The Annals of Probability
2014, Vol. 42, No. 5, 2161–2196
DOI: 10.1214/13-AOP848

ON REGULARITY PROPERTIES AND APPROXIMATIONS OF VALUE FUNCTIONS FOR STOCHASTIC DIFFERENTIAL GAMES IN DOMAINS

BY N. V. KRYLOV[1]

University of Minnesota

We prove that for any constant $K \geq 1$, the value functions for time homogeneous stochastic differential games in the whole space can be approximated up to a constant over K by value functions whose second-order derivatives are *bounded* by a constant times K.

On the way of proving this result we prove that the value functions for stochastic differential games in domains and in the whole space admit estimates of their Lipschitz constants in a variety of settings.

1. Introduction. In this paper we prove that for any constant $K \geq 1$, the value functions for time homogeneous stochastic differential games in the whole space can be approximated up to a constant over K by the value functions whose second-order derivatives are *bounded* by a constant times K (see Theorem 2.4 and Remark 2.4). To prove Theorem 2.4 we needed a few auxiliary facts organized in [12] and [10], so that the goal to prove this theorem was the major driving force of the series of three articles. Along the way some fruitful ideas were developed, leading, in particular, first to understanding from probabilistic point of view and then to proving in purely PDE terms the fact that one can find in $C^{1+\alpha}$ viscosity solutions of the uniformly nondegenerate Isaacs parabolic equations with coefficients measurable in time and VMO in x; see [11]. It would be extremely interesting to find a proof of this fact based on the theory of viscosity solutions in the situation of discontinuous coefficients, although in the case of continuous ones such a proof was given by Święch [15].

In terms of the corresponding Isaacs equations the approximation in Theorem 2.4 is done in such a way that the equations are modified only for large values of the derivatives of the value functions. Such approximation of stochastic games can be useful while evaluating the value functions numerically because one can expect that approximations might be more accurate if the approximating function is more regular.

Received August 2012; revised March 2013.

[1]Supported in part by NSF Grant DMS-1160569.

MSC2010 subject classifications. Primary 35J60, 49N70; secondary 91A15.

Key words and phrases. Stochastic differential games, smoothness of value functions, Isaacs equation.

Two main tools are used. One is the stochastic dynamic principle with randomized stopping times, and another is based on estimates of the Lipschitz constants of the value functions.

The dynamic programming principle we use is proved in [10] and originated in the work by Fleming and Souganidis [3]; see also Kovats [5] and Święch [14].

Here we concentrate on proving the Lipschitz continuity of the value functions for time homogeneous stochastic differential games in domains and in the whole space and on proving the above mentioned approximation result, which is a particular case of a conjecture from [9].

There is an enormous literature treating smoothness properties for *controlled* diffusion processes or, from analytical point of view, for fully nonlinear equations under convexity assumptions. We are going to focus only on stochastic differential games for which there is not much known concerning the regularity of the value function in more or less general case.

Ishii and Lions in [4] prove the Lipschitz continuity for viscosity solutions of fully nonlinear uniformly nondegenerate equations. Earlier Trudinger in [16] proved that the first derivatives are, actually, Hölder continuous. The same result under somewhat more restrictive assumptions can be found in the book [2] by Caffarelli and Cabré. Further results on Lipschitz continuity, still for uniformly nondegenerate case, are contained and referred to in Święch [15], Vitolo [17] and Krylov [11].

We deal with global and local estimates only for the Isaacs equations in contrast with the more general equations in the above mentioned references, which reduce to the Isaacs equations only if the equation is determined by the so-called boundedly inhomogeneous functions. Our methods are also different from the methods of the above cited articles where the authors rely on the theory of viscosity solutions. Our solutions are given as value functions of stochastic differential games, and we use probabilistic methods, with the main tool being based on different probabilistic representations for the value functions at different points. This is very close to using the so-called quasiderivatives of solutions of stochastic equations in the theory of controlled diffusion processes, which can be traced down starting from [7]. We could also use quasiderivatives in this article, but this would require more work, and what we are actually using can be called the method of quasidifferences. In the author's opinion the methods of this article can be also applied to proving interior first derivatives estimates for degenerate equations similar to those in [18] when the boundary data are only Lipschitz continuous and processes are not uniformly nondegenerate. Just in case, observe that there are no global gradient estimates even for the equation $\Delta u = 0$ in a ball if the boundary data are only Lipschitz continuous.

Even though our stochastic differential games are assumed to be uniformly nondegenerate, one of our main results, Theorem 2.3, is about estimates of the Lipschitz constant *independent* of the constant of nondegeneracy. The author is not aware of any analytical proof of it. The only results similar to the one mentioned above that the author is aware of are contained in Barles [1]. We discuss them in detail in Remark 3.5.

We also prove two estimates which do depend on the constant of nondegeneracy: one is global, Theorem 2.1, and another is local, Theorem 2.2. These results are much weaker than the ones in [16]. The emphasis here is to show that probabilistic methods can use nondegeneracy in an efficient way. Of course, Theorem 2.3 contains Theorem 2.1, the proof of the latter is given just because it is short, instructive and requires less machinery.

The main results of the paper are stated in Section 2. Section 3 contains their discussion continued in Section 4 where we describe some ideas behind our arguments. In Section 5 we show that the value function admits many representations. In Section 6 we prove auxiliary results aimed at estimating the difference of value function at close points when different probabilistic representations are taken for those points. The result of Section 5, in a very rough form, is used in Section 7 to prove Theorem 2.1. In Section 8 we prove Theorem 2.2 about interior estimates. A very short Section 9 contains the proof of Theorem 2.3 about estimates independent of the constant of nondegeneracy. It is short because the main ideas have already been given in Section 5. In the final, and again short, Section 10 we prove Theorem 2.4.

The author is very grateful to the referees for their comments which helped improve the presentation of this paper.

2. Main results. Let $\mathbb{R}^d = \{x = (x_1, \dots, x_d)\}$ be a d-dimensional Euclidean space, and let $d_1 \geq d$ be an integer. Denote by $\mathcal{O}$ the set of $d_1 \times d_1$ orthogonal matrices, fix an integer $k \geq 1$ and assume that we are given separable metric spaces A and B and let, for each $\alpha \in A$, $\beta \in B$ and $p \in \mathbb{R}^k$, the following functions on $\mathbb{R}^k \times \mathbb{R}^d$ be given:

(i) $d \times d_1$ matrix-valued $\sigma^{\alpha\beta}(p, x) = (\sigma_{ij}^{\alpha\beta}(p, x))$;

(ii) $\mathcal{O}$-valued function $P^{\alpha\beta}(x, y)$, $\mathbb{R}^k$-valued function $p^{\alpha\beta}(x, y)$ and real-valued function $r^{\alpha\beta}(x, y)$;

(iii) $\mathbb{R}^d$-valued $b^{\alpha\beta}(p, x) = (b_i^{\alpha\beta}(p, x))$;

(iv) real-valued functions $c^{\alpha\beta}(p, x) \geq 0$, $f^{\alpha\beta}(p, x)$ and $g(x)$.

Define

$$a^{\alpha\beta}(p, x) := (1/2)\sigma^{\alpha\beta}(p, x)\bigl(\sigma^{\alpha\beta}(p, x)\bigr)^*.$$

Also set

$$(\sigma, a, b, c, f)^{\alpha\beta}(x) = (\sigma, a, b, c, f)^{\alpha\beta}(0, x),$$

and note that for our first main result, Theorem 2.1, only these values of σ, a, b, c, f are relevant, and the parameters r, p, P are not present. These parameters are important in Theorem 2.3. The role of these parameters is discussed in

Remark 3.1 and Example 4.1 concerning P, in Remarks 3.4, 3.6 and Example 4.2 concerning r and in Remark 3.6 concerning p.

Fix some constants $K_0, K_1 \in [0, \infty)$, and $\delta_0 \in (0, 1]$.

ASSUMPTION 2.1. (i) The functions $(\sigma, a, b, c, f)^{\alpha\beta}(p, x)$ and $p^{\alpha\beta}(x, y)$ are continuous with respect to $\beta \in B$ for each (α, p, x, y) and continuous with respect to $\alpha \in A$ uniformly with respect to $\beta \in B$ for each (p, x, y). Furthermore, they are Borel measurable functions of (p, x, y) for each (α, β) and they are bounded by K_0.

(ii) The functions $r^{\alpha\beta}(x, y)$ and $P^{\alpha\beta}(x, y)$ are bounded by constant K_0, they are Borel measurable with respect to all variables, and along with $p^{\alpha\beta}(x, y)$ they are Lipschitz continuous with respect to x with Lipschitz constant K_1, and

$$r^{\alpha\beta}(x, x) \equiv 1, \qquad p^{\alpha\beta}(x, x) \equiv 0, \qquad P^{\alpha\beta}(x, x) \equiv I,$$

where I is the $d_1 \times d_1$-identity matrix. The function $p^{\alpha\beta}(x, y)$ is uniformly continuous with respect to y uniformly with respect to (α, β, x).

(iii) The functions $\sigma^{\alpha\beta}(p, x)$, $b^{\alpha\beta}(p, x)$, $c^{\alpha\beta}(p, x)$ and $f^{\alpha\beta}(p, x)$ are Lipschitz continuous with respect to (p, x) with Lipschitz constant K_1. We have $\|g\|_{C^2(\mathbb{R}^d)} \le K_1$.

(iv) For any $\alpha \in A$, $\beta \in B$, $x, \lambda \in \mathbb{R}^d$ and $p \in \mathbb{R}^k$, we have

$$a_{ij}^{\alpha\beta}(p, x)\lambda_i\lambda_j \ge \delta_0|\lambda|^2.$$

The reader understands, of course, that the summation convention is adopted throughout the article.

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space, let $\{\mathcal{F}_t, t \ge 0\}$ be an increasing filtration of σ-fields $\mathcal{F}_t \subset \mathcal{F}$ such that each $\mathcal{F}_t$ is complete with respect to $\mathcal{F}, P$ and let $w_t, t \ge 0$, be a standard d_1-dimensional Wiener process given on Ω such that w_t is a Wiener process relative to the filtration $\{\mathcal{F}_t, t \ge 0\}$.

The set of progressively measurable A-valued processes $\alpha_t = \alpha_t(\omega)$ is denoted by $\mathfrak{A}$. Similarly we define $\mathfrak{B}$ as the set of B-valued progressively measurable functions. By $\mathbb{B}$ we denote the set of $\mathfrak{B}$-valued functions $\boldsymbol{\beta}(\alpha_\cdot)$ on $\mathfrak{A}$ such that, for any $T \in (0, \infty)$ and any $\alpha^1_\cdot, \alpha^2_\cdot \in \mathfrak{A}$ satisfying

$$P\big(\alpha^1_t = \alpha^2_t \text{ for almost all } t \le T\big) = 1, \tag{2.1}$$

we have

$$P\big(\boldsymbol{\beta}_t(\alpha^1_\cdot) = \boldsymbol{\beta}_t(\alpha^2_\cdot) \text{ for almost all } t \le T\big) = 1.$$

For $\alpha. \in \mathfrak{A}$, $\beta. \in \mathfrak{B}$ and $x \in \mathbb{R}^d$ introduce $x_t^{\alpha.\beta.x}$ as a unique solution of the Itô equation

$$x_t = x + \int_0^t \sigma^{\alpha_s\beta_s}(x_s)\, dw_s + \int_0^t b^{\alpha_s\beta_s}(x_s)\, ds, \tag{2.2}$$

and denote

$$\phi_t^{\alpha.\beta.x} = \int_0^t c^{\alpha_s\beta_s}\big(x_s^{\alpha.\beta.x}\big)\,ds.$$

Next, fix a domain $D \subset \mathbb{R}^d$, define $\tau^{\alpha.\beta.x}$ as the first exit time of $x_t^{\alpha.\beta.x}$ from D ($\tau^{\alpha.\beta.x} = \infty$ if $D = \mathbb{R}^d$) and introduce

$$v(x) = \inf_{\boldsymbol{\beta}\in\mathbb{B}} \sup_{\alpha.\in\mathfrak{A}} E_x^{\alpha.\boldsymbol{\beta}(\alpha.)}\bigg[\int_0^\tau f(x_t)e^{-\phi_t}\,dt + g(x_\tau)e^{-\phi_\tau}\bigg], \tag{2.3}$$

where the indices $\alpha.$, $\boldsymbol{\beta}$, and x at the expectation sign are written to mean that they should be placed inside the expectation sign wherever and as appropriate, that is,

$$\begin{aligned} E_x^{\alpha.\beta.}&\bigg[\int_0^\tau f(x_t)e^{-\phi_t}\,dt + g(x_\tau)e^{-\phi_\tau}\bigg] \\ &:= E\bigg[g\big(x^{\alpha.\beta.x}_{\tau^{\alpha.\beta.x}}\big)e^{-\phi^{\alpha.\beta.x}_{\tau^{\alpha.\beta.x}}} + \int_0^{\tau^{\alpha.\beta.x}} f^{\alpha_t\beta_t}\big(x_t^{\alpha.\beta.x}\big)e^{-\phi_t^{\alpha.\beta.x}}\,dt\bigg]. \end{aligned}$$

Observe that $v(x) = g(x)$ in $\mathbb{R}^d \setminus D$. Next, introduce

$$L^{\alpha\beta}u(p,x) = a_{ij}^{\alpha\beta}(p,x)D_{ij}u(x) + b_i^{\alpha\beta}(p,x)D_iu(x) - c^{\alpha\beta}(p,x)u(x),$$

where $D_i = \partial/(\partial x_i)$, $D_{ij} = D_iD_j$ and note for orientation that v is a viscosity solution of the corresponding Isaacs equation

$$\sup_{\alpha\in A}\inf_{\beta\in B}\big[L^{\alpha\beta}u(0,x) + f^{\alpha\beta}(x)\big] = 0, \qquad x \in D.$$

This fact which will not play any role here is proved in [5] for bounded domains.

Our first main result is the following.

THEOREM 2.1. *Under the above assumptions also suppose that either D is bounded and satisfies the uniform exterior ball condition, or $D = \mathbb{R}^d$ and there is a constant $\delta_1 > 0$ such that $c^{\alpha\beta}(x) \ge \delta_1$.*

Then v is Lipschitz continuous in $\mathbb{R}^d$ with Lipschitz constant depending only on D, K_0, K_1, δ_0 and δ_1.

The above setting and notation follow [10] and, as there, we convince ourselves that the definition of v makes sense, and v is bounded.

Here is a result about interior smoothness of v.

THEOREM 2.2. *Let D be bounded and in Assumption 2.1*(iii) *replace the requirement $\|g\|_{C^2(\mathbb{R}^d)} \le K_1$ with the requirement that g is continuous. Then v is Lipschitz continuous on any compact set $\Gamma \subset D$.*

As we have pointed out in the Introduction, Theorems 2.1 and 2.2 are known and even in much stronger forms for quite some time and we give them with proofs just to show that there is a probabilistic technique to derive them and also to prepare some necessary tools for proving our next result, which is about Lipschitz continuity of v with constant *independent* of δ_0. As usual, in this case we need the following:

ASSUMPTION 2.2. There exists a $\delta_1 \in (0,1]$ such that for any $\alpha \in A$, $\beta \in B$, $x \in \mathbb{R}^d$ and $p \in \mathbb{R}^k$ we have

$$c^{\alpha\beta}(p,x) \geq \delta_1.$$

REMARK 2.1. Assume that D lies in the ball of radius R centered at the origin. For $\mu > 0$ define $\Psi(x) = \cosh(\mu R) - \cosh(\mu|x|) + 2$. It is easy to check that for μ large enough depending only on δ_0, K_0 and d, the function Ψ is infinitely differentiable on $\mathbb{R}^d$, $\Psi \geq 2$ on D and $(L^{\alpha\beta} + c^{\alpha\beta})\Psi \leq -1$ on D for all α, β. This is a so-called global barrier for D.

