Research Article

Approximating Explicitly the Mean-Reverting CEV Process

N. Halidias and I. S. Stamatiou

Department of Mathematics, University of the Aegean, Karlovassi, 83 200 Samos, Greece

Correspondence should be addressed to N. Halidias; nikoshalidias@hotmail.com

We are interested in the numerical solution of mean-reverting CEV processes that appear in financial mathematics models and are described as nonnegative solutions of certain stochastic differential equations with sublinear diffusion coefficients of the form $(x_t)^q$, where $1/2 < q < 1$. Our goal is to construct explicit numerical schemes that preserve positivity. We prove convergence of the proposed SD scheme with rate depending on the parameter q. Furthermore, we verify our findings through numerical experiments and compare with other positivity preserving schemes. Finally, we show how to treat the two-dimensional stochastic volatility model with instantaneous variance process given by the above mean-reverting CEV process.

1. Introduction

Consider the following stochastic models in Itô form:

$$S_t = S_0 + \int_0^t \mu \cdot S_u du + \int_0^t (V_u)^p \cdot S_u dW_u, \quad t \in [0, T],$$

$$V_t = V_0 + \int_0^t (k_1 - k_2 V_s) ds + \int_0^t k_3 (V_s)^q d\widetilde{W}_s \tag{1}$$

$$t \in [0, T],$$

where S_t represents the underlying financially observable variable, V_t is the instantaneous volatility when $p = 1$ or the instantaneous variance when $p = 1/2$, and the Wiener processes (W_t), $(\widetilde{W}_t)$ have correlation ρ.

We assume that (V_t) is a mean-reverting CEV process of the above form, with the coefficients $k_i > 0$ for $i = 1, 2, 3$ and $q > 1/2$, since the process (V_t) has to be nonnegative. To be more precise the above restriction on q implies that (V_t) is positive; that is, 0 is unattainable, as well as nonexplosive; that is, ∞ is unattainable, as can be verified by Feller's classification of boundaries [17, Proposition 5.22] (see also Appendix A). The steady-state level of V_t is k_1/k_2 and the rate of mean-reversion is k_2.

System (1) for $p = q = 1/2$ is the Heston model. When $q = 1$ we get the Brennan-Schwartz model [2, Section II] that despite its simple form cannot provide analytical expressions for S_t.

Process (V_t) for $q = 1/2$, also known as the CIR process [6, (13)], by the initials of the authors that proposed it for the term structure of interest rates, has received a lot of attention and we just mention the latest two contributions to the study of such processes (see [4, 5] and references therein).

Process (V_t) for $1/2 \le q \le 1$ has been also considered for the dynamics of the short-term interest rate [3, (1)]. The stationary distribution of the process has also been derived in [7, Proposition 2.2].

We aim for a positivity preserving scheme for the process (V_t). The scheme that we propose and denote as semidiscrete (SD) preserves the analytical property of (V_t) staying positive. The idea of the semidiscrete method is that we discretize a part of the original SDE and then apply Itô's formula (cf. [8] where the method originally appeared and [5, 9, 23]). The explicit Euler scheme fails to preserve positivity, as well as the standard Milstein scheme. We intend to apply the semidiscrete method for the numerical approximation of (V_t) in model (1) with $1/2 < q < 1$ and compare with other positivity preserving methods such as the balanced implicit method (BIM) (introduced by [11, (3.2)] with the positivity preserving property [25, Section 5] and its stability properties [13]; see also [14] for an extended balanced method with better stability behavior) and the balanced Milstein method

(BMM) [25, Theorem 5.9] (we give in Appendix B the form of all the above schemes for the approximation of (V_t)). Finally, we approximate the stochastic volatility model (1) with $p = 1/2$. In [15] a thorough treatment can be found, where also another stochastic volatility model is suggested.

Section 2 provides the setting and the main results, Theorems 1 and 2, concerning the $\mathcal{L}^2$-convergence of the proposed semidiscrete method to the true solution of mean-reverting CEV processes of the form of the stochastic volatility in (1). The rate of mean-square convergence in Theorem 1 is logarithmic and in Theorem 2 is polynomial with magnitude $(1/2)(q-1/2)$. The main ingredient of the approach we adopt, inspired by [16], is a change of the initial Brownian motion (W_t) to another Brownian motion $(\widehat{W}_t)$ justified by Lévy's martingale characterization of Brownian motion.

Section 3 is devoted to the logarithmic rate of convergence of the proposed semidiscrete scheme, while Section 4 concerns the proof of the polynomial rate of convergence. In Section 5 we briefly discuss the case where we do not alter the initial Brownian motion (W_t). This approach produces reduced convergence rate. Finally, Section 6 presents illustrative figures where the behavior of the proposed scheme, regarding the order of convergence, is shown and a comparison with BIM and BMM schemes is given. In Section 7 we treat the full model (1) for a special case. Concluding remarks are in Section 8 and in Appendix B we briefly present numerical schemes for the integration of the variance-volatility process (V_t).

2. The Setting and the Main Results

We consider the following SDE:

$$x_t = x_0 + \int_0^t (k_1 - k_2 x_s)\, ds + \int_0^t k_3 (x_s)^q\, dW_s, \tag{2}$$

$$t \in [0, T],$$

where k_1, k_2, k_3 are positive and $1/2 < q < 1$. Then, Feller's test implies that there is a unique strong solution such that $x_t > 0$ a.s. when $x_0 > 0$ a.s. Let

$$f_\theta(x, y)$$

$$= \underbrace{k_1 - k_2 (1 - \theta) x - \left((k_3)^2 / 4 (1 + k_2 \theta \Delta)\right) x^{2q-1} - k_2 \theta y}_{f_1(x,y)} \tag{3}$$

$$+ \underbrace{\left((k_3)^2 / 4 (1 + k_2 \theta \Delta)\right) x^{2q-1}}_{f_2(x)},$$

$$g(x, y) = k_3 x^{q-1/2} \sqrt{y}, \tag{4}$$

where $f(x, x) = a(x) = k_1 - k_2 x$ and $g(x, x) = b(x) = k_3 x^q$.

Let the partition $0 = t_0 < t_1 < \cdots < t_N = T$ with $\Delta = T/N$ and consider the following process:

$$y_t^{\mathrm{SD}}(q) = y_{t_n} + f_1\left(y_{t_n}, y_t\right) \cdot \Delta + \int_{t_n}^t f_2\left(y_{t_n}\right) ds \tag{5}$$

$$+ \int_{t_n}^t \mathrm{sgn}(z_s)\, g\left(y_{t_n}, y_s\right) dW_s,$$

with $y_0 = x_0$ a.s. or more explicitly

$$y_t^{\mathrm{SD}}(q) = y_{t_n} + \left(k_1 - k_2 (1 - \theta) y_{t_n} \right.$$

$$\left. - \frac{(k_3)^2}{4(1 + k_2 \theta \Delta)} \left(y_{t_n}\right)^{2q-1} - k_2 \theta y_t \right) \cdot \Delta \tag{6}$$

$$+ \int_{t_n}^t \frac{(k_3)^2}{4(1 + k_2 \theta \Delta)} \left(y_{t_n}\right)^{2q-1} ds + k_3 \left(y_{t_n}\right)^{q-1/2}$$

$$\cdot \int_{t_n}^t \mathrm{sgn}(z_s) \sqrt{y_s}\, dW_s,$$

for $t \in (t_n, t_{n+1}]$, where $\theta \in [0, 1]$ represents the level of implicitness and

$$z_t = \sqrt{y_n} + \frac{k_3}{2(1 + k_2 \theta \Delta)} \left(y_{t_n}\right)^{q-1/2} \left(W_t - W_{t_n}\right), \tag{7}$$

with

$$y_n := y_{t_n} \left(1 - \frac{k_2 \Delta}{1 + k_2 \theta \Delta} \right) + \frac{k_1 \Delta}{1 + k_2 \theta \Delta}$$

$$- \frac{(k_3)^2}{4(1 + k_2 \theta \Delta)^2} \left(y_{t_n}\right)^{2q-1} \Delta. \tag{8}$$

Process (6) is well defined when $y_n \geq 0$ and this is true when $(1/(1 + k_2 \theta \Delta))(k_3)^2 \leq 4(k_2 \wedge k_1)$ and $\Delta(2 - \theta) \leq 1/k_2$. Furthermore, (6) has jumps at nodes t_n. Solving for y_t, we end up with the following explicit scheme:

$$y_t^{\mathrm{SD}}(q) = y_n + \int_{t_n}^t \frac{(k_3)^2}{4(1 + k_2 \theta \Delta)^2} \left(y_{t_n}\right)^{2q-1} ds$$

$$+ \frac{k_3}{1 + k_2 \theta \Delta} \left(y_{t_n}\right)^{q-1/2} \int_{t_n}^t \mathrm{sgn}(z_s) \sqrt{y_s}\, dW_s, \tag{9}$$

with solution in each step given by [1, (4.39), page 123]

$$y_t^{\mathrm{SD}}(q) = (z_t)^2, \tag{10}$$

which has the pleasant feature $y_t^{\mathrm{SD}}(q) \geq 0$.

