Probab. Theory Relat. Fields 105, 143–158 (1996)

Probability Theory and Related Fields

Existence of strong solutions for Itô's stochastic equations via approximations

István Gyöngy[1,★], **Nicolai Krylov**[2,★★]

[1] Department of Probability and Statistics, Eötvös Loránd University Budapest, Múzeum krt. 6-8, H-1088 Budapest, Hungary
[2] School of Mathematics, University of Minnesota, Minneapolis, MN 55455, USA

Received: 6 January 1994/In revised form: 7 November 1995

Summary. Given strong uniqueness for an Itô's stochastic equation with discontinuous coefficients, we prove that its solution can be constructed on "any" probability space by using, for example, Euler's polygonal approximations. Stochastic equations in $\mathbb{R}^d$ and in domains in $\mathbb{R}^d$ are considered.

Mathematics Subject Classification (1991): 60H10

1. Introduction

We start with an example illustrating the results we present in the paper. Consider the stochastic differential equation

$$
\begin{aligned}
dx(t) &= \left[\tan\left(-\frac{\pi}{2}x(t)\right) + \operatorname{sign} x(t)\right] dt + |1 - |x(t)||^{\alpha}\, dw(t)\,, \\
x(0) &= 0
\end{aligned}
\tag{1.1}
$$

with a given $\alpha > 0$, and a Wiener process w. Note that the drift coefficient is not continuous at $x = 0$, $x = 2k + 1$, for integers k, and it does not satisfy the linear growth condition. Moreover the diffusion coefficient does not satisfy the linear growth condition for $\alpha > \frac{1}{2}$ and it is not Hölder continuous with exponent $1/2$ if $\alpha < \frac{1}{2}$.

The coefficients in the above equations are rather irregular. However, one can define Euler's "polygonal" approximations:

$$
\begin{aligned}
dx_n(t) &= b(x_n(\kappa_n(t)))\, dt + \sigma(x_n(\kappa_n(t)))\, dw(t)\,, \\
x_n(0) &= 0
\end{aligned}
\tag{1.2}
$$

★ Research supported by the Hungarian National Foundation of Scientific Research No. 2990.
★★ Supported in part by NSF Grant DMS-9302516.

for every integer $n > 0$, where $\kappa_n(t) := [nt]/n$, with the corresponding drift and diffusion coefficient, setting for example $b(x)=0$ when x is an odd integer. One expects that x_n converges in probability to a process which solves equation (1.1).

The existence of a local strong solution to equation (1.1) can be seen from the following result of Veretennikov [15]. Let b, σ be bounded measurable functions on $\mathbb{R}_+ \times \mathbb{R}^d$ with values in $\mathbb{R}^d$ and in $\mathbb{R}^{d \times d_1}$, respectively, such that $\sigma\sigma^{\mathrm{T}}$ is uniformly elliptic, σ is Hölder continuous in $x \in \mathbb{R}$ with exponent $\frac{1}{2}$ when $d = 1$, and it is Lipschitz in $x \in \mathbb{R}^d$ in the multidimensional case. Then there exists a unique strong solution to the stochastic differential equation

$$dx(t) = b(t,x(t))\,dt + \sigma(t,x(t))\,dw(t), \quad x_0 \in \mathbb{R}^d \tag{1.3}$$

The method of establishing this existence and uniqueness theorem is rather different from those used in the theory of ordinary differential equations. It is based on a famous result from Yamada and Watanabe [16] stating that the existence of a solution (on some probability space with some Wiener process) and the pathwise uniqueness imply the existence of a strong solution. (See also [17] and the references therein on this topic.) We emphasize that the proof of this result involves no construction of the solution.

The existence of a solution to equation (1.3) with bounded measurable coefficients is known under the additional condition that either $\sigma(t,x)$ and $b(t,x)$ are continuous in x (Skorokhod [14], Stroock and Varadhan [13]), or $\sigma\sigma^{\mathrm{T}}$ is uniformly elliptic (Krylov [6, 9]). Hence Veretennikov [15] establishes the existence of a strong solution by proving the pathwise uniqueness. His proof raises the following questions. Is it possible to construct the strong solutions in some classical way? Define, for example, Euler's approximations (1.2) to equation (1.3). Do these approximations converge to a stochastic process in probability and can one construct a strong solution in this way? Suppose the coefficients in the equation (1.3) are approximated by smooth ones. Do the strong solutions of the corresponding equations converge in probability to the strong solution of equation (1.3) under the assumptions of the cited existence theorem? More generally, does the strong solution depend continuously, in the topology of convergence in probability, on the initial condition and on the drift and diffusion coefficients? Our aim is to show that the answers to these questions are in the affirmative. We prove that, roughly speaking, the Euler's polygonal approximations converge uniformly in t in bounded intervals, in probability, to a process, which we show to be the strong solution. The basic assumption is that the pathwise uniqueness for the equation holds. In particular, applying Corollary 2.9 to equation (1.1) with $D := (-1, 1)$, $D_k := (-1 + 2^{-k}, 1 - 2^{-k})$ and with $V(t,x) := (2 - x^2)/(1 - x^2)$, we get that Euler's approximations $x_n(t)$, defined by (1.2) converge in probability, uniformly in t in bounded intervals to some stochastic process, which is the strong solution of equation (1.1).

The possibility of showing the convergence in probability of different approximations to solutions of stochastic equations is based on the following simple observation.

Lemma 1.1. *Let Z_n be a sequence of random elements in a Polish space $(\mathbb{E}, \rho)$ equipped with the Borel σ-algebra. Then Z_n converges in probability to an $\mathbb{E}$-valued random element if and only if for every pair of subsequences*

Z_l and Z_m there exists a subsequence $v_k := (Z_{l(k)}, Z_{m(k)})$ converging weakly to a random element v supported on the diagonal $\{(x, y) \in \mathbb{E} \times \mathbb{E}: x = y\}$.

To prove the sufficiency of the above condition for the convergence in probability it is enough to note that for the continuous function $f(x, y) = \rho(x, y)$ the random variables $f(v_k)$ converge to $f(v) = 0$ weakly and hence, $f(v_k) \to 0$ *in probability*. This implies that $\{Z_n\}$ is a Cauchy sequence in the space of random $\mathbb{E}$-valued elements with the metric corresponding to convergence in probability. Since this space is complete, our assertion holds indeed. The necessity of our condition is obvious.

In our applications of the lemma Skorokhod's embedding method and the assumption of pathwise uniqueness will allow us to check that the limiting random element v takes values in $\{(x, y) \in \mathbb{E} \times \mathbb{E}: x = y\}$.