We modify it for $|x| \geq R$ in such a way that it will be still infinitely differentiable on $\mathbb{R}^d$, have bounded derivatives and be such that $\Psi \geq 1$ on $\mathbb{R}^d$. We keep the same notation for the modified function. By Remark 2.3 of [10] if we construct $\check{v}$ from

$$\check{\sigma}^{\alpha\beta}(x) = \Psi^{1/2}(x)\sigma^{\alpha\beta}(x), \qquad \check{b}^{\alpha\beta}(x) = \Psi(x)b^{\alpha\beta}(x) + 2a^{\alpha\beta}(x)D\Psi(x),$$
$$\check{c}^{\alpha\beta}(x) = -L^{\alpha\beta}\Psi(x), \qquad \check{f}^{\alpha\beta}(x) = f^{\alpha\beta}(x), \qquad \check{g}(x) = \Psi^{-1}(x)g(x),$$

where $D\Psi$ is the gradient of Ψ (a column vector), in the same way as v was constructed from the original σ, b, c, f and g, then $\check{v} = \Psi^{-1}v$. By no means the above transformation is something new; see, for instance, Sections 1.2 and 2.5 in [13]. Just in case, observe that now $c^{\alpha\beta}$ influences $\check{v}$ through $\check{c}^{\alpha\beta}$, which is bigger than one (remember that $c^{\alpha\beta} \geq 0$). This shows that without restricting generality we could have supposed that Assumption 2.2 is satisfied even in Theorems 2.1 and 2.2.

Introduce

$$\hat{\sigma}^{\alpha\beta}(x,y) = r^{\alpha\beta}(x,y)\sigma^{\alpha\beta}\big(p^{\alpha\beta}(x,y),x\big)P^{\alpha\beta}(x,y),$$
$$(\hat{a},\hat{b},\hat{c},\hat{f})^{\alpha\beta}(x,y) = \big[r^{\alpha\beta}(x,y)\big]^2(a,b,c,f)^{\alpha\beta}\big(p^{\alpha\beta}(x,y),x\big),$$

and for unit $\xi \in \mathbb{R}^d$ introduce a convex function $\|\sigma\|_\xi^2$ on the set of $d \times d_1$ matrices by

$$\|\sigma\|_\xi^2 := \|\sigma\|^2 - |\xi^*\sigma|^2 = \big\|(I - \xi\xi^*)\sigma\big\|^2, \qquad \|\sigma\|^2 = \sum_{i,j}\sigma_{ij}^2, \tag{2.4}$$

where I is the unit $d \times d$ matrix.

ASSUMPTION 2.3. For all $\alpha \in A$, $\beta \in B$ and $x, y \in \mathbb{R}^d$

$$\delta_1^{-1} \geq r^{\alpha\beta}(x, y) \geq \delta_1.$$

ASSUMPTION 2.4. There exist constants $\delta \geq 2\delta_1$, $\varepsilon_0 > 0$ and $\mu \geq 1$ such that for all $\alpha \in A$, $\beta \in B$ and $x, y \in \mathbb{R}^d$, for which $|x - y| \leq \varepsilon_0$, we have

$$\begin{aligned} &\|\hat{\sigma}^{\alpha\beta}(x, y) - \sigma^{\alpha\beta}(y)\|_{\xi}^2 + 2\langle x - y, \hat{b}^{\alpha\beta}(x, y) - b^{\alpha\beta}(y)\rangle \\ &\quad \leq 2(c^{\alpha\beta}(y) - \delta)|x - y|^2 + 4\mu\langle x - y, a^{\alpha\beta}(x)(x - y)\rangle, \end{aligned} \tag{2.5}$$

where $\xi = (x - y)/|x - y|$.

REMARK 2.2. If $d = 1$, then for any $d \times d_1$-matrix σ and unit $\xi \in \mathbb{R}^d$, we have $\|\sigma\| = |\xi^*\sigma|$, so that in that case the term involving σ in (2.5) disappears. Also notice that if σ and b are independent of p, and $r \equiv 1$, $p \equiv 0$, and $P \equiv I$, then $(\hat{a}, \hat{\sigma}, \hat{b}, \hat{c})^{\alpha\beta}(x, y) = (a, \sigma, b, c)^{\alpha\beta}(x)$, and condition (2.5) becomes

$$\begin{aligned} &\|\sigma^{\alpha\beta}(x) - \sigma^{\alpha\beta}(y)\|_{\xi}^2 + 2\langle x - y, b^{\alpha\beta}(x) - b^{\alpha\beta}(y)\rangle \\ &\quad \leq 2(c^{\alpha\beta}(y) - \delta)|x - y|^2 + 4\mu\langle x - y, a^{\alpha\beta}(x)(x - y)\rangle, \end{aligned} \tag{2.6}$$

which is satisfied with any δ on the account of choosing a sufficiently large μ (depending on δ_0 and K_1) since σ and b are Lipschitz continuous. Therefore, Theorem 2.1 is a particular case of Theorem 2.3. It is also worth noting that if $d = 1$, condition (2.6) is satisfied with $\mu = 0$ when $b^{\alpha\beta}(x)$ are decreasing functions of x and $c^{\alpha\beta} \geq \delta$.

In Section 3 we give more examples when one can check Assumption 2.3.

Introduce

$$\begin{aligned} &H(p, x, u, (u_i), (u_{ij})) \\ &\quad = \sup_{\alpha \in A} \inf_{\beta \in B} [a_{ij}^{\alpha\beta}(p, x)u_{ij} + b_i^{\alpha\beta}(p, x)u_i - c^{\alpha\beta}(p, x)u + f^{\alpha\beta}(p, x)]. \end{aligned}$$

ASSUMPTION 2.5. The set of $(x, u, (u_i), (u_{ij}))$ such that

$$H(p, x, u, (u_i), (u_{ij})) \leq 0 \tag{2.7}$$

is *independent* of p and the same is true if we reverse the sign of the inequality.

Note that the next result does not cover Theorem 2.2 and by “the above assumptions” we mean all assumptions which are stated above in this section.

THEOREM 2.3. *Under the above assumptions also suppose that either* $D=\mathbb{R}^d$, *or* D *is bounded and there exists a nonnegative function* $G\in C^{0,1}(\bar{D})\cap C^2_{\mathrm{loc}}(D)$ *such that* $G=0$ *on* ∂D *and*

$$L^{\alpha\beta}G(p,x)\le -1$$

in D *for any* p.

Then v *is Lipschitz continuous in* $\mathbb{R}^d$ *with Lipschitz constant* independent *of* δ_0.

REMARK 2.3. If D is bounded and satisfies the uniform exterior ball condition, the function G always exists since the operators $L^{\alpha\beta}$ are uniformly nondegenerate, have bounded coefficients and $c^{\alpha\beta}\ge 0$. However, the proof of this well-known fact relies on the uniform nondegeneracy and gives a function G depending on δ_0. The reader should understand that there are plenty of cases when this assumption is satisfied, even for degenerate operators; see, for instance, Example 3.1 with $\delta_0=0$.

Finally, we state one more result, which was actually the main motivation of writing the whole series consisting of [10, 12] and the present article, as we have pointed out in the Introduction. We take $D=\mathbb{R}^d$ and suppose that all above assumptions are satisfied and σ, b, c, f are independent of p.

Set

$$A_1=A,$$

and let A_2 be a separable metric space having no common points with A_1.

ASSUMPTION 2.6. The functions $\sigma^{\alpha\beta}(x)$, $b^{\alpha\beta}(x)$, $c^{\alpha\beta}(x)$ and $f^{\alpha\beta}(x)$ are also defined on $A_2\times B\times\mathbb{R}^d$ in such a way that they are *independent* of β and satisfy Assumptions 2.1(i), (iii), (iv) with the same constants K_0, K_1 and, of course, with A_2 in place of A.

Define

$$\hat{A}=A_1\cup A_2.$$

Then we introduce $\hat{\mathfrak{A}}$ as the set of progressively measurable $\hat{A}$-valued processes and $\hat{\mathbb{B}}$ as the set of $\mathfrak{B}$-valued functions $\boldsymbol{\beta}(\alpha_\cdot)$ on $\hat{\mathfrak{A}}$ such that, for any $T\in[0,\infty)$ and any $\alpha^1_\cdot,\alpha^2_\cdot\in\hat{\mathfrak{A}}$ satisfying

$$P\bigl(\alpha^1_t=\alpha^2_t \text{ for almost all } t\le T\bigr)=1,$$

we have

$$P\bigl(\boldsymbol{\beta}_t(\alpha^1_\cdot)=\boldsymbol{\beta}_t(\alpha^2_\cdot) \text{ for almost all } t\le T\bigr)=1.$$

For a constant $K \ge 0$, set

$$v_K(x) = \inf_{\boldsymbol{\beta} \in \hat{\mathbb{B}}} \sup_{\alpha. \in \hat{\mathfrak{A}}} v_K^{\alpha.\boldsymbol{\beta}(\alpha.)}(x),$$

where

$$v_K^{\alpha.\beta.}(x) = E_x^{\alpha.\beta.} \int_0^\infty f_K(x_t) e^{-\phi_t}\, dt =: v^{\alpha.\beta.}(x) - K E_x^{\alpha.\beta.} \int_0^\gamma I_{\alpha_t \in A_2} e^{-\phi_t}\, dt,$$

$$f_K^{\alpha\beta}(x) = f^{\alpha\beta}(x) - K I_{\alpha \in A_2}.$$

The above formula extends $v^{\alpha.\beta.}(x)$, initially defined for $\alpha. \in \mathfrak{A}$ and $\beta. \in \mathfrak{B}$, on the set $\hat{\mathfrak{A}} \times \mathfrak{B}$. Of course, (2.3) is preserved with $\tau = \infty$, and no g is involved.

THEOREM 2.4. *There is a constant N, depending only on the constants in all above assumptions (but not on K), such that $|v_K(x) - v(x)| \le N/K$ for all $x \in \mathbb{R}^d$ and $K \ge 1$.*

REMARK 2.4. In one of the main cases of interest v_K turns out to have second-order derivatives bounded by a constant times K if $K \ge 1$; see Section 7 in [10]. From the point of view of finite-difference approximations it should be easier to approximate "smooth" functions v_K than v. However, the author has no idea how to prove a fact similar to Theorem 2.4 for finite-difference equations.

In this connection it would be very interesting to find any proof of Theorem 2.4 not using probability theory, of course, defining v_K and v as viscosity solutions of the corresponding Isaacs equations.

3. Comments and examples.

REMARK 3.1. Let σ and b be independent of α and β, and consider a particular case where $d_1 = d$, and equation (2.2) is

$$x_t = x + \int_0^t \sigma(x_s)\, dw_s, \tag{3.1}$$

where σ is an $\mathcal{O}$-valued Lipschitz continuous function. Then the left-hand side of (2.5) vanishes for $r \equiv 1$ and $P(x, y) = \sigma^*(x)\sigma(y)$. Of course, this is not a big surprise since x_t is just a Brownian motion starting at x. Still one can see that the parameters P take care of rotations of the increments of the original Wiener process and basically show that (2.5) is a condition on a rather than σ.

REMARK 3.2. The function v will not change if we change σ, b, c, f outside D. In connection with this it is worth noting that in Assumption 2.4 we may restrict x and y to D_{ε_0} which is the ε_0 neighborhood of D. Indeed, if only thus restricted Assumption 2.4 is satisfied we could just change c outside D so that it will be bigger than the original one and become any large constant outside D_{ε_0}. Then Assumption 2.4 will be satisfied in the form it is stated.

REMARK 3.3. For later discussion we show that Assumption 2.4 can be replaced with a slightly more transparent one. We will be only concerned with Assumption 2.4 leaving other assumptions aside.

Denote by $\mathcal{S}k$ the set of $d_1 \times d_1$ skew-symmetric matrices and assume that for each $\alpha \in A$, $\beta \in B$ and $\xi \in \mathbb{R}^d$, the following functions on $\mathbb{R}^d$ are also given: $\mathcal{S}k$-valued function $\Theta^{\alpha\beta}(x,\xi)$, $k \times d$ matrix-valued function $p^{\alpha\beta}(x)$, and $\mathbb{R}^d$-valued function $r^{\alpha\beta}(x)$.

For a differentiable function $u(p,x)$ and $\xi \in \mathbb{R}^d$, introduce

$$\partial_\xi u^{\alpha\beta}(x) = \xi_i u_{x_i}(0,x) + \big(p^{\alpha\beta}(x)\xi\big)_j u_{p_j}(0,x).$$

Also denote $\operatorname{Conv}(D)$ the open convex hull of D.

ASSUMPTION 3.1. (i) For $|\xi| \le 1$ the above functions are bounded by K_0 and $\Theta^{\alpha\beta}(x,y)$ is a linear function of y [in particular $\Theta^{\alpha\beta}(x,0)=0$].

(ii) For any $\alpha \in A$ and $\beta \in B$ the functions $\sigma^{\alpha\beta}(p,x)$ and $b^{\alpha\beta}(p,x)$ are continuously differentiable with respect to $(p,x) \in \mathbb{R}^k \times \mathbb{R}^d$, and their first-order derivatives are bounded by K_1. Furthermore, their derivatives are uniformly continuous with respect to (p,x) uniformly with respect to $(\alpha,\beta) \in A \times B$.

(iii) There are constants $\mu \ge 1$ and $\delta \ge 2\delta_1$ such that for any unit $\xi \in \mathbb{R}^d$ and $(\alpha,\beta,x) \in A \times B \times \operatorname{Conv}(D)$, we have

$$\begin{aligned} &\big\|\partial_\xi \sigma^{\alpha\beta}(x) + \langle r^{\alpha\beta}(x),\xi\rangle \sigma^{\alpha\beta}(x) + \sigma^{\alpha\beta}(x)\Theta^{\alpha\beta}(x,\xi)\big\|_\xi^2 \\ &\quad + 2\big\langle \xi, \partial_\xi b^{\alpha\beta}(x) + 2\langle r^{\alpha\beta}(x),\xi\rangle b^{\alpha\beta}(x)\big\rangle \\ &\qquad \le 2\big(c^{\alpha\beta}(x) - \delta_1 - \delta\big) + 4\mu\langle \xi, a^{\alpha\beta}(x)\xi\rangle. \end{aligned} \tag{3.2}$$

Introduce

$$\begin{aligned} &r^{\alpha\beta}(x,y) = 1 + \langle r^{\alpha\beta}(y), x-y\rangle, \qquad p^{\alpha\beta}(x,y) = p^{\alpha\beta}(y)(x-y), \\ &P^{\alpha\beta}(x,y) = \exp \Theta^{\alpha\beta}(y, x-y). \end{aligned} \tag{3.3}$$

We claim that there exists an $\varepsilon_0 > 0$, depending only on K_0, K_1, δ_1, d, and the moduli of continuity in (p,x) of the derivatives of $\sigma^{\alpha\beta}(p,x)$ and $b^{\alpha\beta}(p,x)$ with respect to (p,x), such that Assumption 2.4 is satisfied with x, y restricted to D.

To prove the claim, fix $y \in D$ and a unit $\xi \in \mathbb{R}^d$, and for $t \ge 0$ introduce $x(t) = y + t\xi$, so that (2.5) becomes

$$\begin{aligned} &\big\|\hat\sigma^{\alpha\beta}\big(x(t),y\big) - \sigma^{\alpha\beta}(y)\big\|_\xi^2 + 2t\big\langle \xi, \hat b^{\alpha\beta}\big(x(t),y\big) - b^{\alpha\beta}(y)\big\rangle \\ &\qquad \le 2\big(c^{\alpha\beta}(y) - \delta\big)t^2 + 4\mu\big\langle \xi, a^{\alpha\beta}\big(x(t)\big)\xi\big\rangle t^2, \end{aligned} \tag{3.4}$$

which we want to prove for $t \in (0,\varepsilon_0]$. For simplicity of notation we will drop the superscripts α, β in a few lines below.

