Numerical methods for SDEs with drift discontinuous on a set of positive reach

G. Leobacher and M. Szölgyenyi

Numerical methods for SDEs with drift discontinuous on a set of positive reach

Gunther Leobacher, Michaela Szölgyenyi

KFU Graz, ETH Zürich

Abstract. For time-homogeneous stochastic differential equations (SDEs) it is enough to know that the coefficients are Lipschitz to conclude existence and uniqueness of a solution, as well as the existence of a strongly convergent numerical method for its approximation. Here we introduce a notion of piecewise Lipschitz functions and study SDEs with a drift coefficient satisfying only this weaker regularity condition. For these SDEs we can construct a strongly convergent approximation scheme, if the set of discontinuities is a sufficiently smooth hypersurface satisfying the geometrical property of being of positive reach. We then arrive at similar conclusions as in the Lipschitz case. We will see that, although SDEs are in the center of our interest, we will talk surprisingly little about probability theory here.

1   Introduction

Stochastic differential equations (SDEs) are essential for many models in mathematical finance, risk theory, biology, physics, and chemistry. Usually, these equations cannot be solved explicitly. Hence, we are interested in finding numerical methods with positive convergence speed for solving them.

We consider general SDEs on the $\mathbb{R}^d$, which are of the form

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

with initial value $x \in \mathbb{R}^d$, drift coefficient $\mu : \mathbb{R}^d \longrightarrow \mathbb{R}^d$, diffusion coefficient $\sigma : \mathbb{R}^d \longrightarrow \mathbb{R}^{d \times m}$, and m-dimensional standard Brownian motion W (thus adding

noise to the ordinary differential equation). Little generality is lost if we assume $m = d$, and we will do so throughout this article.

By a (strong) solution we mean a continuous stochastic process X that is adapted to the filtration generated by W and that satisfies

$$X_t = x + \int_0^t \mu(X_s)ds + \int_0^t \sigma(X_s)dW_s \tag{2}$$

for all $t \geq 0$ almost surely. The solution X is unique, if the paths of any other solution to (2) coincide with those of X almost surely.

The second integral in (2) is Itô's stochastic integral, the construction of which we will not repeat here. Suffice it to mention that for K from a suitable class of stochastic processes it holds that

$$\int_0^t K_s \, dW_s = \lim_{n \to \infty} \sum_{k=1}^{\lfloor 2^n t \rfloor - 1} K_{k2^{-n}} \left(W_{(k+1)2^{-n}} - W_{k2^{-n}} \right),$$

reminding us of the Riemann integral (but with evaluation of the integrand only in the left boundary of small intervals). A particularity of Itô's integral is that there appears a correction term in the fundamental theorem of calculus, that is, for $X_t = X_0 + \int_0^t H_s ds + \int_0^t K_s dWs$ and for a sufficiently regular function $f : \mathbb{R} \longrightarrow \mathbb{R}$,

$$f(X_t) = f(X_0) + \int_0^t f'(X_s)H_s ds + \int_0^t f'(X_s)K_s dW_s + \frac{1}{2}\int_0^t f''(X_s)K_s^2 ds.$$

This is known as Itô's formula. The rigorous construction of the stochastic integral gave meaning to the concept of a solution of an SDE. In addition to that Itô [4] proved that a unique solution to (1) exists, whenever μ and σ are Lipschitz-continuous.

Under the same assumptions Maruyama [12] proved that the Euler-Maruyama (EM) scheme

$$X_t^{\delta} = x + \int_0^t \mu(X_{\underline{s}}^{\delta})ds + \int_0^t \sigma(X_{\underline{s}}^{\delta})dW_s,$$

with $\underline{s} = j\delta$ for $s \in [j\delta, (j+1)\delta)$, $j = 0, \ldots, (T - \delta)/\delta$, (which reminds us of the Euler scheme for ordinary differential equations, but with an additional term corresponding to the stochastic integral) converges with strong order $1/2$. In general we say that a numerical approximation X^{δ} converges with strong order γ, if for any fixed $T > 0$, there exists a constant C such that for sufficiently small step-size $\delta > 0$ it holds that

$$\mathbb{E}\left(\sup_{0 \leq t \leq T} \|X_t - X_t^{\delta}\|^2 \right)^{1/2} \leq C\delta^{\gamma}.$$

2

Higher order algorithms exist under stronger regularity conditions on the coefficients, most notably the Milstein method and stochastic Runge-Kutta schemes, see Kloeden and Platen [7].

The question of how to solve SDEs with irregular (non-globally Lipschitz) coefficients approximately is a very active topic of research. There is still a big gap between the assumptions on the coefficients of these equations under which strong convergence with convergence rate has been proven in the scientific literature, and the assumptions that equations in real-world applications satisfy.

In contrast to that, several delimiting results have been proven recently, stating that a certain SDE with relatively well-behaved (infinitely often differentiable) coefficients cannot be solved approximately in finite time, cf. Hairer et al. [3], Jentzen et al. [5], Müller-Gronbach and Yaroslavtseva [13], Yaroslavtseva [24]. However, there is still a big discrepancy between the assumptions on the coefficients under which convergence with strong convergence rate has been proven and the properties of the coefficients of the SDE presented in Hairer et al. [3].

