Properties of the EMCEL scheme for approximating irregular diffusions

Stefan Ankirchner∗ Thomas Kruse† Wolfgang Löhr‡ Mikhail Urusov§

December 16, 2021

We prove several properties of the EMCEL scheme, which is capable of approximating one-dimensional continuous strong Markov processes in distribution on the path space (the scheme is briefly recalled). Special cases include irregular stochastic differential equations and processes with sticky features. In particular, we highlight differences from the Euler scheme in the case of stochastic differential equations and discuss a certain “stabilizing” behavior of the EMCEL scheme like “smoothing and tempered growth behavior”.

Keywords: one-dimensional Markov process; speed measure; sticky point; Markov chain approximation; numerical scheme.

2010 MSC: Primary: 60J25; 60J60. Secondary: 60J22; 60H35.

Introduction

The aim of this paper is to prove several desirable properties of the EMCEL approximation scheme (the idea behind the scheme is briefly recalled below in the introduction, and the scheme is formally described in Section [1]), which is capable of approximating all one-dimensional continuous strong Markov processes (abbreviated as general diffusions in what follows).

∗Stefan Ankirchner, Institute of Mathematics, University of Jena, Ernst-Abbe-Platz 2, 07745 Jena, Germany. Email: s.ankirchner@uni-jena.de, Phone: +49 (0)3641 946275.
†Thomas Kruse, Institute of Mathematics, University of Giessen, Arndtstr. 2, 35392 Gießen, Germany. Email: thomas.kruse@math.uni-giessen.de, Phone: +49 (0)641 9932102.
‡Wolfgang Löhr, Faculty of Mathematics, University of Duisburg-Essen, Thea-Leymann-Str. 9, 45127 Essen, Germany. Email: wolfgang.loehr@uni-due.de, Phone: +49 (0)201 1834729.
§Mikhail Urusov, Faculty of Mathematics, University of Duisburg-Essen, Thea-Leymann-Str. 9, 45127 Essen, Germany. Email: mikhail.urusov@uni-due.de, Phone: +49 (0)201 1837428.
The set of general diffusions includes strong and weak solutions of one-dimensional stochastic differential equations (SDEs) with possibly irregular coefficients whenever uniqueness in law holds for the SDE. For the case of SDEs, there are many different approaches for approximating their solutions, e.g., the Euler scheme. There are, however, many general diffusions that cannot be written as solutions to SDEs. This is, in particular, true for general diffusions with sticky features, where a sticky point is located in the interior of the state space. A related interesting phenomenon is sticky reflection, where the sticky point is located at the boundary of the state space. Recent years have witnessed an increased interest in general diffusions with sticky features, see [8], [11], [12], [13], [16], [17], [18], [20], [24], [28], [29] and references therein. All such (and other) general diffusions can be approximated via the EMCEL scheme.

Let us briefly illustrate the central idea behind the construction of the EMCEL scheme. To this end let \( Y = (Y_t)_{t \in [0, \infty)} \) be a general diffusion in natural scale. For simplicity we assume throughout this paragraph that the state space of the general diffusion \( Y \) is equal to the whole real line \( \mathbb{R} \). Let \( (\xi_k)_{k \in \mathbb{N}} \) be an iid sequence of random variables, on a probability space with a measure \( P \), satisfying \( P(\xi_k = \pm 1) = \frac{1}{2} \). Given an initial value \( y_0 \in \mathbb{R} \) and a discretization parameter \( h \in (0, \infty) \), we recursively define a Markov chain on the time grid \( \{kh : k \in \mathbb{N}_0\} \) by the formula

\[
\hat{X}^h_0 = y_0 \quad \text{and} \quad \hat{X}^h_{(k+1)h} = \hat{X}^h_{kh} + \hat{a}_h(\hat{X}^h_{kh})\xi_{k+1}, \quad \text{for } k \in \mathbb{N}_0.
\]

Here the function \( \hat{a}_h : \mathbb{R} \to [0, \infty) \), termed scale factor in the sequel (as it is used to scale the incoming random variables \( \xi_{k+1} \) in \( \hat{Y}^h \)), is chosen in such a way that the expected time it takes \( Y \) started in \( y \in \mathbb{R} \) to leave the interval \( (y - \hat{a}_h(y), y + \hat{a}_h(y)) \) is equal to \( h \), i.e., \( \hat{a}_h \) satisfies for all \( y \in \mathbb{R} \) that \( E_y[H_y-y, \hat{a}_h(y), y+\hat{a}_h(y)](Y) = h \) where \( H_{b,c}(Y) \) is the first exit time of \( Y \) from the interval \( (b, c) \). Next, let \( \tau^h_0 = 0 \) and then recursively define \( \tau^h_{k+1} \) as the first time \( Y \) exits the interval \( (Y_{\tau^h_k} - \hat{a}_h(Y_{\tau^h_k}), Y_{\tau^h_k} + \hat{a}_h(Y_{\tau^h_k})) \) after \( \tau^h_k \). It follows that the discrete-time process \( (Y_{\tau^h_k})_{k \in \mathbb{N}_0} \) has the same law as the Markov chain \((\hat{X}^h_{kh})_{k \in \mathbb{N}_0} \) defined in \( \hat{Y}^h \). We say that the Markov chain \((\hat{X}^h_{kh})_{k \in \mathbb{N}_0} \) is embedded into \( Y \) with the sequence of stopping times \((\tau^h_k)_{k \in \mathbb{N}_0}\). Moreover, the stopping times satisfy that \( E_y[\tau^h_{k+1} - \tau^h_k] = h \). This explains why we refer to \((\hat{X}^h_{kh})_{k \in \mathbb{N}_0} \) as Embeddable Markov Chain with Expected time Lag \( h \) and write \( \hat{X}^h \in \text{EMCEL}(h) \) as a shorthand. A key observation is that the requirement \( E_y[H_y-y, \hat{a}_h(y), y+\hat{a}_h(y)](Y) = h \) can be transformed into the analytic condition

\[
\int_{(y-\hat{a}_h(y), y+\hat{a}_h(y))} (\hat{a}_h(y) - |u - y|) m(du) = 2h
\]

(see Remark 1.2 in [3]), where \( m \) denotes the speed measure of \( Y \). This condition is used to define the scale factors \( \hat{a}_h, h \in (0, \infty) \), which determine the scheme. For more general state spaces than \( \mathbb{R} \) this condition has to be adjusted appropriately. This is done in [8] below, where the EMCEL scheme is formally introduced in the general case. For the discussion of the embedding stopping times in the general situation we refer to Section 3 and, in particular, Proposition 3.1 in [3].
The properties of the EMCEL scheme discussed in this paper fall into the following three categories:

(i) asymptotic properties of the scale factor as $h \downarrow 0$ (Section 2.1);

(ii) stability properties (Section 2.2);

(iii) ODE characterization of the scale factor (Section 2.3).

In category (i) we, in fact, discuss many different properties. Some of them highlight relationships and differences between the EMCEL and the Euler schemes in the SDE case. The main result in category (ii) is Theorem 2.9 (a so-called comparison principle), which implies some “stabilizing” behavior of the EMCEL scheme like “smoothing and tempered growth behavior”. The main result in category (iii) is Theorem 2.14 (a certain ODE for the scale factors), which gives understanding of what regularity the scale factors do have in general and helps implementing the EMCEL scheme in specific situations.

We refrain from a further discussion in the introduction, as such a discussion would require a thorough description of the setting. Much more details, also to the meaning of our results, are present in Section 2.

It remains to describe relations with the literature and to state our contributions. For classical results on the approximation of SDEs with *globally* Lipschitz coefficients via the Euler scheme we refer to the monographs [26] and [37] and to references therein. In the case of SDEs with *locally* Lipschitz coefficients the Euler scheme is known to converge almost surely (see [19]), but, unless the Lipschitz condition is global, neither in strong nor in numerically weak sense (see [22] for the terminology and a precise result of this kind). There are, however, results in positive direction: e.g., see [27] and [36] and references therein for results on the weak and strong convergence for the Euler-type approximations of SDEs with discontinuous coefficients. But, in contrast to the EMCEL scheme, the Euler scheme is defined only for SDEs (not for *all* general diffusions) and may fail to converge even in the stochastically weak sense when the SDE coefficients are irregular (see Section 5.4 in [3]).

It is necessary to say that there exist other schemes capable of approximating interesting subclasses of general diffusions, e.g., SDEs with discontinuous coefficients; see [15], [33], [34], [35] and references therein. However, the properties discussed in our paper are specific for the EMCEL scheme only. The EMCEL scheme appears as an important example in [5], [6] and [30], but the objects of study in those papers are certain classes of schemes for general diffusions. The properties studied in the present paper are not shared by the mentioned classes of schemes. Therefore, such properties are not at all discussed in the aforementioned papers.

The approach to approximate solutions of one-dimensional driftless SDEs via sequences of embedding stopping times appears in [3] and [1], where [1] proves a functional limit theorem for irregular SDEs, and [3] elaborates a scheme for approximating irregular diffusions. The latter is the EMCEL scheme in the SDE setting, although the name “EMCEL” does not appear in [3]. As compared to the present paper, the scheme
of [3] is described in somewhat different terms, the discussion of the properties in [3] is far less complete than that in the present paper and is performed under more restrictive assumptions (precisely: subclass of Example 1.1 below with locally bounded on $I^o$ functions $|\eta|$ and $\frac{1}{|\eta|}$), which are, however, essential for the proofs in [3]. To be more specific, [3] does not contain properties of category (i) (except for some very basic statements), but it contains the mentioned comparison principle and the ODE for the scale factors in its more restrictive setting (in particular, the ODE in [3] is simpler). Thus, our contribution essentially includes the properties of category (i), but also in categories (ii) and (iii) our contribution in comparison to [3] consists not in mere generalizations of the respective proofs but rather in finding completely new ones (and the right formulation of the ODE), as those ideas from [3] do not work in our present generality. More precisely, those proofs from [3] heavily rely on the implicit function theorem, which is not applicable in our situation (contrary to [3], the involved functions have kinks in general; an it is worth noting that the latter statement also follows from our general description in Theorem 2.14 below). For further details, see Section 2.

1 The EMCEL scheme

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, (P_y)_{y \in I}, (Y_t)_{t \geq 0})$ be a one-dimensional continuous strong Markov process in the sense of Section VII.3 in [39]. We refer to this class of processes as general diffusions in the sequel. We assume that the state space is an open, half-open or closed interval $I \subseteq \mathbb{R}$. We denote by $I^o = (l, r)$ the interior of $I$, where $-\infty \leq l < r \leq \infty$, and we set $\overline{I} = [l, r]$. Recall that by the definition we have $P_y[Y_0 = y] = 1$ for all $y \in I$.

We further assume that $Y$ is regular. This means that for every $y \in I^o$ and $x \in I$ we have that $P_y[H_x(Y) < \infty] > 0$, where $H_x(Y) = \inf\{t \geq 0 : Y_t = x\}$ (with the usual convention $\inf\emptyset = \infty$). Moreover, for $a < b$ in $\overline{I}$ we denote by $H_{a,b}(Y)$ the first exit time of $Y$ from $(a, b)$, i.e., $H_{a,b}(Y) = H_a(Y) \wedge H_b(Y)$. Without loss of generality we suppose that the general diffusion $Y$ is conservative (i.e., with infinite life time) and in natural scale. While the setting in Section VII.3 in [39] allows for finite life times, the process can be killed only at the endpoints of $I$ that do not belong to $I$, in which case we can add such an endpoint to $I$ and make it absorbing; such a procedure gives us a conservative process. If $Y$ is not in natural scale, then there exists a strictly increasing continuous function $s: I \rightarrow \mathbb{R}$, the so-called scale function, such that $s(Y)$ is in natural scale.

Let $m$ be the speed measure of the Markov process $Y$ on $I^o$ (see VII.3.7 in [39]). Recall that for all $a < b$ in $I^o$ we have

$$0 < m([a, b]) < \infty. \quad (2)$$

We also recall that a boundary point $b$ ($b \in \{l, r\}$) is called accessible if $P_y[H_b(Y) < \infty] > 0$ for some, hence, for all, $y \in I^o$ (such a definition because $Y$ is conservative). Due to our assumption that $Y$ is regular and conservative, $b \in \{l, r\}$ is an accessible boundary if and only if $b \in I$. For both boundaries $l$ and $r$, we assume that if a boundary
point is accessible, then it is absorbing\(^1\).

For what follows, we briefly recall Feller’s test for explosions (see, e.g., Theorem 23.12 in \[23\] or Lemma 2.1 in \[1\] or Theorem 3.3 in \[2\]): the left boundary \(l\) is accessible if and only if
\[
l > -\infty \quad \text{and} \quad \int_{(l,y)} (u - l) m(du) < \infty
\]  
for some, equivalently, for all, \(y \in I^o\) (recall that \(Y\) is in natural scale). Symmetric statement holds of course for the right boundary point \(r\).

**Example 1.1** (Driftless SDE with possibly irregular diffusion coefficient). Consider the case, where inside \(I^o\) the process \(Y\) is driven by the SDE
\[
dY_t = \eta(Y_t) dW_t,
\]
where \(W\) is a Brownian motion and \(\eta: I^o \to \mathbb{R}\) is a Borel function satisfying the Engelbert-Schmidt conditions
\[
\eta(x) \neq 0 \quad \forall x \in I^o,
\]
\[
\eta^{-2} \in L^1_{\text{loc}}(I^o)
\]
(\(L^1_{\text{loc}}(I^o)\) denotes the set of Borel functions locally integrable on \(I^o\)). Under (5)–(6) SDE (4) has a unique in law (possibly exiting \(I^o\) in finite time) weak solution; see [14] or Theorem 5.5.7 in [25]. We make the convention that \(Y\) remains constant after reaching \(l\) or \(r\) in finite time, which makes the boundary points absorbing whenever accessible. This is a particular case of our setting, where the speed measure of \(Y\) on \(I^o\) is given by the formula
\[
m(dx) = \frac{2}{\eta^2(x)} \, dx.
\]
This completes the description in Example 1.1.