We note that our approach is very close in spirit to the celebrated result of Yamada and Watanabe on the existence of strong solutions via pathwise uniqueness. We assume somewhat more and in return we can get more. From our approach it is clear that the strong solution depends continuously on the initial condition and on the drift and diffusion coefficient. In particular, in parallel with the proofs of Theorems 2.4 and 2.8 the strong solution can be constructed by smooth approximation of the coefficients. One can construct the strong solution by Euler's approximations and simultaneously approximating the coefficients and the initial condition. Clearly, we immediately get the convergence of Euler's approximations (or of the other approximations we mentioned) in probability in every metric space V, in which these approximations are tight. (See [3], were the convergence in probability of Wong–Zakai type approximations are proved in suitable Banach spaces.)

We also note that the convergence of Euler's approximations under various conditions is proved by many authors. It is shown in Krylov [8] that under the monotonicity condition Euler's polygonal line method can be adjusted to prove (strong) solvability (for equations even with random coefficients). Earlier this was known from Maruyama [10] if the drift and diffusion coefficients are Lipschitz continuous. The method of [8] was afterward used in Alyushina [1] in a short proof of existence of strong solutions under monotonicity and linear growth. A short and simple proof of (strong) solvability is presented in Krylov [7] under monotonicity and under a condition which is weaker than the usual linear growth. Moreover, the continuous dependence of the strong solution on the coefficients is obtained.

It is worth mentioning that the fact that the pathwise uniqueness implies the possibility of effective constructing the solutions has already been noticed in Zvonkin and Krylov [17] (see, for instance, Lemma 3.2 there). Later Kaneko and Nakao [5] exploited this fact without noticing [17]. In [5] the authors consider equation (1.3) in $\mathbb{R}^d$ and they assume that it admits a unique strong solution $x(t)$. They show that $x(t)$ can be constructed by approximating the coefficients and also by Euler's polygonal approximation. In what concerns Euler's approximations they only consider equations in the whole space with continuous coefficients satisfying the linear growth condition. We consider equations also in domains of $\mathbb{R}^d$ and with discontinuous coefficients as well. We construct the strong solution without assuming its existence. Our basic idea of proving convergence in probability is an extension of the idea of an-other result of Yamada and Watanabe saying that pathwise uniqueness implies uniqueness in law. Essentially the same idea is used in [5]. Due to our above lemma this idea becomes more apparent and its range of applicability becomes evident.

The paper is organized as follows. In the next section we formulate our results in Theorems 2.4, 2.8 and their corollaries. By Lemma 1.1 the proof of Theorem 2.4 is simple, we present it in Sect. 3. To prove Theorem 2.8 we need an estimate of the distribution for Euler's approximations. Since such estimates play an important role not only in the subject of the paper, we present our estimate (Theorem 4.2 below) separately in Sect. 4. We prove the main result, Theorem 2.8, in the last section.

2. Formulation of the results

On a given stochastic basis $(\Omega, \mathscr{F}, P, (\mathscr{F}_t)_{t \geqq 0})$ we consider the stochastic differential equation

$$\begin{aligned} dx(t) &= b(t,x(t))\,dt + \sigma(t,x(t))\,dw(t)\,, \\ x(0) &= \xi \end{aligned} \tag{2.1}$$

in a domain D of $\mathbb{R}^d$, where $(w(t), \mathscr{F}_t)$ is a d_1-dimensional Wiener process, ξ is an $\mathscr{F}_0$-measurable random vector with values in D, b and σ are Borel functions on $\mathbb{R}_+ \times D$ taking values in $\mathbb{R}^d$ and in $\mathbb{R}^{d \times d_1}$, respectively. For equation (2.1) to have sense we need the coefficients to be defined for any $x \in \mathbb{R}^d$. Actually under our future assumptions solutions of (2.1) will never leave D so the values of σ and b outside D are irrelevant and just for convenience we define $\sigma(t,x) = 0$, $b(t,x) = 0$ for $x \notin D$, $t \geqq 0$. Let

$$0 = t_0^n < t_1^n < t_2^n < \cdots < t_i^n < t_{i+1}^n < \cdots$$

be a sequence of partitions of $\mathbb{R}_+$ such that for every $T > 0$

$$d_n(T) := \sup_{i:\, t_{i+1} \leqq T} |t_{i+1}^n - t_i^n| \to 0$$

as $n \to \infty$. We define Euler's "polygonal" approximations as the process $(x_n(t))$ satisfying

$$\begin{aligned} dx_n(t) &= b(t,x_n(\kappa_n(t)))\,dt + \sigma(t,x_n(\kappa_n(t)))\,dw(t)\,, \\ x_n(0) &= \xi \end{aligned} \tag{2.2}$$

where $\kappa_n(t) := t_i^n$ for $t \in [t_i^n, t_{i+1}^n)$.

In the whole article $M(t) > 0$ and $M_1(t) > 0$, $M_2(t) > 0, \ldots$ are fixed locally integrable functions on $[0, \infty)$. We will use the following assumptions:

(i) there exists an increasing sequence of bounded domains $\{D_k\}_{k=1}^{\infty}$ such that $\bigcup_{k=1}^{\infty} D_k = D$, and for every k, $t \in [0,k]$

$$\sup_{x \in D_k} |b(t,x)| \leqq M_k(t), \qquad \sup_{x \in D_k} |\sigma(t,x)|^2 \leqq M_k(t)\,;$$

(ii) there exists a non-negative function $V \in C^{1,2}(\mathbb{R}_+ \times D)$ such that

$$LV(t,x) \leqq M(t)V(t,x), \quad \forall t \geqq 0,\ x \in D\,,$$

$$V_k(T) := \inf_{x \in \partial D_k,\, t \leqq T} |V(t,x)| \to \infty$$

as $k \to \infty$ for every finite T, where ∂D_k denotes the boundary of D_k and L is the differential operator

$$L := \frac{\partial}{\partial t} + \sum_i b_i(t,x)\frac{\partial}{\partial x^i} + \frac{1}{2}\sum_{i,j}(\sigma\sigma^{\mathrm{T}})_{ij}(t,x)\frac{\partial^2}{\partial x^i \partial x^j}\,;$$

(iii) $P(\xi \in D) = 1$.

Note that by (i) and by our definition of σ and b outside D, Euler's approximations $x_n(t)$ are well defined for all $t \geqq 0$.

Definition 2.1. *By solution of equation* (2.1) *we mean an $\mathscr{F}_t$-adapted process $x(t)$ which does not ever leave D and satisfies* (2.1).

An explanation of the definition can be found in the following statement.

