Optimal control of stochastic functional differential equations with a bounded memory*

Mou-Hsiung (Harry) Chang
Mathematics Division
U. S. Army Research Office
P. O. Box 12211
RTP, NC 27709, USA
Email: mouhsiung.chang@us.army.mil

Tao Pang†
Department of Mathematics
Center for Research in Scientific Computation
North Carolina State University
Raleigh, NC 27695-8205 USA
Email: tpang@unity.ncsu.edu

Moustapha Pemy
Department of Mathematics
Towson University
Towson, MD 21252-0001 USA
Email: mpemy@towson.edu

January 2, 2006 (Revised July 26, 2007)

*This research is partially supported by a grant W911NF-04-D-0003 from the U. S. Army Research Office

†Corresponding author.

Abstract

This paper treats a finite time horizon optimal control problem in which the controlled state dynamics is governed by a general system of stochastic functional differential equations with a bounded memory. An infinite-dimensional HJB equation is derived using a Bellman-type dynamic programming principle. It is shown that the value function is the unique viscosity solution of the HJB equation.

Keywords: Stochastic control, stochastic functional differential equations, viscosity solutions, Hamilton-Jacobi-Bellman equation.

AMS 2000 subject classifications: primary 93E20, 60G20; secondary 49L20, 49L25

1 Introduction

Optimal control of Itô's diffusion processes have been extensively studied in the literature, see Fleming and Rishel [8] for the classical theory, Fleming and Soner [9] and the references contained herein for the viscosity solution approach.

However, in many real world applications (see Kolmanovskii and Shaikhet [12]), these physical systems can only be modeled by stochastic dynamical systems whose evolutions depend on the past history of the states. Such models are referred to as stochastic (retarded) functional differential equations (see Mohammed [17], [18] for an introduction of these models). The linear-quadratic regulatory problem involving stochastic delay equations was first studied in Kolmanovskii and Maizenberg [11], and optimal control problems for a class of nonlinear stochastic equations that involve a continuous delay of the following type

$$\begin{aligned} dX(s) &= \alpha(s, X(s), Y(s), u(s))ds \\ &\quad + \beta(s, X(s), Y(s), u(s))dW(s), \quad s \in [t, T], \end{aligned} \tag{1}$$

have been studied in recent literature (see e.g. Elsanousi [6], Elsanousi et al [7], and Larssen [13], Oksendal and Sulem [19]), in which $Y(s) = \int_{-r}^{0} e^{-\delta\theta} X(s+\theta)d\theta$.

The purpose of this paper is to investigate a finite time horizon optimal control problem for a general system of stochastic functional differential equations that include (1) as a special case. We use the viscosity solution concept introduced by Crandall and Lions [4], [14], [16] in order to characterize the value function as the unique viscosity solution of the associated HJB equation.

This paper is organized as follows. Notations and the formulation of the problem are contained in Section 2. In Section 3, the infinite dimensional Hamilton-Jacobi-Bellman (HJB) equation for the value function is heuristically derived using Bellman's type dynamic programming principle, which was first obtained in Larssen [13]. In Section 4, the continuity of the value function is proved. It is shown in Section 4 that the value function is actually a viscosity solution of the HJB equation. The uniqueness result for viscosity solution of the HJB equation is given in Section 5.

2 Problem Formulation

Let $T > 0$ denote a fixed terminal time, and let $t \in [0, T]$ denote an initial time. We study the finite time horizon optimal control problem for a general system of stochastic functional differential equations on the interval $[t, T]$.

Let $r > 0$ be a fixed constant, and let $\mathbb{J} = [-r, 0]$ denote the duration of the bounded memory of the equations considered in this paper. For the sake of simplicity, we denote $C(\mathbb{J}; \Re^n)$, the space of continuous functions $\phi : \mathbb{J} \to \Re^n$, by $\mathbf{C}$. Note that $\mathbf{C}$ is a real separable Banach space under the supremum norm defined by

$$\|\phi\| = \sup_{t \in \mathbb{J}} |\phi(t)|, \quad \phi \in \mathbf{C},$$

where $|\cdot|$ is the Euclidean norm in $\Re^n$.

We denote by $(\,\cdot\,|\,\cdot\,)$ the inner product in $L^2(\mathbb{J}, \Re^n)$ as the following

$$(\phi|\psi) = \int_{-r}^{0} \langle \phi(s), \psi(s) \rangle ds, \qquad \text{and } \|\phi\|_2 = (\phi|\phi)^{\frac{1}{2}}, \quad \forall \phi, \psi \in \mathbf{C},$$

where $\langle \cdot\,, \cdot \rangle$ is the inner product in $\Re^n$.

Note that the space $\mathbf{C}$ can be continuously embedded into $L^2(\mathbb{J}; \Re^n)$.

Convention 2.1 *In this paper, we use the following conventional notation for functional differential equations (see Hale [10]):*

If $\psi \in C([-r, \infty); \Re^n)$ and $t \in \Re_+$, let $\psi_t \in \mathbf{C}$ be defined by $\psi_t(\theta) = \psi(t + \theta)$, $\theta \in \mathbb{J}$.

Let $\{W(t), t \geq 0\}$ be a certain m-dimensional standard Brownian motion defined on a complete filtered probability space $(\Omega, \mathcal{F}, P; \mathbf{F})$, where $\mathbf{F} = \{\mathcal{F}(t), t \geq 0\}$ is the P-augmentation of the natural filtration $\{\mathcal{F}^W(t), t \geq 0\}$ generated by the Brownian motion $\{W(t), t \geq 0\}$, *i.e.*, if $t \geq 0$,

$$\mathcal{F}^W(t) = \sigma\{W(s), 0 \leq s \leq t\}$$

and

$$\mathcal{F}(t) = \mathcal{F}^W(t) \vee \{A \subset \Omega | \exists B \in \mathcal{F} \text{ such that } A \subset B \text{ and } P(B) = 0\}$$

where the operator $\vee$ denotes that $\mathcal{F}(t)$ is the smallest σ-algebra such that $\mathcal{F}^W(t) \subset \mathcal{F}(t)$ and

$$\{A \subset \Omega | \exists B \in \mathcal{F} \text{ such that } A \subset B \text{ and } P(B) = 0\} \subset \mathcal{F}(t).$$

Let $L^2(\Omega, \mathbf{C})$ be the space of $\mathbf{C}$-valued random variables $\Xi : \Omega \to \mathbf{C}$ such that

$$\|\Xi\|_{L^2} = \left\{ \int_\Omega \|\Xi(\omega)\|^2 dP(\omega) \right\}^{\frac{1}{2}} < \infty.$$

In addition, let $L^2(\Omega, \mathbf{C}; \mathcal{F}(t))$ be those $\Xi \in L^2(\Omega, \mathbf{C})$ which are $\mathcal{F}(t)$-measurable.

We consider the following system of controlled stochastic functional differential equations with a bounded memory:

$$dX(s) = f(s, X_s, u(s))ds + g(s, X_s, u(s))dW(s), \quad s \in [t, T], \tag{2}$$

with the initial condition $X_t = \psi_t$, where $\psi_t \in L^2(\Omega, \mathbf{C}; \mathcal{F}(t))$ and $u(\cdot) = \{u(s), s \in [t, T]\}$ is a control process taking values in a compact set U (of an Euclidean space). The functions, $f : [0, T] \times \mathbf{C} \times U \to \Re^n$ and $g : [0, T] \times \mathbf{C} \times U \to \Re^{n \times m}$ are given deterministic functions.

The definitions of solutions of (2) and the admissible controls are given as follows.

Definition 2.2 *Given the m-dimensional standard Brownian motion $\{W(s), s \in [0, T]\}$ and the control process $\{u(s), s \in [t, T]\}$, a process $\{X(s; t, \psi_t, u(\cdot)), s \in [t - r, T]\}$ is said to be a (strong) solution of the controlled equation (2) on the interval $[t - r, T]$ and through the initial datum $(t, \psi_t) \in [0, T] \times L^2(\Omega, \mathbf{C}; \mathcal{F}(t))$ if it satisfies the following conditions:*

1. $X_t(\cdot; t, \psi_t, u(\cdot)) = \psi_t$;
2. $X(s; t, \psi_t, u(\cdot))$ *is $\mathcal{F}(s)$-measurable for each $s \in [t, T]$;*
3. *The process $\{X(s; t, \psi_t, u(\cdot)), s \in [t, T]\}$ is continuous and it satisfies the following stochastic integral equation P-a.s.*

$$X(s) = \psi_t(0) + \int_t^s f(\lambda, X_\lambda, u(\lambda))d\lambda + \int_t^s g(\lambda, X_\lambda, u(\lambda))dW(\lambda). \tag{3}$$

In addition, the solution process $\{X(s; t, \psi_t, u(\cdot)), s \in [t - r, T]\}$ is said to be (strongly) unique if $\{\tilde{X}(s; t, \psi_t, u(\cdot)), s \in [t - r, T]\}$ is also a solution of (2) on $[t - r, T]$ with the control process $u(\cdot)$ and through the same initial datum (t, ψ_t), then

$$P\{X(s; t, \psi_t, u(\cdot)) = \tilde{X}(s; t, \psi_t, u(\cdot)), \forall s \in [t, T]\} = 1.$$

Definition 2.3 *For each $t \in [0, T]$, a 5-tuples $\alpha = (\Omega, \mathcal{F}, P, W(\cdot), u(\cdot))$ is said to be an admissible control if it satisfies the following conditions:*

1. *$(\Omega, \mathcal{F}, P)$ is a complete probability space;*
2. *$W(\cdot) = \{W(s), s \in [t, T]\}$ is an m-dimensional standard Brownian motion on $(\Omega, \mathcal{F}, P)$ over $[t, T]$ with $W(t) = 0$ a.s., and $\mathcal{F}(t, s) = \sigma\{W(\tau), t \leq \tau \leq s\}$ augmented by the P-null sets in $\mathcal{F}$;*
3. *$u : [t, T] \times \Omega \to U$ is an $\{\mathcal{F}(t, s), s \in [t, T]\}$-adapted process on $(\Omega, \mathcal{F}, P)$ that is right-continuous at the initial time t;*
4. *Under the control process $u(\cdot) = \{u(s), s \in [t, T]\}$, equation (2) admits a unique strong solution $X^{t,\psi,u(\cdot)}(\cdot) = \{X(s; t, \psi, u(\cdot)), s \in [t, T]\}$ defined on $(\Omega, \mathcal{F}, P; \{\mathcal{F}(t, s), s \in [t, T]\})$ and through each initial datum $(t, \psi) \in [0, T] \times \mathbf{C}$.*

5. *The* $\mathbf{C}$*-valued process* $\{X_s(\cdot;t,\psi,u(\cdot)), s\in[t,T]\}$ *defined by*

$$X_s(\theta;t,\psi,u(\cdot)) = X(s+\theta;t,\psi,u(\cdot)), \quad \theta\in[-r,0],$$

is a strong Markov process with respect to the Brownian stochastic basis $(\Omega,\mathcal{F},P,\mathbf{F},W(\cdot))$*;*

6. *The control process* $u(\cdot)$ *is such that*

$$\mathbb{E}\Big[\int_t^T |L(s,X_s(t,\psi,u(\cdot)),u(s))|ds + |\Psi(X_T(t,\psi,u(\cdot)))|\Big] < \infty,$$

where $L:[0,T]\times\mathbf{C}\times U\to\Re$ *and* $\Psi:\mathbf{C}\to\Re$ *represent the running and terminal cost functions, respectively.*

The collection of admissible controls $\alpha = (\Omega,\mathcal{F},P,W(\cdot),u(\cdot))$ *over the interval* $[t,T]$ *is denoted by* $\mathcal{U}[t,T]$*.*

Remark 2.4 *To use Dynkin's formula for the the* $\mathbf{C}$*-valued process* $\{X_s(\cdot;t,\psi,u(\cdot)), s\in[t,T]\}$ *in the derivation of the infinite dimensional HJB equation, we make additional requirement that it is a* $\mathbf{C}$*-valued Markov process. This requirement is not a stringent one. In fact, the class of admissible controls* $\mathcal{U}[t,T]$ *includes all feedback controls* $u(s)=\eta(s,X_s)$*, where* $\eta:[0,T]\times\mathbf{C}\to U$ *is Lipschitz with respect to the segment variable, i.e., there exists a constant* $K>0$ *such that*

$$|\eta(t,\phi)-\eta(t,\varphi)| \le K\|\phi-\varphi\|, \quad \forall (t,\phi),(t,\varphi)\in[0,T]\times\mathbf{C}.$$

We will write $u(\cdot)\in\mathcal{U}[t,T]$ or $\alpha=(\Omega,\mathcal{F},P,W(\cdot),u(\cdot))\in\mathcal{U}[t,T]$ interchangeably, whenever there is no danger of ambiguity.

We assume that the functions $f:[0,T]\times\mathbf{C}\times U\to\Re^n$, and $g:[0,T]\times\mathbf{C}\times U\to\Re^{n\times m}$ are continuous and they satisfy the following linear growth and Lipschitz conditions (See also Mohammed [17, 18]).

Assumption 2.5 *There exists a constant* $\Lambda>0$ *such that*

$$|f(t,\varphi,u)-f(t,\phi,u)| + |g(t,\varphi,u)-g(t,\phi,u)| \le \Lambda\|\varphi-\phi\|,$$
$$\forall (t,\varphi,u),(t,\phi,u)\in[0,T]\times\mathbf{C}\times U.$$

Assumption 2.6 *There exists a constant* $K>0$ *such that*

$$|f(t,\phi,u)|+|g(t,\phi,u)| \le K(1+\|\phi\|),\ \forall (t,\phi,u)\in[0,T]\times\mathbf{C}\times U.$$

Given an admissible control $u(\cdot)\in\mathcal{U}[t,T]$, let $X^{t,\psi,u(\cdot)}(\cdot)=\{X(s;t,\psi,u(\cdot)), s\in[t,T]\}$ be the solution of (2) through the initial datum $(t,\psi)\in[0,T]\times\mathbf{C}$. We again consider the corresponding $\mathbf{C}$-valued process $\{X_s(t,\psi,u(\cdot)), s\in[t,T]\}$ defined by

$$X_s(\theta;t,\psi,u(\cdot)) = X(s+\theta;t,\psi,u(\cdot)), \quad \theta\in\mathbb{J}. \tag{4}$$

For notational simplicity, we often write $X(s) = X(s; t, \psi, u(\cdot))$ and $X_s = X_s(t, \psi, u(\cdot))$ for $s \in [t, T]$ whenever there is no danger of ambiguity.

