

## On the regularisation of the noise for the Euler-Maruyama scheme with irregular drift

Konstantinos Dareiotis\*      Máté Gerencsér†

### Abstract

The strong rate of convergence of the Euler-Maruyama scheme for nondegenerate SDEs with irregular drift coefficients is considered. In the case of  $\alpha$ -Hölder drift in the recent literature the rate  $\alpha/2$  was proved in many related situations. By exploiting the regularising effect of the noise more efficiently, we show that the rate is in fact arbitrarily close to  $1/2$  for all  $\alpha > 0$ . The result extends to Dini continuous coefficients, while in  $d = 1$  also to all bounded measurable coefficients.

**Keywords:** stochastic differential equations; Euler-Maruyama scheme; quadrature estimates.

**AMS MSC 2010:** 60H35; 60H10; 65C30.

Submitted to EJP on June 23, 2019, final version accepted on June 8, 2020.

Supersedes arXiv:1812.04583.

## 1 Introduction and main results

We consider stochastic differential equations (SDEs)

$$dX_t = b(X_t) dt + dW_t, \quad X_0 = x_0, \quad (1.1)$$

on a fixed time horizon  $[0, T]$ , driven by a  $d$ -dimensional Wiener process  $W = (W^1, \dots, W^d)$  on a filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, \mathbb{P})$  with  $(\mathcal{F}_t)_{t \in [0, T]}$  satisfying the usual conditions, where  $x_0$  is an  $\mathbb{R}^d$ -valued  $\mathcal{F}_0$ -measurable random variable with finite variance, and the drift coefficient  $b = (b^1, \dots, b^d)$  is a measurable function on  $\mathbb{R}^d$  with values in  $\mathbb{R}^d$ . It is well known that equation (1.1) has the remarkable property that even if  $b$  is only known to be bounded and measurable then a unique strong solution exists [Zvo74, Ver80], for further developments see among others [GM01, KR05, Dav07, FF11, Sha16].

It has recently been of interest to study such regularising effect of the noise in the discretisation of SDEs with irregular drift. The most common approximation method is the Euler-Maruyama scheme

$$dX_t^n = b(X_{k_n(t)}^n) dt + dW_t, \quad X_0^n = x_0^n, \quad (1.2)$$

---

\*Max Planck Institute for Mathematics in the Sciences, Germany.

E-mail: [konstantinos.dareiotis@mis.mpg.de](mailto:konstantinos.dareiotis@mis.mpg.de)

†IST Austria.

E-mail: [mate.gerencser@ist.ac.at](mailto:mate.gerencser@ist.ac.at)

where  $k_n(t) = \lfloor nt \rfloor / n$ . When  $b$  is Lipschitz continuous, then the analysis of the scheme is quite standard. Going beyond (locally) Lipschitz  $b$ , the first result goes back to [GK96], who established convergence in probability (without rate) of the Euler-Maruyama scheme. In the recently revived interest in quantifying the rate of convergence, there has been essentially two classes of approaches:

- (A) In [MP91, NT17a, NT17b, NT16, PT17, BS18, BHY18, HL18, MX18] mild assumptions on the modulus of continuity of  $b$  are imposed. This usually means  $\alpha$ -Hölder continuity (although [BHY18] also discusses the Dini continuous case). While  $\alpha > 0$  is allowed to be arbitrarily low, the drawback is that the convergence rates obtained become increasingly worse as  $\alpha \rightarrow 0$ . In [PT17] for example, whose setting is similar to ours, the rate of  $L_2$ -convergence is  $\alpha/2$ , which becomes negligible for small  $\alpha$ .
- (B) In [LS17, NSS18, LS18, LS15, MY18, MY19]  $b$  is allowed to have discontinuities on a ‘small’ (lower dimensional) set. In dimension  $d = 1$ , for example, the set of discontinuities is a discrete set of points. Outside of this exceptional set the usual Lipschitz condition is assumed. Under this condition, rate  $1/2$  is achieved for Euler-type schemes, and in the scalar case recently the improved rate  $3/4$  was shown [MY19] for a modified Milstein-type scheme.

Our contribution is twofold:

- Note that the results for (A) seem to suggest that the rate of convergence is arbitrarily small for  $\alpha$ -Hölder drift  $b$ , when  $\alpha$  is small. Here we show that this is *not* the case: in Theorem 1.1 we show that (almost)  $1/2$  rate of convergence holds for all Dini continuous coefficients.
- In dimension 1, we unify the type of irregularities considered (A) and (B) in the following sense: on one hand the irregularities do not have to be restricted on exceptional sets (as in (A)), on the other hand they are allowed to be discontinuous (as in (B)). More precisely, in Theorem 1.2 we prove that for any  $\varepsilon > 0$ , the Euler-Maruyama scheme has rate  $1/2 - \varepsilon$  in  $L_2$  for all bounded and measurable  $b$ . To the best of our knowledge this is the first result on the rate of convergence (and as far as  $L_2$  is concerned, even merely on convergence) without posing any continuity assumption whatsoever on  $b$ .

We now proceed by stating our main results. Recall the definition of Dini continuity: fix a continuous increasing function  $\vartheta : [0, 1] \rightarrow [0, \infty)$  such that

$$\int_0^1 \frac{\vartheta(r)}{r} dr < \infty. \quad (1.3)$$

We then denote by  $\mathcal{D}$  the space of all continuous functions  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ , such that

$$\|f\|_{\mathcal{D}} := \sup_{x \in \mathbb{R}^d} |f(x)| + \sup_{\substack{x, y \in \mathbb{R}^d \\ |x-y| \leq 1 \\ x \neq y}} \frac{|f(x) - f(y)|}{\vartheta(|x-y|)} < \infty.$$

For vector-valued  $f$ ,  $\|f\|_{\mathcal{D}}$  denotes  $\max_i \|f^i\|_{\mathcal{D}}$  (and similarly for other norms); since the context will always make it clear whether we mean scalar- or vector-valued functions, this abuse of notation will not cause any confusion. The main results then read as follows.

**Theorem 1.1.** Assume  $b \in \mathcal{D}$ . Let  $\varepsilon \in (0, 1)$  and suppose that for some constant  $C$  one has  $\mathbb{E}|x_0^n - x_0|^2 \leq Cn^{-1+\varepsilon}$  for all  $n \in \mathbb{N}$ . Then for all  $n \in \mathbb{N}$  one has the bound

$$\sup_{t \in [0, T]} \mathbb{E}|X_t^n - X_t|^2 \leq Nn^{-1+\varepsilon} \quad (1.4)$$

with some  $N = N(C, \varepsilon, \vartheta, T, d, \|b\|_{\mathcal{D}})$ .

Let us mention that the generality of allowing different initial condition  $x_0^n$  of the discretisation than that of (1.1) is a convenient setup for the proof, see Section 3. In the next theorem we consider the scalar case where even the assumption of Dini continuity can be lifted.

**Theorem 1.2.** Assume  $d = 1$ , and let  $b : \mathbb{R} \rightarrow \mathbb{R}$  be a bounded measurable function. Let  $\varepsilon \in (0, 1)$  and suppose that for some constant  $C$  one has  $\mathbb{E}|x_0^n - x_0|^2 \leq Cn^{-1+\varepsilon}$  for all  $n \in \mathbb{N}$ . Then for all  $n \in \mathbb{N}$  one has the bound

$$\sup_{t \in [0, T]} \mathbb{E}|X_t^n - X_t|^2 \leq Nn^{-1+\varepsilon} \quad (1.5)$$

with some  $N = N(C, \varepsilon, T, \sup |b|)$ .

**Remark 1.3.** The choice of looking at  $L_2$ -convergence is also motivated by recent results [HHJ15, JMY16, GJS17, Yar17] that show that in  $L_2$  very slow convergence rate (in particular, worse than any polynomial) of arbitrary approximation schemes may occur even in cases when the rate of convergence in probability is known to be  $1/2$  [Gyö98]. Our results can also be seen as extending the class of equations where the slow convergence phenomena can not happen.

**Remark 1.4.** A few months after the current work was available on the ArXiv, the article [NS19] also appeared on the ArXiv. In the one-dimensional setting the authors show an interesting complementary result to Theorem 1.2: if in addition to our assumption  $b$  possesses regularity of order  $\alpha \in (0, 1)$  in a Sobolev-Slobodecki scale, then the rate of convergence also improves. The technique in [NS19] is essentially the same as the one of the present work in  $d = 1$ : after a transformation by the scale function (3.8) a key estimate of the type (2.11) leads to the result; illustrating that, the method is reasonably robust. We would also like to thank the authors of [NS19] for bringing to our knowledge that a time-homogeneous version of the estimate (2.1) below has appeared in [Alt17].

## 2 Auxiliary results

### 2.1 Quadrature bounds

In this section we formulate some lemmata that will be used in the proofs of Theorems 1.1-1.2. In the proofs the constants  $N$  may change from line to line. Our first lemma is similar in spirit to [Dav07, Proposition 2.1], which is the “basic estimate” in Davie’s proof of the path-by-path uniqueness of the solution of (1.1). A time-homogeneous variant of the estimate below without stopping times and further random dependence also recently appeared in [Alt17], following a different proof. In the following,  $\mathcal{Y}$  is a Polish space.