Observe that

$$\hat{\sigma}\big(x(t), y\big) - \sigma(y) = \int_0^t \xi_i \hat{\sigma}_{x_i}\big(x(s), y\big)\, ds,$$

where

$$\begin{aligned}
&\xi_i \hat{\sigma}_{x_i}\big(x(s), y\big) \\
&\quad = \langle r(y), \xi\rangle \sigma\big(sp(y)\xi, x(s)\big) P\big(x(s), y\big) \\
&\qquad + r\big(x(s), y\big)\big[\xi_i \sigma_{x_i}\big(sp(y)\xi, x(s)\big) + \big(p(y)\xi\big)_j \sigma_{p_j}\big(sp(y)\xi, x(s)\big)\big] \\
&\qquad \times P\big(x(s), y\big) \\
&\qquad + r\big(x(s), y\big)\sigma\big(sp(y)\xi, x(s)\big)\Theta(y, \xi) P\big(x(s), y\big) \\
&\quad =: \langle r(y), \xi\rangle \sigma(y) + \partial_\xi \sigma(y) + \sigma(y)\Theta(y, \xi) + R(s),
\end{aligned}$$

and $R(s)$ is introduced by the above equality.

Owing to the convexity of function (2.4) and Assumption 3.1, there exists an $\varepsilon_0 > 0$ such that for all $t \in (0, \varepsilon_0]$ and all values of other arguments, we have

$$\begin{aligned}
&\big\|\hat{\sigma}^{\alpha\beta}\big(x(t), y\big) - \sigma^{\alpha\beta}(y)\big\|_\xi^2 - 4\mu\langle \xi, a^{\alpha\beta}\big(x(t)\big)\xi\rangle t^2 \\
&\qquad \le t^2 \big\|\partial_\xi \sigma^{\alpha\beta}(y) + \langle r^{\alpha\beta}(y), \xi\rangle \sigma^{\alpha\beta}(y) + \sigma^{\alpha\beta}(y)\Theta^{\alpha\beta}(y, \xi)\big\|_\xi^2 \\
&\qquad\quad - 4\mu\langle \xi, a^{\alpha\beta}(y)\xi\rangle t^2 + t^2\delta_1.
\end{aligned}$$

It is even easier to prove that, by reducing ε_0 if necessary, we have that for $t \in (0, \varepsilon_0]$ and all values of other arguments

$$\begin{aligned}
&t\langle \xi, \hat{b}^{\alpha\beta}\big(x(t), y\big) - b^{\alpha\beta}(y)\rangle \\
&\qquad \le t^2 \langle \xi, \partial_\xi b^{\alpha\beta}(y) + 2\langle r^{\alpha\beta}(y), \xi\rangle b^{\alpha\beta}(y)\rangle + t^2\delta_1.
\end{aligned}$$

Hence, by assumption, the left-hand side of (3.4) is less than

$$t^2\big[2\big(c^{\alpha\beta}(y) - \delta_1 - \delta\big) + 4\mu\langle \xi, a^{\alpha\beta}(y)\xi\rangle\big] + 2t^2\delta_1,$$

which is the right-hand side of (3.4).

REMARK 3.4. Consider the case that σ and b are independent of α and β. Let $d = 1$, $c > 0$ and $D = (-1, 1)$. Assume that $a = a_0 + \delta_0$, where $a_0 \ge 0$. In that case, as it follows from the arguments in Remarks 2.2 and 3.3, we do not need to assume that σ' is continuous. We still assume that a, b' and c are continuous. Then by Remark 2.2 Assumption 2.4 is satisfied with μ depending on δ_0, among other things.

However, assume additionally that at every point $x \in [-2, 2]$ where

$$a_0(x) = b(x) = 0$$

we have

$$b'(x) < c(x). \tag{3.5}$$

We claim that then Assumption 2.4 is satisfied with x, y restricted to $[-2, 2]$ with some $\delta, \delta_1, \varepsilon_0$, and μ *independent* of δ_0 and hence, by Remark 3.2, it will be satisfied in the original form, making the assertion of Theorem 2.3 valid in case $D = (-1, 1)$.

To prove the claim, we use Remark 3.3 and observe that for $r = -nb/2$, $\delta_1 + \delta = 1/n$, $\mu = n$ and $|\xi| = 1$ condition (3.2) is satisfied if

$$b'(x) \le c(x) - \frac{1}{n} + n\big(a_0(x) + |b(x)|^2\big). \tag{3.6}$$

Suppose that for any $n = 1, 2, \dots$ we can find a point $x_n \in [-2, 2]$ at which the inequality converse to (3.6) holds. Then we can extract from the sequence x_n a subsequence that converges to an $x_0 \in [-2, 2]$. Clearly, for large n,

$$a_0(x_n) + |b(x_n)|^2 \le N n^{-1},$$

where $N = \sup b' + 1$. Therefore, $a_0(x_0) + |b(x_0)|^2 = 0$ and

$$b'(x_n) \ge c(x_n) - 1/n, \qquad b'(x_0) \ge c(x_0).$$

We have obtained a contradiction to (3.5), so inequality (3.6) holds in $[-2, 2]$ for some n independent of δ_0 thus proving our claim.

EXAMPLE 3.1. Consider the one-dimensional equation

$$\delta_0 v'' + bxv' - v = 0 \tag{3.7}$$

on $[-1, 1]$ with data 1 at ± 1, where constant $b > 0$. This is, of course, a simple example of the Isaacs equation in a differential "game" with the value function v. Here the assumption stated in Theorem 2.3 concerning G is satisfied with $G(x) = (1 - x^2)\max(1, 1/(2b))$.

If we assume that the solution $v = v_{\delta_0}$ admits an estimate of its Lipschitz constant independent of δ_0, then, as is easy to understand, say from the probabilistic representation of v_δ, the function

$$v_0(x) = Ee^{-\tau_x}$$

would be Lipschitz continuous, where τ_x is the first exit time of the solution of

$$x_t = x + \int_0^t bx_s\, ds$$

from $(-1, 1)$. Since $x_t = xe^{bt}$, $\tau_x = -b^{-1}\ln|x|$ for $|x| < 1$ and $v_0(x) = |x|^{1/b}$, which is Lipschitz continuous only if $b \le 1$.

This example shows that in the situation of Remark 3.4, if one has $b'(x) > c(x)$ at least at one point at which $a_0(x) = b(x) = 0$, the assertion of Theorem 2.3 may be no longer true. In this respect, requiring condition (3.5) at those points is close to being optimal and it is, actually, necessary for v to be *continuously* differentiable.

REMARK 3.5. Barles in [1] derived first-order derivatives estimates for viscosity solutions of nonlinear equations

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

in domains, where $Du=(D_iu)$ is the gradient of u, and $D^2u=(D^2_{ij}u)$ is its Hessian. Our value functions are viscosity solutions of the corresponding Isaacs equations. This fact is proved in [5] for bounded domains. The Isaacs equations in this paper are included in the framework of [1] and many of the equations in [1] do not fit into our scheme. Yet it is worth comparing our conditions with the ones from [1] in the *simplest* example of *linear* equations with

$$H(x,u_0,u',u'')=a_{ij}(x)u''_{ij}+b_i(x)u'_i-c(x)u_0+f(x)$$

for which solutions have probabilistic representations (with no α and β involved).

One of the assumptions in [1] reads as follows: For any $R>0$ and all large enough L,

$$\begin{aligned}&c\sum_{i=1}^d|u'_i|^2+g\operatorname{tr}u''au''\\ &\qquad-\big[u'_kD_ka_{ij}u''_{ij}+u'_kD_kb_i(x)u'_i-u'_kD_kc(x)u_0+u'_kD_kf(x)\big]\\ &\qquad\ge h,\end{aligned}\tag{3.8}$$

where $g,h>0$ are some constants >0, provided that

$$|u_0|\le R,\qquad \sum_{i=1}^d|u'_i|^2\ge L,\qquad H(x,u_0,u',u'')=0,\qquad u''_{ij}u'_j=0\ \forall i.\tag{3.9}$$

If $c\equiv 0$, $b\equiv 0$, and both f and Df vanish at a point x_0, so that $H(x_0,0)=0$, then for $u''=0$ inequality (3.8) at x_0 becomes $0\ge h$, which cannot hold even in the one-dimensional case. Therefore, the one-dimensional equation

$$D^2u+x^2=0$$

in $(-1,1)$ with zero boundary condition does not fit in the scheme of [1].

Equation

$$\delta_0D^2u+(b_1x+b_0)Du-cu+x^2=0$$

in $(-1,1)$ with zero boundary condition and constant $c>0,b_0,b_1$ does not fit in either if $c\le b_1$.

Indeed, if we take $x=0$, $u''=0$, $u_0=0$, and u' bigger by magnitude than L, (3.8) becomes

$$(c-b_1)|u'|^2\ge h,$$

which for large $|u'|$ can only hold if $b_1 < c$. Remark 3.4 shows that one always has an estimate of the Lipschitz constant of v. This estimate is even *independent* of δ_0, provided that either $b_1 x + b_0 \neq 0$ for $x \in [-1, 1]$ or $b_1 < c$.

It looks like the methods of [1] are not adapted to use uniform nondegeneracy and even in the above examples lead to the requirement that c be sufficiently large.

REMARK 3.6. Above we saw that the parameters μ, r and P can play a role while checking Assumption 2.4. We now show how the external parameters p can be used. Here we consider the situation in which σ, b, c and f depend only on x and α so that we are dealing with controlled diffusion processes rather than differential games. Our interest is in obtaining estimates independent of δ_0, and therefore, from the start in this remark we focus on *degenerate* processes.

Let $A = \mathbb{R}$ and consider a one-dimensional process defined by the equation

$$x_t = x + \int_0^t \sigma(x_s)\,dw_s + \int_0^t \tanh(x_s + 2\cos\alpha_s)\,ds, \tag{3.10}$$

where w_t is a one-dimensional Wiener process, $\sigma(x)$ is a smooth nonnegative even function satisfying $\sigma(x) > 0$ for $x \in (1, 3)$ and vanishing outside $(1, 3)$ (and α_t is a progressively measurable A-valued process). We also take a sufficiently regular function $c(x) \geq \delta_2$ (independent of α and β), where $\delta_2 > 0$, and take $D = \mathbb{R}$.

If we want to satisfy (3.2) for $|x| \notin [1, 3]$ with $r(x) = 0$ (and $\Theta \equiv 0$ for having no other options) and some δ's we obviously need to have

$$c(x) > 1 \qquad \text{for } |x| \leq 1, \qquad c(x) > \cosh^{-2}\big(|x| - 2\big) \qquad \text{for } |x| \geq 3. \tag{3.11}$$

The inequalities in (3.11) extend for $|x| \notin (1 + \varepsilon, 3 - \varepsilon)$ with some $\varepsilon > 0$, and one can find $\mu \geq 1$ such that (3.2) is satisfied (with some δ's) for $|x| \in (1 + \varepsilon, 3 - \varepsilon)$ with $r(x) = 0$. Therefore, if we do not use parameter r, then (3.2) reduces to (3.11).

However, if we take

$$r^\alpha(x) = -2 I_{|x+2\cos\alpha|>\varepsilon} \sinh^{-1}(2x + 4\cos\alpha), \tag{3.12}$$

then the left-hand side of (3.2) becomes

$$2 I_{|x+2\cos\alpha|\leq\varepsilon} \cosh^{-2}(x + 2\cos\alpha) \leq 2 I_{|x+2\cos\alpha|\leq\varepsilon},$$

and for $|x| \notin (1 + \varepsilon, 3 - \varepsilon)$ this is strictly less than $2c(x)$ if

$$c(x) > 1 \qquad \text{for } |x| \leq 1 + \varepsilon. \tag{3.13}$$

Hence, with the so specified r^α condition, (3.2) reduces to (3.13), which is a significant improvement over (3.11).

Next we take f independent of α, say $f \equiv 1$, and instead of

$$b^\alpha(x) = \tanh(x + 2\cos\alpha)$$

consider

$$b^\alpha(p, x) = \tanh\big(x + 2\cos(\alpha + p)\big),$$

where $p \in \mathbb{R}$. Obviously, Assumption 2.5 is satisfied.

Take $r^\alpha(x)$ from (3.12) and

$$p^\alpha(x) = (1/2)I_{|x+2\cos\alpha|\le\varepsilon} I_{|\sin\alpha|>\varepsilon} \sin^{-1}\alpha. \tag{3.14}$$

Then the left-hand side of (3.2) becomes

$$\begin{aligned}
2I_{|x+2\cos\alpha|\le\varepsilon} &\cosh^{-2}(x+2\cos\alpha) - 2I_{|x+2\cos\alpha|\le\varepsilon} I_{|\sin\alpha|>\varepsilon} \cosh^{-2}(x+2\cos\alpha) \\
&= 2I_{|x+2\cos\alpha|\le\varepsilon} I_{|\sin\alpha|\le\varepsilon} \cosh^{-2}(x+2\cos\alpha) \le 2I_{|x+2\cos\alpha|\le\varepsilon} I_{|\sin\alpha|\le\varepsilon},
\end{aligned}$$

and the latter is zero if $|x| \le 1+\varepsilon$ and ε is sufficiently small. Thus adding $p^\alpha(x)$ into the picture eliminates condition (3.13) entirely, and there is nothing more than $c(x) \ge \delta_2$ required of $c(x)$ in order for (3.2) to be satisfied with $r^\alpha(x)$ from (3.12) and $p^\alpha(x)$ from (3.14).

By the way, the Isaacs (Bellman) equation in this case is

$$a(x)D^2v(x) + \big(Dv(x)\big)\tanh\big[x + 2\,\mathrm{sign}\big(Dv(x)\big)\big] - c(x)v(x) + f(x) = 0,$$

where $a = (1/2)\sigma^2$. This equation suggests a different representation of the value function with $A = \{\pm 1\}$ when using parameters p becomes unnecessary (and impossible) but using r will suffice. In this connection it is worth mentioning that much more sophisticated use of the external parameters p can be found in [7], where in an example of (degenerate) complex Monge–Ampère equation they are shown to be indispensable in proving the global $C^{1,1}$ regularity of solutions.

4. Some underlying ideas. This article is written for probabilists and the translation of the proof of the central Theorem 2.4 in PDE terms or in terms of the theory of viscosity solutions is unknown to the author. On the other hand, such a translation may exist for Theorem 2.3 and the interested, more PDE oriented, reader can find in Section 8.5 of [6] analytical tools allowing one to prove an analog of Theorem 2.3 for Bellman's equations.

However, for probabilists the following explanation of ideas behind the proof of Theorem 2.3 might be helpful. The main idea is that while differentiating $v(x)$ with respect to x we can take different representation for v at different points. We explain how various terms in (3.2) appear naturally on two examples of stochastic equations without games.

EXAMPLE 4.1. In Remark 3.1, take a smooth bounded $f(x)$ and define

$$v(x) = E\int_0^\infty e^{-t} f\big(x_t^x\big)\,dt, \tag{4.1}$$

where we use the same stipulation about indices as before and do not write α and β because nothing is depending on these parameters. One can formally differentiate $v(x)$ and obtain that for any $\xi \in \mathbb{R}^d$,

$$v_{(\xi)}(x) = E\int_0^\infty e^{-t} f_{(\xi_t)}\big(x_t^x\big)\,dt, \tag{4.2}$$

where ξ_t is defined as the solution of

$$d\xi_t = \sigma_{(\xi_t)}\big(x_t^x\big)\,dw_t, \qquad \xi_0 = \xi.$$

Actually, it is not hard to see that (4.1) is indeed true, provided that

$$E|\xi_t| \le Ne^{\gamma t}, \tag{4.3}$$

where N is a constant and a constant $\gamma < 1$. In that case the right-hand side of (4.2) is well defined. This may not happen if the derivatives of σ are big.

