



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Existence of strong solutions for $It^{\wedge}o$'s stochastic equations via approximations. Revisited

Citation for published version:

Gyongy, I & Krylov, NV 2021, 'Existence of strong solutions for $It^{\wedge}o$'s stochastic equations via approximations. Revisited', *Stochastics and Partial Differential Equations: Analysis and Computations*. <<https://arxiv.org/abs/2107.14384>>

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Stochastics and Partial Differential Equations: Analysis and Computations

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



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

I. Gyöngy · N.V. Krylov

Received: date / Accepted: date

Abstract Given strong uniqueness for an Itô's stochastic equation, 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. This is almost a copy of an old article in which we correct errors in the original proof of Lemma 4.1 found by Martin Dieckmann in 2013. We present also a new result on the convergence of “tamed Euler approximations” for SDEs with locally unbounded drifts, which we achieve by proving an estimate for appropriate exponential moments.

Keywords Stochastic differential equations · strong solutions · Euler's approximations

Mathematics Subject Classification (2000) 60H10 · 60H17

1 Introduction

We start with two examples illustrating the results we present in the paper. Consider the stochastic differential equation

$$dx(t) = \left[\tan\left(-\frac{\pi}{2}x(t)\right) + 1 \right] dt + |1 - |x(t)||^\alpha (x_+(t))^{\frac{1}{2}} dw(t), \quad x(0) = 0 \quad (1.1)$$

with a given $\alpha > 0$, where w is a Wiener process. Note that the drift coefficient $\tan(-\frac{\pi}{2}x)$ is not continuous at $x = 2k + 1$, for integers k , and it does not satisfy the linear growth condition. Note moreover that the diffusion coefficient $|1 - |x||^\alpha (x_+)^{\frac{1}{2}}$ does not satisfy the linear growth condition for $\alpha > \frac{1}{2}$ and it

I. Gyöngy
School of Mathematics and Maxwell Institute, University of Edinburgh, Scotland, UK
E-mail: I.Gyongy@ed.ac.edu

N.V. Krylov
School of Mathematics, University of Minnesota, Minneapolis, MN, 55455, USA
E-mail: nkrylov@umn.edu

is not Hölder continuous with exponent $1/2$ if $\alpha < 1/2$. Consider also the equation

$$dx(t) = \left[\tan\left(-\frac{\pi}{2}x(t)\right) + \text{sign } x(t) \right] dt + |1 - |x(t)||^\alpha dw(t), \quad x(0) = 0, \quad (1.2)$$

and note that here the drift is discontinuous also at 0.

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

$$dx_n(t) = b(x_n(\kappa_n(t))) dt + \sigma(x_n(\kappa_n(t))) dw(t), \quad x_n(0) = 0 \quad (1.3)$$

for every integer $n > 0$, where $\kappa_n(t) := \lfloor nt \rfloor / n$, with the corresponding drift and diffusion coefficient, setting for example $b(x) = 0$ when x is an odd integer. One expects that in each of these examples x_n converges in probability to a process which solves the corresponding equation (1.1) and (1.2), respectively.

In fact, the drift in equation (1.1) is Lipschitz continuous and the diffusion coefficient is Hölder continuous with exponent $\frac{1}{2}$ at $x = 0$. Therefore by a well-known result of Yamada and Watanabe [23] one knows the existence (at least of a local) strong solution to equation (1.1). In the case of equation (1.2) one can say the same due to a result of Veretennikov [22], stating the existence of a unique strong solution to the stochastic differential equation

$$dx(t) = b(t, x(t)) dt + \sigma(t, x(t)) dw(t), \quad x_0 \in \mathbb{R}^d \quad (1.4)$$

in \mathbb{R}^d , with a given d_1 -dimensional Wiener process w , if b, σ are 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; $\sigma\sigma^*$ 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.

The method of establishing these existence and uniqueness theorems is rather different from those used in the theory of ordinary differential equations. It is based on a famous result from Yamada and Watanabe [23] which reads as follows. If any two solutions to equation (1.4) on the same probability space with the same Wiener process almost surely coincide, and if there is a solution to the equation on some probability space with some suitable Wiener process, then there exists a strong solution to the equation with the given Wiener process. Shortly speaking, the existence of a solution and the pathwise uniqueness imply the existence of the unique strong solution. (See also Zvonkin and Krylov [26] and the references therein on this topic.) We emphasize that by this approach one gets only pure existence result, without presenting any construction of the solution.

The existence of a solution to equation (1.4) with bounded measurable coefficients is known under the additional condition that $\sigma(t, x), b(t, x)$ are continuous in x (Skorokhod [21], Stroock and Varadhan [20]), or $\sigma\sigma^*$ is uniformly elliptic (Krylov [9], [12]), regarding recent progress in the case of singular b see [13] and the references therein. Hence Veretennikov, Yamada and Watanabe establish the existence of a strong solution (in [22] and in [23], respectively) by proving the pathwise uniqueness. Their proofs raise the following questions.

Is it possible to construct the strong solutions in some classical way under the conditions of their theorems? Define for example Euler's approximations (1.3) to equation (1.2). Do these approximations converge to a stochastic process in probability and can one construct a strong solution in this way? Let us approximate the coefficients in the equation (1.4) by smooth ones. Do the strong solutions of the corresponding equations converge in probability to the strong solution of equation (1.4) 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, roughly speaking, that Euler's polygonal approximations converge uniformly in t in bounded intervals, in probability, to a process, which we show to be the strong solution, if the pathwise uniqueness for the equation holds, provided the drift and diffusion coefficients have some additional property permitting the passage to the limit. Such additional property is their continuity in the space variable, or the strong ellipticity of the diffusion coefficient. (See Theorems 2.1 and 2.2 below.) In particular, applying Corollaries 2.2 and 2.3 to equations (1.1) and (1.2), respectively with $D := (-1, 1)$, $D_k := (-1+2^{-k}, 1-2^{-k})$ and with $V(t, x) := \frac{2-x^2}{1-x^2}$, we get that Euler's approximations $x_n(t)$, defined by (1.3) converge uniformly in t in bounded intervals in probability to some stochastic processes, which are the strong solutions of equations (1.1) and (1.2) respectively.

Let us finally consider the following example of an SDE with singular drift

$$dx(t) = |x(t)|^{-\alpha} I_{|x(t)| < 1} dt + (2 + \sin(x(t))) dw(t), \quad x(0) = 0. \quad (1.5)$$

By results on SDEs with locally unbounded drifts (see, e.g., [16], [5], [24], [25], [18]) one can see that this equation has a unique strong solution for any $\alpha \in (0, 1)$. Note that the Euler approximations (1.3) are not meaningful, but we can define the "tamed" Euler approximations

$$dx_n(t) = b_n(x_n(\kappa_n(t))) dt + (2 + \sin(x_n(\kappa_n(t)))) dt, \quad x_n(0) = 0$$

for example with $b_n(x) = (|x|^{-\alpha} I_{|x| < 1}) \wedge \lambda_n$ for a sequence of positive constants λ_n converging to infinity. Applying our result, Theorem 2.3 below, we get that if $\alpha \in (1, 1/2)$ and λ_n converges to infinity sufficiently slowly, then the tamed Euler approximations converge to the solution $x(t)$ of equation (1.5), in probability, uniformly in t in bounded intervals.

The possibility to show convergence 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}, ρ) 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 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 $\{(x, y) \in \mathbb{E} \times \mathbb{E} : x = y\}$.*

The necessity of the condition is obvious. To prove the sufficiency 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, $f(v_k) \rightarrow 0$ 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.