Inspired by [16] we remove the term $\mathrm{sgn}(z_s)$ from (6) by considering the process

$$\widetilde{W}_t = \int_0^t \mathrm{sgn}(z_s)\, dW_s, \tag{11}$$

which is a martingale with quadratic variation $\langle \widetilde{W}_t, \widetilde{W}_t \rangle = t$ and thus a standard Brownian motion with respect to its own

filtration, justified by Lévy's theorem [17, Theorem 3.16, page 157]. Therefore, the compact form of (6) becomes

$$
\begin{aligned}
y_t^{\text{SD}} = x_0 &+ \int_0^t \left(k_1 - k_2 \left(1 - \theta \right) y_{\widehat{s}} - k_2 \theta y_{\widetilde{s}} \right) ds \\
&+ \int_t^{t_{n+1}} \left(k_1 - k_2 \left(1 - \theta \right) y_{t_n} \right. \\
&\quad \left. - \frac{\left(k_3 \right)^2}{4 \left(1 + k_2 \theta \Delta \right)} \left(y_{t_n} \right)^{2q-1} - k_2 \theta y_t \right) ds \\
&+ k_3 \int_0^t \left(y_{\widehat{s}} \right)^{q-1/2} \sqrt{y_s} d\widetilde{W}_s,
\end{aligned}
\tag{12}
$$

for $t \in (t_n, t_{n+1}]$ where

$$
\widehat{s} = t_j, \quad s \in \left(t_j, t_{j+1} \right], \quad j = 0, \ldots, n,
$$

$$
\widetilde{s} = \begin{cases} t_{j+1}, & \text{for } s \in \left[t_j, t_{j+1} \right], \\ t, & \text{for } s \in \left[t_n, t \right] \end{cases} \quad j = 0, \ldots, n-1.
\tag{13}
$$

Consider also the process

$$
\widetilde{x}_t = x_0 + \int_0^t \left(k_1 - k_2 \widetilde{x}_s \right) ds + \int_0^t k_3 \left(\widetilde{x}_s \right)^q d\widetilde{W}_s,
\tag{14}
$$

$$
t \in [0, T].
$$

The process (x_t) of (2) and the process $(\widetilde{x}_t)$ of (14) have the same distribution. We show in the following that $\mathbb{E} \sup_{0 \le t \le T} |y_t^{\text{SD}}(q) - \widetilde{x}_t|^2 \to 0$ as $\Delta \downarrow 0$; thus the same holds for the unique solution of (2); that is, $\mathbb{E} \sup_{0 \le t \le T} |y_t^{\text{SD}}(q) - x_t|^2 \to 0$ as $\Delta \downarrow 0$. To simplify notation we write $\widetilde{W}, (\widetilde{x}_t)$ as $W, (x_t)$. We end up with the following explicit scheme:

$$
\begin{aligned}
y_t^{\text{SD}}(q) = y_n &+ \int_{t_n}^t \frac{\left(k_3 \right)^2}{4 \left(1 + k_2 \theta \Delta \right)^2} \left(y_{t_n} \right)^{2q-1} ds \\
&+ \frac{k_3}{1 + k_2 \theta \Delta} \left(y_{t_n} \right)^{q-1/2} \int_{t_n}^t \sqrt{y_s} dW_s,
\end{aligned}
\tag{15}
$$

where y_n is as in (8).

Assumption A. Let the parameters k_1, k_2, k_3 be positive such that $(1/(1 + k_2 \theta \Delta))(k_3)^2 \le 4(k_2 \wedge k_1)$ and consider $\Delta > 0$ such that $\Delta(2 - \theta) < 1/k_2$, for $\theta \in [0, 1]$. Moreover assume $x_0 \ge 0$ a.s. and $\mathbb{E}(x_0)^p < A$ for some $p \ge 4$.

Theorem 1 (logarithmic rate of convergence). *Let Assumption A hold. The semidiscrete scheme (15) converges to the true solution of (2) in the mean-square sense with rate given by*

$$
\mathbb{E} \sup_{0 \le t \le T} \left| y_t^{\text{SD}}(q) - x_t \right|^2 \le \frac{C}{\sqrt{\ln (\Delta)^{-1}}},
\tag{16}
$$

where C is independent of Δ and given by

$$
C := 32 \sqrt{\frac{6}{\epsilon}} \left(k_3 \right)^4 T^2 e^{6T^2(k_2)^2 + k_2 T},
\tag{17}
$$

where $0 < \epsilon < q - 1/2$.

Assumption B. Let Assumption A hold where now $x_0 \in \mathbb{R}$ and $x_0 > 0$.

Theorem 2 (polynomial rate of convergence). *Let Assumption B hold. Then the semidiscrete scheme (15) converges to the true solution of (2) in the mean-square sense with rate given by*

$$
\mathbb{E} \sup_{0 \le t \le T} \left| y_t^{\text{SD}}(q) - x_t \right|^2 \le C\Delta^{(q-1/2)},
\tag{18}
$$

where

$$
\begin{aligned}
C := 16 \left(k_3 \right)^2 T \sqrt{A_2 \left(x_0 + k_1 T \right)^2} \sqrt{\widehat{A}_{4q-2}} &\left(2e^{6(k_2)^2 T^2} \right. \\
&\left. + \frac{C_{HK}}{\epsilon - 1} \left(x_0 \right)^{(1-q)\nu(\lambda)} \right),
\end{aligned}
\tag{19}
$$

and C_{HK} is the constant described in (83) and λ is an appropriately chosen positive parameter which satisfies (84) and always exists, $\nu(\lambda) := \lambda/2(1 - q)^2(k_3)^2 - 1$, and $\epsilon > 1$.

In the following sections we write for simplicity y_t^{SD} or y_t for $y_t^{\text{SD}}(q)$.

3. Logarithmic Rate of Convergence

3.1. Moment Bounds

Lemma 3 (moment bound for SD approximation). *It holds that*

$$
\mathbb{E} \sup_{0 \le t \le T} \left(y_t \right)^p \le A_p \mathbb{E} \left(x_0 + k_1 T \right)^p,
\tag{20}
$$

for any $p > 2$, where $A_p := \exp\{(p(p - 1)/2)(k_3)^2((p - 1)/2p + 2^{p-1}/p)T\}$.