Here we narrow the gap described above by settling convergence with positive convergence speed of a numerical method for d-dimensional SDEs with discontinuous drift and degenerate diffusion coefficient. First steps in this direction have previously been made by Ngo and Taguchi [16], who proved convergence of order up to $1/4$ of the Euler-Maruyama method for d-dimensional SDEs which have a discontinuous, bounded drift that satisfies a one-sided Lipschitz condition and a Hölder continuous, bounded, and uniformly non-degenerate diffusion coefficient. In Ngo and Taguchi [14, 15] they do not need the one-sided Lipschitz condition any more, but the result only works for one-dimensional SDEs and relies on uniform non-degeneracy of the diffusion coefficient.

SDEs with discontinuous drift appear naturally when studying stochastic optimal control problems with bang-bang type optimal strategies, that is with strategies of the form $\mathbf{1}_{\mathcal{S}}(X)$ for a measurable set $\mathcal{S} \subseteq \mathbb{R}^d$. If in addition only a noisy signal of the underlying state process X is available, then filtering this signal leads to a degenerate diffusion coefficient and increases the dimension substantially. Examples can be found in Sass and Haussmann [20], Rieder and Bäuerle [18], Frey et al. [2], Leobacher et al. [11], Szölgyenyi [23], Shardin and Szölgyenyi [21], Shardin and Wunderlich [22].

The idea for tackling the problem is illustrated in Figure 1: to overcome the issues caused by a discontinuous drift coefficient, we want to find a transform G with the property that the coefficients of the transformed SDE for $G(X)$ are Lipschitz. Then we want to apply the EM scheme to that SDE, which converges with strong order $1/2$, to obtain an approximation of the solution to the transformed SDE. In the end, we want to transform back to obtain an approximation of the solution to the original SDE (1). In Figure 1, the set of discontinuities of the drift is illustrated by a smooth curve. Indeed, we need to make some assumptions to that end so that we can carry through our idea.
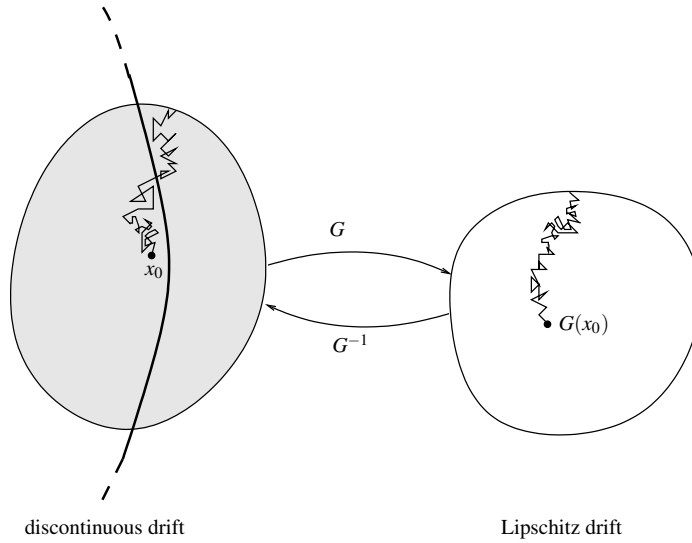
Figure 1 – A sketch of the idea for the construction of our numerical method.

Thus, we have to solve the following tasks:

1. construct G and all prove necessary properties;
2. prove an existence and uniqueness result;
3. construct a numerical method using G (called GM) and prove convergence and convergence rate;
4. prove convergence and convergence rate for EM starting from GM.

We will start with presenting our results in dimension one, and subsequently we will show how these ideas can be extended to general dimension.

This is a review article; the results and examples presented here, can be found in Leobacher and Szölgyenyi [8, 9, 10].

2   Result in dimension one

In order to construct an appropriate $G : \mathbb{R} \longrightarrow \mathbb{R}$, we have to know how such a transform acts on the coefficients: assuming existence of a solution X and also validity of Itô's formula for X and G we get

$$G(X_t) = G(x) + \int_0^t G'(X_s)\mu(X_s)ds + \int_0^t G'(X_s)\sigma(X_s)dW_s + \int_0^t \frac{1}{2}G''(X_s)\sigma(X_s)^2 ds.$$

4

Thus $Z = G(X)$ is the solution of an SDE with coefficients

$$\tilde{\mu}(Z) = G'(G^{-1}(Z))\mu(G^{-1}(Z)) + \frac{1}{2}G''(G^{-1}(Z))\sigma(G^{-1}(Z))^2,$$

$$\tilde{\sigma}(Z) = G'(G^{-1}(Z))\sigma(G^{-1}(Z)).$$

Hence, G maps $X \mapsto Z$ and it transforms μ, σ into $\tilde{\mu}, \tilde{\sigma}$.