In our applications of the lemma Skorokhod's embedding method and the assumption about 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 its 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, it can be seen in the same way as Theorems 2.1 and 2.2 are proved that, the strong solution can be constructed by approximating the coefficients by smooth ones. One can construct the strong solution by Euler's approximations and approximating simultaneously the coefficients and the initial condition. We remark, that clearly we immediately get the convergence of Euler's approximations (or of the other approximations we mentioned) in probability in every metric space V , where these approximations are tight. (See Gyöngy, Nualart and Sanz-Solé [6], where the convergence in probability of Wong-Zakai type approximations are proved in the modulus spaces introduced there.)

Finally we remark that the convergence of Euler's approximations under various conditions is proved by many authors. It is shown in Krylov [11] that under the monotonicity condition Euler's polygonal line method can be adjusted to prove (strong) solvability. Earlier this was known from Maruyama [15] when the drift and diffusion coefficients are Lipschitz continuous. The method of [11] was afterward used in Alyushina [1] in a short proof of existence of strong solutions under the monotonicity and the linear growth conditions. Later a short and simple proof of (strong) solvability is presented in Krylov [10] under the monotonicity condition and under a growth condition which is weaker than the usual linear growth condition. Moreover, the continuous dependence of the strong solution on the coefficients is also obtained.

It is also worth mentioning that the fact that the pathwise uniqueness implies the possibility of effectively constructing the solutions has already been noticed in Zvonkin and Krylov [26] (see, for instance, Lemma 3.2 there). Later Kaneko and Nakao [8] exploited this fact without noticing [26]. In [8] the authors consider equation (1.4) 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 another result of Yamada and Watanabe saying that pathwise uniqueness implies uniqueness in law. Essentially the same idea is used in [9]. Due to our above lemma this idea becomes more apparent and its range of applicability becomes evident.

We remark that since the original version of the present paper was published there has been a growing interest in studying the convergence of Euler's approximations for SDEs with irregular coefficients. For recent results on the rate of convergence we refer to [2], [14], [17], and the references therein.

The paper is organized as follows. In the next section we formulate our results Theorems 2.1, 2.2, 2.3 and their corollaries. By Lemma 1.1 the proof of Theorem 2.1 is simple, we present it in Section 3. To prove Theorem 2.2 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 4.1 below) separately in Section 4. We prove our main results, Theorem 2.2 and 2.3, in the last two sections.

2 Formulation of the results

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

$$dx(t) = b(t, x(t)) dt + \sigma(t, x(t)) dw(t), \quad x(0) = \xi \quad (2.1)$$

in a domain D of \mathbb{R}^d , where $(w(t), \mathcal{F}_t)$ is a d_1 -dimensional Wiener process, ξ is an \mathcal{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 that values of σ and b outside D are irrelevant and just for convenience we define $\sigma(t, x) = 0, b(t, x) = 0$ for $x \notin D, t \geq 0$. Let

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

be a sequence of partitions of \mathbb{R}_+ such that $\lim_{i \rightarrow \infty} t_i^n = \infty$ for every $n \geq 1$, and for every $T > 0$

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

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

$$dx_n(t) = b(t, x_n(\kappa_n(t))) dt + \sigma(t, x_n(\kappa_n(t))) dw(t), \quad x_n(0) = \xi \quad (2.2)$$

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 $M_1(t) > 0, M_2(t) > 0, \dots$ 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 $\cup_{k=1}^\infty D_k = D$, and for every $k, t \in [0, k]$

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

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

$$LV(t, x) \leq M(t)V(t, x), \quad \forall t \geq 0, x \in D,$$

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

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

$$L := \frac{\partial}{\partial t} + \sum_i b_i(t, x)D_i + \frac{1}{2} \sum_{i,j} (\sigma\sigma^*)_{ij}(t, x)D_{ij} \quad D_i = \frac{\partial}{\partial x^i}, \quad D_{ij} = D_i D_j,$$

where σ^* denotes the transpose of the matrix σ ;

(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 $t \geq 0$.

Definition 1 By solution of equation (2.1) we mean an \mathcal{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 2.1 *Let $x(t)$ be an \mathcal{F}_t -adapted process defined for all $t \geq 0$. Assume that $x(t)$ satisfies (2.1) for $t < \tau := \inf\{t : x(t) \notin D\}$, and assume (i)–(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 $\delta, T > 0$ we have

$$P(\tau^k \leq T) \leq P(\xi \notin D_k) + P(V(0, \xi) \geq \log(1/\delta)) + \frac{1}{\delta V_k(T)} \exp \int_0^T M(t) dt. \quad (2.3)$$

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

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

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

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

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 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 $w'(t), \xi'$ in place of $w(t), \xi$ cannot have more than one solution.

Theorem 2.1 *Assume (i)–(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 $\mathcal{F}_t^w \vee \sigma(\xi)$ -adapted.*

Remark 2.1 Note that taking $V(t, x) := (|x|^2 + 1) \exp(-\int_0^t M(s) ds)$ in the case $D = \mathbb{R}^d$, $D_k := \{x \in \mathbb{R}^d : |x| < k\}$, conditions (i)–(ii) can be restated as follows:

- (1) $\sup_{|x| < k} \{|b(t, x)| + |\sigma(t, x)|^2\} \leq M_k(t)$ for every $t \geq 0$ and integer $k \geq 1$;
 - (2) $2(x, b(t, x)) + \|\sigma(t, x)\|^2 \leq M(t)(|x|^2 + 1)$ for every $t \geq 0$ and $x \in \mathbb{R}^d$,
- where $\|\alpha\|$ denotes the Hilbert–Schmidt norm for matrices α and (x, y) is the scalar product of $x, y \in \mathbb{R}^d$.

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

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

Corollary 2.1 (c.f. [10]) *Assume (i)–(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 2.1 are satisfied and that the monotonicity condition is satisfied for $D_k = \{x \in \mathbb{R}^d : |x| < k\}$. Assume moreover that b is continuous in $x \in D$. Then the conclusions of Theorem 2.1 hold.*

Proof One can easily show that the monotonicity condition implies the pathwise uniqueness (see e.g. Krylov [11]). Hence this corollary is immediate from Theorem 2.1.

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

Corollary 2.2 *Let $d = 1$. Assume (i)–(iii) and let b be continuous in x in D for any t . Assume moreover that for every k and $t \geq 0$, $x, y \in D_k$ we have*

$$(x - y, b(t, x) - b(t, y)) \leq M_k(t)|x - y|^2, \quad |\sigma(t, x) - \sigma(t, y)|^2 \leq M_k(t)\rho_k(|x - y|),$$

where $\rho_k \geq 0$ is an increasing function on \mathbb{R}_+ such that

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

for some $\varepsilon > 0$. Then the conclusions of Theorem 2.1 hold.

Proof For any given $k = 1, 2, \dots$ 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 [23] (see also Ikeda and Watanabe [7]) 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 2.1.

If we are dealing with nondegenerate equations, the continuity condition on b in Theorem 2.1 can be dropped. To state this more precisely, in addition to the conditions (i)–(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

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

for every $t \in [0, k]$, $x \in D_k$, $\lambda \in \mathbb{R}$, 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 $t \geq 0$, $x, y \in D_k$

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

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

Theorem 2.2 *Assume (i)–(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 $\mathcal{F}_t^w \vee \sigma(\xi)$ -adapted.*

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

Corollary 2.3 *Let $d = 1$ and assume (i)–(iv). Suppose that σ is locally Hölder in x in D with some exponent $\alpha \in (0, 1]$. Assume moreover that for every k*

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

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

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

for some $\varepsilon > 0$. Then the conclusions of Theorem 2.2 hold.

