A class of stochastic optimal control problems with state constraint: an extended version

Hitoshi Ishii * and Paola Loreti*

July 16, 2001

Abstract We investigate, via the dynamic programming approach, optimal control problems of infinite horizon with state constraint, where the state X_t is given as a solution of a controlled stochastic differential equation and the state constraint is described either by the condition that $X_t \in \overline{G}$ for all $t > 0$ or by the condition that $X_t \in G$ for all $t > 0$, where G be a given open subset of $\mathbf{R}^N$. Under the assumption that for each $z \in \partial G$ there exists a continuous map $a_z : \mathbf{R}^N \to A$, where A denotes the control set, such that the diffusion matrix $\sigma(x, a)$, with $a = a_z(x)$, vanishes for $x \in \partial G$ in a neighborhood of z and the drift vector $b(z, a)$, with $a = a_z(z)$, directs inside of G at z together with some other mild assumptions, we establish the unique existence of a continuous viscosity solution of the state constraint problem for the associated Hamilton-Jacobi-Bellman equation, prove that the value functions V associated with the constraint $\overline{G}$, $V_{\#}$ of the problem associated with the constraint $\overline{G}$, where only a finite number of selection is allowed for the controller, V_r of the relaxed problem associated with the constraint $\overline{G}$, and V_0 associated with the constraint G, satisfy in the viscosity sense the state constraint problem, and establish Lipschitz or Hölder regularity results for the viscosity solution of the state constraint problem.

Key Words: Stochastic optimal control, State constraint, Hamilton-Jacobi-Bellman equations, Viscosity solutions, Degenerate elliptic equations.

AMS Classification: 93E20, 70H20, 35E70, 49L25.

1. Introduction

*Department of Mathematics, School of Education, Waseda University, Nishi-Waseda 1-6-1, Shinjuku-ku, Tokyo 169-8050, Japan.

*Dipartimento di Metodi e Modelli Matematici per le Scienze Applicate, Università di Roma “La Sapienza”, Via Scarpa n. 16, 00161 Roma, Italy.

We investigate optimal control problems of infinite horizon with state constraint via the dynamic programming approach.

To explain our control problems, we first introduce the controlled systems α at $x \in \mathbf{R}^N$ as the collections

$$\alpha \equiv (\Omega^\alpha, \mathcal{F}^\alpha, P^\alpha, \{\mathcal{F}_t^\alpha\}_{t\geq 0}, \{W_t^\alpha\}_{t\geq 0}, \{u_t^\alpha\}_{t\geq 0}, \{X_t^\alpha\}_{t\geq 0}),$$

where $(\Omega^\alpha, \mathcal{F}^\alpha, P^\alpha, \{\mathcal{F}_t^\alpha\}_{t\geq 0})$ is a filtered probability space satisfying the usual condition (see e.g. [YZ]), $\{W_t^\alpha\}_{t\geq 0}$ is a standard l-dimensional Brownian motion on this filtered probability space, $\{u_t^\alpha\}_{t\geq 0}$ is an $\mathcal{F}_t^\alpha$–progressively measurable stochastic process taking values in a given control set A, and $\{X_t^\alpha\}_{t\geq 0}$ is a strong solution of the stochastic differential equation (SDE for short)

$$dX_t = b(X_t, u_t^\alpha)dt + \sigma(X_t, u_t^\alpha)dW_t^\alpha, \quad X_0 = x, \tag{1}$$

where $b : \mathbf{R}^N \times A \to \mathbf{R}^N$ and $\sigma : \mathbf{R}^N \times A \to \mathbf{R}^{N\times l}$ are given functions. The set of controlled systems at x will be denoted by $\mathcal{C}(x)$. Having regard to the dependence of $\mathcal{C}(x)$ on A, b and σ, we call $\alpha \in \mathcal{C}(x)$ a controlled system at x *associated with* A, b and σ as well.

Let $G \subset \mathbf{R}^N$ be a given open set. For each $x \in \overline{G}$, $\mathcal{A}(x)$ denotes the set of those $\alpha \in \mathcal{C}(x)$ for which

$$X_t^\alpha \in \overline{G} \qquad \forall t \geq 0 \quad P^\alpha\text{–a.s.}$$

Now let $\lambda > 0$ be a given constant. The cost functional and value function are defined, respectively, as

$$J(x,\alpha) = E^\alpha \int_0^\infty e^{-\lambda t} f(X_t^\alpha, u_t^\alpha)dt \tag{2}$$

for all $x \in \mathbf{R}^N$ and $\alpha \in \mathcal{C}(x)$, where E^α denotes the mathematical expectation with respect to P^α, and

$$V(x) = \inf_{\alpha\in\mathcal{A}(x)} J(x,\alpha) \tag{3}$$

for any $x \in \overline{G}$.

In the dynamic programming approach, one of most important aspects is the identification of the value function as a solution u of the associated Hamilton-Jacobi-Bellman (HJB for short) equation, i.e., the equation

$$\lambda u(x) + H(x, Du(x), D^2u(x)) = 0,$$

where

$$H(x,p,X) := \sup_{a\in A}\{-\frac{1}{2}\operatorname{tr}\sigma\sigma^T(x,a)X - b(x,a)\cdot p - f(x,a)\}.$$

As is well-known, the value function V is not so smooth in general that the HJB equation above makes the classical sense. It is nowadays well recognized that the best way to interpret the HJB equation above is to adapt the notion of viscosity solution. In this paper we mostly study our control problems in this line.

The study of optimal control with state constraint in this framework goes back to P.-L. Lions [LB], where the case of all possible states being confined in a given bounded set was studied for deterministic control problems, i.e. the case when $\sigma = 0$. Later, H. M. Soner [S] developed the theory of optimal control with state constraint in the deterministic case, especially introducing a sufficient condition for the continuity of the value function, introducing an appropriate boundary problem for the corresponding HJB equation and identifying the value function as the unique continuous viscosity solution of this boundary value problem. Many other authors contributed to develop further in this direction. Here we refer in particular to the formulation in H. Ishii and S. Koike [IK], which is a modification of the value boundary problem introduced by Soner, which has the advantage to have uniqueness of viscosity solutions among bounded (and possibly discontinuous) functions, and which we rely on in this paper.

In the stochastic case, the first contribution was due to J.-M. Lasry and P.-L. Lions [LL] and in their paper they dealt with the case of nondegenerate diffusion (i.e., the case where $\sigma =$ the identity matrix) and unbounded drift b so that the value function behaves singularly near the boundary ∂G. M. Katsoulakis [KA] initiated to study the case where diffusion depends on the control and degenerates on the boundary. G. Barles and J. Burdeau [BB] studied the Dirichlet problem for degenerate elliptic equations, obtaining a continuity result for the value functions under the assumption that the diffusion coefficient depends only on the state variable but not on the control (i.e., $\sigma = \sigma(x)$). The study of the Dirichlet problem was further developed by G. Barles and E. Rouy [BR].

The main results of this paper concern: (i) the identification of different kinds of value functions for the control problem described above as the viscosity solution of

$$\begin{cases} \lambda u(x) + H(x, Du(x), D^2u(x)) \geq 0 & \text{in } \overline{G}, \\ \lambda u(x) + H_{in}(x, Du(x), D^2u(x)) \leq 0 & \text{in } \overline{G}, \end{cases} \tag{4}$$

where

$$H_{in}(x,p,X) = \sup_{a\in A(x)}\{-\frac{1}{2}\operatorname{tr}\sigma\sigma^T(x,a)X - b(x,a)\cdot p - f(x,a)\},$$

with $A(x)$ the subset of A consisting of those a such that $\sigma(x,a)=0$ and $b(x,a)$ directs inside of G at x (see the next section for the precise definition of $A(x)$), and (ii) the Lipschitz or Hölder regularity of the value functions.

Regarding the identification, our results are close to those obtained by [BB] and the new feature beyond [BB] in our result is dependence of σ in a. Related results can be found in [BR] in the framework of the Dirichlet problem. On the other hand our degeneracy assumption on σ on the boundary is stronger than those in [BB] and [KA]. [KA] studies a different case from ours, at least, for the continuity result of value functions. In our results we consider four kinds of value functions, the identification of which is new in the setting of stochastic control. For the deterministic case, we refer to [L]. Again the Lipschitz or Hölder continuity results are new in the setting of stochastic control. To our knowledge, in the literature just the continuity of value functions is studied. For the deterministic case, we refer to [CDL], [LT], and [IK]. Many results of this paper can be extended to the case of differential games problems, we will not pursue this here in order to make the paper concise.

2. Statements of the problems and main results

Let A be a compact metric space with metric d_A and let

$$\sigma:\mathbf{R}^N\times A\to\mathbf{R}^{N\times l},\qquad b:\mathbf{R}^N\times A\to\mathbf{R}^N,\qquad f:\mathbf{R}^N\times A\to\mathbf{R}$$

be given continuous functions which satisfy:

(A1) There is a constant $M>0$ such that for $\phi=\sigma,b,f$,

$$\sup_{a\in A}\|\phi(\cdot,a)\|_{W^{1,\infty}(\mathbf{R}^N)}\le M.$$

(A2) There exists a continuous function $m:[0,+\infty)\to[0,+\infty)$, with $m(0)=0$, such that for $\phi=\sigma,b,f$,

$$|\phi(x,a)-\phi(x,a')|\le m(d_A(a,a'))\qquad \forall a,a'\in A,\forall x\in\mathbf{R}^N.$$

We introduce the notation: for $x,b\in\mathbf{R}^N$ and $r>0$ let $TC(x,b,r)$ denote the truncated cone

$$TC(x,b,r)=\bigcup_{0\le t\le r}B(x+tb,r).$$

Note that $TC(x,b,r)$ has the shape of a truncated cone if $|b|>r$ and it is a ball if $|b|\le r$.

Other assumptions we use are:

(A3) G is an open, bounded subset of $\mathbf{R}^N$.

(A4) For any $z \in \partial G$, there exist a constant $r > 0$ and a continuous map $\hat{a} : B(z,r) \cap \overline{G} \to A$ such that

(i) $\sigma(x, \hat{a}(x)) = 0 \qquad \forall x \in B(z,r) \cap \partial G$,

(ii) $TC(x, b(x, \hat{a}(x)), r) \subset \overline{G} \qquad \forall x \in B(z,r) \cap \overline{G}$.

(A5) For each $z \in \partial G$ there are a constant $r > 0$ and a continuous map $\hat{a} : B(z,r) \to A$ such that

(i) the maps $x \mapsto \sigma(x, \hat{a}(x))$ and $x \mapsto b(x, \hat{a}(x))$ are Lipschitz continuous on $B(z,r)$,

(ii) $\sigma(x, \hat{a}(x)) = 0 \qquad \forall x \in B(z,r) \cap \partial G$,

(iii) $TC(x, b(x, \hat{a}(x)), r) \subset \overline{G} \qquad \forall x \in B(z,r) \cap \overline{G}$.

(A6) For each $x \in \mathbf{R}^N$, the set $\{(\sigma\sigma^T(x,a), b(x,a), f(x,a)) \mid a \in A\}$ is convex.

(A7) The set A is a convex subset of a normed space with norm $|\cdot|$ and there is a constant $M > 0$ such that for any $\phi = \sigma, b$ and for all $x \in \mathbf{R}^N$, $a, a' \in A$,

$$|\phi(x,a) - \phi(x,a')| \leq M|a - a'|,$$

and moreover there are a Lipschitz continuous function $\hat{a} : \mathbf{R}^N \to A$ and a constant $r > 0$ such that

(i) $\sigma(x, \hat{a}(x)) = 0 \qquad \forall x \in \partial G$,

(ii) $TC(x, b(x, \hat{a}(x)), r) \subset \overline{G} \qquad \forall x \in \overline{G}$.

Remark. It is clear that (A7) implies (A5) and that (A5) implies (A4). The boundedness assumption in (A3) could be replaced by the uniformity in z in (A4) in the results of this paper. Moreover the Lipschitz continuity of f in (A1) is only needed to obtain the Lipschitz property of the solution of (4), and it can be replaced by the Hölder continuity or just the continuity of f in x in the assertion of Hölder continuity of solution of (4) or in other results, respectively.

First of all we consider the problem (4). To be precise, we let $A(x)$, in the definition of H_{in}, denote the subset of A consisting of those a such that there are a constant $r > 0$ and a continuous map $\hat{a} : B(x,r) \cap \overline{G} \to A$ satisfying $\hat{a}(x) = a$ for which (i), (ii) of (A4) hold with x in place of z.

Remark. We note that $H_{in}(x,p,X) = H(x,p,X)$ if $x \in G$.

Theorem 1. *Assume (A1), (A3), and (A4). Then there exists a unique viscosity solution $U \in C(\overline{G})$ of the problem (4).*

In what follows, under the assumptions of Theorem 1, we write U for the unique viscosity solution of (4).

The theorem above extends the existence and uniqueness result in [IK]. We refer the reader to [BR] for results closely related to the above, in which the "strong" comparison principle has been established under a similar but more general assumption.

Next we consider the identification of value functions with the solution of (4).

First Control Problem : state constraint in $\overline{G}$.

The control problem with state constraint in $\overline{G}$ is already described in the introduction. The value function V associated with this control problem is defined by (3) with help of the sets $\mathcal{A}(x)$, $x \in \overline{G}$. For each $x \in \overline{G}$ we call a controlled system $\alpha \in \mathcal{A}(x)$ admissible at $x \in \overline{G}$ with respect to $\overline{G}$, i.e., $\alpha \in \mathcal{C}(x)$ is admissible with respect to $\overline{G}$ if X_t^α satisfies

$$X_t^\alpha \in \overline{G} \qquad \forall t \geq 0 \quad P^\alpha\text{–a.s.}$$

Second Control Problem : finite selection.

Henceforth we let $I(n)$ denote the set $\{1, ..., n\}$ for $n \in \mathbf{N}$. Let $\mathcal{E}$ denote the family of all finite subsets of $C(\mathbf{R}^N, A)$. Let $E \equiv \{a_1, ..., a_n\} \in \mathcal{E}$. Define $\sigma^* \in C(\mathbf{R}^N \times I(n), \mathbf{R}^{N\times l})$, $b^* \in C(\mathbf{R}^N \times I(n), \mathbf{R}^N)$, and $f^* \in C(\mathbf{R}^N \times I(n), \mathbf{R})$ by

$$\sigma^*(x,i) = \sigma(x, a_i(x)), \quad b^*(x,i) = b(x, a_i(x)), \quad \text{and} \quad f^*(x,i) = f(x, a_i(x)).$$

Define $\mathcal{C}_E(x)$, $\mathcal{A}_E(x)$, $J_E(x,\alpha)$, and $V_E(x)$ in the same way as $\mathcal{C}(x)$, $\mathcal{A}(x)$, $J(x,\alpha)$, and $V(x)$, respectively, but replacing σ, b, f, and A by σ^*, b^*, f^*, and $I(n)$. For this control problem the controller has only n choices at each moment.

If $x \in \overline{G}$ and

$$\alpha \equiv (\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\geq 0}, \{W_t\}_{t\geq 0}, \{i_t\}_{t\geq 0}, \{X_t\}_{t\geq 0}) \in \mathcal{A}_E(x)$$

and if we set

$$u_t = a_{i_t}(X_t) \qquad \forall t \geq 0,$$

then it is immediate to see that

$$\beta \equiv (\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\geq 0}, \{W_t\}_{t\geq 0}, \{u_t\}_{t\geq 0}, \{X_t\}_{t\geq 0}) \in \mathcal{A}(x),$$

and $J_E(x,\alpha) = J(x,\beta)$. This shows that

(5)
$$V(x) \leq V_E(x) \qquad \forall E \in \mathcal{E},\ x \in \overline{G}.$$

Finally define $V_\# : \overline{G} \to \mathbf{R}$ by

$$V_\#(x) = \inf_{E\in\mathcal{E}} V_E(x).$$

Third Control Problem : relaxation.

We introduce the space $\Gamma \equiv \Gamma_A$ of those Borel measures γ on $[0,\infty) \times A$ for which

$$\gamma([0,t] \times A) = t \qquad \forall t \geq 0.$$

By virtue of Prohorov thereoem, the space Γ can be equipped with metric ρ with which it is a compact metric space and the convergence in ρ is equivalent with the weak convergence of measures on each $[0,T]\times A$, with $T>0$ (see [IW] and [SV]).

Define

$$\mathcal{B}_t(\Gamma)=\sigma\Big\{\{\gamma\in\Gamma\mid\gamma(B)\in C\}\ \Big|\ B\subset[0,t]\times A\ \text{ and }\ C\subset\mathbf{R}\text{ are Borel subsets}\Big\}.$$

This family $\{\mathcal{B}_t(\Gamma)\}$ provides a natural filtration on Γ.

As before we define the set $\mathcal{C}_r(x)$ for $x\in\mathbf{R}^N$ as the set of the collections

$$\alpha\equiv(\Omega^\alpha,\mathcal{F}^\alpha,P^\alpha,\{\mathcal{F}_t^\alpha\}_{t\ge0},\gamma^\alpha,\{X_t^\alpha\}_{t\ge0}),$$

where $(\Omega^\alpha,\mathcal{F}^\alpha,P^\alpha,\{\mathcal{F}_t^\alpha\}_{t\ge0})$ is a filtered probability space satisfying the usual condition, γ^α is an $\mathcal{F}_t^\alpha$–progressively measurable (i.e., $\mathcal{F}_t^\alpha/\mathcal{B}_t(\Gamma)$–measurable) Γ–valued random variable, and $\{X_t^\alpha\}_{t\ge0}$ is an $\mathcal{F}_t^\alpha$–adapted continuous stochastic process satisfying $X_0^\alpha=x$ such that for any $\phi\in C_b^2(\mathbf{R}^N)$,

$$\phi(X_t^\alpha)-\int_{[0,t]\times A}\mathcal{L}^a\phi(X_s^\alpha)\gamma^\alpha(dsda)$$

is an $\mathcal{F}_t^\alpha$–martingale with $\mathcal{L}^a$ denoting the linear operator on $C^2(\mathbf{R}^N)$ defined by

$$\mathcal{L}^a\phi(x)=\frac{1}{2}\operatorname{tr}\sigma\sigma^T(x,a)D^2\phi(x)+b(x,a)\cdot D\phi(x).$$

We define $\mathcal{A}_r(x)$ for $x\in\overline{G}$ as the set of those $\alpha\in\mathcal{C}_r(x)$ for which

$$X_t^\alpha\in\overline{G}\quad\forall t>0\qquad P^\alpha\text{–a.s.}$$

Finally we define the (relaxed) value function V_r by

$$V_r(x)=\inf_{\alpha\in\mathcal{A}_r(x)}E^\alpha\int_{[0,\infty)\times A}e^{-\lambda t}f(X_t^\alpha,a)\gamma^\alpha(dtda)\qquad\forall x\in\overline{G}.$$

A standard remark here is that $V(x)\ge V_r(x)$ for all $x\in\overline{G}$. Indeed, for $x\in\overline{G}$, if

$$\alpha\equiv(\Omega^\alpha,\mathcal{F}^\alpha,P^\alpha,\{\mathcal{F}_t^\alpha\}_{t\ge0},\{W_t^\alpha\}_{t\ge0},\{u_t^\alpha\}_{t\ge0},\{X_t^\alpha\}_{t\ge0})\in\mathcal{C}(x)$$

and if we set

$$\gamma=dt\otimes\delta_{u_t^\alpha},$$
$$\beta\equiv(\Omega^\alpha,\mathcal{F}^\alpha,P^\alpha,\{\mathcal{F}_t^\alpha\}_{t\ge0},\gamma,\{X_t^\alpha\}_{t\ge0}),$$

then $\beta \in \mathcal{C}_r(x)$ and $J(x, \alpha) = J_r(x, \beta) \geq V_r(x)$.

Fourth Control Problem: state constraint in G.

We are as well interested in state constraint problems where the states are required to stay in G for $t > 0$. For each $x \in \overline{G}$ we call a controlled system α admissible with respect to G at $x \in \overline{G}$ if

$$X_t^\alpha \in G \qquad \forall t > 0 \quad P^\alpha\text{–a.s.}$$

The set of admissible systems α with respect to G at x will be denoted by $\mathcal{A}_0(x)$.

The value function corresponding to $\mathcal{A}_0(x)$ is defined as

$$V_0(x) = \inf_{\alpha \in \mathcal{A}_0(x)} J(x, \alpha) \tag{6}$$

for any $x \in \overline{G}$.

Theorem 2. *Assume (A1)–(A3) and (A5). Then:* (i)

$$U(x) = V(x) = V_\#(x) = V_r(x) \qquad \text{in } \overline{G}.$$

(ii) *If ∂G is of class C^2 and (A7) is satisfied, then*

$$U(x) = V_0(x) \qquad \text{in} \quad \overline{G}.$$

Recall that U denotes the unique viscosity solution of (4).

Theorem 2, of course, says that, under the assumptions above, the value functions V, $V_\#$, V_r, and V_0 are all equal to the unique viscosity solutions of (4).

Regarding the existence of optimal control we have

Theorem 3. *Assume (A1)–(A3), (A5). Then:* (i) *For each $x \in \overline{G}$ there exits a (relaxed) controlled system $\alpha^* \in \mathcal{A}_r(x)$ such that $V_r(x) = J_r(x, \alpha^*)$.* (ii) *If (A6) is satisfied, then there exists a controlled system $\alpha^* \in \mathcal{A}(x)$ such that $V(x) = J(x, \alpha^*)$.*

Now we are concerned with the regularity of solutions of (4).

Theorem 4. *Assume (A1), (A3), and (A4). There is a constant $k > 0$ such that for each $\gamma \in (0, 1]$, if $\lambda \geq k\gamma$ then the viscosity solution U of (4) is Hölder continuous with exponent γ.*

Theorems 2 and 3 immediately yield Lipschitz or Hölder estimates of the value functions V, V_0, and V_r under appropriate assumptions.

Here we give an example which shows that the Lipschitz continuity of value functions is a rather optimal regularity result.

This example is the deterministic optimal control problem, and we show that the value function is neither semiconcave nor of class C^1 independently how large the discount factor $\lambda > 0$ is.

Example. Consider the case where

$$N = 2, \qquad A = [-1, 1] \subset \mathbf{R},$$
$$G = \{x^2 + y^2 < 2\} \cup \{x > 1,\ x - 2 < y < -x + 2\},$$

$$\begin{aligned} \sigma(x, y, a) \equiv& 0, \\ b(x, y, a) =& (a, \phi(x)), \\ f(x, y, a) =& -(a+1)x, \end{aligned}$$

where $\phi \in C^\infty(\mathbf{R})$ is a function satisfying

$$\phi(y) = 0 \quad \text{if } |y| \le 1 \quad \text{and} \quad \phi(y) < 0 \quad \text{if } |y| > 1.$$

The set G can be described as the union of the open disk with radius $\sqrt{2}$ and center $(0, 0)$ and the open triangle with vertices $(1, 1)$, $(0, 2)$, and $(-1, -1)$.

Since $\sigma = 0$, the control problem is equivalent to the corresponding deterministic one. Observe that, because of our choice of ϕ, if the initial state $(x(0), y(0))$ is in the strip $|y| \le 1$, the any state $(x(t), y(t)$, which is by definition a solution of

$$\frac{d}{dt}(x(t), y(t)) = b(x(t), y(t), u(t)) \qquad \text{for } t > 0,$$

with measurable $u : [0, \infty) \to A$, stays in the strip $|y| \le 1$.

Fix $0 \le y \le 1$. It is easy to see that the best way to minimize the cost functional

$$J(x, y, u) = \int_0^\infty e^{-\lambda t} f(x(t), y(t), u(t))dt,$$

under state-constraint

$$(x(t), y(t)) \in \overline{G},$$

is to choose $u : [0, \infty) \to A$ to be

$$u(t) = \begin{cases} 1 & \text{if } 0 \le t < 2 - y, \\ 0 & \text{if } t \ge 2 - y. \end{cases}$$

Thus the value function at $(0, y)$ is given by

$$V(0, y) = -2 \int_0^{2-y} te^{-\lambda t} dt - \int_{2-y}^\infty e^{-\lambda t} dt.$$

If we set

$$g(x) = -2\int_0^x te^{-\lambda t}dt - \int_x^\infty e^{-\lambda t}dt \quad \forall x \in \mathbf{R},$$

then $V(0, y) = g(2 - y)$ and

$$V_y(0, +0) = -g'(2) = 3e^{-2\lambda}.$$

By symmetry, we get

$$V_y(0, -0) = -3e^{-2\lambda}.$$

These together show that V cannot be either semiconcave or differentiable at $(0, 0)$.

3. Proof of the main results

We begin with the preparations for the proof of Theorem 1.

Let $(\xi_0, \eta_0) \in C(\partial G, \mathbf{R}^N \times \mathbf{R})$ be a continuous function such that

$$(\xi_0(x), \eta_0(x)) \in \operatorname{co}\{(b(x,a), f(x,a)) \mid a \in A(x)\} \qquad \forall x \in \partial G,$$

and such that for some constant $r > 0$,

$$TC(x, \xi_0(x), r) \subset \overline{G} \qquad \forall z \in \partial G,\ x \in B(z, r) \cap \overline{G}.$$

By an argument utilizing partition of unity, we see (see e.g. [IK]) that under the assumptions (A1), (A3), and (A4) there is a function (ξ_0, η_0) satisfying these requirements. Assuming (A1), (A3), and (A4), we fix such a function (ξ_0, η_0) in what follows.

We consider the problem

$$(7) \qquad \begin{cases} \lambda u(x) + H(x, Du(x), D^2u(x)) \geq 0 & \text{in } \overline{G}, \\ \lambda u(x) + H_0(x, Du(x), D^2u(x)) \leq 0 & \text{in } \overline{G}, \end{cases}$$

where

$$H_0(x, p, X) := \begin{cases} H(x, p, X) & \text{if } x \in G, \\ -\xi_0(x) \cdot p - \eta_0(x) & \text{if } x \in \partial G. \end{cases}$$

Since $H_0(x, p, X) \leq H_{in}(x, p, X)$ for all $(x, p, X) \in \overline{G} \times \mathbf{R}^N \times \mathcal{S}^N$, it follows that any viscosity subsolution of (4) is a viscosity subsolution of (7).