However, observe that for any $d \times d$-valued skew-symmetric progressively measurable process Θ_t and any ε we also have

$$v(x + \varepsilon\xi) = E\int_0^\infty e^{-t} f\big(x_t^x(\varepsilon)\big)\,dt, \tag{4.4}$$

where $x_t^x(\varepsilon)$ is defined as a unique solution of

$$dx_t = \sigma(x_t)e^{\varepsilon\Theta_t}\,dw_t, \qquad x_0 = x.$$

Formula (4.4) is indeed true because

$$e^{\varepsilon\Theta_t}\,dw_t = db_t,$$

where b_t is a Wiener process and the distributions of solutions of (3.1) are independent of which Wiener process is involved. Now let us formally differentiate (4.4) through with respect to ε at $\varepsilon = 0$. We again obtain (4.2), but this time ξ_t satisfies

$$d\xi_t = \big[\sigma_{(\xi_t)}\big(x_t^x\big) + \sigma\big(x_t^x\big)\Theta_t\big]\,dw_t, \qquad \xi_0 = \xi. \tag{4.5}$$

Here the coefficient of dw_t vanishes if we take $\Theta_t = -\sigma^*(x_t^x)\sigma_{(\xi_t)}(x_t^x)$, so that $\xi_t \equiv \xi$ and nothing like (4.3) is an issue any longer. The reader may object that one cannot take $\Theta_t = -\sigma^*(x_t^x)\sigma_{(\xi_t)}(x_t^x)$ before solving (4.5). Then take $\Theta_t = -\sigma^*(x_t^x)\sigma_{(\xi)}(x_t^x)$ and use that $\xi_t \equiv \xi$ satisfies (4.5).

For any Θ_t we have from (4.5) that

$$d|\xi_t|^2 = \big\|\sigma_{(\xi_t)}\big(x_t^x\big) + \sigma\big(x_t^x\big)\Theta_t\big\|^2\,dt + dm_t,$$

where m_t is a local martingale. This shows the origin of $\sigma^{\alpha\beta}(x)\Theta^{\alpha\beta}(x,\xi)$ in (3.2). The subscript ξ appears there after we compute $d|\xi_t|$.

EXAMPLE 4.2. Consider the one-dimensional Itô equation

$$dx_t = \sigma(x_t)\,dw_t + b(x_t)\,dt, \qquad x_0 = x$$

with one-dimensional w_t, and introduce $v(x)$ as in (4.1), so that $c = 1$. Then we again have (4.2) provided that (4.3) holds with a $\gamma < 1$ and ξ_t defined as a unique solution of

$$d\xi_t = \sigma_{(\xi_t)}\big(x_t^x\big)\,dw_t + b_{(\xi_t)}\big(x_t^x\big)\,dt, \qquad \xi_0 = \xi. \tag{4.6}$$

The solution of (4.6) is known to be

$$\xi_t = \xi m_t \exp \int_0^t b'(x_s^x)\,ds,$$

where

$$m_t = \exp\left(\int_0^t \sigma'(x_s^x)\,dw_s - (1/2)\int_0^t |\sigma'(x_s^x)|^2\,ds\right)$$

is at least a supermartingale. Hence (4.3) becomes

$$Em_t \exp \int_0^t b'(x_s^x)\,ds \le Ne^{\gamma t}$$

and a sufficient condition for that to happen is $b' \le \gamma c$ (since $Em_t \le 1$).

However, one can use a random time change and get a different representation for v. Namely, take any progressively measurable real-valued bounded process r_t and for ε such that $1 + 2\varepsilon r_t \ge 1/2$ introduce $x_t^x(\varepsilon)$ as a unique solution of

$$dx_t = \sqrt{1+2\varepsilon r_t}\,\sigma(x_t)\,dw_t + (1+2\varepsilon r_t)b(x_t)\,dt, \qquad x_0 = x. \tag{4.7}$$

Then it is well known that

$$v(x) = E\int_0^\infty f(x_t^x(\varepsilon))(1+2\varepsilon r_t)\exp\left(-\int_0^t (1+2\varepsilon r_s)\,ds\right)dt. \tag{4.8}$$

We substitute $x + \varepsilon\xi$ in place of x in (4.8) and differentiate with respect to ε at $\varepsilon = 0$. Then instead of (4.2) we obtain

$$v_{(\xi)}(x) = E\int_0^\infty \left[f_{(\xi_t)}(x_t^x) + 2r_t f(x_t^x) - 2f(x_t^x)\int_0^t r_s\,ds\right]e^{-t}\,dt, \tag{4.9}$$

where ξ_t is defined by the equation

$$d\xi_t = [\sigma_{(\xi_t)} + r_t\sigma](x_t^x)\,dw_t + [b_{(\xi_t)} + 2r_t b](x_t^x)\,dt, \qquad \xi_0 = \xi. \tag{4.10}$$

After formula (4.9) is obtained for bounded processes r_t, it can be extended for a wider class and we plug $r_t = \xi_t\alpha(x_t^x)$, where $\alpha(x)$ will be specified later, into (4.10) solve it and use the solution in (4.9). Similarly to what was said before, these manipulations can be easily justified if

$$b' + 2\alpha b \le \gamma c.$$

This is what (3.2) becomes in our case with $\mu = 0$.

We described the way how the parameters Θ and r appear. One can also use a change of probability measure based on Girsanov's theorem and then one includes in (3.2) an additional helping term $(a\xi, \xi)$ with as big factor as one likes.

More details in a more difficult case of controlled diffusion processes can be found in [18]. Note that in the above explanation in both cases in (4.5) and (4.10) we first found Θ and r in the form we like, then solved these equations and used thus specified Θ and r in (4.4) and (4.9). The same procedure works for controlled diffusion processes because it is known that one can use any progressively measurable Θ and r without affecting the value function. This property is unknown, however, for stochastic differential games. We can only use $\Theta=\Theta(x_t)$ and $r=r(x_t)$, which would not lead to any good result even in the above examples where Θ and r depend linearly on ξ. Therefore, what we actually do is that we consider the couple consisting of our processes issued from two different points and define Θ and r as functions of this couple. When the starting points are close we can almost recover the derivative of the initial process with respect to the initial data. Of course, the couple is a degenerate process and that is why in [12] and [10] we paid a special attention not to impose the nondegeneracy condition whenever it is not necessary.

In contrast with controlled diffusion processes, no version of random time change rule, change of Wiener process and Girsanov's theorem is known, and instead we can only rely on what the results of [10] allow one to extract from inspecting the corresponding Isaacs equations.

5. On equivalent representations of value functions. Here we suppose that Assumptions 2.1, 2.2, 2.3 and 2.5 are satisfied.

ASSUMPTION 5.1. There exists a nonnegative $G\in C(\bar D)\cap C^2_{\mathrm{loc}}(D)$ such that $G=0$ on ∂D (if $D\neq\mathbb{R}^d$) and

$$L^{\alpha\beta}G(p,x)\le -1$$

in D for all $p\in\mathbb{R}^k$, $\alpha\in A$ and $\beta\in B$.

Suppose that we are also given an $\mathbb{R}^{d_1}$-valued function $\pi^{\alpha\beta}(x,y)$ defined for $x,y\in\mathbb{R}^d$, $\alpha\in A$ and $\beta\in B$, which is bounded by K_0, Borel measurable, and Lipschitz continuous with respect to x with Lipschitz constant K_1.

Then for $\alpha_\cdot\in\mathfrak{A}$, $\beta_\cdot\in\mathfrak{B}$, $x,y\in\mathbb{R}^d$ introduce $y_t^{\alpha.\beta.y}=y_t^{\alpha.\beta.x,y}$ as a unique solution of the Itô equation

$$y_t=y+\int_0^t\sigma^{\alpha_s\beta_s}(y_s)\,dw_s+\int_0^t b^{\alpha_s\beta_s}(y_s)\,ds \tag{5.1}$$

and introduce $x_t^{\alpha.\beta.x,y}$ as a unique solution of the Itô equation (recall that $\hat\sigma$, $\hat b$, $\hat c$, $\hat f$ are introduced before Assumption 2.3)

$$x_t=x+\int_0^t\hat\sigma^{\alpha_s\beta_s}(x_s,y_s)\,dw_s+\int_0^t(\hat b-\hat\sigma\pi)^{\alpha_s\beta_s}(x_s,y_s)\,ds, \tag{5.2}$$

where, of course, $y_s=y_s^{\alpha.\beta.x,y}$. We emphasize that (5.1) has a unique solution since the coefficients are Lipschitz continuous in y and are bounded, and for given $y_\cdot$, equation (5.2) has a unique solution since its coefficients are Lipschitz continuous in x and are bounded. It follows that, in the terminology of [12], system (5.1)–(5.2) satisfies the usual hypothesis [although the coefficients in (5.2) may not be Lipschitz continuous with respect to the y variable].

With the above y_s and $x_s = x_s^{\alpha.\beta.x,y}$ also define

$$\phi_t^{\alpha.\beta.x,y} = \int_0^t \hat{c}^{\alpha_s\beta_s}(x_s, y_s)\, ds,$$

and for $z \in \mathbb{R}$ introduce $z_s^{\alpha.\beta.x,y,z}$ as a unique solution of

$$z_t = z + \int_0^t z_s[\pi^{\alpha_s\beta_s}(x_s, y_s)]^* \, dw_s. \tag{5.3}$$

Next, for $X = (x, y, z)$, $x, y \in \mathbb{R}^d$, $z \in \mathbb{R}$ denote

$$x_t^{\alpha.\beta.X} = x_t^{\alpha.\beta.x,y}, \qquad y_t^{\alpha.\beta.X} = y_t^{\alpha.\beta.y}, \qquad \phi_t^{\alpha.\beta.X} = \phi_t^{\alpha.\beta.x,y}$$
$$X_t^{\alpha.\beta.X} = (x_t, y_t, z_t)^{\alpha.\beta.X},$$

fix a number $M \in (1, \infty)$, for $X = (x, y, z)$ define $\tau^{\alpha.\beta.X}$ as the first exit time of $(x, z)_t^{\alpha.\beta.X}$ from $D \times (M^{-1}, M)$ and set

$$v^{\alpha.\beta.}(X) = E_X^{\alpha.\beta.}\left[\int_0^\tau \hat{f}(X_t)e^{-\phi_t}\, dt + z_\tau v(x_\tau)e^{-\phi_\tau}\right],$$

where $\hat{f}^{\alpha\beta}(x, y, z) = z\hat{f}^{\alpha\beta}(x, y)$, and v is taken as in Theorem 2.1 and is at least bounded and continuous according to the results of [10] and owing to Assumption 5.1. Finally, introduce

$$v(X) = \inf_{\boldsymbol{\beta}\in\mathbb{B}} \sup_{\alpha.\in\mathfrak{A}} v^{\alpha.\boldsymbol{\beta}(\alpha.)}(X).$$

The fact that $v^{\alpha\beta}(X)$ and $v(X)$ are well defined and bounded will be seen from the proof of the following.

THEOREM 5.1. *Under the above notation for $X = (x, y, z)$ we have*

$$v(X) = zv(x). \tag{5.4}$$

Furthermore, if we are given stopping times $\gamma^{\alpha.\beta.X} \le \tau^{\alpha.\beta.X}$, then

$$zv(x) = \inf_{\boldsymbol{\beta}\in\mathbb{B}} \sup_{\alpha.\in\mathfrak{A}} E_X^{\alpha.\boldsymbol{\beta}(\alpha.)}\left[\int_0^\gamma \hat{f}(X_t)e^{-\phi_t}\, dt + z_\gamma v(x_\gamma)e^{-\phi_\gamma}\right]. \tag{5.5}$$

PROOF. Introduce

$$(\mathbf{a}, \boldsymbol{\sigma}, \mathbf{b}, \mathbf{c}, \mathbf{f})^{\alpha\beta}(x, y) = (a, \sigma, b, c, f)^{\alpha\beta}\big(p^{\alpha\beta}(x, y), x\big) \tag{5.6}$$

(specifying the value of p transforms the letters to their boldface options). Also denote by $\mathcal{P}$ the set of triples $\check{p} = (r, \pi, P)$, where $r \in [\delta_1, \delta_1^{-1}]$, $\pi \in \mathbb{R}^{d_1}$ with $|\pi| \le K_0$ and $P \in \mathcal{O}$. For $\check{p} = (r, \pi, P) \in \mathcal{P}$ define

$$\check{\sigma}^{\alpha\beta}(\check{p}, x, y) = r\boldsymbol{\sigma}^{\alpha\beta}(x, y)P, \qquad \check{b}^{\alpha\beta}(\check{p}, x, y) = r^2\mathbf{b}^{\alpha\beta}(x, y) - r\boldsymbol{\sigma}^{\alpha\beta}(x, y)P\pi,$$
$$\check{c}^{\alpha\beta}(\check{p}, x, y, z) = r^2\mathbf{c}^{\alpha\beta}(x, y), \qquad \check{f}^{\alpha\beta}(\check{p}, x, y, z) = r^2z\mathbf{f}^{\alpha\beta}(x, y)$$

and also write

$$r = r(\check{p}), \qquad \pi = \pi(\check{p}), \qquad P = P(\check{p}).$$

We thus freed the coefficients of (5.2) of the particular values of r, π, P.

For each $\check{p} \in \mathcal{P}$ there is a natural operator $\check{L}^{\alpha\beta}$ acting on smooth functions $u(x, y, z)$ and mapping them to

$$\check{L}^{\alpha\beta} u(\check{p}, x, y, z)$$

associated with the matrix of second-order coefficients

$$\frac{1}{2} \begin{pmatrix} \check{\sigma}^{\alpha\beta}(\check{p}, x, y) \\ \sigma^{\alpha\beta}(y) \\ z\pi^*(\check{p}) \end{pmatrix} \begin{pmatrix} \check{\sigma}^{\alpha\beta}(\check{p}, x, y) \\ \sigma^{\alpha\beta}(y) \\ z\pi^*(\check{p}) \end{pmatrix}^*,$$

the drift term

$$\begin{pmatrix} \check{b}^{\alpha\beta}(\check{p}, x, y) \\ \check{b}^{\alpha\beta}(y) \\ 0 \end{pmatrix}$$

and the zeroth-order (killing) coefficient $-\check{c}^{\alpha\beta}(\check{p}, x, y, z)$. Introduce $\bar{p} = (1, 0, I)$ and

$$\bar{L}^{\alpha\beta} u(x, y, z) = \check{L}^{\alpha\beta} u(\bar{p}, x, y, z), \qquad \bar{f}^{\alpha\beta}(x, y, z) = \check{f}^{\alpha\beta}(\bar{p}, x, y, z).$$

We also need the operator $\mathbf{L}$ acting on functions $u(x, y)$ by the formula

$$\mathbf{L}^{\alpha\beta} u(x, y) = \mathbf{a}_{ij}^{\alpha\beta}(x, y) D_{ij} u(x, y) + \mathbf{b}_i^{\alpha\beta}(x, y) D_i u(x, y) - \mathbf{c}^{\alpha\beta}(x, y) u(x, y)$$

(no differentiation with respect to y is involved). Notice that, if $u = u(x)$ is a smooth function on $\mathbb{R}^d$ and $\check{u}(x, y, z) := zu(x)$, then as is easy to check

$$\check{L}^{\alpha\beta} \check{u}(\check{p}, x, y, z) = zr^2(\check{p})\big(\bar{L}^{\alpha\beta} u\big)(x, y, z) = zr^2(\check{p}) \mathbf{L}^{\alpha\beta} u(x, y). \tag{5.7}$$

One of consequences of Assumption 5.1 and (5.7) is that in $D \times \mathbb{R}^d \times (M^{-1}, M)$ we have

$$\check{L}^{\alpha\beta} \check{G}(\check{p}, x, y, z) \le -1$$

for all $\check{p}$, where $\check{G}(x, y, z) = M\delta_1^{-2} zG(x)$. In particular, this implies that $v^{\alpha\beta}(X)$ and $v(X)$ are well defined and are bounded.