**Lemma 2.1.** There exists a constant  $N = N(T, d)$  such that for all stopping times  $\tau \leq \tau' \leq T$ , all  $\mathcal{F}_\tau$ -measurable  $\mathcal{Y}$ -valued random variables  $Y$ , all bounded measurable functions  $f : [0, T] \times \mathbb{R}^d \times \mathcal{Y} \rightarrow \mathbb{R}$ , and all  $n \in \mathbb{N}$ , one has the bound

$$\mathbb{E} \left| \int_{\tau}^{\tau'} (f(s, W_s, Y) - f(s, W_{k_n(s)}, Y)) ds \right|^2 \leq N (\sup |f|^2) n^{-1} \log(n+1). \quad (2.1)$$

*Proof.* Introduce the notation  $W_{s,t} = W_t - W_s$ . We will first show that for  $f : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$  and for  $\alpha \in [0, T - 2/n]$  we have the bound

$$2I_n := \mathbb{E} \left| \int_{\alpha+1/n}^T (f(s, W_{\alpha,s}) - f(s, W_{\alpha,k_n(s)})) ds \right|^2 \leq N(\sup |f|^2) n^{-1} \log(n+1). \quad (2.2)$$

We make two simplifying assumptions that will be removed later on: (i)  $f(t, \cdot)$  is compactly supported and twice continuously differentiable for all  $t \in [0, T]$ ; (ii)  $d = 1$ . Notice that

$$2I_n = \mathbb{E} \left| \int_{\alpha+1/n}^T (f(s, W_{s-\alpha}) - f(s, W_{k_n(s)-\alpha})) ds \right|^2, \quad (2.3)$$

which gives,

$$\begin{aligned} I_n &= \mathbb{E} \int_{\alpha+1/n}^T \int_{\alpha+1/n}^t (f(s, W_{s-\alpha}) - f(s, W_{k_n(s)-\alpha})) (f(t, W_{t-\alpha}) - f(t, W_{k_n(t)-\alpha})) ds dt \\ &= \int_{\alpha+2/n}^T \int_{\alpha+1/n}^{k_n(t)-1/n} F_n(s, t) ds dt \\ &\quad + \int_{\alpha+2/n}^T \int_{k_n(t)-1/n}^t F_n(s, t) ds dt + \int_{\alpha+1/n}^{\alpha+2/n} \int_{\alpha+1/n}^t F_n(s, t) ds dt \\ &=: I_n^1 + I_n^2 + I_n^3, \end{aligned} \quad (2.4)$$

where we have used the notation

$$F_n(s, t) := \mathbb{E} \left[ (f(s, W_{s-\alpha}) - f(s, W_{k_n(s)-\alpha})) (f(t, W_{t-\alpha}) - f(t, W_{k_n(t)-\alpha})) \right].$$

We deal first with  $I_n^1$ . We will show that for  $(s, t)$  as in the integrand of  $I_n^1$  we have the bound

$$|F_n(s, t)| \leq N n^{-1} (\sup |f|^2) ((k_n(s) - \alpha)(k_n(t) - s))^{-1/2} + (k_n(t) - s)^{-1}. \quad (2.5)$$

By the fundamental theorem of calculus

$$\begin{aligned} F_n(s, t) &= \mathbb{E} \int_0^1 \int_0^1 (W_{s-\alpha} - W_{k_n(s)-\alpha}) f'(s, W_{s-\alpha} - \theta(W_{s-\alpha} - W_{k_n(s)-\alpha})) \\ &\quad \times (W_{t-\alpha} - W_{k_n(t)-\alpha}) f'(t, W_{t-\alpha} - \lambda(W_{t-\alpha} - W_{k_n(t)-\alpha})) ds dt d\theta d\lambda \end{aligned}$$

Notice that for  $s < k_n(t)$ , we have  $k_n(s) \leq s < k_n(t) \leq t$ . We can express the integrand in terms of random variables whose joint density is relatively simple: Setting

$$Y_1 := W_{k_n(s)-\alpha}, \quad Y_2 := W_{s-\alpha} - W_{k_n(s)-\alpha}, \quad Y_3 := W_{k_n(t)-\alpha} - W_{s-\alpha}, \quad Y_4 := W_{t-\alpha} - W_{k_n(t)-\alpha}$$

and introducing the notations, for  $x = (x_1, x_2, x_3, x_4) \in \mathbb{R}^4$ ,

$$\bar{f}(s, x) = f(s, (1 - \theta)x_2 + x_1), \quad \tilde{f}(t, x) = f(t, (1 - \lambda)x_4 + x_3 + x_2 + x_1),$$

we can write

$$\begin{aligned} f'(s, W_{s-\alpha} - \theta(W_{s-\alpha} - W_{k_n(s)-\alpha})) &= \bar{f}_{x_1}'(s, Y), \\ f'(t, W_{t-\alpha} - \lambda(W_{t-\alpha} - W_{k_n(t)-\alpha})) &= \tilde{f}_{x_3}'(s, Y). \end{aligned}$$

Moreover  $Y_i = Y_i(s, t)$  are independent, Gaussian, with mean 0 and variance  $\sigma_i^2 = \sigma_i^2(s, t)$ , where

$$\sigma_1^2(s, t) = k_n(s) - \alpha, \quad \sigma_2^2(s, t) = s - k_n(s), \quad \sigma_3^2(s, t) = k_n(t) - s, \quad \sigma_4^2(s, t) = t - k_n(t).$$

Their joint density is therefore given by

$$\rho(s, t, x) = \exp\left(-\sum_{i=1}^4 \frac{|x_i|^2}{2\sigma_i^2(s, t)}\right) \frac{1}{(2\pi)^2 \prod_{i=1}^4 \sigma_i(s, t)},$$

for  $x = (x_1, x_2, x_3, x_4) \in \mathbb{R}^4$ . Hence we can write

$$F_n(s, t) = \int_0^1 \int_0^1 \int_{\mathbb{R}^4} x_4 x_2 \bar{f}_{x_1}(s, x) \tilde{f}_{x_3}(t, x) \rho(s, t, x) dx d\theta d\lambda,$$

where

$$\bar{f}(s, x) = f(s, (1 - \theta)x_2 + x_1), \quad \tilde{f}(t, x) = f(t, (1 - \lambda)x_4 + x_3 + x_2 + x_1).$$

After integration by parts with respect to  $x_1$  and  $x_3$  we get

$$\begin{aligned} F_n(s, t) &= \int_0^1 \int_0^1 \int_{\mathbb{R}^4} x_4 x_2 \bar{f}(s, x) \tilde{f}(t, x) \rho_{x_1 x_3}(s, t, x) dx d\theta d\lambda \\ &\quad + \int_0^1 \int_0^1 \int_{\mathbb{R}^4} x_4 x_2 \bar{f}(s, x) \tilde{f}_{x_1}(t, x) \rho_{x_3}(s, t, x) dx d\theta d\lambda \\ &=: F_n^1(s, t) + F_n^2(s, t). \end{aligned}$$

By explicitly differentiating the Gaussian density  $\rho$ , we get

$$\begin{aligned} |F_n^1(s, t)| &= \left| \int_0^1 \int_0^1 \int_{\mathbb{R}^4} x_4 x_2 \bar{f}(s, x) \tilde{f}(t, x) \rho(s, t, x) x_1 x_3 \sigma_1^{-2}(s, t) \sigma_3^{-2}(s, t) dx d\theta d\lambda \right| \\ &\leq (\sup |f|^2) \left( \prod_{i=1}^4 \mathbb{E}|Y_i(s, t)| \right) \sigma_1^{-2}(s, t) \sigma_3^{-2}(s, t) \end{aligned}$$

Therefore, since  $\mathbb{E}|Y_i(s, t)| \leq \sigma_i(s, t)$  and  $\sigma_2(s, t), \sigma_4(s, t) \leq n^{-1/2}$ , we can write

$$|F_n^1(s, t)| \leq N n^{-1} K^2 ((k_n(s) - \alpha)(k_n(t) - s))^{-1/2}. \tag{2.6}$$

For  $F_n^2(s, t)$ , notice that  $\tilde{f}_{x_1} = \tilde{f}_{x_3}$ , so we can integrate by parts with respect to  $x_3$  again to get

$$\begin{aligned} |F_n^2(s, t)| &= \left| \int_0^1 \int_0^1 \int_{\mathbb{R}^4} x_4 x_2 \bar{f}(s, x) \tilde{f}(t, x) \rho_{x_3 x_3}(s, t, x) dx d\theta d\lambda \right| \\ &\leq \left| \int_0^1 \int_0^1 \int_{\mathbb{R}^4} x_4 x_2 \bar{f}(s, x) \tilde{f}(t, x) |x_3|^2 \sigma_3^{-4}(s, t) \rho(s, t, x) dx d\theta d\lambda \right| \\ &\quad + \left| \int_0^1 \int_0^1 \int_{\mathbb{R}^4} x_4 x_2 \bar{f}(s, x) \tilde{f}(t, x) \sigma_3^{-2}(s, t) \rho(s, t, x) dx d\theta d\lambda \right| \\ &\leq N n^{-1} (\sup |f|^2) (k_n(t) - s)^{-1}, \end{aligned}$$

which combined with (2.6) shows (2.5). Consequently, we have

$$I_n^1 \leq N n^{-1} K^2 \int_{\alpha+2/n}^T \int_{\alpha+1/n}^{k_n(t)-1/n} (((k_n(s) - \alpha)(k_n(t) - s))^{-1/2} + (k_n(t) - s)^{-1}) ds dt.$$

Note that one has  $k_n(s) - \alpha \geq s - \alpha - 1/n$  and that by a change of variables one sees

$$\int_{\alpha+1/n}^{k_n(t)} ((s - \alpha - n^{-1})(k_n(t) - s))^{-1/2} ds = \int_0^1 (s(1 - s))^{-1/2} ds = \pi.$$

Next, we see that

$$\int_{\alpha+2/n}^T \int_{\alpha+1/n}^{k_n(t)-1/n} (k_n(t) - s)^{-1} ds dt = \int_{\alpha+2/n}^T \log(k_n(t) - \alpha - n^{-1}) - \log(n^{-1}) dt \leq N \log(n + 1).$$

