Probab. Theory Relat. Fields 117, 1–16 (2000)
Digital Object Identifier (DOI) 10.1007/s004409900044

N.V. Krylov

On the rate of convergence of finite-difference approximations for Bellman's equations with variable coefficients

Received: 24 August 1998 / Revised version: 11 May 1999 /
Published online: 30 March 2000

Abstract. The estimates presented here for parabolic Bellman's equations with variable coefficients extend the ones earlier obtained for constant coefficients.

Introduction

The main purpose of the article is to present some estimates for the rate of convergence of finite-difference approximations in the problem of finding viscosity or probabilistic solutions to degenerate Bellman's equations. Specifically, we are dealing with the problem

$$F(D_t u, u_{x^i x^j}, u_{x^i}, u, t, x) = 0 \quad \text{in} \quad (0,T) \times \mathbb{R}^d, \quad u(T,x) = g(x) \quad \text{in} \quad \mathbb{R}^d \ , \tag{0.1}$$

where $\mathbb{R}^d$ is a d-dimensional Euclidean space of points $x = (x^1, \ldots, x^d)$, T is a fixed finite positive number, $D_t u = \partial u/\partial t$, $u_{x^i} = \partial u/\partial x^i$, $u_{x^i x^j} = \partial^2 u/\partial x^i \partial x^j$, and

$$\begin{aligned} &F(u_0, u_{ij}, u_i, u, t, x) \\ &\quad = \sup_{\alpha \in A}\{u_0 + a^{ij}(\alpha,t,x)u_{ij} + b^i(\alpha,t,x)u_i - c^\alpha(t,x)u + f^\alpha(t,x)\} \end{aligned}$$

with A, a, b, c, and f to be specified later. Such equations arise as dynamic programming equations for value functions in control problems of diffusion processes. Indeed, under our conditions the corresponding value function turns out to be a viscosity solution of (0.1).

Numerical approximations for problem (0.1) has been considered for quite long time (see [10], [3] and the references therein). There are known two methods for

---

N.V. Krylov: School of Mathematics, University of Minnesota, Minneapolis, MN 55455, USA. e-mail: `krylov@math.umn.edu`

The work was partially supported by an NSF Grant.

Mathematics Subject Classification (1991): 65M12, 35J60, 93E20

Key words and phrases: Finite-difference approximations – Bellman's equations – Fully nonlinear equations

proving the convergence of numerical approximations to the true solution. One of them, introduced by Kushner (see [10]), relies on constructing finite difference schemes, which generate controlled Markov chains, and proving that the chains converge weakly to the underlying controlled diffusion process. This turns out to be not an easy task because generally the set of the weak limit points of controlled Markov chains is wider than the set of the processes which can be obtained by controlling the diffusion process. Therefore, in [10] (for controlled drift) and in [8] (for controlled drift and diffusion) one has either to "convexify" the set of infinitesimal characteristics of controlled processes or relax the notion of controlled diffusion process as suggested in [2].

Another method is based on the uniqueness of viscosity solutions of (0.1). This method is introduced in [1] in an abstract setting and discussed in detail in [3] for our case.

Although both methods are quite general and allow one to treat a large variety of control problems, neither of them is suitable for estimating the rate of convergence. In our opinion, the main reason for this is that these methods do not use analytic properties of solutions to Bellman's equations. In this paper we present a different method of proving the convergence, which also provides an estimate of the rate of convergence. The idea of the method is explained in [5] for the case of controlled processes with time and state independent coefficients, where we use some quite elementary analytic properties of the value function. The arguments here for the more general case are a little bit more involved and rely also on mean value theorems for stochastic integrals (see [7]), which again are obtained on the basis of simple properties of Bellman's equations. By the way, if A is a singleton, then the value function is just the expectation of a certain well known functional of a solution to Itô's equation and the results of [7] are not needed.

Finally, it is worth noticing that unlike [10] and [3] we only consider problem (0.1) and unlike [5] we do not know anything about sharpness of our estimates even if A is a singleton. To the best of our knowledge, even in this case our estimates seem to be the first in their kind.

The article is organized as follows. In Sec. 1 we state our main results, which we prove in Sec. 3. The proofs are based on the results of Sec. 2 dealing with approximate smooth solutions of (degenerate) Bellman's equations. These results make sense even if the equation is linear and, in our opinion, are of independent interest. In Appendix we present the proof of Hölder continuity of the value function. This fact can be considered as well known and we give the proof only for the sake of completeness.

Throughout the paper $T \in (0, \infty)$, $K, K_1 \in [1, \infty)$, and $\delta_0, \delta \in (0, 1]$ are fixed constants. Everywhere, apart from Sec. 2, by N we denote various constants depending only on K, K_1, and d unless explicitly stated otherwise.

1. The main results

Let $(\Omega, \mathscr{F}, P)$ be a complete probability space, $\{\mathscr{F}_t; t \geq 0\}$ be an increasing filtration of σ-algebras $\mathscr{F}_t \subset \mathscr{F}$ which are complete with respect to $\mathscr{F}, P$. Assume that on $(\Omega, \mathscr{F}, P)$ a d_1-dimensional Wiener process w_t is defined for $t \geq 0$. We suppose that w_t is a Wiener process with respect to $\{\mathcal{F}_t\}$, or in other terms, that $\{w_t, \mathcal{F}_t\}$ is a Wiener process.

Let A be a separable metric space (the set of all admissible controls).

Definition 1.1. An A-valued function $\alpha_t = \alpha_t(\omega)$ defined for all $t \geq 0$ and $\omega \in \Omega$ is called *a policy* if it is measurable with respect to the product of $\mathcal{F}$ and Borel σ-field of $[0,\infty)$ and, in addition, $\mathcal{F}_t$-measurable with respect to ω for each $t \geq 0$. The set of all policies is denoted by $\mathfrak{A}$.

For $r_1, r_2 \in [0,\infty)$ and real-valued functions $f = f(t,x)$ given on $[r_1, r_2] \times \mathbb{R}^d$ we define

$$|f|_{0,[r_1,r_2]} = \sup_{[r_1,r_2]\times\mathbb{R}^d} |f(t,x)|, \quad [f]_{,\delta,[r_1,r_2]} = \sup_{\substack{t\in[r_1,r_2]\\ x\neq y}} \frac{|f(t,x)-f(t,y)|}{|x-y|^\delta} \ ,$$

$$[f]_{\delta/2,,[r_1,r_2]} = \sup_{x\in\mathbb{R}^d} \sup_{\substack{s,t\in[r_1,r_2]\\ s\neq t}} \frac{|f(s,x)-f(t,x)|}{|s-t|^{\delta/2}} \ ,$$

$$[f]_{\delta,[r_1,r_2]} = [f]_{,\delta,[r_1,r_2]} + [f]_{\delta/2,,[r_1,r_2]}, \quad |f|_{\delta,[r_1,r_2]} = |f|_{0,[r_1,r_2]} + [f]_{\delta,[r_1,r_2]} \ .$$

If $f = f(\alpha,t,x)$ is defined on $A \times [r_1,r_2] \times \mathbb{R}^d$ we write

$$|f|_{0,[r_1,r_2]} = \sup_{\alpha\in A} |f(\alpha,\cdot,\cdot)|_{0,[r_1,r_2]}, \quad [f]_{,\delta,[r_1,r_2]} = \sup_{\alpha\in A} [f(\alpha,\cdot,\cdot)]_{,\delta,[r_1,r_2]} \ ,$$

$$[f]_{\delta/2,,[r_1,r_2]} = \sup_{\alpha\in A} [f(\alpha,\cdot,\cdot)]_{\delta/2,,[r_1,r_2]} \ ,$$

$$[f]_{\delta,[r_1,r_2]} = [f]_{,\delta,[r_1,r_2]} + [f]_{\delta/2,,[r_1,r_2]}, \quad |f|_{\delta,[r_1,r_2]} = |f|_{0,[r_1,r_2]} + [f]_{\delta,[r_1,r_2]} \ .$$

The same notation is applied for vector-valued and matrix-valued functions in which case by $|\cdot|$ we mean the square root of the sum of squares of all entries. Actually, in the case of matrices $\sigma = (\sigma^{ij})$ instead of $|\sigma|$ we use a different notation $||\sigma||$, which is defined by

$$||\sigma||^2 = \sum_{i,j} |\sigma^{ij}|^2 \ .$$

Sometimes we take $r_2 = \infty$ in which case instead of $[r_1, r_2]$ in the above definitions we write $[r_1,\infty)$. If $r_1 = 0$ and $r_2 = \infty$, we drop $[r_1,r_2]$, so that, for instance, $|f|_0 := |f|_{0,[0,\infty)}$.

