*The Annals of Probability*
2001, Vol. 29, No. 1, 385–410

# MEAN VALUE THEOREMS FOR STOCHASTIC INTEGRALS[1]

BY N. V. KRYLOV

*University of Minnesota*

The distributions of stochastic integrals are approximated by the distributions of stochastic integrals of piece-wise constant processes. The rate of approximation in some negative Sobolev spaces is estimated. Generalizations are given for problems arising in control theory.

**1. Introduction.** This article deals with approximating Itô's stochastic integrals of more or less arbitrary integrands with the integrals of piece-wise constant ones, more precisely being constant on each of the intervals $[0, 1/n)$, $[1/n, 2/n)$, and so on, where $n \in \{1, 2, \ldots\}$ is a number *fixed* throughout the article. This is the sense in which we understand mean value theorems for stochastic integrals. The goal is to design the approximations and estimate the error of approximation in terms revealing the dependence on $n$ so as to show that the approximations become better as $n \to \infty$. Regarding the terminology related to Itô's stochastic integrals we refer the reader to [6] or [17].

Very often, if we are given an Itô's stochastic integral

$$\xi := \int_0^1 \sigma_t \, dw_t$$

of a random process $\sigma_t$ against a one-dimensional Wiener process $w_t$, we want to replace it with a finite sum

$$\eta := \sum_{i=0}^{n-1} (w_{(i+1)/n} - w_{i/n}) \sigma_{n,i}. \tag{1.1}$$

If we can write

$$\int_0^1 \sigma_t \, dw_t = (w_1 - w_0) b, \tag{1.2}$$

where $b$ is in some sense a value of the process $\sigma_t$, then we have a mean value theorem for our stochastic integral. Of course, one can always define $b$ by (1.2) (a.s.) and then (1.2) holds by definition. However, usually this $b$ will depend on $w_t$ for $t \in [0, 1]$, and this is not what is usually wanted from expressions like the right-hand side of (1.2). Also, generally, we want $\sigma_{n,i}$ in (1.1) to be independent of $w_t - w_{i/n}$, $t \geq i/n$. Therefore, one cannot hope to have the equality $\xi = \eta$ but rather some kind of approximation.

Received January 1999; revised May 2000.

[1]Supported in part by NSF Grant DMS-96-25483.

*AMS* 2000 *subject classifications*. Primary 93E20; secondary 65M12.

*Key words and phrases*. Stochastic integrals, stochastic control, numerical approximations.

If $\sigma_t$ and $\sigma_{n,i}$ are nonrandom, then

$$E|\xi - \eta|^2 = \int_0^1 \left|\sigma_t - \sigma_t^{(n)}\right|^2 dt, \tag{1.3}$$

where $\sigma_t^{(n)} = \sigma_{n,i}$ for $t \in [i/n, (i+1)/n)$, $i = 0, 1, \dots, n-1$. One can easily check that if $\sigma_t = \sin(2n\pi t)$, then the right-hand side of (1.3) is equal to

$$\begin{aligned}\sum_{i=0}^{n-1} \int_{i/n}^{(i+1)/n} &|\sin(2n\pi t) - \sigma_{n,i}|^2\, dt \\ &= \sum_{i=0}^{n-1} \int_{i/n}^{(i+1)/n} [\sin^2(2n\pi t) + \sigma_{n,i}^2]\, dt \\ &\geq \sum_{i=0}^{n-1} \int_{i/n}^{(i+1)/n} \sin^2(2n\pi t)\, dt = \int_0^1 \sin^2(2\pi t)\, dt,\end{aligned}$$

and the last expression is independent of $n$. This shows that we cannot achieve our goal in the mean-square sense even if we only consider all deterministic $\sigma$ satisfying $|\sigma| \leq 1$.

Therefore, if we want the approximations of $\xi$ by $\eta$ to become better and better as $n \to \infty$, we have either to *restrict the set of processes* $\sigma_t$ under consideration by imposing some kind of control on their continuity, or to relax the sense in which we want to approximate $\xi$. We choose the latter idea, because we have in mind applications to control theory where $\sigma_t$ can play the role of control and any restrictions on its smoothness are inadmissible.

EXAMPLE 1.1 (cf. Exercise 5.3.20 of [8]). One of our results, Theorem 2.9, is applicable to the following situation. Consider the problem of maximizing $Ex_1^2$ over the set of all solutions of $dx_t = \sigma(x_t + \alpha_t)\, dw_t$, $x_0 = 0$, where $w_t$ is again a one-dimensional Wiener process, $\alpha_t$ is an arbitrary $[-1, 1]$-valued Itô integrable process, and $\sigma(y) = (-1) \vee y \wedge 1$. Since $\sigma^2 \leq 1$ and

$$Ex_1^2 = \int_0^1 E\sigma^2(x_t + \alpha_t)\, dt,$$

it is natural to take $\alpha_t$ so that $\sigma^2(x_t + \alpha_t) \equiv 1$, which is achieved if $|x_t + \alpha_t| \geq 1$, say $\alpha_t = \operatorname{sign} x_t$ ($\operatorname{sign} 0 := 1$). In this way naturally we come to the equation

$$dx_t^o = \sigma(x_t^o + \operatorname{sign} x_t^o)\, dw_t, \qquad x_0^o = 0. \tag{1.4}$$

If $x_t^o$ solves this equation, then $\alpha_t^o = \operatorname{sign} x_t^o$ is an optimal process since its response defined by $dx_t = \sigma(x_t + \alpha_t^o)\, dw_t$, $x_0 = 0$, coincides with $x_t^o$ due to uniqueness and satisfies

$$\sigma(x_t + \alpha_t^o) = \sigma(x_t^o + \operatorname{sign} x_t^o) = \operatorname{sign} x_t^o.$$

Notice that $|\operatorname{sign} x_t^o| = 1$, so that by Lévy's theorem $x_t^o$ is a Wiener process and $\alpha_t^o = \operatorname{sign} x_t^o$ has no regularity in time indeed. It is also worth noticing that equation (1.4) may not have solutions at all if the probability space is not rich enough. This is the famous Tanaka's example.

Let us discuss the possibility of approximating *the distribution* of $\xi$ by distributions of $\eta$. If again $\sigma_t$ is nonrandom, then $\xi$ is normal with mean 0 and variance $\int_0^1 \sigma_t^2\,dt$. By taking a nonrandom $b$ such that

$$b^2 = \int_0^1 \sigma_t^2\,dt,$$

we get that $\xi$ and $bw_1$ have just the same distribution. This is again a kind of mean value theorem, since on many occasions $b$ can be taken as a value of $\sigma_t$ at some $t \in [0, 1]$.

The situation becomes much more complicated if we consider *random* processes $\sigma_t$. It turns out (see [4] or [16]) that one can find a process $\sigma_t$, satisfying $1 \le \sigma_t \le 2$ for all $\omega$ and $t$, for which $\xi$ is well defined and the distribution of $\xi$ is singular with respect to Lebesgue measure. In the same time, obviously, the distribution of $\eta$ is absolutely continuous.

Therefore, it is natural to try to approximate the distribution of $\xi$ not in variation norm but in some weaker sense. One of our results (see Remark 2.6) says that if $A$ is a bounded subset of $\mathbb{R}$ and $\sigma_t$ is an $A$-valued process, then, for any integer $n \ge 1$, there exists an $A$-valued process $\sigma_t^{(n)}$ which is constant on each interval $[i/n, (i+1)/n)$ and such that, for any function $g$,

$$|Eg(\xi) - Eg(\eta)| \le Nn^{-1/4} \sup_{x \ne y} \frac{|g(x) - g(y)|}{|x - y|}, \tag{1.5}$$

where $\eta := \int_0^1 \sigma_t^{(n)}\,dw_t$ and the constant $N$ depends only on $A$. Results of such kind can be considered as mean value theorems because the values of $\sigma_t^{(n)}$ are taken from the same set as those of $\sigma_t$.

One can reformulate (1.5) by introducing a norm in the set of all finite (signed) measures $\mu$ on $\mathbb{R}$, satisfying $\mu(\mathbb{R}) = 0$, by the formula

$$||\mu|| := \sup\left\{\int_{\mathbb{R}} g(x)\,\mu(dx) : |g(x) - g(y)| \le |x - y| \quad \forall x, y\right\}.$$

Then (1.5) means that

$$||p_\xi - p_\eta|| \le Nn^{-1/4}, \tag{1.6}$$

where $p_\xi$ and $p_\eta$ are the distributions of $\xi$ and $\eta$, respectively.

One can also rewrite (1.5) in terms of negative Sobolev spaces losing only a little bit of information. The point is that any function from $H_p^\gamma$ (the space of Bessel potentials; see [19]) is Lipschitz continuous, provided $p \in (1, \infty)$ and $(\gamma - 1)p > 1$. Hence,

$$||\mu|| \ge \sup\left\{\int_{\mathbb{R}} g(x)\,\mu(dx) : ||g||_{H_p^\gamma} \le N_1\right\},$$

where $N_1$ is certain constant depending only on $\gamma$ and $p$. This and (1.6) imply that

$$||p_\xi - p_\eta||_{H_q^{-\gamma}} \le Nn^{-1/4},$$

if $\gamma > 1$, $q \in (1, \infty)$ and $(\gamma - 1)q/(q - 1) > 1$.

For the author, the main motivation for proving results like (1.5) is the problem of numerical approximations in control theory. There we need to design a numerical method of finding $\sup Eg(\xi)$, where $\xi$ is defined as before and sup is taken over the set of all nonanticipating $A$-valued processes $\sigma_t$. Knowing how much we cannot lose restricting ourselves to the piece-wise constant processes like $\sigma_t^{(n)}$, we make the problem considerably easier. In [11] and [12] the author presents results of such approach to estimating the rate of convergence of numerical approximations in finding value functions in control problems or solutions of fully nonlinear elliptic and parabolic equations.

In this connection, it is worth mentioning that many authors have been dealing with approximations of value functions or of viscosity or probabilistic solutions for fully nonlinear elliptic and parabolic equations for about twenty five years. There are many books and articles on this subject. We will only mention few of them in which the reader can find further information: [2], [5], [13], [14] and [15]. In almost all papers and books on the subject, the fact of convergence is obtained on the basis of uniqueness of solution to Bellman's equations or on the basis of weak convergence results for stochastic processes. Along these lines, it seems impossible to get any estimate of the rate of convergence, which is of some importance in real applications. The only exceptions, known to the author, are articles [9] and [10] where the rate of convergence is estimated in the case of Bellman's equations with "constant" coefficients. The results of the present article are aimed at Bellman's equations with variable coefficients, whereas the results like (1.5) come out as byproducts of, perhaps, wider interest. Speaking about [9], it is also worth saying that, in contrast with the situation there, we do not know anything about sharpness of the results in this article.

The article is organized as follows. In Section 2 we present main results. Theorems 2.4 and 2.8 are generalizations of (1.5) for multidimensional case. Their proofs, based on Theorem 2.7 and the minimax theorem, are given in the same Section 2. Theorem 2.7 is the central result of the paper. Its proof, presented in Section 3, is based on some quite elementary ideas from control theory and is close to some arguments from [9]. Although, Theorem 2.7 relates to processes with "constant" coefficients it easily allows one to prove Theorem 2.9 in Section 4 by using a penalization method (in the spirit of [7]).