It can be shown that, under Assumptions 2.5-2.6, the $\mathbf{C}$-valued process $\{X_s(t, \psi, u(\cdot)), s \in [t, T]\}$ is a Markov process (see Mohammed [17], [18]).

Let L and Ψ be two continuous real-valued functions on $[0, T] \times \mathbf{C} \times U$ and $[0, T] \times \mathbf{C}$, respectively. Moreover, we assume that they both have at most polynomial growth in $L^2(\mathbb{J}; \Re)$. In other words, there exist a constant $\Lambda > 0$ and an integer $k > 0$ such that

$$|L(t, \phi, u)| \leq \Lambda(1 + \|\phi\|_2)^k, \qquad \text{and} \;\; |\Psi(t, \phi)| \leq \Lambda(1 + \|\phi\|_2)^k,$$

for all $(t, \phi, u) \in [0, T] \times \mathbf{C} \times U$.

Given any initial data $(t, \psi) \in [0, T] \times \mathbf{C}$ and any admissible control $u(\cdot) \in \mathcal{U}[t, T]$, we define the objective function

$$\begin{aligned} J(t, \psi; u(\cdot)) \;\; &\equiv \;\; \mathbb{E}\bigg[\int_t^T e^{-\rho(s-t)} L(s, X_s(t, \psi, u(\cdot)), u(s)) ds \\ &\qquad\qquad + e^{-\rho(T-t)} \Psi(X_T(t, \psi, u(\cdot))) \bigg], \end{aligned} \tag{5}$$

where $\rho > 0$ denotes a discount factor. For each initial datum $(t, \psi) \in [0, T] \times \mathbf{C}$, the optimal control problem is to find $u(\cdot) \in \mathcal{U}[t, T]$ so as to maximize the objective function J. In this case, the value function $V : [0, T] \times \mathbf{C} \to \Re$ is defined to be

$$V(t, \psi) = \sup_{u(\cdot) \in \mathcal{U}[t, T]} J(t, \psi; u(\cdot)). \tag{6}$$

3 The Hamilton-Jacobi-Bellman Equation

In this section, we will heuristically derive the infinite dimensional Hamilton-Jacobi-Bellman (HJB) equation for the value function defined by (6) using Bellman's type dynamic programming principle, which was first obtained in Larssen [13].

3.1 The Infinitesimal Generator

Let $\mathbf{C}^*$ and $\mathbf{C}^\dagger$ be the space of bounded linear functionals $\Phi : \mathbf{C} \to \Re$ and bounded bilinear functionals $\tilde{\Phi} : \mathbf{C} \times \mathbf{C} \to \Re$, of the space $\mathbf{C}$, respectively. They are equipped with the operator norms which will be, respectively, denoted by $\|\cdot\|^*$ and $\|\cdot\|^\dagger$.

Let $\mathbf{B} = \{v\mathbf{1}_{\{0\}}, v \in \Re^n\}$, where $\mathbf{1}_{\{0\}} : [-r, 0] \to \Re$ is defined by

$$\mathbf{1}_{\{0\}}(\theta) = \begin{cases} 0 & \text{for } \theta \in [-r, 0), \\ 1 & \text{for } \theta = 0. \end{cases}$$

We form the direct sum

$$\mathbf{C} \oplus \mathbf{B} = \{\phi + v\mathbf{1}_{\{0\}} \mid \phi \in \mathbf{C}, \;\, v \in \Re^n\}$$

and equip it with the norm $\|\cdot\|$ defined by

$$\|\phi + v\mathbf{1}_{\{0\}}\| = \sup_{\theta\in[-r,0]} |\phi(\theta)| + |v|, \quad \phi \in \mathbf{C},\ v \in \Re^n.$$

Note that for each sufficiently smooth function $\Phi : \mathbf{C} \to \Re$, its first order Fréchet derivative (with respect to $\phi \in \mathbf{C}$), $D\Phi(\varphi) \in \mathbf{C}^*$, has a unique and continuous linear extension $\overline{D\Phi(\varphi)} \in (\mathbf{C} \oplus \mathbf{B})^*$. Similarly, its second order Fréchet derivative, $D^2\Phi(\varphi) \in \mathbf{C}^\dagger$, has a unique and continuous linear extension $\overline{D^2\Phi(\varphi)} \in (\mathbf{C} \oplus \mathbf{B})^\dagger$. In above, $(\mathbf{C} \oplus \mathbf{B})^*$ and $(\mathbf{C} \oplus \mathbf{B})^\dagger$ are spaces of bounded linear and bilinear functionals of $\mathbf{C} \oplus \mathbf{B}$, respectively. (See Lemma (3.1) and Lemma (3.2) on pp 79-83 of Mohammed [17] for details).

For a Borel measurable function $\Phi : \mathbf{C} \to \Re$, we also define

$$\mathcal{S}(\Phi)(\phi) = \lim_{h\to 0+} \frac{1}{h}\left[\Phi(\tilde{\phi}_h) - \Phi(\phi)\right] \tag{7}$$

for all $\phi \in \mathbf{C}$, where $\tilde{\phi} : [-r, T] \to \Re^n$ is an extension of ϕ defined by

$$\tilde{\phi}(t) = \begin{cases} \phi(t) & \text{if } t \in [-r, 0) \\ \phi(0) & \text{if } t \geq 0, \end{cases}$$

and $\tilde{\phi}_t \in \mathbf{C}$ is defined by

$$\tilde{\phi}_t(\theta) = \tilde{\phi}(t + \theta), \quad \theta \in [-r, 0].$$

Let $\hat{\mathcal{D}}(\mathcal{S})$, the domain of the operator $\mathcal{S}$, be the set of $\Phi : \mathbf{C} \to \Re$ such that the above limit exists for each $\phi \in \mathbf{C}$. Define $\mathcal{D}(\mathcal{S})$ as the set of all functions $\Phi : [0, T] \times \mathbf{C} \to \Re$ such that $\Phi(t, \cdot) \in \hat{\mathcal{D}}(\mathcal{S}), \forall t \in [0, T]$.

In addition, let $C^{1,2}_{lip}([0, T] \times \mathbf{C})$ be the space of functions $\Phi : [0, T] \times \mathbf{C} \to \Re$ such that $\frac{\partial\Phi}{\partial t} : [0, T] \times \mathbf{C} \to \Re$ and $D^2\Phi : [0, T] \times \mathbf{C} \to \mathbf{C}^\dagger$ exist and are continuous and satisfy the following Lipschitz condition:

$$\|D^2\Phi(t, \phi) - D^2\Phi(t, \varphi)\|^\dagger \leq K\|\phi - \varphi\| \quad \forall t \in [0, T],\ \phi, \varphi \in \mathbf{C}.$$

We have the following results.

Theorem 3.1 *Suppose that $\Phi \in C^{1,2}_{lip}([0, T] \times \mathbf{C}) \cap \mathcal{D}(\mathcal{S})$. Let $u(\cdot) \in \mathcal{U}[t, T]$ and $\{X_s, s \in [t, T]\}$ be the $\mathbf{C}$-valued Markov solution process of equation (2) with the initial data $(t, \varphi_t) \in [0, T] \times \mathbf{C}$. Then*

$$\begin{aligned} & \lim_{\epsilon\downarrow 0} \frac{E[\Phi(t + \epsilon, X_{t+\epsilon})] - \Phi(t, \varphi_t)}{\epsilon} \\ = \quad & \frac{\partial}{\partial t}\Phi(t, \varphi_t) + \mathcal{S}(\Phi)(t, \varphi_t) + \overline{D\Phi(t, \varphi_t)}(f(t, \varphi_t, u(t))\mathbf{1}_{\{0\}}) \\ & + \frac{1}{2}\sum_{j=1}^{m} \overline{D^2\Phi(t, \varphi_t)}(g(t, \varphi_t, u(t))\mathbf{e}_j\mathbf{1}_{\{0\}}, g(t, \varphi_t, u(t))\mathbf{e}_j\mathbf{1}_{\{0\}}), \end{aligned} \tag{8}$$

where $\mathbf{e}_j, j = 1, 2, \cdots, m$, is the jth unit vector of the standard basis in $\Re^m$.

Proof. One can refer to Mohammed [17], [18]. $\square$

3.2 Heuristic Derivation of the HJB Equation

Before we heuristically derive the HJB equation for V, we re-state the dynamic programming principle described in Larssen [13] which will be used later.

Theorem (Larssen) *Let Assumptions 2.5-2.6 hold. Then for any $(t, \psi) \in [0, T] \times \mathbf{C}$ and $\mathbf{F}(t)$-stopping time $\tau \in [t, T]$,*

$$\begin{aligned} V(t,\psi) &= \sup_{u(\cdot)\in\mathcal{U}[t,T]} \mathbf{E}\Big[\int_t^\tau e^{-\rho(s-t)} L(s, X_s(t,\psi,u(\cdot)), u(s))ds \\ &\qquad + e^{-\rho(\tau-t)} V(\tau, X_\tau(t,\psi,u(\cdot)))\Big]. \end{aligned} \tag{9}$$

Let $v \in U$. We define:

$$\begin{aligned} \mathcal{A}^v V(t,\psi) &\equiv \mathcal{S}(V)(t,\psi) + \overline{DV(t,\psi)}(f(t,\psi,v)\mathbf{1}_{\{0\}}) \\ &+ \frac{1}{2}\sum_{i=1}^m \overline{D^2V(t,\psi)}(g(t,\psi,v)\mathbf{e}_i\mathbf{1}_{\{0\}}, g(t,\psi,v)\mathbf{e}_i\mathbf{1}_{\{0\}}). \end{aligned}$$

$\square$

We assume that for every $v \in U$, the domain of the generator $\mathcal{A}^v$ is large enough to contain $C^{1,2}_{lip}([0,T] \times \mathbf{C}) \cap \mathcal{D}(\mathcal{S})$.

From the dynamic programming principle (Theorem (Larssen)), if we take a constant control $u(\cdot) \equiv v$, then $\forall \delta \geq 0$

$$V(t,\psi) \geq \mathbb{E}\left[\int_t^{t+\delta} e^{-\rho(s-t)} L(s, X_s(t,\psi,v), v)ds + e^{-\rho\delta} V(t+\delta, X_{t+\delta}(t,\psi,v))\right].$$

Combined with Theorem 3.1, the above inequality implies

$$\begin{aligned} 0 &\geq \lim_{\delta\downarrow 0} \frac{1}{\delta}\mathbb{E}\Big[\int_t^{t+\delta} e^{-\rho(s-t)} L(s, X_s(t,\psi,v), v)ds \\ &\qquad + e^{-\rho\delta} V(t+\delta, X_{t+\delta}(t,\psi,v)) - V(t,\psi)\Big] \\ &= -\rho V(t,\psi) + \frac{\partial V}{\partial t}(t,\psi) + \mathcal{A}^v V(t,\psi) + L(t,\psi,v) \end{aligned} \tag{10}$$

for all $(t,\psi) \in [0,T] \times \mathbf{C}$, provided that $V \in C^{1,2}_{lip}([0,T] \times \mathbf{C}) \cap \mathcal{D}(\mathcal{S})$.

Moreover, if $u^*(\cdot) \in \mathcal{U}[t,T]$ is the optimal control policy which satisfies $\lim_{s\downarrow t} u^*(s) = v^*$, we should have that $\forall \delta \geq 0$

$$V(t,\psi) = \mathbb{E}\left[\int_t^{t+\delta} e^{-\rho(s-t)} L(s, X^*_s, u^*(s))ds + e^{-\rho\delta} V(t+\delta, X^*_{t+\delta})\right], \tag{11}$$

where $X_s^* = X_s^*(t, \psi, u^*(\cdot))$ is the $\mathbf{C}$-valued solution process corresponding to the initial datum (t, ψ) and the optimal control $u^*(\cdot) \in \mathcal{U}[t, T]$. Similarly, under strong assumption on $u^*(\cdot)$ (including the right-continuity at the initial time t), we can get

$$0 = -\rho V(t, \psi) + \frac{\partial V}{\partial t}(t, \psi) + \mathcal{A}^{v^*} V(t, \psi) + L(t, \psi, v^*). \tag{12}$$

The inequality (10) and (12) are equivalent to the HJB equation

$$0 = -\rho V(t, \psi) + \frac{\partial V}{\partial t}(t, \psi) + \max_{v \in U} \; [\mathcal{A}^v V(t, \psi) + L(t, \psi, v)]. \tag{13}$$

We therefore have the following result.

Theorem 3.2 *Suppose V is the value function defined by (6) and atisfies $V \in C_{lip}^{1,2}([0, T] \times \mathbf{C}) \cap \mathcal{D}(\mathcal{S})$. Then the value function V satisfies the following HJB equation:*

$$\rho V(t, \psi) - \frac{\partial V}{\partial t}(t, \psi) - \max_{v \in U} \; [\mathcal{A}^v V(t, \psi) + L(t, \psi, v)] = 0 \tag{14}$$

on $[0, T] \times \mathbf{C}$, and $V(T, \psi) = \Psi(\psi), \ \forall \psi \in \mathbf{C}$.

Note that it is not known that the value function V satisfies the necessary smoothness condition $V \in C_{lip}^{1,2}([0, T] \times \mathbf{C}) \cap \mathcal{D}(\mathcal{S})$. Therefore, in general we need to consider viscosity solution instead of a classical solution for HJB equation (14). In fact, it will be shown that the value function is a unique viscosity solution of the HJB equation (14). These results shall be given in the next two sections.