Theorem 3.1. *Assume (A1), (A3), and (A4). Let u and v be a viscosity subsolution and a viscosity supersolution of (7), respectively. Then $u \leq v$ on $\overline{G}$.*

For the proof of this theorem, we adapt the arguments from [IK] to our case.

Theorem 3.1 and the remark preceding the theorem immediately yield the following

Theorem 3.2. *Assume (A1), (A3), and (A4). Let u and v be a viscosity subsolution and a viscosity supersolution of (4), respectively. Then $u \leq v$ on $\overline{G}$.*

The following two lemmas are needed for our proof of Theorems 3.1 and 3.

Lemma 3.3. *Assume (A1), (A3), and (A4). Then there exists a function* $\psi \in C^\infty(\overline{G})$ *such that*

$$\xi_0(x) \cdot D\psi(x) \geq 1 \qquad \forall x \in \partial G.$$

This lemma is similar to [IK, Lemma 3.3] in which ξ_0 is assumed to be Lipschitz continuous. We give a proof of Lemma 3.3 in the Appendix.

Lemma 3.4. *Assume (A1), (A3), and (A4). Then there exist* $w \in C^{1,1}(\overline{G} \times \overline{G})$ *and constants* $C > 0$, $r > 0$ *such that*

$$\xi_0(x) \cdot D_x w(x,y) \leq 0 \qquad \forall x \in \partial G,\ y \in \overline{G} \cap B(x,r),$$

and for all $x, y \in \overline{G}$,

$$\begin{gathered} |x-y|^2 \leq w(x,y) \leq C|x-y|^2, \\ \max\{|D_x w(x,y)|, |D_y w(x,y)|\} \leq C|x-y|, \\ |D_x w(x,y) + D_y w(x,y)| \leq C|x-y|^2, \\ D^2 w(x,y) \leq C \begin{pmatrix} I & -I \\ -I & I \end{pmatrix} + C|x-y|^2 \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}, \end{gathered}$$

where $D^2 w$ *in the last inequality should be understood in the distributional sense.*

This lemma is similar to [IK, Lemma 3.4], but we need here the stronger version of the above form. A sketch of proof of this lemma can be found in the Appendix.

The proof of Theorem 3.1 is a combination of the proof of [IK, Theorem 3.1], which is a comparison proof for first-order PDE, and the standard techniques for second-order PDE, which can be found e.g. in [CIL]. However, we give the proof of Theorem 3.1 for the interested reader in the Appendix.

Proof of Theorem 1. The uniqueness of viscosity solutions of (4) is a direct consequence of Theorem 3.2.

Let $M > 0$ be the constant from (A1). Define $g^\pm : \overline{G} \to \mathbf{R}$ by

$$g^\pm(x) = \pm M/\lambda.$$

Clearly, g^+ and g^- are a viscosity supersolution and a viscosity subsolution of

$$\lambda u(x) + H(x, Du(x), D^2u(x)) = 0 \qquad \text{in } G,$$

respectively.

Let (ξ_0, η_0) be as above. Let $x \in \partial G$ and $\varphi \in C^2(\overline{G})$, and assume that $g^+ - \varphi$ attains a minimum at x. Noting that the function:

$$t \mapsto (g^+ - \varphi)(x + t\xi_0(x))$$

on an interval $[0, \varepsilon)$ attains a minimum at $t = 0$ for some $\varepsilon > 0$, we see that

$$D(g^+ - \varphi)(x) \cdot \xi_0(x) \geq 0.$$

Hence, we have

$$\lambda g^+(x) + H(x, D\varphi(x), D^2\varphi(x)) \geq M - \xi_0(x) \cdot D\varphi(x) - \eta_0(x) \geq 0,$$

which proves that g^+ is a viscosity supersolution of (4). Similarly, we see that g^- is a viscosity subsolution of (4).

Now, the standard Perron method (see e.g. [CIL]) yields a viscosity solution U of (4) such that $U \in C(\overline{G})$ and $g^- \leq U \leq g^+$ on $\overline{G}$. QED

For the proof of Theorems 2 and 3 we need the following three propositions.

Theorem 3.5. *Assume (A1)–(A3) and (A5). Then $U(x) \leq V_r(x)$ for all $x \in \mathbf{R}^N$.*

Recall that U is the unique viscosity solution of (4) in the theorem above.

Theorem 3.6. *Assume (A1)–(A3) and (A5). Then $V_\#(x) \leq U(x)$ for all $x \in \overline{G}$.*

Theorem 3.7. *Assume (A1)–(A3) and (A5). Then:* (i) *For each $x \in \overline{G}$ there exists a controlled system $\alpha^* \in \mathcal{A}_r(x)$ such that $V_r(x) = J_r(x, \alpha^*)$.* (ii) *If, in addition, (A6) holds, then for each $x \in \overline{G}$ there exists a controlled system $\alpha^* \in \mathcal{A}(x)$ such that $V(x) = J(x, \alpha^*)$.*

For each $\varepsilon \in (0, 1)$ we set

$$G_\varepsilon = \{x \in G \mid \text{dist}\,(x, G^c) > \varepsilon\}.$$

Assume (A7) for the time being. Let $\hat{a}$ be the function given by (A7). For each $\varepsilon > 0$ choose a function $\chi_\varepsilon \in C^1(\mathbf{R}^N)$ so that

$$0 \leq \chi_\varepsilon(x) \leq 1 \quad \forall x \in \mathbf{R}^N, \qquad \chi_\varepsilon(x) = 1 \quad \forall x \in G_\varepsilon, \qquad \chi_\varepsilon(x) = 0 \quad \forall x \in \mathbf{R}^N \setminus G_{\varepsilon/2},$$

and define functions $\sigma_\varepsilon, b_\varepsilon, f_\varepsilon$ on $\mathbf{R}^N \times A$ by

$$\begin{aligned} \sigma_\varepsilon(x, a) =& \sigma(x, \chi_\varepsilon(x)a + (1 - \chi_\varepsilon(x))\hat{a}(x)), \\ b_\varepsilon(x, a) =& b(x, \chi_\varepsilon(x)a + (1 - \chi_\varepsilon(x))\hat{a}(x)), \\ f_\varepsilon(x, a) =& f(x, \chi_\varepsilon(x)a + (1 - \chi_\varepsilon(x))\hat{a}(x)). \end{aligned}$$

Note that the functions σ_ε and b_ε are Lipschitz continuous and the function f_ε is uniformly continuous on $\mathbf{R}^N \times A$.

For $x \in \mathbf{R}^N$, $\mathcal{C}_\varepsilon(x)$ denotes the set of controlled systems associated with σ_ε and b_ε and for $x \in \overline{G}$, $\mathcal{A}_\varepsilon(x)$ denotes the set of admissible $\alpha \in \mathcal{C}_\varepsilon(x)$ at $x \in \overline{G}$ with respect to $\overline{G}$.

The value function U_ε is defined by

$$(8) \qquad U_\varepsilon(x) = \inf_{\alpha \in \mathcal{A}_\varepsilon(x)} E^\alpha \int_0^\infty e^{-\lambda t} f_\varepsilon(X_t^\alpha, u_t^\alpha) dt \qquad \forall x \in \overline{G}.$$

Theorem 3.8. *Assume (A1)–(A3) and (A7). Then*

$$U_\varepsilon(x) \to U(x) \quad \textit{uniformly for } x \in \overline{G} \quad \textit{as } \varepsilon \to 0.$$

Theorem 3.9. *Assume (A1)–(A3), (A5), (A7), and that $\partial G \in C^2$. Then*

$$U_\varepsilon(x) \geq V_0(x) \qquad \forall x \in \overline{G}.$$

Admitting Theorems 3.5 – 3.9 for the moment, the proof of which will be given in sections 4–8, we complete the proof of Theorems 2 and 3.

Proof of Theorem 2. Assume (A1)–(A3) and (A5). From (5), Theorems 3.5 and 3.6, we see that for all $x \in \overline{G}$,

$$U(x) \leq V_r(x) \leq V(x) \leq V_\#(x) \leq U(x),$$

and hence

$$V(x) = V_r(x) = V_\#(x) = U(x).$$

Next, assume (A7) and that $\partial G \in C^2$. Then Theorem 3.9 asserts that that $U_\varepsilon(x) \geq V_0(x)$ for all $x \in \overline{G}$ and $\varepsilon > 0$, and hence $U_\varepsilon(x) \geq V_0(x) \geq V(x)$ for all $x \in \overline{G}$ and $\varepsilon > 0$. Now, since $U = V$ on $\overline{G}$ and for all $x \in \overline{G}$, $U_\varepsilon(x) \to U(x)$ as $\varepsilon \to 0$ by Theorem 3.8, we conclude that $V_0(x) = U(x) = V(x)$ for all $x \in \overline{G}$ and thus, the claim (ii) is valid. QED

Proof of Theorem 3. Theorems 3.7 is nothing but Theorem 3. QED

Proof of Theorem 4. Fix $\gamma > 0$. We choose a function $(\xi_0, \eta_0) \in C(\overline{G})$ and a constant $r_0 > 0$ so that

$$(9) \qquad (\xi_0(x), \eta_0(x)) \in \operatorname{co}\{(b(x,a), f(x,a)) \mid a \in A(x)\} \quad \forall x \in \partial G,$$

$$(10) \qquad TC(y, \xi_0(x), r_0) \subset \overline{G} \quad \forall x \in \partial G.$$

Let $\psi \in C^2(\overline{G})$ be the function from Lemma 3.3.

Recall that we showed in the proof of Theorem 1 that $|U(x)| \le M/\lambda$ for all $x \in \overline{G}$.

If we replace U by the function

$$\widetilde{U} = U + (2M+1)\psi,$$

then we formally have

$$-\xi_0(x) \cdot D\widetilde{U}(x) \le \eta_0(x) + \lambda \|U\|_\infty - (2M+1)\xi_0(x) \cdot D\psi(x) \le -1 \quad \forall x \in \partial G.$$

Indeed, it is not hard to see that $\widetilde{U}$ satisfies (7) in the viscosity sense with H and η_0 replaced by the functions

$$\widetilde{H}(x, p, X) := H(x, p - 2(M+1)D\psi(x), X - 2(M+1)D^2\psi(x))$$

and $\tilde{\eta}_0(x) := -1$, respectively. To prove that U is Lipschitz continuous on $\overline{G}$, it is enough to show that $\widetilde{U}$ is Lipschitz continuous on $\overline{G}$. So, we may assume by replacing U, H, and η_0 by $\widetilde{U}$, $\widetilde{H}$, and $\tilde{\eta}_0$, respectively, that U satisfies (7) with $\eta_0(x) = -1$ in the viscosity sense.

Let $w \in C^{1,1}(\overline{G} \times \overline{G})$, $r > 0$, and $C > 0$ be from Lemma 3.4. Set

$$v(x, y) := w(x, y)^{1/2}.$$

Note that if $x \ne y$,

$$Dv(x, y) = \frac{1}{2v(x, y)} Dw(x, y),$$

$$D^2 v(x, y) = \frac{1}{2v(x, y)} D^2 w(x, y) - \frac{1}{4v^3(x, y)} Dw(x, y) \otimes Dw(x, y) \le \frac{1}{2v(x, y)} D^2 w(x, y).$$

Hence, for all $x, y \in \overline{G}$ with $x \ne y$, we have

$$\xi_0(x) \cdot Dv(x, y) \le 0 \qquad \text{if } x \in \partial G,$$

and

$$|x - y| \le v(x, y) \le C^{1/2} |x - y|,$$

$$\max\{|D_x v(x, y)|, |D_y v(x, y)|\} \le C/2,$$

$$|D_x v(x, y) + D_y v(x, y)| \le (C/2)|x - y|,$$

$$D^2 v(x, y) \le (C/2) \Big\{ |x - y|^{-1} \begin{pmatrix} I & -I \\ -I & I \end{pmatrix} + |x - y| \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \Big\}.$$

We write K for $\max\{C/2, C^{1/2}\}$.

Fix $L > 0$ and consider the function

$$\Phi(x,y) \equiv U(x) - U(y) - Lv(x,y)^\gamma$$

on the set $\overline{G} \times \overline{G}$. We suppose that

$$\sup\{\Phi(x,y) \mid x,y \in \overline{G},\ |x-y| \le r\} > 0.$$

We select $\hat{x}, \hat{y} \in \overline{G}$ so that $|\hat{x}-\hat{y}| \le r$ and

$$\Phi(\hat{x},\hat{y}) = \sup\{\Phi(x,y) \mid x,y \in \overline{G},\ |x-y| \le r\}.$$

By choosing L large enough we may assume that

$$\sup\{\Phi(x,y) \mid x,y \in \overline{G},\ |x-y| = r\} \le 0,$$

and hence that $|\hat{x}-\hat{y}| < r$. By the continuity of U we see that $\hat{x} \ne \hat{y}$.

Now suppose that $\hat{x} \in \partial G$. This yields that

$$\xi_0(\hat{x}) \cdot D_x v(\hat{x},\hat{y}) \le 0,$$

$$-\xi_0(\hat{x}) \cdot D_x v(\hat{x},\hat{y}) \le -1.$$

These are contradictory. That is, this case never arises.

Since

$$\gamma v(\hat{x},\hat{y})^{\gamma-1}\big(D_x v(\hat{x},\hat{y}), D_y v(\hat{x},\hat{y}),$$
$$K\Big(|\hat{x}-\hat{y}|^{-1}\begin{pmatrix} I & -I \\ -I & I \end{pmatrix} + |\hat{x}-\hat{y}|\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}\Big)\big) \in J^{2,+}v^\gamma(\hat{x},\hat{y}),$$

there are matrices $X, Y \in \mathcal{S}^N$ such that

$$(L\gamma v(\hat{x},\hat{y})^{\gamma-1} D_x v(\hat{x},\hat{y}), X) \in \overline{J}^{2,+}U(\hat{x}),$$

$$(-L\gamma v(\hat{x},\hat{y})^{\gamma-1} D_y v(\hat{x},\hat{y}), -Y) \in \overline{J}^{2,-}U(\hat{y}),$$

$$\begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix} \le LK\gamma v(\hat{x},\hat{y})^{\gamma-1}\Big(|\hat{x}-\hat{y}|^{-1}\begin{pmatrix} I & -I \\ -I & I \end{pmatrix} + |\hat{x}-\hat{y}|\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}\Big).$$

Since U is a solution of (7) and $\hat{x} \in G$, we have

$$\lambda U(\hat{x}) + H(\hat{x}, \gamma Lv(\hat{x},\hat{y})^{\gamma-1} D_x v(\hat{x},\hat{y}), X) \le 0,$$

$$\lambda U(\hat{y}) + H(\hat{y}, -\gamma Lv(\hat{x},\hat{y})^{\gamma-1} D_y v(\hat{x},\hat{y}), -Y) \ge 0.$$

Compute that for any $a \in A$, if we set $\sigma_i(x) = (\sigma_{1i}(x,a), ..., \sigma_{Ni}(x,a))^T$ then

$$\begin{aligned}
&\begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix} \begin{pmatrix} \sigma_i(\hat{x}) \\ \sigma_i(\hat{y}) \end{pmatrix} \cdot \begin{pmatrix} \sigma_i(\hat{x}) \\ \sigma_i(\hat{y}) \end{pmatrix} \\
&\leq 2KL\gamma v(\hat{x},\hat{y})^{\gamma-1}|\hat{x}-\hat{y}|^{-1}|\sigma_i(\hat{x})-\sigma_i(\hat{y})|^2 + 2KL\gamma v(\hat{x},\hat{y})^{\gamma-1}|\hat{x}-\hat{y}||\sigma_i(\hat{y})|^2 \\
&\leq 4KLM^2\gamma v(\hat{x},\hat{y})^{\gamma-1}|\hat{x}-\hat{y}|.
\end{aligned}$$

Summing over all $i \in \{1, ..., N\}$, for any $a \in A$ we get

$$\operatorname{tr}\sigma\sigma^T(\hat{x},a)X + \operatorname{tr}\sigma\sigma^T(\hat{y},a)Y \leq 4NKLM^2\gamma v(\hat{x},\hat{y})^{\gamma-1}|\hat{x}-\hat{y}|.$$

Compute also that

$$\begin{aligned}
&b(\hat{x},a)\cdot D_x v(\hat{x},\hat{y}) + b(\hat{y},a)\cdot D_y v(\hat{x},\hat{y}) \\
&= (b(\hat{x},a) - (b(\hat{y},a)) \cdot D_x v(\hat{x},\hat{y}) + b(\hat{y},a)\cdot(D_x v(\hat{x},\hat{y}) + D_y v(\hat{x},\hat{y})) \\
&\leq M|\hat{x}-\hat{y}||D_x v(\hat{x},\hat{y})| + M|D_x v(\hat{x},\hat{y}) + D_y v(\hat{x},\hat{y})| \\
&\leq 2MK|\hat{x}-\hat{y}|.
\end{aligned}$$

Combining these together we obtain

$$\begin{aligned}
0 \geq &\lambda(U(\hat{x}) - U(\hat{y})) + H(\hat{x}, \gamma v(\hat{x},\hat{y})^{\gamma-1}D_x v(\hat{x},\hat{y}), X) - H(\hat{y}, -\gamma v(\hat{x},\hat{y})^{\gamma-1}D_y v(\hat{x},\hat{y}), -Y) \\
> &\lambda L v(\hat{x},\hat{y})^\gamma + \inf_{a\in A}\{-\frac{1}{2}\operatorname{tr}\sigma\sigma^T(\hat{x},a)X - \frac{1}{2}\operatorname{tr}\sigma\sigma^T(\hat{y},a)Y \\
&- \gamma v(\hat{x},\hat{y})^{\gamma-1}b(\hat{x},a)\cdot D_x v(\hat{x},\hat{y}) - \gamma v(\hat{x},\hat{y})^{\gamma-1}b(\hat{y},a)\cdot D_y v(\hat{x},\hat{y}) - f(\hat{x},a) + f(\hat{y},a)\} \\
\geq &\lambda L|\hat{x}-\hat{y}|^\gamma - 4\gamma NKLM^2|\hat{x}-\hat{y}|^\gamma - 2\gamma MKL|\hat{x}-\hat{y}|^\gamma - Mr^{1-\gamma}|\hat{x}-\hat{y}|^\gamma \\
= &(\lambda L - 4\gamma NKLM^2 - 2\gamma MKL - M)|\hat{x}-\hat{y}|^\gamma.
\end{aligned}$$

If we set $k = 4NKM^2 + 2MK$ and assume that $\lambda > k\gamma$, then by choosing L large enough we have

$$\lambda L - 4\gamma NKLM^2 - 2\gamma MKL - Mr^{1-\gamma} > 0,$$

which contradicts with the previous inequality, i.e., we have

$$U(x) - U(y) \leq Lv(x,y)^\gamma \quad (x, y \in \overline{G}).$$

This inequality yields

$$|U(x) - U(y)| \leq KL\gamma|x-y|^\gamma \quad (x, y \in \overline{G}),$$

proving the Lipschitz continuity of U under the assumption that $\lambda > k\gamma$. QED

4. Proof of Theorem 3.5

Assume (A1)–(A3) and (A5). Define

$$(11)\quad V_r^n(x) = \inf_{\alpha\in\mathcal{C}_r(x)} E^\alpha \int_{[0,\infty)\times A} e^{-\lambda t}[f(X_t^\alpha, a) + nd(X_t^\alpha)]\gamma^\alpha(dtda) \qquad \forall x \in \mathbf{R}^N,$$

for all $n \in \mathbf{N}$, where

$$d(x) = \operatorname{dist}(x, G) \wedge 1.$$

By [EHJ, KU], we know that V_r^n is a viscosity solution of

$$(12)\qquad \lambda u(x) + H(x, Du(x), D^2u(x)) = nd(x) \quad \text{in } \mathbf{R}^N.$$

Then we have:

Proposition 4.1 *Under the assumptions (A1)–(A3) and (A5), V_r^n is a subsolution of* (4).

Proof. Fix $n \in \mathbf{N}$. We need to show that if $z \in \partial G$ and $a \in A(z)$ and if $\varphi \in C^2(\overline{G})$ and $V_r^n - \varphi$ has a maximum at z, then

$$\lambda V_r^n(z) - \frac{1}{2}\operatorname{tr}\sigma\sigma^T(z,a)D^2\varphi(z) - b(z,a)\cdot D\varphi(z) - f(z,a) \le 0.$$

Fix $z \in \partial G$, $a \in A(z)$. Then there exist $r > 0$ and $\hat{a} \in C(\mathbf{R}^N, A)$ such that (i), (ii), and (iii) of (A5) hold with these z, $\hat{a}$, and r. Set $U = G \cap \operatorname{Int} B(z,r)$. Choose a C^1 function ζ on $\mathbf{R}^N$ so that

$$\zeta \ge 0 \quad \text{in } \mathbf{R}^N, \qquad \zeta = 0 \quad \text{in } \mathbf{R}^N \setminus B(z, r/2), \qquad \zeta \equiv 1 \quad \text{in } B(z, r/3).$$

Note that if we set

$$\tilde{\sigma}(x) = \zeta(x)\sigma(x,a), \quad \tilde{b}(x) = \zeta^2(x)b(x,a),$$

$$\tilde{f}(x) = \zeta^2(x)f(x,a) + (1-\zeta^2(x))\lambda V_r^n(x) + n\zeta^2(x)d(x),$$

then V_r^n satisfies

$$\lambda V_r^n(x) - \frac{1}{2}\operatorname{tr}\tilde{\sigma}\tilde{\sigma}^T D^2 V_r^n(x) - \tilde{b}\cdot DV_r^n(x) - \tilde{f} \le 0 \quad \text{in } \mathbf{R}^N$$

in the viscosity sense.

We now invoke [ILT, Theorem 2.1] (more precisely, its proof). It is easy to check that (2.2) and (MP) of [ILT] with $K = \overline{U}$ and with

$$F(x,r,p,X) = \lambda r - \frac{1}{2}\operatorname{tr}\tilde{\sigma}\tilde{\sigma}^T(x)X - \tilde{b}(x)\cdot p - \tilde{f}(x)$$

are satisfied. Therefore we see that V_r^n is a viscosity subsolution of

$$\lambda V_r^n(x) - \frac{1}{2}\operatorname{tr}\tilde\sigma\tilde\sigma^T(x)D^2V_r^n(x) - \tilde b(x)\cdot DV_r^n(x) - \tilde f(x) \le 0 \quad \text{in } \overline{U}.$$

This shows that if $\varphi \in C^2(\overline{G})$ and $V_r^n - \varphi$ has a maximum over $\overline{G}$ at z then

$$\begin{aligned} 0 \ge& \lambda V_r^n(z) - \frac{1}{2}\operatorname{tr}\tilde\sigma\tilde\sigma^T(z)D^2\varphi(z) - \tilde b(z)\cdot D\varphi(z) - \tilde f(z) \\ \ge& \lambda V_r^n(z) - \frac{1}{2}\operatorname{tr}\sigma\sigma^T(z)D^2\varphi(z) - b(z)\cdot D\varphi - f(z,a), \end{aligned}$$

which completes the proof. QED

By comparison we have

$$V_r^n(x) \le U(x) \quad \forall x \in \overline{G},\ \forall n \in \mathbf{N},$$

where as before $U \in C(\overline{G})$ denotes the solution of (4).

From the definition of V_r^n, it is immediate to see that

$$V_r^n(x) \le V_r^{n+1}(x) \quad \forall x \in \overline{G},\ \forall n \in \mathbf{N}.$$

Set

$$V^+(x) = \sup_n V_r^n(x) \quad \forall x \in \overline{G}.$$

Theorem 4.2. *Under the assumptions (A1)–(A3) and (A5), we have*

$$V^+ = U \qquad \text{on } \overline{G}. \tag{13}$$

Proof. Since

$$V_r^n(x) \le U(x) \quad \forall x \in \overline{G},\ \forall n \in \mathbf{N},$$

we have

$$V^+(x) \le U(x) \quad \forall x \in \overline{G},\ \forall n \in \mathbf{N}.$$

Next, we show that V^+ is a supersolution of (4). Indeed, if we set

$$W(x) = \lim_{r\searrow 0} \inf\{V_r^n(y) \mid |y-x| < r,\ n > r^{-1}\},$$

then, since V_r^n are supersolutions of

$$\lambda u(x) + H(x, Du(x), D^2u(x)) \ge 0 \quad \text{in } \mathbf{R}^N, \tag{14}$$

we see that W is a supersolution of (14) in the sense that if $\varphi \in C^2(\mathbf{R}^N)$, $z \in \mathbf{R}^N$, $W(z) < \infty$, and $W - \varphi$ attains its minimum at z, then we have

$$\lambda W(z) + H(z, D\varphi(z), D^2\varphi(z)) \geq 0.$$

It is not hard to see that $V^+(x) = W(x)$ for $x \in \overline{G}$. Thus, in order to conclude that V^+ is a supersolution of (4), it is enough to show that

$$W(x) = \infty \quad \forall x \in \mathbf{R}^N \setminus \overline{G}. \tag{15}$$

Fix $z \in \mathbf{R}^N \setminus \overline{G}$ and choose $R > 0$ so that $B(z, 2R) \subset \mathbf{R}^N \setminus \overline{G}$. We select a function $\zeta : \mathbf{R}^N \to \mathbf{R}$ such that $\zeta \in C^2(\mathbf{R}^N)$, $\zeta(z) > 0$ and $\zeta \leq -1$ on $\partial B(z, R)$. (This can be achieved by taking for instance $\zeta(x) = 1 - \frac{2}{R^2}|x - z|^2$.) We fix an upper bound $K > 0$ of

$$\sup_{a \in A}\{-\frac{1}{2}\operatorname{tr}\sigma\sigma^T(x,a)D^2\zeta(x) - b(x,a)\cdot D\zeta(x) - f(x,a)\} + \lambda\zeta(x)$$

over $B(z, R)$.

