Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Seminar for Applied Mathematics

An adaptive Euler-Maruyama scheme for stochastic differential equations with discontinuous drift and its convergence analysis

A. Neuenkirch and M. Szölgyenyi and L. Szpruch

Research Report No. 2018-06
February 2018

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

An adaptive Euler-Maruyama scheme for stochastic differential equations with discontinuous drift and its convergence analysis

Andreas Neuenkirch Michaela Szölgyenyi Lukasz Szpruch

Preprint, February 2018

Abstract

We study the strong approximation of stochastic differential equations with discontinuous drift coefficients and (possibly) degenerate diffusion coefficients. To account for the discontinuity of the drift coefficient we construct an adaptive step sizing strategy for the explicit Euler-Maruyama scheme. As a result, we obtain a numerical method which has – up to logarithmic terms – strong convergence order 1/2 with respect to the average computational cost. We support our theoretical findings with several numerical examples.

Keywords: stochastic differential equations, discontinuous drift, degenerate diffusion, adaptive Euler-Maruyama scheme, strong convergence order
Mathematics Subject Classification (2010): 60H10, 65C30, 65C20, 65L20

1 Introduction and Main Results

In this manuscript, we consider the strong approximation of time-homogeneous Itō-stochastic differential equations (SDEs) of the form

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t, \quad t \geq 0, \qquad X_0 = x, \tag{1}$$

where $x \in \mathbb{R}^d$ is the initial value, $\mu\colon \mathbb{R}^d \to \mathbb{R}^d$ is the drift coefficient, $\sigma\colon \mathbb{R}^d \to \mathbb{R}^{d,d}$ is the diffusion coefficient and $W = (W_t)_{t\geq 0}$ is a d-dimensional Brownian motion. In contrast to most of the analysis in the literature, we allow

(i) the drift coefficient μ to be discontinuous on a hypersurface Θ,

(ii) and the diffusion coefficient σ to be degenerate outside Θ.

Our aim is to construct an easily implementable numerical scheme which has root mean square convergence order 1/2 in terms of the computational cost for a large class of SDEs. So far only the transformation-based Euler-Maruyama scheme given in [23] for SDE (1) is known to have this property.

To state our main results denote the distance to the exceptional set Θ by

$$d(x, \Theta) = \inf\{\|x - y\| : y \in \Theta\}, \quad x \in \mathbb{R}^d,$$

and for every $\varepsilon > 0$ define

$$\Theta^{\varepsilon} := \{x \in \mathbb{R}^d : d(x, \Theta) < \varepsilon\}.$$

We consider the adaptive Euler-Maruyama scheme given by

$$X_0^h = x, \qquad \tau_0 = 0,$$

and

$$X^h_{\tau_{k+1}} = X^h_{\tau_k} + \mu(X^h_{\tau_k})(\tau_{k+1} - \tau_k) + \sigma(X^h_{\tau_k})(W_{\tau_{k+1}} - W_{\tau_k}), \qquad \tau_{k+1} = \tau_k + h(X^h_{\tau_k}, \delta),$$

with $k \in \mathbb{N}_0$, and step size function $h\colon \mathbb{R}^d \times (0,1) \to (0,1)$,

$$h(x,\delta) = \begin{cases} \delta^2, & x \in \Theta^{\varepsilon_2}, \\ \frac{1}{\sup_{x\in\Theta^{\varepsilon_0}} \|\sigma(x)\|^2} \left(\frac{d(x,\Theta)}{\log(1/\delta)}\right)^2, & x \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}, \\ \delta, & x \notin \Theta^{\varepsilon_1}, \end{cases}$$

where

$$\varepsilon_1 = \sup_{x\in\Theta^{\varepsilon_0}} \|\sigma(x)\| \log(1/\delta)\sqrt{\delta}, \qquad \varepsilon_2 = \sup_{x\in\Theta^{\varepsilon_0}} \|\sigma(x)\| \log(1/\delta)\delta,$$

with $\delta \in (0,1)$ and $\varepsilon_0 > 4\varepsilon_1 > 4\varepsilon_2$. Note that τ depends on $h(\cdot,\delta)$, but to simplify the notation we suppress this dependence. For mathematical convenience we will work with the continuous time Euler-Maruyama scheme, i.e. between discretization points we set

$$X^h_t = X^h_{\tau_k} + \mu(X^h_{\tau_k})(t - \tau_k) + \sigma(X^h_{\tau_k})(W_t - W_{\tau_k}), \quad t \in [\tau_k, \tau_{k+1}]. \tag{2}$$

Obviously this scheme uses smaller steps close to the discontinuities, has maximal step size δ, minimal step size δ^2, and interpolates both step sizes in an intermediate regime. The step sizing strategy arises from optimally balancing Gaussian tail estimates and occupation time estimates of the Euler-Maruyama scheme, which in particular accounts for the log-terms.

The computational cost of X^h on $[0,T]$, i.e. the number of arithmetic operations, function evaluations, and random numbers, is proportional to the number of steps which are needed to reach time T, that is

$$N(h,\delta) = \min\{k \in \mathbb{N}:\ \tau_k \geq T\}.$$

We will use this quantity as a proxy for the computational cost of the scheme.

We will work under mild assumptions, i.e.

(i) μ is supposed to be piecewise Lipschitz and its discontinuity set Θ is a sufficiently regular hypersurface,

(ii) σ is globally Lipschitz,

(iii) μ and σ satisfy a geometric smoothness and boundedness condition close to Θ,

see Assumption 2.1.

For fixed $T > 0$ we will show that

$$\left(\mathbb{E}\left[\sup_{s\in[0,T]} \|X_s - X^h_s\|^2\right]\right)^{1/2} \leq C_{\mathrm{rmse}} \cdot \sqrt{1 + \log(1/\delta)}\sqrt{\delta},$$

see Theorem 4.1, and

$$\mathbb{E}[N(h,\delta)] \leq C_{\mathrm{cost}} \cdot (1 + \log(1/\delta))\delta^{-1},$$

see Theorem 5.1, for some constants $C_{\mathrm{rmse}}, C_{\mathrm{cost}} > 0$ depending only on μ, σ, Θ, T, x. f So up to logarithmic terms the adaptive Euler-Maruyama scheme recovers the classical order 1/2 with respect to the average computational cost.

The remainder of this article is structured as follows: in the following subsections we briefly review recent results on the approximation of SDEs with discontinuous coefficients and adaptive numerical methods for SDEs. Sections 2 and 3 contain preliminary and auxiliary results, while Sections 4 and 5 contain the error and cost analysis. Section 6 provides some numerical examples.

1.1 Numerical methods for SDEs with discontinuous coefficients

Typically, existence and uniqueness results for SDEs only allow discontinuities in the drift coefficient, but not in the diffusion coefficient, see, e.g., [35] and the recent works [25, 34]. Thus – unless otherwise mentioned – the diffusion coefficient is globally Lipschitz for the following methods and results.

Up to the best of our knowledge the first contribution is [9]. In this work, the almost sure convergence for an SDE with possibly discontinuous drift coefficient is established, as long the drift coefficient is still one-sided Lipschitz, the diffusion coefficient is locally Lipschitz and there exists a Lyapunov function for the SDE. For SDEs with additive noise the results of [10] provide strong convergence of the Euler-Maruyama scheme for discontinuous, but monotone drift coefficients.

Recently several contributions for strong approximation have been given in a series of articles of Ngo and Taguchi [31, 32, 30] and Leobacher and Szölgyenyi [22, 23, 24]. For the multi-dimensional SDE (1) these works provide

(i) the L^1-convergence order $1/2$ for the equidistant Euler-Maruyama scheme, if μ is one-sided Lipschitz and an appropriate limit of smooth functions, and σ is bounded and uniformly non-degenerate, see [30],

(ii) the L^2-convergence order $1/4 - \epsilon$ for arbitrarily small $\epsilon > 0$ for the equidistant Euler-Maruyama scheme under Assumption 2.1 and additionally the boundedness of μ and σ, see [24],

(iii) the L^2-convergence order $1/2$ for a transformation based Euler-Maruyama method under Assumption 2.1, see [23]. However, this transformation is in general difficult to compute, which limits its applicability.

The weak approximation of SDEs with discontinuous coefficients has been studied in [18], where an Euler-type scheme based on an SDE with mollified drift coefficient is analyzed, and in [5], where an error bound for the density of the Euler-Maruyama scheme for (skew) diffusions is obtained. Finally, for scalar SDEs with additive noise, [2] provides a simulation scheme based on an approximation by skew perturbed SDEs.

1.2 Adaptive step sizing procedures for strong approximation of SDEs

Adaptive timestepping strategies have turned out to be very effective in probabilistic numerical analysis, though their error analysis typically provides mathematical challenges.

- The early works on adaptive methods propose strategies based on local error estimators in analogy to numerical methods for ordinary differential equations, see, e.g., [6], [1], and [20].
- Adaptive methods have also been used to preserve ergodic properties of the underlying SDE, see, e.g., [21, 27]. In fact, to recover ergodicity by adaptivity has already been proposed in [33].

- For optimal approximation of SDEs in the Information-Based-Complexity framework adaptive methods have been exhaustively analyzed in a series of articles [29, 28, 13, 12]. In these works, optimal convergence rates and asymptotically optimal schemes have been established for various error criteria.
- Finally, step adaptation strategies for SDEs with non-globally Lipschitz coefficients have been studied in [16, 3, 4]. This research has been partially motivated for the purpose of multilevel Monte Carlo simulations [26, 11].

2 Preliminaries

In this section we present some notions from differential geometry and analysis, state the assumptions we make on the coefficients of SDE (1) together with an existence and uniqueness result, and present a Krylov-type estimate for Itō processes.

All stochastic variables introduced in the following are assumed to be defined on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, where $\mathbb{F} = (\mathbb{F}_t)_{t\geq 0}$ is a normal filtration. In particular, W is a d-dimensional $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$-Brownian motion.

2.1 Definitions from differential geometry and analysis

In order to allow for discontinuities of the drift, we replace the usual global Lipschitz condition by the *piecewise Lipschitz condition*, which was first introduced in [23]. For this, we recall two definitions.

Definition 2.1 ([23, Definitions 3.1 and 3.2])**.** Let $A \subseteq \mathbb{R}^d$.

1. For a continuous curve $\gamma\colon [0,1] \to \mathbb{R}^d$, let $\ell(\gamma)$ denote its length, i.e.

$$\ell(\gamma) = \sup_{n\in\mathbb{N},\, 0\leq t_1<\ldots<t_n\leq 1} \sum_{k=1}^{n} \|\gamma(t_k) - \gamma(t_{k-1})\|.$$

   The *intrinsic metric* ρ on A is given by

$$\rho(x,y) := \inf\{\ell(\gamma)\colon \gamma\colon [0,1] \to A \text{ is a continuous curve satisfying } \gamma(0) = x,\, \gamma(1) = y\},$$

   where $\rho(x,y) := \infty$, if there is no continuous curve from x to y.

2. Let $f\colon A \to \mathbb{R}^m$ be a function. We say that f is *intrinsic Lipschitz*, if it is Lipschitz w.r.t. the intrinsic metric on A, i.e. if there exists a constant $L > 0$ such that

$$\forall x, y \in A\colon \|f(x) - f(y)\| \leq L\rho(x,y).$$

Of course every Lipschitz function is intrinsic Lipschitz, but the reverse does not hold.

Definition 2.2 ([23, Definition 3.4])**.** A function $f\colon \mathbb{R}^d \to \mathbb{R}^m$ is *piecewise Lipschitz*, if there exists a hypersurface Θ with finitely many connected components and with the property, that the restriction $f|_{\mathbb{R}^d\setminus\Theta}$ is intrinsic Lipschitz. We call Θ an *exceptional set* for f, and we call

$$\sup_{x,y\in\mathbb{R}^d\setminus\Theta} \frac{\|f(x) - f(y)\|}{\rho(x,y)}$$

the *piecewise Lipschitz constant* of f.

The following example shows, why it is necessary to resort to the intrinsic metric in the definition of the piecewise Lipschitz condition.

Example 2.3. Consider a function $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ which is piecewise Lipschitz and discontinuous at $\Theta = \{(x_1, x_2) \in \mathbb{R}^2 \colon x_1 < 0\}$. To obtain a Lipschitz estimate of, e.g., $\|f(-x_1, x_1) - f(-x_1, -x_1)\|$ for $x_1 \in \mathbb{R}$, the Euclidean metric cannot be used, since the direct connection of $(-x_1, x_1)$ and $(-x_1, -x_1)$ crosses Θ. The intrinsic metric provides a Lipschitz estimate with a connecting curve that lies in $\mathbb{R}^2 \backslash \Theta$.