Proof Using the result obtained in Veretennikov [22] on pathwise uniqueness for stochastic Itô's equations in one dimension (which generalizes the corresponding results in Yamada and Watanabe [23] and in Nakao [16]), we can repeat the argument from the proof of Corollary 2.2.

Finally we present a result on Euler's approximations for the equation

$$dx(t) = b(t, x(t)) dt + \sigma(t, x(t)) dw(t), \quad x(0) = \xi \quad (2.4)$$

with locally unbounded drift $b = b(t, x)$ and bounded uniformly non-degenerate $\sigma = \sigma(t, x)$, which is Hölder continuous in x , where ξ is and \mathcal{F}_0 -measurable random vector in \mathbb{R}^d .

We will be dealing with tamed Euler approximations for (2.4) defined as

$$dx_n(t) = b_n(t, x_n(\kappa_n(t))) dt + \sigma(t, x_n(\kappa_n(t))) dw(t), \quad x_n(0) = \xi, \quad (2.5)$$

where b_n are certain Borel functions.

To formulate our conditions, for $p, q \in [1, \infty]$ we introduce the notation $L_p = L_p(\mathbb{R}^d)$ for the usual space of Borel functions on \mathbb{R}^d summable to the power p with norm $\|\cdot\|_p$ and use $L_{p,q}(T) = L_{p,q}([0, T] \times \mathbb{R}^d)$ for the space of Borel functions $f = f(t, x)$ on $[0, T] \times \mathbb{R}^d$ such that

$$\|f\|_{p,q,T} = \left(\int_0^T \|f(t)\|_p^q dt \right)^{1/q} < \infty \quad \text{when } p, q \in [1, \infty),$$

$$\|f\|_{\infty,q,T} = \left(\int_0^T \sup_{\mathbb{R}^d} |f(t, x)|^q dt \right)^{1/q} < \infty \quad \text{when } p = \infty, q \in [1, \infty) \quad (2.6)$$

and $\|f\|_{p,\infty,T} = \lim_{q \rightarrow \infty} \|f\|_{p,q,T} < \infty$ when $q = \infty, p \in [1, \infty]$.

Assumption 2.1 (1) *The diffusion coefficient σ is a Borel function on $[0, \infty) \times \mathbb{R}^d$ such that for each $T \in [0, \infty)$ there are constants $\varepsilon > 0$, $K < \infty$ and $\alpha \in (0, 1)$ such that*

$$\varepsilon I \leq (\sigma\sigma^*)(s, x) \leq KI, \quad |\sigma(s, x) - \sigma(s, y)| \leq K|x - y|^\alpha \quad (2.7)$$

for $s \in [0, T]$ and $x, y \in \mathbb{R}^d$, where I is the $d \times d$ unit matrix.

(2) *For each $T > 0$ we have $|b| \in L_{2p,2q,T}$ for some $q \in (1, \infty)$ and $p \in (\frac{d}{\alpha}, \infty)$, such that*

$$\frac{d}{p} + \frac{2}{q} < 2. \quad (2.8)$$

For each $T \in (0, \infty)$ we have $b_n \rightarrow b$ in $L_{2p,2q,T}$.

(3) *For each $T > 0$ there is a constant $\delta(T) > 0$ such that*

$$\min_{i:t_{i+1} \leq T} (t_{i+1}^n - t_i^n) / d_n(T) \geq \delta(T) \quad \text{for all } n \geq 1, \quad (2.9)$$

and for a $\gamma = \gamma(T) \in (0, (q-1)/q)$

$$B(T) := \sup_{n \geq 1} d_n^{\gamma/2}(T) \|b_n\|_{\infty,2q,T} < \infty.$$

Theorem 2.3 *Under Assumption 2.1 suppose that for equation (2.4) the pathwise uniqueness holds. Then the tamed Euler approximations converge in probability, uniformly on finite time intervals, to a continuous $\mathcal{F}_t^w \vee \sigma(\xi)$ -adapted process $x(t)$, which is the unique solution of (2.4).*

3 Proof of Theorem 2.1

For every positive integers k, n define the stopping time

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

Then

$$|b(t, x_n(\kappa_n(t)))| \leq M_k(t), \quad |\sigma(t, x_n(\kappa_n(t)))|^2 \leq M_k(t)$$

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

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

is tight in $C([0, T])$ for every k and $T \geq 0$. We want to deduce from this the tightness in $C([0, T])$ of

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

Clearly, it suffices to show that

$$\lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} P(\tau_n^k \leq T) = 0. \quad (3.2)$$

At first fix k and apply Skorokhod's embedding theorem. Then by virtue of the tightness of distributions of $x_n^k(t)$ in $C([0, T])$ for every $T \geq 0$, we can find a subsequence $n(j)$ and a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$, 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

$$(\tilde{x}_{n(j)}^k, \tilde{w}_j) \text{ and } (x_{n(j)}^k, w)$$

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

$$\sup_{t \leq T} |\tilde{x}_{n(j)}^k(t) - \tilde{x}^k(t)| \rightarrow 0, \quad \sup_{t \leq T} |\tilde{w}_j(t) - \tilde{w}(t)| \rightarrow 0, \quad (3.3)$$

as $j \rightarrow \infty$, 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

$$\liminf_{j \rightarrow \infty} \tilde{\tau}_{n(j)}^k \geq \tilde{\tau}^k \quad (a.s.). \quad (3.4)$$

Next define

$$\tilde{\mathcal{F}}_t^j := \sigma(\tilde{x}_{n(j)}^k(s), \tilde{w}_j(s) : s \leq t), \quad \tilde{\mathcal{F}}_t := \sigma(\tilde{x}^k(s), \tilde{w}(s) : s \leq t).$$

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

$$\tilde{x}_{n(j)}^k(t) = \tilde{x}_{n(j)}^k(0) + \int_0^t b(s, \tilde{x}_{n(j)}^k(\kappa_{n(j)}(s))) ds + \int_0^t \sigma(s, \tilde{x}_{n(j)}^k(\kappa_{n(j)}(s))) d\tilde{w}_j(s), \quad (3.5)$$

almost surely. Now we make use of the following lemma which is just an adaptation of a result of Skorokhod [21].

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

$$\begin{aligned} \int_0^t f(s, \tilde{x}_{n(j)}^k(s)) ds &\rightarrow \int_0^t f(s, \tilde{x}^k(s)) ds, \\ \int_0^t f(s, \tilde{x}_{n(j)}^k(\kappa_{n(j)}(s))) ds &\rightarrow \int_0^t f(s, \tilde{x}^k(s)) ds, \\ \int_0^t f(s, \tilde{x}_{n(j)}^k(s)) d\tilde{w}_j^i(s) &\rightarrow \int_0^t f(s, \tilde{x}^k(s)) d\tilde{w}^i(s), \\ \int_0^t f(s, \tilde{x}_{n(j)}^k(\kappa_{n(j)}(s))) d\tilde{w}_j^i(s) &\rightarrow \int_0^t f(s, \tilde{x}^k(s)) d\tilde{w}^i(s) \end{aligned} \quad (3.6)$$

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_n^k have the same distributions as $\tilde{\tau}_n^k$,

$$\begin{aligned} \lim_{k \rightarrow \infty} \limsup_{j \rightarrow \infty} P(\tau_{n(j)}^k \leq T) &= \lim_{k \rightarrow \infty} \limsup_{j \rightarrow \infty} P(\tilde{\tau}_{n(j)}^k \leq T) \\ &\leq \lim_{k \rightarrow \infty} P(\tilde{\tau}^k \leq T) = 0. \end{aligned}$$

The 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 tight.