Now, for $n \in \mathbf{N}$ we set

$$\Phi_n(x) = \frac{n(R \wedge 1)}{2K}\zeta(x).$$

Then we get

$$\lambda\Phi_n(x) + H(x, D\Phi_n(x), D^2\Phi_n(x)) - nd(x) \leq \frac{n(R \wedge 1)}{2} - n(R \wedge 1) < 0.$$

Since $\zeta \leq -1$ on $\partial B(z, R)$ and by the definition of V_r^n, $V_r^n(x) \geq -M/\lambda$ for all $x \in \mathbf{R}^N$, we have

$$\Phi_n(x) \leq V_r^n(x) \quad \forall x \in \partial B(z, R)$$

if n is large enough. Hence by comparison we get the inequality

$$W(x) \geq \Phi_n(x) \qquad \text{in } B(z, R)$$

for n large enough, and we conclude (15) by letting $n \to \infty$.

Now, by using Theorem 3.2 we see that $V^+(x) \geq U(x)$ for all $x \in \overline{G}$, and conclude the proof. QED

Completion of the proof of Theorem 3.5. Fix $x \in \overline{G}$. Since

$$J_r(x, \alpha) = E^\alpha \int_{[0,\infty)\times A} e^{-\lambda t}[f(X_t^\alpha, a) + nd(X_t^\alpha)]\gamma^\alpha(dtda)$$

for all $\alpha \in \mathcal{A}_r(x)$ and $x \in \mathbf{R}^N$, it is immediate to see that for all $x \in \overline{G}$, $V_r^n(x) \leq V_r(x)$ and hence $V^+(x) \leq V_r(x)$. As a consequence of (13), we get $U(x) \leq V_r(x)$ for all $x \in \overline{G}$. QED

5. Proof of Theorem 3.6

Throughout this section we assume (A1)–(A3) and (A5). We adapt some arguments from the proof of [EHJ, Theorem 4.9]. The main part of proof is divided into several lemmas.

The first lemma concerns a control problem with $I(p)$ as its control set.

Let $p \in \mathbf{N}$ and let

$$\hat{b} : \mathbf{R}^N \times I(p) \to \mathbf{R}^N, \quad \hat{\sigma} : \mathbf{R}^N \times I(p) \to \mathbf{R}^{N\times l}, \quad \hat{f} : \mathbf{R}^N \times I(p) \to \mathbf{R}.$$

Assume that $\hat{b}$ and $\hat{\sigma}$ are Lipschitz continuous and $\hat{f}$ is uniformly continuous on $\mathbf{R}^N \times I(p)$.

We denote by $\hat{\mathcal{C}}(\xi)$ the set of all controlled systems at ξ associated with $I(p),\ \hat{b},\ \hat{\sigma}$.

Set

$$C(p) = \{\gamma \equiv (\gamma^1, \gamma^2, \ldots, \gamma^p) \in \mathbf{R}^p \mid \gamma^i \geq 0,\ \sum_{i\in I(p)} \gamma^i = 1\}$$

For $\gamma \equiv (\gamma^1, \gamma^2, \ldots, \gamma^p) \in C(p)$ we set

$$\begin{aligned}
\overline{b}(x,\gamma) &= \sum_{i\in I(p)} \gamma^i \hat{b}(x,i),\\
\overline{\sigma}(x,\gamma) &= ((\gamma^1)^{\frac{1}{2}}\hat{\sigma}(x,1), \ldots, (\gamma^p)^{\frac{1}{2}}\hat{\sigma}(x,p)) \in \mathbf{R}^{N\times lp},\\
\overline{f}(x,\gamma) &= \sum_{i\in I(p)} \gamma^i \hat{f}(x,i).
\end{aligned}$$

Let $v \in \mathrm{BUC}(\mathbf{R}^N)$. Let $\varepsilon > 0, \xi \in \mathbf{R}^N$, and

$$(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\geq 0}, \{W_t\}, \{\gamma_t\}_{t\geq 0}, \{X_t\}_{t\geq 0}) \in \overline{\mathcal{C}}(\xi),$$

where $\overline{\mathcal{C}}(\xi)$ denotes the set of all controlled systems at ξ associated with $C(p),\ \overline{b},\ \overline{\sigma}$.

Let $h > 0,\ q \in \mathbf{R}^N$ and suppose that

$$\gamma_t = \gamma_{[qh^{-1}t]q^{-1}h} \qquad \forall t \in [0,\infty),$$

where we used the notation: $[x]$ denotes the largest integer $\leq x$ for $x \in \mathbf{R}$. That is, we assume that $\{\gamma_t\}$ is a step process with step length $q^{-1}h$.

Lemma 5.1. *There exist an extension* $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P}, \{\hat{\mathcal{F}}_t\}_{t\geq 0})$ *of a filtered probability space* $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\geq 0})$ *and a controlled system*

$$(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P}, \{\hat{\mathcal{F}}_t\}_{t\geq 0}, \{\hat{W}_t\}, \{\hat{u}_t\}_{t\geq 0}, \{\hat{X}_t\}_{t\geq 0}) \in \hat{\mathcal{C}}(x)$$

such that

$$E\left[\int_0^h \overline{f}(X_t,\gamma_t)e^{-\lambda t}dt + e^{-\lambda h}v(X_h)\right] + \varepsilon > \hat{E}\left[\int_0^h \hat{f}(\hat{X}_t,\hat{u}_t)e^{-\lambda t}dt + e^{-\lambda h}v(\hat{X}_h)\right],$$

where $\hat{E}$ *denote the mathematical expectation with respect to* $\hat{P}$*. In addition, if* $\tau > 0$*,* $\omega \in \Omega$*, and* $\gamma_t^i(\omega) \equiv 0$ *on* $[0,\tau)$*, then*

$$\hat{u}_t(\omega) \neq i \qquad \forall t \in [0,\tau).$$

Proof. Let $n \in \mathbf{N}$ and set for $(i,\omega) \in I(p) \times \Omega$,

$$I_i^n(\omega) = \bigcup_{j=0}^{q-1}\bigcup_{m=0}^{n-1}[jhq^{-1} + mhq^{-1}n^{-1}\sum_{k<i}\gamma^k_{jhq^{-1}}(\omega), jhq^{-1} + mh(qn)^{-1} + h(qn)^{-1}\sum_{k\le i}\gamma^k_{jhq^{-1}}(\omega)).$$

Also set

$$I_i^n = \{(t,\omega) \mid t \in I_i^n(\omega)\}.$$

Note that

$$\sum_{i\in I(p)} \mathbf{1}_{I_i^n}(t,\omega) = 1$$

and

(16) $$(\mathbf{1}_{I_1^n}(t,\omega),\ldots,\mathbf{1}_{I_p^n}(t,\omega)) \to \gamma_t(\omega) \qquad \text{weakly-star in } (L^\infty[0,h])^p$$

for all $\omega \in \Omega$.

By El Karoui et al. [ENJ] (see also Proposition A.1 in the apendix), there are $\mathcal{F}_{t+h(nq)^{-1}}$–adapted processes

$$\{M_t^{n,1}\}_{t\ge 0},\ldots,\{M_t^{n,p}\}_{t\ge 0}$$

which are orthogonal martingales with respect to the natural filtration,

$$\mathcal{F}_t^n := \sigma(M_s^{n,i} \mid 0 \le s \le t, i \in I(p)) \vee \mathcal{F}_{[qh^{-1}t]q^{-1}h},$$

having the quadratic variational processes

$$\langle M^{n,i}\rangle_t = \int_0^t \mathbf{1}_{I_i^n}(s,\omega)ds.$$

Moreover, we have

$$M_t^n := (M_t^{n,1}, \dots, M_t^{n,p}) \to M_t = (M_t^1, \dots, M_t^p)$$

uniformly on $[0, h]$, as $n \to \infty$, P–a.s., where

$$M_t^i = \int_0^t (\gamma_s^i)^{\frac{1}{2}} dW_s^i, \quad W_t = (W_t^1, \dots, W_t^p).$$

Note that by (16), as $n \to \infty$,

$$\gamma_t^n := \sum_{i \in I(p)} \mathbf{1}_{I_i^n}(\cdot, \omega) dt \otimes \delta_{e_i} \to \gamma_t := \sum_{i \in I(p)} \gamma^i(\cdot, \omega) dt \otimes \delta_{e_i}$$

in $\Gamma_{C(p)}$ for all $\omega \in \Omega$.

Define the probability measures Q^n, with $n \in N$, and Q on $W^N \times \Gamma_{C(p)} \times W^{lp}$, with its natural filtration, by

$$Q^n(B) = P((X^n, \gamma^n, M^n) \in B),$$

$$Q(B) = P((X, \gamma, M) \in B),$$

where X_t^n is the strong solution of

$$X_t^n = \xi + \sum_{i \in I(p)} \Big(\int_0^t \hat{b}(X_s^n, i) \mathbf{1}_{I_i^n}(s, \omega) ds + \int_0^t \hat{\sigma}(X_s^n, i) dM_s^{n,i} \Big). \tag{17}$$

Note that X_t is a strong solution of

$$X_t = \xi + \sum_{i \in I(p)} \Big(\int_0^t \hat{b}(X_s, i) \gamma^i(s, \omega) ds + \int_0^t \hat{\sigma}(X_s, i) dM_s^i \Big). \tag{18}$$

Since the sequence $\{Q^n\}_{t \geq 0}$ is tight, we may assume that

$$Q^n \to Q^\infty$$

in the weak convergence of measures for some probability measure Q^∞.

For $\eta \in C_b(\Gamma_{C(p)} \times W^{lp})$ we have

$$\int \eta(\gamma^n, M^n) P(d\omega) \to \int \eta(\gamma, M) P(d\omega),$$

Since

$$\int \eta(y,z)Q^n(dxdydz) = \int \eta(\gamma^n, M^n)P(d\omega),$$
$$\int \eta(y,z)Q(dxdydz) = \int \eta(\gamma, M)P(d\omega),$$

and, as $n \to \infty$,

$$\int \eta(y,z)Q^n(dxdydz) \to \int \eta(y,z)Q^\infty(dxdydz),$$

we have

$$\int \eta(y,z)Q^\infty(dxdydz) = \lim_{n\to\infty} \int \eta(\gamma^n, M^n)P(d\omega)$$
$$= \int \eta(\gamma, M)P(d\omega) = \int \eta(y,z)Q(dxdydz).$$

For any open set $U \subset \Gamma_{C(p)} \times W^{lp}$, define

$$\phi_k(y,z) := 1 \wedge k \operatorname{dist}((y,z), U^c) \quad \text{ for } k \in \mathbf{N}.$$

Then

$$\phi_k(y,z) \nearrow \mathbf{1}_U(y,z) \quad \text{ as } k \to \infty,$$

and hence

$$\int \mathbf{1}_U(y,z)Q^\infty(dxdydz) = \lim_{k\to\infty} \int \phi_k(y,z)Q^\infty(dxdydz)$$
$$= \lim_{k\to\infty} \int \phi_k(y,z)Q(dxdydz) = \int \mathbf{1}_U(y,z)Q(dxdydz).$$

That is,

$$Q^\infty(W^N \times U) = Q(W^N \times U).$$

Therefore, for any Borel $B \subset \Gamma_{C(p)} \times W^{lp}$,

$$Q^\infty(W^N \times B) = Q(W^N \times B). \tag{19}$$

Since $\{M_t^n\}_{t\geq 0}$ is an $\mathcal{F}_{t+h(nq)^{-1}}$–martingale, (M_t^n) is an $\mathcal{F}_t^n$–martingale, where

$$\mathcal{F}_t^n := \sigma(X_s^n, M_s^n \mid 0 \leq s \leq t,\ i \in I(p)) \wedge (\gamma^n)^{-1}(\mathcal{B}_t(\Gamma_{C(p)})).$$

Fix any countable family $\{f_j\} \subset C([0,\infty) \times C(p))$. Fix any $0 \leq s \leq t < \infty,\ k, m \in N$, and any $g \in C_b(R^{(N+lp+m)k})$. We have

$$Eg(Y)(M_t^n - M_s^n) = 0,$$

where

$$Y := (X^n_{t_1}, \ldots, X^n_{t_k}, M^n_{t_1}, \ldots, M^n_{t_k}, \gamma^n_{t_1}(f_1), \ldots, \gamma^n_{t_1}(f_m), \ldots, \gamma^n_{t_k}(f_1), \ldots, \gamma^n_{t_k}(f_m))$$

and $0 \le t_1 < t_2 < \cdots < t_n \le s$. Hence

$$E^n g(\eta)(z(t) - z(s)) = 0 \tag{20}$$

where

$$\eta := (x(t_1, \dot{;} x(t_k), z(t_1), \ldots, z(t_k), y_{t_1}(f_1), \ldots, y_{t_1}(f_m), \ldots, y_{t_k}(f_1), \ldots, y_{t_k}(f_m))$$

and E^n denotes the mathematical expectation with respect to Q^n.

This is, on the probability space

$$(W^N \times \Gamma_{C(p)} \times W^{lp}, \mathcal{B}(W^N \times \Gamma_{C(p)} \times W^{lp}), Q^n),$$

with its natural filtration $\{\mathcal{F}^n_t\}_{t\ge 0}$, the process $\{z(t)\}_{t\ge 0}$ is an $\mathcal{F}^n_t$-martingale. From (17), we see that the process $\{x(t)\}_{t\ge 0}$ is the strong solution of

$$x(t) = \xi + y_t(\bar{b} \circ x) + \sum_{i\in I(p)} \int_0^t \hat{\sigma}(x(s), i) dz^i(s) \tag{21}$$

on the filtered probability space

$$(W^N \times \Gamma_{C(p)} \times W^{lp}, \mathcal{B}(W^N \times \Gamma_{C(p)} \times W^{lp}), Q^n, \{\mathcal{F}^n_t\}_{t\ge 0}),$$

where

$$\bar{b} \circ x(t, \gamma) = \bar{b}(x(t), \gamma)$$

and

$$y_t(\bar{b} \circ x) = \int_{[0,t]\times C(p)} \bar{b} \circ x(s, \gamma) y(ds d\gamma).$$

In (20) we send $n \to \infty$, to conclude that $\{z(t)\}_{t\ge 0}$ is an $\mathcal{F}^\infty_t$–martingale on the filtered probability space

$$(W^N \times \Gamma_{C(p)} \times W^{lp}, \mathcal{B}(W^N \times \Gamma_{C(p)} \times W^{lp}), Q^\infty, \{\mathcal{F}^\infty_t\}_{t\ge 0}),$$

where $\{\mathcal{F}^\infty_t)\}_{t\ge 0}$ is the natural filtration on the probability space

$$(W^N \times \Gamma_{C(p)} \times W^{lp}, \mathcal{B}(W^N \times \Gamma_{C(p)} \times W^{lp}), Q^\infty).$$

Now, from (21), we see that the process $\{x(t)\}_{t\geq 0}$ is a strong solution of

$$x(t) = \xi + y_t(\overline{b} \circ x) + \sum_{i\in I(p)} \int_0^t \hat{\sigma}(x(s),i)dz^i(s) \tag{22}$$

$$\Big(= \xi + \int_{[0,t]\times C(p)} \overline{b}(x(s),\gamma)y(dsd\gamma) + \sum_{i\in I(p)} \int_0^t \hat{\sigma}(x(s),i)dz^i(s)\Big)$$

on the filtered probability space

$$(W^N \times \Gamma_{C(p)} \times W^{lp}, \mathcal{B}(W^N \times \Gamma_{C(p)} \times W^{lp}), Q^\infty, \{\mathcal{F}_t^\infty\}_{t\geq 0}).$$

Similarly we see that $x(t)$ is also a strong solution of

$$x(t) = \xi + \int_{[0,t]\times C(p)} \overline{b}(x(s),\gamma)y(dsd\gamma) + \sum_{i\in I(p)} \int_0^t \hat{\sigma}(x(s),i)dz^i(s), \tag{23}$$

on the filtered probability space

$$(W^N \times \Gamma_{C(p)} \times W^{lp}, \mathcal{B}(W^N \times \Gamma_{C(p)} \times W^{lp}), Q, \{\mathcal{F}_t\}_{t\geq 0}),$$

where $\{\mathcal{F}_t\}_{t\geq 0}$ is the natural filtration on

$$(W^N \times \Gamma_{C(p)} \times W^{lp}, \mathcal{B}(W^N \times \Gamma_{C(p)} \times W^{lp}), Q).$$

Recalling the construction of a solution of (22) or (23), in the standard iteration process and using (19), we see that the distribution of (x, y, z) on the probability spaces

$$(W^N \times \Gamma_{C(p)} \times W^{lp}, \mathcal{B}(W^N \times \Gamma_{C(p)} \times W^{lp}), Q^\infty)$$

and

$$(W^N \times \Gamma_{C(p)} \times W^{lp}, \mathcal{B}(W^N \times \Gamma_{C(p)} \times W^{lp}), Q)$$

are the same. That is,

$$Q^\infty = Q.$$

Hence, as $n \to \infty$,

$$Q^n \to Q$$

in the weak convergence of measures. In particular we have

$$E\Big[\sum_{i\in I(p)} \int_0^h e^{-\lambda t}\hat{f}(X_t^n, i)\mathbf{1}_{I_i^n}(t,\omega)dt + e^{-\lambda h}v(X_h^n)\Big] \to E\Big[\int_0^h \overline{f}(X_t,\gamma_t)dt + e^{-\lambda h}v(X_h)\Big].$$

Define the map :

$$(t, \omega) \mapsto \hat{u}_t^n(\omega)$$

on $[0, \infty) \times \Omega$ by

$$\hat{u}_t^n(\omega) = i \quad \text{if} \quad \text{and} \quad \text{only} \quad \text{if} \quad (t, \omega) \in I_i^n.$$

This is well-defined since

$$\sum_{i \in I(p)} \mathbf{1}_{I_i^n} = 1$$

on $[0, \infty) \times \Omega$.

With this notation, we have

$$\sum_{i \in I(p)} \int_0^h e^{-\lambda t} \hat{f}(X_t^n, i) \mathbf{1}_{I_i^n} dt = \int_0^h e^{-\lambda t} \hat{f}(X_t^n, \hat{u}_t^n) dt$$

and

$$X_t^n = \xi + \int_0^t \hat{b}(X_s^n, \hat{u}_s^n) ds + \sum_{i \in I(p)} \int_0^t \hat{\sigma}(X_s^n, i) dM_s^{ni}.$$

Writing

$$M_t^{ni} = (N_t^{i1}, \ldots, N_t^{il}),$$

we have

$$\langle N^{ij}, N^{nm} \rangle_t = 0$$

if $(i, j) \neq (n, m)$, and

$$\langle N^{ij} \rangle_t = \int_0^t \mathbf{1}_{I_i^n} ds.$$

By the representation formula for martingale (see, for instance, [IW]), we find an extension

$$(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P}, \{\hat{\mathcal{F}}_t\}_{t \geq 0})$$

of

$$(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t \geq 0})$$

and a standard l–dimensional Brownian motion $\{\hat{W}_t\}_{t \geq 0}$ such that

$$(N_t^{i1}, \ldots, N_t^{il}) = \int_0^t \mathbf{1}_{I_i^n} d\hat{W}_s \qquad \forall i \in I(p).$$

Now, choosing n large enough, we see that

$$E\left[\int_0^h \overline{f}(X_t, \gamma_t) e^{-\lambda t} dt + e^{-\lambda h} v(X_h)\right] + \varepsilon > \hat{E}\left[\int_0^h \hat{f}(X_t^n, \hat{u}_t^n) dt + e^{-\lambda h} v(X_h^n)\right]$$

and

$$X_t^n = \xi + \int_0^t \hat{b}(X_s^n, \hat{u}_s^n) ds + \sum_{i \in I(p)} \int_0^t \hat{\sigma}(X_s^n, i) \mathbf{1}_{I_i^n} d\hat{W}_s$$
$$= \xi + \int_0^t \hat{b}(X_s^n, \hat{u}_s^n) ds + \int_0^t \hat{\sigma}(X_s^n, \hat{u}_s^n) d\hat{W}_s.$$

Now, fix $\tau > 0$, $\omega \in \hat{\Omega}$ and $i \in C(p)$, and assume that

$$\gamma_t^i(\omega) = 0 \qquad \forall t \in [0, \tau).$$

This implies that

$$\mathbf{1}_{I_i^n}(t, \omega) = 0 \qquad \forall t \in [0, \tau),$$

and hence

$$\hat{u}_t^n(\omega) \neq i \qquad \forall t \in [0, \tau),$$

which completes the proof. QED

By (A5) and the compactness of ∂G, there are a constant $r > 0$, a sequence $\{z_j\}_{j \in I(q)}$, with $q \in \mathbf{N}$, of points in ∂G, sequences $\{\hat{\zeta}_j\}_{j \in I(q)}$, $\{\eta_j\}_{j \in I(q)}$ of C^∞ functions on $\mathbf{R}^N$, and a sequence $\{\hat{a}_j\}_{j \in I(q)}$ of a continuous maps : $\mathbf{R}^N \to A$ such that

(24) $0 \le \hat{\zeta}_j \le 1, \quad 0 \le \eta_j \le 1 \quad$ on $\mathbf{R}^N \quad \forall j \in I(q)$;

(25) $\eta_j = 1 \quad$ on $(\operatorname{spt} \hat{\zeta}_j)_{2r} \quad \forall j \in I(q)$;

(26) $\sum_{j \in I(q)} \hat{\zeta}_j(x) = 1 \qquad \forall x \in (\partial G)_{4r}$;

(27) the functions $x \mapsto b(x, \hat{a}_j(x))$ and $x \mapsto \sigma(x, \hat{a}_j(x))$ are Lipschitz continuous on a neighborhood of $\operatorname{spt} \eta_j$ for any $j \in I(q)$;

(28) $\sigma(x, \hat{a}_j(x)) = 0 \qquad \forall x \in (\operatorname{spt} \hat{\zeta}_j)_{2r} \cap \partial G, \ \forall j \in I(q)$;

(29) $TC(x, b(x, \hat{a}_j(x)), r) \subset \overline{G} \qquad \forall x \in (\operatorname{spt} \zeta_j)_{2r} \cap \overline{G}, \ \forall j \in I(q)$.

Fix $\rho \in (0, r]$ and choose a C^∞ function ζ_0 on $\mathbf{R}^N$ so that

$$0 \le \zeta_0 \le 1 \quad \text{on } \mathbf{R}^N,$$
$$\zeta_0(x) = \begin{cases} 0 & \text{if} \quad x \in (\partial G)_{3\rho}, \\ 1 & \text{if} \quad x \in \mathbf{R}^N \setminus (\partial G)_{4\rho}. \end{cases}$$

Then we have

$$\zeta_0(x) + (1 - \zeta_0(x)) \sum_{j \in I(q)} \hat{\zeta}_j(x) = 1 \qquad \forall x \in \mathbf{R}^N.$$

We define C^∞ functions ζ_j on $\mathbf{R}^N$, with $j \in I(q)$, by

$$\zeta_j(x) = (1 - \zeta_0(x)) \hat{\zeta}_j(x).$$

Of course, we have

$$\sum_{j\in I_0(q)} \zeta_j(x) = 1 \qquad \text{on } \mathbf{R}^N; \qquad \operatorname{spt}\zeta_j \subset \operatorname{spt}\hat{\zeta}_j \qquad \forall j \in I(q).$$

Choose a dense subset $\{a_i\}_{i\in\mathbf{N}}$ of A, fix any $p \in \mathbf{N}$ and define functions $\tilde{b}$, $\tilde{\sigma}$, $\tilde{f}$ on $\mathbf{R}^N \times I(p)$ by

$$\begin{aligned}
\tilde{b}(x,i) &= \zeta_0(x)b(x,a_i) + \sum_{j\in I(q)} \zeta_j(x)b(x,\hat{a}_j(x)),\\
\tilde{\sigma}(x,i) &= \big(\zeta_0(x)^{\frac{1}{2}}\sigma(x,a_i), \zeta_1(x)^{\frac{1}{2}}\sigma(x,\hat{a}_1(x)), ..., \zeta_q(x)^{\frac{1}{2}}\sigma(x,\hat{a}_q(x))\big),\\
\tilde{f}(x,i) &= \zeta_0(x)f(x,a_i) + \sum_{j\in I(q)} \zeta_j(x)f(x,\hat{a}_j(x)).
\end{aligned}$$

Observe here that since $\zeta_j \in C^2(\mathbf{R}^N)$, we have

$$x \mapsto \zeta_j(x)^{\frac{1}{2}}\sigma(x,\hat{a}_j(x)) \in W^{1,\infty}(\mathbf{R}^N, \mathbf{R}^{N\times l}).$$

In order to see this, we only need to recall (see, for instance, [SV, the proof of Lemma 3.2.3]) that if $\psi \in C^2(\mathbf{R}^N)$ satisfies $\psi \geq 0$ in $\mathbf{R}^N$ and $|D^2\psi(x)| \leq M$ for all $x \in \mathbf{R}^N$ and for some constant $M > 0$ then the following inequality holds:

$$|\psi^{1/2}(x) - \psi^{1/2}(y)| \leq \Big(\frac{M}{2}\Big)^{1/2}|x-y| \qquad \forall x,y \in \mathbf{R}^N. \tag{30}$$

Select a C^∞ function η_0 on $\mathbf{R}^N$ so that

$$\begin{aligned}
&0 \leq \eta_0 \leq 1 \quad \text{on } \mathbf{R}^N,\\
&\eta_0(x) = \begin{cases} 0 & \text{if} \quad x \in \partial G,\\ 1 & \text{if} \quad x \in \mathbf{R}^N \setminus (\partial G)_{2\rho}.\end{cases}
\end{aligned}$$

Note that for all $(x,i) \in \mathbf{R}^N \times I(p)$,

$$\begin{aligned}
\tilde{b}(x,i) &= \zeta_0(x)\eta_0(x)b(x,a_i) + \sum_{j\in I(q)} \zeta_j(x)\eta_j(x)b(x,\hat{a}_j(x)),\\
\tilde{\sigma}(x,i) &= \big(\zeta_0(x)^{\frac{1}{2}}\eta_0(x)\sigma(x,a_i), \zeta_1(x)^{\frac{1}{2}}\eta_1(x)\sigma(x,\hat{a}_1(x)), ..., \zeta_q(x)^{\frac{1}{2}}\eta_q(x)\sigma(x,\hat{a}_q(x))\big),\\
\tilde{f}(x,i) &= \zeta_0(x)\eta_0(x)f(x,a_i) + \sum_{j\in I(q)} \zeta_j(x)\eta_j(x)f(x,\hat{a}_j(x)).
\end{aligned}$$

