

MEAN VALUE THEOREMS FOR STOCHASTIC INTEGRALS

N.V. KRYLOV

SUMMARY. 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

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

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

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

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

If we can write

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

where b is in some sense a value of the process σ_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.

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

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

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

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

1991 *Mathematics Subject Classification*. Primary 93E20, Secondary 65M12.

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

The work was partially supported by NSF Grant DMS-9625483.

and the last expression is independent of n .

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

Let us discuss the possibility of approximating the distribution of ξ by distributions of η . If again σ_t is nonrandom, then ξ 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 ξ 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 σ_t at some $t \in [0, 1]$.

The situation becomes much more complicated if we consider *random* processes σ_t . It turns out (see [2] or [11]) that one can find a process σ_t , satisfying $1 \leq \sigma_t \leq 2$ for all ω and t , for which ξ is well defined and the distribution of ξ is singular with respect to Lebesgue measure. In the same time, obviously, the distribution of η is absolutely continuous.

Therefore, it is natural to try to approximate the distribution of ξ 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 σ_t is an A -valued process, then, for any integer $n \geq 1$, 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 ,

$$|Eg(\xi) - Eg(\eta)| \leq Nn^{-1/4} \sup_{x \neq y} \frac{|g(x) - g(y)|}{|x - y|}, \quad (1.4)$$

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 σ_t^n are taken from the same set as those of σ_t .

One can reformulate (1.4) by introducing a norm in the set of all finite measures μ on \mathbb{R} , satisfying $\mu(\mathbb{R}) = 0$, by the formula

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

Then (1.4) means that

$$\|p_\xi - p_\eta\| \leq Nn^{-1/4}, \quad (1.5)$$

where p_ξ and p_η are the distributions of ξ and η , respectively.

One can also rewrite (1.4) in terms of negative Sobolev spaces losing only a little bit of information. The point is that any function from H_p^γ (the space of Bessel potentials, see [13]) is Lipschitz continuous, provided $p \in (1, \infty)$ and $(\gamma - 1)p > 1$. Hence,

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

where N_1 is certain constant depending only on γ and p . This and (1.5) imply that

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

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

For the author, the main motivation for proving results like (1.4) is the problem of numerical approximations in control theory. There we need to design a numerical method of

finding $\sup Eg(\xi)$, where ξ is defined as before and \sup is taken over the set of all nonanticipating A -valued processes σ_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 one of subsequent articles the author plans to present 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: [1], [3], [8], [9], and [10]. 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 [6] and [7] 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 (1.4) come out as byproducts of, perhaps, wider interest. Speaking about [6], 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 Sec. 2 we present main results. Theorems 2.4 and 2.8 are generalizations of (1.4) for multidimensional case. Their proofs, based on Theorem 2.7 and the minimax theorem, are given in the same Sec. 2. Theorem 2.7 is the central result of the paper. Its proof, presented in Sec. 3, is based on some quite elementary ideas from control theory and is close to some arguments from [6]. Although, Theorem 2.7 relates to processes with "constant" coefficients it easily allows one to prove Theorem 2.9 in Sec. 4 by using a penalization method (in the spirit of [4]).

2. MAIN RESULTS

Let (Ω, \mathcal{F}, P) be a complete probability space, $\{\mathcal{F}_t; t \geq 0\}$ be an increasing filtration of σ -algebras $\mathcal{F}_t \subset \mathcal{F}$ which are complete with respect to \mathcal{F}, P . Assume that on (Ω, \mathcal{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 in control theory) with metric ρ 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 \mathcal{F}_t -admissible if it is $\mathcal{F} \otimes \mathcal{B}([0, \infty))$ -measurable with respect to (ω, t) and \mathcal{F}_t -measurable with respect to ω for each $t \geq 0$. The set of all \mathcal{F}_t -admissible processes is denoted by $\mathfrak{A} = \mathfrak{A}(\mathcal{F})$. Let \mathfrak{A}_n be the subset of \mathfrak{A} consisting of all processes α_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 α and for any $\alpha \in A$, $t \in [0, T]$, and $x, y \in \mathbb{R}^d$

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

(ii) On (Ω, \mathcal{F}, P) there exists a uniformly distributed random variable $\xi(\omega)$ independent of $\{\mathcal{F}_t\}$.