In the following we consider piecewise Lipschitz functions with exceptional set Θ, where Θ is a fixed, sufficiently regular hypersurface, see Assumption 2.1.2 below. We denote the Lipschitz constant of a function f if it is finite, and otherwise its piecewise Lipschitz constant, by L_f. For a function $f\colon \mathbb{R}^d \to \mathbb{R}$ we denote $\|f\|_{\infty,\Theta^{\varepsilon_0}} := \sup_{x\in\Theta^{\varepsilon_0}} \|f(x)\|$.

Since $\Theta \in C^3$, locally there exists a unit normal vector, that is a continuously differentiable function $n\colon U \subseteq \Theta \to \mathbb{R}^d$ such that for every $\zeta \in U$, $\|n(\zeta)\| = 1$, and $n(\zeta)$ is orthogonal to the tangent space of Θ in ζ.

Recall the following definition from differential geometry.

Definition 2.4. Let $\Theta \subseteq \mathbb{R}^d$.

1. An environment Θ^ε is said to have the *unique closest point property*, if for every $x \in \mathbb{R}^d$ with $d(x, \Theta) < \varepsilon$ there is a unique $p \in \Theta$ with $d(x, \Theta) = \|x - p\|$. Therefore, we can define a mapping $p\colon \Theta^\varepsilon \to \Theta$ assigning to each x the point $p(x)$ in Θ, which is closest to x.

2. A set Θ is said to be of *positive reach*, if there exists $\varepsilon > 0$ such that Θ^ε has the unique closest point property. The *reach* r_Θ of Θ is the supremum over all such ε if such an ε exists, and 0 otherwise.

2.2 Existence of a unique strong solution

The main results in this paper require the following set of assumptions.

Assumption 2.1. We assume for the coefficients $\mu\colon \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma\colon \mathbb{R}^d \to \mathbb{R}^{d,d}$ of SDE (1):

1. the diffusion coefficient σ is Lipschitz;
2. the drift coefficient μ is a piecewise Lipschitz function; its exceptional set Θ is a C^3-hypersurface with reach $r_\Theta > \varepsilon_0$ for some $\varepsilon_0 > 0$;
3. the coefficients μ and σ satisfy
$$\sup_{x\in\Theta^{\varepsilon_0}} (\|\mu(x)\| + \|\sigma(x)\|) < \infty;$$
4. *(non-parallelity condition)* there exists a constant $c_0 > 0$ such that $\|\sigma(\xi)^\top n(\xi)\| \geq c_0$ for all $\xi \in \Theta$;
5. the function
$$\alpha : \Theta \to \mathbb{R}^d, \qquad \alpha(\xi) = \lim_{h\to 0} \frac{\mu(\xi - hn(\xi)) - \mu(\xi + hn(\xi))}{2\|\sigma(\xi)^\top n(\xi)\|^2}$$
is well defined, bounded, and belongs to $C_b^3(\Theta; \mathbb{R}^d)$.

Note that μ and σ satisfy a linear growth condition due to Assumptions 2.1.1, 2.1.2, and 2.1.3.

Remark on Assumption 2.1:

1. Assumption 2.1.2 is needed to be able to locally flatten Θ to a plane in a regular way. Furthermore, it guarantees that $n'(\xi)$ is bounded for all $\xi \in \Theta$, see [23, Lemma 3.10].
2. Assumption 2.1.4 ensures that $\sigma(\xi)$ has a component orthogonal to Θ for all $\xi \in \Theta$. It is significantly weaker than the uniform ellipticity condition which is usually required in the literature on SDEs with discontinuous drift.
3. Assumption 2.1.5 is a technical condition that is required for the transformation method from [23], which is the basis of our convergence proof, to work. Assumption 2.1.5 is satisfied if, e.g.,

   (i) the exceptional set Θ is a compact set. Note that then its complement satisfies $\mathbb{R}^d \backslash \Theta = A_1 \cup \cdots \cup A_n$ where $A_1, \ldots, A_n$ are open and connected subsets of $\mathbb{R}^d$,

   (ii) there exist Lipschitz C^3-functions $\mu_1, \ldots, \mu_n \colon \mathbb{R}^d \longrightarrow \mathbb{R}^d$ such that $\mu = \sum_{k=1}^n \mathbf{1}_{A_k} \mu_k$ and σ is Lipschitz and C^3,

   (iii) and Assumptions 2.1.3 and 2.1.4 hold.

   Compare Example 2.6 in [24].

Theorem 2.5 ([23, Theorem 3.21])**.** *Let Assumption 2.1 hold. Then SDE* (1) *has a unique strong solution.*

The proof of the above theorem and also the proof of our convergence result rely on a mapping $G \colon \mathbb{R}^d \to \mathbb{R}^d$, which transforms the SDE for X in another SDE which has Lipschitz coefficients. More precisely, we define

$$G(x) = \begin{cases} x + \varphi(x)\alpha(p(x)), & x \in \Theta^{\varepsilon_0}, \\ x, & x \in \mathbb{R}^d \backslash \Theta^{\varepsilon_0}, \end{cases}$$

with $r_\Theta > \varepsilon_0 > 0$, see Assumption 2.1.2, α as in Assumption 2.1.5, and

$$\varphi(x) = n(p(x))^\top (x - p(x)) \|x - p(x)\| \phi \left(\frac{\|x - p(x)\|}{c} \right),$$

with a constant $c > 0$ and $\phi \colon \mathbb{R} \to \mathbb{R}$,

$$\phi(u) = \begin{cases} (1+u)^3(1-u)^3, & |u| \leq 1, \\ 0, & |u| > 1. \end{cases}$$

The map G has the following properties:

Lemma 2.6. *Let Assumption 2.1 hold. Then we have*

(i) $G \in C^1(\mathbb{R}^d, \mathbb{R}^d)$*;*

(ii) G'' *is piecewise Lipschitz with exceptional set* Θ*;*

(iii) G' *and* G'' *are bounded;*

(iv) for c *sufficiently small, see [23, Lemma 3.18],* G *is globally invertible;*

(v) G *and* G^{-1} *are Lipschitz continuous;*

(vi) Itō's formula holds for G *and* G^{-1}*.*

Proof. For (i) and (iv) see [23, Theorem 3.14]. Assertion (v) follows from the proof of [23, Theorem 3.20] and for (vi) see [23, Theorem 3.19]. Moreover, assertion (ii) follows again from the proof of [23, Theorem 3.20].

The boundedness of G' follows from (v), so it remains to show that G'' is bounded on $\mathbb{R}^d \setminus \Theta$. By construction we have that $G''(x) = 0$ on $\mathbb{R}^d \setminus \Theta^{\varepsilon_0}$. Moreover, ϕ, ϕ', ϕ'' are bounded as α, α' and α''. So, for the boundedness of G'' we have to study the projection map $p : \Theta^{\varepsilon_0} \to \Theta$ and the signed distance

$$D : \Theta^{\varepsilon_0} \to [0, \infty), \quad D(x) = n(p(x))^\top (x - p(x)),$$

since we have the identity

$$G(x) = x + D(x)|D(x)|\big(1 - D(x)^2\big)^3, \quad x \in \Theta^{\varepsilon_0}.$$

The projection map satisfies

$$p'(x) = \mathrm{id}_{\mathbb{R}^d} - n(p(x))n(p(x))^\top,$$

see [23, Proof of Theorem 3.14, Equation (7)], and consequently p' and p'' are bounded due to the boundedness of $n'(\xi)$ for all $\xi \in \Theta$, see [23, Lemma 3.10]. Since

$$\frac{d}{dx}\big(n(p(x))^\top (x - p(x))\big) = n(p(x))^\top$$

by [24, Lemma 2.8], we also have that $D \in C^2(\Theta^{\varepsilon_0}; \mathbb{R})$, and

$$\sup_{x \in \Theta^{\varepsilon_0}} \|D'(x)\| + \sup_{x \in \Theta^{\varepsilon_0}} \|D''(x)\| < \infty.$$

With this, the boundedness of G'' follows from the product rule and the chain rule. ☐

Now, define the coefficients

$$\begin{aligned} \mu_G(z) &= G'(G^{-1}(z))\mu(G^{-1}(z)) + \frac{1}{2}\operatorname{tr}\Big[\sigma(G^{-1}(z))^\top G''(G^{-1}(z))\sigma(G^{-1}(z))\Big], \\ \sigma_G(z) &= G'(G^{-1}(z))\sigma(G^{-1}(z)), \end{aligned} \tag{3}$$

for $z \in \mathbb{R}^d$.

Lemma 2.7 ([23, Theorem 3.20])**.** *The functions $\mu_G \colon \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma_G \colon \mathbb{R}^d \to \mathbb{R}^{d,d}$ are globally Lipschitz.*

Thus, the SDE

$$dZ_t = \mu_G(Z_t)dt + \sigma_G(Z_t)dW_t, \quad t \geq 0, \qquad Z_0 = G(x), \tag{4}$$

has a unique strong solution $Z = (Z_t)_{t \geq 0}$. Moreover, $X = (X_t)_{t \geq 0}$ given by $X_t = G^{-1}(Z_t)$ solves SDE (1), which follows from an application of Itō's formula.

2.3 Occupation time estimates for Itō processes

In this subsection we study the occupation time of an Itō process close to a C^3-hypersurface. The following result is a slight extension of [24, Theorem 2.7] and is sometimes refereed to as Krylov's estimate. While classically Krylov-estimates are derived for non-degenerate diffusions, see [19], the following result only assumes non-degeneracy of the diffusion coefficient at a normal direction within an environment Θ^{ε_0} of the hypersurface Θ.

Theorem 2.8. *Let Θ be a C^3-hypersurface of positive reach and let $r_\Theta > \epsilon_0 > 0$. Let further $A = (A_t)_{t\geq 0}$, $B = (B_t)_{t\geq 0}$ be $\mathbb{R}^d$, respectively $\mathbb{R}^{d,d}$-valued progressively measurable processes such that*

$$\int_0^t \mathbb{E}\left[\|A_s\| + \|B_s\|^2\right] ds < \infty, \quad t \geq 0.$$

Moreover, let $X = (X_t)_{t\geq 0}$ be the $\mathbb{R}^d$-valued Itō process given by

$$X_t = X_0 + \int_0^t A_s ds + \int_0^t B_s dW_s, \quad t \geq 0,$$

with $X_0 \in \mathbb{R}^d$. Assume finally that

(i) there exists a constant $c_{AB} > 0$ such that for almost all $\omega \in \Omega$ we have

$$\forall t \in [0,T]: \ X_t(\omega) \in \Theta^{\epsilon_0} \Longrightarrow \max(\|A_t(\omega)\|, \|B_t(\omega)\|) \leq c_{AB};$$

(ii) there exists a constant $c_0 > 0$ such that for almost all $\omega \in \Omega$ we have

$$\forall t \in [0,T]: X_t(\omega) \in \Theta^{\epsilon_0} \Longrightarrow n(p(X_t(\omega)))^\top B_t(\omega) B_t(\omega)^\top n(p(X_t(\omega))) \geq c_0^2.$$

Then there exists a constant $C > 0$ such that for all $0 < \epsilon < \epsilon_0/2$ and any measurable function $f: [0,\infty) \to [0,\infty)$ we have

$$\mathbb{E}\left[\int_0^T f(d(X_s,\Theta)) \mathbf{1}_{\{X_s \in \Theta^\epsilon\}} ds\right] \leq C \int_0^\epsilon f(x) dx.$$

Proof. The proof relies on [24, Theorem 2.7], where a scalar process Y is constructed such that the occupation time of Y in an environment of $\{0\}$ is the same as the occupation time of X in an environment of Θ. More precisely, there exists a bounded real-valued Itō process

$$Y_t = Y_0 + \int_0^t \hat{A}_s ds + \int_0^t \hat{B}_s dW_s, \quad t \geq 0,$$

where $\hat{A}$, $\hat{B}$ are uniformly bounded, progressively measurable $\mathbb{R}$, respectively $\mathbb{R}^{1,d}$-valued processes, such that

$$Y_t \cdot \mathbf{1}_{\{X_t \in \Theta^{\epsilon_1}\}} = \lambda(D(X_t)) \cdot \mathbf{1}_{\{X_t \in \Theta^{\epsilon_1}\}}, \quad t \geq 0,$$

where $D(x) = n(p(x))^\top (x - p(x))$ for $x \in \Theta^{\epsilon_0}$, $\epsilon_1 = \epsilon_0/2$, and $\lambda: \mathbb{R} \to \mathbb{R}$ is given by

$$\lambda(z) = \begin{cases} z - \frac{2}{3\epsilon_1^2} z^3 + \frac{1}{5\epsilon_1^4} z^5, & |z| \leq \epsilon_1, \\ \frac{8\epsilon_1}{15}, & z > \epsilon_1, \\ -\frac{8\epsilon_1}{15}, & z < -\epsilon_1. \end{cases}$$

Since $|D(X_t)| = d(X_t,\Theta)$, the value of Y corresponds to the λ-transformed signed distance of X to Θ. Since $\lambda'(\pm\epsilon_1) = \lambda''(\pm\epsilon_1) = 0$, it holds that $\lambda \in C^2$. Moreover, $\lambda\colon [-\epsilon_1, \epsilon_1] \to [\lambda(-\epsilon_1), \lambda(\epsilon_1)]$ is invertible.

