

## AN ADAPTIVE EULER–MARUYAMA SCHEME FOR STOCHASTIC DIFFERENTIAL EQUATIONS WITH DISCONTINUOUS DRIFT AND ITS CONVERGENCE ANALYSIS\*

ANDREAS NEUENKIRCH<sup>†</sup>, MICHAELA SZÖLGYENYI<sup>‡</sup>, AND LUKASZ SZPRUCH<sup>§</sup>

**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.

**Key words.** stochastic differential equations, discontinuous drift, degenerate diffusion, strong convergence order, adaptive Euler–Maruyama scheme

**AMS subject classifications.** 60H10, 65C30, 65C20, 65L20

**DOI.** 10.1137/18M1170017

**1. Introduction and main results.** In this manuscript, we consider the strong approximation of time-homogeneous Itô stochastic differential equations (SDEs) of the form

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

where  $x \in \mathbb{R}^d$  is the initial value,  $\mu: \mathbb{R}^d \rightarrow \mathbb{R}^d$  is the drift coefficient,  $\sigma: \mathbb{R}^d \rightarrow \mathbb{R}^{d,d}$  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  $\mu$  to be discontinuous on a hypersurface  $\Theta$
- (ii) and the diffusion coefficient  $\sigma$  to be degenerate outside  $\Theta$ .

Our aim is to construct a numerical scheme, which is relatively easy to implement and 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 [22] for SDE (1) is known to have this property for multidimensional SDEs.

To state our main results, denote the distance to the exceptional set  $\Theta$  by

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

---

\*Received by the editors February 12, 2018; accepted for publication (in revised form) December 6, 2018; published electronically February 19, 2019.

<http://www.siam.org/journals/sinum/57-1/M117001.html>

**Funding:** The work of the second author was supported by the AXA Research Fund grant “Numerical Methods for Stochastic Differential Equations with Irregular Coefficients with Applications in Risk Theory and Mathematical Finance” and by the Vienna Science and Technology Fund (WWTF): project MA14-031. The first and second authors also acknowledge the support of the Erwin Schrödinger International Institute (ESI) for Mathematics and Physics in Vienna.

<sup>†</sup>Institut für Mathematik, Universität Mannheim, B6, 26, 68131 Mannheim, Germany (neuenkirch@math.uni-mannheim.de).

<sup>‡</sup>Department of Statistics, University of Klagenfurt, Universitätsstraße 65–67, 9020 Klagenfurt, Austria (michaela.szoelgyenyi@aau.at).

<sup>§</sup>School of Mathematics, University of Edinburgh, Peter Guthrie Tait Road, Edinburgh EH9 3FD, Great Britain (l.szpruch@ed.ac.uk).

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

$$(2) \quad X_0^h = x, \quad \tau_0 = 0,$$

and

$$(3) \quad X_{\tau_{k+1}}^h = X_{\tau_k}^h + \mu(X_{\tau_k}^h)(\tau_{k+1} - \tau_k) + \sigma(X_{\tau_k}^h)(W_{\tau_{k+1}} - W_{\tau_k}), \quad \tau_{k+1} = \tau_k + h(X_{\tau_k}^h, \delta),$$

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

$$(4) \quad h(x, \delta) = \begin{cases} \delta^2, & x \in \Theta^{\varepsilon_2}, \\ \frac{1}{\sup_{x \in \Theta^{\varepsilon_0}} \|\sigma(x)\|^2} \left( \frac{d(x, \Theta)}{\log(1/\delta)} \right)^2, & x \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}, \\ \delta, & x \notin \Theta^{\varepsilon_1}, \end{cases}$$

where

$$(5) \quad \varepsilon_1 = \sup_{x \in \Theta^{\varepsilon_0}} \|\sigma(x)\| \log(1/\delta) \sqrt{\delta}, \quad \varepsilon_2 = \sup_{x \in \Theta^{\varepsilon_0}} \|\sigma(x)\| \log(1/\delta) \delta$$

with  $\delta \in (0, 1)$  and  $\varepsilon_0 > 4\varepsilon_1 > 4\varepsilon_2$ . Note that  $\tau$  depends on  $h(\cdot, \delta)$  and also  $\varepsilon_1, \varepsilon_2$  depend on  $\delta$ , 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

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

Obviously this scheme uses smaller steps close to the discontinuities, has maximal step size  $\delta$ , minimal step size  $\delta^2$ , 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, in particular of  $\mu$ ,  $\sigma$ , and  $h$ , and random numbers, is proportional to the number of steps which are needed to reach time  $T$ , that is,

$$(7) \quad N(h, \delta) = \min\{k \in \mathbb{N} : \tau_k \geq T\}.$$

We will use this quantity as a proxy for the computational cost of the scheme. Clearly, the evaluation of  $h$  might be a nontrivial problem, which is however out of the scope of the present work.

We will work under mild assumptions, i.e.,

- (i)  $\mu$  is supposed to be piecewise Lipschitz and its discontinuity set  $\Theta$  is a sufficiently regular hypersurface;
- (ii)  $\sigma$  is globally Lipschitz;
- (iii)  $\mu$  and  $\sigma$  satisfy a geometric smoothness and boundedness condition close to  $\Theta$ ;

see Assumption 2.1.

For fixed  $T > 0$  we will show that

$$(8) \quad \left( \mathbb{E} \left[ \sup_{s \in [0, T]} \|X_s - X_s^h\|^2 \right] \right)^{1/2} \leq C_{\text{rmse}} \cdot \sqrt{1 + \log(1/\delta)} \sqrt{\delta}$$

(see Theorem 4.1), and

$$(9) \quad \mathbb{E}[N(h, \delta)] \leq C_{\text{cost}} \cdot (1 + \log(1/\delta)) \delta^{-1}$$

(see Theorem 5.1), for some constants  $C_{\text{rmse}}, C_{\text{cost}} > 0$  depending only on  $\mu, \sigma, \Theta, T, x$ .

So up to logarithmic terms the adaptive Euler–Maruyama scheme recovers the classical order 1/2 with respect to the average computational cost. Note that both theorems remain valid if the quantity  $\sup_{x \in \Theta^{\varepsilon_0}} \|\sigma(x)\|$  is replaced by an upper bound in the definition of  $\varepsilon_1, \varepsilon_2$ , and  $h$ .

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., [36] and the recent works [25, 35]. 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 [8]. 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 [9] provide strong convergence of the Euler–Maruyama scheme for discontinuous, but monotone drift coefficients.

Recently several contributions for the strong approximation have been given in a series of articles of Ngo and Taguchi [32, 33, 31] and Leobacher and Szölgyenyi [21, 22, 24]. For the multidimensional SDE (1) these works provide

- (i) the  $L^1$ -convergence order 1/2 for the equidistant Euler–Maruyama scheme, if  $\mu$  is one-sided Lipschitz and an appropriate limit of smooth functions, and  $\sigma$  is bounded and uniformly nondegenerate; see [31];
- (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  $\mu$  and  $\sigma$ ; see [24];
- (iii) the  $L^2$ -convergence order 1/2 for a transformation-based Euler–Maruyama method under Assumption 2.1; see [22]. However, this transformation is in general difficult to compute, which limits its applicability.

After our work was prepared and submitted, Müller-Gronbach and Yaroslavtseva obtained for scalar SDEs and the equidistant Euler scheme the  $L^p$ -convergence order 1/2 for any  $p \geq 1$ ; see [30]. Whether such an improvement of (ii) is also possible in the multidimensional case, presently remains an open question.