We now recall the EMCEL approximation scheme introduced in [5]. Fix an arbitrary \(\bar{h} \in (0, \infty)\). For each \(h \in (0, \bar{h}]\) we define the function \(\hat{a}_h: \bar{T} \to [0, \infty)\) by the formulas
\[
\hat{a}_h(l) = \hat{a}_h(r) = 0 \quad \text{and, for } y \in I^o,
\]
\[
\hat{a}_h(y) = \sup \left\{ a \geq 0 : y \pm a \in I \quad \text{and} \quad \frac{1}{2} \int_{(y-a,y+a)} (a - |u - y|) m(du) \leq h \right\}.
\]
For now, fix some \(h \in (0, \bar{h}]\). We consider \(h\) as a discretization parameter. In what follows, the function \(\hat{a}_h\) is referred to as the EMCEL scale factor. Notice that, for all \(y \in \bar{T}\), we have
\[
y \pm \hat{a}_h(y) \in \bar{T}.
\]

---
\(^1\)The remaining case that \(Y\) has at least one reflecting boundary (both instantaneous and sticky reflections are allowed here), is reduced to the case of inaccessible or absorbing boundaries by a suitable symmetrization (see Section 6 in [5]). In this sense, that assumption does not result in a loss of generality. It is only convenient for the exposition.
We next construct an approximation \( \hat{X}^h \) of \( Y \) associated to the scale factor \( \hat{a}_h \). To this end, we fix a starting point \( y \in I^o \) of \( Y \). Let \((\xi_k)_{k \in \mathbb{N}}\) be an iid sequence of random variables, on a probability space with a measure \( P \), satisfying \( P(\xi_k = \pm 1) = \frac{1}{2} \). We denote by \( (\hat{X}^h_{kh})_{k \in \mathbb{N}_0} \) the Markov chain given by the formula

\[
\hat{X}^h_0 = y \quad \text{and} \quad \hat{X}^h_{(k+1)h} = \hat{X}^h_{kh} + \hat{a}_h(\hat{X}^h_{kh})\xi_{k+1}, \quad k \in \mathbb{N}_0,
\]

which is well-defined due to (3). We extend \( (\hat{X}^h_{kh})_{k \in \mathbb{N}_0} \) to a continuous-time process by linear interpolation, i.e., for all \( t \in [0, \infty) \), we set

\[
\hat{X}^h_t = \hat{X}^h_{\lfloor t/h \rfloor h} + (t/h - \lfloor t/h \rfloor)(\hat{X}^h_{\lfloor t/h \rfloor h} - \hat{X}^h_{\lfloor t/h \rfloor h}).
\]

To highlight the dependence of \( \hat{X}^h = (\hat{X}^h_t)_{t \in [0, \infty)} \) on the starting point \( y \in I^o \) we also sometimes write \( \hat{X}^h_{y,h} \).

The process \( \hat{X}^h \) is referred to as Embeddable Markov Chain with Expected time Lag \( h \) (we write shortly \( \hat{X}^h \in \text{EMCEL}(h) \)). The whole family \( (\hat{X}^h)_{h \in (0, \tilde{h}]} \) is referred to as the EMCEL approximation scheme.

It is worth noting that (9) can be made more precise:

\[
\text{for all } y \in I, \quad \text{we have } y \pm \hat{a}_h(y) \in I.
\]

This can be deduced from (8) with the help of Feller’s test. Indeed if, e.g., \( y - \hat{a}_h(y) = l \), then \( l > -\infty \) and \( \int_{(z,y)}(u - z) m(du) \leq 2h \) for all \( z \in (l, y) \) (see (8)). Letting \( z \downarrow l \) we obtain (3) by monotone convergence. Feller’s test implies that, in this case, \( l \) is accessible, and hence \( l \in I \). Observe that (12) translates into the fact that the EMCEL scheme never reaches inaccessible boundary points of \( I \) (cf. (10)–(11)), which is, in fact, a desirable property for a scheme.

**Remark 1.2.** In more detail, what is included in (8) is as follows. For all \( y \in I^o \), define \( a_f(y) = \min\{y - l, r - y\} \in (0, \infty] \) and notice that, for \( a \geq 0 \), it holds \( y \pm a \in I^o \) if and only if \( a < a_f(y) \). Fix \( y \in I^o \). It follows from (2) that \( \int_{(y-a,y+a)}(a - |u - y|) m(du) < \infty \) whenever \( a \in [0, a_f(y)) \). Therefore, the function

\[
a \mapsto \int_{(y-a,y+a)}(a - |u - y|) m(du) \equiv \int_I (a - |u - y|)^+ m(du)
\]

is strictly increasing and continuous on \( [0, a_f(y)) \) (by the dominated convergence theorem). The number \( \hat{a}_h(y) \) is thus a unique positive root of the equation (in \( a \in [0, a_f(y)) \))

\[
\frac{1}{2} \int_{(y-a,y+a)}(a - |u - y|) m(du) = h
\]

whenever \( \sup_{a \in [0,a_f(y))} \int_{(y-a,y+a)}(a - |u - y|) m(du) > h \), i.e., when \( y \) is “not too close” to an accessible boundary point of \( I \).
We now make the last statement more precise with the help of the following notations. If \( l > -\infty \), we define, for all \( h \in (0, h] \),
\[
l_h = l + \inf \left\{ a \in \left(0, \frac{r - l}{2}\right] : a < \infty \quad \text{and} \quad \frac{1}{2} \int_{(l,l+2a)} (a - |u - (l + a)|) m(du) \geq h \right\},
\]
where we use the convention \( \inf \emptyset = \frac{r - l}{2} \). If \( l = -\infty \), we set \( l_h = -\infty \). Similarly, if \( r < \infty \), then we define, for all \( h \in (0, h] \),
\[
r_h = r - \inf \left\{ a \in \left(0, \frac{r - l}{2}\right] : a < \infty \quad \text{and} \quad \frac{1}{2} \int_{(r-r, 2a, r)} (a - |u - (r - a)|) m(du) \geq h \right\}
\]
with the same convention \( \inf \emptyset = \frac{r - l}{2} \). If \( r = \infty \), we set \( r_h = \infty \). In any case, it holds \( l_h \leq r_h \) and, moreover, \( l_h \leq \frac{r - l}{2} \leq r_h \) whenever \( l \) and \( r \) are finite.

Using Feller’s test for explosions once again we see that \( l \) is inaccessible if and only if \( l_h = l \) for all \( h \in (0, h] \). Similarly, \( r \) is inaccessible if and only if \( r_h = r \) for all \( h \in (0, h] \). Notice that, if \( l \) or \( r \) are accessible, it holds that \( l_h \nearrow l \) or \( r_h \nearrow r \), respectively, as \( h \searrow 0 \).

Thus, the definitions of \( l_h \) and \( r_h \) yield that, for \( y \in (l_h, r_h) \), the number \( \hat{a}_h(y) \) is a unique positive root of (13), while, for \( y \in (l, l_h) \) (resp., \( y \in (r_h, r) \)), \( \hat{a}_h(y) \) is chosen to satisfy
\[
y - \hat{a}_h(y) = l \quad \text{(resp.,} \quad y + \hat{a}_h(y) = r \text{).} \tag{14}
\]
This concludes the detailed description of what is included in [8].

We emphasize that the EMCEL scheme is capable of weakly approximating every general diffusion \( Y \). For an illustration, we now recall a couple of results from [5], [6] and [30] and refer to those papers for more detail.

To formulate the results, we need to equip \([0, \infty] \times [0, \infty] \times C([0, \infty], \mathbb{R}) \) with a suitable topology. On \([0, \infty] \) we use the topology generated by the metric
\[
d(s, t) = \sqrt{\frac{s}{1 + s} - \frac{t}{1 + t}}, \quad s, t \in [0, \infty]
\]
with the convention that \( \frac{s}{1 + s} = 1 \). We equip \( C([0, \infty], \mathbb{R}) \) with the topology of uniform convergence on compact intervals, which is generated, e.g., by the metric
\[
\rho(x, y) = \sum_{n=1}^{\infty} 2^{-n} \left( \|x - y\|_{C[0, n]} \wedge 1 \right), \quad x, y \in C([0, \infty], \mathbb{R}),
\]
where \( \| \cdot \|_{C[0, n]} \) denotes the sup norm on \( C([0, n], \mathbb{R}) \). Finally, we use the standard product topology on the product space \([0, \infty] \times [0, \infty] \times C([0, \infty], \mathbb{R}) \).

The following result is a consequence of Theorem 2.1 in [30].

**Proposition 1.3.** For any speed measure \( m \) and for any \( y \in \mathcal{I}^0 \), the distributions of the random elements \( (H_l(\hat{X}^{h,y}), H_r(\hat{X}^{h,y}), \hat{X}^{h,y}) \) under \( P \) converge weakly to the distribution of \( (H_l(Y), H_r(Y), Y) \) under \( P_y \), as \( h \to 0 \); i.e., for every bounded and continuous functional \( F: [0, \infty] \times [0, \infty] \times C([0, \infty], \mathbb{R}) \to \mathbb{R} \), it holds that
\[
E[F(H_l(\hat{X}^{h,y}), H_r(\hat{X}^{h,y}), \hat{X}^{h,y})] \to E_y[F(H_l(Y), H_r(Y), Y)], \quad h \to 0. \tag{15}
\]
We remark that the weak convergence in (15) holds jointly for paths and exit times (i.e., hitting times $H_l$ and $H_r$ of the boundary points of the state space), which is a stronger statement than the weak convergence in the path space because the exit times are, in general, essentially discontinuous path functionals (e.g., $H_l$ is discontinuous with positive probability whenever the boundary $l$ is accessible).

The following result about convergence rates is a consequence of Theorem 1.7 in [6].

**Proposition 1.4.** Suppose that the speed measure $m$ satisfies

$$m(dx) \geq \frac{2}{k(1+x^2)} dx \quad \text{on } I^0$$

with some $k \in (0, \infty)$. Let $T \in (0, \infty)$ and let $F: C([0,T], I) \to \mathbb{R}$ be a locally Lipschitz continuous path functional with polynomially growing Lipschitz constant, i.e., there exist $L, \alpha \in [0, \infty)$ such that for all $x_1, x_2 \in C([0,T], I)$ it holds

$$|F(x_1) - F(x_2)| \leq L \{ 1 + (\|x_1\|_{C[0,T]} \vee \|x_2\|_{C[0,T]})^\alpha \} \|x_1 - x_2\|_{C[0,T]}.$$  

Then for every $\epsilon \in (0, \frac{1}{4})$ and $y \in I^0$ there exist a constant $C \in [0, \infty)$ such that for all $h \in (0, \frac{1}{h})$ it holds

$$\left| E\left[ F(\widehat{X}_t^{h,y}; t \in [0,T]) \right] - E_y [F(Y_t; t \in [0,T])] \right| \leq Ch^{\frac{1}{4} - \epsilon}.$$  

Moreover, in the case $F(x) = f(x(T))$ (with some $f: I \to \mathbb{R}$) when the functional $F$ depends only on the terminal value $x(T)$ of $x \in C([0,T], I)$, the rate is $\frac{1}{4}$, i.e., (18) holds with $\epsilon = 0$.

The role of assumption (16) is to ensure that the expected values in (18) exist. We remark that (16) does not exclude sticky features mentioned above, as these are modeled via atoms in $m$. Notice that in the case, where $Y$ is a solution of a driftless SDE of the form $dY_t = \eta(Y_t) dW_t$ (cf. Example 1.1), assumption (16) means that $\eta$ has at most linear growth; however, $\eta$ can be arbitrarily irregular (just a Borel function satisfying (5)–(6)). Finally, we stress that the rate $\frac{1}{4}$ in Proposition 1.4 cannot be considered as too slow because it holds in arbitrarily irregular cases (as discussed above, the only assumption (16) is not a regularity condition) and refer to [6] for more detail.

We, finally, mention that [6] contains also results about convergence rates of the EMCEL scheme in the Wasserstein distances.

### 2 Properties of the EMCEL scheme

In this section we gather several properties of the approximating Markov chain (10), which are encoded in the functions (scale factors) $\widehat{a}_h$, $h \in (0, \frac{1}{h}]$. 

8
2.1 Dependence on the discretization parameter

We first discuss properties from category (i) of the introduction. Specifically, here we study the asymptotic behavior of the EMCEL scale factors \( \hat{a}_h \) as \( h \downarrow 0 \). We also discuss relationships and differences with the Euler scheme in the SDE case (cf. Example 1.1).

**Proposition 2.1.** For any \( y \in I^\circ \), the function

\[
(0, \overline{h}] \ni h \mapsto \hat{a}_h(y)
\]

is strictly positive and nondecreasing. Moreover, for any \( y \in I^\circ \), there is \( h_0(y) \in (0, \overline{h}] \) such that the function in (19) is strictly increasing on \((0, h_0(y))\).

**Proof.** Both claims follow from the detailed description of the EMCEL scale factors \( \hat{a}_h \), \( h \in (0, \overline{h}] \), presented in Remark 1.2. Specifically, \( h_0(y) \) in the second claim can be defined as \( \sup \{ h \in (0, \overline{h}] : y \in [l_h, r_h] \} \).

**Proposition 2.2.** For any \( m \in \mathbb{N} \), defining \( K_m = I \cap [-m, m] \), we have

\[
\lim_{h \to 0} \sup_{y \in K_m} \hat{a}_h(y) = 0. 
\]

Notice that this result is a bit stronger than \( \lim_{h \to 0} \sup_{y \in K} \hat{a}_h(y) = 0 \) for all compact subsets \( K \) of \( I^\circ \): suprema over one-sided neighborhoods of finite boundary points of \( I^\circ \) are also included in (20).