We see that if $G \in C^2$ – the classical assumption for Itô's formula – then $\tilde{\mu}, \tilde{\sigma}$ are continuous, if and only if μ, σ are continuous. However if $G \in C^1$ and σ is continuous and non-zero, then we can offset jumps of μ with jumps of G''. So we can get continuous $\tilde{\mu}$ from discontinuous μ with a less smooth transform. Note that $\tilde{\sigma}$ is continuous in either case. Hence, we choose $G \in C^1$ to be able to eliminate the discontinuities from the drift. Note that we will have to verify that the heuristic application of Itô's formula above is valid, since the classical Itô formula holds for C^2 functions.

With this, we are able to relax the Lipschitz condition on the drift.

Definition 2.1. A function $\mu : \mathbb{R} \longrightarrow \mathbb{R}$ is called *piecewise Lipschitz*, if there are finitely many points $\xi_1 < \cdots < \xi_m$ such that the restriction of μ to each of the intervals $(-\infty, \xi_1), (\xi_m, \infty)$ and (ξ_k, ξ_{k+1}), $k = 1, \ldots, m-1$, is Lipschitz.

For the presentation here, we now assume that μ is piecewise Lipschitz with only one jump in ξ, but note that our result also holds for multiple jumps. Let

- μ be Lipschitz on $(-\infty, \xi)$ and (ξ, ∞);
- $\sigma : \mathbb{R} \longrightarrow \mathbb{R}$ be Lipschitz with $\sigma(\xi) \neq 0$.

Note that the last condition is by far weaker than uniform non-degeneracy, as for non-degeneracy one would need σ to be bounded away from 0 on the whole of $\mathbb{R}$. We define the transform $G : \mathbb{R} \longrightarrow \mathbb{R}$ by

$$G(x) = x + \alpha(x - \xi)|x - \xi|\phi\left(\frac{x - \xi}{c}\right) =: x + \alpha\bar{\phi}(x), \tag{3}$$

where α, c are appropriate constants, and

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

localizes the impact of G. If $0 < c < 1/6|\alpha|$, then $G' > 0$, and hence G is globally invertible. Furthermore, we can prove that G and G^{-1} are Lipschitz.

Setting $Z = G(X)$, we have

$$dZ_t = \tilde{\mu}(Z_t)dt + \tilde{\sigma}(Z_t)dW_t,$$

where

$$\tilde{\mu}(z) = \mu(G^{-1}(z)) + \frac{1}{2}\alpha\bar{\phi}''(G^{-1}(z))\sigma(G^{-1}(z))^2 + \alpha\bar{\phi}'(G^{-1}(z))\mu(G^{-1}(z)),$$

$$\tilde{\sigma}(z) = \sigma(G^{-1}(z)) + \alpha\bar{\phi}'(G^{-1}(z))\sigma(G^{-1}(z)).$$

In order to offset the jump of μ in ξ by the jump of G'' (by construction also in ξ), we choose α as

$$\mu(\xi+) + \frac{1}{2}\alpha\bar{\phi}''(\xi+)\sigma(\xi)^2 = \mu(\xi-) + \frac{1}{2}\alpha\bar{\phi}''(\xi-)\sigma(\xi)^2 \implies \alpha = \frac{\mu(\xi-) - \mu(\xi+)}{2\sigma(\xi)^2}.$$

With this choice of α we have that $\tilde{\mu}$ is continuous.

Lemma 2.2 (Elementary but essential). *Let $\tilde{\mu} : \mathbb{R} \longrightarrow \mathbb{R}$ be a function satisfying*

1. *$\tilde{\mu}$ is continuous;*
2. *$\tilde{\mu}$ is piecewise Lipschitz.*

Then $\tilde{\mu}$ is Lipschitz.

Altogether we have that the coefficients of the SDE for Z are Lipschitz.

Now, we are ready to prove the following theorem.

Theorem 2.3 (Leobacher and Szölgyenyi [8]). *Let μ be piecewise Lipschitz and let σ be Lipschitz and $\mu(\xi+) \neq \mu(\xi-) \implies \sigma(\xi) \neq 0$.*
Then there exists a unique strong solution to the one-dimensional version of (1).

The proof works as follows:

- show that the SDE for $Z = G(X)$ has Lipschitz coefficients using Lemma 2.2;
- then by Itô's theorem, there exists a unique strong solution to this SDE;
- set $X = G^{-1}(Z)$ and apply Itô's formula to it, to see that

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t.$$

So we have constructed a process X that solves our SDE. There is one issue that we have already mentioned above: $G^{-1} \notin C^2$. But in 1D, Itô's formula holds nevertheless, see [6, Problem 7.3].

As sketched in Figure 1 above, the transformation method in a natural way also leads to the following numerical scheme.

6

Algorithm 2.4 (Leobacher and Szölgyenyi [8]). Given μ, σ, x, T, and the step-size $\delta > 0$,

1. precompute $G, G^{-1}, \tilde{\mu}, \tilde{\sigma}$;
2. solve $dZ = \tilde{\mu}(Z)dt + \tilde{\sigma}(Z)dW$, $Z_0 = G(x)$ on $[0,T]$ using the EM method to obtain the EM approximation Z^δ;
3. compute the numerical approximation $\bar{X}_t = G^{-1}(Z_t^\delta)$, for $t \in [0,T]$.