The weak approximation of SDEs with discontinuous coefficients has been studied in [17], where an Euler-type scheme based on an SDE with mollified drift coefficient is analyzed, and in [4], 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 time stepping 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., [5, 1, 19].
- Adaptive methods have also been used to preserve ergodic properties of the underlying SDE; see, e.g., [20, 27]. In fact, to recover ergodicity by adaptivity has already been proposed in [34].
- For optimal approximation of SDEs in the information-based-complexity framework adaptive methods have been exhaustively analyzed in a series of articles [29, 28, 12, 11]. In these works, optimal convergence rates and asymptotically optimal schemes have been established for various error criteria.
- Finally, step adaptation strategies for SDEs with nonglobally Lipschitz coefficients have been studied in [15, 3]. This research has been partially motivated for the purpose of multilevel Monte Carlo simulations [26, 10].

**2. Preliminaries.** In this section we present some notions from differential geometry and analysis, state the assumptions we make on the coefficients of the SDE (1) together with an existence and uniqueness result, and present a Krylov-type estimate for Itô 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 [22]. For this, we recall two definitions.

DEFINITION 2.1 (see [22, Definitions 3.1 and 3.2]). *Let  $A \subseteq \mathbb{R}^d$ .*

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

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

*The intrinsic metric  $\rho$  on  $A$  is given by*

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

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

2. *Let  $f: A \rightarrow \mathbb{R}^m$  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*

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

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

**DEFINITION 2.2** (see [22, Definition 3.4]). *A function  $f: \mathbb{R}^d \rightarrow \mathbb{R}^m$  is piecewise Lipschitz, if there exists a hypersurface  $\Theta$  with finitely many connected components and with the property, that the restriction  $f|_{\mathbb{R}^d \setminus \Theta}$  is intrinsic Lipschitz. We call  $\Theta$  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  $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  which is piecewise Lipschitz and discontinuous at  $\Theta = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 < 0\}$ . 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  $\Theta$ . The intrinsic metric provides a Lipschitz estimate with a connecting curve that lies in  $\mathbb{R}^2 \setminus \Theta$ .

In the following, we consider piecewise Lipschitz functions with exceptional set  $\Theta$ , where  $\Theta$  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  $f: \mathbb{R}^d \rightarrow \mathbb{R}$  we denote  $\|f\|_{\infty, \Theta^{\varepsilon_0}} := \sup_{x \in \Theta^{\varepsilon_0}} \|f(x)\|$ .

If  $\Theta \in C^4$ , locally there exists a unit normal vector that is a continuously differentiable  $C^3$  function  $n: U \subseteq \Theta \rightarrow \mathbb{R}^d$  such that for every  $\zeta \in U$ ,  $\|n(\zeta)\| = 1$  and  $n(\zeta)$  is orthogonal to the tangent space of  $\Theta$  in  $\zeta$ .

Recall the following definition from differential geometry.

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

1. An environment  $\Theta^\varepsilon$  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  $p: \Theta^\varepsilon \rightarrow \Theta$  assigning to each  $x$  the point  $p(x)$  in  $\Theta$ , which is closest to  $x$ .
2. A set  $\Theta$  is said to be of positive reach, if there exists  $\varepsilon > 0$  such that  $\Theta^\varepsilon$  has the unique closest point property. The reach  $r_\Theta$  of  $\Theta$  is the supremum over all such  $\varepsilon$  if such an  $\varepsilon$  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.* For the coefficients  $\mu: \mathbb{R}^d \rightarrow \mathbb{R}^d$  and  $\sigma: \mathbb{R}^d \rightarrow \mathbb{R}^{d,d}$  of SDE (1) we assume that

1. the diffusion coefficient  $\sigma$  is Lipschitz;
2. the drift coefficient  $\mu$  is a piecewise Lipschitz function; its exceptional set  $\Theta$  is a  $C^4$ -hypersurface with reach  $r_\Theta > \varepsilon_0$  for some  $\varepsilon_0 > 0$  and every unit normal vector  $n$  of  $\Theta$  has bounded second and third derivative;
3. the coefficients  $\mu$  and  $\sigma$  satisfy

$$\sup_{x \in \Theta^{\varepsilon_0}} (\|\mu(x)\| + \|\sigma(x)\|) < \infty;$$

4. (nonparallelity 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

$$(10) \quad \alpha: \Theta \rightarrow \mathbb{R}^d, \quad \alpha(\xi) = \lim_{h \rightarrow 0+} \frac{\mu(\xi - hn(\xi)) - \mu(\xi + hn(\xi))}{2\|\sigma(\xi)^\top n(\xi)\|^2}$$

is well defined, bounded, and belongs to  $C_b^3(\Theta; \mathbb{R}^d)$ .

Note that  $\mu$  and  $\sigma$  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  $\Theta$  to a plane in a regular way. Furthermore, it guarantees that  $n'$  is bounded; see [22, Lemma 3.10].
2. Assumption 2.1.4 ensures that  $\sigma(\xi)$  has a component orthogonal to  $\Theta$  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 [22], which is the basis of our convergence proof, to work.
4. Assumption 2.1 is satisfied if, e.g.,
  - (i) the exceptional set  $\Theta$  is a compact set. Note that then its complement satisfies  $\mathbb{R}^d \setminus \Theta = A_1 \cup \dots \cup A_n$ , where  $A_1, \dots, A_n$  are open and connected subsets of  $\mathbb{R}^d$ ;
  - (ii) there exist Lipschitz  $C^3$ -functions  $\mu_1, \dots, \mu_n: \mathbb{R}^d \rightarrow \mathbb{R}^d$  such that  $\mu = \sum_{k=1}^n \mathbf{1}_{A_k} \mu_k$  and  $\sigma$  is Lipschitz and  $C^3$ ;
  - (iii) Assumptions 2.1.3 and 2.1.4 hold.

Compare [24, Example 2.6].

**THEOREM 2.5** (see [22, 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  $G: \mathbb{R}^d \rightarrow \mathbb{R}^d$ , which transforms the SDE for  $X$  into another SDE which has Lipschitz coefficients. More precisely, we define

$$(11) \quad G(x) = \begin{cases} x + \varphi(x)\alpha(p(x)), & x \in \Theta^{\varepsilon_0}, \\ x, & x \in \mathbb{R}^d \setminus \Theta^{\varepsilon_0}, \end{cases}$$

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

$$(12) \quad \varphi(x) = n(p(x))^\top (x - p(x)) \|x - p(x)\| \phi\left(\frac{\|x - p(x)\|}{c}\right)$$

with a constant  $c > 0$  and  $\phi: \mathbb{R} \rightarrow \mathbb{R}$ ,

$$\phi(u) = \begin{cases} (1+u)^4(1-u)^4, & |u| \leq 1, \\ 0, & |u| > 1. \end{cases}$$

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 Lipschitz,  $G''$  exists on  $\mathbb{R}^d \setminus \Theta$  and is piecewise Lipschitz with exceptional set  $\Theta$ ;
- (iii)  $G'$  and  $G''$  are bounded;
- (iv) for  $c$  sufficiently small (see [23, Lemma 1]),  $G$  is globally invertible;
- (v)  $G$  and  $G^{-1}$  are Lipschitz continuous;
- (vi) Itô's formula holds for  $G$  and  $G^{-1}$ .

*Proof.* For (i) and (iv) see [22, Theorem 3.14]. Assertion (v) follows from the proof of [22, Theorem 3.20] and for (vi) see [22, Theorem 3.19]. The boundedness of  $G'$  follows from (v), and the boundedness of  $G''$  is proven in [23, Lemma 4], showing (iii).