Next, fix $x_0 \in D$, $y_0 \in \mathbb{R}^d$, and set

$$\check{p}_t^{\alpha.\beta.} = (r, \pi, P)^{\alpha_t \beta_t} (x, y)_t^{\alpha.\beta.x_0, y_0}.$$

As is easy to see, $\check{p}_t^{\alpha.\beta.}$ is a control adapted process in terminology of [12]; see Remark 2.3 there. For $\alpha. \in \mathfrak{A}$ and $\beta. \in \mathfrak{B}$, consider the following system of Itô's equations:

$$\begin{aligned} d\check{x}_t &= \check{\sigma}^{\alpha_t \beta_t}\big(\check{p}_t^{\alpha.\beta.}, \check{x}_t, \check{y}_t\big)\, dw_t + \check{b}^{\alpha_t \beta_t}\big(\check{p}_t^{\alpha.\beta.}, \check{x}_t, \check{y}_t\big)\, dt, \\ d\check{y}_t &= \sigma^{\alpha_t \beta_t}(\check{y}_t)\, dw_t + b^{\alpha_t \beta_t}(\check{y}_t)\, dt, \\ d\check{z}_t &= \check{z}_t \pi^*\big(\check{p}_t^{\alpha.\beta.}\big)\, dw_t. \end{aligned} \tag{5.8}$$

Its solution with initial condition $X = (x, y, z)$ will be denoted by

$$\check{X}_t^{\alpha.\beta.X} = (\check{x}, \check{y}, \check{z})_t^{\alpha.\beta.X}.$$

Observe that by uniqueness,

$$\check{X}_t^{\alpha.\beta.x_0,y_0,z} = X_t^{\alpha.\beta.x_0,y_0,z} \tag{5.9}$$

for any z. Also define

$$\check{\phi}_t^{\alpha.\beta.X} = \int_0^t \check{c}^{\alpha_t\beta_t}(\check{p}_s^{\alpha.\beta.}, \check{X}_s^{\alpha.\beta.X})\,ds,$$

$$\check{v}(X) = \inf_{\boldsymbol{\beta}\in\mathbb{B}} \sup_{\alpha.\in\mathfrak{A}} E_X^{\alpha.\boldsymbol{\beta}(\alpha.)}\left[\int_0^{\check{\tau}} \check{f}(\check{p}_t, \check{X}_t)e^{-\check{\phi}_t}\,dt + \check{z}_{\check{\tau}} v(\check{x}_{\check{\tau}})e^{-\check{\phi}_{\check{\tau}}}\right],$$

where $\check{\tau}^{\alpha.\beta.X}$ is the first exit time of $\check{X}_t^{\alpha.\beta.X}$ from $\check{D} = D \times \mathbb{R}^d \times (M^{-1}, M)$.

It turns out that, in the terminology of [12], for any $C^2_{\mathrm{loc}}(D)$ function $u = u(x)$, the function $zu(x)$ is p-insensitive in $\check{D}$ relative to $(zr^2(\check{p}), \check{L}^{\alpha\beta})$. This follows from the fact that, if $X \in \check{D}$, then by Itô's formula and (5.7), for $t < \check{\tau}^{\alpha.\beta.X}$,

$$\begin{aligned} &d\big(u(\check{x}_t^{\alpha.\beta.X})\check{z}_t^{\alpha.\beta.X}e^{-\check{\phi}_t^{\alpha.\beta.X}}\big) \\ &\quad = e^{-\check{\phi}_t}\check{z}_t^{\alpha.\beta.X} r^2(\check{p}_t^{\alpha.\beta.})(\bar{L}^{\alpha_t\beta_t}u)(\check{x}_t^{\alpha.\beta.X}, \check{y}_t^{\alpha.\beta.X}, \check{z}_t^{\alpha.\beta.X})\,dt + dm_t, \end{aligned}$$

where m_t is a local martingale starting at zero, and $zr^2(\check{p}) \in [M^{-1}\delta_1^2, M\delta_1^{-2}]$.

Furthermore, it turns out that equation (5.7) and Assumption 2.5 also imply that for smooth $u = u(x)$, if at a particular point x it holds that

$$J(x) := \sup_{\alpha\in A}\inf_{\beta\in B}\big[a_{ij}^{\alpha\beta}(x)D_{ij}u(x) + b_i^{\alpha\beta}(x)D_iu(x) - c^{\alpha\beta}(x)u(x) + f^{\alpha\beta}(x)\big] \le 0,$$

then with the same x, any y and $z > 0$, we also have

$$I(x, y, z) := \sup_{\alpha\in A}\inf_{\beta\in B}\big[\bar{L}^{\alpha\beta}\check{u}(x, y, z) + \bar{f}^{\alpha\beta}(x, y, z)\big] \le 0,$$

where $\check{u}(x, y, z) := zu(x)$. Indeed, since

$$\begin{aligned} J(x) = \sup_{\alpha\in A}\inf_{\beta\in B}\big[&\mathbf{a}_{ij}^{\alpha\beta}(x, x)D_{ij}u(x) + \mathbf{b}_i^{\alpha\beta}(x, x)D_iu(x) - \mathbf{c}^{\alpha\beta}(x, x)u(x) \\ &+ \mathbf{f}^{\alpha\beta}(x, x)\big], \end{aligned}$$

the inequality $J(x) \le 0$ implies by Assumption 2.5 that

$$\begin{aligned} \sup_{\alpha\in A}\inf_{\beta\in B}\big[&\mathbf{a}_{ij}^{\alpha\beta}(x, y)D_{ij}u(x) + \mathbf{b}_i^{\alpha\beta}(x, y)D_iu(x) - \mathbf{c}^{\alpha\beta}(x, y)u(x) \\ &+ \mathbf{f}^{\alpha\beta}(x, y)\big] \le 0, \end{aligned}$$

and it only remains to notice that the left-hand side is just $z^{-1}I(x,y,z)$. Similarly, $J(x)\geq 0$ implies that $I(x,y,z)\geq 0$.

These facts combined imply by Theorems 2.3 and 3.1 of [10] that for all $x\in\bar{D}$, $y\in\mathbb{R}^d$ and $z\in[M^{-1},M]$ we have

$$\check{v}(x,y,z)=zv(x)$$

and, for any stopping times $\gamma^{\alpha.\beta.X}\leq\check{\tau}^{\alpha.\beta.X}$,

$$zv(x)=\inf_{\boldsymbol{\beta}\in\mathbb{B}}\sup_{\alpha.\in\mathfrak{A}}E_X^{\alpha.\boldsymbol{\beta}(\alpha.)}\left[\int_0^\gamma\check{f}(\check{p}_t,\check{X}_t)e^{-\check{\phi}_t}\,dt+\check{z}_\gamma v(\check{x}_\gamma)e^{-\check{\phi}_\gamma}\right].\tag{5.10}$$

By (5.9) for $X_0=(x_0,y_0,z_0)$, $z_0\in[M^{-1},M]$, we have

$$\check{X}^{\alpha.\beta.X_0}=X^{\alpha.\beta.X_0},\qquad \check{f}^{\alpha_t\beta_t}(\check{p}_t^{\alpha.\beta.},\check{X}_t^{\alpha.\beta.X_0})=\hat{f}(X_t^{\alpha.\beta.X_0}),$$
$$\check{\phi}^{\alpha.\beta.X_0}=\phi^{\alpha.\beta.X_0},$$

so that $v(x_0,y_0,z_0)=\check{v}(x_0,y_0,z_0)$. It follows that (5.4) holds at $(x_0,y_0,z_0)\in D^{\check{}}$. Outside $D^{\check{}}$ the equality is obvious. Finally, (5.5) follows from (5.10), and the theorem is proved. □

REMARK 5.1. One of assumptions in Theorems 2.3 and 3.1 of [10] is that the coefficients satisfy Assumption 2.1(i) without $p^{\alpha\beta}(x,y)$ there. Since p is involved in (5.6) we needed to include it in Assumption 2.1(i) in contrast with the parameters $r^{\alpha\beta}(x,y)$ and $P^{\alpha\beta}(x,y)$. The same reasons caused the last requirement in Assumption 2.1(ii). Recall that in Theorems 2.3 and 3.1 of [10] the coefficients of Itô equations are not supposed to be Lipschitz, but rather uniformly continuous.

6. Estimating the difference of solutions of stochastic equations whose coefficients are close.

Suppose that on $\Omega\times(0,\infty)\times\mathbb{R}^d\times\mathbb{R}^d$ we are given the following functions: $d\times d_1$ matrix-valued $\sigma_t(x,y)$, $\mathbb{R}^d$-valued $b_t(x,y)$ and real-valued functions $c_t(x,y)\geq\delta_1$, $f_t(x,y)$, where $\delta_1>0$ is a fixed constant.

Introduce

$$(\sigma_t,b_t,c_t,f_t)(x)=(\sigma_t,b_t,c_t,f_t)(x,x),\qquad a_t(x)=(1/2)\sigma_t\sigma_t^*(x).$$

ASSUMPTION 6.1. (i) All the above functions are measurable with respect to the product of $\mathcal{F}$ and Borel σ-algebras on $(0,\infty)$, $\mathbb{R}^d$ and $\mathbb{R}^d$, and they are progressively measurable as functions of (ω,t) for each (x,y).

(ii) All the above functions are bounded by a constant K_0.

(iii) For any $t>0$, $x',x'',y\in\mathbb{R}^d$ and

$$\xi_t=(\sigma_t,b_t)(x,y),\qquad \eta_t=(\sigma_t,b_t)(x),$$

we have

$$|\xi_t(x',y)-\xi_t(x'',y)|+|\eta_t(x')-\eta_t(x'')|\leq K_1|x'-x''|,$$

where K_1 is a fixed constant. Also there exists a constant $\varepsilon_0 > 0$ such that for any $t > 0$ and $x, y \in \mathbb{R}^d$ with $|x - y| \le \varepsilon_0$, we have

$$\left|c_t(x, y) - c_t(y)\right| + \left|f_t(x, y) - f_t(y)\right| \le K_1|x - y|.$$

Observe that Assumption 6.1(iii) implies, in particular, that $|b_t(x, y) - b_t(y)| \le K_1|x - y|$.

ASSUMPTION 6.2. There exist constants $\mu \ge 1$ and $\delta \ge 2\delta_1$ such that for all $x, y \in \mathbb{R}^d$ satisfying $|x - y| \le \varepsilon_0$ we have

$$\begin{aligned} R_t(x, y) :=& \left\|\sigma_t(x, y) - \sigma_t(y)\right\|_\xi^2 + 2\langle x - y, b_t(x, y) - b_t(y)\rangle \\ & - 4\mu\langle x - y, a_t(x)(x - y)\rangle \\ \le& \ 2\bigl(c_t(y) - \delta\bigr)|x - y|^2, \end{aligned} \tag{6.1}$$

where $\xi = (x - y)/|x - y|$.

Fix a unit $\xi \in \mathbb{R}^d$, and for $\varepsilon \in [0, \varepsilon_0]$ introduce x_t^ε as a unique solution of

$$x_t = \varepsilon\xi + \int_0^t \sigma_s(x_s, y_s)\, dw_s + \int_0^t \bigl[b_s(x_s, y_s) - 2\mu a_s(x_s)(x_s - y_s)\bigr]\, ds,$$

where y_s is a unique solution of

$$y_t = \int_0^t \sigma_s(y_s)\, dw_s + \int_0^t b_s(y_s)\, ds.$$

Observe that owing to uniqueness,

$$x_t^0 = y_t.$$

For $\varepsilon > 0$ define

$$\xi_t^\varepsilon = \frac{1}{\varepsilon}\bigl(x_t^\varepsilon - x_t^0\bigr), \qquad \phi_t = \int_0^t c_s\bigl(x_s^0\bigr)\, ds,$$

and for $\lambda > 0$ let

$$\kappa_\varepsilon(\lambda) = \inf\bigl\{t \ge 0 : \bigl|x_t^\varepsilon - x_t^0\bigr| \ge \lambda\bigr\}.$$

Notice that $\kappa_\varepsilon(\lambda) = 0$ if $\lambda \le \varepsilon$, and start with the following:

LEMMA 6.1. *For any* $\lambda \in (0, \varepsilon_0]$

$$J_\varepsilon := E\int_0^{\kappa_\varepsilon(\lambda)} \bigl|\xi_t^\varepsilon\bigr| e^{-\phi_t + \delta t/2}\, dt \le 2/\delta, \tag{6.2}$$

$$I_\varepsilon := E \sup_{t < \kappa_\varepsilon(\lambda)} \bigl|\xi_t^\varepsilon\bigr| e^{-\phi_t + \delta t/2} \le N, \tag{6.3}$$

where N is a constant depending only on K_1 and δ.

PROOF. We have

$$d\xi_t^\varepsilon = \varepsilon^{-1}[\sigma_t(x_t^\varepsilon, x_t^0) - \sigma_t(x_t^0)]\,dw_t \tag{6.4}$$
$$+ \varepsilon^{-1}[b_t(x_t^\varepsilon, x_t^0) - b(x_t^0) - 2\mu a_t(x_t^\varepsilon)(x_t^\varepsilon - x_t^0)]\,dt,$$

where the magnitudes of the coefficients of dw_t and dt are dominated by constants times $|\xi_t^\varepsilon|$. This allows us to use Itô's formula (cf. the proof of Theorem 5.8.7 of [8]) and obtain that $(0/0 := 0)$

$$\begin{aligned} d|\varepsilon\xi_t^\varepsilon|e^{-\phi_t+\delta t/2} &= \frac{1}{2|x_t^\varepsilon - x_t^0|}[R_t(x_t^\varepsilon, x_t^0) - 2(c_t(x_t^0) - \delta/2)|x_t^\varepsilon - x_t^0|^2]e^{-\phi_t+\delta t/2}\,dt \\ &\quad + S_t(x_t^\varepsilon, x_t^0)e^{-\phi_t+\delta t/2}\,dw_t, \end{aligned}$$

where

$$S_t(x_t^\varepsilon, x_t^0) = \frac{1}{|\xi_t^\varepsilon|}\xi_t^{\varepsilon *}[\sigma_t(x_t^\varepsilon, x_t^0) - \sigma_t(x_t^0)].$$

By assumption, for $t < \kappa_\varepsilon(\lambda)$ we have

$$R_t(x_t^\varepsilon, x_t^0) - 2(c_t(x_t^0) - \delta/2)|x_t^\varepsilon - x_t^0|^2 \le -\delta|x_t^\varepsilon - x_t^0|^2.$$

It follows that for $t < \kappa_\varepsilon(\lambda)$,

$$d|\xi_t^\varepsilon|e^{-\phi_t+\delta t/2} \le -(\delta/2)|\xi_t^\varepsilon|e^{-\phi_t+\delta t/2}\,dt + \varepsilon^{-1}S_t(x_t^\varepsilon, x_t^0)e^{-\phi_t+\delta t/2}\,dw_t. \tag{6.5}$$

In particular, (6.2) holds. Furthermore,

$$|\varepsilon^{-1}S_t(x_t^\varepsilon, x_t^0)| \le K_1|\xi_t^\varepsilon|, \tag{6.6}$$

and by Davis's inequality,

$$\begin{aligned} I_\varepsilon &\le 3K_1 E\left(\int_0^{\kappa_\varepsilon(\lambda)} |\xi_t^\varepsilon|^2 e^{-2\phi_t+\delta t}\,dt\right)^{1/2} \\ &\le 3K_1 E\Big(\sup_{s<\kappa_\varepsilon(\lambda)} |\xi_s^\varepsilon|e^{-\phi_s+\delta s/2}\Big)^{1/2}\left(\int_0^{\kappa_\varepsilon(\lambda)} |\xi_t^\varepsilon|e^{-\phi_t+\delta t/2}\,dt\right)^{1/2} \le N I_\varepsilon^{1/2} J_\varepsilon^{1/2}, \end{aligned}$$

which, due to (6.2), proves (6.3) and the lemma. □

COROLLARY 6.2. *For $\lambda > 0$ we have*

$$Ee^{-\phi_{\kappa_\varepsilon(\lambda)}+\kappa_\varepsilon(\lambda)\delta/2} I_{\kappa_\varepsilon(\lambda)<\infty} \le N\varepsilon/\lambda.$$

Indeed, if $\lambda \le \varepsilon$, the estimate is obvious since $\kappa_\varepsilon(\lambda) = 0$ and for $\lambda > \varepsilon$

$$\lambda Ee^{-\phi_{\kappa_\varepsilon(\lambda)}+\kappa_\varepsilon(\lambda)\delta/2} I_{\kappa_\varepsilon(\lambda)<\infty} = \varepsilon E|\xi_{\kappa_\varepsilon(\lambda)}^\varepsilon|e^{-\phi_{\kappa_\varepsilon(\lambda)}+\kappa_\varepsilon(\lambda)\delta/2} I_{\kappa_\varepsilon(\lambda)<\infty} \le N\varepsilon.$$

REMARK 6.1. If $\delta \ge K_1^2$, then it follows from (6.5) and (6.6) that for $t < \kappa_\varepsilon(\lambda)$ we have

$$d|\xi_t^\varepsilon|^2 e^{-2\phi_t+\delta t} \le dm_t,$$

where m_t is a local martingale. Hence, for any stopping time $\gamma \le \kappa_\varepsilon(\lambda)$,

$$E|\xi_\gamma^\varepsilon|^2 e^{-2\phi_\gamma+\delta\gamma} \le 1.$$

Psychologically, the condition $\delta \ge K_1^2$ may look artificial. However, in the proof of Theorem 2.2 the parameter δ will be, basically, sent to infinity.