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 $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P})$, carrying sequences of continuous processes $\hat{x}_{l(j)}, \bar{x}_{m(j)}, \hat{w}_j$, such that for every positive integer j the finite dimensional distributions of

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

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

$$\begin{aligned} \sup_{t \leq T} |\hat{x}_{l(j)}(t) - \hat{x}(t)| &\rightarrow 0, & \sup_{t \leq T} |\bar{x}_{l(j)}(t) - \bar{x}(t)| &\rightarrow 0, \\ \sup_{t \leq T} |\hat{w}_j(t) - \hat{w}(t)| &\rightarrow 0, \end{aligned}$$

as $j \rightarrow \infty$ 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 $[0, \tilde{\tau}^k)$ 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 $\hat{\tau}^k, \bar{\tau}^k \rightarrow \infty$, 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 2.1.

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 the 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 2.2.

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 $KtI \geq a = a^* \geq \varepsilon tI$. Also let $g(x)$ be a real-valued function such that $|g(x) - g(y)| \leq K|x - y|^\alpha$ for all x, y . Let ξ and η be independent d -dimensional Gaussian vectors with zero means. Assume $\xi \sim \mathcal{N}(0, I)$ and $\kappa_i \leq K\kappa_j$ for $i, j = 1, \dots, d$, where the κ_i 's are the eigenvalues of the covariance matrix of η . 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*

$$\left| g(x)E[D_{ij}Tf](x + \eta) - E[D_{ij}T(gf)](x + \eta) \right| \leq Nt^{-d/(2p)-1+\alpha/2}\|f\|_p, \quad (4.1)$$

$$\left\{ \int_{\mathbb{R}^d} |g(x)E[D_{ij}Tf](x + \eta) - E[D_{ij}T(gf)](x + \eta)|^p dx \right\}^{1/p} \leq Nt^{-1+\alpha/2}\|f\|_p, \quad (4.2)$$

where the constants N depend only on K, ε, d, p .

Proof First observe that for our matrix-valued a with $c := (2\pi)^{-d/2}$ and a_1 being the covariance matrix of η we have

$$Tf(x) = c \int_{\mathbb{R}^d} (\det a(y))^{-1/2} f(y) \exp\{-(a^{-1}(y)(y-x), y-x)/2\} dy, \quad (4.3)$$

$$\begin{aligned} E[D_{ij}Tf](x+\eta) &= D_{ij}ETf(x+\eta), \quad E[D_{ij}T(gf)](x+\eta) = D_{ij}ET(gf)(x+\eta), \\ & cE(\det a)^{-1/2} \exp\{-(a^{-1}(y-x-\eta), y-x-\eta)/2\} \\ &= c(\det(a+a_1))^{-1/2} \exp\{-((a+a_1)^{-1}(y-x), y-x)/2\} =: p_a(x, y). \end{aligned}$$

Note also that

$$\begin{aligned} KtI + a_1 &\geq a(y) + a_1 \geq \varepsilon tI + a_1 \geq (\varepsilon/K)(KtI + a_1), \\ (KtI + a_1)^{-1} &\leq (a(y) + a_1)^{-1} \leq (\varepsilon tI + a_1)^{-1} \leq (K/\varepsilon)(KtI + a_1)^{-1}, \\ \det(a(y) + a_1) &\geq \det(\varepsilon tI + a_1) \geq (\varepsilon/K)^d \det(KtI + a_1). \end{aligned}$$

It follows that $p_a(x, y) \leq r(x - y)$, where

$$r(z) := c(K/\varepsilon)^{d/2} (\det(KtI + a_1))^{-1/2} \exp \left\{ -((KtI + a_1)^{-1}z, z)/2 \right\}.$$

Next, let $A(y) = (a(y) + a_1)^{-1}$, then

$$\begin{aligned} &ED_{ij}Tf(x + \eta) \\ &= \int_{\mathbb{R}^d} f(y) [(A(y)(y - x))_i (A(y)(y - x))_j - A_{ij}(y)] p_{a(y)}(x, y) dy. \end{aligned}$$

This allows us to deal with (we adopt the summation convention over repeated indices)

$$I_\lambda(x) := \lambda^i \lambda^j \left(g(x) E[D_{ij}Tf](x + \eta) - E[D_{ij}T(gf)](x + \eta) \right),$$

where λ is a fixed vector in \mathbb{R}^d . By the above

$$I_\lambda(x) = \int_{\mathbb{R}^d} [g(x) - g(y)] f(y) [(\lambda, A(y)(y - x))^2 - \lambda^i \lambda^j A_{ij}(y)] p_{a(y)}(x, y) dy.$$

By ordering the eigenvalues of a_1 as $\kappa_1 \leq \dots \leq \kappa_d$ we have

$$\begin{aligned} 0 &\leq \lambda^i \lambda^j A_{ij}(y) \leq (\varepsilon t + \kappa_1)^{-1} |\lambda|^2, \\ |(\lambda, A(y)z)| &\leq (\lambda, A(y)\lambda)^{1/2} (z, A(y)z)^{1/2} \\ &\leq (\varepsilon t + \kappa_1)^{-1/2} |\lambda| (K/\varepsilon)^{1/2} (z, (KtI + a_1)^{-1}z)^{1/2}. \end{aligned}$$

Using this and making the change of variables $x - y = \sqrt{KtI + a_1}z$ we find that $|I_\lambda(x)|$ is less than or equal to

$$\begin{aligned} &N(\varepsilon t + \kappa_1)^{-1} |\lambda|^2 \int_{\mathbb{R}^d} |f(y)| |x - y|^\alpha \left[(x - y, (KtI + a_1)^{-1}(x - y)) + 1 \right] r(x - y) dy \\ &= N(\varepsilon t + \kappa_1)^{-1} |\lambda|^2 \int_{\mathbb{R}^d} |\sqrt{KtI + a_1}z|^\alpha |f(x - \sqrt{KtI + a_1}z)| [|z|^2 + 1] e^{-|z|^2/2} dz \\ &\leq N(\varepsilon t + \kappa_1)^{\alpha/2-1} |\lambda|^2 \int_{\mathbb{R}^d} |f(x - \sqrt{KtI + a_1}z)| |z|^\alpha [|z|^2 + 1] e^{-|z|^2/2} dz. \end{aligned}$$

The arbitrariness of λ implies that

$$\begin{aligned} &\left| g(x) E[D_{ij}Tf](x + \eta) - E[D_{ij}T(gf)](x + \eta) \right| \\ &\leq N(\varepsilon t + \kappa_1)^{\alpha/2-1} \int_{\mathbb{R}^d} |f(x - \sqrt{KtI + a_1}z)| |z|^\alpha [|z|^2 + 1] e^{-|z|^2/2} dz \\ &\leq N(\varepsilon t + \kappa_1)^{\alpha/2-d/(2p)-1} \|f\|_p \leq Nt^{-d/(2p)-1+\alpha/2} \|f\|_p. \end{aligned}$$