By construction the quadratic variation of Y satisfies $\mathbb{P}$-a.s. that

$$\int_0^t \mathbf{1}_{\{Y_s \in (-\lambda(\epsilon), \lambda(\epsilon))\}} d[Y]_s = \int_0^t \mathbf{1}_{\{Y_s \in (-\lambda(\epsilon), \lambda(\epsilon))\}} \lambda'\left(D(X_t)\right)^2 n(p(X_s))^\top B_s B_s^\top n(p(X_s)) ds, \ t \geq 0,$$

see the proof of [24, Theorem 2.7]. Since

$$|\lambda^{-1}(Y_t)| \cdot \mathbf{1}_{\{X_t \in \Theta^{\epsilon_1}\}} = \big|D(X_t)\big| \cdot \mathbf{1}_{\{X_t \in \Theta^{\epsilon_1}\}} = d(X_t, \Theta) \cdot \mathbf{1}_{\{X_t \in \Theta^{\epsilon_1}\}}, \quad t \geq 0,$$

and $\lambda'(z) \geq \left(\frac{3}{4}\right)^2$ for all $|z| \leq \epsilon \leq \epsilon_0/2$, Assumption (ii) assures that

$$\mathbb{E}\left[\int_0^T f(d(X_s,\theta))\mathbf{1}_{\{X_s \in \Theta^\epsilon\}} ds\right] \leq \left(\frac{4}{3c_0}\right)^2 \mathbb{E}\left[\int_0^T f(|\lambda^{-1}(Y_s)|)\mathbf{1}_{\{Y_s \in (-\lambda(\epsilon),\lambda(\epsilon))\}} d\,[Y]_s\right].$$

Therefore, the occupation time formula [15, Chapter 3, 7.1 Theorem] for one-dimensional continuous semimartingales yields

$$\begin{aligned}\mathbb{E}\left[\int_0^T f(d(X_s,\theta))\mathbf{1}_{\{X_s \in \Theta^\epsilon\}} ds\right] &\leq \left(\frac{4}{3c_0}\right)^2 \mathbb{E}\left[\int_0^T f(|\lambda^{-1}(Y_s)|)\mathbf{1}_{\{Y_s \in (-\lambda(\epsilon),\lambda(\epsilon))\}} d\,[Y]_s\right] \\ &= \left(\frac{4}{3c_0}\right)^2 \mathbb{E}\left[\int_{\mathbb{R}} f(|\lambda^{-1}(y)|)\mathbf{1}_{(-\lambda(\epsilon),\lambda(\epsilon))}(y) L_T^y\,(Y)\,dy\right].\end{aligned}$$

The local time $L_T^y(Y), y \in \mathbb{R}$ of a real-valued Itō process with bounded coefficients satisfies

$$\sup_{y\in\mathbb{R}} \mathbb{E}\left[L_T^y\,(Y)\right] < \infty.$$

Thus, it follows that

$$\mathbb{E}\left[\int_0^T f(d(X_s,\theta))\mathbf{1}_{\{X_s \in \Theta^\epsilon\}} ds\right] \leq \frac{2^5}{3^2 c_0^2} \sup_{y\in\mathbb{R}} \mathbb{E}\left[L_T^y\,(Y)\right] \int_0^{\lambda(\epsilon)} f(\lambda^{-1}(x))dx.$$

Since $0 \leq \lambda'(x) \leq 1$ for all $x \in \mathbb{R}$, substitution yields

$$\mathbb{E}\left[\int_0^T f(d(X_s,\theta))\mathbf{1}_{\{X_s \in \Theta^\epsilon\}} ds\right] \leq \frac{2^5}{3^2 c_0^2} \sup_{y\in\mathbb{R}} \mathbb{E}\left[L_T^y\,(Y)\right] \int_0^{\epsilon} f(x)dx,$$

which shows the assertion. □

3 Properties of the adaptive Euler-Maruyama scheme

3.1 General properties

Recall that our Euler-Maruyama scheme is given by

$$\tau_0 = 0, \qquad X_0^h = x \in \mathbb{R}^d, \tag{5}$$

and

$$\tau_{k+1} = \tau_k + h(X_{\tau_k}^h, \delta), \qquad X_{\tau_{k+1}}^h = X_{\tau_k}^h + \mu(X_{\tau_k}^h)(t-\tau_k) + \sigma(X_{\tau_k}^h)(W_t - W_{\tau_k}), \quad t \in (\tau_k, \tau_{k+1}], \tag{6}$$

with $k \in \mathbb{N}_0$, and $h\colon \mathbb{R}^d \times (0,1) \to (0,1)$,

$$h(x,\delta) = \begin{cases} \delta^2, & x \in \Theta^{\varepsilon_2}, \\ \frac{1}{\|\sigma\|^2_{\infty,\Theta^{\varepsilon_0}}} \left(\frac{d(x,\Theta)}{\log(1/\delta)}\right)^2, & x \in \Theta^{\varepsilon_1} \backslash \Theta^{\varepsilon_2}, \\ \delta, & x \notin \Theta^{\varepsilon_1}, \end{cases} \tag{7}$$

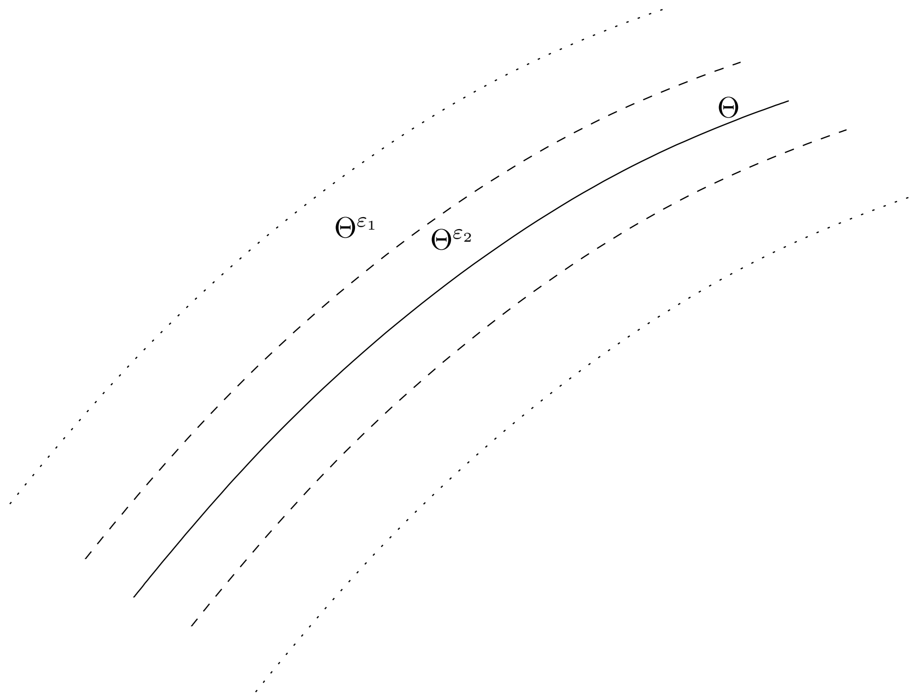

Figure 1: The three step size regimes.

where

$$\varepsilon_1 = \|\sigma\|_{\infty,\Theta^{\varepsilon_0}} \log(1/\delta)\sqrt{\delta}, \qquad \varepsilon_2 = \|\sigma\|_{\infty,\Theta^{\varepsilon_0}} \log(1/\delta)\delta. \tag{8}$$

Moreover, set $\underline{t} = \max\{\tau_k \colon \tau_k \leq t\}$.

Figure 1 illustrates the different step size regimes. The solid line is the set of discontinuities Θ of the drift, the area between the two dashed lines is Θ^{ε_2}, and the area between the dotted lines is Θ^{ε_1}.

Framework 3.1. We assume that

(i) Assumption 2.1 holds,

(ii) the step size δ is sufficiently small such that

$$\varepsilon_1 = \|\sigma\|_{\infty,\Theta^{\varepsilon_0}} \log(1/\delta)\sqrt{\delta} < \varepsilon_0/4,$$

(iii) c is sufficiently small, see [23, Lemma 3.18], in particular, $c < \varepsilon_0$, such that G is globally invertible.

For fixed $\delta \in (0,1)$ define the mapping

$$\Phi : \mathbb{R}^d \times C([0,\infty);\mathbb{R}^d) \to C([0,\infty);\mathbb{R}^d), \qquad \Phi(x,W) = (X^h_s)_{s\geq 0}.$$

Note that by construction Φ is $\mathcal{B}(\mathbb{R}^d \times C([0,\infty);\mathbb{R}^d)) - \mathcal{B}(C([0,\infty);\mathbb{R}^d))$ measurable. Also by construction our discretization points τ_k are stopping times, i.e. they satisfy

$$\{\tau_k \leq u\} \in \mathbb{F}_u, \quad u \geq 0, \quad k \in \mathbb{N}_0.$$

This can be shown by induction since τ_{k+1} is $X^h_{\tau_k}$ measurable and X_0 is deterministic. The strong Markov property of Brownian motion then implies that the process

$$(W^{\tau_k}_t)_{t\geq 0} = (W_{t+\tau_k} - W_{\tau_k})_{t\geq 0}$$

is again a Brownian motion and independent of X_{τ_k} for all $k \in \mathbb{N}_0$.

Lemma 3.1. *Assume Framework 3.1, let $F \in \mathcal{B}(C([0,\infty);\mathbb{R}^d))$, and $k \in \mathbb{N}_0$. Then we have*

$$\mathbb{P}((X^h_{t+\tau_k})_{t\geq 0} \in F | X^h_{\tau_k} = y) = \mathbb{P}(\Phi(y,W) \in F) \quad \textit{for} \quad \mathbb{P}^{X_{\tau_k}}\textit{-almost all } y \in \mathbb{R}^d.$$

Proof. Using Φ and W^{τ_k} we can write

$$X^h_{t+\tau_k} = X^h_{\tau_k} + \mu(X^h_{\tau_k})t + \sigma(X^h_{\tau_k})W^{\tau_k}_t = \Phi(X^h_{\tau_k}, W^{\tau_k})(t), \quad t \in [0, \tau_{k+1} - \tau_k].$$

Proceeding iteratively we obtain

$$(X^h_{t+\tau_k})_{t\geq 0} = (\Phi(X^h_{\tau_k}, W^{\tau_k}))_{t\geq 0}.$$

Since W^{τ_k} is a Brownian motion which is independent of $X^h_{\tau_k}$, the Factorization Lemma for conditional expectations concludes the proof. □

Lemma 3.2. *Assume Framework 3.1 and let $T > 0$. For any $p \geq 2$ we have*

$$\mathbb{E}\Big[\sup_{t\in[0,T]} \|X^h_t\|^p\Big] < \infty$$

and there exists a constant $C > 0$ such that

$$\mathbb{E}[\|X^h_t - X^h_s\|^p] \leq C \cdot |t-s|^{p/2}, \quad 0 \leq s < t \leq T.$$

Proof. For the first statement note that

$$\sup_{t\in[0,s]} \|X^h_t\|^p \leq 3^{p-1}\|x\|^p + 3^{p-1} \sup_{t\in[0,s]} \left\|\int_0^t \mu(X^h_{\underline{u}})du\right\|^p + 3^{p-1} \sup_{t\in[0,s]} \left\|\int_0^t \sigma(X^h_{\underline{u}})dW_u\right\|^p, \quad s \in [0,T].$$

Jensen's inequality for the Riemann-integral, the Burkholder-Davis-Gundy inequality for the Itō-integral, and the linear growth condition on μ and σ ensure the existence of a constant $c > 0$, which depends on $T > 0$, μ, σ, p, and x, such that

$$\mathbb{E}\Big[\sup_{t\in[0,s]} \|X^h_t\|^p\Big] \leq c + c\int_0^s \mathbb{E}\Big[\sup_{t\in[0,u]} \|X^h_t\|^p\Big] du.$$

Gronwall's Lemma now yields the assertion. For the second statement note that

$$\|X^h_t - X^h_s\|^p \leq 2^{p-1} \left\|\int_s^t \mu(X^h_{\underline{s}})ds\right\|^p + 2^{p-1} \left\|\int_s^t \sigma(X^h_{\underline{s}})dW_s\right\|^p, \quad s,t \in [0,T].$$

Jensen's inequality, the Burkholder-Davis-Gundy inequality, and the linear growth condition of the coefficients together with the first assertion yield the statement. □

Applying Theorem 2.8 to our Euler-Maruyama scheme we obtain:

Lemma 3.3. *Assume Framework 3.1 and let $T > 0$. Moreover, let $\varepsilon < \varepsilon_0/2$ and $f\colon [0,\infty) \to [0,\infty)$ be a measurable function. Then there exists a constant $C > 0$ such that*

$$\mathbb{E}\left[\int_0^T f(d(X^h_t, \Theta))\mathbf{1}_{\{X^h_t \in \Theta^\varepsilon\}}dt\right] \leq C\int_0^\varepsilon f(x)dx.$$

In particular, we have

$$\int_0^T \mathbb{P}(X^h_t \in \Theta^\varepsilon)dt \leq C \cdot \varepsilon.$$

3.2 Exit time estimates for the adaptive Euler-Maruyama scheme

In this subsection we present exit time estimates for our Euler-Maruyama scheme. For this, we require the following Lemma.