**Proof.** Assume that (20) does not hold for some \( m \in \mathbb{N} \), i.e., there exist \( \varepsilon > 0 \) and sequences \( \{h_n\} \subset (0, \overline{h}], \{y_n\} \subset K_m \) such that \( h_n \to 0 \) and \( \hat{a}_{h_n}(y_n) \geq \varepsilon \) for all \( n \). By considering a suitable subsequence we assume without loss of generality that \( \{y_n\} \) is monotone and \( y_n \to y_\infty \) with some \( y_\infty \in K_m \). Notice that \( y_\infty \in I^\circ \) and its distance from the boundary of \( I^\circ \) is at least \( \varepsilon \), otherwise (9) would be violated for \( y_n \) with sufficiently large \( n \). Then we get

\[
\liminf_{n \to \infty} \frac{1}{2} \int_{(y_n - \hat{a}_{h_n}(y_n), y_n + \hat{a}_{h_n}(y_n))} (\hat{a}_{h_n}(y_n) - |u - y_n|) \, m(du) \\
\geq \liminf_{n \to \infty} \frac{1}{2} \int_{(y_n - \frac{1}{2}\hat{a}_{h_n}(y_n), y_n + \frac{1}{2}\hat{a}_{h_n}(y_n))} \frac{\hat{a}_{h_n}(y_n)}{2} \, m(du) \\
\geq \varepsilon \liminf_{n \to \infty} m \left( (y_n - \varepsilon/2, y_n + \varepsilon/2) \right) \geq \varepsilon \frac{1}{4} m \left( (y_\infty - \varepsilon/2, y_\infty + \varepsilon/2) \right) > 0.
\]

Observing that \( \{y_n\} \subset I^\circ \) (due to \( \hat{a}_{h_n}(y_n) > 0 \)), we get from (8)

\[
\frac{1}{2} \int_{(y_n - \hat{a}_{h_n}(y_n), y_n + \hat{a}_{h_n}(y_n))} (\hat{a}_{h_n}(y_n) - |u - y_n|) \, m(du) \leq h_n \to 0, \quad n \to \infty.
\]

The obtained contradiction concludes the proof. \( \square \)
Our next aim is to discuss the speed of convergence of \( \hat{a}_h(y) \) to zero, as \( h \to 0 \), for any fixed \( y \in I^\circ \). The next result helps establishing the order of convergence (in \( h \)) in many specific situations.

**Lemma 2.3.** For any \( y \in I^\circ \), there exists \( h_0 \in (0, \bar{h}] \) such that, for \( h \in (0, h_0] \), we have the inequalities

\[
\hat{a}_h(y) \sup_{\lambda \in [0,1]} \left\{ (1 - \lambda)m([y - \lambda\hat{a}_h(y), y + \lambda\hat{a}_h(y)]) \right\} \leq 2h \leq \hat{a}_h(y)m((y - \hat{a}_h(y), y + \hat{a}_h(y)))
\]  

(21)

and, in particular,

\[
\frac{\hat{a}_h(y)}{2} m([y - \hat{a}_h(y)/2, y + \hat{a}_h(y)/2]) \leq 2h \leq \hat{a}_h(y)m((y - \hat{a}_h(y), y + \hat{a}_h(y))).
\]  

(22)

**Proof.** Fix \( y \in I^\circ \). Choose a sufficiently small \( h_0 \in (0, \bar{h}] \) such that \( y \in (l_{h_0}, r_{h_0}) \). Remark \( \text{Remark 1.2} \) implies that, for all \( h \in (0, h_0] \), we have

\[
\int_{(y - \hat{a}_h(y), y + \hat{a}_h(y))} (\hat{a}_h(y) - |u - y|) \, m(du) = 2h.
\]

For any \( \lambda \in [0,1] \), we have

\[
(1 - \lambda)\hat{a}_h(y)m([y - \lambda\hat{a}_h(y), y + \lambda\hat{a}_h(y)]) \leq \int_{(y - \hat{a}_h(y), y + \hat{a}_h(y))} (\hat{a}_h(y) - |u - y|) \, m(du)
\]

\[
\leq \hat{a}_h(y)m((y - \hat{a}_h(y), y + \hat{a}_h(y))).
\]

This implies both claims.

We will extensively use the following terminology and notations. Let \( z \in \mathbb{R} \), \( \varepsilon > 0 \) and \( f, g : (z - \varepsilon, z + \varepsilon) \setminus \{z\} \to \mathbb{R} \) be two real functions defined in a deleted neighborhood of \( z \). We say that \( f \) and \( g \) are of the same order, as \( x \to z \), and write

\[
f(x) \sim g(x), \quad x \to z,
\]  

(23)

if \( \limsup_{x \to z} \left| \frac{f(x)}{g(x)} \right| < \infty \) and \( \liminf_{x \to z} \left| \frac{f(x)}{g(x)} \right| > 0 \) (with the convention \( \frac{0}{0} := 1 \)). We also use the same notation \( \sim \) in the case when \( f \) and \( g \) are only defined in a one-sided neighborhood of \( z \), i.e., \((z, z + \varepsilon) \) or \((z - \varepsilon, z) \).

**Remark 2.4.** We fix \( y \in I^\circ \) and discuss straightforward consequences of (22) is several specific situations.

(a) If the speed measure has an atom at \( y \), i.e., \( m(\{y\}) > 0 \), then we get \( \hat{a}_h(y) \sim h \), \( h \to 0 \).

(b) In the setting of Example \( \text{Example 1.1} \) with \( \eta(x) \sim 1 \), \( x \to y \), we have \( \hat{a}_h(y) \sim \sqrt{h}, h \to 0 \).

(c) More generally, in the setting of Example \( \text{Example 1.1} \) with \( \eta(x) \sim |x - y|^\alpha \), \( x \to y \), for some \( \alpha \in (-\infty, \frac{1}{2}) \) (this restriction on \( \alpha \) is to ensure (6)), we obtain \( \hat{a}_h(y) \sim h^{\frac{1}{2 - \alpha}} \), \( h \to 0 \).

It is worth noting that, by varying \( \alpha \in (-\infty, \frac{1}{2}) \) in (c), we can obtain orders \( h^\beta \) for all \( \beta \in (0, 1) \).
At this point, it is instructive to compare the EMCEL scheme and the Euler scheme. The latter is defined only in the SDE case. More precisely, in the setting of Example 1.1, we define the Euler scale factors $a_h^{Eu}(y) = \eta(y) \sqrt{h}$, $y \in I^o$, $h \in (0, [\bar{R}])$ (and, to treat jumping out of the state space, we extend the functions $a_h^{Eu}$, $h \in (0, [\bar{R}])$, to be zero in $\mathbb{R} \setminus I^o$). The (linearly interpolated, weak) Euler scheme $X_{Eu,h} = (X_{Eu,h}^t)_{t \in [0,\infty)}$ is defined through $a_h^{Eu}$ in the same way as the EMCEL scheme $\hat{X}^h$ is defined through $\hat{a}_h$ in (10)–(11). The properties of the Euler scheme are thus encoded in the Euler scale factors $a_h^{Eu}$, $h \in (0, [\bar{R}])$, and, by [5], for any $y \in I^o$, we have $a_h^{Eu}(y) \sim \sqrt{h}$, $h \to 0$. This is like what we have for the EMCEL scale factors in Remark 2.4 (b) and different from what we have in Remark 2.4 (c). While the EMCEL scheme always converges (Proposition 1.3), the Euler scheme can fail to converge when $\eta$ is irregular. In Example 5.4 of [3] it is proved that the Euler scheme does not converge (even weakly) in the case $\eta(x) = \frac{1}{\lambda(y)}1_{\mathbb{R}\setminus\{0\}} + 1_{\{0\}}(x)$, $x \in I := \mathbb{R}$. Contrary to the Euler scheme, for the EMCEL scheme, in the latter case we have $\hat{a}_h(0) \sim h^{1/4}$, $h \to 0$ (Remark 2.4 (c)).

Next let $y \in I^o$ and assume that $\hat{a}_h(y) \sim h^\beta$, $h \to 0$, for some $\beta > 0$. A natural question is then what is the limit of $\frac{\hat{a}_h(y)}{h^\beta}$ as $h \to 0$ (and if it exists at all). The claims in Remark 2.4 do not say anything about this. We now provide several more precise statements of this kind (in particular, improving the claims in Remark 2.4).

**Corollary 2.5.** Let $y \in I^o$ and $m(\{y\}) > 0$. Then $\lim_{h \to 0} \frac{\hat{a}_h(y)}{h} = \frac{2}{m(\{y\})}$.

**Proof.** This result still follows from Lemma 2.3. In more detail, the second inequality in (21) implies

$$\liminf_{h \to 0} \frac{\hat{a}_h(y)}{h} \geq \frac{2}{m(\{y\})}.$$ 

Next, the first inequality in (21) yields

$$\limsup_{h \to 0} \frac{\hat{a}_h(y)}{h} \leq \limsup_{h \to 0} \frac{2}{m(\{y\})} \leq \sup_{\lambda \in [0,1]} \{ (1-\lambda)m([y - \lambda \hat{a}_h(y), y + \lambda \hat{a}_h(y)]) \} \leq \frac{2}{m(\{y\})} = \frac{2}{m(\{y\})}.$$

This completes the proof. \qed

Below in Propositions 2.6 and 2.7 we discuss the case, where the speed measure $m$ has the following structure in a neighborhood of some point $y \in I^o$: there is $\alpha \in (-\infty, \frac{1}{2})$

---

2We do not compare the EMCEL and the Euler schemes in the situation of Remark 2.4 (a) because the latter falls out of the SDE case, and hence the Euler scheme is not defined.

3In this connection, it is worth mentioning that Theorem 2.1 in [11] states an equivalent condition (and Theorem 2.2 there provides a sufficient condition) in a setting with possibly discontinuous $\eta$ for the Euler scheme to converge weakly to the diffusion $Y$, provided uniqueness in law holds for the SDE (which we have in Example 1.1 due to the Engelbert-Schmidt conditions) and $\eta$ is locally bounded and has at most linear growth. Example 5.4 of [3] does not fall into the setting of [11] (and, indeed, the Euler scheme fails to converge) because, in that example, $\eta$ is not locally bounded.
and a non-vanishing Borel function \( \varphi \) such that
\[
m(dx) = \frac{2}{\varphi^2(x)} |x - y|^{-2\alpha} \, dx \quad \text{in some neighborhood of } y. \tag{24}
\]

On the one hand, this allows to improve the claims in Remark 2.4 (b) and (c). On the other hand, this allows to complement the above comparison with the Euler scheme (see Remark 2.8).

We first need the following notation. Let \( z \in \mathbb{R}, \varepsilon > 0 \) and \( f: (z - \varepsilon, z + \varepsilon) \setminus \{z\} \to \mathbb{R} \) be a real function defined in a deleted neighborhood of \( z \). We set
\[
|f|^*(z) = \limsup_{x \to z} |f(x)|,
|f|_*(z) = \liminf_{x \to z} |f(x)|.
\]

**Proposition 2.6.** Let \( y \in I^\circ \). Assume that there exist \( \alpha \in (-\infty, \frac{1}{2}) \) and a non-vanishing Borel function \( \varphi \) such that the speed measure \( m \) has structure (24). Then it holds
\[
[(1 - 2\alpha)(1 - \alpha)]^{\frac{1}{1 - \alpha}} |\varphi|_*(y) \frac{1}{(1 + \varepsilon)} \leq \liminf_{h \to 0} \frac{\bar{a}_h(y)}{h^{\frac{2\alpha}{1 - \alpha}}} \leq \limsup_{h \to 0} \frac{\bar{a}_h(y)}{h^{\frac{2\alpha}{1 - \alpha}}} \leq [(1 - 2\alpha)(1 - \alpha)]^{\frac{1}{1 - \alpha}} |\varphi|^*(y) \frac{1}{(1 + \varepsilon)}. \tag{25}
\]

**Proof.** We prove only the last inequality in (25). The first one is proved in a similar way. If \(|\varphi|_*(y) = \infty\), then there is nothing to prove. Below we assume that \(|\varphi|^*(y) < \infty\). Choose \( h_0 \in (0, \bar{h}) \) such that \( y \in (l_{h_0}, r_{h_0}) \) and \((y - \hat{a}_{h_0}(y), y + \hat{a}_{h_0}(y))\) is included in the neighborhood, where (24) holds. Remark 1.2 and Proposition 2.1 imply that, for all \( h \in (0, h_0) \), we have
\[
\int_{y - \hat{a}_h(y)}^{y + \hat{a}_h(y)} (\hat{a}_h(y) - |u - y|) |u - y|^{-2\alpha} \frac{2}{\varphi^2(u)} \, du = 2 h. \tag{26}
\]
Consider an arbitrary \( \varepsilon > 0 \). Then choose \( h_1 \in (0, h_0) \) such that
\[
|\varphi(u)| \leq (1 + \varepsilon)|\varphi|^*(y) \quad \text{for all } u \in (y - \hat{a}_{h_1}(y), y + \hat{a}_{h_1}(y)) \tag{27}
\]
(this is possible due to Proposition 2.2). It follows from (26) and (27) that
\[
\frac{2}{(1 + \varepsilon)^2 |\varphi|^*(y)^2} \int_{y - \hat{a}_h(y)}^{y + \hat{a}_h(y)} (\hat{a}_h(y) - |u - y|) |u - y|^{-2\alpha} \, du \leq 2 h \quad \text{for all } h \in (0, h_1).
\]
The integral is explicitly computable, and we get
\[
\frac{2}{(1 + \varepsilon)^2 |\varphi|^*(y)^2} \frac{\bar{a}_h(y)^{2 - 2\alpha}}{(1 - 2\alpha)(1 - \alpha)} \leq 2 h \quad \text{for all } h \in (0, h_1),
\]
hence
\[
\limsup_{h \to 0} \frac{\bar{a}_h(y)}{h^{\frac{2\alpha}{1 - \alpha}}} \leq [(1 - 2\alpha)(1 - \alpha)]^{\frac{1}{1 - \alpha}} |\varphi|^*(y)(1 + \varepsilon) \frac{1}{1 - \alpha}.
\]
As \( \varepsilon > 0 \) is arbitrary, we obtain the result. \( \square \)
The next result provides a sufficient condition for the \( \liminf \) and \( \limsup \) in Proposition 2.6 to coincide.