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

$$x_n(t) = x_0 + \int_0^t \sigma(s, x_n(\kappa_n(s))) dw(s), \quad (4.4)$$

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

$$\varepsilon I \leq (\sigma \sigma^*)(s, x) \leq KI, \quad |\sigma(s, x) - \sigma(s, y)| \leq K|x - y|^\alpha \quad (4.5)$$

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 notation. 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 4.1 (a) *There exists a constant N_0 depending only on $d, \alpha, K, \varepsilon, q$ such that if $1 \leq q < \frac{d}{d-\alpha}$, then for all $t > 0$, $n = 1, 2, 3, \dots$*

$$\left(\int_{\mathbb{R}^d} p_n^q(t, x) dx \right)^{1/q} \leq N_0(t^{-d/(2p)} + 1) \quad (p = q/(q-1)). \quad (4.6)$$

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

$$m_n(t) \leq N_0(t^{-d/2} + 1) \quad (4.7)$$

for any $t > 0$, $n = 1, 2, 3, \dots$ and (4.6) holds for any $q \in [1, \infty]$, $t > 0$, $n = 1, 2, 3, \dots$

Proof The last assertion in (b) is true since $p_n^q \leq p_n(m_n)^{q-1}$ and $\int p_n dx = 1$. To prove (a) for $0 \leq s \leq t < \infty$ 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 t the function $T_{s,t} f(x)$ is infinitely differentiable for $s < t$ and

$$\frac{\partial}{\partial s} T_{s,t} f(x) = -D_{ij} T_{s,t} a^{ij}(s, \cdot) f(\cdot)(x), \quad (4.8)$$

where $a_{ij} := \frac{1}{2}(\sigma\sigma^*)_{ij}$. For the sake of simplicity of notations we drop the subscripts n , and from (4.8) by the Newton-Leibnitz and Itô's formulas for any $r \in [0, t]$ we obtain

$$Ef(x(t)) = \int_r^t \frac{d}{ds} ET_{s,t}f(x(s)) ds + ET_{r,t}f(x(r)) = ET_{r,t}f(x(r)) \\ + \int_r^t E \left[a^{ij}(s, x(\kappa(s))) D_{ij} T_{s,t} f(x(s)) - D_{ij} T_{s,t} a^{ij}(s, \cdot) f(\cdot)(x(s)) \right] ds.$$

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

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

we get

$$Ef(x(t)) = ET_{r,t}f(x(r)) + \int_r^t EH(s, t, x(\kappa(s))) ds, \quad (4.9)$$

where

$$H(s, t, x) = a_{ij}(s, x) E[D_{ij} T_{s,t} f](x + \eta(s, x)) \\ - E[D_{ij} T_{s,t} a_{ij}(s, \cdot) f(\cdot)](x + \eta(s, x)).$$

Note that by Lemma 4.1

$$|H(s, t, x)| \leq N(t-s)^{-d/(2p)-1+\alpha/2} \|f\|_p, \\ \int_{\mathbb{R}^d} |H(s, t, x)| dx \leq N(t-s)^{-1+\alpha/2} \|f\|_1. \quad (4.10)$$

This and (4.9) with $r = 0$ give us (4.6) for $p > d/\alpha$ and for $t \in (0, T]$ with a constant N_0 depending only on $d, \alpha, K, \varepsilon, q$ and T . Indeed (cf. (4.3)),

$$T_{0,t}f(x_0) \leq Nt^{-d/2} \int_{\mathbb{R}^d} f(y) \exp\left\{-\frac{1}{Nt}(x-y)^2\right\} dy \leq Nt^{-d/(2p)} \|f\|_p, \\ \int_0^t (t-s)^{-d/(2p)-1+\alpha/2} ds = Nt^{-d/(2p)+\alpha/2}.$$

To prove (4.6) and (4.7) with a constant N_0 independent of T we need a longer argument. Fix a $T \in (0, \infty)$, and define γ_T as the smallest number γ such that $m(s) \leq \gamma(s^{-d/2} + 1)$ 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 [20] for the same purposes. Such a number γ_T does exist since $m(t)$ is bounded on $[t_1^n, T]$ and $m(t) \leq N(d, K, \varepsilon)t^{-d/2}$ 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 . We use (4.9) with $r = t_1^n$ and $t \in [t_1^n, T]$ and observe that the first term on the right can be easily estimated if we take into account (4.3) and use that the convolution of

Gaussian densities is again Gaussian. We also use (4.10) and the inequality $\kappa(s) \geq \varepsilon s$ ($s \geq t_1^n$) and we obtain for $t \in [t_1^n, T]$

$$\begin{aligned} Ef(x(t)) &\leq Nt^{-d/2} \|f\|_1 + \int_{t_1^n}^t \left[\gamma_T \left(\frac{1}{\kappa^{d/2}(s)} + 1 \right) \|H(s, t, \cdot)\|_1 \right] \wedge \sup_x |H(s, t, x)| ds \\ &\leq \left\{ Nt^{-d/2} + N \int_{t_1^n}^t \left[\gamma_T \left(\frac{1}{\kappa^{d/2}(s)} + 1 \right) \frac{1}{(t-s)^{1-\alpha/2}} \right] \wedge \frac{1}{(t-s)^{d/2+1-\alpha/2}} ds \right\} \|f\|_1, \\ m(t) &\leq Nt^{-d/2} + N \int_0^t \left[\gamma_T \left(\frac{1}{s^{d/2}} + 1 \right) \frac{1}{(t-s)^{1-\alpha/2}} \right] \wedge \frac{1}{(t-s)^{d/2+1-\alpha/2}} ds. \end{aligned} \quad (4.11)$$

Next, as is easy to see after the substitution $s = u\gamma_T^{-2/d}$,

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

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

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

Thus from (4.11) for $t \in [t_1^n, T]$ we conclude

$$m(t) \leq N(1 + \gamma_T^{1-\alpha/d})(t^{-d/2} + 1). \quad (4.12)$$

As we observed above this estimate is also true for $t \in (0, t_1^n]$. By definition of γ_T estimate (4.12) means that

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

We emphasize that the last constant N , as well as all constants called N in the above proof of (4.7), 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 (4.6) can be taken to be the same for all $t > 0$. The theorem is proved.

Corollary 4.1 *Assume the conditions of Theorem 2.2. Let $x_n(t)$ be the Euler approximation defined by (1.2) and let τ_n^k be the first exit time of $x_n(t)$ from D_k . Then for every $t > 0$ the measure $P(x_n(t) \in \Gamma, t < \tau_n^k)$ has a density $p_n^k(t, x)$, and for any $0 < t_0 < T < \infty$, $1 \leq q < \frac{d}{d-\alpha}$ and $k = 1, 2, \dots$ we have*

$$\sup_n \sup_{t \in [t_0, T]} \int_{\mathbb{R}^d} [p_n^k(t, x)]^q dx < \infty. \quad (4.13)$$

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

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

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 (4.6) if we notice the obvious relation between Euler's approximations for fixed initial value and for random one. \square

Remark 4.1 One knows from Fabes and Kenig [4] and Safonov [19] that none of the estimates (4.6), (4.7) and (4.13) remains valid if the Hölder continuity of σ in x is replaced by the assumption of uniform continuity of σ in (t, x) ,

Remark 4.2 We derived Theorem 4.1 for approximations starting at time zero at a fixed point. Obviously, the approximations can start at any t_k^n and then we get a "conditional" estimate

$$E\{f(x_n(t) \mid \mathcal{F}_{t_k^n}\} \leq N_0((t - t_k^n)^{-d/(2p)} + 1)\|f\|_p, \quad (4.14)$$

with the same N_0 as in (4.6), whenever $p > d/\alpha$, $k = 0, 1, 2, \dots$, $n = 1, 2, \dots$, $t > t_k^n$, and f is Borel.