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

By Itô's theorem for any $\alpha \in \mathfrak{A}(\mathcal{G}_t)$ 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 (α, x) for any t , we define

$$\begin{aligned} |f|_0 &= \sup_{\alpha \in A} \sup_{t \in [0, T]} \sup_{x \in \mathbb{R}^d} |f(\alpha, t, x)|, \\ [f]_\delta &:= \sup_{\alpha \in A} \sup_{t \in [0, T]} \sup_{\substack{x, y \in \mathbb{R}^d \\ x \neq y}} \frac{|f(\alpha, t, x) - f(\alpha, t, y)|}{|x - y|^\delta}, \end{aligned}$$

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

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 $\alpha^{(n)} \in \mathfrak{A}_n(\mathcal{G}_t)$ (also depending on K_1, T, \dots) such that with a constant $N = N(K, d)$ the inequality

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

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

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

Remark 2.5. Condition (2.3) is automatically satisfied in two cases: if f is independent of α ; 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 (1.4) 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*

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

where the constant N_1 depends only on K and d .

This theorem is proved in Sec. 3.

Proof of Theorem 2.4. Take $n \geq 1$, $K_1 \in [0, \infty)$, and $\alpha \in \mathfrak{A}$. From Lemma 3.2.6 of [5] we know that there exists a sequence of functions $\beta(m) \in \mathfrak{A}$, each taking only finitely many values in A , such that

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

As in the proof of Lemma 3.2.7 of [5] we have

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

Obviously

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

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

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

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

$$(1+T) E \sup_{t \leq T} |x_t^\alpha - x_t^{\beta(m)}|^\delta + K_1 E \int_0^T \rho(\alpha_t, \beta_t(m)) dt \leq N(1+T)n^{-\delta/4}.$$

By fixing an appropriate m we see that we only need to find $\alpha(n) \in \mathfrak{A}_n$ such that (2.2) holds with $\beta(m)$ in place of α . 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 $\sup_{t \leq T} |x_t^\beta|$ uniform with respect to $\beta \in \mathfrak{A}(\mathcal{G})$. 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 . Obviously, F is a compact convex closed set in the space

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

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 . It is well known that the unit ball in Y with X -topology is a compact metrizable space.

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

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

Obviously,

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

Hence 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 $\text{Conv}(G)$ the least convex set containing G and by H the closure of $\text{Conv}(G)$ in X -topology of Y . By the above, H is a compact metric space. Observe also that for any $\gamma \in H$ there exist $p^{ki} \geq 0$, $i = 1, \dots, k$ such that $\sum_i p^{ki} = 1$ and $\alpha_{ki} \in \mathfrak{A}_n$ such that

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

as $k \rightarrow \infty$ 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 γ_1 and γ_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 [12])

$$\begin{aligned} &\min_{\gamma \in H} \max_{(f, g) \in F} \left[\int f(\beta, t, x) \mu^\alpha(d\beta dt dx) + \int g(x) \nu^\alpha(dx) \right. \\ &\quad \left. - \int (f + g) \gamma(d\beta dt dx dx) \right] \\ &= \max_{(f, g) \in F} \min_{\gamma \in H} \left[\int f(\beta, t, x) \mu^\alpha(d\beta dt dx) + \int g(x) \nu^\alpha(dx) \right. \\ &\quad \left. - \int (f + g) \gamma(d\beta dt dx dx) \right]. \end{aligned}$$

By (2.4), the last minimum equals

$$\begin{aligned} & \inf_{\alpha(n) \in \mathfrak{A}_n} \left[\int f(\beta, t, x) \mu^\alpha(d\beta dt dx) + \int g(x) \nu^\alpha(dx) \right. \\ & \left. - \int f(\beta, t, x) \mu^{\alpha(n)}(d\beta dt dx) - \int g(x) \nu^{\alpha(n)}(dx) \right], \end{aligned}$$

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 $(f, g) \in F$,

$$\begin{aligned} & \int f(\beta, t, x) \mu^\alpha(d\beta dt dx) + \int g(x) \nu^\alpha(dx) \\ & \leq \int (f + g) \gamma(d\beta dt dx) + N_1(1+T)n^{-\delta/4}. \end{aligned}$$