Therefore, we get

$$|I_n^1| \leq NK^2 n^{-1} \log(n + 1). \tag{2.7}$$

Both  $I_n^2$  and  $I_n^3$  are integrals over domains whose size is bounded by  $2Tn^{-1}$  with an integrand that is bounded by  $\sup |f|^2$ . Consequently,

$$|I_n^2| + |I_n^3| \leq Nn^{-1}(\sup |f|^2). \tag{2.8}$$

Combining (2.4) with (2.7)–(2.8) we obtain (2.2) under the additional assumptions (i)–(ii). We proceed by removing first the additional assumption (ii). Suppose now that we have (2.2) for  $d = k$  and let  $W = (W^1, \dots, W^{k+1})$  be a  $k + 1$ -dimensional Wiener process. Setting  $\tilde{W} = (W^2, \dots, W^{k+1})$ , we have

$$\begin{aligned} & \mathbb{E} \left| \int_{\alpha+1/n}^T (f(s, W_{\alpha,s}) - f(s, W_{\alpha,k_n(s)})) ds \right|^2 \\ &= \mathbb{E} \left| \int_{\alpha+1/n}^T (f(s, W_{\alpha,s}^1, \tilde{W}_{\alpha,s}) - f(s, W_{\alpha,k_n(s)}^1, \tilde{W}_{\alpha,k_n(s)})) ds \right|^2 \\ &\leq \mathbb{E} \left( \mathbb{E} \left[ \left| \int_{\alpha+1/n}^T (f(s, W_{\alpha,s}^1, \tilde{W}_{\alpha,s}) - f(s, W_{\alpha,s}^1, \tilde{W}_{\alpha,k_n(s)})) ds \right|^2 \middle| W^1 \right] \right) \\ &\quad + \mathbb{E} \left( \mathbb{E} \left[ \left| \int_{\alpha+1/n}^T (f(s, W_{\alpha,s}^1, \tilde{W}_{\alpha,k_n(s)}) - f(s, W_{\alpha,k_n(s)}^1, \tilde{W}_{\alpha,k_n(s)})) ds \right|^2 \middle| \tilde{W} \right] \right) \\ &\leq N(\sup |f|^2)n^{-1} \log(n + 1), \end{aligned}$$

where we have used the fact that  $W^1, \tilde{W}$  are independent and the induction hypothesis.

Removing the additional regularity assumption (i) on  $f$  follows from a standard approximation argument. For a bounded measurable  $f$  take a sequence  $(f_m)_{m \in \mathbb{N}}$  satisfying (i), such that for each  $t \in [0, T]$ ,  $f_m(t, \cdot) \rightarrow f(t, \cdot)$  as  $m \rightarrow \infty$  almost everywhere on  $\mathbb{R}^d$ , and  $\sup |f_m| \leq \sup |f|$  (one can construct such  $f_m$  by, for example, mollification). Since for positive times the law of the Brownian motion is absolutely continuous with respect to the Lebesgue measure, we have that for all  $s \in [\alpha + 2/n, T]$ ,  $f_m(s, W_s - W_\alpha) \rightarrow f(s, W_s - W_\alpha)$  and  $f_m(s, W_{k_n(s)} - W_\alpha) \rightarrow f(s, W_{k_n(s)} - W_\alpha)$  almost surely. Let  $\alpha \in [0, T - 2/n]$ . By the first part of the proof we have the inequality

$$\begin{aligned} \mathbb{E} \left| \int_{\alpha+2/n}^T (f_m(s, W_s - W_\alpha) - f_m(s, W_{k_n(s)} - W_\alpha)) ds \right|^2 &\leq N \sup |f_m|^2 n^{-1} \log(n + 1) \\ &\leq N \sup |f|^2 n^{-1} \log(n + 1), \end{aligned}$$

and by Fatou's lemma we get

$$\mathbb{E} \left| \int_{\alpha+2/n}^T (f(s, W_s - W_\alpha) - f(s, W_{k_n(s)} - W_\alpha)) ds \right|^2 \leq N(\sup |f|^2)n^{-1} \log(n + 1).$$

It then immediately follows that for any  $\alpha \in [0, T]$  we have the bound

$$\mathbb{E} \left| \int_\alpha^T (f(s, W_s - W_\alpha) - f(s, W_{k_n(s)} - W_\alpha)) ds \right|^2 \leq N(\sup |f|^2)n^{-1} \log(n + 1). \tag{2.9}$$

Now let us consider a simple stopping time  $\tau \leq T$ , that is, one that takes only finitely many values  $\{\alpha_1, \dots, \alpha_m\}$ . Take an  $\mathcal{F}_\tau$ -measurable random variable  $Y$ , and a measurable bounded function  $f : [0, T] \times \mathbb{R}^d \times \mathcal{Y} \rightarrow \mathbb{R}$ . We have

$$\begin{aligned} & \mathbb{E} \left| \int_\tau^T (f(s, W_s, Y) - f(s, W_{k_n(s)}, Y)) ds \right|^2 \\ &= \sum_{i=1}^m \mathbb{E} \left( \mathbf{1}_{\tau=\alpha_i} \mathbb{E} \left( \left| \int_{\alpha_i}^T (f(s, W_s, Y) - f(s, W_{k_n(s)}, Y)) ds \right|^2 \middle| \mathcal{F}_{\alpha_i} \right) \right). \end{aligned}$$

Since

$$\mathbf{1}_{\tau=\alpha_i} \mathbb{E} \left( \left| \int_{\alpha_i}^T (f(s, W_s, Y) - f(s, W_{k_n(s)}, Y)) ds \right|^2 \middle| \mathcal{F}_{\alpha_i} \right) = \mathbf{1}_{\tau=\alpha_i} g(W_{\alpha_i}, Y),$$

with

$$g(x, y) = \mathbb{E} \left| \int_\alpha^T (f(s, W_s - W_\alpha + x, y) - f(s, W_{k_n(s)} - W_\alpha + x, y)) ds \right|^2,$$

it follows from (2.9) that

$$\mathbb{E} \left| \int_\tau^T (f(s, W_s, Y) - f(s, W_{k_n(s)}, Y)) ds \right|^2 \leq N \sup |f|^2 n^{-1} \log(n+1). \quad (2.10)$$

If  $\tau \leq T$  is an arbitrary stopping time, one can find a sequence of simple stopping times  $\tau_m$  converging to  $\tau$  from above which by Fatou's lemma implies that (notice that if  $Y$  is  $\mathcal{F}_\tau$ -measurable then it is  $\mathcal{F}_{\tau_m}$ -measurable as well) (2.10) holds for arbitrary stopping times  $\tau \leq T$ . Finally, (2.1) follows from (2.10) since  $\int_\tau^{\tau'} = \int_\tau^T - \int_{\tau'}^T$ .  $\square$

**Lemma 2.2.** *Let  $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$  and  $f : [0, T] \times \mathbb{R}^d \times \mathcal{Y} \rightarrow \mathbb{R}$  be bounded measurable functions and define  $X^n$  as in (1.2). Let  $\tau \leq \tau' \leq T$  be stopping times and  $Y$  be a  $\mathcal{F}_\tau$ -measurable  $\mathcal{Y}$ -valued random variable. Then for all  $\varepsilon \in (0, 1)$  one has the bound*

$$\mathbb{E} \left| \int_\tau^{\tau'} (f(s, X_s^n, Y) - f(s, X_{k_n(s)}^n, Y)) ds \right|^2 \leq N n^{-1+\varepsilon}, \quad (2.11)$$

for all  $n \in \mathbb{N}$ ,  $r \in [0, T]$ , where  $N = N(\varepsilon, T, d, \sup |b|, \sup |f|)$ .

*Proof.* For an  $\mathbb{R}^d$ -valued stochastic process  $Z$  let us denote

$$h(Z) = \left| \int_\tau^{\tau'} (f(s, Z_s, Y) - f(s, Z_{k_n(s)}, Y)) ds \right|.$$

In other words, we aim to bound  $\mathbb{E}h(X^n)^2$ . Notice also that (2.1) provides a bound for  $\mathbb{E}h(W+x)^2$  whenever  $W$  is an  $(\mathcal{F}_t)_{t \in [0, T]}$ -Wiener process and  $x \in \mathbb{R}^d$ , as well as the fact that one has a trivial bound  $h \leq N$ . Let us set

$$\rho_n = \exp \left( - \int_0^T \sum_{i=1}^n b^i(X_{k_n(s)}^n) dW_s^i - \frac{1}{2} \int_0^T \sum_{i=1}^n |b^i(X_{k_n(s)}^n)|^2 ds \right).$$

By Girsanov's theorem, under the measure  $d\mathbb{P}^n = \rho_n d\mathbb{P}$ ,  $X^n - x$  is an  $(\mathcal{F}_t)_{t \in [0, T]}$ -Wiener process, therefore by the above properties and Hölder's inequality we have

$$\begin{aligned} \mathbb{E}h(X^n)^2 &\leq N \mathbb{E}(h(X^n)^{2-\varepsilon} \rho_n^{(2-\varepsilon)/2} \rho_n^{(\varepsilon-2)/2}) \\ &\leq N [\mathbb{E}(h(X^n)^2 \rho_n)]^{(2-\varepsilon)/2} [\mathbb{E} \rho_n^{(\varepsilon-2)/\varepsilon}]^{\varepsilon/2} \\ &\leq N [n^{-1} \log(n+1)]^{(2-\varepsilon)/2} [\mathbb{E} \rho_n^{(\varepsilon-2)/\varepsilon}]^{\varepsilon/2}. \end{aligned}$$

Notice that  $[n^{-1} \log(n+1)]^{(2-\varepsilon)/2} \leq N n^{-1+\varepsilon}$ , and we are done since

$$\sup_n \mathbb{E} \rho_n^{(\varepsilon-2)/\varepsilon} \leq N \exp \left( \frac{(\varepsilon-2)^2}{\varepsilon^2} T \sup |b| \right) \leq N. \quad \square$$

## 2.2 Regularity of the Kolmogorov equation

The final ingredient concerns the regularity of the associated Kolmogorov equation. While the result is possibly known, we did not find a reference in this form, so we provide a short proof. We note that (2.13) is the only instance where the Dini continuity assumption is used. We denote by  $C^{1,2}(Q_T)$  the space of all bounded continuous functions  $f$  on  $Q_T = (0, T) \times \mathbb{R}^d$  such that the derivatives  $\partial_t f$ ,  $\nabla f$ ,  $\nabla^2 f$  exist, are continuous and bounded, and we use the norm

$$\|f\|_{C^{1,2}(Q_T)} := \|f\|_{L^\infty(Q_T)} + \|\nabla f\|_{L^\infty(Q_T)} + \|\nabla^2 f\|_{L^\infty(Q_T)} + \|\partial_t f\|_{L^\infty(Q_T)}.$$

**Lemma 2.3.** *Let  $f, g \in \mathcal{D}$  and suppose that  $\|f\|_{\mathcal{D}} \leq K$  for some  $K \geq 0$ . There exists  $T_0 > 0$  depending only on  $K$  and  $d$ , such that there exists a unique bounded classical solution  $u$  of*

$$\partial_t u = \frac{1}{2} \Delta u + f \cdot \nabla u + g, \quad u(0, \cdot) = 0. \quad (2.12)$$

on  $Q_{T_0}$ . Moreover, for all  $T \in (0, T_0]$ , the solution of (2.12) on  $[0, T]$  satisfies the bounds

$$\|u\|_{C^{1,2}(Q_T)} \leq N \|g\|_{\mathcal{D}}, \quad (2.13)$$

$$\|\nabla u^i\|_{L^\infty(Q_T)} \leq \sqrt{T} N \|g\|_{L^\infty(\mathbb{R}^d)}, \quad (2.14)$$

where  $N = N(d, \vartheta, K)$ .