Next introduce

$$\pi_s(x, y) = \mu\sigma_s^*(x)(x - y)$$

and introduce ρ_t^ε as a unique solution of

$$\rho_t = 1 + \int_0^t \rho_s \pi_s^*(x_s^\varepsilon, x_s^0)\, dw_s + \int_0^t \rho_s\big[c_s(x_s^0) - c_s(x_s^\varepsilon, x_s^0)\big]\, ds.$$

Take a constant $M > 1$ and define

$$\gamma_\varepsilon(M)$$

as the first exit time of ρ_t^ε from (M^{-1}, M).

Recall that $c \ge \delta_1$.

LEMMA 6.3. *There exists $\lambda_1 \in (0, \varepsilon_0]$, depending only on ε_0, K_0, K_1 and δ_1, and there exists a constant N, depending only on K_1 and δ_1, such that for $\lambda = \lambda_1/\mu$ and $\mu \ge 1$ we have*

$$I := E \sup_{t<\gamma_\varepsilon(M)\wedge\kappa_\varepsilon(\lambda)} |\rho_t^\varepsilon - 1| e^{-\phi_t+\delta_1 t/2} \le N\big(M\mu^2 + 1\big)^{1/2}\delta^{-1/2}\varepsilon. \tag{6.7}$$

PROOF. Denote $C_t(x_t^\varepsilon, x_t^0) = c_t(x_t^0) - c_t(x_t^\varepsilon, x_t^0)$ and $\eta_t = (\rho_t^\varepsilon - 1)^2$. Then

$$\begin{aligned} d\eta_t &= 2(\rho_t^\varepsilon - 1)\rho_t^\varepsilon \pi_t^*(x_t^\varepsilon, x_t^0)\, dw_t + 2(\rho_t^\varepsilon - 1)\rho_t^\varepsilon C_t(x_t^\varepsilon, x_t^0)\, dt \\ &\quad + |\rho_t^\varepsilon|^2 |\pi_t(x_t^\varepsilon, x_t^0)|^2\, dt, \\ d\eta_t e^{-2\phi_t+\delta_1 t} &= e^{-2\phi_t+\delta_1 t}\big[2\eta_t C_t(x_t^\varepsilon, x_t^0) + 2(\rho_t^\varepsilon - 1)C_t(x_t^\varepsilon, x_t^0) \\ &\quad + \eta_t|\pi_t(x_t^\varepsilon, x_t^0)|^2 + 2(\rho_t^\varepsilon - 1)|\pi_t(x_t^\varepsilon, x_t^0)|^2 \\ &\quad + |\pi_t(x_t^\varepsilon, x_t^0)|^2 - \eta_t(2c_t(x_t^0) - \delta_1)\big]\, dt + dm_t, \end{aligned}$$

where m_t is a local martingale starting at zero, and for $t < \gamma_\varepsilon(M)$, the expression in the square brackets is less than

$$\begin{aligned} &\eta_t\big[2C_t(x_t^\varepsilon, x_t^0) + \delta_1/2 + |\pi_t(x_t^\varepsilon, x_t^0)|^2 - (2c_t(x_t^0) - \delta_1)\big] \\ &\quad + (2/\delta_1)C_t^2(x_t^\varepsilon, x_t^0) + (2M - 1)|\pi_t(x_t^\varepsilon, x_t^0)|^2. \end{aligned}$$

We have that $|G_t| \le K_1|x_t^\varepsilon - x_t^0|$, $|\pi_t| \le \mu K_0|x_t^\varepsilon - x_t^0|$, $c \ge \delta_1$ and $\mu \ge 1$ and, therefore, one can find $\lambda_1 \in (0, \varepsilon_0]$ such that, for $\lambda = \lambda_1/\mu$ and $t < \kappa_\varepsilon(\lambda)$,

$$2C_t(x_t^\varepsilon, x_t^0) + \delta_1/2 + |\pi_t(x_t^\varepsilon, x_t^0)|^2 - (2c_t(x_t^0) - \delta_1) \le 0$$

and then

$$d\eta_t e^{-2\phi_t+\delta_1 t} \le N_1(M\mu^2+1)\varepsilon^2|\xi_t^\varepsilon|^2 e^{-2\phi_t+\delta_1 t}\,dt + dm_t.$$

Hence, for any bounded stopping time τ it holds that

$$\begin{aligned} &E\eta_{\tau\wedge\gamma_\varepsilon(M)\wedge\kappa_\varepsilon(\lambda)} e^{-2\phi_{\tau\wedge\gamma_\varepsilon(M)\wedge\kappa_\varepsilon(\lambda)}+\delta_1(\tau\wedge\gamma_\varepsilon(M)\wedge\kappa_\varepsilon(\lambda))} \\ &\qquad \le N_1(M\mu^2+1)\varepsilon^2 E\int_0^{\tau\wedge\gamma_\varepsilon(M)\wedge\kappa_\varepsilon(\lambda)} |\xi_t^\varepsilon|^2 e^{-2\phi_t+\delta_1 t}\,dt, \end{aligned}$$

which owing to well-known properties of such inequalities (see, e.g., Theorem 3.6.8 in [8]) implies that

$$\begin{aligned} &E\sup_{t\le\gamma_\varepsilon(M)\wedge\kappa_\varepsilon(\lambda)} \eta_t^{1/2} e^{-\phi_t+\delta_1 t/2} \\ &\qquad \le 3N_1(M\mu^2+1)^{1/2}\varepsilon E\left(\int_0^{\kappa_\varepsilon(\lambda)} |\xi_t^\varepsilon|^2 e^{-2\phi_t+\delta_1 t}\,dt\right)^{1/2}. \end{aligned}$$

Owing to (6.3) and the assumption that $\delta \ge 2\delta_1$, the last expectation is dominated by

$$N\left(\int_0^\infty e^{(\delta_1-\delta)t}\,dt\right)^{1/2} \le N\delta^{-1/2}.$$

The lemma is proved. □

COROLLARY 6.4. *There is a constant N, depending only on K_1 and δ_1, such that for any $M \ge 2$ and $\lambda = \lambda_1/\mu$*

$$Ee^{-\phi_{\gamma_\varepsilon(M)\wedge\kappa_\varepsilon(\lambda)}} \le N[\mu + (M\mu^2+1)^{1/2}\delta^{-1/2}]\varepsilon. \tag{6.8}$$

To prove (6.8), it suffices to notice that

$$\begin{aligned} Ee^{-\phi_{\gamma_\varepsilon(M)\wedge\kappa_\varepsilon(\lambda)}} I_{\gamma_\varepsilon(M)<\kappa_\varepsilon(\lambda)} &\le M(M-1)^{-1}E|\rho^\varepsilon_{\gamma_\varepsilon(M)} - 1|e^{-\phi_{\gamma_\varepsilon(M)}} I_{\gamma_\varepsilon(M)<\kappa_\varepsilon} \\ &\le M(M-1)^{-1}E\sup_{t<\gamma_\varepsilon(M)\wedge\kappa_\varepsilon(\lambda)} |\rho_t^\varepsilon - 1|e^{-\phi_t} \end{aligned}$$

and then to use Corollary 6.2 and to recall that $c \ge \delta_1$.

Now for $\lambda = \lambda_1/\mu$, $\varepsilon \in (0, \varepsilon_0]$, and $M \ge 2$ take a stopping time

$$\tau \le \gamma_\varepsilon(M) \wedge \kappa_\varepsilon(\lambda).$$

Also take a function $g_t(x)$, which is measurable in (ω, t, x) and such that $|g| \le K_0$ and introduce

$$v^\varepsilon = E\biggl[\int_0^\tau z_t^\varepsilon f(x_t^\varepsilon, x_t^0) e^{-\phi_t^\varepsilon}\, dt + z_\tau^\varepsilon g_\tau(x_\tau^\varepsilon) e^{-\phi_\tau^\varepsilon}\biggr],$$

where

$$\phi_t^\varepsilon = \int_0^t c_s(x_s^\varepsilon, x_s^0)\, ds$$

and z_t^ε is defined as a unique solution of

$$z_t = 1 + \int_0^t z_s \pi_s^*(x_s^\varepsilon, x_s^0)\, dw_s.$$

Finally, define

$$v^0 = E\biggl[\int_0^\tau f(x_t^0) e^{-\phi_t}\, dt + g_\tau(x_\tau^0) e^{-\phi_\tau}\biggr].$$

THEOREM 6.5. *Suppose that there is a constant N_0 such that*

$$E\bigl|g_\tau(x_\tau^\varepsilon) - g_\tau(x_\tau^0)\bigr| e^{-\phi_\tau} I_{\tau < \gamma_\varepsilon(M) \wedge \kappa_\varepsilon(\lambda)} \le N_0 \varepsilon. \tag{6.9}$$

Then there exists a constant N, depending only on K_0, K_1 and δ_1, such that for $\lambda = \lambda_1/\mu$ we have

$$\bigl|v^\varepsilon - v^0\bigr| \le N_0\varepsilon + N\bigl[\mu + (M\mu^2 + 1)^{1/2}\delta^{-1/2} + \delta^{-1}\bigr]\varepsilon.$$

PROOF. First notice that

$$z_t^\varepsilon e^{-\phi_t^\varepsilon} = \rho_t^\varepsilon e^{-\phi_t},$$

so that

$$\biggl|\int_0^\tau \bigl[z_t^\varepsilon f(x_t^\varepsilon, x_t^0) e^{-\phi_t^\varepsilon} - f(x_t^0) e^{-\phi_t}\bigr]\, dt\biggr| \le I_\varepsilon + J_\varepsilon,$$

where

$$I_\varepsilon = \int_0^\tau |\rho_t^\varepsilon - 1| |f(x_t^\varepsilon, x_t^0)| e^{-\phi_t}\, dt,$$

$$J_\varepsilon = \int_0^\tau |f(x_t^\varepsilon, x_t^0) - f(x_t^0)| e^{-\phi_t}\, dt.$$

By Lemma 6.3,

$$\begin{aligned} E I_\varepsilon &\le N E \sup_{s\le\tau} |\rho_s^\varepsilon - 1| e^{-\phi_s + \delta_1 s/2} \int_0^\infty e^{-\delta_1 t/2}\, dt \\ &\le N(M\mu^2 + 1)^{1/2}\delta^{-1/2}\varepsilon. \end{aligned}$$

By Lemma 6.1,

$$E J_\varepsilon \le N\varepsilon E\int_0^\tau |\xi_t^\varepsilon| e^{-\phi_t}\,dt \le N\varepsilon/\delta.$$

Next

$$\begin{aligned} E\big|z_\tau^\varepsilon g_\tau(x_\tau^\varepsilon)e^{-\phi_\tau^\varepsilon} - g_\tau(x_\tau^0)e^{-\phi_\tau}\big| &= E\big|\rho_\tau^\varepsilon g_\tau(x_\tau^\varepsilon) - g_\tau(x_\tau^0)\big|e^{-\phi_\tau} \\ &\le K_0 E|\rho_\tau^\varepsilon - 1|e^{-\phi_\tau} + E\big|g_\tau(x_\tau^\varepsilon) - g_\tau(x_\tau^0)\big|e^{-\phi_\tau}, \end{aligned}$$

where the first term is estimated as above and, owing to (6.9), the second term is dominated by

$$\begin{aligned} & N_0\varepsilon + E\big|g_\tau(x_\tau^\varepsilon) - g_\tau(x_\tau^0)\big|e^{-\phi_\tau} I_{\tau=\gamma_\varepsilon(M)\wedge\kappa_\varepsilon(\lambda)} \\ &\qquad \le N_0\varepsilon + 2K_0 E e^{-\phi_{\gamma_\varepsilon(M)\wedge\kappa_\varepsilon(\lambda)}} \le N_0\varepsilon + N\big[\mu + (M\mu^2+1)^{1/2}\delta^{-1/2}\big]\varepsilon, \end{aligned}$$

with the second inequality following from Corollary 6.4. The theorem is proved. □

7. Proof of Theorem 2.1. According to Remark 2.1, in the proof of Theorem 2.1 we may assume that $c^{\alpha\beta}(x) \ge \delta_1$.

First, we estimate the Lipschitz constant of v on the boundary when $D \ne \mathbb{R}^d$.

LEMMA 7.1. *Let D be bounded and satisfy the uniform exterior ball condition. Let $x \in \mathbb{R}^d$ and $y \notin D$. Then there is a constant N depending only on D, K_0 and $\|g\|_{C^2(\mathbb{R}^d)}$, such that*

$$\big|v(x) - v(y)\big| \le N|x-y|.$$

PROOF. If $x \notin D$, then $|v(x) - v(y)| = |g(x) - g(y)| \le N|x-y|$. Therefore in the rest of the proof we assume that $x \in D$. Then observe that by Itô's formula we have

$$v(x) = g(x) + \inf_{\boldsymbol{\beta}\in\mathbb{B}}\sup_{\alpha.\in\mathfrak{A}} E_x^{\alpha.\boldsymbol{\beta}(\alpha.)}\int_0^\tau \big[Lg(x_t) + f(x_t)\big]e^{-\phi_t}\,dt. \tag{7.1}$$

It is well known that, in light of the boundedness of $L^{\alpha\beta}g + f^{\alpha\beta}$ and D and the uniform exterior ball condition, the expectations in (7.1) by magnitude are dominated by a constant times $\operatorname{dist}(x, \partial D) \le |x-y|$. This proves the lemma since $v(y) = g(y)$ and $|g(x) - g(y)| \le N|x-y|$. □

PROOF OF THEOREM 2.1. In Section 5 take

$$r \equiv 1, \qquad p \equiv 0, \qquad P \equiv I, \qquad \pi^{\alpha\beta}(x,y) = \mu\big[\sigma^{\alpha\beta}(x)\big]^*(x-y),$$

where the constant $\mu \ge 1$ is chosen to be such that (6.1) with $\delta = 1$ and

$$(\sigma_t, b_t)(x,y) = (\sigma, b)^{\alpha_t\beta_t}(x)$$

holds for all $\alpha. \in \mathfrak{A}$, $\beta. \in \mathfrak{B}$, x and y. This is possible since σ and b are Lipschitz continuous, and a is uniformly nondegenerate. In Section 5 we required $\pi^{\alpha\beta}(x, y)$ to be bounded and Lipschitz continuous with respect to x. Since we will be only concerned with its values for $|x - y| \le 1$, we can appropriately modify the above $\pi^{\alpha\beta}(x, y)$ for $|x - y| \ge 1$ keeping the same notation.