By $C^{2+\delta}([r_1,r_2])$ we denote the space of all functions $u = u(s,x)$ defined on $[r_1,r_2] \times \mathbb{R}^d$ with finite norm

$$|u|_{2+\delta,[r_1,r_2]} := |u|_{0,[r_1,r_2]} + |D_x u|_{0,[r_1,r_2]} + |D_t u|_{\delta,[r_1,r_2]} + |D^2_{xx} u|_{\delta,[r_1,r_2]} \ .$$

Suppose that on $A \times [0,\infty) \times \mathbb{R}^d$ we are given a $d \times d_1$ matrix-valued function $\sigma(\alpha,t,x)$, an $\mathbb{R}^d$-valued function $b(\alpha,t,x)$ and real-valued functions $c^\alpha(t,x) \geq 0$, $f^\alpha(t,x)$, and $g(x)$.

Assumption 1.2. The functions $\sigma(\alpha,t,x)$, $b(\alpha,t,x)$, $c^\alpha(t,x)$, and $f^\alpha(t,x)$ are Borel in all variables, continuous with respect to α, and

$$|\sigma|_0^2 + |b|_0 + |\sigma|_{,1} + |b|_{,1} + |\sigma|_{\delta_0/2,} + |b|_{\delta_0/2,} + |c|_\delta + |f|_\delta \leq K, \quad |g|_\delta \leq K_1 \ .$$

Since σ and b are bounded and Lipschitz in x, by Itô's theorem for any $\alpha \in \mathfrak{A}$, $s \in [0, \infty)$, and $x \in \mathbb{R}^d$ there exists a unique solution $x_t = x_t^{\alpha,s,x}$, $t \in [0, \infty)$, of the following equation

$$x_t = x + \int_0^t \sigma(\alpha_r, s + r, x_r)\, dw_r + \int_0^t b(\alpha_r, s + r, x_r)\, dr \ .$$

For $s \in [0, T]$ and $x \in \mathbb{R}^d$ define

$$\varphi_t^\alpha = \varphi_t^{\alpha,s,x} = \int_0^t c^{\alpha_r}(s + r, x_r^{\alpha,s,x})\, dr,$$

$$v^\alpha(s, x) = E[\int_0^{T-s} f^{\alpha_r}(s + r, x_r^{\alpha,s,x})e^{-\varphi_r^{\alpha,s,x}}\, dr + g(x_{T-s}^{\alpha,s,x})e^{-\varphi_{T-s}^{\alpha,s,x}}] \ .$$

We follow usual abbreviations putting arguments α, s, x around the symbols of expectation and probability to indicate that they should be placed inside in appropriate places, so that, for instance,

$$v^\alpha(s, x) = E_{s,x}^\alpha[\int_0^{T-s} f^{\alpha_r}(s + r, x_r)e^{-\varphi_r}\, dr + g(x_{T-s})e^{-\varphi_{T-s}}] \ .$$

Let

$$\begin{aligned} v(s, x) &:= \sup_{\alpha\in\mathfrak{A}} v^\alpha(s, x), \\ a(\alpha, s, x) &= \tfrac{1}{2}\sigma(\alpha, s, x)\sigma^*(\alpha, s, x), \\ L^\alpha u(s, x) &= a^{ij}(\alpha, s, x)u_{x^ix^j}(s, x) + b^i(\alpha, s, x)u_{x^i}(s, x) \\ &\quad - c^\alpha(s, x)u(s, x) + D_s u(s, x), \\ F[u] &= \sup_{\alpha\in A}[L^\alpha u + f^\alpha] \ . \end{aligned}$$

By definition v is the probabilistic solution of the problem

$$F[u] = 0 \quad \text{in} \quad (0, T) \times \mathbb{R}^d, \quad u(T, \cdot) = g \ . \tag{1.1}$$

The function v is also a viscosity solution of (1.1) (see, for instance, [3]).

Now we describe the approximating scheme for solving (1.1), which we are going to study. Let $\mathbb{R}_+^{d+1} = [0, \infty) \times \mathbb{R}^d$ and $\mathscr{B} = \mathscr{B}(\mathbb{R}_+^{d+1})$ be the set of all bounded functions on $\mathbb{R}_+^{d+1}$. For any $h \in (0, 1)$ let a number $p_h \in [1, \infty)$ and an operator $F_h : u \in \mathscr{B} \to F_h[u] \in \mathscr{B}$ be defined.

Assumption 1.3. (i) F_h is an h^2-local operator in t, that is, for any $t \in [0, \infty)$ and $u_1, u_2 \in \mathscr{B}$, we have $F_h[u_1](t, x) = F_h[u_2](t, x)$ for all $x \in \mathbb{R}^d$ whenever $u_1(s, y) = u_2(s, y)$ for all $s \in [t, t + h^2]$ and $y \in \mathbb{R}^d$;

(ii) F_h is locally consistent with F in the sense that, for any $t \in [0, \infty)$, $x \in \mathbb{R}^d$, and $u \in C^{2+\delta_0}([t, t + h^2])$, we have

$$|F_h[u](t, x) - F[u](t, x)| \le Kh^{\delta_0}|u|_{2+\delta_0,[t,t+h^2]} \ ;$$

(iii) the operator $u \to \Phi_h[u] := F_h[u] + p_h u$ is monotone, by which we mean that if $u_1, u_2 \in \mathscr{B}$ and $u_1 \geq u_2$, then $\Phi_h[u_1] \geq \Phi_h[u_2]$; moreover,

(iv) let $\ell := \ell(t) := e^{-2t}$, then for any constant $M \geq 0$ and $u_1, u_2 \in \mathscr{B}$ such that $u_1 + M\ell \geq u_2$, we have

$$\Phi_h[u_1] + M\ell(p_h - 1) \geq \Phi_h[u_1 + M\ell] \geq \Phi_h[u_2] \ .$$

Remark 1.4. As an example of approximating operators F_h one may take operators constructed on the basis of approximations of L^α in the following way. For any $h \in (0,1)$ and $\alpha \in A$, let a linear bounded operator $L_h^\alpha : \mathscr{B} \to \mathscr{B}$ and a function $q_h^\alpha(t,x) > 0$, $(t,x) \in \mathbb{R}_+^{d+1}$, be defined. Assume that

(a) for any $t \in [0,\infty)$ and $u \in \mathscr{B}$, we have $L_h^\alpha u(t,x) = 0$ for all $x \in \mathbb{R}^d$ whenever $u(s,y) = 0$ for all $s \in [t, t+h^2]$ and $y \in \mathbb{R}^d$;

(b) for any $u \in C^{2+\delta_0}$, $\alpha \in A$, $t \geq 0$, and $x \in \mathbb{R}^d$, we have

$$|L_h^\alpha u - L^\alpha u|(t,x) \leq K h^{\delta_0} |u|_{2+\delta_0,[t,t+h^2]} \ ;$$

(c) we have

$$p_h := \sup_{\alpha \in A, t, x} q_h^\alpha(t,x) + 1 < \infty \ ;$$

(d) the operator $u \to L_h^\alpha u + q_h^\alpha u$ maps nonnegative functions into nonnegative ones.