To prove Lemma 2.3, since replacing  $\vartheta(r)$  by  $\vartheta(r) \vee \sqrt{r}$  can only decrease the  $\|\cdot\|_{\mathcal{D}}$  norm of any function, without loss of generality we assume that  $\vartheta(r) \geq \sqrt{r}$  for all  $r \in [0, 1]$ . We denote by  $\mathcal{D}_T$  the space of all continuous functions on  $[0, T] \times \mathbb{R}^d$  such that

$$\|f\|_{\mathcal{D}_T} := \sup_{t \in [0, T]} \|f(t)\|_{\mathcal{D}} < \infty.$$

First consider the simpler equation

$$\partial_t u = \frac{1}{2} \Delta u + f, \quad u(0, \cdot) = 0. \quad (2.15)$$

We will use the following well-known properties of (2.15). While these are well-known in the PDE folklore (for an elliptic counterpart, see e.g. [GT83]), we did not find a reference for this exact form, so we sketch the proof in the Appendix for the interested reader.

**Lemma 2.4.** *Let  $T \in (0, 1]$  and  $f \in \mathcal{D}_T$ . Let  $u$  be a bounded distributional solution of (2.15). Then  $u$  is a classical solution of (2.15) which moreover satisfies for  $\alpha \in (0, 1)$*

$$\|u\|_{C^{1,2}(Q_T)} \leq N_0 \|f\|_{\mathcal{D}_T}, \quad (2.16)$$

$$\|u\|_{L^\infty(Q_T)} \leq N_1 T \|f\|_{L^\infty(Q_T)}, \quad (2.17)$$

$$\sup_{t \in [0, T]} \|\nabla u(t)\|_{C^\alpha(\mathbb{R}^d)} \leq N_2 T^{(1-\alpha)/2} \|f\|_{L^\infty(Q_T)} \quad (2.18)$$

where  $N_0 = N_0(d, \vartheta)$ ,  $N_1 = N_1(d)$ , and  $N_2 = N_2(d)$ .

*Proof of Lemma 2.3.* Uniqueness easily follows, see e.g. [Kry96]. As for existence, let  $g_n, f_n$  be bounded functions with bounded derivatives of any order such that

$$\|g_n - g\|_{L^\infty(\mathbb{R}^d)} + \|f_n - f\|_{L^\infty(\mathbb{R}^d)} \rightarrow 0, \quad \text{as } n \rightarrow \infty,$$

and

$$\|g_n\|_{L^\infty(\mathbb{R}^d)} \leq \|g\|_{L^\infty(\mathbb{R}^d)}, \quad \|f_n\|_{L^\infty(\mathbb{R}^d)} \leq \|f\|_{L^\infty(\mathbb{R}^d)}.$$

The problem

$$\partial_t u_n = \frac{1}{2} \Delta u_n + f_n \cdot \nabla u_n + g_n, \quad u(0, \cdot) = 0$$

has a classical solution  $u_n \in C^{1,2}(Q_{T_0})$  for any  $T_0 > 0$  such that  $u_n, \nabla u_n \in C(\overline{Q_{T_0}})$  (see e.g. [Kry96]). By (2.18) we have

$$\sup_{t \in [0, T_0]} \|\nabla u_n(t)\|_{C^{1/2}(\mathbb{R}^d)} \leq N_2(d) T_0^{1/4} (\|f\|_{L_\infty(\mathbb{R}^d)} \|\nabla u_n\|_{L_\infty(Q_T)} + \|g\|_{L_\infty(\mathbb{R}^d)}).$$

Hence, for sufficiently small  $T_0 > 0$  depending only on  $d$  and  $K$ , we have for all  $n \in \mathbb{N}$

$$\sup_{t \in [0, T_0]} \|\nabla u_n(t)\|_{C^{1/2}(\mathbb{R}^d)} \leq \|g\|_{L_\infty(\mathbb{R}^d)}. \tag{2.19}$$

Similarly, by (2.17)–(2.18), for  $T_0$  sufficiently small, we see that

$$\begin{aligned} & \|u_n - u_m\|_{L_\infty(Q_{T_0})} + \sup_{t \in [0, T_0]} \|\nabla u_n(t) - \nabla u_m(t)\|_{C^{1/2}(\mathbb{R}^d)} \\ & \leq \|g_n - g_m\|_{L_\infty(\mathbb{R}^d)} + \|f_n - f_m\|_{L_\infty(\mathbb{R}^d)} \|\nabla u_m\|_{L_\infty(Q_{T_0})} \\ & \leq \|g_n - g_m\|_{L_\infty(\mathbb{R}^d)} + \|f_n - f_m\|_{L_\infty(\mathbb{R}^d)} \|g\|_{L_\infty(\mathbb{R}^d)}, \end{aligned}$$

where for the last inequality we used (2.19). Consequently,  $u_n$  converges to a limit  $u \in C(\overline{Q_{T_0}})$  with  $\nabla u \in C(\overline{Q_{T_0}})$ , which is a distributional solution (2.12). Moreover, it satisfies

$$\sup_{t \in [0, T_0]} \|\nabla u(t)\|_{C^{1/2}(\mathbb{R}^d)} \leq \|g\|_{L_\infty(\mathbb{R}^d)}. \tag{2.20}$$

Consequently,  $f \nabla u \in \mathcal{D}_{T_0}$ , which implies by Lemma 2.3 that  $u$  is a classical solution of (2.12) and satisfies

$$\begin{aligned} \|u\|_{C^{1,2}(Q_{T_0})} & \leq N (\|f \nabla u\|_{\mathcal{D}_{T_0}} + \|g\|_{\mathcal{D}}) \\ & \leq N (\|f\|_{\mathcal{D}} \|\nabla u\|_{\mathcal{D}_{T_0}} + \|g\|_{\mathcal{D}}) \\ & \leq N \left( \|f\|_{\mathcal{D}} \sup_{t \in [0, T_0]} \|\nabla u\|_{C^{1/2}(\mathbb{R}^d)} + \|g\|_{\mathcal{D}} \right), \end{aligned}$$

where  $N = N(\vartheta, d)$  and we have used  $\vartheta(r) \geq \sqrt{r}$  for the last inequality. By (2.20) and (2.16) we obtain

$$\|u\|_{C^{1,2}(Q_{T_0})} \leq N \|g\|_{\mathcal{D}},$$

with  $N = N(\vartheta, d, K)$ , so the estimates (2.13)–(2.14) hold for  $T = T_0$ .

The bound (2.13) trivially extends to  $T < T_0$ . As for (2.14), with  $\lambda = \sqrt{T}/\sqrt{T_0} \leq 1$  let us set  $u_\lambda(t, x) := u(\lambda^2 t, \lambda x)$ . Clearly  $u_\lambda$  satisfies on  $Q_{T_0}$

$$\partial_t u_\lambda = \frac{1}{2} \Delta u_\lambda + \lambda f_\lambda \cdot \nabla u_\lambda + \lambda^2 g_\lambda, \quad u_\lambda(0, \cdot) = 0,$$

with  $f_\lambda(x) = f(\lambda x)$  and  $g_\lambda(x) = g(\lambda x)$ . Since

$$\|\lambda f_\lambda\|_{\mathcal{D}} \leq \|f_\lambda\|_{\mathcal{D}} \leq \|f\|_{\mathcal{D}} = K,$$

we get by applying (2.14) with  $T = T_0$

$$\lambda \|\nabla u\|_{L_\infty(Q_T)} = \|\nabla u_\lambda\|_{L_\infty(Q_{T_0})} \leq N \lambda^2 \|g_\lambda\|_{L_\infty(\mathbb{R}^d)} = N \lambda^2 \|g\|_{L_\infty(\mathbb{R}^d)},$$

thus obtaining (2.14) for arbitrary  $T \in (0, T_0]$ . □

### 3 Proofs of the main results

*Proof of Theorem 1.1.* First we prove for sufficiently small (but  $n$ -independent) time horizon  $T$ . Let  $T_0 = T_0(\|b\|_{\mathcal{D}}, d)$  be given by Lemma 2.3. We invoke the idea of [FGP10] (see also [PT17]): for any  $T \in (0, T_0]$  and for each  $i = 1, \dots, d$ , by Lemma 2.3 and a simple time reversal, there exists a classical solution  $u^i \in C^{1,2}(Q_T) \cap C(\overline{Q_T})$  to

$$\partial_t u^i + \frac{1}{2} \Delta u^i + b \cdot \nabla u^i = -b^i, \quad u(T, \cdot) = 0,$$

which satisfies the bounds

$$\|u^i\|_{C^{1,2}} \leq N' \|b\|_{\mathcal{D}} \quad (3.1)$$

$$\|\nabla u^i\|_{L^\infty(Q_T)} \leq \sqrt{T} N' \|b\|_{L^\infty(\mathbb{R}^d)}, \quad (3.2)$$

where  $N' = N(d, \vartheta, \|b\|_{\mathcal{D}})$ . We emphasise that  $N'$  is independent of  $T$ , which is a convention that we keep for the rest of the proof. Denote the  $j$ -th partial derivative of  $u$  by  $u_{x_j}$ . Applying Itô's formula for  $u^i(t, X_t)$  we have

$$\int_0^t b^i(X_s) ds = u^i(0, x_0) - u^i(t, X_t) + \sum_{j=1}^d \int_0^t u_{x_j}^i(s, X_s) dW_s^j. \quad (3.3)$$