Lemma 2.2. *Let $x(t)$ be an $\mathscr{F}_t$-adapted process defined for all $t \geqq 0$. Assume that $x(t)$ satisfies* (2.1) *for $t < \tau := \inf\{t\colon x(t) \notin D\}$, and assume* (i) *through* (iii). *Then $\tau = \infty$ (a.s.).*

Proof. Define τ^k as the first exit time of $x(t)$ from D_k. Obviously $\tau^k \uparrow \tau$. Therefore to prove the lemma it suffices to show that for any k and δ, $T > 0$ we have

(2.3)
$$P(\tau^k \leqq T) \leqq P(\xi \notin D_k) + P\left(V(0,\xi) \geqq \log\frac{1}{\delta}\right) + \frac{1}{\delta V_k(T)}\exp\int_0^T M(t)\,dt\,.$$

Apply Itô's formula to $\gamma(t)V(t,x(t))$ where

$$\gamma(t) := \exp\left[-\int_0^t M(s)\,ds - V(0,\xi)\right],$$

and use assumption (ii). Then it follows that for all t

$$\gamma(t)V(t \wedge \tau^k, x(t \wedge \tau^k))\chi_{\tau^k > 0} \leqq \gamma(0)V(0,\xi) + m^k(t)\,,$$

where $m^k(t)$ is a continuous local martingale starting from 0. Hence for any $R > 0$

$$P\left\{\sup_{t \leqq \tau^k} \gamma(t)V(t,x^k(t))\chi_{\tau^k > 0} \geqq R\right\} \leqq \frac{1}{R}E(\gamma(0)V(0,\xi)) \leqq \frac{1}{R}\,,$$

and this gives (2.3) almost immediately. The lemma is proved.

In order to state our main results we need one more notion.

Definition 2.3. *We say that the pathwise uniqueness holds for equation* (2.1) *if for any stochastic basis carrying a d_1-dimensional Wiener process $w'(\cdot)$ and a random variable ξ' such that the joint distribution of $(w'(\cdot),\xi')$ is the same as that of the given $(w(\cdot),\xi)$, equation* (2.1) *with $w'(t),\xi'$ instead of $w(t),\xi$ cannot have more than one solution.*

Theorem 2.4. (cf. [5]) *Assume* (i) *through* (iii). *Suppose moreover that b and σ are continuous in $x \in D$ and that for equation* (2.1) *the pathwise uniqueness holds. Then $x_n(t)$ converges in probability to a process $x(t)$, uniformly in t in bounded intervals, and $x(t)$ is the unique solution of equation* (2.1). *Furthermore, $x(t)$ is $\mathscr{F}_t^w \vee \sigma(\xi)$-adapted.*

Remark. 2.5. Note that taking $V(t,x) := (|x|^2+1)\exp(-\int_0^t M(s)\,ds)$ in the case $D = \mathbb{R}^d$, $D_k := \{x \in \mathbb{R}^d\colon\ |x| < k\}$, conditions (i) and (ii) can be restated as follows:

(1) $\sup_{|x|<k}\{|b(t,x)| + |\sigma(t,x)|^2\} \leqq M_k(t)$ for every $t \geqq 0$ and positive integer k;
(2) $2xb(t,x) + \|\sigma(t,x)\|^2 \leqq M(t)(|x|^2+1)$ for every $t \geqq 0$ and $x \in \mathbb{R}^d$,

where $\|\alpha\|$ denotes the Hilbert–Schmidt norm for matrices α.

We say that the coefficients b, σ satisfy the monotonicity condition on D if for every k and $t \geqq 0$, $x,y \in D_k$ we have

$$2(x-y)(b(t,x)-b(t,y)) + \|\sigma(t,x)-\sigma(t,y)\|^2 \leqq M_k(t)|x-y|^2 .$$

Corollary 2.6. (cf. [7]) *Assume* (i) *through* (iii) *and let the coefficients b, σ satisfy the monotonicity condition on D. Or in case $D = \mathbb{R}^d$ we may assume that the conditions* (1) *and* (2) *from Remark* 2.5 *are satisfied and that the monotonicity condition is satisfied for $D_k = \{x \in \mathbb{R}^d\colon\ |x| < k\}$. Assume moreover that b is continuous in $x \in D$. Then the conclusions of Theorem* 2.4 *hold.*

Proof. One can easily show that the monotonicity condition implies the pathwise uniqueness (see e.g. Krylov [8]). Hence this corollary is immediate from Theorem 2.4.

In the one-dimensional case (i.e. when $d = 1$) we have the following result.

Corollary 2.7. *Let $d = 1$. Assume* (i) *through* (iii) *and let b be continuous in x in D for any t. Assume moreover that for every k and $t \geqq 0$, $x,y \in D_k$ we have*

$$(x-y)(b(t,x)-b(t,y)) \leqq M_k(t)|x-y|^2$$

$$|\sigma(t,x)-\sigma(t,y)|^2 \leqq M_k(t)\rho_k(|x-y|)\,,$$

where ρ_k is an increasing non negative function on $\mathbb{R}_+$ such that

$$\int_0^1 1/\rho_k(r)\,dr = \infty\,.$$

Then the conclusions of Theorem 2.4 *hold.*

Proof. For any given $k = 1,2,\ldots$ we can make a nonrandom time change which reduces the general case to the case $M_k \equiv 1$. In this case one can see by a straightforward modification of the well-known method from Yamada and Watanabe [16] (see also [4]) that the above conditions imply the pathwise uniqueness for solutions of equation (2.1) until they leave D_k. Of course, after this we see that even without time change we have the pathwise uniqueness for solutions until they leave D_k. Since this is true for any k we have the pathwise uniqueness in D, and this is the only thing we need to apply Theorem 2.4.

If we are dealing with nondegenerate equations, the continuity condition on b in Theorem 2.4 can be dropped. To state this more precisely, in addition to the conditions (i) through (iii) let us introduce the following non-degeneracy condition on the diffusion coefficient σ:

(iv) For every k the domain D_k is bounded and convex, and

$$\sum_{i,j}(\sigma\sigma^{\mathrm{T}})_{ij}(t,x)\lambda^i\lambda^j \geqq \varepsilon_k M_k(t)\sum_i |\lambda^i|^2$$

for every $t \in [0,k]$, $x \in D_k$, $\lambda^i \in \mathbb{R}$, where $\varepsilon_k > 0$ are some constants.

We say that a function f on $\mathbb{R}_+ \times D$ is locally Hölder in x in D (with exponent $\alpha \in (0,1]$) if for every k and $t \geqq 0$, $x, y \in D_k$

$$|f(t,x) - f(t,y)|^2 \leqq M_k(t)|x-y|^{2\alpha} .$$

If $\alpha = 1$, then we say that f is locally Lipschitz in x in D.