Under these conditions, the operator

$$F_h[u] := \sup_{\alpha \in A} (L_h^\alpha u + f^\alpha)$$

obviously satisfies conditions (i) through (iii) of Assumption 1.3. Condition (iv) is also satisfied for $h \in (0, h_0]$ where h_0 is a constant depending only on K and δ_0.

Indeed, owing to (b)

$$L_h^\alpha \ell \leq L^\alpha \ell + N(K) h^{\delta_0} \ell \leq -2\ell + N(K) h^{\delta_0} \ell \ ,$$

so that

$$\begin{aligned} \Phi_h[u + M\ell] &= F_h[u + M\ell] + p_h(u + M\ell) \leq F_h[u] + p_h u + M \sup_{\alpha \in A} L_h^\alpha \ell + M p_h \ell \\ &= \Phi_h[u] + M \sup_{\alpha \in A} L_h^\alpha \ell + M p_h \ell \leq \Phi_h[u] + M\ell(p_h - 2 + N(K) h^{\delta_0}) \ . \end{aligned}$$

Hence, upon defining h_0 so that $N(K) h^{\delta_0} \leq 1$, we see that the operator $F_{h \wedge h_0}$ satisfies all conditions of Assumption 1.3.

Remark 1.5. One can use many different ways to construct the operators L_h^α satisfying the conditions in Remark 1.4 (see [10], [8], [9]). For instance, take $d = 1$ and

$$\begin{aligned} L_h^\alpha u(t,x) = a(\alpha,t,x) h^{-2}[u(t,x+h) - 2u(t,x) + u(t,x-h)] \\ + b_+(\alpha,t,x) h^{-1}[u(t,x+h) - u(t,x)] \\ + b_-(\alpha,t,x) h^{-1}[u(t,x-h) - u(t,x)] \\ - c^\alpha(t,x) u(t,x) + h^{-2}[u(t+h^2,x) - u(t,x)] \ , \\ q_h^\alpha = [2a(\alpha) + 1] h^{-2} + |b(\alpha)| h^{-1} + K \ , \end{aligned}$$

where $b_\pm = (|b| \pm b)/2$. Then the requirements (a), (c), and (d) of Remark 1.4 are trivially satisfied and (b) is satisfied owing to the mean value theorem. Defining $F_h[u]$ as in Remark 1.4 and solving $F_h[u] = 0$ provides an implicit scheme of solving (1.1).

By the way, the equation $F_h[u] = 0$ can be considered only on the lattice $Q_h := \{t = ih^2, x = jh : i = 0, 1, \dots, t \le T, j = 0, \pm 1, \pm 2, ...\}$, and this explains the title of the article. However, considering this equation for all (t, x) turns out to be very helpful in proving the convergence results we are after, even if we were only interested in convergence at points which belong to all lattices $Q_{2^{-n}}$.

To discuss explicit schemes take a number $\gamma > 0$ and define $L_h^\alpha u(t, x)$ as

$$\begin{aligned}
&a(\alpha, t, x)(\gamma h)^{-2}[u(t + h^2, x + \gamma h) - 2u(t + h^2, x) + u(t + h^2, x - \gamma h)]\\
&\quad + b_+(\alpha, t, x)(\gamma h)^{-1}[u(t + h^2, x + \gamma h) - u(t + h^2, x)]\\
&\quad + b_-(\alpha, t, x)(\gamma h)^{-1}[u(t + h^2, x - \gamma h) - u(t + h^2, x)]\\
&\quad - c^\alpha(t, x)u(t + h^2, x) + h^{-2}[u(t + h^2, x) - u(t, x)] \ .
\end{aligned}$$

This time $F_h[u](t, x) = \Psi_h[u](t, x) - h^{-2}u(t, x)$, where $\Psi[u](t, x)$ is determined by the values of $u(t + h^2, \cdot)$, so that the equation $F_h[u] = 0$ gives an explicit formula $u(t, x) = h^2\Psi_h[u](t, x)$ and allows one to go down to smaller values of t just iterating this formula starting from the given value $u(T, x) = g(x)$. For such L_h^α the requirements (a) and (b) in Remark 1.4 are satisfied by the same reason as above. However, this time to satisfy (d) we obviously need to have

$$1 - 2a(\alpha, t, x)\gamma^{-2} - \gamma^{-1}h|b(\alpha, t, x)| - h^2c^\alpha(t, x) \ge 0 \ , \tag{1.2}$$

in which case one may take $q_h = h^{-2}$. Of course (1.2) is satisfied for all small h if γ is sufficiently large. This time after finding appropriate γ one can confine oneself to solving the equation $F_h[u] = 0$ on the lattice $\{t = ih^2, x = j\gamma h : i = 0, 1, \dots, t \le T, j = 0, \pm 1, \pm 2, \dots\}$.

Remark 1.6. It is worth noting that if one wants to follow the approximating method of Remark 1.4, in general, one need not approximate *all* operators L_h^α. The following Bellman's equation

$$D_t u + \sup_{\alpha \in B_1} \alpha^i\alpha^j u_{x^ix^j} = 0 \ , \tag{1.3}$$

where $B_1 = \{x \in \mathbb{R}^d : |x| < 1\}$, involves infinitely many operators $L^\alpha = D_t + \alpha^i\alpha^j D_{x^i}D_{x^j}$. Each of L^α can be approximated, say by

$$\begin{aligned}
L_h^\alpha u(t, x) := \ &h^{-2}[u(t + h^2, x) - u(t, x)]\\
&+ h^{-2}[u(t, x + \alpha h) - 2u(t, x) + u(t, x - \alpha h)]
\end{aligned}$$

and then, as easy to see, the requirements of Remark 1.4 are satisfied. However, we do not need to consider the approximations for $|\alpha| < 1$ since the sup in (1.3) is always equal to the sup over $S_1 = \{|\alpha| = 1\}$.

Furthermore, for any $h \in (0,1]$ one can find $n(h) \sim h^{(1-d)\delta_0}$ points α_k on the unit sphere S_1 such that they form an h^{δ_0}-net. Then, for any $\alpha \in S_1$, $t \geq 0$, and $u \in C^{2+\delta_0}([t, t+h^2])$ it holds that

$$\inf_k \left\{ \alpha^i \alpha^j u_{x^i x^j}(t,x) - \alpha_k^i \alpha_k^j u_{x^i x^j}(t,x) \right\} \leq N h^{\delta_0} |u|_{2+\delta_0,[t,t+h^2]} \ ,$$

$$|F[u](t,x) - \sup_{k=1,\dots,n(h)} L_h^{\alpha_k} u(t,x)| \leq N h^{\delta_0} |u|_{2+\delta_0,[t,t+h^2]} \ ,$$

so that we only have to deal with approximating finite number of operators, although increasing when $h \downarrow 0$.

Lemma 1.7. *(i) Let $\xi(t,x)$ and $\zeta(t,x)$ be bounded functions on $[0, T+h^2] \times \mathbb{R}^d$. Then the problem*

$$\begin{aligned} F_h[\eta](t,x) &= -\xi(t,x) \quad \forall t \in [0,T], x \in \mathbb{R}^d; \\ \eta(t,x) &= \zeta(t,x) \quad \forall t \in (T, T+h^2], x \in \mathbb{R}^d \end{aligned} \tag{1.4}$$

has a unique bounded solution $\eta_h = \eta_h(\xi, \zeta)$.