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

$$\begin{aligned} & \int f(\beta, t, x) \mu^\alpha(d\beta dt dx) + \int g(x) \nu^\alpha(dx) \\ & \leq \sum_{i=1}^k p^i \left[\int f(\beta, t, x) \mu^{\alpha_i}(d\beta dt dx) + \int g(x) \nu^{\alpha_i}(dx) \right] + 2N_1(1+T)n^{-\delta/4} \end{aligned}$$

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

$$\alpha_t(n, \omega) = \alpha_{it}(\omega) \quad \text{if} \quad \sum_{j < i} p^j \leq \xi(\omega) < \sum_{j \leq i} p^j \quad \left(\sum_{\emptyset} := 0 \right).$$

Since ξ and α_i 's are independent, it is easy to see that

$$\begin{aligned} & \sum_{i=1}^k p^i \left[\int f(\beta, t, x) \mu^{\alpha_i}(d\beta dt dx) + \int g(x) \nu^{\alpha_i}(dx) \right] \\ & = \sum_{i=1}^k p^i E \left[T^{-1} \int_0^T f(\alpha_{it}, t, x_t^{\alpha_i}) dt + g(x_T^{\alpha_i}) \right] \\ & = E \left[T^{-1} \int_0^T f(\alpha_t(n), t, x_t^{\alpha(n)}) dt + g(x_T^{\alpha(n)}) \right]. \end{aligned}$$

Thus,

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

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 α .

Theorem 2.8. *In Theorem 2.4 replace $C^{\delta/2, \delta}$ with C^δ . Then its assertion will remain to hold if we consider only the functions f independent of α 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

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

where by Fubini's theorem the last integral equals

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

where

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

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

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

Hence,

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

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 $\alpha(n) \in \mathfrak{A}(\mathcal{G}_.)$ such that

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

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

$$\left| E \left[\int_0^T f(s, x_s^\alpha) ds + g(x_T^\alpha) \right] - E \left[\int_0^T f(s, x_s^{\alpha(n)}) ds + g(x_T^{\alpha(n)}) \right] \right|$$

$$\leq N(T[f]_\delta \varepsilon^{\delta/2} + |f|_0 \varepsilon) + N(1+T)n^{-\delta/4}(|f^{(\varepsilon)}|_{\delta/2,\delta} + [g]_\delta). \quad (2.5)$$

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

$$\begin{aligned} |f^{(\varepsilon)}(s, x) - f^{(\varepsilon)}(t, x)| &\leq |t-s| \sup |D_t f^{(\varepsilon)}| \\ &\leq N|t-s|\varepsilon^{-1}|f|_0 \leq N|t-s|^{\delta/2}\varepsilon^{-\delta/2}|f|_0, \end{aligned}$$

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

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

This implies that

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

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 σ 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

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

Theorem 2.9. *Assume that for any $|s| \leq 2$*

$$\begin{aligned} &\int_0^T \sup_{\alpha \in \mathfrak{A}} \sup_{x \in \mathbb{R}^d} (|\sigma(\alpha, s+r, x) - \sigma(\alpha, r, x)| \\ &+ |b(\alpha, s+r, x) - b(\alpha, r, x)|) dr \leq K(1+T)|s|^{\delta_0/2}, \end{aligned}$$

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

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

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

We prove Theorem 2.9 in Sec. 4.

Remark 2.10. If $f \equiv c \equiv 0$, then Corollary 4.3 says that (2.6) is valid for $\delta \leq \delta_0$ as well (remember we assume $\delta_0 \leq \delta$). 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 $\gamma \leq \delta_0$.

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

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

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

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

Naturally, $p v^{\beta(n)} + (1-p)v^{\alpha(n)}$ 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 $\alpha \in \mathfrak{A}$, $n \geq 1$, and constant $K_1 \in [0, \infty)$ there exists $\alpha(n) \in \mathfrak{A}_n(\mathcal{G})$ such that, with a constant N depending only on K, d , and $(1+T)^{-1}|c|_{\delta/2, \delta}$, the inequality*

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

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 $\alpha \in \mathfrak{A}$, $n \geq 1$ there exists $\alpha(n) \in \mathfrak{A}_n(\mathcal{G})$ 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*

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

holds for all $p \geq \delta$.