Theorem 2.8. *Assume* (i) *through* (iv) *and suppose that* σ *is locally Hölder in* x *in* D *with some exponent* $\alpha \in (0,1]$. *In the case* $\alpha \neq 1$ *assume in addition that the pathwise uniqueness holds for equation* (2.1). *Then Euler's approximations* $x_n(t)$ *converge to a process* $x(t)$ *in probability, uniformly in* t *in bounded intervals, and* $x(t)$ *is the unique solution of equation* (2.1). *Furthermore,* $x(t)$ *is* $\mathscr{F}_t^w \vee \sigma(\xi)$*-adapted.*

In the one-dimensional case one can state a condition on pathwise uniqueness differently.

Corollary 2.9. *Let* $d = 1$ *and assume* (i) *through* (iv). *Suppose that* σ *is locally Hölder in* x *in* D *with some exponent* $\alpha \in (0,1]$. *Assume moreover that for every* k

$$|\sigma(t,x) - \sigma(t,y)|^2 \leqq M_k(t)(\rho_k(|x-y|) + |v_k(x) - v_k(y)|)$$

for every $t \geqq 0$, $x, y \in D_k$, *where* v_k *is a real function of locally bounded variation and* ρ_k *is an increasing continuous function satisfying*

$$\int_0^1 1/(r \vee \rho_k(r))\, dr = \infty .$$

Then the conclusions of Theorem 2.8 *hold.*

Proof. Using the result obtained in Veretennikov [15] on pathwise uniqueness for stochastic Itô's equations in one dimension (which generalizes the corresponding results in Yamada and Watanabe [16] and in Nakao [11]), we can repeat the argument from the proof of Corollary 2.7.

3. The Proof of Theorem 2.4

For every positive integers k, n define the stopping time

$$\tau_n^k := \inf\{t \geqq 0 \colon x_n(t) \notin D_k\} .$$

Then

$$|b(t,x_n(\kappa_n(t)))| \leqq M_k(t), \qquad |\sigma(t,x_n(\kappa_n(t)))|^2 \leqq M_k(t)$$

for $t \leqq \tau_n^k$, and clearly the family of stochastic processes $\{x_n^k : n = 1,2,\ldots\}$ defined by

$$x_n^k(t) := x_n(t \wedge \tau_n^k) ,$$

is weakly compact in $C([0,T])$ for every k and $T \geqq 0$. We want to deduce from this the weak compactness in $C([0,T])$ of

$$\{(x_n(t))_{t\in[0,T]} : n = 1,2,\ldots\} . \tag{3.1}$$

Clearly it suffices to show that

$$\lim_{k\to\infty} \limsup_{n\to\infty} P(\tau_n^k \leqq T) = 0 . \tag{3.2}$$

At first fix k and apply Skorokhod's embedding theorem. Then by virtue of the weak compactness of distributions of $x_n^k(t)$ in $C([0,T])$ for every $T \geqq 0$, we can find a subsequence $n(j)$ and a probability space $(\tilde\Omega,\tilde{\mathscr{F}},\tilde P)$, carrying the sequences of continuous processes $\tilde x_{n(j)}^k$, $\tilde w_j$, such that for every positive integer j finite dimensional distributions of

$$(\tilde x_{n(j)}^k, \tilde w_j) \quad \text{and} \quad (x_{n(j)}^k, w)$$

coincide, and for any $T < \infty$ for $\tilde P$-almost every $\tilde\omega \in \tilde\Omega$

$$\sup_{t\leqq T} |\tilde x_{n(j)}^k(t) - \tilde x^k(t)| \to 0, \qquad \sup_{t\leqq T} |\tilde w_j(t) - \tilde w(t)| \to 0 , \tag{3.3}$$

as $j \to \infty$, where $\tilde x$, $\tilde w$ are some stochastic processes. Define $\tilde\tau_{n(j)}^k, \tilde\tau^k$ as the first exit times from D_k of the processes $\tilde x_{n(j)}^k$, $\tilde x^k$, respectively. It follows from (3.3) that

$$\liminf_{j\to\infty} \tilde\tau_{n(j)}^k \geqq \tilde\tau^k \quad \text{(a.s.)} . \tag{3.4}$$

Next define

$$\tilde{\mathscr{F}}_t^j := \sigma(\tilde x_{n(j)}^k(s), \tilde w_j(s) : s \leqq t), \qquad \tilde{\mathscr{F}}_t := \sigma(\tilde x^k(s), \tilde w(s) : s \leqq t) .$$

Then it is easy to see that for every j the process $(\tilde w_j(t), \tilde{\mathscr{F}}_t^j)$ and $(\tilde w(t), \tilde{\mathscr{F}}_t)$ are Wiener processes, and for all $t \in [0, \tilde\tau_{n(j)}^k)$

(3.5)

$$\tilde x_{n(j)}^k(t) = \tilde x_{n(j)}^k(0) + \int_0^t b(s, \tilde x_{n(j)}^k(\kappa_{n(j)}(s)))\, ds + \int_0^t \sigma(s, \tilde x_{n(j)}^k(\kappa_{n(j)}(s)))\, d\tilde w_j(s) ,$$

almost surely. Now we make use of the following lemma which is just an adaptation of a result of Skorokhod [14].

Lemma 3.1. *Let $f(s,x)$ be continuous in x and Borel in t bounded function defined on $\mathbb{R}_+ \times \mathbb{R}^d$. Then for any $i = 1,\dots,d_1$*

$$
(3.6) \qquad
\begin{aligned}
\int_0^t f(s,\tilde{x}^k_{n(j)}(s))\,ds &\to \int_0^t f(s,\tilde{x}^k(s))\,ds\,,\\
\int_0^t f(s,\tilde{x}^k_{n(j)}(\kappa_{n(j)}(s)))\,ds &\to \int_0^t f(s,\tilde{x}^k(s))\,ds\,,\\
\int_0^t f(s,\tilde{x}^k_{n(j)}(s))\,d\tilde{w}^i_j(s) &\to \int_0^t f(s,\tilde{x}^k(s))\,d\tilde{w}^i(s)\,,\\
\int_0^t f(s,\tilde{x}^k_{n(j)}(\kappa_{n(j)}(s)))\,d\tilde{w}^i_j(s) &\to \int_0^t f(s,\tilde{x}^k(s))\,d\tilde{w}^i(s)
\end{aligned}
$$

uniformly in $t \in [0,T]$ in probability for any $T < \infty$.

