



HAL
open science

Numerical approximation of irregular SDEs via Skorokhod embeddings

Stefan Ankirchner, Thomas Kruse, Mikhail Urusov

► **To cite this version:**

Stefan Ankirchner, Thomas Kruse, Mikhail Urusov. Numerical approximation of irregular SDEs via Skorokhod embeddings. 2015. hal-01205690

HAL Id: hal-01205690

<https://hal.science/hal-01205690v1>

Preprint submitted on 26 Sep 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Numerical approximation of irregular SDEs via Skorokhod embeddings

Stefan Ankirchner Thomas Kruse Mikhail Urusov *

September 25, 2015

We provide a new algorithm for approximating the law of a one-dimensional diffusion M solving a stochastic differential equation with possibly irregular coefficients. The algorithm is based on the construction of Markov chains whose laws can be embedded into the diffusion M with a sequence of stopping times. The algorithm does not require *any* regularity or growth assumption; in particular it applies to SDEs with coefficients that are nowhere continuous and that grow superlinearly. We show that if the diffusion coefficient is bounded and bounded away from 0, then our algorithm has a weak convergence rate of order $1/4$. Finally, we illustrate the algorithm's performance with several examples.

2010 MSC : 60F17, 60J60, 65C30.

Keywords : stochastic differential equations, irregular diffusion coefficient, superlinearly growing diffusion coefficient, weak convergence of processes, Skorokhod embedding problem.

*Stefan Ankirchner, Institute for Mathematics, University of Jena, Ernst-Abbe-Platz 2, 07745 Jena, Germany. *Email*: s.ankirchner@uni-jena.de, *Phone*: +49 (0)3641 946275; Thomas Kruse, Faculty of Mathematics, University of Duisburg-Essen, Thea-Leymann-Str. 9, 45127 Essen, Germany. *Email*: thomas.kruse@uni-due.de, *Phone*: +49 (0)201 183 3911; Mikhail Urusov, Faculty of Mathematics, University of Duisburg-Essen, Thea-Leymann-Str. 9, 45127 Essen, Germany; and Steklov Mathematical Institute, Russian Academy of Sciences, ul. Gubkina 8, 119991 Moscow, Russia. *Email*: mikhail.urusov@uni-due.de, *Phone*: +49 (0)201 183 7428. Stefan Ankirchner and Thomas Kruse acknowledge the financial support from the French Banking Federation through the *Chaire Markets in Transition*. The work of Mikhail Urusov is supported by the Russian Science Foundation under grant 14-21-00162 in Steklov Mathematical Institute of Russian Academy of Sciences. We are grateful to seminar participants in Beijing, Evry, Essen, Le Mans, London, Lyon, Milan, Paris for helpful comments.

Introduction

Consider a one-dimensional stochastic differential equation (SDE) of the form

$$dM_t = b(M_t) dt + \eta(M_t) dW_t, \quad M_0 = m. \quad (1)$$

Suppose that b and η are Borel functions such that (1) possesses a weak solution (M, W) that is unique in law (cf. Section 5.5.C in [17]). In general the law of the solution process M is not known explicitly, and therefore one has to fall back on numerical methods for its approximation. The most used approximation method is the Euler scheme. If the coefficients b and η are Lipschitz continuous, then the Euler approximations are known to converge to the solution (see e.g. [18]). In many applications, however, one has to deal with non-Lipschitz SDEs. For example in economics frequently risk factors are modeled as non-Lipschitz SDEs (e.g. the CIR process, or the price process in CEV models and quadratic normal volatility models (see [3])). Diffusion models in mathematical genetics make use of non-Lipschitz SDE, e.g. the Wright-Fisher-diffusion (see [5]). Finally, many physical phenomena, e.g. the movement of a particle between two different media, can be modeled with SDEs possessing a discontinuous diffusion coefficient (see introduction in [22]).

If the coefficients of the SDE are not globally Lipschitz continuous, then usually the Euler scheme is numerically unstable. Indeed, one can even prove that if the coefficients grow superlinearly, then the Euler approximations do not converge in a weak L^1 sense (see [13]). Therefore, for the simulation of non-Lipschitz SDEs alternative approximation methods are needed.

In this article we provide a new approximation method that does not require any regularity or growth assumption on the coefficients b and η . In particular, the method applies to SDEs with coefficients that are nowhere continuous and that grow superlinearly.

The *basic idea* of our approximation method is to construct Markov chains that can be embedded in distribution into M with a sequence of stopping times. For the construction one first applies a standard change of variables (see e.g. Section 5.5.B in [17]) in order to transform the SDE (1) into an SDE without drift. Therefore, without loss of generality we can assume that $b \equiv 0$. Now let (X_k) be an i.i.d. sequence of centered random variables and (Y_k) a Markov chain with dynamics

$$Y_{k+1} = Y_k + a(Y_k)X_{k+1}, \quad k \in \mathbb{Z}_+. \quad (2)$$

We construct the Borel function $a: \mathbb{R} \rightarrow \mathbb{R}_+$ in a way that one can find stopping times $0 = \tau_0 < \tau_1 < \dots$ such that the conditional expectation between two consecutive stopping times is equal to a constant $h \in (0, \infty)$, and such that the discrete-time process $(M_{\tau_k})_{k \geq 0}$ has the same distribution as $(Y_k)_{k \geq 0}$. One can extend (Y_k) to a continuous-time process by setting $Y_t = Y_{[t]} + (t - [t])(Y_{[t]+1} - Y_{[t]})$ for all $t \in \mathbb{R}_+$. A functional limit theorem implies that the distribution of $(Y_{t/h})_{t \in [0, T]}$ converges to the distribution of $(M_t)_{t \in [0, T]}$ as $h \downarrow 0$ (see [2]). One can thus use (Y_k) for approximating the law of M . In other words, by simulating (Y_k) one can estimate the distribution of M . Notice that a realization of (Y_k) can be interpreted as an *exact* simulation of M along a sequence of stopping times.

Our idea works, because there is a simple integral formula for the minimal expected time needed for embedding a distribution into M (see [1]). The integral formula allows to determine the scale factor a in (2) such that the transition probabilities of (Y_k) can be embedded into M with stopping times having expectation h .

We now give a brief overview on some alternative methods for approximating non-Lipschitz SDEs and compare them with our proposed scheme.

A related numerical scheme for approximating diffusions with discontinuous coefficients is provided in [6]. The authors construct a *continuous-time* process on a given finite grid by embedding, in every time step, a weighted sum of the Dirac measures of the two neighboring grid points. The average time for attaining one of the neighboring points is determined by solving a PDE. In contrast to [6], in our approach the average time lag is fixed, whereas the scale factor, and hence the state distribution, is determined endogeneously.

The literature comprises also several papers analyzing the convergence of Euler-type approximation schemes for SDEs with irregular and/or quickly growing coefficients. [7] shows almost sure convergence of the Euler scheme for SDEs with a drift satisfying a certain *monotonicity* condition (a local *one-sided Lipschitz* condition in space, uniformly in time) and with a locally Lipschitz diffusion coefficient. [26] proves weak convergence of the Euler scheme in the case where the diffusion coefficient is allowed to be discontinuous on a set of Lebesgue measure zero and has at most linear growth. There are also results on the rate of convergence in [26], but only for Hölder continuous coefficients. In [23] it is shown that, for the Euler scheme, the strong rates obtained in [26] hold also for certain discontinuous (but one-sided Lipschitz) drifts; however, the diffusion coefficient needs to be bounded and Hölder continuous. As far as modifications of the Euler scheme are concerned, the analyses of the *implicit Euler scheme* (also known as *backward Euler scheme*) and of the related *split-step backward Euler scheme* allow some irregularities in the drift, but assume the diffusion coefficient to be Lipschitz (see [8] and [10]). The *tamed Euler scheme* is an explicit scheme that dampens superlinear growth, in [14] of the drift coefficient, and in [12] also for the diffusion coefficient. The scheme is shown to converge strongly if the coefficients are locally Lipschitz continuous and satisfy a Lyapunov-type growth condition (see [12, Theorem 3.15]).

The paper [19] provides convergence rates for the weak approximation error, assuming again that only the diffusion coefficient is regular (the method regularizes the drift and uses the Euler scheme for the regularized equation). [24] studies certain explicit Euler-type schemes that exhibit strong convergence also in certain cases when drift and diffusion coefficients are growing superlinearly and compensate each other. However, in the case of a driftless SDE the assumptions in [24] imply that the diffusion coefficient is locally Lipschitz and of linear growth. [11] provides a general framework, which contains the Euler scheme (but not our one) as a special case, and proves strong convergence rates, where again the diffusion coefficient is regular and the drift satisfies a certain non-global monotonicity condition. For further literature on approximation schemes we refer to the overview [16].

The paper is organized in the following way. In Section 1 we describe in detail the new

approximation scheme. We explain why the scheme converges, and why no regularity and growth assumptions on the diffusion coefficient are needed.

In Section 2 we show some properties of the scheme that are desirable from a numerical point of view. In particular, we prove, under some nice conditions on (X_k) , that the approximating process (Y_k) satisfies a certain comparison principle. This excludes divergent oscillations of the scheme for quickly growing η . Moreover, we show that the scale factor a in (2) is Lipschitz continuous and hence grows at most linearly. This smoothing and tempered growth behavior accounts for a good numerical performance. Furthermore we show that a in (2) is characterized by an ordinary differential equation and can therefore be efficiently computed.

Section 3 deals with the convergence rate. We show that if η is bounded and bounded away from zero, then our scheme has a weak convergence rate of order $1/4$. We stress here that the convergence rate holds true without any regularity assumptions on the diffusion coefficient.

Section 4 shapes our approximation scheme into a concise simulation algorithm. This algorithm is then illustrated in Section 5 with several examples. Moreover, results from some numerical experiments are reported.

1 A scheme that is exact along stopping times

In this section we describe a method for approximating the law of a one-dimensional homogeneous stochastic differential equation (SDE). We start by specifying the class of SDEs the method can be applied to. We first restrict ourselves to SDEs without drift; in Section 5.5 we explain how one can extend the method to SDEs with drift.

Let $I = (l, r)$ with $l \in [-\infty, \infty)$ and $r \in (-\infty, \infty]$. Let $\eta : \mathbb{R} \rightarrow \mathbb{R}$ be a Borel-measurable function satisfying

$$\eta(x) \neq 0 \text{ for all } x \in I, \quad (3)$$

$$\frac{1}{\eta^2} \in L^1_{\text{loc}}(I), \quad (4)$$

$$\eta(x) = 0 \text{ for all } x \in \mathbb{R} \setminus I, \quad (5)$$

where $L^1_{\text{loc}}(I)$ denotes the set of functions that are locally integrable on I . Consider the SDE

$$dM_t = \eta(M_t)dW_t, \quad M_0 = m \in (l, r). \quad (6)$$

The assumptions (3)–(5) imply that (6) possesses a weak solution that is unique in law (see e.g. [4] or Theorem 5.5.7 in [17]). This means that there exists a pair of processes (M, W) on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$, with (\mathcal{F}_t) satisfying the usual conditions, such that W is an (\mathcal{F}_t) -Brownian motion and (M, W) satisfies the SDE (6). Let us note that M stays in l (resp. r) once it hits l (resp. r).