4 Viscosity Solution of the HJB Equation

In this section, we shall show that the value function V defined by (6) is actually a viscosity solution of the HJB equation (14). First, let us define the viscosity solution of (14) as follows.

Definition 4.1 *Let $w \in C([0, T] \times \mathbf{C})$. We say that w is a viscosity subsolution of (14) if, for every $\Gamma \in C_{lip}^{1,2}([0, T] \times \mathbf{C}) \cap \mathcal{D}(\mathcal{S})$, for $(t, \psi) \in [0, T] \times \mathbf{C}$ satisfying $\Gamma \geq w$ on $[0, T] \times \mathbf{C}$ and $\Gamma(t, \psi) = w(t, \psi)$, we have*

$$\rho \Gamma(t, \psi) - \frac{\partial \Gamma}{\partial t}(t, \psi) - \max_{v \in U} \; [\mathcal{A}^v \Gamma(t, \psi) + L(t, \psi, v)] \leq 0.$$

We say that w is a viscosity super solution of (14) if, for every $\Gamma \in C_{lip}^{1,2}([0, T] \times \mathbf{C}) \cap \mathcal{D}(\mathcal{S})$, and for $(t, \psi) \in [0, T] \times \mathbf{C}$ satisfying $\Gamma \leq w$ on $[0, T] \times \mathbf{C}$ and $\Gamma(t, \psi) = w(t, \psi)$, we have

$$\rho \Gamma(t, \psi) - \frac{\partial \Gamma}{\partial t}(t, \psi) - \max_{v \in U} \; [\mathcal{A}^v \Gamma(t, \psi) + L(t, \psi, v)] \geq 0.$$

We say that w is a viscosity solution of (14) if it is both a viscosity supersolution and a viscosity subsolution of (14).

For our value function V defined by (6), we now show that it has the following property.

Lemma 4.2 *The value function V defined in (6) is continuous and there exists a constant $\Lambda > 0$ and a positive integer k such that, for every $(t, \phi) \in [0, T] \times \mathbf{C}$,*

$$|V(t, \phi)| \leq \Lambda(1 + \|\phi\|_2)^k. \tag{15}$$

and there exists a constant $K > 0$ such that

$$|V(s, \phi) - V(s, \varphi)| \leq K\|\phi - \varphi\|, \ \forall (s, \phi), (s, \varphi) \in [0, T] \times \mathbf{C}. \tag{16}$$

Proof. It is clear that V has at most polynomial growth, because L and Φ have at most polynomial growth. The fact that the value function V satisfies (16) follows from Lemma 5.1 (pp. 11) of Larssen [13].

We next show the continuity of $V(t, \psi)$ with respect to t. Let $\Xi_1(s) = X_s(t_1, \psi, u(\cdot))$ and $\Xi_2(s) = X_s(t_2, \psi, u(\cdot))$, $s \in [t, T]$, be two $\mathbf{C}$-valued solutions of (2) with $u(\cdot) \in \mathcal{U}[0, T]$ and the initial data (t_1, ψ) and (t_2, ψ), respectively.

Without loss of generality, let us assume that $0 \leq t_1 < t_2 \leq T$ then

$$\begin{aligned} & J(t_1, \psi; u(\cdot)) - J(t_2, \psi; u(\cdot)) \qquad (17) \\ = \ & \mathbb{E}\bigg[\int_{t_1}^{t_2} e^{-\rho(\xi - t_1)}[L(\xi, \Xi_1(\xi), u(\xi))]d\xi \\ + \ & \int_{t_2}^{T} e^{-\rho(\xi - t_2)}[L(\xi, \Xi_1(\xi), u(\xi)) - L(\xi, \Xi_2(\xi), u(\xi))]d\xi \\ + \ & e^{-\rho(T - t_1)}\Psi(\Xi_1(T)) - e^{-\rho(T - t_2)}\Psi(\Xi_2(T))\bigg] \end{aligned}$$

Therefore, there exists a constant $\Lambda > 0$ such that

$$\begin{aligned} & |J(t_1, \psi, u(\cdot)) - J(t_2, \psi; u(\cdot))| \\ \leq \ & \Lambda\bigg(|t_1 - t_2|\mathbb{E}\|\Xi_1(T)\| + \mathbb{E}\|\Xi_1(T) - \Xi_2(T)\| \bigg) \end{aligned} \tag{18}$$

Now let $\varepsilon > 0$ be a small enough constant. Using the compactness of $[0, T]$ and the uniform continuity of the trajectory map in t, we know that there exists $\eta > 0$ such that if $|t_1 - t_2| < \eta$ then $\mathbb{E}\|\Xi_1(s) - \Xi_2(s)\| \leq \frac{\varepsilon}{2\Lambda}$ for all $s \in [t_2, T]$. Therefore, for

$$|t_1 - t_2| < \min \ \left(\eta, \frac{\varepsilon}{2\Lambda\mathbb{E}[\sup_{s \in [0,T]} \|\Xi_1(s)\|]} \right)$$

we have

$$|J(t_1, \psi, u(\cdot)) - J(t_2, \psi, u(\cdot))| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Consequently,

$$|V(t_1, \psi) - V(t_2, \psi)| \leq \varepsilon.$$

□

We have the following result:

Theorem 4.3 *The value function V is a viscosity solution of the HJB equation*

$$\rho V(t,\psi)-\frac{\partial V}{\partial t}(t,\psi)-\max_{v\in U}\ [\mathcal{A}^v V(t,\psi)+L(t,\psi,v)]=0 \tag{19}$$

on $[0,T]\times\mathbf{C}$, *and* $V(T,\psi)=\Psi(\psi),\ \forall\psi\in\mathbf{C}$.

Proof. Let $\Gamma\in C^{1,2}_{lip}([0,T]\times\mathbf{C})\cap\mathcal{D}(\mathcal{S})$. For $(t,\psi)\in[0,T]\times\mathbf{C}$ such that $\Gamma\le V$ on $[0,T]\times\mathbf{C}$ and $\Gamma(t,\psi)=V(t,\psi)$, we want to prove the viscosity supersolution inequality, *i.e.*,

$$\rho\Gamma(t,\psi)-\frac{\partial\Gamma}{\partial t}(t,\psi)-\max_{v\in U}[\mathcal{A}^v\Gamma(t,\psi)+L(t,\psi,v)]\ge 0. \tag{20}$$

Let $u(\cdot)\in\mathcal{U}[t,T]$. Since $\Gamma\in C^{1,2}_{lip}([0,T]\times\mathbf{C})\cap\mathcal{D}(\mathcal{S})$, (by virtue of Theorem 3.1 pp. 78 of Mohammed [17]) for $t\le s\le T$, we have

$$\begin{aligned}
&\mathbb{E}\Big[e^{-\rho(s-t)}\Gamma(s,X_s(t,\psi,u(\cdot)))\Big]-\Gamma(t,\psi)\\
=\ &\mathbb{E}\Big[\int_t^s e^{-\rho(\xi-t)}\Big(\frac{\partial\Gamma(\xi,X_\xi(t,\psi,u(\cdot)))}{\partial\xi}\\
&\qquad+\mathcal{A}^{u(\xi)}\Gamma(\xi,X_\xi(t,\psi,u(\cdot)))-\rho\Gamma(\xi,X_\xi(t,\psi,u(\cdot)))\Big)d\xi\Big].
\end{aligned} \tag{21}$$

On the other hand, for any $s\in[t,T]$, the dynamic programming principle (Theorem (Larssen)) gives,

$$\begin{aligned}
V(t,\psi)\ =\ &\max_{u(\cdot)\in\mathcal{U}[t,T]}\mathbb{E}\Big\{\int_t^s e^{-\rho(\xi-t)}L(\xi,X_\xi(t,\psi,u(\cdot)),u(\xi))d\xi\\
&+e^{-\rho(s-t)}V(s,X_s(t,\psi,u(\cdot)))\Big\}.
\end{aligned} \tag{22}$$

Therefore, we have

$$\begin{aligned}
V(t,\psi)\ \ge\ &\mathbb{E}\Big[\int_t^s e^{-\rho(\xi-t)}L(\xi,X_\xi(t,\psi,u(\cdot)),u(\xi))d\xi\Big]\\
&+\mathbb{E}\Big[e^{-\rho(s-t)}V(s,X_s(t,\psi,u(\cdot)))\Big].
\end{aligned} \tag{23}$$

By virtue of (21) and using $\Gamma\le V,\ \Gamma(t,\psi)=V(t,\psi)$, we can get

$$\begin{aligned}
0\ \ge\ &\mathbb{E}\Big[\int_t^s e^{-\rho(\xi-t)}L(\xi,X_\xi(t,\psi,u(\cdot)),u(\xi))d\xi\Big]\\
&+\mathbb{E}\Big[e^{-\rho(s-t)}V(s,X_s(t,\psi,u(\cdot)))\Big]-V(t,\psi)\\
\ge\ &\mathbb{E}\Big[\int_t^s e^{-\rho(\xi-t)}L(\xi,X_\xi(t,\psi,u(\cdot)),u(\xi))d\xi\Big]\\
&+\mathbb{E}\Big[e^{-\rho(s-t)}\Gamma(s,X_s(t,\psi,u(\cdot)))\Big]-\Gamma(t,\psi)\\
\ge\ &\mathbb{E}\int_t^s e^{-\rho(\xi-t)}\Big[-\rho\Gamma(\xi,X_\xi(t,\psi,u(\cdot)))+\frac{\partial\Gamma(\xi,X_\xi(t,\psi,u(\cdot)))}{\partial\xi}\\
&\qquad+\mathcal{A}^u\Gamma(\xi,X_\xi(t,\psi,u(\cdot)))+L(\xi,X_\xi(t,\psi,u(\cdot)),u(\xi))\Big]d\xi.
\end{aligned}$$

Dividing both sides by $(s-t)$, we have

$$\begin{aligned}
0 \quad \leq \quad & \mathbb{E}\bigg[\frac{1}{s-t}\int_t^s e^{-\rho(\xi-t)}\Big(\rho\Gamma(\xi, X_\xi(t,\psi,u(\cdot))) \\
& \quad -\frac{\partial\Gamma(\xi, X_\xi(t,\psi,u(\cdot)))}{\partial\xi} - \mathcal{A}^{u(\xi)}\Gamma(\xi, X_\xi(t,\psi,u(\cdot))) \\
& \quad -L(\xi, X_\xi(t,\psi,u(\cdot)), u(\xi))\Big)d\xi\bigg]. \qquad (24)
\end{aligned}$$

Now let $s \downarrow t$ in (24). Noting that $\lim_{s\downarrow t} u(s) = v$, we can obtain

$$\rho\Gamma(t,\psi) - \frac{\partial\Gamma}{\partial t}(t,\psi) - [\mathcal{A}^v\Gamma(t,\psi) + L(t,\psi,v)] \geq 0. \qquad (25)$$

Since $v \in U$ is arbitrary, we now can derive the inequality (20).

Next we want to prove that V is a viscosity subsolution. Let $\Gamma \in C^{1,2}_{lip}([0,T]\times\mathbf{C})\cap\mathcal{D}(\mathcal{S})$. For $(t,\psi) \in [0,T]\times\mathbf{C}$ satisfying $\Gamma \geq V$ on $[0,T]\times\mathbf{C}$ and $\Gamma(t,\psi) = V(t,\psi)$, we want to prove that

$$\rho\Gamma(t,\psi) - \frac{\partial\Gamma}{\partial t}(t,\psi) - \max_{v\in U}[\mathcal{A}^v\Gamma(t,\psi) + L(t,\psi,v)] \leq 0. \qquad (26)$$

We assume the contrary, and we will find a contradiction.

Let us assume that, there exist $(t,\psi) \in [0,T]\times\mathbf{C}$, $\Gamma \in C^{1,2}_{lip}([0,T]\times\mathbf{C})\cap\mathcal{D}(\mathcal{S})$, with $\Gamma \geq V$ on $[0,T]\times\mathbf{C}$ and $\Gamma(t,\psi) = V(t,\psi)$, and $\delta > 0$ such that for all control $u(\cdot) \in \mathcal{U}[t,T]$ with $\lim_{s\downarrow t} u(s) = v$,

$$\rho\Gamma(\tau,\phi) - \frac{\partial\Gamma}{\partial t}(\tau,\phi) - \mathcal{A}^v\Gamma(\tau,\phi) - L(\tau,\phi,v) \geq \delta \qquad (27)$$

for all $(\tau,\phi) \in N(t,\psi)$, where $N(t,\psi)$ is a neighborhood of (t,ψ).

