[5] Di Nunno, G., Øksendal, B. and Proske, F. 2009. *Malliavin Calculus for Lévy Processes with Applications to Finance*, Springer, Berlin Heidelberg.

[6] Dahl, K., Mohammed, S., Øksendal, B. and Røse, E. 2016. Optimal control of systems with noisy memory and BSDEs with Malliavin derivatives, *Journal of Functional Analysis*, 2 (271): 289-329.

[7] Ivanov, A. F., Kazmerchuk Y. I. and Swishchuk A. V., Theory, Stochastic Stability and Applications of Stochastic Delay Differential Equations: a Survey of Recent Results, Research report, http://www.math.yorku.ca/~aswishch/sddesurvey.pdf

[8] Krylov, N. V. 2009. *Controlled Diffusion Processes*, Springer, Berlin Heidelberg.

[9] Krylov, N. V. 1972. Control of a Solution of a Stochastic Integral Equation *Theory of Probability and its Applications*, 1 (17): 114-130.

[10] Ma, J., Yin, H. and Zhang, J. 2012. On non-Markovian forward-backward SDEs and backward stochastic PDEs. *Stochastic Processes and their Applications*, 12 (122): 3980-4004.

[11] Nualart, D. 2006. *The Malliavin Calculus and Related Topics*, Springer, Berlin Heidelberg.

[12] Øksendal, B. 2013. *Stochastic differential equations: an introduction with applications*, Springer, Berlin Heidelberg.

[13] Øksendal, B. and Sulem, A. 2007. *Applied Stochastic Control of Jump Diffusions*, Springer, Berlin Heidelberg.

[14] Øksendal, B. and Sulem, A. 2014. Forward-backward stochastic differential games and stochastic control under model uncertainty, *Journal of Optimization Theory and Applications*, 22 (161): 22-55.
[15] Øksendal, B. and Sulem, A. 2000. A maximum principle for optimal control of stochastic systems with delay, with applications to finance, Eds. J.M. Menaldi, E. Rofman and A. Sulem: *Optimal Control and Partial Differential Equations - Innovations and Applications*, IOS Press, Amsterdam.

[16] Øksendal, B. and Sulem, A. 2016. Optimal control of predictive mean-field equations and applications to finance, Eds. Benth F., Di Nunno G.: *Stochastics of Environmental and Financial Economics*, Springer Proceedings in Mathematics & Statistics, 138, Springer, Cham.

[17] Øksendal, B. and Sulem, A. 2014. Risk minimization in financial markets modeled by Itô Lévy processes. Research report, Department of Mathematics, University of Oslo.

[18] Øksendal, B., Sulem, A. and Zhang, T. 2011. Optimal control of stochastic delay equations and time-advanced backward stochastic differential equations, *Advances in Applied Probability*, 2 (43): 572-596.

[19] Pham, H. 2009. Continuous-time Stochastic Control and Optimization with Financial Applications, *Stochastic Modelling and Applied Probability*, 61: 139-169.

[20] Pardoux, E. and Peng, S. 1990. Adapted Solutions of Backward stochastic differential equation *Systems and Control Letters*, 1(14): 55-61.

[21] Rockafellar, R. T. 1970. *Convex Analysis*, Princeton University Press, Princeton.

[22] Sanz-Solè, M. 2005. *Malliavin Calculus*. EPFL Press, Lausanne.

[23] Wang, S. and Wu, Z. 2016. Stochastic maximum principle for optimal control problems of forward-backward delay systems involving impulse controls, *Journal of Systems Science and Complexity*, 2 (30): 280-306.