Proof of Lemma 3. We first observe that (y_t) is bounded in the following way:

$$
\begin{aligned}
0 \le y_t &\le x_0 + \int_0^t k_1 ds + k_3 \int_0^t \left(y_{\widehat{s}} \right)^{q-1/2} \sqrt{y_s} dW_s \\
&\le x_0 + k_1 T + k_3 \int_0^t \left(y_{\widehat{s}} \right)^{q-1/2} \sqrt{y_s} dW_s := u_t,
\end{aligned}
\tag{21}
$$

a.s., where the lower bound comes from the construction of (y_t) and the upper bound follows from a comparison theorem. We will bound (u_t) and therefore (y_t), since $0 \le y_t \le u_t$ a.s. Set the stopping time $\tau_R := \inf\{t \in [0, T] : u_t > R\}$, for $R > 0$ with the convention $\inf \emptyset = \infty$. Application of Itô's formula on $(u_{t \wedge \tau_R})^p$ implies

$$
\begin{aligned}
\left(u_{t \wedge \tau_R} \right)^p = \left(x_0 + k_1 T \right)^p &+ \frac{p \left(p - 1 \right)}{2} \left(k_3 \right)^2 \\
&\cdot \int_0^{t \wedge \tau_R} \left(u_s \right)^{p-2} \left(y_{\widehat{s}} \right)^{2q-1} y_s ds \\
&+ pk_3 \int_0^{t \wedge \tau_R} \left(u_s \right)^{p-1} \left(y_{\widehat{s}} \right)^{q-1/2} \sqrt{y_s} dW_s
\end{aligned}
$$

$$\leq \left(x_0 + k_1 T\right)^p + \frac{p\left(p-1\right)}{2}\left(k_3\right)^2$$

$$\cdot \int_0^{t \wedge \tau_R} \left(u_s\right)^{p-1}\left(y_{\bar{s}}\right)^{2q-1} ds + M_t \leq \left(x_0 + k_1 T\right)^p$$

$$+ \frac{p\left(p-1\right)}{2}\left(k_3\right)^2$$

$$\cdot \int_0^{t \wedge \tau_R} \left(\frac{p-1}{2p}\left(u_s\right)^p + \frac{2^{p-1}}{p}\left(y_{\bar{s}}\right)^{(2q-1)p}\right) ds + M_t$$

$$\leq \left(x_0 + k_1 T\right)^p + \frac{p\left(p-1\right)}{2}\left(k_3\right)^2 \left(\frac{p-1}{2p} + \frac{2^{p-1}}{p}\right)$$

$$\cdot \int_0^{t \wedge \tau_R} \left(u_s\right)^p ds + M_t, \tag{22}$$

where in the second step we have used the fact that $0 \leq y_t \leq u_t$, in the third step the inequality $x^{p-1} y \leq \epsilon((p-1)/p)x^p + (1/p\epsilon^{p-1})y^p$, valid for $x \wedge y \geq 0$ and $p > 1$ with $\epsilon = 1/2$, and in the final step the fact that $1/2 < q < 1$ and $M_t := pk_3 \int_0^{t \wedge \tau_R} (u_s)^{p-1}(y_{\bar{s}})^{q-1/2} \sqrt{y_s} dW_s$. Taking expectations in the above inequality and using that M_t is a local martingale vanishing at 0, we get

$$\mathbb{E}\left(u_{t \wedge \tau_R}\right)^p \leq \mathbb{E}\left(x_0 + k_1 T\right)^p + \frac{p\left(p-1\right)}{2}\left(k_3\right)^2$$

$$\cdot \left(\frac{p-1}{2p} + \frac{2^{p-1}}{p}\right) \int_0^t \mathbb{E}\left(u_{s \wedge \tau_R}\right)^p ds$$

$$\leq \mathbb{E}\left(x_0 + k_1 T\right)^p \tag{23}$$

$$\cdot \exp\left\{\frac{p\left(p-1\right)}{2}\left(k_3\right)^2 \left(\frac{p-1}{2p} + \frac{2^{p-1}}{p}\right) T\right\}$$

$$\leq A_p \mathbb{E}\left(x_0 + k_1 T\right)^p,$$

where we have applied the Gronwall inequality [18, (7)]. We have that

$$\left(y_{t \wedge \tau_R}\right)^p = \left(y_{\tau_R}\right)^p \mathbb{I}_{\{\tau_R \leq t\}} + \left(y_t\right)^p \mathbb{I}_{\{t < \tau_R\}} \geq \left(y_t\right)^p \mathbb{I}_{\{t < \tau_R\}}. \tag{24}$$

Thus taking expectations in the above inequality and using the estimated upper bound for $\mathbb{E}(u_{t \wedge \tau_R})^p$ we arrive at

$$\mathbb{E}\left(y_t\right)^p \mathbb{I}_{\{t < \tau_R\}} \leq \mathbb{E}\left(y_{t \wedge \tau_R}\right)^p \leq \mathbb{E}\left(u_{t \wedge \tau_R}\right)^p$$

$$\leq A_p \mathbb{E}\left(x_0 + k_1 T\right)^p, \tag{25}$$

and, taking the limit as $R \to \infty$, we get

$$\lim_{R \to \infty} \mathbb{E}\left(y_t\right)^p \mathbb{I}_{\{t < \tau_R\}} \leq A_p \mathbb{E}\left(x_0 + k_1 T\right)^p. \tag{26}$$

Let us fix t. The sequence of stopping times τ_R is increasing in R and $t \wedge \tau_R \to t$ as $R \to \infty$, and thus the sequence $(y_t)^p \mathbb{I}_{\{t < \tau_R\}}$ is nondecreasing in R and $(y_t)^p \mathbb{I}_{\{t < \tau_R\}} \to (y_t)^p$ as

$R \to \infty$. Application of the monotone convergence theorem implies

$$\mathbb{E}\left(y_t\right)^p \leq A_p \mathbb{E}\left(x_0 + k_1 T\right)^p, \tag{27}$$

for any $p > 2$. Using again Itô's formula on $(u_t)^p$, taking the supremum, and then using Doob's martingale inequality on the diffusion term we bound $\mathbb{E}\sup_{0 \leq t \leq T}(u_t)^p$ and thus $\mathbb{E}\sup_{0 \leq t \leq T}(y_t)^p$. $\qquad\square$

Lemma 4 (error bound for SD scheme). *Let n_s be an integer such that $s \in [t_{n_s}, t_{n_s+1}]$. Then*

$$\mathbb{E}\left|y_s - y_{\bar{s}}\right|^p \leq \widehat{A}_p \Delta^{p/2},$$

$$\mathbb{E}\left|y_s - y_{\bar{s}}\right|^p < \widetilde{A}_p \Delta^{p/2}, \tag{28}$$

for any $p > 0$, where the positive quantities $\widehat{A}_p$, $\widetilde{A}_p$ do not depend on Δ.

Proof of Lemma 4. First we take a $p > 2$. We get that

$$|y_s - y_{\bar{s}}|^p = \left| \int_{t_{n_s}}^s \left(k_1 - k_2\left(1-\theta\right) y_{\bar{u}} - k_2\theta y_{\bar{u}}\right) du \right.$$

$$+ \int_{t_{n_s}}^{t_{n_s+1}} k_2\theta y_{\bar{s}} du - \int_s^{t_{n_s+1}} k_2\theta y_s du$$

$$+ \int_s^{t_{n_s}} \left(k_1 - k_2\left(1-\theta\right) y_{t_{n_s}} - \frac{\left(k_3\right)^2}{4\left(1 + k_2\theta\Delta\right)}\left(y_{t_{n_s}}\right)^{2q-1}\right) du$$

$$\left. + k_3 \int_{t_{n_s}}^s \left(y_{\bar{u}}\right)^{q-1/2} \sqrt{y_u} dW_u \right|^p$$

$$\leq 5^{p-1} \left(\left| \int_{t_{n_s}}^s \left(k_1 - k_2\left(1-\theta\right) y_{\bar{u}} - k_2\theta y_{\bar{u}}\right) du \right|^p \right.$$

$$+ \left(k_2\right)^p \theta^p \left(y_{\bar{s}}\right)^p \left(t_{n_s+1} - t_{n_s}\right)^p + \left(k_2\right)^p \theta^p \left(y_s\right)^p \left(t_{n_s+1} - s\right)^p$$

$$+ \left| \int_s^{t_{n_s}} \left(k_1 - k_2\left(1-\theta\right) y_{t_{n_s}} - \frac{\left(k_3\right)^2}{4\left(1 + k_2\theta\Delta\right)}\left(y_{t_{n_s}}\right)^{2q-1}\right) du \right|^p \tag{29}$$