Let $u(\cdot) \in \mathcal{U}[t,T]$ with $\lim_{s\downarrow t} u(s) = v$, and let t_1 be such that for $t \leq s \leq t_1$, we can get that the solution $X(s;t,\psi,u(\cdot)) \in N(t,\psi)$. Therefore, for any $s \in [t,t_1]$, we have almost surely

$$\begin{aligned}
&\rho\Gamma(s, X_s(t,\psi,u(\cdot))) - \frac{\partial\Gamma}{\partial t}(s, X_s(t,\psi,u(\cdot))) \\
&\quad -\mathcal{A}^v\Gamma(s, X_s(t,\psi,u(\cdot))) - L(s, X_s(t,\psi,u(\cdot)), u(s)) \geq \delta. \qquad (28)
\end{aligned}$$

On the other hand, since $\Gamma \geq V$, using the definition of J and V, we can get

$$\begin{aligned}
J(t,\psi;u(\cdot)) \quad \leq \quad & \mathbb{E}\left[\int_t^{t_1} e^{-\rho(s-t)}L(s,X_s,u(s))ds + e^{-\rho(t_1-t)}V(t_1,X_{t_1})\right], \\
\leq \quad & \mathbb{E}\bigg[\int_t^{t_1} e^{-\rho(s-t)}L(s,X_s,u(s))ds \\
& \qquad + e^{-\rho(t_1-t)}\Gamma(t_1, X_{t_1}(t,\psi,u(\cdot)))\bigg].
\end{aligned}$$

Using (28), we have

$$\begin{aligned} J(t,\psi;u(\cdot)) &\leq \mathbb{E}\bigg[\int_t^{t_1} e^{-\rho(s-t)}\Big(-\delta+\rho\Gamma(s,X_s(t,\psi,u(\cdot))) \\ &\qquad -\frac{\partial\Gamma}{\partial t}(s,X_s(t,\psi,u(\cdot)))-\mathcal{A}^{u(s)}\Gamma(s,X_s(t,\psi,u(\cdot)))\Big)ds \\ &\qquad + e^{-\rho(t_1-t)}\Gamma(t_1,X_{t_1}(t,\psi,u(\cdot)))\bigg]. \end{aligned} \tag{29}$$

In addition, similar to (21), we can get

$$\begin{aligned} &\mathbb{E}\left[e^{-\rho(t_1-t)}\Gamma(t_1,X_{t_1}(t,\psi,u(\cdot)))\right]-\Gamma(t,\psi) \\ &=\mathbb{E}\bigg[\int_t^{t_1} e^{-\rho(s-t)}\Big(\frac{\partial\Gamma(s,X_s(t,\psi,u(\cdot)))}{\partial s}+\mathcal{A}^{u(s)}\Gamma(s,X_s(t,\psi,u(\cdot))) \\ &\qquad -\rho\Gamma(s,X_s(t,\psi,u(\cdot)))\Big)ds\bigg]. \end{aligned} \tag{30}$$

Therefore, by virtue of (29), (30), we can get

$$\begin{aligned} J(t,\psi;u(\cdot)) &\leq -\int_t^{t_1} e^{-\rho(s-t)}\delta ds+\Gamma(t,\psi) \\ &= -\frac{\delta}{\rho}(1-e^{-\rho(t_1-t)})+V(t,\psi). \end{aligned} \tag{31}$$

Taking the supremum over all admissible control $u(\cdot)\in\mathcal{U}[t,T]$, we have

$$V(t,\psi)\leq -\frac{\delta}{\rho}(1-e^{-\rho(t_1-t)})+V(t,\psi).$$

This contradicts the fact that $\delta>0$. Therefore $V(t,\psi)$ is a viscosity subsolution. This completes the proof of the theorem. □

5 Uniqueness

In this section, we will show the uniqueness result for the viscosity solution of the HJB equation (14). Given the uniqueness result, combining the results we obtained in last section, we can say that the value function $V(t,\psi)$ is the only viscosity solution of the HJB equation (14).

Since a viscosity solution is both a subsolution and a supersolution, the uniqueness result will follow immediately after we establish the following comparison principle:

Theorem 5.1 (Comparison Principle) *Assume that $V_1(t,\psi)$ and $V_2(t,\psi)$ are both continuous with respect to the argument (t,ψ) and are respectively viscosity subsolution and supersolution of (14) with at most a polynomial growth. In other terms, there exists a real number $\Lambda>0$ and a positive integer $k>0$ such that,*

$$|V_i(t,\psi)|\leq\Lambda(1+\|\psi\|_2)^k,\ \ for\ \ (t,\psi)\in[0,T]\times\mathbf{C},\ i=1,2.$$

Then

$$V_1(t,\psi) \le V_2(t,\psi) \quad \text{for all } (t,\psi) \in [0,T] \times \mathbf{C}. \tag{32}$$

Before we proceed to the proof of Theorem 5.1, we shall use the following result proven in Ekeland and Lebourg [5] and also in a general form in Stegall [21] and Bourgain [2]. The readers are also referred to Crandall et al [4] and Lions [15] for an application example of this result in a setting similar to ours.

Lemma 5.2 *Let Φ be a bounded and upper semicontinuous real-valued function on a closed ball B of a Banach space $\mathcal{X}$ which has the Radon-Nikodym property, then for any $\epsilon > 0$ there exists an element $u^* \in \mathcal{X}^*$ with norm at most ϵ, where $\mathcal{X}^*$ is the dual of $\mathcal{X}$, such that $\Phi + u^*$ attains its maximum on B.*

Note that every separable Hilbert space $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ satisfy the Radon-Nikodym property (see e.g. Ekeland and Lebourg [5]). In order to apply Lemma 5.2, we shall therefore restrict ourself to a subspace $\mathcal{X}$ of $\mathbf{C} = C([-r,0];\Re^n)$ which is a separable Hilbert space and dense in $\mathbf{C}$. One of the good candidates is the Sobolev space $W^{1,2}((-r,0);\Re^n)$, where

$$\begin{aligned} W^{1,2}((-r,0);\Re^n) &= \{\phi \in \mathbf{C} \mid \phi \text{ is absolutely continuous} \\ &\qquad \text{and } \|\phi\|_{1,2} < \infty\}, \end{aligned}$$

$$\|\phi\|_{1,2}^2 \equiv \|\phi\|_2^2 + \|\dot{\phi}\|_2^2,$$

and $\dot{\phi}$ is the derivative of ϕ in the distributional sense.

From the Sobolev embedding theorems (see e.g. Adams [1]), it is known that $W^{1,2}((-r,0);\Re^n) \subset \mathbf{C}$ and that $W^{1,2}((-r,0);\Re^n)$ is dense in $\mathbf{C}$. For more about Sobolev spaces and corresponding results, one can refer to Adams [1].

Let V_1 and V_2 be a viscosity subsolution and supersolution of (14), respectively. For any $0 < \delta, \gamma < 1$, and for all $\psi, \phi \in \mathbf{C}$ and $t, s \in [0,T]$, define

$$\begin{aligned} \Theta_{\delta\gamma}(t,s,\psi,\phi) &\equiv \frac{1}{\delta}\Big[\|\psi-\phi\|_2^2 + \|\psi^0-\phi^0\|_2^2 + |t-s|^2\Big] \\ &+ \gamma\left[\exp(1+\|\psi\|_2^2+\|\psi^0\|_2^2) + \exp(1+\|\phi\|_2^2+\|\phi^0\|_2^2)\right], \end{aligned} \tag{33}$$

and

$$\Phi_{\delta\gamma}(t,s,\psi,\phi) \equiv V_1(t,\psi) - V_2(s,\phi) - \Theta_{\delta\gamma}(t,s,\psi,\phi), \tag{34}$$

where $\psi^0, \phi^0 \in W^{1,2}((-r,0);\Re^n)$ with $\psi^0(\theta) = \frac{\theta}{-r}\psi(-r-\theta)$, $\phi^0(\theta) = \frac{\theta}{-r}\phi(-r-\theta)$ for $\theta \in [-r,0]$.

Moreover, using the polynomial growth condition for V_1 and V_2, we have

$$\lim_{\|\psi\|_2+\|\phi\|_2\to\infty} \Phi_{\delta\gamma}(t,s,\psi,\phi) = -\infty. \tag{35}$$

The function $\Phi_{\delta\gamma}$ is a real-valued function. In addition, if we have $\psi, \phi \in W^{1,2}((-r,0);\Re^n)$, then it is easy to verify that $\Phi_{\delta\gamma}$ is bounded above and continuous on $[0,T] \times [0,T] \times W^{1,2}((-r,0);\Re^n) \times W^{1,2}((-r,0);\Re^n)$. Therefore, from Lemma 5.2 (which is applicable by virtue of (35)), for any $1 > \epsilon > 0$ there exits a continuous linear functional T_ϵ in the topological dual of $W^{1,2}((-r,0);\Re^n) \times W^{1,2}((-r,0);\Re^n)$, with norm at most ϵ, such that the function $\Phi_{\delta\gamma}+T_\epsilon$ attains it maximum in $[0,T]\times[0,T]\times W^{1,2}((-r,0);\Re^n)\times W^{1,2}((-r,0);\Re^n)$ (see Lemma 5.2). Let us denote by $(t_{\delta\gamma\epsilon}, s_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon})$ the global maximum of $\Phi_{\delta\gamma} + T_\epsilon$ on $[0,T] \times [0,T] \times W^{1,2}((-r,0);\Re^n) \times W^{1,2}((-r,0);\Re^n)$.

Without loss of generality, we assume that for any given δ, γ and ϵ, there exists a constant $M_{\delta\gamma\epsilon}$ such that the maximum value $\Phi_{\delta\gamma} + T_\epsilon + M_{\delta\gamma\epsilon}$ is zero. In other words, we have

$$\Phi_{\delta\gamma}(t_{\delta\gamma\epsilon}, s_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) + T_\epsilon(\psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) + M_{\delta\gamma\epsilon} = 0. \tag{36}$$

We have the following lemmas.

Lemma 5.3 *$(t_{\delta\gamma\epsilon}, s_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon})$ is the global maximum of $\Phi_{\delta\gamma} + T_\epsilon$ over $[0,T] \times [0,T] \times \mathbf{C} \times \mathbf{C}$.*

Proof. Let $(t, s, \psi, \phi) \in [0,T]\times[0,T]\times\mathbf{C}\times\mathbf{C}$. By virtue of the density of $W^{1,2}((-r,0);\Re^n)$ in $\mathbf{C}$, we can find a sequence $(t_k, s_k, \psi_k, \phi_k)$ in $[0,T]\times[0,T]\times W^{1,2}((-r,0);\Re^n)\times W^{1,2}((-r,0);\Re^n)$ such that

$$(t_k, s_k, \psi_k, \phi_k) \to (t, s, \psi, \phi) \text{ as } k \to \infty.$$

Since $(t_{\delta\gamma\epsilon}, s_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon})$ is the global maximum of $\Phi_{\delta\gamma}+T_\epsilon$ on $[0,T]\times[0,T]\times W^{1,2}((-r,0);\Re^n)\times W^{1,2}((-r,0);\Re^n)$, we have

$$\begin{aligned} & \Phi_{\delta\gamma}(t_k, s_k, \psi_k, \phi_k) + T_\epsilon(\psi_k, \phi_k) \\ \leq \; & \Phi_{\delta\gamma}(t_{\delta\gamma\epsilon}, s_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) + T_\epsilon(\psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}). \end{aligned} \tag{37}$$

Taking the limit as k goes to ∞ in the last inequality, we obtain

$$\begin{aligned} & \Phi_{\delta\gamma}(t, s, \psi, \phi) + T_\epsilon(\psi, \phi) \\ \leq \; & \Phi_{\delta\gamma}(t_{\delta\gamma\epsilon}, s_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) + T_\epsilon(\psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}). \end{aligned} \tag{38}$$

This shows that $(t_{\delta\gamma\epsilon}, s_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon})$ is the global maximum over $[0,T] \times [0,T] \times \mathbf{C} \times \mathbf{C}$. □

Lemma 5.4 *For each fixed $\gamma > 0$, we can find a constant $\Lambda_\gamma > 0$ such that*

$$\|\psi_{\delta\gamma\epsilon}\|_2 + \|\psi^0_{\delta\gamma\epsilon}\|_2 + \|\phi_{\delta\gamma\epsilon}\|_2 + \|\phi^0_{\delta\gamma\epsilon}\|_2 \leq \Lambda_\gamma, \tag{39}$$

and

$$\lim_{\epsilon\downarrow 0, \delta\downarrow 0} \Big(\|\psi_{\delta\gamma\epsilon} - \phi_{\delta\gamma\epsilon}\|_2^2 + \|\psi^0_{\delta\gamma\epsilon} - \phi^0_{\delta\gamma\epsilon}\|_2^2 + |t_{\delta\gamma\epsilon} - s_{\delta\gamma\epsilon}|^2 \Big) = 0, \tag{40}$$

Proof. Noting that $(t_{\delta\gamma\epsilon}, s_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon})$ is the global maximum of $\Phi_{\delta\gamma} + T_\epsilon$, we can get that

$$\begin{aligned} & \Phi_{\delta\gamma}(t_{\delta\gamma\epsilon}, t_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}) + T_\epsilon(\psi_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}) \\ & + \Phi_{\delta\gamma}(s_{\delta\gamma\epsilon}, s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) + T_\epsilon(\phi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) \\ \leq\ & 2\Phi_{\delta\gamma}(t_{\delta\gamma\epsilon}, s_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) + 2T_\epsilon(\psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}). \end{aligned} \tag{41}$$

It implies

$$\begin{aligned} & V_1(t_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}) - V_2(t_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}) - 2\gamma(\exp(1 + \|\psi_{\delta\gamma\epsilon}\|_2^2 + \|\psi^0_{\delta\gamma\epsilon}\|_2^2)) \\ & + T_\epsilon(\psi_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}) + V_1(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) - V_2(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) \\ & - 2\gamma(\exp(1 + \|\phi_{\delta\gamma\epsilon}\|_2^2 + \|\phi^0_{\delta\gamma\epsilon}\|_2^2)) + T_\epsilon(\phi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) \\ \leq\ & 2V_1(t_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}) - 2V_2(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) - \frac{2}{\delta}\Big[\|\psi_{\delta\gamma\epsilon} - \phi_{\delta\gamma\epsilon}\|_2^2 \\ & + \|\psi^0_{\delta\gamma\epsilon} - \phi^0_{\delta\gamma\epsilon}\|_2^2 + |t_{\delta\gamma\epsilon} - s_{\delta\gamma\epsilon}|^2\Big] - 2\gamma\Big(\exp(1 + \|\psi_{\delta\gamma\epsilon}\|_2^2 \\ & + \|\psi^0_{\delta\gamma\epsilon}\|_2^2) + \exp(1 + \|\phi_{\delta\gamma\epsilon}\|_2^2 + \|\phi^0_{\delta\gamma\epsilon}\|_2^2)\Big) \\ & + 2T_\epsilon(\psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}). \end{aligned} \tag{42}$$

