3

Optimal Classical Control

Let $0 < T < \infty$ be the given and fixed terminal time and let $t \in [0, T]$ be a variable initial time for the optimal classical control problem treated in this chapter. The main theme of this chapter is to consider the infinite-dimensional optimal classical control problem over a finite time horizon $[t, T]$. The dynamics of the process $x(\cdot) = \{x(s), s \in [t-r, T]\}$ being controlled are governed by a stochastic hereditary differential equation (SHDE) with a bounded memory of duration $0 < r < \infty$ and are taking values in the Banach space $\mathbf{C} = C([-r, 0]; \Re^n)$. The formulation of the control problem is given in Section 3.1. The value function $V : [0, T] \times \mathbf{C} \to \Re$ of the optimal classical control problem is written as a function of the initial datum $(t, \psi) \in [0, T] \times \mathbf{C}$. The existence of optimal control is proved in Section 3.2. In there, we consider an optimizing sequence of stochastic relaxed control problems with its corresponding sequence of value functions that converges to the value function of our original optimal control problem. Since the regular optimal control is a special case of optimal relaxed control, the existence of optimal control is therefore established. The Bellman-type dynamic programming principle (DPP) originally due to Larssen [Lar02] is derived and proved in Section 3.3. Based on the DDP, an infinite-dimensional Hamilton-Jacobi-Bellman equation (HJBE) is heuristically derived in Section 3.4 for the value function under the condition that it is sufficiently smooth. This HJBE involves a first- and second-order Fréchet derivatives with respect to spatial variable $\psi \in \mathbf{C}$ as well as an $\mathcal{S}$-operator that is unique only to SHDE. However, it is known in most optimal control problems, deterministic or stochastic, that the value functions, although can be proven to be continuous, do not meet these smoothness conditions and, therefore, cannot be a solution to the HBJE in the classical sense. To overcome this difficulty, the concept of viscosity solution to the infinite-dimensional HJBE is introduced in Section 3.5. Section 3.6 concerns the comparison principle between a super-viscosity solution and a sub-viscosity solution. Based on this comparison principle, it is shown that the value function is the unique viscosity solution to the HJBE. Due to the lack of smoothness of the value function, a classical verification theorem will not be useful in characterizing the optimal control. A generalized verification theorem in the framework of a viscosity solution is stated without a proof in Section 3.7. In Section 3.8, we prove that, under some special conditions on the controlled SHDE and the value function, the HJBE can take a finite-dimensional form in which only regular partial derivatives but not Fréchet derivatives are involved. Two application examples in this special form are also illustrated in this section.

We give the following example as a motivation for studying optimal classical control problems. Two other completely worked-out examples are given in Subsection 3.8.3.

Example. (Optimal Advertising Problem) (see Gossi and Marinelli [GM04] and Gossi et al. [GMS06])

Let $y(\cdot) = \{y(s), s \in [0,T]\}$ denote the stock of advertising goodwill of the product to be launched. The process $y(\cdot)$ is described by the following one-dimensional controlled stochastic hereditary differential equation:

$$dy(s) = \Big[a_0 y(s) + \int_{-r}^{0} a_1(\theta) y(s+\theta)\, d\theta + b_0 u(s) + \int_{-r}^{0} b_1(\theta) u(s+\theta)\, d\theta\Big] ds + \sigma dW(s), \quad s \in [0,T],$$

with the initial conditions $y_0 = \psi \in C[-r,0]$ and $u_0 = \phi \in L^2([-r,0])$ at initial time $t = 0$.

In the above $(\Omega, \mathcal{F}, P, \mathbf{F}, W(\cdot))$ denotes an one-dimensional standard Brownian motion and the control process $u(\cdot) = \{u(s), s \in [0,T]\}$ denotes the advertising expenditures as a process in $L^2([0,T], \Re_+; \mathbf{F})$, the space of square integrable non-negative processes adapted to $\mathbf{F}$. Moreover, it is assumed that the following conditions are satisfied:

(i) $a_0 \le 0$ denotes a constant factor of image deterioration of the product in absence of advertising.

(ii) $a_1(\cdot) \in L^2([-r,0], \Re)$ is the distribution of the forgetting time.

(iii) $b_0 \ge 0$ denotes the effective constant of instantaneous advertising effect.

(iv) $b_1(\cdot) \in L^2([-r,0], \Re_+)$ is the density function of the time lag between the advertising expenditure $u(\cdot)$ and the corresponding effect on the goodwill level.

(v) $\psi(\cdot)$ and $\phi(\cdot)$ are non-negative and represent, respectively, the histories of goodwill level and the advertising expenditure before time zero.

The objective of this optimal advertising problem is to seek an advertising strategy $u(\cdot)$ that maximizes the objective functional

$$J(\psi, \phi; u(\cdot)) = E\left[\Psi(y(T)) - \int_0^T L(u(s))\, ds\right],$$

where $\Psi : [0,\infty) \to [0,\infty)$ is a concave utility function with polynomial growth at infinity, $L : [0,\infty) \to [0,\infty)$ is a convex cost function which is superlinear at infinity, that is,

$$\lim_{u\to\infty}\frac{L(u)}{u}=\infty.$$

The above objective functional accounts for the balance between an utility of terminal goodwill $\Psi(y(T))$ and overall functional of advertising expenditures $\int_0^T L(u(s))\,ds$ over the period. Note that this model example involves the histories of both the state and control processes. The general theory for optimal control of stochastic systems with delays in both state and control processes has yet to be developed. If $b_1(\cdot)=0$, then there is no aftereffect of previous advertising expenditures on the goodwill level. In this case, it is a special case of what to be developed in this chapter.

3.1 Problem Formulation

3.1.1 The Controlled SHDE

In the following, let $(\Omega,\mathcal{F},P,\mathbf{F},W(\cdot))$ be a certain m-dimensional Brownian stochastic basis.

Consider the following controlled SHDE with a bounded memory (or delay) of duration $0<r<\infty$:

$$dx(s)=f(s,x_s,u(s))\,ds+g(s,x_s,u(s))\,dW(s),\quad s\in[t,T], \tag{3.1}$$

with the given initial data $(t,x_t)=(t,\psi)\in[0,T]\times\mathbf{C}$ and defined on a certain m-dimensional Brownian stochastic basis $(\Omega,\mathcal{F},P,\mathbf{F},W(\cdot))$ that is yet to be determined.

In (3.1), the following is understood:
(i) The drift $f:[0,T]\times\mathbf{C}\times U\to\Re^n$ and the diffusion coefficient $g:[0,T]\times\mathbf{C}\times U\to\Re^{n\times m}$ are deterministic continuous functions.
(ii) U, the control set, is a complete metric space and is typically a subset of an Euclidean space.
(iii) $u(\cdot)=\{u(s),s\in[t,T]\}$ is a U-valued $\mathbf{F}$-progressively measurable process that satisfies the following conditions:

$$E\left[\int_t^T|u(s)|^2\,ds\right]<\infty. \tag{3.2}$$

Note that the control process $u(\cdot)=\{u(s),s\in[t,T]\}$ defined on

$$(\Omega,\mathcal{F},P,\mathbf{F},W(\cdot))$$

is said to be progressively measurable if $u(\cdot)=\{u(s),t\le s\le T\}$ in U is $\mathbf{F}$-adapted (i.e., $u(s)$ is $\mathcal{F}(s)$-measurable for every $s\in[t,T]$), and for each $a\in[t,T]$ and $A\in\mathcal{B}(U)$, the set $\{(s,\omega)\mid t\le s\le a,\omega\in\Omega,u(s,\omega)\in A\}$ belongs to the product σ-field $\mathcal{B}([t,a])\otimes\mathcal{F}(a)$; that is, if the mapping

$$(s,\omega) \mapsto u(s,\omega) : ([t,a] \times \Omega, \mathcal{B}([t,a]) \otimes \mathcal{F}(a)) \to (U, \mathcal{B}(U))$$

is measurable, for each $a \in [t,T]$.

Let $L : [0,T] \times \mathbf{C} \times U \to \Re$ and $\Psi : \mathbf{C} \to \Re$ be two deterministic continuous functions that represent the instantaneous and terminal reward functions, respectively, for the optimal classical control problem.

Assumption 3.1.1 *The assumptions on the functions f, g, L, and Ψ are stated as follows:*
(A3.1.1) (Lipschitz Continuity) *The maps $f(t,\phi,u)$, $g(t,\phi,u)$, $L(t,\phi,u)$, and $\Psi(\phi)$ are Lipschitz on $[0,T] \times \mathbf{C} \times U$ and Hölder continuous in $t \in [0,T]$: There is a constant $K_{lip} > 0$ such that*

$$\begin{aligned} &|f(t,\phi,u) - f(s,\varphi,v)| + |g(t,\phi,u) - g(s,\varphi,v)| \\ &+ |L(t,\phi,u) - L(s,\varphi,v)| + |\Psi(\phi) - \Psi(\varphi)| \\ &\le K_{lip}(\sqrt{|t-s|} + \|\phi - \varphi\| + |u - v|), \end{aligned}$$

$$\forall s,t \in [0,T], u,v \in U, \text{ and } \phi,\varphi \in \mathbf{C}.$$

(A3.1.2) (Linear Growth) *There exists a constant $K_{grow} > 0$ such that*

$$|f(t,\phi,u)| + |g(t,\phi,u)| \le K_{grow}(1 + \|\phi\|)$$

and

$$|L(t,\phi,u)| + |\Psi(\phi)| \le K_{grow}(1 + \|\phi\|_2)^k, \quad \forall (t,\phi) \in [0,T] \times \mathbf{C},\ u \in U.$$

(A3.1.3) The initial function ψ belongs to the space $L^2(\Omega, \mathbf{C}; \mathcal{F}(t))$ of $\mathcal{F}(t)$-measurable elements in $L^2(\Omega, \mathbf{C})$ such that

$$\|\psi\|^2_{L^2(\Omega;\mathbf{C})} \equiv E[\|\psi\|^2] < \infty.$$

Condition (A3.1.2) in Assumption 3.1.1 stipulates that both L and Ψ satisfy a polynomial growth in $\phi \in \mathbf{C}$ under the norm L^2-norm $\|\cdot\|_2$ instead of the sup-norm $\|\cdot\|$. This stronger requirement is needed in order to show that the uniqueness of the viscosity solution of the HJBE in Section 3.6.

The solution process of the controlled SHDE (3.1) is given next.

Definition 3.1.2 *Given an m-dimensional Brownian stochastic basis $(\Omega, \mathcal{F}, P, \mathbf{F}, W(\cdot))$ and the control process $u(\cdot) = \{u(s), s \in [t,T]\}$, a process $x(\cdot; t, \psi_t, u(\cdot)) = \{x(s; t, \psi_t, u(\cdot)), s \in [t-r, T]\}$ is said to be a (strong) solution of the controlled SHDE (3.1) on the interval $[t-r, T]$ and through the initial datum $(t,\psi) \in [0,T] \times L^2(\Omega, \mathbf{C}; \mathcal{F}(t))$ if it satisfies the following conditions:*

1. *$x_t(\cdot; t, \psi_t, u(\cdot)) = \psi_t(\cdot)$, P-a.s.*
2. *$x(s; t, \psi_t, u(\cdot))$ is $\mathcal{F}(s)$-measurable for each $s \in [t,T]$;*

3. *For* $1 \leq i \leq n$ *and* $1 \leq j \leq m$*, we have*

$$P\Big[\int_t^T \Big(|f_i(s, x_s, u(s))| + g_{ij}^2(s, x_s, u(s))\Big) ds < \infty\Big] = 1.$$

4. *The process* $\{x(s; t, \psi_t, u(\cdot)), s \in [t, T]\}$ *is continuous and satisfies the following stochastic integral equation* P*-a.s.:*

$$x(s) = \psi_t(0) + \int_t^s f(\lambda, x_\lambda, u(\lambda))\, d\lambda + \int_t^s g(\lambda, x_\lambda, u(\lambda))\, dW(\lambda).$$

In addition, the solution process $\{x(s; t, \psi_t, u(\cdot)), s \in [t - r, T]\}$ is said to be (strongly) unique if $\{y(s; t, \psi, u(\cdot)), s \in [t - r, T]\}$ is also a solution of (3.1) on $[t - r, T]$ with the control process $u(\cdot)$ and through the same initial datum (t, ψ_t); then

$$P\{x(s; t, \psi, u(\cdot)) = y(s; t, \psi, u(\cdot)),\ \forall s \in [t, T]\} = 1.$$

Theorem 3.1.3 *Let* $(\Omega, \mathcal{F}, P, \mathbf{F}, W(\cdot))$ *be an* m*-dimensional Brownian motion and let* $u(\cdot) = \{u(s), s \in [t, T]\}$ *be a control process. Then for each initial datum* $t, \psi_t) \in [0, T] \times L^2(\Omega, \mathbf{C}; \mathcal{F}(t))$*, the controlled SHDE (3.1) has a unique strong solution process*

$$x(\cdot; t, \psi_t, u(\cdot)) = \{x(s; t, \psi_t, u(\cdot)), \quad s \in [t, T]\}$$

under Assumption 3.1.1. The following holds:

1. *The map* $(s, \omega) \mapsto x(s; t, \psi_t, u(\cdot))$ *belongs to the space* $L^2(\Omega, C([t - r, T]; \Re^n)); \mathcal{F}(s))$*, and the map* $(t, \omega) \mapsto x_s(\cdot; t, \psi_t, u(\cdot))$ *belongs to the space* $L^2(\Omega, \mathbf{C}; \mathcal{F}(s))$*. Moreover, there exists constants* $K_b > 0$ *and* $k \geq 1$ *such that*

$$E[\|x_s(\cdot; t, \psi_t, u(\cdot))\|^2] \leq K_b(1 + E[\|\psi_t\|^2])^k, \quad \forall s \in [t, T] \text{ and } u(\cdot) \in \mathcal{U}[t, T]. \tag{3.3}$$

2. *The map* $\psi_t \mapsto x_s(\cdot; t, \psi_t, u(\cdot))$ *is Lipschitz; that is, there is a constant* $K > 0$ *such that for all* $s \in [t, T]$ *and* $\psi_t^{(1)}, \psi_t^{(2)} \in L^2(\Omega, \mathbf{C}; \mathcal{F}(t))$*,*

$$E[\|x_s(\cdot; t, \psi_t^{(1)}, u(\cdot)) - x_s(\cdot; t, \psi_t^{(2)}, u(\cdot))\|] \leq KE[\|\psi_t^{(1)} - \psi_t^{(2)}\|]. \tag{3.4}$$

Proof. Let the random functions $\tilde{f} : [0, T] \times \mathbf{C} \times \Omega \to \Re^n$ and $\tilde{g} : [0, T] \times \mathbf{C} \times \Omega \to \Re^{n \times m}$ be defined as follows:

$$\tilde{f}(s, \phi, \omega) = f(s, \phi, u(s, \omega))$$

and

$$\tilde{g}(s, \phi, \omega) = g(s, \phi, u(s, \omega)),$$

where $u(\cdot) = \{u(s), s \in [t, T]\}$ is the control process. If the functions f and g satisfy Assumption 3.1.1, then the random functions $\tilde{f}$ and $\tilde{g}$ defined above satisfy Assumptions 1.3.6 and 1.3.7 of Chapter 1 and, therefore, the controlled system (3.1) has a unique *strong* solution process on $[t-r, T]$ and through the initial datum $(t, \psi) \in [0, T] \times L^2(\Omega, \mathbf{C}; \mathcal{F}(t))$, which is denoted by

$$x(\cdot; t, \psi_t, u(\cdot)) = \{x(s; t, \psi_t, u(\cdot)), s \in [t, T]\}$$

(or simply $x(\cdot)$ when there is no danger of ambiguity) according to Theorem 1.3.12 of Chapter 1. □

Remark 3.1.4 *It is clear from the appearance of (3.1) that the use of the term "classical control" (as opposed to "impulse control" in Chapter 7) becomes self-explanatory. This is due to the fact that an application of the control $u(s)$ at time $s \in [t, T]$ will only change the rate of the drift and the diffusion coefficient and, therefore, the pathwise continuity of controlled state process $x(\cdot) = \{x(s), s \in [t-r, T]\}$ will not be affected by this action.*

Definition 3.1.5 *The (classical) control process $u(\cdot)$ is an Markov (or feedback) control if there exist a Borel measurable function $\eta : [0, T] \times \mathbf{C} \to U$ such that*

$$u(s) = \eta(s, x_s),$$

where $\{x_s, s \in [t, T]\}$ is the $\mathbf{C}$-valued Markov process corresponding to the solution process $\{x(s), s \in [t-r, T]\}$ of the following feedback equation:

$$dx(s) = f(s, x_s, \eta(s, x_s))\, ds + g(s, x_s, \eta(s, x_s))\, dW(s) \tag{3.5}$$

with the initial data $(t, x_t) = (t, \psi) \in [0, T] \times \mathbf{C}$.

3.1.2 Admissible Controls

Definition 3.1.6 *For each $t \in [0, T]$, a six-tuple $(\Omega, \mathcal{F}, P, \mathbf{F}, W(\cdot), u(\cdot))$ is said to be an admissible control if it satisfies the following conditions:*

1. *$(\Omega, \mathcal{F}, P, \mathbf{F}, W(\cdot))$ is a certain m-dimensional Brownian stochastic basis;*
2. *$u : [t, T] \times \Omega \to U$ is an $\mathbf{F}$-adapted and is right-continuous at the initial time t; that is, $\lim_{s \downarrow t} u(s) = u(t)$ (say $= u \in U$).*
3. *Under the control process $u(\cdot) = \{u(s), s \in [t, T]\}$, (3.1) admits a unique strong solution $x(\cdot; t, \psi, u(\cdot)) = \{x(s; t, \psi, u(\cdot)), s \in [t, T]\}$ on $(\Omega, \mathcal{F}, P, \mathbf{F}, W(\cdot))$ and through each initial datum $(t, \psi) \in [0, T] \times \mathbf{C}$.*
4. *The $\mathbf{C}$-valued segment process $\{x_s(t, \psi, u(\cdot)), s \in [t, T]\}$ defined by*

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

   is a strong Markov process with respect to the Brownian stochastic basis $(\Omega, \mathcal{F}, P, \mathbf{F}, W(\cdot))$.
5. *The control process $u(\cdot)$ is such that*

   $$E\Big[\int_t^T |L(s, x_s(\cdot; t, \psi, u(\cdot)), u(s))|\, ds + |\Psi(x_T(\cdot; t, \psi, u(\cdot)))|\Big] < \infty,$$

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

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

We will write $u(\cdot) \in \mathcal{U}[t,T]$ or formally the 6-tuple $(\Omega, \mathcal{F}, P, \mathbf{F}, W(\cdot), u(\cdot)) \in \mathcal{U}[t,T]$ interchangeably, whenever there is no danger of ambiguity.

Remark 3.1.7 *Definition 3.1.6 defines a weak formulation of an admissible control in that the Brownian stochastic basis* $(\Omega, \mathcal{F}, P, \mathbf{F}, W(\cdot))$ *is not predetermined and in fact is a part of the ingredients that constitute an admissible control. This is contrary to the strong formulation of an admissible control in which the Brownian stochastic basis* $(\Omega, \mathcal{F}, P, \mathbf{F}, W(\cdot))$ *is predetermined and given.*

Remark 3.1.8 *To avoid using the yet-to-be-developed Itô formula for the the* $\mathbf{C}$*-valued process* $\{x_s(t,\psi,u(\cdot)), s \in [t,T]\}$ *in the development of the HJB theory, we make additional requirement in Condition 3 of Definition 3.1.6 that it is a strong Markov process. This requirement is not a stringent one. In fact, the class of admissible controls* $\mathcal{U}[t,T]$ *defined in Definition 3.1.6 includes all Markov (or feedback) control (see Definition 3.1.5), where* $\eta : [0,T] \times \mathbf{C} \to U$ *is Lipschitz with respect to the segment variable; that is, there exists a constant* $K > 0$ *such that*

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

Throughout, we assume that the functions $f : [0,T] \times \mathbf{C} \times U \to \Re^n$, $g : [0,T] \times \mathbf{C} \times U \to \Re^{n\times m}$, $L : [0,T] \times \mathbf{C} \times U \to \Re$, and $\Psi : \mathbf{C} \to \Re$ satisfy Assumption 3.1.1.

Given an admissible control $u(\cdot) \in \mathcal{U}[t,T]$, let

$$x(\cdot; t, \psi, u(\cdot)) = \{x(s; t, \psi, u(\cdot)), s \in [t-r, T]\}$$

be the solution of (3.1) through the initial datum $(t,\psi) \in [0,T] \times \mathbf{C}$. We again consider the corresponding $\mathbf{C}$-valued segment process $\{x_s(\cdot; t, \psi, u(\cdot)), s \in [t,T]\}$. For notational simplicity, we often write $x(s) = x(s; t, \psi, u(\cdot))$ and $x_s = x_s(\cdot; t, \psi, u(\cdot))$ for $s \in [t,T]$ whenever there is no danger of ambiguity.

3.1.3 Statement of the Problem

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

$$\begin{aligned} J(t,\psi; u(\cdot)) \equiv E\bigg[\int_t^T & e^{-\alpha(s-t)} L(s, x_s(\cdot; t, \psi, u(\cdot)), u(s))\, ds \\ & + e^{-\alpha(T-t)} \Psi(x_T(\cdot; t, \psi, u(\cdot))) \bigg], \end{aligned} \tag{3.6}$$

where $\alpha \geq 0$ 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 functional $J(t, \psi; u(\cdot))$ defined above. In this case, the value function $V : [0, T] \times \mathbf{C} \to \Re$ is defined to be

$$V(t, \psi) = \sup_{u(\cdot) \in \mathcal{U}[t,T]} J(t, \psi; u(\cdot)). \tag{3.7}$$

The control process $u^*(\cdot) = \{u^*(s), s \in [t, T]\} \in \mathcal{U}[t, T]$ is said to be an optimal control for the optimal classical control problem if

$$V(t, \psi) = J(t, \psi; u^*(\cdot)). \tag{3.8}$$

The (strong) solution process

$$x^*(\cdot; t, \psi, u^*(\cdot)) = \{x^*(s; t, \psi, u^*(\cdot)), s \in [t - r, T]\}$$

of (3.1) corresponding to the optimal control $u^*(\cdot))$ will be called the optimal state process corresponding to $u^*(\cdot)$. The pair $(u^*(\cdot), x^*(\cdot))$ will be called the optimal control-state pair.

The characterizations of the value function $V : [0, T] \times \mathbf{C} \to \Re$ and the optimal control-state pair $(u^*(\cdot), x^*(\cdot))$ will normally constitute an solution to the control problem. The optimal classical control problem, Problem (OCCP), is now formally formulated as follows.

Problem (OCCP). For each initial datum $(t, \psi) \in [0, T] \times \mathbf{C}$:

1. Find an $u^*(\cdot) \in \mathcal{U}[t, T]$ that maximizes $J(t, \psi; u(\cdot))$ defined in (3.6) among $\mathcal{U}[t, T]$.
2. Characterize the value function $V : [0, T] \times \mathbf{C} \to \Re$ defined in (3.7).
3. Identify the optimal control-state pair $(u^*(\cdot), x^*(\cdot))$.

3.2 Existence of Optimal Classical Control

In the class $\mathcal{U}[t, T]$ of admissible controls it may happen that there does not exist an optimal control. The following artificial example of Kushner and Dupuis [KD01, p.86] shows that an optimal control does not exist even for a controlled deterministic equation without a delay.

Example. Consider the following one-dimensional controlled deterministic equation:

$$\dot{x}(s) = f(x(s), u(s)) \equiv u(s), \quad s \geq 0$$

with the control set $U = [-1, 1]$. Starting from the initial state $x(0) = x \in \Re$, the objective is to find an admissible (deterministic) control $u(\cdot) = \{u(s), s \geq 0\}$ that minimizes the following discounted cost functional over the infinite time horizon:

$$J(x; u(\cdot)) = \int_0^\infty e^{-\alpha s}[x^2(s) + (u^2(s) - 1)^2]\, ds.$$

Again, let $V : \Re \to \Re$ be the value function of the control problem defined by

$$V(x) = \inf_{u(\cdot) \in \mathcal{U}[0,\infty)} J(x; u(\cdot)).$$

Note that $V(0) = 0$ and, in general, $V(x) = x^2/\alpha$ for all $x \in \Re$. To see this, define the sequence of controls $u^{(k)}(\cdot)$ by

$$u^{(k)}(s) = (-1)^j \text{ on the half-open interval } [j/k, (j+1)/k),\ j = 0, 1, 2, \ldots.$$

It is not hard to see that $J(0; u^{(k)}(\cdot)) \to 0$ as $k \to \infty$. In a sense, when $x(0) = 0$, the optimal control $u^*(\cdot)$ wants to take values ± 1. However, it is easy to check that $u^*(\cdot)$ does not satisfy Definition 3.1.6. Therefore, an optimal control $u^*(\cdot)$ does not exist as defined.

In order to establish the existence of an optimal control for Problem (OCCP), we will enlarge the class of controls, allowing the so-called relaxed controls, so that the existence of an optimal (relaxed) control is guaranteed, and the supremum of the expected objective functional over this new class of controls coincides with the value function $V : [0,T] \times \mathbf{C} \to \Re$ of the original optimal classical control problem defined by (3.7). The idea to show the existence of an optimal relaxed control is to consider a maximizing sequence of admissible relaxed controls $\{\mu^{(k)}(\cdot,\cdot)\}_{k=1}^\infty$ on the Borel measurable space $([0,T] \times U, \mathcal{B}([0,T] \times U))$ and the corresponding sequence of objective functionals $\{\hat{J}(t,\psi; \mu^{(k)}(\cdot,\cdot))\}_{k=1}^\infty$. By the fact that $[0,T] \times U$ is compact (and hence the maximizing sequence of admissible relaxed controls $\{\mu^{(k)}(\cdot,\cdot)\}_{k=1}^\infty$ is compact in the Prohorov metric) and the fact that $\hat{J}(t,\psi;\mu(\cdot,\cdot))$ is upper semicontinuous in admissible relaxed controls $\mu(\cdot,\cdot)$, one can show that the sequence $\{\mu^{(k)}(\cdot,\cdot)\}_{k=1}^\infty$ converges weakly to an admissible relaxed control $\mu^*(\cdot,\cdot)$. This $\mu^*(\cdot,\cdot)$ can be shown to be optimal among the class of admissible relaxed controls and that its value function coincides with the value function of Problem (OCCP). We also prove that an optimal (classical) control exists if the value function $V(t,\psi)$ is finite for each initial datum $(t,\psi) \in [0,T] \times \mathbf{C}$.

We recall the concept and characterizations of weak convergence of probability measures without proofs as follows. The detail can be found in Billingsley [Bil99].

Let (Ξ, d) be a generic metric space with the Borel σ-algebra denoted by $\mathcal{B}(\Xi)$. Let $\mathcal{P}(\Xi)$ (or simply $\mathcal{P}$ whenever there is no ambiguity) be the collection of probability measures defined on $(\Xi, \mathcal{B}(\Xi))$. We will equip the space $\mathcal{P}(\Xi)$ with the Prohorov metric $\pi : \mathcal{P}(\Xi) \times \mathcal{P}(\Xi) \to [0,\infty)$ defined by

$$\pi(\mu,\nu) = \inf\{\epsilon > 0 \mid \mu(A^\epsilon) \le \nu(A) + \epsilon \text{ for all closed } A \in \mathcal{B}(\Xi)\}, \tag{3.9}$$

where A^ϵ is the ϵ-neighborhood of $A \in \mathcal{B}(\Xi)$, that is,

$$A^\epsilon = \{\mathbf{y} \in \Xi \mid d(\mathbf{x},\mathbf{y}) < \epsilon \text{ for some } \mathbf{x} \in \Xi\}.$$

If $\mu^{(k)}, k = 1, 2, \ldots$, is a sequence in $\mathcal{P}(\Xi)$, we say that the sequence $\mu^{(k)}, k = 1, 2, \ldots$, converges weakly to $\mu \in \mathcal{P}(\Xi)$ and is to be denoted by $\mu^{(k)} \Rightarrow \mu$ as $k \to \infty$ if

$$\lim_{k\to\infty} \int_{\Xi} \Phi(\mathbf{x})\mu^{(k)}(d\mathbf{x}) = \int_{\Xi} \Phi(\mathbf{x})\mu(d\mathbf{x}), \quad \forall \Phi \in C_b(\Xi), \tag{3.10}$$

where $C_b(\Xi)$ is the space of bounded and continuous functions $\Phi : \Xi \to \Re$ equipped with the sup-norm:

$$\|\Phi\|_{C_b(\Xi)} = \sup_{\mathbf{x}\in\Xi} |\Phi(\mathbf{x})|, \quad \Phi \in C_b(\Xi).$$

In the case where $\mu^{(k)}$ and $\mu \in \mathcal{P}(\Xi)$ are probability measures induced by Ξ-valued random variables $X^{(k)}$ and X, respectively, we often say that $X^{(k)}$ converges weakly to X and is to be denoted by $X^{(k)} \Rightarrow X$ as $k \to \infty$. A direct consequence of the definition of weak convergence is that $X^{(k)} \Rightarrow X$ implies that $\Phi(X^{(k)}) \Rightarrow \Phi(X)$ for any continuous function Φ from Ξ to another metric space.

We state the following results without proofs. The details of these results can be found in [Bil99].

Theorem 3.2.1 *If Ξ is complete and separable, then $\mathcal{P}(\Xi)$ is complete and separable under the Prohorov metric. Furthermore, if Ξ is compact, then $\mathcal{P}(\Xi)$ is compact.*

Let $\{\mu^{(\lambda)}, \lambda \in \Lambda\} \subset \mathcal{P}(\Xi)$, where Λ is an arbitrary index set.