(ii) If $\xi_i(t,x)$ and $\zeta_i(t,x)$ are bounded functions on $[0, T+h^2] \times \mathbb{R}^d$, $i = 1, 2$, then

$$\begin{aligned} &\eta_h(\xi_1, \zeta_1) \\ &\quad \geq \eta_h(\xi_2, \zeta_2) - e^{2(T+h^2)} \sup_{(T,T+h^2]\times\mathbb{R}^d} (\zeta_2 - \zeta_1)_+ - e^{2(T+h^2)} \sup_{[0,T]\times\mathbb{R}^d} (\xi_2 - \xi_1)_+ \ . \end{aligned}$$

Proof. We introduce $\tilde\eta(t,x) := e^{2t}\eta(t,x)$ and $\tilde\xi(t,x) := p_h^{-1} e^{2t} \xi(t,x)$ and rewrite the first equation in (1.4) as

$$\Phi_h[\eta] = p_h \eta - \xi, \quad \Phi_h[\ell\tilde\eta] = p_h \ell \tilde\eta - p_h \ell \tilde\xi, \quad \tilde\eta = \tilde\Phi_h[\tilde\eta] + \tilde\xi \ ,$$

where

$$\tilde\Phi_h[u] := p_h^{-1} \ell^{-1} \Phi_h[\ell u] \ .$$

Hence, upon denoting $\tilde\zeta(t,x) := e^{2t}\zeta(t,x)$, we see that (1.4) is equivalent to the equation

$$\tilde\eta = \Psi_h[\tilde\eta] := \Psi_h[\tilde\eta, \tilde\xi, \tilde\zeta] := I_{t\in[0,T]}\{\tilde\Phi_h[\tilde\eta] + \tilde\xi\} + I_{t\in(T,T+h^2]}\tilde\zeta \ .$$

Owing to Assumption 1.3 (i) the operator $\tilde\eta \to \Psi_h[\tilde\eta, \tilde\xi, \tilde\zeta]$ can be regarded as an operator in the space $\mathscr{B}([0, T+h^2] \times \mathbb{R}^d)$ of all bounded functions on $[0, T+h^2] \times \mathbb{R}^d$. It turns out that this operator is a contraction.

Indeed, if $\tilde\eta_2 \leq \tilde\eta_1 + M$ with a positive constant M, then

$$\begin{aligned} \Psi_h[\tilde\eta_2] - \Psi_h[\tilde\eta_1] &= I_{t\in[0,T]}\{\tilde\Phi_h[\tilde\eta_2] - \tilde\Phi_h[\tilde\eta_1]\} \\ &= I_{t\in[0,T]} p_h^{-1} \ell^{-1} \{\Phi_h[\ell\tilde\eta_2] - \Phi_h[\ell\tilde\eta_1]\} \ , \end{aligned}$$

where, by Assumption 1.3 (iv), the last factor is less than $M\ell(p_h - 1)$. Therefore,

$$\Psi_h[\tilde\eta_2] - \Psi_h[\tilde\eta_1] \leq M(1 - p_h^{-1}) \ .$$

Taking here $M = |\tilde\eta_1 - \tilde\eta_2|_{0,[0,T+h^2]}$ and also interchanging $\tilde\eta_i$, we see that Ψ_h is a contraction with coefficient $1 - p_h^{-1} < 1$. By Banach's theorem we get the first assertion of the lemma.

To prove the second one, first notice that if $\tilde\xi_1 \geq \tilde\xi_2$ and $\tilde\zeta_1 \geq \tilde\zeta_2$, then, for $\tilde\eta_i(n)$ defined by $\tilde\eta_i(n+1) = \Psi_h[\tilde\eta_i(n), \tilde\xi_i, \tilde\zeta_i]$, $n \geq 0$, with $\tilde\eta_i(0) = 0$, we have $\tilde\eta_1(n) \geq \tilde\eta_2(n)$ for all n because of monotonicity of $\Psi_h[\tilde\eta, \tilde\xi, \tilde\zeta]$ in $\tilde\eta, \tilde\xi$, and $\tilde\zeta$. It follows that assertion (ii) holds true if $\xi_1 \geq \xi_2$ and $\zeta_1 \geq \zeta_2$. In other words, for any bounded functions η_1 and η_2 given in $[0, T+h^2] \times \mathbb{R}^d$, we see that if $F_h[\eta_1] \leq F_h[\eta_2]$ in $[0,T] \times \mathbb{R}^d$ and $\eta_1 \geq \eta_2$ in $(T, T+h^2] \times \mathbb{R}^d$, then $\eta_1 \geq \eta_2$ in $[0, T+h^2] \times \mathbb{R}^d$.

In the general case, denote

$$M = e^{2T+2h^2}\Big[\sup_{[0,T]\times\mathbb{R}^d}(\xi_2 - \xi_1)_+ + \sup_{(T,T+h^2]\times\mathbb{R}^d}(\zeta_2 - \zeta_1)_+\Big] .$$

Observe that by Assumption 1.3 (iv)

$$\begin{aligned} F_h[\eta_1 + M\ell] = \Phi_h[\eta_1 + M\ell] - p_h(\eta_1 + M\ell) &\leq \Phi_h[\eta_1] \\ +M\ell(p_h - 1) - p_h(\eta_1 + M\ell) \\ = F_h[\eta_1] - M\ell = -\xi_1 - M\ell &\leq -\xi_2 = F_h[\eta_2] \end{aligned}$$

in $[0,T] \times \mathbb{R}^d$ and obviously $\eta_1 + M\ell \geq \eta_2$ in $(T, T+h^2] \times \mathbb{R}^d$. Hence, $\eta_1 + M \geq \eta_2$ in $[0, T+h^2] \times \mathbb{R}^d$ and the lemma is proved.

Remark 1.8. We will apply Lemma 1.7 not only to the interval $[0,T]$ but also to its subintervals changing the origin of t-axis if needed.

In the sequel by v_h we denote the function defined by Lemma 1.7 for $\xi = 0$ and $\zeta = g$.

Here are our main results.

Theorem 1.9. *For any $h \in (0,1]$, in $[0,T] \times \mathbb{R}^d$ we have*

$$v_h \leq v + Ne^{NT}h^{\delta_1} ,$$

where the constant N depends only on K, K_1, d and $\delta_1 = \delta\delta_0^2/(2 + \delta_0 + \delta\delta_0 - \delta)$, which is $1/3$ if $\delta_0 = \delta = 1$.

Corollary 1.10. *If A is a singleton and $F_h[u] = L_h u + f_h$ where L_h is a linear operator, then for any $h \in (0,1]$,*

$$|v_h - v| \leq Ne^{NT}h^{\delta\delta_0^2/(2+\delta_0+\delta\delta_0-\delta)} .$$

Theorem 1.11. *Let $\delta_0 \leq \delta$. Then for any $h \in (0,1]$, in $[0,T] \times \mathbb{R}^d$ we have*

$$v \leq v_h + Ne^{NT}h^{\delta_1} , \tag{1.5}$$

where the constant N depends only on K, K_1, d and $\delta_1 = \delta^2\delta_0^3(8 + \delta\delta_0)^{-1}(2 + \delta_0 + \delta\delta_0 - \delta)^{-1}$, which is $1/27$ if $\delta_0 = \delta = 1$.