Set

$$C(p) = \{\gamma \equiv (\gamma^1, ..., \gamma^p) \in \mathbf{R}^p \mid \gamma^i \geq 0, \ \sum_{i\in I(p)} \gamma^i = 1\}$$

and define functions $\overline{b},\ \overline{\sigma},\ \overline{f}$ on $\mathbf{R}^N \times C(p)$ by

(31)
$$\begin{cases} \overline{b}(x,\gamma) = \sum_{i\in I(p)} \gamma^i \tilde{b}(x,i), \\ \overline{\sigma}(x,\gamma) = ((\gamma^1)^{\frac{1}{2}}\tilde{\sigma}(x,1), (\gamma^2)^{\frac{1}{2}}\tilde{\sigma}(x,2), ..., (\gamma^p)^{\frac{1}{2}}\tilde{\sigma}(x,p)), \\ \overline{f}(x,\gamma) = \sum_{i\in I(p)} \gamma^i \tilde{f}(x,i). \end{cases}$$

Set

$$K = I(p) \times I_0(q) \equiv \{(i,j) \mid i = 1, ..., p,\ \ j = 0, 1, ..., q\}.$$

and define $\hat{b},\ \hat{\sigma},\ \hat{f}$ on $\mathbf{R}^N \times K$ by

$$\begin{aligned} \hat{b}(x,i,0) &= \eta_0(x) b(x,a_i), \\ \hat{b}(x,i,j) &= \eta_j(x) b(x,\hat{a}_j(x)) \qquad \text{if } j \geq 1, \\ \hat{\sigma}(x,i,0) &= \eta_0(x)\sigma(x,a_i), \\ \hat{\sigma}(x,i,j) &= \eta_j(x)\sigma(x,\hat{a}_j(x)) \qquad \text{if } j \geq 1, \\ \hat{f}(x,i,j) &= \eta_0(x) f(x,a_i), \\ \hat{f}(x,i,j) &= \eta_j(x) f(x,\hat{a}_j(x)) \qquad \text{if } j \geq 1. \end{aligned}$$

Note that for all $(x,\gamma) \in \mathbf{R}^N \times C(p)$,

$$\begin{aligned} \overline{b}(x,\gamma) &= \sum_{i\in I(p)} \gamma^i \sum_{j\in I_0(q)} \zeta_j(x)\hat{b}(x,i,j), \\ \overline{\sigma}(x,\gamma) &= ((\gamma^1)^{\frac{1}{2}}\zeta_0(x)^{\frac{1}{2}}\hat{\sigma}(x,1,0), (\gamma^1)^{\frac{1}{2}}\zeta_1(x)^{\frac{1}{2}}\hat{\sigma}(x,1,1), ..., (\gamma^1)^{\frac{1}{2}}\zeta_q(x)^{\frac{1}{2}}\hat{\sigma}(x,1,q), \\ &\qquad ..., (\gamma^p)^{\frac{1}{2}}\zeta_0(x)^{\frac{1}{2}}\hat{\sigma}(x,p,0), (\gamma^p)^{\frac{1}{2}}\zeta_1(x)^{\frac{1}{2}}\hat{\sigma}(x,p,1), ..., (\gamma^p)^{\frac{1}{2}}\zeta_q(x)^{\frac{1}{2}}\hat{\sigma}(x,p,q)), \\ \overline{f}(x,\gamma) &= \sum_{i\in I(p)} \gamma^i \sum_{j\in I_0(q)} \zeta_j(x)\hat{f}(x,i,j). \end{aligned}$$

Note also that for all $(x,\gamma) \in \mathbf{R}^N \times C(p)$,

$$\begin{aligned} \Psi(\gamma,x) := (&\gamma^1\zeta_0(x), \gamma^1\zeta_1(x), ..., \gamma^1\zeta_q(x), \gamma^2\zeta_0(x), \gamma^2\zeta_1(x), ..., \gamma^2\zeta_q(x), \\ &..., \gamma^p\zeta_0(x), \gamma^p\zeta_1(x), ..., \gamma^p\zeta_q(x)) \in C(p(1+q)). \end{aligned}$$

Associated with the collection $C(p)$, $\overline{b}$, $\overline{\sigma}$, and $\overline{f}$, we define the set $\overline{\mathcal{C}}(x)$ of all controlled systems at $x \in \mathbf{R}^N$, the cost functional $J(x,\alpha)$ and the value function $\overline{V}(x)$.

Lemma 5.2. *For any $\varepsilon > 0$ there are $\rho \in (0,r]$ and $p \in \mathbf{N}$ such that*

$$U(x) + \varepsilon > \overline{V}(x) \qquad \forall x \in \overline{G}.$$

Proof. Note that conditions (A1) and (A2) are obviously satisfied with $\overline{\sigma}$, $\overline{b}$, $\overline{f}$, and $C(p)$ in place of σ, b, f, and A. Moreover, a condition much stronger than (A5) is satisfied for $\overline{\sigma}$, $\overline{b}$, $\overline{f}$, and $C(p)$. Indeed, we have:

(A8) For any $z \in \partial G$, there exists a constant $r_z > 0$ such that

(i) $\overline{\sigma}(x,\gamma) = 0 \qquad \forall x \in B(z,r_z) \cap \partial G,\ \forall \gamma \in C(p)$,

(ii) $TC(x, \overline{b}(x,\gamma), r_z) \subset \overline{G} \qquad \forall x \in B(z,r_z) \cap \overline{G},\ \gamma \in C(p)$.

We prove that (A8) is valid. Let $z \in \partial G$. Observe that if $\operatorname{spt}\zeta_j \cap B(z,\rho) \neq \emptyset$ for some $j \in I(q)$, then $B(z,\rho) \subset (\operatorname{spt}\zeta_j)_{2\rho}$. From this we see immediately that (i) of (A8) is satisfied.

Also, we see that if $x \in \operatorname{spt}\zeta_j \cap B(z,\rho) \cap \overline{G}$ for some $j \in I(q)$, then we have

$$TC(x, \zeta_j(x) b(x, \hat{a}_j(x)), \zeta_j(x)\rho) \subset TC(x, b(x, \hat{a}_j(x)), \rho) \subset \overline{G}.$$

If $x \in B(z,\rho) \cap \overline{G} \setminus \operatorname{spt}\zeta_j$ for some $j \in I(q)$, then it is clear that

$$TC(x, \zeta_j(x) b(x, \hat{a}_j(x)), \zeta_j(x)\rho) = \{x\} \subset \overline{G}.$$

Therefore, if $x \in B(z,\rho) \cap \overline{G}$, then we have

$$TC(x, \zeta_j(x) b(x, \hat{a}_j(x)), \zeta_(x)\rho) \subset \overline{G}.$$

Note also that $\zeta_0(x) = 0$ for all $x \in B(z,\rho)$ and therefore,

$$\overline{b}(x,\gamma) = \sum_{j \in I(q)} \zeta_j(x) b(x, \hat{a}_j(x)) \qquad \forall (x,\gamma) \in B(z,\rho) \cap C(p).$$

It is easily seen that for some $j \in I(q)$ we have

$$\zeta_j(z) \geq \frac{1}{q}.$$

We may assume by relabelling $(\hat{a}_j, \zeta_j)$'s that q is one of such j's. Reselecting ρ if necessary, we may assume that

$$\zeta_q(x) \geq \frac{1}{2q} \qquad \forall x \in B(z,\rho).$$

Fix $x \in B(z, \frac{\rho}{2})$. Choose $\varepsilon \in (0, \rho]$ sufficiently small so that $\varepsilon M \leq \frac{\rho}{2}$. Then, if $x \in B(z, \frac{\rho}{2})$, $0 \leq t \leq \varepsilon$, and $I \subset I(q)$, we have

$$x + t \sum_{j \in I} \zeta_j(x) b(x, \hat{a}_j(x)) \in B(z, \frac{\rho}{2} + \varepsilon M) \subset B(z,\rho).$$

By induction we see easily that if $x \in B(z, \frac{\rho}{2})$ and $0 \le t \le \varepsilon$, then

$$B(x + t \sum_{j \in I(q)} \zeta_j(x) b(x, \hat{a}_j(x)), \zeta_q(x)\rho) \subset \overline{G},$$

and hence

$$B(x + t \sum_{j \in I(q)} \zeta_j(x) b(x, \hat{a}_j(x)), \frac{\rho}{2q}) \subset \overline{G} \qquad \forall \gamma \in C(p).$$

This yields immediately (ii) of (A8).

We define

$$\overline{H}(x, v, X) = \sup_{\gamma \in C(p)} \{-\frac{1}{2} \operatorname{tr} \overline{\sigma\sigma}^T(x, \gamma) X - \overline{b}(x, \gamma) \cdot v - \overline{f}(x, \gamma)\}$$

for $(x, v, X) \in \mathbf{R}^N \times \mathbf{R}^N \times \mathcal{S}^N$. Observe that for all $\gamma \in C(p)$,

$$\operatorname{tr} \overline{\sigma\sigma}^T(x, \gamma) X = \sum_{i \in I(p)} \gamma_i \operatorname{tr} \tilde{\sigma}\tilde{\sigma}^T(x, i),$$

and that if $x \in \partial G$, then

$$\overline{H}(x, v, X) \ge -\xi(x) \cdot v - \chi(x),$$

where

$$\xi(x) = \sum_{j \in I(q)} \zeta_j(x) b(x, \hat{a}_j(x)),$$
$$\chi(x) = \sum_{j \in I(q)} \zeta_j(x) f(x, \hat{a}_j(x)).$$

Note that for all $x \in \partial G$,

$$(\xi(x), \chi(x)) \in \operatorname{co}\{(b(x, a), f(x, a)) \mid a \in A(x)\}.$$

Define

$$W_{\rho,p}(x) = \inf_{\alpha \in \overline{C}(x)} E^\alpha \int_0^\infty e^{-\lambda t} \overline{f}(X_t^\alpha, u_t^\alpha) dt \qquad \forall x \in \mathbf{R}^N.$$

Since $C(p)$ is convex, by a classical result (see, e.g., [NA, NI1, NI2]), we know that $u := W_{\rho,p}$ satisfies

$$\lambda u(x) + \overline{H}(x, Du(x), D^2 u(x)) = 0 \qquad \text{in } \mathbf{R}^N$$

in the viscosity sense, where $\overline{H} : \mathbf{R}^N \times \mathbf{R}^N \times \mathcal{S}^N \to \mathbf{R}$ is given by

$$\overline{H}(x, v, X) = \max_{a \in C(p)} \{-\frac{1}{2} \operatorname{tr} \overline{\sigma\sigma}^T(x, a) X - \overline{b}(x, a) \cdot v - \overline{f}(x, a)\}.$$

Since $F := \overline{H}$ satisfies condition (MP) of [ILT] with $K = \overline{G}$ by virtue of (A8), we see from ([ILT, Corollary 2.3]) that $u := W_{\rho,p}$ is also a viscosity solution of

$$\lambda u(x) + \overline{H}(x, Du(x), D^2u(x)) = 0 \qquad \text{in } \overline{G} \tag{32}$$

and that if $x \in \overline{G}$ and $\alpha \in \overline{\mathcal{C}}(x)$, then

$$X_t^\alpha \in \overline{G} \qquad \forall t \geq 0 \qquad P^\alpha\text{–a.s.},$$

i.e., any element of $\mathcal{C}_\delta(x)$, with $x \in \overline{G}$, is admissible with respect to $\overline{G}$. This shows that

$$\overline{V}(x) = W_{\rho,p}(x) \qquad \forall x \in \overline{G}.$$

Thus we need only to show that the viscosity solution $U_{\rho,p} := \overline{V}$ of (32) converges to U uniformly on $\overline{G}$ as $\rho \searrow 0$ and $p \to \infty$.

For this, we see that M/λ and $-M/\lambda$ are a supersolution and a subsolution of (32), respectively, and therefore by comparison that $|U_{\rho,p}(x)| \leq M/\lambda$ for $x \in \overline{G}$. If we define

$$U^*(x) = \lim_{gd \searrow 0} \sup\{U_{\rho,p}(y) \mid 0 < \rho < \delta,\ p > \delta^{-1},\ y \in \overline{G},\ |y - x| < \delta\},$$

and

$$U_*(x) = \lim_{\delta \searrow 0} \inf\{U_{\rho,p}(y) \mid 0 < \rho < \delta,\ p > \delta^{-1},\ y \in \overline{G},\ |y - x| < \delta\},$$

then U^* and U_* are upper and lower semicontinuous on $\overline{G}$, respectively, and $u := U^*$ and $w := U_*$ are a subsolution and a supersolution of the problems

$$\lambda u(x) + H_*(x, Du(x), D^2u(x)) = 0 \qquad \text{in } \overline{G},$$
$$\lambda w(x) + H^*(x, Dw(x), D^2w(x)) = 0 \qquad \text{in } \overline{G},$$

respectively, where H_* and H^* are defined for $(x, v, X) \in \overline{G} \times \mathbf{R}^N \times \mathcal{S}^N$ by,

$$H_*(x,v,X) = \lim_{r \searrow 0} \inf\{H_\delta(y,q,Y) \mid (y,q,Y) \in \overline{G} \times \mathbf{R}^N \times \mathcal{S}^N,\ |y-x| + |q-v| + |Y-X| < r\},$$

$$H^*(x,v,X) = \lim_{r \searrow 0} \sup\{H_\delta(y,q,Y) \mid (y,q,Y) \in \overline{G} \times \mathbf{R}^N \times \mathcal{S}^N,\ |y-x| + |q-v| + |Y-X| < r\}.$$

Note that for all $(x, v, X) \in G \times \mathbf{R}^N \times \mathcal{S}^N$,

$$H^*(x, v, X) = H_*(x, v, X) = H(x, v, X),$$

and for all $(x, v, X) \in \partial G \times \mathbf{R}^N \times \mathcal{S}^N$,

$$H^*(x, v, X) = H(x, v, X);$$

$$H_*(x,v,X) \geq -\xi(x)\cdot v - \chi(x).$$

The comparison result, Theorem 3.1, applied to the problem

$$\begin{cases} \lambda u(x) + H^*(x, Du(x), D^2u(x)) \geq 0 & \text{in } \overline{G}, \\ \lambda u(x) + H_*(x, Du(x), D^2u(x)) \leq 0 & \text{in } \overline{G}, \end{cases} \tag{33}$$

guarantees that $U^* \leq U_*$ in $\overline{G}$ and so $U^* = U_*$ in $\overline{G}$. Since U is a viscosity solution of (33), we conclude by the same comparison result that $U = U^* = U_*$ on $\overline{G}$. This immediately yields the uniform convergence of $U_{\rho,p}(x)$ to $U(x)$ for all $x \in \overline{G}$. QED

We fix $\varepsilon > 0$ and select $p \in \mathbf{N}$ so that

$$U(x) + \varepsilon > \overline{V}(x) \qquad \forall x \in \overline{G}.$$

Let $h > 0$ be a constant to be fixed later.

Lemma 5.3. *For any $x \in \mathbf{R}^N$ and $\delta > 0$ there is a controlled system $\overline{\alpha} \in \overline{\mathcal{C}}(x)$ such that*

$$\overline{V}(x) + \delta > \overline{E}\Big(\int_0^h e^{-\lambda t}\overline{f}(\overline{X}_t, \overline{\gamma}_t)dt + e^{-\lambda h}\overline{V}(\overline{X}_h)\Big)$$

and such that $\{\overline{\gamma}_t\}_{t\geq 0}$ is a continuous stochastic process, i.e., the function $t \mapsto \overline{\gamma}_t$ is continuous on $[0,\infty)$ almost surely, where

$$\overline{\alpha} = (\overline{\Omega}, \overline{\mathcal{F}}, \overline{P}, \{\overline{\mathcal{F}}_t\}_{t\geq 0}, \{\overline{W}_t\}_{t\geq 0}, \{\overline{\gamma}_t\}_{t\geq 0}, \{\overline{X}_t\}_{t\geq 0})$$

and $\overline{E}$ denotes the mathematical expectation with respect to $\overline{P}$.

Proof. Fix any $x \in \overline{G}$ and select $\alpha \in \overline{\mathcal{C}}(x)$ so that

$$E^\alpha \int_0^h e^{-\lambda t}\overline{f}(X_t^\alpha, u_t^\alpha)dt + e^{-\lambda h}\overline{V}(X_h^\alpha) < \overline{V}(x) + \frac{\delta}{2}.$$

Define the stochastic process $\{v_t\}_{t\in\mathbf{R}}$ with values in $C(p)$ by

$$v_t = u_t^\alpha \quad \forall t \geq 0 \quad \text{and} \quad v_t = (1,0,...,0) \in C(p) \quad \forall t < 0.$$

(Here the choice "$(1,0,...,0)$" is not important and any fixed $e \in C(p)$ can be used instead.)

For $k \in \mathbf{N}$ we define the $\mathcal{F}_t^\alpha$–progressively measurable continuous stochastic process $\{w_t^k\}_{t\geq 0}$ by

$$w_t^k = k\int_{t-1/k}^t v_s ds.$$

Let $\{X_t^k\}_{t\geq 0}$ be the strong solution of

$$dX_t^k = \overline{b}(X_t^k, w_t^k)dt + \overline{\sigma}(X_t^k, w_t^k)dW_t^\alpha, \quad X_0^k = x.$$

By standard estimates for solutions of SDE, we get

$$\begin{aligned} E^\alpha[\sup_{0\leq t\leq h} |X_t^\alpha - X_t^k|^4] \leq & C_1 E^\alpha \int_0^h (|\overline{b}(X_t^\alpha, u_t^\alpha) - \overline{b}(X_t^\alpha, w_t^k)|^4 \\ & + |\overline{\sigma}(X_t^\alpha, u_t^\alpha) - \overline{\sigma}(X_t^\alpha, w_t^k)|^4)dt \\ \leq & C_2 E^\alpha \int_0^h |u_t^\alpha - w_t^k|^2 dt, \end{aligned}$$

where C_1 and C_2 are positive constants independent of k. Here we have used the fact that the functions $\overline{\sigma}(x,\cdot)$ are 1/2–Hölder continuous on $C(p)$ uniformly in $x \in \mathbf{R}^N$ (see (30)).

Noting that as $k \to \infty$,

$$E^\alpha \int_0^h |w_t^k - u_t^\alpha|^2 dt \to 0,$$

we deduce that as $k \to \infty$,

$$|E^\alpha \int_0^h e^{-\lambda t}\overline{f}(X_t^\alpha, u_t^\alpha)dt - E^\alpha \int_0^h e^{-\lambda t}\overline{f}(X_t^k, w_t^k)dt| \to 0;$$
$$E^\alpha|\overline{V}(X_h^\alpha) - \overline{V}(X_h^k)| \to 0.$$

Hence we conclude that for sufficiently large k,

$$E^\alpha \int_0^h e^{-\lambda t}\overline{f}(X_t^k, w_t^k)dt < \overline{V}(x) + \delta. \qquad \text{QED}$$

Define functions $b^*,\ \sigma^*,\ f^*$ on $\mathbf{R}^N \times C(p(1+q))$ by

$$\begin{aligned} b^*(x,\mu) &= \sum_{i\in I(p)} \sum_{j\in I_0(q)} \mu^{ij}\hat{b}(x,i,j), \\ \sigma^*(x,\mu) &= \big((\mu^{10})^{\frac{1}{2}}\hat{\sigma}(x,1,0), (\mu^{11})^{\frac{1}{2}}\hat{\sigma}(x,1,1), ..., (\mu^{pq})^{\frac{1}{2}}\hat{\sigma}(x,p,q)\big), \\ f^*(x,\mu) &= \sum_{i\in I(p)} \sum_{j\in I(q)} \mu^{ij}\hat{f}(x,i,j), \end{aligned}$$

where $\mu = (\mu^{10}, \mu^{11}, ..., \mu^{pq})$. For any $x \in \mathbf{R}^N$ let $\mathcal{C}^*(x)$ denote the set of controlled systems at x associated with $C(p(1+q)),\ b^*$, and σ^*.

Lemma 5.4. *For any $x \in \mathbf{R}^N$ and any $\delta > 0$ there is a controlled system*

$$(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\geq 0}, \{W_t\}_{t\geq 0}, \{\mu_t\}_{t\geq 0}, \{X_t\}_{t\geq 0}) \in \mathcal{C}^*(x)$$

such that

$$\overline{V}(x) + \delta > E\Big(\int_0^h e^{-\lambda t} f^*(X_t, \mu_t)dt + e^{-\lambda h}\overline{V}(X_h)\Big);$$
$$\mu_t = \Psi(\gamma_{[mh^{-1}t]m^{-1}h}, X_{[mh^{-1}t]m^{-1}h}) \qquad \forall t \geq 0$$

for some $m \in \mathbf{N}$ and some $\mathcal{F}_t$–progressively measurable $C(p)$–valued stochastic process $\{\gamma_t\}_{t\geq 0}$.

Proof. By definition, we have

$$\overline{b}(x,\gamma) = b^*(x, \Psi(\gamma, x)), \qquad \overline{\sigma}(x,\gamma) = \sigma^*(x, \Psi(\gamma, x)), \qquad \overline{f}(x,\gamma) = f^*(x, \Psi(\gamma, x))$$

for all $(x, \gamma) \in \mathbf{R}^N \times C(p)$.

Fix $x \in \mathbf{R}^N$ and $\delta > 0$. According to Lemma 5.3, there is a controlled system $\alpha \in \overline{\mathcal{C}}(x)$ such that

$$\overline{V}(x) + \frac{\delta}{2} > E^\alpha\Big(\int_0^h e^{-\lambda t}\overline{f}(X_t^\alpha, \gamma_t^\alpha)dt + e^{-\lambda h}\overline{V}(X_h^\alpha)\Big)$$

and such that $\{\gamma_t^\alpha\}_{t\geq 0}$ is a continuous stochastic process.

For each $k \in \mathbf{N}$, we define the $\mathcal{F}_t^\alpha$–progressively measurable stochastic processes $\{\mu_t\}_{t\geq 0}$ and $\{\mu_t^k\}_{t\geq 0}$ with values in $C(p(1+q))$ by

$$\mu_t = \Psi(\gamma_t^\alpha, X_t^\alpha), \qquad \mu_t^k = \mu_{[kh^{-1}t]k^{-1}h}.$$

Recall that for $t \in \mathbf{R}$, $[t]$ denotes the largest integer less than or equal to t.

Note that $\{\mu_t^k\}_{t\geq 0}$ is a step process with step length $k^{-1}h$.

Let $\{X_t^k\}_{t\geq 0}$ be the strong solution of

$$dX_t^k = b^*(X_t^k, \mu_t^k)dt + \sigma^*(X_t^k, \mu_t^k)dW_t^\alpha, \qquad X_0^k = x.$$

Since X_t^α is a strong solution of

$$dX_t^\alpha = b^*(X_t^\alpha, \mu_t)dt + \sigma^*(X_t^\alpha, \mu_t)dW_t^\alpha, \qquad X_0^\alpha = x,$$

as in the previous proof, we have

$$E^\alpha[\sup_{0\leq t\leq h} |X_t^\alpha - X_t^k|^4] \leq C_3 E^\alpha \int_0^h |\mu_t^\alpha - \mu_t^k|^2 dt,$$

where C_3 is a positive constant independent of k.

Since

$$E^\alpha \int_0^h e^{-\lambda t}\overline{f}(X_t^\alpha, \gamma_t^\alpha)dt = E^\alpha \int_0^h e^{-\lambda t} f^*(X_t^\alpha, \mu_t)dt$$

and

$$\mu_t^k \to \mu_t \quad \text{in } C([0,h]) \qquad P^\alpha\text{–a.s.}$$

as $k \to \infty$, we see that for sufficiently large k,

$$E^\alpha\Big(\int_0^h e^{-\lambda t} f^*(X_t^k, \mu_t^k)dt + e^{-\lambda h}\overline{V}(X_h^\alpha)\Big) < \overline{V}(x) + \delta. \qquad \text{QED}$$

For $x \in \mathbf{R}^N$ let $\hat{\mathcal{C}}(x)$ denote the set of controlled systems at x associated with K, $\hat{b}$, $\hat{\sigma}$.

Observe that the pair of functions b^*, σ^* satisfies condition (MP) of [ILT] with respect to $\overline{G}$. Indeed, it is easy to check that

$$\sigma^*(x,\mu) = 0 \qquad \forall (x,\mu) \in \mathbf{R}^N \times C(p(1+q)),$$

and arguments similar to the proof of (A8) work to show that for each $(z,\mu) \in \overline{G} \times C(p(1+q))$ there is a constant $c > 0$ such that

$$z + tb^*(z,\mu) \in \overline{G} \qquad \forall t \in [0,c].$$

Therefore, for any $x \in \overline{G}$ and any controlled system

$$(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\geq 0}, \{W_t\}_{t\geq 0}, \{\mu_t\}_{t\geq 0}, \{X_t\}_{t\geq 0}) \in \mathcal{C}^*(x),$$

we have

$$X_t \in \overline{G} \quad \forall t \geq 0 \qquad P\text{–a.s.}$$

Moreover, for any $x \in \mathbf{R}^N$ and any

$$(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\geq 0}, \{W_t\}_{t\geq 0}, \{u_t\}_{t\geq 0}, \{X_t\}_{t\geq 0}) \in \hat{\mathcal{C}}(x),$$

if we set

$$\mu_t(\omega) = \hat{e}_{\hat{u}_t(\omega)} \qquad \forall (t,\omega) \in [0,\infty) \times \hat{\Omega},$$

where $\hat{e}_u$, with $u \equiv (i,j) \in K$, denotes the unit vector of $\mathbf{R}^{p(1+q)}$ with unity as its (i,j) component, then

$$(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\geq 0}, \{W_t\}_{t\geq 0}, \{\mu_t\}_{t\geq 0}, \{X_t\}_{t\geq 0}) \in \mathcal{C}^*(x).$$

Hence, if $x \in \overline{G}$ and

$$(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\geq 0}, \{W_t\}_{t\geq 0}, \{u_t\}_{t\geq 0}, \{X_t\}_{t\geq 0}) \in \hat{\mathcal{C}}(x),$$

then

$$X_t \in \overline{G} \quad t \geq 0 \qquad P\text{–a.s.}$$

The following Lemma is an easy consequence of Lemmas 5.1 and 5.4.