**Proposition 2.7.** Let \( y \in \Gamma^0 \). Assume that there exist \( \alpha \in (-\infty, \frac{1}{2}) \) and a non-vanishing Borel function \( \varphi \) such that the speed measure \( m \) has structure \([24] \). Further assume that

\[
\lim_{x \searrow y} |\varphi(x)| = |\varphi|(y-) \in (0, \infty) \quad \text{and} \quad \lim_{x \nearrow y} |\varphi(x)| = |\varphi|(y+) \in (0, \infty).
\]

Then

\[
\lim_{h \to 0} \frac{\hat{a}_h(y)}{h^{2-2\alpha}} = \left[ \frac{(1 - 2\alpha)(2 - 2\alpha)}{\frac{1}{|\varphi|^2(y-)} + \frac{1}{|\varphi|^2(y+)}} \right]^{\frac{1}{2-2\alpha}}.
\]

**Proof.** A formal proof is obtained along the lines of the proof of Proposition 2.6. We only show some technical steps that need to be elaborated differently. For notational convenience we set \( c_1 = |\varphi|(y-) \) and \( c_2 = |\varphi|(y+) \). Consider an arbitrary \( \varepsilon \in (0, 1) \). By \([28] \), for sufficiently small \( a > 0 \), we have

\[
\int_y^{y+a} (a - |u - y|)|u - y|^{-2\alpha} \frac{2}{\varphi^2(u)} \, du \in [B_{\text{low}}, B_{\text{up}}],
\]

where

\[
B_{\text{low}} = \frac{a^{2-2\alpha}}{(1 - 2\alpha)(2 - 2\alpha)} \left( \frac{1}{c_1^2} + \frac{2}{c_2^2} \right)(1 - \varepsilon),
\]

\[
B_{\text{up}} = \frac{a^{2-2\alpha}}{(1 - 2\alpha)(2 - 2\alpha)} \left( \frac{2}{c_1^2} + \varepsilon \right).
\]

It is worth noting that \( \varepsilon \) appears in the lower and upper bounds in a non-symmetric way because we need nonnegative bounds and we need to include the possibility \( c_2 = \infty \). With similar bounds for the integral from \( y - a \) to \( y \), we obtain, for sufficiently small \( h > 0 \),

\[
\frac{\hat{a}_h(y)^{2-2\alpha}}{(1 - 2\alpha)(2 - 2\alpha)} \left[ \frac{2}{c_1^2} + \frac{2}{c_2^2} \right](1 - \varepsilon) \leq 2h \leq \frac{\hat{a}_h(y)^{2-2\alpha}}{(1 - 2\alpha)(2 - 2\alpha)} \left[ \frac{2}{c_1^2} + \frac{2}{c_2^2} + 2\varepsilon \right].
\]

On the one hand, \([29] \) yields

\[
\liminf_{h \to 0} \frac{\hat{a}_h(y)}{h^{2-2\alpha}} \geq \left[ \frac{(1 - 2\alpha)(2 - 2\alpha)}{\frac{1}{c_1^2} + \frac{1}{c_2^2} + \varepsilon} \right]^{\frac{1}{2-2\alpha}},
\]

which already implies the result in the case \( c_1 = c_2 = \infty \), as \( \varepsilon \in (0, 1) \) is arbitrary. In the case \( \min\{c_1, c_2\} < \infty \), we also get from \([29] \)

\[
\limsup_{h \to 0} \frac{\hat{a}_h(y)}{h^{2-2\alpha}} \leq \left[ \frac{(1 - 2\alpha)(2 - 2\alpha)}{\frac{1}{c_1^2} + \frac{1}{c_2^2}} \right]^{\frac{1}{2-2\alpha}} \left( \frac{1}{1 - \varepsilon} \right)^{\frac{1}{2-2\alpha}},
\]

which, together with \([30] \), concludes the proof. \( \square \)
Remark 2.8. Let $y \in I^\circ$. In the setting of Example 1.1 Proposition 2.6 implies

$$|\eta|_*(y) \leq \liminf_{h \to 0} \frac{\hat{a}_h(y)}{\sqrt{h}} \leq \limsup_{h \to 0} \frac{\hat{a}_h(y)}{\sqrt{h}} \leq |\eta|^*(y)$$

(because (24) holds with $\eta$ in place of $\phi$ and $\alpha = 0$). In particular, if $\eta$ is continuous at point $y$, then

$$\lim_{h \to 0} \frac{\hat{a}_h(y)}{\sqrt{h}} = |\eta|(y),$$

which has a clear interpretation that, for small $h > 0$, the EMCEL scheme is close to the Euler one at points, where $\eta$ is continuous.

Furthermore, if in the setting of Example 1.1 we have (28) with $\eta$ in place of $\phi$, then Proposition 2.7 improves (31) by establishing

$$\lim_{h \to 0} \frac{\hat{a}_h(y)}{\sqrt{h}} = \sqrt{\frac{2}{\frac{1}{|\eta|^2(y^-)} + \frac{1}{|\eta|^2(y^+)}}},$$

(32)

i.e., the limit $\lim_{h \to 0} \frac{\hat{a}_h(y)}{\sqrt{h}}$ is equal to the power mean with exponent $-2$ of the left and the right limits of $|\eta|$ at $y$.

To illustrate this observation consider the SDE $dY_t = \sigma(Y_t)dW_t$ with periodic diffusion coefficient

$$\sigma(y) = \begin{cases} 1 & \text{if } y \in \cup_{k \in \mathbb{Z}}[2k, 2k + 1), \\ 2 & \text{if } y \in \cup_{k \in \mathbb{Z}}[2k + 1, 2k + 2). \end{cases}$$

The SDE describes a diffusion in a medium that is periodically compounded with two types of layers. Equation (32) implies that

$$\lim_{h \to 0} \frac{\hat{a}_h(y)}{\sqrt{h}} = \begin{cases} 1 & \text{if } y \in \cup_{k \in \mathbb{Z}}(2k, 2k + 1), \\ \sqrt{\frac{8}{5}} & \text{if } y \in \mathbb{Z}, \\ 2 & \text{if } y \in \cup_{k \in \mathbb{Z}}(2k + 1, 2k + 2) \end{cases}$$

with no need to actually compute the EMCEL scale factors. We remark that the scale factor of $\sqrt{8/5}$ at the boundaries of the layers also appears in the diffusion’s homogenization limit: the distribution of the solution $Y^\epsilon$ of the SDE $dY_t = \sigma(\frac{Y_t}{\epsilon})dW_t$, $Y_0 = 0$, converges, as $\epsilon \downarrow 0$, to the distribution of a BM scaled by $\sqrt{8/5}$ (see [10] for an introduction into the homogenization theory for periodic SDEs).

2.2 Dependence on the state

Next we discuss properties from category (ii) of the introduction. Specifically, here we examine how the EMCEL scale factors depend on the state variable, i.e., we study the functions $I \ni y \mapsto \hat{a}_h(y)$, and observe that the results translate into some good stability features of the scheme.
Theorem 2.9 (Comparison principle). For every \( h \in (0, h] \) and \( z \in \{-1, 1\} \), the mapping \( y \mapsto y + \hat{a}_h(y)z \) is nondecreasing on \( I \).

Let us discuss the meaning of this result. If at some time \( kh \) the EMCEL scheme at time \( t \) is in position \( y \in I \), then it will be either in \( y + \hat{a}_h(y) \) or in \( y - \hat{a}_h(y) \) at time \((k + 1)h\). Consider two points \( y_1 < y_2 \) in \( I \). Theorem 2.9 suggests to compare two situations, where at time \( kh \) we are in \( y_1 \) (the 1st situation) or in \( y_2 \) (the 2nd one), and asserts that, if we use the same realized \( \xi_{k+1} \) in both situations (recall \( (10) \)), then, in the 1st situation, we end up in the smaller position at time \((k + 1)h\) than in the 2nd situation. Notice that this property need not hold for the Euler scheme\(^4\), which may result in a kind of “diverging oscillations” in the Euler scheme.

We quote Figure 1 from \([3]^6\) as an example of what can happen when the SDE coefficients are of superlinear growth (also see Theorem 2.1 in \([22]\) for a related quantitative statement regarding the Euler scheme). More precisely, Figure 1 considers the numerical performance of the weak Euler scheme\(^5\) and the EMCEL scheme for the SDE \( dY_t = \cosh(Y_t)\,dW_t \). To understand why for fixed time step \( h > 0 \) the weak Euler scheme \( X_{Eu,h} \) exhibits “diverging oscillations” in this example assume that at some time \( kh \) the scheme has reached a point \( y > 0 \) large enough so that \( 2y/\cosh(y) < \sqrt{h} \). If we next have an upward jump (i.e., \( \xi_{k+1} = 1 \)) then clearly \( X_{Eu,1} > y \). In the other case where we next have a downward jump (i.e., \( \xi_{k+1} = -1 \)) the condition \( 2y/\cosh(y) < \sqrt{h} \) entails that \( X_{Eu,1} < -y \). So with probability one we have that the absolute value of the Euler scheme at time \((k + 1)h\) is bigger than its absolute value at time \( kh \). In fact, the exponential growth of \( \cosh \) entails that almost surely \( \{k, k + 1, \ldots \} \ni n \mapsto |X_{Eu,h}^{nkh}| \in \mathbb{R} \) increases super-exponentially, which ultimately leads to the “diverging oscillations”. Let us next justify why Theorem 2.9 ensures that such explosions cannot happen for the EMCEL scheme. To this end suppose again that at some time \( kh \) the scheme has reached a high value \( y > 0 \). If there is a streak of further subsequent upward jumps, then Theorem 2.9 ensures that the size of each jump \( \hat{a}_h(\hat{X}_l) \), \( l \geq k \), is at most linear in \( \hat{X}_l \) (cf. also Corollary 2.11 below). Eventually, at some time \( nh \) there will be a downward jump. In this case Theorem 2.9 ensures that \( \hat{X}_{(n+1)h} = \sup_{y \leq \hat{X}_{kh}} (y - \hat{a}_h(y)) \geq -\hat{a}_h(0) \) so that the EMCEL scheme opposed to the Euler scheme does not overjump a large region around 0 but rather jumps back into a “stable” region around 0 (in the setting of the bottom figure of Figure 1 a downward jump from any level \( y \geq 3 \) leads to a value close to 1 after the jump).

It is instructive to discuss relations and differences between the comparison principle of Theorem 2.9 and comparison theorems for solutions of SDEs (see, e.g., Theorem 1.4 in \([32]\) for Itô SDEs and Theorem 4.2 in \([9]\) for Stratonovich ones):

\(^4\)When we speak about the Euler scheme, we restrict ourselves to the SDE case.

\(^5\)We remark that Theorem 2.9, which holds for all possible speed measures \( m \), is a much stronger result than the comparison principle in \([3]\) and that the proof in \([3]\) is based on the implicit function theorem, which cannot be used for all all possible speed measures \( m \), as the involved functions, in particular, \( \hat{a}_h \), are, in general, not in \( C^1 \) (the latter claim follows from Theorem 2.14 below). Thus, the the proof of Theorem 2.9 uses ideas that are not present in \([3]\).

\(^6\)See the text preceding Corollary 2.5 for the definition of the weak Euler scheme.

\(^7\)The same reasoning applies to the case \( y < 0 \) with straightforward modifications.
Figure 1: The figure on top shows two realizations of discrete approximations to the SDE $dY_t = \cosh(Y_t) \, dW_t$ with $Y_0 = 0$. The dashed line depicts the realization based on the EMCEL scheme. The crosses show the realization obtained with the Euler scheme. Both use the same realized increments $(\xi_k)$. 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 bottom figure the solid and dashed lines are the graphs of the functions $y \mapsto y - \hat{a}_h(y)$ and $y \mapsto y + \hat{a}_h(y)$. The dash-dotted line indicates level zero. As outlined above the monotonicity of both functions implies that divergent oscillations are impossible in the EMCEL scheme.
• Comparison theorems for SDEs are pathwise results, and they apply in the situations when the SDE has the pathwise uniqueness property. And this makes perfect sense, as if an SDE that has a solution does not satisfy pathwise uniqueness, then one can find two different solutions to it with the same driving Brownian motion. That is, there are no such pathwise comparison results beyond the case of pathwise uniqueness. It is worth noting that pathwise uniqueness can fail even in the situation of Example 1.1 even with a continuous \( \eta \) (see [7]).

• The comparison principle of Theorem 2.9 is a pathwise property of the EMCEL scheme only, which, as discussed above, translates into a good stability feature of the scheme. However, it does not imply any pathwise comparison result in the spirit of Theorem 1.4 in [32], as the EMCEL scheme approximates general diffusions only in the weak sense. On the other hand, Theorem 2.9 applies to the EMCEL approximation of every general diffusion \( Y \) and thus goes far beyond the SDE case under pathwise uniqueness (e.g., \( Y \) can be a solution to an SDE from [7], \( Y \) can have sticky features, etc.).

Proof of Theorem 2.9. For \( y \in I^o \) and \( a > 0 \) such that \( y \pm a \in I \), we use the notation

\[
G(y,a) = \int_{(y-a,y+a)} (a - |u-y|) m(du). \tag{33}
\]

We fix \( h \in (0,\bar{h}] \) and elements \( y_1 < y_2 \) of \( I \). Define \( a_i = \hat{a}_h(y_i), i = 1,2 \). We need to show that

\[
y_1 + a_1 \leq y_2 + a_2 \quad \text{and} \quad y_1 - a_1 \leq y_2 - a_2. \tag{34}
\]

This is clear whenever \( y_1 \in I \setminus I^o \) or \( y_2 \in I \setminus I^o \) (recall that, by construction, \( \hat{a}_h(l) = \hat{a}_h(r) = 0 \) and, for all \( y \in I^o \), we have \( y \pm \hat{a}_h(y) \in [l,r] \), see [9]). Below we, therefore, assume \( y_1, y_2 \in I^o \) and consider four cases.

1. Let both endpoints \( l \) and \( r \) be inaccessible. Then we have

\[
G(y_1,a_1) = 2h = G(y_2,a_2). \tag{35}
\]

The expression for \( G(y,a) \) on the right-hand side of (33) together with (35) imply that neither of the intervals \((y_1 - a_1, y_1 + a_1)\) and \((y_2 - a_2, y_2 + a_2)\) contains the other. This yields (34).