Theorem 2.5 (Leobacher and Szölgyenyi [8]). *Let μ be piecewise Lipschitz and let σ be Lipschitz and $\mu(\xi+) \neq \mu(\xi-) \Longrightarrow \sigma(\xi) \neq 0$.*

Then Algorithm 2.4 converges with strong order $1/2$.

The proof is straightforward: Maruyama [12] showed that for sufficiently small step-size $\delta > 0$,

$$\mathbb{E}\left(\sup_{0 \leq t \leq T} |Z_t - Z_t^\delta|^2 \right)^{1/2} \leq C\delta^{1/2}.$$

Denote by $L_{G^{-1}}$ the Lipschitz constant of G^{-1}. We get

$$\mathbb{E}\left(\sup_{0 \leq t \leq T} |X_t - \bar{X}_t|^2 \right)^{1/2} = \mathbb{E}\left(\sup_{0 \leq t \leq T} |G^{-1}(Z_t) - G^{-1}(Z_t^\delta)|^2 \right)^{1/2}$$

$$\leq L_{G^{-1}} \mathbb{E}\left(|Z_t - Z_t^\delta|^2 \right)^{1/2} \leq L_{G^{-1}} C\delta^{1/2}.$$

The following theorem is of particular relevance for the practical implementation and efficiency of our algorithm and shows that in 1D our result is already quite satisfactory:

Theorem 2.6 (Leobacher and Szölgyenyi [8]). *We can define an alternative transform $\widehat{G}$ which fulfills all the necessary properties and which is piecewise cubic.*

The relevance of this theorem lies in the fact that for a piecewise cubic function the inverse can easily be computed explicitly.

3 Result in general dimension

Extending our results to the multidimensional setting poses several challenges:

1. introduce a notion of *piecewise Lipschitz* functions;
2. prove that piecewise Lipschitz + continuous implies Lipschitz;
3. find a transform G that makes the drift continuous;
4. show that G has a global inverse;
5. show that Itô's formula holds for G^{-1}.

In this section we will sketch how these challenges were addressed.

3.1 Piecewise Lipschitz functions on the $\mathbb{R}^d$

There is no unique or universally accepted notion of a piecewise Lipschitz function on a subset of the $\mathbb{R}^d$. Below we propose such a definition that generalizes the one-dimensional notion.

We call a continuous function $\gamma : [0,1] \longrightarrow A \subseteq \mathbb{R}^d$ a *curve in A from $\gamma(0)$ to $\gamma(1)$* and we denote by

$$\ell(\gamma) := \sup \left\{ \sum_{k=1}^{n} |\gamma(t_k) - \gamma(t_{k-1})| : n \in \mathbb{N}, 0 = t_0 < \cdots < t_n = 1 \right\}$$

its (possibly infinite) length.

Definition 3.1. Let $\emptyset \neq A \subseteq \mathbb{R}^d$. Define the *intrinsic metric* on A by

$$\rho(x,y) := \inf\{\ell(\gamma) : \gamma \text{ a curve in } A \text{ from } x \text{ to } y\}.$$

Here, the infimum over an empty set is defined as ∞.

Definition 3.2. A function $\mu : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ is *piecewise Lipschitz*, if there exists a hypersurface Θ with finitely many connected components such that the restriction $\mu|_{\mathbb{R}^d \setminus \Theta}$ is Lipschitz w.r.t. the intrinsic metric on $\mathbb{R}^d \setminus \Theta$, and w.r.t. the Euclidean metric on $\mathbb{R}^m$.

In that case we call Θ an *exceptional set* for μ.

Note that the definition coincides with Definition 2.1 for $d = 1$. It shares also some basic and well-known properties with the elementary definition.

Proposition 3.3. *Let Θ be a hypersurface in $\mathbb{R}^d$ and let $\mu : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ be a function such that $\mu|_{\mathbb{R}^d \setminus \Theta}$ is differentiable with bounded derivative.*

Then μ is piecewise Lipschitz with exceptional set Θ and

$$\sup_{x,y \in \mathbb{R}^d \setminus \Theta : \rho(x,y) > 0} \frac{\|\mu(x) - \mu(y)\|}{\rho(x,y)} = \sup_{x \in \mathbb{R}^d \setminus \Theta} \|\mu'(x)\|.$$

The following lemma is almost trivial in dimension one, but not so in general dimension:

Lemma 3.4. *Let $\mu : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ be a function such that*

1. *μ is continuous;*
2. *μ is piecewise Lipschitz with exceptional set Θ;*
3. *Θ is such that for all $x, y \in \mathbb{R}^d \setminus \Theta$ and all $\eta > 0$ there exists a curve γ in the $\mathbb{R}^d$ from x to y such that $\ell(\gamma) < \|y - x\| + \eta$ and $\#(\gamma \cap \Theta) < \infty$.*
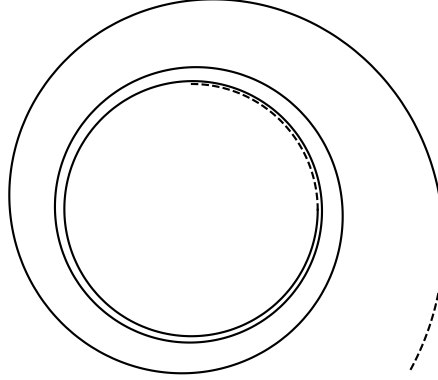
Figure 2 – An example for a hypersurface with bounded derivative of the unit normal vector which is not of positive reach.

