A Class of Stochastic Optimal Control Problems with State Constraint

HITOSHI ISHII AND PAOLA LORETI

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 $a_z \in A$, where A denotes the control set, such that the diffusion matrix $\sigma(x, a)$ vanishes for $a = a_z$ and for $x \in \partial G$ in a neighborhood of z and the drift vector $b(x, a)$ directs inside of G at z for $a = a_z$ and $x = z$ as well as 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_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 Hölder regularity results for the viscosity solution of the state constraint problem.

1. INTRODUCTION

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

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

Indiana University Mathematics Journal ©, Vol. 51, No. 5 (2002)

where $(\Omega^\alpha, \mathcal{F}^\alpha, \{\mathcal{F}_t^\alpha\}_{t\geq 0}, P^\alpha)$ is a filtered probability space satisfying the usual condition (see e.g. [19]), $\{W_t^\alpha\}_{t\geq 0}$ is a standard l-dimensional Brownian motion on this filtered probability space, $\{u_t^\alpha\}_{t\geq 0}$ is an $\{\mathcal{F}_t^\alpha\}_{t\geq 0}$–progressively measurable 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

$$dX_t = b(X_t, u_t^\alpha)\,dt + \sigma(X_t, u_t^\alpha)dW_t^\alpha, \quad X_0 = x, \tag{1.1}$$

where $b : \mathbf{R}^N \times A \to \mathbf{R}^N$ and $\sigma : \mathbf{R}^N \times A \to \mathbf{R}^N \otimes \mathbf{R}^l$ are given functions. The set of controlled systems at x will be denoted by C^x.

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 $\alpha \in C^x$ for which

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

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

$$J(x, \alpha) = E^\alpha \int_0^\infty e^{-\lambda t} f(X_t^\alpha, u_t^\alpha)\,dt \tag{1.2}$$

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

$$V(x) = \inf_{\alpha \in \mathcal{A}(x)} J(x, \alpha) \tag{1.3}$$

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 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 Hamilton-Jacobi-Bellman equation above makes the classical sense. It is nowadays well recognized that the best way to interpret the Hamilton-Jacobi-Bellman equation above is to adapt the notion of viscosity solutions. 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 [12], 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 [18]] 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 Hamilton-Jacobi-Bellman 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 [5], which is a modification of the boundary value 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 is due to J.-M. Lasry and P.-L. Lions [11] and in their paper they dealt with the case of nondegenerate diffusion (i.e., the case where σ = the identity matrix) and unbounded drift b so that the value function behaves singularly near the boundary ∂G. M. Katsoulakis [10] initiated to study the case where diffusion depends on the control and degenerates on the boundary. G. Barles and J. Burdeau [1] 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 [2].

The main results of this paper concern:

(i) the identification of value functions of different control problems as the viscosity solution of

$$\begin{cases} \lambda u(x) + H(x, Du(x), D^2u(x)) \geq 0 & (x \in \overline{G}), \\ \lambda u(x) + H_{in}(x, Du(x), D^2u(x)) \leq 0 & (x \in \overline{G}), \end{cases} \tag{1.4}$$

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 $A(x)$);

(ii) the Hölder regularity of the value functions.

Here it may be useful for those readers who are not familiar with viscosity solutions defined on closed sets to recall the definition of viscosity solution of (1.4). An upper (resp., lower) semicontinuous function $u : \overline{G} \to \mathbf{R}$ is called a viscosity subsolution (resp., supersolution) of (1.4) provided whenever $\varphi \in C^2(\overline{G})$ and $u - \varphi$ attains a maximum (resp., minimum) over $\overline{G}$ at $y \in \overline{G}$ then

$$\lambda u(y) + H_{in}(y, D\varphi(y), D^2\varphi(y)) \le 0$$
$$(\text{ resp.,} \quad \lambda u(y) + H(y, D\varphi(y), D^2\varphi(y)) \ge 0 \text{ }).$$

Then as usual, a continuous function $u : \overline{G} \to \mathbf{R}$ is called a viscosity solution of (1.4) if it is both a viscosity sub- and supersolution of (1.4). See also [4] for a general review of the theory of viscosity solutions.

Regarding the identification, our results are close to those obtained by [1] and the new feature beyond [1] in our result is dependence of σ in a. Related results can be found in [2] in the framework of the Dirichlet problem. On the other hand our degeneracy assumption on σ on the boundary is stronger than those in [1] and [10]. [10] studies a different case from ours, at least, for the continuity result of value functions. In our results we consider three kinds of value functions, the identification of which is new in the setting of stochastic control. For the deterministic case, we refer to [13]. Again the Hölder continuity is 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 [3], [14], and [5]. 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.

Finally, we remark that we have recently improved Theorem 2.4 below, for which we refer to [20].

2. Statements of the Problems and Main Results

Let A be a *compact, convex* subset of a Euclidean space $\mathbf{R}^m$ and let

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

be given functions which satisfy:

(A1) there is a constant $M > 0$ such that for all $x, y \in \mathbf{R}^N$ and $a \in A$,

$$\begin{aligned}&\max\{|\sigma(x,a)|, |b(x,a)|, |f(x,a)|\} \le M;\\ &\max\{|\sigma(x,a) - \sigma(y,a)|, |b(x,a) - b(y,a)|, |f(x,a) - f(y,a)|\}\\ &\qquad \le M|x - y|.\end{aligned}$$

(A2) there exists a continuous function $\omega : [0, +\infty) \to [0, +\infty)$, with $\omega(0) = 0$, such that

$$\begin{aligned}\max\{|\sigma(x,a) - \sigma(x,a')|, |b(x,a) - b(x,a')|, |f(x,a) - f(x,a')|\}\\ \le \omega(|a - a'|), \quad \text{for all } a, a' \in A, \ x \in \mathbf{R}^N.\end{aligned}$$

Other assumptions we need are:

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

(A4) for any $z \in \partial G$, there exist $r_z > 0$ and $a_z \in A$ such that

$$\sigma(x, a_z) = 0 \quad (x \in B(z, r_z) \cap \partial G), \tag{2.1}$$

$$B(x + tb(x, a_z), r_z t) \subset \overline{G} \quad (x \in B(z, r_z) \cap \overline{G},\ 0 \le t \le r_z); \tag{2.2}$$

(A5) 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.