Owing to (3.4) and (3.6) we then conclude that for $t < \tilde{\tau}^k$ (a.s.)

$$\tilde{x}^k(t) = \tilde{x}^k(0) + \int_0^t b(s,\tilde{x}^k(s))\,ds + \int_0^t \sigma(s,\tilde{x}^k(s))\,d\tilde{w}(s)\,.$$

In the proof of estimate (2.3) we have used only that $x(t)$ satisfies equation (2.1) until it hits ∂D_k. Therefore estimate (2.3) holds for our $\tilde{\tau}^k$, and since τ^k_n have the same distributions as $\tilde{\tau}^k_n$,

$$
\begin{aligned}
\lim_{k\to\infty}\limsup_{j\to\infty} P(\tau^k_{n(j)} \leqq T) &= \lim_{k\to\infty}\limsup_{k\to\infty} P(\tilde{\tau}^k_{n(j)} \leqq T)\\
&\leqq \lim_{k\to\infty} P(\tilde{\tau}^k \leqq T) = 0\,.
\end{aligned}
$$

Arbitrariness in the choice of the subsequence $n(j)$ allows us to assert that (3.2) holds, and thus the family (3.1) is indeed weakly compact. On our way of applying Lemma 1.1 we now take two subsequences x_l, x_m of the approximations $\{x_n\}_{n=1}^{\infty}$. Then obviously $\{(x_l,x_m)\}$ is a tight family of processes in $C([0,T];\mathbb{R}^{2d})$ for any $T<\infty$. Again by Skorokhod's embedding theorem there exist subsequences $l(j), m(j)$, a probability space $(\hat{\Omega}, \hat{\mathscr{F}}, \hat{P})$, carrying sequences of continuous processes $\hat{x}_{l(j)}$, $\bar{x}_{m(j)}$, $\hat{w}_j$, such that for every positive integer j finite dimensional distributions of

$$(\hat{x}_{l(j)}, \bar{x}_{m(j)}, \hat{w}_j) \quad \text{and} \quad (x_{l(j)}, x_{m(j)}, w)$$

coincide, and for $\hat{P}$-almost every $\hat{\omega} \in \hat{\Omega}$

$$\sup_{t\leqq T} |\hat{x}_{l(j)}(t) - \hat{x}(t)| \to 0, \qquad \sup_{t\leqq T} |\bar{x}_{l(j)}(t) - \bar{x}(t)| \to 0\,,$$

$$\sup_{t\leqq T} |\hat{w}_j(t) - \hat{w}(t)| \to 0\,,$$

as $j \to \infty$ for any $T < \infty$, where $\hat{x}$, $\bar{x}$, $\hat{w}$ are some stochastic processes. In the same way as above we get that for any k the processes $\hat{x}(t)$ and $\bar{x}(t)$ satisfy equation (2.1) on the time intervals $[0,\hat{\tau}^k)$ and $[0,\bar{\tau}^k)$, respectively, with $\hat{w}$

instead of w, where $\hat{\tau}^k$ and $\bar{\tau}^k$ are defined in an obvious way. Again as above $\hat{\tau}^k, \bar{\tau}^k \to \infty$, so that actually $\hat{x}(t)$ and $\bar{x}(t)$ satisfy the corresponding equation on $[0,\infty)$. Since the initial condition in both cases is the same ($\hat{x}_{l(j)}(0) = \bar{x}_{m(j)}(0)$ because $x_l(0)=x_m(0)=\xi$) and since the joint distribution of the initial value and $\hat{w}$ coincides with the distribution of ξ, w, by the pathwise uniqueness we conclude that $\hat{x}(t) = \bar{x}(t)$ for all t (a.s.). Hence, by applying Lemma 1.1 we finish the proof of Theorem 2.4.

4. An estimate of densities for Euler's approximations

In the case when the coefficients of equation (2.1) are not supposed to be continuous, in order to apply the above scheme we need a counterpart of Lemma 3.1 for measurable f. The proof of the corresponding assertion is based on an estimate on densities of distribution of the Euler approximation $x_n(t)$. Since such estimates can be applied in other situations, the result we prove below is stronger than we actually need in the proof of Theorem 2.8. First of all we need the following lemma.

Lemma 4.1. *Let $K, t, \varepsilon > 0$, $\alpha \in (0,1)$ be fixed numbers, and let $a(x)$ be a $d \times d$ matrix-valued function such that $KtI \geqq a = a^* \geqq \varepsilon tI$, where I is the $d \times d$ unit matrix. Also let $g(x)$ be a real-valued function such that $|g(x) - g(y)| \leqq K|x-y|^\alpha$ for all x, y. Let ξ and η be independent d-dimensional Gaussian vectors with zero means. Assume $\xi \sim \mathcal{N}(0,I)$. Define an operator T^* by the formula $T^*f(y) = Ef(y + \sqrt{a(y)}\xi)$ and let T be the conjugate for T^* in L_2-sense. Then for any $i,j = 1,\dots,d$, $x \in \mathbb{R}^d$, $p \in [1,\infty]$, and bounded Borel f*

$$\left| g(x)E\left[\frac{\partial^2}{\partial x^i \partial x^j} Tf\right](x+\eta) - E\left[\frac{\partial^2}{\partial x^i \partial x^j} T(gf)\right](x+\eta)\right| \leqq N t^{-d/(2p)-1+\alpha/2} \|f\|_p , \tag{4.1}$$

$$\left\{ \int_{\mathbb{R}^d} \left| g(x)E\left[\frac{\partial^2}{\partial x^i \partial x^j} Tf\right](x+\eta) - E\left[\frac{\partial^2}{\partial x^i \partial x^j} T(gf)\right](x+\eta)\right|^p dx \right\}^{1/p} \leqq N t^{-1+\alpha/2} \|f\|_p , \tag{4.2}$$

where the constants N depend only on K, ε, d, p and $E|\eta|^2$.

Proof. First observe that

$$\frac{\partial^2}{\partial x^i \partial x^j} Tf(x) = \int_{\mathbb{R}^d} (2\pi \det a(y))^{-d/2} f(y) \left.\frac{\partial^2}{\partial y^i \partial y^j} \exp\{-(a^{-1}(y-x), y-x)/2\}\right|_{a=a(y)} dy ,$$

$$E(2\pi \det a)^{-d/2} \exp\{-(a^{-1}(y-x-\eta), y-x-\eta)/2\} = (2\pi \det(a+a_1))^{-d/2} \exp\{-((a+a_1)^{-1}(y-x), y-x)/2\} =: p_a(x,y) ,$$

where a_1 is the covariance matrix of η. Let $A(y) = (a(y) + a_1)^{-1}$, then

$$E\frac{\partial^2}{\partial x^i \partial x^j} Tf(x+\eta)$$
$$= \int_{\mathbb{R}^d} f(y)[(A(y)(y-x))_i (A(y)(y-x))_j - A_{ij}(y)]\, p_{a(y)}(x,y)\, dy .$$

Thus the expression on the left in (4.1) equals

$$\left| \int_{\mathbb{R}^d} [g(x)-g(y)] f(y)[(A(y)(y-x))_i (A(y)(y-x))_j - A_{ij}(y)]\, p_{a(y)}(x,y)\, dy \right|$$
$$\leqq \frac{N}{t^{d/2}} \int_{\mathbb{R}^d} |f(y)||x-y|^{\alpha} \left[\frac{|x-y|^2}{t^2} + \frac{1}{t}\right] \exp\left\{-\frac{|x-y|^2}{Nt}\right\} dy$$
$$\leqq \|f\|_p \frac{N}{t^{d/2}} \left\{ \int_{\mathbb{R}^d} |y|^{q\alpha} \left[\frac{|y|^2}{t^2} + \frac{1}{t}\right]^q \exp\left\{-q\frac{|y|^2}{Nt}\right\} dy \right\}^{1/q}$$
$$= N t^{-d/2p - 1 + \alpha/2} \|f\|_p .$$

Here we have used the Hölder inequality. To prove (4.2) we apply instead the Minkowski inequality. The lemma is proved.