Lemma 3.4. *There exists a constant $C_{tail} > 0$ such that*

$$\mathbb{P}\left(\sup_{0\le s\le t}\|W_s\| \ge \varepsilon\right) \le C_{tail}\cdot\exp\left(-\frac{\varepsilon}{\sqrt{t}}\right), \quad \varepsilon > 0, \quad t \ge 0.$$

Proof. Using the scaling property of Brownian motion and applying Doob's submartingale inequality we obtain

$$\begin{aligned}\mathbb{P}\left(\sup_{0\le s\le t}\|W_s\| \ge \varepsilon\right) &= \mathbb{P}\left(\sup_{0\le s\le 1}\|W_s\| \ge \frac{\varepsilon}{\sqrt{t}}\right)\\ &= \mathbb{P}\left(\exp\left(\sup_{0\le s\le 1}\|W_s\|\right) \ge \exp\left(\frac{\varepsilon}{\sqrt{t}}\right)\right) \le \mathbb{E}\left[\exp(\|W_1\|)\right]\exp\left(-\frac{\varepsilon}{\sqrt{t}}\right).\end{aligned}$$

□

The next lemma controls the probabilities that the Euler-Maruyama scheme has increments that are relatively large compared to its distance from Θ.

Lemma 3.5. *Assume Framework 3.1 and let $T > 0$. Then there exists a constant $C > 0$, independent of δ, such that*

(i) $\int_0^T \mathbb{P}\big(X_t^h \notin \Theta^{2\varepsilon_2}; X_{\underline{t}}^h \in \Theta^{\varepsilon_2}\big)dt \le C\cdot\delta,$

(ii) $\int_0^T \mathbb{P}\big(\|X_t^h - X_{\underline{t}}^h\| \ge d(X_{\underline{t}}^h,\Theta); X_{\underline{t}}^h \in \Theta^{\varepsilon_1}\setminus\Theta^{\varepsilon_2}\big)dt \le C\cdot\delta,$

(iii) $\int_0^T \mathbb{P}\big(\|X_t^h - X_{\underline{t}}^h\| \ge \varepsilon_1; X_{\underline{t}}^h \in \Theta^{\varepsilon_0}\setminus\Theta^{\varepsilon_1}\big)dt \le C\cdot\delta.$

Proof. (i) Note that

$$\left\{\omega\in\Omega : X_t^h(\omega)\notin\Theta^{2\varepsilon_2}; X_{\underline{t}}^h(\omega)\in\Theta^{\varepsilon_2}\right\} \subseteq \left\{\omega\in\Omega : \|X_t^h(\omega) - X_{\underline{t}}^h(\omega)\| \ge \varepsilon_2; X_{\underline{t}}^h(\omega)\in\Theta^{\varepsilon_2}\right\}$$

and that

$$\begin{aligned}\int_0^T \mathbb{P}\big(\|X_t^h - X_{\underline{t}}^h\| \ge \varepsilon_2; X_{\underline{t}}^h\in\Theta^{\varepsilon_2}\big)dt &\le \int_0^T \mathbb{P}\left(\sup_{t\in[\underline{t},\underline{t}+h(X_{\underline{t}}^h,\delta)]}\|X_t^h - X_{\underline{t}}^h\| \ge \varepsilon_2; X_{\underline{t}}^h\in\Theta^{\varepsilon_2}\right)dt\\ &= \int_0^T \mathbb{P}\left(\sup_{t\in[\underline{t},\underline{t}+h(X_{\underline{t}}^h,\delta)]}\|X_t^h - X_{\underline{t}}^h\| \ge \varepsilon_2 \,\middle|\, X_{\underline{t}}^h\in\Theta^{\varepsilon_2}\right)\mathbb{P}(X_{\underline{t}}^h\in\Theta^{\varepsilon_2})dt.\end{aligned}$$

Here we set the value of the above conditional probability to zero, if $\mathbb{P}(X_{\underline{t}}^h \in \Theta^{\varepsilon_2}) = 0$. By Lemma 3.1 we obtain that

$$\begin{aligned}&\int_0^T \mathbb{P}\left(\sup_{t\in[\underline{t},\underline{t}+h(X_{\underline{t}}^h,\delta)]}\|X_t^h - X_{\underline{t}}^h\| \ge \varepsilon_2 \,\middle|\, X_{\underline{t}}^h\in\Theta^{\varepsilon_2}\right)\mathbb{P}(X_{\underline{t}}^h\in\Theta^{\varepsilon_2})dt\\ &\quad= \int_0^T\int_{\Theta^{\varepsilon_2}} \mathbb{P}\left(\sup_{t\in[0,\delta^2]}\|\Phi(y,W^{\underline{t}})(t) - y\| \ge \varepsilon_2\right)\mathbb{P}(X_{\underline{t}}^h\in dy)dydt\\ &\quad\le \int_0^T\int_{\Theta^{\varepsilon_2}} \mathbb{P}\left(\sup_{t\in[0,1]}\|W_t\| \ge \frac{\varepsilon_2 - \delta^2\|\mu\|_{\infty,\Theta^{\varepsilon_0}}}{\delta\|\sigma\|_{\infty,\Theta^{\varepsilon_0}}}\right)\mathbb{P}(X_{\underline{t}}^h\in dy)dydt.\end{aligned}$$

Recall that $\varepsilon_2 = \|\sigma\|_{\infty,\Theta^{\varepsilon_0}} \log(1/\delta)\delta$, and hence

$$\frac{\varepsilon_2 - \delta^2\|\mu\|_{\infty,\Theta^{\varepsilon_0}}}{\delta\|\sigma\|_{\infty,\Theta^{\varepsilon_0}}} = -\log(\delta) - \frac{\|\mu\|_{\infty,\Theta^{\varepsilon_0}}}{\|\sigma\|_{\infty,\Theta^{\varepsilon_0}}}\delta \geq -\log(\delta) - \frac{\|\mu\|_{\infty,\Theta^{\varepsilon_0}}}{\|\sigma\|_{\infty,\Theta^{\varepsilon_0}}}. \tag{9}$$

An application of Lemma 3.4 together with (9) now yields

$$\begin{aligned}\int_0^T \int_{\Theta^{\varepsilon_2}} &\mathbb{P}\left(\sup_{t\in[0,1]} \|W_t\| \geq \frac{\varepsilon_2 - \delta^2\|\mu\|_{\infty,\Theta^{\varepsilon_0}}}{\delta\|\sigma\|_{\infty,\Theta^{\varepsilon_0}}}\right)\mathbb{P}(X^h_{\underline{t}} \in dy)dydt \\ &\leq C_{\text{tail}}\int_0^T \exp\left(\log(\delta)\right)\exp\left(\frac{\|\mu\|_{\infty,\Theta^{\varepsilon_0}}}{\|\sigma\|_{\infty,\Theta^{\varepsilon_0}}}\right)\int_{\Theta^{\varepsilon_2}} \mathbb{P}(X^h_{\underline{t}} \in dy)dydt \leq C\cdot\delta,\end{aligned}$$

with

$$C = C_{\text{tail}} T \exp\left(\frac{\|\mu\|_{\infty,\Theta^{\varepsilon_0}}}{\|\sigma\|_{\infty,\Theta^{\varepsilon_0}}}\right).$$

(ii) Observe that

$$\begin{aligned}&\int_0^T \mathbb{P}\big(\|X^h_t - X^h_{\underline{t}}\| \geq d(X^h_{\underline{t}},\Theta); X^h_{\underline{t}} \in \Theta^{\varepsilon_1}\setminus\Theta^{\varepsilon_2}\big)dt \\ &\quad\leq \int_0^T \mathbb{P}\left(\sup_{t\in[\underline{t},\underline{t}+h(X^h_{\underline{t}},\delta)]} \|X^h_t - X^h_{\underline{t}}\| \geq d(X^h_{\underline{t}},\Theta); X^h_{\underline{t}} \in \Theta^{\varepsilon_1}\setminus\Theta^{\varepsilon_2}\right)dt \\ &\quad= \int_0^T \mathbb{P}\left(\sup_{t\in[\underline{t},\underline{t}+h(X^h_{\underline{t}},\delta)]} \|X^h_t - X^h_{\underline{t}}\| \geq d(X^h_{\underline{t}},\Theta)\Big| X^h_{\underline{t}} \in \Theta^{\varepsilon_1}\setminus\Theta^{\varepsilon_2}\right)\mathbb{P}(X^h_{\underline{t}} \in \Theta^{\varepsilon_1}\setminus\Theta^{\varepsilon_2})dt.\end{aligned}$$

Using again Lemma 3.1 we obtain

$$\begin{aligned}\int_0^T &\mathbb{P}\left(\sup_{t\in[\underline{t},\underline{t}+h(X^h_{\underline{t}},\delta)]} \|X^h_t - X^h_{\underline{t}}\| \geq d(X^h_{\underline{t}},\Theta)\Big| X^h_{\underline{t}} \in \Theta^{\varepsilon_1}\setminus\Theta^{\varepsilon_2}\right)\mathbb{P}(X^h_{\underline{t}} \in \Theta^{\varepsilon_1}\setminus\Theta^{\varepsilon_2})dt \\ &= \int_0^T\int_{\Theta^{\varepsilon_1}\setminus\Theta^{\varepsilon_2}} \mathbb{P}\left(\sup_{t\in[0,h(y,\delta)]} \|\Phi(y,W^{\underline{t}})(t) - y\| \geq d(y,\Theta)\right)\mathbb{P}\big(X^h_{\underline{t}} \in dy\big)dydt \\ &\leq \int_0^T\int_{\Theta^{\varepsilon_1}\setminus\Theta^{\varepsilon_2}} \mathbb{P}\left(\sup_{t\in[0,1]} \|W_t\| \geq \frac{d(y,\Theta) - \|\mu\|_{\infty,\Theta^{\varepsilon_0}}h(y,\delta)}{h(y,\delta)^{1/2}\|\sigma\|_{\infty,\Theta^{\varepsilon_0}}}\right)\mathbb{P}(X^h_{\underline{t}} \in dy)dydt.\end{aligned}$$

Since

$$h(x,\delta) = \frac{1}{\|\sigma\|^2_{\infty,\Theta^{\varepsilon_0}}}\left(\frac{d(x,\Theta)}{\log(1/\delta)}\right)^2, \qquad x \in \Theta^{\varepsilon_1}\setminus\Theta^{\varepsilon_2},$$

we arrive at

$$\frac{d(y,\Theta) - \|\mu\|_{\infty,\Theta^{\varepsilon_0}}h(y,\delta)}{h(y,\delta)^{1/2}\|\sigma\|_{\infty,\Theta^{\varepsilon_0}}} = -\log(\delta) - \frac{\|\mu\|_{\infty,\Theta^{\varepsilon_0}}}{\|\sigma\|_{\infty,\Theta^{\varepsilon_0}}}h(y,\delta)^{1/2} \geq -\log(\delta) - \frac{\|\mu\|_{\infty,\Theta^{\varepsilon_0}}}{\|\sigma\|_{\infty,\Theta^{\varepsilon_0}}}. \tag{10}$$

Lemma 3.4 together with (10) yields that

$$\int_0^T\int_{\Theta^{\varepsilon_1}\setminus\Theta^{\varepsilon_2}} \mathbb{P}\left(\sup_{t\in[0,1]} \|W_t\| \geq \frac{d(y,\Theta) - \|\mu\|_{\infty,\Theta^{\varepsilon_0}}h(y,\delta)}{h(y,\delta)^{1/2}\|\sigma\|_{\infty,\Theta^{\varepsilon_0}}}\right)\mathbb{P}(X^h_{\underline{t}} \in dy)dt \leq C\cdot\delta.$$

(iii) This can be shown along the same lines as (ii) taking into account that the Euler-Maruyama scheme under the condition $X^h_{\underline{t}} \notin \Theta^{\varepsilon_1}$ has step size δ and that $\varepsilon_1 = \|\sigma\|_{\infty,\Theta^{\varepsilon_0}}\log(1/\delta)\sqrt{\delta}$.
□

Using the previous lemmas we obtain that the following lemma holds.