Lemma 5.5. *For any $\delta > 0$ and any $x \in \mathbf{R}^N$ there are controlled systems*

$$(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\geq 0}, \{W_t\}_{t\geq 0}, \{\mu_t\}_{t\geq 0}, \{X^*_t\}_{t\geq 0}) \in \mathcal{C}^*(x),$$
$$(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\geq 0}, \{W_t\}_{t\geq 0}, \{u_t\}_{t\geq 0}, \{\hat{X}_t\}_{t\geq 0}) \in \hat{\mathcal{C}}(x)$$

such that

$$\overline{V}(x) + \delta > E\Big[\int_0^h e^{-\lambda t}\hat{f}(\hat{X}_t, u_t)dt + e^{-\lambda h}\overline{V}(\hat{X}_h)\Big]$$

and such that for any $\omega \in \Omega$, $j \in I_0(q)$, and $\tau \in (0, h]$, if

$$\zeta_j(X^*_t(\omega)) \equiv 0 \qquad on \ \ [0, \tau),$$

then

$$u_t(\omega) \neq (i, j) \qquad \forall (t, i) \in [0, \tau) \times I(p).$$

Proof. Fix $x \in \overline{G}$ and $\delta > 0$. By Lemma 5.4, there is a controlled system

$$(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\geq 0}, \{W_t\}_{t\geq 0}, \{\mu_t\}_{t\geq 0}, \{X_t\}_{t\geq 0}) \in \mathcal{C}^*(x)$$

such that

$$\overline{V}(x) + \frac{\delta}{2} > E\Big(\int_0^h e^{-\lambda t} f^*(X_t, \mu_t)dt + e^{-\lambda h}\overline{V}(X_h)\Big);$$
$$\mu_t = \Psi(\gamma_{[mh^{-1}t]m^{-1}h}, X_{[mh^{-1}t]m^{-1}h}) \qquad \forall t \geq 0 \tag{34}$$

for some $m \in \mathbf{N}$ and some $\mathcal{F}_t$–progressively measurable stochastic process $\{\gamma_t\}_{t\geq 0}$.

Next, by Lemma 5.1, there is a controlled system

$$(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P}, \{\hat{\mathcal{F}}_t\}_{t\geq 0}, \{\hat{W}_t\}_{t\geq 0}, \{\hat{u}_t\}_{t\geq 0}, \{\hat{X}_t\}_{t\geq 0}) \in \hat{\mathcal{C}}(x),$$

where $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P}, \{\hat{\mathcal{F}}_t\}_{t\geq 0})$ is an extension of the filtered probability space $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\geq 0})$, such that

$$E\Big[\int_0^h e^{-\lambda t}\overline{f}(X_t, \mu_t)dt + e^{-\lambda h}\overline{V}(X_h)\Big] + \frac{\delta}{2} > \hat{E}\Big[\int_0^h \hat{f}(\hat{X}_t, \hat{u}_t)e^{-\lambda t}dt + e^{-\lambda h}\overline{V}(\hat{X}_h)\Big],$$

where $\hat{E}$ denote the mathematical expectation with respect to $\hat{P}$, and such that for any $\tau > 0$, $\omega \in \hat{\Omega}$, and $(i,j) \in C(p(1+q))$, if $\mu_t^{ij}(\omega) \equiv 0$ on $[0,\tau)$, then

$$\hat{u}_t(\omega) \neq (i,j) \qquad \forall t \in [0,\tau).$$

It is immediate to see that

$$\overline{V}(x) + \delta > \hat{E}\Big[\int_0^h \hat{f}(\hat{X}_t, \hat{u}_t) e^{-\lambda t} dt + e^{-\lambda h} \overline{V}(\hat{X}_h)\Big].$$

Let $\omega \in \Omega$, $\tau \in (0,h]$, and $j \in I_0(q)$. Suppose that $\zeta_j(X_t(\omega)) \equiv 0$ on $[0,\tau)$. By (34), we have

$$\mu_t^{ij}(\omega) = 0 \qquad \forall (t,i) \in [0,\tau) \times I(p),$$

and so,

$$\hat{u}_t(\omega) \neq (i,j) \qquad \forall t \in [0,\tau).$$

The proof is complete. QED

For

$$\alpha \equiv (\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\geq 0}, \{W_t\}_{t\geq 0}, \{u_t\}_{t\geq 0}, \{X_t\}_{t\geq 0}) \in \hat{\mathcal{C}}(x),$$

we set

$$\hat{J}_h(x,\alpha) = E\Big(\int_0^h e^{-\lambda t} \hat{f}(X_t, u_t) dt + e^{-\lambda h} \overline{V}(X_h)\Big).$$

Let $\hat{m}$ and $\overline{m}$ denotes the moduli of continuity of $\hat{f}$ and $\overline{V}$, respectively.

Lemma 5.6. *For any $x, y \in \mathbf{R}^N$ and any*

$$\alpha_x \equiv (\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\geq 0}, \{W_t\}_{t\geq 0}, \{u_t\}_{t\geq 0}, \{X_t\}_{t\geq 0}) \in \hat{\mathcal{C}}(x),$$
$$\alpha_y \equiv (\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\geq 0}, \{W_t\}_{t\geq 0}, \{u_t\}_{t\geq 0}, \{Y_t\}_{t\geq 0}) \in \hat{\mathcal{C}}(y),$$

we have

$$|\hat{J}_h(x,\alpha_x) - \hat{J}_h(y,\alpha_y)| \leq \lambda^{-1} \hat{m}(C|x-y|) + \overline{m}(C|x-y|)$$

for some constant $C \equiv C(h, M, \eta_0, ..., \eta_q) > 0$.

Proof. Note that $\hat{m}$ and $\overline{m}$ are non-decreasing concave functions. We calculate that

$$\begin{aligned}
&|\hat{J}_h(x,\alpha_x) - \hat{J}_h(y,\alpha_y)| \leq E\int_0^h e^{-\lambda t} \hat{m}(|X_t - Y_t|) dt + E e^{-\lambda h} \overline{m}(|X_h - Y_h|) \\
&\leq \hat{m}(\sup_{t\leq h} |X_t - Y_t|) E\int_0^h e^{-\lambda t} dt + e^{-\lambda h} E\overline{m}(\sup_{t\leq h} |X_t - Y_t|) \\
&\leq \lambda^{-1} \hat{m}(E\sup_{t\leq h} |X_t - Y_t|) + E\overline{m}(\sup_{t\leq h} |X_t - Y_t|) \\
&\leq \lambda^{-1} \hat{m}((E\sup_{t\leq h} |X_t - Y_t|^2)^{\frac{1}{2}}) + \overline{m}((E\sup_{t\leq h} |X_t - Y_t|^2)^{\frac{1}{2}}) \\
&\leq \lambda^{-1} \hat{m}(C|x-y|) + \overline{m}(C|x-y|),
\end{aligned}$$

and finish the proof. QED

Completion of the proof of Theorem 3.6. Assume that $h \le 1$. Fix $T > 0$ and $m \in \mathbf{N}$ so that

$$\lambda^{-1} e^{-\lambda T} M < \varepsilon \quad \text{and} \quad T \le mh < T + 1.$$

By using Lemmas 5.5 and 5.6 we infer that there is a constant

$$\delta \equiv \delta(h, M, \eta_0, ..., \eta_q) > 0$$

such that for each $x \in \mathbf{R}^N$ there are

$$\alpha_x^* \equiv (\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\ge 0}, \{W_t\}_{t\ge 0}, \{\mu_t\}_{t\ge 0}, \{X_t^*\}_{t\ge 0}) \in \mathcal{C}^*(x),$$
$$\hat{\alpha}_x \equiv (\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\ge 0}, \{W_t\}_{t\ge 0}, \{u_t\}_{t\ge 0}, \{X_t\}_{t\ge 0}) \in \hat{\mathcal{C}}(x)$$

having the properties:

(i) if $y \in B(x, \delta)$ and Y_t is the strong solution of

$$\begin{cases} dY_t = \hat{b}(Y_t, u_t)dt + \hat{\sigma}(Y_t, u_t)dW_t & \text{for } t > 0, \\ Y_0 = y, \end{cases}$$

then

$$\overline{V}(y) + h^2 > E\Big(\int_0^h e^{-\lambda t} \hat{f}(Y_t, u_t)dt + e^{-\lambda h}\overline{V}(Y_h)\Big);$$

(ii) for any $\omega \in \Omega$, if

$$\zeta_j(X_t^*)) \equiv 0 \qquad \text{on} \qquad [0,\ \tau)$$

for some $j \in I_0(q)$ and $0 < \tau \le h$, then

$$u_t(\omega) \ne (i, j) \qquad \forall (t, i) \in [0,\ \tau) \times I(q).$$

We may assume that $2\delta < \rho$.

We can choose a finite family $\{B_k\}_{k\in I(\nu)}$ of disjoint Borel subsets of $\mathbf{R}^N$ such that

$$\bigcup_{k\in I(\nu)} B_k = \overline{G},$$
$$B_k \subset B(x_k, \delta) \qquad \forall k \in I(\nu)$$

for some $x_1, ..., x_\nu \in \overline{G}$.

For each $k \in I(\nu)$ we fix

$$\alpha_{x_k}^* \equiv (\Omega^k, \mathcal{F}^k, P^k, \{\mathcal{F}_t^k\}_{t\ge 0}, \{W_t^k\}_{t\ge 0}, \{\mu_t^k\}_{t\ge 0}, \{X_t^{*k}\}_{t\ge 0}) \in \mathcal{C}^*(x_k),$$
$$\hat{\alpha}_{x_k} \equiv (\Omega^k, \mathcal{F}^k, P^k, \{\mathcal{F}_t^k\}_{t\ge 0}, \{W_t^k\}_{t\ge 0}, \{u_t^k\}_{t\ge 0}, \{X_t^k\}_{t\ge 0}) \in \hat{\mathcal{C}}(x_k)$$

so that (i) and (ii) above hold with $x = x_k$.

We define stopping times $\tau_1^*, ..., \tau_\nu^*$ by

$$\tau_k^* = \inf\{t > 0 \mid |X_t^{*k} - x_k| \geq \rho\} \wedge h.$$

Set

$$B_0 = \mathbf{R}^N \setminus \overline{G},$$

and fix

$$(\Omega^0, \mathcal{F}^0, P^0, \{\mathcal{F}_t^0\}_{t\geq 0}, \{W_t^0\}_{t\geq 0}, \{u_t^0\}_{t\geq 0})$$

arbitrarily so that $(\Omega^0, \mathcal{F}^0, P^0, \{\mathcal{F}_t^0\}_{t\geq 0}$ is a filtered probability space satisfying the usual condition, $\{W_t^0\}_{t\geq 0}$ is an l–dimensional standard Brownian motion on $(\Omega^0, \mathcal{F}^0, P^0, \{\mathcal{F}_t^0\}_{t\geq 0}$, and $\{u_t^0\}_{t\geq 0}$ is an $\mathcal{F}_t^0$–progressively measurable process taking values in K.

Fix $x \in \overline{G}$. We intend to define a sequence $\{\alpha_n\}_{n\in I(m)}$ of controlled systems

$$\alpha_n \equiv (\hat{\Omega}^n, \hat{\mathcal{F}}^n, \hat{P}^n, \{\hat{\mathcal{F}}_t^n\}_{t\geq 0}, \{\hat{W}_t^n\}_{t\geq 0}, \{\hat{u}_t^n\}_{t\geq 0}, \{X_t^n\}_{t\geq 0}) \in \hat{\mathcal{C}}(x),$$

together with a sequence $\{\hat{\tau}_n\}_{n\in I(m)}$ of stopping times.

We may assume by relabelling $\{B_k\}$ that $x \in B_1$. Let $\hat{X}_t^1$ be the strong solution of

$$d\hat{X}_t^1 = \hat{b}(\hat{X}_t^1, u_t^1)dt + \hat{\sigma}(\hat{X}_t^1, u_t^1)dW_t^1, \qquad \hat{X}_0^1 = x.$$

Set

$$\hat{\Omega}^1 = \Omega^1, \quad \hat{\mathcal{F}}^1 = \mathcal{F}^1, \quad \hat{P}^1 = P^1, \quad \hat{\mathcal{F}}_t^1 = \mathcal{F}_t^1, \quad \hat{W}_t^1 = W_t^1, \quad \hat{u}_t^1 = u_t^1.$$

Define the stopping time $\hat{\tau}^1$ by

$$\hat{\tau}^1 = \inf\{t > 0 \mid |\hat{X}_t^1 - x_1| \geq \rho\} \wedge \tau_1^*(\omega).$$

Of course, we have

$$(\hat{\Omega}^1, \hat{\mathcal{F}}^1, \hat{P}^1, \{\hat{\mathcal{F}}_t^1\}_{t\geq 0}, \{\hat{W}_t^1\}_{t\geq 0}, \{\hat{u}_t^1\}_{t\geq 0}, \{\hat{X}_t^1\}_{t\geq 0}) \in \hat{\mathcal{C}}(x),$$

and

$$\overline{V}(x) + h^2 > \hat{E}^1\Big(\int_0^h e^{-\lambda t}\hat{f}(\hat{X}_t^1, \hat{u}_t^1)dt + e^{-\lambda h}\overline{V}(\hat{X}_h^1)\Big).$$

Here and later $\hat{E}^i$ denotes the mathematical expectation with respect to $\hat{P}^i$.

Let $\omega \in \hat{\Omega}^1$ and $t \in [0, \hat{\tau}_1(\omega)]$. Then

$$\sup_{0\leq s\leq t} |X_s^{*1}(\omega) - x_1| \leq \rho, \qquad |\hat{X}_t^1(\omega) - x_1| < \rho.$$

We see from (ii) that if $\hat{u}_t^1(\omega) = (i, j)$ for some $(i, j) \in K$, then

$$\zeta_j(X_s^{*1}(\omega)) > 0 \qquad \text{for some} \quad s \in [0,\ t],$$

and hence,

$$|\hat{X}_t^1(\omega) - X_s^{*1}(\omega)| \le |\hat{X}_t^1(\omega) - x_1| + |x_1 - X_s^{*1}(\omega)| < 2\rho.$$

Since $\eta_j = 1$ on $(\operatorname{spt}\zeta_j)_{2\rho}$, we have

$$\begin{aligned}
\hat{b}(\hat{X}_t^1(\omega), \hat{u}_t^1(\omega)) &= b_0(\hat{X}_t^1(\omega), \hat{u}_t^1(\omega)),\\
\hat{\sigma}(\hat{X}_t^1(\omega), \hat{u}_t^1(\omega)) &= \sigma_0(\hat{X}_t^1(\omega), \hat{u}_t^1(\omega)),\\
\hat{f}(\hat{X}_t^1(\omega), \hat{u}_t^1(\omega)) &= f_0(\hat{X}_t^1(\omega), \hat{u}_t^1(\omega)),
\end{aligned}$$

where functions $b_0,\ \sigma_0,\ f_0$ on $\mathbf{R}^N \times K$ are defined by

$$\begin{aligned}
b_0(y, i, j) &= \begin{cases} b(y, a_i) & \text{if } j = 0,\\ b(y, \hat{a}_j(y)) & \text{if } j \ge 1,\end{cases}\\
\sigma_0(y, i, j) &= \begin{cases} \sigma(y, a_i) & \text{if } j = 0,\\ \sigma(y, \hat{a}_j(y)) & \text{if } j \ge 1,\end{cases}\\
f_0(y, i, j) &= \begin{cases} f(y, a_i) & \text{if } j = 0,\\ f(y, \hat{a}_j(y)) & \text{if } j \ge 1.\end{cases}
\end{aligned}$$

By using a standard estimate for solutions of stochastic differential equations, we get

$$E^k \sup_{0 \le t \le h} |X_t^{*k} - x_k|^4 \le Ch^2,$$

where $C \equiv C(\hat{b}, \hat{\sigma})$ is a positive constant and E^k denotes the mathematical expectation with respect to P^k. From this we get

$$P^k(\sup_{0 \le t \le h} |X_t^{*k} - x_k| \ge \rho) \le \frac{Ch^2}{\rho^4},$$

i.e.,

$$P^k(\tau_k^* < h) \le \frac{Ch^2}{\rho^4}.$$

Similarly, setting

$$\tau_1(\omega) = \inf\{t > 0 \mid |\hat{X}_t^1(\omega) - x| > \frac{\rho}{2}\},$$

we have

$$P^1(\tau_1 < h) \le C\Big(\frac{2}{\rho}\Big)^4 h^2.$$

Since

$$|\hat{X}_t^1(\omega) - x_1| \ge \rho \Longrightarrow |\hat{X}_t^1(\omega) - x| \ge \frac{\rho}{2},$$

we have

$$\hat{P}^1(\hat{\tau}_1 < h) \le P^1(\tau_1 < h) + P^1(\tau_1^* < h) \le 2C_\rho h^2,$$

where $C_\rho := C(\frac{2}{\rho})^4$.

Set

$$\mathcal{G}_t = \mathcal{F}_t^0 \otimes \cdots \otimes \mathcal{F}_t^\nu, \qquad \Sigma = \Omega^\times \cdots \times \Omega^\nu, \qquad Q = P^0 \otimes \cdots \otimes P^\nu.$$

Set

$$\hat{\Omega}^2 = \hat{\Omega}^1 \times \Sigma, \qquad \hat{P}^2 = P^1 \otimes Q, \qquad \hat{\mathcal{F}}_t^2 = \mathcal{F}_t^1 \otimes \mathcal{G}_{(t-h)_+}.$$

For $\omega \in \hat{\Omega}^2$ we write

$$\omega = (\omega^0, \omega^1) = (\omega^0, \omega^{10}, ..., \omega^{1\nu}).$$

Define $\hat{\mathcal{F}}_t^2$–progressively measurable processes $\{\hat{u}_t^2\}_{t\ge0}$ and $\{\hat{W}_t^2\}_{t\ge0}$ by

$$\hat{u}_t^2(\omega) = \begin{cases} u_t^1(\omega^0) & \text{if} \quad 0 \le t < h, \\ u_{t-h}^k(\omega^{1k} & \text{if} \quad t \ge h \text{ and } \hat{X}_h^1(\omega^0) \in B_k, \end{cases}$$
$$\hat{W}_t^2(\omega) = \begin{cases} W_t^1(\omega^0) & \text{if} \quad 0 \le t < h, \\ W_{t-h}^k(\omega^{1k}) & \text{if} \quad t \ge h \text{ and } \hat{X}_h^1(\omega^0) \in B_k, \end{cases}$$

respectively. It is easily seen that $\hat{W}_t^2$ is an l–dimensional standard Brownian motion on $(\hat{\Omega}^2, \hat{\mathcal{F}}^2, \hat{P}^2, \{\hat{\mathcal{F}}_t^2\}_{t\ge0})$, where

$$\hat{\mathcal{F}}^2 := \bigvee_{t\ge0} \hat{\mathcal{F}}_t^2.$$

Let $\hat{X}_t^2$ be the strong solution of

$$d\hat{X}_t^2 = \hat{b}(\hat{X}_t^2, \hat{u}_t^2)dt + \hat{\sigma}(\hat{X}_t^2, \hat{u}_t^2)d\hat{W}_t^2, \qquad \hat{X}_0^2 = x. \tag{35}$$

We then have

$$(\hat{\Omega}^2, \hat{\mathcal{F}}^2, \hat{P}^2, \{\hat{\mathcal{F}}_t^2\}_{t\ge0}, \{\hat{W}_t^2\}_{t\ge0}, \{\hat{u}_t^2\}_{t\ge0}, \{\hat{X}_t^2\}_{t\ge0}) \in \hat{\mathcal{C}}(x).$$

We set

$$I = \{k \in I_0(\nu) \mid \hat{P}^1(\hat{X}_h^1(\omega) \in B_k) > 0\}.$$

Note that, since

$$\hat{X}_t^1(\omega) \in \overline{G} \quad \forall t \ge 0 \qquad \hat{P}^1\text{–a.s.},$$

we have $0 \notin I$. For $k \in I$ we set

$$C^k = \{\omega \in \hat{\Omega}^1 \mid \hat{X}_h^1(\omega) \in B_k\},$$
$$Q^k(B) = \frac{\hat{P}^1(B \cap C^k)}{\hat{P}^1(C^k)} \qquad \forall B \in \hat{\mathcal{F}}^1.$$

Then Q^k is a probability measure on $(\hat{\Omega}^1, \hat{\mathcal{F}}^1)$ and we have

$$\hat{P}^1(B \cap C^k) = \hat{P}^1(C^k) Q^k(B) \qquad \forall B \in \hat{\mathcal{F}}^1$$

and

$$\hat{P}^1(B) = \sum_{k \in I} \hat{P}^1(C^k) Q^k(B) \qquad \forall B \in \hat{\mathcal{F}}^1.$$

For $k \in I$ we set

$$\hat{\Omega}^{2k} = \hat{\Omega}^1 \times \Omega^k, \quad \hat{P}^{2k} = Q^k \otimes P^k, \quad \hat{\mathcal{F}}_t^{2k} = \mathcal{F}_h^1 \otimes \mathcal{F}_t, \quad \hat{\mathcal{F}}^{2k} = \bigvee_{t \geq 0} \hat{\mathcal{F}}_t^{2k}.$$

For $k \in I$ let $\hat{X}_t^{2k}$ be the strong solution of

$$d\hat{X}_t^{2k} = \hat{b}(\hat{X}_t^{2k}, u_t^k)dt + \hat{\sigma}(\hat{X}_t^{2k}, u_t^k) d\hat{W}_t^{2k}, \qquad \hat{X}_0^{2k} = \hat{X}_h^1$$

on the filtered probability space

$$(\hat{\Omega}^{2k}, \hat{\mathcal{F}}^{2k}, \hat{P}^{2k}, \{\hat{\mathcal{F}}_t^{2k}\}_{t \geq 0}).$$

In view of the uniqueness of strong solutions of (35), we see that

$$\hat{X}_t^2(\omega) = \begin{cases} \hat{X}_t^1(\omega^0) & \text{if} \quad 0 \leq t < h, \\ \hat{X}_{t-h}^2(\omega^0, \omega^{1k}) & \text{if} \quad t \geq h \text{ and } \hat{X}_h^1(\omega^0) \in B_k, k \in I. \end{cases}$$

We calculate that

$$\begin{aligned}
&\hat{E}^2\Big(\int_0^{2h} e^{-\lambda t} \hat{f}(\hat{X}_t^2, \hat{u}_t^2)dt + e^{-2\lambda h} \overline{V}(\hat{X}_{2h}^2)\Big) \\
&= \hat{E}^2\Big(\int_0^{h} e^{-\lambda t} \hat{f}(\hat{X}_t^2, \hat{u}_t^2)dt + e^{-\lambda h} \int_0^h e^{-\lambda t} \hat{f}(\hat{X}_{t+h}^2, \hat{u}_{t+h}^2)dt + e^{-2\lambda h} \overline{V}(\hat{X}_{2h}^2)\Big) \\
&= \hat{E}^1 \int_0^h e^{-\lambda t} \hat{f}(\hat{X}_t^1, u_t^1)dt \\
&\quad + e^{-\lambda h} \sum_{k \in I} \hat{P}^1(C^k) \tilde{E}^k \Big(\int_0^h e^{-\lambda t} \hat{f}(\hat{X}_t^{2k}(\omega^0, \omega^{1k}), u_t^k(\omega^{1k})dt + e^{-\lambda h} \overline{V}(\hat{X}_h^{2k}(\omega^0, \omega^{1k})\Big)
\end{aligned}$$

(where $\tilde{E}^k$ denotes the mathematical expectation with respect to probability $Q^k \otimes P^k$,)

$$\leq \hat{E}^1 \int_0^h e^{-\lambda t} \hat{f}(\hat{X}_t^1, u_t^1)dt + e^{-\lambda h} \sum_{k \in I} \hat{P}^1(C^k) \tilde{E}^k \left(\overline{V}(\hat{X}_h^2) + h^2\right)$$

(by (i) since $Q^k \otimes P^k$-almost surely we have $\hat{X}_h^{2k} \in B_k$,)

$$
\begin{aligned}
&< \hat{E}^1 \int_0^h e^{-\lambda t} \hat{f}(\hat{X}_t^1, u_t^1) dt + e^{-\lambda h} \sum_{k \in I} \hat{P}^1(C^k) \int \overline{V}(\hat{X}_h^1) Q^k(d\omega^0) + h^2 \\
&= \hat{E}^1 \int_0^h e^{\lambda t} \hat{f}(\hat{X}_t^1, u_t^1) dt + e^{-glh} \sum_{k \in I} \int \overline{V}(\hat{X}_h^1) \mathbf{1}_{C^k} \hat{P}^1(d\omega^0) + h^2 \\
&= \hat{E}^1 \Big(\int_0^h e^{-\lambda t} \hat{f}(\hat{X}_t^1, u_t^1) dt + e^{-\lambda h} \overline{V}(\hat{X}_h^1) \Big) + h^2 \\
&< \overline{V}(x) + 2h^2.
\end{aligned}
$$

Thus we obtain

$$
\overline{V}(x) + 2h^2 > \hat{E}^2 \Big(\int_0^{2h} e^{-\lambda t} \hat{f}(\hat{X}_t^2, \hat{u}_t^2) dt + e^{-2\lambda h} \overline{V}(\hat{X}_{2h}^2) \Big).
$$

Define the stopping time $\hat{\tau}_2$ by

$$
\hat{\tau}_2(\omega) = \begin{cases} \hat{\tau}_1(\omega^0) & \text{if} \quad \hat{\tau}_1(\omega^0) < h, \\ h + \inf\{t > 0 \mid |\hat{X}_t^{2k}(\omega^0, \omega^{1k} - x_k| \geq \rho\} & \text{if} \quad \hat{\tau}_1(\omega^0) \geq h \text{ and} \\ & \qquad \hat{X}_h^1(\omega^0) \in B_k, \ k \in I. \end{cases}
$$