We will apply Lemma 4.1 to prove some estimates for distributions of the process $x_n(t)$ defined as

$$x_n(t) = x_0 + \int_0^t \sigma(s, x_n(\kappa_n(s)))\, dw(s) ,$$

where $x_0 \in \mathbb{R}^d$ is nonrandom and $\sigma : \mathbb{R}_+ \times \mathbb{R}^d :\longrightarrow \mathbb{R}^{d \times d_1}$ is Borel measurable and satisfies the condition

$$\varepsilon I \leqq (\sigma\sigma^{\mathrm{T}})(s,x) \leqq KI, \qquad |\sigma(s,x) - \sigma(s,y)| \leqq K|x-y|^{\alpha}$$

for some constants $\alpha \in (0,1)$, $K, \varepsilon > 0$ and all $x, y \in \mathbb{R}^d$, $s > 0$. Before stating the main result of this section we introduce some notations. For fixed n and $t > 0$ a very cumbersome expression can be found explicitly in an obvious way for the distribution density $p_n(t,x)$ of $x_n(t)$. We do not know if it is possible to estimate the density analyzing this expression, but at least it shows that the density is bounded on $[\delta, \delta^{-1}] \times \mathbb{R}^d$ for any $\delta > 0$. We denote by $m_n(t)$ the supremum of $p_n(t,x)$ over $x \in \mathbb{R}^d$. The function $m_n(t)$ is bounded on $[\delta, \delta^{-1}]$ for any $\delta > 0$ and any n.

Theorem 4.2. (a) *There exists a constant N_0 depending only on $d, \alpha, K, \varepsilon, q$ such that if $1 \leqq q < \frac{d}{d-\alpha}$, then for all $t > 0$, $n = 1, 2, 3, \ldots$*

$$\left(\int_{\mathbb{R}^d} p_n^q(t,x)\, dx \right)^{1/q} \leqq N_0(t^{-d/(2p)} + 1) \quad (p = q/(q-1)) . \tag{4.3}$$

(b) *If the partitions $\{0 = t_0^n < t_1^n < \cdots\}$ satisfy the additional condition $\kappa_n(s) \geqq \varepsilon s$ for all n and $s > t_1^n$, then there exists a constant N_0 depending only* *on* $d, \alpha, K, \varepsilon$ *such that*

$$m_n(t) \leqq N_0(t^{-d/2} + 1), \quad t > 0, \ n = 1, 2, \dots . \tag{4.4}$$

Also (4.3) *holds, for any* $q \in [1, \infty]$, $t > 0$, $n = 1, 2, 3, \dots$.

Proof. The last assertion in (b) is true since $p_n^q \leqq p_n(m_n)^{q-1}$ and $\int p_n \, dx = 1$. To prove (a) for $0 \leqq s \leqq t < \infty$ and bounded measurable $f(x)$ let

$$T_{s,t}^* f(y) := Ef\left(y + \int_s^t \sigma(r, y) \, dw(r)\right),$$

and let the operator $T_{s,t}$ be conjugate to $T_{s,t}^*$ in L_2-sense. The expression $T_{s,t}f(x)$ can be written as an integral with respect to a Gaussian-like density, and from this formula it is not hard to see that for any $s < t$ the function $T_{s,t}f(x)$ is infinitely differentiable and for $s < t$

$$\frac{\partial}{\partial s} T_{s,t} f(x) = -\frac{\partial^2}{\partial x^i \partial x^j} T_{s,t} a^{ij}(s, \cdot) f(\cdot)(x), \tag{4.5}$$

where $a_{ij} := \frac{1}{2}(\sigma\sigma^{\mathrm{T}})_{ij}$. For the sake of simplicity of notations we drop the subscripts n, and from (4.5) by the Newton–Leibnitz and Itô's formulas for any $r \in [0, t]$ we obtain

$$\begin{aligned} Ef(x(t)) &= \int_r^t \frac{d}{ds} ET_{s,t} f(x(s)) \, ds + ET_{r,t} f(x(r)) \\ &= ET_{r,t} f(x(r)) + \int_r^t E[a^{ij}(s, x(\kappa(s))) \frac{\partial^2}{\partial x^i \partial x^j} T_{s,t} f(x(s)) \\ &\quad - \frac{\partial^2}{\partial x^i \partial x^j} T_{s,t} a^{ij}(s, \cdot) f(\cdot)(x(s))] \, ds . \end{aligned}$$

We take conditional expectations given $x(\kappa(s))$, and after denoting

$$\eta(s, x) = \int_{\kappa(s)}^s \sigma(r, x) \, dw(r)$$

we get

$$Ef(x(t)) = ET_{r,t} f(x(r)) + \int_r^t EH(s, t, x(\kappa(s))) \, ds , \tag{4.6}$$

where

$$\begin{aligned} H(s, t, x) &= a_{ij}(s, x) E\left[\frac{\partial^2}{\partial x^i \partial x^j} T_{s,t} f\right](x + \eta(s, x)) \\ &\quad - E\left[\frac{\partial^2}{\partial x^i \partial x^j} T_{s,t} a_{ij}(s, \cdot) f(\cdot)\right](x + \eta(s, x)) . \end{aligned}$$

By Lemma 4.1

$$\begin{aligned} &|H(s, t, x)| \leqq N(t - s)^{-d/(2p) - 1 + \alpha/2} \|f\|_p, \\ &\int_{\mathbb{R}^d} |H(s, t, x)| dx \leqq N(t - s)^{-1 + \alpha/2} \|f\|_1 . \end{aligned} \tag{4.7}$$