Similarly, we have

$$\begin{aligned} \int_0^t b^i(X_s^n) ds &= u^i(0, x_0^n) - u^i(t, X_t^n) + \sum_{j=1}^d \int_0^t u_{x_j}^i(s, X_s^n) dW_s^j \\ &\quad + \sum_{j=1}^d \int_0^t u_{x_j}^i(s, X_s^n) (b^j(X_{k_n(s)}^n) - b^j(X_s^n)) ds. \end{aligned} \quad (3.4)$$

Moreover

$$\begin{aligned} \mathbb{E}|X_t - X_t^n|^2 &\leq 9\mathbb{E}|x_0 - x_0^n|^2 + 9\mathbb{E} \left| \int_0^t b(X_s) ds - \int_0^t b(X_s^n) ds \right|^2 \\ &\quad + 9\mathbb{E} \left| \int_0^t b(X_s^n) ds - \int_0^t b(X_{k_n(s)}^n) ds \right|^2 \\ &=: I_n^1 + I_n^2 + I_n^3. \end{aligned}$$

For  $I_n^2$  we have by (3.3)–(3.4)

$$\begin{aligned} I_n^2 &\leq N' \sum_{i=1}^d \mathbb{E}|u^i(0, x_0) - u^i(0, x_0^n)|^2 + N' \sum_{i=1}^d \mathbb{E}|u^i(t, X_t) - u^i(t, X_t^n)|^2 \\ &\quad + \sum_{i,j=1}^d N' \mathbb{E} \left| \int_0^t u_{x_j}^i(s, X_s) - u_{x_j}^i(s, X_s^n) dW_s^j \right|^2 \\ &\quad + N' \sum_{i,j=1}^d \mathbb{E} \left| \int_0^t u_{x_j}^i(s, X_s^n) b^j(X_s^n) - u_{x_j}^i(s, X_{k_n(s)}^n) b^j(X_{k_n(s)}^n) ds \right|^2 \\ &\quad + N' \sum_{i,j=1}^d \mathbb{E} \int_0^t |b^j(X_{k_n(s)}^n)|^2 |u_{x_j}^i(s, X_s^n) - u_{x_j}^i(s, X_{k_n(s)}^n)|^2 ds \\ &=: \sum_{i=1}^5 I_n^{2i}. \end{aligned} \quad (3.5)$$

One then has the following estimates: by (3.1) and the assumption on  $x_0, x_0^n$ , we have

$$I_n^1 + I_n^{21} \leq Nn^{-1+\varepsilon};$$

by (3.2) we have

$$I_n^{22} \leq N'\sqrt{T}\mathbb{E}|X_t - X_t^n|^2;$$

by Itô's isometry and (3.1) we have

$$I_n^{23} = N' \sum_{i,j=1}^d \int_0^t |u_{x_j}^i(s, X_s) - u_{x_j}^i(s, X_s^n)|^2 ds \leq N \int_0^t \mathbb{E}|X_s - X_s^n|^2 ds;$$

by the boundedness of  $b$  and by (3.1) we have

$$I_n^{25} \leq N \int_0^t \mathbb{E}|X_s^n - X_{\kappa_n(s)}^n|^2 ds \leq Nn^{-1}.$$

The terms  $I_n^{24}$  and  $I_n^3$  both can be estimated using Lemma 2.2, with  $f = u_{x_j}^i b^j$  and  $f = b$ , respectively, to get

$$I_n^{24} + I_n^3 \leq Nn^{-1+\varepsilon}.$$

Combining the above bounds we get

$$\mathbb{E}|X_t - X_t^n|^2 \leq Nn^{-1+\varepsilon} + N'\sqrt{T}\mathbb{E}|X_t - X_t^n|^2 + N \int_0^t \mathbb{E}|X_s - X_s^n|^2 ds.$$

For sufficiently small  $T$  (say  $\sqrt{T} < 1/(2N')$ ) the second term on the right-hand side can be omitted at the price of a constant factor. The conclusion then follows from Gronwall's lemma applied to the function  $t \mapsto \mathbb{E}|X_t - X_t^n|^2$ .

We now remove the smallness assumption on  $T$ . By the above, for some  $\delta = \delta(C, \varepsilon, \vartheta, d, \|b\|_{\mathcal{D}}) > 0$ , the statement is proved for  $T'$  in place of  $T$ , for arbitrary  $T' \in [0, \delta]$ . Now fix  $n$  and define  $k$  by  $k/n = \kappa_n(\delta)$ . Denote  $m = \lceil \frac{T}{\kappa_n(\delta)} \rceil$  and note that  $m \leq \frac{2T}{\delta} \leq N$ . Using the above proven bound, we have

$$\sup_{t \in [0, k/n]} \mathbb{E}|X_t^n - X_t|^2 \leq Nn^{-1+\varepsilon}. \tag{3.6}$$

On the time interval  $[k/n, 2k/n]$ ,  $X_{t+k/n}$  is the solution of

$$dY_t = b(Y_t) dt + dW_{t+k/n}, \quad Y_0 = X_{k/n},$$

while  $X_{t+k/n}^n$  is the solution of

$$dY_t^n = b(Y_{\kappa_n(t)}^n) dt + dW_{t+k/n}, \quad Y_0^n = X_{k/n}^n.$$

Moreover, the initial conditions  $X_{k/n}, X_{k/n}^n$  satisfy the condition of Theorem 1.1, by (3.6). Since the theorem is now proven to hold for small times, we can conclude

$$\sup_{t \in [k/n, 2k/n]} \mathbb{E}|X_t^n - X_t|^2 = \sup_{t \in [0, k/n]} \mathbb{E}|Y_t^n - Y_t|^2 \leq Nn^{-1+\varepsilon},$$

and combining with (3.6),

$$\sup_{t \in [0, 2k/n]} \mathbb{E}|X_t^n - X_t|^2 \leq Nn^{-1+\varepsilon}. \tag{3.7}$$

Iterating this procedure  $m$  times we reach the full time horizon  $[0, T]$ . □

*Proof of Theorem 1.2.* We again prove the statement for  $T \in (0, T_0]$  for some sufficiently small  $T_0 > 0$  (to be determined later), the general case can then be deduced exactly as above. Correspondingly, from here on  $N = N(C, \varepsilon, \sup |b|)$ .

For  $z \in \mathbb{R}$  let us define

$$\phi_z(x) := \int_0^x \exp\left(-2 \int_z^r I_{|z-s| \leq 2} b(s) ds\right) dr, \quad \psi_z(x) := \phi_z^{-1}(x).$$

Notice that  $\phi_z \in W^{2,\infty}(\mathbb{R})$  and for almost everywhere on  $[z - 2, z + 2]$  we have

$$\frac{1}{2}\phi_z'' + b\phi_z' = 0. \tag{3.8}$$

Moreover, it is straightforward to check that

$$\sup_{z,x} (|\phi_z'(x)| + |\phi_z''(x)| + |\psi_z'(x)| + |(\phi_z' \circ \psi_z)'(x)|) \leq N, \tag{3.9}$$

with  $N = N(\sup |b|)$ . We define inductively the stopping times

$$\bar{\tau}_0 := 0, \quad \bar{\tau}_{m+1} := \inf\{t \geq \tau_m : |X_t - X_{\tau_m}| > 1\}, \quad \tau_m := \bar{\tau}_m \wedge T.$$

First we aim to obtain a bound for

$$\sup_{t \in [0, T]} \mathbb{E}(\mathbf{1}_{\tau_m \leq t} |X_{t \wedge \tau_{m+1}} - X_{t \wedge \tau_{m+1}}^n|^2). \tag{3.10}$$

By definition of  $\phi_z$ , we can find a function  $g = g(z, x)$  such that for each  $z \in \mathbb{R}$ ,  $g(z, \cdot)$  is a version of  $\phi_z''$ ,  $\frac{1}{2}g(z, x) + b(x)\phi_z'(x) = 0$  for all  $x \in [z - 2, z + 2]$ , and for each  $x \in \mathbb{R}$ ,  $g(\cdot, x)$  is continuous in  $z$ . We choose this function as a representative of  $\phi_z''(x)$  for now on and by Itô's formula (which, although  $\phi_z''$  is not necessarily continuous, holds, see [Kal02, Thm. 22.5]) we have on  $[[\tau_m, \tau_{m+1}]] := \{(\omega, t) \in \Omega \times [0, T] : t \in [\tau + m, \tau_{m+1}]\}$

$$\phi_z(X_t) = \phi_z(X_{\tau_m}) + \int_{\tau_m}^t \frac{1}{2}\phi_z''(X_s) + b(X_s)\phi_z'(X_s) ds + \int_{\tau_m}^t \phi_z'(X_s) dW_s.$$

Let  $(\rho_\varepsilon)_{\varepsilon > 0}$  be a mollification family. We multiply the above equality with  $\rho_\varepsilon(z - X_{\tau_m})$ , we integrate over  $z \in \mathbb{R}$  and we let  $\varepsilon \rightarrow 0$ , to obtain by virtue of the continuity of  $\phi_z, \phi_z'$  and  $\phi_z''$  in  $z$  and the dominated convergence theorem

$$\phi_{X_{\tau_m}}(X_t) = \phi_{X_{\tau_m}}(X_{\tau_m}) + \int_{\tau_m}^t \frac{1}{2}\phi_{X_{\tau_m}}''(X_s) + b(X_s)\phi_{X_{\tau_m}}'(X_s) ds + \int_{\tau_m}^t \phi_{X_{\tau_m}}'(X_s) dW_s.$$