From the above inequality, it is easy to get that

$$\begin{aligned} & \frac{2}{\delta}\Big[\|\psi_{\delta\gamma\epsilon} - \phi_{\delta\gamma\epsilon}\|_2^2 + \|\psi^0_{\delta\gamma\epsilon} - \phi^0_{\delta\gamma\epsilon}\|_2^2 + |t_{\delta\gamma\epsilon} - s_{\delta\gamma\epsilon}|^2\Big] \\ \leq\ & [V_1(t_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}) - V_1(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon})] + [V_2(t_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}) - V_2(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon})] \\ & + 2T_\epsilon(\psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) - [T_\epsilon(\psi_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}) + T_\epsilon(\phi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon})]. \end{aligned} \tag{43}$$

From the polynomial growth condition of V_1, V_2, and the fact that the norm of T_ϵ is $\epsilon \in (0,1)$, we can find a constant Λ and a positive integer $k \geq 2$ such that

$$\begin{aligned} & \frac{2}{\delta}\Big[\|\psi_{\delta\gamma\epsilon} - \psi_{\delta\gamma\epsilon}\|_2^2 + \|\psi^0_{\delta\gamma\epsilon} - \phi^0_{\delta\gamma\epsilon}\|_2^2 + |t_{\delta\gamma\epsilon} - s_{\delta\gamma\epsilon}|^2\Big] \\ \leq\ & \Lambda(1 + \|\psi_{\delta\gamma\epsilon}\|_2 + \|\phi_{\delta\gamma\epsilon}\|_2)^k. \end{aligned} \tag{44}$$

So

$$\begin{aligned} & \|\psi_{\delta\gamma\epsilon} - \phi_{\delta\gamma\epsilon}\|_2^2 + \|\psi^0_{\delta\gamma\epsilon} - \phi^0_{\delta\gamma\epsilon}\|_2^2 + |t_{\delta\gamma\epsilon} - s_{\delta\gamma\epsilon}|^2 \\ \leq\ & \delta\Lambda(1 + \|\psi_{\delta\gamma\epsilon}\|_2 + \|\phi_{\delta\gamma\epsilon}\|_2)^k, \end{aligned} \tag{45}$$

On the other hand, because $(t_{\delta\gamma\epsilon}, s_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon})$ is the global maximum of $\Phi_{\delta\gamma} + T_\epsilon$, we can get

$$\begin{aligned} & \Phi_{\delta\gamma}(t_{\delta\gamma\epsilon}, s_{\delta\gamma\epsilon}, 0, 0) + T_\epsilon(0,0) \\ \leq\ & \Phi_{\delta\gamma}(t_{\delta\gamma\epsilon}, s_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) + T_\epsilon(\psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}), \end{aligned} \tag{46}$$

In addition, by the definition of $\Phi_{\delta\gamma}$ and the polynomial growth condition of V_1, V_2, we can get a $\Lambda > 0$ and a positive integer $k \geq 2$ such that

$$|\Phi_{\delta\gamma}(t_{\delta\gamma\epsilon}, s_{\delta\gamma\epsilon}, 0, 0) + T_\epsilon(0,0)| \leq \Lambda(1 + \|\psi_{\delta\gamma\epsilon}\|_2 + \|\phi_{\delta\gamma\epsilon}\|_2)^k,$$

and

$$V_1(t_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}) - V_2(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) \leq \Lambda(1 + \|\psi_{\delta\gamma\epsilon}\|_2 + \|\phi_{\delta\gamma\epsilon}\|_2)^k.$$

Therefore, by virtue of (46) and the definition of $\Phi_{\delta\gamma}$, we have

$$\begin{aligned}
& \gamma\left[\exp(1 + \|\psi_{\delta\gamma\epsilon}\|_2^2 + \|\psi^0_{\delta\gamma\epsilon}\|_2^2) + \exp(1 + \|\phi_{\delta\gamma\epsilon}\|_2^2 + \|\phi^0_{\delta\gamma\epsilon}\|_2^2)\right] \\
\leq\ & V_1(t_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}) - V_2(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) \\
& -\frac{1}{\delta}\left[\|\psi_{\delta\gamma\epsilon} - \phi_{\delta\gamma\epsilon}\|_2^2 + \|\psi^0_{\delta\gamma\epsilon} - \phi^0_{\delta\gamma\epsilon}\|_2^2 + |t_{\delta\gamma\epsilon} - s_{\delta\gamma\epsilon}|^2\right] \\
& - \Phi_{\delta\gamma}(t_{\delta\gamma\epsilon}, s_{\delta\gamma\epsilon}, 0, 0) + T_\epsilon(\psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) - T_\epsilon(0,0) \\
\leq\ & 3\Lambda(1 + \|\psi_{\delta\gamma\epsilon}\|_2 + \|\phi_{\delta\gamma\epsilon}\|_2)^k. \qquad (47)
\end{aligned}$$

It follows that

$$\frac{\gamma\left[\exp(1 + \|\psi_{\delta\gamma\epsilon}\|_2^2 + \|\psi^0_{\delta\gamma\epsilon}\|_2^2) + \exp(1 + \|\phi_{\delta\gamma\epsilon}\|_2^2 + \|\phi^0_{\delta\gamma\epsilon}\|_2^2)\right]}{(1 + \|\psi_{\delta\gamma\epsilon}\|_2 + \|\phi_{\delta\gamma\epsilon}\|_2)^k} \leq 3\Lambda.$$

Consequently, there exists $\Lambda_\gamma < \infty$ such that

$$\|\psi_{\delta\gamma\epsilon}\|_2 + \|\psi^0_{\delta\gamma\epsilon}\|_2 + \|\phi_{\delta\gamma\epsilon}\|_2 + \|\phi^0_{\delta\gamma\epsilon}\|_2 \leq \Lambda_\gamma. \qquad (48)$$

In order to obtain (40), we send δ to zero in (45) using the above inequality. $\square$

Now let us introduce a functional $F : \mathbf{C} \to \Re$ defined by

$$F(\psi) \equiv \|\psi\|_2^2, \qquad (49)$$

and the linear map $H : \mathbf{C} \to \mathbf{C}$ defined by

$$H(\psi)(\theta) \equiv \frac{\theta}{-r}\psi(-r-\theta) = \psi^0(\theta), \qquad \theta \in [-r, 0]. \qquad (50)$$

Note that $H(\psi)(0) = \psi^0(0) = 0$ and $H(\psi)(-r) = \psi^0(-r) = -\psi(0)$. It is not hard to show that the map F is Fréchet differentiable and its derivative is given by $DF(u)h = 2(u|h)$. This comes from the fact that

$$\|\psi + h\|_2^2 - \|\psi\|_2^2 = 2(\psi|h) + \|h\|_2^2,$$

and we can always find a constant $\Lambda > 0$ such that

$$\frac{|\|\psi + h\|_2^2 - \|\psi\|_2^2 - 2(\psi|h)}{\|h\|} = \frac{\|h\|_2^2}{\|h\|} \leq \frac{\Lambda\|h\|^2}{\|h\|} = \Lambda\|h\|. \qquad (51)$$

Moreover, we have

$$2(\psi + h|\cdot) - 2(\psi|\cdot) = 2(h|\cdot).$$

We deduce that F is twice differentiable and $D^2F(u)(h,k) = 2(h|k)$.

In addition, the map H is linear, thus twice Fréchet differentiable. Therefore, we can get $DH(\psi)(h) = H(h)$ and $D^2H(\psi)(h,k) = 0$, for all $\psi, h, k \in \mathbf{C}$.

From the definition of $\Theta_{\delta\gamma}$ and the definition of F, we have

$$\begin{aligned}\Theta_{\delta\gamma}(t,s,\psi,\phi) &= \frac{1}{\delta}\Big[F(\psi-\phi) + F(\psi^0-\phi^0) + |t-s|^2\Big] \\ &+ \gamma[e^{1+F(\psi)+F(H(\psi))} + e^{1+F(\phi)+F(H(\phi))}].\end{aligned}$$

The following chain rule, quoted from [20] (Theorem 5.2.5 on page 208), is needed to get the Fréchet derivatives of $\Theta_{\delta\gamma}$:

Theorem 5.5 (Chain Rule.) *Let $\mathcal{X}, \mathcal{Y}$, and $\mathcal{Z}$ be real Banach spaces. If $S : \mathcal{X} \to \mathcal{Y}$ and $T : \mathcal{Y} \to \mathcal{Z}$ are Fréchet differentiable at $\mathbf{x}$ and $\mathbf{y} = S(\mathbf{x}) \in \mathcal{Y}$, respectively, then $U = T \circ S$ is Fréchet differentiable at $\mathbf{x}$ and its Fréchet derivative is given by*

$$D_{\mathbf{x}}U(\mathbf{x}) = D_{\mathbf{y}}T(S(\mathbf{x}))D_{\mathbf{x}}S(\mathbf{x}).$$

Given the above chain rule, we can say that $\Theta_{\delta\gamma}$ is Fréchet differentiable. Actually, for $h, k \in \mathbf{C}$, we can get

$$\begin{aligned}&D_\psi\Theta_{\delta\gamma}(t,s,\psi,\phi)(h) \\ =\ &\frac{2}{\delta}\Big[(\psi-\phi|h) + (H(\psi-\phi)|H(h))\Big] \\ &+ 2\gamma e^{1+F(\psi)+F(H(\psi))}[(\psi|h) + (H(\psi)|H(h))].\end{aligned} \tag{52}$$

Similarly,

$$\begin{aligned}&D_\phi\Theta_{\delta\gamma}(t,s,\psi,\phi)(k) \\ =\ &\frac{2}{\delta}\Big[(\phi-\psi|k) + (H(\phi-\psi)|H(k))\Big] \\ &+ 2\gamma e^{1+F(\phi)+F(H(\phi))}[(\phi|k) + (H(\phi)|H(k))].\end{aligned} \tag{53}$$

Furthermore,

$$\begin{aligned}&D^2_\psi\Theta_{\delta\gamma}(t,s,\psi,\phi)(h,k) \\ =\ &\frac{2}{\delta}\Big[(h|k) + (H(h)|H(k))\Big] \\ &+ 2\gamma e^{1+F(\psi)+F(H(\psi))}\Big[2((\psi|k) + (H(\psi)|H(k)))((\psi|h) + (H(\psi)|H(h))) \\ &+ (k|h) + (H(k)|H(h))\Big].\end{aligned} \tag{54}$$

Similarly,

$$
\begin{aligned}
& D^2_\phi \Theta_{\delta\gamma}(t,s,\psi,\phi)(h,k) \\
= \;& \frac{2}{\delta}\Big[(h|k) + (H(h)|H(k))\Big] \\
& + 2\gamma e^{1+F(\phi)+F(H(\phi))}\Big[2((\phi|k) + (H(\phi)|H(k)))((\phi|h) + (H(\phi)|H(h))) \\
& + (k|h) + (H(k)|H(h))\Big]. \qquad (55)
\end{aligned}
$$

On the other hand, by the Hahn-Banach theorem (see e.g. [20]), we can extend the continuous linear functional T_ϵ to the space $\mathbf{C}\times\mathbf{C}$ and keep its norm preserved. Thus, the first order Fréchet derivatives of T_ϵ is just T_ϵ, *i.e.,* $\forall \psi,\phi,h,k\in\mathbf{C}$, we have

$$
\begin{aligned}
D_\psi T_\epsilon(\psi,\phi)h &= T_\epsilon(h,\phi), \qquad (56)\\
D_\phi T_\epsilon(\psi,\phi)k &= T_\epsilon(\psi,k). \qquad (57)
\end{aligned}
$$

Moreover, for the second derivatives, $\forall \psi,\phi,h,k\in\mathbf{C}$, we have

$$
\begin{aligned}
D^2_\psi T_\epsilon(\psi,\phi)(h,k) &= 0, \qquad (58)\\
D^2_\phi T_\epsilon(\psi,\phi)(h,k) &= 0. \qquad (59)
\end{aligned}
$$

Observe that we can extend $D_\psi\Theta_{\delta\gamma}(t,s,\psi,\phi)$ and $D^2_\psi\Theta_{\delta\gamma}(t,s,\psi,\phi)$, the first and second order Fréchet derivatives of $\Theta_{\delta\gamma}$ with respect to ψ, to the space $\mathbf{C}\oplus\mathbf{B}$ (see Section 3 of this paper or Lemma (3.1) and Lemma (3.2) on pp 79-83 of Mohammed [17]) by setting

$$
\begin{aligned}
& \overline{D_\psi\Theta_{\delta\gamma}(t,s,\psi,\phi)}(h+v\mathbf{1}_{\{0\}}) \\
= \;& \frac{2}{\delta}\Big[(\psi-\phi|h+v\mathbf{1}_{\{0\}}) + (H(\psi-\phi)|H(h+v\mathbf{1}_{\{0\}})\Big] \\
& + 2\gamma e^{1+F(\psi)+F(H(\psi))}[(\psi|h+v\mathbf{1}_{\{0\}}) + (H(\psi)|H(h+v\mathbf{1}_{\{0\}}))]. \qquad (60)
\end{aligned}
$$

and

$$
\begin{aligned}
& \overline{D^2_\psi\Theta_{\delta\gamma}(t,s,\psi,\phi)}(h+v\mathbf{1}_{\{0\}},k+w\mathbf{1}_{\{0\}}) \\
= \;& \frac{2}{\delta}\Big[(h+v\mathbf{1}_{\{0\}}|k+w\mathbf{1}_{\{0\}}) + (H(h+v\mathbf{1}_{\{0\}})|H(k+w\mathbf{1}_{\{0\}})\Big] \\
& + 2\gamma e^{1+F(\psi)+F(H(\psi))}\Big[2((\psi|k+w\mathbf{1}_{\{0\}}) + (H(\psi)|H(k+w\mathbf{1}_{\{0\}}))) \\
& \cdot((\psi|h+v\mathbf{1}_{\{0\}}) + (H(\psi)|H(h+v\mathbf{1}_{\{0\}}))) \\
& + (k+w\mathbf{1}_{\{0\}}|h+v\mathbf{1}_{\{0\}}) + (H(k+w\mathbf{1}_{\{0\}})|H(h+v\mathbf{1}_{\{0\}}))\Big], \qquad (61)
\end{aligned}
$$

for $v,w\in\Re^n$ and $h,k\in\mathbf{C}$.