In one of the main results we need additional regularity assumptions on σ and b as well as an assumption similar to but slightly different from (A4):

(A6)
$$\sup_{a \in A} \|\sigma(\cdot, a)\|_{W^{2,\infty}(\mathbf{R}^N)} < \infty;$$

(A7) there is a constant $M > 0$ such that for all $x \in \mathbf{R}^N$ and $a, a' \in A$,

$$\max\{|\sigma(x, a) - \sigma(x, a')|, |b(x, a) - b(x, a')|\} \le M|a - a'|.$$

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

$$\sigma(x, \hat{a}(x)) = 0 \quad (x \in \partial G);$$
$$B(x + tb(x, \hat{a}(x)), rt) \subset \overline{G} \quad (x \in \overline{G} \cap \bigcup_{z \in \partial G} B(z, r)).$$

Remark 2.1. 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 (1.4), 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 (1.4) or in other results, respectively. Regarding the latter half of (A7), under assumption (A2) this condition with the Lipschitz continuity requirement on $\hat{a}$ replaced by just the continuity is weaker than (A4). This weaker assumption can be used in this paper in place of (A4) although we did not do so.

First of all we consider the problem (1.4). To be precise, we let $A(x)$, in the definition of H_{in}, be the subset of A consisting of those a such that there is $r > 0$ for which (2.1) and (2.2) hold with x, r, and a in place of z, r_z, and a_z.

Remark 2.2. $H_{in}(x, p, X) = H(x, p, X)$ if $x \in G$.

Theorem 2.3. *Assume* (A1)*,* (A2)*,* (A3)*, and* (A4)*. Then there exists a unique viscosity solution $U \in C(\overline{G})$ of the problem* (1.4)*.*

The theorem above extends the existence and uniqueness result in [5]. We refer the reader to [2] for results closely related to the above, in which the "strong" comparison principle has been established under a similar but more general assumption.

We now describe three kinds of control problems with state constraint whose value functions represent the viscosity solution of (1.4) under appropriate hypotheses.

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 (1.3) 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 $x \in \overline{G}$, i.e., $\alpha \in C^x$ is admissible if X_t^α satisfies

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

Second Control Problem : Relaxation. Let $M(A)$ denote the set of probability (Radon) measures on A. Define $s : \mathbf{R}^N \times M(A) \to \mathrm{SL}^N$ by

$$s(x,\mu) := \left[\frac{1}{2}\iint_{A\times A}(\sigma(x,a)-\sigma(x,a'))(\sigma(x,a)-\sigma(x,a'))^T\mu(da)\mu(da')\right]^{1/2}.$$

We set

$$\begin{aligned}
\hat{\sigma}(x,\mu) &:= \left(\int_A \sigma(x,a)\mu(da), s(x,\mu)\right) \in \mathbf{R}^N \otimes \mathbf{R}^{l+N},\\
\hat{b}(x,\mu) &:= \int_A b(x,a)\mu(da),\\
\hat{f}(x,\mu) &:= \int_A f(x,a)\mu(da)
\end{aligned}$$

for $(x,\mu) \in \mathbf{R}^N \times M(A)$.

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

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

where $(\Omega^\alpha, \mathcal{F}^\alpha, \{\mathcal{F}_t^\alpha\}_{t\geq 0}, P^\alpha)$ is a filtered probability space satisfying the usual condition, $\{W_t^\alpha\}_{t\geq 0}$ is a standard $(l+N)$–dimensional Brownian motion on this filtered space, $\{\mu_t^\alpha\}_{t\geq 0}$ is an $\{\mathcal{F}_t^\alpha\}_{t\geq 0}$–progressively measurable process taking values in $M(A)$, and $\{X_t^\alpha\}_{t\geq 0}$ is the unique strong solution of

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

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

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

Finally we define the (relaxed) value function V_r by

$$V_r(x) := \inf_{\alpha \in \widehat{\mathcal{A}}(x)} E^\alpha \int_0^\infty e^{-\lambda t} \hat{f}(X_t^\alpha, \mu_t^\alpha)\, dt \quad (x \in \overline{G}).$$

Third Control Problem: State constraint in G. We are as well interested in state constraint problems where the trajectories 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

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

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

$$V_0(x) = \inf_{\alpha \in \mathcal{A}_0(x)} J(x, \alpha) \tag{2.3}$$

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

Theorem 2.4. *Assume* (A1)*,* (A2)*,* (A3)*, and* (A4)*. Then:*

(i) *If either* (A5) *or* (A7) *is satisfied, then*

$$U(x) = V(x) \quad \textit{in} \quad \overline{G}.$$

(ii) *If* (A6) *is satisfied, then the unique viscosity solution U from problem* (1.4) *has the representation*

$$U(x) = V_r(x) \quad \textit{in} \quad \overline{G}.$$

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

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

In other words, under the assumptions above, the value functions V, V_r, and V_0 are viscosity solutions of (1.4) and therefore, by the uniqueness of viscosity solutions of (1.4), they are the same function.

In a recent work [20], we have shown that assertion (i) above holds without assuming any of (A5), (A7), and the convexity of A.

Theorem 2.5. *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* (1.4) *is Hölder continuous with exponent γ.*

Theorems 2.4 and 2.5 immediately yield Hölder estimates of the value functions V, V_0, and V_r under appropriate assumptions.

3. PROOF OF THE MAIN RESULTS

We begin with the preparations for the proof of Theorem 2.3.

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

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

such that ξ_0 is Lipschitz continuous on $\overline{G}$, and such that for some constant $r > 0$,

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

By an argument utilizing partition of unity, we see (see e.g. [5]) 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

$$\begin{cases} \lambda u(x) + H(x, Du(x), D^2u(x)) \ge 0 & (x \in \overline{G}), \\ \lambda u(x) + H_0(x, Du(x), D^2u(x)) \le 0 & (x \in \overline{G}), \end{cases} \tag{3.1}$$

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 $H_0(x,p,X) \le H_{in}(x,p,X)$ for all $(x,p,X) \in \overline{G} \times \mathbf{R}^N \times \mathrm{SL}^N$, it follows that any viscosity subsolution of (1.4) is a viscosity subsolution of (3.1).

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

For the proof of this theorem, we adapt the arguments from [5] to our case.

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

Theorem 3.2. *Assume* (A1), (A3), *and* (A4). *Let u and v be a viscosity subsolution and a viscosity supersolution of* (1.4), *respectively. Then $u \le v$ on $\overline{G}$.*

The following two lemmas are needed for our proof of Theorems 3.1 and 2.5.

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

$$\xi_0(x) \cdot D\psi(x) \ge 1 \quad (x \in \partial G).$$

For a proof of the lemma above see [5, Lemma 3.4].