Moreover this, [22, Lemmas 3.6, 3.8, and 3.11] assure that  $G'$  is Lipschitz. Hence, assertion (ii) follows again from the proof of [22, Theorem 3.20] and the fact  $G'''$  is bounded on  $\mathbb{R}^d \setminus \Theta$ , see [23, Lemma 4].  $\square$

Now, define the coefficients

$$(13) \quad \begin{aligned} \mu_G(z) &= G'(G^{-1}(z))\mu(G^{-1}(z)) + \frac{1}{2} \operatorname{tr} [\sigma(G^{-1}(z))^{\top} G''(G^{-1}(z))\sigma(G^{-1}(z))], \\ \sigma_G(z) &= G'(G^{-1}(z))\sigma(G^{-1}(z)) \end{aligned}$$

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

LEMMA 2.7 (see [22, Theorem 3.20]). *The functions  $\mu_G: \mathbb{R}^d \rightarrow \mathbb{R}^d$  and  $\sigma_G: \mathbb{R}^d \rightarrow \mathbb{R}^{d,d}$  are globally Lipschitz.*

Thus, the SDE

$$(14) \quad dZ_t = \mu_G(Z_t)dt + \sigma_G(Z_t)dW_t, \quad t \geq 0, \quad Z_0 = G(x),$$

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 Itô's formula and [22, Theorem 3.19].

**2.3. Occupation time estimates for Itô processes.** In this subsection we study the occupation time of an Itô process close to a  $C^4$ -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 nondegenerate diffusions (see [18]), the following result only assumes nondegeneracy of the diffusion coefficient at a normal direction within an environment of the hypersurface  $\Theta$ .

THEOREM 2.8. *Let  $\Theta$  be a  $C^4$ -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  $\mathbb{R}^d$ -valued, respectively,  $\mathbb{R}^{d,d}$ -valued progressively measurable processes such that*

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

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

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

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

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

Then there exists a constant  $C > 0$  such that for all  $0 < \epsilon < \epsilon_0/2$  and any measurable function  $f: [0, \infty) \rightarrow [0, \infty)$  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  $\Theta$ . More precisely, there exists a bounded real-valued Itô 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,  $\mathbb{R}$ -valued processes, such that

$$(15) \quad Y_t \cdot \mathbf{1}_{\{X_t \in \Theta^{\epsilon_1}\}} = \lambda(D(X_t)) \cdot \mathbf{1}_{\{X_t \in \Theta^{\epsilon_1}\}}, \quad t \geq 0,$$

where  $D(x) = n(p(x))^\top (x - p(x))$  for  $x \in \Theta^{\epsilon_0}$ ,  $\epsilon_1 = \epsilon_0/2$ , and  $\lambda : \mathbb{R} \rightarrow \mathbb{R}$  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  $\lambda$ -transformed signed distance of  $X$  to  $\Theta$ . Since  $\lambda'(\pm\epsilon_1) = \lambda''(\pm\epsilon_1) = 0$ , it holds that  $\lambda \in C^2$ . Moreover,  $\lambda : [-\epsilon_1, \epsilon_1] \rightarrow [\lambda(-\epsilon_1), \lambda(\epsilon_1)]$  is invertible.

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

$$\begin{aligned} & \int_0^t \mathbf{1}_{\{Y_s \in (-\lambda(\epsilon), \lambda(\epsilon))\}} d[Y]_s \\ &= \int_0^t \mathbf{1}_{\{Y_s \in (-\lambda(\epsilon), \lambda(\epsilon))\}} |\lambda'(D(X_s))|^2 n(p(X_s))^\top B_s B_s^\top n(p(X_s)) ds, \quad t \geq 0; \end{aligned}$$

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

$$(16) \quad |\lambda^{-1}(Y_t)| \cdot \mathbf{1}_{\{X_t \in \Theta^{\epsilon_1}\}} = |D(X_t)| \cdot \mathbf{1}_{\{X_t \in \Theta^{\epsilon_1}\}} = d(X_t, \Theta) \cdot \mathbf{1}_{\{X_t \in \Theta^{\epsilon_1}\}}, \quad t \geq 0,$$

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

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

Therefore, the occupation time formula [14, Chapter 3, 7.1 Theorem] for one-dimensional continuous semimartingales yields

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

By the Tanaka–Meyer formula (see, e.g, [14, Chapter 3, Theorem 7.1]), we have

$$2\mathbb{E}[L_T^y] = \mathbb{E}[|Y_T - y| - |Y_0 - y|] - \int_0^T \mathbb{E}[\text{sign}(Y_s - y)\hat{B}_s] ds$$

and therefore

$$\sup_{y \in \mathbb{R}} \mathbb{E}[L_T^y] \leq \mathbb{E}[|Y_T - Y_0|] + \int_0^T \mathbb{E}[|\hat{B}_s|] ds \leq \mathbb{E}[|Y_T - Y_0|] + T\|\hat{B}\|_\infty < \infty,$$

since  $Y$  is a real-valued Itô process with bounded coefficients. Thus, it follows that

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

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

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

which shows the assertion.  $\square$

### 3. Properties of the adaptive Euler–Maruyama scheme.

#### 3.1. General properties.

Recall that our Euler–Maruyama scheme is given by

$$(17) \quad \tau_0 = 0, \quad X_0^h = x \in \mathbb{R}^d,$$

and

$$(18) \quad \begin{aligned} \tau_{k+1} &= \tau_k + h(X_{\tau_k}^h, \delta), \\ X_t^h &= X_{\tau_k}^h + \mu(X_{\tau_k}^h)(t - \tau_k) + \sigma(X_{\tau_k}^h)(W_t - W_{\tau_k}), \quad t \in (\tau_k, \tau_{k+1}], \end{aligned}$$

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

$$(19) \quad h(x, \delta) = \begin{cases} \delta^2, & x \in \Theta^{\varepsilon_2}, \\ \frac{1}{\|\sigma\|_{\infty, \Theta^{\varepsilon_0}}} \left( \frac{d(x, \Theta)}{\log(1/\delta)} \right)^2, & x \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}, \\ \delta, & x \notin \Theta^{\varepsilon_1}, \end{cases}$$

where

$$(20) \quad \varepsilon_1 = \|\sigma\|_{\infty, \Theta^{\varepsilon_0}} \log(1/\delta) \sqrt{\delta}, \quad \varepsilon_2 = \|\sigma\|_{\infty, \Theta^{\varepsilon_0}} \log(1/\delta) \delta.$$

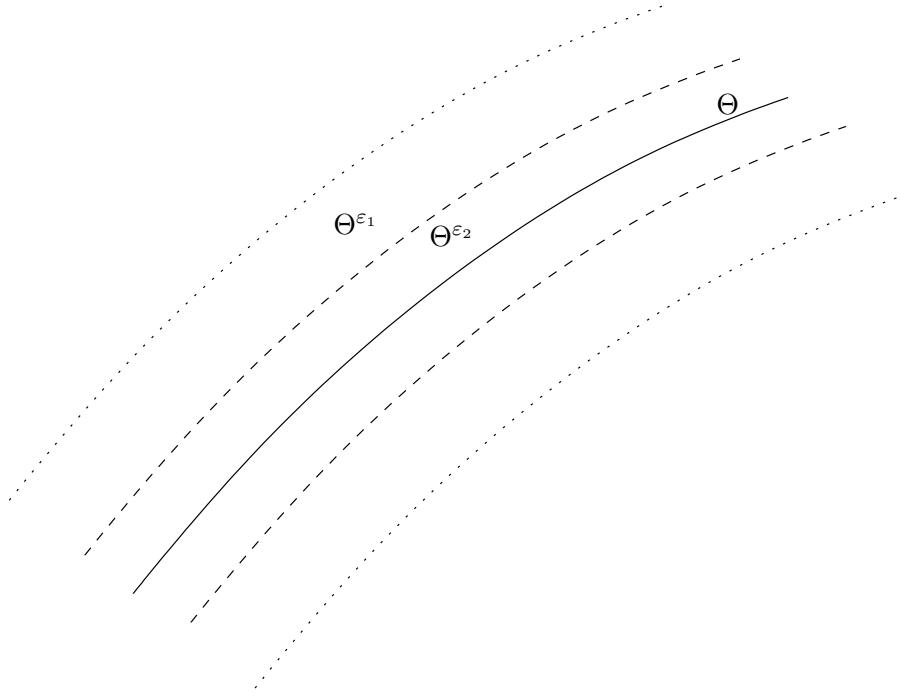
Moreover, set  $\underline{t} = \max\{\tau_k : \tau_k \leq t\}$ .