2. Let \( l \) be accessible and \( r \) inaccessible. We first show the second statement in (34). If \( y_1 - a_1 = l \), then the statement is clear, as, by construction, for all \( y \in I^o \), it holds that \( y - \hat{a}_h(y) \geq l \). If \( y_1 - a_1 > l \), then we have

\[
G(y_1,a_1) = 2h \geq G(y_2,a_2). \tag{36}
\]

If we now assume that \( y_1 - a_1 > y_2 - a_2 \), then the interval \((y_2 - a_2, y_2 + a_2)\) strictly contains the interval \((y_1 - a_1, y_1 + a_1)\). Notice that the integrand in (33) is strictly positive, the integrand corresponding to the bigger interval \((y_2 - a_2, y_2 + a_2)\) dominates
the one corresponding to the smaller interval \((y_1 - a_1, y_1 + a_1)\) on that interval and that \(m\) has full support by \([2]\). This implies \(G(y_2, a_2) > G(y_1, a_1)\) and hence contradicts \((36)\).

Next we show the first statement in \((34)\). In the case \(y_2 - a_2 = l\), the statement follows from the fact that \(y_1 - a_1 \geq l\) and hence \(a_1 < a_2\). If \(y_2 - a_2 > l\), then we have
\[
G(y_1, a_1) \leq 2h = G(y_2, a_2).
\]
Assume that \(y_1 + a_1 > y_2 + a_2\). Then the interval \((y_1 - a_1, y_1 + a_1)\) strictly contains the interval \((y_2 - a_2, y_2 + a_2)\), which, together with \((33)\), implies \(G(y_1, a_1) > G(y_2, a_2)\) and hence contradicts \((37)\).

3. The case, where \(l\) is inaccessible and \(r\) is accessible, is symmetric to case 2.

4. Let finally both \(l\) and \(r\) be accessible. We prove only the first statement in \((34)\), as the second one is symmetric. If \(y_2 + a_2 = r\), then the statement follows from the fact that \(y_1 + a_1 \leq r\). In the case \(y_2 - a_2 = l\), the statement follows from the fact that \(y_1 - a_1 \geq l\) and hence \(a_1 < a_2\). In the remaining case \(y_2 \pm a_2 \in I^0\), we have
\[
G(y_1, a_1) \leq 2h = G(y_2, a_2),
\]
and the argument after \((37)\) yields the desired statement.

\[\square\]

**Remark 2.10.** Analyzing the proof of Theorem 2.9 in more detail, we obtain that the following more precise version of the comparison principle holds true:

Let \(h \in (0, \overline{h}]\). Then the mapping
\[
y \mapsto y + \hat{a}_h(y) \text{ is strictly increasing on } (l, r_h) \text{ and constant on } [r_h, r),
\]
and the mapping
\[
y \mapsto y - \hat{a}_h(y) \text{ is constant on } (l, l_h] \text{ and strictly increasing on } (l_h, r).
\]

**Corollary 2.11** (Smoothing and tempered growth behavior). For any \(h \in (0, \overline{h}]\), the function \(I \ni y \mapsto \hat{a}_h(y)\) is Lipschitz continuous on \(I\) with Lipschitz constant 1, i.e.,
\[
|\hat{a}_h(y_1) - \hat{a}_h(y_2)| \leq |y_1 - y_2| \quad \text{for all } y_1, y_2 \in I \text{ and } h \in (0, \overline{h}].
\]
Moreover, there exists a constant \(C_0 \in [0, \infty)\) such that
\[
\hat{a}_h(y) \leq C_0 + |y| \quad \text{for all } y \in I \text{ and } h \in (0, \overline{h}].
\]

**Proof.** The first statement is an immediate consequence of the comparison principle. The second statement easily follows from the first one with, e.g., \(C_0 = |y_0| + \sup_{h \in (0, \overline{h}]} \hat{a}_h(y_0) = |y_0| + \hat{a}_\overline{h}(y_0) < \infty\) (recall Proposition 2.1), where \(y_0\) is an arbitrary point in \(I\).

\[\square\]

**Remark 2.12.** Corollary 2.11 is named as above to stress the difference with the Euler scheme in the SDE case, where the Euler scale factors \(a_{Eu}^h(y) = \eta(y)\sqrt{h}\), as functions of \(y\), inherit irregularities and the growth from \(\eta\). On the contrary, the EMCEL scale factors \(\hat{a}_h\) are as described in Corollary 2.11 no matter how irregular \(\eta\) is in the SDE case and also beyond the SDE case.
Corollary \[2.11\] provides the functional bound \( C_0 + |y| \) (independent of \( h \)) for all functions \( \hat{a}_h, h \in (0, \bar{h}] \). We also know that \( \hat{a}_h(y) \to 0, h \to 0 \), for all fixed \( y \in I \) (recall Proposition \[2.2\] and \( \hat{a}_h(l) = \hat{a}_h(r) = 0 \)). A natural question is now to find a functional bound for \( \hat{a}_h, h \in (0, \bar{h}] \), that depends on \( h \) and vanishes as \( h \to 0 \). However, this does not seem to be feasible in general, as the order of convergence (in \( h \)) of \( \hat{a}_h(y) \) to zero can be different in different points \( y \). The discussion following Remark \[2.4\] suggests that the precise forms of such functional bounds have to depend on the structure of the speed measure \( m \). We, finally, present a result of such kind.

**Proposition 2.13.** Suppose that we have

\[
m(dx) \geq \frac{dx}{g(|x|)} \quad \text{on } I^c
\]

(understood in the integral form) with some positive nondecreasing function \( g : [0, \infty) \to (0, \infty) \). Then, with any constant \( C_0 \in [0, \infty) \) satisfying \( (39) \), we obtain

\[
\hat{a}_h(y) \leq \sqrt{2g(C_0 + 2|y|)h} \quad \text{for all } y \in I \text{ and } h \in (0, \bar{h}].
\]

We list a couple of specific functional bounds for \( \hat{a}_h \) implied by Proposition \[2.13\].

(a) Let \( (40) \) be satisfied with \( g(x) = c(1 + x^p), x \in [0, \infty) \), for some \( c, p \in (0, \infty) \). Then there exists \( C \in (0, \infty) \) such that

\[
\hat{a}_h(y) \leq C(1 + |y|^{p/2})\sqrt{h} \quad \text{for all } y \in I \text{ and } h \in (0, \bar{h}].
\]

(b) Let \( (40) \) be satisfied with \( g(x) = c \exp\{px\}, x \in [0, \infty) \), for some \( c, p \in (0, \infty) \). Then there exists \( C \in (0, \infty) \) such that

\[
\hat{a}_h(y) \leq C \exp\{|y|\} \sqrt{h} \quad \text{for all } y \in I \text{ and } h \in (0, \bar{h}].
\]

**Proof of Proposition 2.13.** It follows from \( (8) \) and \( (40) \) that, for all \( y \in I^c \) and \( h \in (0, \bar{h}] \), it holds

\[
2h \geq \int_{\{y-\hat{a}_h(y), y+\hat{a}_h(y)\}} (\hat{a}_h(y) - |u - y|) m(du)
\]

\[
\geq \int_{y-\hat{a}_h(y)}^{y+\hat{a}_h(y)} \frac{\hat{a}_h(y) - |u - y|}{g(|u|)} \, du = \hat{a}_h(y)^2 \int_{-1}^1 \frac{1 - |z|}{g(|y + z\hat{a}_h(y)|)} \, dz
\]

\[
\geq \frac{\hat{a}_h(y)^2}{\sup_{z \in [-1, 1]} g(|y + z\hat{a}_h(y)|)} = \frac{\hat{a}_h(y)^2}{g(|y| + |\hat{a}_h(y)|)}.
\]

The claim now follows from \( (39) \). \( \Box \)
2.3 ODE for the scale factors

We, finally, turn to properties from category iii of the introduction, i.e., the properties that help implementing the EMCEL scheme in specific situations. It follows from the discussion in Remark 1.2 that the main challenge in implementing the EMCEL scheme is to determine the scale factor \( \hat{\alpha}_h \) (for a fixed \( h \in (0, \tilde{h}] \)) inside \((l_h, r_h)\) because it requires to solve the nonlinear equation (13) for all \( y \in (l_h, r_h) \). On the contrary, there is no problem to determine \( \hat{\alpha}_h \) outside \((l_h, r_h)\) (recall (14)).

Theorem 2.14 below shows that \( \hat{\alpha}_h \) is a unique solution to ODE (44) inside \((l_h, r_h)\). Thus, in order to implement the scheme, it is enough to solve (13) numerically only for some \( y_0 \in (l_h, r_h) \) (not for all \( y \in (l_h, r_h) \)), which provides the initial condition for the ODE, and then to apply an appropriate ODE solver. It is worth noting that the ODE itself does not depend on the discretization parameter \( h \). Dependence on \( h \) comes into the picture through the initial condition (solving (13) for some \( y_0 \in (l_h, r_h) \)).

In fact, a solution to ODE (44) is understood in the sense that it is an absolutely continuous function satisfying (44) almost everywhere (with respect to the Lebesgue measure). In general, we cannot require (44) everywhere, as we treat possible speed measures (they can, e.g., have atoms). Essentially, the first part of Theorem 2.14 deals with existence and the second with uniqueness for ODE (44). The minimal requirement for the uniqueness is exactly the one mentioned above: the solution must be absolutely continuous and satisfy (44) almost everywhere. But in the existence part of Theorem 2.14 we provide more detail about what de facto holds for the EMCEL scale factor \( \hat{\alpha}_h \). For instance, it turns out that \( \hat{\alpha}_h \) is differentiable everywhere on \((l_h, r_h)\) except at most on a countable set, regardless of how “irregular” the speed measure \( m \) is.

Below we use the notation \( \mu_L \) for the Lebesgue measure.

**Theorem 2.14.** Let \( h \in (0, \tilde{h}] \).

(i) For all \( y \in (l_h, r_h) \) the right derivative \( \partial_+ \hat{\alpha}_h(y) = \lim_{\varepsilon \searrow 0} \frac{\hat{\alpha}_h(y + \varepsilon) - \hat{\alpha}_h(y)}{\varepsilon} \) and the left derivative \( \partial_- \hat{\alpha}_h(y) = \lim_{\varepsilon \nearrow 0} \frac{\hat{\alpha}_h(y) - \hat{\alpha}_h(y - \varepsilon)}{\varepsilon} \) exist and it holds that

\[
\partial_+ \hat{\alpha}_h(y) = \frac{m((y - \hat{\alpha}_h(y), y]) - m((y, y + \hat{\alpha}_h(y))]}{m((y - \hat{\alpha}_h(y), y + \hat{\alpha}_h(y)]})
\]

and

\[
\partial_- \hat{\alpha}_h(y) = \frac{m([y - \hat{\alpha}_h(y), y]) - m([y, y + \hat{\alpha}_h(y)])}{m([y - \hat{\alpha}_h(y), y + \hat{\alpha}_h(y)])}.
\]

Moreover, \( \partial_+ \hat{\alpha}_h \) is càdlàg, \( \partial_- \hat{\alpha}_h \) is càglàd and, for all \( y \in (l_h, r_h) \), we have \( \partial_+ \hat{\alpha}_h(y) - := \lim_{\varepsilon \searrow 0} \partial_+ \hat{\alpha}_h(y - \varepsilon) = \partial_- \hat{\alpha}_h(y) \). In particular, there exists an at most countable set \( \mathcal{N} \subset (l_h, r_h) \) such that \( \hat{\alpha}_h \) exists on \((l_h, r_h) \setminus \mathcal{N} \) and for all \( y \in (l_h, r_h) \setminus \mathcal{N} \) we have \( \hat{\alpha}_h(y) = \partial_+ \hat{\alpha}_h(y) = \partial_- \hat{\alpha}_h(y) \).

(ii) Let \( y_0 \in (l_h, r_h) \) and let \( a: (l_h, r_h) \to (0, \infty) \) be an absolutely continuous function on compact subintervals of \((l_h, r_h)\) that satisfies, for \( \mu_L \)-almost all \( y \in (l_h, r_h) \),

\[
a'(y) = F(y, a(y))
\]
and \( a(y_0) = \hat{a}_h(y_0) \). Here, the function \( F: (l_h, r_h) \times (0, \infty) \to \mathbb{R} \) is defined by the formulas

\[
F(y, a) = \frac{m((y - a, y]) - m((y, y + a]), \quad (y, a) \in \mathcal{G},
\]

\[
F(y, a) = 0, \quad (y, a) \in (l_h, r_h) \times (0, \infty) \setminus \mathcal{G},
\]

where the domain \( \mathcal{G} \) is defined as

\[
\mathcal{G} = \{(y, a) \in (l_h, r_h) \times (0, \infty) : y \pm a \in I^o\}.
\]

Then we have \( a(y) = \hat{a}_h(y) \) for all \( y \in (l_h, r_h) \).

To explain ODE (44) in more detail we now make several comments:

- For \( y \in (l_h, r_h) \), we have \( y \pm \hat{a}_h(y) \in I^o \), i.e., \((y, \hat{a}_h(y)) \in \mathcal{G} \). In other words, for the EMCEL scale factor, the right-hand side of (44) reduces to the right-hand side of (45) only, i.e., (46) is not needed.

- In (46) we extend the function \( F \) beyond the domain \( \mathcal{G} \) only because we need a well-defined right-hand side of (44) for functions \( a \) that \textit{a priori} need not coincide with the EMCEL scale factor.\(^8\)

- On the other hand, it will become clear from the proof that it does not matter how to define \( F \) outside \( \mathcal{G} \). In this sense, (46) is not important.

The proof of Theorem 2.14 is based on the following lemma.