The estimates in Theorems 1.9 and 1.11 are of different order and it is interesting to know to what extent the estimate in Theorem 1.11 can be improved. In [5] we considered processes with time and space independent coefficients, so that δ_0 can be any number in $[0, 1]$. The result of [5] says that $|v - v_h|$ goes to zero as $h \downarrow 0$ at least as $h^{\delta/3}$. Taking $\delta_0 = 1$, from Theorem 1.9 we get the same estimate of order $h^{\delta/3}$. However, this estimate is only for $v_h - v$ from above. We believe that under additional and quite general conditions on F_h one can improve the estimate in Theorem 1.11 to be of the same order as in Theorem 1.9. In [6] we prove that the rate in Theorem 1.11 is not less than $1/21$ if $\delta = \delta_0 = 1$ and we also give some approximations with rate $1/3$. By the way, the examples in [5] show that generally one cannot do better than $h^{1/2}$.

2. Auxiliary results

The proofs of our main results are based on the following theorem in which and everywhere in this section by N we denote various constants depending only on K and d.

Theorem 2.1. *For any* $\varepsilon \in (0, 1]$ *there exists a function* u *defined in* $[0, T + \varepsilon^2] \times \mathbb{R}^d$ *such that*

$$|u(t, x) - g(x)| \le N K_1 \varepsilon^{\delta} \quad for \quad t \in [T, T + \varepsilon^2] \ ; \tag{2.1}$$

$$\sup_{\alpha \in A} [L^{\alpha} u + f^{\alpha}] \le 0, \quad |u - v| \le N e^{NT} K_1 \varepsilon^{\delta\delta_0} \quad in \quad [0, T] \times \mathbb{R}^d \ , \tag{2.2}$$

$$|u|_{2+\delta_0, [0, T+\varepsilon^2]} \le N e^{NT} K_1 \varepsilon^{\delta - 2 - \delta_0} \ . \tag{2.3}$$

Proof. First observe that we may assume $|g|_{\delta} \le 1$ and $K_1 = 1$. Indeed if $|g|_{\delta} > 1$ we can replace f and g with $f(1 + |g|_{\delta})^{-1}$ and $g(1 + |g|_{\delta})^{-1}$, respectively, and then after getting an appropriate function u make the inverse transformation.

Next, define $\tilde{A} = A \times \{(\tau, \beta) : \tau \in (-1, 0), \beta \in B_1\}$. Extend σ, b, c, f for negative t following the example $\sigma(\alpha, t, x) = \sigma(\alpha, 0, x)$ and for a fixed $\varepsilon \in (0, 1]$ and any $\tilde{\alpha} = (\alpha, \tau, \beta) \in \tilde{A}$ let

$$\sigma(\tilde{\alpha}, t, x) = \sigma_{\varepsilon}(\tilde{\alpha}, t, x) = \sigma(\alpha, t + \varepsilon^2 \tau, x + \varepsilon \beta)$$

and similarly define $b(\tilde{\alpha}, t, x)$, $c^{\tilde{\alpha}}(t, x)$, and $f^{\tilde{\alpha}}(t, x)$. We denote by $\tilde{\mathfrak{A}}$ the set of all measurable $\mathscr{F}_t$-adapted $\tilde{A}$-valued processes. As usual starting with these objects defined on $\tilde{A}$, for any $\tilde{\alpha} \in \tilde{\mathfrak{A}}$, $s \ge 0$, and $x \in \mathbb{R}^d$, we define the controlled diffusion process $x_t^{\tilde{\alpha}, s, x}$ and the value functions

$$u^{\tilde{\alpha}}(s, x) = E_{s,x}^{\tilde{\alpha}} [\int_0^{S-s} f^{\tilde{\alpha}_t}(s + t, x_t) e^{-\varphi_t} \, dt + g(x_{S-s}) e^{-\varphi_{S-s}}],$$

$$u(s, x) = \sup_{\tilde{\alpha} \in \tilde{A}} v^{\tilde{\alpha}}(s, x) \ ,$$

which we consider for $s \le S$, where $S = T + \varepsilon^2$. We will keep in mind that $u^{\tilde\alpha}(s,x)$ and $u(s,x)$ also depend on ε which is not explicitly shown just for simplicity of notation.

Next, take a nonnegative function $\zeta \in C_0^\infty((-1,0)\times B_1)$ with unit integral, and for $\varepsilon > 0$ define $\zeta_\varepsilon(t,x) = \varepsilon^{-d-2}\zeta(t/\varepsilon^2, x/\varepsilon)$. We also use the notation

$$u^{(\varepsilon)}(t,x) = u(t,x) * \zeta_\varepsilon(t,x).$$

Our goal is to prove that $u^{(\varepsilon)}$ is a function for which the assertions of the theorem are true.

To prove (2.3) it suffices to use well-known properties of mollifiers and notice that (see Appendix and remember $|g|_\delta \le 1$)

$$|u(t,x) - u(s,y)| \le Ne^{NT}(|t-s|^{\delta/2} + |x-y|^\delta) \ . \tag{2.4}$$

Next assume temporarily that σ and b are twice continuously differentiable in x with the derivatives being bounded. Then by Lemma 4.1.5 of [4], for any $\tilde\alpha = (\alpha,\tau,\beta) \in \tilde A$,

$$\begin{aligned}\tilde L^{\tilde\alpha} u(t,x) + f^{\tilde\alpha}(t,x) := {} & D_t u(t,x) + a^{ij}(\alpha, t+\varepsilon^2\tau, x+\varepsilon\beta) u_{x^i x^j}(t,x)\\ & + b^i(\alpha, t+\varepsilon^2\tau, x+\varepsilon\beta) u_{x^i}(t,x)\\ & - c^\alpha(t+\varepsilon^2\tau, x+\varepsilon\beta) u(t,x)\\ & + f^\alpha(t+\varepsilon^2\tau, x+\varepsilon\beta) \le 0\end{aligned} \tag{2.5}$$

in the sense of generalized functions of $(t,x) \in (0,S)\times\mathbb{R}^d$. The same inequality holds in the sense of generalized functions of $(t,x,\tau,\beta) \in (0,S)\times\mathbb{R}^d\times(-1,0)\times B_1$. After "changing variables", we easily get that the inequality

$$\begin{aligned}& D_t u(t-\varepsilon^2\tau, x-\varepsilon\beta) + a^{ij}(\alpha,t,x) u_{x^i x^j}(t-\varepsilon^2\tau, x-\varepsilon\beta)\\ & + b^i(\alpha,t,x) u_{x^i}(t-\varepsilon^2\tau, x-\varepsilon\beta) - c^\alpha(t,x) u(t-\varepsilon^2\tau, x-\varepsilon\beta)\\ & + f^\alpha(t,x) \le 0\end{aligned} \tag{2.6}$$

holds in the sense of generalized functions of $(t,x,\tau,\beta) \in (0,T)\times\mathbb{R}^d\times(-1,0)\times B_1$. Furthermore, since u, a, b, c, f are continuous in (t,x), inequality (2.6) holds in the sense of generalized functions of $(\tau,\beta) \in (-1,0)\times B_1$ for any $(t,x) \in [0,T]\times\mathbb{R}^d$. By multiplying (2.6) by $\zeta(\tau,\beta)$ and integrating (that is using the definition of the words "in the sense of generalized functions"), we conclude that $L^\alpha u^{(\varepsilon)} + f^\alpha \le 0$ in $[0,T]\times\mathbb{R}^d$ for any $\alpha \in A$. This proves the first inequality in (2.2) in the particular case of smooth σ, b. The general case is obtained by mollifying σ, b with respect to x and passing to the limit, when the kernel tends to the delta function, on the basis of Theorem 3.1.13 of [4] which says that the corresponding value functions converge uniformly to u on any bounded subset of $[0,S]\times\mathbb{R}^d$.