Our basic idea is to approximate M with a discrete-time process that coincides in law with M along a sequence of stopping times. To explain the method, let X_1, X_2, \dots be

a sequence of i.i.d. real-valued random variables. We assume that each X_i is integrable with $E(X_i) = 0$, and we denote the distribution of X_i by μ . Let us also suppose $\mu \neq \delta_0$. We define a process $(Y_k)_{k \in \mathbb{Z}_+}$ by setting

$$Y_0 = m \quad \text{and} \quad Y_{k+1} = Y_k + a(Y_k)X_{k+1}, \quad k \in \mathbb{Z}_+, \quad (7)$$

where $a: \mathbb{R} \rightarrow \mathbb{R}_+$ is a Borel function. In the following we refer to (Y_k) as to the *scaled random walk* generated by (X_k) with *scale factor* a . Moreover, we say that (Y_k) is *embeddable into* M with expected time step $h \in (0, \infty)$ if there exists an increasing sequence of stopping times (τ_k) such that $(M_{\tau_k}) \stackrel{d}{=} (Y_k)$ (i.e. both discrete-time processes have the same law) and

$$\tau_0 = 0 \quad \text{and} \quad E[\tau_{k+1} - \tau_k | \mathcal{F}_{\tau_k}] = h, \quad k \in \mathbb{Z}_+. \quad (8)$$

For the remainder of this section we suppose that for all $h \in (0, \infty)$ we can find a unique scale factor a such that the random walk (Y_k) with scale factor a is embeddable into M with expected time step h . In Section 2 we recall a sufficient condition from [2] guaranteeing that this assumption is satisfied (see Proposition 2.1). We next explain why the embeddable scaled random walks can be used for approximating the law of the diffusion M .

Let $T \in \mathbb{R}_+$ be a finite time horizon and $h_N = \frac{T}{N}$, $N \in \mathbb{N}$. Let (Y_k^N) be the scaled random walk that is embeddable into M with expected time step h_N and let (τ_k^N) be a sequence of stopping times satisfying (8) such that $(M_{\tau_k^N}) \stackrel{d}{=} (Y_k^N)$. Besides, let $(M_t^N)_{t \in [0, T]}$ and $(\widetilde{M}_t^N)_{t \in [0, T]}$ be continuous-time processes defined by $M_{kh_N}^N := Y_k^N$ and $\widetilde{M}_{kh_N}^N := M_{\tau_k^N}$ on the grid $\pi^N = \{0, h_N, 2h_N, \dots, Nh_N \equiv T\}$ and via linear interpolation between the grid points:

$$M_t^N = M_{[t]_N}^N + \frac{1}{h_N}(t - [t]_N)(M_{[t]_N + h_N}^N - M_{[t]_N}^N)$$

and similarly for \widetilde{M}^N , where $[t]_N := \sup\{[0, t] \cap \pi^N\}$. For all $k \geq 1$ let $\rho_k^N = \tau_k^N - \tau_{k-1}^N$. Equation (8) entails that the family (ρ_k^N) , $k \geq 1$, is uncorrelated. Moreover, we have $E(\rho_k^N) = h_N$. If in addition (ρ_k^N) satisfies some nice uniform integrability condition, then a certain weak law of large numbers for arrays applies, and hence $\lim_{N \rightarrow \infty} \tau_{[Nt/T]}^N = t$ in probability, for all $t \in \mathbb{R}_+$. Since M has continuous sample paths, this further indicates that $\lim_{N \rightarrow \infty} \widetilde{M}^N = M$ in probability uniformly in the space $\mathcal{C}([0, T])$ of continuous functions $[0, T] \rightarrow \mathbb{R}$; in other words, we have convergence of the distributions of M^N on the path space. The previous line of argument is made precise in [2].

Theorem 1.1 (Theorem 3.6 in [2]). *Suppose that $|\eta|$ and $\frac{1}{|\eta|}$ are locally bounded on I , μ has a compact support and that the following implications hold true:*

$$\text{if } l > -\infty, \text{ then } \mu(\{\inf \text{supp } \mu\}) > 0, \quad (9)$$

$$\text{if } r < \infty, \text{ then } \mu(\{\sup \text{supp } \mu\}) > 0. \quad (10)$$

Then $(M_t^N)_{t \in [0, T]}$ converges in distribution to $(M_t)_{t \in [0, T]}$ as $N \rightarrow \infty$.

Let us notice that Theorem 3.6 in [2] contains, in fact, a weaker but more technical assumption than that in Theorem 1.1 (compare (9)–(10) above with (19), (27) and (29) as well as Propositions 2.12 and 2.20 in [2]).

Remark 1.2. If $|\eta|$ and $\frac{1}{|\eta|}$ are globally bounded, then one can prove convergence also for μ without compact support (see Theorem 3.1 in [2]).

The preceding theorem shows that we can use M^N for approximating the law of M . More precisely, let $f: \mathcal{C}([0, T]) \rightarrow \mathbb{R}$ be a bounded functional that is continuous with respect to the sup norm. Then Theorem 1.1 means that $E[f(M^N)] \approx E[f(M)]$. Thus, by simulating the scaled random walk (Y_k^N) one can construct Monte Carlo estimators for $E[f(M)]$.

The approximation method requires to compute the scale factor guaranteeing that (Y_k^N) is embeddable into M with a given expected time step h . Recent results on the Skorokhod embedding problem (see [1] or [9]) imply that the scale factor is uniquely determined by a simple integral equation making use of the function

$$q(y, x) = \int_y^x \int_y^u \frac{2}{\eta^2(z)} dz du, \quad y \in I, x \in \mathbb{R}. \quad (11)$$

The assumptions (3)–(5) imply that for all $y \in I$ the nonnegative function $q(y, \cdot)$ is finite on I and equal to ∞ on $\mathbb{R} \setminus [l, r]$. Besides, $q(y, \cdot)$ is strictly convex on I , strictly decreasing to zero on (l, y) and strictly increasing from zero on (y, r) . Moreover, for all $y, \bar{y} \in I$ and $x \in \mathbb{R}$ we have

$$q(y, x) = q(\bar{y}, x) - q(\bar{y}, y) - q_x(\bar{y}, y)(x - y), \quad (12)$$

where q_x denotes the partial derivative of q with respect to the second argument.

Consider an integrable distribution ν with $\int x\nu(dx) = m$. If $\int q(m, x)\nu(dx) < \infty$, then there exists an integrable stopping time τ such that $M_\tau \sim \nu$ and $E[\tau] = \int q(m, x)\nu(dx)$ (see [1]). To provide an intuition why this formula for $E[\tau]$ holds true, observe that by Itô's formula the process $(q(M_0, M_t) - t)$ is a local martingale starting in 0. Therefore, if $M_\tau \sim \nu$ and the optional sampling theorem applies, we have $E[\tau] = E[q(m, M_\tau)] = \int q(m, x)\nu(dx)$.

In order to make the scaled random walk (Y_k) embeddable into M with expected time step h , we choose the scale factor a such that $\mu(y, a(y), dx)$, the distribution of $y + a(y)X_i$, satisfies $\int q(y, x)\mu(y, a(y), dx) = h$ for all $y \in I$. The latter condition is equivalent to the integral equation

$$\int q(y, y + a(y)x) \mu(dx) = h, \quad y \in I. \quad (13)$$

It is shown in [2] that if the scale factor a satisfies (13), then the associated scaled random walk (Y_k) is indeed embeddable into M . In the case where at least one of the state space boundaries l or r is finite and absorbing, Equality (13) is not necessary for (Y_k) to be embeddable into M with expected time step h . Indeed, let τ be a stopping

time with $E[\tau] < h$ such that M_τ has positive mass in l or r . On the event where M_τ is at the boundary one can enlarge τ by a constant amount of time such that τ has expectation h , without changing the distribution of M_τ . Therefore, in case of attainable boundaries, the appropriate condition for the scale factor is given by

$$a(y) = \sup\{a \geq 0 : \int q(y, y + ax) \mu(dx) \leq h\}, \quad y \in I. \quad (14)$$

We can summarize the approximation scheme as follows. Choose $T \in \mathbb{R}_+$ and $N \in \mathbb{N}$, and set $h = T/N$. Choose a centered probability distribution $\mu \neq \delta_0$ and a scale factor a such that (13) resp. (14) is satisfied. Then compute the random walk (Y_k) generated by (X_k) with scale factor a .

2 Properties of the approximation scheme

In this section we collect properties of the approximating scheme described in Section 1. We first state the precise assumptions under which the subsequent statements hold true.

We use the setting and notation of Section 1. In particular, we are given a state space $I = (l, r)$, a Borel function $q: \mathbb{R} \rightarrow \mathbb{R}$ satisfying (3)–(5) and a centered probability measure $\mu \neq \delta_0$ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, which is the distribution of the random variables X_i . As usual, we denote by \bar{I} the closure of I in \mathbb{R} .

Throughout the whole paper we assume that the following condition on μ is satisfied:

Condition (A)

- (1) If $l = -\infty$, then there exists $y \in I$ such that the integral over the negative real line $\int_{\mathbb{R}_-} q(y, y + ax) \mu(dx) < \infty$ for all $a > 0$.
- (2) If $l > -\infty$, then $\inf \text{supp } \mu > -\infty$ and $\mu(\{\inf \text{supp } \mu\}) > 0$.
- (3) If $r = \infty$, then there exists $y \in I$ such that the integral over the positive real line $\int_{\mathbb{R}_+} q(y, y + ax) \mu(dx) < \infty$ for all $a > 0$.
- (4) If $r < \infty$, then $\sup \text{supp } \mu < \infty$ and $\mu(\{\sup \text{supp } \mu\}) > 0$.

Let us remark that Condition (A) is satisfied whenever μ has a compact support, $\mu(\{\inf \text{supp } \mu\}) > 0$ and $\mu(\{\sup \text{supp } \mu\}) > 0$. A simple example, which will be often considered below, is $\mu = \frac{1}{2}(\delta_{-1} + \delta_1)$. Let us also notice that Condition (A) is satisfied under the assumptions of Theorem 1.1 above (in fact, the role of (9)–(10) is to ensure Condition (A) only). We stress that we require Condition (A) in all statements below, but do not repeat it explicitly any longer.

We define a function $G: I \times \mathbb{R}_+ \rightarrow [0, \infty]$ by

$$G(y, a) = \int q(y, y + ax) \mu(dx), \quad (15)$$

where q is defined as in (11). Monotone convergence entails that, for every $y \in I$, the mapping $a \mapsto G(y, a)$ is increasing, left-continuous, vanishing at zero and, by Condition (A), finite in a sufficiently small right neighborhood of zero. Moreover, by the

dominated convergence theorem, this mapping is continuous possibly except for at the point $a_\infty(y) = \inf\{a \in \mathbb{R}_+ : G(y, a) = \infty\}$ ($\inf \emptyset := \infty$). Also, this mapping is *strictly* increasing on $[0, a_\infty(y)]$ for every $y \in I$. We next define (cf. (14))

$$a(y) = \sup\{a \geq 0 : G(y, a) \leq h\}, \quad y \in I. \quad (16)$$

Then the scaled random walk (Y_k) with scale factor a is embeddable into M with expected time step $h \in (0, \infty)$. More precisely, we have the following result.

Proposition 2.1. *Let $h \in (0, \infty)$ and define the scale factor $a: \bar{I} \rightarrow \mathbb{R}$ as in (16) for $y \in I$, and set $a(\cdot) = 0$ on $\bar{I} \setminus I$. Then the scaled random walk (Y_k) with scale factor a is \bar{I} -valued and embeddable into M with expected time step h .*

Proof. The statements follow from Theorems 2.3, 2.7, 2.18 and Propositions 2.12, 2.20 of [2]. \square

Remark 2.2. (i) In the case when the state space I has a finite inaccessible¹ boundary point for M , the sufficient conditions in Section 2 of [2] are even weaker than Condition (A); see formulas (19), (27), (29) in [2]. Under those weaker conditions, Proposition 2.1 can be adjusted as follows: there is $h_0 > 0$ such that, for all $h \in (0, h_0)$, the scaled random walk (Y_k) with scale factor a given by (16) is embeddable into M with expected time step h .

(ii) If with probability one M does not attain the boundary points l and r in finite time, then the scale factor $a(y)$ satisfies not only (16) but also $G(y, a(y)) = h$ (cf. (13)).