Lemma 3.4. *Assume* (A1), (A2), (A3), *and* (A4). *Then there exist* $w \in C^{1,1}(\overline{G} \times \overline{G})$ *and constants* $C > 0$, $r > 0$ *such that*

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

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

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

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

This lemma is similar to [5, 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 [5, 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 [4]. However, we give the proof of Theorem 3.1 for the interested reader in the Appendix.

Proof of Theorem 2.3. The uniqueness of viscosity solutions of (1.4) 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

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

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

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

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 (1.4). Similarly, we see that g^{-} is a viscosity subsolution of (1.4).

Now, the standard Perron's method (see e.g. [4]) yields a viscosity solution U of (1.4) such that $U \in C(\overline{G})$ and $g^{-} \leq U \leq g^{+}$ on $\overline{G}$. □

For the proof of Theorem 2.4 we need the following three theorems.

Theorem 3.5. *Assume* (A1), (A2), (A3), *and* (A4). *Let* U *be the unique viscosity solution of* (1.4). *Then* $U(x) \leq V(x)$ *for all* $x \in \overline{G}$. *If in addition* (A5) *holds, then* $U = V$.

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

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

Here and henceforth we use the notation: $G^c = \mathbf{R}^N \setminus G$.

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

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

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

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

Recall here that A is convex, and thus the functions σ_ε, b_ε, and f_ε are well-defined. Note that the functions σ_ε and b_ε are Lipschitz continuous and the function f_ε is uniformly continuous on $\mathbf{R}^N \times A$.

As before we call any collection

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

a controlled system at $x \in \mathbf{R}^N$ associated with σ_ε and b_ε, if $(\Omega^\alpha, \mathcal{F}^\alpha, \{\mathcal{F}_t^\alpha\}_{t \geq 0}, P^\alpha)$ is a filtered probability space satisfying the usual condition, $\{W_t^\alpha\}_{t \geq 0}$ is a standard l-dimensional Brownian motion on this filtered probability space, $\{u_t^\alpha\}_{t \geq 0}$ is an $\{\mathcal{F}_t^\alpha\}_{t \geq 0}$– progressively measurable process taking values in A, and $\{X_t^\alpha\}_{t \geq 0}$ is the unique strong solution of the stochastic differential equation

$$dX_t = b_\varepsilon(X_t, u_t^\alpha)\, dt + \sigma_\varepsilon(X_t, u_t^\alpha) dW_t^\alpha, \quad X_0 = x. \tag{3.2}$$

For $x \in \mathbf{R}^N$, C_ε^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 $\alpha \in C_\varepsilon^x$ for $x \in \overline{G}$, i.e., those of $\alpha \in C_\varepsilon^x$ such that

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

The value functions U_ε are defined by

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

Theorem 3.6. *Assume* (A1), (A2), (A3), (A4), *and* (A7). *Then* $V(x) = U(x)$ *for all* $x \in \overline{G}$. *Moreover*

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

Theorem 3.7. *Assume that* $\partial G \in C^2$, (A1), (A2), (A3), (A4), *and* (A7). *Then*

$$U_\varepsilon(x) \geq V_0(x) \quad \text{for all } x \in \overline{G}.$$

Conceding Theorems 3.5, 3.6, and 3.7 for the moment, the proof of which will be given in Sections 4, 5, and 6, we complete the proof of Theorem 2.4.

Proof of Theorem 2.4. Assume (A1), (A2), (A3), and (A4). If, in addition, we assume (A5), then we see immediately from Theorem 3.5 that $U(x) = V(x)$ for all $x \in \overline{G}$.

Now assume that (A6) is satisfied. It is well-known that (A6) yields

$$\sup_{\mu \in M(A)} \|s(\cdot, \mu)\|_{W^{1,\infty}(\mathbf{R}^N)} < \infty.$$

Also, by (A2), s is uniformly continuous on $\mathbf{R}^N \times M(A)$. Therefore, conditions (A1) and (A2) are satisfied with σ, b, f replaced by $\hat{\sigma}$, $\hat{b}$, $\hat{f}$.

Observe by the definition of s that

$$s(x, \mu)^2 = \int_A \sigma\sigma^T(x, a)\mu(da) - \int_A \sigma(x, a)\mu(da)\left(\int_A \sigma(x, a')\mu(da')\right)^T.$$

Therefore,

$$\hat{\sigma}\hat{\sigma}^T = \int_A \sigma\mu(da) \int_A \sigma^T\mu(da) + s^2 = \int_A \sigma\sigma^T\mu(da),$$

and

$$\begin{aligned}\{(\hat{\sigma}\hat{\sigma}^T(x, \mu), \hat{b}(x, \mu), \hat{f}(x, \mu)) \mid \mu \in M(A)\}& \\ = \overline{\mathrm{co}}\,\{(\sigma\sigma^T(x, a), b(x, a), f(x, a)) \mid a \in A\}&.\end{aligned}$$

In particular, for any $x \in \mathbf{R}^N$ the set $\{(\hat\sigma\hat\sigma^T(x,\mu), \hat b(x,\mu), \hat f(x,\mu)) \mid \mu \in M(A)\}$ is a closed convex set in $\mathrm{SL}^N \times \mathbf{R}^N \times \mathbf{R}$, and

$$\max_{\mu\in M(A)} \left\{-\frac{1}{2}\operatorname{tr}\hat\sigma\hat\sigma^T(x,\mu)X - \hat b(x,\mu)\cdot p - \hat f(x,\mu)\right\} = H(x,p,X)$$

for all $(x,p,X) \in \mathbf{R}^N \times \mathbf{R}^N \times \mathrm{SL}^N$.

Recall that the set A can be regarded as a subset of $M(A)$ by identifying $a \in A$ with the Dirac measure $\delta_a \in M(A)$, and note that $s(x,\delta_a) = 0$ for all $x \in \mathbf{R}^N$, $a \in A$. It is then clear that (A4) is satisfied with σ and b replaced by $\hat\sigma$ and $\hat b$, respectively. If we define $\hat A(x) \subset M(A)$ in the same way as $A(x)$ but with A replaced by $M(A)$, then we have

$$\{\delta_a \mid a \in A(x)\} \subset \hat A(x) \quad \forall x \in \overline{G},$$

and hence

$$\sup_{\mu\in\hat A(x)} \left\{-\frac{1}{2}\operatorname{tr}\hat\sigma\hat\sigma^T(x,\mu)X - \hat b(x,\mu)\cdot p - \hat f(x,\mu)\right\} \geq H_{in}(x,p,X)$$

for all $(x,p,X) \in \mathbf{R}^N \times \mathbf{R}^N \times \mathrm{SL}^N$.