To prove (2.1) observe that obviously

$$-N(S-s) + \inf_{\tilde\alpha\in\tilde{\mathfrak{A}}} E_x^{\tilde\alpha} g(x_{S-s}) \le u(s,x) \le \sup_{\tilde\alpha\in\tilde{\mathfrak{A}}} E_x^{\tilde\alpha} g(x_{S-s}) + N(S-s) \ .$$

For $S-s\le\varepsilon^2$, it follows that

$$|u(s,x)-g(x)|\le N\varepsilon^2+\sup_{\tilde\alpha\in\tilde{\mathfrak{A}}}E^{\tilde\alpha}_{s,x}|g(x_{S-s})-g(x)|$$
$$\le N\varepsilon^2+N\sup_{\tilde\alpha\in\tilde{\mathfrak{A}}}E^{\tilde\alpha}_{s,x}|x_{S-s}-x|^\delta\le N\varepsilon^\delta\ .$$

This and (2.4) easily yield (2.1) for $u^{(\varepsilon)}$ in place of u.

Thus, to finish the proof of the theorem, it only remains to prove that the second relation in (2.2) holds for $u^{(\varepsilon)}$ in place of u.

Fix $s\in[0,T]$ and $x\in\mathbb{R}^d$, take $\tilde\alpha=(\alpha,\tau,\beta)\in\tilde{\mathfrak{A}}$, and notice that by comparing equations defining $x_t^{\tilde\alpha,s,x}$ and $x_t^{\alpha,s,x}$ one can easily get (see, for instance, Theorem 2.5.9 in [4]) that

$$E\sup_{t\le T-s}|x_t^{\tilde\alpha,s,x}-x_t^{\alpha,s,x}|^2\le Ne^{NT}\sup\{||\sigma(\alpha,t+\varepsilon^2\tau,x+\varepsilon\beta)-\sigma(\alpha,t,x)||^2$$
$$+|b(\alpha,t+\varepsilon^2\tau,x+\varepsilon\beta)-b(\alpha,t,x)|^2\}\ ,$$

where the sup is taken over

$$\alpha\in A,\quad t\le T,\quad \tau\in(-1,0),\quad x\in\mathbb{R}^d,\quad \beta\in B_1\ .$$

It follows by our assumptions and Hölder's inequality that

$$E\sup_{t\le T-s}|x_t^{\tilde\alpha,s,x}-x_t^{\alpha,s,x}|^2\le Ne^{NT}\varepsilon^{2\delta_0}\ ,$$
$$E\sup_{t\le T-s}|x_t^{\tilde\alpha,s,x}-x_t^{\alpha,s,x}|^\delta\le Ne^{NT}\varepsilon^{\delta\delta_0}\ ,$$

$$E\int_0^{S-s}|f^{\tilde\alpha_t}(s+t,x_t^{\tilde\alpha,s,x})-f^{\alpha_t}(s+t,x_t^{\alpha,s,x})|\,dt$$
$$\le\ NT\varepsilon^\delta+\ E\int_0^{T-s}|f^{\alpha_t}(s+t,x_t^{\tilde\alpha,s,x})-f^{\alpha_t}(s+t,x_t^{\alpha,s,x})|\,dt\le Ne^{NT}\varepsilon^{\delta\delta_0}\ ,$$
$$E\int_0^{S-s}|c^{\tilde\alpha_t}(s+t,x_t^{\tilde\alpha,s,x})-c^{\alpha_t}(s+t,x_t^{\alpha,s,x})|\,dt\le\ Ne^{NT}\varepsilon^{\delta\delta_0}\ ,$$
$$E^{\tilde\alpha}_{s,x}|g(x_{S-s})-g(x_{T-s})|\le N\varepsilon^\delta\ .$$

Next, upon using the inequality

$$|f_1e^{-c_1}-f_2e^{-c_2}|\le|f_1-f_2|+(|f_1|+|f_2|)|c_1-c_2|,\quad c_1,c_2\ge0\ ,\qquad(2.7)$$

we get that, for $s\in[0,S]$ and $x\in\mathbb{R}^d$,

$$|u^{\tilde\alpha}(s,x)-v^\alpha(s\wedge T,x)|\le Ne^{NT}\varepsilon^{\delta\delta_0},\quad |u(s,x)-v(s\wedge T,x)|\le Ne^{NT}\varepsilon^{\delta\delta_0},$$
$$|u^{(\varepsilon)}(s,x)-[v(\cdot\wedge T,\cdot)]^{(\varepsilon)}(s,x)|\le Ne^{NT}\varepsilon^{\delta\delta_0}\ .$$

Now, to get the second inequality in (2.2) for $u^{(\varepsilon)}$ in place of u, it only remains to notice that (2.4) holds with v in place of u, so that $|[v(\cdot\wedge T,\cdot)]^{(\varepsilon)}-v|\le Ne^{NT}\varepsilon^\delta$. The theorem is proved.

Remark 2.2. In the second inequality in (2.2) only the estimate of $u - v$ from above is of interest. Indeed, the following estimate from below is trivially derived by using Itô's formula from the first relation in (2.2) and from (2.1):

$$u \geq v - NK_1 \varepsilon^{\delta} \ .$$

When A consists of only one point, replacing f, g with $-f, -g$ leads to replacing v with $-v$ and easily leads to the following.

Corollary 2.3. *For any $\alpha \in A$ and $\varepsilon \in (0, 1]$ there exist functions $u_{\pm}(s, x)$ defined in $[0, T + \varepsilon^2] \times \mathbb{R}^d$ such that*

$$|u_{\pm}(t, x) - g(x)| \leq NK_1 \varepsilon^{\delta} \quad \textit{for} \quad t \in [T, T + \varepsilon^2],$$

$$\pm(L^{\alpha} u_{\pm} + f^{\alpha}) \leq 0, \quad |u_{\pm} - v^{\alpha}| \leq Ne^{NT} K_1 \varepsilon^{\delta\delta_0} \quad \textit{in} \quad [0, T] \times \mathbb{R}^d,$$

$$|u_{\pm}|_{2+\delta_0,[0,T+\varepsilon^2]} \leq Ne^{NT} K_1 \varepsilon^{\delta-2-\delta_0} \ ,$$

where N depends only on K and d and v^{α} is generated by the constant policy $\alpha_t \equiv \alpha$.

3. Proof of Theorems 1.9 and 1.11

Proof of Theorem 1.9. Take $h \in (0, 1]$, $\varepsilon \in [h, 1]$, and a function u from Theorem 2.1. Then by Assumption 1.3 (ii) we have that

$$F_h[u] \leq Ne^{NT} h^{\delta_0} \varepsilon^{\delta-2-\delta_0}$$

in $[0, T] \times \mathbb{R}^d$ and $u(t, x) \geq g(x) - N\varepsilon^{\delta} = v_h(t, x) - N\varepsilon^{\delta}$ in $(T, T + h^2]$. It follows from Lemma 1.7 that, in $[0, T] \times \mathbb{R}^d$,

$$v_h \leq u + Ne^{NT}(\varepsilon^{\delta} + h^{\delta_0}\varepsilon^{\delta-2-\delta_0}) \ .$$

By applying the second equation in (2.2) we conclude

$$v_h \leq v + Ne^{NT}(\varepsilon^{\delta\delta_0} + h^{\delta_0}\varepsilon^{\delta-2-\delta_0}) \ .$$

It only remains to set $\varepsilon = h^{\xi}$, where $\xi = \delta_0(2 + \delta_0 + \delta\delta_0 - \delta)^{-1}$ (<1). The theorem is proved.