In what follows, let $T \in \mathbb{R}_+$, $N \in \mathbb{N}$ and $h = T/N$. Let a be the scale factor satisfying (16) and (Y_k) the random walk generated by (X_k) with scale factor a . If we want to stress the dependence on N , we write a_N and (Y_k^N) .

2.1 (Y_k) is a generalized martingale

Consistently with Section VII.1 in [25] we say that a discrete-time process $(Z_k)_{k \in \mathbb{Z}_+}$ adapted to a filtration $(\mathcal{G}_k)_{k \in \mathbb{Z}_+}$ is a *generalized martingale* if $E(Z_{k+1} | \mathcal{G}_k) = Z_k$ a.s. for all $k \in \mathbb{Z}_+$ (in particular, all $E(Z_{k+1} | \mathcal{G}_k)$ need to be well defined). In contrast to a martingale, a generalized martingale does not need to be integrable.

Lemma 2.3. *The process $(Y_k)_{k \in \mathbb{Z}_+}$ is a generalized martingale with respect to its natural filtration. In particular, it is a martingale whenever $E(Y_k)^- < \infty$ for all $k \in \mathbb{Z}_+$ (or $E(Y_k)^+ < \infty$ for all $k \in \mathbb{Z}_+$).*

Proof. It follows from the very definition (7) that (Y_k) is a generalized martingale. If $E(Y_k)^- < \infty$ for all $k \in \mathbb{Z}_+$, then all $EY_k > -\infty$ are well defined, and all expectations are equal to each other by the tower property of conditional expectations. The only way for $(Y_k)_{k \in \mathbb{Z}_+}$ to violate the martingale property is to have $EY_k = \infty$ for all $k \in \mathbb{Z}_+$. This is impossible, because $Y_0 = m$. \square

¹Recall that the boundary point l (resp. r) is inaccessible for M if and only if $q(y, l+) = \infty$ (resp. $q(y, r-) = \infty$) and that this condition does *not* depend on $y \in I$. In particular, infinite boundary points are always inaccessible for M driven by (6).

2.2 Asymptotic equivalence with the Euler scheme

In this subsection we analyze the asymptotic behavior of $a_N(y)$ as N tends to ∞ . We show in the next theorem that $a_N(y)$ is asymptotically dominated by $\frac{\text{const}}{\sqrt{N}}\eta^*(y)$, where η^* denotes the upper semicontinuous envelope of $|\eta|$:

$$\eta^*(y) = \limsup_{z \rightarrow y} |\eta(z)|, \quad y \in I. \quad (17)$$

We first prove that a_N converge pointwise to zero:

Lemma 2.4. *For any $y \in I$ we have $\lim_{N \rightarrow \infty} a_N(y) = 0$.*

Proof. Fix $y \in I$. Since $\int q(y, y + a_N(y)x) \mu(dx) \leq \frac{T}{N}$, Fatou's lemma yields $\int q(y, y + \liminf_{N \rightarrow \infty} a_N(y)x) \mu(dx) = 0$. Since $\mu \neq \delta_0$ and, clearly, the sequence $\{a_N(y)\}_{N \in \mathbb{N}}$ is decreasing, we obtain the result. \square

Theorem 2.5. (i) *For any $y \in I$, we have*

$$\limsup_{N \rightarrow \infty} \sqrt{\frac{N}{T}} a_N(y) \leq \frac{\eta^*(y)}{\sqrt{\int x^2 \mu(dx)}}, \quad (18)$$

where η^* is defined in (17) and $\frac{\infty}{\infty}$ is understood as ∞ .

(ii) *If $|\eta(y)| \geq \varepsilon > 0$ for all $y \in I$ and η is continuous at some point $y_0 \in I$, then*

$$\lim_{N \rightarrow \infty} \sqrt{\frac{N}{T}} a_N(y_0) = \frac{|\eta(y_0)|}{\sqrt{\int x^2 \mu(dx)}}. \quad (19)$$

Proof. (i) Fix $y \in I$. Using Condition (A) and Lemma 2.4 one can show that there exists $N_1 \in \mathbb{N}$, which in general depends on y , such that for all $N \geq N_1$ we have

$$\int q(y, y + a_N(y)x) \mu(dx) = \frac{T}{N}.$$

By performing two changes of variables we obtain

$$T = N a_N^2(y) \int \int_0^x \int_0^u \frac{2}{\eta^2(y + a_N(y)r)} dr du \mu(dx). \quad (20)$$

Then Fatou's lemma together with Lemma 2.4 yields

$$\limsup_{N \rightarrow \infty} \frac{N}{T} a_N^2(y) \leq \frac{1}{\int \int_0^x \int_0^u \liminf_{N \rightarrow \infty} \frac{2}{\eta^2(y + a_N(y)r)} dr du \mu(dx)} \leq \frac{(\eta^*(y))^2}{\int x^2 \mu(dx)}. \quad (21)$$

(ii) If $\int x^2 \mu(dx) = \infty$, then the right-hand side of (19) is zero and, therefore, (19) follows from (18). Assume that $\int x^2 \mu(dx) < \infty$. Then using Lebesgue's dominated convergence theorem instead of Fatou's lemma, we get (19) from (20) in the same way as (21) was obtained. \square

Remark 2.6. Observe that if η is *not* locally bounded at some point $y \in I$ (i.e. $\eta^*(y) = \infty$), the term $\sqrt{\frac{N}{T}}a_N(y)$ might tend to ∞ as $N \rightarrow \infty$ (cf. Section 5.4).

The following corollary will be useful in the sequel.

Corollary 2.7. *Assume that $|\eta|$ is locally bounded on I . Then, for any compact subinterval $J \subset I$, we have $\lim_{N \rightarrow \infty} \sup_{y \in J} a_N(y) = 0$.*

Proof. The statement follows from (18). □

2.3 Comparison principle

Suppose that the i.i.d. sequence (X_k) satisfies $P(X_k = \pm 1) = \frac{1}{2}$. Then the scale factor is Lipschitz continuous with Lipschitz constant one. Moreover, the scaled random walk generated by (X_k) satisfies the following comparison principle: if $Y_k \geq \hat{Y}_k$, then $Y_{k+1} = Y_k + a_N(Y_k)X_{k+1}$ dominates the random variable $\hat{Y}_{k+1} = \hat{Y}_k + a_N(\hat{Y}_k)X_{k+1}$. Both properties follow from the next result.

Theorem 2.8 (Comparison principle and Lip(1)). *Let $\mu = \frac{1}{2}(\delta_{-1} + \delta_1)$. Then for every $N \in \mathbb{N}$ the mappings $y \mapsto y + a_N(y)z$ are nondecreasing on I for $z \in \{-1, 1\}$. Moreover, $y \mapsto a_N(y)$ is Lipschitz continuous on I with Lipschitz constant 1.*

Proof. Fix $N \in \mathbb{N}$. Since

$$G(y, a) = \frac{q(y, y + a) + q(y, y - a)}{2},$$

we get that the function $(y, a) \mapsto G(y, a)$ is C^1 in both arguments with $\partial_a G(y, a) > 0$ in the interior of the set $\{(y, a) \in I \times \mathbb{R}_+ : G(y, a) < \infty\}$. ($\partial_a G$ denotes the derivative of G with respect to the second argument.)

We need to consider four cases. We start with Case 1: $l = -\infty$, $r = +\infty$. In this case Proposition 2.1 and Remark 2.2 (ii) ensure that the scale factor a_N satisfies $G(y, a_N(y)) = \frac{T}{N}$ for all $y \in I$. By the implicit function theorem, the mapping $y \mapsto a_N(y)$ is C^1 in I . Differentiating the equation

$$\frac{T}{N} = \frac{q(y, y + a_N(y)) + q(y, y - a_N(y))}{2}$$

with respect to y yields

$$(1 + a'_N(y)) \int_y^{y+a_N(y)} \frac{2}{\eta^2(z)} dz = (1 - a'_N(y)) \int_{y-a_N(y)}^y \frac{2}{\eta^2(z)} dz.$$

Since $a_N(y) > 0$ for all $y \in I$, it follows that $1 + a'_N(y)$ and $1 - a'_N(y)$ have the same sign. But this implies $1 + a'_N(y) \geq 0$ and $1 - a'_N(y) \geq 0$, which proves both statements of the theorem for Case 1.

Case 2: $l > -\infty$, $r = \infty$. First consider the subcase $q(m, l+) = \infty$, i.e. the boundary point l is inaccessible. Then, again by Remark 2.2 (ii), the scale factor a_N satisfies

$G(y, a_N(y)) = \frac{T}{N}$ for all $y \in I$ and $N \in \mathbb{N}$. The claim now follows from the same reasoning as for the first case. If $q(m, l+) < \infty$, we set $\bar{a}(y) = y - l$ and $h(y) = G(y, \bar{a}(y))$ for $y \in I$. A calculation reveals that $h'(y) = \int_y^{2y-l} \frac{2}{\eta^2(z)} dz > 0$ for all $y \in I$. In particular, the function h is strictly increasing with $h(l+) = 0$. Then there exists $\bar{y} \in (l, \infty]$ such that $a_N(y) = y - l$ for $y \leq \bar{y}$ and a_N satisfies $G(y, a_N(y)) = \frac{T}{N}$ for $y > \bar{y}$. Clearly the mappings $y \mapsto y + a_N(y)z$, $z \in \{-1, 1\}$, are nondecreasing on (l, \bar{y}) . On (\bar{y}, ∞) monotonicity of $y \mapsto y + a_N(y)z$ follows as for Case 1. Monotonicity of $y \mapsto y + a_N(y)$ for $y > \bar{y}$ implies that the limit $a_N(\bar{y}+) = \lim_{y \searrow \bar{y}} a_N(y)$ exists. From the continuity of G on $\{(y, a) \in I \times \mathbb{R}_+ : a < y - l\}$ we deduce that $T/N = \lim_{y \searrow \bar{y}} G(y, a_N(y)) = G(\bar{y}, a_N(\bar{y}+))$. Consequently, we have $a_N(\bar{y}+) = \bar{y} - l$ and thus a_N is continuous at \bar{y} , which establishes the claims.

Case 3: $l = -\infty$, $r < \infty$: Case 3 can be reduced to Case 2 by considering $-M$.

Case 4: $l > -\infty$, $r < \infty$. In this case set $\bar{a}(y) = (y - l) \wedge (r - y)$, $h(y) = G(y, \bar{a}(y))$ for $y \in I$ and $\bar{y} = \frac{l+r}{2}$. The considerations of Case 2 and Case 3 imply that $y \mapsto y + a_N(y)z$, $z \in \{-1, 1\}$, is nondecreasing on (l, \bar{y}) and (\bar{y}, r) . Moreover, h is strictly increasing on (l, \bar{y}) and strictly decreasing on (\bar{y}, r) . Observe that h is continuous at \bar{y} . Indeed, at \bar{y} we have $h(\bar{y}) = \frac{1}{2}(q(\bar{y}, r) + q(\bar{y}, l))$. For $y < \bar{y}$ we have by (12)

$$h(y) = \frac{q(y, 2y - l) + q(y, l)}{2} = \frac{q(\bar{y}, l) + q(\bar{y}, 2y - l)}{2} - q(\bar{y}, y).$$

This implies left-continuity: $\lim_{y \nearrow \bar{y}} h(y) = h(\bar{y})$. Right-continuity is verified similarly. If $h(\bar{y}) \leq T/N$, then $a_N(y) = \bar{a}(y)$ for all $y \in I$ and the claims follow directly. If $h(\bar{y}) > T/N$, then it follows that $G(y, a_N(y)) = 1/N$ for all y in a neighborhood of \bar{y} . Then the arguments from Case 1 yield that $y \mapsto y + a_N(y)z$, $z \in \{-1, 1\}$, is also nondecreasing around \bar{y} . \square

From the preceding theorem we can directly deduce that the scale factors grow at most linearly, uniformly in N .