Fix any $\omega \in \hat{\Omega}^2$ and $0 \leq t \leq \hat{\tau}_2(\omega)$. Consider the case when $t < h$. Then we have

$$
t \leq \hat{\tau}_1(\omega^0), \quad \hat{X}_t^2(\omega) = \hat{X}_t^1(\omega^0), \quad \hat{u}_t^2(\omega) = \hat{u}_t^1(\omega^0),
$$

and hence

$$
\begin{aligned}
\hat{b}(\hat{X}_t^2(\omega), \hat{u}_t^2(\omega)) &= b_0(\hat{X}_t^2(\omega), \hat{u}_t^2(\omega)), \\
\hat{\sigma}(\hat{X}_t^2(\omega), \hat{u}_t^2(\omega)) &= \sigma_0(\hat{X}_t^2(\omega), \hat{u}_t^2(\omega)), \\
\hat{f}(\hat{X}_t^2(\omega), \hat{u}_t^2(\omega)) &= f_0(\hat{X}_t^2(\omega), \hat{u}_t^2(\omega)).
\end{aligned}
$$

Next we consider the case when $t \geq h$. For some $k \in I$ we have

$$
\hat{X}_h^1(\omega^0) \in B_k \quad \text{and} \quad \hat{\tau}_1(\omega^0) \geq h.
$$

By the definition of $\hat{\tau}_2$, we see that

$$
\sup_{0 \leq s \leq t-h} |X_s^{*k}(\omega^{1k}) - x_k| \leq \rho, \qquad |X_s^{2k}(\omega^0, \omega^{1k}) - x_k| < \rho.
$$

Using the fact that $\eta_j = 1$ on $(\operatorname{spt} \zeta_j)_{2\rho}$, we deduce as before that

$$
(\hat{b}, \hat{\sigma}, \hat{f})(\hat{X}_t^2(\omega), \hat{u}^2(\omega)) = (b_0, \sigma_0, f_0)(\hat{X}_t^2(\omega), \hat{u}^2(\omega)).
$$

We have
$$\begin{aligned}\hat{P}^2(\hat{\tau}_1 < h) &= \hat{P}^2(\hat{\tau}^2 < h) + \hat{P}^2(h \le \hat{\tau}^2 < 2h) \\ &\le \hat{P}^1(\hat{\tau}_1 < h) + \sum_{k \in I} \hat{P}^2(\omega^0 \in C^k,\ \tau_k^* < h) \\ &\quad + \sum_{k\in I} \hat{P}^2(\omega^0 \in C^k,\ \tau^{2k}(\omega^0, \omega^{1k}) < h),\end{aligned}$$
where
$$\tau^{2k} := \inf\{t \ge 0 \mid |\hat{X}_t^{2k} - x_k| \ge \rho\}.$$
Note that
$$\sum_{k\in I} \hat{P}^2(\omega^0,\ \tau_k^*(\omega^{1k}) < h) = \sum_{k \in I} \hat{P}^1(C^k) P^k(\tau_k^k < h) \le C_\rho h^2 \sum_{k\in I} \hat{P}^1(C^k) = C_\rho h^2;$$
$$\begin{aligned}\sum_{k\in I} \hat{P}^2(\omega^0 \in C^k,\ \tau^{2k}(\omega^0, \omega^{2k}) < h) &= \sum_{k\in I} \hat{P}^1 \otimes P^k(\omega^0 \in C^k,\ \tau^{2k}(\omega^0,\omega^{2k}) < h) \\ &= \sum_{k\in I} \hat{P}^1(C^k) P^{2k}(\tau^{2k}(\omega^0, \omega^{2k}) < h).\end{aligned}$$
Note also that since
$$\hat{X}_0^{2k} \in B_k \qquad \hat{P}^{2k}\text{–a.s.},$$
we have
$$\tilde{E}^k \sup_{0 \le t \le h} |\hat{X}_t^{2k} - \hat{X}_0^{2k}| \le Ch^2.$$
Hence, as before we have
$$\hat{P}^{2k}(\tau^{2k} < h) \le C_\rho h^2.$$
Therefore,
$$\sum_{k\in I} \hat{P}^2(\omega^0 \in C^k,\ \tau^{2k}(\omega^0, \omega^{1k}) < h) \le C_\rho h \sum_{k\in I} \hat{P}^1(C^k) = C_\rho h^2,$$
and so
$$\hat{P}^2(\hat{\tau}^2 < 2h) \le 2C_\rho h^2 + 2C_\rho h^2 = 4C_\rho h^2.$$
Inductively we find a controlled system
$$(\hat{\Omega}^m, \hat{\mathcal{F}}^m, \hat{P}^m, \{\hat{\mathcal{F}}_t^m\}_{t\ge 0}, \{\hat{W}_t^m\}_{t\ge0}, \{\hat{u}_t^m\}_{t\ge0}, \{\hat{X}_t^m\}_{t\ge0}) \in \hat{\mathcal{C}}(x)$$
and a stopping time $\hat{\tau}^m$ such that
$$\overline{V}(x) + mh^2 > \hat{E}^m\Big(\int_0^{mh} e^{-\lambda t} \hat{f}(\hat{X}_t^m, \hat{u}_t^m) dt + e^{-m\lambda h} \overline{V}(\hat{X}_{mh}^m)\Big);$$
$$(\hat{b}, \hat{\sigma}, \hat{f})(\hat{X}_t^m, \hat{u}_t^m) = (b_0, \sigma_0, f_0)(\hat{X}_t^m, \hat{u}_t^m) \quad \forall t \in [0,\ \hat{\tau}^m] \qquad \hat{P}\text{–a.s.};$$
$$\hat{P}^m(\hat{\tau}^m < mh) \le 2mC_\rho h^2.$$

In view of Lemma 5.1, we find stochastic processes $\{\mu_t\}_{t\geq 0}$ and $\{X_t\}_{t\geq 0}$ so that

$$\mu_t(\omega) = \hat{u}_t(\omega) \qquad \forall t \in [0, \hat{\tau}^m(\omega));$$
$$(\hat{\Omega}^m, \hat{\mathcal{F}}^m, \hat{P}^m, \{\hat{\mathcal{F}}_t^m\}_{t\geq 0}, \{\hat{W}_t^m\}_{t\geq 0}, \{\mu_t\}_{t\geq 0}, \{X_t\}_{t\geq 0}) \in \mathcal{C}_0(x),$$

where $\mathcal{C}_0(x)$ denotes the set of controlled systems at x associated with $K,\ b_0,\ \sigma_0$;

$$X_t(\omega) \in \overline{G} \quad \forall t \geq 0 \qquad \hat{P}^m\text{–a.s.}$$

Of course, we have

$$X_t(\omega) = \hat{X}_t^m(\omega) \quad \forall t \leq \hat{\tau}^m(\omega) \qquad \hat{P}^m\text{–a.s.}$$

Setting

$$\Omega_1 = \{\omega \in \hat{\Omega}^m \mid \hat{\tau}^m(\omega) < mh\}, \Omega_2 = \{\omega \in \hat{\Omega}^m \mid \hat{\tau}^m(\omega) \geq mh\},$$

we compute that

$$\begin{aligned}
&\overline{V}(x) + (T+1)h > \hat{E}^m \mathbf{1}_{\Omega_1} \int_0^{mh} e^{-\lambda t} \hat{f}(\hat{X}_t^m, \hat{u}_t^m)dt \\
&+ \hat{E}^m \mathbf{1}_{\Omega_2} \int_0^{mh} e^{-\lambda t} f_0(X_t, \mu_t)dt - e^{-\lambda T}\|\overline{V}\|_\infty \\
&\geq -2M(T+1)\hat{P}^m(\Omega_1) + \hat{E}^m \int_0^{mh} e^{-\lambda t} f_0(X_t, \mu_t)dt - \varepsilon \\
&\geq \hat{E}^m \int_0^{\infty} f_0(X_t, \mu_t)dt - 2\varepsilon - 2M(T+1)\hat{P}^m(\Omega_1) \\
&\geq \hat{E}^m \int_0^{\infty} f_0(X_t, \mu_t)dt - 2\varepsilon - 4MC_\rho(T+1)^2 h.
\end{aligned}$$

Choosing $h > 0$ small enough, we get

$$\hat{E}^m \int_0^\infty e^{-\lambda t} f_0(X_t, \mu_t)dt < \overline{V}(x) + 3\varepsilon.$$

In conclusion, there is

$$(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\geq 0}, \{W_t\}_{t\geq 0}, \{\mu_t\}_{t\geq 0}, \{X_t\}_{t\geq 0}) \in \mathcal{C}_0(x)$$

such that

$$X_t \in \overline{G} \quad \forall t \geq 0 \qquad P\text{–a.s.};$$
$$\overline{V}(x) + 3\varepsilon > E \int_0^\infty e^{-\lambda t} f_0(X_t, \mu_t)dt.$$

We set

$$u_t = \begin{cases} a_i & \text{if} \quad \mu_t = (i,j) \text{ and } j = 0, \\ \hat{a}_j(X_t) & \text{if} \quad \mu_t = (i,j) \text{ and } j \geq 1, \end{cases}$$

and observe that

$$(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\geq 0}, \{W_t\}_{t\geq 0}, \{u_t\}_{t\geq 0}, \{X_t\}_{t\geq 0}) \in \mathcal{C}(x);$$
$$E\int_0^\infty e^{-\lambda t} f_0(X_t, \mu_t)dt = E\int_0^\infty e^{-\lambda t} f(X_t, u_t)dt.$$

If we set $E = \{a_i \mid i \in I(p)\} \cup \{\hat{a}_j \mid j \in I_0(q)\} \in \mathcal{E}$, then $V_{\mathcal{E}}(x) \leq V_E < U(x) + 4\varepsilon$. QED

6. Proof of Theorem 3.7

We prove (ii). Assume (A1)–(A3) and (A5)–(A6).

Define

$$V_n(x) = \inf_{\alpha\in\mathcal{C}(x)} E^\alpha \int_0^\infty e^{-\lambda t}[f(X_t^\alpha, u_t^\alpha) + nd(X_t^\alpha)]dt \qquad \forall x \in \mathbf{R}^N, \tag{36}$$

for all $n \in \mathbf{N}$, where $d(x) = \operatorname{dist}(x, G) \wedge 1$.

By a result of [NA, NI1, NI2], we see that V_n is a viscosity solution of

$$\lambda u(x) + H(x, Du(x), D^2u(x)) = nd(x) \quad \text{in } \mathbf{R}^N. \tag{37}$$

As in the proof of Theorem 3.5, we see that

$$V_+(x) := \sup_{n\in\mathbf{N}} V_n(x) = \lim_{n\to\infty} V_n(x) = U(x) \qquad \forall x \in \overline{G}. \tag{38}$$

Fix $x \in \overline{G}$ and let

$$\alpha_n \equiv (\Omega^n, \mathcal{F}^n, P^n, \{\mathcal{F}_t^n\}_{t\geq 0}, \{W_t^n\}_{t\geq 0}, \{u_t^n\}_{t\geq 0}, \{X_t^n\}_{t\geq 0})$$

be a controlled system at x such that

$$V_n(x) + n^{-1} > E^n \int_0^\infty e^{-\lambda t}[f(X_t^n, u_t^n) + nd(X_t^n)]dt,$$

where E^n denotes the mathematical expectation with respect to P^n.

Following for instance the argument of [YZ, Theorem 5.3], we can find a controlled system $\alpha \in \mathcal{C}(x)$ for which we have

$$\liminf_{n\to\infty} J(x, \alpha_n) = J(x, \alpha),$$

$$\liminf_{n\to\infty} E^n \int_0^\infty e^{-\lambda t} d(X_t^n)dt = E^\alpha \int_0^\infty e^{-\lambda t} d(X_t^\alpha)dt.$$

Note that we needed the convexity assumption (A6) in the above assertion. We see from the latter identity that

$$E^\alpha \int_0^\infty e^{-\lambda t} d(X_t^\alpha)dt = 0,$$

which ensures that $\alpha \in \mathcal{A}(x)$. On the other hand, the former guarantees that

$$J(x,\alpha) \le \lim_{n\to\infty} V_n(x) = V_+(x).$$

Since $V_+(x) = U(x) = V(x)$, we have $J(x,\alpha) \le V(x)$ and therefore,

$$J(x,\alpha) = V(x).$$

The proof of (i) is more or less the same as (ii). We use the observations in [EHJ, KU] instead of those in [NA, NI1, NI2]. QED

7. Proof of Theorem 3.8

In what follows the symbols $\mathcal{C}_\varepsilon(x)$, $\mathcal{A}_\varepsilon(x)$, and U_ε denote those defined in Section 3. Let $\hat{a}$ be the Lipschitz function from (A7). Let $\varepsilon > 0$ and $x \in \overline{G}$. If

$$\alpha \equiv (\Omega^\alpha, \mathcal{F}^\alpha, P^\alpha, \{\mathcal{F}_t^\alpha\}_{t\ge0}, \{W_t^\alpha\}_{t\ge0}, \{u_t^\alpha\}_{t\ge0}, \{X_t^\alpha\}_{t\ge0}) \in \mathcal{A}_\varepsilon(x)$$

and if we set

$$v_t^\alpha := \chi_\varepsilon(X_t^\alpha)u_t^\alpha + (1-\chi_\varepsilon(X_t^\alpha))\hat{a}(X_t^\alpha),$$

then

$$(\Omega^\alpha, \mathcal{F}^\alpha, P^\alpha, \{\mathcal{F}_t^\alpha\}_{t\ge0}, \{W_t^\alpha\}_{t\ge0}, \{v_t^\alpha\}_{t\ge0}, \{X_t^\alpha\}_{t\ge0}) \in \mathcal{A}(x).$$

Hence,

$$U_\varepsilon(x) \ge V(x) \qquad \forall x \in \overline{G}.$$

By Theorem 3.5, we have $U(x) \le V(x)$ for all $x \in \overline{G}$. Hence we have

$$V(x) \le U_\varepsilon(x) \qquad \forall x \in \overline{G}. \tag{39}$$

Next, arguments similar to the proof of Lemma 5.2 yield that as $\varepsilon \searrow 0$, $U_\varepsilon(x)$ converge to the unique viscosity solution of

$$\begin{cases} \lambda u(x) + H(x, Du(x), D^2u(x)) \ge 0 & \text{in } \overline{G}, \\ \lambda u(x) + H_*(x, Du(x), D^2u(x)) \le 0 & \text{in } \overline{G}, \end{cases} \tag{40}$$

where

$$H_*(x,p,X) = \begin{cases} H(x,p,X) & \text{if } x \in G, \\ -\frac{1}{2}\operatorname{tr}\sigma\sigma^T(x,\hat{a}(x))X - b(x,\hat{a}(x))\cdot p - f(x,\hat{a}(x)) & \text{if } x \in \partial G. \end{cases}$$

Noting that U is the unique viscosity solution of this problem, we conclude the proof. QED

Remark. The same arguments as the above proof but with minor changes yield that the assertion of Theorem 3.8, with (A7) replaced by the following (A9) holds.

(A9) There are a continuous functions $\phi : \mathbf{R}^N \times A \times [0,1) \to A$ and $\phi_0 : \mathbf{R}^N \to A$ such that for any $\lambda \in [0,1)$ the functions: $(x,a) \mapsto b(x,\phi(x,a))$, $(x,a) \mapsto \sigma(x,\phi(x,a))$, $(x,a) \mapsto f(x,\phi(x,a))$ on $\mathbf{R}^N \times A$ satisfy (A1) and (A2), such that for any $\lambda \in [0,1)$,

$$\phi(x,a,\lambda) = a \qquad \forall (x,a) \in (G \setminus (\partial G)_{2\lambda}) \times A,$$
$$\phi(x,a,\lambda) = \phi_0(x) \qquad \forall (x,a) \in (\overline{G} \cap (\partial G)_\lambda) \times A,$$

and such that for some $r > 0$,

$$\sigma(x,\phi_0(x)) = 0 \qquad \forall x \in \partial G,$$
$$TC(x, b(x,\phi_0(x)), r) \subset \overline{G} \qquad \forall x \in \overline{G} \cap (\partial G)_r.$$

Here we used the notation: $E_r = \{x \in \mathbf{R}^N \mid \operatorname{dist}(x,E) < r\}$ for $E \subset \mathbf{R}^N$ and $r > 0$.

8. Proof of Theorem 3.9

Let $\varepsilon > 0$. We utilize the functions U_ε, σ_ε, etc defined in the formulation of Theorem 3.8.

We choose a function $d \in C^2(\mathbf{R}^N)$ so that

$$\begin{cases} d(x) > 0 & \text{in } G, \\ d(x) \le 0 & \text{in } G^c, \\ d(x) = \operatorname{dist}(x,G^c) - \operatorname{dist}(x,G) & \text{in a neighborhood of } \partial G. \end{cases}$$

For each $n \in \mathbf{N}$ we choose a function $\zeta_n \in C^2(\mathbf{R})$ so that

$$\zeta_n'(r) \ge 0, \qquad \zeta_n''(r) \le 0,$$
$$\zeta_n(r) = \begin{cases} r & (r \le n-1), \\ n & (r \ge n+1). \end{cases}$$

We may assume that the sequence of $\{\zeta_n\}$ satisfies

$$\sup_{n\in\mathbf{N}} (\|\zeta_n'\|_\infty + \|\zeta_n''\|_\infty) < \infty.$$

Then define the function $\psi_n \in C^2(\mathbf{R})$ by

$$\psi_n(r) = \begin{cases} \zeta_n(-\log r) & (r > 0), \\ n & (r \le 0). \end{cases}$$

Fix $\gamma > 0$ and choose a function $\rho \in C^2(\mathbf{R})$ so that

$$\rho(t) = 0 \quad \forall t \ge \gamma, \qquad 0 < \rho(t) \le 1 \quad \forall t < \gamma, \qquad \|\rho'\|_\infty \le 1.$$

For each $n \in \mathbf{N}$, $\gamma > 0$, and $\delta \in (0, 1)$, we set

$$g(x, t) \equiv g_{n,\gamma,\delta}(x, t) := \psi_n(d(x) + \delta\rho(t)).$$

Clearly, $g \in C^2(\mathbf{R}^{N+1})$.

In a neighborhood $\mathcal{N}$ of ∂G, we have

$$\frac{1}{2} \operatorname{tr} \sigma\sigma^T(x, \hat{a}(x))D^2 d(x) + b(x, \hat{a}(x)) \cdot Dd(x) \ge \beta$$

for some constant $\beta > 0$.

Note that for $r > 0$,

$$\begin{aligned} \psi_n'(r) =& - \zeta_n'(-\log r)\frac{1}{r} \le 0, \\ \psi_n''(r) =& \zeta_n''(-\log r)\left(\frac{1}{r}\right)^2 + \zeta_n'(-\log r)\frac{1}{r^2} \\ \le& \zeta_n'(-\log r)\frac{1}{r^2} = -\psi_n'(r)\frac{1}{r}. \end{aligned}$$

We may assume that $d(x) = -1$ for $x \in \mathbf{R}^N$ with large $|x|$. Choose $\delta > 0$ and $\mu > 0$ so small that if $x \in \mathbf{R}^N$ satisfies

$$-\delta < d(x) < \mu, \quad \text{then } x \in \mathcal{N}.$$

We may assume as well that $\mathcal{N} \subset \mathbf{R}^N \setminus G_{\varepsilon/2}$.

Let $(x, t) \in \mathbf{R}^{N+1}$. We now divide our considerations into three cases:

Case 1. Consider the case where $d(x) + \delta\rho(t) < e^{-(n+1)}$, i.e., the case where we have either

$$d(x) + \delta\rho(t) \le 0 \quad \text{or} \quad -\log(d(x) + \delta\rho(t)) > n + 1.$$

Then we have $g(y, s) = n$ near the point (x, t) and hence

$$g_t(x, t) + \frac{1}{2} \operatorname{tr} \sigma_\varepsilon\sigma_\varepsilon^T(x, a)D^2 g(x, t) + b_\varepsilon(x, t) \cdot Dg(x, t) = 0.$$

Case 2. Consider the case where $0 < d(x) + \delta\rho(t) < \mu$. We have $-\delta < d(x) < \mu$. We compute that

$$
\begin{aligned}
& g_t(x,t) + \frac{1}{2}\operatorname{tr}\sigma_\varepsilon\sigma_\varepsilon^T(x,a)D^2 g(x,t) + b_\varepsilon(x,a)\cdot Dg(x,t) \\
& = \psi_n'(d(x)+\delta\rho(t))\left[\delta\rho'(t) + \frac{1}{2}\operatorname{tr}\sigma\sigma^T(x,\hat{a}(x))D^2 d(x) + b(x,\hat{a}(x))\cdot Dd(x)\right] \\
& + \frac{1}{2}\psi_n''(d(x)+\delta\rho(t))\operatorname{tr}\sigma\sigma^T(x,\hat{a}(x))Dd(x)\otimes Dd(x) \\
& \le \psi_n'(d(x)+\delta\rho(t))(\beta-\delta) \\
& - \frac{1}{2d(x)}\psi_n'(d(x)+\delta\rho(t))\operatorname{tr}\sigma\sigma^T(x,\hat{a}(x))Dd(x)\otimes Dd(x) \\
& \le \psi_n'(d(x)+\delta\rho(t))\left(\beta-\delta-\frac{1}{2d(x)}L^2 d(x)^2\|Dd\|_\infty^2\right),
\end{aligned}
$$

where L is the Lipschitz constant of the function $x \mapsto \sigma(x,\hat{a}(x))$. We may assume by replacing δ and $\mathcal{N}$ by smaller ones if necessary that

$$\delta + \frac{1}{2}L^2 d(x)\|Dd\|_\infty^2 \le \beta.$$

We thus have

$$g_t(x,t) + \frac{1}{2}\operatorname{tr}\sigma_\varepsilon\sigma_\varepsilon^T(x,a)D^2 g(x,t) + b_\varepsilon(x,a)\cdot Dg(x,t) \le 0.$$

Case 3. Now consider the case where $d(x) + \delta\rho(t) > \mu/2$. In what follows we assume that n is so large that $-\log(\mu/2) \le n-1$. Hence,

$$-\log(d(x)+\delta\rho(t)) \le n-1$$

and therefore $g(x,t) = -\log(d(x)+\delta\rho(t))$. Thus there is a constant $C > 0$ independent of n such that

$$g_t(x,t) + \frac{1}{2}\operatorname{tr}\sigma_\varepsilon\sigma_\varepsilon^T(x,a)D^2 g(x,t) + b_\varepsilon(x,a)\cdot Dg(x,t) \le C.$$

This way we conclude that for any $(x,t) \in \mathbf{R}^{N+1}$, we have

$$g_t(x,t) + \frac{1}{2}\operatorname{tr}\sigma_\varepsilon\sigma_\varepsilon^T(x,a)D^2 g(x,t) + b_\varepsilon(x,a)\cdot Dg(x,t) \le C.$$

Let

$$\alpha \equiv (\Omega^\alpha, \mathcal{F}^\alpha, P^\alpha, \{\mathcal{F}_t^\alpha\}_{t\ge 0}, \{W_t^\alpha\}_{t\ge 0}, \{u_t^\alpha\}_{t\ge 0}, \{X_t^\alpha\}_{t\ge 0})$$

be an admissible controlled system at z associated with σ_ε and b_ε. I.e., $\alpha \in \mathcal{A}_\varepsilon(z)$. We apply the Itô formula, to obtain

$$\begin{aligned} g(X_t^\alpha, t) =& g(z,0) \\ &+ \int_0^t \Big(g_s(X_s^\alpha, s) + \frac{1}{2} \operatorname{tr} \sigma_\varepsilon \sigma_\varepsilon^T (X_s^\alpha, u_s^\alpha) D^2 g(X_s^\alpha, s) \\ &+ b_\varepsilon(X_s, u_s^\alpha) \cdot Dg(X_s^\alpha, s) \Big) ds \\ &+ \int_0^t Dg(X_s^\alpha, s) \cdot \sigma_\varepsilon(X_s^\alpha, u_s^\alpha) dW_s^\alpha. \end{aligned}$$

From this, we get

$$\begin{aligned} g(X_{\tau\wedge t}^\alpha, \tau \wedge t) =& g(z,0) \\ &+ \int_0^{\tau\wedge t} \Big(g_s(X_s^\alpha, s) + \frac{1}{2} \operatorname{tr} \sigma_\varepsilon \sigma_\varepsilon^T (X_s^\alpha, u_s^\alpha) D^2 g(X_s^\alpha, s) \\ &\qquad + b_\varepsilon(X_s^\alpha, u_s^\alpha) \cdot Dg(X_s^\alpha, s) \Big) ds \\ &+ \int_0^{\tau\wedge t} Dg(X_s^\alpha, s) \cdot \sigma_\varepsilon(X_s^\alpha, u_s^\alpha) dW_s^\alpha, \end{aligned}$$

where τ is the first hitting time of X_t^α after time γ to the closed set, ∂G, i.e.,

$$\tau := \inf\{t \geq \gamma \mid X_t^\alpha \in \partial G\}.$$

Hence, if n is large enough, we have

$$E^\alpha g(X_{\tau\wedge t}, \tau \wedge t) \leq g(z,0) + Ct \qquad \forall t > 0.$$

If n is large enough, then

$$g(z,0) = -\log(d(z) + \delta\rho(0)) \leq -\log(\delta\rho(0)).$$

We may assume that for each $r > 0$,

$$\psi_n(r) \nearrow -\log r \qquad \text{as } n \to \infty.$$

Now the monotone convergence theorem implies that

$$E^\alpha[-\log(d(X_{\tau\wedge t}^\alpha) + \delta\rho(\tau \wedge t))] \leq -\log(\delta\rho(0)) + Ct \qquad \forall t > 0.$$

This implies that $\tau = \infty$ P^α–a.s. and by the arbitrariness of $\gamma > 0$ that

$$X_t^\alpha \in G \qquad \forall t > 0 \quad P^\alpha\text{–a.s.}$$

If we define

$$v_t^\alpha := \chi_\varepsilon(X_t^\alpha)u_t^\alpha + (1-\chi_\varepsilon(X_t^\alpha))\hat{a}(X_t^\alpha),$$

then

$$(\Omega^\alpha, \mathcal{F}^\alpha, P^\alpha, \{\mathcal{F}_t^\alpha\}_{t\ge 0}, \{W_t^\alpha\}_{t\ge 0}, \{v_t^\alpha\}_{t\ge 0}, \{X_t^\alpha\}_{t\ge 0}) \in \mathcal{A}(z).$$

Therefore,

$$U_\varepsilon(x) \ge V_0(x) \ge V(x) \qquad \forall x \in \overline{G}.$$

Thus in the limit as $\varepsilon \searrow 0$, using Theorem 3.8, we get:

$$U(x) \ge V_0(x) \ge V(x) \qquad \forall x \in \overline{G}. \qquad \text{QED}$$

Remark. As the proof above shows, the assumption (A7) of Theorem 3.9 can be replaced by (A9). Hence the assertion (ii) of Theorem 2, with (A9) in place of (A9), is valid.