**2. Main results.** Let $(\Omega, \mathscr{F}, P)$ be a complete probability space, $\{\mathscr{F}_t; t \geq 0\}$ be an increasing filtration of $\sigma$-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 $\{\mathscr{F}_t\}$, or in other terms, that $\{w_t, \mathscr{F}_t\}$ is a Wiener process.

Let $A$ be a separable metric space (the set of all admissible controls in control theory) with metric $\rho$ satisfying $\rho \leq 1$ and let $T \in (0, \infty)$, $K \in [1, \infty)$, and $\delta_0, \delta \in (0, 1]$ be some constants satisfying $\delta_0 \leq \delta$.

DEFINITION 2.1. An $A$-valued process $\alpha_t = \alpha_t(\omega)$ defined for all $t \geq 0$ and $\omega \in \Omega$ is called *$\mathscr{F}_t$-admissible* if it is $\mathscr{F} \otimes \mathscr{B}([0, \infty))$-measurable with respect to $(\omega, t)$ and $\mathscr{F}_t$-measurable with respect to $\omega$ for each $t \geq 0$. The set of all $\mathscr{F}_t$-admissible processes is denoted by $\mathfrak{A} = \mathfrak{A}(\mathscr{F}_\cdot)$. Let $\mathfrak{A}_n$ be the subset of $\mathfrak{A}$ consisting of all processes $\alpha_t$ which are constant on intervals $[0, 1/n)$, $[1/n, 2/n)$, and so on.

Fix an integer $d \geq 1$ and suppose that on $A \times [0, T] \times \mathbb{R}^d$ we are given a $d \times d_1$ matrix-valued function $\sigma(\alpha, t, x)$ and an $\mathbb{R}^d$-valued function $b(\alpha, t, x)$. We assume that these functions are Borel measurable. For any matrix $\sigma = (\sigma^{ij})$ denote

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

ASSUMPTION 2.2. (i) *The functions $\sigma(\alpha, t, x)$ and $b(\alpha, t, x)$ are continuous with respect to $\alpha$ and for any $\alpha \in A$, $t \in [0, T]$ and $x, y \in \mathbb{R}^d$,*

$$\begin{aligned} ||\sigma(\alpha, t, x)|| + |b(\alpha, t, x)| &\leq K, \\ ||\sigma(\alpha, t, x) - \sigma(\alpha, t, y)|| + |b(\alpha, t, x) - b(\alpha, t, y)| &\leq K|x - y|. \end{aligned} \tag{2.1}$$

(ii) *On $(\Omega, \mathscr{F}, P)$ there exists a $[0, 1]$-valued uniformly distributed random variable $\xi(\omega)$ independent of $\{\mathscr{F}_\cdot\}$.*

DEFINITION 2.3. Denote $\mathscr{G}_t = \mathscr{F}_t \vee \sigma(\xi)$ and let $\mathfrak{A}_n(\mathscr{G}_\cdot)$ be the subset of $\mathfrak{A}(\mathscr{G}_\cdot)$ consisting of all processes $\alpha_t$ which are constant on intervals $[0, 1/n)$, $[1/n, 2/n)$ and so on.

By Itô's theorem, for any $\alpha \in \mathfrak{A}(\mathscr{G}_\cdot)$ there exists a unique solution $x_t = x_t^\alpha$, $t \in [0, T]$, of the following equation

$$x_t = \int_0^t \sigma(\alpha_s, s, x_s)\, dw_s + \int_0^t b(\alpha_s, s, x_s)\, ds.$$

For Borel functions $f(\alpha, t, x)$ defined on $A \times [0, T] \times \mathbb{R}^d$, which are *continuous* with respect to $(\alpha, x)$ for any $t$, we define

$$|f|_0 = \sup_{\alpha \in A} \sup_{t \in [0,T]} \sup_{x \in \mathbb{R}^d} |f(\alpha, t, x)|,$$

$$[f]_\delta := \sup_{\alpha \in A} \sup_{t \in [0,T]} \sup_{\substack{x, y \in \mathbb{R}^d \\ x \neq y}} \frac{|f(\alpha, t, x) - f(\alpha, t, y)|}{|x - y|^\delta},$$

$$[|f|]_{\delta/2} = \sup_{|s| \leq 2} |s|^{-\delta/2} \int_0^T \sup_{\alpha \in A} \sup_{x \in \mathbb{R}^d} |f(\alpha, s + r, x) - f(\alpha, r, x)|\, dr,$$

where in the last expression we put $f(\alpha, r, x) = f(\alpha, T, x)$ if $r \geq T$ and $f(\alpha, r, x) = f(\alpha, 0, x)$ if $r \leq 0$. We also write $f \in C^\delta$ if

$$|f|_\delta := |f|_0 + [f]_\delta < \infty$$

and we write $f \in C^{\delta/2,\delta}$ if

$$|f|_{\delta/2,\delta} := (1+T)|f|_\delta + [|f|]_{\delta/2} < \infty.$$

Notice that we measure the continuity in time by using integral norms in order to be able to treat the case of data which are just piecewise constant in time.

One of our main results is the following theorem, in which and everywhere in the article by $N$, $N_i$ we denote various constants depending only on $K$ and $d$ unless explicitly stated otherwise.

THEOREM 2.4. *Let $\sigma(\alpha, s, x)$ and $b(\alpha, s, x)$ be* independent *of $(s, x)$. Then for any $\alpha \in \mathfrak{A}$, $n \geq 1$ and constant $K_1 \in [0, \infty)$ there exists $\alpha(n) \in \mathfrak{A}_n(\mathscr{G})$ (also depending on $K_1$, $T$,...) such that with a constant $N = N(K, d)$ the inequality*

$$\begin{aligned} &\left| E\left[\int_0^T f(\alpha_t, t, x_t^\alpha)\,dt + g(x_T^\alpha)\right] \right.\\ (2.2)\qquad &\left. -E\left[\int_0^T f\left(\alpha_t(n), t, x_t^{\alpha(n)}\right)\,dt + g\left(x_T^{\alpha(n)}\right)\right]\right| \\ &\leq N(1+T)n^{-\delta/4}(|f|_{\delta/2,\delta} + [g]_\delta) \end{aligned}$$

*holds for any functions $f = f(\alpha, t, x)$ and $g = g(x)$ of class $C^{\delta/2,\delta}$ satisfying*

$$\begin{aligned} &\sup_{t\in(0,T)}\sup_{x\in\mathbb{R}^d}\left|f(\alpha, t, x) - f(\beta, t, x)\right| \\ (2.3)\qquad &\leq K_1\rho(\alpha, \beta)(|f|_{\delta/2,\delta} + [g]_\delta) \qquad \forall \alpha, \beta \in A. \end{aligned}$$

REMARK 2.5. Condition (2.3) is automatically satisfied in two cases: if $f$ is independent of $\alpha$; or if the set $A$ is finite. In the first case this is obvious. In the second case notice that $\rho(\alpha, \beta)$ is either zero or bounded away from zero by a constant, say $\kappa > 0$. Then (2.3) is satisfied with $K_1 = 2\kappa^{-1}$ since $|f(\alpha, t, x) - f(\beta, t, x)| \leq 2|f|_0$.

REMARK 2.6. We obtain (1.5) from (2.2) if we take $\delta = 1$, $\sigma(\alpha) = \alpha$, $b = f = 0$.

We derive this theorem from the following result which looks much weaker at first sight.

THEOREM 2.7. *Let $\sigma(\alpha, s, x)$ and $b(\alpha, s, x)$ be* independent *of $(s, x)$. Then for any $f \in C^{\delta/2,\delta}$, $g = g(x) \in C^\delta$, $n \geq 1$, and $\alpha \in \mathfrak{A}$ there exists $\alpha(n) \in \mathfrak{A}_n$*

*such that*

$$E\left[\int_0^T f(\alpha_t,t,x_t^\alpha)\,dt+g(x_T^\alpha)\right]$$
$$\le E\left[\int_0^T f(\alpha_t(n),t,x_t^{\alpha(n)})\,dt+g(x_T^{\alpha(n)})\right]$$
$$+N_1(1+T)n^{-\delta/4}([|f|]_{\delta/2}+(1+T)[f]_\delta+n^{(\delta-2)/4}|f|_0+[g]_\delta),$$

*where the constant $N_1$ depends only on $K$ and $d$.*

This theorem is proved in Section 3.

PROOF OF THEOREM 2.4. Take $n\ge 1$, $K_1\in[0,\infty)$, and $\alpha\in\mathfrak{A}$. From Lemma 3.2.6 of [8] we know that there exists a sequence of functions $\beta(m)\in\mathfrak{A}$, each taking only finitely many values in $A$, such that

$$\lim_{m\to\infty} E\int_0^T \rho(\alpha_t,\beta_t(m))\,dt=0.$$

As in the proof of Lemma 3.2.7 of [8] we have

$$\lim_{m\to\infty} E\sup_{t\le T}|x_t^\alpha-x_t^{\beta(m)}|^\delta=0.$$

Obviously,

$$\left|E\left[\int_0^T f(\alpha_t,t,x_t^\alpha)\,dt+g(x_T^\alpha)\right]-E\left[\int_0^T f\left(\beta_t(m),t,x_t^{\beta(m)}\right)\,dt+g\left(x_T^{\beta(m)}\right)\right]\right|$$
$$\le(1+T)([f]_\delta+[g]_\delta)E\sup_{t\le T}\left|x_t^\alpha-x_t^{\beta(m)}\right|^\delta$$
$$+E\int_0^T|f(\alpha_t,t,x_t^\alpha)-f(\beta_t(m),t,x_t^\alpha)|\,dt,$$

where owing to (2.3), the last term is less than

$$K_1(|f|_{\delta/2,\delta}+[g]_\delta)E\int_0^T\rho(\alpha_t,\beta_t(m))\,dt.$$

For any $N,T,n$ one can find $m$ such that

$$(1+T)E\sup_{t\le T}\left|x_t^\alpha-x_t^{\beta(m)}\right|^\delta+K_1E\int_0^T\rho(\alpha_t,\beta_t(m))\,dt\le N(1+T)n^{-\delta/4}.$$

By fixing an appropriate $m$ we see that we only need to find $\alpha(n)\in\mathfrak{A}_n(\mathscr{G}_\cdot)$ such that (2.2) holds with $\beta(m)$ in place of $\alpha$. Since $\beta_t(m)$ takes only finitely many values in $A$, we may and will assume in the remaining part of the proof that the set $A$ is finite. By the way, by Remark 2.5 in this case there always exists a constant $K_1$ such that (2.3) holds for all $f$.

To make further simplifications, observe that one has estimates of moments of $\sup_{t\le T}|x_t^\beta|$ uniform with respect to $\beta\in\mathfrak{A}(\mathscr{G}_\cdot)$. It follows easily that we need to prove (2.2) only for $f$ and $g$ vanishing for large $x$. Fix $R \in (0, \infty)$ and only consider $(f, g)$ such that $f(\beta, t, x) = g(x) = 0$ for $|x| \geq R$. The set of such couples $(f, g)$ satisfying the additional condition

$$T^{-1}|f|_{\delta/2,\delta} + [g]_\delta \leq 1$$

will be denoted by $F$. By Arzelà's theorem (see [3]), $F$ is a compact convex closed set in the Banach space

$$X = L_2^{|A|}([0, T], C(B_R)) \times C(B_R),$$

with norm of $(f, g) \in X$ defined by

$$||(f, g)||_X = \left( \sum_{\beta \in A} \int_0^T \sup_{x \in B_R} |f(\beta, t, x)|^2 \, dt \right)^{1/2} + \sup_{x \in B_R} |g(x)|,$$

where $|A|$ is the number of elements in $A$ and $B_R = \{|x| \leq R\}$. Observe that the space $X^*$ dual to $X$ is naturally identified with

$$Y = L_2^{|A|}([0, T], M(B_R)) \times M(B_R),$$

where $M(B_R)$ is the space of all measures on $B_R$ (see, e.g., Chapter 6, Section 2, exercise 21 of [1]). By Alaoglu's theorem (see [3]) and due to separability of $X$, the unit ball in $Y$ with $X$-topology is a compact metrizable space.