Moreover, it is easy to see that these extensions are continuous in that there exists a constant $\Lambda>0$ such that

$$
\begin{aligned}
|(\psi-\phi|h+v\mathbf{1}_{\{0\}})| &\le \|\psi-\phi\|_2\cdot\|h+v\mathbf{1}_{\{0\}}\|_2 \\
&\le \Lambda\|\psi-\phi\|_2(\|h\|+|v|); \qquad (62)
\end{aligned}
$$

$$
\begin{aligned}
|(\psi|h+v\mathbf{1}_{\{0\}})| &\leq \|\psi\|_2\cdot\|h+v\mathbf{1}_{\{0\}}\|_2 \\
&\leq \Lambda\|\psi\|_2(\|h\|+|v|);
\end{aligned} \tag{63}
$$

$$
\begin{aligned}
|(\psi|k+w\mathbf{1}_{\{0\}})| &\leq \|\psi\|_2\cdot\|k+w\mathbf{1}_{\{0\}}\|_2 \\
&\leq \Lambda\|\psi\|_2(\|k\|+|w|);
\end{aligned} \tag{64}
$$

and

$$
\begin{aligned}
|(k+w\mathbf{1}_{\{0\}}|h+v\mathbf{1}_{\{0\}})| &\leq \|k+w\mathbf{1}_{\{0\}}\|_2\|h+v\mathbf{1}_{\{0\}}\|_2 \\
&\leq \Lambda(\|k\|+|w|)(\|h\|+|v|).
\end{aligned} \tag{65}
$$

Similarly, we can extend the first and second order Fréchet derivatives of $\Theta_{\delta\gamma}$ with respect to ϕ, to the space $\mathbf{C}\oplus\mathbf{B}$ and obtain similar expressions for $\overline{D_\phi\Theta_{\delta\gamma}(t,s,\psi,\phi)}(k+w\mathbf{1}_{\{0\}})$ and $\overline{D_\phi^2\Theta_{\delta\gamma}(t,s,\psi,\phi)}(h+v\mathbf{1}_{\{0\}},k+w\mathbf{1}_{\{0\}})$.

The same is also true for the bounded linear functional T_ϵ whose extension is still written as T_ϵ.

In addition, it is easy to verify that for any $\phi\in\mathbf{C}$ and $v,w\in\Re^n$ we have

$$
\begin{aligned}
(\phi|v\mathbf{1}_{\{0\}}) &= \int_{-r}^0\langle\phi(s),v\mathbf{1}_{\{0\}}(s)\rangle ds = 0, && (66)\\
(w\mathbf{1}_{\{0\}}|v\mathbf{1}_{\{0\}}) &= \int_{-r}^0\langle w\mathbf{1}_{\{0\}}(s),v\mathbf{1}_{\{0\}}(s)\rangle ds = 0, && (67)\\
H(v\mathbf{1}_{\{0\}}) &= v\mathbf{1}_{\{-r\}}, && (68)\\
(H(\psi)|H(v\mathbf{1}_{\{0\}})) &= 0, \quad (H(w\mathbf{1}_{\{0\}})|H(v\mathbf{1}_{\{0\}})) = 0. && (69)
\end{aligned}
$$

These observations will be used later.

Next we will prove several lemmas about the operator $\mathcal{S}$, which will be used later.

Lemma 5.6 *Given $\phi\in\mathbf{C}$, we have*

$$
\mathcal{S}(F)(\phi) = |\phi(0)|^2 - |\phi(-r)|^2, \tag{70}
$$

$$
\mathcal{S}(F)(\phi^0) = -|\phi(0)|^2, \tag{71}
$$

where F is the functional defined in (49) and $\mathcal{S}$ is the operator defined in (7).

Proof. Recall that

$$
\mathcal{S}(F)(\phi) = \lim_{t\downarrow 0}\frac{1}{t}\left[F(\tilde{\phi}_t) - F(\phi)\right] \tag{72}
$$

for all $\phi\in\mathbf{C}$, where $\tilde{\phi}:[-r,T]\to\Re^n$ is an extension of ϕ defined by

$$
\tilde{\phi}(t) = \begin{cases} \phi(t) & \text{if } t\in[-r,0) \\ \phi(0) & \text{if } t\geq 0, \end{cases} \tag{73}
$$

and again $\tilde{\phi}_t \in \mathbf{C}$ is defined by

$$\tilde{\phi}_t(\theta) = \tilde{\phi}(t+\theta), \quad \theta \in [-r, 0].$$

Therefore, we have

$$\begin{aligned}
\mathcal{S}(F)(\phi) &= \lim_{t\to 0+} \frac{1}{t}\left[\|\tilde{\phi}_t\|_2^2 - \|\phi\|_2^2\right] \\
&= \lim_{t\to 0+} \frac{1}{t}\left[\int_{-r}^0 |\tilde{\phi}_t(\theta)|^2 d\theta - \int_{-r}^0 |\phi(\theta)|^2 d\theta\right] \\
&= \lim_{t\to 0+} \frac{1}{t}\left[\int_{-r}^0 |\tilde{\phi}(\theta+t)|^2 d\theta - \int_{-r}^0 |\phi(\theta)|^2 d\theta\right] \\
&= \lim_{t\to 0+} \frac{1}{t}\left[\int_{-r+t}^t |\tilde{\phi}(\theta)|^2 d\theta - \int_{-r}^0 |\phi(\theta)|^2 d\theta\right] \\
&= \lim_{t\to 0+} \frac{1}{t}\left[\int_{-r+t}^0 |\tilde{\phi}(\theta)|^2 d\theta + \int_0^t |\tilde{\phi}(\theta)|^2 d\theta - \int_{-r}^0 |\phi(\theta)|^2 d\theta\right] \\
&= \lim_{t\to 0+} \frac{1}{t}\left[\int_{-r+t}^0 |\phi(\theta)|^2 d\theta + \int_0^t |\phi(0)|^2 d\theta - \int_{-r}^0 |\phi(\theta)|^2 d\theta\right] \\
&= \lim_{t\to 0+} \frac{1}{t}\left[\int_{-r+t}^0 |\phi(\theta)|^2 d\theta - \int_{-r}^0 |\phi(\theta)|^2 d\theta\right] \\
&\qquad + \lim_{t\to 0+} \frac{1}{t}\int_0^t |\phi(0)|^2 d\theta \\
&= \lim_{t\to 0+} \frac{1}{t}\int_0^t |\phi(0)|^2 d\theta - \lim_{t\to 0+} \frac{1}{t}\left[\int_{-r}^{-r+t} |\phi(\theta)|^2 d\theta\right] \\
&= |\phi(0)|^2 - |\phi(-r)|^2. \qquad (74)
\end{aligned}$$

Similarly, we have

$$\mathcal{S}(F)(\phi^0) = |\phi^0(0)|^2 - |\phi^0(-r)|^2 = -|\phi(0)|^2. \qquad \square$$

Let $\mathcal{S}_\psi$ and $\mathcal{S}_\phi$ denote the operator $\mathcal{S}$ applied to ψ and ϕ, respectively. We have the following lemma:

Lemma 5.7 *Given $\phi, \psi \in \mathbf{C}$,*

$$\mathcal{S}_\psi(F)(\phi - \psi) + \mathcal{S}_\phi(F)(\phi - \psi) = |\psi(0) - \phi(0)|^2 - |\psi(-r) - \phi(-r)|^2, \qquad (75)$$

and

$$\mathcal{S}_\psi(F)(\phi^0 - \psi^0) + \mathcal{S}_\phi(F)(\phi^0 - \psi^0) = -|\psi(0) - \phi(0)|^2, \qquad (76)$$

where F is the functional defined in (49) and $\mathcal{S}$ is the operator defined in (7).

Proof. To prove the lemma, we need the following result which can be easily proved by definition provided that $\psi \in \mathcal{D}(\tilde{\mathcal{S}})$.

$$\mathcal{S}(F)(\psi) = DF(\psi)\tilde{\mathcal{S}}(\psi), \qquad (77)$$

where $DF(\psi)$ is the Fréchet derivative of $F(\psi)$ and $\tilde{\mathcal{S}} : \mathcal{D}(\tilde{\mathcal{S}}) \subset \mathbf{C} \to \mathbf{C}$ is defined by

$$\tilde{\mathcal{S}}(\psi) = \lim_{t\downarrow 0} \frac{\tilde{\psi}_t - \psi}{t}. \tag{78}$$

We first assume that $\psi \in \mathcal{D}(\tilde{\mathcal{S}})$, the domain of the operator $\tilde{\mathcal{S}}$ that consists of those $\psi \in \mathbf{C}$ for which the above limit exists. It can be shown that

$$\mathcal{D}(\tilde{\mathcal{S}}) = \{\psi \in \mathbf{C} \mid \psi \text{ is absolutely continuous and } \dot{\psi}(0+) = 0\}.$$

In this case, we have

$$\mathcal{S}(F)(\psi) = DF(\psi)\tilde{\mathcal{S}}(\psi) = 2(\psi|\tilde{\mathcal{S}}(\psi)). \tag{79}$$

On the other hand, by virtue of Lemma 5.6, we have

$$\mathcal{S}(F)(\psi) = |\psi(0)|^2 - |\psi(-r)|^2. \tag{80}$$

Therefore, we can get

$$(\psi|\tilde{\mathcal{S}}(\psi)) = \frac{1}{2}\Big[|\psi(0)|^2 - |\psi(-r)|^2\Big]. \tag{81}$$

Since $\tilde{\mathcal{S}}$ is a linear operator, we have

$$\begin{aligned} & (\psi - \phi|\tilde{\mathcal{S}}(\psi) - \tilde{\mathcal{S}}(\phi)) \\ = \; & (\psi - \phi|\tilde{\mathcal{S}}(\psi - \phi)) \\ = \; & \frac{1}{2}\Big[|\psi(0) - \phi(0)|^2 - |\psi(-r) - \phi(-r)|^2\Big]. \end{aligned} \tag{82}$$

Given the above results, now we can get

$$\begin{aligned} & \mathcal{S}_\psi(F)(\psi - \phi) + \mathcal{S}_\phi(F)(\psi - \phi) \\ = \; & \lim_{t\downarrow 0} \frac{1}{t}\Big[\|\tilde{\psi}_t - \phi\|_2^2 - \|\psi - \phi\|_2^2 + \|\psi - \tilde{\phi}_t\|_2^2 - \|\psi - \phi\|_2^2\Big] \\ = \; & \lim_{t\downarrow 0} \frac{1}{t}\Big[\|\tilde{\psi}_t\|_2^2 - \|\psi\|_2^2 + \|\tilde{\phi}_t\|_2^2 - \|\phi\|_2^2 \\ & \qquad - 2[(\tilde{\psi}_t|\phi) - (\psi|\phi) + (\psi|\tilde{\phi}_t) - (\psi|\phi)]\Big] \\ = \; & \mathcal{S}(F)(\psi) + \mathcal{S}(F)(\phi) - 2[(\tilde{\mathcal{S}}(\psi)|\phi) + (\psi|\tilde{\mathcal{S}}(\phi))] \\ = \; & 2(\psi|\tilde{\mathcal{S}}(\psi)) + 2(\phi|\tilde{\mathcal{S}}(\phi)) - 2[(\tilde{\mathcal{S}}(\psi)|\phi) + (\psi|\tilde{\mathcal{S}}(\phi))] \\ = \; & 2(\psi - \phi|\tilde{\mathcal{S}}(\psi - \phi)) \\ = \; & [|\psi(0) - \phi(0)|^2 - |\psi(-r) - \phi(-r)|^2], \end{aligned}$$

provided that $\psi, \phi \in \mathcal{D}(\tilde{\mathcal{S}})$.

For any $\psi, \phi \in \mathbf{C}$, one can construct sequences $\{\psi_k\}_{k=1}^\infty$ and $\{\phi_k\}_{k=1}^\infty$ in $\mathcal{D}(\tilde{\mathcal{S}})$ such that

$$\lim_{k\to\infty} \|\psi_k - \psi\| = 0 \quad \text{and} \quad \lim_{k\to\infty} \|\phi_k - \phi\| = 0.$$

Consequently, by the linearity of the $\mathcal{S}$ operator and continuity of $F : \mathbf{C} \to \Re$, we have

$$\begin{aligned}
& \mathcal{S}_{\psi}(F)(\psi - \phi) + \mathcal{S}_{\phi}(F)(\psi - \phi) \\
= & \lim_{k\to\infty} \Big(\mathcal{S}_{\psi}(F)(\psi_k - \phi_k) + \mathcal{S}_{\phi}(F)(\psi_k - \phi_k) \Big) \\
= & \lim_{k\to\infty} [|\psi_k(0) - \phi_k(0)|^2 - |\psi_k(-r) - \phi_k(-r)|^2] \\
= & [|\psi(0) - \phi(0)|^2 - |\psi(-r) - \phi(-r)|^2].
\end{aligned}$$

By similar argument as we used earlier, we have

$$\begin{aligned}
& \mathcal{S}_{\psi}(F)(\psi^0 - \phi^0) + \mathcal{S}_{\phi}(F)(\psi^0 - \phi^0) \\
= & |\psi^0(0) - \phi^0(0)|^2 - |\psi^0(-r) - \phi^0(-r)|^2 \\
= & -|\psi(0) - \phi(0)|^2.
\end{aligned}$$

□

Lemma 5.8 *Given $\phi \in \mathbf{C}$, we define a new operator G as follows*

$$G(\phi) = e^{1+F(\phi)+F(\phi^0)}. \tag{83}$$

We have the following result:

$$\mathcal{S}(G)(\phi) = (-|\phi(-r)|^2)e^{1+F(\phi)+F(\phi^0)}, \tag{84}$$

where F is the functional defined in (49) and $\mathcal{S}$ is the operator defined in (7).