Using (3.8) (which now holds for all  $x \in [z - 2, z + 2]$ ) we have for  $Y_t^m = \phi_{X_{\tau_m}}(X_t)$  on  $[[\tau_m, \tau_{m+1}]]$  that

$$dY_t^m = \phi_{X_{\tau_m}}(X_{\tau_m}) + \int_{\tau_m}^t \phi_{X_{\tau_m}}' \circ \psi_{X_{\tau_m}}(Y_s^m) dW_s. \tag{3.11}$$

On  $[[\tau_m, \tau_{m+1}]]$  we define

$$Y_t^{m,n} = \phi_{X_{\tau_m}}(X_{\tau_m}^n) + \int_{\tau_m}^t \phi_{X_{\tau_m}}' \circ \psi_{X_{\tau_m}}(Y_{k_n(s) \vee \tau_m}^{m,n}) dW_s.$$

Then by (3.9) and the triangle inequality we can bound the quantity in (3.10) as

$$\begin{aligned} \mathbb{E}(\mathbf{1}_{\tau_m \leq t} |X_{t \wedge \tau_{m+1}} - X_{t \wedge \tau_{m+1}}^n|^2) &\leq N \mathbb{E}(\mathbf{1}_{\tau_m \leq t} |Y_{t \wedge \tau_{m+1}}^m - \phi_{X_{\tau_m}}(X_{t \wedge \tau_{m+1}}^n)|^2) \\ &\leq N \mathbb{E}(\mathbf{1}_{\tau_m \leq t} |Y_{t \wedge \tau_{m+1}}^m - Y_{t \wedge \tau_{m+1}}^{m,n}|^2) + N \mathbb{E}(\mathbf{1}_{\tau_m \leq t} |Y_{t \wedge \tau_{m+1}}^{m,n} - \phi_{X_{\tau_m}}(X_{t \wedge \tau_{m+1}}^n)|^2). \end{aligned} \tag{3.12}$$

The first term on the right-hand side is simply the error of an Euler scheme for the SDE (3.11) with Lipschitz coefficient  $\phi'_{X_{\tau_m}} \circ \psi_{X_{\tau_m}}$ . This is standard to estimate (for the convenience of the reader we include a short proof in the Appendix):

$$\begin{aligned} \mathbb{E}(\mathbf{1}_{\tau_m \leq t} |Y_{t \wedge \tau_{m+1}}^{m,n} - Y_{t \wedge \tau_{m+1}}^{m,n}|^2) &\leq N \mathbb{E}|\phi_{X_{\tau_m}}(X_{\tau_m}) - \phi_{X_{\tau_m}}(X_{\tau_m}^n)|^2 + Nn^{-1} \\ &\leq N \mathbb{E}|X_{\tau_m} - X_{\tau_m}^n|^2 + Nn^{-1}, \end{aligned} \tag{3.13}$$

where we have used again (3.9) for the last inequality. Moving on the second term on the right-hand side of (3.12), by Itô's formula again we have on  $[[\tau_m, \tau_{m+1}]]$

$$\begin{aligned} \phi_{X_{\tau_m}}(X_t^n) - Y_t^{m,n} &= \int_{\tau_m}^t \phi'_{X_{\tau_m}}(X_s^n) b(X_{k_n(s)}^n) + \frac{1}{2} \phi''_{X_{\tau_m}}(X_s^n) ds \\ &\quad + \int_{\tau_m}^t (\phi'_{X_{\tau_m}}(X_s^n) - \phi'_{X_{\tau_m}} \circ \psi(Y_{k_n(s) \wedge \tau_m}^{m,n})) dW_s \\ &= \int_0^t \mathbf{1}_{\tau_m \leq s} \mathbf{1}_{|X_s - X_s^n| < 1} (\phi'_{X_{\tau_m}}(X_s^n) b(X_{k_n(s)}^n) - \phi'_{X_{\tau_m}}(X_s^n) b(X_s^n)) ds \\ &\quad + \int_0^t \mathbf{1}_{\tau_m \leq s} \mathbf{1}_{|X_s - X_s^n| \geq 1} (\phi'_{X_{\tau_m}}(X_s^n) b(X_{k_n(s)}^n) + \frac{1}{2} \phi''_{X_{\tau_m}}(X_s^n)) ds \\ &\quad + \int_0^t \mathbf{1}_{\tau_m \leq s} (\phi'_{X_{\tau_m}}(X_s^n) - \phi'_{X_{\tau_m}} \circ \psi(Y_{k_n(s) \vee \tau_m}^n)) dW_s \\ &=: I_t^1 + I_t^2 + I_t^3, \end{aligned}$$

where for the second equality we have used (3.8). It immediately follows from (3.9) that

$$\mathbb{E} \mathbf{1}_{\tau_m \leq t} |I_{t \wedge \tau_{m+1}}^2|^2 \leq N \mathbb{E} \int_0^{t \wedge \tau_{m+1}} \mathbf{1}_{\tau_m \leq s} \mathbf{1}_{|X_s - X_s^n| \geq 1} ds =: J. \tag{3.14}$$

Next, one can write

$$\begin{aligned} \mathbb{E} \mathbf{1}_{\tau_m \leq t} |I_{t \wedge \tau_{m+1}}^1|^2 &\leq N \mathbb{E} \int_{\tau_m}^{t \wedge \tau_{m+1}} \mathbf{1}_{|X_s - X_s^n| \geq 1} ds \\ &\quad + 2 \mathbb{E} \left| \int_{\tau_m}^{(\tau_m \vee t) \wedge \tau_{m+1}} (\phi'_{X_{\tau_m}}(X_s^n) b(X_{k_n(s)}^n) - \phi'_{X_{\tau_m}}(X_s^n) b(X_s^n)) ds \right|^2 \\ &\leq J + N \int_0^T \mathbb{E} |X_s^n - X_{k_n(s)}^n|^2 ds \\ &\quad + N \mathbb{E} \left| \int_{\tau_m}^{(\tau_m \vee t) \wedge \tau_{m+1}} (\phi'_{X_{\tau_m}}(X_{k_n(s)}^n) b(X_{k_n(s)}^n) - \phi'_{X_{\tau_m}}(X_s^n) b(X_s^n)) ds \right|^2 \\ &\leq J + Nn^{-1+\varepsilon}, \end{aligned} \tag{3.15}$$

where for the last term we have used Lemma 2.2. Concerning  $I^3$ , by Itô's isometry and (3.9) we have

$$\begin{aligned} \mathbb{E} \mathbf{1}_{\tau_m \leq t} |I_{t \wedge \tau_{m+1}}^3|^2 &\leq N \mathbb{E} \int_0^{t \wedge \tau_{m+1}} \mathbf{1}_{\tau_m \leq s} |\phi_{X_{\tau_m}}(X_s^n) - Y_s^{m,n}| ds \\ &\quad + N \mathbb{E} \int_{\tau_m}^{\tau_{m+1}} |Y_s^{m,n} - Y_{k_n(s) \wedge \tau_m}^{m,n}|^2 ds \\ &\leq N \int_0^t \mathbb{E} \mathbf{1}_{\tau_m \leq s} |\phi_{X_{\tau_m}}(X_{s \wedge \tau_{m+1}}^n) - Y_{s \wedge \tau_{m+1}}^{m,n}|^2 ds + Nn^{-1}. \end{aligned} \tag{3.16}$$

Putting (3.14)–(3.15)–(3.16) together, we arrive at

## Euler-Maruyama scheme with irregular drift

$$\mathbb{E}\mathbf{1}_{\tau_m \leq t} |\phi_{X_{\tau_m}}(X_{t \wedge \tau_{m+1}}^n) - Y_{t \wedge \tau_{m+1}}^{m,n}|^2 \leq N \int_0^t \mathbb{E}\mathbf{1}_{\tau_m \leq s} |\phi_{X_{\tau_m}}(X_{s \wedge \tau_{m+1}}^n) - Y_{s \wedge \tau_{m+1}}^{m,n}|^2 ds + NJ + Nn^{-1+\varepsilon}.$$

Applying Gronwall's lemma to the function  $t \mapsto \mathbb{E}\mathbf{1}_{\tau_m \leq t} |\phi_{X_{\tau_m}}(X_{t \wedge \tau_{m+1}}^n) - Y_{t \wedge \tau_{m+1}}^{m,n}|^2$ , we get

$$\mathbb{E}\mathbf{1}_{\tau_m \leq t} |\phi_{X_{\tau_m}}(X_{t \wedge \tau_{m+1}}^n) - Y_{t \wedge \tau_{m+1}}^{m,n}|^2 \leq NJ + Nn^{-1+\varepsilon}. \quad (3.17)$$

Putting (3.12)–(3.13)–(3.17) together, we get

$$\sup_{t \in [0, T]} \mathbb{E}(\mathbf{1}_{\tau_m \leq t} |X_{t \wedge \tau_{m+1}} - X_{t \wedge \tau_{m+1}}^n|^2) \leq N\mathbb{E}|X_{\tau_m} - X_{\tau_m}^n|^2 + Nn^{-1+\varepsilon} + NJ. \quad (3.18)$$

Notice that

$$J \leq NT \sup_{t \in [0, T]} \mathbb{E}(\mathbf{1}_{\tau_m \leq t} |X_{t \wedge \tau_{m+1}} - X_{t \wedge \tau_{m+1}}^n|^2).$$

Therefore, provided that  $T$  is sufficiently small, gives the last term on the right-hand side of (3.18) can be absorbed at a price of a constant factor. We then choose  $T$  in place of  $t$ , which implies  $\mathbf{1}_{\tau_m \leq T} = 1$  and  $T \wedge \tau_{m+1} = \tau_{m+1}$ , yielding

$$\mathbb{E}|X_{\tau_{m+1}} - X_{\tau_{m+1}}^n|^2 \leq N\mathbb{E}|X_{\tau_m} - X_{\tau_m}^n|^2 + Nn^{-1+\varepsilon}.$$