Definition 3.2.2 *The collection of probability measure $\{\mu^{(\lambda)}, \lambda \in \Lambda\}$ is called tight if for each $\epsilon > 0$, there exists a compact set $K_\epsilon \subset \Xi$ such that*

$$\inf_{\lambda\in\Lambda} \mu^{(\lambda)}(K_\epsilon) \geq 1 - \epsilon. \tag{3.11}$$

If the measures $\mu^{(\lambda)}$ are the induced measures defined by some random variables $X^{(\lambda)}$, then we also refer to the collection $\{X^{(\lambda)}\}$ as tight. Condition (3.11) then reads (in the special case where all the random variables are defined on the same space)

$$\inf_{\lambda\in\Lambda} P\{X^{(\lambda)} \in K_\epsilon\} \geq 1 - \epsilon.$$

Theorem 3.2.3 (Prohorov's Theorem) *If Ξ is complete and separable, then a set $\{\mu^{(\lambda)}, \lambda \in \Lambda\} \subset \mathcal{P}(\Xi)$ has compact closure in the Prohorov metric if and only if $\{\mu^{(\lambda)}, \lambda \in \Lambda\}$ is tight.*

Remark 3.2.4 *Let Ξ_1 and Ξ_2 be two complete and separable metric spaces, and consider the space $\Xi = \Xi_1 \times \Xi_2$ with the usual product space topology. For $\{\mu^{(\lambda)}, \lambda \in \Lambda\} \subset \mathcal{P}(\Xi)$, let $\{\mu_i^{(\lambda)}, \lambda \in \Lambda\} \subset \mathcal{P}(\Xi_i)$, for $i = 1, 2$, be defined by taking $\mu_i^{(\lambda)}$ to be the marginal distribution of $\mu^{(\lambda)}$ on Ξ_i. Then $\{\mu^{(\lambda)}, \lambda \in \Lambda\}$ is tight if and only if $\{\mu_1^{(\lambda)}, \lambda \in \Lambda\}$ and $\{\mu_2^{(\lambda)}, \lambda \in \Lambda\}$ are tight.*

Theorem 3.2.5 *Let Ξ be a metric space and let $\mu^{(k)}, k = 1, 2, \ldots,$ and μ be probability measures in $\mathcal{P}(\Xi)$ satisfying $\mu^{(k)} \Rightarrow \mu$. Let Φ be a real-valued measurable function on Ξ and define $\mathbf{D}(\Phi)$ to be the measurable set of points at which Φ is not continuous. Let $X^{(k)}$ and X be Ξ-valued random variables defined on a probability space $(\Omega, \mathcal{F}, P)$ that induce the measures $\mu^{(k)}$ and μ on Ξ, respectively. Then $\Phi(X^{(k)}) \Rightarrow \Phi(X)$ whenever $P\{X \in \mathbf{D}(\Phi)\} = 0$.*

Theorem 3.2.6 (Aldous Criterion) *Let $X^{(k)}(\cdot) = \{X^{(k)}(t), t \in [0, T]\}, k = 1, 2, \ldots,$ be a sequence of Ξ-valued continuous processes (defined on the same filtered probability space $(\Omega, \mathcal{F}, P, \mathbf{F})$). Then the sequence $\{X^{(k)}(\cdot)\}_{k=1}^{\infty}$ converges weakly if and only if the following condition is satisfied: Given $k = 1, 2, \ldots,$ any bounded $\mathbf{F}$-stopping time τ, and $\delta > 0$, we have*

$$E^{(k)}[\|X^{(k)}(\tau + \delta) - X^{(k)}(\tau)\|_{\Xi}^2 \mid \mathcal{F}^{(k)}(\tau)] \leq 2K^2\delta(\delta + 1).$$

We also recall the following *Skorokhod representation theorem* that is often used to prove convergence with probability 1. The proof can be found in Ethier and Kurtz [EK86].

Theorem 3.2.7 (Skorokhod Representation Theorem) *Let Ξ be a separable metric space and assume the probability measures $\{\mu^{(k)}\}_{k=1}^{\infty} \subset \mathcal{P}(\Xi)$ converges weakly to $\mu \in \mathcal{P}(\Xi)$. Then there exists a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ on which there are defined Ξ-valued random variables $\{\tilde{X}^{(k)}\}_{k=1}^{\infty}$ and $\tilde{X}$ such that for all Borel sets $B \in \mathcal{B}(\Xi)$ and all $k = 1, 2, \ldots,$*

$$\tilde{P}\{\tilde{X}^{(k)} \in B\} = \mu^{(k)}(B), \qquad \tilde{P}\{\tilde{X} \in B\} = \mu(B)$$

and such that

$$\tilde{P}\{\lim_{k \to \infty} \tilde{X}^{(k)} = \tilde{X}\} = 1.$$

3.2.1 Admissible Relaxed Controls

We first define a *deterministic relaxed control* as follows.

Definition 3.2.8 *A deterministic relaxed control is a positive measure μ on the Borel σ-algebra $\mathcal{B}([0, T] \times U)$ such that*

$$\mu([0, t] \times U) = t, \quad t \in [0, T]. \tag{3.12}$$

For each $G \in \mathcal{B}(U)$, the function $t \mapsto \mu([0, t] \times G)$ is absolutely continuous with respect to Lebesque measure on $([0, T], \mathcal{B}([0, T])$ by virtue of (3.12). Denote by $\dot{\mu}(\cdot, G) = \frac{d}{dt}\mu([0, t] \times G)$ any Lebesque density of $\mu([0, t] \times G)$. The family of densities $\{\dot{\mu}(\cdot, G), G \in \mathcal{B}(U)\}$ is a probability measure on $\mathcal{B}(U)$ for each $t \in [0, T]$, and

$$\begin{aligned} \mu_{u(\cdot)}(B) &= \int_0^T \int_U \mathbf{1}_{\{(t,u) \in B\}} \mu(dt, du) \\ &= \int_0^T \int_U \mathbf{1}_{\{(t,u) \in B\}} \dot{\mu}(t, du)\, dt, \quad \forall B \in \mathcal{B}([0, T] \times U). \end{aligned} \tag{3.13}$$

Denote by $\mathcal{R}$ the space of deterministic relaxed controls that is equipped with the *weak compact topology* induced by the following notion of convergence: A sequence $\{\mu^{(k)}, k = 1, 2, \ldots\}$ of relaxed controls converges (weakly) to $\mu \in \mathcal{R}$ if

$$\lim_{k\to\infty} \int_{[0,T]\times U} \gamma(t,u)\, d\mu^{(k)}(t,u) = \int_{[0,T]\times U} \gamma(t,u)\, d\mu(t,u), \tag{3.14}$$
$$\forall \gamma \in C_c([0,T]\times U),$$

where $C_c([0,T]\times U)$ is the space of all real-valued continuous functions on $[0,T]\times U$ having compact support. Note that if U is compact then $C_c([0,T]\times U) = C([0,T]\times U)$. Under the *weak-compact* topology defined above, $\mathcal{R}$ is a (sequentially) compact space; that is, every sequence in $\mathcal{R}$ has a subsequence that converges to an element in $\mathcal{R}$ in the sense of (3.14).

Now, we introduce a suitable filtration for $\mathcal{R}$ as follows. We first identify each $\mu \in \mathcal{R}$ as a linear functional on $C([0,T]\times U)$ in the following way:

$$\mu(\varsigma) \equiv \int_0^T \int_U \varsigma(t,u)\mu(dt,du), \quad \forall \varsigma \in C([0,T]\times U).$$

For any $\varsigma \in C([0,T]\times U)$ and $t \in [t,T]$, define $\varsigma^t \in C([0,T]\times U)$ by

$$\varsigma^t(s,u) \equiv \varsigma(s\wedge t, u).$$

Since $C([0,T]\times U)$ is separable (and therefore has a countable dense subset), we may let $\{\varsigma^{(k)}\}_{k=1}^\infty$ be countable dense subset (with respect to the uniform norm). It is easy to see that $\{\varsigma^{(k),t}\}_{k=1}^\infty$ is dense in the set $\{\varsigma^t \mid \varsigma \in C([0,T]\times U)\}$. Define

$$\mathcal{B}_s(\mathcal{R}) \equiv \sigma\{\{\mu \in \mathcal{R} \mid \mu(\varsigma^t) \in B\} \mid \varsigma \in C([0,T]\times U), t \in [0,s], B \in \mathcal{B}(\Re)\}.$$

One can easily show that $\mathcal{B}_s(\mathcal{R})$ can be generated by cylinder sets of the following form:

$$\sigma\{\{\mu \in \mathcal{R} \mid \mu(\varsigma^{(k),t}) \in (a,b)\} \mid s \ge t \in \mathbf{Q}, k = 1, 2, \ldots, a, b \in \mathbf{Q}\}. \tag{3.15}$$

Definition 3.2.9 *A relaxed control process is an $\mathcal{R}$-valued random variable μ, defined on a Brownian stochastic basis $(\Omega, \mathcal{F}, P, \mathbf{F}, W(\cdot))$, such that the mapping $\omega \mapsto \mu([0,t]\times G)(\omega)$ is $\mathcal{F}(t)$-measurable for all $t \in [0,T]$ and $G \in \mathcal{B}(U)$.*

Using the relaxed control process $\mu(\cdot,\cdot) \in \mathcal{R}$, the controlled state equation can be written as

$$dx(s) = \int_U f(s, x_s, u)\dot{\mu}(s,du)\, ds + \int_U g(s, x_s, u)\dot{\mu}(s,du)\, dW(s), \quad s \in [t,T],$$

or, equivalently, in the form of the stochastic integral equation:

$$x(s) = \psi(0) + \int_t^s \int_U f(\lambda, x_\lambda, u)\dot{\mu}(\lambda, du)\, d\lambda \qquad (3.16)$$
$$+ \int_t^s \int_U g(\lambda, x_\lambda, u)\dot{\mu}(\lambda, du)\, dW(\lambda), \quad s \in [t, T],$$

with the initial datum $(t, \psi) \in [0, T] \times \mathbf{C}$. The objective functional can be written as

$$\hat{J}(t, \psi; \mu(\cdot, \cdot)) = E\Big[\int_U \int_t^T L(s, x_s(\cdot; t, \psi, \mu(\cdot, \cdot)), u)\dot{\mu}(s, du)\, ds$$
$$+ \Psi(x_T(\cdot; t, \psi, \mu(\cdot, \cdot))\Big], \qquad (3.17)$$

where $\{x(s; t, \psi, \mu(\cdot, \cdot)), s \in [t, T]\}$ is the solution process of (3.16) when the relaxed control process $\mu(\cdot, \cdot) \in \mathcal{R}$ is applied.

We now define an admissible relaxed control as follows.

Definition 3.2.10 *For each initial datum $(t, \psi) \in [0, T] \times \mathbf{C}$, a six-tuple $(\Omega, \mathcal{F}, P, \mathbf{F}, W(\cdot), \mu(\cdot, \cdot))$ is said to be an admissible relaxed control at $(t, \psi) \in [0, T] \times \mathbf{C}$ if it satisfies the following conditions:*

1. *$(\Omega, \mathcal{F}, P, \mathbf{F}, W(\cdot))$ is a certain m-dimensional Brownian stochastic basis.*
2. *$\mu(\cdot, \cdot) \in \mathcal{R}$ is a relaxed control defined on the Brownian stochastic basis $(\Omega, \mathcal{F}, P, \mathbf{F}, W(\cdot))$.*
3. *Under the relax control process $\mu(\cdot, \cdot)$, (3.16) admits a unique strong solution $x(\cdot; t, \psi, \mu(\cdot, \cdot)) = \{x(s; t, \psi, \mu(\cdot, \cdot)), s \in [t, T]\}$ and through each initial datum $(t, \psi) \in [0, T] \times \mathbf{C}$.*
4. *The control process $\mu(\cdot, \cdot)$ is such that*

$$E\Big[\int_U \int_t^T |L(s, x_s(\cdot; t, \psi, \mu(\cdot, \cdot)), u)|\dot{\mu}(du, s)\, ds$$
$$+ |\Psi(x_T(\cdot; t, \psi, \mu(\cdot, \cdot))|\Big] < \infty,$$

The collection of admissible relaxed controls $(\Omega, \mathcal{F}, P, \mathbf{F}, W(\cdot), \mu(\cdot, \cdot))$ over the interval $[t, T]$ will be denoted by $\hat{\mathcal{U}}[t, T]$. Again, when there is no ambiguity, we often write $\mu(\cdot, \cdot) \in \hat{\mathcal{U}}[t, T]$ instead of $(\Omega, \mathcal{F}, P, \mathbf{F}, W(\cdot), \mu(\cdot, \cdot))$.

The optimal relaxed control problem can be stated as follows.
Problem (ORCP) Find an optimal relaxed control $\mu^*(\cdot, \cdot) \in \hat{\mathcal{U}}[t, T]$ that maximizes (3.17) subject to (3.16).

We again define the value function $\hat{V} : [0, T] \times \mathbf{C} \to \Re$ for the Problem (ORCP) by

$$\hat{V}(t, \psi) = \sup_{\mu(\cdot, \cdot) \in \hat{\mathcal{U}}[t, T]} \hat{J}(t, \psi; \mu(\cdot, \cdot)). \qquad (3.18)$$

We have the following existence theorem for Problem (ORCP).

Theorem 3.2.11 *Let Assumption 3.1.1 hold. Given any initial datum $(t, \psi) \in [0, T] \times \mathbf{C}$, then Problem (ORCP) admits an optimal relaxed control $\mu^*(\cdot, \cdot)$, and its value function $\hat{V}$ coincides with the value function V of Problem (OCCP).*

We will postpone proof of Theorem 3.2.11 until the end of the next subsection.

3.2.2 Existence Result

For the existence of an optimal control for Problem (OCCP), we need the following Roxin condition:
(**Roxin's Condition**). For every $(t, \psi) \in [0, T] \times \mathbf{C}$, the set

$$\left(f, gg^\top, L\right)(t, \psi, U) \equiv \Big\{ \left(f_i(t, \psi, u), (gg^\top)_{ij}(t, \psi, u), L(t, \psi, u)\right) \mid \\ u \in U, i, j = 1, 2, \ldots, n\Big\}$$

is convex in $\Re^{n+nn+1}$.

The main purpose of this subsection is to prove the existence theorem.

Theorem 3.2.12 *Let Assumption 3.1.1 and the Roxin condition hold. Given any initial datum $(t, \psi) \in [0, T] \times \mathbf{C}$, then Problem (OCCP) admits an optimal classical control $u^*(\cdot) \in \mathcal{U}[t, T]$ if the value function $V(t, \psi)$ is finite.*

Proof. Without loss of generality, we can and will assume that $t = 0$ in the following for notational simplicity. The proof is similar to that of Theorem 2.5.3 in Yong and Zhou [YZ99] and will be carried out by the following steps:
Step 1. Since $V(0, \psi)$ is finite, we can find a sequence of maximizing admissible controls in $\mathcal{U}[t, T]$,

$$\{(\Omega^{(k)}, \mathcal{F}^{(k)}, P^{(k)}, \mathbf{F}^{(k)}, W^{(k)}(\cdot), u^{(k)}(\cdot))\}_{k=1}^{\infty},$$

such that

$$\lim_{k \to \infty} J(0, \psi; u^{(k)}(\cdot)) = V(0, \psi). \tag{3.19}$$

Let $x^{(k)}(\cdot) = \{x(\cdot; 0, \psi, u^{(k)}(\cdot)), s \in [0, T]\}$ be the solution of (3.1) corresponding to $u^{(k)}(\cdot)$. Define

$$X^{(k)}(\cdot) \equiv (x^{(k)}(\cdot), F^{(k)}(\cdot), G^{(k)}(\cdot), L^{(k)}(\cdot), W^{(k)}(\cdot)), \tag{3.20}$$

where the processes $F^{(k)}(\cdot)$, $G^{(k)}(\cdot)$, and $L^{(k)}(\cdot)$ are defined as follows:

$$F^{(k)}(s) = \int_0^s f(t, x_t^{(k)}(\cdot; 0, \psi, u^{(k)}(\cdot)), u^{(k)}(t))\, dt,$$

$$G^{(k)}(s) = \int_0^s g(t, x_t^{(k)}(\cdot; 0, \psi, u^{(k)}(\cdot)), u^{(k)}(t))\, dW^{(k)}(t),$$

and

$$L^{(k)}(s) = \int_0^s e^{-\alpha t} L(t, x_t^{(k)}(\cdot; 0, \psi, u^{(k)}(\cdot)), u^{(k)}(t))\, dt, \quad s \in [0, T].$$

Step 2. We prove the following lemma:

Lemma 3.2.13 *Assume Assumption 3.1.1 holds. Then there exists a constant $K > 0$ such that*

$$E^{(k)} \left[\left| X^{(k)}(s_1) - X^{(k)}(s_2) \right|^4 \right] \leq K|s_1 - s_2|^2, \quad \forall s_1, s_2 \in [0, T], \ \forall k = 1, 2, \ldots,$$

where $E^{(k)}[\cdots]$ is the expectation with respect to the probability measure $P^{(k)}$.

Proof of the Lemma. Let us fix k, $0 \leq s_1 \leq s_2 \leq T$, and consider

$$\begin{aligned}
&E^{(k)} \left[|F^{(k)}(s_1) - F^{(k)}(s_2)|^4 \right] \\
&\leq E^{(k)} \left[\left| \int_{s_1}^{s_2} f(t, x_t^{(k)}(\cdot; 0, \psi, u^{(k)}(\cdot)), u^{(k)}(t)) dt \right|^4 \right] \\
&\leq |s_1 - s_2|^2 E^{(k)} \left[\left(\int_{s_1}^{s_2} \left| f(t, x_t^{(k)}(\cdot; 0, \psi, u^{(k)}(\cdot)), u^{(k)}(t)) \right|^2 dt \right)^2 \right] \\
&\leq |s_1 - s_2|^2 \int_{s_1}^{s_2} K_{grow}^2 E^{(k)} \left[\left(1 + \|x_t^{(k)}(\cdot; t, \psi, u^{(k)}(\cdot))\| \right)^2 \right] dt.
\end{aligned}$$

Since $x_t^{(k)}(\cdot; 0, \psi, u^{(k)}(\cdot))$ is continuous $P^{(k)}$-a.s. in t on the compact interval $[0, T]$, it can be shown that there exists a constant $K > 0$ such that

$$\int_{s_1}^{s_2} K_{grow}^2 E^{(k)} \left[\left(1 + \|x_t^{(k)}(\cdot; 0, \psi, u^{(k)}(\cdot))\| \right)^2 \right] dt < K.$$

Therefore,

$$E^{(k)} \left[|F^{(k)}(s_1) - F^{(k)}(s_2)|^4 \right] \leq K|s_1 - s_2|^2, \quad \forall s_1, s_2 \in [0, T], \ \forall k = 1, 2, \ldots.$$

Similar conclusion holds for $L^{(k)}$; that is,

$$E^{(k)} \left[|L^{(k)}(s_1) - L^{(k)}(s_2)|^4 \right] \leq K|s_1 - s_2|^2, \quad \forall s_1, s_2 \in [0, T], \ \forall k = 1, 2, \ldots.$$

We consider

$$\begin{aligned}
&E^{(k)} \left[|G^{(k)}(s_1) - G^{(k)}(s_2)|^4 \right] \\
&\leq E^{(k)} \left[\left| \int_{s_1}^{s_2} g(t, x_t^{(k)}(\cdot; 0, \psi, u^{(k)}(\cdot)), u^{(k)}(t)) dW^{(k)}(t) \right|^4 \right]
\end{aligned}$$

$$\leq K_1(s_2 - s_1) \left(\int_{s_1}^{s_2} E^{(k)} \left[\left| g(t, x_t^{(k)}(\cdot; 0, \psi, u^{(k)}(\cdot)), u^{(k)}(t)) \right|^4 \right] dt \right)^2$$

for some constant $K_1 > 0$

$$\leq |s_1 - s_2|^2 \int_{s_1}^{s_2} K_1 K_{grow}^2 \left(1 + \|x_t^{(k)}(\cdot; 0, \psi, u^{(k)}(\cdot))\| \right)^2 dt.$$

$$\leq K|s_1 - s_2|^2 \text{ for some constant } K > 0.$$

It is clear that

$$E^{(k)} \left[|W^{(k)}(s_1) - W^{(k)}(s_2)|^4 \right] = E^{(k)} \left[|W^{(k)}(s_2 - s_1)|^4 \right] \leq K|s_1 - s_2|^2,$$

since $W^{(k)}(s_2 - s_1)$ is Gaussian with mean zero and variance $I^{(m)}(s_2 - s_1)$.

The above estimates give

$$E^{(k)} \left[\left| X^{(k)}(s_1) - X^{(k)}(s_2) \right|^4 \right] \leq K|s_1 - s_2|^2, \quad \forall s_1, s_2 \in [0, T], \ \forall k = 1, 2, \ldots.$$

This completes the proof of the lemma. □

From the above lemma, we use the following well-known results to conclude that $\{(X^{(k)}(\cdot), \mu_{u^{(k)}}(\cdot, \cdot))\}_{k=1}^\infty$ is tight as a sequence of $C([0, T], \Re^{3n+m+1})$, since $\mathcal{R}$ is compact.

Proposition 3.2.14 *Let $\{\zeta^{(k)}(\cdot)\}_{k=1}^\infty$ be a sequence of d-dimensional continuous processes over $[0, T]$ on a probability space $(\Omega, \mathcal{F}, P)$ satisfying the following conditions:*

$$\sup_{k \geq 1} E[|\zeta^{(k)}(0)|^c] < \infty$$

and

$$\sup_{k \geq 1} E[|\zeta^{(k)}(t) - \zeta^{(k)}(s)|^a] \leq K|t - s|^{1+b}, \quad \forall t, s \in [0, T],$$

for some constants $a, b, c > 0$. Then $\{\zeta^{(k)}(\cdot)\}_{k=1}^\infty$ is tight as $C([0, T], \Re^d)$-valued random variables. As a consequence there exists a subsequence $\{k_j\}$, d-dimensional continuous processes $\{\hat{\zeta}^{(k_j)}(\cdot)\}_{j=1}^\infty$ and $\hat{\zeta}(\cdot)$ defined on a probability space $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P})$ such that

$$P(\zeta^{(k_j)}(\cdot) \in A) = \hat{P}(\hat{\zeta}^{(k_j)}(\cdot) \in A), \quad \forall A \in \mathcal{B}(C([0, T], \Re^d))$$

and

$$\lim_{j \to \infty} \hat{\zeta}^{(k_j)}(\cdot) \to \hat{\zeta}(\cdot) \ \textit{in} \ C([0, T], \Re^d), \ \hat{P}\text{-a.s.}$$

Corollary 3.2.15 *Let $\zeta(\cdot)$ be a d-dimensional process over $[0, T]$ such that*

$$E[|\zeta(t) - \zeta(s)|^a] \leq K|t - s|^{1+b}, \quad \forall t, s \in [0, T],$$

for some constants $a, b > 0$. Then there exists a d-dimensional continuous process that is stochastically equivalent to $\zeta(\cdot)$.

We refer the readers to Ikeda and Watanabe [IW81, pp.17-20] for a proof of the above proposition and corollary.

Step 3. Since

$$\{(X^{(k)}(\cdot), \mu_{u^{(k)}}(\cdot,\cdot))\}_{k=1}^{\infty}$$

is tight as a sequence in $C([0,T],\Re^{3n+m+1})$, by the Skorokhod representation theorem (see Theorem 3.2.7), one can choose a subsequence (still labeled as $\{k\}$) and have

$$\{(\bar{X}^{(k)}(\cdot), \bar{\mu}^{(k)}(\cdot,\cdot))\} \equiv \{(\bar{x}^{(k)}(\cdot), \bar{F}^{(k)}(\cdot), \bar{G}^{(k)}(\cdot), \bar{L}^{(k)}(\cdot), \bar{W}^{(k)}(\cdot), \bar{\mu}^{(k)}(\cdot,\cdot))\}$$

and

$$(\bar{X}(\cdot), \bar{\mu}(\cdot,\cdot)) \equiv (\bar{x}(\cdot), \bar{F}(\cdot), \bar{G}(\cdot), \bar{L}(\cdot), \bar{W}(\cdot), \bar{\mu}(\cdot,\cdot))$$

on a suitable common probability space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{P})$ such that

$$\text{law of } (\bar{X}^{(k)}(\cdot), \bar{\mu}^{(k)}(\cdot,\cdot)) = \text{law of } (X^{(k)}(\cdot), \mu^{(k)}(\cdot,\cdot)), \quad \forall k \geq 1, \tag{3.21}$$

and $\bar{P}$-a.s.,

$$\bar{X}^{(k)}(\cdot) \to \bar{X}(\cdot) \text{ uniformly on } [0,T] \tag{3.22}$$

and

$$\bar{\mu}^{(k)}(\cdot,\cdot) \to \bar{\mu}(\cdot,\cdot) \text{ weakly on } \mathcal{R}. \tag{3.23}$$

Step 4. Construct the filtration $\bar{\mathbf{F}}^{(k)} = \{\bar{\mathcal{F}}^{(k)}(s), s \geq 0\}$ and $\bar{\mathbf{F}} = \{\bar{\mathcal{F}}, s \geq 0\}$, where

$$\bar{\mathcal{F}}^{(k)}(s) = \sigma\{(\bar{W}^{(k)}(t), \bar{x}^{(k)}(t)), t \leq s\} \vee (\bar{\mu}^{(k)})^{-1}(\mathcal{B}_s(\mathcal{R}))$$

and

$$\bar{\mathcal{F}}(s) = \sigma\{(\bar{W}(t), \bar{x}(t)), t \leq s\} \vee (\bar{\mu})^{-1}(\mathcal{B}_s(\mathcal{R})), \quad s \geq 0.$$

By the definition of $\mathcal{B}_s(\mathcal{R})$ and the fact that the σ-algebra generated by the cylinder sets of $C([0,T];\Re^d)$ coincides with $\mathcal{B}(C([0,T];\Re^d))$, $\bar{\mathcal{F}}^{(k)}(s)$ is the σ-algebra generated by

$$\bar{W}^{(k)}(t_1), \ldots, \bar{W}^{(k)}(t_l), \bar{x}^{(k)}(t_1), \ldots, \bar{x}^{(k)}(t_l), \bar{\mu}^{(k)}(\varsigma^{(j),t_1}), \ldots, \bar{\mu}^{(k)}(\varsigma^{(j),t_l}),$$

$$t_1, \ldots, t_l \leq s, \ \varsigma^{(j)} \in C([0,T],U) \text{ and } j, l = 1, 2, \ldots.$$

A similar statement can be made for $\bar{\mathcal{F}}(s)$.