Figure 1 illustrates the different step size regimes. The solid line is the set of discontinuities  $\Theta$  of the drift, the area between the two dashed lines is  $\Theta^{\varepsilon_2}$ , and the area between the dotted lines is  $\Theta^{\varepsilon_1}$ .

FRAMEWORK 3.1. *We assume that*

- (i) *Assumption 2.1 holds,*
- (ii) *the step size  $\delta$  is sufficiently small such that*

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

FIG. 1. *The three step size regimes.*

- (iii) *the constant  $c$  from (12) is sufficiently small (see [22, Lemma 3.18]), in particular,  $c < \varepsilon_0$ , such that  $G$  is globally invertible.*

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

$$\Phi : \mathbb{R}^d \times C([0, \infty); \mathbb{R}^d) \rightarrow C([0, \infty); \mathbb{R}^d), \quad \Phi(x, W) = (X_s^h)_{s \geq 0}.$$

Note that by construction  $\Phi$  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  $\tau_k$  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  $\tau_{k+1}$  is  $X_{\tau_k}^h$  measurable and  $X_0^h$  is deterministic. The strong Markov property of Brownian motion then implies that for all  $k \in \mathbb{N}_0$  the process

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

is again a Brownian motion and independent of  $X_{\tau_k}$ .

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*

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

*Proof.* Using  $\Phi$  and  $W^{\tau_k}$  we can write

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

Proceeding iteratively we obtain

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

Since  $W^{\tau_k}$  is a Brownian motion which is independent of  $X_{\tau_k}^h$ , the factorization lemma for conditional expectations concludes the proof.  $\square$

**LEMMA 3.2.** *Assume Framework 3.1 and let  $T > 0$ . For any  $p \geq 2$  there exist constants  $C_p, K_p > 0$  such that*

$$\mathbb{E} \left[ \sup_{t \in [0, T]} \|X_t^h\|^p \right] \leq C_p$$

and

$$\mathbb{E}[\|X_t^h - X_s^h\|^p] \leq K_p \cdot |t - s|^{p/2}, \quad 0 \leq s < t \leq T.$$

*Proof.* As preparation for the first statement note that  $W_{\tau_{k+1}} - W_{\tau_k}$  is independent of  $X_{\tau_k}^h$  and satisfies

$$\mathbb{E}[\|W_{\tau_{k+1}} - W_{\tau_k}\|^p] = \int_{\mathbb{R}} \mathbb{E}[\|W_{h(y, \delta)}^{\tau_k}\|^p] d\mathbb{P}^{X_{\tau_k}^h}(y) \leq \kappa_p \cdot \delta^{p/2}$$

for some constant  $\kappa_p > 0$ . Clearly, we have

$$\|X_{\tau_{k+1}}^h\|^p \leq 3^{p-1} \|X_{\tau_k}^h\|^p + 3^{p-1} \|\mu(X_{\tau_k}^h)\|^p \delta^p + 3^{p-1} \|\sigma(X_{\tau_k}^h)\|^p \|W_{\tau_{k+1}} - W_{\tau_k}\|^p.$$

Hence, the preparations together with the linear growth of  $\mu$  and  $\sigma$  imply the existence of a constant  $c_1 > 0$  independent of  $k$  such that

$$\mathbb{E}[\|X_{\tau_{k+1}}^h\|^p] \leq c_1 (1 + \mathbb{E}[\|X_{\tau_k}^h\|^p]), \quad k \in \mathbb{N}_0.$$

Since  $X_0 = x$  it follows that

$$(21) \quad \mathbb{E}[\|X_{\tau_k}^h\|^p] < \infty, \quad k \in \mathbb{N}_0.$$

We also have

$$\begin{aligned} \sup_{t \in [0, s]} \|X_t^h\|^p &\leq 3^{p-1} \|x\|^p + 3^{p-1} \sup_{t \in [0, s]} \left\| \int_0^t \mu(X_u^h) du \right\|^p \\ &\quad + 3^{p-1} \sup_{t \in [0, s]} \left\| \int_0^t \sigma(X_u^h) dW_u \right\|^p, \quad s \in [0, T]. \end{aligned}$$

The Cauchy–Schwarz inequality for the Riemann integral, the Burkholder–Davis–Gundy inequality for the Itô integral, and the linear growth condition on  $\mu$  and  $\sigma$  ensure the existence of a constant  $c_2 > 0$ , which depends on  $T > 0$ ,  $\mu$ ,  $\sigma$ ,  $p$ , and  $x$ , such that

$$(22) \quad \mathbb{E} \left[ \sup_{t \in [0, s]} \|X_t^h\|^p \right] \leq c_2 \left( 1 + \int_0^s \mathbb{E}[\|X_u^h\|^p] du \right), \quad s \in [0, T].$$

By (21) we now obtain that

$$(23) \quad \mathbb{E} \left[ \sup_{t \in [0, T]} \|X_t^h\|^p \right] < \infty.$$

Equation (22) also yields

$$\mathbb{E} \left[ \sup_{t \in [0, s]} \|X_t^h\|^p \right] \leq c_2 + c_2 \int_0^s \mathbb{E} \left[ \sup_{t \in [0, u]} \|X_t^h\|^p \right] du.$$

Since (23) holds, Gronwall's lemma now yields the first assertion.

For the second statement note that

$$\|X_t^h - X_s^h\|^p \leq 2^{p-1} \left\| \int_s^t \mu(X_u^h) du \right\|^p + 2^{p-1} \left\| \int_s^t \sigma(X_u^h) dW_u \right\|^p, \quad s, t \in [0, T].$$

The Cauchy–Schwarz inequality, the Burkholder–Davis–Gundy inequality, and the linear growth condition of the coefficients together with the first assertion yield the statement.  $\square$

Applying Theorem 2.8 to our Euler–Maruyama scheme we obtain the following.

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

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

In particular, we have

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

**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*

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

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

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

The next lemma controls the probabilities that the Euler–Maruyama scheme has increments that are relatively large compared to its distance from  $\Theta$ .

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

- (i)  $\int_0^T \mathbb{P}(X_t^h \notin \Theta^{2\varepsilon_2}; X_{\underline{t}}^h \in \Theta^{\varepsilon_2}) dt \leq C \cdot \delta,$
- (ii)  $\int_0^T \mathbb{P}(\|X_t^h - X_{\underline{t}}^h\| \geq d(X_{\underline{t}}^h, \Theta); X_{\underline{t}}^h \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}) dt \leq C \cdot \delta,$
- (iii)  $\int_0^T \mathbb{P}(\|X_t^h - X_{\underline{t}}^h\| \geq \varepsilon_1; X_{\underline{t}}^h \in \Theta^{\varepsilon_0} \setminus \Theta^{\varepsilon_1}) dt \leq C \cdot \delta.$

*Proof.* (i) Note that

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

and that

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

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

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

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

$$(24) \quad \frac{\varepsilon_2 - \delta^2 \|\mu\|_{\infty, \Theta^{\varepsilon_0}}}{\delta \|\sigma\|_{\infty, \Theta^{\varepsilon_0}}} = -\log(\delta) - \frac{\|\mu\|_{\infty, \Theta^{\varepsilon_0}}}{\|\sigma\|_{\infty, \Theta^{\varepsilon_0}}} \delta \geq -\log(\delta) - \frac{\|\mu\|_{\infty, \Theta^{\varepsilon_0}}}{\|\sigma\|_{\infty, \Theta^{\varepsilon_0}}}.$$

An application of Lemma 3.4 together with (24) now yields

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

with

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

(ii) Observe that

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

Using again Lemma 3.1 we obtain the following.