This by induction gives for  $m \geq 1$

$$\mathbb{E}|X_{\tau_m} - X_{\tau_m}^n|^2 \leq N_0^m n^{-1+\varepsilon},$$

for some constant  $N_0 = N_0(C, \varepsilon, \sup |b|)$  that we fix for the remainder of the proof. It remains to remove the stopping time from the estimate. We write

$$\begin{aligned} \mathbb{E}|X_T - X_T^n|^2 &\leq \mathbb{E}(|X_{\tau_m} - X_{\tau_m}^n|^2) + \mathbb{E}(\mathbf{1}_{\tau_m < T} |X_T - X_T^n|^2) \\ &\leq N_1^m n^{-1+\varepsilon} + N(\mathbb{P}(\bar{\tau}_m < T))^{1/2}. \end{aligned}$$

Define a new set of stopping times by

$$\hat{\tau}_0 := 0, \quad \hat{\tau}_{m+1} := \inf\{t \geq \hat{\tau}_m : |W_t - W_{\tau_m}| > 1/2\}.$$

For  $T < (2 \sup |b|)^{-1}$ , the contribution of the drift to  $X$  is strictly smaller than  $1/2$ , and hence  $\hat{\tau}_m \leq \bar{\tau}_m$ . Also, by the strong Markov property we have that the random variables  $\{\hat{\tau}_m - \hat{\tau}_{m-1}\}_{m=1}^\infty$  are independent and identically distributed. Consequently, we have

$$\mathbb{P}(\hat{\tau}_m \leq T) \leq \prod_{i=1}^m \mathbb{P}(\hat{\tau}_i - \hat{\tau}_{i-1} \leq T) = (\mathbb{P}(\hat{\tau}_1 \leq T))^m.$$

As  $T \rightarrow 0$ , one has  $\lim_{T \rightarrow 0} \mathbb{P}(\hat{\tau}_1 \leq T) = 0$ , and therefore  $\alpha_T := -\log \mathbb{P}(\hat{\tau}_1 \leq T) \rightarrow \infty$ . Hence, we obtain for all  $m \in \mathbb{N}$

$$\mathbb{E}|X_T - X_T^n|^2 \leq N_0^m n^{-1+\varepsilon} + Ne^{-\alpha_T m},$$

which upon choosing  $m = \left\lceil \varepsilon \frac{\ln n}{\ln N_0} \right\rceil$  and assuming that  $T$  is sufficiently small (so that  $\alpha_T$  is sufficiently large) gives

$$\mathbb{E}|X_T - X_T^n|^2 \leq Nn^{-1+2\varepsilon},$$

which brings the proof to an end. (Note that since  $T$  was arbitrary in the interval  $(0, T_0]$ , the form (1.5) with the supremum outside of the expectation follows.)  $\square$

## A Appendix

*Proof of Lemma 2.4.* It is well known that the function

$$u(t, x) = \int_0^t P_{t-s} f(s, x) ds, \quad (\text{A.1})$$

where  $(P_t)_{t \geq 0}$  is the heat semigroup, is the unique bounded distributional solution of equation (2.16). One can write the action of the heat semigroup with its kernel in a convolution form  $P_t g(x) = (p(t, \cdot) * g)(x)$  and recall that for any  $k = 0, 1, 2$  and  $t \in (0, 1]$ , one has

$$\|\partial_x^k p(t, \cdot)\|_{L_1(\mathbb{R}^d)} \leq N_0 t^{-k/2} \quad (\text{A.2})$$

It then follows immediately that one has the bounds

$$\|u\|_{L_\infty(Q_T)} + \|\nabla u\|_{L_\infty(Q_T)} \leq N_0 \|f\|_{L_\infty(Q_T)} \quad (\text{A.3})$$

We now show that  $\nabla^2 u \in C((0, T) \times \mathbb{R}^d)$ . First let us remark that from (1.3) it follows by a simple change of variables that  $\int_0^1 \vartheta(r^\gamma) r^{-1} dr < \infty$  for any  $\gamma \in (0, 1)$ . Notice that for any  $t > 0$ ,  $g \in \mathcal{D}$ , we have

$$\begin{aligned} |\partial_{x_i x_j} P_t g(x)| &= \left| \int_{\mathbb{R}^d} p_{x_i x_j}(t, x - y) (g(y) - g(x)) dy \right| \\ &\leq N(d, \gamma) \left( \|g\|_{\mathcal{D}} \vartheta(t^{1/4}) t^{-1} + \|g\|_{L_\infty} \left| \int_{|x| > t^{1/4}} p_{x_i x_j}(t, x) dx \right| \right). \end{aligned}$$

Since on  $|x| > t^{1/4}$  we have  $|x|^2/t > T^{-1/2}$ , we see that

$$\begin{aligned} \left| \int_{|x| > t^{1/4}} p_{x_i x_j}(t, x) dx \right| &\leq N(d) \int_{|x| > t^{1/4}} \frac{|x|^2}{t^{d/2+2}} \exp\left(\frac{-|x|^2}{2t}\right) dx \\ &\leq N(d, T) \int_{t^{1/4}}^\infty r^{d-1} \frac{r^2}{t^{d/2+2}} \exp\left(\frac{-r^2}{2t}\right) dr \\ &\leq N(d, T) t^{-3/2} \int_{t^{1/4}}^\infty \exp\left(\frac{-r^2}{4t}\right) dr \\ &\leq N(d, T) t^{-1} \exp\left(\frac{-t^{-1/2}}{4}\right) \end{aligned}$$

Clearly, the right-hand side is integrable in  $t$ . We can conclude

$$\|\partial_{x_i x_j} P_t g\|_{L_\infty(\mathbb{R}^d)} \leq N(d, \gamma) \lambda(t) \|g\|_{\mathcal{D}}, \quad (\text{A.4})$$

where  $\lambda \in L_1(0, T)$ . It follows then that for each  $t \in (0, T)$ ,  $u(t)$  is twice differentiable in space. Moreover, since

$$\partial_{x_i x_j} u(t, x) = \int_0^t \int_{\mathbb{R}^d} \partial_{x_i x_j} p(s, y) f(t - s, x - y) dy ds,$$

the continuity of  $\partial_{x_i x_j} u$  in  $(t, x)$  follows easily by the continuity of  $f$  and (A.4), by virtue of Lebesgue's theorem on dominated convergence. The estimate

$$\|\nabla^2 u\|_{L_\infty(Q_T)} \leq N_0 \sup_{t \leq T} \|f(t)\|_{\mathcal{D}}$$

follows from (A.4). The fact that  $\partial_t u \in C((0, T) \times \mathbb{R}^d)$  now follows from the equation. Finally, (2.16) follows from (A.3), (A.4), and the equation.

We proceed with the proof for the remaining two estimates. Estimate (2.17) follows immediately from (A.1) and (A.2). Finally for the last estimate we have

$$|\partial_{x_i} u(t, x) - \partial_{x_i} u(t, y)| \leq \|f\|_{L_\infty(Q_T)} \int_0^t \int_{\mathbb{R}^d} |p_{x_i}(t-s, x-z) - p_{x_i}(t-s, y-z)| dz ds. \tag{A.5}$$

By (A.2) we have

$$\int_{\mathbb{R}^d} |p_{x_i}(t-s, x-z) - p_{x_i}(t-s, y-z)| dz \leq N(d)|t-s|^{-1/2},$$

and by the fundamental theorem of calculus and (A.2) we have

$$\begin{aligned} \int_{\mathbb{R}^d} |p_{x_i}(t-s, x-z) - p_{x_i}(t-s, y-z)| dz &\leq N(d)|x-y| \|\nabla p_{x_i}(t-s)\|_{L^1(\mathbb{R}^d)} \\ &\leq N(d)|x-y| |t-s|^{-1}, \end{aligned}$$

which imply that for all  $\alpha \in [0, 1]$

$$\int_{\mathbb{R}^d} |p_{x_i}(t-s, x-z) - p_{x_i}(t-s, y-z)| dz \leq N(d)|x-y|^\alpha |t-s|^{-(1+\alpha)/2}.$$

This combined with (A.5) shows (2.18). The lemma is proved.  $\square$

*Proof of (3.13).* By definition of  $\phi_z$  and (3.9) we have a uniform (in  $n$  and  $m$ ) bound on the initial condition for  $Y^{n,m}$ :

$$\mathbb{E}|\phi_{X_{\tau_m}}(X_{\tau_m}^n)|^2 \leq N(1 + \mathbb{E}|X_{\tau_m}^n|^2) \leq N(1 + \mathbb{E} \sup_{t \in [0, T]} |X_t^n|^2) \leq N.$$

Combined with the (uniform in  $z$ ) linear growth of the coefficient  $\phi'_z \circ \psi_z$ , this implies by standard arguments

$$\mathbb{E} \sup_{t \in [\tau_m, \tau_{m+1}]} |Y_t^{m,n}|^2 \leq N.$$

From this we get

$$\begin{aligned} \mathbb{E} \sup_{t \in [\tau_m, \tau_{m+1}]} |Y_{k_n(t) \vee \tau_m}^{m,n} - Y_t^{m,n}|^2 &\leq \mathbb{E} \left| \int_{k_n(t) \vee \tau_m}^t \mathbf{1}_{s \leq \tau_{m+1}} \phi'_{X_{\tau_m}} \circ \psi_{X_{\tau_m}}(Y_{k_n(s) \vee \tau_m}^{m,n}) ds \right|^2 \\ &\leq n^{-1} N(1 + \mathbb{E} \sup_{t \in [\tau_m, \tau_{m+1}]} |Y_t^{m,n}|^2) \leq Nn^{-1}. \end{aligned}$$