Theorem 4.2 *Let $x_n(t)$ be the Euler approximation defined by (4.4), $q \geq 1$, $p > d/\alpha$. Assume that condition (4.5) is satisfied and (2.9), (2.8) hold. Then for any $T \in [0, \infty)$ and Borel functions f on $\mathbb{R}_+ \times \mathbb{R}^d$ we have*

$$E \exp \left(\int_0^T f(r, x_n(\kappa_n(r))) dr \right) \leq 2 \exp \left(N (\|f\|_{p,q,T}^q + d_n^{q-1}(T) \|f\|_{\infty,q,T}^q) \right), \quad (4.15)$$

where N depends only on $d, \alpha, K, \varepsilon, p, q, \delta(T), T$.

Proof We may assume that $f \geq 0$ and f is bounded. Then fix T, n and for $t \leq T$ introduce

$$\psi(r) = f(r, x_n(\kappa_n(r))), \quad \phi(t) = \int_t^T \psi(r) dr, \quad \Phi(t) = \text{esssup } E\{e^{\phi(t)} \mid \mathcal{F}_{\kappa_n(t)}\}.$$

Observe that

$$e^{\phi(t)} = 1 + \int_t^T \psi(r) e^{\phi(r)} dr.$$

Hence, by taking into account that $\psi(r)$ is $\mathcal{F}_{\kappa_n(r)}$ -measurable, we get

$$\begin{aligned} E\{e^{\phi(t)} \mid \mathcal{F}_{\kappa_n(t)}\} &= 1 + \int_t^T E\{\psi(r) e^{\phi(r)} \mid \mathcal{F}_{\kappa_n(t)}\} dr \\ &= 1 + \int_t^T E\{\psi(r) E\{e^{\phi(r)} \mid \mathcal{F}_{\kappa_n(r)}\} \mid \mathcal{F}_{\kappa_n(t)}\} dr \end{aligned}$$

$$\begin{aligned}
&\leq 1 + \int_t^T \Phi(r) E\{\psi(r) \mid \mathcal{F}_{\kappa_n(t)}\} dr \\
&\leq 1 + \int_t^{\bar{\kappa}_n(t) \wedge T} \Phi(r) \psi(r) dr + \int_{\bar{\kappa}_n(t) \wedge T}^T \Phi(r) E\{\psi(r) \mid \mathcal{F}_{\kappa_n(t)}\} dr, \quad (4.16)
\end{aligned}$$

where $\bar{\kappa}_n(r)$ is defined to be equal to t_{k+1}^n if $\kappa_n(r) = t_k^n$. Note that for any Borel $h = h(r) \geq 0$ and $t \leq r \leq T$ by Hölder's inequality (observe that, if $r < t_1^n$, $x_n(\kappa_n(r))$ may not have density and this is why we use sup in notation (2.6))

$$\int_r^{\bar{\kappa}_n(r)} h(s) \psi(s) ds \leq (\bar{\kappa}_n(r) - r)^{(q-1)/q} \left(\int_r^{\bar{\kappa}_n(r)} h^q(s) \|f(s, \cdot)\|_\infty^q ds \right)^{1/q}. \quad (4.17)$$

Also, in light of Remark 4.2 and the fact that $\kappa_n(r) - \kappa_n(t) \geq \delta(T)(r - t)/2$ for $r \geq \bar{\kappa}_n(t)$, the last term in (4.16) is dominated by

$$\begin{aligned}
&N_0 2^{d/(2p)} \delta^{-d/(2p)}(T) \int_t^T \Phi(r) \|f(r, \cdot)\|_p \left((r - t)^{-d/(2p)} + 1 \right) dr \\
&\leq N \left(\int_t^T \Phi^q(r) \|f(r, \cdot)\|_p^q dr \right)^{1/q},
\end{aligned}$$

where the last N is finite in light of (2.8). By using this and (4.17) we conclude

$$\Phi(t) \leq 1 + N \left(\int_t^T \Phi^q(r) \xi(r) dr \right)^{1/q}, \quad \Phi^q(t) \leq 2 + N \int_t^T \Phi^q(r) \xi(r) dr,$$

where $\xi(r) = \|f(r, \cdot)\|_p^q + d_n^{q-1}(T) \|f(r, \cdot)\|_\infty^q$. Gronwall's inequality yields

$$\Phi^q(t) \leq 2 \exp \left(N \int_t^T \xi(r) dr \right),$$

which for $t = 0$ implies (4.15) and proves the theorem.

For fixed $T > 0$ and $n \geq 1$ we denote by $\gamma_n(T)$ the Girsanov exponent

$$\begin{aligned}
\gamma_n(T) &= \exp \left(- \int_0^T (\sigma^{-1} b_n)(s, x_n(\kappa_n(s))) dw(s) \right. \\
&\quad \left. - \frac{1}{2} \int_0^T |(\sigma^{-1} b_n)(s, x_n(\kappa_n(s)))|^2 ds \right),
\end{aligned}$$

where $x_n(t)$ is the Euler approximation defined by (4.4) and σ^{-1} is the right inverse of σ ($\sigma^{-1} = \sigma^*(\sigma\sigma^*)^{-1}$). Let \tilde{P} denote the probability measure defined by $d\tilde{P}/dP = \gamma_n(T)$ and use the notation \tilde{E} for the expectation under \tilde{P} .

Proposition 4.1 *Let $x_n(t)$ be the Euler approximation defined by (2.5), $q \geq 1$, $p > d/\alpha$. Assume that conditions (4.5), (2.9), (2.8) hold. Then for any $\rho \in \mathbb{R}$*

$$\tilde{E}\gamma_n^\rho(T) \leq 2 \exp\left(N(\|b_n\|_{2p,2q,T}^{2q} + d_n^{q-1}(T)\|b_n\|_{\infty,2q,T}^{2q})\right),$$

where the constant N depends only on $q, p, T, d, K, \alpha, \varepsilon, \delta(T)$ and ρ .

Proof We may assume that $x_n(0) = \xi$ is non random. Notice that

$$dx_n(t) = \sigma(t, x_n(\kappa_n(t)))d\tilde{w}(t)$$

with

$$\tilde{w}(t) = \int_0^t (\sigma^{-1}b_n)(s, x_n(\kappa_n(s)))ds + w(t), \quad t \in [0, T],$$

which is a Wiener process under \tilde{P} by Girsanov's theorem. Thus setting $h_s = (\sigma^{-1}b_n)(s, x_n(\kappa_n(s)))$, by simple calculations we obtain

$$\tilde{E}\gamma_n^\rho(T) = \tilde{E}H \exp\left(-\rho \int_0^T h_s d\tilde{w}_s - \rho^2 \int_0^T |h_s|^2 ds\right)$$

with

$$H = \exp\left(\left(\frac{\rho}{2} + \rho^2\right) \int_0^T |h_s|^2 ds\right).$$

Hence by Cauchy-Bunjakovski-Schwarz inequality and using Theorem 4.2 we obtain

$$\begin{aligned} \tilde{E}\gamma_n^\rho(T) &\leq (\tilde{E}H^2)^{1/2} \leq \left(\tilde{E} \exp\left(|\rho + 2\rho^2| \int_0^T |b(s, x_n(\kappa_n(s)))|^2 ds\right)\right)^{1/2} \\ &\leq \sqrt{2} \exp\left(N(\|b_n\|_{2p,2q,T}^{2q} + d_n^{q-1}(T)\|b_n\|_{\infty,2q,T}^{2q})\right) \end{aligned}$$

with N depending only on $q, p, T, d, K, \alpha, \varepsilon, \delta(T)$ and ρ .