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

Since

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

we arrive at

$$(25) \quad \frac{d(y, \Theta) - \|\mu\|_{\infty, \Theta^{\varepsilon_0}} h(y, \delta)}{h(y, \delta)^{1/2} \|\sigma\|_{\infty, \Theta^{\varepsilon_0}}} = -\log(\delta) - \frac{\|\mu\|_{\infty, \Theta^{\varepsilon_0}}}{\|\sigma\|_{\infty, \Theta^{\varepsilon_0}}} h(y, \delta)^{1/2} \geq -\log(\delta) - \frac{\|\mu\|_{\infty, \Theta^{\varepsilon_0}}}{\|\sigma\|_{\infty, \Theta^{\varepsilon_0}}}.$$

Lemma 3.4 together with (25) yields that

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

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

Using the previous lemmas we obtain the following.

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

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

and

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

*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 Lemmas 3.5(i) and 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 Lemmas 3.5(ii) and 3.3.  $\square$

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

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

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.  $\square$

**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  $h: \mathbb{R}^d \times (0, 1) \rightarrow (0, 1)$  be given by (19) and (20), and let  $X^h$  be given by (17) and (18). Then there exists a constant  $C_{\text{rmse}} > 0$  such that*

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

*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  $\delta$  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  $Z = (Z_t)_{t \geq 0}$  by  $Z_t = G(X_t)$ . It solves

$$(26) \quad dZ_t = \mu_G(Z_t)dt + \sigma_G(Z_t)dW_t, \quad t \geq 0, \quad Z_0 = G(x),$$

where  $\mu_G, \sigma_G$  are given by (13) and are globally Lipschitz by Lemma 2.7. Moreover,  $G$  and  $G^{-1}$  are globally Lipschitz by Lemma 2.6(v). So we obtain

$$(27) \quad \left( \mathbb{E} \left[ \sup_{0 \leq t \leq T} \|X_t - X_t^h\|^2 \right] \right)^{1/2} \leq L_{G^{-1}} \left( \mathbb{E} \left[ \sup_{0 \leq t \leq T} \|Z_t - G(X_t^h)\|^2 \right] \right)^{1/2}.$$

Now, denote by  $Z^h$  the Euler–Maruyama approximation of the process  $Z$  based on the step sizing function  $h$  given by (19) and (20). Using this scheme we can split the error as follows:

$$(28) \quad \begin{aligned} & \left( \mathbb{E} \left[ \sup_{0 \leq t \leq T} \|Z_t - G(X_t^h)\|^2 \right] \right)^{1/2} \\ & \leq \left( \mathbb{E} \left[ \sup_{0 \leq t \leq T} \|Z_t - Z_t^h\|^2 \right] \right)^{1/2} + \left( \mathbb{E} \left[ \sup_{0 \leq t \leq T} \|Z_t^h - G(X_t^h)\|^2 \right] \right)^{1/2}. \end{aligned}$$

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

$$(29) \quad \mathbb{E} \left[ \sup_{0 \leq t \leq T} \|Z_t - Z_t^h\|^2 \right] \leq c_1 \cdot \delta.$$

This can be shown by a simple modification of the error analysis of the Euler–Maruyama scheme for deterministic nonequidistant discretizations as, e.g., in [16, Theorem 10.2.2].

The second error term in (28) is the difference between the transformation applied to the time-continuous Euler–Maruyama approximation of  $X$  defined in (6) and the Euler–Maruyama approximation of the transformed process  $Z$  defined in (14). For all  $\tau \in [0, T]$  set

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

The Lipschitz continuity of  $G$ ,  $\mu_G$ , and  $\sigma_G$  and Lemma 3.2 imply that  $\sup_{\tau \in [0, T]} u(\tau) < \infty$ . For all  $x_1, x_2 \in \mathbb{R}^d$  define

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

and notice that  $\nu(x, x) = \mu_G(G(x))$ ,  $\sigma_G(G(x)) = G'(x)\sigma(x)$ .

By Itô'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

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

With the Cauchy–Schwarz inequality and the  $d$ -dimensional Burkholder–Davis–Gundy inequality (see, e.g., [14, Theorem III.3.28] or [13, Lemma 3.7]), we obtain

$$(30) \quad u(\tau) \leq 4T E_1(\tau) + 8d E_2(\tau) + 4T E_3(\tau) + 8d E_4(\tau)$$

with

$$\begin{aligned} (31) \quad E_1(\tau) &= \mathbb{E} \left[ \int_0^\tau \left\| \nu(X_s^h, X_{\underline{s}}^h) - \nu(X_{\underline{s}}^h, X_{\underline{s}}^h) \right\|^2 ds \right], \\ E_2(\tau) &= \mathbb{E} \left[ \int_0^\tau \left\| G'(X_s^h) \sigma(X_{\underline{s}}^h) - G'(X_{\underline{s}}^h) \sigma(X_{\underline{s}}^h) \right\|^2 ds \right], \\ E_3(\tau) &= \mathbb{E} \left[ \int_0^\tau \left\| \mu_G(G(X_{\underline{s}}^h)) - \mu_G(Z_{\underline{s}}^h) \right\|^2 ds \right], \\ E_4(\tau) &= \mathbb{E} \left[ \int_0^\tau \left\| \sigma_G(G(X_{\underline{s}}^h)) - \sigma_G(Z_{\underline{s}}^h) \right\|^2 ds \right]. \end{aligned}$$

*Step 2:* Now we estimate the above error terms. For  $E_1$ , using the linear growth property of  $\mu$  and  $\sigma$  and the properties of  $G$  we have

$$(32) \quad \begin{aligned} &\|\nu(x_1, x_2) - \nu(x_2, x_2)\|^2 \\ &\leq \begin{cases} K_1 \cdot (1 + \|x_2\|^4) \cdot \|x_1 - x_2\|^2, & \|x_1 - x_2\| = \rho(x_1, x_2), \\ K_2 \cdot (1 + \|x_2\|^4), & \text{otherwise,} \end{cases} \end{aligned}$$

with

$$(33) \quad K_1 = 2L_{G'}^2 C_\mu^2 + \frac{1}{2} L_{G''}^2 C_\sigma^4, \quad K_2 = 4C_\mu^2 \|G'\|_\infty^2 + C_\sigma^4 \|G''\|_\infty^2,$$

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  $\Theta$ . Further, set

$$(34) \quad K_3 = \sup \{ \|\nu(x_1, x_2)\|^2 : x_2 \in \Theta^{\varepsilon_0}, x_1 \in \mathbb{R}^d \}.$$

The latter quantity is finite due to our assumptions.