We need to show that $\bar{W}^{(k)}(\cdot) = \{\bar{W}^{(k)}(s), s \geq 0\}$ is an $\bar{\mathbf{F}}^{(k)}$ Brownian motion. We first note that $W^{(k)}(\cdot)$ is a Brownian motion with respect to

$$\left\{\sigma\{(W^{(k)}(t), x^{(k)}(t)), t \leq s\} \vee (\mu_{u^{(k)}}^{-1}(\mathcal{B}_s(\mathcal{R})), s \geq 0\right\}.$$

This implies that for any $0 \leq t \leq s \leq T$ and any bounded function H on $\Re^{(m+n+b)l}$ (b is a positive integer), we have

$$E^{(k)}\left[H(y^{(k)})(W^{(k)}(s) - W^{(k)}(t))\right] = 0,$$

where

$$y^{(k)} = \{W^{(k)}(t_i), x^{(k)}(t_i), \mu^{(k)}(\varsigma^{j_a,t_i})\}$$
$$\forall \le 0 \le t_1 \le t_2 \le \cdots \le t_l \le t,\ a = 1, 2, \ldots, b.$$

We have

$$\bar{E}^{(k)}\left[H(\bar{y}^{(k)})(\bar{W}^{(k)}(s) - \bar{W}^{(k)}(t))\right] = 0,$$

where

$$\bar{y}^{(k)} = \{\bar{W}^{(k)}(t_i), \bar{x}^{(k)}(t_i), \bar{\mu}^{(k)}(\varsigma^{j_a,t_i})\}$$
$$\forall 0 \le t_1 \le t_2 \le \cdots \le 0 \le t_l \le t, a = 1, 2, \ldots, b,$$

since

$$\text{law of } (\bar{X}^{(k)}(\cdot), \bar{\mu}^{(k)}(\cdot,\cdot)) = \text{law of } (X^{(k)}(\cdot), \mu^{(k)}(\cdot,\cdot)), \quad \forall k \ge 1.$$

We therefore have

$$E^{(k)}[(W^{(k)}(s) - W^{(k)}(t)) \mid \mathcal{F}^{(k)}(t)] = 0.$$

In order to show $\bar{W}^{(k)}(\cdot)$ is an $\bar{\mathbf{F}}^{(k)}$ Brownian motion, we need

$$E^{(k)}[(W^{(k)}(s) - W^{(k)}(t))(W^{(k)}(s) - W^{(k)}(t))^\top \mid \mathcal{F}^{(k)}(t)] = (s-t)I^{(m)}.$$

This can be shown similarly.
Step 5. Again, since

$$\text{law of } (\bar{X}^{(k)}(\cdot), \bar{\mu}^{(k)}(\cdot,\cdot)) = \text{law of } (X^{(k)}(\cdot), \mu^{(k)}(\cdot,\cdot)), \quad \forall k \ge 1,$$

the following stochastic integral equation (defined on $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbf{F}}^{(k)}, \bar{P})$) holds:

$$\begin{aligned}\bar{x}^{(k)}(s) = \psi(0) &+ \int_0^s \tilde{f}(t, \bar{x}_t^{(k)}, \bar{\mu}^{(k)}) \quad dt \\ &+ \int_0^s \tilde{g}(t, \bar{x}_t^{(k)}, \bar{\mu}^{(k)})\, d\bar{W}^{(k)}(t),\end{aligned}$$

where

$$\tilde{f}(t, \bar{x}_t^{(k)}, \bar{\mu}^{(k)}) = \int_U f(t, \bar{x}_t^{(k)}, u)\dot{\bar{\mu}}^{(k)}(t, du)$$

and

$$\tilde{g}(t, \bar{x}_t^{(k)}, \bar{\mu}^{(k)}) = \int_U g(t, \bar{x}_t^{(k)}, u)\dot{\bar{\mu}}^{(k)}(t, du).$$

Note that the above integrals are well defined, since $\bar{W}^{(k)}(\cdot)$ is a $\bar{\mathbf{F}}^{(k)}$ Brownian motion. Moreover, for each $s \in [0, T]$,

$$\lim_{k\to\infty} \int_0^s \tilde{f}(t, \bar{x}_t^{(k)}, \bar{\mu}^{(k)})\, dt = \int_0^s \tilde{f}(t, \bar{x}_t, \bar{\mu})\, dt$$

$$\lim_{k\to\infty}\int_0^s e^{-\alpha t}\tilde{L}(t,\bar{x}_t^{(k)},\bar{\mu}^{(k)})\,dt = \int_0^s e^{-\alpha t}\tilde{L}(t,\bar{x}_t,\bar{\mu})\,dt, \qquad \bar{P}\text{-a.s.},$$

$$e^{-\alpha(T-t)}\Psi(\bar{x}_T^{(k)}) \to e^{-\alpha(T-t)}\Psi(\bar{x}_T),\ \ \bar{P}\text{-a.s.},$$

and

$$\begin{aligned}&\lim_{k\to\infty}\int_0^s \tilde{g}(t,\bar{x}_t^{(k)},\bar{\mu}^{(k)})\,d\bar{W}^{(k)}(t)\\ &= \int_0^s\int_U \tilde{g}(t,\bar{x}_t,\bar{\mu})\,d\bar{W}(t), \qquad \bar{P}\text{-a.s.},\end{aligned}$$

where

$$\tilde{f}(t,\bar{x}_t,\bar{\mu}) = \int_U f(t,\bar{x}_t,u)\dot{\bar{\mu}}(t,du),$$

$$\tilde{L}(t,\bar{x}_t^{(k)},\bar{\mu}^{(k)}) = \int_U L(t,\bar{x}_t^{(k)},u)\dot{\bar{\mu}}^{(k)}(t,du),$$

$$\tilde{L}(t,\bar{x}_t,\bar{\mu}) = \int_U L(t,\bar{x}_t,u)\dot{\bar{\mu}}(t,du),$$

and

$$\tilde{g}(t,\bar{x}_t,\bar{\mu}) = \int_U g(t,\bar{x}_t,u)\dot{\bar{\mu}}(t,du).$$

We have by taking the limit $k\to\infty$

$$\begin{aligned}\bar{x}(s) &= \psi(0) + \int_0^s \tilde{f}(t,\bar{x}_t,\bar{\mu})\,dt\\ &\quad + \int_0^s \tilde{g}(t,\bar{x}_t,\bar{\mu})\,d\bar{W}(t), \quad \forall s\in[0,T],\ \bar{P}-a.s.\end{aligned}$$

Moreover,

$$\begin{aligned}J(0,\psi;u^{(k)}(\cdot)) &= \bar{E}\left[\int_0^T e^{-\alpha t}\tilde{L}(t,\bar{x}_t,\bar{\mu}^{(k)})\,dt + e^{-\alpha T}\Psi(\bar{x}_T^{(k)})\right]\\ &\to \sup_{u(\cdot)\in\mathcal{U}[0,T]} J(0,\psi;u(\cdot))\\ &\quad\ \text{as } k\to\infty.\end{aligned}$$

Step 6. Let us consider the sequence

$$\tilde{A}^{(k)}(s) \equiv (\tilde{g}\tilde{g}^\top)(s,\bar{x}_s^{(k)},\bar{\mu}^{(k)}), \quad s\in[0,T].$$

By the Lipschitz continuity and linear growth conditions on f and g, it is tedious but straight forward to show that

$$\sup_k \bar{E}\left[\int_0^T |\tilde{A}^{(k)}(s)|^2 ds\right] < \infty.$$

Hence the sequence $\{\tilde{A}^{(k)}\}_{k=1}^{\infty}$ is weakly relatively compact in the space $L^2([0,T] \times \bar{\Omega}, \mathcal{S}^n)$, where $\mathcal{S}^n$ is the space of symmetric $n \times n$ matrices. We can then find a subsequence (still labelled by $\{k\}$) and a function $\tilde{A} \in L^2([0,T] \times \bar{\Omega}, \mathcal{S}^n)$ such that

$$\tilde{A}^{(k)} \to \tilde{A}, \quad \text{weakly on } L^2([0,T] \times \bar{\Omega}, \mathcal{S}^n). \tag{3.24}$$

Denoting by $\tilde{A}_{ij}$ the (ij) entry of the matrix $\tilde{A}$, we claim that for almost all (s,ω),

$$\begin{aligned} \underline{\lim}_{k\to\infty} \tilde{A}_{ij}^{(k)}(s,\omega) &\leq \tilde{A}_{ij}(s,\omega) \\ &\leq \overline{\lim}_{k\to\infty} \tilde{A}_{ij}^{(k)}(s,\omega), \quad i,j = 1,2,\ldots,n. \end{aligned} \tag{3.25}$$

Indeed, if (3.25) is not true and on a set $A \subset [0,T] \times \bar{\Omega}$ of positive measure,

$$\underline{\lim}_{k\to\infty} \tilde{A}_{ij}^{(k)}(s,\omega) > \tilde{A}_{ij}(s,\omega),$$

then, by Fatou's lemma, we have

$$\underline{\lim}_{k\to\infty} \int_A \tilde{A}_{ij}^{(k)}(s,\omega)dsd\bar{P}(\omega) > \int_A \tilde{A}_{ij}(s,\omega)dsd\bar{P}(\omega),$$

which is a contradiction to (3.24). The same can be said for $\overline{\lim}$, which proves (3.25). Moreover, by the Lipschitz continuity and linear growth of f and g and the fact that $\bar{X}^{(k)}(\cdot) \to \bar{X}(\cdot)$ uniformly on $[0,T]$, for almost all (s,ω), we have

$$\begin{aligned} &\underline{\lim}_{k\to\infty} \tilde{A}^{(k)}(s) \\ &= \underline{\lim}_{k\to\infty} (\tilde{g}\tilde{g}^\top)(s, \bar{x}_s^{(k)}, \bar{\mu}^{(k)}), \quad (s,\omega) \in [0,T] \times \bar{\Omega}, \end{aligned} \tag{3.26}$$

and

$$\begin{aligned} &\overline{\lim}_{k\to\infty} \tilde{A}^{(k)}(s) \\ &= \overline{\lim}_{k\to\infty} (\tilde{g}\tilde{g}^\top)(s, \bar{x}_s^{(k)}, \bar{\mu}^{(k)}), \quad (s,\omega) \in [0,T] \times \bar{\Omega}, \end{aligned} \tag{3.27}$$

Then, combining (3.25), (3.26), (3.27) and the Roxin condition, we have

$$\tilde{A}(s,\omega) \in (gg^\top)(s, \bar{x}_s(\omega), U), \quad a.e. (s,\omega). \tag{3.28}$$

Modify $\tilde{A}$ on a null set, if necessary, so that (3.28) holds for all $(s,\omega) \in [0,T] \times \bar{\Omega}$. One can similarly prove that there are $\tilde{f}$ and $\tilde{L} \in L^2([0,T] \times \bar{\Omega}, \Re)$ such that

$$\tilde{f}^{(k)} \to \tilde{f}, \tilde{L}^{(k)} \to \tilde{L}, \quad \text{weakly on } L^2([0,T] \times \bar{\Omega}, \Re), \tag{3.29}$$

and

$$\tilde{f}(s,\omega) \in f(s, \bar{x}_s(\omega), U), \quad \tilde{L}(s,\omega) \in L(s, \bar{x}_s(\omega), U), \tag{3.30}$$

$$\forall (s,\omega) \in [0,T] \times \bar{\Omega}.$$

By (3.28), (3.30) the Roxin condition, and a measurable selection theorem (see Corollary 2.26 of Li and Yong [LY91, p102]), there is a U-valued, $\bar{\mathbf{F}}$-adapted process $\bar{u}(\cdot)$ such that

$$(\tilde{f}, \tilde{A}, \tilde{L})(s,\omega) = (f, gg^\top, L)(s, \bar{x}_s(\omega), \bar{u}(s,\omega)), \tag{3.31}$$
$$\forall (s,\omega) \in [0,T] \times \bar{\Omega}.$$

Step 7. The last step is to use Roxin condition to show that there exists an m-dimensional Brownian motion defined on the filtered space $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P}, \hat{\mathbf{F}})$ which extends the filtered probability space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{P}, \bar{\mathbf{F}})$. We, then, conclude that

$$(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P}, \hat{\mathbf{F}}, \hat{W}(\cdot), \bar{u}(\cdot)) \in \mathcal{U}[0,T]$$

is an optimal control.

We next prove that the Itô's integral process $\bar{I}(\tilde{g})(\cdot) = \{\bar{I}(\tilde{g})(s), s \in [0,T]\}$ is an $\bar{\mathbf{F}}$-martingale, where

$$\bar{I}(\tilde{g})(s) = \int_0^s \tilde{g}(t, \bar{x}_t, \bar{\mu})\, d\bar{W}(t), \quad s \in [0,T].$$

To see this, once again, let $0 \leq t \leq s \leq T$, and define

$$\bar{y}^{(k)} \equiv \{\bar{W}^{(k)}(t_i), \bar{x}^{(k)}(t_i), \bar{\mu}^{(k)}(\varsigma^{j_a,t_i})\},$$

and

$$\bar{y} \equiv \{\bar{W}(t_i), \bar{x}(t_i), \bar{\mu}(\varsigma^{j_a,t_i})\},$$
$$0 \leq t_1 \leq t_2 \leq \cdots \leq t_l \leq s, \quad a = 1, 2, \ldots, b.$$

Since $\bar{I}^{(k)}(\tilde{g})(\cdot)$ is a $\bar{\mathbf{F}}^{(k)}$-martingale, for any bounded continuous function $H : \Re^{(m+n+b)l} \to \Re$, we have

$$\begin{aligned} 0 &= \bar{E}[\Phi(\bar{y}^{(k)}(\bar{I}^{(k)}(\tilde{g})(s) - \bar{I}^{(k)}(\tilde{g})(t))] \\ &\to \bar{E}[\Phi(\bar{y}(\bar{I}(\tilde{g})(s) - \bar{I}(\tilde{g})(t))], \end{aligned} \tag{3.32}$$

since $\bar{X}^{(k)}(\cdot) \to \bar{X}(\cdot)$ uniformly on $[0,T]$ and $\bar{\mu}^{(k)} \to \bar{\mu}$ weakly on $\mathcal{R}$ and by the dominated convergence theorem. This proves that $\bar{I}(\tilde{g})(\cdot)$ is an $\bar{\mathbf{F}}$-martingale. Furthermore,

$$\langle \bar{I}^{(k)}(\tilde{g}) \rangle (s) = \int_0^s \tilde{A}^{(k)}(t)dt,$$

where $\langle \bar{I}^{(k)}(\tilde{g}) \rangle$ is the quadratic variation of $\bar{I}^{(k)}(\tilde{g})(\cdot)$. Hence,

$$(\bar{I}^{(k)}(\tilde{g}))(\bar{I}^{(k)}(\tilde{g}))^\top - \int_0^s \tilde{A}^{(k)}(t)dt$$

is an $\bar{\mathbf{F}}^{(k)}$-martingale. Recalling $\tilde{A}^{(k)}(\cdot) \to \tilde{A}(\cdot)$ weakly on $L^2([0,T] \times \bar{\Omega})$, we have for any $t, s \in [0,T]$,

$$\int_t^s \tilde{A}^{(k)}(\lambda)d\lambda \to \int_t^s (gg^\top)(\lambda, \bar{x}_\lambda, \bar{u}(\lambda))d\lambda, \quad \text{weakly on } L^2(\Omega).$$

On the other hand, by the dominated convergence theorem, we have, for real-valued function H of appropriate dimension,

$$H(\bar{y}^{(k)}) \to H(\bar{y}), \quad \text{strongly on } L^2(\Omega).$$

Thus,

$$\bar{E}\left[H(\bar{y}^{(k)}) \int_t^s \tilde{A}^{(k)}(\lambda)d\lambda\right] \to \bar{E}\left[H(\bar{y}) \int_t^s (gg^\top)(\lambda, \bar{x}_\lambda, \bar{u}(\lambda))d\lambda\right].$$

Therefore, using an argument similar to the above, we obtain that $\bar{M}(\cdot) = \{\bar{M}(s), s \in [0,T]\}$ is an $\bar{\mathbf{F}}$-martingale, where

$$\bar{M}(s) \equiv (\bar{I}(g))(\bar{I}(g))^\top(s) - \int_0^s (gg^\top)(t, \bar{x}_t, \bar{u}(t))dt.$$

This implies that

$$\bar{I}(g)(s) = \int_0^s (gg^\top)(t, \bar{x}_t, \bar{u}(t))dt.$$

By a martingale representation theorem (see Subsection of Chapter 1), there is an extension $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbf{F}}, \hat{P})$ of $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbf{F}}, \hat{P})$ on which lives an m-dimensional $\hat{\mathbf{F}}$ Brownian motion $\hat{W}(\cdot) = \{\hat{W}(s), s \geq 0\}$, such that

$$\bar{I}(g)(s) = \int_0^s g(t, \bar{x}_t, \bar{u}(t))\, d\hat{W}(t).$$

Similarly, one can show that

$$\bar{F}(s) = \int_0^s f(t, \bar{x}_t, \bar{u}(t))\, dt.$$

Putting into

$$\bar{x}(s) = \psi(0) + \bar{F}(s) + \bar{I}(g)(s), \quad \forall s \in [0,T], \ \bar{P} - a.s.,$$

with

$$\bar{E}\Big[\int_0^T e^{-\alpha t}(t, \bar{x}_t, \bar{u}(t))\, dt \quad + e^{-\alpha T}\Psi(\bar{x}_T)\Big] = \inf_{u(\cdot)\in\mathcal{U}[0,T]} J(0, \psi; u(\cdot)),$$

we arrive at the conclusion that

$$(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbf{F}}, \hat{P}, \hat{W}(\cdot), \bar{u}(\cdot)) \in \mathcal{U}[0,T]$$

is an optimal control. This prove the theorem. $\square$

Proof of Theorem 3.2.11**.**

The idea in proving the existence of an optimal relaxed control $\mu^*(\cdot,\cdot)$ is to (1) observe that the space of relaxed control $\mathcal{R}$ is sequentially compact, since $[0,T]\times U$ is compact (see Theorem 3.2.1), and hence every sequence in $\mathcal{R}$ has a convergent subsequence under the weak compact topology defined by (3.14); (2) check that $\hat{J}(t,\psi;\cdot):\mathcal{R}\to\Re$ is a (sequentially) upper semi-continuous function defined on the sequentially compact space $\mathcal{R}$; 3) provoke a classical theorem (see, e.g., Rudin [Rud71]) that states any (sequentially) upper semicontinuous real-valued function defined on a (sequentially) compact space attends a maximum in the space, and hence $\hat{J}(t,\psi;\cdot)$ attains its maximum at some point $\mu^*(\cdot,\cdot)\in\mathcal{R}$ (see, e.g., Yong and Zhou [YZ99, p.65]); and (4) show that the value function for the Problem (ORCP) coincides with that of the original Optimal Classical Control Problem (OCCP).

First, the following proposition is analogous to Theorem 10.1.1 of Kushner and Dupuis [KD01, pp.271-275] for our setting. The detail of the proof is very similar to that of Theorem 3.2.12 and, therefore, only a sketch is provided here.

Proposition 3.2.16 *Let Assumption 3.1.1 hold. Let*

$$\{(\Omega^{(k)},\mathcal{F}^{(k)},P^{(k)},\mathbf{F}^{(k)},W^{(k)}(\cdot),\mu^{(k)}(\cdot,\cdot))\}_{k=1}^{\infty}$$

be any sequence of admissible relaxed controls in $\hat{\mathcal{U}}[t,T]$*. For each* $k=1,2,\ldots$*, let* $\{x^{(k)}(s;t,\psi,\mu^{(k)}(\cdot,\cdot)), s\in[t-r,T]\}$ *be the corresponding strong solution of (3.16) through the initial datum* $(t,\psi^{(k)})\in[0,T]\times\mathbf{C}$*. Assume that the sequence of initial functions* $\{\psi^{(k)},k=1,2,\cdots\}$ *converges to* $\psi\in\mathbf{C}$*. The sequence*

$$\{(x^{(k)}(\cdot),W^{(k)}(\cdot),\mu^{(k)}(\cdot,\cdot)),k=1,2,\ldots\}$$

is tight. Denote by $(x(\cdot),W(\cdot),\mu(\cdot))$ *the limit point of the sequence*

$$\{(x^{(k)}(\cdot),W^{(k)}(\cdot),\mu^{(k)}(\cdot,\cdot)),k=1,2,\ldots\}$$

Define the filtration $\{\mathcal{H}(s),s\in[t,T]\}$ *by*

$$\mathcal{H}(s)=\sigma((x(\lambda),W(\lambda),\mu(\lambda,G)),t\le\lambda\le s,G\in\mathcal{B}(U)).$$

Then $W(\cdot)$ *is is an* $(\mathcal{H}(t,s),s\in[t,T])$*-adapted Brownian motion, the six-tuple*

$$\{(\Omega,\mathcal{F},P,\mathbf{F},W(\cdot),\mu(\cdot,\cdot))\}$$

is an admissible relaxed control and the process

$$x(\cdot)=\{x(s;t,\psi,\mu(\cdot,\cdot)),s\in[t-r,T]\}$$

is the strong solution process to (3.16) defined on

$$\{(\Omega,\mathcal{F},P,\mathbf{F},W(\cdot),\mu(\cdot,\cdot))\}$$

and with the initial datum $(t,\psi)\in[0,T]\times\mathbf{C}$*.*

A Sketch of Proof. Without loss of generality, we can and will assume that

$$\{(\Omega^{(k)}, \mathcal{F}^{(k)}, P^{(k)}, \mathbf{F}^{(k)}, W^{(k)}(\cdot), \mu^{(k)}(\cdot,\cdot)), k = 1, 2, \ldots\}$$

is the maximizing sequence of *admissible relaxed controls* for Problem (ORCP). We claim that the sequence of triplets

$$\{(W^{(k)}(\cdot), \mu^{(k)}(\cdot,\cdot), x^{(k)}(\cdot)), k = 1, 2, \ldots\} \tag{3.33}$$

is tight and therefore has a subsequence that is also to be denoted by (3.33), which converges weakly to some triplet $(W(\cdot), \mu(\cdot,\cdot), x(\cdot))$, where $W(\cdot)$ is a standard Brownian motion, $\mu(\cdot,\cdot)$ is the optimal relaxed control process, and $x(\cdot)$ is the optimal state process (corresponding to the optimal relaxed control process) that satisfies (3.16). Componentwise tightness implies tightness of the products (cf. Remark 3.2.4 or [Bil99,p.65]). We therefore prove the following componentwise results.

First, we observe that the sequence $\{W^{(k)}(\cdot)\}_{k=1}^{\infty}$ is tight. This is because they all have the same (Wiener) probability measure. Note that $W^{(k)}(\cdot)$ is continuous for each $k = 1, 2, \ldots$, so is its limit $W(\cdot)$. To show that $W(\cdot)$ is an m-dimensional standard Brownian motion, we will use the martingale characterization theorem in Section 1.2.1 of Chapter 1 by showing that if $\Phi \in C_0^2(\Re^m)$ (the space of real-valued twice continuously differentiable functions on $\Re^m$ and with compact support), then $M_\Phi(\cdot)$ is a $\mathbf{F}$-martingale, where

$$M_\Phi(s) \equiv \Phi(W(s)) - \Phi(0) - \int_0^s \mathcal{L}_w \Phi(W(t))\, dt, \quad s \geq 0,$$

and $\mathcal{L}_w$ is the differential operator defined by (1.8). To prove this, we have by the fact that $W^{(k)}(\cdot)$ is $\mathbf{F}^{(k)}$-Brownian motion,

$$\begin{aligned} E\Bigg[& H(x^{(k)}(t_i), W^{(k)}(t_i), \mu^{(k)}(t_i), i \leq p) \\ & \times \left(\Phi(W^{(k)}(t+\lambda)) - \Phi(W^{(k)}(t)) - \int_t^{t+\lambda} \mathcal{L}_w \Phi(W^{(k)}(s)) ds \right) \Bigg] = 0. \end{aligned} \tag{3.34}$$

By the probability 1 convergence which is implied by the Skorokhod representation theorem,

$$E\left[\left| \int_t^{t+\lambda} \mathcal{L}_w \Phi(W^{(k)}(s))\, ds - \int_t^{t+\lambda} \mathcal{L}_w \Phi(W(s))\, ds \right| \right] \to 0.$$

Using this result and taking limits in (3.34) yields

$$\begin{aligned} E\Bigg[& H(x(t_i), W(t_i), \mu(t_i), i \leq p) \\ & \times \left(\Phi(W(t+\lambda)) - \Phi(W(t)) - \int_t^{t+\lambda} \mathcal{L}_w \Phi(W(s))\, ds \right) \Bigg] = 0. \end{aligned} \tag{3.35}$$

The set of random variables $H(x(t_i), W(t_i), \mu(t_i), i \leq p)$, as $H(\cdot)$, p, and t_i vary over all possibilities, induces the σ-algebra $\mathcal{F}(t)$. Thus, (3.35) implies that

$$(\Phi(W(s) - \Phi(0) - \int_0^s \mathcal{L}_w \Phi(W(t))\, dt$$

is an **F**-martingale for all Φ of chosen class. Thus $W(\cdot)$ is a standard **F**-Brownian motion.

Second, the sequence $\{(\mu^{(k)}(\cdot,\cdot), k = 1, 2, \ldots\}$ is tight, because the space $\mathcal{R}$ is (sequentially) weak compact. Furthermore, its weak limit $\mu(\cdot) \in \mathcal{R}$ and $\mu([0,t]; U) = t$ for all $t \in [0, T]$.

Third, the tightness of the sequence of processes $\{x^{(k)}(\cdot), k = 1, 2, ...\}$ follows from the Aldous criterion (cf. Theorem 3.2.6 or [Bil99, pp. 176-179]: Given $k = 1, 2, \ldots$, any bounded **F**-stopping time τ, and $\delta > 0$, we have

$$E^{(k)}[|x^{(k)}(\tau + \delta) - x^{(k)}(\tau)|^2 \mid \mathcal{F}^{(k)}(\tau)] \leq 2K^2\delta(\delta + 1)$$

as a consequence of Assumption 3.1.1 and Itô's isometry. To show that its limit process $x(\cdot) = \{x(s), s \in [t, T]\}$ satisfies (3.16), we note that the weak limit $(x(\cdot), W(\cdot), \mu(\cdot,\cdot))$ is continuous on the time interval $[t, T]$. This is because it has been shown in the proof of Theorem 3.2.12 that both the pathwise convergence of the Lebesque integral

$$\lim_{k\to\infty} \int_t^s \int_U f(\lambda, x_\lambda^{(k)}, u)\dot{\mu}^{(k)}(\lambda, du)\, d\lambda = \int_t^s \int_U f(\lambda, x_\lambda, u)\dot{\mu}(\lambda, du)\, d\lambda, \quad P\text{-a.s.},$$

and of the stochastic integral

$$\begin{aligned} &\lim_{k\to\infty} \int_t^s \int_U g(\lambda, x_\lambda^{(k)}, u) dW^{(k)}(\lambda)\dot{\mu}^{(k)}(\lambda, du) d\lambda \\ &= \int_t^s \int_U g(\lambda, x_\lambda, u)\dot{\mu}(\lambda, du) dW(\lambda), \qquad P\text{-a.s.} \end{aligned}$$

We assume that the probability spaces are chosen as required by the Skorokhod representation (Theorem 3.2.7), so that we can suppose that the convergence of

$$\{(W^{(k)}(\cdot), \mu^{(k)}(\cdot,\cdot), x^{(k)}(\cdot))\}_{k=1}^{\infty}$$

to its limit is with probability 1 in the topology of the path spaces of the processes. Thus,

$$\int_t^s \int_U f(\lambda, x_\lambda^{(k)}, u)\dot{\mu}^{(k)}(\lambda, du)\, d\lambda \to \int_t^s \int_U f(\lambda, x_\lambda, u)\dot{\mu}(\lambda, du) d\,\lambda$$

and

$$\int_t^s \int_U g(\lambda, x_\lambda^{(k)}, u)\dot{\mu}^{(k)}(\lambda, du)\, dW^{(k)}(\lambda) \to \int_t^s \int_U g(\lambda, x_\lambda, u)\dot{\mu}(\lambda, du)\, dW(\lambda)$$

as $k \to \infty$ uniformly on $[t, T]$ with probability 1. The sequence $\{\mu^{(k)}(\cdot, \cdot)\}_{k=1}^{\infty}$ converges weakly. In particular, for any $\Phi \in C_b([0, T] \times U)$,

$$\int_t^T \int_U \Phi(\lambda, u)\mu^{(k)}(d\lambda, du) \to \int_t^T \int_U \Phi(\lambda, u)\mu(d\lambda, du).$$

Now, the Skorokhod representation theorem 3.2.7 and weak convergence imply that

$$\int_t^s \int_U f(\lambda, x_\lambda^{(k)}, u)\dot{\mu}^{(k)}(\lambda, du)\, d\lambda \to \int_t^s \int_U f(\lambda, x_\lambda, u)\dot{\mu}(\lambda, du)\, d\lambda$$

and

$$\int_t^s \int_U g(\lambda, x_\lambda^{(k)}, u)\dot{\mu}^{(k)}(\lambda, du)\, dW^{(k)}(\lambda) \to \int_t^s \int_U g(\lambda, x_\lambda, u)\dot{\mu}(\lambda, du)\, dW(\lambda)$$

as $k \to \infty$ uniformly on $[t, T]$ with probability 1. Since $\psi^{(k)} \in \mathbf{C}$ converges to $\psi \in \mathbf{C}$, we therefore prove that

$$\begin{aligned} x(s) = \psi(0) &+ \int_t^s \int_U f(\lambda, x_\lambda, u)\dot{\mu}(\lambda, du)\, d\lambda \\ &+ \int_t^s \int_U g(\lambda, x_\lambda, u)\dot{\mu}(\lambda, du)\, dW(\lambda). \quad s \in [t, T], \end{aligned} \tag{3.36}$$

We next claim that the weak limit $(x(\cdot), W(\cdot), \mu(\cdot, \cdot))$ is continuous on the time interval $[t, T]$. First, $x(\cdot)$ is a continuous process; this is because both the pathwise Lebesque integral

$$\lim_{k\to\infty} \int_t^s \int_U f(\lambda, x_\lambda^{(k)}, u)\dot{\mu}^{(k)}(\lambda, du)\, d\lambda = \int_t^s \int_U f(\lambda, x_\lambda, u)\dot{\mu}(\lambda, du)\, d\lambda, \quad P\text{-a.s.},$$

and the stochastic integral

$$\begin{aligned} &\lim_{k\to\infty} \int_t^s \int_U g(\lambda, x_\lambda^{(k)}, u)\dot{\mu}^{(k)}(\lambda, du)\, dW^{(k)}(\lambda) \\ &= \int_t^s \int_U g(\lambda, x_\lambda, u)\dot{\mu}(\lambda, du)\, dW(\lambda), \qquad P\text{-a.s.} \end{aligned}$$

Similarly,

$$\int_t^T \int_U e^{\alpha(s-t)} L(s, x_s, u)\dot{\mu}^{(k)}(s, du)\, ds \to \int_t^T \int_U e^{\alpha(s-t)} L(s, x_s, u)\dot{\mu}(s, du)\, ds,$$

and

$$e^{-\alpha(T-t)}\Psi(x_T^{(k)}) \to e^{-\alpha(T-t)}\Psi(x_T)$$

as $k \to \infty$ with probability 1. $\square$

We have therefore proved the following two propositions.

Proposition 3.2.17 *Let Assumption 3.1.1 hold. Suppose the sequence of initial segment functions* $\{\psi^{(k)}\}_{k=1}^{\infty} \subset \mathbf{C}$ *converges to* $\psi \in \mathbf{C}$*. Then*

$$\lim_{k\to\infty} \hat{V}(t, \psi(k)) = \hat{V}(t, \psi).$$

Proposition 3.2.18 *Let Assumption 3.1.1 hold. Let* $\mu(\cdot,\cdot) \in \hat{U}[t,T]$ *be the relaxed control representation of* $u(\cdot) \in \mathcal{U}[t,T]$ *via (3.13). Then*

$$V(t,\psi) := \sup_{u(\cdot)\in\mathcal{U}[0,T]} J(t,\psi;u(\cdot)) = \hat{V}(t,\psi) := \sup_{\mu(\cdot,\cdot)\in\mathcal{U}[0,T]} \hat{J}(t,\psi;\mu(\cdot,\cdot)).$$

Proof of Theorem 3.2.11 The theorem follows immediately from Propositions 3.2.17 and 3.2.18. □

3.3 Dynamic Programming Principle

3.3.1 Some Probabilistic Results

To establish and prove the dynamics programming principle (DDP), we need some probabilistic results as follows.

First, we recall that if O is a nonempty set and if $\mathcal{O}$ is a collection of subsets of O, the collection $\mathcal{O}$ is called a π-system if it is closed under the finite intersection; that is, $A, B \in \mathcal{O}$ imply that $A \cap B \in \mathcal{O}$. It is a λ-system if the following three conditions are satisfied: (i) $O \in \mathcal{O}$; (ii) $A, B \in \mathcal{O}$ and $A \subset B$ imply that $B - A \in \mathcal{O}$; and (iii) $A_i \in \mathcal{O}$, $A_i \uparrow A, i = 1, 2, \ldots$, implies that $A \in \mathcal{O}$.

The following lemmas will be used later.

Lemma 3.3.1 *Let* $\mathcal{O}$ *and* $\tilde{\mathcal{O}}$ *be two collections of subsets of* O *with* $\mathcal{O} \subset \tilde{\mathcal{O}}$*. Suppose* $\mathcal{O}$ *is a* π*-system and* $\tilde{\mathcal{O}}$ *is a* λ*-system. Then* $\sigma(\mathcal{O}) \subset \tilde{\mathcal{O}}$*, where* $\sigma(\mathcal{O})$ *is the smallest* σ*-algebra containing* $\mathcal{O}$*.*

Proof. This is the well-known *monotone class theorem*, the proof of which can be found in Lemma 1.1.2 of [YZ99].