Proof. Recall that

$$\mathcal{S}(G)(\phi) = \lim_{t\downarrow 0} \frac{1}{t} \Big[G(\tilde{\phi}_t) - G(\phi) \Big] \tag{85}$$

for all $\phi \in \mathbf{C}$, where $\tilde{\phi} : [-r, T] \to \Re^n$ defined by (73) is an extension of ϕ. We have,

$$\begin{aligned}
& \mathcal{S}(G)(\phi) \\
= & \lim_{t\to 0+} \frac{1}{t} [e^{1+\int_{-r}^{0} |\tilde{\phi}_t(\theta)|^2 d\theta + \int_{-r}^{0} |\tilde{\phi}_t^0(\theta)|^2 d\theta} \\
& \quad - e^{1+\int_{-r}^{0} |\phi(\theta)|^2 d\theta + \int_{-r}^{0} |\phi^0(\theta)|^2 d\theta}] \\
= & \lim_{t\to 0+} \frac{1}{t} [e^{1+\int_{-r}^{0} |\tilde{\phi}(\theta+t)|^2 d\theta + \int_{-r}^{0} |\tilde{\phi}^0(\theta+t)|^2 d\theta} \\
& \quad - e^{1+\int_{-r}^{0} |\phi(\theta)|^2 d\theta + \int_{-r}^{0} |\phi^0(\theta)|^2 d\theta}] \\
= & \lim_{t\to 0+} \frac{1}{t} [e^{1+\int_{-r+t}^{t} |\tilde{\phi}(\theta)|^2 d\theta + \int_{-r+t}^{t} |\tilde{\phi}^0(\theta)|^2 d\theta} \\
& \quad - e^{1+\int_{-r}^{0} |\phi(\theta)|^2 d\theta + \int_{-r}^{0} |\phi^0(\theta)|^2 d\theta}] \\
= & \lim_{t\to 0+} \frac{1}{t} [e^{1+\int_{-r+t}^{0} |\phi(\theta)|^2 d\theta + \int_{0}^{t} |\phi(0)|^2 d\theta + \int_{-r+t}^{0} |\phi^0(\theta)|^2 d\theta + \int_{0}^{t} |\phi^0(0)|^2 d\theta} \\
& \quad - e^{1+\int_{-r}^{0} |\phi(\theta)|^2 d\theta + \int_{-r}^{0} |\phi^0(\theta)|^2 d\theta}] \\
= & \lim_{t\to 0+} \frac{1}{t} [e^{1+\int_{-r+t}^{0} |\phi(\theta)|^2 d\theta + t|\phi(0)|^2 + \int_{-r+t}^{0} |\phi^0(\theta)|^2 d\theta + t|\phi^0(0)|^2} \\
& \quad - e^{1+\int_{-r}^{0} |\phi(\theta)|^2 d\theta + \int_{-r}^{0} |\phi^0(\theta)|^2 d\theta}].
\end{aligned} \tag{86}$$

Using the L'Hospital rule on the last equality, we obtain

$$\begin{aligned}
& \mathcal{S}(G)(\phi) \\
= & \lim_{t\to 0+} e^{1+\int_{-r+t}^{0}|\phi(\theta)|^2d\theta+t|\phi(0)|^2+\int_{-r+t}^{0}|\phi^0(\theta)|^2d\theta+t|\phi^0(0)|^2}\Big(|\phi(0)|^2 \\
& -|\phi(-r+t)|^2+|\phi^0(0)|^2-|\phi^0(-r+t)|^2\Big) \\
= & (|\phi(0)|^2-|\phi(-r)|^2-|\phi^0(-r)|^2)e^{1+\int_{-r}^{0}|\phi(\theta)|^2d\theta)+\int_{-r}^{0}|\phi^0(\theta)|^2d\theta} \\
= & (|\phi(0)|^2-|\phi(-r)|^2-|\phi(0)|^2)e^{1+\int_{-r}^{0}|\phi(\theta)|^2d\theta+\int_{-r}^{0}|\phi^0(\theta)|^2d\theta} \\
= & -|\phi(-r)|^2e^{1+F(\phi)+F(\phi^0)}. \qquad (87)
\end{aligned}$$

□

Lemma 5.9 *For any $\psi, \phi \in \mathbf{C}$, we have*

$$\lim_{\epsilon\downarrow 0}|\mathcal{S}_\psi(T_\epsilon)(\psi,\phi)| = 0 \quad \textit{and} \quad \lim_{\epsilon\downarrow 0}|\mathcal{S}_\phi(T_\epsilon)(\psi,\phi)| = 0. \qquad (88)$$

Proof. We shall only prove the first equality in (88), since the second one can be proved similarly.

We first assume that $\psi \in \mathcal{D}(\tilde{\mathcal{S}})$, where the operator $\tilde{\mathcal{S}} : \mathcal{D}(\tilde{\mathcal{S}}) \subset \mathbf{C} \to \mathbf{C}$ and $\mathcal{D}(\tilde{\mathcal{S}})$ are defined in (78).

In this case,

$$\begin{aligned}
\lim_{\epsilon\downarrow 0}|\mathcal{S}_\psi(T_\epsilon)(\psi,\phi)| &= \lim_{\epsilon\downarrow 0}\left|\lim_{t\downarrow 0}\frac{T_\epsilon(\tilde{\psi}_t,\phi)-T_\epsilon(\psi,\phi)}{t}\right| \\
&= \lim_{\epsilon\downarrow 0}\left|(T_\epsilon)\lim_{t\downarrow 0}\left(\frac{\tilde{\psi}_t-\psi}{t},\phi\right)\right| \\
&\leq \lim_{\epsilon\downarrow 0}\|T_\epsilon\|\Big(\Big\|\lim_{t\downarrow 0}\frac{\tilde{\psi}_t-\psi}{t}\Big\|+\|\phi\|\Big) \\
&\leq \lim_{\epsilon\downarrow 0}\epsilon\Big(\|\tilde{\mathcal{S}}\psi\|+\|\phi\|\Big) = 0, \qquad (89)
\end{aligned}$$

because T_ϵ is a bounded linear functional on $\mathbf{C}\times\mathbf{C}$ with norm equal to ϵ.

For any $\psi, \phi \in \mathbf{C}$, one can construct a sequence of

$$\psi_k \in \mathcal{D}(\tilde{\mathcal{S}}), \quad k = 1, 2, \cdots,$$

such that

$$\lim_{k\to\infty}\|\psi_k-\psi\| = 0.$$

We have

$$\lim_{\epsilon\downarrow 0}|\mathcal{S}_\psi(T_\epsilon)(\psi_k,\phi)| = 0 \quad \forall k = 1, 2, \cdots.$$

Consequently, we can get

$$\lim_{\epsilon\downarrow 0}|\mathcal{S}_\psi(T_\epsilon)(\psi,\phi)| = 0.$$

$\square$

Given all above results, now we are ready to prove Theorem 5.1.

Proof Theorem 5.1**.** Define

$$\Gamma_1(t,\psi) \equiv V_2(s_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon}) + \Theta_{\delta\gamma}(t,s_{\delta\gamma\epsilon},\psi,\phi_{\delta\gamma\epsilon}) - T_\epsilon(\psi,\phi_{\delta\gamma\epsilon}) - M_{\delta\gamma\epsilon}, \tag{90}$$

and

$$\Gamma_2(s,\phi) \equiv V_1(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon}) - \Theta_{\delta\gamma}(t_{\delta\gamma\epsilon},s,\psi_{\delta\gamma\epsilon},\phi) + T_\epsilon(\psi_{\delta\gamma\epsilon},\phi) + M_{\delta\gamma\epsilon} \tag{91}$$

for all $s,t\in[0,T]$ and $\psi,\phi\in\mathbf{C}$.

Recall that

$$\Phi_{\delta\gamma}(t,s,\psi,\phi) = V_1(t,\psi) - V_2(s,\phi) - \Theta_{\delta\gamma}(t,s,\psi,\phi),$$

and $\Phi_{\delta\gamma}+T_\epsilon+M_{\delta\gamma\epsilon}$ reaches its maximum value zero at $(t_{\delta\gamma\epsilon},s_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon})$ in $[0,T]\times[0,T]\times\mathbf{C}\times\mathbf{C}$. By the definition of Γ_1 and Γ_2, it is easy to verify that, for all ϕ and ψ we have,

$$\Gamma_1(t,\psi)\ge V_1(t,\psi),\quad \Gamma_2(s,\phi)\le V_2(s,\phi),\quad \forall t,s\in[0,T] \text{ and } \phi,\psi\in\mathbf{C},$$

and

$$V_1(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon}) = \Gamma_1(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon}) \text{ and } V_2(s_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon}) = \Gamma_2(s_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon}).$$

Using the definitions of the viscosity subsolution of V_1 and Γ_1, we have

$$\rho\Gamma_1(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon}) - \frac{\partial\Gamma_1}{\partial t}(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon}) - \max_{u\in U}\ [\mathcal{A}^u\Gamma_1(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon}) + L(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon},u)] \le 0.$$

Using (56), (57), we can rewrite the above inequality as

$$\begin{aligned}
&\rho V_1(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon}) - \mathcal{S}(\Gamma_1)(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon}) - \frac{\partial\Gamma_1}{\partial t}(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon})\\
&- \max_{u\in U}\Bigg[\frac{1}{2}\sum_{j=1}^m \overline{D^2_\psi\Theta_{\delta\gamma}(t,s,\psi,\phi)}\Big(g(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon},u)\mathbf{e}_j\mathbf{1}_{\{0\}},\\
&\qquad\qquad g(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon},u)\mathbf{e}_j\mathbf{1}_{\{0\}}\Big)\\
&\quad + \overline{D_\psi\Theta_{\delta\gamma}(t,s,\psi,\phi)}(f(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon},u)\mathbf{1}_{\{0\}}) + L(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon},u)\\
&\quad - \overline{D_\psi T_\epsilon(\psi_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon})}(f(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon},u)\mathbf{1}_{\{0\}})\Bigg] \le 0.
\end{aligned} \tag{92}$$

By the definitions of the operator $\mathcal{A}$ and Γ_1, and using (52),(53), (54), (55), (58), (59), (60), (61), (66), (67), and (69), we can get

$$\begin{aligned}
&\mathcal{A}^u(\Gamma_1)(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon})\\
=\ & \mathcal{S}(\Gamma_1)(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon}) + \overline{D_\psi\Theta_{\delta\gamma}(\cdots)}(f(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon},u)\mathbf{1}_{\{0\}})\\
&+ \frac{1}{2}\sum_{j=1}^m \overline{D^2_\psi\Theta_{\delta\gamma}(\cdots)}\Big(g(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon},u)\mathbf{e}_j\mathbf{1}_{\{0\}}, g(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon},u)\mathbf{e}_j\mathbf{1}_{\{0\}}\Big)\\
&- \overline{D_\psi T_\epsilon(\psi_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon})}(f(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon},u)\mathbf{1}_{\{0\}})\\
=\ & \mathcal{S}(\Gamma_1)(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon}) - T_\epsilon(f(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon},u)\mathbf{1}_{\{0\}},\phi_{\delta\gamma\epsilon}).
\end{aligned}$$

Note that $\Theta_{\delta\gamma}(\cdots)$ is an abbreviation for $\Theta_{\delta\gamma}(t_{\delta\gamma\epsilon}, s_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon})$ in the above equation.

Together with (92), the above inequality yields that

$$\begin{aligned} &\rho V_1(t_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}) - \mathcal{S}(\Gamma_1)(t_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}) - \frac{\partial \Gamma_1}{\partial t}(t_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}) \\ &- \max_{u\in U}\left[L(t_{\delta\gamma}, \psi_{\delta\gamma}, u) - T_\epsilon(f(t, \psi_{\delta\gamma\epsilon}, u)\mathbf{1}_{\{0\}}, \phi_{\delta\gamma\epsilon})\right] \le 0. \end{aligned} \tag{93}$$

Similarly, using the definitions of the viscosity supersolution of V_2 and Γ_2 and by the virtue of the same techniques we used earlier to obtain (93), we have

$$\begin{aligned} &\rho V_2(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) - \mathcal{S}(\Gamma_2)(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) - \frac{\partial \Gamma_2}{\partial s}(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) \\ &- \max_{u\in U}\Bigg[\frac{1}{2}\sum_{j=1}^m \overline{D_\phi^2\Theta_{\delta\gamma}(t,s,\psi,\phi)}\Big(g(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}, u)\mathbf{e}_j\mathbf{1}_{\{0\}}, \\ &\qquad\qquad g(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}, u)\mathbf{e}_j\mathbf{1}_{\{0\}}\Big) \\ &\quad + \overline{D_\phi\Theta_{\delta\gamma}(t,s,\psi,\phi)}(f(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}, u)\mathbf{1}_{\{0\}}) + L(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}, u) \\ &\quad - \overline{D_\phi T_\epsilon(\psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon})}(f(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}, u)\mathbf{1}_{\{0\}})\Bigg] \ge 0. \end{aligned} \tag{94}$$

The above inequality implies

$$\begin{aligned} &\rho V_2(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) - \mathcal{S}(\Gamma_2)(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) - \frac{\partial \Gamma_2}{\partial s}(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) \\ &- \max_{u\in U}\left[L(s_{\delta\gamma}, \phi_{\delta\gamma}, u) - T_\epsilon(f(s, \phi_{\delta\gamma\epsilon}, u)\mathbf{1}_{\{0\}}, \phi_{\delta\gamma\epsilon})\right] \ge 0. \end{aligned} \tag{95}$$

On the other hand, by virtue of (93) and (95), we can obtain

$$\begin{aligned} & \rho[V_1(t_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}) - V_2(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon})] \\ \le\ & \mathcal{S}(\Gamma_1)(t_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}) - \mathcal{S}(\Gamma_2)(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) + 4(t_{\delta\gamma\epsilon} - s_{\delta\gamma\epsilon}) \\ & + \max_{u\in U}\Big[(L(t_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}, u) - T_\epsilon(f(t_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}, u)\mathbf{1}_{\{0\}}, \phi_{\delta\gamma\epsilon})) \\ & - (L(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}, u) + T_\epsilon(\psi_{\delta\gamma\epsilon}, f(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon})\mathbf{1}_{\{0\}}))\Big]. \end{aligned} \tag{96}$$