Then μ is Lipschitz (w.r.t. the Euclidean norm) with Lipschitz constant

$$L_\mu = \sup_{x,y \in \mathbb{R}^d \setminus \Theta : \rho(x,y) > 0} \frac{\|\mu(x) - \mu(y)\|}{\rho(x,y)}.$$

Lemma 3.4 differs from Lemma 2.2 essentially by item 3, which is trivially satisfied in dimension one by our definition of 'piecewise Lipschitz'.

Figure 2 shows an example of a two-dimensional C^∞-hypersurface (i.e. a curve) for which item 3 of Lemma 3.4 is not satisfied. The following notion will prove useful for this issue:

Definition 3.5. A subset $\Theta \subseteq \mathbb{R}^d$ is of *positive reach*, if there exists $\varepsilon > 0$ such that for every $x \in \mathbb{R}^d$ with $d(x, \Theta) < \varepsilon$ there is a unique $p \in \Theta$ with $\|x - p\| = d(x, \Theta) := \inf\{\|x - \xi\| : \xi \in \Theta\}$.

If Θ has positive reach, then the projection map p which assigns to x its closest point $p(x)$ on Θ is a well-defined single-valued map on $\Theta^\varepsilon := \{x \in \mathbb{R}^d : d(x, \Theta) < \varepsilon\}$ for some $\varepsilon > 0$. Examples of hypersurfaces having this property include hyperplanes and all compact C^2-hypersurfaces, which follows from the lemma in Foote [1], where it is also shown that the projection map p is in C^{k-1} if Θ is in C^k. We will always assume that the set of discontinuities of the drift coefficient is of positive reach.

The projection map p will play a prominent role in the construction of the multivariate transform G.

One consequence of the positive reach property for Θ is that item 3 of Lemma 3.4 is automatically satisfied. This is the assertion of Leobacher and Szölgyenyi [9,

Lemma 3.11], the proof of which is surprisingly technical. Another useful consequence is that the derivative of the unit normal vector is bounded, see Leobacher and Szölgyenyi [9, Lemma 3.10].

3.2 Definition of the transform and main results

Our choice of the transform G is

$$G(x) = x + \alpha(p(x))\tilde{\phi}(x),$$

where

$$\tilde{\phi}(x) = (x - p(x)) \cdot n(p(x)) \|x - p(x)\| \phi\left(\frac{\|x - p(x)\|}{c}\right).$$

This should be compared to the 1D analog, equation (3). In Leobacher and Szölgyenyi [9, Theorem 3.14 and Lemma 3.18] it is proven that under the assumptions of Theorem 3.7 below, c can always be chosen sufficiently small, so that G has a global inverse by Hadamard's global inverse function theorem [19, Theorem 2.2]. In 1D the constant α had the purpose of making sure that the jump of μ is offset by the jump of G''. In general dimension, α is defined on the hypersurface Θ:

$$\alpha(\xi) = \lim_{h \to 0} \frac{\mu(\xi - hn(\xi)) - \mu(\xi + hn(\xi))}{2\|\sigma(\xi)^\top n(\xi)\|^2}, \qquad \xi \in \Theta. \tag{4}$$

Although α depends on the choice of the normal unit vector, it is readily checked that G does not.

We will need to make additional assumptions on μ and σ to guarantee existence and sufficient regularity of α and, a fortiori, of G.

It remains to show that Itô's formula holds for G^{-1}. This follows from the following special case of [17, Theorem 2.1].

Theorem 3.6 (Itô's formula)**.** *Let X be a d-dimensional Itô process and let $b :$ $\mathbb{R}^{d-1} \longrightarrow \mathbb{R}$ be a C^2-function. Let furthermore $f_1, f_2 : \mathbb{R}^d \longrightarrow \mathbb{R}$ be C^2-functions such that the function $f : \mathbb{R}^d \longrightarrow \mathbb{R}$ defined by*

$$f(x) = f_1(x)\mathbf{1}_{x_d \leq b(x_1,\dots,x_{d-1})} + f_2(x)\mathbf{1}_{x_d > b(x_1,\dots,x_{d-1})}$$

is in C^1. Then Itô's formula holds for X and f.

We have the following existence and uniqueness result.

Theorem 3.7 (Leobacher and Szölgyenyi [9]). *Let the following assumptions hold:*

- *$\mu : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ is piecewise Lipschitz with exceptional set Θ;*
- *$\Theta \in C^3$ and has positive reach;*
- *$\sigma : \mathbb{R}^d \longrightarrow \mathbb{R}^{d \times d}$ is Lipschitz and $\|\sigma(\xi)^\top n(\xi)\|^2 \geq c_0 > 0$ for all $\xi \in \Theta$;*
- *μ, σ are bounded on Θ^ε for some $\varepsilon > 0$;*
- *μ, σ are such that α, as described in (4), is well-defined and has bounded derivatives up to order 3.*

Then there exists a unique strong solution to (1).