Lemma 3.3.2 *Let* $\mathcal{O}$ *be a* π*-system on* O*. Let* $\mathcal{H}$ *be a linear space of functions from* O *to* $\Re$ *such that*

$$\mathbf{1} \in \mathcal{H}; \quad I_A \in \mathcal{H}, \quad \forall A \in \mathcal{O},$$

and

$$\varphi^{(i)} \in \mathcal{H} \text{ with } 0 \le \varphi^{(i)} \uparrow \varphi, \varphi \text{ is finite} \Rightarrow \varphi \in \mathcal{H}.$$

Then $\mathcal{H}$ *contains all* $\sigma(\mathcal{O})$*-measurable functions from* O *to* $\Re$*, where* $\mathbf{1}$ *is the constant function of value* 1 *and* I_A *is the indicator function of* A*.*

Proof. Let

$$\tilde{\mathcal{O}} = \{A \subset O \mid I_A \in \mathcal{H}\}.$$

Then $\tilde{\mathcal{O}}$ is a λ-system containing $\mathcal{O}$. From Lemma 3.3.1 it can be shown that $\sigma(\mathcal{O}) \subset \tilde{\mathcal{O}}$.

Now, for any $\sigma(\mathcal{O})$-measurable function $\varphi : O \to \Re$, we set for $i = 1, 2, \ldots$

$$\varphi^{(i)} = \sum_{j \geq 0} j2^{-i} I_{\{j2^{-i} \leq \varphi^+(\omega) < (j+1)2^{-i}\}},$$

where $\varphi^+ \equiv \max\{\varphi, 0\}$ denotes the positive part of φ. Clearly, $\varphi^{(i)} \in \mathcal{H}$ and $0 \leq \varphi^{(i)} \uparrow \varphi^+$. Hence, by our assumption, $\varphi^+ \in \mathcal{H}$. Similarly, $\varphi^- \in \mathcal{H}$, where $\varphi^+ \equiv \max\{-\varphi, 0\}$ denotes the negative part of φ. Thus, $\varphi \in \mathcal{H}$. This proves the lemma. $\Box$

First, let us introduce some notation. Define

$$\begin{aligned}
&C_t([0,T];\Re^m) := \{\eta(\cdot \wedge t) \mid \eta \in C([0,T];\Re^m)\}, \\
&\mathcal{B}_t(C([0,T];\Re^m)) := \sigma(\mathcal{B}(C_t([0,T];\Re^m))), \\
&\mathcal{B}_{t+}(C([0,T];\Re^m)) := \cap_{s>t}\mathcal{B}_s(C([0,T];\Re^m))), \ t \in [0,T],
\end{aligned}$$

where $\sigma(\mathcal{B}(C_t([0,T];\Re^m)))$ denotes the smallest σ-algebra in $C([0,T];\Re^m)$ that contains $\mathcal{B}(C_t([0,T];\Re^m))$. Clearly, both of the following are filtered measurable spaces:

$$(C([0,T];\Re^m), \mathcal{B}(C([0,T];\Re^m)), \{\mathcal{B}_t(C([0,T];\Re^m))\}_{t\geq 0})$$

and

$$(C([0,T];\Re^m), \mathcal{B}(C([0,T];\Re^m)), \{\mathcal{B}_{t+}(C([0,T];\Re^m))\}_{t\geq 0}).$$

However,

$$\mathcal{B}_{t+}(C([0,T];\Re^m)) \neq \mathcal{B}_t(C([0,T];\Re^m)), \quad \forall t \in [0,T].$$

A set $B \subset C([0,T];\Re^m)$ is called a *Borel cylinder* if there exists a partition $\pi = \{0 \leq t_1 < t_2 < \cdots < t_j \leq T\}$ of $[0,T]$ and $A \in \mathcal{B}((\Re^m)^j)$ such that

$$B = \pi^{-1}(A) \equiv \{\xi \in C([0,T];\Re^m) \mid (\xi(t_1), \xi(t_2), \ldots, \xi(t_j)) \in A\}. \tag{3.37}$$

For $s \in [0,T]$, let $\mathbf{C}(s)$ be the set of all Borel cylinder in $C_s([0,T];\Re^m)$ of the form (3.37) with partition $\pi \subset [0,s]$.

Lemma 3.3.3 *The σ-algebra $\sigma(\mathbf{C}(T))$ generated by $\mathbf{C}(T)$ coincides with the Borel σ-algebra $\mathcal{B}(C[0,T];\Re^m))$ of $C([0,T];\Re^m)$.*

Proof. Let the partition $\pi = \{0 \leq t_1 < t_2 < \cdots < t_j \leq T\}$ of $[0,T]$ be given. We define a point-mass projection map $\Pi : C([0,T];\Re^m) \to (\Re^m)^j$ associated with the partition π as follows:

$$\Pi(\xi) = (\xi(t_1), \xi(t_2), \ldots, \xi(t_j)), \quad \forall \xi \in C([0,T];\Re^m).$$

Clearly, Π is continuous. Consequently, for any $A \in \mathcal{B}((\Re^m)^j)$, it follows that $\Pi^{-1}(A) \in \mathcal{B}(C([0,T];\Re^m))$. This implies that

$$\sigma(\mathbf{C}(T)) \subset \mathcal{B}((C([0,T];\Re^m)). \tag{3.38}$$

Next, for any $\hat{\xi} \in C([0,T];\Re^m)$ and $\epsilon > 0$, the closed ϵ-ball $B(\hat{\xi};\epsilon)$ in $C([0,T];\Re^n)$ can be written as

$$\begin{aligned} B(\hat{\xi};\epsilon) &\equiv \{\xi \in C([0,T];\Re^m) \mid \sup_{t\in[0,T]} |\xi(t) - \hat{\xi}(t)| \leq \epsilon\} \qquad (3.39) \\ &= \bigcap_{t\in\mathbf{Q}\cap[0,T]} \{\xi \in C([0,T];\Re^m) \mid |\xi(t) - \hat{\xi}(t)| \leq \epsilon\} \in \sigma(\mathbf{C}(T)), \end{aligned}$$

since $\{\xi \in C([0,T];\Re^m) \mid |\xi(t) - \hat{\xi}(t)| \leq \epsilon\}$ is a Borel cylinder and $\mathbf{Q}$ is the set of all rational numbers (which is countable). Because the set of all sets in the form of the left-hand side of (3.39) is a basis of the closed (and therefore open) sets in $C([0,T];\Re^m)$, we have

$$\mathcal{B}(C([0,T];\Re^m)) \subset \sigma(\mathbf{C}(T)). \tag{3.40}$$

Combining (3.38) and (3.40), we obtain the conclusion of the lemma. □

Lemma 3.3.4 *Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $\zeta : [0,T]\times\Omega \to \Re^m$ be a continuous process. Then there exists an $\Omega_0 \in \mathcal{F}$ with $P(\Omega_0) = 1$ such that $\zeta : \Omega_0 \to C([0,T];\Re^m)$, and for any $s \in [0,T]$,*

$$\Omega_0 \bigcap \mathcal{F}^\zeta(s) = \Omega_0 \bigcap \zeta^{-1}(\mathcal{B}_s(C([0,T];\Re^m)), \tag{3.41}$$

where $\mathbf{F}^\zeta = \{\mathcal{F}^\zeta(s), s \in [0,T]\}$ is the filtration of sub-σ-algebras generated by the process $\zeta(\cdot)$; that is, for all $s \in [0,T]$,

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

Proof. Let $t \in [0,s]$ and $A \in \mathcal{B}(\Re^m)$ be fixed. Then

$$B(t) \equiv \{\xi \in C([0,T];\Re^m) \mid \xi(t) \in A\} \in \mathbf{C}(s)$$

and

$$\begin{aligned} \omega \in \zeta^{-1}(B(t)) &\Longleftrightarrow \zeta(\cdot,\omega) \in B(t) \Longleftrightarrow \zeta(t,\omega) \in A \\ &\Longleftrightarrow \omega \in \zeta^{-1}(t,\cdot)(A). \end{aligned}$$

Thus,

$$\zeta^{-1}(t,\cdot)(A) = \zeta^{-1}(B(t)).$$

Then by Lemma 3.3.3, we obtain (3.41). □

Lemma 3.3.5 *Let* $(\Omega, \mathcal{F})$ *and* $(\tilde{\Omega}, \tilde{\mathcal{F}})$ *be two measurable spaces and let* (Ξ, d) *be a Polish (complete and separable) metric space. Let* $\zeta : \Omega \to \tilde{\Omega}$ *and* $\varphi : \Omega \to \Xi$ *be two random variables. Then* φ *is* $\sigma(\zeta)$*-measurable; that is,*

$$\varphi^{-1}(\mathcal{B}(\Xi)) \subset \zeta^{-1}(\tilde{\mathcal{F}}) \tag{3.42}$$

if and only if there exists a measurable map $\eta : \tilde{\Omega} \to \Xi$ *such that*

$$\varphi(\omega) = \eta(\zeta(\omega)), \quad \forall \omega \in \Omega. \tag{3.43}$$

Proof. We only need to prove the necessity. First, we assume that $\Xi = \Re$. For this case, set

$$\mathcal{H} \equiv \{\eta(\zeta(\cdot)) \mid \eta : \tilde{\Omega} \to \Xi \text{ is measurable}\}.$$

Then $\mathcal{H}$ is a linear space and $\mathbf{1} \in \mathcal{H}$. We note here that $\zeta^{-1}(\tilde{\mathcal{F}})$ forms a π-system; that is, it is closed under finite intersections. Also, if $A \in \sigma(\zeta) \equiv \zeta^{-1}(\tilde{\mathcal{F}})$, then for some $B \in \tilde{\mathcal{F}}$, $I_A(\cdot) = I_B(\zeta(\cdot)) \in \mathcal{H}$. Now, suppose $\eta^{(i)} : \tilde{\Omega} \to \Xi$ is measurable for $i = 1, 2, \ldots$ and $\eta^{(i)}(\zeta(\cdot)) \in \mathcal{H}$ such that $0 \leq \eta^{(i)}(\zeta(\cdot)) \uparrow \xi(\cdot)$, which is finite. Set

$$A = \{\tilde{\omega} \in \tilde{\Omega} \mid \sup_i \eta^{(i)}(\tilde{\omega}) < \infty\}.$$

Then $A \in \tilde{\mathcal{F}}$ and $\zeta(\Omega) \subset A$. Define

$$\eta(\tilde{\omega}) = \begin{cases} \sup_i \eta^{(i)}(\tilde{\omega}) & \text{if } \tilde{\omega} \in A \\ 0 & \text{if } \tilde{\omega} \in \tilde{\Omega} - A. \end{cases}$$

Clearly, $\eta : \tilde{\Omega} \to \Xi$ is measurable and $\xi(\cdot) = \eta(\zeta(\cdot))$. Thus, $\xi(\cdot) \in \mathcal{H}$. By Lemma 3.3.2, $\mathcal{H}$ contains all $\sigma(\zeta)$-measurable random variables, in particular, $\varphi \in \mathcal{H}$, which lead to (3.43). This proves our conclusion for the case $U = \Re$.

Now, let (Ξ, d) be an uncountable Polish space. Then it is known that (Ξ, d) is Borel isomorphic to the Borel measurable space of real numbers $(\Re, \mathcal{B}(\Re))$; that is, there exists a bijection $h : \Xi \to \Re$ such that $h(\mathcal{B}(\Xi)) = \mathcal{B}(\Re)$. Consider the map $\tilde{\varphi} = h \circ \varphi : \Omega \to \Re$, which satisfies

$$\tilde{\varphi}^{-1}(\mathcal{B}(\Re)) = \varphi^{-1} \circ h^{-1}(\mathcal{B}(\Re)) = \varphi^{-1}(\mathcal{B}(\Xi)) \subset \zeta^{-1}(\tilde{\mathcal{F}}).$$

Thus, there exists an $\tilde{\eta} : \tilde{\Omega} \to \Re$ such that

$$\tilde{\varphi}(\omega) = \tilde{\eta}(\zeta(\omega)), \quad \forall \omega \in \Omega.$$

By taking $\eta = h^{-1} \circ \tilde{\eta}$, we obtain the desired result.

Finally, if (Ξ, d) is countable or finite, we can prove the result by replacing $\Re$ in the above by the set of natural numbers $\mathbf{N}$ or $\{1, 2, \ldots, n\}$. $\square$

Later we will take Ξ to be the control set U. Let $\mathcal{A}_T^m(U)$ be the set of all $\mathcal{B}_{t+}(C([0,T];\Re^m))$-progressively measurable processes

$$\eta : [0,T] \times C([0,T];\Re^m) \to U,$$

where U is the control set, which is assumed to be a Polish (complete and separable) metric space.

Proposition 3.3.6 *Let $(\Omega, \mathcal{F}, P)$ be a complete probability space and let U be a Polish space. Let $\zeta : [0,T] \times \Omega \to \Re^m$ be a continuous process and $\mathcal{F}^\zeta(s) = \sigma(\zeta(t); 0 \le t \le s)$. Then $\varphi : [0,T] \times \Omega \to U$ is $\{\mathcal{F}^\zeta(s)\}$-adapted if and only if there exists an $\eta \in \mathcal{A}_T^m(U)$ such that*

$$\varphi(t,\omega) = \eta(t, \zeta(\cdot \wedge t, \omega)),\ t \in [0,T],\ P-a.s.\ \omega \in \Omega.$$

Proof. We prove only the "only if" part. The "if" part is clear.

For any $s \in [0,T]$, we consider a mapping

$$\theta^s(t,\omega) \equiv (t \wedge s, \zeta(\cdot \wedge s, \omega) : [0,T] \times \Omega \to [0,s] \times C_t([0,T];\Re^m).$$

By Lemma 3.3.4, we have $\mathcal{B}([0,s]) \times \mathcal{F}^\zeta(s) = \sigma(\theta^s)$. On the other hand, $(t,\omega) \mapsto \varphi(t \wedge s, \omega)$ is $(\mathcal{B}([0,s]) \times \mathcal{F}^\zeta(s))/\mathcal{B}(U)$-measurable. Thus, by Lemma 3.3.5, there exists a measurable map

$$\eta^s : ([0,T] \times C_s([0,T];\Re^m), \mathcal{B}([0,s]) \times \mathcal{B}_s(C([0,T];\Re^m)) \to U$$

such that

$$\varphi(t \wedge s, \omega) = \eta^s(t \wedge s, \zeta(\cdot \wedge s, \omega)), \quad \forall \omega \in \Omega, t \in [0,T]. \tag{3.44}$$

Now, for any $i \ge 1$, let $0 = t_0^i < t_1^i < \cdots$ be a partition of $[0,T]$ (with the mesh $\max_{j \ge 1} |t_j^i - t_{j-1}^i| \to 0$ as $i \to \infty$) and define

$$\begin{aligned} \eta^{(i)}(t,\xi) &= \eta^0(0, \xi(\cdot \wedge 0)) I_{\{0\}}(t) \\ &\quad + \sum_{j \ge 1} \eta^{t_j^i}(t, \xi(\cdot \wedge t_j^i)) I_{(t_{j-1}^i, t_j^i]}(t), \quad \forall (t,\xi) \in [0,T] \times C([0,T];\Re^m). \end{aligned} \tag{3.45}$$

For any $t \in [0,T]$, there exists j such that $t_{j-1}^i < t \le t_j^i$. Then

$$\eta^{(i)}(t, \zeta(\cdot \wedge t_j^i, \omega)) = \eta^{t_j^i}(t, \zeta(\cdot \wedge t_j^i, \omega)) = \varphi(t,\omega). \tag{3.46}$$

Now, in the case $U = \Re$, $\mathbf{N}$, $\{1, 2, \ldots, n\}$, we may define

$$\eta(t,\xi) = \overline{\lim}_{i \to \infty} \eta^{(i)}(t,\xi) \tag{3.47}$$

to get the desired result. In the case where U is a general Polish space, we need to amend the proof in the same fasion as in that of Lemma 3.3.5. □

3.3.2 Continuity of the Value Function

For each $t \in [0, T]$, the following lemma shows that the value function $V(t, \cdot) : \mathbf{C} \to \Re$ is Lipschitz.

Lemma 3.3.7 *Assume Assumptions (A3.1.1)-(A3.1.3) hold. The value function V satisfies the following properties: There is a constant $K_V \geq 0$ not greater than $3K_{lip}(T+1)e^{3T(T+4m)K_{lip}^2}$ such that for all $t \in [0, T]$ and $\phi, \tilde{\phi} \in \mathbf{C}$, we have*

$$|V(t, \phi) - V(t, \varphi)| \leq K_V \|\phi - \varphi\|.$$

Proof. Let $t \in [0, T]$ and $\phi, \varphi \in \mathbf{C}$. We have, by definition,

$$|V(t, \phi) - V(t, \varphi)| \leq \sup_{u(\cdot) \in \mathcal{U}[t,T]} |J(t, \phi; u(\cdot)) - J(t, \varphi; u(\cdot))|.$$

Let $x(\cdot)$ and $y(\cdot)$ be the solution processes of (3.1) under the control process $u(\cdot)$ but with two different initial data: (t, ϕ) and $(t, \varphi) \in [0, T] \times \mathbf{C}$, respectively. Then by (A3.1.3) of Assumption 3.1.1, we have for all $u(\cdot) \in \mathcal{U}[t, T]$,

$$\begin{aligned}
&|J(t, \phi; u(\cdot) - J(t, \varphi; u(\cdot))| \\
&\leq E\left[\int_t^T |L(s, x_s, u(s)) - L(s, y_s, u(s))|\, ds + |\Psi(x_T) - \Psi(y_T)|\right] \\
&\leq K_{kip}(1 + T - t) E\left[\sup_{s \in [-r,T]} |x(s) - y(s)|^2\right].
\end{aligned}$$

Now,

$$E\left[\sup_{s \in [-r,T]} |x(s) - y(s)|^2\right] \leq 2E\left[\sup_{s \in [0,T]} |x(s) - y(s)|^2\right] + 2\|\phi - \varphi\|^2,$$

while Hölder's inequality, Doob's maximal inequality, Itô's isometry, Fubini's theorem and (A3.1.3) of Assumption 3.1.1 together yield

$$\begin{aligned}
&E\left[\sup_{s \in [0,T]} |x(s) - y(s)|^2\right] \\
&\leq 3|\phi(0) - \varphi(0)|^2 + 3TE\left[\int_t^T |f(s, x_s, u(s)) - f(s, y_s, u(s))|^2\, ds\right] \\
&\quad + 3m \sum_{i=1}^n \sum_{j=1}^m E\left[\int_t^T |(g_{ij}(s, x_s, u(s)) - g_{ij}(s, y_s, u(s)))\, dW_j(s)|^2\right] \\
&\leq 3|\phi(0) - \varphi(0)|^2 + 3TK_{lip}^2 \int_t^T E[\|x_s - y_s\|^2\, ds
\end{aligned}$$

$$+ 12mE\left[\int_t^T \sum_{i=1}^n \sum_{j=1}^m |(g_{ij}(s,x_s,u(s)) - g_{ij}(s,y_s,u(s)))|^2\, ds\right]$$

$$\le 3|\phi(0)-\varphi(0)|^2 + 3(T+4m)K_{lip}^2 \int_t^T E\left[\sup_{\lambda\in[t-r,s]} \|x(\lambda)-y(\lambda)\|^2\right] ds.$$

Since $|\phi(0)-\varphi(0)| \le \|\phi-\varphi\|$, Gronwall's lemma gives

$$E\left[\sup_{s\in[t-r,T]} \|x(s)-y(s)\|^2\right] \le 8\|\phi-\varphi\|^2 e^{6T(T+4m)K_{lip}^2}.$$

Combining the above estimates, we obtain the assertion. □

3.3.3 The DDP

The advantage of the weak formulation of the control problem will be apparent in the following lemmas and its use in the proof of the DPP. Let $t \in [0,T]$ and $u(\cdot) \in \mathcal{U}[t,T]$. Then under Assumption 3.1.1, for any $\mathbf{F}(t)$-stopping time $\tau \in [t,T)$ and $\mathcal{F}(t,\tau)$-measurable random variable $\xi : \Omega \to \mathbf{C}$, we can solve the following controlled SHDE:

$$dx(s) = f(s,x_s,u(s))\,ds + g(s,x_s,u(s))\,dW(s),\ s\in[\tau,T], \tag{3.48}$$

with the initial function $x_\tau = \xi$ at the stopping time τ.

Lemma 3.3.8 *Let $t\in[0,T]$ and $(\Omega,\mathcal{F},P,\mathbf{F},W(\cdot),u(\cdot)) \in \mathcal{U}[t,T]$. Then for any $\mathbf{F}$-stopping time, $\tau\in[t,T)$, and any $\mathcal{F}(\tau)$- measurable random variable $\xi:\Omega\to\mathbf{C}$,*

$$J(\tau,\xi(\omega);u(\cdot)) = E\Big[\int_\tau^T e^{\alpha(s-\tau)} L(s,x_s(\cdot;\tau,\xi,u(\cdot)),u(s))ds$$
$$+ e^{-\alpha(T-\tau)}\Psi(x_T(\cdot;\tau,\xi,u(\cdot))|\mathcal{F}(\tau)\Big](\omega) \quad P\text{-a.s. } \omega.$$

Proof. Since $u(\cdot)$ is $\mathbf{F}$-adapted, where $\mathbf{F}$ is the P-augmented natural filtration generated by $W(\cdot)$, by Propostion 3.3.6 there is a function $\eta\in\mathcal{A}_T^m(U)$ such that

$$u(s,\omega) = \eta(s,W(\cdot\wedge s,\omega)),\ P-a.s.\ \omega\in\Omega,\ \forall s\in[t,T].$$

Then (3.48) can be written as

$$dx(s) = f(s,x_s,\eta(s,W(\cdot\wedge s)))\,ds$$
$$+g(s,x_s,\eta(s,W(\cdot\wedge s)))\,dW(s),\ s\in[\tau,T], \tag{3.49}$$

with $x_\tau=\xi$. Due to Assumption 3.1.1, Theorem 1.3.12, and Theorem 1.3.16, this equation has a strongly unique strong solution and, therefore, weak uniqueness holds. In addition, we may write, for $s\ge\tau$,

$$u(s,\omega) = \eta(s, W(\cdot \wedge s, \omega)) = \eta(s, \tilde{W}(\cdot \wedge s, \omega) + W(\tau, \omega))$$

where $\tilde{W}(s) = W(s) - W(\tau)$. Since τ is random, $\tilde{W}(\cdot)$ is not a Brownian motion under the probability measure P. However, we may, under the weak formulation of the control problem, change the probability measure P as follows. Note first that

$$\begin{aligned}
P\{\omega' \mid \tau(\omega') = \tau(\omega)|\mathcal{F}(\tau)\}(\omega) \\
&= E[\mathbf{1}_{\{\omega' : \tau(\omega') = \tau(\omega)\}}|\mathcal{F}(\tau)](\omega) \\
&= \mathbf{1}_{\{\omega' : \tau(\omega') = \tau(\omega)\}}(\omega) \\
&= 1, \; P - a.s. \; \omega \in \Omega.
\end{aligned}$$

This means that there is an $\Omega_0 \in \mathcal{F}$ with $P(\Omega_0) = 1$, so that for any fixed $\omega_0 \in \Omega_0$, τ becomes a deterministic time $\tau(\omega_0)$; that is, $\tau = \tau(\omega_0)$ almost surely in the new probability space $(\Omega, \mathcal{F}, P(\cdot|\mathcal{F}(\tau)))$, where $P(\cdot|\mathcal{F}(\tau)))$ denotes the probability measure P restricted to the σ-sub-algebra $\mathcal{F}(\tau)$. A similar argument shows that $W(\tau)$ almost surely equals a constant $W(\tau(\omega_0), \omega)$ and also that ξ almost surely equals a constant $\xi(\omega_0)$ when we work in the probability space $(\Omega, \mathcal{F}, P(\cdot|\mathcal{F}(\tau))(\omega_0))$. So, under the measure $P(\cdot|\mathcal{F}(\tau))(\omega_0)$, for $s \geq \tau(\omega_0)$, the process $\tilde{W}(\cdot)$ will be a standard Brownian motion

$$\tilde{W}(s) = W(s) - W(\tau(\omega_0)),$$

and for any $s \geq \tau(\omega_0)$,

$$u(s,\omega) = \eta(s, \tilde{W}(\cdot \wedge s, \omega) + W(\tau(\omega_0), \omega)).$$

It follows then that $u(\cdot)$ is adapted to the filtration $\mathbf{F}(\tau(\omega_0))$ generated by the standard Brownian motion $\tilde{W}(s)$ for $s \geq \tau(\omega_0)$. Hence, by the definition of admissible controls,

$$(\Omega, \mathcal{F}, P(\cdot|\mathcal{F}(\tau))(\omega_0), \tilde{W}(\cdot), u|_{[\tau(\omega_0),T]}) \in \mathcal{U}[\tau(\omega_0), T].$$

Note that for $A \in \mathcal{B}(\mathbf{C})$,

$$P[\xi \in A|\mathcal{F}(\tau)](\omega_0) = E[\mathbf{1}_{\{\xi \in A\}}|\mathcal{F}(\tau)](\omega_0) = E[\mathbf{1}_{\{\xi \in A\}}] = P\{\xi \in A\}.$$

This means that the two weak solutions

$$(\Omega, \mathcal{F}, P, \mathbf{F}, x(\cdot), W(\cdot)) \text{ and } (\Omega, \mathcal{F}, P(\cdot|\mathcal{F}(\tau)(\omega_0)), \tilde{\mathbf{F}}, x(\cdot), \tilde{W}(\cdot))$$

of (3.1) have the same initial distribution. Then, by the weak uniqueness,

$$\begin{aligned}
J(\tau, \xi(\omega); u(\cdot)) &= E^{\tau, \xi(\omega), u(\cdot)}\Big[\int_\tau^T e^{-\alpha(s-\tau)} L(s, x_s, u(s))\, ds + e^{-\alpha(T-\tau)} \Psi(x_T)\Big] \\
&= E\Big[\int_\tau^T e^{-\alpha(s-\tau)} L(s, x_s(\tau, \xi, u(\cdot)), u(s))\, ds \\
&\quad + e^{-\alpha(T-\tau)} \Psi(x_T(\tau, \xi, u(\cdot))\Big|\mathcal{F}(\tau)\Big](\omega), \; P - a.s. \; \omega. \qquad \square
\end{aligned}$$

The following dynamic programming principle (DDP) is due to Larssen [Lar02].