$$\left. + \left(k_3\right)^p \left| \int_{t_{n_s}}^s \left(y_{\bar{u}}\right)^{q-1/2} \sqrt{y_u} dW_u \right|^p \right)$$

$$\leq 5^{p-1} \left(\left|t_{n_s} - s\right|^{p-1} \int_{t_{n_s}}^s \left|k_1 - k_2\left(1-\theta\right) y_{\bar{u}} - k_2\theta y_{\bar{u}}\right|^p du \right.$$

$$+ \left(k_2\right)^p \theta^p \left(\left(y_{\bar{s}}\right)^p + \left(y_s\right)^p\right) \Delta^p$$

$$+ \left|k_1 - k_2\left(1-\theta\right) y_{t_{n_s}} - \frac{\left(k_3\right)^2}{4\left(1 + k_2\theta\Delta\right)}\left(y_{t_{n_s}}\right)^{2q-1}\right|^p \Delta^p$$

$$\left. + \left(k_3\right)^p \left| \int_{t_{n_s}}^s \left(y_{\bar{u}}\right)^{q-1/2} \sqrt{y_u} dW_u \right|^p \right),$$

where we have used the Cauchy-Schwarz inequality. Taking expectations in the above inequality and using Lemma 3 and Doob's martingale inequality on the diffusion term we conclude

$$\mathbb{E}\left|y_s - y_{\bar{s}}\right|^p \leq \widehat{A}_p \Delta^{p/2}, \tag{30}$$

where the positive quantity $\widehat{A}_p$, except on p, depends also on the parameters k_1, k_2, k_3, θ, q, but not on Δ. Now, for $0 < p < 2$, we get

$$\mathbb{E}\left|y_s - y_{\bar{s}}\right|^p \leq \left(\mathbb{E}\left|y_s - y_{\bar{s}}\right|^2\right)^{p/2} \leq \widehat{A}_p \Delta^{p/2}, \qquad (31)$$

where we have used Jensen's inequality for the concave function $\phi(x) = x^{p/2}$. Following the same lines, we can show that

$$\mathbb{E}\left|y_s - y_{\bar{s}}\right|^p \leq \widetilde{A}_p \Delta^{p/2}, \qquad (32)$$

for any $0 < p$, where the positive quantity $\widetilde{A}_p$, except on p, depends also on the parameters k_1, k_2, k_3, θ, q, but not on Δ. □

For the rest of this section we rewrite again the compact form of (12) in the following way:

$$y_t^{\mathrm{SD}} = \underbrace{x_0 + \int_0^t f_\theta\left(y_{\bar{s}}, y_{\bar{s}}\right) ds + \int_0^t g\left(y_{\bar{s}}, y_s\right) dW_s}_{h_t}$$

$$+ \int_t^{t_{n+1}} f_1\left(y_{t_n}, y_t\right) ds, \qquad (33)$$

where $f_\theta(\cdot, \cdot)$ is given by (3) and the auxiliary process (h_t) is close to (y_t) as shown in the next result.

Lemma 5 (moment bounds involving the auxiliary process). *For any $s \in [0, T]$ it holds that*

$$\mathbb{E}\left|h_s - y_s\right|^p \leq C_p \Delta^p,$$

$$\mathbb{E}\left|h_s\right|^p \leq C_h, \qquad (34)$$

and for $s \in [t_n, t_{n+1}]$ one has that

$$\mathbb{E}\left|h_s - y_{\bar{s}}\right|^p \leq \widehat{C}_p \Delta^{p/2},$$

$$\mathbb{E}\left|h_s - y_{\bar{s}}\right|^p \leq \widetilde{C}_p \Delta^{p/2}, \qquad (35)$$

for any $p > 0$, where the positive quantities C_p, $\widehat{C}_p$, $\widetilde{C}_p$, C_h do not depend on Δ.

Proof of Lemma 5. We have that

$$\left|h_s - y_s\right|^p = \left|\int_s^{t_{n+1}} f_1\left(y_{t_n}, y_s\right) du\right|^p$$

$$\leq \left|t_{n+1} - s\right|^p \left|f_1\left(y_{t_n}, y_s\right)\right|^p, \qquad (36)$$

for any $p > 0$, where we have used (33). Using Lemma 3 we get the left part of (34). Now for $p > 2$ and noting that

$$\mathbb{E}\left|h_s\right|^p \leq 2^{p-1}\mathbb{E}\left|h_s - y_s\right|^p + 2^{p-1}\mathbb{E}\left|y_s\right|^p$$

$$\leq 2^{p-1}C_p\Delta^p + 2^{p-1}A_p\mathbb{E}\left(x_0 + k_1 T\right)^p \leq C_h, \qquad (37)$$

we get the right part of (34), where we have used Lemma 3. The case $0 < p < 2$ follows by Jensen's inequality as in Lemma 4.

Furthermore, for $s \in [t_n, t_{n+1}]$ and $p > 2$, we derive that

$$\mathbb{E}\left|h_s - y_{\bar{s}}\right|^p \leq 2^{p-1}\mathbb{E}\left|h_s - y_s\right|^p + 2^{p-1}\mathbb{E}\left|y_s - y_{\bar{s}}\right|^p$$

$$\leq 2^{p-1}C_p\Delta^p + 2^{p-1}\widehat{A}_p\Delta^{p/2} \leq \widehat{C}_p\Delta^{p/2}, \qquad (38)$$

where we have used (30) and in the same manner

$$\mathbb{E}\left|h_s - y_{\bar{s}}\right|^p \leq 2^{p-1}C_p\Delta^p + 2^{p-1}\widetilde{A}_p\Delta^{p/2} \leq \widetilde{C}_p\Delta^{p/2}. \qquad (39)$$

The case $0 < p < 2$ follows by Jensen's inequality. □

3.2. Convergence of the Auxiliary Process (h_t) to (x_t) in $\mathscr{L}^1$. We will use the representation (33) and write

$$h_t - x_t = \int_0^t \left(f_\theta\left(y_{\bar{s}}, y_{\bar{s}}\right) - f_\theta\left(x_s, x_s\right)\right) ds$$

$$+ \int_0^t \left(g\left(y_{\bar{s}}, y_s\right) - g\left(x_s, x_s\right)\right) dW_s. \qquad (40)$$

Proposition 6. *Let Assumption A hold. Then one has*

$$\sup_{0 \leq t \leq T} \mathbb{E}\left|h_t - x_t\right| \leq \left(J_3 \frac{\Delta^{q-1/2}}{m e_m} + 2\left(k_3\right)^2 T \frac{1}{m}\right) e^{k_2 T}, \qquad (41)$$

for any $m > 1$, where $e_m = e^{-m(m+1)/2}$ and

$$J_3 := 2\left(k_3\right)^2 T \sqrt{A_2 \mathbb{E}\left(x_0 + k_1 T\right)^2} \sqrt{\widehat{A}_{4q-2}}. \qquad (42)$$

Proof of Proposition 6. Let the nonincreasing sequence $\{e_m\}_{m \in \mathbb{N}}$ with $e_m = e^{-m(m+1)/2}$ and $e_0 = 1$. We introduce the following sequence of smooth approximations of $|x|$ (method of Yamada and Watanabe [19]):

$$\phi_m\left(x\right) = \int_0^{|x|} dy \int_0^y \psi_m\left(u\right) du, \qquad (43)$$

where the existence of the continuous function $\psi_m(u)$ with $0 \leq \psi_m(u) \leq 2/(mu)$ and support in (e_m, e_{m-1}) is justified by $\int_{e_m}^{e_{m-1}} (du/u) = m$. The following relations hold for $\phi_m \in \mathscr{C}^2(\mathbb{R}, \mathbb{R})$ with $\phi_m(0) = 0$,

$$|x| - e_{m-1} \leq \phi_m\left(x\right) \leq |x|,$$

$$\left|\phi_m'\left(x\right)\right| \leq 1,$$

$$x \in \mathbb{R}, \qquad (44)$$

$$\left|\phi_m''\left(x\right)\right| \leq \frac{2}{m|x|}, \quad \text{when } e_m < |x| < e_{m-1},$$

$$\left|\phi_m''\left(x\right)\right| = 0 \quad \text{otherwise.}$$

We have that

$$\mathbb{E}\left|h_t - x_t\right| \leq e_{m-1} + \mathbb{E}\phi_m\left(h_t - x_t\right). \qquad (45)$$