Next, for $\beta \in \mathfrak{A}(\mathscr{G})$ and Borel sets $\Gamma \subset A \times [0, T] \times \mathbb{R}^d$ and $\Pi \subset \mathbb{R}^d$ denote

$$\mu^\beta(\Gamma) = \frac{1}{T} E \int_0^T I_\Gamma(\beta_t, t, x_t^\beta) \, dt, \quad \nu^\beta(\Pi) = E I_\Pi(x_T^\beta),$$
$$G := \{\mu^\beta \times \nu^\beta : \beta \in \mathfrak{A}_n\}.$$

Obviously,

$$\int f(\gamma, t, x) \, \mu^\beta(d\gamma dt dx) = \frac{1}{T} E \int_0^T f(\beta_t, t, x_t^\beta) \, dt,$$
$$\int g(x) \, \nu^\beta(dx) = E g(x_T^\beta).$$

It follows that each element of $G$ defines a bounded linear functional on $X$, so that $G \subset Y$. Moreover, the set $G$ is bounded in $Y$. Denote by $\mathrm{Conv}\,(G)$ the least convex set containing $G$ and by $H$ the closure of $\mathrm{Conv}\,(G)$ in $X$-topology of $Y$. By the above, $H$ is a compact metric space. Observe also that for any $\gamma \in H$ there exist $p^{ki} \geq 0$, $i = 1, \ldots, k$ such that $\sum_i p^{ki} = 1$ and $\alpha_{ki} \in \mathfrak{A}_n$ such that

$$\sum_{i=1}^k p^{ki} \mu^{\alpha_{ki}} \times \nu^{\alpha_{ki}} \to \gamma \tag{2.4}$$

as $k \to \infty$ in $X$-topology of $Y$. To finish with auxiliary observations notice that each element $\gamma \in Y$ acts on couples $(f, g) \in X$ as a couple of measures

$(\gamma, (f, g)) = \int f\, d\gamma_1 + \int g\, d\gamma_2$. For $\gamma \in H$, both $\gamma_1$ and $\gamma_2$ are probability measures due to (2.4), so that we can write

$$(\gamma, (f, g)) = \int (f + g)\, \gamma(d\beta dt dx dx).$$

Now, due to the convexity and compactness of the sets $F$ and $H$, by the minimax theorem (see Theorem 1 on page 220 of [18]),

$$\min_{\gamma\in H} \max_{(f,g)\in F} \Big[ \int f(\beta,t,x)\mu^{\alpha}(d\beta dt dx) + \int g(x)\nu^{\alpha}(dx) - \int (f+g)\gamma(d\beta dt dx dx)\Big]$$

$$= \max_{(f,g)\in F} \min_{\gamma\in H} \Big[ \int f(\beta,t,x)\mu^{\alpha}(d\beta dt dx) + \int g(x)\nu^{\alpha}(dx)$$

$$- \int (f+g)\gamma(d\beta dt dx dx)\Big].$$

By (2.4), the last minimum equals

$$\inf_{\alpha(n)\in\mathfrak{A}_n} \Big[ \int f(\beta, t, x)\, \mu^{\alpha}(d\beta dt dx) + \int g(x)\, \nu^{\alpha}(dx)$$

$$- \int f(\beta, t, x)\, \mu^{\alpha(n)}(d\beta dt dx) - \int g(x)\, \nu^{\alpha(n)}(dx)\Big],$$

which is less than $N_1(1 + T)n^{-\delta/4}$ by Theorem 2.7.

Hence, there exists $\gamma \in H$ such that, for any $(f, g) \in F$,

$$\int f(\beta, t, x)\, \mu^{\alpha}(d\beta dt dx) + \int g(x)\, \nu^{\alpha}(dx)$$

$$\le \int (f + g)\, \gamma(d\beta dt dx dx) + N_1(1 + T)n^{-\delta/4}.$$

Again by (2.4) and compactness of $F$, there is $k \ge 1$, $p^i \ge 0$, $i = 1, \ldots, k$ such that $\sum_i p^i = 1$, and $\alpha_i \in \mathfrak{A}_n$ such that

$$\int f(\beta, t, x)\, \mu^{\alpha}(d\beta dt dx) + \int g(x)\, \nu^{\alpha}(dx)$$

$$\le \sum_{i=1}^{k} p^i \Big[ \int f(\beta, t, x)\, \mu^{\alpha_i}(d\beta dt dx) + \int g(x)\, \nu^{\alpha_i}(dx)\Big] + 2N_1(1 + T)n^{-\delta/4}$$

for any $(f, g) \in F$. On the basis of $p^i$ and $\alpha_i$ we now define a process $\alpha(n) \in \mathfrak{A}_n(\mathscr{G}_\cdot)$ by the formula

$$\alpha_t(n, \omega) = \alpha_{it}(\omega) \qquad \text{if } \sum_{j<i} p^j \le \xi(\omega) < \sum_{j\le i} p^j \quad \Big(\sum_{\varnothing} := 0\Big).$$

Since $\xi$ and $\alpha_i$'s are independent, it is easy to see that

$$\sum_{i=1}^{k} p^i \left[ \int f(\beta, t, x)\, \mu^{\alpha_i}(d\beta dt dx) + \int g(x)\, \nu^{\alpha_i}(dx) \right]$$

$$= \sum_{i=1}^{k} p^i E \left[ T^{-1} \int_0^T f(\alpha_{it}, t, x_t^{\alpha_i})\, dt + g(x_T^{\alpha_i}) \right]$$

$$= E \left[ T^{-1} \int_0^T f(\alpha_t(n), t, x_t^{\alpha(n)})\, dt + g(x_T^{\alpha(n)}) \right].$$

Thus,

$$E \left[ T^{-1} \int_0^T f(\alpha_t, t, x_t^{\alpha})\, dt + g(x_T^{\alpha}) \right]$$

$$\leq E \left[ T^{-1} \int_0^T f(\alpha_t(n), t, x_t^{\alpha(n)})\, dt + g(x_T^{\alpha(n)}) \right] + 2N_1(1+T)n^{-\delta/4}$$

for any $(f, g) \in F$. This yields (2.2) by homogeneity and by replacing $f$ and $g$ with $-f$ and $-g$. The theorem is proved. □

The following is a version of Theorem 2.4 when no kind of continuity of $f$ with respect to $t$ is required but instead we assume that $f$ is independent of $\alpha$.

THEOREM 2.8. *In Theorem* 2.4 *replace* $C^{\delta/2,\delta}$ *with* $C^{\delta}$. *Then its assertion will remain to hold if we consider only the functions* $f$ *independent of* $\alpha$ *and in* (2.2) *replace the right-hand side with*

$$N(1+T^2)(n^{-\delta/8}|f|_{\delta} + n^{-\delta/4}[g]_{\delta}).$$

PROOF. Take a nonnegative function $\zeta \in C_0^{\infty}(\mathbb{R}_+) = C_0^{\infty}((0,\infty))$ vanishing for $t \geq 1$ and with unit integral. For $\varepsilon \in (0,1)$ define $\zeta_{\varepsilon}(t) = \varepsilon^{-1}\zeta(t/\varepsilon)$ and let $f^{(\varepsilon)}(t,x) = \zeta_{\varepsilon}(t) * f(t,x)$, where the convolution is taken with respect to $t$ and we define $f(t,x) = f(0,x)$ for $t \leq 0$ and $f(t,x) = f(T,x)$ for $t \geq T$.

Define $x_t^{\alpha} = 0$ for $t \leq 0$ and notice that

$$\int_0^T f^{(\varepsilon)}(t, x_t^{\alpha})\, dt = \int_0^T \int \zeta_{\varepsilon}(t-s) f(s, x_t^{\alpha})\, ds dt$$

$$\geq -[f]_{\delta} \int_0^T \int \zeta_{\varepsilon}(t-s)|x_t^{\alpha} - x_s^{\alpha}|^{\delta}\, ds dt$$

$$+ \int_0^T \int \zeta_{\varepsilon}(t-s) f(s, x_s^{\alpha})\, ds dt,$$

where by Fubini's theorem the last integral equals

$$\begin{aligned}
&\int_{-\infty}^{T} f(s, x_s^\alpha)\left(\int_{s_+}^{T} \zeta_\varepsilon(t-s)\,dt\right) ds \\
&\quad = \int_{-\varepsilon}^{0} f(0,0)\left(\int_{s_+}^{T} \zeta_\varepsilon(t-s)\,dt\right) ds \\
&\qquad + \int_0^T f(s, x_s^\alpha)\,ds - \int_0^T \left(1 - \int_s^T \zeta_\varepsilon(t-s)\,dt\right) f(s, x_s^\alpha)\,ds \\
&\quad =: I_1 + I_2 + I_3,
\end{aligned}$$

where

$$|I_1| \le |f|_0 \varepsilon, \quad |I_3| \le |f|_0 \int_{T-\varepsilon}^{T} \left(1 - \int_s^T \zeta_\varepsilon(t-s)\,dt\right)\, ds \le |f|_0 \varepsilon.$$

Also, by using the fact that $\zeta_\varepsilon(t-s) = 0$ if $|t-s| \ge \varepsilon$, we get

$$\begin{aligned}
&E\int_0^T \int \zeta_\varepsilon(t-s)|x_t^\alpha - x_s^\alpha|^\delta\, ds dt \\
&\quad = \int_0^T \int \zeta_\varepsilon(t-s) E|x_t^\alpha - x_s^\alpha|^\delta\, ds dt \\
&\quad \le N \int_0^T \int \zeta_\varepsilon(t-s)|t-s|^{\delta/2}\, ds dt \le N T \varepsilon^{\delta/2}.
\end{aligned}$$

Hence,

$$E\int_0^T f(s, x_s^\alpha)\,ds \le E\int_0^T f^{(\varepsilon)}(t, x_t^\alpha)\,dt + N(T[f]_\delta \varepsilon^{\delta/2} + |f|_0 \varepsilon).$$

Below we also use that obviously one can interchange the expectations in the last inequality.