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

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

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

(i) From (32) we get that

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

An application of the Cauchy–Schwarz inequality and Lemma 3.2 now yield

$$(35) \quad \mathbb{E} \left[ \int_0^\tau \mathbf{1}_{\{\|X_{\underline{s}}^h - X_s^h\| = \rho(X_{\underline{s}}^h, X_s^h)\}} \left\| \nu(X_s^h, X_{\underline{s}}^h) - \nu(X_{\underline{s}}^h, X_s^h) \right\|^2 ds \right] \leq c_2 \cdot \delta.$$

(ii) Now consider the case that  $\|X_{\underline{s}}^h - X_s^h\| \neq \rho(X_{\underline{s}}^h, X_s^h)$  and that  $X_{\underline{s}}^h$  is more than  $\varepsilon_0$  away from  $\Theta$ . Here (32) gives

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

Since

$$\begin{aligned} \left\{ \|X_{\underline{s}}^h - X_s^h\| \neq \rho(X_{\underline{s}}^h, X_s^h) \right\} \cap \left\{ X_{\underline{s}}^h \notin \Theta^{\varepsilon_0} \right\} \\ \subseteq \left\{ X_{\underline{s}}^h \notin \Theta^{\varepsilon_0} \right\} \cap \left\{ \|X_{\underline{s}}^h - X_s^h\| \geq \varepsilon_0 \right\}, \end{aligned}$$

the Cauchy–Schwarz inequality together with (32) yields

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

Markov's inequality, i.e.,

$$\mathbb{P}(\|X_{\underline{s}}^h - X_s^h\| \geq \varepsilon_0) \leq \frac{\mathbb{E}[\|X_{\underline{s}}^h - X_s^h\|^4]}{\varepsilon_0^4},$$

and Lemma 3.2 now give

$$(36) \quad \begin{aligned} & \mathbb{E} \left[ \int_0^\tau \mathbf{1}_{\{\|X_{\underline{s}}^h - X_s^h\| \neq \rho(X_{\underline{s}}^h, X_s^h)\}} \mathbf{1}_{\{X_{\underline{s}}^h \notin \Theta^{\varepsilon_0}\}} \left\| \nu(X_s^h, X_{\underline{s}}^h) - \nu(X_{\underline{s}}^h, X_s^h) \right\|^2 ds \right] \\ & \leq c_3 \cdot \delta. \end{aligned}$$

(iii) The next case is that  $\|X_{\underline{s}}^h - X_s^h\| \neq \rho(X_{\underline{s}}^h, X_s^h)$  and  $X_{\underline{s}}^h$  lies in  $\Theta^{\varepsilon_0} \setminus \Theta^{\varepsilon_1}$ . Since

$$\nu(X_s^h, X_{\underline{s}}^h)^2 \cdot \mathbf{1}_{\{X_{\underline{s}}^h \in \Theta^{\varepsilon_0} \setminus \Theta^{\varepsilon_1}\}} \leq K_3,$$

we obtain

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

Since

$$\begin{aligned} & \left\{ X_{\underline{s}}^h \in \Theta^{\varepsilon_0} \setminus \Theta^{\varepsilon_1} \right\} \cap \left\{ \|X_{\underline{s}}^h - X_s^h\| \neq \rho(X_{\underline{s}}^h, X_s^h) \right\} \\ & \subseteq \left\{ X_{\underline{s}}^h \in \Theta^{\varepsilon_0} \setminus \Theta^{\varepsilon_1} \right\} \cap \left\{ \|X_{\underline{s}}^h - X_s^h\| \geq \varepsilon_1 \right\}, \end{aligned}$$

Lemma 3.5(iii) gives

$$(37) \quad \begin{aligned} & \mathbb{E} \left[ \int_0^\tau \mathbf{1}_{\{X_{\underline{s}}^h \in \Theta^{\varepsilon_0} \setminus \Theta^{\varepsilon_1}\}} \mathbf{1}_{\{\|X_{\underline{s}}^h - X_s^h\| \neq \rho(X_{\underline{s}}^h, X_s^h)\}} \left\| \nu(X_s^h, X_{\underline{s}}^h) - \nu(X_{\underline{s}}^h, X_s^h) \right\|^2 ds \right] \\ & \leq c_4 \cdot \delta. \end{aligned}$$

(iv) For the next case observe that

$$\begin{aligned} & \left\{ X_{\underline{s}}^h \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2} \right\} \cap \left\{ \|X_{\underline{s}}^h - X_s^h\| \neq \rho(X_{\underline{s}}^h, X_s^h) \right\} \\ & \subseteq \left\{ X_{\underline{s}}^h \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2} \right\} \cap \left\{ \|X_{\underline{s}}^h - X_s^h\| \geq d(X_{\underline{s}}^h, \Theta) \right\}, \end{aligned}$$

and so (34) and Lemma 3.5(ii) yield

$$(38) \quad \begin{aligned} & \mathbb{E} \left[ \int_0^\tau \mathbf{1}_{\{X_{\underline{s}}^h \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}\}} \mathbf{1}_{\{\|X_{\underline{s}}^h - X_s^h\| \neq \rho(X_{\underline{s}}^h, X_s^h)\}} \left\| \nu(X_s^h, X_{\underline{s}}^h) - \nu(X_{\underline{s}}^h, X_s^h) \right\|^2 ds \right] \\ & \leq 2K_3 \int_0^\tau \mathbb{P}(X_{\underline{s}}^h \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}; \|X_{\underline{s}}^h - X_s^h\| \neq \rho(X_{\underline{s}}^h, X_s^h)) ds \\ & \leq c_5 \cdot \delta. \end{aligned}$$

(v) For the final case, the boundedness of the coefficients on  $\Theta^{\varepsilon_0}$  and the fact that

$$\left\{ \|X_{\underline{s}}^h - X_s^h\| \neq \rho(X_{\underline{s}}^h, X_s^h) \right\} \cap \left\{ X_{\underline{s}}^h \in \Theta^{\varepsilon_2} \right\} \subseteq \left\{ X_{\underline{s}}^h \in \Theta^{\varepsilon_2} \right\}$$

together with the first statement of Lemma 3.6 yield that

$$(39) \quad \begin{aligned} & \mathbb{E} \left[ \int_0^T \mathbf{1}_{\{\|X_{\underline{s}}^h - X_s^h\| \neq \rho(X_{\underline{s}}^h, X_s^h)\}} \mathbf{1}_{\{X_{\underline{s}}^h \in \Theta^{\varepsilon_2}\}} \left\| \nu(X_s^h, X_{\underline{s}}^h) - \nu(X_{\underline{s}}^h, X_{\underline{s}}^h) \right\|^2 ds \right] \\ & \leq c_6 \cdot \log(1/\delta) \delta. \end{aligned}$$

Combining (31) with (35), (36), (37), (38), and (39) yields

$$(40) \quad E_1(\tau) \leq c_7 \cdot (1 + \log(1/\delta)) \delta.$$

For estimating  $E_2$  in (30), we exploit that  $G'$  is globally Lipschitz,  $\sigma$  satisfies a linear growth condition, and use Lemma 3.2 to obtain

$$(41) \quad \begin{aligned} E_2(\tau) & \leq L_{G'}^2 C_\sigma^2 \int_0^T \mathbb{E} \left[ (1 + \|X_{\underline{s}}^h\|^2) \|X_s^h - X_{\underline{s}}^h\|^2 \right] ds \\ & \leq L_{G'}^2 C_\sigma^2 \int_0^T \left( \mathbb{E} \left[ |1 + \|X_{\underline{s}}^h\|^2|^2 \right] \right)^{1/2} \left( \mathbb{E} [\|X_s^h - X_{\underline{s}}^h\|^4] \right)^{1/2} ds \leq c_8 \cdot \delta. \end{aligned}$$

For the remaining two terms in (30) we use the fact that  $\mu_G, \sigma_G$  are globally Lipschitz by Lemma 2.7. This gives

$$(42) \quad E_3(\tau) \leq L_{\mu_G}^2 \int_0^\tau \mathbb{E} [\|G(X_{\underline{s}}^h) - Z_{\underline{s}}^h\|^2] ds \leq L_{\mu_G}^2 \int_0^\tau u(s) ds,$$

$$(43) \quad E_4(\tau) \leq L_{\sigma_G}^2 \int_0^\tau \mathbb{E} [\|G(X_{\underline{s}}^h) - Z_{\underline{s}}^h\|^2] ds \leq L_{\sigma_G}^2 \int_0^\tau u(s) ds.$$

*Step 3:* Combining (30) with the estimates (40), (41), (42), and (43) we obtain

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

Gronwall's inequality yields

$$(44) \quad u(\tau) \leq c_{10} \exp(c_9 \tau) \cdot \delta(1 + \log(1/\delta)), \quad \tau \in [0, T].$$

Finally, combining (28) with (29) and (44), and the result with (27) concludes the proof.  $\square$

**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  $\mu$  and  $\sigma$  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  $\mu$  and  $\sigma$ .