Moreover we find that

$$f_\theta(y_{\hat{s}}, y_s) - f_\theta(x_s, x_s) = (k_1 - k_2(1-\theta)y_{\hat{s}} - k_2\theta y_{\bar{s}})$$
$$- (k_1 - k_2 x_s) = -k_2(1-\theta)(y_{\hat{s}} - x_s)$$
$$- k_2\theta(y_{\bar{s}} - x_s) = k_2(1-\theta)(h_s - y_{\hat{s}})$$
$$+ k_2\theta(h_s - y_{\bar{s}}) - k_2(h_s - x_s),$$
(46)

$$|g(y_{\hat{s}}, y_s) - g(x_s, x_s)|^2$$
$$= \left|k_3(y_{\hat{s}})^{q-1/2}\sqrt{y_s} - k_3(x_s)^q\right|^2 \le (k_3)^2$$
$$\cdot \left(\sqrt{y_s}\left((y_{\hat{s}})^{q-1/2} - (y_s)^{q-1/2}\right) + \left((y_s)^q - (x_s)^q\right)\right)^2$$
$$\le 2(k_3)^2$$
(47)
$$\cdot \left(y_s\left((y_{\hat{s}})^{q-1/2} - (y_s)^{q-1/2}\right)^2 + \left((y_s)^q - (x_s)^q\right)^2\right)$$
$$\le 2(k_3)^2\left(y_s|y_{\hat{s}} - y_s|^{2q-1} + \left(\sqrt{|y_s - x_s|}\right)^2\right)$$
$$\le 2(k_3)^2\left(y_s|y_{\hat{s}} - y_s|^{2q-1} + |h_s - y_s| + |h_s - x_s|\right),$$

where we have used properties of Hölder continuous functions and, namely, the fact that x^q is q-Hölder continuous for $q \le 1$, that is, $|x^q - y^q| \le |x-y|^q$, and that x^q is $1/2$-Hölder continuous since $q > 1/2$. Application of Itô's formula to the sequence $\{\phi_m\}_{m\in\mathbb{N}}$ implies

$$\phi_m(h_t - x_t) = \int_0^t \phi_m'(h_s - x_s)\left(f_\theta(y_{\hat{s}}, y_{\bar{s}})\right)$$
$$- f_\theta(x_s, x_s)) ds + M_t + \frac{1}{2}\int_0^t \phi_m''(h_s - x_s)$$
$$\cdot (g(y_{\hat{s}}, y_{\bar{s}}) - g(x_s, x_s))^2 ds$$
$$\le \int_0^t (k_2(1-\theta)|h_s - y_{\hat{s}}| + k_2\theta|h_s - y_{\bar{s}}|$$
$$+ k_2|h_s - x_s|) ds + M_t + 2(k_3)^2$$
$$\cdot \int_0^t \frac{1}{m|h_s - x_s|}\left(y_s|y_{\hat{s}} - y_s|^{2q-1} + |h_s - y_s|\right.$$
(48)
$$+ |h_s - x_s|) ds \le k_2(1-\theta)\int_0^t |h_s - y_{\hat{s}}| ds$$
$$+ k_2\theta\int_0^t |h_s - y_{\bar{s}}| ds + \frac{2(k_3)^2}{me_m}\int_0^t |h_s - y_s| ds$$
$$+ k_2\int_0^t |h_s - x_s| ds + M_t + \frac{2(k_3)^2}{me_m}\int_0^t y_s|y_{\hat{s}}$$
$$- y_s|^{2q-1} ds + \frac{2(k_3)^2 T}{m},$$

where in the second step we have used (46) and (47) and the properties of ϕ_m and

$$M_t := \int_0^t \phi_m'(h_u - x_u)(g(y_{\hat{u}}, y_{\bar{u}}) - g(x_u, x_u)) dW_u. \quad (49)$$

Taking expectations in the above inequality yields

$$\mathbb{E}\phi_m(h_t - x_t)$$
$$\le k_2(1-\theta)\int_0^t \mathbb{E}|h_s - y_{\hat{s}}| ds$$
$$+ k_2\theta\int_0^t \mathbb{E}|h_s - y_{\bar{s}}| ds$$
$$+ \frac{2(k_3)^2}{me_m}\int_0^t \mathbb{E}|h_s - y_s| ds$$
$$+ \frac{2(k_3)^2}{me_m}\int_0^t \mathbb{E}y_s|y_{\hat{s}} - y_s|^{2q-1} ds + \frac{2(k_3)^2 T}{m}$$
$$+ k_2\int_0^t \mathbb{E}|h_s - x_s| ds$$
$$\le k_2(1-\theta)T\widehat{C}_1\sqrt{\Delta} + k_2\theta T\widetilde{C}_1\sqrt{\Delta}$$
$$+ \frac{2(k_3)^2 TC_1}{me_m}\Delta + k_2\int_0^t \mathbb{E}|h_s - x_s| ds$$
(50)
$$+ \frac{2(k_3)^2}{me_m}\int_0^t \sqrt{\mathbb{E}(y_s)^2}\sqrt{\mathbb{E}|y_{\hat{s}} - y_s|^{4q-2}} ds$$
$$+ \frac{2(k_3)^2 T}{m}$$
$$\le k_2 T\left((1-\theta)\widehat{C}_1 + \theta\widetilde{C}_1\right)\sqrt{\Delta} + \frac{2(k_3)^2 TC_1}{me_m}\Delta$$
$$+ k_2\int_0^t \mathbb{E}|h_s - x_s| ds$$
$$+ \frac{2(k_3)^2 T}{me_m}\sqrt{A_2\mathbb{E}(x_0 + k_1 T)^2}\sqrt{\widehat{A}_{4q-2}}\Delta^{q-1/2}$$
$$+ \frac{2(k_3)^2 T}{m},$$

where we have used Lemma 5 in the second step and the Hölder inequality and Lemmas 3 and 4 in the third step and the fact that $\mathbb{E}M_t = 0$ (the function $d(u) = \phi_m'(h_u - x_u)(g(y_{\hat{u}}, y_{\bar{u}}) - g(x_u, x_u))$ belongs to the space $\mathcal{M}^2([0,t];\mathbb{R})$ of real-valued measurable $\mathscr{F}_t$-adapted processes such that $\mathbb{E}\int_0^t |d(u)|^2 du < \infty$; thus [20, Theorem 1.5.8] implies $\mathbb{E}M_t = 0$). Thus (45) becomes

$$\mathbb{E}|h_t - x_t| \le e_{m-1} + J_1\sqrt{\Delta} + 2(k_3)^2 TC_1\frac{\Delta}{me_m}$$

$$+ J_3\frac{\Delta^{q-1/2}}{me_m} + 2(k_3)^2 T\frac{1}{m}$$

$$+ k_2 \int_0^t \mathbb{E} \left| h_s - x_s \right| ds$$

$$\leq J_3 \frac{\Delta^{q-1/2}}{m e_m} + 2 \left(k_3 \right)^2 T \frac{1}{m}$$

$$+ k_2 \int_0^t \mathbb{E} \left| h_s - x_s \right| ds$$

$$\leq \left(J_3 \frac{\Delta^{q-1/2}}{m e_m} + 2 \left(k_3 \right)^2 T \frac{1}{m} \right) e^{k_2 t}, \tag{51}$$

where in the second step we have used the asymptotic relations $\Delta^\kappa = o(\Delta^{q-1/2})$ as $\Delta \downarrow 0$ for any $\kappa \geq 1/2$, $e_{m-1} = o(1/m)$ as $m \to \infty$, $\sqrt{\Delta} = o(\Delta^\kappa / m e_m)$ for any $\kappa \leq 1$ as $m \to \infty$; in the last step we have used the Gronwall inequality and J_3 is as defined in Proposition 6 while

$$J_1 := k_2 T \left((1 - \theta) \widehat{C}_1 + \theta \widetilde{C}_1 \right). \tag{52}$$