To prove Theorem 1.11 we need two more auxiliary results before which we introduce the following notation. For any $\beta \in A$, $t \geq s \geq 0$, and Borel function $u = u(x)$ define

$$G^{\beta}_{s,t} u(x) = E^{\alpha}_{s,x}\Big[\int_0^{t-s} f^{\beta}(s + r, x_r) e^{-\varphi_r} \, dr + u(x_{t-s}) e^{-\varphi_{t-s}}\Big] \ ,$$

where $\alpha_t \equiv \beta$. Also let $G_{s,t} u = \sup_{\beta} G^{\beta}_{s,t} u$.

Lemma 3.1. *Let $u = u(x)$ be a function such that $|u|_\delta < \infty$. Take $0 \le s \le t \le T$ and $h \in (0, 1]$ and assume that $u(y) \le v_h(r, y) + q$ for $r \in (t, t + h^2]$ and $y \in \mathbb{R}^d$, where q is a positive constant. Then*

$$G_{s,t}u(x) \le v_h(s, x) + qe^{2(t-s)+2h^2} + Ne^{N(t-s)}Mh^{\delta_1} \quad \forall\, x \in \mathbb{R}^d \ ,$$

where $\delta_1 := \delta\delta_0^2/(2 + \delta_0 + \delta\delta_0 - \delta)$ and $M = 1 + |u|_\delta$.

Proof. We need to prove that

$$G^\beta_{s,t}u(x) \le v_h(s, x) + NMe^{N(t-s)}h^{\delta_1} + qe^{2(t-s)+2h^2}$$

for any $\beta \in A$, which we therefore fix. Take $h \in (0, 1]$ and $\varepsilon \in [h, 1]$ and notice that by Corollary 2.3 there exists a function $z(r, x)$ defined in $[s, t + \varepsilon^2] \times \mathbb{R}^d$ such that

$$|z(r, x) - u(x)| \le NM\varepsilon^\delta \quad \text{for} \quad r \in [t, t + \varepsilon^2], x \in \mathbb{R}^d,$$

$$L^\beta z + f^\beta \ge 0, \quad z(r, x) \ge G^\beta_{r,t}u(x) - NMe^{N(t-r)}\varepsilon^{\delta\delta_0} \tag{3.1}$$

in $[s, t] \times \mathbb{R}^d$ and $|z|_{2+\delta_0,[s,t+\varepsilon^2]} \le Ne^{N(t-s)}M\varepsilon^{\delta-2-\delta_0}$.

In particular, $F[z] \ge 0$, so that by Assumption 1.3 (ii)

$$F_h[z] \ge -Ne^{N(t-s)}Mh^{\delta_0}\varepsilon^{\delta-2-\delta_0} \quad \text{in} \quad [s, t] \times \mathbb{R}^d \ .$$

In addition,

$$\sup_{(t,t+h^2]\times\mathbb{R}^d} (z - v_h)_+ \le \sup_{(t,t+h^2]\times\mathbb{R}^d} (u - v_h)_+ + NM\varepsilon^\delta \le NM\varepsilon^\delta + q \ ,$$

$$\begin{aligned} F_h[v_h] &= 0 \quad \text{in} \quad [s, t] \times \mathbb{R}^d \\ v_h &= v_h \quad \text{in} \quad (t, t + h^2] \times \mathbb{R}^d \ , \end{aligned}$$

which by Lemma 1.7 (ii) and Remark 1.8 implies that

$$z(s, x) \le v_h(s, x) + Ne^{N(t-s)}M(\varepsilon^\delta + h^{\delta_0}\varepsilon^{\delta-2-\delta_0}) + qe^{2(t-s)+2h^2} \ .$$

Combining this with (3.1) we get

$$G^\beta_{s,t}u(x) \le v_h(s, x) + Ne^{N(t-s)}M(\varepsilon^{\delta\delta_0} + h^{\delta_0}\varepsilon^{\delta-2-\delta_0}) + qe^{2(t-s)+2h^2} \ .$$

By taking the same ε as in the preceding proof, we get the result. The lemma is proved.

For $s \in [0, T]$ and integers $n \ge 1$ denote by $\mathfrak{A}_n(s)$ the set of all $\alpha \in \mathfrak{A}$ satisfying $\alpha_t = \alpha_{(\kappa_n(t+s)-s)_+}$ for $t \ge 0$, where $\kappa_n(r) = [rn]/n$. The policies of class $\mathfrak{A}_n(s)$ are constant on the intervals $[t_{nk}(s), t_{n,k+1}(s))$ with $t_{n0}(s) = 0$, $t_{nk}(s) = \kappa_n(s) - s + k/n$ for $k \ge 1$. Also let

$$v^n(s, x) = \sup_{\alpha\in\mathfrak{A}_n(s)} v^\alpha(s, x) \ .$$

From Lemma 3.2.14 and the proof of Lemma 3.3.1 in [4], we infer that the following dynamic programming equation holds

$$v^n(s,x) = G_{s,\tau_{n,k+1}} v^n(\tau_{n,k+1},\cdot)(x) \quad \text{if} \quad s \in \Delta(n,k), k = 0,\dots,k(n)-1 \ , \tag{3.2}$$

where $k(n)$ is the number of intervals $(i/n, (i+1)/n]$, $i \geq 0$, intersecting $[0,T]$, $\Delta(n,k) = [\tau_{nk}, \tau_{n,k+1}]$, $\tau_{nk} = (k/n) \wedge T$ (so that $\tau_{nk(n)} = T$ and $\tau_{nk} < T$ if $k < k(n)$).

In the following few lines we introduce several different constants N_0 depending only on K, K_1. We will use that (see Appendix)

$$|v^n|_{,\delta,[s,T]} \leq N_0 e^{N_0(T-s)} \ , \tag{3.3}$$

which implies (see Appendix) that, for any $\beta \in A$ and $s \leq t \leq T$, we have

$$[G^{\beta}_{\cdot,t} v^n(t,\cdot)]_{\delta/2,,[s,t]} \leq N_0 e^{N_0(T-s)} \ , \tag{3.4}$$

and on account of (3.2),

$$|v^n(s,x) - v^n(s+r,x)| \leq N_0 e^{N_0(T-s)} r^{\delta/2} \tag{3.5}$$

if s and $s+r$ belong to the same interval $\Delta(n,k)$ and $r \geq 0$.

Lemma 3.2. *For any $n \geq 1$ and $h^2 \leq 1/n$, we have*

$$v^n(s,x) \leq v_h(s,x) + N e^{N(T-s)} n h^{\delta_1} \ , \tag{3.6}$$

where δ_1 is the same as in Lemma 3.1.

Proof. Lemma 3.1 applied with $u = g$ immediately implies that, for $s \in \Delta_{n,k(n)-k}$ with $k = 1$

$$v^n(s,x) \leq v_h(s,x) + M_k e^{N_k/n} h^{\delta_1} \ , \tag{3.7}$$

where M_1 and N_1 are some constants depending only on K, K_1, d. Bearing in mind the induction on k, assume that (3.7) holds for $s \in \Delta_{n,k(n)-k}$ and some $k \geq 1$. We are going to prove then that, if $k < k(n)$, then (3.7) also holds for $s \in \Delta_{n,k(n)-k-1}$ with appropriate constants N_{k+1} and M_{k+1} in place of N_k and M_k. Furthermore, we will see that one may take

$$N_{k+1} \leq N_k \vee (Nk + N), \quad M_{k+1} \leq N + e^{4/n} M_k \ .$$

Then by induction $N_{k+1} \leq Nk + N$ and $M_k \leq Nne^{5k/n}$, which allows us to replace M_k and N_k in (3.7) with $Nne^{5(T-s)}$ and $N(T-s+1)n$, respectively, and leads us to (3.6).