Appendix

Proof of Lemma 3.3

For $\delta > 0$ we introduce the set

$$G_\delta := \{x \in \mathbf{R}^N \mid \operatorname{dist}(x, G) < \delta\}.$$

Our strategy for constructing a function ψ with the required properties is to define ψ as a mollification of the distance function

$$d(x) := \operatorname{dist}(x, \mathbf{R}^N \setminus G_\delta)$$

with $\delta > 0$ small enough.

Fix $r > 0$ so that

$$(A.1) \qquad B(x + t\xi_0(x), rt) \subset \overline{G} \qquad (z \in \partial G,\ x \in B(z,r) \cap \overline{G},\ 0 \le t \le r).$$

Let $0 < r_0 \le r/3$. Let $0 < \delta < r_0$, $z \in \partial G_\delta$, and $x \in B(z, r_0) \cap \overline{G_\delta}$. Then there are points $\hat{z} \in \partial G$ and $\hat{x} \in \overline{G}$ such that $|\hat{z} - z| = \delta$ and $|\hat{x} - x| \le \delta$. Since

$$|\hat{z} - \hat{x}| \le 2\delta + |z - x| \le 3r_0 \le r,$$

from (A.1) we get

$$B(\hat{x} + t\xi_0(\hat{x}), rt) \subset \overline{G} \qquad (0 \le t \le r).$$

Hence,

$$B(\hat{x} + t\xi_0(\hat{x}), rt) + B(0, \delta) \subset \overline{G_\delta} \qquad (0 \le t \le r),$$

and so,

$$(A.2) \qquad B(x + t\xi_0(\hat{x}), rt) \subset \overline{G_\delta} \qquad (0 \le t \le r).$$

We may assume that $r \le 1$ and furthermore by choosing $r_0 > 0$ small enough that if $y, \eta \in \overline{G}_1$ and $|y - \eta| \le r_0$, then $|\xi_0(y) - \xi_0(\eta)| \le r/2$. Now, from (A.2) we see that

$$B(x + t\xi_0(x), (r/2)t) \subset \overline{G_\delta} \qquad (0 \le t \le r).$$

From this we conclude that for any $0 < \delta < r_0$,

$$B(x + \xi_0(x), r_0 t) \subset \overline{G}_\delta \qquad (z \in \partial G_\delta,\ x \in B(z, r_0) \cap \overline{G_\delta},\ 0 \le t \le r_0).$$

Next, for $\delta > 0$ we define

$$e_\delta := \sup_{x \in \partial G} \operatorname{dist}(x, \partial(G_\delta)).$$

Note that $\delta \le e_\delta < \infty$ and that $e_\delta \to 0$ as $\delta \to 0$. Fix $0 < \delta < r_0$ so that if $y, \eta \in \overline{G_1}$ and $|y - \eta| \le e_\delta + \delta$, then $|\xi_0(y) - \xi_0(\eta)| \le r_0/2$.

Since the function $d(x) := \operatorname{dist}(x, \mathbf{R}^N \setminus G_\delta)$ is Lipschitz continuous on $\mathbf{R}^N$, it is differentiable a.e. in $\mathbf{R}^N$. Let $x \in (\partial G)_\delta$ be a point where d is differentiable, which means that there is a unique point $\hat{x} \in \partial G_\delta$ for which

$$d(x) = |x - \hat{x}|.$$

Moreover we have

$$n := Dd(x) = \frac{x - \hat{x}}{|x - \hat{x}|}.$$

We want to prove that

$$\xi_0(x) \cdot Dd(x) \ge r_0/2.$$

For this we first note that

$$B(x, d(x)) \subset \overline{G_\delta},$$
$$B(\hat{x} + t\xi_0(\hat{x}), r_0 t) \subset \overline{G_\delta} \qquad (0 \le t \le r_0).$$

Then we define the half space

$$S := \{z \in \mathbf{R}^N \mid z \cdot n \ge 0\},$$

and claim that

$$(A.3) \qquad B(\xi_0(\hat{x}), r_0) \subset S.$$

To prove this claim, we argue by contradiction. Thus we assume that

$$B(\xi_0(\hat{x}), r_0) \setminus S \neq \emptyset.$$

This implies that

$$\operatorname{Int} B(\xi_0(\hat{x}), r_0) \setminus S \neq \emptyset.$$

Fix $p \in \operatorname{Int} B(\xi_0(\hat{x}), r_0) \setminus S$. Clearly we have

$$p \cdot n < 0.$$

We may assume that $|p| = 1$. Let $\varepsilon > 0$ be a small constant to be specified later, and set

$$x_\varepsilon := \hat{x} - \varepsilon p.$$

Note that for some $\varepsilon_0 > 0$, if $0 < \varepsilon \leq \varepsilon_0$, then

$$\begin{aligned} |x - x_\varepsilon|^2 =& |x - x_\varepsilon + \varepsilon p|^2 = |d(x)n + \varepsilon p|^2 \\ =& d(x)^2 + \varepsilon^2 |p|^2 + 2\varepsilon d(x) p \cdot n \leq d(x)^2, \end{aligned}$$

and hence, $x_\varepsilon \in \overline{G_\delta}$. Furthermore, if $0 < \varepsilon \leq \min\{\varepsilon_0, r_0\}$, then we have $|x_\varepsilon - \hat{x}| \leq r_0$ and therefore,

$$\begin{aligned} \overline{G_\varepsilon} \supset& B(x_\varepsilon + \varepsilon \xi_0(\hat{x}), \varepsilon r_1) = x_\varepsilon + \varepsilon B(\xi_0(\hat{x}), r_1) \\ \ni& \hat{x} - \varepsilon p + \varepsilon(p + B(0, \rho)) = \hat{x} + B(0, \varepsilon \rho) \end{aligned}$$

for some constant $\rho > 0$. This is a contradiction, which shows that (A.3) holds.

Now, from (A.3) we see that

$$(\xi_0(\hat{x}) - r_0 n) \cdot n \geq 0,$$

and hence,

$$\xi_0(\hat{x}) \cdot n \geq r_0.$$

Since $|x - \hat{x}| \leq e_\delta + \delta$, we have $|\xi_0(x) - \xi_0(x)| \leq r_0/2$ and therefore,

$$\xi_0(x) \cdot Dd(x) \geq r_0/2,$$

proving that

$$\xi_0(x) \cdot Dd(x) \geq r_0/2 \quad \text{a.e. in } (\partial G)_\delta.$$

Finally choose a function $\phi \in C_0^\infty(\mathbf{R}^N)$ so that $\operatorname{spt} \phi \subset B(0,1)$, $\phi \geq 0$, and $\int_{\mathbf{R}^N} \phi(x) dx = 1$. For $\varepsilon > 0$ define $\psi_\varepsilon \in C^\infty(\mathbf{R}^N)$ by $\psi_\varepsilon = \phi_\varepsilon * d$ with $\phi_\varepsilon(x) = \varepsilon^{-N} \phi(x/\varepsilon)$. If we choose $\varepsilon > 0$ small enough, then ψ_ε has the required properties. QED

Sketch of proof of Lemma 3.4

We give here comments on the proof of Lemma 3.4. As already noted, the same construction of w in the proof of [IK, Lemma 3.4] yields a function having the desired properties in our lemma. Unfortunately, the assertion of [IK, Lemma 3.4] is apparently different and weaker than our Lemma 3.4. The differences are that the function w in [IK, Lemma 3.4] has only the C^1 regularity as it stated and that the last inequality

$$D^2w(x,y) \le C\begin{pmatrix} I & -I \\ -I & I \end{pmatrix} + C|x-y|^2\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \tag{A.4}$$

is not in [IK, Lemma 3.4].

We explain how to show these properties of w, the function constructed in the proof of [IK, Lemma 3.4].

The function w constructed in the proof of [IK, Lemma 3.4] has the form:

$$w(x,y) = v((x-y)\cdot\xi(y), |x-y-((x-y)\cdot\xi(y))\xi(y)|^2)^2,$$

where $\xi : \mathbf{R}^N \to \mathbf{R}^N$ is a C^∞ function and $v : \mathbf{R}^2 \to \mathbf{R}$ is a function in $C(\mathbf{R}^2) \cap C^{1,1}(\mathbf{R}^2 \setminus \{0\})$ satisfying the homogeneity property

$$v(tx) = tv(x) \qquad \forall x \in \mathbf{R}^2,\ t \ge 0,$$

and the property

$$D_{x_2}v(x_1, 0) = 0 \qquad \forall x_1 \in \mathbf{R}.$$

To see the $C^{1,1}$ regularity of w, define the functions $f : \mathbf{R}^2 \to \mathbf{R}$ and $g : \mathbf{R}^{1+N} \to \mathbf{R}$, respectively, by

$$f(x) = v(x)^2 \qquad \forall x \in \mathbf{R}^2,$$
$$g(x,y) = f(x,|y|) \qquad \forall (x,y) \in \mathbf{R}\times\mathbf{R}^N,$$

and observe that $f \in C^{1,1}(\mathbf{R}^2)$ and that for all $x, y \in \mathbf{R}^N$,

$$w(x,y) = g((x-y)\cdot\xi(y), Q(\xi(y))(x-y)), \tag{A.5}$$

where

$$Q(\xi) = I - \xi\otimes\xi.$$

For the moment we assume that $f \in C^2(\mathbf{R}^2)$ and calculate that for all $x \in \mathbf{R}$ and $y \in \mathbf{R}^N \setminus \{0\}$,

$$D_x g(x,y) = D_{x_1} f(x,|y|), \qquad D_y g(x,y) = D_{x_2} f(x,|y|)\frac{y}{|y|},$$
$$D_x^2 g(x,y) = D_{x_1}^2 f(x,|y|), \qquad D_y D_x g(x,y) = D_{x_2}D_{x_1} f(x,|y|)\frac{y}{|y|},$$
$$D_y^2 g(x,y) = D_{x_2} f(x,|y|)\frac{y\otimes y}{|y|^2} + D_{x_2} f(x,|y|)\frac{I}{|y|} - D_{x_2} f(x,|y|)\frac{y\otimes y}{|y|^3}.$$

If we fix $R > 0$, then, since $D_{x_2} v(x_1, 0) = 0$,

$$|D_{x_2} f(x, |y|)| \le C_R |y| \quad \text{if } |x| + |y| \le R,$$

for some constant $C_R > 0$, and hence, from the calculations above we have

$$\sup_{|x|+|y|\le R} |D^2 g(x,y)| < \infty.$$

A simple approximation argument and the above calculations show that $g \in C^{1,1}(\mathbf{R}^{1+N})$. Now (A.5) shows that $w \in C^{1,1}(\mathbf{R}^{2N})$.

Next we set

$$h(\xi, p) = g(\xi \cdot p, Q(\xi)p) \qquad (\xi, p \in \mathbf{R}^N).$$

With an approximation argument in mind we may assume that $h \in C^2(\mathbf{R}^{2N})$. Note that

$$w(x,y) = h(\xi(y), x - y) \qquad (x, y \in \mathbf{R}^N),$$

and that

$$h(\xi, tp) = t^2 h(\xi, p) \qquad (\xi, p \in \mathbf{R}^N,\ t \ge 0).$$

By this homogeneity, for each bounded subset B of $\mathbf{R}^N$ there is a constant $C_B > 0$ such that for any $\xi \in B$ and $p \in \mathbf{R}^N$,

$$\begin{gathered} \max\{h(\xi,p), |D_\xi h(\xi,p)|, |D_\xi^2 h(\xi,p)|\} \le C_B |p|^2, \\ \max\{|D_p h(\xi,p)|, |D_p D_\xi h(\xi,p)| \le C_B |p|, \\ |D_p^2 h(\xi,p)| \le C_B. \end{gathered}$$

Compute that

$$\begin{aligned} D_x^2 w(x,y) &= D_p^2 h(\xi(y), x-y), \\ D_y D_x w(x,y) &= (D\xi(y))^T D_\xi D_p h(\xi(y), x-y) - D_p^2 h(\xi(y), x-y), \end{aligned}$$

$$\begin{aligned} D_y^2 w(x,y) =& (D\xi(y))^T D_\xi^2 h(\xi(y), x-y) D\xi(y) + D_p^2 h(\xi(y), x-y) \\ &- (D_\xi D_p h(\xi(y), x-y) D\xi(y) - (D\xi(y))^T D_p D_\xi h(\xi(y), x-y), \end{aligned}$$

to obtain

$$\text{(A.6)} \qquad D^2 w(x,y) = \begin{pmatrix} A_1 & -A_1 \\ -A_1 & A_1 \end{pmatrix} + \begin{pmatrix} 0 & A_2 \\ A_2^T & -A_2 - A_2^T \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & A_3 \end{pmatrix},$$

where

$$\begin{gathered} A_1 = D_p^2 h(\xi(y), x-y), \qquad A_2 = (D_\xi D_p h(\xi(y), x-y)) D\xi(y), \\ A_3 = (D\xi(y))^T D_\xi^2 h(\xi(y), x-y) D\xi(y). \end{gathered}$$

For instance, for each bounded $B \subset \mathbf{R}^N$ and for all $x, y, p, q \in \mathbf{R}^N$, if $y \in B$, then we have

$$\begin{aligned}\begin{pmatrix} 0 & A_2 \\ A_2^T & -A_2 - A_2^T \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix} \cdot \begin{pmatrix} p \\ q \end{pmatrix} =& 2A_2 q \cdot (p-q) \le 2C_B|x-y||q||p-q| \\ \le& C_B(|p-q|^2 + |x-y|^2|q|^2)\end{aligned}$$

for some constant $C_B > 0$, i.e.,

$$\begin{pmatrix} 0 & A_2 \\ A_2^T & -A_2 - A_2^T \end{pmatrix} \le C_B\Big(\begin{pmatrix} I & -I \\ -I & I \end{pmatrix} + |x-y|^2 \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}\Big).$$

This way we see from (A.6) that for each bounded $B \subset \mathbf{R}^N$, there is a constant $C_B > 0$ for which

$$D^2 w(x,y) \le C_B\Big(\begin{pmatrix} I & -I \\ -I & I \end{pmatrix} + |x-y|^2 \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}\Big) \qquad \forall x \in \mathbf{R}^N,\ y \in B. \qquad \text{QED}$$

Proof of Theorem 3.1

We argue by contradiction, and hence assume that

$$\max_{x \in \overline{G}} (u(x) - v(x)) \ge \theta$$

for some constant $\theta > 0$.

If necessary, by replacing u by

$$\tilde{u}(x) := u(x) + (2M+1)\psi(x),$$

where ψ is the function from Lemma 3.3, we may assume that u satisfies

$$-\xi_0(x) \cdot Du(x) \le -1 \qquad \forall x \in \partial G$$

in the viscosity sense.

Let $w \in C^{1,1}(\overline{G} \times \overline{G})$, $r > 0$, and $C > 0$ be from Lemma 3.4.

Let $L > 0$ and consider the function

$$\Phi(x,y) \equiv u(x) - v(y) - Lw(x,y)$$

on the set $\overline{G} \times \overline{G}$. We select $\hat{x}, \hat{y} \in \overline{G}$ so that $|\hat{x} - \hat{y}| \le r$ and

$$\Phi(\hat{x}, \hat{y}) = \sup\{\Phi(x,y) \mid x, y \in \overline{G},\ |x-y| \le r\};$$

note that $\Phi(\hat{x}, \hat{y}) \ge \theta$. By choosing L large enough we may assume that

$$\sup\{\Phi(x,y) \mid x, y \in \overline{G},\ |x-y| = r\} \le 0,$$

and hence that $|\hat{x}-\hat{y}|<r$.

Now suppose for the moment that $\hat{x}\in\partial G$, which immediately yields that

$$\xi_0(\hat{x})\cdot D_x w(\hat{x},\hat{y})\le 0;$$
$$-\xi_0(\hat{x})\cdot D_2 w(\hat{x},\hat{y})\le -1.$$

This a contradiction, which shows that $\hat{x}\in G$.

Since

$$\big(D_x w(\hat{x},\hat{y}), D_y w(\hat{x},\hat{y}), C\Big(\begin{pmatrix} I & -I \\ -I & I\end{pmatrix}+|\hat{x}-\hat{y}|^2\begin{pmatrix} I & 0 \\ 0 & I\end{pmatrix}\Big)\big)\in J^{2,+}w(\hat{x},\hat{y}),$$

there are matrices $X,Y\in\mathcal{S}^N$ such that

$$(LD_x w(\hat{x},\hat{y}),X)\in\overline{J}^{2,+}u(\hat{x}),$$
$$(-LD_y w(\hat{x},\hat{y}),-Y)\in\overline{J}^{2,-}v(\hat{y}),$$
$$\begin{pmatrix} X & 0 \\ 0 & Y\end{pmatrix}\le LC\Big(\begin{pmatrix} I & -I \\ -I & I\end{pmatrix}+|\hat{x}-\hat{y}|^2\begin{pmatrix} I & 0 \\ 0 & I\end{pmatrix}\Big).$$

Since u and v are a subsolution and a supersolution of (7) and $\hat{x}\in G$, we have

$$\lambda u(\hat{x})+H(\hat{x},LD_x w(\hat{x},\hat{y}),X)\le 0,$$
$$\lambda v(\hat{y})+H(\hat{y},-LD_y w(\hat{x},\hat{y}),-Y)\ge 0.$$

Computations parallel to those in the proof of Theorem 4 yield

$$\operatorname{tr}\sigma\sigma^T(\hat{x},a)X+\operatorname{tr}\sigma\sigma^T(\hat{y},a)Y\le 4NCLM^2|\hat{x}-\hat{y}|^2,$$

and

$$b(\hat{x},a)\cdot D_x w(\hat{x},\hat{y})+b(\hat{y},a)\cdot D_y w(\hat{x},\hat{y})\le 2MC|\hat{x}-\hat{y}|^2.$$

Thus we obtain

$$\begin{aligned}0\ge&\lambda(u(\hat{x})-v(\hat{y}))+H(\hat{x},D_x w(\hat{x},\hat{y}),X)-H(\hat{y},D_y w(\hat{x},\hat{y}),-Y)\\ \ge&\lambda\theta-4NCLM^2|\hat{x}-\hat{y}|^2-2MC|\hat{x}-\hat{y}|^2-M|\hat{x}-\hat{y}|.\end{aligned}$$

Now, noting (see e.g. [CIL]) that $L|\hat{x}-\hat{y}|^2\to 0$ as $L\to\infty$, we get a contradiction, $0>\theta$, from the above inequality as we let $L\to\infty$. QED

An approximation of Ito integrals

Let $\{\lambda_t^i\}_{t\ge 0}$, with $i\in I(m)$, be stochastic processes with values in $[0,1]$ on a filtered probability space $(\Omega,\mathcal{F},P,\{\mathcal{F}_t\}_{t\ge 0})$ satisfying the usual conditions. Assume that the processes $\{\lambda_t^i\}_{t\ge 0}$ are $\mathcal{F}_t$–adapted, that

$$\sum_{i\in I(m)}\lambda_t^i(\omega)=1\qquad\forall t\ge 0,\ \forall\omega\in\Omega,$$

and that for all $i \in I(m)$,

$$\lambda_t^i(\omega) = \lambda_{[t]}^i(\omega) \qquad \forall t \geq 0,\ \forall \omega \in \Omega,$$

i.e.,

$$\lambda_t^i(\omega) = \lambda_k^i(\omega) \qquad \text{if}\ \ k \in \mathbf{N} \cup \{0\}\ \ \text{and}\ \ k \leq t < k+1.$$

Fix $n \in \mathbf{N}$. Define intervals depending on $\omega \in \Omega$ as

$$I_{k,j,i}^n(\omega) = \Big[k + \frac{j}{n} + \frac{1}{n}\sum_{l<i} \lambda_k^l(\omega),\ k + \frac{j}{n} + \frac{1}{n}\sum_{l\leq i} \lambda_k^l(\omega)\Big),$$

where $k \in \mathbf{N} \cup \{0\}$, $j \in \{0, ..., n-1\}$, and $i \in I(m)$. Set

$$\theta_{k,j,i}^n(\omega) = k + \frac{j}{n} + \frac{1}{n}\sum_{l<i} \lambda_k^l(\omega)$$

for $k \in \mathbf{N} \cup \{0\}$, $j \in \{0, ..., n-1\}$, and $i \in I(m+1)$. Note that for each $\omega \in \Omega$, the family of intervals $I_{k,j,i}^n$, with $k \in \mathbf{N} \cup \{0\}$, $j \in \{0, ..., n-1\}$, and $i \in I(m)$, are mutually disjoint,

$$I_{k,j,i}^n(\omega) = [\theta_{k,j,i}^n(\omega),\ \theta_{k,j,i+1}^n(\omega)),$$

and

$$\bigcup_{i \in I(m)} I_{k,j,i}^n(\omega) = \Big[k + \frac{j}{n}, k + \frac{j+1}{n}\Big).$$

For $k \in \mathbf{N} \cup \{0\}$, $j \in \{0, ..., n-1\}$, and $i \in I(m)$ we define $\psi^{n,k,j,i} : [0, \infty) \times \Omega \to [0, 1/n]$ and $\varphi^{n,i} : [0, \infty) \times \Omega \to [0, \infty)$ by

$$\psi_t^{n,k,j,i}(\omega) = \begin{cases} (\lambda_k^i(\omega))^{-1} \displaystyle\int_0^t \mathbf{1}_{I_{k,j,i}^n(\omega)}(s)ds & \text{if } \lambda_k^i(\omega) > 0, \\ 0 & \text{if }\ \lambda_k^i(\omega) = 0\ \ \text{and}\ \ t < \theta_{k,j,i}^n(\omega), \\ \frac{1}{n} & \text{if }\ \lambda_k^i(\omega) = 0\ \ \text{and}\ \ t \geq \theta_{k,j,i}^n(\omega), \end{cases}$$

and

$$\varphi_t^{n,i}(\omega) = \sum_{k=0}^{\infty}\sum_{j=0}^{n-1} \psi_t^{n,k,j,i}(\omega).$$

It is clear that for each $\omega \in \Omega$, the function: $t \mapsto \varphi_t^{n,i}(\omega)$ is non-decreasing and right-continuous on $[0, \infty)$. Note that if $t \geq 0$ and $(k, j) \in \mathbf{N} \cup \{0\} \times \{0, ..., n-1\}$ satisfy $k + \frac{j}{n} \leq t < k + \frac{j+1}{n}$, then

$$\varphi_t^{n,i}(\omega) = k + \frac{j}{n} + \psi_t^{n,k,j,i}(\omega).$$

In particular, we deduce that

$$\lim_{t \nearrow k+\frac{j}{n}} \varphi_t^{n,i}(\omega) = k + \frac{j}{n} \qquad \text{if } k \geq 1,$$

and

$$(1) \qquad |\varphi_t^{n,k,j,i}(\omega) - t| \leq \frac{1}{n} \qquad \forall t \geq 0,\ \forall \omega \in \Omega.$$

Observe that for each $t \geq 0$, the random variable $\varphi_t^{n,i}$ is a stopping (or Markov) time relative to $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\geq 0})$. To see this, let $k \in \mathbf{N} \cup \{0\}$ and $j \in \{0, ..., n-1\}$ satisfy $k + \frac{j}{n} \leq t < k + \frac{j+1}{n}$.

Fix $T \geq 0$. Consider first the case when $T < k + \frac{j}{n}$. Then we have

$$T < \varphi_t^{n,i}(\omega) \qquad \forall \omega \in \Omega,$$

and hence,

$$B := \{\omega \in \Omega \mid \varphi_t^{n,i}(\omega) \leq T\} = \emptyset \in \mathcal{F}_T.$$

Next consider the case when $k + \frac{j}{n} \leq T < k + \frac{j+1}{n}$. By the definition of $\psi_t^{n,k,j,i}$, we see that if $\lambda_k^i(\omega) = 0$, then

$$\varphi_t^{n,i}(\omega) \leq T \qquad \Longleftrightarrow \qquad t < \theta_{k,j,i}^n(\omega),$$

and if $\lambda_k^i(\omega) > 0$, then

$$\varphi_t^{n,i}(\omega) \leq T \qquad \Longleftrightarrow \qquad t - \theta_{k,j,i}^n(\omega) \leq \lambda_k^i(\omega)\left(T - k - \frac{j}{n}\right).$$

Hence, we have

$$\begin{aligned} B = &\{\omega \in \Omega \mid \lambda_k^i(\omega) = 0,\ t < \theta_{k,j,i}^n(\omega)\} \\ &\cup \left\{\omega \in \Omega \mid \lambda_k^i(\omega) > 0,\ t - \theta_k^i(\omega) \leq \lambda_k^i(\omega)\left(T - k - \frac{j}{n}\right)\right\} \in \mathcal{F}_k \subset \mathcal{F}_T. \end{aligned}$$

Finally, if $T \geq k + \frac{j+1}{n}$, then we have

$$\varphi_t^{n,i}(\omega) \leq k + \frac{j+1}{n} \leq T \qquad \forall \omega \in \Omega,$$

and so, $B = \Omega \in \mathcal{F}_T$, concluding that $\varphi_t^{n,i}$ is a stopping time relative to $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\geq 0})$.