Thus we see from Theorem 3.5 (see the remark just before Proposition 4.1), that $U = V_r$ on $\overline{G}$.

Next assume that (A7) is satisfied. Let $\hat a$ be the Lipschitz function from (A7). Let $\varepsilon > 0$ and $x \in \overline{G}$. If

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

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, since A is convex, $v_t^\alpha \in A$ for all $t \geq 0$ and moreover,

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

Hence,

$$U_\varepsilon(x) \geq V(x) \quad \forall x \in \overline{G}.$$

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

$$U(x) \leq V(x) \leq U_\varepsilon(x) \quad \forall x \in \overline{G}.$$

Now Theorem 3.6 tells us that

$$V(x) = U(x) \quad \forall x \in \overline{G}.$$

In what follows we assume (A7) and that $\partial G \in C^2$. Theorem 3.7 guarantees that $U_\varepsilon(x) \ge V_0(x)$ for all $x \in \overline{G}$, $\varepsilon > 0$ and hence $U_\varepsilon(x) \ge V_0(x) \ge V(x)$ for all $x \in \overline{G}$, $\varepsilon > 0$. This together with the above considerations yields that $V_0(x) = U(x) = V(x)$ for all $x \in \overline{G}$. □

We continue to prove Theorem 2.5. Fix $y > 0$. We choose a function $(\xi_0, \eta_0) \in C^{0,1}(\overline{G})$ and a constant $r_0 > 0$ so that

$$(\xi_0(x), \eta_0(x)) \in \mathrm{co}\,\{(b(x,a), f(x,a)) \mid a \in A(x)\} \quad (x \in \partial G); \tag{3.3}$$

$$B(y + t\xi_0(x), tr_0) \subset \overline{G} \quad (x \in \partial G,\ 0 \le t \le r_0). \tag{3.4}$$

Let $\psi \in C^2(\overline{G})$ and $w \in C^{1,1}(\overline{G})$ be functions from Lemmas 3.3 and 3.4, respectively.

Recall that we showed in the proof of Theorem 2.3 that $|U(x)| \le M/\lambda$ for all $x \in \overline{G}$.

If we replace U by the function $\tilde{U} = U + (2M+1)\psi$, then we have

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

in the viscosity sense. To see that U is Hölder continuous on $\overline{G}$, it is enough to show that $\tilde{U}$ is Hölder continuous on $\overline{G}$. So, we may assume by replacing U by $\tilde{U}$ that U satisfies

$$-\xi_0(x) \cdot DU(x) \le -1 \quad (x \in \partial G)$$

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. Set

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

Note that if $x \neq y$,

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

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

$$\begin{aligned}
&\xi_0(x) \cdot Dv(x,y) \le 0 \quad \text{if } x \in \partial G \text{ and } |x-y| \le r,\\
&|x-y| \le v(x,y) \le C^{1/2}|x-y|,\\
&\max\{|D_x v(x,y)|, |D_y v(x,y)|\} \le C/2,\\
&|D_x v(x,y) + D_y v(x,y)| \le (C/2)|x-y|,\\
&D^2 v(x,y) \le (C/2)\left\{|x-y|^{-1}\begin{pmatrix} I & -I \\ -I & I\end{pmatrix} + |x-y|\begin{pmatrix} I & 0 \\ 0 & I\end{pmatrix}\right\}.
\end{aligned}$$

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

Let $L > 0$ and $\gamma \in (0,1]$, 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 $\hat{x} \neq \hat{y}$.

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

$$\xi_0(\hat{x}) \cdot D_x v(\hat{x},\hat{y}) \le 0 \quad \text{and} \quad -\xi_0(\hat{x}) \cdot D_x v(\hat{x},\hat{y}) \le -1.$$

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

$$\gamma v(\hat{x},\hat{y})^{\gamma-1}\left(Dv(\hat{x},\hat{y}), K\left(|\hat{x}-\hat{y}|^{-1}\begin{pmatrix} I & -I \\ -I & I\end{pmatrix} + |\hat{x}-\hat{y}|\begin{pmatrix} I & 0 \\ 0 & I\end{pmatrix}\right)\right) \in J^{2,+}v^\gamma(\hat{x},\hat{y}),$$

there are matrices $X, Y \in \mathrm{SL}^N$ such that

$$\begin{gathered}
(L\gamma v(\hat{x},\hat{y})^{\gamma-1} D_x v(\hat{x},\hat{y}), X) \in \overline{J}^{2,+}U(\hat{x}),\\
(-L\gamma v(\hat{x},\hat{y})^{\gamma-1} D_y v(\hat{x},\hat{y}), -Y) \in \overline{J}^{2,-}U(\hat{y}),\\
\begin{pmatrix} X & 0 \\ 0 & Y\end{pmatrix} \le LK\gamma v(\hat{x},\hat{y})^{\gamma-1}\left(|\hat{x}-\hat{y}|^{-1}\begin{pmatrix} I & -I \\ -I & I\end{pmatrix} + |\hat{x}-\hat{y}|\begin{pmatrix} I & 0 \\ 0 & I\end{pmatrix}\right).
\end{gathered}$$

Since U is a solution of (3.1) and $\hat{x} \in G$, we have

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

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

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

$$\begin{aligned}
&\begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix} \begin{pmatrix} \sigma_i(\hat{x}) \\ \sigma_i(\hat{y}) \end{pmatrix} \cdot \begin{pmatrix} \sigma_i(\hat{x}) \\ \sigma_i(\hat{y}) \end{pmatrix} \\
&\leq KL\gamma v(\hat{x}, \hat{y})^{\gamma-1} |\hat{x} - \hat{y}|^{-1} |\sigma_i(\hat{x}) - \sigma_i(\hat{y})|^2 \\
&\qquad + KL\gamma v(\hat{x}, \hat{y})^{\gamma-1} |\hat{x} - \hat{y}| (|\sigma_i(\hat{x})|^2 + |\sigma_i(\hat{y})|^2) \\
&\leq 3KLM^2 \gamma v(\hat{x}, \hat{y})^{\gamma-1} |\hat{x} - \hat{y}|.
\end{aligned}$$