Then for a unit $\xi \in \mathbb{R}^d$, $\varepsilon \ge 0$, $\alpha. \in \mathfrak{A}$ and $\beta. \in \mathfrak{B}$ introduce $x_t^{\alpha.\beta.0}(\varepsilon)$ as a unique solution of

$$x_t = \varepsilon\xi + \int_0^t \sigma^{\alpha_s\beta_s}(x_s)\,dw_s + \int_0^t \bigl[b^{\alpha_s\beta_s}(x_s) - \sigma^{\alpha_s\beta_s}(x_s)\pi^{\alpha_s\beta_s}(x_s, y_s)\bigr]\,ds,$$

where

$$y_s = x_s^{\alpha.\beta.0}.$$

Next introduce

$$\phi_t^{\alpha.\beta.0}(\varepsilon) = \int_0^t c^{\alpha_s\beta_s}\bigl(x_s^{\alpha.\beta.0}(\varepsilon)\bigr)\,ds,$$

and let $z_t^{\alpha.\beta.0}(\varepsilon)$ be a unique solution of

$$z_t = 1 + \int_0^t z_s\bigl[\pi^{\alpha_s,\beta_s}\bigl(x_s^{\alpha.\beta.0}(\varepsilon), x_s^{\alpha.\beta.0}(0)\bigr)\bigr]^*\,dw_s.$$

Keeping in mind that μ is already fixed, set $\delta_1 := \varepsilon_1 = 1$, take λ from Lemma 6.3, fix $\varepsilon \in (0, 1]$ and introduce

$$\begin{aligned}
\tau_\varepsilon^{\alpha.\beta.0} &= \inf\{t \ge 0 : x_t^{\alpha.\beta.0}(\varepsilon) \notin D\},\\
\gamma_\varepsilon^{\alpha.\beta.0} &= \inf\{t \ge 0 : z_t^{\alpha.\beta.0}(\varepsilon)e^{\phi_t^{\alpha.\beta.0}(0) - \phi_t^{\alpha.\beta.0}(\varepsilon)} \notin (1/2, 2)\},\\
\kappa_\varepsilon^{\alpha.\beta.0} &= \inf\{t \ge 0 : \bigl|x_t^{\alpha.\beta.0}(\varepsilon) - x_t^{\alpha.\beta.0}(0)\bigr| \ge \lambda\},\\
\gamma^{\alpha.\beta.0} &= \tau_\varepsilon^{\alpha.\beta.0} \wedge \tau_0^{\alpha.\beta.0} \wedge \kappa_\varepsilon^{\alpha.\beta.0} \wedge \gamma_\varepsilon^{\alpha.\beta.0}.
\end{aligned}$$

By Theorem 5.1,

$$\begin{aligned}
v(\varepsilon\xi) = \inf_{\boldsymbol{\beta}\in\mathbb{B}}\sup_{\alpha.\in\mathfrak{A}} E_0^{\alpha.\boldsymbol{\beta}(\alpha.)}\biggl[\int_0^\gamma z_t(\varepsilon) f\bigl(x_t(\varepsilon)\bigr)e^{-\phi_t(\varepsilon)}\,dt\\
+ z_\gamma(\varepsilon)v\bigl(x_\gamma(\varepsilon)\bigr)e^{-\phi_\gamma(\varepsilon)}\biggr].
\end{aligned} \tag{7.2}$$

Next we fix $\alpha. \in \mathfrak{A}$ and $\beta. \in \mathfrak{B}$, and in Section 6 use the functions

$$(\sigma_t, b_t, c_t, f_t)(x, y) = (\sigma, b, c, f)^{\alpha_t\beta_t}(x).$$

Observe that in the expectation

$$E_0^{\alpha.\beta.}\biggl[\int_0^\gamma z_t(\varepsilon) f\bigl(x_t(\varepsilon)\bigr)e^{-\phi_t(\varepsilon)}\,dt + z_\gamma(\varepsilon)v\bigl(x_\gamma(\varepsilon)\bigr)e^{-\phi_\gamma(\varepsilon)}\biggr],$$

one can replace $x_s^{\alpha.\beta.0}(\varepsilon)$ with x_t^{ε} since both satisfy the same equation on $[0, \gamma^{\alpha.\beta.0}]$, and by Theorem 6.5 we get that

$$\begin{aligned}\bigg| E_0^{\alpha.\beta.}\bigg[\int_0^{\gamma} z_t(\varepsilon) f\big(x_t(\varepsilon)\big)e^{-\phi_t(\varepsilon)}\,dt + z_\gamma(\varepsilon)v\big(x_\gamma(\varepsilon)\big)e^{-\phi_\gamma(\varepsilon)}\bigg] \\ - E_0^{\alpha.\beta.}\bigg[\int_0^{\gamma} f(x_t)e^{-\phi_t}\,dt + v(x_\gamma)e^{-\phi_\gamma}\bigg]\bigg| \\ \le N\varepsilon + E_0^{\alpha.\beta.}\big|v\big(x_\gamma(\varepsilon)\big) - v\big(x_\gamma(0)\big)\big|e^{-\phi_\gamma} I_{\gamma<\gamma_\varepsilon\wedge\kappa_\varepsilon}.\end{aligned} \tag{7.3}$$

If $t = \gamma^{\alpha.\beta.0} < \gamma_\varepsilon^{\alpha.\beta.0} \wedge \kappa_\varepsilon^{\alpha.\beta.0}$, then ($D \ne \mathbb{R}^d$ and) at least one of $x_t^{\alpha.\beta.0}(\varepsilon)$ and $x_t^{\alpha.\beta.0}(0)$ is outside D, and by Lemma 7.1 we obtain

$$\begin{aligned} &E_0^{\alpha.\beta.}\big|v\big(x_\gamma(\varepsilon)\big) - v\big(x_\gamma(0)\big)\big|e^{-\phi_\gamma} I_{\gamma<\gamma_\varepsilon\wedge\kappa_\varepsilon} \\ &\quad\le N E_0^{\alpha.\beta.}\big|x_\gamma(\varepsilon) - x_\gamma(0)\big|e^{-\phi_\gamma} I_{\gamma<\gamma_\varepsilon\wedge\kappa_\varepsilon} \\ &\quad= N\varepsilon E_0^{\alpha.\beta.}\big|\xi_\gamma(\varepsilon)\big|e^{-\phi_\gamma} I_{\gamma<\gamma_\varepsilon\wedge\kappa_\varepsilon} \le N\varepsilon E_0^{\alpha.\beta.}\sup_{t<\kappa_\varepsilon}\big|\xi_t(\varepsilon)\big|e^{-\phi_t},\end{aligned}$$

where $\varepsilon\xi_t^{\alpha.\beta.0}(\varepsilon) = x_t^{\alpha.\beta.0}(\varepsilon) - x_t^{\alpha.\beta.0}(0)$. By using Lemma 6.1, equation (7.3), and the fact that $\alpha.$ and $\beta.$ in the above argument are arbitrary, we see that $|v(\varepsilon\xi) - v(0)| \le N\varepsilon$. Similarly one proves that $|v(x + \varepsilon\xi) - v(x)| \le N\varepsilon$ for any x, which is what we need. The theorem is proved. □

8. Proof of Theorem 2.2. In contrast with Section 7, where we used $\delta = 1$, here δ will be chosen large. We begin with the following.

LEMMA 8.1. *Let D be a bounded domain satisfying the uniform exterior ball condition, and let $\|g\|_{C^2(\mathbb{R}^d)} < \infty$. For $R \in (0, 1]$ let $B_R = \{x : |x| \le R\}$. Assume that for an R we have $B_R \subset D$ and denote by L_R the Lipschitz constant of v in B_R (finite by Theorem 2.1). Finally assume that $|v| \le K_0$ in B_R.*

Then for any $\delta \ge K_1^2 + 4K_0^2 + 2$ we have

$$\varlimsup_{x\to 0}\frac{|v(x) - v(0)|}{|x|} \le N\delta R^{-1} + Ne^{-\nu\sqrt{\delta}}L_R, \tag{8.1}$$

where N and $\nu > 0$ depend only on d, K_0, K_1 and δ_0.

PROOF. First suppose that $R = 1$. Observe that by the dynamic programming principle

$$v(x) = \inf_{\boldsymbol{\beta}\in\mathbb{B}}\sup_{\alpha.\in\mathfrak{A}} E_x^{\alpha.\boldsymbol{\beta}(\alpha.)}\bigg[\int_0^{\tau_1} f(x_t)e^{-\phi_t}\,dt + v\big(x_{\tau_1}(\varepsilon)\big)e^{-\phi_{\tau_1}}\bigg], \tag{8.2}$$

where $\tau_1^{\alpha.\beta.x}$ is the first exit time of $x_t^{\alpha.\beta.x}$ from B_1.

Remark 2.1 allows us to rewrite (8.2) by using a global barrier for B_1 for a slightly modified v. Obviously, if we can prove (8.1) with $R=1$ for such modification, then we will have it also for the original function. Hence, concentrating on (8.2) *and the case* $R=1$, without losing generality we may assume that $c^{\alpha\beta}\geq 1$.

Set $\mu=\delta_0^{-1}\delta+N_0$, where N_0 depending only on K_1, δ_0, and d is chosen in such a way that (6.1) is satisfied with

$$(\sigma_t,b_t)(x,y)=(\sigma,b)^{\alpha_t\beta_t}(x)$$

for all $\alpha.\in\mathfrak{A}$, $\beta.\in\mathfrak{B}$, x, y and $\delta>0$.

We use the notation from the proof of Theorem 2.1 in Section 7 and write (7.2) with

$$\gamma^{\alpha.\beta.0}=\tau_1^{\alpha.\beta.0}(\varepsilon)\wedge\tau_1^{\alpha.\beta.0}(0)\wedge\kappa_\varepsilon^{\alpha.\beta.0}\wedge\gamma_\varepsilon^{\alpha.\beta.0},$$

where $\tau_1^{\alpha.\beta.0}(\varepsilon)$ is the first exit time of $x_t^{\alpha.\beta.0}(\varepsilon)$ from B_1.

As in the proof of Theorem 2.1, by Theorem 6.5 (with $\tau=\gamma^{\alpha.\beta.0}$ there), we get that (recall that $M=2$ and μ is of order δ if $\delta\geq 1$)

$$|v(\varepsilon\xi)-v(0)|\leq N\delta\varepsilon+S_\varepsilon, \tag{8.3}$$

where N depends only on K_0, K_1 and δ_0 (recall that $\delta_1=1$) and

$$\begin{aligned}S_\varepsilon&:=\sup_{\alpha.,\beta.}E_0^{\alpha.\beta.}|v(x_\gamma(\varepsilon))-v(x_\gamma(0))|e^{-\phi_\gamma}I_{\tau_1(\varepsilon)\wedge\tau_1(0)<\gamma_\varepsilon\wedge\kappa_\varepsilon}\\&\leq\varepsilon L_1\sup_{\alpha.,\beta.}E_0^{\alpha.\beta.}|\xi_{\tau_1(\varepsilon)\wedge\tau_1(0)}(\varepsilon)|e^{-\phi_{\tau_1(\varepsilon)\wedge\tau_1(0)}}I_{\tau_1(\varepsilon)\wedge\tau_1(0)<\kappa_\varepsilon}.\end{aligned}$$

Observe that for any $T>0$ by Lemma 6.1 and Remark 6.1 ($\delta\geq K_1^2$),

$$\begin{aligned}&E_0^{\alpha.\beta.}|\xi_{\tau_1(\varepsilon)}(\varepsilon)|e^{-\phi_{\tau_1(\varepsilon)}}I_{\tau_1(\varepsilon)<\kappa_\varepsilon}\\&\quad=E_\varepsilon^{\alpha.\beta.}|\xi_{\tau_1(\varepsilon)}(\varepsilon)|e^{-\phi_{\tau_1(\varepsilon)}}I_{\tau_1(\varepsilon)<\kappa_\varepsilon\wedge T}\\&\qquad+E_0^{\alpha.\beta.}|\xi_{\tau_1(\varepsilon)}(\varepsilon)|e^{-\phi_{\tau_1(\varepsilon)}}I_{\tau_1(\varepsilon)<\kappa_\varepsilon}I_{\tau_1(\varepsilon)\geq T}\\&\quad\leq(E_0^{\alpha.\beta.}I_{\tau_1(\varepsilon)<T})^{1/2}+e^{-\delta T/2}E_0^{\alpha.\beta.}\sup_{t<\kappa_\varepsilon}|\xi_t(\varepsilon)|e^{-\phi_t+\delta t/2}\\&\quad\leq Ne^{-\delta T/2}+(E_0^{\alpha.\beta.}I_{\tau_1(\varepsilon)<T})^{1/2}.\end{aligned}$$

Similarly,

$$E_0^{\alpha.\beta.}|\xi_{\tau_1(0)}(\varepsilon)|e^{-\phi_{\tau_1(0)}}I_{\tau_1(0)<\kappa_\varepsilon}\leq Ne^{-\delta T/2}+(E_0^{\alpha.\beta.}I_{\tau_1(0)<T})^{1/2}.$$

One knows that if the starting point of a diffusion process with coefficients bounded by K_0 is in the ball of radius $\varepsilon<1/2$, then the probability that the process will exit from B_1 before time T is less than $N\exp(-\nu/T)$ if $K_0T\leq 1/2$, where N

and ν depend only on K_0 and d. This result is easily obtained by using the McKeen estimate (see, e.g., Corollary IV.2.9 of [8]) for each coordinate of the process from which one subtracts the drift term. Hence (with another ν)

$$S_\varepsilon \le \varepsilon L_1\big(Ne^{-\delta T/2} + Ne^{-\nu/T}\big).$$

For $T = \delta^{-1/2}$ (so that $K_0 T \le 1/2$ since $\delta \ge 4K_0^2$) we get that (yet with another ν)

$$S_\varepsilon \le \varepsilon L_1 N e^{-\nu\sqrt{\delta}},$$

and the result follows in case $R = 1$.

Once (8.1) is proved for $R = 1$, for $R \in (0, 1)$ it follows by using dilations (see Remark 2.5 of [10]), which allow us to keep the constants δ_0, K_0 and K_1 (actually, after dilations the constant K_1 can be taken even smaller then the original one). The lemma is proved. $\square$

PROOF OF THEOREM 2.2. First suppose that $\|g\|_{C^2(\mathbb{R}^d)} < \infty$ and that for an $R_0 > 0$ we have $B_{2R_0} \subset D$. Estimate (8.1) can be applied to any point rather than only 0, and it shows that for any $R' < R'' \le 2R_0$ and $\delta \ge K_1^2 + 4K_0^2 + 2$ we have

$$L_{R'} \le N\delta/\big(R'' - R'\big) + N_1 e^{-\nu\sqrt{\delta}} L_{R''}.$$

We apply this inequality to $R' = R_n$ and $R'' = R_{n+1}$, where R_n, $n \ge 1$, are defined by

$$R_n = R_0 + R_0 \sum_{i=1}^{n} \frac{\chi}{i^2},$$

and χ is such that $R_n \to 2R_0$ as $n \to \infty$. We also take and fix $\delta \ge K_1^2 + 4K_0^2 + 2$ so large that $N_1 e^{-\nu\sqrt{\delta}} \le 1/2$. Then for a constant N_0 depending only on δ_0, K_0, K_1 and d and all $n \ge 0$, we get that

$$L_{R_n} \le N_0 R_0^{-1}(n+1)^2 + 2^{-1} L_{R_{n+1}},$$

$$2^{-n} L_{R_n} \le 2^{-n} N_0 R_0^{-1}(n+1)^2 + 2^{-(n+1)} L_{R_{n+1}},$$

$$\sum_{n=0}^{\infty} 2^{-n} L_{R_n} \le N_0 R_0^{-1} \sum_{n=0}^{\infty} 2^{-n}(n+1)^2 + \sum_{n=0}^{\infty} 2^{-(n+1)} L_{R_{n+1}}$$

and $L_{R_0} \le N_0 I R_0^{-1}$, where

$$I = 2\sum_{n=1}^{\infty} 2^{-n} n^2.$$

One can do the same estimate for any ball inside D not necessarily centered at the origin, and this yields the desired result in case $\|g\|_{C^2(\mathbb{R}^d)} < \infty$. In the general case where g is only continuous, it suffices to use appropriate approximations of it by smooth functions. The theorem is proved. $\square$

9. Proof of Theorem 2.3. First of all we point out that the assertion of Lemma 7.1 continues to hold true with only one difference that N depends only on K_0, G, d and $\|g\|_{C^2(\mathbb{R}^d)}$. The proof remains the same with Itô's formula showing that the expectations in (7.1) are bounded by $NG(x)$. The remaining arguments follow the ones from Section 7 almost word for word.