Theorem 3.3.9 (Dynamic Programming Principle) *Let Assumption 3.1.1 hold. Then for any initial datum* $(t, \psi) \in [0, T] \times \mathbf{C}$ *and* $\mathbf{F}$*-stopping time* $\tau \in [t, T]$,

$$V(t, \psi) = \sup_{u(\cdot) \in \mathcal{U}[t,T]} E\Big[\int_t^\tau e^{-\alpha(s-t)} L(s, x_s(\cdot; t, \psi, u(\cdot)), u(s))\, ds \\ + e^{-\alpha(\tau-t)} V(\tau, x_\tau(\cdot; t, \psi, u(\cdot)))\Big]. \tag{3.50}$$

Proof. Denote the right-hand side of (3.50) by $\bar{V}(t, \psi)$. Given any $\epsilon > 0$, there exists an $(\Omega, \mathcal{F}, P, \mathbf{F}, W(\cdot), u(\cdot)) \in \mathcal{U}[t, T]$ such that

$$V(t, \psi) - \epsilon \le J(t, \psi; u(\cdot)).$$

Equivalently,

$$\begin{aligned} V(t, \psi) - \epsilon &\le J(t, \psi; u(\cdot)) \\ &= E\Big[\int_t^T e^{-\alpha(s-t)} L(s, x_s(\cdot; t, \psi, u(\cdot)), u(s))\, ds \\ &\quad + e^{-\alpha(T-t)} \Psi(x_T(\cdot; t, \psi, u(\cdot)))\Big] \\ &= E\Big[\int_t^\tau e^{-\alpha(s-t)} L(s, x_s(\cdot; t, \psi, u(\cdot)), u(s))\, ds \\ &\quad + \int_\tau^T e^{-\alpha(s-t)} L(s, x_s(\cdot; t, \psi, u(\cdot)), u(s))\, ds \\ &\quad + e^{-\alpha(T-t)} \Psi(x_T(\cdot; t, \psi, u(\cdot)))\Big]. \end{aligned}$$

Therefore,

$$\begin{aligned} V(t, \psi) - \epsilon \le\ & E\Big[\int_t^\tau e^{-\alpha(s-t)} L(s, x_s(\cdot; t, \psi, u(\cdot)), u(s))\, ds \\ & + E\Big[\int_\tau^T e^{-\alpha(s-\tau)} L(s, x_s(\cdot; t, \psi, u(\cdot)), u(s))\, ds \\ & + e^{-\alpha(T-t)} \Psi(x_T(\cdot; t, \psi, u(\cdot)))\Big| \mathcal{F}(\tau)\Big]\Big] \\ =\ & E\Big[\int_t^\tau e^{-\alpha(s-t)} L(s, x_s(\cdot; t, \psi, u(\cdot)), u(s))\, ds \\ & + e^{-\alpha(\tau-t)} J(\tau, x_\tau(\cdot; t, \psi, u(\cdot)); u(\cdot))\Big] \text{ (by Lemma 3.3.11)} \\ \le\ & E\Big[\int_t^\tau e^{-\alpha(s-t)} L(s, x_s(\cdot; t, \psi, u(\cdot)), u(s))\, ds \\ & + e^{-\alpha(\tau-t)} V(\tau, x_\tau(\cdot; t, \psi, u(\cdot))\Big], \end{aligned}$$

so by taking the supremum over $u(\cdot) \in \mathcal{U}[t,T]$, we have

$$V(t,\psi) - \epsilon \leq \bar{V}(t,\psi), \quad \forall \epsilon > 0.$$

This shows that

$$V(t,\psi) \leq \bar{V}(t,\psi), \quad \forall (t,\psi) \in [0,T] \times \mathbf{C}. \tag{3.51}$$

Conversely, we want to show that

$$V(t,\psi) \geq \bar{V}(t,\psi), \quad \forall (t,\psi) \in [0,T] \times \mathbf{C}.$$

Let $\epsilon > 0$; by Lemma 3.3.7 and its proof, there is a $\tilde{\delta} = \tilde{\delta}(\epsilon)$ such that whenever

$$\|\psi - \hat{\psi}\| < \tilde{\delta},$$

$$|J(\tau,\psi;u(\cdot)) - J(\tau,\hat{\psi};u(\cdot))| + |V(\tau,\psi) - V(\tau,\hat{\psi})| \leq \epsilon, \ \forall u(\cdot) \in \mathcal{U}[\tau,T]. \tag{3.52}$$

Now, let $\{D_j\}_{j\geq 1}$ be a *Borel partition* of $\mathbf{C}$. This means that

$$D_j \in \mathcal{B}(\mathbf{C}) \text{ for each } j, \ \cup_{j\geq 1} D_j = \mathbf{C}, \text{ and } D_i \cap D_j = \text{ if } i \neq j.$$

We also assume that the D_j are chosen so that $\|\phi - \varphi\| < \tilde{\delta}$ whenever ϕ and φ are both in D_j. Choose $\psi^{(j)} \in D_j$. For each j, there exists an $(\Omega^{(j)}, \mathcal{F}^{(j)}, P^{(j)}, W^{(j)}(\cdot), u^{(j)}(\cdot)) \in \mathcal{U}[\tau,T]$ such that

$$J(\tau,\psi^{(j)};u^{(j)}(\cdot)) \leq V(\tau,\psi^{(j)}) + \epsilon. \tag{3.53}$$

For any $\psi \in D_j$, (3.52) implies in particular that

$$J(\tau,\psi;u^{(j)}(\cdot)) \geq J(\tau,\psi^{(j)};u^{(j)}(\cdot)) - \epsilon \text{ and } V(\tau,\psi^{(j)}) \geq V(\tau,\psi) - \epsilon. \tag{3.54}$$

Combining the above two inequalities, we see that

$$J(\tau,\psi;u^{(j)}(\cdot)) \geq J(\tau,\psi^{(j)};u^{(j)}(\cdot)) - \epsilon \geq V(\tau;\psi^{(j)}) - 2\epsilon \geq V(\tau,\psi) - 3\epsilon. \tag{3.55}$$

By the definition of the five-tuple

$$(\Omega^{(j)}, \mathcal{F}^{(j)}, P^{(j)}, \mathbf{F}^{(j)}, W^{(j)}(\cdot), u^{(j)}(\cdot)) \in \mathcal{U}[\tau,T],$$

there is a function $\varphi_j \in \mathcal{A}_T(U)$ such that

$$u^{(j)}(s,\omega) = \varphi^{(j)}(s, W^{(j)}(\cdot \wedge s, \omega)), \quad P^{(j)} - a.s.\ \omega \in \Omega^{(j)}, \ \forall s \in [\tau,T].$$

Now, let $(\Omega, \mathcal{F}, P, \mathbf{F}, W(\cdot), u(\cdot)) \in \mathcal{U}[t,T]$ be arbitrary. Define the new control $\tilde{u}(\cdot) = \{\tilde{u}(s), s \in [t,T]\}$, where

$$\tilde{u}(s,\omega) = u(s,\omega), \text{if } s \in [t,\tau),$$

and

$$\tilde{u}(s,\omega) = \varphi^{(j)}(s, W^{(j)}(\cdot \wedge s, \omega) - W(\tau,\omega)) \text{ if } s \in [\tau, T] \text{ and } x_s(t,\psi,u(\cdot)) \in D_j.$$

Then $(\Omega, \mathcal{F}, P, \mathbf{F}, W(\cdot), \tilde{u}(\cdot)) \in \mathcal{U}[t,T]$. Thus,

$$\begin{aligned}
V(t,\psi) &\geq J(t,\psi;\tilde{u}(\cdot)) \\
&= E\bigg[\int_t^T e^{-\alpha(s-t)} L(s, x_s(\cdot;t,\psi,\tilde{u}(\cdot)), \tilde{u}(s))\,ds \\
&\quad + e^{-\alpha(T-t)} \Psi(x_T(\cdot;t,\psi,\tilde{u}(\cdot)))\bigg] \\
&\geq E\bigg[\int_t^\tau e^{-\alpha(s-t)} L(s, x_s(\cdot;t,\psi,u(\cdot)), u(s))\,ds \\
&\quad + E\bigg[\int_\tau^T e^{-\alpha(s-\tau)} L(s, x_s(\cdot;\tau, x_s(t,\psi,u(\cdot)), \tilde{u}(s))\,ds \\
&\quad + e^{-\alpha(T-t)} \Psi(x_T(\cdot;\tau, x_\tau(t,\psi,u(\cdot)), \tilde{u}(\cdot)))\Big|\mathcal{F}(\tau)\bigg]\bigg].
\end{aligned}$$

Therefore,

$$\begin{aligned}
V(t,\psi) &\geq E\bigg[\int_t^\tau e^{-\alpha(s-t)} L(s, x_s(\cdot;t,\psi,u(\cdot)), u(s))\,ds \\
&\quad + J(\tau, x_\tau(\cdot;t,\psi,u(\cdot)); \tilde{u}(\cdot))\bigg] \text{ (by Lemma reflem:3.3.8)} \\
&\geq E\bigg[\int_t^\tau L(s, x_s(\cdot;t,\psi,u(\cdot)), u(s))\,ds \\
&\quad + V(\tau, x_\tau(\cdot;t,\psi,u(\cdot)))\bigg] - 3\epsilon \text{ (by (3.55).}
\end{aligned}$$

Since this holds for arbitrary $(\Omega, \mathcal{F}, P, \mathbf{F}(t), W(\cdot), \tilde{u}(\cdot)) \in \mathcal{U}[t,T]$, by taking the supremum over $\mathcal{U}[t,T]$ we obtain

$$V(t,\psi) \geq \bar{V}(t,\psi) - 3\epsilon, \quad \forall \epsilon > 0. \tag{3.56}$$

Letting $\epsilon \downarrow 0$, we conclude that

$$V(t,\psi) \geq \bar{V}(t,\psi), \quad \forall (t,\psi) \in [0,T] \times \mathbf{C}.$$

This proves the DDP. $\square$

Assumption 3.3.10 *There exists a constant* $K > 0$ *such that*

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

In addition to its continuity in the initial function $\psi \in \mathbf{C}$ as proved in Lemma 3.3.7, the value function V has some regularity in the time variable as well. By using Theorem 3.3.9, it can be shown below that the value function V is Hölder continuous in time with respect to a parameter γ for any $\gamma \le \frac{1}{2}$, provided that the initial segment is at least γ-Hölder continuous. Notice that the coefficients f, g, and L need not be Hölder continuous in time. Except for the role of the initial segment, the statement and proof of the following lemma are analogous to the nondelay case (see, e.g., Krylov [Kry80, p.167]). See also Proposition 2 of Fischer and Nappo [FN06] for delay case.

Lemma 3.3.11 *Assume Assumptions 3.1.1 and 3.8.1 hold. Let the initial function* $\psi \in \mathbf{C}$. *If* ψ *is* γ*-Hölder continuous with* $\gamma \le K(H)$, *then the function* $V(\cdot,\psi) : [0,T] \to \Re$ *is Hölder continuous; that is, there is a constant* $K(V) > 0$ *depending only on* $K(H)$, K *(the Lipschitz constant in Assumption 3.1.1),* T, *and the dimensions such that for all* $t, \tilde{t} \in [0,T]$

$$|V(t,\psi) - V(\tilde{t},\psi)| \le K(V)\left(|t-\tilde{t}|^{\gamma} \vee \sqrt{|t-\tilde{t}|}\right).$$

Proof. Let the initial function $\psi \in \mathbf{C}$ be γ-Hölder continuous with $\gamma \le K(H)$. Without loss of generality, we assume that $s = t + h$ for some $h > 0$. We may also assume that $h \le \frac{1}{2}$, because we can choose $K(V) \ge 4K(T+1)$ so that the asserted inequality holds for $|t-s| > \frac{1}{2}$. By the DDP (Theorem 3.3.9), we see that

$$\begin{aligned}
&|V(t,\psi) - V(s,\psi)| \\
&= |V(t,\psi) - V(t+h,\psi)| \\
&= \Big| \sup_{u(\cdot)\in\mathcal{U}[t,T]} E\Big[\int_t^{t+h} e^{-\alpha(s-t)} L(s, x_s(\cdot;t,\psi,u(\cdot)), u(s))\,ds \\
&\quad + e^{-\alpha h} V(t+h, x_{t+h}(t,\psi,u(\cdot))\Big] \\
&\quad - V(t+h,\psi)\Big| \\
&\le \sup_{u(\cdot)\in\mathcal{U}[t,T]} E\left[\int_t^{t+h} e^{-\alpha(s-t)} L(s, x_s(\cdot;t,\psi,u(\cdot)), u(s))\,ds\right] \\
&\quad + \sup_{u(\cdot)\in\mathcal{U}[t,T]} E\left[|e^{-\alpha h} V(t+h, x_{t+h}(\cdot;t,\psi,u(\cdot)) - V(t+h,\psi)|\right] \\
&\le Kh + \sup_{u(\cdot)\in\mathcal{U}[t,T]} L(V) E[\|x_{t+h}(\cdot;t,\psi,u(\cdot)) - \psi\|],
\end{aligned}$$

where K is the largest Lipschitz constant from Assumption 3.1.1 and $L(V)$ is the Lipschitz constant for V in the segment variable according to Lemma 3.3.7. Notice that $\psi = x_t(t, \psi, u(\cdot))$ for all $u(\cdot) \in \mathcal{U}[t, T]$. By the linear growth condition (Assumption 3.1.1) of f and g, the Hölder inequality, Doob-Davis-Gundy's maximal inequality (Theorem 1.2.11), and Itó's isometry, we have for arbitrary $u(\cdot)$,

$$\begin{aligned} E[\|x_{t+h}(\cdot; t, \psi, u(\cdot)) - \psi\|] \leq & \sup_{\theta \in [-r,-h]} |\psi(\theta + h) - \psi(\theta)| + \sup_{\theta \in [-r,0]} |\psi(0) - \psi(\theta)| \\ & + E\left[\int_t^{t+h} |f(s, x_s(\cdot; t, \psi, u(\cdot)), u(s))| \, ds\right] \\ & + E\left[\left|\int_t^{t+h} g(s, x_s(\cdot; t, \psi, u(\cdot)), u(s)) \, dW(s)\right|^2\right]^{\frac{1}{2}} \\ \leq & \, 2K(H)h^\gamma + Kh + 4Km\sqrt{h}. \end{aligned}$$

Putting everything together, we have the assertion of the lemma. □

3.4 The Infinite-Dimensional HJB Equation

We will use the dynamic programming principle (Theorem 3.3.9) to derive the HJBE. Recall that $\{x_s(\cdot; t, \psi, u(\cdot)), s \in [t, T]$ is a $\mathbf{C}$-valued (strong) Markov process whenever $u(\cdot) \in \mathcal{U}[t, T]$. Therefore, using Theorem 2.4.1, we have the following.

Theorem 3.4.1 *Suppose that $\Phi \in C^{1,2}_{lip}([0, T] \times \mathbf{C}) \cap \mathcal{D}(\mathcal{S})$. Let $u(\cdot) \in \mathcal{U}[t, T]$ with $\lim_{s \downarrow t} u(s) = u \in U$ and $\{x_s(\cdot; t, \psi, u(\cdot)), s \in [t, T]\}$ be the $\mathbf{C}$-valued Markov process of (3.1) with the initial datum $(t, \psi) \in [0, T] \times \mathbf{C}$. Then*

$$\lim_{\epsilon \downarrow 0} \frac{E[\Phi(t + \epsilon, x_{t+\epsilon}(t, \psi, u(\cdot)))] - \Phi(t, \psi)}{\epsilon} = \partial_t \Phi(t, \psi) + \mathbf{A}^u \Phi(t, \psi), \quad (3.57)$$

where

$$\begin{aligned} \mathbf{A}^u \Phi(t, \psi) = & \, \mathcal{S}\Phi(t, \psi) + \overline{D\Phi(t, \psi)}(f(t, \psi, u)\mathbf{1}_{\{0\}}) \quad (3.58) \\ & + \frac{1}{2} \sum_{j=1}^m \overline{D^2\Phi(t, \psi)}(g(t, \psi, u)(\mathbf{e}_j)\mathbf{1}_{\{0\}}, g(t, \psi, u)(\mathbf{e}_j)\mathbf{1}_{\{0\}}), \end{aligned}$$

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

Heuristic Derivation of the HJB Equation

Let $u \in U$. We define the Fréchet differential operator $\mathcal{A}^u$ as follows:

$$\begin{aligned}\mathcal{A}^u\Phi(t,\psi) &\equiv \mathcal{S}(\Phi)(t,\psi) + \overline{D\Phi(t,\psi)}(f(t,\psi,u)\mathbf{1}_{\{0\}}) \\ &\quad + \frac{1}{2}\sum_{j=1}^m \overline{D^2\Phi(t,\psi)}(g(t,\psi,u)\mathbf{e}_j\mathbf{1}_{\{0\}}, g(t,\psi,u)\mathbf{e}_j\mathbf{1}_{\{0\}}),\end{aligned} \tag{3.59}$$

for any $\Phi \in C^{1,2}_{lip}([0,T]\times\mathbf{C})\cap\mathcal{D}(\mathcal{S})$, where $\mathbf{e}_j$ is the jth vector of the standard basis in $\Re^m$.

To recall the meaning of the terms involved in this differential operator, we remind the readers of the following definitions given earlier in this volume.

First, $\mathcal{S}(\Phi)(t,\psi)$ is defined (see (2.11) of Chapter 2) as

$$\mathcal{S}(\Phi)(t,\psi) = \lim_{\epsilon\downarrow 0}\frac{\Phi(t,\tilde{\psi}_{t+\epsilon}) - \Phi(t,\psi)}{\epsilon} \tag{3.60}$$

and $\tilde{\psi} : [-r,T] \to \Re^n$ is the extension of $\psi \in \mathbf{C}$ from $[-r,0]$ to $[-r,T]$ and is defined by

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

Second, $D\Phi(t,\psi) \in \mathbf{C}^*$ and $D^2\Phi(t,\psi) \in \mathbf{C}^\dagger$ are the first- and second-order Fréchet derivatives of Φ with respect to its second argument $\psi \in \mathbf{C}$. In addition, $\overline{D\Phi(t,\psi)} \in (\mathbf{C}\oplus\mathbf{B})^*$ is the extension of of $D\Phi(t,\psi)$ from $\mathbf{C}^*$ to $(\mathbf{C}\oplus\mathbf{B})^*$ (see Lemma 2.2.3 of Chapter 2) and $\overline{D^2V(t,\psi)} \in (\mathbf{C}\oplus\mathbf{B})^\dagger$ is the extension of $D^2V(t,\psi)$ from $\mathbf{C}^\dagger$ to $(\mathbf{C}\oplus\mathbf{B})^\dagger$ (see Lemma 2.2.4 of Chapter 2).

Finally, the function $\mathbf{1}_{\{0\}} : [-r,0] \to \Re$ is defined by

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

Without loss of generality, we can and will assume that for every $u \in U$, the domain of the generator $\mathbf{A}^u$ is large enough to contain $C^{1,2}_{lip}([0,T]\times\mathbf{C})\cap\mathcal{D}(\mathcal{S})$.

From the DDP (Theorem 3.3.9), if we take a constant control $u(\cdot) \equiv u \in \mathcal{U}[t,T]$, then for $\forall\delta \geq 0$,

$$\begin{aligned}V(t,\psi) \geq E\Big[\int_t^{t+\delta} &e^{-\alpha(s-t)}L(s,x_s(\cdot;t,\psi,v),v)\,ds \\ &+ e^{-\alpha\delta}V(t+\delta,x_{t+\delta}(t,\psi,v))\Big].\end{aligned}$$

From this principle, we have

$$
\begin{aligned}
0 &\geq \lim_{\delta \downarrow 0} \frac{1}{\delta} E\Big[\int_t^{t+\delta} e^{-\alpha(s-t)} L(s, x_s(\cdot; t, \psi, u), u)\, ds \\
&\quad + e^{-\alpha\delta} V(t+\delta, x_{t+\delta}(\cdot; t, \psi, u)) - V(t, \psi) \Big] \\
&= \lim_{\delta \downarrow 0} \frac{1}{\delta} E\Big[\int_t^{t+\delta} e^{-\alpha(s-t)} L(s, x_s(\cdot; t, \psi, u), u)\, ds \\
&\quad + \lim_{\delta \downarrow 0} \frac{1}{\delta} E[e^{-\alpha\delta} V(t+\delta, x_{t+\delta}(\cdot; t, \psi, u)) - e^{-\alpha\delta} V(t + x_{t+\delta}(\cdot; t, \psi, u))] \\
&\qquad + \lim_{\delta \downarrow 0} \frac{1}{\delta} E[e^{-\alpha\delta} V(t, x_{t+\delta}(t, \psi, u)) - e^{-\alpha\delta} V(t, \psi))] \\
&\qquad + \lim_{\delta \downarrow 0} \frac{1}{\delta} [(e^{-\alpha\delta} - 1) V(t, \psi)] \\
&\qquad = -\alpha V(t, \psi) + \partial_t V(t, \psi) + \mathbf{A}^u V(t, \psi) + L(t, \psi, u) \qquad (3.61)
\end{aligned}
$$

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 that satisfies $\lim_{s \downarrow t} u^*(s) = v^*$, we should have, $\forall \delta \geq 0$, that

$$
\begin{aligned}
V(t, \psi) = E\Big[&\int_t^{t+\delta} e^{-\alpha(s-t)} L(s, x^*_s(\cdot; t, \psi, u^*(\cdot)), u^*(s))\, ds \\
&+ e^{-\alpha\delta} V(t+\delta, x^*_{t+\delta}(\cdot; t, \psi, u^*(\cdot)) \Big], \qquad (3.62)
\end{aligned}
$$

where $x^*_s(t, \psi, u^*(\cdot))$ 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 the strong assumption on $u^*(\cdot)$ (including the right-continuity at the initial time t), we can get

$$
0 = -\alpha V(t, \psi) + \partial_t V(t, \psi) + \mathcal{A}^{v^*} V(t, \psi) + L(t, \psi, v^*). \qquad (3.63)
$$

Inequalities (3.61) and (3.62) are equivalent to the HJBE

$$
0 = -\alpha V(t, \psi) + \partial_t V(t, \psi) + \max_{u \in U} \; [\mathcal{A}^u V(t, \psi) + L(t, \psi, u)].
$$

We therefore have the following result.

Theorem 3.4.2 *Let $V : [0, T] \times \mathbf{C} \to \Re$ be the value function defined by (3.7). Suppose $V \in C^{1,2}_{lip}([0, T] \times \mathbf{C}) \cap \mathcal{D}(\mathcal{S})$. Then the value function V satisfies the following HJBE:*

$$
\alpha V(t, \psi) - \partial_t V(t, \psi) - \max_{u \in U} \; [\mathbf{A}^u V(t, \psi) + L(t, \psi, u)] = 0 \qquad (3.64)
$$

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

$$\mathcal{A}^u V(t,\psi) \equiv \mathcal{S}V(t,\psi) + \overline{DV(t,\psi)}(f(t,\psi,u)\mathbf{1}_{\{0\}})$$
$$+\frac{1}{2}\sum_{j=1}^{m} \overline{D^2V(t,\psi)}(g(t,\psi,u)\mathbf{e}_j\mathbf{1}_{\{0\}}, g(t,\psi,u)\mathbf{e}_j\mathbf{1}_{\{0\}}).$$

Note that it is not known that the value function V satisfies the necessary smoothness condition $V \in C^{1,2}_{lip}([0,T] \times \mathbf{C}) \cap \mathcal{D}(\mathcal{S})$. In fact, the following simple example shows that the value function does not possess the smoothness condition for it to be a classical solution of the HJBE (3.64) even for a very simple deterministic control problem.

The following is an example taken from Example 2.3 of [FS93, p.57].

Example. Consider a one-dimensional simple deterministic control problem described by $\dot{x}(s) = u(s), s \in [0,1)$, with the running cost function $L \equiv 0$ and the terminal cost function $\Psi(x) = x \in \Re$ and a control set $U = [-a,a]$ with some constant $a > 0$. Since the boundary data are increasing and the running cost is zero, the optimal control is $u^*(s) \equiv -a$. Hence, the value function is

$$V(t,x) = \begin{cases} -1 & \text{if } x + at \leq a - 1 \\ x + at - a & \text{if } x + at \geq a - 1 \end{cases} \tag{3.65}$$

for $(t,x) \in [0,1] \times \Re$. Note that the value function is differentiable except on the set $\{(t,x) \mid x + at = a - 1\}$ and it is a generalized solution of

$$-\partial_t V(t,x) + a\,|\partial x V(t,x)| = 0, \tag{3.66}$$

with the corresponding terminal boundary condition given by

$$V(1,x) = \Psi(x) = x, \quad x \in [-1,1]. \tag{3.67}$$

The above example shows that in general we need to seek a weaker condition for the HJBE (3.64) such as a viscosity solution instead of a solution for HJBE (3.64) in the classical sense. In fact, it will be shown that the value function is a unique viscosity solution of the HJBE (3.64). These results will be given in Sections 3.5 and 3.6.

3.5 Viscosity Solution

In this section, we shall show that the value function $V : [0,T] \times \mathbf{C} \to \Re$ defined by Equation (3.7) is actually a viscosity solution of the HJBE (3.64).

Definitions of Viscosity Solution

First, let us define the viscosity solution of (3.64) as follows.

Definition 3.5.1 *An upper semicontinuous (respectively, lower semicontinuous) function $w : (0, T] \times \mathbf{C}) \to \Re$ is said to be a viscosity subsolution (respectively, supersolution) of the HJBE (3.64) if*

$$w(T, \phi) \leq (\geq) \Psi(\phi), \quad \forall \phi \in \mathbf{C},$$

and if, for every $\Gamma \in C^{1,2}_{lip}([0, T] \times \mathbf{C}) \cap \mathcal{D}(\mathcal{S})$ and for every $(t, \psi) \in [0, T] \times \mathbf{C}$ satisfying $\Gamma \geq (\leq) w$ on $[0, T] \times \mathbf{C}$ and $\Gamma(t, \psi) = w(t, \psi)$, we have

$$\alpha \Gamma(t, \psi) - \partial_t \Gamma(t, \psi) - \max_{v \in U} \left[\mathbf{A}^v \Gamma(t, \psi) + L(t, \psi, v) \right] \leq (\geq)\ 0. \tag{3.68}$$

We say that w is a viscosity solution of the HJBE (3.64) if it is both a viscosity supersolution and a viscosity subsolution of the HJBE (3.64).

Definition 3.5.2 *Let a function $\Phi : [0, T] \times \mathbf{C} \to \Re$ be given. We say that $(p, q, Q) \in \Re \times \mathbf{C}^* \times \mathbf{C}^\dagger$ belongs to $D^{1,2,+}_{t+,\phi} \Phi(t, \phi)$, the second-order superdifferential of Φ at $(t, \phi) \in (0, T) \times \mathbf{C}$, if, for all $\psi \in \mathbf{C}$, $s \geq t$,*

$$\begin{aligned} \Phi(s, \psi) \leq \Phi(t, \phi) &+ p(s - t) + q(\psi - \phi) \\ &+ \frac{1}{2} Q(\psi - \phi, \psi - \phi) + o(s - t + \|\psi - \phi\|^2). \end{aligned} \tag{3.69}$$

The second-order one-sided parabolic subdifferential of Φ at $(t, \phi) \in (0, T) \times \mathbf{C}$, $D^{1,2,-}_{t+,\phi} \Phi(t, \phi)$, is defined as those $(p, q, Q) \in \Re \times \mathbf{C}^* \times \mathbf{C}^\dagger$ such that the above inequality is reversed; that is,

$$\Phi(s, \psi) \geq \Phi(t, \phi) + p(s - t) + q(\psi - \phi) \tag{3.70}$$

$$+ \frac{1}{2} Q(\psi - \phi, \psi - \phi) + o(s - t + \|\psi - \phi\|^2). \tag{3.71}$$

Remark 3.5.3 *The second-order one-sided parabolic subdifferential and second-order one-sided parabolic superdifferential have the following relationship:*

$$D^{1,2,-}_{t+,\phi} \Phi(t, \phi) = -D^{1,2,+}_{t+,\phi} (-\Phi)(t, \phi).$$

Lemma 3.5.4 *Let $w : (0, T] \times \mathbf{C} \to \Re$ and let $(t, \psi) \in (0, T) \times \mathbf{C}$. Then $(p, q, Q) \in D^{1,2,+}_{t+,\phi} w(t, \psi)$ if and only if there exists a function $\Phi \in C^{1,2}((0, T) \times \mathbf{C})$ such that $w - \Phi$ attains a strict global maximum at (t, ψ) relative to the set of (s, ϕ) such that $s \geq t$ and*

$$(\Phi(t, \psi), \partial_t \Phi(t, \psi), D\Phi(t, \psi), D^2 \Phi(t, \psi)) = (w(t, \psi), p, q, Q). \tag{3.72}$$

Moreover, if w has polynomial growth (i.e., there exist positive constants K_p and $k \geq 1$ such that

$$|w(s, \phi)| \leq K_p (1 + \|\phi\|_2)^k, \quad \forall (s, \phi) \in (0, T) \times \mathbf{C}), \tag{3.73}$$

then Φ can be chosen so that Φ $\partial_t \Phi$, $D\Phi$, and $D^2 \Phi$ satisfy (3.73) (with possibly different constants K_p).

Lemma 3.5.5 *Let v be an upper-semicontinuous function on $(0,T)\times\mathbf{C}$ and let $(\bar t,\bar\psi)\in(0,T)\times\mathbf{C}$. Then $(p,q,Q)\in\bar D^{1,2}_{t+,\psi}v(\bar t,\bar\psi)$ if and only if there exists a function $\Phi\in C^{1,2}((0,T)\times\mathbf{C})$ such that $v-\Phi\in C((0,T)\times\mathbf{C})$ attains a strict global maximum at $(\bar t,\bar\psi)$ relative to the set of $(t,\psi)\in[\bar t,T)\times\mathbf{C}$ and*

$$(\Phi(\bar t,\bar\psi),\partial_t\Phi(\bar t,\bar\psi),D\Phi(\bar t,\bar\psi),D^2\Phi(\bar t,\bar\psi))=(v(\bar t,\bar\psi),p,q,Q). \tag{3.74}$$

Moreover, if v has polynomial growth (i.e., if there exists a constant $k\ge 1$ such that

$$|v(t,\psi)|\le C(1+\|\psi\|^k)\ \forall(t,\psi)\in[0,T]\times\mathbf{C}), \tag{3.75}$$

then Φ can be chosen so that Φ, $\partial_t\Phi$, $D\Phi$, and $D^2\Phi$ satisfy (3.69) under the appropriate norms and with possibly different constants C.

Proof. The proof of this lemma is an extension of Lemma 5.4 in Yong and Zhou [YZ99, Chap.4] from an Euclidean space to the infinite-dimensional space $\mathbf{C}$.

Suppose $(p,q,Q)\in\bar D^{1,2}_{t+,\psi}v(\bar t,\bar\psi)$. Define the function $\gamma:(0,T)\times\mathbf{C}\to\Re$ as

$$\begin{aligned}\gamma(t,\psi)=&\frac{1}{t-\bar t+\|\psi-\bar\psi\|^2}[v(t,\psi)-v(\bar t,\bar\psi)\\&-p(t-\bar t)-q(\psi-\bar\psi)-\frac12 Q(\phi-\psi,\phi-\psi)]\vee 0\\&\text{for }(t,\psi)\ne(\bar t,\bar\psi),\end{aligned}$$

and $\gamma(t,\psi)=0$ otherwise.

We also define the function $\epsilon:\Re\to\Re$ by

$$\epsilon(r)=\sup\{\gamma(t,\psi)\mid(t,\psi)\in(\bar t,T]\times\mathbf{C},\ t-\bar t+\|\psi-\bar\psi\|^2\le r\}\text{ if }r>0$$

and $\epsilon(r)=0$ if $r\le 0$. Then it follows from the definition of $\bar D^{1,2}_{t+,\psi}v(t,\psi)$ that

$$\begin{aligned}&v(t,\psi)-[v(\bar t,\bar\psi)+p(t-\bar t)+q(\psi-\bar\psi)+\frac12 Q(\phi-\psi,\phi-\psi)]\\&\le(t-\bar t+\|\psi-\bar\psi\|^2)\epsilon(t-\bar t+\|\psi-\bar\psi\|^2)\ \forall(t,\psi)\in[\bar t,T]\times\mathbf{C}.\end{aligned}$$

Define the function $\alpha:\Re_+\to\Re$ by

$$\alpha(\rho)=\frac{2}{\rho}\int_0^{2\rho}\int_0^r\epsilon(\theta)\,d\theta\,dr,\rho>0. \tag{3.76}$$

Then it is easy to see that its first-order derivative

$$\dot\alpha(\rho)=-\frac{2}{\rho^2}\int_0^{2\rho}\int_0^r\epsilon(\theta)\,d\theta\,dr+\frac{4}{\rho}\int_0^{2\rho}\epsilon(\theta)\,d\theta$$

and its second order-derivative

$$\ddot{\alpha}(\rho) = \frac{4}{\rho^3}\int_0^{2\rho}\int_0^r \epsilon(\theta)\,d\theta\,dr - \frac{8}{\rho^2}\int_0^{2\rho}\epsilon(\theta)\,d\theta + \frac{8}{\rho}\epsilon(2\rho).$$

Consequently,

$$|\alpha(\rho)| \le 4\rho\epsilon(2\rho), \ |\dot{\alpha}(\rho)| \le 12\rho\epsilon(2\rho), \text{ and } |\ddot{\alpha}(\rho)| \le \frac{32\rho\epsilon(2\rho)}{\rho}.$$

Now, we define the function $\beta : [0,T] \times \mathbf{C} \to \Re$ by

$$\beta(t,\psi) = \begin{cases} \alpha(\rho(t,\psi)) + \rho^2(t,\psi) & \text{if } (t,\psi) \neq (\bar{t},\bar{\psi}) \\ 0 & \text{otherwise,} \end{cases}$$

where $\rho(t,\psi) = t - \bar{t} + \|\psi - \bar{\psi}\|_2$.

Finally, we define the function $\Phi : [\bar{t},T] \times \mathbf{C} \to \Re$ by

$$\begin{aligned}\Phi(t,\psi) &= v(\bar{t},\bar{\psi}) + p(t-\bar{t}) + q(\psi - \bar{\psi}) \\ &\quad + \frac{1}{2}Q(\phi - \psi, \phi - \psi) + \beta(t,\psi), \quad \forall (t,\psi) \in [0,T] \times \mathbf{C}. \quad (3.77)\end{aligned}$$

We claim that $\Phi \in C^{1,2}_{lip}([0,T] \times \mathbf{C})$ and it satisfies the following three conditions:
(i) $v(\bar{t},\bar{\psi}) = \Phi(\bar{t},\bar{\psi})$.
(ii) $v(t,\psi) < \Phi(t,\psi)$ for all $(t,\psi) \neq (\bar{t},\bar{\psi})$.
(iii) $(\Phi(\bar{t},\bar{\psi}), \partial_t\Phi(\bar{t},\bar{\psi}), D\Phi(\bar{t},\bar{\psi}), D^2\Phi(\bar{t},\bar{\psi})) = (v(\bar{t},\bar{\psi}), p, q, Q)$.