It follows from (3.5) and (3.7) that, if $\tau_{n,k(n)-k} + r \in \Delta(n, k(n)-k)$ and $r \leq h^2$, then

$$\begin{aligned} v^n(\tau_{n,k(n)-k},x) &\leq v^n(\tau_{n,k(n)-k}+r,x) + N_0 e^{N_0 k/n} h^{\delta} \\ &\leq v_h(\tau_{n,k(n)-k}+r,x) + N_0 e^{N_0 k/n} h^{\delta} + M_k e^{N_k/n} h^{\delta_1} \\ &=: v_h(\tau_{n,k(n)-k}+r,x) + q_k \ . \end{aligned} \tag{3.8}$$

Here, if $k > 1$, one can obviously take any $r \in [0, h^2]$ due to the assumption $h^2 \le 1/n$. It turns out that for $k = 1$ the inequality between extreme terms in (3.8) holds for any $r \in [0, h^2]$ as well. Indeed, for $T - \tau_{n,k(n)-1} \le r \le h^2$ (if there are such r at all), we have

$$v_h(\tau_{n,k(n)-1} + r, x) = g(x) = v^n(T, x) \ge v^n(\tau_{n,k(n)-1}, x) - N_0 h^\delta \ .$$

Hence owing to (3.3), (3.2), and Lemma 3.1 we get that, for $s \in \Delta_{n,k(n)-k-1}$,

$$v^n(s, x) \le v_h(s, x) + e^{4/n} q_k + N(1 + N_0) e^{N/n + N_0 k/n} h^{\delta_1} \ .$$

This yields (3.7) with

$$N_{k+1} = (N + N_0 k) \vee N_k, \quad M_{k+1} = N + e^{4/n} M_k$$

in place of N_k and M_k. As explained above, the lemma is thus proved.

Proof of Theorem 1.11. By applying Theorem 2.9 of [7] and taking into account that $\delta_0 \le \delta$, we have

$$v(0, x) \le v^n(0, x) + N e^{NT} n^{-\delta\delta_0/8} \ .$$

Combining this with (3.6) and choosing integer n to be of order h^{-q} with $q = 8\delta_1(8 + \delta\delta_0)^{-1}$ ($q < 2$, so that $n \le h^{-2}$ as required in Lemma 3.2), we get (1.5) at points $(0, x)$. By the way, we have only considered the points $(0, x)$ because in [7] the intervals, on which the policies of type $\mathfrak{A}_n$ are constant, are assumed to have the same length $1/n$. Other points (s, x) are considered by shifting the time variable. The theorem is proved.

Appendix

We closely follow the arguments from Sections 3.2 and 3.3 of [4] and prove that

$$|v(s, x) - v(t, y)| \le N e^{NT} N_1 (|t - s|^{\delta/2} + |x - y|^\delta) \ , \qquad \text{(A.1)}$$

where $N_1 = (1 + |c|_{,\delta})(|f|_{,\delta} + |g|_{,\delta})$ and N depends only on $|\sigma|_{,1}$ and $|b|_{,1}$.

First, according to Theorem 2.5.9 of [4], for any $\alpha \in \mathfrak{A}$, $s \in [0, T]$, and $x, y \in \mathbb{R}^d$, we have

$$E \sup_{t \le T-s} |x_t^{\alpha,s,x} - x_t^{\alpha,s,y}|^2 \le N e^{N(T-s)} |x - y|^2 \ ,$$

$$E \sup_{t \le T-s} |x_t^{\alpha,s,x} - x_t^{\alpha,s,y}|^\delta \le N e^{N(T-s)} |x - y|^\delta \ .$$

Hence, by remembering (2.7), we get

$$\begin{aligned} |v^\alpha(s, x) - v^\alpha(s, y)| \le N_2 E \int_0^{T-s} & |c^{\alpha_t}(s + t, x_t^{\alpha,s,x}) - c^{\alpha_t}(s + t, x_t^{\alpha,s,y})| \, dt \\ + E \int_0^{T-s} & |f^{\alpha_t}(s + t, x_t^{\alpha,s,x}) - f^{\alpha_t}(s + t, x_t^{\alpha,s,y})| \, dt \\ + E & |g(x_{T-s}^{\alpha,s,x}) - g(x_{T-s}^{\alpha,s,y})| \ , \end{aligned}$$

$$|v^\alpha(s, x) - v^\alpha(s, y)| \le N e^{N(T-s)} N_1 |x - y|^\delta \ ,$$

where $N_2 = 2T|f|_0 + 2|g|_0$. This certainly implies (3.3) and

$$|v(s,x) - v(s,y)| \le Ne^{NT} N_1 |x-y|^{\delta} \ . \tag{A.2}$$

Furthermore, by Bellman's principle (Theorem 3.3.6 of [4]), for $s \le t \le T$,

$$v(s,x) = \sup_{\alpha \in \mathfrak{A}} E^{\alpha}_{s,x} \left[\int_0^{t-s} f^{\alpha_r}(s+r, x_r) e^{-\varphi_r}\, dr + v(t, x_{t-s}) e^{-\varphi_{t-s}} \right] \ .$$

Combining this with (A.2), for $t - s \le 1$, we conclude

$$\begin{aligned} |v(s,x) - v(t,x)| &\le (|f|_0 + 2|c|_0 \sup_{[s,T]\times\mathbb{R}^d} |v|)(t-s) \\ &\quad + \sup_{\alpha \in \mathfrak{A}} E^{\alpha}_{s,x} |v(t, x_{t-s}) - v(t,x)| \\ &\le Ne^{NT} N_1 \{|t-s|^{\delta/2} + \sup_{\alpha \in \mathfrak{A}} E^{\alpha}_{s,x} |x_{t-s} - x|^{\delta}\} \\ &\le Ne^{NT} N_1 |t-s|^{\delta/2} \ , \end{aligned}$$

where in the last step we have used again Theorem 2.5.9 of [4]. The same inequality trivially holds if $|t-s| \ge 1$. This and (A.2) yield (A.1).

Applying this argument to the case $A = \{\beta\}$ and using (3.3) we also get (3.4).

Acknowledgement. The author is grateful to the referee for carefully reading and pointing out some obscurities in the original draft.

References

1. Barles, G., Souganidis, P.E.: Convergence of approximation schemes for fully nonlinear second order equations, Asymp. Anal. **4** (3), 271–283 (1991)
2. Fleming, W., Nisio, M.: On stochastic relaxed control for partially observable diffusions, Nagoya Math. J., **93**, 71–108 (1984)
3. Fleming, W., Soner, M.: "Controlled Markov processes and viscosity solutions", Springer Verlag, 1993
4. Krylov, N.V.: "Controlled diffusion processes", Nauka, Moscow, 1977 in Russian; English translation: Springer, 1980
5. Krylov, N.V.: On the rate of convergence of finite-difference approximations for Bellman's equations, Алгебра и Анализ, St. Petersburg Math. J. **9** (3), 245–256 (1997)
6. Krylov, N.V.: Approximating value functions for controlled degenerate diffusion processes by using piece-wise constant policies, Electronic Journal of Probability, **4** (2), 1–19 (1999)
   http://www.math.washington.edu/ ejpecp/EjpVol4/paper2.abs.html
7. Krylov, N.V.: Mean value theorems for stochastic integrals, submitted to the Annals of Probability
8. Kushner, H.J.: Numerical methods for stochastic control problems in continuous time, SIAM J. Control and Optimization, **28** (5), 999–1048 (1990)
9. Kushner, H.J.: Numerical methods for variance control, with applications to optimization in finance, Preprint, April 1998
10. Kushner, H.J., Dupuis, P.G.: "Numerical methods for stochastic control problems in continuous time", Springer Verlag, 1992