We remark that the assumptions of Theorem 3.7 impose extra regularity on μ, σ only close to, and on Θ. Away from Θ we basically have the classical Lipschitz requirements. In analogy to the one-dimensional result, we have the following:

Theorem 3.8 (Leobacher and Szölgyenyi [9]). *Let the assumptions of Theorem 3.7 hold. Then, also in the multidimensional setting, Algorithm 2.4 converges with strong order $1/2$.*

3.3 Example

We apply our Algorithm 2.4 to solve an example of an SDE where the drift is discontinuous on the unit circle in the $\mathbb{R}^2$, i.e. the exceptional set $\Theta = \{x \in \mathbb{R}^2 | x_1^2 + x_2^2 = 1\}$, and the diffusion coefficient is degenerate. Let

$$dX_t = \mu(X_t^1, X_t^2)dt + \sigma(X_t^1, X_t^2)dW_t,$$

where

$$\mu(x_1, x_2) = \begin{cases} (1,1)^\top, & x_1^2 + x_2^2 > 1 \\ (-x_1, x_2)^\top, & x_1^2 + x_2^2 \leq 1, \end{cases}$$

$$\sigma(x_1, x_2) = \frac{1}{1 + x_1^2 + x_2^2} \begin{pmatrix} x_1 & 0 \\ x_2 & 0 \end{pmatrix}.$$

Figure 3 shows the estimated L^2-error of GM for this example. We observe that GM shows the convergence behaviour we expect from our theoretical result, namely it converges as fast as $\delta^{1/2}$, i.e. the purple dotted line has the same slope as the yellow line. So in principle we could be satisfied. We have constructed the first numerical method that is proven to converge for a rather general class of SDEs with discontinuous drift and we have established its convergence speed.
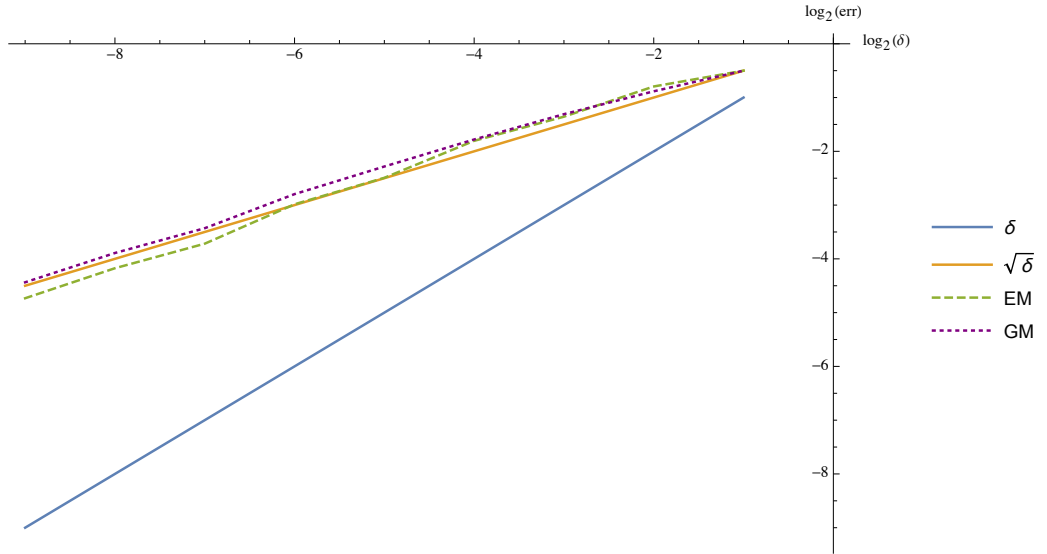
Figure 3 – Estimated L^2-error of GM and EM.

However, GM has two shortcomings. First, it needs the geometrical structure of the set of discontinuities of the drift as an input. However, if for example the discontinuity stems from a discontinuous control policy in a stochastic optimal control problem, then this geometric structure for the optimal control is not explicitly known. Finding the discontinuity of a function numerically is a problem of high complexity on its own. Second, our method requires inversion of G in each step. In 1D the inverse can be calculated explicitly, see Theorem 2.6, but in general dimension, we have to resort to numerical inversion, which makes the calculation of a single path rather costly.

However, Figure 3 tells us even more. We observe that the green dashed line, which corresponds to the convergence speed of the EM method applied to our example, also has roughly the same slope as the yellow line. This means that for our example and our range of δ, the EM method seems to converge, too. To deal with the issues raised above, it would be desirable to prove a positive strong convergence rate for the EM method. This is what we are going to study in the next section.

4 Convergence of the EM method

We seek to estimate the mean square error of the EM approximation by considering the difference between GM and EM. Here, we only sketch the idea of the proof.