We prove this corollary in Sec. 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 α . 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 $s \leq 0$ define $f(\alpha, s, x) = f(\alpha, 0, x)$ and for $s \leq T$, $x \in \mathbb{R}^d$, and $\alpha \in \mathfrak{A}$ denote

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

Obviously to prove Theorem 2.7, it suffices to prove

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

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

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

$$\begin{aligned} |v^\alpha(s, x) - v^\alpha(s, y)| &\leq |x - y|^\delta (T[f]_\delta + [g]_\delta) \\ |v_n(s, x) - v_n(s, y)| &\leq |x - y|^\delta (T[f]_\delta + [g]_\delta). \end{aligned}$$

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

$$\begin{aligned} |v^\alpha(t, x) - v^\alpha(s, x)| &\leq |t - s| |f|_0 + N|t - s|^{\delta/2} (|f|_{\delta/2} + [g]_\delta), \\ |v_n(t, x) - v_n(s, x)| &\leq |t - s| |f|_0 + N|t - s|^{\delta/2} (|f|_{\delta/2} + [g]_\delta). \end{aligned}$$

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

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

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^β defined as x_t^α for $\alpha_t \equiv \beta$.

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

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

Assertion (iii) is a particular case of Exercise 3.2.1 of [5], a solution to which can be easily obtained from Lemma 3.2.14 and the proof of Lemma 3.3.1 of [5]. 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

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

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

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

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 $t - \varepsilon^2 \leq s \leq t$, so that for $t \leq T - 1/n$ 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

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

We estimate the last expression from below. We have

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

with

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

Hence

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

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

$$\begin{aligned} v_n^{(\varepsilon)}(t, x) & \geq n^{-1} f_{n\varepsilon}(\beta, t, x) + E v_n^{(\varepsilon)}(t + 1/n, x + x_{1/n}^\beta) \\ & \quad - N[f]_\delta n^{-1}(n^{-\delta/2} + \varepsilon^\delta). \end{aligned} \quad (3.3)$$

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

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

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

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

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

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

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 $s \leq T$

$$\begin{aligned} |D_t^2 v_n^{(\varepsilon)}(s, x)| & = \varepsilon^{-6} \left| \int_{\mathbb{R}} \eta''(\varepsilon^{-2}r) (v^\varepsilon(s-r, x) - v^\varepsilon(s, x)) dr \right| \\ & \leq N K_0 \varepsilon^{-6} \int_{\mathbb{R}} |\eta''(\varepsilon^{-2}r)| |r|^{\delta/2} dr = N K_0 \varepsilon^{\delta-4}. \end{aligned}$$

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

$$\begin{aligned} |D^i v_n^{(\varepsilon)}(s, x)| & \leq \sup_{s \leq T} |D^i v_n^\varepsilon(s, x)| \\ & = \sup_{s \leq T} \varepsilon^{-i-d} \left| \int_{\mathbb{R}^d} (D^i \xi)(\varepsilon^{-1}y) [v_n(s, x-y) - v_n(s, x)] dy \right| \leq N K_0 \varepsilon^{\delta-i}, \end{aligned}$$

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

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

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

At this point we choose ε . 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

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

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

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

By combining this with the inequalities

$$|E \int_{T-1/n}^T f_{n\varepsilon}(\alpha_t, t, x_t^\alpha) dt| \leq |f|_0 n^{-1} \leq |f|_0 \varepsilon^{2-\delta} n^{-\delta/4} \leq K_0 n^{-\delta/4},$$

$$\begin{aligned} |v_n(T-1/n, x) - g(y)| &= |v_n(T-1/n, x) - v_n(T, y)| \\ &\leq n^{-1} |f|_0 + Nn^{-\delta/2} (|[f]_\delta]_{\delta/2} + [g]_\delta) + |x - y| (T[f]_\delta + [g]_\delta) \\ &\leq NK_0 n^{-\delta/4} + |x - y|^\delta (T[f]_\delta + [g]_\delta), \end{aligned}$$

$$\begin{aligned} E v_n(T-1/n, x_{T-1/n}^\alpha) &\geq E g(x_T^\alpha) - NK_0 n^{-\delta/4} \\ &\quad - (T[f]_\delta + [g]_\delta) E |x_{T-1/n}^\alpha - x_T^\alpha|^\delta \geq E g(x_T^\alpha) - NK_0 n^{-\delta/4}, \end{aligned}$$