Corollary 2.9 (Linear growth uniformly in N). *Let $\mu = \frac{1}{2}(\delta_{-1} + \delta_1)$. Then there is a finite constant C such that, for all $N \in \mathbb{N}$, we have*

$$0 < a_N(y) \leq C + |y|, \quad y \in I, \tag{22}$$

that is, the mapping $y \mapsto a_N(y)$ has linear growth, while the bound in (22) is uniform in N .

Proof. The statement follows from the Lip(1) property. We can take $C = |y_0| + \sup_N |a_N(y_0)|$ for any fixed $y_0 \in I$, where the sup is finite due to Lemma 2.4. \square

Finally we remark that the scale factor vanishes at finite state boundaries. Indeed, we have the following more general result that applies also for measures μ other than $\frac{1}{2}(\delta_1 + \delta_{-1})$ whenever at least one of the boundaries is finite.

Proposition 2.10 (Vanishing scale factors at finite boundaries). *Suppose that $r < \infty$. Then for all $N \in \mathbb{N}$*

$$0 < a_N(y) \leq \frac{r - y}{\sup \text{supp } \mu}, \quad y \in I.$$

Similarly, if $l > -\infty$, we have $0 < a_N(y) < \frac{l - y}{\inf \text{supp } \mu}$.

Proof. Consider the case $r < \infty$. The statement follows from the fact that $G(y, a) = \infty$ whenever $a > \frac{r - y}{\sup \text{supp } \mu}$. \square

Notice that the bounds in Proposition 2.10 are uniform in N .

2.4 Convergence of expectations

In this section we show that the expectation $E[f(M_T)]$ can be approximated by $E[f(Y_N^N)]$ not only for bounded continuous functions f , but also for functions that are unbounded at the boundary of the state space.

Proposition 2.11. *Suppose that $|\eta|$ and $\frac{1}{|\eta|}$ are locally bounded on $I = (l, r)$ and that μ has a compact support. Assume that l and r are inaccessible. Let $f: I \rightarrow \mathbb{R}$ be continuous satisfying*

$$\lim_{y \searrow l} \frac{|f(y)|}{q(m, y)} = 0 \quad \text{and} \quad \lim_{y \nearrow r} \frac{|f(y)|}{q(m, y)} = 0.$$

Then $Ef(Y_N^N) \rightarrow Ef(M_T)$ as $N \rightarrow \infty$.

Proof. We first show that the family $(f(Y_N^N))_{N \in \mathbb{N}}$ is uniformly integrable. To this end we introduce the localizing sequence of stopping times

$$\sigma_k = k \wedge \inf\{t \geq 0 : \int_0^t |q_x(m, M_s)|^2 \eta(M_s)^2 ds \geq k\}.$$

Using Itô's formula we obtain $E[q(m, M_{\sigma_k \wedge \tau_N^N})] = E[\sigma_k \wedge \tau_N^N]$. Then Fatou's lemma and the monotone convergence theorem imply

$$E[q(m, Y_N^N)] = E[q(m, M_{\tau_N^N})] \leq \lim_{k \rightarrow \infty} E[q(m, M_{\sigma_k \wedge \tau_N^N})] = \lim_{k \rightarrow \infty} E[\sigma_k \wedge \tau_N^N] = E[\tau_N^N] = T.$$

Fix $\varepsilon > 0$. By assumption there exist $\bar{l}, \bar{r} \in I$ with $\bar{l} < \bar{r}$ such that

$$|f(y)| \leq \frac{\varepsilon}{T} q(m, y) \quad \text{for all } y \in I \setminus [\bar{l}, \bar{r}].$$

Let $K = \sup_{y \in [\bar{l}, \bar{r}]} |f(y)|$, which is finite by continuity of f . Then the following inequality holds uniformly for all $N \in \mathbb{N}$

$$E \left[|f(Y_N^N)| \mathbf{1}_{\{|f(Y_N^N)| > K\}} \right] \leq E \left[|f(Y_N^N)| \mathbf{1}_{\{Y_N^N \in I \setminus [\bar{l}, \bar{r}]\}} \right] \leq \frac{\varepsilon}{T} E[q(m, Y_N^N)] \leq \varepsilon.$$

This implies uniform integrability of $(f(Y_N^N))_{N \in \mathbb{N}}$. By Theorem 1.1 the sequence (Y_N^N) converges weakly to M_T and therefore $(f(Y_N^N))$ converges weakly to $f(M_T)$. It follows that $Ef(Y_N^N) \rightarrow Ef(M_T)$ as $N \rightarrow \infty$, which completes the proof. \square

Remark 2.12. If the left boundary l is accessible, then we can replace the condition $\lim_{y \searrow l} \frac{|f(y)|}{q(m,y)} = 0$ by the condition that f is continuous on $[l, r)$ and the conclusion $Ef(Y_N^N) \rightarrow Ef(M_T)$ as $N \rightarrow \infty$ still holds true. Similarly, $\lim_{y \nearrow r} \frac{|f(y)|}{q(m,y)} = 0$ can be replaced by continuity of f on $(l, r]$ in the case where r is accessible.

We next show that our scheme provides appropriate approximations of $Ef(M_T)$ for all functions f with sublinear growth — regardless of the growth behaviour of η . We say that a function $f: \bar{I} \rightarrow \mathbb{R}$ has *sublinear growth* if $\lim_{y \rightarrow -\infty} \frac{|f(y)|}{|y|} = 0$ whenever $l = -\infty$ and $\lim_{y \rightarrow \infty} \frac{|f(y)|}{|y|} = 0$ whenever $r = \infty$.

Corollary 2.13. *Suppose that $|\eta|$ and $\frac{1}{|\eta|}$ are locally bounded on $I = (l, r)$ and that μ has compact support. Let $f: \bar{I} \rightarrow \mathbb{R}$ be continuous having sublinear growth. Then $Ef(Y_N^N) \rightarrow Ef(M_T)$ as $N \rightarrow \infty$.*

Proof. If $-\infty < l, r < \infty$ the claim follows directly from Theorem 1.1 since f is bounded in this case. If $l = -\infty$, convexity of q in the second argument implies $\limsup_{y \rightarrow -\infty} \frac{|y|}{q(m,y)} < \infty$. Therefore we have $\lim_{y \rightarrow -\infty} \frac{|f(y)|}{q(m,y)} = 0$. Similarly, we have $\lim_{y \rightarrow \infty} \frac{|f(y)|}{q(m,y)} = 0$ if $r = \infty$. Then the claim follows from Proposition 2.11 and Remark 2.12. \square

Remark 2.14. The conclusion of Corollary 2.13 is sharp in the sense that without imposing further assumptions on η we cannot expect convergence of $Ef(Y_N^N)$ to $Ef(M_T)$ for a broader class of functions than the functions with sublinear growth. Indeed, if $l = -\infty$ and $\int_{-\infty}^a \frac{|x|}{\eta^2(x)} dx < \infty$ for some $a \in \mathbb{R}$ or $r = \infty$ and $\int_b^\infty \frac{|x|}{\eta^2(x)} dx < \infty$ for some $b \in \mathbb{R}$, then it follows from [20] that M is a strict local martingale. In particular, there exist initial values $M_0 = m \in I$ and terminal times T such that $E[M_T] \neq m$ (see Example 2.15 below for such a case). If $l > -\infty$ or $r < \infty$, it follows from Lemma 2.3 that for all $N \in \mathbb{N}$ the process (Y_k^N) is a discrete-time martingale and therefore satisfies $E[Y_N^N] = m$ for all $N \in \mathbb{N}$. Hence, $Ef(Y_N^N)$ cannot be used to approximate $Ef(M_T)$ for linear functions f in this case.

Example 2.15 (Johnson-Helms example of a strict local martingale). Let $I = (0, \infty)$ and $\eta(x) = x^2$ for $x \in I$. We have $q(l+) = q(r-) = \infty$. Note that the assumption of Theorem 1.1 is satisfied for any centered $\mu \neq \delta_0$ with a compact support, and, in particular, the sequence (Y_N^N) converges weakly to M_1 .

It is well known that the process (M_t) has the same distribution as $(1/|B_t|)$, where B is a three-dimensional Brownian motion starting in B_0 with $|B_0| = 1/m$, and $|\cdot|$ denotes the Euclidean norm in \mathbb{R}^3 . Moreover, one can show that $EM_1 < m$ (see e.g. [15]). On the contrary, by Lemma 2.3, the processes $(Y_k^N)_{k \in \mathbb{Z}_+}$ are martingales, hence $EY_N^N = m$ for all N .

We now present a generalization of Proposition 2.11 to certain path functionals. To this end, let us recall the notation $(M_t^N)_{t \in [0, T]}$ of Section 1 for the processes that approximate $(M_t)_{t \in [0, T]}$ in law: (M_t^N) is a continuous process that satisfies $M_{kT/N}^N = Y_k^N$, $k = 0, \dots, N$, and is defined via linear interpolation between the points of the grid $\{kT/N : k = 0, \dots, N\}$.

Proposition 2.16. *Suppose that $|\eta|$ and $\frac{1}{|\eta|}$ are locally bounded on $I = (l, r)$ and that μ has a compact support. Assume that l and r are inaccessible. Let $f: I \rightarrow \mathbb{R}$ be a Borel-measurable function satisfying*

$$\lim_{y \searrow l} \frac{|f(y)|}{q(m, y)} = 0 \quad \text{and} \quad \lim_{y \nearrow r} \frac{|f(y)|}{q(m, y)} = 0$$

and $G: \mathcal{C}([0, T]) \rightarrow \mathbb{R}$ a bounded Borel-measurable path functional. Assume that

$$P(M_T \in C_f) = 1 \quad \text{and} \quad P(M. \in C_G) = 1,$$

where C_f (resp. C_G) denotes the set of points $x \in I$ (resp. $x. \in \mathcal{C}([0, T])$) such that f is continuous at x (resp. G is continuous at $x.$). Then $E[f(M_T^N)G(M.^N)] \rightarrow E[f(M_T)G(M.)]$ as $N \rightarrow \infty$.

The proof is basically the same as that of Proposition 2.11. A remark similar to Remark 2.12 applies also in the situation of Proposition 2.16. Finally, to motivate the somewhat technical formulation of Proposition 2.16, we notice that path functionals with such a structure appear in some applications: e.g. think about a down-and-in barrier call option, where we have $f(x) = (x - K)^+$ and $G(x.) = 1_{\{\inf_{s \in [0, T]} x_s < B\}}$.

2.5 Advantages over the Euler scheme

Let us compare our approximation scheme with the (weak) explicit Euler scheme that approximates the process M with

$$\tilde{Y}_{k+1}^N = \tilde{Y}_k^N + \tilde{a}_N(\tilde{Y}_k^N)\xi_{k+1}, \quad \tilde{Y}_0^N = m, \quad (23)$$

where

$$\tilde{a}_N(y) = \eta(y)\sqrt{\frac{T}{N}} \quad (24)$$

is the *Euler scale factor*, and ξ_1, ξ_2, \dots are independent identically distributed random variables with $E\xi_k = 0$ and $E\xi_k^2 = 1$ (the classical Euler scheme uses Gaussian ξ_k). Theorem 2.5 (ii) implies that, under nice conditions on η , for a fixed $y \in I$ our scale factors $a_N(y)$ are close to the Euler scale factors $\tilde{a}_N(y)$ as N tends to infinity (here $E\xi_k^2 = 1$ entails that the denominator in (19) is equal to one). Notice, however, that in general this convergence is only local in the y -variable: For fixed $N \in \mathbb{N}$ the function $y \mapsto a_N(y)$ is qualitatively different from the Euler scale factor $y \mapsto \tilde{a}_N(y)$. In contrast to the Euler scale factor, the factor (16) involves *smoothing* (a_N is $\text{Lip}(1)$, while \tilde{a}_N inherits all irregularities of η) and has a *tempered growth behavior* (linear growth regardless of the growth of η). Moreover, a_N ensures that the associated process (Y_k^N) satisfies the comparison principle when using $\mu = \frac{1}{2}(\delta_{-1} + \delta_1)$, while it is not the case for the Euler scheme (also see Section 5.3 for a further pertinent discussion of this point).