**Lemma 2.15.** Fix \( h \in (0, \bar{h}] \). Let \( a: (l_h, r_h) \to (0, \infty) \) be a function such that, for all \( y_1, y_2 \in (l_h, r_h) \), it holds \( y_1 \pm a(y_1) \in I^o \) and \( |a(y_2) - a(y_1)| \leq |y_2 - y_1| \). Define the function \( H: (l_h, r_h) \to (0, \infty) \) by the formula \( H(y) = \int_{(y-a(y), y+a(y))} (a(y) - |u - y|) m(du) \). Then \( H \) is locally Lipschitz continuous and, for all \( y_1 < y_2 \) in \((l_h, r_h)\) sufficiently close to each other,\(^7\) we have the representation

\[
H(y_2) - H(y_1) = (a(y_2) - a(y_1))m([y_2 - a(y_2), y_1 + a(y_1)])
\]

\[
+ (y_2 - y_1)(m([y_2, y_1 + a(y_1)]) - m([y_2 - a(y_2), y_1])) + R(y_1, y_2),
\]

where, for the remainder term \( R(y_1, y_2) \) in (48), it holds

\[
R(y, y + \Delta y) \in o(\Delta y), \quad \Delta y \searrow 0,
\]

\[
R(y - \Delta y, y) \in o(\Delta y), \quad \Delta y \searrow 0,
\]

for all \( y \in (l_h, r_h) \).

\(^8\)Notice that the right-hand side of (45) can fail to be well-defined outside \( \mathcal{G} \) because the speed measure \( m \) can be infinite near the boundary points of \( I \); cf. (2).

\(^7\)In the sense \( y_2 - a(y_2) \leq y_1 < y_2 \leq y_1 + a(y_1) \), which is needed for the right ordering between the endpoints of the intervals involved in (48).
Proof of Theorem 2.14. (i) First note that it follows from Corollary 2.11 that \( \hat{a}_h \) is Lipschitz continuous on \((l_h,r_h)\) with Lipschitz constant 1. Therefore we are in a position to apply Lemma 2.15 with \( a = \hat{a}_h \) in the notation of Lemma 2.15 and we denote by \( \hat{H} \) the associated function \( H \). It follows from Remark 1.2 that for all \( y \in (l_h,r_h) \) we have \( \hat{H}(y) = 2h \). Therefore, we obtain from (48) and (49) that, for all \( y \in (l_h,r_h) \) and \( \Delta y > 0 \),

\[
\frac{\hat{a}_h(y + \Delta y) - \hat{a}_h(y)}{\Delta y} \to m(\{y + \Delta y - \hat{a}_h(y + \Delta y), y\}) - m(\{y + \Delta y, y + \hat{a}_h(y)\}) + o(1), \quad \Delta y \searrow 0.
\]

This yields (42) (recall that the functions \( y \mapsto y \pm \hat{a}_h(y) \) are strictly increasing on \((l_h,r_h)\) by Remark 2.10). In a similar way, using (50) instead of (49), we obtain (43). The further claims in (i) follow from these two formulas.

(ii) We first prove the claim under the additional assumption \( y \pm a(y) \in \mathcal{I}^o \) for all \( y \in (l_h,r_h) \) (notice that the EMCEL scale factor \( \hat{a}_h \) satisfies this assumption). It follows from (41) that for \( \mu_L \)-almost all \( y \in (l_h,r_h) \) we have \( |a'(y)| < 1 \) (recall (2)). As \( a \) is absolutely continuous, for \( y_1 < y_2 \) in \((l_h,r_h)\), we have

\[
|a(y_2) - a(y_1)| = \left| \int_{y_1}^{y_2} a'(y) \, dy \right| \leq \int_{y_1}^{y_2} |a'(y)| \, dy < y_2 - y_1
\]

and, in particular, the functions \( y \mapsto y \pm a(y) \) are strictly increasing on \((l_h,r_h)\). Notice that the latter implies that the set

\[
\{y \in (l_h,r_h) : m(\{y - a(y)\} \cup \{y\} \cup \{y + a(y)\}) > 0\}
\]

is at most countable, hence \( \mu_L \)-negligible. Due to (52) we can apply Lemma 2.15. We obtain that the function \( H \), as defined in Lemma 2.15, is locally Lipschitz continuous and hence absolutely continuous on every compact subinterval of \((l_h,r_h)\). Moreover, it follows from (48), (49), (50) and the fact that the set in (53) is \( \mu_L \)-negligible that, for \( \mu_L \)-almost all \( y \in (l_h,r_h) \), we have

\[
H'(y) = a'(y)m((y - a(y), y + a(y))) + m((y, y + a(y))) - m((y - a(y), y)) = 0.
\]

Consequently, \( H \) is constant on \((l_h,r_h)\) and it follows for all \( y \in (l_h,r_h) \) that

\[
\int_{(y-a(y),y+a(y))} (a(y) - |u - y|) \, m(du) = H(y) = H(y_0) = 2h.
\]

The claim that \( a \) and \( \hat{a}_h \) are identical now follows from Remark 1.2.

It remains to drop the assumption \( y \pm a(y) \in \mathcal{I}^o \) for all \( y \in (l_h,r_h) \). As \( a(y_0) = \hat{a}_h(y_0) \), it holds \( y_0 \pm a(y_0) \in \mathcal{I}^o \). By the continuity of \( a \), we get \( y \pm a(y) \in \mathcal{I}^o \) in a sufficiently small neighborhood of \( y_0 \). Hence, by the considerations above, \( a \) and \( \hat{a}_h \) coincide in this neighborhood of \( y_0 \). Define

\[
\tilde{t} = \inf\{y \in (l_h,y_0) : a(z) = \hat{a}_h(z) \text{ for all } z \in [y,y_0]\} \quad (\in [l_h,y_0)),
\]

\[
\tilde{r} = \sup\{y \in [y_0,r_h) : a(z) = \hat{a}_h(z) \text{ for all } z \in [y_0,y]\} \quad (\in (y_0,r_h]).
\]
If $\bar{t} > l_h$, we repeat the preceding argumentation with $y_0$ replaced by $\bar{t}$ and conclude that $a$ and $\tilde{a}$ coincide in some neighborhood of $\bar{t}$, which contradicts the definition of $\bar{t}$. Thus, $\bar{t} = l_h$. Similarly, $\bar{r} = r_h$. This completes the proof.

**Proof of Lemma 2.15** Throughout the proof we work with various choices of $y_1 < y_2$ in $(l_h, r_h)$. Note that if $y_2 - y_1$ is small enough we have that $y_2 - a(y_2) \leq y_1 < y_2 \leq y_1 + a(y_1)$. Moreover, it follows from the assumption that $a$ is Lipschitz continuous on $(l_h, r_h)$ with Lipschitz constant 1 that $y_1 - a(y_1) \leq y_2 - a(y_2)$ and $y_1 + a(y_1) \leq y_2 + a(y_2)$. To summarize, we have for $y_2 - y_1$ small enough that

$$y_1 - a(y_1) \leq y_2 - a(y_2) \leq y_1 < y_2 \leq y_1 + a(y_1) \leq y_2 + a(y_2).$$

Therefore, we have that

$$H(y_2) - H(y_1) = -\int_{(y_1-a(y_1),y_2-a(y_2))} (a(y_1) - |u - y_1|) m(du)$$

$$+ \int_{[y_2-a(y_2),y_1+a(y_1)]} (a(y_2) - |u - y_2| - a(y_1) + |u - y_1|) m(du)$$

$$+ \int_{(y_1+a(y_1),y_2+a(y_2))} (a(y_2) - |u - y_2|) m(du)$$

$$= -\int_{(y_1-a(y_1),y_2-a(y_2))} (a(y_1) + u - y_1) m(du)$$

$$+ (a(y_2) - a(y_1))m([y_2 - a(y_2), y_1 + a(y_1)])$$

$$+ \int_{[y_2-a(y_2),y_1+a(y_1)]} (|u - y_1| - |u - y_2|) m(du)$$

$$+ \int_{(y_1+a(y_1),y_2+a(y_2))} (a(y_2) - u + y_2) m(du).$$

For each fixed $y_1$ and moving $y_2$ such that $y_2 - y_1 \searrow 0$ we have

$$\left| \int_{(y_1-a(y_1),y_2-a(y_2))} (a(y_1) + u - y_1) m(du) \right| = \int_{(y_1-a(y_1),y_2-a(y_2))} (a(y_1) + u - y_1) m(du)$$

$$\leq (y_2 - y_1 + a(y_1) - a(y_2)) m((y_1 - a(y_1), y_2 - a(y_2)))$$

$$\leq 2(y_2 - y_1)m((y_1 - a(y_1), y_2 - a(y_2))) \in o(y_2 - y_1)$$

and similarly

$$\left| \int_{(y_1+a(y_1),y_2+a(y_2))} (a(y_2) - u + y_2) m(du) \right|$$

$$\leq 2(y_2 - y_1)m((y_1 + a(y_1), y_2 + a(y_2))) \in o(y_2 - y_1).$$

$$23$$
Moreover, it holds that

\[
\int_{[y_2-a(y_2),y_1+a(y_1)]} (|u-y_1|-|u-y_2|) \, m(du) \\
= (y_2 - y_1)(m([y_2, y_1 + a(y_1)]) - m([y_2 - a(y_2), y_1])) + \int_{(y_1,y_2)} (2u - y_1 - y_2) \, m(du)
\]

(57)

and that, again for a fixed \( y_1 \) and moving \( y_2 \) such that \( y_2 - y_1 \searrow 0 \),

\[
\left| \int_{(y_1,y_2)} (2u - y_1 - y_2) \, m(du) \right| \leq (y_2 - y_1)m((y_1, y_2)) \in o(y_2 - y_1).
\]

(58)

Combining (54)–(58) we obtain, for a fixed \( y_1 \) and moving \( y_2 \) such that \( y_2 - y_1 \searrow 0 \),

\[
H(y_2) - H(y_1) = (a(y_2) - a(y_1))m([y_2 - a(y_2), y_1 + a(y_1)]) \\
+ (y_2 - y_1)(m([y_2, y_1 + a(y_1)]) - m([y_2 - a(y_2), y_1])) + o(y_2 - y_1),
\]

(59)

which is (48) and (49). Property (50) follows from similar considerations, only with fixed \( y_2 \) and moving \( y_1 \) such that \( y_2 - y_1 \searrow 0 \). Moreover, the preceding calculations imply that, for \( y_2 - y_1 \) small enough,

\[
\frac{|H(y_2) - H(y_1)|}{y_2 - y_1} \leq m([y_2 - a(y_2), y_1 + a(y_1)]) \frac{|a(y_2) - a(y_1)|}{y_2 - y_1} \\
+ 2(m((y_1 - a(y_1), y_2 - a(y_2))) + m((y_1 + a(y_1), y_2 + a(y_2)))) \\
+ (m([y_2, y_1 + a(y_1)]) - m([y_2 - a(y_2), y_1])) + m((y_1, y_2)) \\
\leq 2m((y_1 - a(y_1), y_2 + a(y_2))) < \infty
\]

because, by the assumptions, \( y_1 - a(y_1), y_2 + a(y_2) \in I^\circ \) (also recall (2)). This implies that \( H \) is Lipschitz continuous on every compact subinterval of \((l_h, r_h)\). This completes the proof of Lemma 2.15.

We now mention the following immediate consequence of Theorem 2.14 for the case when the speed measure \( m \) does not have atoms in \( I^\circ \).

**Corollary 2.16.** Assume \( m(\{y\}) = 0 \) for any \( y \in I^\circ \). Then, for any \( h \in (0, \overline{h}] \), the restriction of the EMCEL scale factor \( \hat{\alpha}_h \) to \((l_h, r_h)\) is a \( C^1 \) function that coincides with the unique solution to the initial value problem

\[
a'(y) = F(y, a(y)), \quad y \in (l_h, r_h), \quad a(y_0) = \hat{\alpha}_h(y_0),
\]

(61)

for any fixed \( y_0 \in (l_h, r_h) \).

We stress that existence and uniqueness on \((l_h, r_h)\) for the initial value problem (61) are also claimed in Corollary 2.16. It is interesting to compare this with what we can get concerning the existence and uniqueness for (61) from standard results on ODEs:
The assumption \( m(\{y\}) = 0 \) for any \( y \in I^\circ \) of Corollary 2.16 implies that the function \( F \) is continuous in the domain \( G \) (recall (45) and (47)). Therefore, Peano’s existence theorem (e.g., see Theorem II.2.1 in [21] or Theorem 2.19 in [40]) yields that there is a solution to the initial value problem (61), and it can be extended up to a boundary of \( G \) (e.g., see Theorem II.3.1 in [21] or § 11 in [38]). This is weaker than the existence on the whole \((l_h,r_h)\) because the boundary of \( G \) can be achieved at points with \( y \)-coordinates that are strictly between \( l_h \) and \( r_h \).

Peano’s existence theorem does not say anything about uniqueness. Under the assumption of Corollary 2.16 only, the function \( a \mapsto F(y,a) \) (for fixed \( y \)) need not be locally Lipschitz inside \( G \); on the contrary, it can have a quite unpleasant behavior (e.g., think about the case when \( m \) is singular with respect to the Lebesgue measure). Therefore, standard results on ODEs do not provide uniqueness for (61).

This discussion raises the question of whether there is a numerical algorithm in order to compute the unique solution of the initial value problem (61). In what follows, we present such an algorithm. As in Corollary 2.16 we assume that \( m(\{y\}) = 0 \) for all \( y \in I^\circ \).

(a) \( F \) is continuous in \( G \).

(b) \( |F(y,a)| < 1 \) for all \((y,a) \in G\).

(c) Notice, however, that \( F \) is, in general, not extendable to a continuous function on \( \overline{G} \), where \( \overline{G} \) denotes the closure of \( G \) in \( \mathbb{R}^2 \).