$$\begin{aligned} |v_n^{(\varepsilon)}(t, x) - v_n(t, x)| &\leq \int_{\mathbb{R}^{d+1}} \zeta_\varepsilon(r, y) |v_n(t-r, x-y) - v_n(t, x)| dr dy \\ &\leq N \int_{\mathbb{R}^{d+1}} \zeta_\varepsilon(r, y) \{ |f|_0 r + r^{\delta/2} (|[f]_\delta]_{\delta/2} + [g]_\delta) \\ &\quad + |y|^\delta (T[f]_\delta + [g]_\delta) \} dr dy \leq N |f|_0 \varepsilon^2 + N (|[f]_\delta]_{\delta/2} + T[f]_\delta + [g]_\delta) \varepsilon^\delta \\ &\leq NK_0 n^{-\delta/4}, \end{aligned}$$

we get

$$\begin{aligned} v_n(0, 0) + NK_0 n^{-\delta/4} &\geq v_n^{(\varepsilon)}(0, 0) \\ &\geq E \int_0^T f_{n\varepsilon}(\alpha_t, t, x_t^\alpha) dt + E g(x_T^\alpha) - N(1+T)K_0 n^{-\delta/4}. \quad (3.6) \end{aligned}$$

Finally,

$$\begin{aligned} &E \int_0^T |f_{n\varepsilon}(\alpha_t, t, x_t^\alpha) - f(\alpha_t, t, x_t^\alpha)| dt \\ &\leq \int_0^T \sup_{\alpha \in \mathfrak{A}} \sup_{x \in \mathbb{R}^d} |f_{n\varepsilon}(\alpha, t, x) - f(\alpha, t, x)| dt \end{aligned}$$

$$\begin{aligned} &\leq \int_0^T \int_{-\varepsilon^2 \leq s \leq 0} \eta_{\varepsilon^2}(s) n \int_0^{1/n} \sup_{\alpha \in A} \sup_{x \in \mathbb{R}^d} |f(\alpha, t-s+r, x) - f(\alpha, t, x)| dr ds dt \\ &\leq \|f\|_{\delta/2} (\varepsilon^\delta + n^{-\delta/2}) \leq K_0 n^{-\delta/4}, \end{aligned}$$

which along with (3.6) easily yield (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*

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

hold for any $x, y \in \mathbb{R}^d$, $t \in [0, T]$, and $|s| \leq 2$, 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*

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

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

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

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

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

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 $r \leq 0$, $\sigma(\alpha, r, x) = \sigma(\alpha, T, x)$ and $b(\alpha, r, x) = b(\alpha, T, x)$ for $r \geq T$, and define

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

Every element $\tilde{\alpha}$ of the set \tilde{A} is a couple (σ, 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

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

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

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

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

Hence,

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

where $|s| \leq 2$.

Next, let $\tilde{\mathfrak{A}}$ be the set of all \tilde{A} -valued \mathcal{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

$$\begin{aligned} \tilde{x}_t^{\tilde{\alpha}} &= \int_0^t \sigma(\tilde{\alpha}_r) dw_r + \int_0^t b(\tilde{\alpha}_r) dr, \\ v &= \sup_{\alpha \in \tilde{\mathfrak{A}}} E g(x_T^\alpha), \quad \tilde{v} = \sup_{\tilde{\alpha} \in \tilde{\mathfrak{A}}} E [g(\tilde{x}_T^{\tilde{\alpha}}) - \int_0^T \gamma(\tilde{\alpha}, t, \tilde{x}_t^{\tilde{\alpha}}) dt], \\ v_n &= \sup_{\alpha \in \tilde{\mathfrak{A}}_n} E g(x_T^\alpha), \quad \tilde{v}_n = \sup_{\tilde{\alpha} \in \tilde{\mathfrak{A}}_n} E [g(\tilde{x}_T^{\tilde{\alpha}}) - \int_0^T \gamma(\tilde{\alpha}, t, \tilde{x}_t^{\tilde{\alpha}}) dt]. \end{aligned}$$