Note that (i) is trivial by the definition of Φ. The proofs for (ii) and (iii) are very similar to those of Lemma 2.7 and Lemma 2.8 of Yong and Zhou [YZ99] and are omitted here. □

Proposition 3.5.6 *A function $w \in C([0,T] \times \mathbf{C})$ is a viscosity solution of the HJBE (3.64) if*

$$-p - \sup_{u \in U} G(t,\phi,u,q,Q) \le 0, \quad \forall (p,q,Q) \in D^{1,2,+}_{t+,\phi}w(t,\phi), \forall (t,\phi) \in [0,T) \times \mathbf{C},$$

$$-p - \sup_{u \in U} G(t,\phi,u,q,Q) \ge 0, \quad \forall (p,q,Q) \in D^{1,2,-}_{t+,\phi}w(t,\phi), \forall (t,\phi) \in [0,T) \times \mathbf{C},$$

and

$$w(T,\phi) = \Psi(\phi), \quad \forall \phi \in \mathbf{C},$$

where the function G is defined as

$$\begin{aligned}G(t,\phi,u,q,Q) &= \mathcal{S}(\Phi)(t,\psi) + \overline{q}(f(t,\phi,v)\mathbf{1}_{\{0\}}) \\ &\quad + \frac{1}{2}\sum_{j=1}^m \overline{Q}(g(t,\psi,v)(\mathbf{e}_j)\mathbf{1}_{\{0\}}, g(t,\psi,v)(\mathbf{e}_j)\mathbf{1}_{\{0\}}). \quad (3.78)\end{aligned}$$

Proof. The proposition follows immediately from Lemma 3.5.4 and Lemma 3.5.5. □

For our value function V defined by (3.7), we now show that it has the following property.

Lemma 3.5.7 *Let $V : [0,T] \times \mathbf{C} \to \Re$ be the value function defined in (3.7). Then there exists a constant $k > 0$ and a positive integer p such that for every $(t,\psi) \in [0,T] \times \mathbf{C}$,*

$$|V(t,\psi)| \leq K(1 + \|\psi\|_2)^k. \tag{3.79}$$

Proof. It is clear that V has at most a polynomial growth, since L and Φ have at most a polynomial growth. This proves the lemma. □

Theorem 3.5.8 *The value function $V : [0,T] \times \mathbf{C} \to \Re$ defined in (3.7) is a viscosity solution of the HJBE:*

$$\alpha V(t,\psi) - \partial_t V(t,\psi) - \max_{v \in U} \; [\mathbf{A}^v V(t,\psi) + L(t,\psi,v)] = 0 \tag{3.80}$$

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

$$\begin{aligned} \mathbf{A}^v V(t,\psi) = \mathcal{S}V(t,\psi) &+ \overline{DV(t,\psi)}(f(t,\psi,v)\mathbf{1}_{\{0\}}) \\ &+ \frac{1}{2}\sum_{j=1}^{m} \overline{D^2V(t,\psi)}(g(t,\psi,v)(\mathbf{e}_j)\mathbf{1}_{\{0\}}, g(t,\psi,v)(\mathbf{e}_j)\mathbf{1}_{\{0\}}), \end{aligned} \tag{3.81}$$

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

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 $\Gamma \leq V$ on $[0,T] \times \mathbf{C}$ and $\Gamma(t,\psi) = V(t,\psi)$, we want to prove the viscosity supersolution inequality, that is,

$$\alpha \Gamma(t,\psi) - \partial_t \Gamma(t,\psi) - \max_{v \in U} \; [\mathbf{A}^v \Gamma(t,\psi) + L(t,\psi,v)] \geq 0. \tag{3.82}$$

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 3.4.1) for $t \leq s \leq T$, we have

$$\begin{aligned} &E\left[e^{-\alpha(s-t)}\Gamma(s, x_s(\cdot; t,\psi,u(\cdot)))\right] - \Gamma(t,\psi) \\ &= E\Bigg[\int_t^s e^{-\alpha(\xi-t)}\Bigg(\partial_\xi \Gamma(\xi, x_\xi(\cdot;t,\psi,u(\cdot))) \\ &\quad + \mathbf{A}^{u(\xi)}\Gamma(\xi, x_\xi(\cdot;t,\psi,u(\cdot))) - \alpha\Gamma(\xi, x_\xi(\cdot;t,\psi,u(\cdot)))\Bigg)d\xi\Bigg]. \end{aligned} \tag{3.83}$$

On the other hand, for any $s \in [t,T]$, the DDP (Theorem 3.3.9) gives,

$$\begin{aligned} V(t,\psi) = \max_{u(\cdot)\in\mathcal{U}[t,T]} E\Bigg\{ &\int_t^s e^{-\alpha(\xi-t)} L(\xi, x_\xi(\cdot;t,\psi,u(\cdot)), u(\xi))d\xi \\ &+ e^{-\alpha(s-t)}V(s, x_s(\cdot;t,\psi,u(\cdot)))\Bigg\}. \end{aligned} \tag{3.84}$$

Therefore, we have

$$V(t,\psi) \geq E\left[\int_t^s e^{-\alpha(\xi-t)}L(\xi, x_\xi(\cdot;t,\psi,u(\cdot)),u(\xi))\,d\xi\right] + E\left[e^{-\alpha(s-t)}V(s, x_s(\cdot;t,\psi,u(\cdot)))\right]. \tag{3.85}$$

By virtue of (3.177) and using $\Gamma \leq V, \Gamma(t,\psi) = V(t,\psi)$, we can get

$$\begin{aligned}
0 &\geq E\left[\int_t^s e^{-\alpha(\xi-t)}L(\xi, x_\xi(\cdot;t,\psi,u(\cdot)),u(\xi))d\xi\right] \\
&\quad + E\Big[e^{-\alpha(s-t)}V(s, x_s(\cdot;t,\psi,u(\cdot)))\Big] - V(t,\psi) \\
&\geq E\left[\int_t^s e^{-\alpha(\xi-t)}L(\xi, x_\xi(\cdot;t,\psi,u(\cdot)),u(\xi))\,d\xi\right]
\end{aligned}$$

$$\begin{aligned}
&+ E\Big[e^{-\alpha(s-t)}\Gamma(s, x_s(\cdot;t,\psi,u(\cdot)))\Big] - \Gamma(t,\psi) \\
&\geq E\int_t^s e^{-\alpha(\xi-t)}\Big[- \alpha\Gamma(\xi, x_\xi(\cdot;t,\psi,u(\cdot))) + \partial_\xi\Gamma(\xi, x_\xi(\cdot;t,\psi,u(\cdot))) \\
&\quad + \mathbf{A}^u\Gamma(\xi, x_\xi(\cdot;t,\psi,u(\cdot))) + L(\xi, x_\xi(\cdot;t,\psi,u(\cdot)),u(\xi))\Big]\,d\xi.
\end{aligned}$$

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

$$\begin{aligned}
0 \leq E\Bigg[\frac{1}{s-t}\int_t^s e^{-\alpha(\xi-t)}\Big(&\alpha\Gamma(\xi, x_\xi(\cdot;t,\psi,u(\cdot))) \\
&- \partial_\xi\Gamma(\xi, x_\xi(\cdot;t,\psi,u(\cdot))) - \mathbf{A}^{u(\xi)}\Gamma(\xi, x_\xi(\cdot;t,\psi,u(\cdot))) \\
&-L(\xi, x_\xi(\cdot;t,\psi,u(\cdot)),u(\xi))\Big)\,d\xi\Bigg].
\end{aligned} \tag{3.86}$$

Now, let $s \downarrow t$ in (3.86) and $\lim_{s\downarrow t} u(s) = v$, and we obtain

$$\alpha\Gamma(t,\psi) - \partial_t\Gamma(t,\psi) - [\mathbf{A}^v\Gamma(t,\psi) + L(t,\psi,v)] \geq 0. \tag{3.87}$$

Since $v \in U$ is arbitrary, we prove that V is a viscosity supersolution.

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

$$\alpha\Gamma(t,\psi) - \partial_t\Gamma(t,\psi) - \max_{v\in U}\ [\mathbf{A}^v\Gamma(t,\psi) + L(t,\psi,v)] \leq 0. \tag{3.88}$$

We assume the contrary and try to obtain a contradiction. Let suppose that there exit $(t,\psi) \in [0,T]\times\mathbf{C}$, $\Gamma \in C^{1,2}_{lip}([0,T]\times\mathbf{C}) \cap \mathcal{D}(\mathcal{S})$, 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 $\lim_{s\downarrow t} u(s) = v$,

$$\alpha \Gamma(\tau, \phi) - \partial_t \Gamma(\tau, \phi) - \mathbf{A}^v \Gamma(\tau, \phi) - L(\tau, \phi, v) \geq \delta \tag{3.89}$$

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 t_1 such that for $t \leq s \leq t_1$, the solution $x(s; t, \psi, u(\cdot)) \in N(t, \psi)$. Therefore, for any $s \in [t, t_1]$, we have almost surely

$$\begin{aligned} &\alpha \Gamma(s, x_s(t, \psi, u(\cdot))) - \partial_t \Gamma(s, x_s(\cdot; t, \psi, u(\cdot))) \\ &\quad - \mathbf{A}^v \Gamma(s, x_s(\cdot; t, \psi, u(\cdot))) - L(s, x_s(\cdot; t, \psi, u(\cdot)), u(s)) \geq \delta. \end{aligned} \tag{3.90}$$

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

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

Using (3.90), we have

$$\begin{aligned} J(t, \psi; u(\cdot)) \leq E\bigg[\int_t^{t_1} & e^{-\alpha(s-t)} \Big(- \delta + \alpha \Gamma(s, x_s(\cdot; t, \psi, u(\cdot))) \\ & - \partial_t \Gamma(s, x_s(\cdot; t, \psi, u(\cdot))) - \mathbf{A}^{u(s)} \Gamma(s, x_s(\cdot; t, \psi, u(\cdot)))\Big)\, ds \\ & + e^{-\alpha(t_1 - t)} \Gamma(t_1, x_{t_1}(t, \psi, u(\cdot)))\bigg]. \end{aligned} \tag{3.91}$$

In addition, similar to (3.86), we can get

$$\begin{aligned} &E\left[e^{-\alpha(t_1 - t)} \Gamma(t_1, x_{t_1}(\cdot; t, \psi, u(\cdot)))\right] - \Gamma(t, \psi) \\ &= E\bigg[\int_t^{t_1} e^{-\alpha(s-t)} \Big(\partial_s \Gamma(s, x_s(\cdot; t, \psi, u(\cdot))) + \mathbf{A}^{u(s)} \Gamma(s, x_s(\cdot; t, \psi, u(\cdot))) \\ &\quad - \alpha \Gamma(s, x_s(\cdot; t, \psi, u(\cdot)))\Big) ds\bigg]. \end{aligned} \tag{3.92}$$

Therefore, we can get

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

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

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

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

3.6 Uniqueness

In this section, we will show that the value function $V : [0,T] \times \mathbf{C} \to \Re$ of Problem (OCCP) is the unique viscosity solution of the HJBE (3.64). We first need the following comparison principle.

Theorem 3.6.1 (Comparison Principle) *Assume that $V_1(t,\psi)$ and $V_2(t,\psi)$ are both continuous with respect to the argument $(t,\psi) \in [0,T] \times \mathbf{C}$ and are respectively the viscosity subsolution and supersolution of (3.64) with at most a polynomial growth (Lemma 3.5.7). In other words, there exists a real number $\Lambda > 0$ and a positive integer $k \geq 1$ such that*

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

Then

$$V_1(t,\psi) \leq V_2(t,\psi), \quad \forall (t,\psi) \in [0,T] \times \mathbf{C}. \tag{3.93}$$

Before we proceed to the proof of Theorem 3.6.1, we will use the following result proven in Ekeland and Lebourg [EL76] and also in a general form in Stegall [Ste78] and Bourgain [Bou79]. The reader is also referred to Crandall et al. [CIL92] and Lions [Lio82, Lio89] for an application example of this result in a setting similar but significantly different in details from what are to be presented below. A similar proof of uniqueness of viscosity solution is also done in Chang et al. [CPP07a].

Lemma 3.6.2 *Let Φ be a bounded and upper-semicontinuous real-valued function on a closed ball B of a Banach space Ξ that has the Radon-Nikodym property. Then for any $\epsilon > 0$, there exists an element $u^* \in \Xi^*$ with norm at most ϵ, where Ξ^* is the topological dual of Ξ, such that $\Phi + u^*$ attains its maximum on B.*

Note that every separable Hilbert space $(\Xi, \|\cdot\|_\Xi)$ satisfies the Radon-Nikodym property (see, e.g., [EL76]). In order to apply Lemma 3.6.2, we will therefore restrict ourself to a subspace Ξ of the product space $\mathbf{C} \times \mathbf{C}$ ($\mathbf{C} = C([-r,0];\Re^n)$), which is a separable Hilbert space and dense in $\mathbf{C} \times \mathbf{C}$. A good candidates is the product space $W^{1,2}((-r,0);\Re^n) \times W^{1,2}((-r,0);\Re^n)$, where $W^{1,2}((-r,0);\Re^n)$ is the Sobolev space defined by

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

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

and $\dot{\phi}$ is the derivative of ϕ in the distributional sense. Note that it can be shown that the Hilbertian norm $\|\cdot\|_{1,2}$ is weaker than the sup-norm $\|\cdot\|$; that is, there exists a constant $K > 0$ such that

$$\|\phi\|_{1,2} \leq K\|\phi\|, \quad \forall \phi \in W^{1,2}((-r,0);\Re^n).$$

From the Sobolev embedding theorems (see, e.g., Adams [Ada75]), 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 [Ada75].

Before we proceed to the proof of the Comparison Principle, first let us establish some results that will be needed in the proof.

Let V_1 and V_2 be respectively a viscosity subsolution and supersolution of (3.64). For any $0 < \delta, \gamma < 1$, and for all $\psi, \phi \in \mathbf{C}$ and $t, s \in [0,T]$, define

$$\begin{aligned} \Theta_{\delta\gamma}(t,s,\psi,\phi) \equiv \frac{1}{\delta}\left[\|\psi-\phi\|_2^2 + \|\psi^0 - \phi^0\|_2^2 + |t-s|^2\right] \\ +\gamma\left[\exp(1+\|\psi\|_2^2 + \|\psi^0\|_2^2) + \exp(1+\|\phi\|_2^2+\|\phi^0\|_2^2)\right], \end{aligned} \tag{3.94}$$

and

$$\Phi_{\delta\gamma}(t,s,\psi,\phi) \equiv V_1(t,\psi) - V_2(s,\phi) - \Theta_{\delta\gamma}(t,s,\psi,\phi), \tag{3.95}$$

where $\psi^0, \phi^0 \in \mathbf{C}$ with $\psi^0(\theta) = \frac{\theta}{-r}\psi(-r-\theta)$ and $\phi^0(\theta) = \frac{\theta}{-r}\phi(-r-\theta)$ for $\theta \in [-r,0]$.

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

$$\lim_{\|\psi\|_2, \|\phi\|_2 \to \infty} \Phi_{\delta\gamma}(t,s,\psi,\phi) = -\infty. \tag{3.96}$$

The function $\Phi_{\delta\gamma}$ is a real-valued function that is bounded above and continuous on $[0,T] \times [0,T] \times W^{1,2}((-r,0);\Re^n) \times W^{1,2}((-r,0);\Re^n)$ (since the Hilbertian norm $\|\cdot\|_{1,2}$ is weaker than the sup-norm $\|\cdot\|$). Therefore, from Lemma 3.6.2 (which is applicable by virtue of (3.96)), 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 3.6.2). 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

$$\Phi_{\delta\gamma}(t_{\delta\gamma\epsilon}, s_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) + T_\epsilon(\psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) + M_{\delta\gamma\epsilon} = 0. \tag{3.97}$$

We have the following lemmas.

Lemma 3.6.3 *$(t_{\delta\gamma\epsilon}, s_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon})$ is the global maximum of $\Phi_{\delta\gamma} + T_\epsilon$ in $[0,T] \times [0,T] \times \mathbf{C} \times \mathbf{C}$.*

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

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

It is known that

$$\Phi_{\delta\gamma}(t_k, s_k, \psi_k, \phi_k) + T_\epsilon(\psi_k, \phi_k) \le \Phi_{\delta\gamma}(t_{\delta\gamma\epsilon}, s_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) + T_\epsilon(\psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}).$$

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

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

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

$$\|\psi_{\delta\gamma\epsilon}\|_2 + \|\psi^0_{\delta\gamma\epsilon}\|_2 + \|\phi_{\delta\gamma\epsilon}\|_2 + \|\phi^0_{\delta\gamma\epsilon}\|_2 \le \Lambda_\gamma \tag{3.98}$$

and

$$\lim_{\epsilon\downarrow 0, \delta\downarrow 0} \Big(\|\psi_{\delta\gamma\epsilon} - \phi_{\delta\gamma\epsilon}\|_2^2 + \|\psi^0_{\delta\gamma\epsilon} - \phi^0_{\delta\gamma\epsilon}\|_2^2 + |t_{\delta\gamma\epsilon} - s_{\delta\gamma\epsilon}|^2 \Big) = 0, \tag{3.99}$$

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 obtain

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

It implies that

$$\begin{aligned}
&V_1(t_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}) - V_2(t_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}) - 2\gamma(\exp(1 + \|\psi_{\delta\gamma\epsilon}\|_2^2 + \|\psi^0_{\delta\gamma\epsilon}\|_2^2)) \\
&+ T_\epsilon(\psi_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}) + V_1(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) - V_2(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) \\
&- 2\gamma(\exp(1 + \|\phi_{\delta\gamma\epsilon}\|_2^2 + \|\phi^0_{\delta\gamma\epsilon}\|_2^2)) + T_\epsilon(\phi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) \\
\le\ & 2V_1(t_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}) - 2V_2(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) \\
&- \frac{2}{\delta}\Big[\|\psi_{\delta\gamma\epsilon} - \phi_{\delta\gamma\epsilon}\|_2^2 + \|\psi^0_{\delta\gamma\epsilon} - \phi^0_{\delta\gamma\epsilon}\|_2^2 + |t_{\delta\gamma\epsilon} - s_{\delta\gamma\epsilon}|^2 \Big] \\
&- 2\gamma\Big(\exp(1 + \|\psi_{\delta\gamma\epsilon}\|_2^2 + \|\psi^0_{\delta\gamma\epsilon}\|_2^2) + \exp(1 + \|\phi_{\delta\gamma\epsilon}\|_2^2 + \|\phi^0_{\delta\gamma\epsilon}\|_2^2) \Big) \\
&+ 2T_\epsilon(\psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}).
\end{aligned} \tag{3.100}$$

From the above inequality, it is easy to obtain

$$\begin{aligned}
&\frac{2}{\delta}\Big[\|\psi_{\delta\gamma\epsilon} - \phi_{\delta\gamma\epsilon}\|_2^2 + \|\psi^0_{\delta\gamma\epsilon} - \phi^0_{\delta\gamma\epsilon}\|_2^2 + |t_{\delta\gamma\epsilon} - s_{\delta\gamma\epsilon}|^2 \Big] \\
&\le [V_1(t_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}) - V_1(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon})] + [V_2(t_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}) - V_2(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon})] \\
&\quad + 2T_\epsilon(\psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) - [T_\epsilon(\psi_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}) + T_\epsilon(\phi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon})].
\end{aligned} \tag{3.101}$$

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

$$\frac{2}{\delta}\left[\|\psi_{\delta\gamma\epsilon}-\psi_{\delta\gamma\epsilon}\|_2^2+\|\psi^0_{\delta\gamma\epsilon}-\phi^0_{\delta\gamma\epsilon}\|_2^2+|t_{\delta\gamma\epsilon}-s_{\delta\gamma\epsilon}|^2\right] \leq \Lambda(1+\|\psi_{\delta\gamma\epsilon}\|_2+\|\phi_{\delta\gamma\epsilon}\|_2)^k. \tag{3.102}$$

So,

$$\|\psi_{\delta\gamma\epsilon}-\phi_{\delta\gamma\epsilon}\|_2^2+\|\psi^0_{\delta\gamma\epsilon}-\phi^0_{\delta\gamma\epsilon}\|_2^2+|t_{\delta\gamma\epsilon}-s_{\delta\gamma\epsilon}|^2 \leq \delta\Lambda(1+\|\psi_{\delta\gamma\epsilon}\|_2+\|\phi_{\delta\gamma\epsilon}\|_2)^k. \tag{3.103}$$

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 obtain

$$\Phi_{\delta\gamma}(t_{\delta\gamma\epsilon}, s_{\delta\gamma\epsilon}, 0, 0) + T_\epsilon(0,0) \leq \Phi_{\delta\gamma}(t_{\delta\gamma\epsilon}, s_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) + T_\epsilon(\psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) \tag{3.104}$$

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 $k \geq 1$ 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 (3.104) and the definition of $\Phi_{\delta\gamma}$, we have

$$\begin{aligned}
&\gamma\left[\exp(1+\|\psi_{\delta\gamma\epsilon}\|_2^2+\|\psi^0_{\delta\gamma\epsilon}\|_2^2)+\exp(1+\|\phi_{\delta\gamma\epsilon}\|_2^2+\|\phi^0_{\delta\gamma\epsilon}\|_2^2)\right] \\
&\leq V_1(t_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}) - V_2(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) \\
&\quad -\frac{1}{\delta}\left[\|\psi_{\delta\gamma\epsilon}-\phi_{\delta\gamma\epsilon}\|_2^2+\|\psi^0_{\delta\gamma\epsilon}-\phi^0_{\delta\gamma\epsilon}\|_2^2+|t_{\delta\gamma\epsilon}-s_{\delta\gamma\epsilon}|^2\right] \\
&\quad -\Phi_{\delta\gamma}(t_{\delta\gamma\epsilon}, s_{\delta\gamma\epsilon}, 0, 0) + T_\epsilon(\psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) - T_\epsilon(0,0) \\
&\leq 3\Lambda(1+\|\psi_{\delta\gamma\epsilon}\|_2+\|\phi_{\delta\gamma\epsilon}\|_2)^k.
\end{aligned} \tag{3.105}$$

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

$$\|\psi_{\delta\gamma\epsilon}\|_2+\|\psi^0_{\delta\gamma\epsilon}\|_2+\|\phi_{\delta\gamma\epsilon}\|_2+\|\phi^0_{\delta\gamma\epsilon}\|_2 \leq \Lambda_\gamma. \tag{3.106}$$

In order to obtain (3.99), we set δ to zero in (3.103) using the above inequality. □

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

$$F(\psi) \equiv \|\psi\|_2^2 \tag{3.107}$$

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

$$H(\psi)(\theta) \equiv \frac{\theta}{-r}\psi(-r-\theta) = \psi^0(\theta), \qquad \theta \in [-r, 0]. \tag{3.108}$$

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

$$DF(\varphi)h = 2\int_{-r}^{0} \varphi(\theta) \cdot h(\theta)\, d\theta \equiv 2\langle u, h\rangle_2,$$

where $\langle \cdot, \cdot \rangle_2$ and $\| \cdot \|_2$ are the inner product and the Hilbertian norm in the Hilbert space $L^2([-r, 0]; \Re^n)$. This comes from the fact that

$$\|\psi + h\|_2^2 - \|\psi\|_2^2 = 2\langle \psi, h\rangle_2 + \|h\|_2^2,$$

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

$$\frac{|\|\psi + h\|_2^2 - \|\psi\|_2^2 - 2\langle \psi, h\rangle_2|}{\|h\|} = \frac{\|h\|_2^2}{\|h\|} \le \frac{\Lambda\|h\|^2}{\|h\|} = \Lambda\|h\|. \tag{3.109}$$

Moreover, we have

$$2\langle \psi + h, \rangle_2 - 2\langle \psi, \cdot\rangle_2 = 2\langle h, \cdot\rangle_2.$$

We deduce that F is twice differentiable and $D^2F(u)(h, k) = 2\langle h, k\rangle_2$.

In addition, the map H is linear, thus twice Fréchet differentiable. Therefore, $DH(\psi)(h) = H(h)$ 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 obtain

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

The following chain rule, quoted in Theorem 5.2.5 in Siddiqi [Sid04], is needed to get the Fréchet derivatives of $\Theta_{\delta\gamma}$:

Theorem 3.6.5 (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

$$\begin{aligned}&D_\psi \Theta_{\delta\gamma}(t,s,\psi,\phi)(h)\\ &= \frac{2}{\delta}\Big[\langle \psi-\phi, h\rangle_2 + \langle H(\psi-\phi), H(h)\rangle_2\Big]\\ &\quad + 2\gamma e^{1+F(\psi)+F(H(\psi))}[\langle \psi, h\rangle_2 + \langle H(\psi), H(h)\rangle_2].\end{aligned} \tag{3.110}$$

Similarly,

$$\begin{aligned}&D_\phi \Theta_{\delta\gamma}(t,s,\psi,\phi)(k)\\ &= \frac{2}{\delta}\Big[\langle \phi-\psi, k\rangle_2 + \langle H(\phi-\psi), H(k)\rangle_2\Big]\\ &\quad + 2\gamma e^{1+F(\phi)+F(H(\phi))}[\langle \phi, k\rangle_2 + \langle H(\phi), H(k)\rangle_2].\end{aligned} \tag{3.111}$$

Furthermore,

$$\begin{aligned}&D^2_\psi \Theta_{\delta\gamma}(t,s,\psi,\phi)(h,k)\\ &= \frac{2}{\delta}\Big[\langle h, k\rangle_2 + \langle H(h), H(k)\rangle_2\Big]\\ &\quad + 2\gamma e^{1+F(\psi)+F(H(\psi))}\Big[2(\langle \psi, k\rangle_2 + \langle H(\psi), H(k)\rangle_2)(\langle \psi, h\rangle_2 + \langle H(\psi), H(h)\rangle_2)\\ &\quad + \langle k, h\rangle_2 + \langle H(k), H(h)\rangle_2\Big].\end{aligned} \tag{3.112}$$

Similarly,

$$D^2_\phi \Theta_{\delta\gamma}(t,s,\psi,\phi)(h,k) \tag{3.113}$$

$$\begin{aligned}&= \frac{2}{\delta}\Big[\langle h, k\rangle_2 + \langle H(h), H(k)\rangle_2\Big]\\ &\quad + 2\gamma e^{1+F(\phi)+F(H(\phi))}\Big[2(\langle \phi, k\rangle_2 + \langle H(\phi), H(k)\rangle_2)(\langle \phi, h\rangle_2 + \langle H(\phi), H(h)\rangle_2)\\ &\quad + \langle k, h\rangle_2 + \langle H(k), H(h)\rangle_2\Big].\end{aligned} \tag{3.114}$$

By the Hahn-Banach theorem (see, e.g., [Sid04]), we can extent the continuous linear functional T_ϵ to the space $\mathbf{C} \times \mathbf{C}$ and its norm is preserved. Thus, the first-order Fréchet derivatives of T_ϵ is just T_ϵ, that is,

$$D_\psi T_\epsilon(\psi,\phi)h = T_\epsilon(h,\phi),$$

$$D_\phi T_\epsilon(\psi,\phi)k = T_\epsilon(\psi,k) \quad \forall \psi,\phi,h,k \in \mathbf{C}.$$

For the second derivative, we have

$$D^2_\psi T_\epsilon(\psi,\phi)(h,k) = 0, \tag{3.115}$$

$$D^2_\phi T_\epsilon(\psi,\phi)(h,k) = 0, \quad \forall \psi,\phi,h,k \in \mathbf{C}.$$

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 Lemma 2.2.3 and Lemma 2.2.4 in Chapter 2) by setting

$$
\begin{aligned}
&\overline{D_\psi \Theta_{\delta\gamma}(t,s,\psi,\phi)}(h+v\mathbf{1}_{\{0\}}) \\
&= \frac{2}{\delta}\Big[\langle \psi-\phi, h+v\mathbf{1}_{\{0\}}\rangle_2 + \langle H(\psi-\phi), H(h+v\mathbf{1}_{\{0\}}\rangle_2\Big] \\
&\quad + 2\gamma e^{1+F(\psi)+F(H(\psi))}[\langle \psi, h+v\mathbf{1}_{\{0\}}\rangle_2 + \langle H(\psi), H(h+v\mathbf{1}_{\{0\}})\rangle_2].
\end{aligned} \tag{3.116}
$$

and

$$
\begin{aligned}
&\overline{D^2_\psi \Theta_{\delta\gamma}(t,s,\psi,\phi)}(h+v\mathbf{1}_{\{0\}}, k+w\mathbf{1}_{\{0\}}) \\
&= \frac{2}{\delta}\Big[\langle h+v\mathbf{1}_{\{0\}}, k+w\mathbf{1}_{\{0\}}\rangle_2 + \langle H(h+v\mathbf{1}_{\{0\}}), H(k+w\mathbf{1}_{\{0\}}\rangle_2\Big] \\
&\quad +2\gamma e^{1+F(\psi)+F(H(\psi))}\Big[2(\langle \psi, k+w\mathbf{1}_{\{0\}}\rangle_2 + \langle H(\psi), H(k+w\mathbf{1}_{\{0\}})\rangle_2) \\
&\quad \times(\langle \psi, h+v\mathbf{1}_{\{0\}}\rangle_2 + \langle H(\psi), H(h+v\mathbf{1}_{\{0\}})\rangle_2) \\
&\quad +\langle k+w\mathbf{1}_{\{0\}}, h+v\mathbf{1}_{\{0\}}\rangle_2 + \langle H(k+w\mathbf{1}_{\{0\}}), H(h+v\mathbf{1}_{\{0\}}\rangle_2)\Big],
\end{aligned} \tag{3.117}
$$

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

$$
\begin{aligned}
|\langle \psi-\phi, h+v\mathbf{1}_{\{0\}}\rangle_2| &\le \|\psi-\phi\|_2 \cdot \|h+v\mathbf{1}_{\{0\}}\|_2 \\
&\le \Lambda\|\psi-\phi\|_2(\|h\|+|v|),
\end{aligned} \tag{3.118}
$$

$$
\begin{aligned}
|\langle \psi, h+v\mathbf{1}_{\{0\}}\rangle_2| &\le \|\psi\|_2 \cdot \|h+v\mathbf{1}_{\{0\}}\|_2 \\
&\le \Lambda\|\psi\|_2(\|h\|+|v|),
\end{aligned} \tag{3.119}
$$

$$
\begin{aligned}
|\langle \psi, k+w\mathbf{1}_{\{0\}}\rangle_2| &\le \|\psi\|_2 \cdot \|k+w\mathbf{1}_{\{0\}}\|_2 \\
&\le \Lambda\|\psi\|_2(\|k\|+|w|),
\end{aligned} \tag{3.120}
$$

and

$$
\begin{aligned}
|\langle k+w\mathbf{1}_{\{0\}}, h+v\mathbf{1}_{\{0\}}\rangle_2| &\le \|k+w\mathbf{1}_{\{0\}}\|_2\|h+v\mathbf{1}_{\{0\}}\|_2 \\
&\le \Lambda(\|k\|+|w|)(\|h\|+|v|).
\end{aligned} \tag{3.121}
$$

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 $\overline{D^2_\phi \Theta_{\delta\gamma}(t,s,\psi,\phi)}(h+v\mathbf{1}_{\{0\}}, k+w\mathbf{1}_{\{0\}})$.