By Theorem 2.4 the above result implies that there is $\alpha(n) \in \mathfrak{A}(\mathscr{G}_\cdot)$ such that

$$\begin{aligned}
&E\left[\int_0^T f(s, x_s^\alpha)\,ds + g(x_T^\alpha)\right] \\
&\quad \le E\left[\int_0^T f^{(\varepsilon)}(s, x_s^{\alpha(n)})\,ds + g(x_T^{\alpha(n)})\right] \\
&\qquad + N(T[f]_\delta \varepsilon^{\delta/2} + |f|_0 \varepsilon) + N(1+T)n^{-\delta/4}(|f^{(\varepsilon)}|_{\delta/2,\delta} + [g]_\delta) \\
&\quad \le E\left[\int_0^T f(s, x_s^{\alpha(n)})\,ds + g(x_T^{\alpha(n)})\right] + N(T[f]_\delta \varepsilon^{\delta/2} + |f|_0 \varepsilon) \\
&\qquad + N(1+T)n^{-\delta/4}(|f^{(\varepsilon)}|_{\delta/2,\delta} + [g]_\delta).
\end{aligned}$$

Upon replacing $f, g$ with $-f, -g$ we see that

$$\left|E\left[\int_0^T f(s, x_s^\alpha)\,ds + g(x_T^\alpha)\right] - E\left[\int_0^T f(s, x_s^{\alpha(n)})\,ds + g(x_T^{\alpha(n)})\right]\right| \tag{2.5}$$
$$\le N(T[f]_\delta \varepsilon^{\delta/2} + |f|_0\varepsilon) + N(1+T)n^{-\delta/4}(|f^{(\varepsilon)}|_{\delta/2,\delta} + [g]_\delta).$$

Finally, if $|t-s| \le \varepsilon$, then

$$|f^{(\varepsilon)}(s, x) - f^{(\varepsilon)}(t, x)| \le |t-s| \sup |D_t f^{(\varepsilon)}|$$
$$\le N|t-s|\varepsilon^{-1}|f|_0 \le N|t-s|^{\delta/2}\varepsilon^{-\delta/2}|f|_0,$$

whereas if $|t-s| \ge \varepsilon$, then

$$|f^{(\varepsilon)}(s, x) - f^{(\varepsilon)}(t, x)| \le 2|f|_0 \le 2|f|_0|t-s|^{\delta/2}\varepsilon^{-\delta/2}.$$

This implies that

$$[|f^{(\varepsilon)}|]_{\delta/2} \le NT|f|_0\varepsilon^{-\delta/2}, \quad |f^{(\varepsilon)}|_{\delta/2,\delta} \le N(1+T)|f|_\delta \varepsilon^{-\delta/2},$$

which implies in turn that the right-hand side of (2.5) is less than

$$N(1+T)|f|_\delta\varepsilon^{\delta/2} + N(1+T^2)n^{-\delta/4}(|f|_\delta\varepsilon^{-\delta/2} + [g]_\delta).$$

Now we get the assertion of the theorem after taking $\varepsilon = n^{-1/4}$ in the last expression. The theorem is proved. □

The following theorem is a generalization of Theorem 2.7 for functions $\sigma$ and $b$ depending on $s, x$. For $\alpha \in \mathfrak{A}$ and some functions $c(\alpha, s, x)$, $f(\alpha, s, x)$ and $g(x)$ to be specified later define

$$\varphi_t^\alpha = \int_0^t c(\alpha_s, s, x_s^\alpha)\,ds,$$
$$v^\alpha = v^\alpha(f, g) = E\left[\int_0^T f(\alpha_s, s, x_s^\alpha)\exp\{-\varphi_s^\alpha\}\,ds + g(x_T^\alpha)\exp\{-\varphi_T^\alpha\}\right].$$

Remember that $\delta, \delta_0$ are some constants satisfying $0 < \delta_0 \le \delta \le 1$.

THEOREM 2.9. *Assume that for any* $|s| \le 2$,

$$\int_0^T \sup_{\alpha\in A}\sup_{x\in\mathbb{R}^d}\big(\|\sigma(\alpha, s+r, x) - \sigma(\alpha, r, x)\| + |b(\alpha, s+r, x) - b(\alpha, r, x))|\big)\,dr$$
$$\le K(1+T)|s|^{\delta_0/2},$$

*where, as usual, we define* $\sigma(\alpha, r, x) = \sigma(\alpha, 0, x)$ *and* $b(\alpha, r, x) = b(\alpha, 0, x)$ *for* $r \le 0$ *and* $\sigma(\alpha, r, x) = \sigma(\alpha, T, x)$ *and* $b(\alpha, r, x) = b(\alpha, T, x)$ *for* $r \ge T$. *Then for any* $c, f \in C^{\delta/2,\delta}$, $g = g(x) \in C^\delta$, $n \ge 1$ *and* $\alpha \in \mathfrak{A}$ *there exists* $\alpha(n) \in \mathfrak{A}_n$ *such that*

$$v^\alpha \le v^{\alpha(n)} + Ne^{NT}(|f|_{\delta/2,\delta} + |g|_\delta)n^{-\delta_0\delta/8}, \tag{2.6}$$

*where the constant* $N$ *depends only on* $K$, $d$ *and* $(1+T)^{-1}|c|_{\delta/2,\delta}$.

We prove Theorem 2.9 in Section 4.

REMARK 2.10. If $f \equiv c \equiv 0$, then Corollary 4.3 says that (2.6) is valid for $\delta \le \delta_0$ as well (remember we assume $\delta_0 \le \delta$). This result is obtained by interpolation, which can also be used to derive other results like Theorem 2.9 when $f \in C^{\gamma/2,\gamma}$ with $\gamma \le \delta_0$.

REMARK 2.11. By replacing $f, g$ with $-f, -g$ in Theorem 2.9 we see that there exists $\beta(n) \in \mathfrak{A}_n$ such that

$$v^\alpha \ge v^{\beta(n)} - Ne^{NT}(|f|_{\delta/2,\delta} + |g|_\delta)n^{-\delta_0\delta/8}.$$

It follows that there exists $p \in [0, 1]$ such that

$$|v^\alpha - [pv^{\beta(n)} + (1-p)v^{\alpha(n)}]| \le Ne^{NT}(|f|_{\delta/2,\delta} + |g|_\delta)n^{-\delta_0\delta/8}.$$

Naturally, $pv^{\beta(n)} + (1-p)v^{\alpha(n)}$ can be interpreted as a result of using $\beta(n)$ with probability $p$ and $\alpha(n)$ with probability $1-p$.

Further development of the idea of "randomizing" from Remark 2.11 leads to the following result, which is derived from Theorem 2.9 in the same way as Theorem 2.4 is derived from Theorem 2.7.

THEOREM 2.12. *Let the assumption of Theorem* 2.9 *be satisfied and let* $c \in C^{\delta/2,\delta}$. *Then for any* $\alpha \in \mathfrak{A}$, $n \ge 1$, *and constant* $K_1 \in [0, \infty)$ *there exists* $\alpha(n) \in \mathfrak{A}_n(\mathscr{G}_\cdot)$ *such that, with a constant* $N$ *depending only on* $K$, $d$ *and* $(1+T)^{-1}|c|_{\delta/2,\delta}$, *the inequality*

$$|v^\alpha(f, g) - v^{\alpha(n)}(f, g)| \le Ne^{NT}(|f|_{\delta/2,\delta} + |g|_\delta)n^{-\delta_0\delta/8}$$

*holds for any functions* $f(\alpha, t, x)$ *and* $g(x)$ *of class* $C^{\delta/2,\delta}$ *satisfying* (2.3).

COROLLARY 2.13. *Let the assumption of Theorem* 2.9 *be satisfied and let* $c, f, g \in C^{\delta/2,\delta}$. *Then for any* $\alpha \in \mathfrak{A}$, $n \ge 1$ *there exists* $\alpha(n) \in \mathfrak{A}_n(\mathscr{G}_\cdot)$ *such that, with a constant* $N$ *depending only on* $K$, $d$, $(1+T)^{-1}|c|_{\delta/2,\delta}$, $(1+T)^{-1}|f|_{\delta/2,\delta}$, *and* $|g|_\delta$, *the inequality*

$$\begin{aligned}\Bigg|E\left|\int_0^T f(\alpha_s, s, x_s^\alpha)\exp\{-\varphi_s^\alpha\}\,ds + g(x_T^\alpha)\exp\{-\varphi_T^\alpha\}\right|^p \\ - E\left|\int_0^T f(\alpha_s(n), s, x_s^\alpha)\exp\{-\varphi_s^{\alpha(n)}\}\,ds + g(x_T^{\alpha(n)})\exp\{-\varphi_T^{\alpha(n)}\}\right|^p\Bigg| \\ \le N^{1+p}\exp\{NT(1+p)\}n^{-\delta_0\delta/8}\end{aligned}$$

*holds for all* $p \ge \delta$.

We prove this corollary in Section 4.

REMARK 2.14. One can also prove a natural version of Theorem 2.9 in the spirit of Theorem 2.8. Theorem 2.12 can also be formulated for functions $f$ independent of $\alpha$. In Corollary 2.13 we can take functions other than powers and consider many $f, c, g$ at once. All these generalizations and extensions are done in the same way as above and are left to the reader.

**3. Proof of Theorem 2.7.** For $s \le 0$ define $f(\alpha, s, x) = f(\alpha, 0, x)$ and for $s \le T$, $x \in \mathbb{R}^d$, and $\alpha \in \mathfrak{A}$ denote

$$v^\alpha(s, x) = E\left[\int_0^{T-s} f(\alpha_t, s+t, x+x_t^\alpha)\,dt + g(x + x_{T-s}^\alpha)\right],$$
$$v(s, x) = \sup_{\alpha\in\mathfrak{A}} v^\alpha(s, x), \qquad v_n(s, x) = \sup_{\alpha\in\mathfrak{A}_n} v^\alpha(s, x).$$

Obviously to prove Theorem 2.7, it suffices to prove

$$\begin{aligned} v^\alpha(0,0) \le v_n(0,0) &+ N(1+T)n^{-\delta/4} \\ &\times([|f|]_{\delta/2} + (1+T)[f]_\delta + n^{(\delta-2)/4}|f|_0 + [g]_\delta). \end{aligned} \tag{3.1}$$

First, we notice some simple properties of the functions $v_n$.

LEMMA 3.1. (i) *For any* $\alpha \in \mathfrak{A}$, $s \le T$, *and* $x, y \in \mathbb{R}^d$,

$$|v^\alpha(s, x) - v^\alpha(s, y)| \le |x-y|^\delta (T[f]_\delta + [g]_\delta),$$

$$|v_n(s, x) - v_n(s, y)| \le |x-y|^\delta (T[f]_\delta + [g]_\delta).$$

(ii) *For any* $\alpha \in \mathfrak{A}$, $s \le t \le T$ *with* $|t-s| \le 1$ *and* $x \in \mathbb{R}^d$,

$$|v^\alpha(t, x) - v^\alpha(s, x)| \le |t-s|\,|f|_0 + N|t-s|^{\delta/2}([|f|]_{\delta/2} + [g]_\delta),$$

$$|v_n(t, x) - v_n(s, x)| \le |t-s|\,|f|_0 + N|t-s|^{\delta/2}([|f|]_{\delta/2} + [g]_\delta).$$

(iii) *For any* $s \le T$ *and* $x \in \mathbb{R}^d$,

$$v_n(s, x) = G_{s,s+1/n} G_{s+1/n,s+2/n} \cdot \dots \cdot G_{s+i/n,T}\, g(x),$$

*where* $i$ *is the integral part of* $n(T-s)$ *and*

$$G_{s,t}u(x) := \sup_{\beta\in A} G_{s,t}^\beta u(x),$$
$$G_{s,t}^\beta u(x) := E\left[\int_0^{t-s} f(\beta, s+r, x+x_r^\beta)\,dr + u(x + x_{t-s}^\beta)\right]$$

*with* $x_t^\beta$ *defined as* $x_t^\alpha$ *for* $\alpha_t \equiv \beta$.