From the definition (see (7)) of $\mathcal{S}$, it is clear that $\mathcal{S}$ is linear and it takes 0 on constants. Recall that

$$\Gamma_1(t,\psi) = V_2(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) + \Theta_{\delta\gamma}(t, s_{\delta\gamma\epsilon}, \psi, \phi_{\delta\gamma\epsilon}) - T_\epsilon(\psi, \phi_{\delta\gamma\epsilon}) - M_{\delta\gamma\epsilon},$$

and

$$\Gamma_2(s,\phi) = V_1(t_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}) - \Theta_{\delta\gamma}(t_{\delta\gamma\epsilon}, s, \psi_{\delta\gamma\epsilon}, \phi) + T_\epsilon(\psi_{\delta\gamma\epsilon}, \phi) + M_{\delta\gamma\epsilon}.$$

Thus we have,

$$\mathcal{S}(\Gamma_1)(t_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}) = \mathcal{S}_\psi(\Theta_{\delta\gamma})(t_{\delta\gamma\epsilon}, s_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) - \mathcal{S}_\psi(T_\epsilon)(\psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}),$$

and

$$\mathcal{S}(\Gamma_2)(s_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon}) = -\mathcal{S}_\phi(\Theta_{\delta\gamma})(t_{\delta\gamma\epsilon},s_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon}) + \mathcal{S}_\phi(T_\epsilon)(\psi_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon}).$$

Therefore,

$$\begin{aligned} & \mathcal{S}(\Gamma_1)(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon}) - \mathcal{S}(\Gamma_2)(s_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon}) \\ = \ & \mathcal{S}_\psi(\Theta_{\delta\gamma})(t_{\delta\gamma\epsilon},s_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon}) + \mathcal{S}_\phi(\Theta_{\delta\gamma})(t_{\delta\gamma\epsilon},s_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon}) \\ & -[\mathcal{S}_\psi(T_\epsilon)(\psi_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon}) + \mathcal{S}_\phi(T_\epsilon)(\psi_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon})]. \end{aligned} \tag{97}$$

Recall that

$$\begin{aligned} & \Theta_{\delta\gamma}(t,s,\psi,\phi) \\ = \ & \frac{1}{\delta}[F(\psi-\phi)+F(\psi^0-\phi^0)+|t-s|^2] + \gamma(G(\psi)+G(\phi)). \end{aligned}$$

Therefore, we can get

$$\begin{aligned} & (\mathcal{S}_\psi(\Theta_{\delta\gamma}) + \mathcal{S}_\phi(\Theta_{\delta\gamma}))(t_{\delta\gamma\epsilon},s_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon}) \\ \equiv \ & \mathcal{S}_\psi(\Theta_{\delta\gamma})(t_{\delta\gamma\epsilon},s_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon}) \\ & + \mathcal{S}_\phi(\Theta_{\delta\gamma})(t_{\delta\gamma\epsilon},s_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon}) \\ = \ & \frac{1}{\delta}[\mathcal{S}_\psi(F)(\psi_{\delta\gamma\epsilon}-\phi_{\delta\gamma\epsilon}) + \mathcal{S}_\phi(F)(\psi_{\delta\gamma\epsilon}-\phi_{\delta\gamma\epsilon}) \\ & + \mathcal{S}_\psi(F)(\psi^0_{\delta\gamma\epsilon}-\phi^0_{\delta\gamma\epsilon}) + \mathcal{S}_\phi(F)(\psi^0_{\delta\gamma\epsilon}-\phi^0_{\delta\gamma\epsilon})] \\ & + \gamma[\mathcal{S}_\psi(G)(\psi_{\delta\gamma\epsilon}) + \mathcal{S}_\phi(G)(\phi_{\delta\gamma\epsilon})]. \end{aligned} \tag{98}$$

Using Lemma 5.7 and Lemma 5.8, we deduce that

$$\begin{aligned} & (\mathcal{S}_\psi(\Theta_{\delta\gamma}) + \mathcal{S}_\phi(\Theta_{\delta\gamma}))(t_{\delta\gamma\epsilon},s_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon}) \\ = \ & \frac{1}{\delta}[-|\psi_{\delta\gamma\epsilon}(-r)-\phi_{\delta\gamma\epsilon}(-r)|^2] \\ & -\gamma\bigg(|\psi_{\delta\gamma\epsilon}(-r)|^2 e^{1+F(\psi_{\delta\gamma\epsilon})+F(\psi^0_{\delta\gamma\epsilon})} \\ & + |\phi_{\delta\gamma\epsilon}(-r)|^2 e^{1+F(\phi_{\delta\gamma\epsilon})+F(\phi^0_{\delta\gamma\epsilon})}\bigg) \\ \leq \ & 0. \end{aligned} \tag{99}$$

Thus, by virtue of (97) and Lemma 5.9, we have

$$\limsup_{\delta\downarrow 0,\epsilon\downarrow 0}\Big[\mathcal{S}(\Gamma_1)(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon}) - \mathcal{S}(\Gamma_2)(s_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon})\Big] \leq 0. \tag{100}$$

Moreover, we know that the norm of T_ϵ is less than ϵ. Thus for any $\gamma>0$ using (97) and taking the $\limsup$ on both sides of (96) as δ and ϵ go to 0, we obtain

$$\limsup_{\epsilon\downarrow 0,\delta\downarrow 0}\rho(V_1(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon}) - V_2(s_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon}))$$

$$
\begin{aligned}
\leq \;& \limsup_{\epsilon\downarrow 0,\delta\downarrow 0}\Big\{\mathcal{S}(\Gamma_1)(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon}) - \mathcal{S}(\Gamma_2)(s_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon}) \\
& + [L(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon},u) - T_\epsilon(f(t,\psi_{\delta\gamma\epsilon},u)\mathbf{1}_{\{0\}},\phi_{\delta\gamma\epsilon})] \\
& - [L(s_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon},u) + T_\epsilon(\psi_{\delta\gamma\epsilon}, f(t,\phi_{\delta\gamma\epsilon},u)\mathbf{1}_{\{0\}})]\Big\} \\
\leq \;& \limsup_{\epsilon\downarrow 0,\delta\downarrow 0}\Big\{\Big|[L(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon},u) - L(s_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon},u)]\Big|\Big\}. \qquad (101)
\end{aligned}
$$

Using the Lipschitz continuity of L and Lemma 5.2, we see that

$$
\begin{aligned}
& \limsup_{\delta\downarrow 0,\epsilon\downarrow 0}|L(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon},u) - L(s_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon},u)| \\
\leq \;& \limsup_{\delta\downarrow 0,\epsilon\downarrow 0} C\Big(|t_{\delta\gamma\epsilon} - s_{\delta\gamma\epsilon}| + \|\psi_{\delta\gamma\epsilon} - \phi_{\delta\gamma\epsilon}\|_2\Big) = 0. \qquad (102)
\end{aligned}
$$

Moreover, by virtue of (102), we get

$$
\limsup_{\epsilon\downarrow 0,\delta\downarrow 0}\rho(V_1(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon}) - V_2(s_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon})) \leq 0. \qquad (103)
$$

Since $(t_{\delta\gamma\epsilon}, s_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon})$ is the maximum of $\Phi_{\delta\gamma} + T_\epsilon$ in $[0,T]\times[0,T]\times\mathbf{C}\times\mathbf{C}$, for all $(t,\psi)\in[0,T]\times\mathbf{C}$, we have

$$
\Phi_{\delta\gamma}(t,t,\psi,\psi) + T_\epsilon(\psi,\psi) \leq \Phi_{\delta\gamma}(t_{\delta\gamma\epsilon}, s_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) + T_\epsilon(\psi_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon}).
$$

Then we can get

$$
\begin{aligned}
& V_1(t,\psi) - V_2(t,\psi) \\
\leq \;& V_1(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon}) - V_2(s_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon}) \\
& - \frac{1}{\delta}\Big[\|\psi_{\delta\gamma\epsilon} - \phi_{\delta\gamma\epsilon}\|_2^2 + \|\psi^0_{\delta\gamma\epsilon} - \phi^0_{\delta\gamma\epsilon}\|_2^2 + |t_{\delta\gamma\epsilon} - s_{\delta\gamma\epsilon}|^2\Big] \\
& + 2\gamma\exp(1 + \|\psi\|_2^2 + \|\psi^0\|_2^2) \\
& - \gamma(\exp(1 + \|\psi_{\delta\gamma\epsilon}\|_2^2 + \|\psi^0_{\delta\gamma\epsilon}\|_2^2) + \exp(1 + \|\phi_{\delta\gamma\epsilon}\|_2^2 + \|\phi^0_{\delta\gamma\epsilon}\|_2^2)) \\
& + T_\epsilon(\psi_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon}) - T_\epsilon(\psi,\psi) \\
\leq \;& V_1(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon}) - V_2(s_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon}) \\
& + 2\gamma\exp(1 + \|\psi\|_2^2 + \|\psi^0\|_2^2) + T_\epsilon(\psi_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon}) - T_\epsilon(\psi,\psi), \qquad (104)
\end{aligned}
$$

where the last inequality comes from the fact that $\delta > 0$ and $\gamma > 0$. By virtue of (103), when we take the $\limsup$ on (104) as δ, ϵ and γ go to zero, we can obtain

$$
\begin{aligned}
& V_1(t,\psi) - V_2(t,\psi) \\
\leq \;& \limsup_{\gamma\downarrow 0,\epsilon\downarrow 0,\delta\downarrow 0}\Big(V_1(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon}) - V_2(s_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon}) \\
& + 2\gamma\exp(1 + \|\psi\|_2^2 + \|\psi^0\|_2^2) + T_\epsilon(\psi_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon}) - T_\epsilon(\psi,\psi)\Big) \\
\leq \;& 0. \qquad (105)
\end{aligned}
$$

Therefore, we have

$$
V_1(t,\psi) \leq V_2(t,\psi), \quad \forall (t,\psi)\in[0,T]\times\mathbf{C} \qquad (106)
$$

This completes the proof of Theorem 5.1. □

Conclusions

This paper investigates an optimal control problem for a general system of stochastic functional differential equations with a bounded memory. An infinite-dimensional HJB equation is derived using a Bellman-type dynamic programming principle. It is shown that the value function is the unique viscosity solution of the HJB equation. Computational issues related to the HJB equation is treated in the authors' another paper [3].

References

[1] Adams, R.A. (1975), *Sobolev Spaces*, Academic Press.

[2] Bourgain, J.(1979), *La propriete de Radon-Nikodym*, Cours polycopie, **36**, universite P.et M. Curie, Paris.

[3] Chang, M.-H., Pang, T. and Pemy, M. (2006), "Finite Difference Approximations for Stochastic Control Systems with Delay", Preprint.

[4] Crandall, M. G., Ishii, H., and P.L. Lions, P.L. (1992), "User's guide to viscosity solutions of second order partial differential equations", *Bull. Amer. Math. Soc.*, **27**, 1-67.

[5] Ekeland, I. and Lebourg, G. (1976), "Generic differentiability and pertubed optimization in Banach spaces", *Trans. Amer. Math. Soc.*, **224**, 193-216.

[6] Elsanousi, I. (2000), *Stochastic control for system with memory*, Dr. Scient. thesis, University of Oslo.

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

[8] Fleming, W.H. and Rishel, R.W. (1975), *Deterministic and Stochastic Optimal Control*, Springer-Verlag.

[9] Fleming, W.H. and Mete Soner H. (1993), *Control Markov Processes and Viscosity Solutions*, Springer Verlag, New York.

[10] Hale, J. (1977), *Theory of Functional Differential Equations*, Springer-Verlag, New York, Heidelberg, Berlin.

[11] Kolmanovskii, V. B. and Maizenberg, T. L. (1973), Optimal control of stochastic systems with aftereffect, *Avtomat. i Telemeh*, **1**, pp. 47-61.

[12] Kolmanovskii, B. V. and L. E. Shaikhet, L. E. (1996), *Control of Systems with Aftereffect*, Translations of Mathematical Monographs Vol. 157, American Mathematical Society.

[13] Larssen, B., (2002), *Dynamic Programming in Stochastic Control of Systems with Delay*, Dr. Scient. thesis, University of Oslo.

[14] Lions, P. L. (1982), *Generalized Solutions of Hamilton-Jacobi Equations, Research Notes in Mathematics, Research Note No. 69, Pitman Publishing, Boston, MA.*

[15] Lions, P. L. (1989), "Viscosity solutions of fully nonlinear second-order equations and optimal stochastic control in infinite dimensions. III. Uniqueness of Viscosity solutions for general second-order equations", *Journal of Functional anlysis*, **62**, 379-396.

[16] Lions, P. L. (1992), "Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations. Part 1: The dynamic programming principle and applications and part 2:Viscosity solutions and uniqueness", *Comm. P.D.E.* **8** 1101-1174 and 1229-1276.

[17] Mohammed, S-E A. (1984), *Stochastic Functional Differential Equations*, Research Notes in Mathematics No. 99, Pitman Publishing, Boston-London-Melbourne.

[18] Mohammed, S-E A. (1998), "Stochastic Differential Equations with memory- theory, examples and applications", *Stochastics Analysis and Related Topics 6.* The Geido workshop, 1996, Progress in probability, Birkhauser, 1998.

[19] Øksendal, B. and Sulem, A. (2000), "A maximum principle for optimal control of stochastic systems with delay, with applications to finance", in *J. M. Mendaldi, E. Rofman and A. Sulem (editors): Optimal Control and Partial Differential Equations-Innovations and Applications*, IOS Press, 1-16.

[20] Siddiqi, A. H. (2004), *Applied Functional Analysis*, Marcel Dekker, New York.

[21] Stegall, C. (1978), "Optimization of functions on certains subsets of Banach spaces", *Math. Ann.*, **236**, 171-176.