By \( \partial G \) we denote the boundary of \( G \) in \( \mathbb{R}^2 \): \( \partial G = \overline{G} \setminus G \). It will be convenient to decompose \( \partial G \) as follows:

\[
\partial G = \partial G_{\text{low}} \cup \partial G_{\text{high}} \cup \partial G_{\text{vertical}},
\]

where

\[
\partial G_{\text{low}} = (l_h,r_h) \times \{0\},
\]

\[
\partial G_{\text{high}} = \{(y,a) \in (l_h,r_h) \times (0,\infty) : a = (y-l) \wedge (r-y)\},
\]

\[
\partial G_{\text{vertical}} = \partial G \setminus (\partial G_{\text{low}} \cup \partial G_{\text{high}})
\]

(cf. with (47)). To provide more detail, we remark that, if \( l = -r = -\infty \) (hence \( l_h = -r_h = -\infty \)), then \( G = \mathbb{R} \times (0,\infty) \), \( \partial G = \partial G_{\text{low}} \), \( \partial G_{\text{high}} = \partial G_{\text{vertical}} = \emptyset \); otherwise we also have non-empty \( \partial G_{\text{high}} \) and \( \partial G_{\text{vertical}} \). The most interesting case \(-\infty < l < r < +\infty\) is illustrated in Figure 2 where we have chosen \( l \) to be accessible and \( r \) inaccessible in order to show different subcases.

\[\text{E.g., the ODE } a'(y) = 2\sqrt{|a|}, \ a(0) = 0, \ \text{has different solutions } a \equiv 0 \text{ and } a(y) = y^2 \text{ sgn } y \text{ (and there are other ones). For illuminating examples of ODEs of the form } a'(y) = F(y,a(y)) \text{ with a continuous function } F \text{ in some region of the } (y,a)\text{-plane such that the initial value problem has more than one solution in any neighborhood of any initial point } (y_0,a_0) \text{ in that region, see [31] and/or Section II.5 in [21]}.\]
Figure 2: Form of the domain $\mathcal{G}$ and its boundary $\partial \mathcal{G}$ in the case $-\infty < l < r < +\infty$ with accessible $l$ and inaccessible $r$ (so that $l_h > l, r_h = r$).

Let us fix a subinterval $(\alpha, \beta)$ of $(l_h, r_h)$ with finite endpoints $\alpha$ and $\beta$ (we can take $\alpha = l_h$ if $l_h > -\infty$, and similarly with $\beta$). We now describe how to approximate the unique solution of the initial value problem (61), which is the EMCEL scale factor $y \mapsto \hat{a}(y)$, uniformly on $[\alpha, \beta]$. Let $\mathcal{Y}$ be a partition of $[\alpha, \beta]$ that contains $y_0$ from (61) as one of its elements, i.e., $\mathcal{Y} = \{y_j\}_{j=K(\mathcal{Y})}^{L(\mathcal{Y})}$ with some $K(\mathcal{Y}), L(\mathcal{Y}) \in \mathbb{Z}, K(\mathcal{Y}) < 0 < L(\mathcal{Y})$ and

$$\alpha = y_{K(\mathcal{Y})} < \cdots < y_{-1} < y_0 < y_1 < \cdots < y_{L(\mathcal{Y})} = \beta.$$ 

We define a piecewise linear function $y \mapsto a^{\mathcal{Y}}(y)$ on $[y_0, \beta]$ as follows.

1. Set $a^{\mathcal{Y}}(y_0) = \tilde{a}_h(y_0)$ and observe that $(y_0, a^{\mathcal{Y}}(y_0)) \in \mathcal{G}$. Set $j = 0$.

2. Define $\tilde{a}^{\mathcal{Y}}(y) = a^{\mathcal{Y}}(y_j) + (y - y_j)F(y_j, a^{\mathcal{Y}}(y_j))$ for all $y \in [y_j, y_{j+1}]$.

3. If $(y_{j+1}, a^{\mathcal{Y}}(y_{j+1})) \in \mathcal{G}$, then
   - set $a^{\mathcal{Y}} = \tilde{a}^{\mathcal{Y}}$ on $(y_j, y_{j+1}]$,
   - observe that, by convexity of $\mathcal{G}$, the graph of $a^{\mathcal{Y}}$ on $[y_j, y_{j+1}]$ lies inside $\mathcal{G}$,
   - set $j = j + 1$,
   - if $j < L(\mathcal{Y})$, then go to step 2 else go to step 5.

4. If $(y_{j+1}, a^{\mathcal{Y}}(y_{j+1})) \notin \mathcal{G}$, then
   - denote by $(\overline{y}, \overline{a})$ the point, where the graph of $\tilde{a}^{\mathcal{Y}}$ on $[y_j, y_{j+1}]$ intersects $\partial \mathcal{G}$,
   - if $\partial \mathcal{G}_{\text{vertical}}$ is intersected, which is only possible in the case $\overline{y} = \beta = r_h$, then set $a^{\mathcal{Y}} = \tilde{a}^{\mathcal{Y}}$ on $(y_j, \overline{y})$ ($\equiv (y_j, \beta]$) and go to step 5.
• if $\partial G_{\text{high}}$ (resp., $\partial G_{\text{low}}$) is intersected, then set $a^\gamma = \tilde{a}^\gamma$ on $(y_j, \bar{y})$, define $a^\gamma$ on $(\bar{y}, \beta)$ such that its graph goes along $\partial G_{\text{high}}$ on $(\bar{y}, \beta)$ (resp., set $a^\gamma \equiv 0$ on $(\bar{y}, \beta)$) and proceed with step 5.

5. We constructed the function $[y_0, \beta] \ni y \mapsto a^\gamma(y)$.

The function $a^\gamma$ is extended to $[\alpha, y_0]$ in a symmetric way. We thus obtain a piecewise linear function

$$[\alpha, \beta] \ni y \mapsto a^\gamma(y), \quad (y, a^\gamma(y)) \in \mathcal{G},$$

which is, in fact, nothing else but Euler’s polygonal approximation for (61) suitably extended to $[\alpha, \beta]$.

**Lemma 2.17.** Assume $m(\{y\}) = 0$ for all $y \in I^o$. Let $(\alpha, \beta)$ be a subinterval of $(l_h, r_h)$ with finite endpoints $\alpha$ and $\beta$ (we can take $\alpha = l_h$ if $l_h > -\infty$, and similarly with $\beta$). Then, for any sequence $\{Y_N\}_{N \in \mathbb{N}}$ of partitions of $[\alpha, \beta]$, $Y_N = \{y_j^{(N)}\}_{j=K_N(Y_N)}$, $K_N(Y_N) < 0 < L_N(Y_N)$,

$$\alpha = y_{K_N(Y_N)}^{(N)} < \cdots < y_{1}^{(N)} < y_0^{(N)} = y_0 < y_1^{(N)} < \cdots < y_{L_N(Y_N)}^{(N)} = \beta$$

with

$$|Y_N| := \max_{K_N(Y_N) < j \leq L_N(Y_N)} (y_j^{(N)} - y_{j-1}^{(N)}) \to 0, \quad N \to \infty,$$

we have

$$\sup_{y \in [\alpha, \beta]} |a^{\gamma_N}(y) - \tilde{a}_h(y)| \to 0, \quad N \to \infty,$$

i.e., the sequence $\{a^{\gamma_N}\}$ of (suitably extended) Euler’s polygonal approximations for (61) converges uniformly on $[\alpha, \beta]$ to the EMCEL scale factor $\tilde{a}_h$, which is the unique solution of (61).

**Proof.** The idea is very similar to that in the proof of Peano’s existence theorem, where it is shown that, on a sufficiently small interval containing $y_0$, the sequence $\{a^{\gamma_N}\}$ has a uniformly convergent subsequence that converges to a solution of (61). Peano’s existence theorem applies in this form due to fact (4) above and the fact that $(y_0, \tilde{a}_h(y_0)) \in \mathcal{G}$. From Corollary 2.16 we know the existence and uniqueness of solution $\tilde{a}_h$ on the whole $[\alpha, \beta]$ and, moreover, due to the properties of the EMCEL scale factor, $(y, \tilde{a}_h(y)) \in \mathcal{G}$ for all $y \in (\alpha, \beta)$ (even on $(l_h, r_h)$). We need to make use of that to prove the result and, in particular, explain that the possibility for the approximations $a^{\gamma_N}$ to leave the “good” region $\mathcal{G}$ on $[\alpha, \beta]$ (see (62) and recall fact (4) above) does not create any problem in our situation.

To this end, we assume that (63) does not hold. Then there exists a small $\varepsilon > 0$ and a subsequence $\{a_1^{(N)}\}$ of $\{a^{\gamma_N}\}$ such that, for every $N$,

$$\sup_{y \in [\alpha, \beta]} |a_1^{(N)}(y) - \tilde{a}_h(y)| \geq \varepsilon.$$
By the construction, the sequence \( \{a_1^{(N)}\} \) is uniformly bounded and equicontinuous on \([\alpha, \beta]\). It is worth noting that the uniform boundedness holds even in the case \( l = -r = -\infty \) (where \( \partial \mathcal{G}_{\text{high}} = \emptyset \) and the upper bound is not seen \textit{a priori}) because of fact [1] above and the fact that \( \alpha \) and \( \beta \) are chosen finite. As for the equicontinuity on the whole \([\alpha, \beta]\), it follows from fact [2] together with the fact that the affine functions constituting \( \partial \mathcal{G}_{\text{low}} \) and \( \partial \mathcal{G}_{\text{high}} \) have bounded slopes \((0, 1 \text{ and } -1)\). By the Arzelà-Ascoli theorem, there exists a uniformly convergent subsequence \( \{a_2^{(N)}\} \) of the sequence \( \{a_1^{(N)}\} \),

\[
\sup_{y \in [\alpha, \beta]} |a_2^{(N)}(y) - \bar{a}(y)| \to 0, \quad N \to \infty,
\]

and the limiting continuous function \([\alpha, \beta] \ni y \mapsto \bar{a}(y)\) is different from \([\alpha, \beta] \ni y \mapsto \hat{a}_h(y)\) because of (64). The standard argumentation in Peano’s existence theorem (see Exercise II.2.1 in [21] or Theorem 2.19 in [40]) yields that \( \bar{a} \) is a solution to (61) in a sufficiently small neighborhood of \( y_0 \). Hence, by the uniqueness for (61), \( \bar{a} = \hat{a}_h \) in a sufficiently small neighborhood of \( y_0 \). Define

\[
\begin{align*}
\bar{a} &= \inf \{ y \in [\alpha, y_0] : \bar{a}(z) = \hat{a}_h(z) \text{ for all } z \in [y, y_0] \} \quad (\in [\alpha, y_0]), \\
\bar{\beta} &= \sup \{ y \in [y_0, \beta] : \bar{a}(z) = \hat{a}_h(z) \text{ for all } z \in [y_0, y] \} \quad (\in (y_0, \beta)).
\end{align*}
\]

As \( \bar{a} \) differs from \( \hat{a}_h \), we either have \( \bar{a} > \alpha \) or \( \bar{\beta} < \beta \). Assume \( \bar{\beta} < \beta \). But, as \( (\bar{\beta}, \bar{\alpha}) = (\hat{a}_h, \hat{a}_h(\beta)) \in \mathcal{G} \), we can again apply the argumentation in Peano’s existence theorem and conclude that \( \bar{a} \) is a solution to (61) in a sufficiently small neighborhood of \( \bar{\beta} \), hence \( \bar{a} = \hat{a}_h \) in that neighborhood. This contradicts the definition of \( \bar{\beta} \) and yields \( \bar{\beta} = \beta \). Similarly, \( \bar{a} = \alpha \). But then \( \bar{a} = \hat{a}_h \) on the whole \([\alpha, \beta]\), which, in turn, contradicts (64) and thus proves (63). This concludes the proof.

In the end we present an example, where we numerically solve the ODE for the EMCEL scale factor and implement our scheme. It is instructive to do this in a situation of an irregular SDE for \( Y \) such that the corresponding Euler scheme does not converge. We first provide a sufficient condition for that.

**Lemma 2.18.** Let \( y \in I^\circ \). Consider the setting of Example 1.1, where the diffusion coefficient \( \eta: I^\circ \to \mathbb{R} \setminus \{0\} \) satisfies\(^{11}\)

\[
\liminf_{y \neq x \to y} \left[ |x - y| |\eta(x)| \right] \in (0, \infty).
\]

Let \( (\zeta_k)_{k \in \mathbb{N}} \) be an iid sequence of random variables with \( E[\zeta_k] = 0, E[\zeta_k^2] = 1 \) and \( P(\zeta_k \neq 0) > 0 \). For each \( h \in (0, \bar{h}) \), we define the Euler scale factor \( a_h: \mathbb{R} \to \mathbb{R} \) by\(^{12}\)

\[
a_h(x) = \eta(x) \sqrt{h}, \quad x \in I^\circ, \quad \text{and} \quad a_h(x) = 0, \quad x \in \mathbb{R} \setminus I^\circ, \quad \text{and define the (linearly interpolated, generalized) Euler scheme } \hat{X}^{h,y} = (X_t^{h,y})_{t \in [0,\infty)} \text{ through } a_h \text{ and } (\zeta_k) \text{ in the same way as the EMCEL scheme } \hat{X}^{h,y} \text{ is defined through } \hat{a}_h \text{ and } (\zeta_k) \text{ in (10)-(11)}. \text{ Then, for any } T \in (0, \infty) \text{ and for any sequence } \{h_n\}_{n \in \mathbb{N}} \subset (0, \infty) \text{ with } h_n \to 0, \text{ the sequence of the laws of the processes } (X_t^{h_n,y})_{t \in [0,T]} \text{ does not converge weakly in } C([0,T]; \mathbb{R}).
\]

\(^{11}\)The notation \( y \neq x \to y \) is to emphasize that we consider \( \liminf \) as \( x \) tends to \( y \) over a deleted neighborhood of \( y \).

\(^{12}\)The extension beyond \( I^\circ \) is needed only to treat the possibility of jumping out of the state space.
Essentially, this result says that Euler-type schemes fail to converge if (65) holds. Notice that (65) is satisfied whenever the diffusion coefficient $\eta$ has in a deleted neighborhood of the point $y$ a singularity proportional to $|x - y|^{-\alpha}$, for some $\alpha \geq 1$.

**Remark 2.19.** (i) We emphasize that not only the convergence of the distributions of $(X_{t}^{h,y})$ to that of $(Y_{t})$ fails; the distributions of $(X_{t}^{h,y})$ do not converge at all.

(ii) Neither it is possible to extract a convergent subsequence (for some $\{h_{n}\} \subset (0, \infty)$ with $h_{n} \to 0$).

(iii) Moreover, the distributions fail to converge on arbitrarily small intervals $[0, T]$ (with $T \in (0, \infty)$).