This and (4.6) with $r=0$ give us (4.3) for $p>d/\alpha$ and for $t\in(0,T]$ with a constant N_0 depending on d,α,K,ε,q and also on T. Indeed,

$$T_{0,t}f(x_0)\leqq Nt^{-d/2}\int_{\mathbb{R}^d} f(y)\exp\left\{-\frac{1}{Nt}(x-y)^2\right\}dy\leqq Nt^{-d/(2p)}\|f\|_p\,,$$

$$\int_0^t (t-s)^{-d/(2p)-1+\alpha/2}\,ds=Nt^{-d/(2p)+\alpha/2}\,.$$

To prove (4.4) and (4.3) with a constant N_0 independent of T we need a longer argument. Fix $T\in(0,\infty)$, and define γ_T as the smallest number γ such that $m(s)\leqq\gamma(s^{-d/2}+1)$ for all $s\in(0,T]$. Introduction of such objects as γ_T is rather common in the theory of PDE. In probability theory they were used for instance in Stroock–Varadhan [13] for the same purposes. Such a number γ_T does exist since $m(t)$ is bounded on $[t_1^n,T]$ and $m(t)\leqq N(d,K,\varepsilon)t^{-d/2}$ for $t\in(0,t_1^n)$ as follows from the explicit formula for the Gaussian density of $x(t)=x_0+\int_0^t\sigma(s,x_0)\,dw(s)$. We want to estimate γ_T. By using (4.6), (4.7) and the inequality $\kappa(s)\geqq\varepsilon s$ for $s\geqq t_1^n$, we obtain

$$\begin{aligned}
&Ef(x(t))\\
&\leqq Nt^{-d/2}\|f\|_1+\int_{t_1^n}^t\left[\gamma_T\left(\frac{1}{\kappa^{d/2}(s)}+1\right)\|H(s,t,\,\cdot\,)\|_1\right]\wedge\sup_x|H(s,t,x)|\,ds\\
&\leqq\left\{Nt^{-d/2}+N\int_{t_1^n}^t\left[\gamma_T\left(\frac{1}{\kappa^{d/2}(s)}+1\right)\frac{1}{(t-s)^{1-\alpha/2}}\right]\right.\\
&\qquad\left.\wedge\frac{1}{(t-s)^{d/2+1-\alpha/2}}\,ds\right\}\|f\|_1\,,
\end{aligned}$$

(4.8)
$$m(t)\leqq Nt^{-d/2}+N\int_0^t\left[\gamma_T\left(\frac{1}{s^{d/2}}+1\right)\frac{1}{(t-s)^{1-\alpha/2}}\right]\wedge\frac{1}{(t-s)^{d/2+1-\alpha/2}}\,ds$$

for $t\in[t_1^n,T]$. Next as easy to see after the substitution $s=u\gamma_T^{-2/d}$,

$$\begin{aligned}
\int_0^t\frac{\gamma_T}{(t-s)^{1-\alpha/2}}\wedge\frac{1}{(t-s)^{d/2+1-\alpha/2}}\,ds&=\int_0^t\frac{\gamma_T}{s^{1-\alpha/2}}\wedge\frac{1}{s^{d/2+1-\alpha/2}}\,ds\\
&=\gamma_T^{1-\alpha/d}\int_0^{t\gamma_T^{2/d}}\frac{1}{u^{1-\alpha/2}}\wedge\frac{1}{u^{d/2+1-\alpha/2}}\,du\\
&\leqq N\gamma_T^{1-\alpha/d}\,.
\end{aligned}$$

Upon setting $u = t\gamma_T^{2/d}(1+\gamma_T^{2/d})^{-1}$, we also have

$$\begin{aligned}
&\int_0^t \frac{\gamma_T}{s^{d/2}(t-s)^{1-\alpha/2}} \wedge \frac{1}{(t-s)^{d/2+1-\alpha/2}}\,ds \\
&\leqq \int_0^u \frac{1}{(t-s)^{d/2+1-\alpha/2}}\,ds + \int_u^t \frac{\gamma_T}{s^{d/2}(t-s)^{1-\alpha/2}}\,ds \\
&\leqq \frac{2}{(d-\alpha)(t-u)^{d/2-\alpha/2}} + \gamma_T u^{-d/2}\frac{2}{\alpha}(t-u)^{\alpha/2} \\
&= Nt^{-(d-\alpha)/2}(1+\gamma_T^{2/d})^{(d-\alpha)/2} \leqq N(1+\gamma_T^{1-\alpha/d})(t^{-d/2}+1)\,.
\end{aligned}$$

Thus from (4.8) for $t \in [t_1^n, T]$ we conclude

$$m(t) \leqq N(1+\gamma_T^{1-\alpha/d})(t^{-d/2}+1)\,. \tag{4.9}$$

As we observed above this estimate is also true for $t \in (0, t_1^n]$. By definition of γ_T estimate (4.9) means that

$$\gamma_T \leqq N(1+\gamma_T^{1-\alpha/d})\,.$$

We emphasize that the last constant N, as well as all constants called N in the above proof of (4.4), depends only on $d, \alpha, K, \varepsilon$. This implies the desired estimate of γ_T, and it remains only to notice that the estimate is independent of T. We can see in the same way that the constant N_0 in the estimate (4.3) can be taken to be the same for all $t > 0$. The theorem is proved.

Corollary 4.3. *Assume the conditions of Theorem* 2.8. *Let $x_n(t)$ be the Euler approximation defined by* (2.2) *and let τ_n^k be the first exit time of $x_n(t)$ from D_k. Then for every $t > 0$ the measure $P(x_n(t) \in \Gamma < \tau_n^k)$ has a density $p_n^k(t,x)$, and for any $0 < t_0 < T < \infty$, $1 \leqq q < \frac{d}{d-\alpha}$ and $k = 1, 2, \ldots$ we have*

$$\sup_n \sup_{t\in[t_0,T]} \int_{\mathbb{R}^d} [p_n^k(t,x)]^q\,dx < \infty\,. \tag{4.10}$$

Proof. By using a nonrandom time change we easily reduce the general case to the one with $M_k(t) \equiv 1$. Next we observe that

$$P(x_n(t) \in \Gamma,\ t < \tau_n^k) \leqq P(x_n^k(t) \in \Gamma)\,,$$

where $x_n^k(t)$ are Euler's approximations for equation (2.1) with coefficients σ, b changed arbitrarily outside D_k. After this an application of the Girsanov theorem allows us to take $b \equiv 0$. Finally we get our assertion from (4.3) if we notice the obvious relation between Euler's approximations for fixed initial value and for random one.

Remark. 4.4. One knows from Fabes and Kenig [2] and Safonov [12] that none of the estimates (4.3), (4.4) and (4.10) remains valid if the Hölder continuity of σ in x is replaced by the assumption of uniform continuity of σ in (t,x).

5. Proof of Theorem 2.8

The reader can easily check that we can repeat the proof of Theorem 2.4 from Sect. 3, if we prove the following version of Lemma 3.1. We use the same notations as in this lemma.