Remark 4.3 Observe that, if Assumption 2.1 is satisfied, then for any ρ and T the sequence $\tilde{E}\gamma_n^\rho(T)$ is bounded.

Proposition 4.2 *Suppose that Assumption 2.1 is satisfied and let $x_n(t)$ be the Euler approximation defined by (2.5). Then for any $\gamma < (q-1)/q$, T, n and bounded Borel functions $f \geq 0$ on $[0, \infty) \times \mathbb{R}^d$ we have*

$$E \int_0^T f(t, x_n(\kappa_n(t))) dt \leq N(\|f\|_{p,q,T} + d_n^\gamma(T)\|f\|_{\infty,q,T}) \quad (4.18)$$

with a constant N depending only on $q, p, \gamma, T, d, K, \alpha, \delta(T), B(T)$ and ε .

Proof First assume $b_n = 0$. Clearly,

$$E \int_0^T f(t, x_n(\kappa_n(t))) dt = E \int_0^{t_1^n \wedge T} f(t, x_n(\kappa_n(t))) dt + E \int_{t_1^n \wedge T}^T f(t, x_n(\kappa_n(t))) dt,$$

where the first term can be estimated from above by $d_n^{(q-1)/q}(T) \|f\|_{\infty, q, T}$. For the second term by Theorem 4.1 and taking into account that $\kappa_n(t) \geq \delta(T)t/2$ for $t \geq t_1^n$, we have

$$\begin{aligned} E \int_{t_1^n \wedge T}^T f(t, x_n(\kappa_n(t))) dt &\leq N_0 \int_{t_1^n \wedge T}^T (\kappa_n^{-d/2p}(t) + 1) \|f(t)\|_p dt \\ &\leq 2^{d/(2p)} \delta^{-d/(2p)}(T) N_0 \int_0^T (t^{-d/2p} + 1) \|f(t)\|_p dt. \end{aligned}$$

Hence by Hölder's inequality we get (4.18). In the general case we use Girsanov's theorem, Hölder's inequality and Proposition 4.1 to get

$$\begin{aligned} E \int_0^T f(t, x_n(\kappa_n(t))) dt &= (\tilde{E} \gamma_n^{-\rho/(\rho-1)}(T))^{1/\rho} \left(\tilde{E} \left(\int_0^T f(t, x_n(\kappa_n(t))) dt \right)^\rho \right)^{1/\rho} \\ &\leq N \left(\tilde{E} \int_0^T f^\rho(t, x_n(\kappa_n(t))) dt \right)^{1/\rho} \leq N' (\|f\|_{\rho p', \rho q', T} + d_n^{(q'-1)/(q'\rho)}(T) \|f\|_{\infty, \rho q', T}) \end{aligned}$$

for any $\rho > 1$, $q' > 1$, $p' > \frac{d}{2} \frac{q'}{q'-1} \vee \frac{d}{\alpha}$ which proves the proposition.

Proposition 4.3 *Under the assumptions of Proposition 4.2 suppose that the processes $x_n(t)$ converge to a process $x(t)$ in probability, uniformly in t in bounded intervals. Then for nonnegative Borel functions h on $[0, \infty) \times \mathbb{R}^d$ for each T we have*

$$E \int_0^T h(t, x(t)) dt \leq N \|h\|_{p, q, T} \quad (4.19)$$

for $q > 1$, $p > (\frac{d}{2} \frac{q}{q-1}) \vee \frac{d}{\alpha}$ with a constant depending on q , p , T , $\delta(T)$, d , K , α and ε .

Proof Letting $n \rightarrow \infty$ in (4.18) we get (4.19) when h is a continuous function with compact support. Hence the general case follows.

Proposition 4.4 *Under the assumptions of Proposition 4.2 for each $T > 0$ the sequence of processes $x_n(t)$ is tight in $C([0, T], \mathbb{R}^d)$.*

Proof By changing measure, then using Hölder's inequality and Proposition 4.1 we have

$$E |x_n(t) - x_n(s)|^4 \leq \left(\tilde{E} \gamma_n^2(T) \right)^{1/2} \left(\tilde{E} |x_n(t) - x_n(s)|^8 \right)^{1/2} \leq N |t - s|^2$$

for all $s, t \in [0, T]$, where the constant N is independent of s, t, n in light of Remark 4.3. This proves the proposition.

5 Proof of Theorem 2.2

The reader can easily check that we can repeat the proof of Theorem 2.1 given in Section 3, if we prove the following version of Lemma 3.1. We use the same notations as in Section 3.

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

Proof We start by proving the last relation in (3.6). Take a continuous in x Borel in t function $g(t, x)$ defined on $\mathbb{R}_+ \times \mathbb{R}^d$ satisfying the same hypotheses as f , and define

$$I_t^{kj}(f) = \int_0^t f(s, \tilde{x}_{n(j)}^k(\kappa_{n(j)}(s))) d\tilde{w}_j^i(s), \quad I_t^k(f) = \int_0^t f(s, \tilde{x}^k(s)) d\tilde{w}^i(s).$$

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

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

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

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

where $\gamma > 0$ and $\eta > 0$ are arbitrary numbers. By Corollary 4.1 we conclude that for p large enough

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

with N independent of g . Since $\tilde{x}_{n(j)}^k(t) \rightarrow \tilde{x}^k(t)$ (a.s.) from Corollary 4.1 we also get an estimate for probability density of $x^k(t)$, and this estimate shows that J_2 can be estimated 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.

The remaining relations in (3.6) are proved in like manner. For instance, to prove the second relation, introduce $I_t^{kj}(f)$ and $I_t^k(f)$ as above but replace

$\tilde{w}_j^i(s)$ and $w^i(s)$ with s . Then use Chebyshev's inequality instead of martingale ones to estimate J_1 through the first power of $|f - g|$ and repeat the rest of the above argument. This brings to an end the proofs of Lemma 5.1 and Theorem 2.2.

6 Proof of Theorem 2.3

To prove Theorem 2.3 notice that by Proposition 4.4 the sequence of Euler approximations $(x_n)_{n=1}^\infty$ is tight in $C([0, T], \mathbb{R}^d)$. Thus to prove the theorem we need only prove the following lemma.