Taking the supremum over all $0 \leq t \leq T$ gives (41). $\qquad \square$

3.3. Convergence of the Auxiliary Process (h_t) to (x_t) in $\mathscr{L}^2$

Proposition 7. *Let Assumption A hold. Then one has*

$$\mathbb{E} \sup_{0 \leq t \leq T} \left| h_t - x_t \right|^2 \leq \frac{C_\epsilon}{\sqrt{\ln (\Delta)^{-1}}}, \tag{53}$$

where C_ϵ is independent of Δ and given by $C_\epsilon := 32 \sqrt{3/2\epsilon} (k_3)^4 T^2 e^{6T^2 (k_2)^2 + k_2 T}$, where $0 < \epsilon < q - 1/2$.

Proof of Proposition 7. We estimate the difference $\left| \mathscr{E}_t \right|^2 := \left| h_t - x_t \right|^2$. It holds that

$$\left| \mathscr{E}_t \right|^2 = \left| \int_0^t \left(f_\theta \left(y_{\widehat{s}}, y_s \right) - f_\theta \left(x_s, x_s \right) \right) ds \right.$$

$$\left. + \int_0^t \left(g \left(y_{\widehat{s}}, y_s \right) - g \left(x_s, x_s \right) \right) dW_s \right|^2$$

$$\leq 2T \int_0^t \left(k_2 \left(1 - \theta \right) \left| h_s - y_s \right| + k_2 \theta \left| h_s - y_{\widehat{s}} \right| \right.$$

$$\left. + k_2 \left| \mathscr{E}_s \right| \right)^2 ds + 2 \left| M_t \right|^2 \leq 6T \left(k_2 \right)^2 \left(1 - \theta \right)^2 \tag{54}$$

$$\cdot \int_0^t \left| h_s - y_s \right|^2 ds + 6T \left(k_2 \right)^2 \theta^2 \int_0^t \left| h_s - y_{\widehat{s}} \right|^2 ds$$

$$+ 6T \left(k_2 \right)^2 \int_0^t \left| \mathscr{E}_s \right|^2 ds + 2 \left| M_t \right|^2,$$

where in the second step we have used the Cauchy-Schwarz inequality and (46) and

$$M_t := \int_0^t \left(g \left(y_{\widehat{u}}, y_u \right) - g \left(x_u, x_u \right) \right) dW_u. \tag{55}$$

Taking the supremum over all $t \in [0, T]$ and then expectations we have

$$\mathbb{E} \sup_{0 \leq t \leq T} \left| \mathscr{E}_t \right|^2 \leq 6T \left(k_2 \right)^2 \left(1 - \theta \right)^2 \int_0^T \mathbb{E} \left| h_s - y_s \right|^2 ds$$

$$+ 6T \left(k_2 \right)^2 \theta^2 \int_0^T \mathbb{E} \left| h_s - y_{\widehat{s}} \right|^2 ds$$

$$+ 6T \left(k_2 \right)^2 \int_0^T \mathbb{E} \sup_{0 \leq l \leq s} \left| \mathscr{E}_l \right|^2 ds$$

$$+ 2 \mathbb{E} \sup_{0 \leq t \leq T} \left| M_t \right|^2 \tag{56}$$

$$\leq 6T^2 \left(k_2 \right)^2 \left(1 - \theta \right)^2 \widehat{A}_2 \Delta$$

$$+ 6T^2 \left(k_2 \right)^2 \theta^2 \widetilde{A}_2 \Delta$$

$$+ 6T \left(k_2 \right)^2 \int_0^T \mathbb{E} \sup_{0 \leq l \leq s} \left| \mathscr{E}_l \right|^2 ds$$

$$+ 8 \mathbb{E} \left| M_T \right|^2,$$

where in the second step we have used Lemma 4 and Doob's martingale inequality with $p = 2$, since M_t is an $\mathbb{R}$-valued martingale that belongs to $\mathscr{L}^2$. We find that

$$\mathbb{E} \left| M_T \right|^2 := \mathbb{E} \left| \int_0^T \left| g \left(y_{\widehat{s}}, y_s \right) - g \left(x_s, x_s \right) \right| dW_s \right|^2$$

$$= \mathbb{E} \left(\int_0^T \left| g \left(y_{\widehat{s}}, y_s \right) - g \left(x_s, x_s \right) \right|^2 ds \right) \leq 2 \left(k_3 \right)^2$$

$$\cdot \mathbb{E} \left(\int_0^T \left(y_s \left| y_{\widehat{s}} - y_s \right|^{2q-1} + \left| h_s - y_s \right| \right. \right.$$

$$\left. + \left| h_s - x_s \right| \right) ds \right) \leq 2 \left(k_3 \right)^2 \int_0^T \mathbb{E} \left(y_s \left| y_{\widehat{s}} \right. \right. \tag{57}$$

$$\left. \left. - y_s \right|^{2q-1} \right) ds + 2 \left(k_3 \right)^2 \int_0^T \mathbb{E} \left| h_s - y_s \right| ds$$

$$+ 2 \left(k_3 \right)^2 \int_0^T \mathbb{E} \left| \mathscr{E}_s \right| ds,$$

where we have used (47). Now, Lemmas 3, 4, and 5 imply

$$\mathbb{E} \left| M_T \right|^2 \leq J_6 \sqrt{\Delta^{2q-1}} + 2 \left(k_3 \right)^2 T C_1 \Delta$$

$$+ 2 \left(k_3 \right)^2 \int_0^T \mathbb{E} \left| \mathscr{E}_s \right| ds \tag{58}$$

$$\leq J_6 \Delta^{q-1/2} + 2 \left(k_3 \right)^2 \int_0^T \mathbb{E} \left| \mathscr{E}_s \right| ds,$$

where we have used the asymptotic relations $\Delta^l = o(\Delta^{q-1/2})$ for all $l \geq 1/2$ as $\Delta \downarrow 0$ and the quantity J_6 is given by $J_6 := 2(k_3)^2 T \sqrt{A_2 \mathbb{E}(x_0 + k_1 T)^2} \sqrt{\widehat{A}_{4q-2}}$.

Relation (56) becomes

$$
\mathbb{E} \sup_{0 \le t \le T} |\mathscr{E}_t|^2
$$

$$
\le 8 J_6 \Delta^{q-1/2} + J_5 \Delta + 6T (k_2)^2 \int_0^T \mathbb{E} \sup_{0 \le l \le s} |\mathscr{E}_l|^2 ds
$$

$$
+ 16 (k_3)^2 \int_0^T \mathbb{E} \sup_{0 \le l \le s} |\mathscr{E}_l| ds
$$

$$
\le 8 J_6 \Delta^{q-1/2}
$$

$$
+ 16 (k_3)^2 T \left(J_3 \frac{\Delta^{q-1/2}}{m e_m} + 3 (k_3)^2 T \frac{1}{m} \right) e^{k_2 T} \tag{59}
$$

$$
+ 6T (k_2)^2 \int_0^T \mathbb{E} \sup_{0 \le l \le s} |\mathscr{E}_l|^2 ds
$$

$$
\le 16 (k_3)^2 T J_3 e^{k_2 T + 6T^2 (k_2)^2} \frac{\Delta^{q-1/2}}{m e_m}
$$

$$
+ 32 (k_3)^4 T^2 e^{k_2 T + 6T^2 (k_2)^2} \frac{1}{m},
$$

where we have used Proposition 6 in the second step with the sequence e_m as defined there and Gronwall's inequality in the last step and the asymptotic relation $\Delta^\kappa = o(\Delta^\kappa / m e_m)$ as $m \to \infty$, for any $\kappa > 0$, and J_5 is independent of Δ and given by $J_5 := 6T^2 (k_2)^2 ((1-\theta)^2 \widehat{A}_2 + \theta^2 \widetilde{A}_2)$.