Lemma 3.6. *Assume Framework 3.1 and let $T > 0$. There exists a constant $C > 0$ such that*

$$\int_0^T \mathbb{P}\left(X_{\underline{t}}^h \in \Theta^{\varepsilon_2}\right) dt \leq \ C \cdot (1 + \log(1/\delta))\delta,$$

and

$$\int_0^T \mathbb{P}\left(X_{\underline{t}}^h \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}\right) dt \leq \ C \cdot (1 + \log(1/\delta))\sqrt{\delta}.$$

Proof. To establish the first estimate we write

$$\begin{aligned}\int_0^T \mathbb{P}\left(X_{\underline{t}}^h \in \Theta^{\varepsilon_2}\right) dt &= \int_0^T \mathbb{P}\left(X_{\underline{t}}^h \in \Theta^{\varepsilon_2}; X_t^h \notin \Theta^{2\varepsilon_2}\right) dt + \int_0^T \mathbb{P}\left(X_{\underline{t}}^h \in \Theta^{\varepsilon_2}; X_t^h \in \Theta^{2\varepsilon_2}\right) dt \\ &\leq \int_0^T \mathbb{P}\left(X_{\underline{t}}^h \in \Theta^{\varepsilon_2}; X_t^h \notin \Theta^{2\varepsilon_2}\right) dt + \int_0^T \mathbb{P}\left(X_t^h \in \Theta^{2\varepsilon_2}\right) dt.\end{aligned}$$

We conclude by Lemma 3.5(i) and Lemma 3.3. Analogously the second estimate follows from

$$\begin{aligned}&\int_0^T \mathbb{P}\left(X_{\underline{t}}^h \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}\right) dt \\ &= \int_0^T \mathbb{P}\left(X_{\underline{t}}^h \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}; X_t^h \notin \Theta^{2\varepsilon_1}\right) dt + \int_0^T \mathbb{P}\left(X_{\underline{t}}^h \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}; X_t^h \in \Theta^{2\varepsilon_1}\right) dt \\ &\leq \int_0^T \mathbb{P}\left(\|X_t^h - X_{\underline{t}}^h\| \geq d(X_{\underline{t}}^h, \Theta); X_{\underline{t}}^h \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}\right) + \int_0^T \mathbb{P}\left(X_t^h \in \Theta^{2\varepsilon_1}\right) dt.\end{aligned}$$

We conclude by Lemma 3.5(ii) and Lemma 3.3.

□

Finally we prove the following refinement of Lemma 3.5(ii).

Lemma 3.7. *Assume Framework 3.1, let $T > 0$ and $\alpha \in (0, 1]$. There exists a constant $C > 0$ such that*

$$\int_0^T \mathbb{P}(\|X_t^h - X_{\underline{t}}^h\| \geq \alpha \cdot d(X_{\underline{t}}^h, \Theta); X_{\underline{t}}^h \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}) dt \leq C \cdot (1 + \log(1/\delta))\delta^{\alpha + \frac{1}{2}}.$$

Proof. Here we follow the same lines as in the proof of Lemma 3.5(ii) to obtain

$$\begin{aligned}&\int_0^T \mathbb{P}\left(\|X_t^h - X_{\underline{t}}^h\| \geq \alpha \cdot d(X_{\underline{t}}^h, \Theta); X_{\underline{t}}^h \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}\right) dt \\ &\leq \int_0^T \int_{\Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}} \mathbb{P}\left(\sup_{t \in [0,1]} \|W_t\| \geq -\alpha \log(\delta) - \frac{\|\mu\|_{\infty, \Theta^{\varepsilon_0}}}{\|\sigma\|_{\infty, \Theta^{\varepsilon_0}}}\right) \mathbb{P}(X_{\underline{t}}^h \in dy) dy dt.\end{aligned}$$

Lemma 3.4 now yields that

$$\begin{aligned}&\int_0^T \mathbb{P}\left(\|X_t^h - X_{\underline{t}}^h\| \geq \alpha \cdot d(X_{\underline{t}}^h, \Theta); X_{\underline{t}}^h \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}\right) dt \\ &\leq C_{\text{tail}} \exp\left(\frac{\|\mu\|_{\infty, \Theta^{\varepsilon_0}}}{\|\sigma\|_{\infty, \Theta^{\varepsilon_0}}}\right) \delta^{\alpha} \int_0^T \mathbb{P}(X_{\underline{t}}^h \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}) dt.\end{aligned}$$

The second case from Lemma 3.6 concludes the proof. □

4 Convergence Analysis

We are ready to prove the main convergence result.

Theorem 4.1. *Assume Framework 3.1 and let $T > 0$. Moreover, let $h: \mathbb{R}^d \times (0,1) \to (0,1)$ be given by* (7) *and* (8)*, and let X^h be given by* (5) *and* (6)*. Then there exists a constant $C_{rmse} > 0$ such that*

$$\mathbb{E}\left[\sup_{0\le t\le T} \|X_t - X_t^h\|^2\right] \le C_{rmse}^2 \cdot (1 + \log(1/\delta))\delta.$$

Proof. The proof will be split into three steps. The first one uses the transformation G, the second the exit probability estimates from the previous section and the last step is a Gronwall argument. We denote constants independent of δ by $c_1, c_2, \dots$.

Step 1: Here we follow [24, Theorem 3.1], where the convergence proof is done by means of the transformation G, see Subsection 2.2. We define a process $Z = (Z_t)_{t\ge 0}$ by $Z_t = G(X_t)$. It solves

$$dZ_t = \mu_G(Z_t)dt + \sigma_G(Z_t)dW_t, \quad t \ge 0, \qquad Z_0 = G(x),$$

where μ_G, σ_G are given by (3) and are globally Lipschitz by Lemma 2.7. Moreover, G and G^{-1} are globally Lipschitz by Lemma 2.6(v). So we obtain

$$\Big(\mathbb{E}\Big[\sup_{0\le t\le T} \|X_t - X_t^h\|^2\Big]\Big)^{1/2} \le L_{G^{-1}}\Big(\mathbb{E}\Big[\sup_{0\le t\le T} \|Z_t - G(X_t^h)\|^2\Big]\Big)^{1/2}. \tag{11}$$

Now, denote by Z^h the Euler-Maruyama approximation of the process Z based on the step sizing function h given by (7) and (8). Using this scheme we can split the error as follows:

$$\Big(\mathbb{E}\Big[\sup_{0\le t\le T} \|Z_t - G(X_t^h)\|^2\Big]\Big)^{1/2} \le \Big(\mathbb{E}\Big[\sup_{0\le t\le T} \|Z_t - Z_t^h\|^2\Big]\Big)^{1/2} + \Big(\mathbb{E}\Big[\sup_{0\le t\le T} \|Z_t^h - G(X_t^h)\|^2\Big]\Big)^{1/2}. \tag{12}$$

Since the maximum step size of Z^h is δ, there exists a constant $c_1 > 0$ such that

$$\mathbb{E}\Big[\sup_{0\le t\le T} \|Z_t - Z_t^h\|^2\Big] \le c_1 \cdot \delta. \tag{13}$$

This can be shown by a simple modification of the error analysis of the Euler-Maruyama scheme for deterministic non-equidistant discretizations as, e.g., in [17, Theorem 10.2.2].

The second error term in (12) is the difference between the transformation applied to the time continuous Euler-Maruyama approximation of X defined in (2) and the Euler-Maruyama approximation of the transformed process Z defined in (4). For all $\tau \in [0,T]$ set

$$u(\tau) := \mathbb{E}\Big[\sup_{0\le t\le \tau} \|G(X_t^h) - Z_t^h\|^2\Big].$$

For all $x_1, x_2 \in \mathbb{R}^d$ define

$$\nu(x_1,x_2) := G'(x_1)\mu(x_2) + \frac{1}{2}\operatorname{tr}(\sigma(x_2)^\top G''(x_1)\sigma(x_2)),$$

and notice that $\nu(x,x) = \mu_G(G(x))$. By Itō's formula we have

$$G(X_t^h) = G(X_0^h) + \int_0^t \nu(X_s^h, X_{\underline{s}}^h)ds + \int_0^t G'(X_s^h)\sigma(X_{\underline{s}}^h)dW_s, \quad t \in [0,T].$$

With this, we get that

$$
\begin{aligned}
u(\tau) &= \mathbb{E}\left[\sup_{0\le t\le\tau}\left\|\int_0^t \nu(X_s^h, X_{\underline{s}}^h)ds + \int_0^t G'(X_s^h)\sigma(X_{\underline{s}}^h)dW_s - \int_0^t \mu_G(Z_{\underline{s}}^h)ds - \int_0^t \sigma_G(Z_{\underline{s}}^h)dW_s\right\|^2\right] \\
&\le 4\mathbb{E}\left[\sup_{0\le t\le\tau}\left\|\int_0^t \left(\nu(X_s^h, X_{\underline{s}}^h) - \nu(X_{\underline{s}}^h, X_{\underline{s}}^h)\right) ds\right\|^2\right] \\
&\quad + 4\mathbb{E}\left[\sup_{0\le t\le\tau}\left\|\int_0^t \left(G'(X_s^h)\sigma(X_{\underline{s}}^h) - G'(X_{\underline{s}}^h)\sigma(X_{\underline{s}}^h)\right) dW_s\right\|^2\right] \\
&\quad + 4\mathbb{E}\left[\sup_{0\le t\le\tau}\left\|\int_0^t \left(\mu_G(G(X_{\underline{s}}^h)) - \mu_G(Z_{\underline{s}}^h)\right) ds\right\|^2\right] \\
&\quad + 4\mathbb{E}\left[\sup_{0\le t\le\tau}\left\|\int_0^t \left(\sigma_G(G(X_{\underline{s}}^h)) - \sigma_G(Z_{\underline{s}}^h)\right) dW_s\right\|^2\right].
\end{aligned}
$$

With the Cauchy-Schwarz inequality and the d-dimensional Burkholder-Davis-Gundy inequality, see, e.g., [15, Theorem III.3.28] or [14, Lemma 3.7], we obtain

$$
u(\tau) \le 4T\, E_1(\tau) + 8d\, E_2(\tau) + 4T\, E_3(\tau) + 8d\, E_4(\tau), \tag{14}
$$

with

$$
\begin{aligned}
E_1(\tau) &= \mathbb{E}\left[\int_0^\tau \left\|\nu(X_s^h, X_{\underline{s}}^h) - \nu(X_{\underline{s}}^h, X_{\underline{s}}^h)\right\|^2 ds\right], \\
E_2(\tau) &= \mathbb{E}\left[\int_0^\tau \left\|G'(X_s^h)\sigma(X_{\underline{s}}^h) - G'(X_{\underline{s}}^h)\sigma(X_{\underline{s}}^h)\right\|^2 ds\right], \\
E_3(\tau) &= \mathbb{E}\left[\int_0^\tau \left\|\mu_G(G(X_{\underline{s}}^h)) - \mu_G(Z_{\underline{s}}^h)\right\|^2 ds\right], \\
E_4(\tau) &= \mathbb{E}\left[\int_0^\tau \left\|\sigma_G(G(X_{\underline{s}}^h)) - \sigma_G(Z_{\underline{s}}^h)\right\|^2 ds\right].
\end{aligned} \tag{15}
$$

Step 2: Now we estimate the above error terms. For E_1, using the linear growth property of μ and σ and the properties of G we have

$$
\|\nu(x_1, x_2) - \nu(x_2, x_2)\|^2 \le \begin{cases} K_1 \cdot (1 + \|x_2\|^4) \cdot \|x_1 - x_2\|^2, & \|x_1 - x_2\| = \rho(x_1, x_2), \\ K_2 \cdot (1 + \|x_2\|^4), & \text{otherwise}, \end{cases} \tag{16}
$$

with

$$
K_1 = 2L_{G'}^2 C_\mu^2 + \frac{1}{2} L_{G''}^2 C_\sigma^4, \qquad K_2 = 4C_\mu^2 \|G'\|_\infty^2 + C_\sigma^4 \|G''\|_\infty^2,
$$

where $C_\mu > 0$ and $C_\sigma > 0$ are the linear growth constants of the respective coefficients. Note that $\|x_1 - x_2\| \neq \rho(x_1, x_2)$ means that the direct connection between x_1 and x_2 passes Θ.