In Section 5 for $|x-y| \le 1$ take

$$\pi^{\alpha\beta}(x,y) = \mu\big[\sigma^{\alpha\beta}(y)\big]^*(x-y)$$

and extend it appropriately for $|x-y|>1$.

Then for a unit $\xi \in \mathbb{R}^d$, $\varepsilon \ge 0$, $\alpha_\cdot \in \mathfrak{A}$, and $\beta_\cdot \in \mathfrak{B}$ introduce $x_t^{\alpha.\beta.0}(\varepsilon)$ as a unique solution of equation (5.2) with initial condition $\varepsilon\xi$ and

$$y_s = x_s^{\alpha.\beta.0}.$$

Observe that $x_t^{\alpha.\beta.0}(0) = x_t^{\alpha.\beta.0}$. Then define $z_t^{\alpha.\beta.0}(\varepsilon)$, $\tau_\varepsilon^{\alpha.\beta.0}$, $\gamma_\varepsilon^{\alpha.\beta.0}$, $\kappa_\varepsilon^{\alpha.\beta.0}$ and $\gamma^{\alpha.\beta.0}$ in the same way as in Section 7, and use Theorem 5.1 to get that

$$v(\varepsilon\xi) = \inf_{\boldsymbol{\beta}\in\mathbb{B}} \sup_{\alpha.\in\mathfrak{A}} E_0^{\alpha.\boldsymbol{\beta}(\alpha.)}\bigg[z_\gamma(\varepsilon)v\big(x_\gamma(\varepsilon)\big)e^{-\phi_\gamma(\varepsilon)} + \int_0^\gamma z_t(\varepsilon)\hat{f}\big(x_t(\varepsilon), x_t(0)\big)e^{-\phi_t(\varepsilon)}\,dt\bigg],$$

where

$$\phi_t^{\alpha.\beta.0}(\varepsilon) = \int_0^t \hat{c}^{\alpha_s\beta_s}\big(x_s^{\alpha.\beta.0}(\varepsilon), x_s^{\alpha.\beta.0}(0)\big)\,ds.$$

Fix $\alpha_\cdot \in \mathfrak{A}$ and $\beta_\cdot \in \mathfrak{B}$, and in Section 6 use the functions

$$(\sigma_t, b_t, c_t, f_t)(x,y) = (\hat{\sigma}, \hat{b}, \hat{c}, \hat{f})^{\alpha_t\beta_t}(x,y).$$

Observe that Assumption 6.2 is satisfied owing to Assumption 2.4.

Furthermore, for $t \le \gamma^{\alpha.\beta.}$ the processes x_t^ε and y_t coincide with $x_t^{\alpha.\beta.0}(\varepsilon)$ and $x_t^{\alpha.\beta.0}(0)$, respectively, since they satisfy the same equations, respectively. It follows that in the expectation

$$E_0^{\alpha.\beta.}\bigg[\int_0^\gamma z_t(\varepsilon)f\big(x_t(\varepsilon), x_t(0)\big)e^{-\phi_t(\varepsilon)}\,dt + z_\gamma(\varepsilon)v\big(x_\gamma(\varepsilon)\big)e^{-\phi_\gamma(\varepsilon)}\bigg],$$

one can replace $x_s^{\alpha.\beta.0}(\varepsilon)$ with x_t^ε, and by Theorem 6.5 we get that

$$\begin{aligned} &\bigg|E_0^{\alpha.\beta.}\bigg[\int_0^\gamma z_t(\varepsilon)f\big(x_t(\varepsilon), x_t(0)\big)e^{-\phi_t(\varepsilon)}\,dt + z_\gamma(\varepsilon)v\big(x_\gamma(\varepsilon)\big)e^{-\phi_\gamma(\varepsilon)}\bigg] \\ &\qquad - E_0^{\alpha.\beta.}\bigg[\int_0^\gamma f(x_t)e^{-\phi_t}\,dt + v(x_\gamma)e^{-\phi_\gamma}\bigg]\bigg| \\ &\quad \le N\varepsilon + E_0^{\alpha.\beta.}\big|v\big(x_\gamma(\varepsilon)\big) - v\big(x_\gamma(0)\big)\big|e^{-\phi_\gamma}I_{\gamma<\gamma_\varepsilon\wedge\kappa_\varepsilon}. \end{aligned} \tag{9.1}$$

If $t = \gamma^{\alpha.\beta.0} < \gamma_\varepsilon^{\alpha.\beta.0} \wedge \kappa_\varepsilon^{\alpha.\beta.0}$, then at least one of $x_t^{\alpha.\beta.0}(\varepsilon)$ and $x_t^{\alpha.\beta.0}(0)$ is outside D, and by Lemma 7.1 we obtain

$$\begin{aligned}
&E_0^{\alpha.\beta.}\big|v\big(x_\gamma(\varepsilon)\big) - v\big(x_\gamma(0)\big)\big|e^{-\phi_\gamma} I_{\gamma<\gamma_\varepsilon\wedge\kappa_\varepsilon} \\
&\quad\le N E_0^{\alpha.\beta.}\big|x_\gamma(\varepsilon) - x_\gamma(0)\big|e^{-\phi_\gamma} I_{\gamma<\gamma_\varepsilon\wedge\kappa_\varepsilon} \\
&\quad= \varepsilon E_0^{\alpha.\beta.}\big|\xi_\gamma(\varepsilon)\big|e^{-\phi_\gamma} I_{\gamma<\gamma_\varepsilon\wedge\kappa_\varepsilon} \le \varepsilon E_0^{\alpha.\beta.} \sup_{t<\kappa_\varepsilon}\big|\xi_t(\varepsilon)\big|e^{-\phi_t},
\end{aligned}$$

where $\varepsilon\xi_t^{\alpha.\beta.0}(\varepsilon) = x_t^{\alpha.\beta.0}(\varepsilon) - x_t^{\alpha.\beta.0}(0)$. By using Lemma 6.1, (9.1) and the fact that $\alpha.$ and $\beta.$ in the above argument are arbitrary, we see that $|v(\varepsilon\xi) - v(0)| \le N\varepsilon$. Similarly one proves that $|v(x+\varepsilon\xi) - v(x)| \le N\varepsilon$ for any x, which is what we need. The theorem is proved.

10. Proof of Theorem 2.4. Obviously $v \le v_K$. To estimate $v_K - v$ from above, define

$$d_K = \sup_{\mathbb{R}^d}(v_K - v), \qquad \lambda = \sup_{\alpha,\beta,x} c^{\alpha\beta}(x).$$

By the dynamic programming principle (see Theorem 3.1 in [10]),

$$v_K(x) = \inf_{\boldsymbol\beta\in\hat{B}}\sup_{\alpha.\in\hat{\mathfrak{A}}} E_x^{\alpha.\boldsymbol\beta(\alpha.)}\bigg[v_K(x_1)e^{-\lambda} + \int_0^1 \big\{f_K + (\lambda - c)v_K\big\}(x_t)e^{-\lambda t}\,dt\bigg].$$

Observe that

$$e^{-\lambda} + \int_0^1 \big[\lambda - c^{\alpha_t\beta_t}\big(x_t^{\alpha.\beta.x}\big)\big]e^{-\lambda t}\,dt \le e^{-\lambda} + \int_0^1 (\lambda - \delta_1)e^{-\lambda t}\,dt =: \kappa < 1.$$

Hence,

$$v_K(x) \le \inf_{\boldsymbol\beta\in\hat{B}}\sup_{\alpha.\in\hat{\mathfrak{A}}} E_x^{\alpha.\boldsymbol\beta(\alpha.)}\bigg[v(x_1)e^{-\lambda} + \int_0^1 \big\{f_K + (\lambda - c)v\big\}(x_t)e^{-\lambda t}\,dt\bigg] + \kappa d_K.$$

Now take a sequence x^n maximizing $v_K - v$, and take $\boldsymbol\beta^n \in \mathbb{B}$ such that

$$\begin{aligned}
v\big(x^n\big) \ge{}& \sup_{\alpha.\in\mathfrak{A}} E_{x^n}^{\alpha.\boldsymbol\beta^n(\alpha.)}\bigg[\int_0^1 \big(f + (\lambda - c)v\big)(x_t)e^{-\lambda t}\,dt + e^{-\lambda}v(x_1)\bigg] \\
&- 1/n.
\end{aligned} \tag{10.1}$$

Also define $\pi\alpha = \alpha$ if $\alpha \in A_1$ and $\pi\alpha = \alpha^*$ if $\alpha \in A_1$, where α^* is a fixed element of A_1, and find $\alpha^n_. \in \hat{\mathfrak{A}}$ such that

$$\begin{aligned}
v_K\big(x^n\big) \le{}& E_{x^n}^{\alpha^n_.\boldsymbol\beta^n(\pi\alpha^n_.)}\bigg[v(x_1)e^{-\lambda} + \int_0^1 \big\{f_K + (\lambda - c)v\big\}(x_t)e^{-\lambda t}\,dt\bigg] \\
&+ \kappa d_K + 1/n
\end{aligned} \tag{10.2}$$

$$
\begin{aligned}
&= E_{x^n}^{\alpha^n_\cdot \boldsymbol{\beta}^n(\pi\alpha^n_\cdot)}\left[v(x_1)e^{-\lambda} + \int_0^1 \{f + (\lambda - c)v\}(x_t)e^{-\lambda t}\,dt\right] \\
&\qquad - KR_n + \kappa d_K + 1/n,
\end{aligned}
$$

where

$$
R_n = E\int_0^1 e^{-\lambda t} I_{\alpha^n_t \in A_2}\,dt.
$$

By Lemma 5.3 of [10] for any $\alpha_\cdot \in \hat{\mathfrak{A}}$, $\beta_\cdot \in \mathfrak{B}$ and $x \in \mathbb{R}^d$, we have

$$
E\sup_{t\le 1}\left|x_t^{\pi\alpha_\cdot\beta_\cdot x} - x_t^{\alpha_\cdot\beta_\cdot x}\right| \le N\left(E_x^{\alpha_\cdot\beta_\cdot}\int_0^1 e^{-t} I_{\alpha^n_t\in A_2}\,dt\right)^{1/2},
$$

where the constant N depends only on K_0, K_1 and d. We use this, and since c, f, v are Lipschitz continuous, we get from (10.2) and (10.1),

$$
\begin{aligned}
&v_K(x^n) + (K - N_0)R_n \\
&\quad\le E_{x^n}^{\pi\alpha^n_\cdot \boldsymbol{\beta}^n(\pi\alpha^n_\cdot)}\left[v(x_1)e^{-\lambda} + \int_0^1 \{f + (\lambda - c)v\}(x_t)e^{-\lambda t}\,dt\right] \\
&\qquad + \kappa d_K + 1/n + NR_n^{1/2} \\
&\quad\le v(x^n) + \kappa d_K + 2/n + NR_n^{1/2},
\end{aligned}
$$

where the constant N_0 depends only on the supremums of c, v and f. Hence

$$
v_K(x^n) - v(x^n) - \kappa d_K + (K - N_0)R_n \le 2/n + NR_n^{1/2}. \tag{10.3}
$$

When n is large enough, $v_K(x^n) - v(x^n) - \kappa d_K \ge 0$ because of the way we chose x^n and the fact that $\kappa < 1$. It follows that for n large enough,

$$
(K - N_0)R_n \le 2/n + NR_n^{1/2},
$$

which for $K \ge 2N_0 + 1$ implies that $KR_n \le 4/n + NR_n^{1/2}$, so that, if $KR_n \ge 8/n$, then $KR_n \le NR_n^{1/2}$ and $R_n \le N/K^2$. Thus

$$
R_n \le 8/(nK) + N/K^2,
$$

which after coming back to (10.3) finally yields

$$
\begin{aligned}
&v_K(x^n) - v(x^n) - \kappa d_K \le 2/n + N/\sqrt{n} + N/K, \\
&\qquad (1-\kappa)d_K = \lim_{n\to\infty}\left[v_K(x^n) - v(x^n)\right] - \kappa d_K \le N/K,
\end{aligned}
$$

and the theorem is proved.

REFERENCES

[1] BARLES, G. (1991). A weak Bernstein method for fully nonlinear elliptic equations. *Differential Integral Equations* **4** 241–262. MR1081182

[2] CAFFARELLI, L. A. and CABRÉ, X. (1995). *Fully Nonlinear Elliptic Equations*. *American Mathematical Society Colloquium Publications* **43**. Amer. Math. Soc., Providence, RI. MR1351007

[3] FLEMING, W. H. and SOUGANIDIS, P. E. (1989). On the existence of value functions of two-player, zero-sum stochastic differential games. *Indiana Univ. Math. J.* **38** 293–314. MR0997385

[4] ISHII, H. and LIONS, P.-L. (1990). Viscosity solutions of fully nonlinear second-order elliptic partial differential equations. *J. Differential Equations* **83** 26–78. MR1031377

[5] KOVATS, J. (2009). Value functions and the Dirichlet problem for Isaacs equation in a smooth domain. *Trans. Amer. Math. Soc.* **361** 4045–4076. MR2500878

[6] KRYLOV, N. V. (1985). *Nonlinear Elliptic and Parabolic Equations of Second Order*. Nauka, Moscow [in Russian; English translation (1987). Reidel, Dordrecht].

[7] KRYLOV, N. V. (1989). Smoothness of the payoff function for a controllable diffusion process in a domain. *Izv. Akad. Nauk SSSR Ser. Mat.* **53** 66–96 in Russian; English translation in *Math. USSR Izvestija* **34** (1990) 65–96. MR0992979

[8] KRYLOV, N. V. (1995). *Introduction to the Theory of Diffusion Processes*. *Translations of Mathematical Monographs* **142**. Amer. Math. Soc., Providence, RI. MR1311478

[9] KRYLOV, N. V. (2012). On the existence of smooth solutions for fully nonlinear elliptic equations with measurable "coefficients" without convexity assumptions. *Methods Appl. Anal.* **19** 119–146. MR3011597

[10] KRYLOV, N. V. (2012). On the dynamic programming principle for uniformly nondegenerate stochastic differential games in domains and the Isaacs equations. Available at http://arxiv.org/abs/1205.0050.

[11] KRYLOV, N. V. (2012). On $C^{1+\alpha}$ regularity of solutions of Isaacs parabolic equations with VMO coefficients. Available at http://arxiv.org/abs/1211.4882.

[12] KRYLOV, N. V. (2013). On the dynamic programming principle for uniformly nondegenerate stochastic differential games in domains. *Stochastic Process. Appl.* **123** 3273–3298. MR3062445

[13] PROTTER, M. H. and WEINBERGER, H. F. (1967). *Maximum Principles in Differential Equations*. Prentice-Hall, Englewood Cliffs, NJ. MR0219861

[14] ŚWIĘCH, A. (1996). Another approach to the existence of value functions of stochastic differential games. *J. Math. Anal. Appl.* **204** 884–897. MR1422779

[15] ŚWIĘCH, A. (1997). $W^{1,p}$-interior estimates for solutions of fully nonlinear, uniformly elliptic equations. *Adv. Differential Equations* **2** 1005–1027. MR1606359

[16] TRUDINGER, N. (1989). On regularity and existence of viscosity solutions of nonlinear second order, elliptic equations. In *Partial Differential Equations and the Calculus of Variations, Vol. II. Progr. Nonlinear Differential Equations Appl.* **2** 939–957. Birkhäuser, Boston, MA. MR1034037

[17] VITOLO, A. (2010). Lipschitz constants for positive solutions of second-order elliptic equations. In *Proceedings of the Eighth Mississippi State-UAB Conference on Differential Equations and Computational Simulations*. *Electron. J. Differ. Equ. Conf.* **19** 257–266. Southwest Texas State Univ., San Marcos, TX. MR2754949

[18] ZHOU, W. (2013). A probabilistic approach to interior regularity of fully nonlinear degenerate elliptic equations in smooth domains. *Appl. Math. Optim.* **67** 419–452. MR3047001

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MINNESOTA
127 VINCENT HALL
MINNEAPOLIS, MINNESOTA 55455
USA
E-MAIL: krylov@math.umn.edu