We take $m = \sqrt{\ln \Delta^{-\lambda}}$, with $\lambda > 0$, to be specified soon and note that $e^{\sqrt{\ln \Delta^{-\lambda}}} = o(\Delta^{-\lambda})$ as $\Delta \downarrow 0$, since $e^{\sqrt{\ln n}} = o(n)$ as $n \to \infty$. Moreover we have that

$$
\frac{\Delta^{q-1/2}}{e_m} = \frac{\Delta^{q-1/2}}{e^{-m^2/2}} e^{m/2} = \frac{\Delta^{q-1/2}}{e^{-(\ln \Delta^{-\lambda})/2}} e^{(1/2)\sqrt{\ln \Delta^{-\lambda}}}
$$

$$
= \Delta^{q-1/2-3\lambda/2} \frac{e^{(1/2)\sqrt{\ln \Delta^{-\lambda}}}}{\Delta^{-\lambda}}. \tag{60}
$$

Now, since $q > 1/2$ there is an $\epsilon > 0$ small enough such that $q - 1/2 - \epsilon > 0$. We take $\lambda = 2\epsilon/3$ and conclude that

$$
\frac{\Delta^{q-1/2}}{e_m} = \Delta^{q-1/2-\epsilon} \frac{e^{(1/2)\sqrt{\ln \Delta^{2\epsilon/3}}}}{\Delta^{-2\epsilon/3}} \longrightarrow 0, \tag{61}
$$

as $\Delta \to 0$ which in turn implies the asymptotic relation $\Delta^{q-1/2}/m e_m = o(1/m)$ as $m \to \infty$, with the logarithmic rate stated before. In the same way we can show $\Delta^{1/4}/m e_m = o(1/m)$ as $m \to \infty$, by taking $\epsilon < 1/4$. We finally arrive at

$$
\mathbb{E} \sup_{0 \le t \le T} |\mathscr{E}_t|^2 \le 32 (k_3)^4 T^2 e^{k_2 T + 6T^2 (k_2)^2} \frac{1}{\sqrt{\ln \Delta^{-2\epsilon/3}}}, \tag{62}
$$

by taking $0 < \epsilon < q - 1/2$, which implies (53). □

3.4. Proof of Theorem 1. In order to finish the proof of Theorem 1 we just use the triangle inequality, Lemma 5, and Proposition 7 to get

$$
\mathbb{E} \sup_{0 \le t \le T} |y_t - x_t|^2 \le 2 \mathbb{E} \sup_{0 \le t \le T} |h_t - y_t|^2 + 2 \mathbb{E} \sup_{0 \le t \le T} |\mathscr{E}_t|^2
$$

$$
\le 2 C_2 \Delta^2 + 2 \frac{C_\epsilon}{\sqrt{\ln \Delta^{-1}}} \le \frac{C}{\sqrt{\ln \Delta^{-1}}}, \tag{63}
$$

where $C = C(k_2, k_3, \epsilon, T)$, given in the statement of Theorem 1.

4. Polynomial Rate of Convergence

We work with the stochastic time change inspired by [21]. We define the process

$$
\gamma (t) := \int_0^t \frac{128 (k_3)^2 q^2}{\left[(y_s)^{1-q} + (x_s)^{1-q} \right]^2} ds \tag{64}
$$

and the stopping time

$$
\tau_l := \inf \left\{ s \in [0, T] : 6T (k_2)^2 s + \gamma (s) \ge l \right\}. \tag{65}
$$

The process $\gamma(t)$ is well defined since $x_t > 0$ a.s. and $y_t \ge 0$ (see Section 2).

The difference $|\mathscr{E}_t|^2 := |h_t - x_t|^2$ is estimated as in Section 3 and we get, as in (56), that

$$
\mathbb{E} \sup_{0 \le t \le \tau} |\mathscr{E}_t|^2 \le J_5 \Delta + 6T (k_2)^2 \int_0^\tau \mathbb{E} \sup_{0 \le l \le s} |\mathscr{E}_l|^2 ds
$$

$$
+ 8 \mathbb{E} |M_\tau|^2, \tag{66}
$$

where τ is a stopping time and J_5 independent of Δ is as in the proof of Proposition 7. The main difference here will be the estimation of the last term in (66). The approach in Section 3 resulted in the $\mathscr{L}^1$ estimation $\mathbb{E}|\mathscr{E}_t|$ where we used the Yamada-Watanabe approach. Now, we use the Berkaoui approach. Relation (47) becomes

$$
|g (y_{\hat{s}}, y_s) - g (x_s, x_s)|^2
$$

$$
\le 2 (k_3)^2 \left(y_s |y_{\hat{s}} - y_s|^{2q-1} + \left| (y_s)^q - (x_s)^q \right|^2 \right)
$$

$$
\le 2 (k_3)^2 \left(y_s |y_{\hat{s}} - y_s|^{2q-1} \right)
$$

$$
+ \left| (y_s)^q - (x_s)^q \right|^2 \left((y_s)^{1-q} + (x_s)^{1-q} \right)^2 \frac{(\gamma_s)'}{64 q^2} \tag{67}
$$

$$
\le 2 (k_3)^2 \left(y_s |y_{\hat{s}} - y_s|^{2q-1} \right)
$$

$$
+ \frac{1}{8} \left(|h_s - y_s|^2 + |\mathscr{E}_s|^2 \right) (\gamma_s)',
$$

where we have used the inequality

$$
|a^q - b^q| \left(a^{1-q} + b^{1-q} \right) \le 2q |a - b|, \tag{68}
$$

valid for all $a \geq 0$, $b \geq 0$, and $1/2 \leq q \leq 1$. Consequently, we get the upper bound

$$
\mathbb{E} |M_\tau|^2 := \mathbb{E} \left| \int_0^\tau |g(y_s, y_s) - g(x_s, x_s)| \, dW_s \right|^2
$$

$$
\leq J_6 \Delta^{q-1/2} + \frac{1}{8} \int_0^\tau \mathbb{E} |h_s - y_s|^2 (\gamma_s)' \, ds
$$

$$
+ \frac{1}{8} \int_0^\tau \mathbb{E} |\mathcal{E}_s|^2 (\gamma_s)' \, ds \tag{69}
$$

$$
\leq J_6 \Delta^{q-1/2}
$$

$$
+ \frac{1}{8} \int_0^\tau \sqrt{\mathbb{E} |h_s - y_s|^4} \sqrt{\mathbb{E} \left((\gamma_s)' \right)^2} \, ds
$$

$$
+ \frac{1}{8} \int_0^\tau \mathbb{E} |\mathcal{E}_s|^2 (\gamma_s)' \, ds,
$$

where we used the Hölder inequality; J_6 independent of Δ is as in proof of Proposition 7. Relation (66) becomes

$$
\mathbb{E} \sup_{0 \leq t \leq \tau} |\mathcal{E}_t|^2
$$

$$
\leq 8 J_6 \Delta^{q-1/2} + 6T (k_2)^2 \int_0^\tau \mathbb{E} \sup_{0 \leq l \leq s} |\mathcal{E}_l|^2 \, ds
$$

$$
+ \int_0^\tau \sqrt{\mathbb{E} |h_s - y_s|^4} \sqrt{\mathbb{E} \left((\gamma_s)' \right)^2} \, ds
$$

$$
+ \int_0^\tau \mathbb{E} |\mathcal{E}_s|^2 (\gamma_s)' \, ds
$$

$$
\leq 8 J_6 \Delta^{q-1/2}
$$

$$
+ \sqrt{C_4} \Delta^2 \int_0^\tau \sqrt{\mathbb{E} \left(\frac{128 (k_3)^2 q^2}{\left[(y_s)^{1-q} + (x_s)^{1-q} \right]^2} \right)^2} \, ds \tag{70}
$$