Finally, we remark that these properties allow to use our scheme for approximating expectations of the form $E[f(M_T)]$ in some cases where the Euler scheme does not work.

Recall that our scheme converges for sublinearly growing functions (cf. Corollary 2.13). It follows from [13] that the explicit Euler scheme does not provide an appropriate approximation of $E[f(M_T)]$ for all sublinearly growing functions in the case where η is not globally Lipschitz continuous.

Proposition 2.17 (Theorem 2.1 in [13]). *Let $I = (-\infty, \infty)$. Let the process (\tilde{Y}_k^N) be defined by (23) and assume that the random variables ξ_1, ξ_2, \dots are independent and standard normally distributed. Let $C \geq 1, \beta > 1$ be constants such that $|\eta(x)| \geq \frac{|x|^\beta}{C}$ for all $|x| \geq C$. Then there exists a constant $c > 1$ and a sequence of events Ω_N with $P(\Omega_N) \geq c^{-N^c}$ such that $|\tilde{Y}_N^N| \geq 2^{2^{N-1}}$ on Ω_N for all $N \in \mathbb{N}$. In particular, this implies that $E|\tilde{Y}_N^N|^\alpha \rightarrow \infty$ as $N \rightarrow \infty$ for all $\alpha > 0$.*

This kind of divergence in finite time of the Euler scheme is due to the following unpleasant effect. Assume that $|\tilde{Y}_k^N|$ has attained a large positive value (which happens with sufficiently large probability for the true process M on any time horizon, and this probability increases as a function of the time horizon). Since the coefficient grows superlinearly and the increment $|W_{(k+1)T/N} - W_{kT/N}|$ is “of the order $\sqrt{1/N}$ ”, the value of $|\tilde{Y}_{k+1}^N|$ is likely to be huge, even for small time steps T/N . Note that if \tilde{Y}_k^N is positive (resp. negative) and $W_{(k+1)T/N} - W_{kT/N}$ is negative (resp. positive), the Euler approximation significantly jumps across zero, entailing a zigzag behavior. The properties of Y^N shown in Theorem 2.8 reveal that our scheme is more stable in this regard and prevent the scheme from these divergent oscillations. In Section 5.3 we illustrate these observations for a specific choice of η (see in particular Figure 3).

2.6 Characterizing the scale factor by an ODE

The scale factor a_N embedding the scaled random walk (Y_k^N) into M with expected time step T/N satisfies the equation $G(y, a_N(y)) = T/N$ on I whenever both boundaries l and r are inaccessible (recall Remark 2.2 (ii)). In general, one can show using Condition (A) and Corollary 2.7 that the equation $G(y, a_N(y)) = T/N$ holds at least on some subinterval J of I whenever $|\eta|$ is locally bounded on I (and, moreover, for any compact $J \subset I$ there exists $N_1 \in \mathbb{N}$ such that for all $N \geq N_1$ this equation holds on J). In this section we use the implicit function theorem to derive an ordinary differential equation characterizing a_N on J . In applications this allows for a quick numerical computation of a_N .

Proposition 2.18. *Assume that μ has compact support. Let $N \in \mathbb{N}$ and $J = (\bar{l}, \bar{r}) \subset I$ be such that $G(y, a_N(y)) = \frac{T}{N}$ for all $y \in J$. Then $a \equiv a_N$ satisfies the differential equation*

$$a'(y) = -\frac{\int q_x(y, y + a(y)x)\mu(dx)}{\int xq_x(y, y + a(y)x)\mu(dx)} \quad (25)$$

for all $y \in J$. Moreover, a is the unique solution: Let $\tilde{a}: J \rightarrow \mathbb{R}$ satisfy (25) for all $y \in J$ and $G(y_0, \tilde{a}(y_0)) = \frac{T}{N}$ for some $y_0 \in J$, then $\tilde{a}(y) = a(y)$ for all $y \in J$.

Proof. Since μ has compact support, it follows from the dominated convergence theorem that the function G is C^1 in both arguments in the interior of the set $S = \{(y, a) \in I \times \mathbb{R}_+ : G(y, a) < \infty\}$. In particular the partial derivatives are given by

$$G_a(y, a) = \int x q_x(y, y + ax) \mu(dx)$$

and

$$\begin{aligned} G_y(y, a) &= \frac{\partial}{\partial y} \int (q(\bar{y}, y + ax) - q(\bar{y}, y) - q_x(\bar{y}, y)ax) \mu(dx) \\ &= \int (q_x(\bar{y}, y + ax) - q_x(\bar{y}, y)) \mu(dx) = \int q_x(y, y + ax) \mu(dx), \end{aligned}$$

where we use Equation (12) with an arbitrary $\bar{y} \in I$. Moreover, since $G_a(y, a) > 0$ in the interior of S , it follows from the implicit function theorem that a is C^1 in J and satisfies

$$a'(y) = -\frac{G_y(y, a(y))}{G_a(y, a(y))} = -\frac{\int q_x(y, y + a(y)x) \mu(dx)}{\int x q_x(y, y + a(y)x) \mu(dx)}$$

for all $y \in J$. This yields the first claim. Now let \tilde{a} satisfy (25) on J with $G(y_0, \tilde{a}(y_0)) = \frac{T}{N}$ for some $y_0 \in J$. Then the function $y \mapsto G(y, \tilde{a}(y))$ has vanishing derivative on J , hence is constant on J , that is, $G(y, \tilde{a}(y)) = \frac{T}{N}$ for all $y \in J$. Since for all $y \in J$ the mapping $a \mapsto G(y, a)$ is strictly increasing (for those a where $G(y, a) < \infty$), we conclude that $a = \tilde{a}$ on J . This completes the proof. \square

3 Rate of convergence

In this section we determine rate of convergence for our approximation scheme. Besides the assumptions of the previous sections we make the additional assumption that the diffusion coefficient η is bounded and bounded away from zero:

(C1) $|\eta|$ and $\frac{1}{|\eta|}$ are bounded on I .

Notice that (C1) implies that $q(m, \cdot)$ grows quadratically on I , more precisely, that

$$\underline{c}(x - m)^2 \leq q(m, x) \leq \bar{c}(x - m)^2, \quad x \in I, \quad (26)$$

for some constants $0 < \underline{c} < \bar{c} < \infty$.

Theorem 3.1. *Assume (C1) and $\int x^4 \mu(dx) < \infty$. Then $M_{\tau_k^N}$ and M_t belong to L^2 for all $k \in \mathbb{Z}_+$ and $t \geq 0$, and there exists a constant $C \in \mathbb{R}_+$ such that*

$$\|M_{\tau_k^N} - M_t\|_{L^2} \equiv \sqrt{E|M_{\tau_k^N} - M_t|^2} \leq CN^{-1/4}.$$

In particular, we also have $\|M_{\tau_N^N} - M_T\|_{L^1} \equiv E|M_{\tau_N^N} - M_T| \leq CN^{-1/4}$, and the numerical scheme that consists in approximating M_T via Y_N^N has weak order of convergence 1/4 in the sense that

$$|Ef(Y_N^N) - Ef(M_T)| \leq C\Lambda N^{-1/4} \quad (27)$$

for all Lipschitz continuous functions $f: I \rightarrow \mathbb{R}$ with global Lipschitz constant Λ .

It is interesting to notice that, in contrast to Theorem 1.1, we do not require μ to have a compact support in Theorem 3.1. But this is not surprising in view of Remark 1.2, which applies exactly under (C1).

To prove Theorem 3.1 we first provide an estimate for the second moment of the embedding stopping times. We recall that we assume that Condition (A) is satisfied and that a is the scale factor defined by (16).

Lemma 3.2. *Assume (C1) and $\int x^4 \mu(dx) < \infty$. Then there exists a stopping time τ with*

$$E[\tau] = \int q(m, m + a(m)x) \mu(dx) < \infty \quad (28)$$

such that $(M_\tau - m)/a(m) \sim \mu$. Moreover, any such stopping time satisfies

$$E[\tau^2] \leq 4 \int q^2(m, m + a(m)x) \mu(dx). \quad (29)$$

Proof. Let $L > 0$ be a lower bound for $|\eta|$. Then we have $q(y, x) \leq \frac{(x-y)^2}{L^2}$, $y, x \in I$. Since $\int x^2 \mu(dx) < \infty$, this implies that $\int q(m, m + a(m)x) \mu(dx) < \infty$. The latter integrability condition guarantees the existence of a stopping time τ with $E[\tau] = \int q(m, m + a(m)x) \mu(dx)$ such that $(M_\tau - m)/a(m) \sim \mu$ (see Theorem 3 in [1]).

Itô's formula yields

$$tq(m, M_t) = \int_0^t q(m, M_s) ds + \int_0^t sq_x(m, M_s) dM_s + \frac{1}{2}t^2,$$

where q_x denotes the partial derivative of q with respect to the second argument. Let (τ_n) denote a localizing sequence for the stochastic integral with respect to M such that $E[\tau_n^2] < \infty$ (one might just take any localizing sequence (τ'_n) and consider $\tau_n := n \wedge \tau'_n$). Let us take any stopping time τ satisfying (28). We then have

$$E[(\tau_n \wedge \tau)^2] \leq 2E[(\tau_n \wedge \tau)q(m, M_{\tau_n \wedge \tau})] \leq 2\sqrt{E[(\tau_n \wedge \tau)^2]E[q^2(m, M_{\tau_n \wedge \tau})]},$$

which implies

$$E[(\tau_n \wedge \tau)^2] \leq 4E[q^2(m, M_{\tau_n \wedge \tau})]. \quad (30)$$

By monotone convergence the left-hand side converges to $E[\tau^2]$ as $n \rightarrow \infty$. Since $|\eta|$ is bounded from above, using the Burkholder-Davis-Gundy inequality we get

$$E \sup_{t \geq 0} (M_{t \wedge \tau} - m)^2 \leq \text{const}_1 E \int_0^\tau \eta^2(M_s) ds \leq \text{const}_2 E\tau < \infty,$$

hence $(M_{t \wedge \tau})_{t \geq 0}$ is a uniformly integrable martingale. Applying Jensen's inequality to the convex function $x \mapsto q^2(m, x)$ yields

$$q^2(m, M_{\tau_n \wedge \tau}) = q^2(m, E[M_\tau | \mathcal{F}_{\tau_n \wedge \tau}]) \leq E[q^2(m, M_\tau) | \mathcal{F}_{\tau_n \wedge \tau}] \quad \text{a.s.},$$

which, together with (30), implies

$$E[\tau^2] \leq 4E[q^2(m, M_\tau)].$$

Since $(M_\tau - m)/a(m) \sim \mu$, the latter formula is precisely (29). \square

Proof of Theorem 3.1. Since $|\eta|$ is bounded from above and $E\tau_k^N = \frac{kT}{N} < \infty$, we have $E \int_0^{\tau_k^N} \eta^2(M_s) ds < \infty$, hence $M_{\tau_k^N} \in L^2$ for $k \in \mathbb{Z}_+$. Similarly, $M_t \in L^2$ for $t \geq 0$. Moreover, it holds

$$E|M_{\tau_N^N} - M_T|^2 = E \int_{\tau_N^N \wedge T}^{\tau_N^N \vee T} \eta^2(M_s) ds \leq U^2 E|\tau_N^N - T| \leq U^2 \sqrt{\text{Var}(\tau_N^N)}, \quad (31)$$

where $U < \infty$ is an upper bound for $|\eta|$. With $L > 0$ being a lower bound for $|\eta|$, we have $q(y, x) \leq \frac{(y-x)^2}{L^2}$, $y, x \in I$. By a conditional version of Lemma 3.2 above applied to the $(\mathcal{F}_{\tau_k^N + t})$ -stopping times $\rho_{k+1}^N = \tau_{k+1}^N - \tau_k^N$, we then have a.s.