**Proof of Lemma 2.18** For each $h \in (0, \infty)$, we have $X_{0}^{h,y} = y$, $X_{h}^{h,y} = y + \eta(y)\sqrt{h}\zeta_{1}$ and hence

$$X_{2h}^{h,y} = y + \eta(y)\sqrt{h}\zeta_{1} + \eta \left(y + \eta(y)\sqrt{h}\zeta_{1}\right) \sqrt{h}\zeta_{2}. \quad (66)$$

We fix some $T \in (0, \infty)$ and consider, for all $\varepsilon \in (0, T]$, bounded continuous path functionals $\Phi^{\varepsilon}: C([0, T], \mathbb{R}) \to \mathbb{R}$ defined by the formula

$$\Phi^{\varepsilon} (\omega) = \sup_{s \in [0, \varepsilon]} |\omega(s) - y| \wedge 1, \quad \omega \in C([0, T], \mathbb{R}).$$

It follows from (66) that, for $h \in (0, \frac{\varepsilon}{2}]$,

$$\Phi^{\varepsilon} (X_{t}^{h,y}) \geq \left|\eta(y)\sqrt{h}\zeta_{1} + \eta \left(y + \eta(y)\sqrt{h}\zeta_{1}\right) \sqrt{h}\zeta_{2}\right| \wedge 1.$$

Now take an arbitrary sequence $\{h_{n}\}_{n \in \mathbb{N}} \subset (0, \infty)$ with $h_{n} \to 0$. Denoting by $A \in (0, \infty]$ the $\lim \inf$ in (65) we obtain by the Fatou lemma that

$$\lim_{n \to \infty} \inf E[\Phi^{\varepsilon} (X_{t}^{h_{n},y})] \geq E \left[\lim_{n \to \infty} \inf \eta(y)\sqrt{h_{n}}\zeta_{1} + \eta \left(y + \eta(y)\sqrt{h_{n}}\zeta_{1}\right) \sqrt{h_{n}}\zeta_{2}\wedge 1\right]$$

$$\geq E \left[1\{\zeta_{1} \neq 0\} \left(\left|\frac{A\zeta_{2}}{\eta(y)\zeta_{1}}\right| \wedge 1\right)\right] > 0. \quad (67)$$

Assume that the sequence of the processes $(X_{t}^{h_{n},y})_{t \in [0, T]}$ converges to some process $(X_{t}^{y})_{t \in [0, T]}$ in distribution. Then $(X_{t}^{y})_{t \in [0, T]}$ is a continuous process starting in $y$. By the dominated convergence theorem it follows

$$\lim_{\varepsilon \to 0} E[\Phi^{\varepsilon} (X_{t}^{y})] = 0.$$

This and (67) contradict the weak convergence of the laws of $(X_{t}^{h_{n},y})_{t \in [0, T]}$ to the law of $(X_{t}^{y})_{t \in [0, T]}$. This completes the proof.

Finally, we illustrate the usefulness of the ODE (61) as well as the implications of Proposition 2.7 and Lemma 2.18 in a numerical experiment.

---

13I.e., the schemes corresponding to different distributions of $\zeta_{k}$.
Example 2.20. Let \( b \in (0, \infty) \) and \( I = \mathbb{R} \). Define the function \( \eta : \mathbb{R} \to (0, \infty) \) by the formula
\[
\eta(x) = 1_{\mathbb{R}\setminus\{0,b\}}(x) \frac{1}{|x|} + 1_{\{0,b\}}(x), \quad x \in \mathbb{R},
\]
and let the process \( Y \) solve
\[
dY_t = \eta(Y_t) \, dW_t, \quad Y_0 = x_0 \in \mathbb{R}.
\]
Observe that \( \eta \) satisfies (65) for \( y \in \{0, b\} \) and thus Lemma 2.18 implies that Euler-type schemes fail to converge in the case \( x_0 \in \{0, b\} \). By Proposition 1.3 the EMCEL scheme converges weakly to \( Y \) for all initial values \( x_0 \in \mathbb{R} \) and Corollary 2.16 ensures that for every \( h \in (0, \tilde{h}] \) the EMCEL scale factor \( \tilde{a}_h \) is the unique solution of the initial value problem (61).

Figure 3 depicts numerical solutions \( \tilde{a}_h \) (top) of this ODE as well as normalized solutions \( \tilde{a}_h / \sqrt{h} \) (middle) and solutions relative to the Euler scale factors \( \tilde{a}_h / (\sqrt{h} \eta) \) (bottom) in the case \( b = 2 \) for \( h \in 2 \cdot \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\} \). In particular, the plot at the bottom indicates that \( \lim_{h \to 0} \tilde{a}_h(y) / (\sqrt{h} \eta(y)) = 1 \) for all \( y \in \mathbb{R} \setminus \{0, b\} \) as established in Proposition 2.7. For \( y \in \{0, b\} \) the plot in the middle suggests that \( \tilde{a}_h(y) \) goes slower to 0 than \( \sqrt{h} \) which is confirmed by Proposition 2.7 ensuring that \( \lim_{h \to 0} \tilde{a}_h(y) / (h^{1/4}) = (6/b^2)^{1/4} \) in this case.

Figure 4 illustrates the performance of the weak Euler scheme\(^{15}\) and of the EMCEL scheme in this setting. The upper panel shows two realizations (dotted: weak Euler, solid: EMCEL) with the same random increments \( (\xi_k) \) in the case \( x_0 = 1, \ h = 2 \cdot 10^{-3} \) with time horizon \( T = 2 \). As long as the trajectories move around the level 1 one can barely spot a difference between the trajectories. But as they approach the singularity of \( \eta \) at 0 around time 0.6 the weak Euler scheme is shot to a level bigger than 5, whereas the EMCEL scheme continues smoothly only exhibiting a higher volatility. The reason is that, in contrast to the EMCEL scale factor, the Euler scale factor inherits the singularities of \( \eta \). This observation is also manifested in the lower panel of Figure 4 which depicts the empirical distribution functions of \( X_{2}^{Eu,h} \) (dotted) and \( \tilde{X}_{2}^{h} \) (solid) in this case for a sample size \( M = 10^5 \). One sees that, opposed to the EMCEL scheme, the weak Euler scheme puts a lot of mass outside the interval \([-2, 4]\) indicating that the weak Euler scheme exhibits the behavior portrayed in the upper panel of Figure 4 with high probability.

Acknowledgement We thank three anonymous referees for their constructive comments and suggestions that helped us improve the manuscript. Wolfgang Löhr and Mikhail Urusov acknowledge the support from the German Research Foundation through the project 415705084.

References

[1] S. Ankirchner, N. Kazi-Tani, M. Klein, and T. Kruse. Stopping with expectation constraints: 3 points suffice. Electron. J. Probab., 24:Paper No. 66, 16 pp., 2019.

\(^{14}\)Note that \( \eta \) satisfies (5) and (6), i.e., we are in the setting of Example 1.1.

\(^{15}\)See the text preceding Corollary 2.5 for the definition of the weak Euler scheme.
Figure 3: The figure shows solutions $\tilde{a}_h$ (top), normalized solutions $\tilde{a}_h/\sqrt{h}$ (middle) and solutions relative to the Euler scale factors $\tilde{a}_h/(\sqrt{h}\eta)$ (bottom) of the initial value problem (61) associated to (68) for the case $b = 2$ and for $h \in 2 \cdot \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}\}$. 
Figure 4: The upper panel shows trajectories of the weak Euler scheme $X^{Eu,h}$ (dotted) and the EMCEL scheme $\hat{X}^h$ (solid) in the setting of Example 2.20 in the case $b = 2$, $x_0 = 1$, $h = 2 \cdot 10^{-3}$ with time horizon $T = 2$. The lower panel shows the empirical distribution functions of $X^{Eu,h}_2$ (dotted) and $\hat{X}^h_2$ (solid) in this case for a sample size $M = 10^5$.

[2] S. Ankirchner, M. Klein, T. Kruse, and M. Urusov. On a certain local martingale in a general diffusion setting. Preprint, hal-01700656, 2018.

[3] S. Ankirchner, T. Kruse, and M. Urusov. Numerical approximation of irregular SDEs via Skorokhod embeddings. J. Math. Anal. Appl., 440(2):692–715, 2016.

[4] S. Ankirchner, T. Kruse, and M. Urusov. A functional limit theorem for irregular SDEs. Ann. Inst. Henri Poincaré Probab. Stat., 53(3):1438–1457, 2017.

[5] S. Ankirchner, T. Kruse, and M. Urusov. A functional limit theorem for coin tossing Markov chains. Ann. Inst. Henri Poincaré Probab. Stat., 56(4):2996–3019, 2020.

[6] S. Ankirchner, T. Kruse, and M. Urusov. Wasserstein convergence rates for random bit approximations of continuous Markov processes. J. Math. Anal. Appl., 493(2):124543, 31, 2021.
[7] M. T. Barlow. One-dimensional stochastic differential equations with no strong solution. *J. London Math. Soc. (2)*, 26(2):335–347, 1982.

[8] R. F. Bass. A stochastic differential equation with a sticky point. *Electron. J. Probab.*, 19:no. 32, 22, 2014.

[9] R. F. Bass and Z.-Q. Chen. Stochastic differential equations for Dirichlet processes. *Probab. Theory Related Fields*, 121(3):422–446, 2001.

[10] A. Bensoussan, J.-L. Lions, and G. Papanicolaou. *Asymptotic analysis for periodic structures*. AMS Chelsea Publishing, Providence, RI, 2011. Corrected reprint of the 1978 original [MR0503330].

[11] B. Can and M. Çağlar. Conditional law and occupation times of two-sided sticky Brownian motion. *Statist. Probab. Lett.*, 165:108856, 11, 2020.

[12] A. Eberle and R. Zimmer. Sticky couplings of multidimensional diffusions with different drifts. *Ann. Inst. Henri Poincaré Probab. Stat.*, 55(4):2370–2394, 2019.

[13] H.-J. Engelbert and G. Peskir. Stochastic differential equations for sticky Brownian motion. *Stochastics*, 86(6):993–1021, 2014.

[14] 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.

[15] P. Etoré and A. Lejay. A Donsker theorem to simulate one-dimensional processes with measurable coefficients. *ESAIM: Probability and Statistics*, 11:301–326, 2007.

[16] T. Fattler, M. Grothaus, and R. Voßhall. Construction and analysis of a sticky reflected distorted Brownian motion. *Ann. Inst. Henri Poincaré Probab. Stat.*, 52(2):735–762, 2016.

[17] M. Grothaus and R. Voßhall. Stochastic differential equations with sticky reflection and boundary diffusion. *Electron. J. Probab.*, 22:Paper No. 7, 37, 2017.

[18] M. Grothaus and R. Voßhall. Strong Feller property of sticky reflected distorted Brownian motion. *J. Theoret. Probab.*, 31(2):827–852, 2018.

[19] I. Gyöngy. A note on Euler’s approximations. *Potential Analysis*, 8(3):205–216, 1998.

[20] H. Hajri, M. Caglar, and M. Arnaudon. Application of stochastic flows to the sticky Brownian motion equation. *Electron. Commun. Probab.*, 22:Paper No. 3, 10, 2017.

[21] P. Hartman. *Ordinary differential equations*, volume 38 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002. Corrected reprint of the second (1982) edition [Birkhäuser, Boston, MA; MR0658490 (83e:34002)], With a foreword by Peter Bates.
[22] 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.

[23] O. Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.

[24] I. Karatzas, A. N. Shiryaev, and M. Shkolnikov. On the one-sided Tanaka equation with drift. *Electron. Commun. Probab.*, 16:664–677, 2011.

[25] 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.

[26] P. E. Kloeden and E. Platen. *Numerical solution of stochastic differential equations*, volume 23 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1992.

[27] A. Kohatsu-Higa, A. Lejay, and K. Yasuda. Weak rate of convergence of the Euler-Maruyama scheme for stochastic differential equations with non-regular drift. *J. Comput. Appl. Math.*, 326:138–158, 2017.

[28] V. Konarovskyi. Coalescing-fragmentating Wasserstein dynamics: particle approach. *Preprint, arXiv:1711.03011v3*, 2017.

[29] V. Konarovskyi and M. von Renesse. Reversible coalescing-fragmentating Wasserstein dynamics on the real line. *Preprint, arXiv:1709.02839v2*, 2017.

[30] T. Kruse and M. Urusov. Approximating exit times of continuous Markov processes. *Discrete Contin. Dyn. Syst. Ser. B*, 25(9):3631–3650, 2020.

[31] M. Lavrentieff. Sur une équation différentielle du premier ordre. *Math. Z.*, 23(1):197–209, 1925.

[32] J.-F. Le Gall. One-dimensional stochastic differential equations involving the local times of the unknown process. In *Stochastic analysis and applications (Swansea, 1983)*, volume 1095 of *Lecture Notes in Math.*, pages 51–82. Springer, Berlin, 1984.

[33] A. Lejay, L. Lenôtre, and G. Pichot. An exponential timestepping algorithm for diffusion with discontinuous coefficients. *J. Comput. Phys.*, 396:888–904, 2019.

[34] A. Lejay and M. Martinez. A scheme for simulating one-dimensional diffusion processes with discontinuous coefficients. *Ann. Appl. Probab.*, 16(1):107–139, 2006.

[35] G. N. Milstein and J. Schoenmakers. Uniform approximation of the Cox-Ingersoll-Ross process via exact simulation at random times. *Adv. in Appl. Probab.*, 48(4):1095–1116, 2016.
[36] H.-L. Ngo and D. Taguchi. Strong convergence for the Euler-Maruyama approximation of stochastic differential equations with discontinuous coefficients. *Statist. Probab. Lett.*, 125:55–63, 2017.

[37] G. Pagès. *Numerical probability*. Universitext. Springer, Cham, 2018. An introduction with applications to finance.

[38] I. G. Petrovski˘ı. *Lektsii po teorii obyknovennikh differentsial’nikh uravneni˘ı. (In Russian)*. Moskov. Gos. Univ., Moscow, seventh edition, 1984. Edited and with a preface by A. D. Myshkis and O. A. Ole˘ınik.

[39] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.

[40] G. Teschl. *Ordinary differential equations and dynamical systems*, volume 140 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2012.

[41] L. Yan. The Euler scheme with irregular coefficients. *The Annals of Probability*, 30(3):1172–1194, 2002.