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

$$\operatorname{tr} \sigma\sigma^T(\hat{x}, a) X + \operatorname{tr} \sigma\sigma^T(\hat{y}, a) Y \leq 3NKLM^2 \gamma v(\hat{x}, \hat{y})^{\gamma-1} |\hat{x} - \hat{y}|.$$

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

$$\begin{aligned}
0 &\geq \lambda(U(\hat{x}) - U(\hat{y})) + H(\hat{x}, \gamma v(\hat{x}, \hat{y})^{\gamma-1} D_x v(\hat{x}, \hat{y}), X) \\
&\qquad - H(\hat{y}, -\gamma v(\hat{x}, \hat{y})^{\gamma-1} D_y v(\hat{x}, \hat{y}), -Y) \\
&> \lambda L v(\hat{x}, \hat{y})^\gamma + \inf_{a \in A} \{-\tfrac{1}{2} \operatorname{tr} \sigma\sigma^T(\hat{x}, a) X - \tfrac{1}{2} \operatorname{tr} \sigma\sigma^T(\hat{y}, a) Y \\
&\qquad - \gamma v(\hat{x}, \hat{y})^{\gamma-1} b(\hat{x}, a) \cdot D_x v(\hat{x}, \hat{y}) - \gamma v(\hat{x}, \hat{y})^{\gamma-1} b(\hat{y}, a) \cdot D_y v(\hat{x}, \hat{y}) \\
&\qquad - f(\hat{x}, a) + f(\hat{y}, a)\} \\
&\geq \lambda L |\hat{x} - \hat{y}|^\gamma - \tfrac{3}{2} \gamma NKLM^2 |\hat{x} - \hat{y}|^\gamma - 2\gamma MKL |\hat{x} - \hat{y}|^\gamma - Mr^{1-\gamma} |\hat{x} - \hat{y}|^\gamma \\
&= (\lambda L - \tfrac{3}{2} \gamma NKLM^2 - 2\gamma MKL - Mr^{1-\gamma}) |\hat{x} - \hat{y}|^\gamma.
\end{aligned}$$

If we set $k = \frac{3}{2} NKM^2 + 2MK$ and assume that $\lambda > k\gamma$, then by choosing L large enough we have

$$\lambda L - \tfrac{3}{2} \gamma NKLM^2 - 2\gamma MKL - Mr^{1-\gamma} > 0,$$

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

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

This inequality yields

$$|U(x) - U(y)| \le K^\gamma L|x-y|^\gamma \quad (x, y \in \overline{G}),$$

proving the Hölder continuity of U under the assumption that $\lambda > k\gamma$. □

4. PROOF OF THEOREM 3.5

We divide the first part of the proof of Theorem 3.5 into two propositions.

Define

$$V_n(x) = \inf_{\alpha \in C^x} E^\alpha \int_0^\infty e^{-\lambda t}[f(X_t^\alpha, u_t^\alpha) + nd(X_t^\alpha)]\, dt \quad (x \in \mathbf{R}^N), \tag{4.1}$$

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

Recalling that $A \subset \mathbf{R}^m$ is compact and convex, we observe by [15], [16], [17] that V_n is a solution of

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

We remark at this point that a careful review of [15], [16], [17] and finite dimensional approximation techniques allow us to conclude that the assertion above is still valid when A, σ, b, f, etc are replaced by $M(A)$, $\hat{\sigma}$, $\hat{b}$, $\hat{f}$, etcetera.

Then we have:

Proposition 4.1. *Under the assumptions* (A1), (A2), (A3), *and* (A4), V_n *is a subsolution of* (1.4).

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 $V_n - \varphi$ has a maximum over $\overline{G}$ at z, then

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

Fix $z \in \partial G$, $a \in A(z)$. Then there exists $r > 0$ such that (2.1) and (2.2) hold with these z, a, and r. Set $W = G \cap \operatorname{Int} B(z,r)$. Choose a C^1 function ζ on $\mathbf{R}^N$ so that

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

Note that if we set

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

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

then V_n satisfies

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

in the viscosity sense.

We now invoke [6, Theorem 2.1] (more precisely, its proof). It is easy to check that (2.2) and (MP) of [6], with $K = \overline{W}$ and with

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

are satisfied. Therefore we see that V_n is a viscosity subsolution of

$$\lambda V_n(x) - \frac{1}{2} \operatorname{tr} \widetilde{\sigma}\widetilde{\sigma}^T(x) D^2 V_n(x) - \widetilde{b}(x) \cdot DV_n(x) - \widetilde{f}(x) \le 0 \quad \text{in } \overline{W}.$$

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

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

which completes the proof. □

Let $U \in C(\overline{G})$ be the solution of (1.4). By comparison we have

$$V_n(x) \le U(x) \quad (x \in \overline{G},\ n \in \mathbf{N}).$$

By the definition of V_n, we see that

$$V_n(x) \le V_{n+1}(x) \quad (x \in \overline{G},\ n \in \mathbf{N}).$$

Set

$$V_+(x) = \sup_n V_n(x) \quad (x \in \overline{G}).$$

Theorem 4.2. *Under the assumptions* (A1), (A2), (A3), (A4), *we have*

$$V_+ = U.$$

Proof. Since V_+ is a pointwise supremum of viscosity subsolutions of (1.4), its upper semicontinuous envelope is a viscosity subsolution of (1.4).

We now check that V_+ is a supersolution of (1.4). Indeed, if we set

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

then, since V_n are supersolutions of

$$\lambda u(x) + H(x, Du(x), D^2 u(x)) \geq 0 \quad \text{in } \mathbf{R}^N, \tag{4.2}$$

we see that W is a supersolution of (4.2) in the sense that if $\varphi \in C^2(\mathbf{R}^N)$, $z \in R^N$, $W(z) < \infty$, 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 (1.4), it is enough to show that

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

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 $\zeta(x) = 1 - (2/R^2)|x - z|^2$.) We fix an upper bound $K > 0$ of

$$\sup\{-\tfrac{1}{2} \operatorname{tr} \sigma\sigma(x,a)^T D^2\zeta(x) - b(x,a) \cdot D\zeta(x)\} + \lambda\zeta(x)$$

over $B(z, R)$.