$$E[(\rho_{k+1}^N)^2 | \mathcal{F}_{\tau_k^N}] = 4 \int_{\mathbb{R}} q^2(M_{\tau_k^N}, M_{\tau_k^N} + a_N(M_{\tau_k^N})x) \mu(dx) \leq \frac{4a_N^4(M_{\tau_k^N})}{L^4} \int x^4 \mu(dx).$$

One can prove (see [2, Lemma 3.2]) that, for all $N \in \mathbb{N}$ and $y \in I$, it holds

$$a_N(y) \leq U \sqrt{\frac{1}{\int x^2 \mu(dx)} \frac{T}{N}}$$

with U as above. It follows that a.s. we have

$$E[(\rho_{k+1}^N)^2 | \mathcal{F}_{\tau_k^N}] \leq 4 \frac{T^2}{N^2} \frac{U^4}{L^4} \frac{\int x^4 \mu(dx)}{(\int x^2 \mu(dx))^2}. \quad (32)$$

Since $\text{Var}(\tau_N^N) = \sum_{k=1}^N \text{Var}(\rho_k^N)$, we obtain from (31) and (32) that

$$\|M_{\tau_N^N} - M_T\|_{L^2} \leq CN^{-1/4}$$

with a constant C depending only on $T, L, U, \int x^2 \mu(dx)$ and $\int x^4 \mu(dx)$. \square

4 A simulation algorithm

We now cast our approximation scheme into an explicit simulation algorithm. For simplicity we suppose that the i.i.d. sequence (X_k) generating the scaled random walk (Y_k) satisfies $P(X_k = \pm 1) = \frac{1}{2}$, that is, $\mu = \frac{1}{2}(\delta_{-1} + \delta_1)$. Recall from Section 2 that in this case (Y_k) and the scale factors have numerically desirable properties (comparison principle, Lip(1), linear growth, vanishing at finite boundaries).

We first formulate the algorithm for diffusions M that a.s. do *not* attain the boundaries l and r in finite time.

- Algorithm 4.1.**
1. Determine $q(y, x) = \int_y^x \int_y^u \frac{2}{\eta^2(z)} dz du$.
 2. Choose a time horizon T and the number of time steps $N \in \mathbb{N}$.
 3. Solve with respect to $a \in \mathbb{R}_+$ the equation $q(y, y+a) + q(y, y-a) = 2T/N$ for all $y \in I$, which determines the scale factor $y \mapsto a(y)$.
 4. Simulate $Y_k = Y_{k-1} + a(Y_{k-1})X_k$, $Y_0 = m$, where X_1, X_2, \dots are i.i.d. with distribution $\frac{1}{2}(\delta_{-1} + \delta_1)$.

If the diffusion M attains at least one of its boundaries l and r in finite time with a positive probability, then the third step of Algorithm 4.1 has to be replaced by:

- 3'. For all $y \in I$ find the maximal $a \in \mathbb{R}_+$ such that $q(y, y+a) + q(y, y-a) \leq 2T/N$, which determines the scale factor $y \mapsto a(y)$.

Remark 4.2. In the examples presented in the next section the function q and the root of the function $\mathbb{R}_+ \ni a \mapsto q(y, y+a) + q(y, y-a) - 2T/N$ can be calculated explicitly. If no closed-form expression for q is known, one has to fall back on numerical methods. In order to find the root of the function $q(y, y+a) + q(y, y-a) - 2T/N$ one can apply the Newton method, which has excellent convergence properties because, for any y , the function $a \mapsto q(y, y+a) + q(y, y-a)$ is convex. Alternatively, one can solve ODE (25) to obtain a numerical approximation for a_N (to set up the Cauchy problem it is enough to find one point $(y_0, a_0) \in I \times (0, \infty)$ with $q(y_0, y_0 + a_0) + q(y_0, y_0 - a_0) = 2T/N$).

5 Examples

In this section we apply our approximation scheme to several example diffusions and report results from numerical experiments. Throughout we set the time horizon $T = 1$. Moreover, we assume that the i.i.d. sequence (X_k) generating the scaled random walk (Y_k) satisfies $P(X_k = \pm 1) = \frac{1}{2}$, that is, $\mu = \frac{1}{2}(\delta_{-1} + \delta_1)$. In any experiment we use Algorithm 4.1 resp. its variant with step 3'.

5.1 Diffusion between two media

Let $I = \mathbb{R}$ and, with some $A \in \mathbb{R} \setminus \{0\}$,

$$\eta(x) = 1_{(0,\infty)}(x) + A1_{(-\infty,0]}(x), \quad x \in \mathbb{R}.$$

Notice that we have

$$\begin{aligned} \text{for } y \geq 0: \quad q(y, x) &= \begin{cases} (x - y)^2, & x \geq 0, \\ y^2 - 2xy + \frac{1}{A^2}x^2, & x < 0, \end{cases} \\ \text{for } y \leq 0: \quad q(y, x) &= \begin{cases} \frac{1}{A^2}(x - y)^2, & x < 0, \\ \frac{1}{A^2}y^2 - \frac{2}{A^2}xy + x^2, & x \geq 0. \end{cases} \end{aligned}$$

One can show that the scale factor a_N is given by

$$a_N(y) = \begin{cases} \sqrt{1/N}, & y \in (\sqrt{1/N}, \infty), \\ \frac{1-A^2}{1+A^2}y + \sqrt{\frac{2A^2}{1+A^2} \frac{1}{N} + 2A^2 \frac{A^2-1}{(1+A^2)^2} y^2}, & y \in (0, \sqrt{1/N}], \\ \frac{1-A^2}{1+A^2}y + \sqrt{\frac{2A^2}{1+A^2} \frac{1}{N} + 2 \frac{1-A^2}{(1+A^2)^2} y^2}, & y \in (-|A|\sqrt{1/N}, 0], \\ |A|\sqrt{1/N}, & y \in (-\infty, -|A|\sqrt{1/N}]. \end{cases}$$

If $M_0 = 0$, then the probability for the diffusion to be greater than zero is equal to $A/(1+A)$, for any time $t > 0$ (see e.g. Section 7 in [21]). How well does a scaled random walk with scale factor a_N approximate this probability? Figure 1 depicts the results of a numerical experiment, where $P(M_1 > 0)$ is approximated with its Monte Carlo estimator (denoted by $P_{emp}(M_1 > 0)$). The empirical convergence rate is roughly 1/2. When using the Euler method, the approximation error becomes larger, but has the same empirical rate of convergence.

5.2 Absorbed Brownian motion

Let M be a Brownian motion that is absorbed at zero, i.e. we have $I = (0, \infty)$ and $\eta(x) = 1_{(0,\infty)}(x)$. For $y > 0$, we have

$$q(y, x) = \begin{cases} (x - y)^2, & x \geq 0, \\ \infty, & x < 0, \end{cases}$$

and the scale factor is given by $a_N(y) = \frac{1}{\sqrt{N}} \wedge y$. How well does a scaled random walk with scale factor a_N approximate the distribution of M at time 1? In order to illustrate the accuracy we have performed a numerical experiment estimating the probability $P(M_1 > 0)$ for the starting point $M_0 = 1$ (see Figure 2). We have calculated the empirical probability for the paths to be positive at time 1. We have also calculated the empirical probability from simulations generated with the Euler method. The exact value for this probability is given by $2\Phi(M_0) - 1$, where Φ denotes the distribution function of a standard normal random variable. Figure 2 depicts the estimation error in dependence of the number of time steps. Observe that the estimator based on our method converges faster than the estimator based on the Euler scheme.

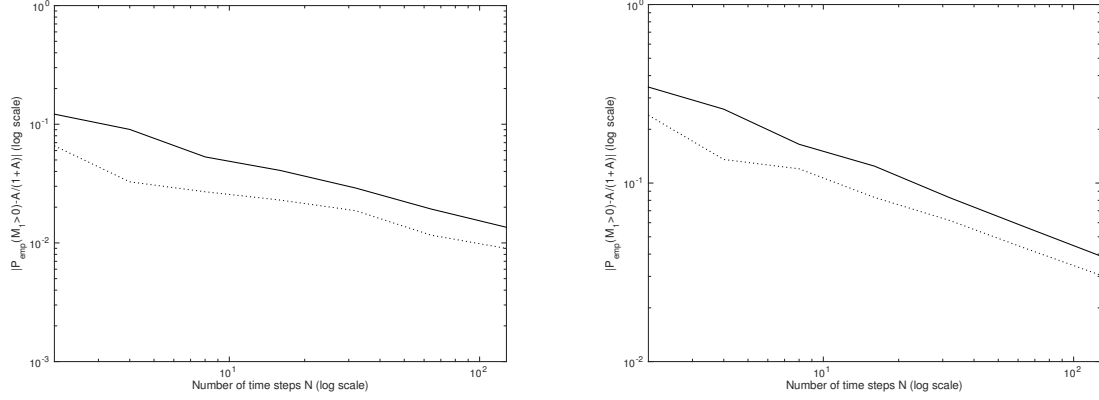


Figure 1: The figures illustrate the dependence of the estimators $P_{emp}(M_1 > 0)$ on the number of time steps N in the situation of Section 5.1. The exact value satisfies $P(M_1 > 0) = \frac{A}{1+A}$. The solid line depicts the logarithmic approximation error based on the Euler method. The dotted line shows the estimated values based on our method. In the left figure $A = 2$, in the right figure $A = 20$.

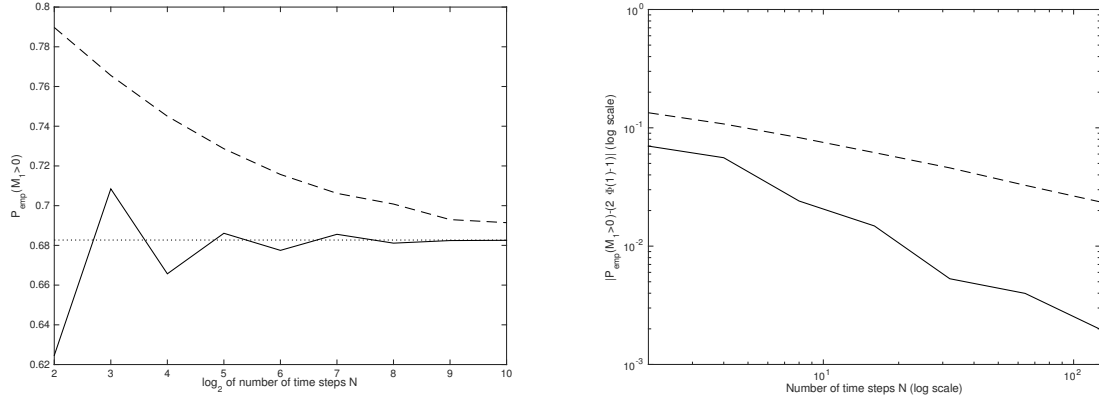


Figure 2: The figures illustrate the dependence of the estimators for $P(M_1 > 0)$ on the number of time steps N in the situation of Section 5.2. In the left figure the dotted line indicates the level $2\Phi(1) - 1$ of the exact value. The dashed line depicts the approximation based on the Euler method. The solid line shows the estimated values based on our method. The right figure shows a loglog plot of the approximation error.