**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.,

$$(45) \quad N(h, \delta) = \inf \{k \in \mathbb{N} : \tau_k \geq T\}.$$

Clearly, we have

$$N(h, \delta) \leq 1 + \int_0^T \frac{1}{h(X_t^h, \delta)} dt,$$

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  $h: \mathbb{R}^d \times (0, 1) \rightarrow (0, 1)$  be given by (19) and (20), and let  $X^h$  be given by (17) and (18). Then there exists a constant  $C_{cost} > 0$  such that*

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

*Proof.* We denote constants independent of  $\delta$  by  $c_1, c_2, \dots$ . We have

$$\begin{aligned} (46) \quad \mathbb{E}[N(h, \delta)] &\leq 1 + \mathbb{E} \left[ \int_0^T \frac{1}{h(X_{\underline{t}}^h, \delta)} dt \right] \\ &\quad + \int_0^T \mathbb{E} \left[ \frac{1}{h(X_{\underline{t}}^h, \delta)} \mathbf{1}_{\{X_{\underline{t}}^h \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}\}} \right] dt \\ &= 1 + I_1 + I_2 + I_3 \end{aligned}$$

with

$$\begin{aligned} (47) \quad I_1 &= \delta^{-2} \int_0^T \mathbb{P}(X_{\underline{t}}^h \in \Theta^{\varepsilon_2}) dt, \\ I_2 &= \int_0^T \mathbb{E} \left[ \frac{1}{h(X_{\underline{t}}^h, \delta)} \mathbf{1}_{\{X_{\underline{t}}^h \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}\}} \right] dt, \\ I_3 &= \delta^{-1} \int_0^T \mathbb{P}(X_{\underline{t}}^h \notin \Theta^{\varepsilon_1}) dt. \end{aligned}$$

We also have

$$(48) \quad I_3 \leq T \cdot \delta^{-1},$$

and by Lemma 3.6

$$(49) \quad I_1 \leq c_1 \cdot (1 + \log(1/\delta))\delta^{-1}.$$

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,

$$(50) \quad A(t) = \left\{ \|X_t^h - X_{\underline{t}}^h\| \leq d(X_{\underline{t}}^h, \Theta)/2 \right\}, \quad t \geq 0.$$

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  $\Theta$  is Lipschitz continuous with Lipschitz constant 1 (see [6, 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

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

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,

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

and so we obtain

$$(51) \quad \frac{1}{h(X_{\underline{t}}^h, \delta)} \mathbf{1}_{A(t) \cap \{X_{\underline{t}}^h \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}\}} \leq \frac{9}{4} \frac{\|\sigma\|_{\infty, \Theta^{\varepsilon_0}}^2 (\log(\delta))^2}{d(X_t^h, \Theta)^2} \mathbf{1}_{\{X_t^h \in \Theta^{\frac{3}{2}\varepsilon_1} \setminus \Theta^{\frac{1}{2}\varepsilon_2}\}}.$$

(ii) Since the minimal step size is  $\delta^2$ , we have

$$\begin{aligned} \frac{1}{h(X_{\underline{t}}^h, \delta)} \mathbf{1}_{A(t)^c \cap \{X_{\underline{t}}^h \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}\}} &\leq \frac{1}{\delta^2} \mathbf{1}_{A(t)^c \cap \{X_{\underline{t}}^h \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}\}} \\ (52) \quad &= \frac{1}{\delta^2} \mathbf{1}_{\{\|X_t^h - X_{\underline{t}}^h\| > d(X_{\underline{t}}^h, \Theta)/2; X_{\underline{t}}^h(\omega) \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}\}}. \end{aligned}$$

Combining (47) with (51) and (52) we obtain

$$\begin{aligned} I_2 &\leq \frac{9}{4} \|\sigma\|_{\infty, \Theta^{\varepsilon_0}}^2 (\log(\delta))^2 \int_0^T \mathbb{E} \left[ \frac{1}{d(X_t^h, \Theta)^2} \mathbf{1}_{\{X_t^h \in \Theta^{\frac{3}{2}\varepsilon_1} \setminus \Theta^{\frac{1}{2}\varepsilon_2}\}} \right] dt \\ &\quad + \delta^{-2} \int_0^T \mathbb{E} \left[ \mathbf{1}_{\{X_t^h \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}; \|X_t^h - X_{\underline{t}}^h\| > d(X_{\underline{t}}^h, \Theta)/2\}} \right] dt \\ (53) \quad &= \frac{9}{4} \|\sigma\|_{\infty, \Theta^{\varepsilon_0}}^2 (\log(\delta))^2 \int_0^T \mathbb{E} \left[ \frac{1}{\max\{\varepsilon_2/2, d(X_t^h, \Theta)\}^2} \mathbf{1}_{\{X_t^h \in \Theta^{\frac{3}{2}\varepsilon_1}\}} \right] dt \\ &\quad + \delta^{-2} \int_0^T \mathbb{E} \left[ \mathbf{1}_{\{X_t^h \in \Theta^{\varepsilon_1} \setminus \Theta^{\varepsilon_2}; \|X_t^h - X_{\underline{t}}^h\| > d(X_{\underline{t}}^h, \Theta)/2\}} \right] dt \\ &=: I_{21} + I_{22}. \end{aligned}$$

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

$$(54) \quad I_{22} \leq c_2 \cdot (1 + \log(1/\delta)) \delta^{-1}.$$

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

$$(55) \quad I_{21} \leq c_3 \left( 2\varepsilon_2^{-1} + \int_{\frac{1}{2}\varepsilon_2}^{\frac{3}{2}\varepsilon_1} \frac{\|\sigma\|_{\infty, \Theta^{\varepsilon_0}}^2 (\log(\delta))^2}{x^2} dx \right) \leq c_4 \cdot \log(1/\delta) \delta^{-1}.$$

Combining (53) with (54) and (55) ensures the existence of a constant  $c_5 > 0$  such that

$$(56) \quad I_2 \leq c_5 \cdot (1 + \log(1/\delta)) \delta^{-1}.$$

Now, (46) together with the estimates (48), (49), and (56) yields the assertion.  $\square$

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

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

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

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

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

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

then also

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

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

For both quantities we choose  $\delta = 2^{-2}, 2^{-3}, \dots, 2^{-10}$ ,  $M = 5 \cdot 10^4$  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

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

and initial value  $x = 1.5$ . 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  $\text{res} = 6.0412 \cdot 10^2$ , 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  $\text{res} = 1.0674 \cdot 10^{-2}$ . This is in good accordance with the predicted asymptotic behavior from Theorem 4.1, respectively, Theorem 5.1.