Let X^δ be the EM approximation of X. Using that $X = G(Z)$, that G^{-1} is Lips-

chitz, and that $(a+b)^2 \leq 2a^2 + 2b^2$, we estimate the mean square error of the EM approximation:

$$\mathbb{E}\left(\sup_{0 \leq t \leq T} \|X_t - X_t^\delta\|^2\right) = \mathbb{E}\left(\sup_{0 \leq t \leq T} \|G^{-1}(Z_t) - G^{-1}(G(X_t^\delta))\|^2\right)$$
$$\leq 2L_{G^{-1}}^2 \mathbb{E}\left(\sup_{0 \leq t \leq T} \|Z_t - Z_t^\delta\|^2\right) + 2L_{G^{-1}}^2 \mathbb{E}\left(\sup_{0 \leq t \leq T} \|Z_t^\delta - G(X_t^\delta)\|^2\right).$$

With this we have decomposed the error into two error terms. The first term is the mean square error of the EM approximation of the solution to the transformed SDE. Since the transformed SDE has Lipschitz coefficients, the EM method converges with strong order $1/2$, i.e.

$$\mathbb{E}\left(\sup_{0 \leq t \leq T} \|Z_t - Z_t^\delta\|^2\right) \leq C\delta.$$

For estimating

$$\mathbb{E}\left(\sup_{0 \leq t \leq T} \|Z_t^\delta - G(X_t^\delta)\|^2\right)$$

the crucial estimate is the one of the drift. For this the main tasks are:

- estimating the probability of the event Ω_ε that during one step the distance between the interpolation of the EM method and the previous EM step becomes greater than some given $\varepsilon > 0$. Lemma 3.3 in [10] states that

$$\mathbb{P}(\Omega_\varepsilon) \leq C\exp\left(-\frac{\varepsilon}{\|\sigma\|_\infty \delta^{1/2}}\right);$$

- estimating the occupation time of the Euler-Maruyama approximation of X close to the hypersurface Θ by constructing a 1D process Y that has the same occupation time close to 0 as X^δ has close to Θ. The process Y is essentially a signed distance of X^δ from Θ. Again we make extensive use of the positive reach property of Θ, which guarantees regularity of a distance function. Theorem 2.7 in [10] says that

$$\int_0^T \mathbb{P}\left(\{X_s^\delta \in \Theta^\varepsilon\}\right) ds \leq C\varepsilon.$$

We are free to choose ε as a function of the step-size δ, and if we do so in an optimal way, we obtain the following convergence rate.

Theorem 4.1 (Leobacher and Szölgyenyi [10]). *Let the assumptions of Theorem 3.7 hold, and let μ, σ be bounded.*

Then the Euler-Maruyama method converges with strong order $1/4 - \zeta$ for arbitrarily small $\zeta > 0$ to the solution of SDE (1).
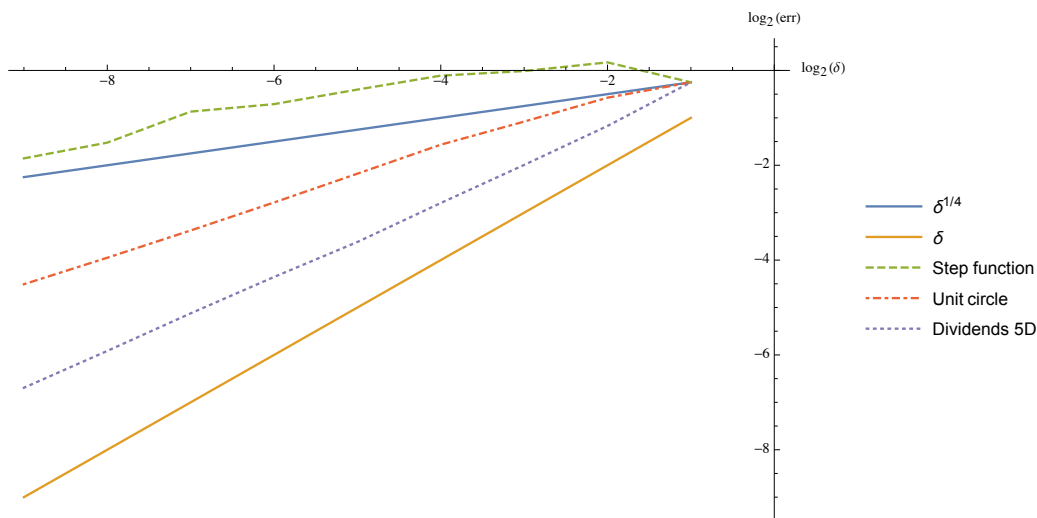
13

Figure 4 – Estimated L^2-error of the EM approximation.

Now the question arises why one would apply EM instead of GM, since GM has a much higher convergence speed. However, as already mentioned at the end of Section 3, the computation of a single path with GM can be so slow, that obtaining comparable errors with GM can take more time for practical purposes. We refer to [10] for more details.