5.3 An exponentially growing diffusion coefficient

Let $I = \mathbb{R}$ and $\eta(x) = \cosh(x)$, $x \in \mathbb{R}$. Then

$$q(y, x) = 2 \left[\log \left(\frac{\cosh(x)}{\cosh(y)} \right) - \tanh(y)(x - y) \right], \quad y, x \in \mathbb{R},$$

$$G(y, a) = 2 \int_{\mathbb{R}} \log \left(\frac{\cosh(y + ax)}{\cosh(y)} \right) \mu(dx), \quad y \in \mathbb{R}, a \geq 0.$$

For the choice $\mu = \frac{1}{2}(\delta_1 + \delta_{-1})$ we have

$$G(y, a) = \log \left(\frac{\cosh(y-a) \cosh(y+a)}{\cosh^2(y)} \right) = \log \left(\frac{\cosh(2y) + \cosh(2a)}{2 \cosh^2(y)} \right).$$

Consequently, for $N \in \mathbb{N}$ the scale factor is given by

$$a_N(y) = \frac{1}{2} \operatorname{arcosh} \left(2(\exp(1/N) - 1) \cosh^2(y) + 1 \right). \quad (33)$$

It follows from Corollary 2.13 that $Ef(Y_N^N) \rightarrow Ef(M_1)$ as $N \rightarrow \infty$ for any continuous function $f: \mathbb{R} \rightarrow \mathbb{R}$ with sublinear growth. By Remark 2.14 the process M is a strict local martingale and therefore $Ef(Y_N^N)$ in general does not converge to $Ef(M_1)$ for linear functions f . The Euler scheme, however, by Proposition 2.17 already fails to provide approximations of expectations like $E|M_1|^\alpha$ with $\alpha \in (0, 1)$. We depict the behaviour of the two schemes in Figure 3.

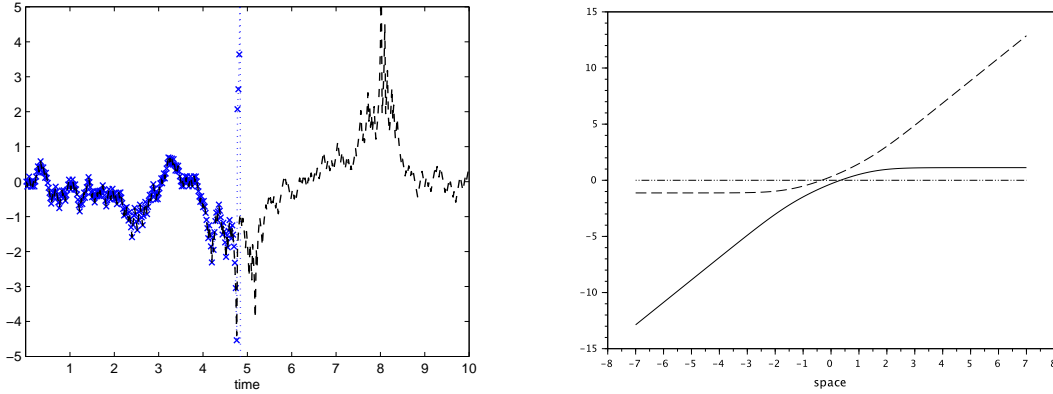


Figure 3: The figure on the left-hand side shows two realizations of discrete approximations of the SDE $dM_t = \cosh(M_t)dW_t$ with $M_0 = 0$. The dashed line depicts the realization based on our method. The crosses show the realization obtained with the Euler method. Both use the same realized binomial increments. Notice that the approximations are nearly identical until shortly before time 5. The large absolute values entail that the Euler approximation explodes and eventually aborts, whereas the dashed approximation easily continues. In the right figure the solid and dashed lines are the graphs of the functions $y \mapsto y - a_N(y)$ and $y \mapsto y + a_N(y)$. The dash-dotted line indicates level zero. Observe that monotonicity and linear growth of both functions implies that such explosions are impossible in our scheme (for a general statement, recall Theorem 2.8).

5.4 A diffusion coefficient that is not locally bounded

Let $I = \mathbb{R}$ and $\eta(x) = 1/|x|$, $x \in \mathbb{R} \setminus \{0\}$, and $\eta(0) = 1$. Then

$$q(y, x) = \frac{1}{6}x^4 - \frac{2}{3}xy^3 + \frac{1}{2}y^4, \quad y, x \in \mathbb{R},$$

$$G(y, a) = a^2 \int (y^2x^2 + \frac{1}{6}a^2x^4) \mu(dx), \quad y \in \mathbb{R}, a \geq 0.$$

For the particular choice $\mu = \frac{1}{2}(\delta_1 + \delta_{-1})$ we have

$$G(y, a) = \frac{1}{6}a^4 + a^2y^2, \quad y \in \mathbb{R}, \quad a \geq 0.$$

The scale factor is then given by

$$a_N(y) = \sqrt{\sqrt{9y^4 + \frac{6}{N}} - 3y^2}.$$

Observe that a_N satisfies

$$a_N^2(y) = \frac{\frac{6}{N}}{\sqrt{9y^4 + \frac{6}{N} + 3y^2}}.$$

It follows that the estimates

$$\sqrt{N}|y|a_N(y) \leq 1 \quad \text{and} \quad a_N(y) \leq a_N(0) = \left(\frac{6}{N}\right)^{1/4} \quad (34)$$

hold for all $y \in \mathbb{R}$.

Observe that η is not locally bounded in this example. However, the assertion of Theorem 1.1 still holds true. For a precise formulation, we as usual extend the discrete-time scaled random walk (Y_k^N) with scale factor a_N to the continuous-time process (Y_t^N) by linear interpolation.

Proposition 5.1. *Let $\mu = \frac{1}{2}(\delta_{-1} + \delta_1)$. The sequence of continuous processes $(Y_{Nt}^N)_{t \geq 0}$ converges in law to the process $(M_t)_{t \geq 0}$, as $N \rightarrow \infty$. Moreover, $Ef(Y_{Nt}^N) \rightarrow Ef(M_1)$ as $N \rightarrow \infty$ for any continuous function $f: \mathbb{R} \rightarrow \mathbb{R}$ with $\lim_{|x| \rightarrow \infty} \frac{|f(x)|}{x^4} = 0$.*

Proof. Let M^y be a solution of $dM_t = \eta(M_t)dW_t$ with initial condition $M_0 = y \in \mathbb{R}$. We can assume that $M_t^y = y + W_{A(t)}$, where (W_t) is a Brownian motion and A is a time change. Indeed, let $\tau(t) = \int_0^t |y + W_s|^2 ds$ and $A(t) = \inf\{s \geq 0 : \tau(s) > t\}$, for $t \geq 0$. Then the process $M_t^y = y + W_{A(t)}$, together with the Brownian motion $B_t^y = \int_0^t |M_s^y| dM_s^y$, is a weak solution of $dM_t = \eta(M_t)dW_t$ with initial condition $M_0 = y$.

Let $\rho^N(y) = \inf\{t \geq 0 : |M_t^y - y| = a_N(y)\}$. We show that the family $\{N\rho^N(y) : y \in \mathbb{R}, N \in \mathbb{N}\}$ is uniformly integrable. Then the first claim follows from the proof of Theorem 3.1 in [2] (the proof relies only on the uniform integrability of the family $\{N\rho^N(y)\}$ and does not directly employ the Condition (C1) that is assumed in the statement of the theorem).

Notice that $M_{\rho^N(y)}^y = y + W_{H^W(-a_N(y), a_N(y))}$, where we denote by $H^X(a, b)$ the first time a process X attains level a or b . Then

$$\begin{aligned} N\rho^N(y) &= N \tau(H^W(-a_N(y), a_N(y))) \\ &= N \int_0^{H^W(-a_N(y), a_N(y))} |y + W_s|^2 ds \\ &= N \int_0^{H^W(-a_N(y), a_N(y))/a_N^2(y)} |y + W_{sa_N^2(y)}|^2 a_N^2(y) ds. \end{aligned} \quad (35)$$

Let $\tilde{W}_s = \frac{1}{a_N(y)} W_s a_N^2(y)$. Observe that $H^W(-a_N(y), a_N(y))/a_N^2(y) = H^{\tilde{W}}(-1, 1)$. Hence (35) further implies

$$\begin{aligned} N\rho^N(y) &= N \int_0^{H^{\tilde{W}}(-1,1)} |y + a_N(y)\tilde{W}_s|^2 a_N^2(y) ds \\ &= \int_0^{H^{\tilde{W}}(-1,1)} |\sqrt{N}a_N(y)y + \sqrt{N}a_N^2(y)\tilde{W}_s|^2 ds \\ &\leq \int_0^{H^{\tilde{W}}(-1,1)} |1 + \sqrt{6}|\tilde{W}_s||^2 ds =: \tilde{\tau}, \end{aligned}$$

where the last inequality follows from (34).

Notice that $\tilde{\tau}$ does not depend on y and N . Moreover, a similar time change argument shows that $\tilde{\tau}$ has the same distribution as the stopping time $H^{\tilde{M}}(-1, 1)$, where $\tilde{M}_t = \int_0^t \tilde{\eta}(\tilde{M}_s) d\tilde{W}_s$ with coefficient $\tilde{\eta}(x) = \frac{1}{1+\sqrt{6}|x|}$. Since $\int \int_0^x \int_0^u \frac{2}{\tilde{\eta}(z)^2} dz du \mu(dx) < \infty$, it follows that $\tilde{\tau}$ is integrable (see e.g. Theorem 4 in [9]).

Finally observe that Proposition 2.11 implies the second claim. \square

In contrast to Proposition 5.1, in the situation of Section 5.4, the Euler scheme does *not* converge in distribution to M in the case $M_0 = 0$. Indeed, let $(\tilde{Y}_k^N)_{k \in \{0, \dots, N\}}$ denote a Euler type approximation of M on the time interval $[0, 1]$ with $N \in \mathbb{N}$ time steps, i.e.

$$\tilde{Y}_{k+1}^N = \tilde{Y}_k^N + \frac{1}{\sqrt{N}} \eta(\tilde{Y}_k^N) \xi_{k+1},$$

where $\tilde{Y}_0^N = 0$ and ξ_1, ξ_2, \dots are independent identically distributed random variables with $E\xi_k = 0$ and $E\xi_k^2 = 1$. For simplicity, assume $P(\xi_k = 0) = 0$. Then we have $\tilde{Y}_1^N = \xi_1/\sqrt{N}$ and $\tilde{Y}_2^N = \xi_1/\sqrt{N} + \xi_2/|\xi_1|$. Denote by $(\tilde{Y}_t^N)_{t \in [0, N]}$ the continuous-time process obtained from $(\tilde{Y}_k^N)_{k \in \{0, \dots, N\}}$ by linear interpolation. For $\varepsilon > 0$ let Φ^ε be the bounded continuous functional on $C[0, 1]$ defined by $\Phi^\varepsilon(\omega) = \sup_{s \in [0, \varepsilon]} |\omega(s)| \wedge 1$ for $\omega \in C[0, 1]$. For $N > 2/\varepsilon$ we have $\Phi^\varepsilon(\tilde{Y}_N^N) \geq |\xi_1/\sqrt{N} + \xi_2/|\xi_1|| \wedge 1$. It follows that $\liminf_{N \rightarrow \infty} E[\Phi^\varepsilon(\tilde{Y}_N^N)] \geq E[|\xi_2/\xi_1| \wedge 1] > 0$. But we have $E[\Phi^\varepsilon(M)] \rightarrow 0$ as $\varepsilon \rightarrow 0$. Consequently, the sequence of continuous processes (\tilde{Y}_{Nt}^N) does not converge in distribution to M . The behaviour of the two schemes is illustrated in Figure 4.