Further, set

$$
K_3 = \sup\left\{\|\nu(x_1, x_2)\|^2 : \, x_2 \in \Theta^{\varepsilon_0}, x_1 \in \mathbb{R}^d\right\}. \tag{17}
$$

The latter quantity is finite due to our assumptions.

For the further analysis we will use the following partitions of 1:

$$
1 = \mathbf{1}_{\{\|X_{\underline{s}}^h - X_s^h\| = \rho(X_{\underline{s}}^h, X_s^h)\}} + \mathbf{1}_{\{\|X_{\underline{s}}^h - X_s^h\| \neq \rho(X_{\underline{s}}^h, X_s^h)\}}
$$

and

$$
\begin{aligned}
\mathbf{1}_{\{\|X^h_{\underline{s}}-X^h_s\|\neq\rho(X^h_{\underline{s}},X^h_s)\}} = \mathbf{1}&_{\{\|X^h_{\underline{s}}-X^h_s\|\neq\rho(X^h_{\underline{s}},X^h_s)\}}\mathbf{1}_{\{X^h_{\underline{s}}\notin\Theta^{\varepsilon_0}\}}\\
&+\mathbf{1}_{\{\|X^h_{\underline{s}}-X^h_s\|\neq\rho(X^h_{\underline{s}},X^h_s)\}}\mathbf{1}_{\{X^h_{\underline{s}}\in\Theta^{\varepsilon_0}\setminus\Theta^{\varepsilon_1}\}}\\
&+\mathbf{1}_{\{\|X^h_{\underline{s}}-X^h_s\|\neq\rho(X^h_{\underline{s}},X^h_s)\}}\mathbf{1}_{\{X^h_{\underline{s}}\in\Theta^{\varepsilon_1}\setminus\Theta^{\varepsilon_2}\}}\\
&+\mathbf{1}_{\{\|X^h_{\underline{s}}-X^h_s\|\neq\rho(X^h_{\underline{s}},X^h_s)\}}\mathbf{1}_{\{X^h_{\underline{s}}\in\Theta^{\varepsilon_2}\}}
\end{aligned}
$$

for a given $s\in[0,T]$, i.e. we split Ω first into the disjoint events that $\|X^h_{\underline{s}}-X^h_s\|=\rho(X^h_{\underline{s}},X^h_s)$ or not, and if not, we split again according to the distance to Θ.

(i) From (16) we get that

$$
\begin{aligned}
&\mathbb{E}\left[\int_0^\tau \mathbf{1}_{\{\|X^h_{\underline{s}}-X^h_s\|=\rho(X^h_{\underline{s}},X^h_s)\}}\left\|\nu(X^h_s,X^h_{\underline{s}})-\nu(X^h_{\underline{s}},X^h_{\underline{s}})\right\|^2 ds\right]\\
&\le K_1\mathbb{E}\left[\int_0^\tau \left(1+\|X^h_{\underline{s}}\|^4\right)\|X^h_{\underline{s}}-X^h_s\|^2 ds\right].
\end{aligned}
$$

An application of the Cauchy-Schwarz inequality and Lemma 3.2 yields

$$
\mathbb{E}\left[\int_0^\tau \mathbf{1}_{\{\|X^h_{\underline{s}}-X^h_s\|=\rho(X^h_{\underline{s}},X^h_s)\}}\left\|\nu(X^h_s,X^h_{\underline{s}})-\nu(X^h_{\underline{s}},X^h_{\underline{s}})\right\|^2 ds\right]\le c_2\cdot\delta. \tag{18}
$$

(ii) Now consider the case that $\|X^h_{\underline{s}}-X^h_s\|\neq\rho(X^h_{\underline{s}},X^h_s)$ and that $X^h_{\underline{s}}$ is more than ε_0 away from Θ. Here (16) gives

$$
\begin{aligned}
&\mathbb{E}\left[\int_0^\tau \mathbf{1}_{\{\|X^h_{\underline{s}}-X^h_s\|\neq\rho(X^h_{\underline{s}},X^h_s)\}}\mathbf{1}_{\{X^h_{\underline{s}}\notin\Theta^{\varepsilon_0}\}}\left\|\nu(X^h_s,X^h_{\underline{s}})-\nu(X^h_{\underline{s}},X^h_{\underline{s}})\right\|^2 ds\right]\\
&\le K_2\mathbb{E}\left[\int_0^\tau \left(1+\|X^h_{\underline{s}}\|^4\right)\mathbf{1}_{\{\|X^h_{\underline{s}}-X^h_s\|\neq\rho(X^h_{\underline{s}},X^h_s)\}}\mathbf{1}_{\{X^h_{\underline{s}}\notin\Theta^{\varepsilon_0}\}}ds\right].
\end{aligned}
$$

Since

$$
\left\{\|X^h_{\underline{s}}-X^h_s\|\neq\rho(X^h_{\underline{s}},X^h_s)\right\}\cap\left\{X^h_{\underline{s}}\notin\Theta^{\varepsilon_0}\right\}\subseteq\left\{X^h_{\underline{s}}\notin\Theta^{\varepsilon_0}\right\}\cap\left\{\|X^h_{\underline{s}}-X^h_s\|\ \ \ge\varepsilon_0\right\},
$$

the Cauchy-Schwarz inequality together with (16) yields

$$
\begin{aligned}
&\mathbb{E}\left[\int_0^\tau \mathbf{1}_{\{\|X^h_{\underline{s}}-X^h_s\|\neq\rho(X^h_{\underline{s}},X^h_s)\}}\mathbf{1}_{\{X^h_{\underline{s}}\notin\Theta^{\varepsilon_0}\}}\left\|\nu(X^h_s,X^h_{\underline{s}})-\nu(X^h_{\underline{s}},X^h_{\underline{s}})\right\|^2 ds\right]\\
&\le 2K_2\int_0^\tau\left(\mathbb{E}[1+\|X^h_{\underline{s}}\|^4]^2\right)^{1/2}\left(\mathbb{P}(\|X^h_{\underline{s}}-X^h_s\|\ge\varepsilon_0)\right)^{1/2}ds.
\end{aligned}
$$

Markov's inequality yields

$$
\mathbb{P}(\|X^h_{\underline{s}}-X^h_s\|\ge\varepsilon_0)\le\frac{\mathbb{E}[\|X^h_{\underline{s}}-X^h_s\|^4]}{\varepsilon_0^4}.
$$

Lemma 3.2 now gives

$$
\mathbb{E}\left[\int_0^\tau \mathbf{1}_{\{\|X^h_{\underline{s}}-X^h_s\|\neq\rho(X^h_{\underline{s}},X^h_s)\}}\mathbf{1}_{\{X^h_{\underline{s}}\notin\Theta^{\varepsilon_0}\}}\left\|\nu(X^h_s,X^h_{\underline{s}})-\nu(X^h_{\underline{s}},X^h_{\underline{s}})\right\|^2 ds\right]\le c_3\cdot\delta. \tag{19}
$$

(iii) The next case is that $\|X^h_{\underline{s}} - X^h_s\| \neq \rho(X^h_{\underline{s}}, X^h_s)$ and $X^h_{\underline{s}}$ lies in $\Theta^{\varepsilon_0}\backslash\Theta^{\varepsilon_1}$. Since

$$\nu(X^h_s, X^h_{\underline{s}}) \cdot \mathbf{1}_{\{X^h_{\underline{s}} \in \Theta^{\varepsilon_0}\backslash\Theta^{\varepsilon_1}\}} \leq K_3,$$

we obtain

$$\begin{aligned}&\mathbb{E}\left[\int_0^\tau \mathbf{1}_{\{X^h_{\underline{s}} \in \Theta^{\varepsilon_0}\backslash\Theta^{\varepsilon_1}\}} \mathbf{1}_{\{\|X^h_{\underline{s}} - X^h_s\| \neq \rho(X^h_{\underline{s}}, X^h_s)\}} \left\|\nu(X^h_s, X^h_{\underline{s}}) - \nu(X^h_{\underline{s}}, X^h_{\underline{s}})\right\|^2 ds\right] \\ &\quad \leq 2K_3 \int_0^\tau \mathbb{P}(X^h_{\underline{s}} \in \Theta^{\varepsilon_0}\backslash\Theta^{\varepsilon_1}; \|X^h_{\underline{s}} - X^h_s\| \neq \rho(X^h_{\underline{s}}, X^h_s))ds.\end{aligned}$$

Since

$$\left\{X^h_{\underline{s}} \in \Theta^{\varepsilon_0}\backslash\Theta^{\varepsilon_1}\right\} \cap \left\{\|X^h_{\underline{s}} - X^h_s\| \neq \rho(X^h_{\underline{s}}, X^h_s)\right\} \subseteq \left\{X^h_{\underline{s}} \in \Theta^{\varepsilon_0}\backslash\Theta^{\varepsilon_1}\right\} \cap \left\{\|X^h_{\underline{s}} - X^h_s\| \ \geq \varepsilon_1\right\},$$

Lemma 3.5(iii) gives

$$\mathbb{E}\left[\int_0^\tau \mathbf{1}_{\{X^h_{\underline{s}} \in \Theta^{\varepsilon_0}\backslash\Theta^{\varepsilon_1}\}} \mathbf{1}_{\{\|X^h_{\underline{s}} - X^h_s\| \neq \rho(X^h_{\underline{s}}, X^h_s)\}} \left\|\nu(X^h_s, X^h_{\underline{s}}) - \nu(X^h_{\underline{s}}, X^h_{\underline{s}})\right\|^2 ds\right] \leq c_4 \cdot \delta. \tag{20}$$

(iv) For the next case observe that

$$\left\{X^h_{\underline{s}} \in \Theta^{\varepsilon_1}\backslash\Theta^{\varepsilon_2}\right\} \cap \left\{\|X^h_{\underline{s}} - X^h_s\| \neq \rho(X^h_{\underline{s}}, X^h_s)\right\} \subseteq \left\{X^h_{\underline{s}} \in \Theta^{\varepsilon_1}\backslash\Theta^{\varepsilon_2}\right\} \cap \left\{\|X^h_{\underline{s}} - X^h_s\| \ \geq d(X^h_{\underline{s}}, \Theta)\right\},$$

and so (17) and Lemma 3.5(ii) yield

$$\begin{aligned}&\mathbb{E}\left[\int_0^\tau \mathbf{1}_{\{X^h_{\underline{s}} \in \Theta^{\varepsilon_1}\backslash\Theta^{\varepsilon_2}\}} \mathbf{1}_{\{\|X^h_{\underline{s}} - X^h_s\| \neq \rho(X^h_{\underline{s}}, X^h_s)\}} \left\|\nu(X^h_s, X^h_{\underline{s}}) - \nu(X^h_{\underline{s}}, X^h_{\underline{s}})\right\|^2 ds\right] \\ &\quad \leq 2K_3 \int_0^\tau \mathbb{P}(X^h_{\underline{s}} \in \Theta^{\varepsilon_1}\backslash\Theta^{\varepsilon_2}; \|X^h_{\underline{s}} - X^h_s\| \neq \rho(X^h_{\underline{s}}, X^h_s))ds \\ &\quad \leq c_5 \cdot \delta.\end{aligned} \tag{21}$$

(v) For the final case, the boundedness of the coefficients on Θ^{ε_0} and the fact that

$$\left\{\|X^h_{\underline{s}} - X^h_s\| \neq \rho(X^h_{\underline{s}}, X^h_s)\right\} \cap \left\{X^h_{\underline{s}} \in \Theta^{\varepsilon_2}\right\} \subseteq \left\{X^h_{\underline{s}} \in \Theta^{\varepsilon_2}\right\}$$

together with the first statement of Lemma 3.6 yield that

$$\mathbb{E}\left[\int_0^T \mathbf{1}_{\{\|X^h_{\underline{s}} - X^h_s\| \neq \rho(X^h_{\underline{s}}, X^h_s)\}} \mathbf{1}_{\{X^h_{\underline{s}} \in \Theta^{\varepsilon_2}\}} \left\|\nu(X^h_s, X^h_{\underline{s}}) - \nu(X^h_{\underline{s}}, X^h_{\underline{s}})\right\|^2 ds\right] \leq c_6 \cdot \log(1/\delta)\delta. \tag{22}$$