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

$$\Phi_n(x) = \frac{Rn}{2K}\zeta(x).$$

By a simple computation we get

$$\begin{aligned} &\lambda\Phi_n(x) + H(x, D\Phi_n(x), D^2\Phi_n(x)) - nd(x) \\ &\qquad \leq \frac{R}{2}n + \sup_{a \in A} -f(x,a) - nR \leq 0 \quad (x \in B(z,R)) \end{aligned}$$

if n is large enough, and, since $\zeta \leq -1$ on $\partial B(z, R)$, we get

$$V_1(x) \geq \Phi_n(x) \quad (x \in \partial B(z, R))$$

if n is large enough. Hence by comparison between V_n and Φ_n, we obtain

$$W(x) \geq \Phi_n(x)$$

in $B(z, R)$ for sufficiently large n and we conclude sending $n \to \infty$.

Therefore, we find that $V_+ = U$ and finish the proof. ◻

Completion of the proof of Theorem 3.5. Assume (A1), (A2), (A3), and (A4). Fix $x \in \overline{G}$. Since

$$J(x,\alpha) = E^{\alpha} \int_0^{\infty} e^{-\lambda t}[f(X_t^{\alpha}, u_t^{\alpha}) + nd(X_t^{\alpha})]\, dt$$

for all $\alpha \in \mathcal{A}(x)$, it is immediate to see that $V_n(x) \le V(x)$ and hence $V_+(x) \le V(x)$. By using Theorem 4.2, we get $U(x) \le V(x)$ for all $x \in \overline{G}$.

We now assume (A5) as well. Let

$$\alpha_n \equiv (\Omega^n, \mathcal{F}^n, \{\mathcal{F}_t^n\}_{t\ge0}, P^n, \{W_t^n\}_{t\ge0}, \{u_t^n\}_{t\ge0}, \{X_t^n\}_{t\ge0})$$

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 the argument of the proof of Theorem 5.3 of [19], we can find a controlled system $\alpha \in C^x$ 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 (A5) in the above assertion. We see from the latter identity above that

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

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

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

Thus we see that $V(x) \le V_+(x)$ and hence $V(x) \le U(x)$ for all $x \in \overline{G}$. Hence we have $U = V$. $\square$

5. Proof of Theorem 3.6

We begin by defining

$$W_\varepsilon(x) = \inf_{\alpha\in C_\varepsilon^x} E^{\alpha} \int_0^{\infty} e^{-\lambda t} f_\varepsilon(X_t^{\alpha}, u_t^{\alpha})\, dt \quad (x \in \mathbf{R}^N).$$

As already explained, we have

$$U_\varepsilon(x) \geq V(x) \quad (x \in \overline{G}). \tag{5.1}$$

Since A is compact and convex, by a classical result (e.g. [15], [16], [17]), we know that $u := W_\varepsilon$ satisfies

$$\lambda u(x) + H_\varepsilon(x, Du(x), D^2u(x)) = 0 \quad (x \in \mathbf{R}^N)$$

in the viscosity sense, where $H_\varepsilon : \mathbf{R}^N \times \mathbf{R}^N \times \mathrm{SL}^N \to \mathbf{R}$ is given by

$$H_\varepsilon(x, p, X) = \max_{a \in A}\{-\frac{1}{2}\operatorname{tr}\sigma_\varepsilon\sigma_\varepsilon^T(x,a)X - b_\varepsilon(x,a)\cdot p - f_\varepsilon(x,a)\}.$$

Since $F := H_\varepsilon$ satisfies condition (MP) of [6] with $K = \overline{G}$, we see from ([6, Cor. 2.3]) that $u := W_\varepsilon$ is a viscosity solution of

$$\lambda u(x) + H_\varepsilon(x, Du(x), D^2u(x)) = 0 \quad (x \in \overline{G})$$

and that if $x \in \overline{G}$ and $\alpha \in C_\varepsilon^x$, then

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

i.e., $C_\varepsilon^x = \mathcal{A}_\varepsilon(x)$ for all $x \in \overline{G}$. This shows that

$$U_\varepsilon(x) = W_\varepsilon(x) \quad (x \in \overline{G}).$$

We see that M/λ and $-M/\lambda$ are super- and subsolutions of (1.4) and therefore by comparison that $|W_\varepsilon(x)| \leq M/\lambda$ for $x \in \overline{G}$. If we define

$$U^*(x) := \lim_{r\searrow 0}\sup\{U_\varepsilon(y) \mid 0 < \varepsilon < r,\ y \in \overline{G},\ |y - x| < r\},$$

and

$$U_*(x) := \lim_{r\searrow 0}\inf\{U_\varepsilon(y) \mid 0 < \varepsilon < r,\ y \in \overline{G},\ |y - x| < r\},$$

then U^* and U_* are upper and lower semicontinuous on $\overline{G}$, respectively, and $u := U^*$ and $v := U_*$ are sub- and supersolutions of the problem

$$\lambda u(x) + H_*(x, Du(x), D^2u(x)) = 0 \quad (x \in \overline{G}),$$
$$\lambda v(x) + H^*(x, Dv(x), D^2v(x)) = 0 \quad (x \in \overline{G}),$$

respectively, where for $(x, p, X) \in \overline{G} \times \mathbf{R}^N \times \mathrm{SL}^N$,

$$H_*(x, p, X) = \lim_{r \searrow 0} \inf\{H_\varepsilon(y, q, Y) \mid (y, q, Y) \in \overline{G} \times \mathbf{R}^N \times \mathrm{SL}^N,$$

$$|y - x| + |q - p| + |Y - X| < r\};$$

$$H^*(x, p, X) = \lim_{r \searrow 0} \sup\{H_\varepsilon(y, q, Y) \mid (y, q, Y) \in \overline{G} \times \mathbf{R}^N \times \mathrm{SL}^N,$$

$$|y - x| + |q - p| + |Y - X| < r\}.$$

Note that for all $(x, p, X) \in G \times \mathbf{R}^N \times \mathrm{SL}^N$

$$H^*(x, p, X) = H_*(x, p, X) = H(x, p, X);$$

for all $(x, p, X) \in \partial G \times \mathbf{R}^N \times \mathrm{SL}^N$

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

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

The comparison result, Theorem 3.1, applied to the problem

$$\lambda u(x) + H_*(x, Du(x), D^2u(x)) \le 0 \quad (x \in \overline{G}), \tag{5.2}$$

$$\lambda u(x) + H(x, Du(x), D^2u(x)) \ge 0 \quad (x \in \overline{G}), \tag{5.3}$$

guarantees that $U^* \le U_*$ in $\overline{G}$ and so $U^* = U_*$ in $\overline{G}$. Since U is a viscosity solution of (5.2), we conclude by the same comparison result that $U = U^* = U_*$ on $\overline{G}$. This immediately implies the uniform convergence of $U_\varepsilon(x)$ to $U(x)$ for all $x \in \overline{G}$. □

6. Proof of Theorem 3.7

Let $\varepsilon > 0$. We utilize the functions U_ε, σ_ε, etc defined in the formulation of Theorem 3.6.