5.5 Diffusions with drift

Finally, we explain in more detail how to approximate solutions of time-homogeneous SDEs *with* drift via our method. Let $J = (\alpha, \beta)$, $-\infty \leq \alpha < \beta \leq \infty$, b and σ are Borel functions $J \rightarrow \mathbb{R}$ satisfying

$$\sigma(x) \neq 0 \text{ for all } x \in J, \tag{36}$$

$$|\sigma| \text{ and } \frac{1}{|\sigma|} \text{ are locally bounded on } J, \tag{37}$$

$$b \in L_{\text{loc}}^1(J). \tag{38}$$

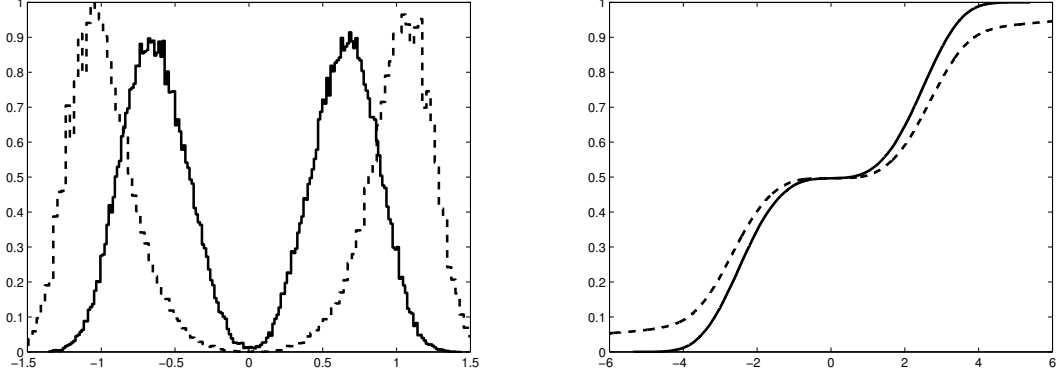


Figure 4: The figures illustrate the numerical performance of our method and the weak Euler scheme with increments drawn from $\mu = \frac{1}{2}(\delta_{-1} + \delta_1)$ for the SDE $dM_t = \eta(M_t)dW_t$ with $\eta(x) = 1/|x|$, $x \in \mathbb{R} \setminus \{0\}$, $\eta(0) = 1$ and $M_0 = 0$. In both cases we use $N = 1000$ time steps and simulate 100000 paths. The figure on the left-hand side depicts the empirical probability density function of Y_N^N (solid) and \tilde{Y}_N^N (dashed) for the time horizon $T = 0.05$. Notice that the two maxima of the density of \tilde{Y}_N^N are centered around -1.1 and 1.1 . This comes from the fact that $\tilde{Y}_2^N = \xi_1/\sqrt{N} + \xi_2/|\xi_1|$ which implies that for large values of N the Euler scheme rather starts in $\frac{1}{2}(\delta_{-1} + \delta_1)$ than in δ_0 . In contrast our method is able to reflect the *short time behavior* of the density of M . The right-hand side shows the empirical distribution function of Y_N^N (solid) and \tilde{Y}_N^N (dashed) for the time horizon $T = 10$. Observe that the Euler scheme compared to our method puts a lot of mass outside the interval $[-6, 6]$. This is due to the fact that in the long run the process (\tilde{Y}_k^N) gets very close to 0 with high probability and then shoots out of the interval $[-6, 6]$, since \tilde{a}_N inherits the singularity of η around zero. Our method does not produce these outliers since it smoothens the singularity.

Let X be a unique in law weak solution of the SDE

$$dX_t = b(X_t) dt + \sigma(X_t) dW_t, \quad X_0 = x_0 \in J, \quad (39)$$

up to the exit time from J . Notice that (39) has a unique in law (possibly exiting J) weak solution, because under (36)–(38) we have $\frac{1+|b|}{\sigma^2} \in L_{\text{loc}}^1(J)$. The scale function of X is given by the formula

$$p(x) = \int_c^x \exp \left\{ - \int_c^y \frac{2b}{\sigma^2}(z) dz \right\} dy, \quad x \in J, \quad (40)$$

where $c \in J$ is arbitrary. We define

$$l := p(\alpha) = \lim_{x \searrow \alpha} p(x) \in [-\infty, \infty) \quad \text{and} \quad r := p(\beta) = \lim_{x \nearrow \beta} p(x) \in (-\infty, \infty].$$

Itô's formula yields that $M = p(X)$ is a driftless diffusion driven by (6) with starting point $m = p(x)$, interior of the state space $I = (l, r) = (p(\alpha), p(\beta))$ and diffusion coefficient $\eta(x) = (p'\sigma) \circ p^{-1}(x)$, $x \in I$ (to ensure (5) we also set $\eta(x) = 0$ for all $x \in \mathbb{R} \setminus I$). Notice that exit of X from J in finite time corresponds to exit of M from I in finite time.

It follows from (36)–(38) that $|\eta|$ and $\frac{1}{|\eta|}$ are locally bounded on I , hence, Theorem 1.1 applies (recall that we are considering $\mu = \frac{1}{2}(\delta_{-1} + \delta_1)$; alternatively, one can take any other centered $\mu \neq \delta_0$ as specified in Theorem 1.1). We again extend the discrete-time scaled random walk (Y_k^N) to the continuous-time process (Y_t^N) by linear interpolation. By Theorem 1.1, the sequence of processes (Y_{Nt}^N) , constructed via our method, converges weakly to (M_t) as $N \rightarrow \infty$. Then $(p^{-1}(Y_{Nt}^N))$ converges weakly to (X_t) . We can, therefore, use our embedding method for approximating the law of X . Again, the convergence holds true without regularity or growth assumptions on the coefficients of the driving SDE (note that (36)–(38) are neither regularity nor growth assumptions).

Let us look at a specific example with drift.

Example 5.2. We consider the time-homogeneous diffusion

$$dX_t = -\frac{1}{2} \tan(X_t) dt + dW_t, \quad X_0 = x_0, \quad (41)$$

with the state space $J = (-\frac{\pi}{2}, \frac{\pi}{2})$ and $x_0 \in J$. The derivative of the scale function (40) is given by

$$p'(x) = \exp \left\{ \int_0^x \tan(u) du \right\} = \frac{1}{\cos(x)}, \quad x \in J,$$

and hence we have

$$p(x) = \log \left(\tan \left(\frac{x}{2} + \frac{\pi}{4} \right) \right), \quad x \in J.$$

In particular, $p(-\frac{\pi}{2}) = -\infty$ and $p(\frac{\pi}{2}) = \infty$, that is, $I = \mathbb{R}$, and the inverse scale function $p^{-1}: \mathbb{R} \rightarrow J$ is given by

$$p^{-1}(x) = 2 \arctan(\exp(x)) - \frac{\pi}{2}. \quad (42)$$

It follows that $M = p(X)$ is a local martingale driven by the driftless SDE with state space $I = \mathbb{R}$ and diffusion coefficient $\eta(x) = p'(p^{-1}(x)) = \cosh(x)$. Thus, we are in the situation of Section 5.3. In particular, the scale factor a_N determining the scaled random walk Y^N is given in closed form by (33) in the case $\mu = \frac{1}{2}(\delta_{-1} + \delta_1)$. Finally, to get a weak approximation for the diffusion (X_t) of (41), we simulate $(p^{-1}(Y_{Nt}^N))$ with p^{-1} given by (42).

References

- [1] S. Ankirchner, D. Hobson, and P. Strack. Finite, integrable and bounded time embeddings for diffusions. *Bernoulli*, 21(2):1067–1088, 2015.
- [2] S. Ankirchner, T. Kruse, and M. Urusov. A functional limit theorem for irregular SDEs. *Preprint, arXiv:1409.7940v3*, 2015.

- [3] P. Carr, T. Fisher, and J. Ruf. Why are quadratic normal volatility models analytically tractable? *SIAM J. Financial Math.*, 4(1):185–202, 2013.
- [4] H. J. Engelbert and W. Schmidt. On solutions of one-dimensional stochastic differential equations without drift. *Z. Wahrsch. Verw. Gebiete*, 68(3):287–314, 1985.
- [5] A. Etheridge. *Some mathematical models from population genetics*, volume 2012 of *Lecture Notes in Mathematics*. Springer, Heidelberg, 2011. Lectures from the 39th Probability Summer School held in Saint-Flour, 2009.
- [6] P. Etoré and A. Lejay. A Donsker theorem to simulate one-dimensional processes with measurable coefficients. *ESAIM: Probability and Statistics*, 11:301–326, 2007.
- [7] I. Gyöngy. A note on Euler’s Approximations. *Potential Analysis*, 8(3):205–216, 1998.
- [8] D. J. Higham, X. Mao, and A. M. Stuart. Strong convergence of Euler-type methods for nonlinear stochastic differential equations. *SIAM J. Numer. Anal.*, 40(3):1041–1063 (electronic), 2002.
- [9] D. Hobson. Integrability of solutions of the Skorokhod embedding problem for diffusions. *Preprint, arXiv:1403.2214*, 2014.
- [10] Y. Hu. Semi-implicit Euler-Maruyama scheme for stiff stochastic equations. In *Stochastic analysis and related topics, V (Silivri, 1994)*, volume 38 of *Progr. Probab.*, pages 183–202. Birkhäuser Boston, Boston, MA, 1996.
- [11] M. Hutzenthaler and A. Jentzen. On a perturbation theory and on strong convergence rates for stochastic ordinary and partial differential equations with non-globally monotone coefficients. *Preprint, arXiv:1401.0295*, 2014.
- [12] M. Hutzenthaler and A. Jentzen. Numerical approximation of stochastic differential equations with non-globally Lipschitz continuous coefficients. *Mem. Amer. Math. Soc.*, 236(1112), 2015.
- [13] M. Hutzenthaler, A. Jentzen, and P. Kloeden. Strong and weak divergence in finite time of Euler’s method for stochastic differential equations with non-globally Lipschitz continuous coefficients. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 467(2130):1563–1576, 2011.
- [14] M. Hutzenthaler, A. Jentzen, and P. E. Kloeden. Strong convergence of an explicit numerical method for SDEs with non-globally Lipschitz continuous coefficients. *Ann. Appl. Probab.*, 22(4):1611–1641, 2012.
- [15] G. Johnson and L. L. Helms. Class D supermartingales. *Bull. Amer. Math. Soc.*, 69:59–62, 1963.

- [16] B. Jourdain and A. Kohatsu-Higa. A review of recent results on approximation of solutions of stochastic differential equations. In *Stochastic analysis with financial applications*, volume 65 of *Progr. Probab.*, pages 141–164. Birkhäuser/Springer Basel AG, Basel, 2011.
- [17] I. Karatzas and S. E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.
- [18] P. E. Kloeden and E. Platen. *Numerical solution of stochastic differential equations*, volume 23. Springer, 1992.
- [19] A. Kohatsu-Higa, A. Lejay, and K. Yasuda. Weak approximation errors for stochastic differential equations with non-regular drift. *Preprint*, 2013.
- [20] S. Kotani. On a condition that one-dimensional diffusion processes are martingales. In *In Memoriam Paul-André Meyer, Séminaire de Probabilités XXXIX*, pages 149–156. Springer, 2006.
- [21] A. Lejay. On the constructions of the skew Brownian motion. *Probab. Surv.*, 3:413–466, 2006.
- [22] A. Lejay and G. Pichot. Simulating diffusion processes in discontinuous media: a numerical scheme with constant time steps. *J. Comput. Phys.*, 231(21):7299–7314, 2012.
- [23] H.-L. Ngo and D. Taguchi. Strong rate of convergence for the Euler-Maruyama approximation of stochastic differential equations with irregular coefficients. *Preprint, arXiv:1311.2725v2*, 2014.
- [24] S. Sabanis. Euler approximations with varying coefficients: the case of superlinearly growing diffusion coefficients. *Preprint, arXiv:1308.1796v3*, 2015.
- [25] A. N. Shiryaev. *Probability*, volume 95 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1996. Translated from the first (1980) Russian edition by R. P. Boas.
- [26] L. Yan. The Euler scheme with irregular coefficients. *The Annals of Probability*, 30(3):1172–1194, 2002.