Combining (15) with (18), (19), (20), (21), and (22) yields

$$E_1(\tau) \leq c_7 \cdot (1 + \log(1/\delta))\delta. \tag{23}$$

For estimating E_2 in (14), we exploit that G' is globally Lipschitz, σ satisfies a linear growth condition, and use Lemma 3.2 to obtain

$$\begin{aligned}E_2(\tau) &\leq L^2_{G'} C^2_\sigma \int_0^T \mathbb{E}\left[(1 + \|X^h_{\underline{s}}\|^2)\|X^h_s - X^h_{\underline{s}}\|^2\right] ds \\ &\leq L^2_{G'} C^2_\sigma \int_0^T \left(\left(\mathbb{E}[1 + \|X^h_{\underline{s}}\|^2]\right)^2\right)^{1/2} \left(\mathbb{E}[\|X^h_s - X^h_{\underline{s}}\|^4]\right)^{1/2} ds \leq c_8 \cdot \delta.\end{aligned} \tag{24}$$

For the remaining two terms in (14) we use the fact that μ_G, σ_G are globally Lipschitz by Lemma 2.7. This gives

$$E_3(\tau) \le L_{\mu_G}^2 \int_0^\tau \mathbb{E}[\|G(X_{\underline{s}}^h) - Z_{\underline{s}}^h\|^2] ds \le L_{\mu_G}^2 \int_0^\tau u(s) ds, \tag{25}$$

$$E_4(\tau) \le L_{\sigma_G}^2 \int_0^\tau \mathbb{E}[\|G(X_{\underline{s}}^h) - Z_{\underline{s}}^h\|^2] ds \le L_{\sigma_G}^2 \int_0^\tau u(s) ds. \tag{26}$$

Step 3: Combining (14) with the estimates (23), (24), (25), and (26) we obtain

$$0 \le u(\tau) \le c_9 \int_0^\tau u(s) ds + c_{10} \cdot \delta(1 + \log(1/\delta)), \quad \tau \in [0, T].$$

Gronwall's inequality yields

$$u(\tau) \le c_{10} \exp(c_9 \tau) \cdot \delta(1 + \log(1/\delta)), \quad \tau \in [0, T]. \tag{27}$$

Finally, combining (12) with (13) and (27), and the result with (11) concludes the proof.

□

Remark 4.2. In [24, Theorem 3.1] the authors prove strong convergence order $1/4 - \epsilon$ for arbitrarily small $\epsilon > 0$ of the equidistant Euler-Maruyama scheme under Assumption 2.1 and under the additional assumption that the coefficients μ and σ are bounded. By applying some of the techniques from the proof of Theorem 4.1 here, [24, Theorem 3.1] can be shown without assuming global boundedness of μ and σ.

5 Cost Analysis

We now turn to the computational cost of our step size procedure. As mentioned, the computational cost of our method is proportional to the number of steps, i.e.

$$N(h, \delta) = \inf\{k \in \mathbb{N} : \tau_k \ge T\}.$$

Clearly, we have

$$N(h, \delta) \le 1 + \int_0^T \frac{1}{h(X_{\underline{t}}^h, \delta)} dt,$$

since

$$\int_{\tau_k}^{\tau_{k+1}} \frac{1}{h(X_{\underline{t}}^h, \delta)} dt = 1.$$

Theorem 5.1. *Assume Framework 3.1 and let $T > 0$. Moreover, let $h \colon \mathbb{R}^d \times (0,1) \to (0,1)$ be given by (7) and (8), and let X^h be given by (5) and (6). Then there exists a constant $C_{cost} > 0$ such that*

$$\mathbb{E}[N(h, \delta)] \le C_{cost} \cdot (1 + \log(1/\delta))\delta^{-1}.$$

Proof. We denote constants independent of δ by $c_1, c_2, \ldots$.

We have

$$\begin{aligned} \mathbb{E}[N(h, \delta)] &\le 1 + \mathbb{E}\left[\int_0^T \frac{1}{h(X_{\underline{t}}^h)} dt\right] \\ &= 1 + \delta^{-2} \int_0^T \mathbb{P}(X_{\underline{t}}^h \in \Theta^{\varepsilon_2}) dt + \int_0^T \mathbb{E}\left[\frac{1}{h(X_{\underline{t}}^h, \delta)} \mathbf{1}_{\{X_{\underline{t}}^h \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}\}}\right] dt \\ &\quad + \delta^{-1} \int_0^T \mathbb{P}(X_{\underline{t}}^h \notin \Theta^{\varepsilon_1}) dt = 1 + I_1 + I_2 + I_3, \end{aligned} \tag{28}$$

with

$$
\begin{aligned}
I_1 &= \delta^{-2} \int_0^T \mathbb{P}(X_{\underline{t}}^h \in \Theta^{\varepsilon_2}) dt, \\
I_2 &= \int_0^T \mathbb{E}\left[\frac{1}{h(X_{\underline{t}}^h, \delta)} \mathbf{1}_{\{X_{\underline{t}}^h \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}\}} \right] dt, \\
I_3 &= \delta^{-1} \int_0^T \mathbb{P}(X_{\underline{t}}^h \notin \Theta^{\varepsilon_1}) dt.
\end{aligned} \tag{29}
$$

Clearly, we have

$$
I_3 \leq T \cdot \delta^{-1}, \tag{30}
$$

and by Lemma 3.6

$$
I_1 \leq c_1 \cdot (1 + \log(1/\delta)) \delta^{-1}. \tag{31}
$$

So, we only need to take care of the remaining term I_2. For this, consider the event that the time-continuous Euler-Maruyama scheme in one step does not move farther than $d(X_{\underline{t}}^h, \Theta)/2$ away from $X_{\underline{t}}^h$, that is

$$
A(t) = \left\{ \|X_t^h - X_{\underline{t}}^h\| \leq d(X_{\underline{t}}^h, \Theta)/2 \right\}, \quad t \geq 0.
$$

We use the following partition of 1:

$$
\mathbf{1}_{\{X_{\underline{t}}^h \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}\}} = \mathbf{1}_{A(t) \cap \{X_{\underline{t}}^h \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}\}} + \mathbf{1}_{A(t)^c \cap \{X_{\underline{t}}^h \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}\}}.
$$

(i) The distance function $d(\cdot, \Theta)$ to the hypersurface Θ is Lipschitz continuous with Lipschitz constant 1, see [7, Equation (14.91)], so we have

$$
d(x, \Theta) - d(y, \Theta) \leq |d(x, \Theta) - d(y, \Theta)| \leq \|x - y\|, \quad x, y \in \Theta^{\varepsilon_0}.
$$

Hence, we observe that

$$
d(x, \Theta) \leq \|x - y\| + d(y, \Theta), \quad x, y \in \Theta^{\varepsilon_0},
$$

which implies

$$
A(t) \cap \{X_{\underline{t}}^h \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}\} \subseteq \left\{ \frac{1}{2} d(X_{\underline{t}}^h, \Theta) \leq d(X_t^h, \Theta) \leq \frac{3}{2} d(X_{\underline{t}}^h, \Theta) \right\} \cap \{X_{\underline{t}}^h \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}\}.
$$

It follows that

$$
\begin{aligned}
\frac{1}{h(X_{\underline{t}}^h, \delta)} \mathbf{1}_{A(t) \cap \{X_{\underline{t}}^h \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}\}} &= \frac{\|\sigma\|_{\infty, \Theta^{\varepsilon_0}}^2 (\log(\delta))^2}{d(X_{\underline{t}}^h, \Theta)^2} \mathbf{1}_{A(t) \cap \{X_{\underline{t}}^h \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}\}} \\
&\leq \frac{9}{4} \frac{\|\sigma\|_{\infty, \Theta^{\varepsilon_0}}^2 (\log(\delta))^2}{d(X_t^h, \Theta)^2} \mathbf{1}_{A(t) \cap \{X_{\underline{t}}^h \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}\}}.
\end{aligned}
$$

Moreover,

$$
\{X_{\underline{t}}^h \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}\} \cap A(t) \subseteq \{X_t^h \in \Theta^{\frac{3}{2}\varepsilon_1} \setminus \Theta^{\frac{1}{2}\varepsilon_2}\}
$$

and so we obtain

$$
\frac{1}{h(X_{\underline{t}}^h, \delta)} \mathbf{1}_{A(t) \cap \{X_{\underline{t}}^h \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}\}} \leq \frac{9}{4} \frac{\|\sigma\|_{\infty, \Theta^{\varepsilon_0}}^2 (\log(\delta))^2}{d(X_t^h, \Theta)^2} \mathbf{1}_{\{X_t^h \in \Theta^{\frac{3}{2}\varepsilon_1} \setminus \Theta^{\frac{1}{2}\varepsilon_2}\}}. \tag{32}
$$

(ii) Since the minimal step size is δ^2, we have

$$\begin{aligned}\frac{1}{h(X^h_{\underline{t}},\delta)}\mathbf{1}_{A(t)^c\cap\{X^h_{\underline{t}}\in\Theta^{\varepsilon_1}\setminus\Theta^{\varepsilon_2}\}} &\le \frac{1}{\delta^2}\mathbf{1}_{A(t)^c\cap\{X^h_{\underline{t}}\in\Theta^{\varepsilon_1}\setminus\Theta^{\varepsilon_2}\}}\\ &= \frac{1}{\delta^2}\mathbf{1}_{\{\|X^h_t-X^h_{\underline{t}}\|>d(X^h_{\underline{t}},\Theta)/2;X^h_{\underline{t}}(\omega)\in\Theta^{\varepsilon_1}\setminus\Theta^{\varepsilon_2}\}}.\end{aligned} \tag{33}$$

Combining (29) with (32) and (33) we obtain

$$\begin{aligned}I_2 &\le \frac{9}{4}\|\sigma\|^2_{\infty,\Theta^{\varepsilon_0}}(\log(\delta))^2\int_0^T\mathbb{E}\left[\frac{1}{d(X^h_t,\Theta)^2}\mathbf{1}_{\{X^h_t\in\Theta^{\frac{3}{2}\varepsilon_1}\setminus\Theta^{\frac{1}{2}\varepsilon_2}\}}\right]dt\\ &\quad+\delta^{-2}\int_0^T\mathbb{E}\left[\mathbf{1}_{\{X^h_{\underline{t}}\in\Theta^{\varepsilon_1}\setminus\Theta^{\varepsilon_2};\|X^h_t-X^h_{\underline{t}}\|>d(X^h_{\underline{t}},\Theta)/2\}}\right]dt\\ &=: I_{21}+I_{22}.\end{aligned} \tag{34}$$

By Lemma 3.7 with $\alpha=1/2$ we obtain that there exists a constant $c_2>0$ such that

$$I_{22}\le c_2\cdot(1+\log(1/\delta))\delta^{-1}. \tag{35}$$

Lemma 3.3 yields that there exist constants $c_3,c_4>0$ such that

$$I_{21}\le c_3\int_{\frac{1}{2}\varepsilon_2}^{\frac{3}{2}\varepsilon_1}\frac{\|\sigma\|^2_{\infty,\Theta^{\varepsilon_0}}(\log(\delta))^2}{x^2}dx\le c_4\cdot\log(1/\delta)\delta^{-1}. \tag{36}$$

Combining (34) with (35) and (36) ensures the existence of a constant $c_5>0$ such that

$$I_2\le c_5\cdot(1+\log(1/\delta))\delta^{-1}. \tag{37}$$

Now, (28) together with the estimates (30), (31), and (37) yields the assertion. □

6 Examples

In this section we present some numerical examples to complement our asymptotic convergence analysis with a study of the non-asymptotic regime. For all examples we choose for simplicity $T=1$. We use

$$\operatorname{cost}(\delta)=\frac{1}{M}\sum_{i=1}^M N(h,\delta)^{(i)},$$

where $M\in\mathbb{N}$ is the sample size, to estimate the computational cost and

$$\operatorname{msq}(\delta)=\frac{1}{M}\sum_{i=1}^M\left\|\left(X_1^{h(\cdot,\delta)}-X_1^{h(\cdot,2\delta)}\right)^{(i)}\right\|^2$$

to estimate the convergence rate. The latter is justified (for dyadic δ) by the fact that if there exist $\beta\in\mathbb{R}$, $\gamma\in(0,\infty)$ such that

$$\limsup_{n\to\infty}\left(\log(n)^\beta(2^n)^\gamma\right)\cdot\mathbb{E}\left\|X_1-X_1^{h(\cdot,2^{-n})}\right\|^2<\infty$$

then also

$$\limsup_{n\to\infty}\left(\log(n)^\beta(2^n)^\gamma\right)\cdot\mathbb{E}\left\|X_1^{h(\cdot,2^{-n-1})}-X_1^{h(\cdot,2^{-n})}\right\|^2<\infty$$

and vice versa. Above we use the standard convention that $Y^{(i)}$ denotes an iid copy of a random variable Y.