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

$$
\langle \phi, v\mathbf{1}_{\{0\}}\rangle_2 = \int_{-r}^0 \phi(\theta)\cdot v\mathbf{1}_{\{0\}}(\theta)d\theta = 0, \tag{3.122}
$$

$$
\langle w\mathbf{1}_{\{0\}}, v\mathbf{1}_{\{0\}}\rangle_2 = \int_{-r}^0 w\mathbf{1}_{\{0\}}(\theta)\cdot v\mathbf{1}_{\{0\}}(\theta)d\theta = 0, \tag{3.123}
$$

$$H(v\mathbf{1}_{\{0\}}) = v\mathbf{1}_{\{-r\}}, \tag{3.124}$$

$$\langle H(\psi), H(v\mathbf{1}_{\{0\}})\rangle_2 = 0, \quad \langle H(w\mathbf{1}_{\{0\}}), H(v\mathbf{1}_{\{0\}})\rangle_2 = 0. \tag{3.125}$$

These observations will be used later.

Next, we need several lemmas about the operator $\mathcal{S}$.

Lemma 3.6.6 *Given $\phi \in \mathbf{C}$, we have*

$$\mathcal{S}(F)(\phi) = |\phi(0)|^2 - |\phi(-r)|^2, \tag{3.126}$$

$$\mathcal{S}(F)(\phi^0) = -|\phi(0)|^2, \tag{3.127}$$

where F is the functional defined in (3.107) and $\mathcal{S}$ is the operator defined in (3.60).

Proof. Recall that

$$\mathcal{S}(F)(\phi) = \lim_{t\downarrow 0} \frac{1}{t}\left[F(\tilde{\phi}_t) - F(\phi)\right] \tag{3.128}$$

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

$$\tilde{\phi}(t) = \begin{cases} \phi(t) & \text{if } t \in [-r, 0) \\ \phi(0) & \text{if } t \geq 0, \end{cases} \tag{3.129}$$

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

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

$$= \lim_{t\to 0+} \frac{1}{t}\int_0^t |\phi(0)|^2\, d\theta - \lim_{t\to 0+} \frac{1}{t}\left[\int_{-r}^{-r+t} |\phi(\theta)|^2\, d\theta\right]$$
$$= |\phi(0)|^2 - |\phi(-r)|^2. \tag{3.130}$$

Similarly, we have

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

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

$$\mathcal{S}_\psi(F)(\phi - \psi) + \mathcal{S}_\phi(F)(\phi - \psi) = |\psi(0) - \phi(0)|^2 - |\psi(-r) - \phi(-r)|^2 \tag{3.131}$$

and

$$\mathcal{S}_\psi(F)(\phi^0 - \psi^0) + \mathcal{S}_\phi(F)(\phi^0 - \psi^0) = -|\psi(0) - \phi(0)|^2, \tag{3.132}$$

where F is the functional defined in (3.107) and $\mathcal{S}$ is the operator defined in (3.60).

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

$$\mathcal{S}(F)(\psi) = DF(\psi)\tilde{\mathcal{S}}(\psi), \tag{3.133}$$

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

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

We first assume that $\psi \in \mathcal{D}(\tilde{\mathcal{S}})$, the domain of the operator $\tilde{\mathcal{S}}$, 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

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

On the other hand, by virtue of Lemma 3.6.6, we have

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

Therefore, we have

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

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

$$(\psi - \phi|\tilde{\mathcal{S}}(\psi) - \tilde{\mathcal{S}}(\phi)) = (\psi - \phi|\tilde{\mathcal{S}}(\psi - \phi))$$
$$= \frac{1}{2}\Big[|\psi(0) - \phi(0)|^2 - |\psi(-r) - \phi(-r)|^2\Big]. \tag{3.134}$$

Given the above results, now we can get

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

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 the same argument, we have

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

Lemma 3.6.8 *Given $\phi\in\mathbf{C}$, we define a new operator G as follows*

$$G(\phi)=e^{1+F(\phi)+F(\phi^0)}. \tag{3.135}$$

We have

$$\mathcal{S}(G)(\phi)=(-|\phi(-r)|^2)e^{1+F(\phi)+F(\phi^0)}, \tag{3.136}$$

where F is the functional defined in (4.88) and $\mathcal{S}$ is the operator defined in (3.60).

Proof. Recall that

$$\mathcal{S}(G)(\phi)=\lim_{t\downarrow 0}\frac{1}{t}\left[G(\tilde{\phi}_t)-G(\phi)\right] \tag{3.137}$$

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

$$\tilde{\phi}(t) = \begin{cases} \phi(t) \text{ if } t \in [-r, 0) \\ \phi(0) \text{ if } t \geq 0, \end{cases} \tag{3.138}$$

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

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

We have

$$\begin{aligned}
&\mathcal{S}(G)(\phi) \\
&= \lim_{t \to 0+} \frac{1}{t} \Big[e^{1+\int_{-r}^0 |\tilde{\phi}_t(\theta)|^2 d\theta + \int_{-r}^0 |\tilde{\phi}_t^0(\theta)|^2 d\theta} \\
&\quad - e^{1+\int_{-r}^0 |\phi(\theta)|^2 d\theta + \int_{-r}^0 |\phi^0(\theta)|^2 d\theta} \Big] \\
&= \lim_{t \to 0+} \frac{1}{t} \Big[e^{1+\int_{-r}^0 |\tilde{\phi}(\theta+t)|^2 d\theta + \int_{-r}^0 |\tilde{\phi}^0(\theta+t)|^2 d\theta} \\
&\quad - e^{1+\int_{-r}^0 |\phi(\theta)|^2 d\theta + \int_{-r}^0 |\phi^0(\theta)|^2 d\theta} \Big] \\
&= \lim_{t \to 0+} \frac{1}{t} \Big[e^{1+\int_{-r+t}^t |\tilde{\phi}(\theta)|^2 d\theta + \int_{-r+t}^t |\tilde{\phi}^0(\theta)|^2 d\theta} \\
&\quad - e^{1+\int_{-r}^0 |\phi(\theta)|^2 d\theta + \int_{-r}^0 |\phi^0(\theta)|^2 d\theta} \Big] \\
&= \lim_{t \to 0+} \frac{1}{t} \Big[e^{1+\int_{-r+t}^0 |\phi(\theta)|^2 d\theta + \int_0^t |\phi(0)|^2 d\theta + \int_{-r+t}^0 |\phi^0(\theta)|^2 d\theta + \int_0^t |\phi^0(0)|^2 d\theta} \\
&\quad - e^{1+\int_{-r}^0 |\phi(\theta)|^2 d\theta + \int_{-r}^0 |\phi^0(\theta)|^2 d\theta} \Big] \\
&= \lim_{t \to 0+} \frac{1}{t} \Big[e^{1+\int_{-r+t}^0 |\phi(\theta)|^2 d\theta + t|\phi(0)|^2 + \int_{-r+t}^0 |\phi^0(\theta)|^2 d\theta + t|\phi^0(0)|^2} \\
&\quad - e^{1+\int_{-r}^0 |\phi(\theta)|^2 d\theta + \int_{-r}^0 |\phi^0(\theta)|^2 d\theta} \Big].
\end{aligned} \tag{3.139}$$

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

$$\begin{aligned}
&\mathcal{S}(G)(\phi) \\
&= \lim_{t \to 0+} e^{1+\int_{-r+t}^0 |\phi(\theta)|^2 d\theta + t|\phi(0)|^2 + \int_{-r+t}^0 |\phi^0(\theta)|^2 d\theta + t|\phi^0(0)|^2} \Big(|\phi(0)|^2 \\
&\quad - |\phi(-r+t)|^2 + |\phi^0(0)|^2 - |\phi^0(-r+t)|^2 \Big) \\
&= (|\phi(0)|^2 - |\phi(-r)|^2 - |\phi^0(-r)|^2) e^{1+\int_{-r}^0 |\phi(\theta)|^2 d\theta) + \int_{-r}^0 |\phi^0(\theta)|^2 d\theta} \\
&= (|\phi(0)|^2 - |\phi(-r)|^2 - |\phi(0)|^2) e^{1+\int_{-r}^0 |\phi(\theta)|^2 d\theta + \int_{-r}^0 |\phi^0(\theta)|^2 d\theta} \\
&= -|\phi(-r)|^2 e^{1+F(\phi)+F(\phi^0)}. \qquad \square
\end{aligned} \tag{3.140}$$

Lemma 3.6.9 *For any $\psi, \phi \in \mathbf{C}$, we have*

$$\lim_{\epsilon \downarrow 0} |\mathcal{S}_\psi(T_\epsilon)(\psi, \phi)| = 0 \quad \text{and} \quad \lim_{\epsilon \downarrow 0} |\mathcal{S}_\phi(T_\epsilon)(\psi, \phi)| = 0. \tag{3.141}$$

Proof. We will only prove the first equality in the Lemma, 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 the proof of Lemma 3.6.8. In this case,

$$\begin{aligned}
\lim_{\epsilon \downarrow 0} |\mathcal{S}_\psi(T_\epsilon)(\psi, \phi)| &= \lim_{\epsilon \downarrow 0} \left| \lim_{t \downarrow 0} \frac{T_\epsilon(\tilde{\psi}_t, \phi) - T_\epsilon(\psi, \phi)}{t} \right| \\
&= \lim_{\epsilon \downarrow 0} \left| (T_\epsilon) \lim_{t \downarrow 0} \left(\frac{\tilde{\psi}_t - \psi}{t}, \phi \right) \right| \\
&\le \lim_{\epsilon \downarrow 0} \|T_\epsilon\| \left(\left\| \lim_{t \downarrow 0} \frac{\tilde{\psi}_t - \psi}{t} \right\| + \|\phi\| \right) \\
&\le \lim_{\epsilon \downarrow 0} \epsilon \left(\|\tilde{\mathcal{S}}\psi\| + \|\phi\| \right) = 0,
\end{aligned} \tag{3.142}$$

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

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

Consequently,

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

by the limit process. $\Box$

Given all of the above results, now we are ready to prove Theorem 3.6.1.

Proof of Theorem 3.6.1. Define

$$\Gamma_1(t, \psi) \equiv V_2(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) + \Theta_{\delta\gamma}(t, s_{\delta\gamma\epsilon}, \psi, \phi_{\delta\gamma\epsilon}) - T_\epsilon(\psi, \phi_{\delta\gamma\epsilon}) - M_{\delta\gamma\epsilon} \tag{3.143}$$

and

$$\Gamma_2(s, \phi) \equiv V_1(t_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}) - \Theta_{\delta\gamma}(t_{\delta\gamma\epsilon}, s, \psi_{\delta\gamma\epsilon}, \phi) + T_\epsilon(\psi_{\delta\gamma\epsilon}, \phi) + M_{\delta\gamma\epsilon} \tag{3.144}$$

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 that $\Phi_{\delta\gamma} + T_\epsilon + M_{\delta\gamma\epsilon}$ 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

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

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

$$\left.\begin{aligned} &\alpha V_1(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon}) - \partial_t \Gamma_1(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon}) \\ &\qquad - \sup_{v\in U}[\mathbf{A}^v(\Gamma_1)(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon}) - L(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon},v)] \end{aligned}\right\} \leq 0. \tag{3.145}$$

By the definitions of the operator $\mathbf{A}^v$ and Γ_1 and the fact that the second-order Fréchet derivatives of $T_\epsilon = 0$, we have, by combining (3.110), (3.111), (3.112), (3.113), (3.116), (3.117), (3.122), (3.123), (3.124), and (3.125),

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

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 and the following.

Inequality (3.145) and the above equation together yield that

$$\begin{aligned} &\alpha V_1(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon}) - \mathcal{S}(\Gamma_1)(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon}) - \partial_t \Gamma_1(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon}) \\ &- \sup_{v\in U}\big[-\overline{T_\epsilon(f(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon},v)\mathbf{1}_{\{0\}},\phi_{\delta\gamma\epsilon})} + L(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon},v)\big] \leq 0. \end{aligned} \tag{3.146}$$

Similarly, using the definitions of the viscosity supersolution of V_2 and Γ_2 and by the virtue of the same techniques similar to (3.146), we have

$$\begin{aligned} &\alpha V_2(s_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon}) - \mathcal{S}(\Gamma_2)(s_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon}) - \partial_s \Gamma_2(s_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon}) \\ &- \sup_{v\in U}\big[\overline{T_\epsilon(\psi_{\delta\gamma\epsilon},f(s_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon},v)\mathbf{1}_{\{0\}})} + L(s_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon},v)\big] \geq 0. \end{aligned} \tag{3.147}$$

Inequality (3.146) is equivalent to

$$\begin{aligned} &\alpha V_1(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon}) - \mathcal{S}(\Gamma_1)(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon}) - 2(t_{\delta\gamma\epsilon} - s_{\delta\gamma\epsilon}) \\ &- \sup_{v\in U}\big[-\overline{T_\epsilon(f(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon},v)\mathbf{1}_{\{0\}},\phi_{\delta\gamma\epsilon})} + L(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon})\big] \leq 0. \end{aligned} \tag{3.148}$$

Similarly, Inequality (3.147) is equivalent to

$$\begin{aligned}
&\alpha V_2(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) - \mathcal{S}(\Gamma_2)(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) - 2(s_{\delta\gamma\epsilon} - t_{\delta\gamma\epsilon}) \\
&- \sup_{v\in U} \left[\overline{T_\epsilon(\psi_{\delta\gamma\epsilon}, f(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}, v)\mathbf{1}_{\{0\}})} + L(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}, v)\right] \geq 0.
\end{aligned} \tag{3.149}$$

By virtue of (3.148) and (3.149), we obtain

$$\begin{aligned}
&\alpha(V_1(t_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}) - V_2(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon})) \\
&\leq \mathcal{S}(\Gamma_1)(t_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}) - \mathcal{S}(\Gamma_2)(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) + 4(t_{\delta\gamma\epsilon} - s_{\delta\gamma\epsilon}) \\
&\quad + \sup_{v\in U} [L(t_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}, v) - \overline{T_\epsilon(f(t_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}, v)\mathbf{1}_{\{0\}}, \phi_{\delta\gamma\epsilon})}] \\
&\quad - \sup_{v\in U} [L(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}, v) + \overline{T_\epsilon(\psi_{\delta\gamma\epsilon}, f(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}, v)\mathbf{1}_{\{0\}})}].
\end{aligned} \tag{3.150}$$

From definition (3.60) of $\mathcal{S}$, it is clear that $\mathcal{S}$ is linear and takes the value zero on constants. Recall that

$$\begin{aligned}
\Gamma_1(t, \psi) = {} & V_2(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) + \Theta_{\delta\gamma}(t, s_{\delta\gamma\epsilon}, \psi, \phi_{\delta\gamma\epsilon}) \\
& - T_\epsilon(\psi, \phi_{\delta\gamma\epsilon}) - M_{\delta\gamma\epsilon}
\end{aligned} \tag{3.151}$$

and

$$\begin{aligned}
\Gamma_2(s, \phi) = {} & V_1(t_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}) - \Theta_{\delta\gamma}(t_{\delta\gamma\epsilon}, s, \psi_{\delta\gamma\epsilon}, \phi) \\
& + T_\epsilon(\psi_{\delta\gamma\epsilon}, \phi) + M_{\delta\gamma\epsilon}.
\end{aligned} \tag{3.152}$$

Thus, we have

$$\begin{aligned}
\mathcal{S}(\Gamma_1)(t_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}) = {} & \mathcal{S}_\psi(\Theta_{\delta\gamma})(t_{\delta\gamma\epsilon}, s_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) \\
& \mathcal{S}_\psi(T_\epsilon)(\psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon})
\end{aligned} \tag{3.153}$$

and

$$\begin{aligned}
\mathcal{S}(\Gamma_2)(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) = {} & -\mathcal{S}_\phi(\Theta_{\delta\gamma})(t_{\delta\gamma\epsilon}, s_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) \\
& + \mathcal{S}_\phi(T_\epsilon)(\psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}).
\end{aligned} \tag{3.154}$$

Therefore,

$$\begin{aligned}
&\mathcal{S}(\Gamma_1)(t_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}) - \mathcal{S}(\Gamma_2)(s_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) \\
&= \mathcal{S}_\psi(\Theta_{\delta\gamma})(t_{\delta\gamma\epsilon}, s_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) + \mathcal{S}_\phi(\Theta_{\delta\gamma})(t_{\delta\gamma\epsilon}, s_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) \\
&\quad - [\mathcal{S}_\psi(T_\epsilon)(\psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) + \mathcal{S}_\phi(T_\epsilon)(\psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon})].
\end{aligned} \tag{3.155}$$

Recall that

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

Therefore, we have

$$\begin{aligned}
&\big(\mathcal{S}_\psi(\Theta_{\delta\gamma})+\mathcal{S}_\phi(\Theta_{\delta\gamma})\big)(t_{\delta\gamma\epsilon},s_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon})\\
&\equiv \mathcal{S}_\psi(\Theta_{\delta\gamma})(t_{\delta\gamma\epsilon},s_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon})\\
&\quad+\mathcal{S}_\phi(\Theta_{\delta\gamma})(t_{\delta\gamma\epsilon},s_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon})\\
&=\frac{1}{\delta}[\mathcal{S}_\psi(F)(\psi_{\delta\gamma\epsilon}-\phi_{\delta\gamma\epsilon})+\mathcal{S}_\phi(F)(\psi_{\delta\gamma\epsilon}-\phi_{\delta\gamma\epsilon})\\
&\quad+\mathcal{S}_\psi(F)(\psi^0_{\delta\gamma\epsilon}-\phi^0_{\delta\gamma\epsilon})+\mathcal{S}_\phi(F)(\psi^0_{\delta\gamma\epsilon}-\phi^0_{\delta\gamma\epsilon})]\\
&\quad+\gamma[\mathcal{S}_\psi(G)(\psi_{\delta\gamma\epsilon})+\mathcal{S}_\phi(G)(\phi_{\delta\gamma\epsilon})].
\end{aligned}\tag{3.156}$$

Using Lemma 3.6.7 and Lemma 3.6.8, we deduce that

$$\begin{aligned}
&\big(\mathcal{S}_\psi(\Theta_{\delta\gamma})+\mathcal{S}_\phi(\Theta_{\delta\gamma})\big)(t_{\delta\gamma\epsilon},s_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon})\\
&=\frac{1}{\delta}\big[-|\psi_{\delta\gamma\epsilon}(-r)-\phi_{\delta\gamma\epsilon}(-r)|^2\big]\\
&\quad-\gamma\bigg(|\psi_{\delta\gamma\epsilon}(-r)|^2e^{1+F(\psi_{\delta\gamma\epsilon})+F(\psi^0_{\delta\gamma\epsilon})}\\
&\quad+|\phi_{\delta\gamma\epsilon}(-r)|^2e^{1+F(\phi_{\delta\gamma\epsilon})+F(\phi^0_{\delta\gamma\epsilon})}\bigg)\\
&\le 0.
\end{aligned}\tag{3.157}$$

Thus, by virtue of (3.155) and Lemma 3.6.9, we have

$$\limsup_{\delta\downarrow 0,\epsilon\downarrow 0}\Big[\mathcal{S}(\Gamma_1)(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon})-\mathcal{S}(\Gamma_2)(s_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon})\Big]\le 0.\tag{3.158}$$

Moreover, we know that the norm of T_ϵ is less than ϵ; thus, for any $\gamma>0$, using (3.155) and taking the lim sup on both sides of (3.150) as δ and ϵ go to zero, we obtain

$$\begin{aligned}
&\limsup_{\epsilon\downarrow 0,\delta\downarrow 0}\alpha(V_1(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon})-V_2(s_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon}))\\
&\le\limsup_{\epsilon\downarrow 0,\delta\downarrow 0}\bigg\{\mathcal{S}(\Gamma_1)(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon})-\mathcal{S}(\Gamma_2)(s_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon})\\
&\quad+\sup_{v\in U}\ [L(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon},v)-\overline{T_\epsilon(f(t,\psi_{\delta\gamma\epsilon},v)\mathbf{1}_{\{0\}},\phi_{\delta\gamma\epsilon})}]\\
&\quad-\sup_{v\in U}\ [L(s_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon},v)+\overline{T_\epsilon(\psi_{\delta\gamma\epsilon},f(t,\phi_{\delta\gamma\epsilon},v)\mathbf{1}_{\{0\}})}]\bigg\}\\
&\le\limsup_{\epsilon\downarrow 0,\delta\downarrow 0}\bigg\{\sup_{v\in U}\ \Big|[L(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon},v)-L(s_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon},v)]\Big|\bigg\}.
\end{aligned}\tag{3.159}$$

Using the Lipschitz continuity of L and Lemma 3.6.2, we see that

$$\begin{aligned}
&\limsup_{\delta\downarrow 0,\epsilon\downarrow 0}\sup_{v\in U}\ |L(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon},v)-L(s_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon},v)|\\
&\le\limsup_{\delta\downarrow 0,\epsilon\downarrow 0}C\Big(|t_{\delta\gamma\epsilon}-s_{\delta\gamma\epsilon}|+\|\psi_{\delta\gamma\epsilon}-\phi_{\delta\gamma\epsilon}\|_2\Big)=0;
\end{aligned}\tag{3.160}$$

moreover, by virtue of (3.160), we get

$$\limsup_{\epsilon\downarrow 0,\delta\downarrow 0} \alpha(V_1(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon}) - V_2(s_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon})) \le 0. \tag{3.161}$$

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

$$\Phi_{\delta\gamma}(t,t,\psi,\psi) + T_\epsilon(\psi,\psi) \le \Phi_{\delta\gamma}(t_{\delta\gamma\epsilon}, s_{\delta\gamma\epsilon}, \psi_{\delta\gamma\epsilon}, \phi_{\delta\gamma\epsilon}) + T_\epsilon(\psi_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon}). \tag{3.162}$$

Then we get

$$\begin{aligned}
& V_1(t,\psi) - V_2(t,\psi) \qquad (3.163)\\
\le\ & V_1(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon}) - V_2(s_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon})\\
& -\frac{1}{\delta}\left[\|\psi_{\delta\gamma\epsilon}-\phi_{\delta\gamma\epsilon}\|_2^2 + \|\psi^0_{\delta\gamma\epsilon}-\phi^0_{\delta\gamma\epsilon}\|_2^2 + |t_{\delta\gamma\epsilon}-s_{\delta\gamma\epsilon}|^2\right]\\
& +2\gamma\exp(1+\|\psi\|_2^2+\|\psi^0\|_2^2)\\
& -\gamma(\exp(1+\|\psi_{\delta\gamma\epsilon}\|_2^2+\|\psi^0_{\delta\gamma\epsilon}\|_2^2) + \exp(1+\|\phi_{\delta\gamma\epsilon}\|_2^2+\|\phi^0_{\delta\gamma\epsilon}\|_2^2))\\
& +T_\epsilon(\psi_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon}) - T_\epsilon(\psi,\psi)\\
\le\ & V_1(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon}) - V_2(s_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon})\\
& +2\gamma\exp(1+\|\psi\|_2^2+\|\psi^0\|_2^2) + T_\epsilon(\psi_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon}) - T_\epsilon(\psi,\psi), \qquad (3.164)
\end{aligned}$$

where the last inequality comes from the fact that $\delta > 0$ and $\gamma > 0$. By virtue of (3.161), when we take the lim sup on (3.164) as δ,ϵ and γ go to zero, we can obtain

$$\begin{aligned}
V_1(t,\psi) - V_2(t,\psi) \le\ & \limsup_{\gamma\downarrow 0,\epsilon\downarrow 0,\delta\downarrow 0}\Big(V_1(t_{\delta\gamma\epsilon},\psi_{\delta\gamma\epsilon}) - V_2(s_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon})\\
& +2\gamma\exp(1+\|\psi\|_2^2+\|\psi^0\|_2^2) + T_\epsilon(\psi_{\delta\gamma\epsilon},\phi_{\delta\gamma\epsilon}) - T_\epsilon(\psi,\psi)\Big)\\
\le\ & 0. \qquad (3.165)
\end{aligned}$$

Therefore, we have

$$V_1(t,\psi) \le V_2(t,\psi), \quad \forall (t,\psi)\in[0,T]\times\mathbf{C}. \tag{3.166}$$

This completes the proof of Theorem 3.6.1. □

The uniqueness of the viscosity solution of (3.64) follows directly from this theorem because any viscosity solution is both the viscosity subsolution and supersolution.

Theorem 3.6.10 *The value function $V:[0,T]\times\mathbf{C}\to\Re$ of Problem (OCCP) defined by (3.7) is the unique viscosity solution of the HJBE (3.64).*

Proof. Suppose $V_1, V_2 : [0,T]\times\mathbf{C}\to\Re$ are two viscosity solutions of the HJBE (3.64). Then they are both the viscosity subsolution and supersolution. By Theorem 3.6.1, we have

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

This shows that

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

Therefore, the value function $V : [0, T] \times \mathbf{C} \to \Re$ of Problem (OCCP) is the unique viscosity solution of the HJBE (3.64). $\square$

3.7 Verification Theorems

In this section, conjecture on a version of the verification theorem in the framework of viscosity solutions is presented. The value function $V : [0, T] \times \mathbf{C} \to \Re$ for Problem (OCCP) has been shown to be the unique viscosity solution of the HJBE (3.64) as shown in Sections 3.5 and 3.6. The remaining question for completely solving Problem (OCCP) is the computation of the optimal state-control pair $(x^*(\cdot), u^*(\cdot))$.

The classical verification theorem reads as follows.

Theorem 3.7.1 *Let $\Phi \in C^{1,2}_{lip}([0, T] \times \mathbf{C}) \cap \mathcal{D}(\mathcal{S})$ be the (classical) solution of the HJBE (3.64). Then the following hold:*
(i) $\Phi(t, \psi) \geq J(t, \psi; u(\cdot))$ for any $(t, \psi) \in [0, T] \times \mathbf{C}$ and any $u(\cdot) \in \mathcal{U}[t, T]$.
(ii) Suppose that a given admissible pair $(x^*(\cdot), u^*(\cdot))$ for the optimal classical control problem $(OCP)(t, \psi)$ satisfies

$$0 = \partial_t \Phi(s, x^*_s) + \mathcal{A}^{u^*(s)} \Phi(s, x^*_s) + L(s, x^*_s, u^*(s)) \;\; P-a.s., \;\; a.e. \; s \in [t, T],$$

then $(x^*(\cdot), u^*(\cdot))$ is an optimal pair for $(OCP)(t, \psi)$.

Define the Hamiltonian function $\mathcal{H} : [0, T] \times \mathbf{C} \times \mathbf{C}^* \times \mathbf{C}^\dagger \times U \to \Re$ as follows:

$$\begin{aligned} \mathcal{H}(t, \phi, q, Q, u) = \frac{1}{2} \sum_{j=1}^{m} \bar{Q}(g(t, \phi, u)\mathbf{1}_{\{0\}}\mathbf{e}_j, g(t, \phi, u)\mathbf{1}_{\{0\}}\mathbf{e}_j) \\ + \bar{q}(f(t, \phi, u)\mathbf{1}_{\{0\}}) + L(t, \phi, u), \end{aligned} \tag{3.167}$$

where $\bar{q} \in (\mathbf{C} \oplus \mathbf{B})$ is the continuous extension of q from $\mathbf{C}^*$ to $(\mathbf{C} \oplus \mathbf{B})^*$ and $\bar{Q} \in (\mathbf{C} \oplus \mathbf{B})^\dagger$ is the continuous extension of Q from $\mathbf{C}^\dagger$ to $(\mathbf{C} \oplus \mathbf{B})^\dagger$. (see Lemma 2.2.3 and Lemma 2.2.4 for details.)

We make the following conjecture on verification theorem in the viscosity framework.

Conjecture. (The Generalized Verification Theorem). Let $\bar{V} \in C((0, T] \times \mathbf{C}, \Re)$ be the viscosity supersolution of the HJBE (3.64) satisfying the following polynomial growth condition

$$|\bar{V}(t, \psi)| \leq C(1 + \|\psi\|_2^k) \text{ for some } k \geq 1, \; (t, \psi) \in (0, T) \times \mathbf{C}. \tag{3.168}$$

and such that $\bar{V}(T,\psi) = \Psi(\psi)$. Then we have the following.
(i) $\bar{V}(t,\psi) \geq J(t,\psi;u(\cdot))$ for any $(t,\psi) \in (0,T] \times \mathbf{C}$ and $u(\cdot) \in \mathcal{U}[t,T]$.
(ii) Fix any $(t,\psi) \in (0,T) \times \mathbf{C}$. Let $(x^*(\cdot),u^*(\cdot))$ be an admissible pair for Problem (OCCP). Suppose that there exists

$$(p^*(\cdot),q^*(\cdot),Q^*(\cdot)) \in L^2_{\mathbf{F}(t)}(t,T;\Re) \times L^2_{\mathbf{F}(t)}(t,T;\mathbf{C}^*) \times L^2_{\mathbf{F}(t)}(t,T;\mathbf{C}^\dagger)$$

such that, for a.e. $s \in [t,T]$,

$$(p^*(s),q^*(s),Q^*(s)) \in \underline{D}^{1,2}_{s+,\psi}\bar{V}(s,x^*_s)),\ P-a.s., \tag{3.169}$$

and

$$E\left[\int_t^T [p^*(s) + \mathcal{H}(s,x^*(s),q^*(s),Q^*(s);u^*(s))]ds\right] \geq 0. \tag{3.170}$$

Then $(x^*(\cdot),u^*(\cdot))$ is an optimal pair for **Problem (OCCP)**.