PROOF. Assertion (i) is a straightforward consequence of our assumptions. To prove (ii) it suffices to observe that

$$\begin{aligned}
&|v^\alpha(t,x) - v^\alpha(s,x)| \\
&\quad \le \left| E\int_0^{T-t} f(\alpha_r, t+r, x+x_r^\alpha)\,dr - E\int_0^{T-s} f(\alpha_r, s+r, x+x_r^\alpha)\,dr \right| \\
&\qquad + E|g(x+x_{T-t}^\alpha) - g(x+x_{T-s}^\alpha)| \\
&\quad \le E\int_0^{T-t} |f(\alpha_r, t+r, x+x_r^\alpha) - f(\alpha_r, s+r, x+x_r^\alpha)|\,dr \\
&\qquad + E\int_{T-t}^{T-s} |f(\alpha_r, s+r, x+x_r^\alpha)|\,dr + [g]_\delta E|x_{T-t}^\alpha - x_{T-s}^\alpha|^\delta.
\end{aligned}$$

Assertion (iii) is a particular case of Exercise 3.2.1 of [8], a solution to which can be easily obtained from Lemma 3.2.14 and the proof of Lemma 3.3.1 of [8]. The lemma is proved. □

To proceed further with the proof of Theorem 2.7, take nonnegative functions $\eta \in C_0^\infty(\mathbb{R}_+)$ and $\xi \in C_0^\infty(\mathbb{R}^d)$ with unit integrals. Assume that $\eta(t) = 0$ for $t$ lying outside $(0,1)$ and for $\varepsilon \in (0,1]$ define

$$\eta_{\varepsilon^2}(t) = \varepsilon^{-2}\eta(t/\varepsilon^2), \quad \xi_\varepsilon(x) = \varepsilon^{-d}\xi(x/\varepsilon), \quad \zeta_\varepsilon(t,x) = \eta_{\varepsilon^2}(t)\xi_\varepsilon(x),$$
$$u^{(\varepsilon)} = u * \zeta_\varepsilon.$$

Observe that, by virtue of Lemma 3.1 for $s \le T - 1/n$, we have

$$v_n(s,y) = G_{s,s+1/n} v_n(s+1/n, \cdot)(y). \tag{3.2}$$

We multiply this equality by $\zeta_\varepsilon(t-s, x-y)$ and integrate with respect to $(s,y)$. Also we use the fact that the integral with respect to $s$ can be restricted to $t-\varepsilon^2 \le s \le t$, so that for $t \le T-1/n$ we are integrating well-defined functions. Obviously, for any $\beta \in A$, the integral of the right-hand side of (3.2) is greater than

$$\int_{\mathbb{R}^{d+1}} \zeta_\varepsilon(t-s, x-y) E\left[\int_0^{1/n} f(\beta, s+r, y+x_r^\beta)\,dr + v_n(s+1/n, y+x_{1/n}^\beta)\right] dsdy.$$

We estimate the last expression from below. We have

$$f(\beta, s+r, y+x_r^\beta) \ge f(\beta, s+r, x) - N[f]_\delta \left[\sup_{r\le 1/n} |x_r^\beta|^\delta + |x-y|^\delta\right],$$

with

$$E\sup_{r\le 1/n} |x_r^\beta|^\delta \le N n^{-\delta/2}, \qquad \int_{\mathbb{R}^{d+1}} \zeta_\varepsilon(t-s, x-y)|x-y|^\delta\,dyds = N\varepsilon^\delta.$$

Hence

$$\int_{\mathbb{R}^{d+1}} \zeta_\varepsilon(t-s, x-y) E\int_0^{1/n} f(\beta, s+r, y+x_r^\beta)\,drdsdy$$
$$\geq \int_{s\leq t} \eta_{\varepsilon^2}(t-s)\int_0^{1/n} f(\beta, s+r, x)\,drds - N[f]_\delta n^{-1}(n^{-\delta/2}+\varepsilon^\delta)$$
$$=: n^{-1} f_{n\varepsilon}(\beta, t, x) - N[f]_\delta n^{-1}(n^{-\delta/2}+\varepsilon^\delta).$$

In this way we get from (3.2) that for $t \leq T - 1/n$ and any $\beta \in A$,

$$v_n^{(\varepsilon)}(t,x) \geq n^{-1} f_{n\varepsilon}(\beta, t, x) + Ev_n^{(\varepsilon)}(t+1/n, x+x_{1/n}^\beta) - N[f]_\delta n^{-1}(n^{-\delta/2}+\varepsilon^\delta). \tag{3.3}$$

Observe that $v_n^{(\varepsilon)}$ is an infinitely differentiable function. Therefore, we can apply Itô's formula on the right of (3.3). Define

$$L^\alpha u(t,x) = a^{ij}(\alpha)u_{x^ix^j}(t,x) + b^i(\alpha)u_{x^i}(t,x),$$

where $(a^{ij}(\alpha)) = \frac{1}{2}\sigma(\alpha)\sigma^*(\alpha)$. From (3.3) we obtain

$$n\int_0^{1/n} E[D_t + L^\beta]v_n^{(\varepsilon)}(t+r, x+x_r^\beta)\,dr + f_{n\varepsilon}(\beta, t, x)$$
$$\leq N[f]_\delta(n^{-\delta/2}+\varepsilon^\delta).$$

We replace the first term on the left with $[D_t + L^\beta]v_n^{(\varepsilon)}(t,x)$ by using again Itô's formula. We get

$$\begin{aligned}&[D_t + L^\beta]v_n^{(\varepsilon)}(t,x)\\ &\quad + n\int_0^{1/n}\int_0^r E[D_t+L^\beta]^2 v_n^{(\varepsilon)}(t+p, x+x_p^\beta)\,dpdr + f_{n\varepsilon}(\beta,t,x)\\ &\quad \leq N[f]_\delta(n^{-\delta/2}+\varepsilon^\delta).\end{aligned} \tag{3.4}$$

To proceed further notice that

$$v_n^{(\varepsilon)}(t,x) = \int_{\mathbb{R}} \eta_{\varepsilon^2}(t-s)v_n^\varepsilon(s,x)\,ds,$$

where

$$v_n^\varepsilon(t,x) = \int_{\mathbb{R}^d} \xi_\varepsilon(x-y)v_n(t,y)\,dy$$

and denote

$$K_0 = [|f|]_{\delta/2} + (1+T)[f]_\delta + \varepsilon^{2-\delta}|f|_0 + [g]_\delta.$$

Use Lemma 3.1 to find that for $s \leq T$,

$$|D_t^2 v_n^{(\varepsilon)}(s,x)| = \varepsilon^{-6}\left|\int_{\mathbb{R}} \eta''(\varepsilon^{-2}r)(v_n^\varepsilon(s-r,x) - v_n^\varepsilon(s,x))\,dr\right|$$
$$\leq NK_0\varepsilon^{-6}\int_{\mathbb{R}} |\eta''(\varepsilon^{-2}r)|\,|r|^{\delta/2}\,dr = NK_0\varepsilon^{\delta-4}.$$

Similarly (here comes the only place where the constants $N$ depend on $d$) for derivatives $D^i v_n^{(\varepsilon)}$ with respect to $x$ of order $i$ we have

$$\begin{aligned}|D^i v_n^{(\varepsilon)}(s,x)| &\le \sup_{s\le T} |D^i v_n^{\varepsilon}(s,x)| \\ &= \sup_{s\le T} \varepsilon^{-i-d} \left| \int_{\mathbb{R}^d} (D^i\xi)(\varepsilon^{-1}y)[v_n(s,x-y)-v_n(s,x)]\,dy \right| \\ &\le NK_0\varepsilon^{\delta-i}, \\ |D_t D^i v_n^{(\varepsilon)}(s,x)| &= \varepsilon^{-4}\left| \int_{\mathbb{R}} \eta'(\varepsilon^{-2}r) D^i v_n^{\varepsilon}(s-r,x)\,dr \right| \\ &\le NK_0\varepsilon^{\delta-i-2}.\end{aligned}$$

Now (3.4) implies that, for $t \le T-1/n$, and any $\beta$, $x$

$$[D_t+L^\beta]v_n^{(\varepsilon)}(t,x)+f_{n\varepsilon}(\beta,t,x) \le NK_0 n^{-1}\varepsilon^{\delta-4}+N[f]_\delta(n^{-\delta/2}+\varepsilon^\delta). \tag{3.5}$$

At this point we choose $\varepsilon$. A little bit later we show that the difference $|v_n-v_n^{(\varepsilon)}|$ is less than $N\varepsilon^\delta$. To equate the order of this difference and that of $NK_0 n^{-1}\varepsilon^{\delta-4}$ we put $\varepsilon = n^{-1/4}$. Then from (3.5) we get

$$[D_t+L^\beta]v_n^{(\varepsilon)}(t,x)+f_{n\varepsilon}(\beta,t,x) \le NK_0 n^{-\delta/4}.$$

Hence, by Itô's formula for any $\alpha \in \mathfrak{A}$,

$$\begin{aligned}v_n^{(\varepsilon)}(0,0) &= -E\int_0^{T-1/n} [D_t+L^{\alpha_t}]v_n^{(\varepsilon)}(t,x_t^\alpha)\,dt + Ev_n^{(\varepsilon)}(T-1/n, x_{T-1/n}^\alpha) \\ &\ge E\int_0^{T-1/n} f_{n\varepsilon}(\alpha_t,t,x_t^\alpha)\,dt + Ev_n^{(\varepsilon)}(T-1/n,x_{T-1/n}^\alpha) - NTK_0 n^{-\delta/4}\end{aligned}$$

By combining this with the inequalities

$$\begin{aligned}\left| E\int_{T-1/n}^T f_{n\varepsilon}(\alpha_t,t,x_t^\alpha)\,dt \right| &\le |f|_0 n^{-1} \le |f|_0 \varepsilon^{2-\delta} n^{-\delta/4} \le K_0 n^{-\delta/4}, \\ |v_n(T-1/n,x)-g(y)| &= |v_n(T-1/n,x)-v_n(T,y)| \\ &\le n^{-1}|f|_0 + Nn^{-\delta/2}([|f|]_{\delta/2}+[g]_\delta) \\ &\quad +|x-y|^\delta(T[f]_\delta+[g]_\delta) \\ &\le NK_0 n^{-\delta/4} + |x-y|^\delta(T[f]_\delta+[g]_\delta), \\ Ev_n(T-1/n,x_{T-1/n}^\alpha) &\ge Eg(x_T^\alpha) - NK_0 n^{-\delta/4} \\ &\quad -(T[f]_\delta+[g]_\delta)E|x_{T-1/n}^\alpha - x_T^\alpha|^\delta \\ &\ge Eg(x_T^\alpha) - NK_0 n^{-\delta/4}, \\ |v_n^{(\varepsilon)}(t,x)-v_n(t,x)| &\le \int_{\mathbb{R}^{d+1}} \zeta_\varepsilon(r,y)|v_n(t-r,x-y)-v_n(t,x)|\,drdy\end{aligned}$$

$$\leq N\int_{\mathbb{R}^{d+1}}\zeta_\varepsilon(r,y)\{|f|_0 r + r^{\delta/2}([|f|]_{\delta/2}+[g]_\delta)$$
$$+|y|^\delta(T[f]_\delta+[g]_\delta)\}\,drdy$$
$$\leq N|f|_0\varepsilon^2 + N([|f|]_{\delta/2}+T[f]_\delta+[g]_\delta)\varepsilon^\delta$$
$$\leq NK_0 n^{-\delta/4},$$

we get

$$\begin{aligned}v_n(0,0)+NK_0n^{-\delta/4} &\geq v_n^{(\varepsilon)}(0,0)\\ &\geq E\int_0^T f_{n\varepsilon}(\alpha_t,t,x_t^\alpha)\,dt + Eg(x_T^\alpha)\\ &\quad -N(1+T)K_0n^{-\delta/4}.\end{aligned}\tag{3.6}$$