For both quantities we choose $\delta = 2^{-2}, 2^{-3}, \ldots, 2^{-10}$, $M = 5 \cdot 10^4$ and perform a regression using the ansatz

$$f(\delta) = c_1 \cdot \log(1/\delta)^{c_2} \cdot \delta^{c_3}$$

to determine $c_1 > 0$ and $c_2, c_3 \in \mathbb{R}$.

Our first test equation is a scalar equation with

$$\mu(x) = -2 \cdot \mathbf{1}_{(-\infty,0)}(x) + x^2 \cdot \mathbf{1}_{[0,1)}(x) + \left(\frac{2}{x} - \frac{3}{x^2}\right) \cdot \mathbf{1}_{[1,\infty)}(x), \qquad \sigma(x) = 0.5 \cdot \left(1 + \frac{1}{1+x^2}\right),$$

and $x = 1.5$. For the cost of the Euler-Maruyama scheme we obtain

$$f_{\text{cost}}(\delta) = 1.2014 \cdot \log(1/\delta)^{0.8936} \cdot \delta^{-1.1218},$$

with a residuum of $\texttt{res} = 6.0412 \cdot 10^2$, while for the mean square error we have

$$f_{\text{msq}}(\delta) = 0.5940 \cdot \log(1/\delta)^{-2.0209} \cdot \delta^{1.1037},$$

with a residuum of $\texttt{res} = 1.0674 \cdot 10^{-2}$. This is in good accordance with the predicted asymptotic behaviour from Theorem 4.1, respectively Theorem 5.1.

The second test equation is again a scalar equation with

$$\mu(x) = -1 \cdot \mathbf{1}_{(-\infty,-1)}(x) + 1 \cdot \mathbf{1}_{[-1,2)}(x) - 2x \cdot \mathbf{1}_{[2,\infty)}(x), \qquad \sigma(x) = 1,$$

and $x = 0$, i.e. an equation with additive noise. Here we have

$$f_{\text{cost}}(\delta) = 0.9148 \cdot \log(1/\delta)^{0.5163} \cdot \delta^{-1.1380},$$

with a residuum of $\texttt{res} = 2.0217 \cdot 10^2$ and

$$f_{\text{msq}}(\delta) = 21.2638 \cdot \log(1/\delta)^{-1.8354} \cdot \delta^{1.5232},$$

with a residuum of $\texttt{res} = 2.2471 \cdot 10^{-2}$. The increase in the observed empirical convergence order is not surprising. Convergence order 3/4 for the equidistant Euler-Maruyama scheme with additive noise has already been indicated by the simulation results in [8]. To determine sharp bounds for the convergence order of the Euler-Maruyama scheme for SDEs with additive noise and discontinuous coefficients will be part of our future research.

The final test equation is two-dimensional with degenerate noise, i.e. we have

$$\mu(x_1, x_2) = \begin{cases} (1,1)^\top, & x_1^2 + x_2^2 \geq 1, \\ (-x_1, x_2)^\top, & x_1^2 + x_2^2 < 1, \end{cases} \qquad \sigma(x_1, x_2) = \frac{1}{2} \begin{pmatrix} x_1 & 0 \\ x_2 & 0 \end{pmatrix},$$

and initial value $x = (0.5, 0.5)^\top$. We obtain

$$f_{\text{cost}}(\delta) = 1.7280 \cdot \log(1/\delta)^{0.7362} \cdot \delta^{-1.0248},$$

with a residuum of $\texttt{res} = 2.3685 \cdot 10^1$ and

$$f_{\text{msq}}(\delta) = 11.9163 \cdot \log(1/\delta)^{-2.2178} \cdot \delta^{1.0389},$$

with a residuum of $\texttt{res} = 1.6205 \cdot 10^{-1}$. This is again in good accordance with our error analysis.

Acknowledgements

The authors are grateful to Mike Giles for numerous discussions. The timestepping strategy of this paper has been inspired by work of M. Giles and K. Ramanan on adaptive Multilevel Monte Carlo timestepping strategies for reflected diffusions. Furthermore, the authors are also grateful to Martin Schmidt for a discussion about geometric aspects of this work.

M. Szölgyenyi is supported by the AXA Research Fund grant "Numerical Methods for Stochastic Differential Equations with Irregular Coefficients with Applications in Risk Theory and Mathematical Finance". A part of this article was written while M. Szölgyenyi was affiliated with the Institute of Statistics and Mathematics, Vienna University of Economics and Business, Welthandelsplatz 1, 1020 Vienna, Austria, and supported by the Vienna Science and Technology Fund (WWTF): Project MA14-031.

Moreover, part of this work was carried out while A. Neuenkirch and M. Szölgyenyi were guests of the Erwin Schrödinger International Institute (ESI) for Mathematics and Physics in Vienna, whose support and hospitality is gratefully acknowledged.

References

[1] P. M. Burrage and K. Burrage. A variable stepsize implementation for stochastic differential equations. *SIAM Journal on Scientific Computing*, 24(3):848–864, 2003.

[2] P. Étoré and M. Martinez. Exact Simulation for Solutions of One-Dimensional Stochastic Differential Equations with Discontinuous Drift. *ESAIM: Probability and Statistics*, 18: 686–702, 2014.

[3] W. Fang and M. B. Giles. Adaptive Euler-Maruyama Method for SDEs with Non-Globally Lipschitz Drift: Part I, Finite Time Interval. *Preprint, arXiv:1609.08101*, 2016.

[4] W. Fang and M. B. Giles. Adaptive Euler-Maruyama Method for SDEs with Non-Globally Lipschitz Drift: Part II, Infinite Time Interval. *Preprint, arXiv:1703.06743*, 2017.

[5] Noufel Frikha. On the weak approximation of a skew diffusion. *Bernoulli*, to appear.

[6] J. G. Gaines and T. J. Lyons. Variable Step Size Control in the Numerical Solution of Stochastic Differential Equations. *SIAM J. Appl. Math.*, 57(5):1455–1484, 1997.

[7] David Gilbarg and Neil S. Trudinger. *Elliptic partial differential equations of second order*, volume 224 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 1983.

[8] S. Göttlich, K. Lux, and A. Neuenkirch. The Euler Scheme for Stochastic Differential Equations with Discontinuous Drift Coefficient: A Numerical Study of the Convergence Rate. 2017. Preprint, arXiv:1705.04562.

[9] I. Gyöngy. A Note on Euler's Approximation. *Potential Analysis*, 8:205–216, 1998.

[10] N. Halidias and P. E. Kloeden. A Note on the Euler–Maruyama Scheme for Stochastic Differential Equations with a Discontinuous Monotone Drift Coefficient. *BIT Numerical Mathematics*, 48(1):51–59, 2008.

[11] H. Hoel, E. Von Schwerin, A. Szepessy, and R. Tempone. Adaptive Multilevel Monte Carlo Simulation. In *Numerical Analysis of Multiscale Computations*, pages 217–234. Springer, 2012.

[12] N. Hofmann, T. Müller-Gronbach, and K. Ritter. Optimal Approximation of Stochastic Differential Equations by Adaptive Step-Size Control. *Math. Comp.*, 69(231):1017–1034, 2000.

[13] N. Hofmann, T. Müller-Gronbach, and K. Ritter. The Optimal Discretization of Stochastic Differential Equations. *J. Complexity*, 17(1):117–153, 2001.

[14] M. Hutzenthaler, A. Jentzen, and P. E. Kloeden. Strong Convergence of an Explicit Numerical Method for SDEs with Nonglobally Lipschitz Continuous Coefficients. *The Annals of Applied Probability*, 22(4):1611–1641, 2012.

[15] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*. Graduate Texts in Mathematics. Springer-Verlag, New York, second edition, 1991.

[16] C. Kelly and G.J. Lord. Adaptive time-stepping strategies for nonlinear stochastic systems. *IMA Journal of Numerical Analysis*, to appear.

[17] P. E. Kloeden and E. Platen. *Numerical Solutions of Stochastic Differential Equations*. Stochastic Modelling and Applied Probability. Springer Verlag, Berlin-Heidelberg, 1992.

[18] A. Kohatsu-Higa, A. Lejay, and K. Yasuda. Weak rate of convergence of the Euler-Maruyama scheme for stochastic differential equations with non-regular drift. *Journal of Computational and Applied Mathematics*, 326:138 – 158, 2017.

[19] N. V. Krylov. *Controlled Diffusion Processes*. Applications of Mathematics. Springer, New York - Berlin, 1980.

[20] H. Lamba. An Adaptive Timestepping Algorithm for Stochastic Differential Equations. *J. Comput. Appl. Math.*, 161(2):417–430, 2003.

[21] H. Lamba, J. C. Mattingly, and A. M. Stuart. An Adaptive Euler-Maruyama Acheme for SDEs: Convergence and Stability. *IMA J. Numer. Anal.*, 27(3):479–506, 2007.

[22] G. Leobacher and M. Szölgyenyi. A Numerical Method for SDEs with Discontinuous Drift. *BIT Numerical Mathematics*, 56(1):151–162, 2016.

[23] G. Leobacher and M. Szölgyenyi. A Strong Order 1/2 Method for Multidimensional SDEs with Discontinuous Drift. *The Annals of Applied Probability*, 27(4):2383–2418, 2017.

[24] G. Leobacher and M. Szölgyenyi. Convergence of the Euler-Maruyama method for multidimensional SDEs with discontinuous drift and degenerate diffusion coefficient. *Numerische Mathematik*, 138(1):219–239, 2018.

[25] G. Leobacher, M. Szölgyenyi, and S. Thonhauser. On the Existence of Solutions of a Class of SDEs with Discontinuous Drift and Singular Diffusion. *Electronic Communications in Probability*, 20(6):1–14, 2015.

[26] C. Lester, C. A. Yates, M. B. Giles, and R. E. Baker. An Adaptive Multi-Level Simulation Algorithm for Stochastic Biological Systems. *The Journal of Chemical Physics*, 142(2): 01B612_1, 2015.

[27] J. C. Mattingly, A. M. Stuart, and D. J. Higham. Ergodicity for SDEs and Approximations: Locally Lipschitz Vector Fields and Degenerate Noise. *Stochastic Process. Appl.*, 101(2): 185–232, 2002.

[28] T. Müller-Gronbach. The Optimal Uniform Approximation of Systems of Stochastic Differential Equations. *Ann. Appl. Probab.*, 12(2):664–690, 2002.

[29] T. Müller-Gronbach. Optimal Pointwise Approximation of SDEs Based on Brownian Motion at Discrete Points. *Ann. Appl. Probab.*, 14(4):1605–1642, 2004.

[30] H. L. Ngo and D. Taguchi. Strong Rate of Convergence for the Euler-Maruyama Approximation of Stochastic Differential Equations with Irregular Coefficients. *Mathematics of Computation*, 85(300):1793–1819, 2016.

[31] 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(4):1864–1883, 2017.

[32] H. L. Ngo and D. Taguchi. Strong Convergence for the Euler-Maruyama Approximation of Stochastic Differential Equations with Discontinuous Coefficients. *Statistics & Probability Letters*, 125:55–63, 2017.

[33] G. O. Roberts and R. L. Tweedie. Exponential Convergence of Langevin Distributions and Their Discrete Approximations. *Bernoulli*, 2(4):341–363, 1996.

[34] A. A. Shardin and M. Szölgyenyi. Optimal Control of an Energy Storage Facility Under a Changing Economic Environment and Partial Information. *International Journal of Theoretical and Applied Finance*, 19(4):1–27, 2016.

[35] A. YU. Veretennikov. On Stochastic Equations with Degenerate Diffusion with Respect to Some of the Variables. *Mathematics of the USSR Izvestiya*, 22(1):173–180, 1984.

Andreas Neuenkirch
Institut für Mathematik, Universität Mannheim A5, 6, 68131 Mannheim, Germany
neuenkirch@math.uni-mannheim.de

Michaela Szölgyenyi
Seminar for Applied Mathematics and RiskLab Switzerland, ETH Zürich, Rämistrasse 101, 8092 Zürich, Switzerland
michaela.szoelgyenyi@sam.math.ethz.ch

Lukasz Szpruch
School of Mathematics, University of Edinburgh, Peter Guthrie Tait Road, Edinburgh EH9 3FD, Great Britain
l.szpruch@ed.ac.uk