3.8 Finite-Dimensional HJB Equation

It is clear that the HJBE as described in (3.64) is infinite dimensional in the sense that it is a generalized differential equation that involve a first- and second-order Fréchet derivatives of a real-valued function defined on the Banach space $\mathbf{C}$ as well as the infinitesimal generator $\mathcal{S}$. The explicit solution of this equation is not well understood in general. In this section, we investigate some special cases of the infinite-dimensional HJBE (3.64) in which only the regular partial derivatives are involved and of which explicit solutions can be found. Much of the material presented in this section can be found in Larssen and Risebro [LR03]. However, they can be shown to be a special case of (3.1) and the general HJBE (3.64) treated in the previous sections.

3.8.1 Special Form of HJB Equation

In (3.1), we consider the one-dimensional case and assume that $m = 1$ and $n = 1$. Let the controlled drift and diffusion $f,g : [0,T] \times \mathbf{C} \times U \to \Re$ be defined as follows:

$$f(t,\phi,u) = b\left(t,\phi(0),\int_{-r}^0 e^{\lambda\theta}\phi(\theta)\,d\theta,\phi(-r),u\right) \tag{3.171}$$

and

$$g(t,\phi,u) = \sigma\left(t,\phi(0),\int_{-r}^0 e^{\lambda\theta}\phi(\theta)\,d\theta,\phi(-r),u\right) \tag{3.172}$$

for all $(t,\phi,u) \in [0,T] \times \mathbf{C} \times U$, where b and σ are some real-valued functions defined on $[0,T] \times \Re \times \Re \times \Re \times U$ that satisfy the following two conditions.

Assumption 3.8.1 *There exist constants* $K_1 > 0$, *and* $K_2 > 0$ *such that for all* $(t, x, y, z), (t, \bar{x}, \bar{y}, \bar{z}, u) \in [0, T] \times \Re \times \Re \times \Re$ *and* $u \in U$,

$$|b(t, x, y, z, u)| + |\sigma(t, x, y, z, u)| \leq K_1(1 + |x| + |y| + |z|)^p$$

and

$$\begin{aligned}&|b(t, x, y, z, u) - b(t, \bar{x}, \bar{y}, \bar{z}, u)| + |\sigma(t, x, y, z, u) - \sigma(t, \bar{x}, \bar{y}, \bar{z}, u)| \\ &\leq K_2(|x - \bar{x}| + |y - \bar{y}| + |z - \bar{z}|).\end{aligned}$$

It is easy to see that if Assumption 3.8.1 holds for b and σ, then Assumption 3.1.1 holds for f and g that are related through (3.171) and (3.172).

Consider the following one-dimensional control stochastic delay equation:

$$\begin{aligned}dx(s) &= b(s, x(s), y(s), z(s), u(s))\, ds \\ &\quad + \sigma(s, x(s), y(s), z(s), u(s))\, dW(s), \quad s \in (t, T],\end{aligned} \tag{3.173}$$

with the initial data $(t, \psi) \in [0, T] \times C[-r, 0]$, where

$$y(s) = \int_{-r}^{0} e^{\lambda\theta} x(s + \theta)\, d\theta \qquad (\lambda > 0 \text{ is a given constant})$$

represents a weighted (by the factor $e^{\lambda\cdot}$) sliding average of $x(\cdot)$ over the time interval $[s-r, s]$, and $z(s) = x(s-r)$ represents the discrete delay of the state process $x(\cdot)$.

The objective of the control problem is to maximize among $\mathcal{U}[t, T]$ the following expected performance index:

$$J(t, \psi; u(\cdot)) = E\Big[\int_t^T l(s, x(s), y(s), u(s))\, ds + h(x(T), y(T))\Big], \tag{3.174}$$

where $l : [0, T]\times\Re\times\Re\times U \to \Re$ and $h : \Re\times\Re \to \Re$ are the instantaneous reward and the terminal reward functions, respectively, that satisfy the following assumptions:

Assumption 3.8.2 *There exist constants* $K, \bar{K} > 0$ *and* $k \geq 1$ *such that*

$$|l(t, x, y, u)| + |h(x, y)| \leq K(1 + |x| + |y|)^k$$

and

$$|l(t, x, y, u) - l(t, \bar{x}, \bar{y}, u)| + |h(x, y) - h(\bar{x}, \bar{y})| \leq \bar{K}(|x - \bar{x}| + |y - \bar{y}|),$$

for all $(t, x, y, u), (t, \bar{x}, \bar{y}, u) \in [0, T] \times \Re \times \Re \times U$.

We again define the value function $V : [0, T]\times C[-r, 0] \to \Re$ for the optimal control problem (3.173) and (3.174) is defined by

$$V(t, \psi) = \sup_{u(\cdot)\in\mathcal{U}[t,T]} J(t, \psi; u(\cdot)). \tag{3.175}$$

As described in Problem (OCCP), the value function V may depend on the initial datum $(t, \psi) \in [0, T] \times C[-r, 0]$ in a very general and complicated way. In this section, we will show that for a certain class of systems of the form (3.173), the value function depends on the initial function only through the functional of $x \equiv \psi(0)$, $y \equiv \int_{-r}^{0} e^{\lambda\theta}\psi(\theta)\, d\theta$. Let us therefore assume with a little abuse of notation that the value function V takes the following form:

$$V(t, \psi) = \Phi\left(t, \psi(0), \int_{-r}^{0} e^{\lambda\theta}\psi(\theta)\, d\theta\right) = \Phi(t, x, y), \tag{3.176}$$

where $\Phi : [0, T] \times \Re \times \Re \to \Re$. Then the DPP (Theorem 3.3.9) takes the form

$$\begin{aligned}\Phi(t, x, y) = \sup_{u(\cdot)\in\mathcal{U}[t,T]} \Big[&\int_t^\tau e^{-\alpha(\tau-t)} l(s, x(s), y(s), u(s))\, ds \\ &+\Phi(\tau, x(\tau), y(\tau))\Big]\end{aligned} \tag{3.177}$$

for all **F**-stopping times $\tau \in \mathcal{T}_t^T$ and initial datum $(t, \psi(0), \int_{-r}^{0} e^{\lambda\theta}\psi(\theta)\, d\theta) \equiv (t, x, y) \in [0, T] \times \Re^2$.

Lemma 3.8.3 (The Itô Formula) *If $\Phi \in C^{1,2,1}([0, T] \times \Re \times \Re)$, then we have the following Itô formula:*

$$\begin{aligned}d\Phi(s, x(s), y(s)) &= \mathcal{L}^u\Phi(s, x(s), y(s)) \\ &\quad +\partial_x\Phi(s, x(s), y(s))\sigma(s, x(s), y(s))\, dW(s),\end{aligned} \tag{3.178}$$

where $\mathcal{L}^u$ is the differential operator defined by

$$\begin{aligned}\mathcal{L}^u\Phi(t, x, y) &= b(t, x, y, z)\partial_x\Phi(t, x, y) \\ &\quad +\frac{1}{2}\sigma^2(t, x, y, z)\partial_x^2\Phi(t, x, y) \\ &\quad +(x - e^{-\lambda r}z - \lambda z)\partial_y\Phi(t, x, y).\end{aligned} \tag{3.179}$$

Proof. Note that if $\phi \in C[t - r, T]$, then $\phi_s \in C[-r, 0]$ for each $s \in [t, T]$. Since

$$y(x_s) = \int_{-r}^{0} e^{\lambda\theta} x(s + \theta)\, d\theta \ (\lambda \text{ constant}),$$

$$\begin{aligned}\partial_s y(x_s) &= \partial_s\left(\int_{-r}^{0} e^{\lambda\theta} x(s + \theta)\, d\theta\right) \\ &= \partial_s\left(\int_{s-r}^{s} e^{\lambda(t-s)} x(t)\, dt\right) \\ &= x(s) - e^{-\lambda r}x(s - r) - \lambda\int_{s-r}^{s} e^{\lambda(t-s)} x(t)\, dt\end{aligned}$$

The result follows from the classical Itô formula (see Theorem 1.2.15). □

We have the following special form of the HJBE (3.64).

Theorem 3.8.4 *If we assume that (3.176) holds and that* $\Phi \in C^{1,2,1}([0,T] \times \Re \times \Re)$*, then* Φ *solves the following HJBE:*

$$\alpha\Phi(t,x,y) - \partial_t\Phi(t,x,y) - \max_{u \in U} \left[\mathcal{L}^u\Phi(t,x,y) + l(t,x,y,u)\right] = 0, \tag{3.180}$$

$$\forall (t,x,y) \in [0,T] \times \Re \times \Re,$$

with the terminal condition $\Phi(T,x,y) = h(x,y)$*, where* $\mathcal{L}^u$ *is the differential operator defined by (3.179).*

Note that (3.180) is also equivalent to the following:

$$\alpha\Phi(t,x,y) - \partial_t\Phi(t,x,y) + \min_{u \in U} \left[-\mathcal{L}^u\Phi(t,x,y) - l(t,x,y,u)\right] = 0, \tag{3.181}$$

$$\forall (t,x,y) \in [0,T] \times \Re \times \Re.$$

Proof. It is clear that $\Phi : [0,T] \times \Re \times \Re$ defined above is a quasi-tame function. From Itô's formula (see Lemma 3.8.3 in Chapter 2), we have

$$\begin{aligned} dV(s,x(s),y(s)) &= \mathcal{L}^u V(s,x(s),y(s))ds \\ &\quad + \sigma(s,x(s),y(s),z(s))\partial_x V(s,x(s),y(s))dW(s), \end{aligned} \tag{3.182}$$

where the differential operator $\mathcal{L}^u$ is as defined in (3.179). We use the DDP (Theorem 3.3.9) and proceed exactly as in Subsection 3.4.1 to obtain (3.180). □

3.8.2 Finite Dimensionality of HJB Equation

Theorem 3.8.4 indicates that under the assumption that the value function takes the form $\Phi \in C^{1,2,1}([0,T] \times \Re \times \Re)$, we have a finite-dimensional HJBE (3.180) in the sense that it only involves regular partial derivatives such as $\partial_x\Phi$, $\partial_y\Phi$, and $\partial_x^2\Phi$ instead of the Fréchet derivatives and the $\mathcal{S}$-operator as required in (3.64). The question that remains to be answered is under what conditions we can have the finite-dimensional HJBE (3.180). This question will be answered in this subsection.

We consider the following one-dimensional controlled SHDE:

$$\begin{aligned} dx(s) &= [\mu(x(s),y(s)) \\ &\quad + \beta(x(s),y(s))z(s) - g(s,x(s),y(s),u(s))]\, ds \\ &\quad + \sigma(x(s),y(s))\, dW(s), \quad s \in (t,T], \end{aligned} \tag{3.183}$$

with the initial datum $(t,\psi) \in [0,T] \times C([-r,0];\Re)$, where, again, $y(s) = \int_{-r}^{0} e^{\lambda\theta}x(s+\theta)\, d\theta$ and $z(s) = x(s-r)$ are described in the previous subsection and $\mu, \beta, \sigma : \Re \times \Re \to \Re$ and $g : [0,T] \times \Re \times \Re \times U \to \Re$ are the given deterministic functions. Assume that the discount rate $\alpha = 0$ for simplicity. It will be shown in this subsection that the HJBE has a solution depending only on (t, x, y) provided that an auxiliary system of four first-order partial differential equations (PDEs) involving μ, β, g, σ, l, and h has a solution. When this is the case, the HJBE (3.180) reduces to an "effective" equation in only one spatial variable in addition to time.

For this model, the HJBE (3.180) takes the form

$$-\partial_t \Phi + \min_{u \in U} F - (x - e^{-\lambda r} z - \lambda y)\partial_y \Phi = 0, \quad \forall z \in \Re, \tag{3.184}$$

where

$$\begin{aligned} F &= F(t, x, y, z, u, \partial_x \Phi, \partial_y \Phi, \partial_x^2 \Phi) \\ &= -\mathcal{L}^u \Phi(t, x, y) - l(t, x, y, u). \end{aligned} \tag{3.185}$$

Assume that $F^* = \inf_{u \in U} F$. Then

$$-\partial_t \Phi + F^* - (x - e^{-\lambda r} z - \lambda y)\partial_y \Phi = 0, \quad \forall z \in \Re. \tag{3.186}$$

Since this holds for all z, we must have

$$\partial_z (F^* - (x - e^{-\lambda r} z - \lambda y)\partial_y \Phi) = 0.$$

Now, $\partial_{u^*} F^* = 0$ since $u^* \in U$ is a minimizer of the function F. With $\partial_z(x - e^{-\lambda r} z - \lambda y) = -e^{-\lambda r}$, this leads to $\partial_z F^* + e^{-\lambda r}\partial_y \Phi = 0$ or

$$\partial_y \Phi = -e^{\lambda r}\partial_z F^*,$$

which we insert into (3.186) to obtain

$$-\partial_t \Phi + F^* + (x - e^{-\lambda r} z - \lambda y)e^{\lambda r}\partial_z F^* = 0. \tag{3.187}$$

Here, $F^* + (x - e^{-\lambda r} z - \lambda y)e^{\lambda r}\partial_z F^*$ should not depend on z. In the following, let H and G denote generic functions that may depend on $t, x, y, u^*, \partial_x \Phi, \partial_y \Phi$, and $\partial_x^2 \Phi$ but not on z. (H and G may change from line to line in a calculation.) Then the following are equivalent:

$$F^* + e^{\lambda r}\xi \partial_z F^* = H,$$

$$e^{-\lambda r} F^* + \xi \partial_z F^* = H,$$

$$\partial_z \left(\frac{F^*}{\xi} \right) = \frac{\xi \partial_z F^* + e^{-\lambda r} F^*}{\xi^2} = \frac{H}{\xi^2},$$

where $\xi \equiv x - e^{-\lambda r} z - \lambda y$. Integrating this yields

$$\frac{F^*}{\xi} = H \int \frac{dz}{\xi^2} = -He^{\lambda r} \int \frac{d\xi}{\xi^2} = \frac{-He^{\lambda r}}{\xi} + G,$$

so that $F^* = H + G\xi$, which implies that F^* is linear in z; that is,

$$F^* = H + Gz,$$

where H and G are functions that do not depend on z.

Motivated by the above reasoning, we investigate more closely a modified version of (3.173) and consider

$$\begin{aligned} dx(s) = [\bar{\mu}(x(s), y(s), z(s)) - g(s, x(s), y(s), u(s))]ds \\ +\bar{\sigma}(x(s), y(s), z(s))dW(s), \quad s \in (t, T], \end{aligned} \tag{3.188}$$

with the initial datum $(t, \psi) \in [0, T] \times C[-r, 0]$. Recall the performance functional (3.174),

$$J(t, \psi; u(\cdot)) = E\Big[\int_t^T l(s, x(s), y(s), u(s))\, ds + h(x(T), y(T))\Big], \tag{3.189}$$

and the value function $\Phi : [0, T] \times C[-r, 0] \to \Re$ defined by (3.176),

$$\Phi(t, \psi) = \Phi\left(t, \psi(0), \int_{-r}^{0} e^{\lambda\theta}\psi(\theta)\, d\theta\right) = \Phi(t, x, y). \tag{3.190}$$

It is known that if $\Phi = \Phi(t, x, y)$, then Φ satisfies the HJBE

$$-\partial_t\Phi - \bar{\mu}\partial_x\Phi - \frac{1}{2}\bar{\sigma}^2\partial_x^2\Phi - (x - e^{-\lambda r}z - \lambda y)\partial_y\Phi + F(\partial_x\Phi, x, y, t) = 0, \tag{3.191}$$

with the terminal condition

$$\Phi(T, x, y) = h(x, y), \tag{3.192}$$

where

$$F(t, x, y, p) = \inf\{(g(t, x, y, u)p - l(t, x, y, u)\}. \tag{3.193}$$

We wish to obtain conditions on $\bar{\mu}, \bar{\sigma}$, and F that ensure that (3.191) has a solution independent of z. Differentiating (3.191) with respect to z, we obtain

$$\partial_y\Phi - e^{\lambda r}\partial_z\bar{\mu}\partial_x\Phi = e^{\lambda r}\partial_z\bar{\gamma}\partial_x^2\Phi, \tag{3.194}$$

where $\bar{\gamma} = \bar{\sigma}^2/2$. Inserting this into (3.191), this equation now takes the form

$$\begin{aligned} &- \partial_t\Phi - [\bar{\mu} - (z - e^{\lambda r}(x - \lambda y))\partial_z\bar{\mu}]\partial_x\Phi \\ &- [\bar{\gamma} - (z - e^{\lambda r}(x - \lambda y))\partial_z\bar{\gamma}]\partial_x^2\Phi + F(\partial_x\Phi, x, y, t) = 0, \end{aligned}$$

If Φ is to be independent of z, then the coefficients of $\partial_x\Phi$ and $\partial_x^2\Phi$ must be independent of z. By arguments analogous to the previous, we see that

$$\bar{\mu}(x, y, z) = \mu(x, y) + \beta(x, y)z$$

and

$$\bar{\gamma}(x, y, z) = \gamma(x, y) + \zeta(x, y)z$$

for some functions μ, β, γ, and ζ depending on x and y only. Now, since $\bar{\gamma} \geq 0$ for all (x, y, z), we must have $\zeta = 0$, and, consequently, $\partial_z \bar{\gamma} = 0$. Also note that $\partial_z \bar{\mu} = \beta$ and that (3.208) takes the form

$$\partial_y \Phi - e^{\lambda r} \beta(x, y) \partial_x \Phi = 0. \tag{3.195}$$

Using this in (3.191) we see that this equation now reads

$$\begin{aligned} &- \partial_t \Phi - [\mu(x, y) + e^{\lambda r}(x - \lambda y)\beta(x, y)]\partial_x \Phi \\ &- \frac{1}{2}\sigma^2(x, y)\partial_x^2 \Phi + F(\partial_x \Phi, x, y, t) = 0. \end{aligned} \tag{3.196}$$

Now, we introduce new variables $\tilde{x}$ and $\tilde{y}$, such that

$$\frac{\partial}{\partial \tilde{y}} = \frac{\partial}{\partial y} - e^{\lambda r} \beta(x, y) \frac{\partial}{\partial x} \quad \text{and} \quad \frac{\partial}{\partial \tilde{x}} = \frac{\partial}{\partial x}. \tag{3.197}$$

Then (3.195) states that $\partial_{\tilde{y}} \Phi = 0$. In order to be compatible with $\partial_{\tilde{y}} \Phi = 0$, the coefficients of (3.196) and the function h must also be constant in $\tilde{y}$, or

$$\partial_y \hat{\mu} - e^{\lambda r} \beta \partial_x \hat{\mu} = 0, \tag{3.198}$$

$$\partial_y \sigma - e^{\lambda r} \beta \partial_x \sigma = 0, \tag{3.199}$$

$$e^{\lambda r} p \partial_p F \partial_x \beta + \partial_y F - e^{\lambda r} \beta \partial_x F = 0, \tag{3.200}$$

$$\partial_y h - e^{\lambda r} \beta \partial_x h = 0, \tag{3.201}$$

where

$$\hat{\mu}(x, y) = \mu(x, y) + e^{\lambda r}(x - \lambda y)\beta(x, y).$$

To see why $\partial_{\tilde{y}} F = 0$ is equivalent to (3.200), note that

$$\begin{aligned} \partial_{\tilde{y}} F &= \partial_p F \partial_y (\partial_x \Phi) + \partial_y F - e^{\lambda r} \beta (\partial_p F \partial_x \Phi_x + \partial_x F) \\ &= \partial_p F (\partial_{yx}^2 \Phi - e^{\lambda r} \beta \partial_x^2 \Phi) + \partial_y F - e^{\lambda r} \beta \partial_x F \\ &= \partial_p F [\partial_x (\partial_y \Phi - e^{\lambda r} \beta \partial_x \Phi) + e^{\lambda r} \partial_x \beta \partial_x \Phi] + \partial_y F - e^{\lambda r} \beta \partial_x F \\ &= e^{\lambda r} p \partial_p F \partial_x \beta + \partial_y F - e^{\lambda r} \beta \partial_x F \quad \text{by (3.195).} \end{aligned} \tag{3.202}$$

Conversely, if $\bar{\mu} = \mu(x, y) + \beta(x, y)z$ and $\bar{\sigma} = \sigma(x, y)$, and (3.198)-(3.200) hold, then we can find a solution of (3.195) that is independent of z.

We collect this in the following theorem.

Theorem 3.8.5 *The HJBE (3.195) with terminal condition $\Phi(T, x, y) = h(x, y)$ has a viscosity solution $\Phi = \Phi(t, x, y)$ if and only if $\bar{\mu} = \mu(x, y) + \beta(x, y)z$ and $\bar{\sigma} = \sigma(x, y)$, and (3.198)-(3.201) hold. In this case, in the coordinates given by (3.197), the HJBE (3.195) reads*

$$-\partial_t \Phi - \hat{\mu}(\tilde{x}) \partial_{\tilde{x}}^2 \Phi + F(\partial_{\tilde{x}} \Phi, \tilde{x}, t) - \frac{1}{2}\sigma^2(\tilde{x}) \partial_{\tilde{x}}^2 \Phi = 0 \tag{3.203}$$

with the terminal condition

$$\Phi(T, \tilde{x}) = h(\tilde{x}). \tag{3.204}$$

3.8.3 Examples

In this subsection we present two examples that satisfy the requirements (3.198)-(3.201), and also indicate why it is difficult to find more general examples that can be completely solvable.

Example 1 (Harvesting with Exponential Growth) Assume that the size $x(\cdot)$ of a population obeys the linear SHDE

$$\begin{aligned} dx(s) &= (ax(s) + by(s) + cz(s) - u(s))\, ds \\ &\quad +(\sigma_1 x(s) + \sigma_2 y(s))dW(s), \quad s \in [t, T], \end{aligned} \tag{3.205}$$

with the initial datum $(t, \psi) \in [0, T] \times C[-r, 0]$. We assume that $x(s) > 0$. The population is harvested at a rate $u(s) \geq 0$, and we are given the performance functional

$$\begin{aligned} J(t, \psi; u(\cdot)) = E^{t,\psi,u(\cdot)} \Big[\int_t^T & \{l_1(x(s), y(s)) + l_2(u(s))\}\, ds \\ & +h(x(T), y(T)) \Big], \end{aligned} \tag{3.206}$$

where T is the stopping time defined by

$$T = \left\{ T_1, \inf_{s>t} \; \{x(s; t, \psi, u(\cdot) = 0\} \right\} \tag{3.207}$$

and $T_1 > t$ is some finite deterministic time. If the value function Φ takes the form $\Phi(t, x, y)$, then Φ satisfies the HJBE

$$\begin{aligned} & - \partial_s \Phi - (x - e^{-\lambda r} z - \lambda y)\partial_y \Phi - \frac{1}{2}(\sigma_1 x + \sigma_2)^2 \partial_x^2 \Phi \\ & + \inf_u \{-(ax + by + cz - u)\partial_x \Phi - l_1(x, y) - l_2(u)\} = 0 \end{aligned} \tag{3.208}$$

Using Theorem 3.8.5, from (3.198) and (3.199) we find that the parameters must satisfy the relations

$$\sigma_2 = \sigma_1 c e^{\lambda r}, \quad b - \lambda c e^{\lambda r} = c e^{\lambda r}(a + c e^{\lambda r}). \tag{3.209}$$

The function F defined in (3.193) now has the form

$$F(p, x, y) = \inf_{u \in U} \{pu - l_2(u)\} - l_1(x, y) = pu^* - l_2(u^*) - l_1(x, y),$$

where u^* is the minimizer in U. Then from (3.200) we see that we must have

$$\partial_y l_1 - c e^{\lambda r} \partial_x l_1 = 0, \tag{3.210}$$

or $l_1 = l_1(x + c e^{\lambda r} y)$. Introducing the variable $\tilde{x} = x + c e^{\lambda r} y$ and the constant $\kappa = a + c e^{\lambda r}$, we find that the "effective" equations (3.203) and (3.204) in this case will be

$$-\partial_s \Phi - (\kappa \tilde{x} - u^*)\partial_{\tilde{x}}\Phi - \frac{1}{2}\sigma_1^2 \tilde{x}^2 \partial_{\tilde{x}}^2 \Phi - l_1(\tilde{x}) - l_2(u^*) = 0, \tag{3.211}$$

with the terminal condition

$$\Phi(T, \tilde{x}) = h(\tilde{x}), \tag{3.212}$$

assuming h satisfies (3.201). This corresponds to the control problem without delay with system dynamics

$$d\tilde{x}(s) = (\kappa \tilde{x}(s) - u)\, ds + \sigma_1 \tilde{x}(s)\, dW(s), \quad s \in (t, T],$$

and $\tilde{x}(t) = \tilde{x} \geq 0$.

To close the discussion of this example, let us be specific and choose

$$l_1(x, y) = -c_0|x + ce^{\lambda r} y - m|, \quad l_2(u) = c_1 u - c_2 u^2, \quad h = 0, \tag{3.213}$$

where c_0, c_1, c_2, and m are positive constants. Then (3.210) and (3.201) hold and

$$F(p, x, y) = \inf_{u \in U} \{c_2 u^2 - (c_1 - p)u\} + c_0|x + ce^{\lambda r} y - m|.$$

We solve for u and find that the optimal harvesting rate is given by

$$u^* = \max\left\{\frac{c_1 - \partial_x \Phi}{2c_2}, 0\right\}. \tag{3.214}$$

Insert (3.213) and (3.214) into the HJBEs (3.211) and (3.212). The resulting equation is a second-order PDE that may be solved numerically and the optimal control can be found provided that the solution of the HJBE really is the value function.

Example 2 (Resource Allocation). Let $x(\cdot) = \{x(s), s \in [t - r, T]\}$ denote a population developing according to (3.205). One can think of $x(\cdot)$ as a wild population that can be caught and bred in captivity and then harvested. The population in captivity, $\hat{x}(\cdot)$, develops according to

$$d\hat{x}(s) = (\gamma \hat{x}(s) + u(s) - v(s))\, ds, \quad s \in (t, T], \tag{3.215}$$

with $\hat{x}(t) = \hat{x} \geq 0$, where v denotes the harvesting rate. The state and control processes for the control problem are $(x(\cdot), \hat{x}(\cdot))$ and $(u(\cdot), v(\cdot))$, respectively. For this case, we consider the gain functional

$$\begin{aligned} J(t, \psi, \hat{x}; u(\cdot), v(\cdot)) = E^{t,\psi,\hat{x};u(\cdot),v(\cdot)} \Big[&\int_t^T (l(v(s)) - c_1 \hat{x}(s) - c_2 u^2(s))\, ds \\ &+ h(x(T), y(T), \hat{x}(T)) \Big], \end{aligned} \tag{3.216}$$

where T is again given by (3.207), $l(v)$ denotes the utility from consumption or sales of the animals, $c_1 \hat{x}$ models the cost of keeping the population, and $c_2 u^2$ models the cost of catch and transfer. Setting

$$V(t, \psi, \hat{x}) = \sup_{u \geq 0, v} J(t, \psi, \hat{x}; u(\cdot), v(\cdot)),$$

we find that if V takes the special form $V = \Phi(t, x, y, \hat{x})$, then

$$\begin{aligned} &- \partial_t \Phi - (ax + by + cz)\partial_x \Phi - \gamma \hat{x} \partial_{\hat{x}} \Phi - \frac{1}{2}(\sigma_1 x + \sigma_2 y + \sigma_3 z)^2 \partial_x^2 \Phi \\ &- (x - e^{-\lambda r} z - \lambda y)\partial_y \Phi + c_1 \hat{x} + F(\partial_x \Phi, \partial_{\hat{x}} \Phi) = 0 \end{aligned} \tag{3.217}$$

and $\Phi(T, x, y, \hat{x}) = h(x, y, \hat{x})$, where

$$F(p, q) = \inf_{u \geq 0, v \leq v_{max}} (c_2 u^2 - u(q - p) + vq - l(v)).$$

Since v is independent of z, we must demand that the parameters satisfy (3.209), and we introduce $\tilde{x}$ as before to find that $\Phi = \Phi(t, \tilde{x}, \hat{x})$ satisfies

$$-\partial_t \Phi - \kappa \tilde{x} \partial_{\tilde{x}} \Phi - \gamma \hat{x} \partial_{\hat{x}} \Phi - \frac{1}{2}\sigma_1^2 \tilde{x}^2 \partial_{\tilde{x}}^2 \Phi + c_1 \hat{x} + F(\partial_{\tilde{x}} \Phi, \partial_{\hat{x}} \Phi) = 0, \quad \text{for } t < T \tag{3.218}$$

and $\Phi(T, \tilde{x}, \hat{x}) = h(\tilde{x}, \hat{x})$. Again, the above PDE with the terminal condition can be solved numerically.

3.9 Conclusions and Remarks

This chapter develops the infinite-dimensional HJBE for the value function of the discounted optimal classical control problem over finite time horizon. The HJBE involves extensions of first- and second-order Fréchet derivatives as well as the shift operator, which are unique in controlled SHDEs. This distinguishes them from all other infinite-dimensional stochastic control problems such as the ones arising from stochastic partial differential equations. The main theme of this chapter is to show under very reasonable assumptions that the value function is the unique viscosity solution of the HJBE. Existence of optimal control as well as special cases that lead to a finite dimensional HJBE are demonstrated. There is no attempt to treat the ergodic controls and/or the combined classical-singular control problem. However, a combined classical-impulse control arising from a hereditary portfolio optimization problem is treated in Chapter 7 in detail.

Springer
http://www.springer.com/978-0-387-75805-3

Stochastic Control of Hereditary Systems and
Applications
Chang, M.-H.
2008, XVIII, 406 p., Hardcover
ISBN: 978-0-387-75805-3