By Itô's formula again and the above inequality we have that

$$\begin{aligned} &\mathbb{E} \mathbf{1}_{\tau_m \leq t} |Y_{t \wedge \tau_{m+1}}^{m,n} - Y_{t \wedge \tau_{m+1}}^m|^2 \\ &\leq \mathbb{E} |\phi_{X_{\tau_m}}(X_{\tau_m}^n) - \phi_{X_{\tau_m}}(X_{\tau_m})|^2 \\ &\quad + \int_0^t \mathbb{E} \mathbf{1}_{\tau_m \leq s \leq \tau_{m+1}} |\phi'_{X_{\tau_m}} \circ \psi_{X_{\tau_m}}(Y_{k_n(s) \vee \tau_m}^{m,n}) - \phi'_{X_{\tau_m}} \circ \psi_{X_{\tau_m}}(Y_s^m)|^2 ds \\ &\leq \mathbb{E} |\phi_{X_{\tau_m}}(X_{\tau_m}^n) - \phi_{X_{\tau_m}}(X_{\tau_m})|^2 \\ &\quad + N \int_0^t \mathbb{E} \mathbf{1}_{\tau_m \leq s \leq \tau_{m+1}} |Y_{k_n(s) \vee \tau_m}^{m,n} - Y_s^{m,n}|^2 + N \mathbb{E} \mathbf{1}_{\tau_m \leq s} |Y_{s \wedge \tau_{m+1}}^{m,n} - Y_{s \wedge \tau_{m+1}}^m|^2 ds \\ &\leq \mathbb{E} |\phi_{X_{\tau_m}}(X_{\tau_m}^n) - \phi_{X_{\tau_m}}(X_{\tau_m})|^2 + Nn^{-1} + N \int_0^t \mathbb{E} \mathbf{1}_{\tau_m \leq s} |Y_{s \wedge \tau_{m+1}}^{m,n} - Y_{s \wedge \tau_{m+1}}^m|^2 ds, \end{aligned}$$

and (3.13) follows from Gronwall's lemma applied to the function  $t \mapsto \mathbb{E} \mathbf{1}_{\tau_m \leq t} |Y_{t \wedge \tau_{m+1}}^{m,n} - Y_{t \wedge \tau_{m+1}}^m|^2$ .  $\square$

## References

- [Alt17] R. Altmeyer. Estimating occupation time functionals. *arXiv e-prints* arXiv:1706.03418.
- [BHY18] J. Bao, X. Huang, and C. Yuan. Convergence rate of Euler–Maruyama scheme for SDEs with Hölder–Dini continuous drifts. *Journal of Theoretical Probability* (2018). doi:10.1007/s10959-018-0854-9. MR-3959630
- [BS18] J. Bao and J. Shao. Weak convergence of path-dependent SDEs with irregular coefficients. *arXiv e-prints* arXiv:1809.03088.
- [Dav07] A. M. Davie. Uniqueness of solutions of stochastic differential equations. *Int. Math. Res. Not. IMRN*, no. 24, (2007), Art. ID rnm124, 26. doi:10.1093/imrn/rnm124. MR-2377011
- [FF11] E. Fedrizzi and F. Flandoli. Pathwise uniqueness and continuous dependence of SDEs with non-regular drift. *Stochastics* **83**, no. 3, (2011), 241–257. doi:10.1080/17442508.2011.553681. MR-2810591
- [FGP10] F. Flandoli, M. Gubinelli, and E. Priola. Well-posedness of the transport equation by stochastic perturbation. *Inventiones Mathematicae* **180**, no. 1, (2010), 1–53. doi:10.1007/s00222-009-0224-4. MR-2593276
- [GJS17] M. Gerencsér, A. Jentzen, and D. Salimova. On stochastic differential equations with arbitrarily slow convergence rates for strong approximation in two space dimensions. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **473**, no. 2207, (2017). doi:10.1098/rspa.2017.0104. MR-3741822
- [GK96] I. Gyöngy and N. Krylov. Existence of strong solutions for Itô’s stochastic equations via approximations. *Probability Theory and Related Fields* **105**, no. 2, (1996), 143–158. doi:10.1007/BF01203833. MR-1392450
- [GM01] I. Gyöngy and T. Martínez. On stochastic differential equations with locally unbounded drift. *Czechoslovak Mathematical Journal* **51**, no. 4, (2001), 763–783. doi:10.1023/A:1013764929351.
- [GT83] D. Gilbarg and N. S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Springer Berlin Heidelberg, 1983. doi:10.1007/978-3-642-61798-0.
- [Gyö98] I. Gyöngy. A note on Euler’s approximations. *Potential Analysis* **8**, no. 3, (1998), 205–216. doi:10.1023/A:1016557804966.
- [HHJ15] M. Hairer, M. Hutzenthaler, and A. Jentzen. Loss of regularity for Kolmogorov equations. *Annals of Probability* **43**, no. 2, (2015), 468–527. doi:10.1214/13-AOP838. MR-3305998
- [HL18] X. Huang and Z.-W. Liao. The Euler–Maruyama method for S(F)DEs with Hölder drift and  $\alpha$ -stable noise. *Stochastic Analysis and Applications* **36**, no. 1, (2018), 28–39. doi:10.1080/07362994.2017.1371037.
- [JMY16] A. Jentzen, T. Müller-Gronbach, and L. Yaroslavtseva. On stochastic differential equations with arbitrary slow convergence rates for strong approximation. *Communications in Mathematical Sciences* **14**, no. 6, (2016), 1477–1500. doi:10.4310/CMS.2016.v14.n6.a1. MR-3538358
- [Kal02] O. Kallenberg. *Foundations of Modern Probability*. Springer New York, 2002. doi:10.1007/978-1-4757-4015-8. MR-1876169
- [KR05] N. V. Krylov and M. Röckner. Strong solutions of stochastic equations with singular time dependent drift. *Probability Theory and Related Fields* **131**, no. 2, (2005), 154–196. doi:10.1007/s00440-004-0361-z.
- [Kry96] N. Krylov. *Lectures on Elliptic and Parabolic Equations in Hölder Spaces*. Graduate Studies in Mathematics. American Mathematical Society, 1996. MR-1406091
- [LS15] G. Leobacher and M. Szölgvényi. A numerical method for SDEs with discontinuous drift. *BIT Numerical Mathematics* **56**, no. 1, (2015), 151–162. doi:10.1007/s10543-015-0549-x. MR-3486457
- [LS17] G. Leobacher and M. Szölgvényi. A strong order 1/2 method for multidimensional SDEs with discontinuous drift. *Ann. Appl. Probab.* **27**, no. 4, (2017), 2383–2418. doi:10.1214/16-AAP1262. MR-3693529

- [LS18] G. Leobacher and M. Szölgényi. Convergence of the Euler-Maruyama method for multidimensional SDEs with discontinuous drift and degenerate diffusion coefficient. *Numerische Mathematik* **138**, no. 1, (2018), 219–239. doi:10.1007/s00211-017-0903-9. MR-3745015
- [MP91] R. Mikulevičius and E. Platen. Rate of convergence of the Euler approximation for diffusion processes. *Math. Nachr.* **151**, (1991), 233–239. doi:10.1002/mana.19911510114. MR-1121206
- [MX18] R. Mikulevičius and F. Xu. On the rate of convergence of strong Euler approximation for SDEs driven by Levy processes. *Stochastics* **90**, no. 4, (2018), 569–604. doi:10.1080/17442508.2017.1381095. MR-3784978
- [MY18] T. Müller-Gronbach and L. Yaroslavtseva. On the performance of the Euler-Maruyama scheme for SDEs with discontinuous drift coefficient. *arXiv e-prints* arXiv:1809.08423. MR-4076779
- [MY19] T. Müller-Gronbach and L. Yaroslavtseva. A strong order  $3/4$  method for SDEs with discontinuous drift coefficient, 2019.
- [NS19] A. Neuenkirch and M. Szölgényi. The Euler-Maruyama scheme for SDEs with irregular drift: convergence rates via reduction to a quadrature problem. *arXiv e-prints* arXiv:1904.07784 (2019).
- [NSS18] A. Neuenkirch, M. Szölgényi, and L. Szpruch. An adaptive Euler-Maruyama scheme for stochastic differential equations with discontinuous drift and its convergence analysis. *arXiv e-prints* arXiv:1802.04521. MR-3914573
- [NT16] H.-L. Ngo and D. Taguchi. Strong rate of convergence for the Euler-Maruyama approximation of stochastic differential equations with irregular coefficients. *Math. Comp.* **85**, no. 300, (2016), 1793–1819. doi:10.1090/mcom3042. MR-3471108
- [NT17a] H.-L. Ngo and D. Taguchi. On the Euler-Maruyama approximation for one-dimensional stochastic differential equations with irregular coefficients. *IMA Journal of Numerical Analysis* **37**, no. 4, (2017), 1864–1883. doi:10.1093/imanum/drw058. MR-3712177
- [NT17b] H.-L. Ngo and D. Taguchi. Strong convergence for the Euler-Maruyama approximation of stochastic differential equations with discontinuous coefficients. *Statistics & Probability Letters* **125**, (2017), 55–63. doi:10.1016/j.spl.2017.01.027. MR-3626068
- [PT17] O. M. Pamen and D. Taguchi. Strong rate of convergence for the Euler-Maruyama approximation of SDEs with Hölder continuous drift coefficient. *Stochastic Processes and their Applications* **127**, no. 8, (2017), 2542–2559. doi:https://doi.org/10.1016/j.spa.2016.11.008. MR-3660882
- [Sha16] A. V. Shaposhnikov. Some remarks on Davie’s uniqueness theorem. *Proc. Edinb. Math. Soc. (2)* **59**, no. 4, (2016), 1019–1035. doi:10.1017/S0013091515000589.
- [Ver80] A. J. Veretennikov. Strong solutions and explicit formulas for solutions of stochastic integral equations. *Mat. Sb. (N.S.)* **111(153)**, no. 3, (1980), 434–452, 480. MR-0568986
- [Yar17] L. Yaroslavtseva. On non-polynomial lower error bounds for adaptive strong approximation of SDEs. *Journal of Complexity* **42**, (2017), 1–18. doi:10.1016/j.jco.2017.04.002. MR-3683255
- [Zvo74] A. K. Zvonkin. A transformation of the phase space of a diffusion process that removes the drift. *Mathematics of the USSR-Sbornik* **22**, no. 1, (1974), 129. doi:10.1070/SM1974v022n01ABEH001689.