Finally,

$$E\int_0^T|f_{n\varepsilon}(\alpha_t,t,x_t^\alpha)-f(\alpha_t,t,x_t^\alpha)|\,dt$$
$$\leq\int_0^T\sup_{\alpha\in A}\sup_{x\in\mathbb{R}^d}|f_{n\varepsilon}(\alpha,t,x)-f(\alpha,t,x)|\,dt$$
$$\leq\int_0^T\int_{-\varepsilon^2\leq s\leq 0}\eta_{\varepsilon^2}(s)n\int_0^{1/n}\sup_{\alpha\in A}\sup_{x\in\mathbb{R}^d}|f(\alpha,t-s+r,x)-f(\alpha,t,x)|\,drdsdt$$
$$\leq[|f|]_{\delta/2}\int_{-\varepsilon^2\leq s\leq 0}\eta_{\varepsilon^2}(s)n\int_0^{1/n}(|s|^{\delta/2}+|r|^{\delta/2})\,drds$$
$$\leq[|f|]_{\delta/2}(\varepsilon^\delta+n^{-\delta/2})\leq K_0n^{-\delta/4},$$

which along with (3.6) easily yields (3.1) and brings the proof of Theorem 2.7 to an end. □

**4. Proof of Theorem 2.9.** The most important step in proving Theorem 2.9 is done in the following lemma, which concentrates on the case $f\equiv c\equiv 0$ and prepares everything needed in the general case. The reader will see that the general case follows from Lemma 4.1 quite easily. It is important to draw the reader's attention to the fact that assumptions (2.1) are not used in Lemma 4.1.

LEMMA 4.1. *Take some constants $K_1\in[0,\infty)$, $\delta_1,\delta_2,\delta_3\in(0,1]$ with $\delta_3\geq\delta_1$ and a function $g=g(x)\in C^{\delta_3}$ and assume the following.*

(i) *The inequalities*

$$\int_0^T\sup_{\alpha\in A}\sup_{x\in\mathbb{R}^d}\{||\sigma(\alpha,s+r,x)-\sigma(\alpha,r,x)||+|b(\alpha,s+r,x)-b(\alpha,r,x)|\}\,dr$$
$$\leq K_1(1+T)|s|^{\delta_1/2},$$

$$||\sigma(\alpha,t,x)||+|b(\alpha,t,x)|\leq K_1,$$
$$||\sigma(\alpha,t,x)-\sigma(\alpha,t,y)||+|b(\alpha,t,x)-b(\alpha,t,y)|\leq K_1|x-y|^{\delta_1}$$

*hold for any* $x, y \in \mathbb{R}^d$, $t \in [0, T]$ *and* $|s| \le 2$, *where we extend* $\sigma(\alpha, r, x)$ *and* $b(\alpha, r, x)$ *outside* $[0, T]$ *as usual as their values at the end points.*

(ii) *If* $x_t$ *and* $y_t$ *are* $d$-*dimensional processes satisfying for* $t \in [0, T]$ *the equation*

$$\begin{aligned} x_t - y_t = &\int_0^t \{\sigma(\alpha_r, r, x_r) - \sigma(\alpha_r, r, y_r)\}\, dw_r \\ &+ \int_0^t \{b(\alpha_r, r, x_r) - b(\alpha_r, r, y_r)\}\, dr + \int_0^t \sigma_r\, dw_r + \int_0^t b_r\, dr, \end{aligned} \tag{4.1}$$

*where* $\alpha \in \mathfrak{A}$ *and* $\sigma_t$ *and* $b_t$ *are appropriately measurable processes of appropriate dimensions, then*

$$|Eg(x_T) - Eg(y_T)| \le |g|_{\delta_3} K_1 e^{K_1 T} \left( E \int_0^T \{||\sigma_t||^2 + |b_t|^2\}\, dt \right)^{\delta_2/2}. \tag{4.2}$$

*We assert that under these assumptions for any* $n \ge 1$ *and* $\alpha \in \mathfrak{A}$ *there exists* $\alpha(n) \in \mathfrak{A}_n$ *such that*

$$Eg(x_T^{\alpha}) \le Eg(x_T^{\alpha(n)}) + Ne^{NT} n^{-\delta_1\delta_2/8} |g|_{\delta_3}, \tag{4.3}$$

*where* $N$ *is a constant depending only on* $K_1$ *and* $d$.

PROOF. Recall that $\sigma(\alpha, r, x) = \sigma(\alpha, 0, x)$ and $b(\alpha, r, x) = b(\alpha, 0, x)$ for $r \le 0$, $\sigma(\alpha, r, x) = \sigma(\alpha, T, x)$ and $b(\alpha, r, x) = b(\alpha, T, x)$ for $r \ge T$, and define

$$\tilde{A}(t, x) = \{\big(\sigma(\alpha, t, x), b(\alpha, t, x)\big) : \alpha \in A\}, \qquad \tilde{A} = \bigcup_{t, x} \tilde{A}(t, x).$$

Every element $\tilde{\alpha}$ of the set $\tilde{A}$ is a couple $(\sigma, b)$, which we denote $(\sigma(\tilde{\alpha}), b(\tilde{\alpha}))$. For $\tilde{\alpha}, \tilde{\beta} \in \tilde{A}$ and $t \in \mathbb{R}$, we define

$$\operatorname{dist}(\tilde{\alpha}, \tilde{\beta}) = ||\sigma(\tilde{\alpha}) - \sigma(\tilde{\beta})|| + |b(\tilde{\alpha}) - b(\tilde{\beta})|,$$

$$\gamma(\tilde{\alpha}, t, x) = \operatorname{dist}(\tilde{\alpha}, \tilde{A}(t, x)).$$

We also define

$$\pi(s, t) = \sup_{\alpha \in A} \sup_{x \in \mathbb{R}^d} (||\sigma(\alpha, s, x) - \sigma(\alpha, t, x)|| + |b(\alpha, s, x) - b(\alpha, t, x)|).$$

Observe that, for any $t, s \in [0, T]$ and $x, y \in \mathbb{R}^d$, we have

$$\begin{aligned} |\gamma(\tilde{\alpha}, t, x) - \gamma(\tilde{\alpha}, s, y)| &\le \sup_{\alpha \in A} \{||\sigma(\alpha, t, x) - \sigma(\alpha, s, y)|| + |b(\alpha, t, x) - b(\alpha, s, y)|\} \\ &\le K_1 |x - y|^{\delta_1} + \pi(s, t). \end{aligned}$$

Hence,

$$\begin{gathered} \big|\gamma(\tilde{\alpha}, t, x) - \gamma(\tilde{\alpha}, t, y)\big| \le K_1 |x - y|^{\delta_1}, \\ \int_0^T \sup_{\tilde{\alpha} \in \tilde{A}} \sup_{x \in \mathbb{R}^d} \big|\gamma(\tilde{\alpha}, s + r, x) - \gamma(\tilde{\alpha}, r, x)\big|\, dr \le K_1 (1 + T) |s|^{\delta_1/2}, \end{gathered} \tag{4.4}$$

where $|s| \le 2$.

Next, let $\tilde{\mathfrak{A}}$ be the set of all $\tilde{A}$-valued $\mathscr{F}_t$-adapted measurable processes. By $\tilde{\mathfrak{A}}_n$ we denote its subset consisting of processes which are constant on each interval $[i/n, (i+1)/n)$.

For $\tilde{\alpha} \in \tilde{\mathfrak{A}}$ we denote

$$\tilde{x}_t^{\tilde{\alpha}} = \int_0^t \sigma(\tilde{\alpha}_r)\, dw_r + \int_0^t b(\tilde{\alpha}_r)\, dr,$$

$$v = \sup_{\alpha\in\mathfrak{A}} Eg(x_T^\alpha), \quad \tilde{v} = \sup_{\tilde{\alpha}\in\tilde{\mathfrak{A}}} E\left[g(\tilde{x}_T^{\tilde{\alpha}}) - \int_0^T \gamma(\tilde{\alpha}, t, \tilde{x}_t^{\tilde{\alpha}})\, dt\right],$$

$$v_n = \sup_{\alpha\in\mathfrak{A}_n} Eg(x_T^\alpha), \quad \tilde{v}_n = \sup_{\tilde{\alpha}\in\tilde{\mathfrak{A}}_n} E\left[g(\tilde{x}_T^{\tilde{\alpha}}) - \int_0^T \gamma(\tilde{\alpha}, t, \tilde{x}_t^{\tilde{\alpha}})\, dt\right].$$

Our plan is as follows. As in Section 3 it suffices to prove that

$$v \le v_n + Ne^{NT}|g|_{\delta_3} n^{-\delta_1\delta_2/8} \tag{4.5}$$

where, and below in the proof by $N$, we denote various constants depending only on $K_1$ and $d$. Observe that obviously, $v \le \tilde{v}$. We claim that, because of that, to prove (4.5) it suffices to prove that

$$\begin{aligned} \tilde{v} &\le \tilde{v}_n + N(1+T)n^{-\delta_1/4}(1+T+[g]_{\delta_1}),\\ \tilde{v}_n &\le v_n + Ne^{NT}(|g|_{\delta_3}^{2/(2-\delta_2)} + |g|_{\delta_3} n^{-\delta_1\delta_2/4}). \end{aligned} \tag{4.6}$$

Indeed, then we have (remember $\delta_3 \ge \delta_1$)

$$v \le v_n + Ne^{NT}(n^{-\delta_1/4} + n^{-\delta_1/4}|g|_{\delta_3} + |g|_{\delta_3}^{2/(2-\delta_2)} + |g|_{\delta_3} n^{-\delta_1\delta_2/4}).$$

Upon taking here $cg$ in place of $g$, where $c$ is any positive constant, we also conclude

$$v \le v_n + Ne^{NT}\left(n^{-\delta_1/4}c^{-1} + n^{-\delta_1/4}|g|_{\delta_3} + |g|_{\delta_3}^{2/(2-\delta_2)} c^{\delta_2/(2-\delta_2)} + |g|_{\delta_3} n^{-\delta_1\delta_2/4}\right),$$

which for $c^{-1} = |g|_{\delta_3} n^{\delta_1(2-\delta_2)/8}$ yields

$$v \le v_n + Ne^{NT}|g|_{\delta_3}\left(n^{-\delta_1\delta_2/8} + n^{-\delta_1/4} + n^{-\delta_1\delta_2/4}\right),$$

and (4.5) follows.

Thus, to prove the lemma, it only remains to prove (4.6). Observe that the first inequality in (4.6) follows directly from Theorem 2.7 and from (4.4).