Figure 4 shows the estimated L^2-error of the EM approximation for three examples: one where the drift is a certain step-function, a five-dimensional example from insurance mathematics (Dividends 5D), and the example from above where the drift is discontinuous on the unit circle. We see that for the step-function example the convergence seems to be approximately as fast as $\delta^{1/4}$ for larger δ, but for smaller δ the slope of the dashed green line seems to become steeper. For the other two examples the EM method clearly converges at a higher rate for this example. This supports the claim from above that in many examples the EM method is the preferred choice.

References

[1] R. L. Foote. Regularity of the Distance Function. *Proceedings of the American Mathematical Society*, 92(1):153–155, 1984.

[2] R. Frey, A. Gabih, and R. Wunderlich. Portfolio Optimization under Partial Information with Expert Opinions. *International Journal of Theoretical and Applied Finance*, 15(1), 2012.

[3] M. Hairer, M. Hutzenthaler, and A. Jentzen. Loss of Regularity for Kolmogorov Equations. *The Annals of Probability*, 43(2):468–527, 2015.

[4] K. Itô. On Stochastic Differential Equations. *Memoirs of the American Mathematical Society*, 4:1–57, 1951.

[5] A. Jentzen, T. Müller-Gronbach, and L. Yaroslavtseva. On Stochastic Differential Equations with Arbitrary Slow Convergence Rates for Strong Approximation. *Communications in Mathematical Sciences*, 14(6):1477–1500, 2016.

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

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

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

[9] G. Leobacher and M. Szölgyenyi. A Strong Order 1/2 Method for Multidimensional SDEs with Discontinuous Drift. *The Annals of Applied Probability*, 2017. Forthcoming, *arXiv:1512.02807*.

[10] G. Leobacher and M. Szölgyenyi. Convergence of the Euler-Maruyama method for multidimensional SDEs with discontinuous drift and degenerate diffusion coefficient. 2017. Submitted, *arXiv:1610.07047*.

[11] G. Leobacher, M. Szölgyenyi, and S. Thonhauser. Bayesian Dividend Optimization and Finite Time Ruin Probabilities. *Stochastic Models*, 30(2):216–249, 2014.

[12] G. Maruyama. Continuous Markov Processes and Stochastic Equations. *Rendiconti del Circolo Matematico di Palermo*, 4(1):48–90, 1955.

[13] T. Müller-Gronbach and L. Yaroslavtseva. On Hard Quadrature Problems for Marginal Distributions of SDEs with Bounded Smooth Coefficients. 2016. *arXiv: 1603.08686*.

[14] H. L. Ngo and D. Taguchi. On the Euler-Maruyama Approximation for One-Dimensional Stochastic Differential Equations with Irregular Coefficients. 2016. *arXiv:1509.06532*.

[15] H. L. Ngo and D. Taguchi. Strong Convergence for the Euler-Maruyama Approximation of Stochastic Differential Equations with Discontinuous Coefficients. 2016. *arXiv:1604.01174*.

[16] 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.

[17] G. Peskir. A Change-of-Variable Formula with Local Time on Surfaces. *Séminaire de probabilités XL*, 1899:69–96, 2007.

[18] U. Rieder and N. Bäuerle. Portfolio Optimization with Unobservable Markov-Modulated Drift Processes. *Journal of Applied Probability*, 42:362–378, 2005.

[19] M. Ruzhansky and M. Sugimoto. On Global Inversion of Homogeneous Maps. *Bulletin of Mathematical Sciences*, 5(1):13–18, 2015.

[20] J. Sass and U. Haussmann. Optimizing the Terminal Wealth under Partial Information: the Drift Process as a Continuous Time Markov Chain. *Finance and Stochastics*, 8:553–577, 2004.

[21] 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.

[22] A. A. Shardin and R. Wunderlich. Partially Observable Stochastic Optimal Control Problems for an Energy Storage. *Stochastics*, 89(1):280–310, 2017.

[23] M. Szölgyenyi. Dividend Maximization in a Hidden Markov Switching Model. *Statistics & Risk Modeling*, 32(3-4):143–158, 2016.

[24] L. Yaroslavtseva. On Non-Polynomial Lower Error Bounds for Adaptive Strong Approximation of SDEs. 2016. *arXiv:1611.08073*.

About the authors. Gunther Leobacher studied mathematics at the University of Salzburg, and finished his PhD in 2001 under the supervision of G. Larcher at Johannes Kepler University Linz. In 2002–2003 he was a Postdoc with L.G.C. Rogers at the University of Cambridge. In 2012 he became an associate professor at JKU Linz. In February 2017 he was appointed full professor of stochastics at the University of Graz.

Michaela Szölgyenyi studied mathematics at the Johannes Kepler University Linz, and finished her PhD in 2015 under the supervision of G. Leobacher. Then she became a Postdoc with R. Frey at Vienna University of Economics and Business. From August 2017 she works at ETH Zürich in the research group of A. Jentzen.

Authors' addresses:

G. Leobacher. Institute for Mathematics and Scientific Computing, University of Graz. Heinrichstraße 36, A 8010 Graz. email gunther.leobacher@uni-graz.at

M. Szölgyenyi. Seminar for Applied Mathematics, ETH Zürich. Rämistraße 101, CH 8092 Zürich. email michaela.szoelgyenyi@sam.math.ethz.ch