Lemma 5.1. *Let $f(s,x)$ be a Borel function defined on $\mathbb{R}_+ \times \mathbb{R}^d$ such that $|f(t,x)| \leqq M_k(t)$ for any k and $x \in D_k$. Then for any $i = 1,\dots,d_1$ the first two convergences in* (3.6) *hold as $j \to \infty$ uniformly in $t \in [0, T \wedge \tilde{\tau}^k)$ in probability for any $T < \infty$. If $|f(t,x)|^2 \leqq M_k(t)$ for any k and $x \in D_k$ then for any $i = 1,\dots,d_1$ the last two convergences* (3.6) *also hold as $j \to \infty$ uniformly in $t \in [0, T \wedge \tilde{\tau}^k)$ in probability for any $T < \infty$.*

Proof. We will prove only the last relation in (3.6). The other ones can be proved similarly. Take a function $g(t,x)$ defined on $\mathbb{R}_+ \times \mathbb{R}^d$ such that it is continuous in x, Borel in t and satisfies the same hypotheses as f. Define

$$I_t^{kj}(g) = \int_0^t g(s, \tilde{x}^k_{n(j)}(\kappa_{n(j)}(s)))\, d\tilde{w}^i_j(s), \qquad I_t^k(g) = \int_0^t g(s, \tilde{x}^k(s))\, d\tilde{w}^i(s)\,.$$

Owing to Lemma 3.1 for any $\delta > 0$ we have

$$\begin{aligned} (5.1) \qquad & \limsup_{j\to\infty} P(\sup\{|I_t^{kj}(f) - I_t^k(f)|:\ t < T \wedge \tilde{\tau}^k\} \geqq 3\delta) \\ & \leqq \limsup_{j\to\infty} P(\sup\{|I_t^{kj}(f-g)|:\ t < T \wedge \tilde{\tau}^k\} \geqq \delta) \\ & \quad + P(\sup\{|I_t^k(f-g)|:\ t < T \wedge \tilde{\tau}^k\} \geqq \delta) =: J_1 + J_2\,. \end{aligned}$$

Now, by virtue of (3.4) and the well-known martingale inequalities

$$\begin{aligned} J_1 & \leqq \gamma^{-1} \limsup_{j\to\infty} E \int_0^{T\wedge\tilde{\tau}^k} |f-g|^2(s, \tilde{x}^k_{n(j)}(\kappa_{n(j)}(s)))\, ds + \frac{\gamma}{\delta^2} \\ & \leqq 4\gamma^{-1} \int_0^\eta M_k(s)\, ds \\ & \quad + \gamma^{-1} \limsup_{j\to\infty} \int_\eta^T E|f-g|^2(s, \tilde{x}^k_{n(j)}(\kappa_{n(j)}(s))) I_{s<\tilde{\tau}^k_{n(j)}}\, ds + \frac{\gamma}{\delta^2}\,, \end{aligned}$$

where $\gamma > 0$ and $\eta > 0$ are arbitrary numbers. By Corollary 4.3 we conclude that for p large enough

$$J_1 \leqq 4\gamma^{-1} \int_0^\eta M_k(s)\, ds + \frac{\gamma}{\delta^2} + N\gamma^{-1} \left[\int_0^T \int_{D_k} |f-g|^{2p}(s,x)\, dx\, ds \right]^{1/p}$$

with N independent of g. Since $\tilde{x}^k_{n(j)}(t) \to \tilde{x}^k(t)$ (a.s.) from Corollary 4.3 we also get an estimate for probability density of $x^k(t)$, which shows that J_2 is bounded from above by the same quantity as J_1. Thus we obtain an estimate for the first limit in (5.1), and this estimate along with the freedom of choice of g, η, γ shows that the limit in question is zero. This brings to the end the proofs of Lemma 5.1 and Theorem 2.8.

Acknowledgements. The authors are sincerely grateful to T. Shiga and to the anonymous referee for drawing their attention to paper [5].

This research was completed during the stay of the first author at the University of Minnesota, and he is grateful for the hospitality, excellent working conditions and for the inspiring scientific atmosphere of this institution.

References

1. Alyushina, L.A.: Euler polygonal lines for Itô's equations with monotone coefficients. Theory Probab. Appl. **32**, 340–346 (1987)
2. Fabes, E.B., Kenig, C.E.: Examples of singular parabolic measures and singular transition probability densities. Duke Mathematical Journal **48**, 848–856 (1981)
3. Gyöngy, I., Nualart, D., Sanz-Solé, M.: Approximation and support theorems in modulus spaces. Probab. Theory Relat. Fields **101**, 495–509 (1995)
4. Ikeda, N., Watanabe, S.: Stochastic differential equations and diffusion processes. Amsterdam Oxford New York: North-Holland 1981
5. Kaneko, H., Nakao, S.: A note on approximation for stochastic differential equations. Séminaire de Probabilités XXII (Lecture Notes in Mathematics **1321**, 155–162) Berlin, Heidelberg: Springer, 1988
6. Krylov, N.V.: On Itô's stochastic integral equations. Theory Probab. Appl. **14**, 330–336 (1969)
7. Krylov, N.V.: A simple proof of the existence of a solution of Itô's equation with monotone coefficients. Theory Probab. Appl. **3**, 583–587 (1990)
8. Krylov, N.V.: Extremal properties of solutions of stochastic equations. Theory Probab. Appl. **29**, 205–214 (1984)
9. Krylov, N.V.: Controlled diffusion processes. New York, Heidelberg, Berlin: Springer, 1980
10. Maruyama, G.: Continuous Markov processes and stochastic equations. Rend Circ. Mat. Palermo **4**, 48–90 (1955)
11. Nakao, S.: On the pathwise uniqueness of solutions of one-dimensional stochastic differential equations. Osaka J. Math. **9**, 513–518 (1972)
12. Safonov, M.V.: An example of a diffusion process with singular distribution at some given time. Abstract of Communications, Third Vilnius Conference on Probability Theory and Mathematical Statistics, pp. 133–134, Vilnius, June 22–27, 1981
13. Stroock, D.W., Vardhan, S.R.S.: Multidimensional diffusion processes. New York: Springer, 1979
14. Skorokhod, A.V.: Studies in the theory of random processes. New York: Dover, 1982
15. Veretennikov, A.Yu.: On strong solution and explicit formulas for solutions of stochastic integral equations. Math. USSR Sb. **39**, 387–403 (1981)
16. Yamada, T., Watanabe S.: On the uniqueness of solutions of stochastic differential equations, I, II. J. Math. Kyoto Univ. **11**, 155–167, 553–563 (1971)
17. Zvonkin, A.K., Krylov, N.V.: On strong solutions of stochastic differential equations. Sel. Math. Sov. **1**, 19–61 (1981)