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

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

Note that we may choose the sequence of ζ_n so that

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

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

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

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 $(x,t) \in \mathbf{R}^{N+1}$. In what follows we assume that n is large enough so that $-\log(\mu/2) \le n - 1$. 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

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

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

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

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

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

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

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

We thus have

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

Case 3. Now consider the case where $d(x)+\delta\rho(t) > \mu/2$. Since $-\log(\mu/2) \le n-1$,

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

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

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

for some constant $C>0$ independent of n.

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

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

Let $z\in\overline{G}$ and

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

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

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

From this, we get

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

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

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

Hence, if n is large enough, we have

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

If n is large enough, then

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

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

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

Now the monotone convergence theorem implies that

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

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

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

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, since A is convex, $v_t^\alpha \in A$ and moreover,

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

Therefore, if we define $\mathcal{A}_0(x) \subset \mathcal{A}(x)$ for each $x \in \overline{G}$ as the set of those $\alpha \in \mathcal{A}(x)$ for which

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

and

$$V_0(x) := \inf_{\alpha\in\mathcal{A}_0(x)} E^\alpha \int_0^\infty e^{-\lambda t} f(X_t^\alpha, u_t^\alpha)\, dt,$$

then

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

Thus in the limit as $\varepsilon \searrow 0$, using Theorem 3.6, we get:

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

Appendix

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 [5, Lemma 3.4] yields a function having the desired properties in our lemma. Unfortunately, the assertion of [5, Lemma 3.4] is apparently different and weaker than our Lemma 3.4. The differences are that the function w in [5, Lemma 3.4] has only the C^1 regularity as it stated and that the last inequality

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

is not in [5, Lemma 3.4].

We explain how to show these properties of w, the function constructed in the proof of [5, Lemma 3.4].

The function w constructed in the proof of [5, 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

$$v(tx) = tv(x) \quad (x \in \mathbf{R}^2,\ t \ge 0),$$

and the property

$$D_{x_2} v(x_1, 0) = 0 \quad (x_1 \in \mathbf{R}).$$

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

$$f(x) = v(x)^2 \quad (x \in \mathbf{R}^2);$$
$$g(x,y) = f(x,|y|) \quad ((x,y) \in \mathbf{R} \times \mathbf{R}^N),$$

and observe that $f \in C^{1,1}(\mathbf{R}^2)$ and that for all $x, y \in \mathbf{R}^N$,

$$w(x,y) = g((x-y)\cdot \xi(y), Q(\xi(y))(x-y)), \tag{A.2}$$

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 $y \in \mathbf{R}^N \setminus \{0\}$,

$$D_x g(x,y) = D_{x_1} f(x,|y|), \qquad D_y g(x,y) = D_{x_2} f(x,|y|)\tfrac{y}{|y|},$$
$$D_x^2 g(x,y) = D_{x_1}^2 f(x,|y|), \quad D_y D_x g(x,y) = D_{x_2} D_{x_1} f(x,|y|)\tfrac{y}{|y|},$$

$$D_y^2 g(x,y) = D_{x_2} f(x,|y|)\frac{y\otimes y}{|y|^2} + D_x(x,|y|)\frac{I}{|y|} - D_{x_2} f(x,|y|)\frac{y \otimes y}{|y|^3}.$$

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.2) shows that $w \in C^{1,1}(\mathbf{R}^{2N})$.

Next we set

$$h(\xi, p) = g(\xi \cdot p, Q(\xi)p) \quad (\xi, p \in \mathbf{R}^N).$$

With an approximation argument in mind we may assume that $h \in C^2(\mathbf{R}^{2N})$. Note that

$$w(x,y) = h(\xi(y), x-y) \quad (x,y \in \mathbf{R}^N),$$

and that

$$h(\xi, tp) = t^2 h(\xi, p) \quad (\xi, p \in \mathbf{R}^N,\ t \ge 0).$$

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 $p \in \mathbf{R}^N$,

$$\max\{h(\xi,p), |D_\xi h(\xi,p)|, \|D_\xi^2 h(\xi,p)\|\} \le C_B |p|^2,$$
$$\max\{|D_p h(\xi,p)|, \|D_p D_\xi h(\xi,p)\|\} \le C_B |p|, \quad \|D_p^2 h(\xi,p)\| \le C_B.$$

Compute that

$$\begin{aligned} D_x^2 w(x,y) &= D_p^2 h(\xi(y),\, x-y), \\ D_y D_x w(x,y) &= (D\xi(y))^T D_\xi D_p h(\xi(y),\, x-y) - D_p^2 h(\xi(y),\, x-y), \\ D_y^2 w(x,y) &= (D\xi(y))^T D_\xi^2 h(\xi(y),\, x-y) D\xi(y) + D_p^2 h(\xi(y),\, x-y) \\ &\quad - (D_\xi D_p h(\xi(y),\, x-y)) D\xi(y) \\ &\quad - (D\xi(y))^T D_p D_\xi h(\xi(y),\, x-y), \end{aligned}$$

to obtain

$$D^2 w(x,y) = \begin{pmatrix} A_1 & -A_1 \\ -A_1 & A_1 \end{pmatrix} + \begin{pmatrix} 0 & A_2 \\ A_2^T & -A_2 - A_2^T \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & A_3 \end{pmatrix}, \tag{A.3}$$

where

$$\begin{aligned} A_1 &= D_p^2 h(\xi(y),\, x-y), \\ A_2 &= (D_\xi D_p h(\xi(y),\, x-y)) D\xi(y), \\ A_3 &= (D\xi(y))^T D_\xi^2 h(\xi(y),\, x-y) D\xi(y). \end{aligned}$$

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.,

$$\begin{pmatrix} 0 & A_2 \\ A_2^T & -A_2 - A_2^T \end{pmatrix} \le C_B \left(\begin{pmatrix} I & -I \\ -I & I \end{pmatrix} + |x-y|^2 \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \right).$$

This way we see from (A.3) that for each bounded $B \subset \mathbf{R}^N$, there is a constant $C_B > 0$ for which

$$D^2 w(x,y) \le C_B \left(\begin{pmatrix} I & -I \\ -I & I \end{pmatrix} + |x-y|^2 \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \right) \quad (x \in \mathbf{R}^N,\ y \in B). \qquad \square$$

Proof of Theorem 3.1. We argue by contradiction, and hence assume that

$$\max_{x\in\overline{G}}(u(x)-v(x))\geq\theta$$

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

$$-\xi_0(x)\cdot Du(x)\leq-1\quad(x\in\partial G)$$

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 $|\hat{x}-\hat{y}|\leq r$ and

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

note that $\Phi(\hat{x},\hat{y})\geq\theta$. By choosing L large enough we may assume that

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

and hence that $|\hat{x}-\hat{y}|<r$.