The second test equation is again a scalar equation with

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

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  $\text{res} = 2.0217 \cdot 10^2$  and

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

with a residuum of  $\text{res} = 2.2471 \cdot 10^{-2}$ . 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 [7]. However, the latter work also indicates that discontinuous coefficients may lead to an unstable behavior of  $\text{msq}(\delta)$ . 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

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

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  $\text{res} = 2.3685 \cdot 10^1$  and

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

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

**Acknowledgments.** The authors are grateful to Mike Giles for numerous discussions. The time stepping strategy of this paper has been inspired by work of M. Giles and K. Ramanan on adaptive multilevel Monte Carlo time stepping strategies for reflected diffusions. Furthermore, the authors are also grateful to Martin Schmidt for a discussion about geometric aspects of this work and to Larisa Yaroslavtseva for pointing out to us an inaccuracy in the proof of a lemma.

A part of this article was written while M. Szölgyenyi was affiliated with the Seminar for Applied Mathematics and the RiskLab Switzerland, ETH Zurich, Rämistrasse 101, 8092 Zurich, Switzerland and with the Institute of Statistics and Mathematics, Vienna University of Economics and Business, Welthandelsplatz 1, 1020 Vienna, Austria.

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 J. Sci. Comput., 24 (2003), pp. 848–864.
- [2] P. ÉTORÉ AND M. MARTINEZ, *Exact simulation for solutions of one-dimensional stochastic differential equations with discontinuous drift*, ESAIM Probab. Stat., 18 (2014), pp. 686–702.
- [3] W. FANG AND M. B. GILES, *Adaptive Euler–Maruyama method for SDEs with non-globally Lipschitz drift*, in Monte Carlo and Quasi-Monte Carlo Methods, Springer, Cham, Switzerland, 2018, pp. 217–234.
- [4] N. FRIKHA, *On the weak approximation of a skew diffusion by an Euler-type scheme*, Bernoulli, 24 (2018), pp. 1653–1691.
- [5] J. G. GAINES AND T. J. LYONS, *Variable step size control in the numerical solution of stochastic differential equations*, SIAM J. Appl. Math., 57 (1997), pp. 1455–1484.
- [6] D. GILBARG AND N. S. TRUDINGER, *Elliptic Partial Differential Equations of Second Order*, 2nd ed., Grundlehren Math. Wiss. 224, Springer, Berlin, 1983.
- [7] 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*, preprint, arXiv:1705.04562, 2017.
- [8] I. GYÖNGY, *A note on Euler’s approximation*, Potential Anal., 8 (1998), pp. 205–216.
- [9] N. HALIDIAS AND P. E. KLOEDEN, *A note on the Euler–Maruyama scheme for stochastic differential equations with a discontinuous monotone drift coefficient*, BIT, 48 (2008), pp. 51–59.
- [10] H. HOEL, E. VON SCHWERIN, A. SZEPESSY, AND R. TEMPONE, *Adaptive multilevel Monte Carlo simulation*, in Numerical Analysis of Multiscale Computations, Springer, Berlin, 2012, pp. 217–234.
- [11] N. HOFMANN, T. MÜLLER-GRONBACH, AND K. RITTER, *Optimal approximation of stochastic differential equations by adaptive step-size control*, Math. Comp., 69 (2000), pp. 1017–1034.
- [12] N. HOFMANN, T. MÜLLER-GRONBACH, AND K. RITTER, *The optimal discretization of stochastic differential equations*, J. Complexity, 17 (2001), pp. 117–153.
- [13] M. HUTZENTHALER, A. JENTZEN, AND P. E. KLOEDEN, *Strong convergence of an explicit numerical method for SDEs with nonglobally Lipschitz continuous coefficients*, Ann. Appl. Probab., 22 (2012), pp. 1611–1641.
- [14] I. KARATZAS AND S. E. SHREVE, *Brownian Motion and Stochastic Calculus*, 2nd ed., Grad. Texts Math., Springer, New York, 1991.
- [15] C. KELLY AND G. J. LORD, *Adaptive time-stepping strategies for nonlinear stochastic systems*, IMA J. Numer. Anal., 38 (2018), pp. 1523–1549.
- [16] P. E. KLOEDEN AND E. PLATEN, *Numerical Solutions of Stochastic Differential Equations*, Stoch. Model. Appl. Probab., Springer, Berlin, 1992.
- [17] 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*, J. Comput. Appl. Math., 326 (2017), pp. 138–158.
- [18] N. V. KRYLOV, *Controlled Diffusion Processes*, Appl. Math., Springer, New York, 1980.
- [19] H. LAMBA, *An adaptive timestepping algorithm for stochastic differential equations*, J. Comput. Appl. Math., 161 (2003), pp. 417–430.
- [20] H. LAMBA, J. C. MATTINGLY, AND A. M. STUART, *An adaptive Euler–Maruyama scheme for SDEs: Convergence and stability*, IMA J. Numer. Anal., 27 (2007), pp. 479–506.
- [21] G. LEOBACHER AND M. SZÖLGYENYI, *A numerical method for SDEs with discontinuous drift*, BIT, 56 (2016), pp. 151–162.
- [22] G. LEOBACHER AND M. SZÖLGYENYI, *A strong order 1/2 method for multidimensional SDEs with discontinuous drift*, Ann. Appl. Probab., 27 (2017), pp. 2383–2418.
- [23] G. LEOBACHER AND M. SZÖLGYENYI, *Correction Note for: A Strong Order 1/2 Method for Multidimensional SDEs with Discontinuous Drift*, manuscript.
- [24] G. LEOBACHER AND M. SZÖLGYENYI, *Convergence of the Euler–Maruyama method for multi-dimensional SDEs with discontinuous drift and degenerate diffusion coefficient*, Numer. Math., 138 (2018), pp. 219–239.
- [25] G. LEOBACHER, M. SZÖLGYENYI, AND S. THONHAUSER, *On the existence of solutions of a class of SDEs with discontinuous drift and singular diffusion*, Electron. Commun. Probab., 20 (2015), pp. 1–14.
- [26] C. LESTER, C. A. YATES, M. B. GILES, AND R. E. BAKER, *An adaptive multi-level simulation algorithm for stochastic biological systems*, J. Chem. Phys., 142 (2015), 024113.

- [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 (2002), pp. 185–232.
- [28] T. MÜLLER-GRONBACH, *The optimal uniform approximation of systems of stochastic differential equations*, Ann. Appl. Probab., 12 (2002), pp. 664–690.
- [29] T. MÜLLER-GRONBACH, *Optimal pointwise approximation of SDEs based on Brownian motion at discrete points*, Ann. Appl. Probab., 14 (2004), pp. 1605–1642.
- [30] T. MÜLLER-GRONBACH AND L. YAROSLAVTSEVA, *On the Performance of the Euler-Maruyama Scheme for SDEs with Discontinuous Drift Coefficient*, preprint, arXiv:1809.08423, 2018.
- [31] 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 (2016), pp. 1793–1819.
- [32] H. L. NGO AND D. TAGUCHI, *On the Euler-Maruyama approximation for one-dimensional stochastic differential equations with irregular coefficients*, IMA J. Numer. Anal., 37 (2017), pp. 1864–1883.
- [33] H. L. NGO AND D. TAGUCHI, *Strong convergence for the Euler-Maruyama approximation of stochastic differential equations with discontinuous coefficients*, Statist. Probab. Lett., 125 (2017), pp. 55–63.
- [34] G. O. ROBERTS AND R. L. TWEEDIE, *Exponential convergence of Langevin distributions and their discrete approximations*, Bernoulli, 2 (1996), pp. 341–363.
- [35] A. A. SHARDIN AND M. SZÖLGYENYI, *Optimal control of an energy storage facility under a changing economic environment and partial information*, Int. J. Theor. Appl. Finance, 19 (2016), pp. 1–27.
- [36] A. YU. VERETENNIKOV, *On stochastic equations with degenerate diffusion with respect to some of the variables*, Math. USSR Izv., 22 (1984), pp. 173–180.