Lemma 6.1 *Let Assumption 2.1 hold. Assume there exists a process $x(t)$ such that for each $T > 0$*

$$\lim_{n \rightarrow \infty} P\left(\sup_{t \in [0, T]} |x_n(t) - x(t)| \geq \varepsilon\right) = 0 \quad \text{for each } \varepsilon > 0.$$

Then for each $T > 0$

$$\lim_{n \rightarrow \infty} E \int_0^T |b_n(t, x_n(\kappa_n(t))) - b(t, x(t))|^2 dt = 0.$$

Proof To ease the notation we write $y_n(t) = x_n(\kappa_n(t))$. For $m = 1, 2, \dots$ introduce

$$b_n^m = b_n m / (m + |b_n|), \quad b^m = b m / (m + |b|).$$

Note that $b_n \rightarrow b$ in measure, hence, $b_n^m \rightarrow b^m$ in measure for each m , and since all these functions are uniformly integrable (in $L_{2p, 2q, T}$ -sense), $b_n^m \rightarrow b^m$ in $L_{2p, 2q, T}$ for each m . Observe also that for γ , which is strictly less than $(q - 1)/q$ but larger than the one in Assumption 2.1 (3), we have

$$\lim_{n \rightarrow \infty} d_n^\gamma(T) \| (b_n - b_n^m)^2 \|_{\infty, q, T} \leq 4 \lim_{n \rightarrow \infty} d_n^\gamma(T) \| b_n^2 \|_{\infty, q, T} = 0.$$

Hence and by Proposition 4.2

$$\begin{aligned} & \limsup_{n \rightarrow \infty} E \int_0^T |b_n(t, y_n(t)) - b_n^m(t, y_n(t))|^2 dt \\ & \leq N \lim_{n \rightarrow \infty} (\| |b_n - b_n^m|^2 \|_{p, q, T} + d_n^\gamma(T) \| |b_n - b_n^m|^2 \|_{\infty, q, T}) = N \| |b - b^m|^2 \|_{p, q, T}, \end{aligned}$$

which can be made arbitrarily small if we choose m large enough. It follows from here and Proposition 4.3 that to prove the lemma it suffices to prove that for each m

$$\lim_{n \rightarrow \infty} E \int_0^T |b_n^m(t, y_n(t)) - b^m(t, x(t))|^2 dt = 0.$$

Furthermore, again by Proposition 4.2

$$\lim_{n \rightarrow \infty} E \int_0^T |b_n^m(t, y_n(t)) - b^m(t, y_n(t))|^2 dt$$

$$\begin{aligned}
&\leq N \lim_{n \rightarrow \infty} (\|b_n^m - b^m\|_{p,q,T}^2 + d_n^\gamma(T) \|b_n^m - b^m\|_{\infty,q,T}^2) \\
&\leq N \lim_{n \rightarrow \infty} 4m^2 T^{1/q} d_n^\gamma(T) = 0,
\end{aligned} \tag{6.1}$$

which reduces the proof to showing that for each m

$$I := \lim_{n \rightarrow \infty} E \int_0^T |b^m(t, y_n(t)) - b^m(t, x(t))|^2 dt = 0.$$

Observe that for any continuous bounded \mathbb{R}^d -valued $\bar{b}(t, x)$ we obviously have

$$I \leq 9 \lim_{n \rightarrow \infty} E \int_0^T |b^m(t, y_n(t)) - \bar{b}(t, y_n(t))|^2 + 9E \int_0^T |b^m(t, x(t)) - \bar{b}(t, x(t))|^2$$

Here the first term on the right is dominated by

$$N \| |b^m - \bar{b}|^2 \|_{p,q,T}, \tag{6.2}$$

which is proved similarly to (6.1), and the second term is dominated by the same expression in light of Proposition 4.3. After that it only remains to notice that (6.2) can be made as small as we wish for an appropriate \bar{b} since $C([0, T], C_0^\infty(\mathbb{R}^d))$ is dense in $L_{2p,2q}(T)$ in light of the condition $p, q < \infty$ (used for the first and the only time). The lemma is proved.

Acknowledgments. The authors are sincerely grateful to T. Shiga and to the anonymous referee of the original paper for drawing their attention to paper [8]. The main part of 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. The authors are also grateful to Martin Dieckmann who found errors in the original proof of Lemma 4.1 in 2013. Finally, we express our gratitude to the referee of this paper for pointing out quite a few glitches which resulted in correcting and improving its presentation.

Conflict of interest

The authors declare that they have no conflict of interest.

Data availability

The document contains no data.

References

1. Alyushina, L.A.: Euler polygonal lines for Itô's equations with monotone coefficients. *Theory Probab. Appl.* **32**, 340–346 (1987)
2. Dareiotis, K. and Gerencsér M.: On the regularisation of the noise for the Euler-Maruyama scheme with irregular drift. *Electron. J. Probab.* **25** (2020), article no. 82, 1–18
3. Dareiotis, K., Máté Gerencsér, and Lê K.: Quantifying a convergence theorem of Gyöngy and Krylov, arXiv:2101.12185v1
4. Fabes, E.B., Kenig, C.E.: Examples of singular parabolic measures and singular transition probability densities. *Duke Mathematical Journal* **48**, 848–856 (1981)
5. Gyöngy, I., Martínez T.: On stochastic differential equations with locally unbounded drift, *Czechoslovak Math. J.* (4) 51 (126) (2001) 763–783
6. Gyöngy, I., Nualart, D., Sanz–Solé, M.: Approximation and support theorems in modulus spaces. *Probab. Theory Related Fields* **101**, 495–509 (1995)
7. Ikeda, N., Watanabe, S.: *Stochastic differential equations and diffusion processes*. Amsterdam Oxford New York: North–Holland Publishing Company 1981
8. 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
9. Krylov, N.V.: On Itô's stochastic integral equations. *Theory Probab. Appl.* **14**, 330–336 (1969)
10. 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)
11. Krylov, N.V.: Extremal properties of solutions of stochastic equations. *Theory Probab. Appl.* **29**, 205–214 (1984)
12. Krylov, N.V.: *Controlled diffusion processes*. New York Heidelberg Berlin: Springer Verlag, 1980
13. Krylov, N.V.: On time inhomogeneous stochastic Itô equations with drift in L_{d+1} . *Ukrains'kyi Matematychnyi Zhurnal* **72**, 1232–1253 (2020)
14. Leobacher, G. and M. Szölgényi M.: A strong order 1/2 method for multidimensional SDEs with discontinuous drift. *Ann. Appl. Probab.* 27, no. 4, (2017), 2383–2418
15. Maruyama, G.: Continuous Markov processes and stochastic equations. *Rend. Circ. Mat. Palermo* **4**, 48–90 (1955)
16. Nakao, S.: On the pathwise uniqueness of solutions of one–dimensional stochastic differential equations. *Osaka J. Math.* **9**, 513–518 (1972)
17. Ngo, H.-L. and Taguchi, D: Strong rate of convergence for the Euler-Maruyama approximation of stochastic differential equations with irregular coefficients. *Math. Comp.* 85, no. 300, (2016), 1793–1819
18. Roekner, M. and Zhao, Guohuan: SDEs with critical time dependent drifts: strong solutions. arXiv:2103.05803
19. 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
20. Stroock, D.W., Varadhan, S.R.S.: *Multidimensional diffusion processes*. New York etc.: Springer Verlag, 1979
21. Skorokhod, A.V.: *Studies in the theory of random processes*. New York: Dover Publications, Inc., 1982
22. Veretennikov, A. Yu.: On strong solution and explicit formulas for solutions of stochastic integral equations. *Math. USSR Sb.* **39**, 387–403 (1981)
23. 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)
24. Zhang, Xicheng: Strong solutions of SDEs with singular drift and Sobolev diffusion coefficients. *Stochastic Processes and their Applications* 115 (2005) 1805–1818
25. Xie, Longjie and Zhang, Xicheng: Ergodicity of stochastic differential equations with jumps and singular coefficients. *Annales de l'Institut Henri Poincaré. Probabilités et Statistiques.* 56 (2020) 175–229
26. Zvonkin, A.K., Krylov, N.V.: On strong solutions of stochastic differential equations. *Sel. Math. Sov.* **1**, 19–61 (1981)