Now suppose for the moment that $\hat{x}\in\partial G$, which immediately yields that

$$\xi_0(\hat{x})\cdot D_xw(\hat{x},\hat{y})\leq 0\quad\text{and}\quad-\xi_0(\hat{x})\cdot D_xw(\hat{x},\hat{y})\leq-1.$$

This is a contradiction, which shows that $\hat{x}\in G$. Since

$$\left(D_xw(\hat{x},\hat{y}),D_yw(\hat{x},\hat{y}),C\left(\begin{pmatrix}I&-I\\-I&I\end{pmatrix}+|\hat{x}-\hat{y}|^2\begin{pmatrix}I&0\\0&I\end{pmatrix}\right)\right)\in J^{2,+}w(\hat{x},\hat{y}),$$

there are matrices $X,Y\in\mathrm{SL}^N$ such that

$$(LD_xw(\hat{x},\hat{y}),X)\in\overline{J}^{2,+}u(\hat{x}),\quad(-LD_yw(\hat{x},\hat{y}),-Y)\in\overline{J}^{2,-}v(\hat{y}),$$
$$\begin{pmatrix}X&0\\0&Y\end{pmatrix}\leq LC\left(\begin{pmatrix}I&-I\\-I&I\end{pmatrix}+|\hat{x}-\hat{y}|^2\begin{pmatrix}I&0\\0&I\end{pmatrix}\right).$$

Since u and v are a subsolution and a supersolution of (3.1) 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 2.5 yield

$$\operatorname{tr} \sigma\sigma^T(\hat{x}, a)X + \operatorname{tr} \sigma\sigma^T(\hat{y}, a)Y \le 4NCLM^2|\hat{x} - \hat{y}|^2,$$
$$b(\hat{x}, a) \cdot D_x w(\hat{x}, \hat{y}) + b(\hat{y}, a) \cdot D_y w(\hat{x}, \hat{y}) \le 2MC|\hat{x} - \hat{y}|^2.$$

Thus we obtain

$$\begin{aligned} 0 &\ge \lambda(u(\hat{x}) - v(\hat{y})) + H(\hat{x}, D_x w(\hat{x}, \hat{y}), X) - H(\hat{y}, D_y w(\hat{x}, \hat{y}), -Y) \\ &\ge \lambda\theta - 2NCLM^2|\hat{x} - \hat{y}|^2 - 2MC|\hat{x} - \hat{y}|^2 - M|\hat{x} - \hat{y}|. \end{aligned}$$

Now, noting (see e.g. [4]) 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 $L \to \infty$. □

Acknowledgements. The first author was supported by Grant-in-Aid for Scientific Research (No. 09440067 and No. 1244044) 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 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 to these institutions for their support and 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

[1] 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.

[2] 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.

[3] I. Capuzzo–Dolcetta and P. L. Lions, *Viscosity solutions of Hamilton–Jacobi equations and state constraints*, Trans. Amer. Math. Soc. **277** (1984), 487–502.

[4] 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.

[5] 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.

[6] H. ISHII, P. LORETI, AND E. TESSITORE, *A PDE approach to stochastic invariance*, Discrete and continuous dynamical systems **6** (2000), no. 3, 651–664.

[7] U. G. HAUSSMANN, *Existence of optimal Markovian controls for degenerate diffusions*, Lecture Notes in Control and Information Sciences Vol. *78*, Springer-Verlag, Berlin, New York, 1986, pp. 171–186.

[8] U. G. HAUSSMANN AND J. P. LEPELTIER, *On the existence of optimal controls*, SIAM J. Control Optim. **28** (1990), no. 4, 851–902.

[9] 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.

[10] 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.

[11] 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.

[12] P.-L. LIONS, *Generalized solutions of Hamilton-Jacobi equations*, Research Notes in Mathematics, Vol. **69**, Pitman (Advanced Publishing Program), Boston, Mass.-London, 1982.

[13] P. LORETI, *Some properties of constrained viscosity solutions of Hamilton-Jacobi-Bellman equations*, SIAM J. Control Optim. **25** (1987), no. 5, 1244–1252.

[14] 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.

[15] H. NAGAI, *Stochastic differential equations*, in Japanese, Kyoritu-Shuppan, Mathematics in 21st Century, Vol. **27**, 1999.

[16] 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., Vol. **550**, Springer, Berlin, 1975.

[17] M. NISIO, *Lectures on stochastic control theory*, ISI Lecture Notes, **9**, MacMillan India Ltd, Delhi, Bombay, 1981.

[18] H. M. Soner, Optimal control with state-space constraint. I. SIAM J. Control Optim. **24** (1986), no. 3, 552–561.

[19] J. YONG AND X. Y. ZHOU, *Stochastic controls: Hamiltonian and HJB equations*, Applications of Mathematics **43**, Springer-Verlag, 1999, New York.

[20] H. ISHII AND P. LORETI, *A class of stochastic optimal control problems with state constraint: an extended version*. Preprint. Available from http://www.edu.waseda.ac.jp/~ishii.

HITOSHI ISHII:
Department of Mathematics
School of Education
Waseda University
1-6-1 Nishi-Waseda, Shinjuku-ku
Tokyo 169-8050, Japan
E-MAIL: `ishii@edu.waseda.ac.jp`

PAOLA LORETI:
Dipartimento di Metodi e Modelli Matematici
per le Scienze Applicate
Università di Roma "La Sapienza"
Via Scarpa n. 16
00161 Roma, Italy
E-MAIL: `loreti@dmmm.uniroma1.it`

KEY WORDS AND PHRASES: stochastic optimal control, state constraint, Hamilton-Jacobi-Bellman equations, viscosity solutions, degenerate elliptic equations.

2000 MATHEMATICS SUBJECT CLASSIFICATION: 93E20, 70H20, 35E70, 49L25.
Received: *July 14th, 2000; revised: October 27th, 2001.*