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

$$v \leq v_n + N e^{NT} |g|_{\delta_3} n^{-\delta_1 \delta_2 / 8}, \quad (4.5)$$

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

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

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

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

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

$$v \leq v_n + N e^{NT} (n^{-\delta_1/4} c^{-1} + n^{-\delta_1/4} |g|_{\delta_3} + |g|_{\delta_3}^{2/(2-\delta_2)} c^{\delta_2/(2-\delta_2)} + |g|_{\delta_3} n^{-\delta_1 \delta_2 / 4}),$$

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

$$v \leq v_n + N e^{NT} |g|_{\delta_3} (n^{-\delta_1 \delta_2 / 8} + n^{-\delta_1/4} + n^{-\delta_1 \delta_2 / 4}),$$

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

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

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

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

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

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

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

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

Moreover, since $\tilde{\alpha}_t$ is piece-wise constant, the appropriate α 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

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

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

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

Hence,

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

This, (4.4), and (4.8) yield

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

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

$$\sigma_t := \sigma(\tilde{\alpha}_r) - \sigma(\alpha_r, r, \tilde{x}_r^{\tilde{\alpha}}), \quad b_t := b(\tilde{\alpha}_r) - b(\alpha_r, r, \tilde{x}_r^{\tilde{\alpha}}).$$

It follows from (4.10) and (4.2) that

$$Eg(\tilde{x}_T^\alpha) - Eg(x_T^\alpha) \leq Ne^{NT}|g|_{\delta_3}(J + \varepsilon + n^{-\delta_1/2})^{\delta_2/2}. \quad (4.11)$$

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)) $\tilde{v}_n \leq Eg(\tilde{x}_T^\alpha) + \varepsilon$, $Eg(x_T^\alpha) \leq v_n$, and $J \leq 2|g|_0$, so that (4.11) implies that

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

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 $Eg(x_T^\alpha) \leq v_n \leq \tilde{v}_n$, so that by definition (4.7) of J , equation (4.11) implies that

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

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

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

In the remaining case and generally

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

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

$$\tilde{v}_n \leq Eg(\tilde{x}_T^\alpha) + \varepsilon \leq Eg(x_T^\alpha) + Ne^{NT}(|g|_{\delta_3}^{2/(2-\delta_2)} + |g|_{\delta_3}\varepsilon^{\delta_2/2} + |g|_{\delta_3}n^{-\delta_1\delta_2/4}) + \varepsilon$$

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,

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

and as well known (apply, for instance, Corollary 2.5.10 of [5] to the process $x_t - y_t$ due to (2.1),

$$E \sup_{t \leq T} |x_t - y_t|^2 \leq Ne^{NT} E \int_0^T \{|\sigma_t|^2 + |b_t|^2\} dt, \quad (4.12)$$

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 $\delta_0 \leq \delta$, our assertion follows from Remark 4.2. For $\delta_0 \geq \delta$, 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

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

$$|g_{x_i}^{(\varepsilon)}(x)| = \varepsilon^{-1-d} \int (g(y) - g(x)) \zeta_{x_i}((y-x)/\varepsilon) dy \leq N[g]_\delta \varepsilon^{\delta-1},$$

so that

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

if $|x - y| \leq \varepsilon$. Furthermore, if $|x - y| \geq \varepsilon$,

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

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

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

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

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

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^α . Therefore, without loss of generality we assume that $c \geq 0$.

First, assume that in addition

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

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

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

where $\xi \in C_0^\infty(\mathbb{R})$, $\xi(y) = e^y$ for $y \leq 0$, $0 \leq \xi \leq 2$, and $|\xi'| \leq 2$. 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

$$\begin{aligned} \bar{g}(x_T^\alpha, y_T^\alpha, z_T^\alpha) &= (-M) \vee \left[\int_0^T f(\alpha_t, t, x_t^\alpha) e^{-\varphi_t^\alpha} dt + g(x_T^\alpha) e^{-\varphi_T^\alpha} \right] \wedge M \\ &= \int_0^T f(\alpha_t, t, x_t^\alpha) e^{-\varphi_t^\alpha} dt + g(x_T^\alpha) e^{-\varphi_T^\alpha}, \\ v^\alpha &= E\bar{g}(x_T^\alpha, y_T^\alpha, z_T^\alpha). \end{aligned} \quad (4.14)$$

Now it is natural to apply Lemma 4.1 to the process $(x_t^\alpha, y_t^\alpha, z_t^\alpha)$.