Let $\{W_t^i\}_{t\geq 0}$, with $i \in I(m)$, be independent d–dimensional standard Brownian motions on $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t\geq 0})$. Fix $i \in I(m)$. Define

$$\mathcal{G}_t^{n,i} = \mathcal{F}_{\varphi_t^{n,i}} \equiv \{B \in \mathcal{F} \mid B \cap \{\varphi_t^{n,i} \leq T\} \in \mathcal{F}_T \ \forall T \geq 0\}.$$

Define $X^{n,i} : [0,\infty) \times \Omega \to \mathbf{R}^d$ and $M^{n,i} : [0,\infty) \times \Omega \to \mathbf{R}^d$, respectively, by

$$X_t^{n,i}(\omega) = W^i_{\varphi_t^{n,i}(\omega)}(\omega),$$

and for any $(t,\omega) \in [0,\infty) \times \Omega$, if $k \in \mathbf{N} \cup \{0\}$ and $k \leq t < k+1$,

$$\begin{aligned} M_t^{n,i}(\omega) = &\sum_{0\leq p<k-1} (\lambda_p^i(\omega))^{\frac{1}{2}} (X_{p+1}^{n,i}(\omega) - X_p^{n,i}(\omega)) \\ &+ (\lambda_k^i(\omega))^{\frac{1}{2}} (X_t^{n,i}(\omega) - X_k^{n,i}(\omega)). \end{aligned} \tag{2}$$

By virtue of the optional sampling theorem due to Doob ([IW, Theorem 6.11]), we see that the process $\{X_t^{n,i}\}_{t\geq 0}$ is a $\mathcal{G}_t^{n,i}$–martingale. It is easily seen that for all $(t,\omega) \in [0,\infty) \times \Omega$,

$$M_t^{n,i}(\omega) = \sum_{k=0}^{\infty} (\lambda_k^i(\omega))^{\frac{1}{2}} (X_{t\wedge(k+1)}^{n,i}(\omega) - X_{t\wedge k}^{n,i}(\omega)).$$

Now observe that $\{M_t^{n,i}\}_{t\geq 0}$ is a $\mathcal{G}_t$–martingale. Let $0 \leq s < t$ and $k \in \mathbf{N} \cup \{0\}$. If $s \leq k \leq t$, then we have

$$\begin{aligned} &E\left((\lambda_k^i)^{\frac{1}{2}} (X_{t\wedge(k+1)}^{n,i} - X_{t\wedge k}^{n,i}) \mid \mathcal{G}_s^{n,i}\right) \\ &= E\left((\lambda_k^i)^{\frac{1}{2}} E\left(X_{t\wedge(k+1)}^{n,i} - X_{t\wedge k}^{n,i} \mid \mathcal{G}_k^{n,i}\right) \mid \mathcal{G}_s^{n,i}\right) \\ &= E\left((\lambda_k^i)^{\frac{1}{2}} (X_k^{n,i} - X_k^{n,i}) \mid \mathcal{G}_s^{n,i}\right) \\ &= 0 = (\lambda_k^i)^{\frac{1}{2}} (X_{s\wedge(k+1)}^{n,i} - X_{s\wedge k}^{n,i}). \end{aligned}$$

If $k < s$, then we have

$$\begin{aligned} &E\left((\lambda_k^i)^{\frac{1}{2}} (X_{t\wedge(k+1)}^{n,i} - X_{t\wedge k}^{n,i}) \mid \mathcal{G}_s^{n,i}\right) \\ &= (\lambda_k^i)^{\frac{1}{2}} \left(E(X_{t\wedge(k+1)}^{n,i} \mid \mathcal{G}_s^{n,i}) - X_k^{n,i}\right) \\ &= (\lambda_k^i)^{\frac{1}{2}} (X_{s\wedge(k+1)}^{n,i} - X_{s\wedge k}^{n,i}). \end{aligned}$$

Next, if $t < k$, then

$$\begin{aligned} &E\left((\lambda_k^i)^{\frac{1}{2}} (X_{t\wedge(k+1)}^{n,i} - X_{t\wedge k}^{n,i}) \mid \mathcal{G}_s^{n,i}\right) \\ &= 0 = (\lambda_k^i)^{\frac{1}{2}} (X_{s\wedge(k+1)}^{n,i} - X_{s\wedge k}^{n,i}). \end{aligned}$$

Thus we see that
$$E\left(M_t^{n,i} \mid \mathcal{G}_s^{n,i}\right) = M_s^{n,i}.$$

Note that the process $\{M_t^{n,i}\}_{t\geq 0}$ is a continuous process. Indeed, for any $\omega \in \Omega$ and $k \in \mathbf{N} \cup \{0\}$ the function

$$f : t \mapsto \lambda_k^i(\omega)^{\frac{1}{2}}(X_t^{n,i}(\omega) - X_k^{n,i}(\omega))$$

is continuous in $[k,\ k+1)$. Moreover, we have

$$f(k) = 0,$$
$$\lim_{t \nearrow k+1} f(t) = (\lambda_k^i(\omega))^{\frac{1}{2}}(X_{k+1}^{n,i}(\omega) - X_k^{n,i}(\omega)),$$

and therefore,
$$\lim_{t \nearrow k+1} M_t^{n,i}(\omega) = M_{k+1}^{n,i}(\omega) = \lim_{t \searrow k+1} M_t^{n,i}(\omega).$$

This shows that for each $\omega \in \Omega$ the function : $t \mapsto M_t^{n,i}(\omega)$ is continuous on $[0, \infty)$.

Note also that the process $\{X_t^{n,i}\}_{t\geq 0}$ is right–continuous. We use the notation:

$$W_t^i(\omega) = (W_t^{i,1}, ..., W_t^{i,d}), \qquad X_t^{n,i}(\omega) = (X_t^{n,i,1}(\omega), ..., X_t^{n,i,d}(\omega)),$$
$$M_t^{n,i}(\omega) = (M_t^{n,i,1}(\omega), ..., M_t^{n,i,d}(\omega)).$$

Fix any $\alpha \in I(d)$ and $i \in I(m)$. Since $\{(W_t^{i,\alpha})^2 - t\}_{t\geq 0}$ is an $\mathcal{F}_t$–martingale, the optional sampling theorem guarantees that $\{(X_t^{n,i,\alpha})^2 - \varphi_t^{n,i}\}_{t\geq 0}$ is a $\mathcal{G}_t^{n,i}$–martingale and hence that
$$\langle X^{n,i,\alpha}\rangle_t = \varphi_t^{n,i}.$$

Next, fix any $k \in \mathbf{N} \cup \{0\}$ and observe that for $k \leq s \leq t < k+1$,

$$\left(M_t^{n,i,\alpha}\right)^2 = \left(M_s^{n,i,\alpha}\right)^2 + \lambda_k^i \left(X_t^{n,i,\alpha} - X_s^{n,i,\alpha}\right)^2 + 2(\lambda_k^i)^{\frac{1}{2}} M_s^{n,i,\alpha}\left(X_t^{n,i,\alpha} - X_s^{n,i,\alpha}\right),$$

and

$$\begin{aligned}
&E(\left(M_t^{n,i,\alpha}\right)^2 \mid \mathcal{G}_s^{n,i}) \\
&= \left(M_s^{n,i,\alpha}\right)^2 + \lambda_k^i \left\{E(\left(X_t^{n,i,\alpha}\right)^2 \mid \mathcal{G}_s^{n,i}) - 2X_s^{n,i,\alpha}E(X_t^{n,i,\alpha} \mid \mathcal{G}_s^{n,i}) + \left(X_s^{n,i,\alpha}\right)^2\right\} \\
&+ 2(\lambda_k^i)^{\frac{1}{2}} M_s^{n,i,\alpha}\left(E(X_s^{n,i,\alpha} \mid \mathcal{G}_s^{n,i}) - X_s^{n,i,\alpha}\right) \\
&= \left(M_s^{n,i,\alpha}\right)^2 + \lambda_k^i \left(E(\varphi_t^{n,i} \mid \mathcal{G}_s^{n,i}) - \varphi_s^{n,i}\right).
\end{aligned}$$

Similarly, for $k \le s \le t < k+1$ and for $\alpha,\ \beta \in I(d)$, with $\alpha \neq \beta$, using the independence of

$$\sigma(W_\tau^{i,\alpha} - W_k^{i,\alpha} \mid k \le \tau \le k+1) \quad \text{and} \quad \sigma(W_\tau^{i,\beta} - W_k^{i,\beta} \mid k \le \tau \le k+1)$$

we have

$$\begin{aligned}
&E\left(M_t^{n,i,\alpha} M_t^{n,i,\beta} \mid \mathcal{G}_s^{n,i}\right) \\
&= M_s^{n,i,\alpha} M_s^{n,i,\beta} \\
&+ (\lambda_k^i)^{\frac{1}{2}} [M_s^{n,i,\alpha} E(X_t^{n,i,\beta} - X_s^{n,i,\beta} \mid \mathcal{G}^{n,i}) + M_s^{n,i,\beta} E(X_t^{n,i,\alpha} - X_s^{n,i,\alpha} \mid \mathcal{G}^{n,i})] \\
&+ \lambda_k^i E[(X_t^{n,i,\alpha} - X_s^{n,i,\alpha})(X_t^{n,i,\alpha} - X_s^{n,i,\alpha}) \mid \mathcal{G}_s^{n,i}] = M_s^{n,i,\alpha} M_s^{n,i,\beta}.
\end{aligned}$$

Therefore, we conclude that

$$\begin{aligned}
\langle M^{n,i,\alpha} \rangle_t &= \int_0^t \chi_s^{n,i}(\omega) ds, \\
\langle M^{n,i,\alpha}, M^{n,i,\beta} \rangle_t &= 0 \quad \text{if } \alpha \neq \beta,
\end{aligned}$$

where

$$\chi_t^{n,i}(\omega) = \sum_{k \in \mathbf{N} \cup \{0\}} \sum_{j=0}^{n-1} \mathbf{1}_{I_{k,j,i}^n(\omega)}(t) = \mathbf{1}_{\bigcup_{k=0}^\infty \bigcup_{j=0}^{n-1} I_{k,j,i}^n(\omega)}(t).$$

To continue, we need the following proposition.

Proposition A.1. *Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $\{\mathcal{A}_i\}_{i=0}^n$ be a family of sub–σ–fields of $\mathcal{F}$. Assume that $\mathcal{A}_0$, $\mathcal{A}_1$, ..., $\mathcal{A}_n$ are independent. Let $\mathcal{F}_i$, with $i \in I(n)$, be sub-σ–fields, respectively, of $\mathcal{A}_0 \vee \mathcal{A}_i$ such that $\mathcal{A}_0 \subset \mathcal{F}_i$. Set*

$$\mathcal{G} = \bigvee_{i \in I(n)} \mathcal{F}_i.$$

Let X be an $\mathcal{A}_0 \vee \mathcal{A}_1$–measurable random variable. Then we have

$$E(X \mid \mathcal{G}) = E(X \mid \mathcal{F}_1) \qquad a.s.$$

We give a proof of this proposition later, and now continue our discussions.

Define

$$\mathcal{H}_t^{n,i,\alpha} = \mathcal{F}_{[t]} \vee \sigma(M_s^{n,i,\alpha} - M_{[t]}^{n,i,\alpha} \mid [t] \le s \le t),$$

for all $n \in \mathbf{N} \cup \{0\}$, $i \in I(m)$, and $\alpha \in I(d)$, By the construction of $M_t^{n,i,\alpha}$ (see (2)), we see that

$$\sigma(M_s^{n,i,\alpha} - M_{[t]}^{n,i,\alpha} \mid [t] \le s \le t) \subset \sigma(W_s^{i,\alpha} - W_{[t]}^{i,\alpha} \mid [t] \le s \le [t]+1),$$

and therefore, that

$$\mathcal{F}_{[t]}, \quad \sigma(M_s^{n,i,\alpha} - M_{[t]}^{n,i,\alpha} \mid [t] \le s \le t), \qquad \text{with } (i,\alpha) \in I(m) \times I(d),$$

are independent. It is clear that $\{\mathcal{H}_t^{n,i,\alpha}\}_{t\ge 0}$ are non-decreasing σ–fields for all $(n,i,\alpha) \in \mathbf{N} \cup \{0\} \times I(m) \times I(d)$.

Since $\mathcal{H}_t^{n,i,\alpha} \subset \mathcal{G}_t^{n,i,\alpha}$ and $M_t^{n,i,\alpha}$ is $\mathcal{H}_t^{n,i,\alpha}$–measurable, it follows that $\{M_t^{n,i,\alpha}\}_{t\ge 0}$ is an $\mathcal{H}^{n,i,\alpha}$–martingale and that

$$\langle M^{n,i,\alpha}\rangle_t = \int_0^t \chi_s^{n,i} ds \quad \text{and} \quad \langle M^{n,i,\alpha}, M^{n,i,\beta}\rangle_t = 0 \quad \text{if } \alpha \ne \beta.$$

Define

$$\mathcal{H}_t^n = \bigvee_{(i,\alpha)\in I(m)\times I(d)} \mathcal{H}_t^{n,i,\alpha}$$

for $t \ge 0$ and $n \in \mathbf{N} \cup \{0\}$. Using Proposition A.1, we deduce that $\{M^{n,i,\alpha}\}_{t\ge 0}$ are all $\mathcal{H}_t^n$–martingales and that for all $t \ge 0$ and $(n,i,\alpha) \in \mathbf{N} \cup \{0\} \times I(m) \times I(d)$,

$$\langle M_t^{n,i,\alpha}\rangle_t = \int_0^t \chi_s^{n,i} ds.$$

In view of (1), we see that as $n \to \infty$,

$$X_t^{n,i}(\omega) \to W_t^i(\omega)$$

almost surely for all $(t,i) \in [0,\infty) \times I(m)$, and hence, from (2) that as $n \to \infty$,

$$M_t^{n,i}(\omega) \to \sum_{0\le p<[t]} (\lambda_p^i(\omega))^{\frac{1}{2}} (W_{p+1}^i(\omega) - W_p^i(\omega)) + (\lambda_{[t]}^i)^{\frac{1}{2}} (W_t^i(\omega) - W_{[t]}^i(\omega))$$

almost surely for all $(t,i) \in [0,\infty) \times I(m)$. Recalling the definition of the Ito integral, we observe that as $n \to \infty$,

$$M_t^{n,i}(\omega) \to \int_0^t (\lambda_s^i(\omega))^{\frac{1}{2}} dW_s^i$$

almost surely for all $(t,i) \in [0,\infty) \times I(m)$.

Proposition A.2. *1) $\{M_t^{n,i}\}_{t\ge 0}$ are $\mathcal{H}_t^n$–martingales for all $i \in I(m)$. 2) $\mathcal{H}_t^n \subset \mathcal{F}_{t+\frac{1}{n}}$. 3) For $t \ge 0$, $(n,i) \in \mathbf{N} \cup \{0\} \times I(m)$, and α, $\beta \in I(d)$,*

$$\langle M^{n,i,\alpha}\rangle_t = \int_0^t \chi_s^{n,i} ds \quad \textit{and} \quad \langle M^{n,i,\alpha}, M^{n,i,\beta}\rangle_t = 0 \quad \textit{if } \alpha \ne \beta,$$

where

$$\chi_t^{n,i} = \sum_{k=0}^{\infty}\sum_{j=0}^{n-1} \mathbf{1}_{I^n_{k,j,i}}(t) = \mathbf{1}_{\bigcup_{k=0}^{\infty}\bigcup_{j=0}^{n-1} I^n_{k,j,i}}(t).$$

4) $M_t^{n,i}(\omega) \to \int_0^t (\lambda_s^i(\omega))^{\frac{1}{2}} dW_s^i$ *for all* $(t,i) \in [0,\infty) \times I(m)$ *as* $n \to \infty$.

We now present a proof of Proposition A.1. We begin with a lemma.

Lemma A.1. *Let* $\mathcal{A}$, $\mathcal{B}$, *and* $\mathcal{C}$ *are complete sub–*σ*–fields of* $\mathcal{F}$. *Then*

$$(\mathcal{A} \vee \mathcal{B}) \cap \mathcal{C} = (\mathcal{A} \cap \mathcal{C}) \vee (\mathcal{B} \cap \mathcal{C}).$$

Proof. Set

$$\mathcal{S} = \{A \cap B \mid A \in \mathcal{A},\ B \in \mathcal{B}\},$$
$$\mathcal{T} = \Big\{ \bigcup_{i\in I(m)} S_i \mid m \in \mathbf{N},\ S_i \in \mathcal{S}\Big\}.$$

Then $\mathcal{T}$ is a subfield of $\mathcal{F}$. That is,

1) $\emptyset, \Omega \in \hat{\mathcal{T}}$,
2) $T \in \mathcal{T} \Longrightarrow T^c \in \mathcal{T}$,
3) $m \in \mathbf{N},\ \ T_i \in \mathcal{T}\ \ \forall i \in I(m) \implies \bigcup_{i\in I(m)} T_i \in \mathcal{T}$.

It is well-known that any set $D \in \mathcal{A} \vee \mathcal{B}$ can be approximated by a set in $\mathcal{A} \cup \mathcal{B}$, i.e., for any $\varepsilon > 0$ and $D \in \mathcal{A} \vee \mathcal{B}$ there is a set $E \in \mathcal{T}$ such that $P(D \triangle E) < \varepsilon$, where $D \triangle E = (D \setminus E) \cup (E \setminus D)$.

Fix $D \in \mathcal{A} \vee \mathcal{B}$. For each $j \in \mathbf{N}$ we choose a set $E \in \mathcal{T}$ so that

$$P(D \triangle E_j) < \frac{1}{j}.$$

The set E_j can be represented as

$$E_j = \bigcup_{i\in I(k)} A_{ji} \cap B_{ji}$$

for some $k \in \mathbf{N}$, where $A_{ji} \in \mathcal{A}$ and $B_{ji} \in \mathcal{B}$, and we may assume that the sets $A_{j1} \cap B_{j1}$, ..., $A_{jk} \cap B_{jk}$ are mutually disjoint.

Set

$$F_j = \bigcup_{i\in I(k)} D \cap A_{ji} \cap B_{ji}.$$

It is clear that $F_j \in (\mathcal{A} \cap \mathcal{C}) \vee (\mathcal{B} \vee \mathcal{C})$ and $F_j \subset D$. Noting that

$$P(D \setminus F_j) \le P(D \triangle E_j) < \frac{1}{j},$$

we observe that if we set

$$F = \bigcup_{j \in \mathbf{N}} F_j,$$

then

$$F \subset D, \quad P(D \setminus F) = 0 \quad \text{and} \quad F \in (\mathcal{A} \cap \mathcal{C}) \vee (\mathcal{B} \cap \mathcal{C}).$$

Setting

$$N = D \setminus F,$$

we have

$$D = N \cup F \in (\mathcal{A} \cap \mathcal{C}) \vee (\mathcal{B} \cap \mathcal{C}),$$

because of the completeness of $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$. QED

Proof of Proposition A.1. By replacing $\mathcal{A}_i$ and $\mathcal{F}_j$ by their completions if necessary, we may assume that $\mathcal{A}_i$ and $\mathcal{F}_j$ are all complete.

For $i \in I(n)$ set

$$\mathcal{B}_i = \mathcal{F}_i \cap \mathcal{A}_i.$$

Then we have

$$\mathcal{F}_i = \mathcal{A}_0 \vee \mathcal{B}_i \quad \forall i \in I(n).$$

Indeed, since $\mathcal{F}_i \subset \mathcal{A}_0 \cup \mathcal{A}_i$, using Lemma A.1, we have

$$\mathcal{F}_i = \mathcal{F}_i \cap (\mathcal{A}_0 \vee \mathcal{A}_i) = (\mathcal{F}_i \cap \mathcal{A}_0) \vee (\mathcal{F}_i \cap \mathcal{A}_i) = \mathcal{A}_0 \vee \mathcal{B}_i.$$

Also we have

$$\mathcal{G} = \mathcal{A}_0 \vee \mathcal{B}_1 \vee \cdots \vee \mathcal{B}_n.$$

Define

$$\mathcal{H} := \bigvee_{i=2}^{n} \mathcal{B}_i.$$

Then we have

$$\mathcal{G} = \mathcal{F}_1 \vee \mathcal{H}.$$

It is easy to see that $\mathcal{F}_1$ and $\mathcal{H}$ are independent, and $\mathcal{A}_0 \vee \mathcal{A}_1$ and $\mathcal{H}$ are independent.

Set

$$\varphi = E(X \mid \mathcal{F}_1).$$

By definition, we have

$$E(\varphi f) = E(Xf) \qquad \text{for all bounded } \mathcal{F}_1\text{–measurable } f.$$

For any bounded $\mathcal{F}_1$–measurable g_1 and any bounded $\mathcal{H}$–measurable g_2, we have

$$E(Xg_1g_2) = E(Xg_1)Eg_2 = E(\varphi g_1)Eg_2 = E(\varphi g_1 g_2).$$

Hence, we have

$$E(Xg) = E(\varphi g) \qquad \text{for all bounded } \mathcal{G}\text{–measurable } \; g.$$

Since φ is $\mathcal{G}$–measurable, we conclude that $\varphi = E(X \mid \mathcal{G})$ almost surely. QED

Acknowledgements. The first author would like to thank H. Nagai and T. Mikami for helping him to understand some of basics in stochastic analysis. The first author was supported by Grant-in-Aid for Scientific Research (No. 09440067 and No. 12440044) and by the Bilateral Program of Scientist Exchanges between JSPS, Japan and CNR, Italy. Major part of this work was done while the first author was a member of Department of Mathematics, Tokyo Metropolitan University , and he is grateful for colleagues there. Part of this work was done while he was visiting Dipartimento di Metodi e Modelli Matematici per le Scienze Applicate and Dipartimento di Matematica, Università di Roma "La Sapienza", and he is grateful for their hospitality. Part of this work was done while the second author was visiting Department of Mathematics, Tokyo Metropolitan University , and she is grateful for its hospitality.

References

[BB] G. Barles and J. Burdeau, The Dirichlet problem for semilinear second-order degenerate elliptic equations and applications to stochastic exit time control problems. Comm. Partial Differential Equations 20 (1995), no. 1-2, 129–178.

[BR] G. Barles and E. Rouy, A strong comparison result for the Bellman equation arising in stochastic exit time control problems and its applications. Comm. Partial Differential Equations **23** (1998), no. 11-12, 1995–2033.

[CDL] I. Capuzzo–Dolcetta and P. L. Lions, Viscosity solutions of Hamilton–Jacobi equations and state constraints, Trans. Amer. Math. Soc., **318** (1990), no. 2, 643–683.

[CIL] M. G. Crandall, H. Ishii, and P.-L. Lions, User's guide to viscosity solutions of second order partial differential equations. Bull. Amer. Math. Soc. (N.S.)**27** (1992), no. 1, 1–67.

[IW] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, North-Holland Math. Library **24**, North-Holland, Kodansha, 1981, Amsterdam, New York, Tokyo.

[IK] H. Ishii and S. Koike, A new formulation of state constraint problems for first-order PDEs. SIAM J. Control Optim. **34** (1996), no. 2, 554–571.

[ILT] H. Ishii, P. Loreti, and E. Tessitore, A PDE approach to stochastic invariance, Discrete and continuous dynamical systems, **6** (2000), no. 3, 651–664.

[H] U. G. Haussmann, Existence of optimal Markovian controls for degenerate diffusions, Lecture Notes in Control and Information Sciences **78**, Springer-Verlag, Berlin, New York, 1986, pp. 171–186.

[HL] U. G. Haussmann and J. P. Lepeltier, On the existence of optimal controls, SIAM J. Control Optim., **28** (1990), no. 4, 851–902.

[EHJ] N. El Karoui, D. Huu Nguyen, and M. Jeanblanc-Picqué, Compactification methods in the control of degenerate diffusions: existence of an optimal control. Stochastics **20** (1987), no. 3, 169–219.

[KA] M. A. Katsoulakis, Viscosity solutions of second order fully nonlinear elliptic equations with state constraints. Indiana Univ. Math. J. **43** (1994), no. 2, 493–519.

[KU] T. G. Kurtz, Martingale problems for controlled processes. Stochastic modelling and filtering (Rome, 1984), 75–90, Lecture Notes in Control and Inform. Sci., **91**, Springer, Berlin-New York, 1987.

[LL] J.-M. Lasry and P.-L. Lions, Nonlinear elliptic equations with singular boundary conditions and stochastic control with state constraints. I. The model problem. Math. Ann. **283** (1989), no. 4, 583–630.

[PLB] P.-L. Lions, *Generalized solutions of Hamilton-Jacobi equations*, Research Notes in Mathematics,**69**, Pitman (Advanced Publishing Program), Boston, Mass.-London, 1982.

[L] P. Loreti, Some properties of constrained viscosity solutions of Hamilton-Jacobi-Bellman equations, SIAM J. Control Optim. **25** (1987), no. 5, 1244–1252.

[LT] P. Loreti and M. E. Tessitore, Approximation and regularity results on constrained viscosity solutions of Hamilton-Jacobi-Bellman equations, J. Math. Systems, Estim. Control, **4** (1994), 467-483.

[NA] H. Nagai, *Stochastic differential equations*, in Japanese, Kyoritu-Shuppan, Mathematics in 21st Century, **27**, 1999.

[NI1] M. Nisio, Some remarks on stochastic optimal controls. Proceedings of the Third Japan-USSR Symposium on Probability Theory (Tashkent, 1975), pp. 446–460. Lecture Notes in Math., **550**, Springer, Berlin, 1975.

[NI2] M. Nisio, *Lectures on stochastic control theory*, ISI Lecture Notes, **9**, MacMillan India Ltd, Delhi, Bombay, 1981.

[S] H. M. Soner, Optimal control with state-space constraint. I. SIAM J. Control Optim. **24** (1986), no. 3, 552–561.

[SV] D. W. Stroock and S. R. S. Varadhan, Multidimensional diffusion processes. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], **233**, Springer-Verlag, Berlin-New York, 1979.

[YZ] J. Yong and X. Y. Zhou, *Stochastic controls: Hamiltonian and HJB equations*, Applications of Mathematics **43**, Springer-Verlag, 1999, New York.