To prove the second inequality in (4.6), fix $\varepsilon > 0$ and take an $\tilde{\alpha} \in \tilde{\mathfrak{A}}_n$ such that

$$E\int_0^T \gamma(\tilde{\alpha}, t, \tilde{x}_t^{\tilde{\alpha}})\, dt \le Eg(\tilde{x}_T^{\tilde{\alpha}}) - \tilde{v}_n + \varepsilon =: J + \varepsilon. \tag{4.7}$$

Notice that by (4.4), for $\kappa_n(t) := n^{-1}[nt]$,

$$n\int_{1/n}^{2/n}\int_0^T \pi(s+\kappa_n(t),t)\,dt\,ds = n\int_0^T\int_{\kappa_n(t)-t+1/n}^{\kappa_n(t)-t+2/n}\pi(s+t,t)\,ds\,dt$$
$$\le n\int_0^{2/n}\int_0^T \pi(s+t,t)\,dt\,ds$$
$$\le N(1+T)n^{-\delta_1/2}.$$

Therefore, we can find and fix an $s\in[0,2/n]$ so that

$$\int_0^T \pi(s+\kappa_n(t),t)\,dt \le N(1+T)n^{-\delta_1/2}. \tag{4.8}$$

Now notice that, for any $\varepsilon > 0$, $t, r \in \mathbb{R}$, $x \in \mathbb{R}^d$ and $\omega \in \Omega$, one can find $\alpha \in A$ such that

$$\|\sigma(\tilde{\alpha}_t) - \sigma(\alpha, r, x)\| + |b(\tilde{\alpha}_t) - b(\alpha, r, x)| \le \varepsilon + \gamma(\tilde{\alpha}_t, r, x).$$

Moreover, since $\tilde{\alpha}_t$ is piece-wise constant, the appropriate $\alpha$ can be chosen the same on each interval of time $[i/n, (i+1)/n)$. By remembering that $A$ is a separable metric space, it is easy to understand that there exists an $\alpha \in \mathfrak{A}_n$ such that

$$\begin{aligned} I &:= E\int_0^T \Big[\Big\|\sigma(\tilde{\alpha}_t) - \sigma(\alpha_t, s+\kappa_n(t), \tilde{x}^{\tilde{\alpha}}_{\kappa_n(t)})\Big\| \\ &\qquad + \Big|b(\tilde{\alpha}_t) - b(\alpha_t, s+\kappa_n(t), \tilde{x}^{\tilde{\alpha}}_{\kappa_n(t)})\Big|\Big]^2\,dt \\ &\le 2\varepsilon + 2E\int_0^T \gamma^2(\tilde{\alpha}_t, s+\kappa_n(t), \tilde{x}^{\tilde{\alpha}}_{\kappa_n(t)})\,dt. \end{aligned} \tag{4.9}$$

To estimate the right-hand side of (4.9), use (4.4), (4.7) and (4.8). Then

$$\begin{aligned} E\int_0^T \gamma^2(\tilde{\alpha}_t, s+\kappa_n(t), \tilde{x}^{\tilde{\alpha}}_{\kappa_n(t)})\,dt &\le NE\int_0^T \gamma(\tilde{\alpha}_t, s+\kappa_n(t), \tilde{x}^{\tilde{\alpha}}_{\kappa_n(t)})\,dt \\ &\le NE\int_0^T \gamma(\tilde{\alpha}_t, s+\kappa_n(t), \tilde{x}^{\tilde{\alpha}}_t)\,dt \\ &\quad + N\int_0^T E|\tilde{x}^{\tilde{\alpha}}_{\kappa_n(t)} - \tilde{x}^{\tilde{\alpha}}_t|^{\delta_1}\,dt \\ &\le NE\int_0^T \gamma(\tilde{\alpha}_t, t, \tilde{x}^{\tilde{\alpha}}_t)\,dt + N(1+T)n^{-\delta_1/2} \\ &\le NJ + N\varepsilon + N(1+T)n^{-\delta_1/2}. \end{aligned}$$

Hence,

$$I \le NJ + N\varepsilon + N(1+T)n^{-\delta_1/2}.$$

This, (4.4) and (4.8) yield

$$\text{(4.10)}\qquad \begin{aligned} &E\int_0^T \left[||\sigma(\tilde\alpha_t)-\sigma(\alpha_t,t,\tilde x_t^{\tilde\alpha})||^2+|b(\tilde\alpha_t)-b(\alpha_t,t,\tilde x_t^{\tilde\alpha})|^2\right]dt\\ &\quad\le 2I+N(1+T)n^{-\delta_1/2}\le NJ+N\varepsilon+N(1+T)n^{-\delta_1/2}.\end{aligned}$$

Now we can estimate $\tilde x_t^{\tilde\alpha}-x_t^{\alpha}$ by observing that $x_t:=\tilde x_t^{\tilde\alpha}$ and $y_t:=x_t^{\alpha}$ satisfy (4.1) with

$$\sigma_r:=\sigma(\tilde\alpha_r)-\sigma(\alpha_r,r,\tilde x_r^{\tilde\alpha}),\quad b_r:=b(\tilde\alpha_r)-b(\alpha_r,r,\tilde x_r^{\tilde\alpha}).$$

It follows from (4.10) and (4.2) that

$$\text{(4.11)}\qquad Eg(\tilde x_T^{\tilde\alpha})-Eg(x_T^{\alpha})\le Ne^{NT}|g|_{\delta_3}\left(J+\varepsilon+n^{-\delta_1/2}\right)^{\delta_2/2}.$$

One can stop here if one is satisfied with a result which is a little bit weaker than (4.3). By this we mean that obviously [cf. (4.7)] $\tilde v_n\le Eg(\tilde x_T^{\tilde\alpha})+\varepsilon$, $Eg(x_T^{\alpha})\le v_n$, and $J\le 2|g|_0$, so that (4.11) implies that

$$\tilde v_n\le v_n+Ne^{NT}(|g|_{\delta_3}^{1+\delta_2/2}+|g|_{\delta_3}n^{-\delta_1\delta_2/4}).$$

This combined with the first inequality in (4.6) leads to (4.3) but with $\delta_1\delta_2/(8+4\delta_2)$ in place of $\delta_1\delta_2/8$.

A better result follows after enhancing the estimate of $J$. Observe that $Eg(x_T^{\alpha})\le v_n\le\tilde v_n$, so that by definition (4.7) of $J$, equation (4.11) implies that

$$J\le Ne^{NT}|g|_{\delta_3}\left(J+\varepsilon+n^{-\delta_1/2}\right)^{\delta_2/2}.$$

Here if $\varepsilon+n^{-\delta_1/2}\le J$, then

$$J\le Ne^{NT}|g|_{\delta_3}J^{\delta_2/2},\qquad J\le Ne^{NT}|g|_{\delta_3}^{2/(2-\delta_2)}.$$

In the remaining case and generally,

$$J\le Ne^{NT}|g|_{\delta_3}^{2/(2-\delta_2)}+\varepsilon+n^{-\delta_1/2}.$$

Plugging this result into (4.11) and remembering (4.7), we get

$$\tilde v_n\le Eg(\tilde x_T^{\tilde\alpha})+\varepsilon\le Eg(x_T^{\alpha})+Ne^{NT}\left(|g|_{\delta_3}^{2/(2-\delta_2)}+|g|_{\delta_3}\varepsilon^{\delta_2/2}+|g|_{\delta_3}n^{-\delta_1\delta_2/4}\right)+\varepsilon$$

This proves the second inequality in (4.6) since $\alpha\in\mathfrak{A}_n$ and $\varepsilon>0$ is arbitrary. The lemma is proved. □

REMARK 4.2. Under the assumptions of Theorem 2.9 [including, of course, assumption (2.1)], the assumptions of Lemma 4.1 are satisfied with $\delta_1=\delta_0$ and $\delta_2=\delta_3=\delta$.

Indeed,

$$|Eg(x_T)-Eg(y_T)|\le[g]_\delta E|x_T-y_T|^\delta\le[g]_\delta(E|x_T-y_T|^2)^{\delta/2},$$

and as well known (apply, e.g., Corollary 2.5.10 of [8] to the process $x_t - y_t$) due to (2.1),

$$E \sup_{t \le T} |x_t - y_t|^2 \le Ne^{NT} E \int_0^T \{||\sigma_t||^2 + |b_t|^2\}\, dt, \tag{4.12}$$

where $N$ depends only on $K$.

COROLLARY 4.3. *Theorem* 2.9 *holds for any* $\delta, \delta_0 \in (0,1]$ *if* $f \equiv c \equiv 0$.

PROOF. If $\delta_0 \le \delta$, our assertion follows from Remark 4.2. For $\delta_0 \ge \delta$, we interpolate. Take a nonnegative function $\zeta \in C_0^\infty(\mathbb{R}^d)$ with unit integral and for $\varepsilon > 0$ define $g^{(\varepsilon)}(x) = \varepsilon^{-d}\zeta(x/\varepsilon) * g(x)$. Then as easy to see

$$\left|g(x) - g^{(\varepsilon)}(x)\right| \le \int |g(x) - g(x - \varepsilon y)|\zeta(y)\, dy \le N[g]_\delta \varepsilon^\delta,$$

$$\left|g^{(\varepsilon)}_{x^i}(x)\right| = \varepsilon^{-1-d} \left| \int (g(y) - g(x))\zeta_{x^i}((y-x)/\varepsilon)\, dy \le N[g]_\delta \varepsilon^{\delta - 1},\right.$$

so that

$$|g^{(\varepsilon)}(x) - g^{(\varepsilon)}(y)| \le N[g]_\delta \varepsilon^{\delta-1}|x-y| \le N[g]_\delta \varepsilon^{\delta-\delta_0}|x-y|^{\delta_0}$$

if $|x-y| \le \varepsilon$. Furthermore, if $|x-y| \ge \varepsilon$,

$$|g^{(\varepsilon)}(x) - g^{(\varepsilon)}(y)| \le [g]_\delta |x-y|^\delta \le [g]_\delta \varepsilon^{\delta-\delta_0}|x-y|^{\delta_0}.$$

Hence, $|g^{(\varepsilon)}|_{\delta_0} \le N|g|_\delta \varepsilon^{\delta-\delta_0}$ and by the result for $\delta \ge \delta_0$ and by the inequalities

$$Eg(x_T^\alpha) \le Eg^{(\varepsilon)}(x_T^\alpha) + N|g|_\delta \varepsilon^\delta, \qquad Eg^{(\varepsilon)}(x_T^{\alpha(n)}) \le Eg(x_T^{\alpha(n)}) + N|g|_\delta \varepsilon^\delta,$$

the last term on the right in (2.6) can be replaced with

$$Ne^{NT}\left(\left|g^{(\varepsilon)}\right|_{\delta_0} n^{-\delta_0^2/8} + |g|_\delta \varepsilon^\delta\right) \le Ne^{NT}\left(|g|_\delta \varepsilon^{\delta-\delta_0} n^{-\delta_0^2/8} + |g|_\delta \varepsilon^\delta\right).$$

It follows that setting $\varepsilon = n^{-\delta_0/8}$ right in the beginning of the proof leads to the desired result. □

PROOF OF THEOREM 2.9. Notice that replacing $f(\alpha, t, x)$, $g(x)$, $c(\alpha, t, x)$ with $f(\alpha, t, x)e^{Mt}$, $g(x)e^{MT}$, $c(\alpha, t, x) + M$, respectively, does not modify $v^\alpha$. Therefore, without loss of generality, we assume that $c \ge 0$.