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

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

Denote

$$I = E \int_0^T \sum_{i=1}^3 \{|\sigma_t^{(i)}|^2 + |b_t^{(i)}|^2\} dt.$$

We get from (4.12) that

$$E \sup_{t \leq T} |x_t - x'_t|^2 \leq Ne^{NT} E \int_0^T \{ \|\sigma_t^{(1)}\|^2 + |b_t^{(1)}|^2 \} dt \leq Ne^{NT} I.$$

Furthermore,

$$\begin{aligned} E \sup_{t \leq T} |y_t - y'_t|^2 &\leq NT^2 [c]_\delta^2 E \sup_{t \leq T} |x_t - x'_t|^{2\delta} + NE \int_0^T \|\sigma_t^{(2)}\|^2 dt \\ &\quad + NTE \int_0^T |b_t^{(2)}|^2 dt \leq Ne^{NT} (I^\delta + I), \\ E |z_T - z'_T|^2 &\leq NT^2 ([f]_\delta^2 E \sup_{t \leq T} |x_t - x'_t|^{2\delta} + |f|_0^2 E \sup_{t \leq T} |y_t - y'_t|^2) \\ &\quad + NE \int_0^T \|\sigma_t^{(3)}\|^2 dt + NTE \int_0^T |b_t^{(3)}|^2 dt \leq Ne^{NT} (I^\delta + I). \end{aligned}$$

By using Hölder's inequality, we obtain

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

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

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

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

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

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

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

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

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

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 $\neq 0$ instead of f and pass to the limit. This finishes 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 $\alpha \in \mathfrak{A}$, $n \geq 1$, and $g(x, y, z) \in C^\delta$ there is an $\alpha(n) \in \mathfrak{A}_n$ such that

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

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 $n \geq 1$ there is an $\alpha(n) \in \mathfrak{A}_n(\mathcal{G})$ such that

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

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

REFERENCES

- [1] Barles, G. and Souganidis, P.E. (1991). Convergence of approximation schemes for fully nonlinear second order equations, *Asymp. Anal.* **4** 271-283.
- [2] Fabes, E.B. and Kenig, C. (1981). Examples of singular parabolic measures and singular transition probability densities, *Duke Math. Journal* **48** 845-856.
- [3] Fleming, W.H. and Soner, H.M. (1993). *Controlled Markov Processes and Viscosity Solutions, Applications in Mathematics*, Vol. 25, Springer-Verlag, New York etc.,
- [4] Krylov, N.V. (1973). On the selection of a Markov process from a system of processes and the construction of quasi-diffusion processes, *Izvestija Akademii Nauk SSSR, serija matematicheskaja* **37** 691-708 in Russian; English translation: *Math. USSR Izvestija* **7** 691-709.
- [5] Krylov, N.V. (1977). *Controlled diffusion processes*, Nauka, Moscow, in Russian; English translation: Springer, 1980.
- [6] Krylov, N.V. (1997). On the rate of convergence of finite-difference approximations for Bellman's equations, *Algebra i Analiz* **9** 245-256.
- [7] Krylov, N.V. Fully nonlinear second order elliptic equations: Recent development, to appear in *Annali Scuola Normale Superiore di Pisa*.
- [8] Kuo, H.J. and Trudinger, N.S. (1992). Discrete methods for fully nonlinear elliptic equations. *SIAM J. on Numer. Anal.* **29** 123-135.
- [9] Kushner, H.J. and Dupuis, P.G. (1992). *Numerical methods for stochastic control problems in continuous time*, Springer Verlag.
- [10] 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.* **235**, No. 1 (1983), 98-108.
- [11] Safonov, M.V. (1981). An example of diffusion process with singular distribution at some given time, *Abstr. Comm. Third Vilnius Conference on Probability Theory and Math. Statistics*. June 22-27, Vilnius, pp. 133-134 (in Russian).
- [12] Szép, J. and Forgó, F. (1983). *Introduction to the theory of games*, Reidel, Dordrecht/Boston/Lancaster
- [13] Triebel, H. (1992). *Theory of function spaces II*. Birkhäuser Verlag, Basel-Boston-Berlin.

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