First, assume that in addition,

$$|f|_{\delta/2,\delta} = 1 + T, \tag{4.13}$$

and add two more coordinates $y$ and $z$ which move according to

$$y_t^\alpha = -\int_0^t c(\alpha_r, r, x_r^\alpha)\, dr, \qquad z_t^\alpha = \int_0^t f(\alpha_r, r, x_r^\alpha)\xi(y_r^\alpha)\, dr,$$

where $\xi \in C^\infty(\mathbb{R})$, $\xi(y) = e^y$ for $y \le 0$, $0 \le \xi \le 2$, and $|\xi'| \le 2$. Also define

$$\bar{g}(x, y, z) = (-M) \vee [\xi(y)g(x) + z] \wedge M,$$

where $M = |g|_0 + T + 1$. Then, owing to (4.13) and $c \geq 0$, we have

$$\bar{g}(x^\alpha_T, y^\alpha_T, z^\alpha_T) = (-M) \vee \left[\int_0^T f(\alpha_t, t, x^\alpha_t)\exp\{-\varphi^\alpha_t\}\,dt + g(x^\alpha_T)\exp\{-\varphi^\alpha_T\}\right] \wedge M$$
$$= \int_0^T f(\alpha_t, t, x^\alpha_t)\exp\{-\varphi^\alpha_t\}\,dt + g(x^\alpha_T)\exp\{-\varphi^\alpha_T\},$$

$$v^\alpha = E\bar{g}(x^\alpha_T, y^\alpha_T, z^\alpha_T). \tag{4.14}$$

Now it is natural to apply Lemma 4.1 to the process $(x^\alpha_t, y^\alpha_t, z^\alpha_t)$.

Assumption (i) of Lemma 4.1 is obviously satisfied with a constant $K_1$ depending only on $K$ and $(1+T)^{-1}|c|_{\delta/2,\delta}$. To check assumption (ii), take $\alpha \in \mathfrak{A}$ and two processes $(x_t, y_t, z_t)$ and $(x'_t, y'_t, z'_t)$ and assume that

$$x_t - x'_t = \int_0^t \{\sigma(\alpha_r, r, x_r) - \sigma(\alpha_r, r, x'_r)\}\,dw_r$$
$$+ \int_0^t \{b(\alpha_r, r, x_r) - b(\alpha_r, r, x'_r)\}\,dr + \int_0^t \sigma^{(1)}_r\,dw_r + \int_0^t b^{(1)}_r\,dr,$$
$$y_t - y'_t = -\int_0^t \{c(\alpha_r, r, x_r) - c(\alpha_r, r, x'_r)\}\,dr + \int_0^t \sigma^{(2)}_r\,dw_r + \int_0^t b^{(2)}_r\,dr,$$
$$z_t - z'_t = \int_0^t \{f(\alpha_r, r, x_r)\xi(y_r) - f(\alpha_r, r, x'_r)\xi(y'_r)\}\,dr$$
$$+ \int_0^t \sigma^{(3)}_r\,dw_r + \int_0^t b^{(3)}_r\,dr.$$

Denote

$$I = E\int_0^T \sum_{i=1}^3 \left\{\left\|\sigma^{(i)}_t\right\|^2 + \left|b^{(i)}_t\right|^2\right\}\,dt.$$

We get from (4.12) that

$$E\sup_{t\leq T}|x_t - x'_t|^2 \leq Ne^{NT}E\int_0^T \left\{\left\|\sigma^{(1)}_t\right\|^2 + |b^{(1)}_t|^2\right\}\,dt \leq Ne^{NT}I.$$

Furthermore,

$$E\sup_{t\leq T}|y_t - y'_t|^2 \leq NT^2[c]^2_\delta E\sup_{t\leq T}|x_t - x'_t|^{2\delta} + NE\int_0^T \left\|\sigma^{(2)}_t\right\|^2 dt + NTE\int_0^T \left|b^{(2)}_t\right|^2 dt$$
$$\leq Ne^{NT}(I^\delta + I),$$
$$E|z_T - z'_T|^2 \leq NT^2\left([f]^2_\delta E\sup_{t\leq T}|x_t - x'_t|^{2\delta} + |f|^2_0 E\sup_{t\leq T}|y_t - y'_t|^2\right)$$
$$+ NE\int_0^T \left\|\sigma^{(3)}_t\right\|^2 dt + NTE\int_0^T \left|b^{(3)}_t\right|^2 dt$$
$$\leq Ne^{NT}(I^\delta + I).$$

By using Hölder's inequality, we obtain

$$\begin{aligned}(4.15)\qquad &|E\bar{g}(x_T, y_T, z_T) - E\bar{g}(x'_T, y'_T, z'_T)|\\ &\le E\{2|g(x_T) - g(x'_T)| + 2|g|_0|y_T - y'_T| + |z_T - z'_T|\}\\ &\le N(1 + |g|_\delta)E\{|x_T - x'_T|^\delta + |y_T - y'_T| + |z_T - z'_T|\}\\ &\le Ne^{NT}(1 + |g|_\delta)(I^{\delta/2} + I^{1/2}).\end{aligned}$$

Since $\bar{g}(x, -\infty, 1) = 1$ and $\bar{g}(x, 0, 0) = g(x)$, we have $1 + |g|_\delta \le 2|\bar{g}|_\delta$, and (4.15) implies

$$(4.16)\qquad |E\bar{g}(x_T, y_T, z_T) - E\bar{g}(x'_T, y'_T, z'_T)| \le Ne^{NT}|\bar{g}|_\delta(I^{\delta/2} + I^{1/2}).$$

Obviously, we can drop $I^{1/2}$ in (4.16) if $I \le 1$ (remember $\delta \in (0, 1]$). On the other hand, if $I \ge 1$, then from

$$|E\bar{g}(x_T, y_T, z_T) - E\bar{g}(x'_T, y'_T, z'_T)| \le 2|\bar{g}|_\delta \le 2|\bar{g}|_\delta I^{\delta/2},$$

it follows that again we can drop $I^{1/2}$ in (4.16). This proves that (4.2) is satisfied and by Lemma 4.1 and (4.14), for any $\alpha \in \mathfrak{A}$ there is $\alpha(n) \in \mathfrak{A}_n$ such that

$$(4.17)\qquad v^\alpha(f, g) \le v^{\alpha(n)}(f, g) + Ne^{NT}n^{-\delta_0\delta/8}|\bar{g}|_\delta.$$

Furthermore, as in (4.15), it is easy to see that $|\bar{g}|_\delta \le N(|g|_\delta + T + 1) = N(|f|_{\delta/2,\delta} + |g|_\delta)$. Therefore, (4.17) implies

$$v^\alpha(f, g) \le v^{\alpha(n)}(f, g) + Ne^{NT}n^{-\delta_0\delta/8}(|f|_{\delta/2,\delta} + |g|_\delta).$$

It only remains to notice that the multiplication by a constant $> 0$ always reduces any $f \not\equiv 0$ to the one satisfying (4.13), whereas if $f \equiv 0$, one can take a small constant $\ne 0$ instead of $f$ and pass to the limit. This completes the proof of Theorem 2.9. □

PROOF OF COROLLARY 2.13. As in the above proof, consider the process $(x_t^\alpha, y_t^\alpha, z_t^\alpha)$ this time without assuming (4.13). We have seen above that the process $(x_t^\alpha, y_t^\alpha, z_t^\alpha)$ satisfies the assumptions of Lemma 4.1 for $\delta_1 = \delta_0$ and $\delta_2 = \delta_3 = \delta$. Hence, for any $\alpha \in \mathfrak{A}$, $n \ge 1$, and $g(x, y, z) \in C^\delta$ there is an $\alpha(n) \in \mathfrak{A}_n$ such that

$$Eg(x_t^\alpha, y_t^\alpha, z_t^\alpha) \le Eg\left(x_t^{\alpha(n)}, y_t^{\alpha(n)}, z_t^{\alpha(n)}\right) + Ne^{NT}n^{-\delta_0\delta/8}|g|_\delta.$$

On the basis of this, similarly to the derivation of Theorem 2.4 from Theorem 2.7, one proves that for any $\alpha \in \mathfrak{A}$ and $n \ge 1$ there is an $\alpha(n) \in \mathfrak{A}_n(\mathcal{G}_\cdot)$ such that

$$|Eg(x_t^\alpha, y_t^\alpha, z_t^\alpha) - Eg(x_t^{\alpha(n)}, y_t^{\alpha(n)}, z_t^{\alpha(n)})| \le Ne^{NT}n^{-\delta_0\delta/8}|g|_\delta.$$

for any $g(x, y, z) \in C^\delta$. It only remains to take $g = |z|^p \wedge M^p$, where $M = (T|f|_0 + |g|_0)\exp(|c|_0T)$. □

**Acknowledgments.** The author is sincerely grateful to the referees of the paper for careful reading and many comments which led to improving the presentation.

## REFERENCES

[1] BOURBAKI, N,. (1959). *Intégration*, Chapitre 6. Hermann, Paris.
[2] BARLES, G. and SOUGANIDIS, P. E. (1991). Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic Anal.* **4** 271–283.
[3] DUNFORD, N. and SCHWARTZ, J. T. (1958). *Linear Operators, Part I. General Theory.* Interscience, New York.
[4] FABES, E. B. and KENIG, C. (1981). Examples of singular parabolic measures and singular transition probability densities. *Duke Math. J.* **48** 845–856.
[5] FLEMING, W. H. and SONER, H. M. (1993). *Controlled Markov Processes and Viscosity Solutions.* Springer, New York.
[6] IKEDA, N. and WATANABE, S. (1981). *Stochastic Differential Equations and Diffusion Processes.* North-Holland, Amsterdam.
[7] KRYLOV, N. V. (1973). On the selection of a Markov process from a system of processes and the construction of quasi-diffusion processes. *Izv. Akad. Nauk SSSR* **37** 691–708 (in Russian). [English translation in *Math. USSR Izv.* (1973) **7** 691–709.]
[8] KRYLOV, N.V. (1977). *Controlled Diffusion Processes.* Nauka, Moscow (in Russian). [English translation published (1980) Springer, New York.]
[9] KRYLOV, N. V. (1997). On the rate of convergence of finite–difference approximations for Bellman's equations. *Algebra i Analiz* **9** 245–256.
[10] KRYLOV, N.V. (1997). Fully nonlinear second order elliptic equations: Recent development, *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* **25** 569–595.
[11] KRYLOV, N.V. (1999). Approximating value functions for controlled degenerate diffusion processes by using piece-wise constant policies. *Electron. J. Probab.* **4** 1–19. Also available at `http://www.math.washington.edu/ ejpecp/EjpVol4/paper2.abs.html`.
[12] KRYLOV, N.V. (2000). On the rate of convergence of finite–difference approximations for Bellman's equations with variable coefficients. *Probab. Theory Related Fields* **117** 1–16.
[13] KUO, H. J. and TRUDINGER, N. S. (1992). Discrete methods for fully nonlinear elliptic equations. *SIAM J. Numer. Anal.* **29** 123–135.
[14] KUSHNER, H. J. and DUPUIS, P. G. (1992). *Numerical Methods for Stochastic Control Problems in Continuous Time.* Springer, New York.
[15] PRAGARAUSKAS, G. (1983). Approximation of controlled solutions of Itô equations by controlled Markov chains. *Lit. Mat. Sbornik*. **23** 175–188 (in Russian). [English translation in *Lithuanian Math. J.* **23** (1983) 98–108.]
[16] SAFONOV, M. V. (1981). An example of diffusion process with singular distribution at some given time. In *Third Vilnius Conference on Probability Theory and Mathematical Statistics* 133–134 (in Russian). VNU Science Press, Vilnius.
[17] STROOCK, D. W. and VARADHAN, S. R. S. (1979). *Multidimensional Diffusion Processes.* Springer, Berlin.
[18] SZÉP, J. and FORGÓ, F. (1983). *Introduction to the Theory of Games.* Reidel, Dordrecht.
[19] TRIEBEL, H. (1992). *Theory of Function Spaces II.* Birkhäuser, Boston.

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