$$
+ \int_0^\tau \mathbb{E} \sup_{0 \leq l \leq s} |\mathcal{E}_l|^2 \left(6T (k_2)^2 s + \gamma_s \right)' \, ds
$$

$$
\leq 8 J_6 \Delta^{q-1/2}
$$

$$
+ \sqrt{C_4} 128 (k_3)^2 q^2 \Delta^2 \int_0^\tau \sqrt{\mathbb{E} \left(\frac{1}{(x_s)^{2-2q}} \right)} \, ds
$$

$$
+ \int_0^\tau \mathbb{E} \sup_{0 \leq l \leq s} |\mathcal{E}_l|^2 \left(6T (k_2)^2 s + \gamma_s \right)' \, ds,
$$

where we have used Lemma 5 in the second step. At this point we want to estimate the inverse moments of (x_t) and to do

so we consider the transformation $v = x^{2-2q}$ and apply Itô's formula to get

$$
v_t = v_0 + \int_0^t \left(\underbrace{\frac{(1-2q)(1-q)(k_3)^2}{K_0}}_{K_0} \right.
$$

$$
+ \underbrace{2(1-q) k_1 (v_s)^{(1-2q)/(2-2q)}}_{K_1} \tag{71}
$$

$$
\left. - \underbrace{2(1-q) k_2 v_s}_{K_2} \right) ds + \int_0^t \underbrace{2 k_3 (1-q)}_{K_3} \sqrt{v_s} \, dW_s,
$$

for $t \in [0, T]$, where $v_0 = (x_0)^{2-2q} > 0$. Denote the drift coefficient of the process (v_t) by $a(v_t)$ and consider the function

$$
\alpha(v) := a(v) - \lambda + K_2 v + \underbrace{\frac{(2q-1)(\lambda + K_0)^{1/(2q-1)}}{(k_1)^{(2-2q)/(2q-1)}}}_{\eta(\lambda)} v, \tag{72}
$$

where $\lambda \geq 0$. Some elementary calculations show that this function attains its minimum at $v^* := (k_1 (2q-1)/\eta(\lambda))^{2-2q}$ and $\alpha(v^*) = 0$, thus

$$
a(v) \geq \lambda - (K_2 + \eta(\lambda)) v. \tag{73}
$$

Consider the process $(\zeta_t(\lambda))$ defined through

$$
\zeta_t(\lambda) = \zeta_0 + \int_0^t (\lambda - (K_2 + \eta(\lambda)) \zeta_s) \, ds
$$

$$
+ \int_0^t K_3 \sqrt{\zeta_s} \, dW_s, \tag{74}
$$

for $t \in [0, T]$ with $\zeta_0(\lambda) = v_0$. Process (74) is a square root diffusion process and when $2\lambda/(K_3)^2 - 1 \geq 0$ or

$$
\lambda \geq 2 (1-q)^2 (k_3)^2 \tag{75}
$$

the process is a CIR process which remains positive if $\zeta_0(\lambda) > 0$. By a comparison theorem [17, Proposition 5.2.18] we obtain that $v_t \geq \zeta_t(\lambda) > 0$ a.s. or $(v_t)^{-1} \leq (\zeta_t(\lambda))^{-1}$ a.s. or equivalently $(x_t)^{2q-2} \leq (\zeta_t(\lambda))^{-1}$ a.s. The inverse moment bounds of $(\zeta_t(\lambda))$ follow [22, (3.1)]:

$$
\sup_{t \in [0,T]} \mathbb{E} (\zeta_t(\lambda))^p < \infty, \quad \text{for } p > -2 \frac{\lambda}{(K_3)^2} \tag{76}
$$

by choosing big enough λ and particularly such that (75) holds strictly. Therefore,

$$
\mathbb{E} \sup_{0 \leq t \leq \tau} |\mathcal{E}_t|^2 \leq 8 J_6 \Delta^{q-1/2}
$$

$$
+ \int_0^\tau \mathbb{E} \sup_{0 \leq l \leq s} |\mathcal{E}_l|^2 \left(6T (k_2)^2 s + \gamma_s \right)' \, ds. \tag{77}
$$

Relation (77) for $\tau = \tau_l$ implies

$$
\mathbb{E} \sup_{0 \le t \le \tau_l} \left(\mathscr{E}_t \right)^2
$$

$$
\le 8 J_6 \Delta^{q-1/2} + \int_0^{\tau_l} \mathbb{E} \sup_{0 \le l \le s} \left(\mathscr{E}_l \right)^2 \left(6 T \left(k_2 \right)^2 s + \gamma_s \right)' ds \quad (78)
$$

$$
\le 8 J_6 \Delta^{q-1/2} + \int_0^l \mathbb{E} \sup_{0 \le j \le u} \left(\mathscr{E}_{\tau_j} \right)^2 du \le 8 J_6 e^l \Delta^{q-1/2},
$$

where in the last step we have used Gronwall's inequality. Using again relation (77) for $\tau = T$ and under the change of variables $u = 6 T (k_2)^2 s + \gamma_s$ we get

$$
\mathbb{E} \sup_{0 \le t \le T} \left(\mathscr{E}_t \right)^2 \le 8 J_6 \Delta^{q-1/2}
$$

$$
+ \int_0^{6(k_2)^2 T^2 + \gamma_T} \mathbb{E} \sup_{0 \le j \le u} \left(\mathscr{E}_{\tau_j} \right)^2 du \le 8 J_6 \Delta^{q-1/2}
$$

$$
+ \int_0^\infty \mathbb{E} \left(\sup_{0 \le j \le u} \left(\mathbb{1}_{\{6(k_2)^2 T^2 + \gamma_T \ge u\}} \mathscr{E}_{\tau_j} \right)^2 \right) du
$$

$$
\le \int_0^{6(k_2)^2 T^2} \mathbb{E} \sup_{0 \le j \le u} \left(\mathscr{E}_{\tau_j} \right)^2 du
$$

$$
+ \int_{6(k_2)^2 T^2}^\infty \mathbb{P} \left(6 \left(k_2 \right)^2 T^2 + \gamma_T \ge u \right) \quad (79)
$$

$$
\cdot \mathbb{E} \left(\sup_{0 \le j \le u} \left(\mathscr{E}_{\tau_j} \right)^2 \mid \left\{ 6 \left(k_2 \right)^2 T^2 + \gamma_T \ge u \right\} \right) du
$$

$$
+ 8 J_6 \Delta^{q-1/2} \le 8 J_6 e^{6(k_2)^2 T^2} \Delta^{q-1/2} + \int_0^\infty \mathbb{P} \left(\gamma_T \ge u \right)
$$

$$
\cdot \mathbb{E} \sup_{0 \le j \le u} \left(\mathscr{E}_{\tau_j} \right)^2 du + 8 J_6 \Delta^{q-1/2}
$$

$$
\le 16 J_6 e^{6(k_2)^2 T^2} \Delta^{q-1/2} + 8 J_6 \Delta^{q-1/2} \int_0^\infty \mathbb{P} \left(\gamma_T \ge u \right)
$$

$$
\cdot e^u du,
$$

where in the last steps we have used (78). We proceed by showing that $u \to \mathbb{P}(\gamma_T \ge u) e^u \in \mathscr{L}^1(\mathbb{R}_+)$. Markov's inequality implies

$$
\mathbb{P} \left(\gamma_T \ge u \right) \le e^{-\epsilon u} \mathbb{E} \left(e^{\epsilon \gamma_T} \right), \quad (80)
$$

for any $\epsilon > 0$. The following bound holds:

$$
\gamma_T = \int_0^T \frac{128 \left(k_3 \right)^2 q^2}{\left[\left(y_s \right)^{1-q} + \left(x_s \right)^{1-q} \right]^2} ds
$$

$$
\le 128 \left(k_3 \right)^2 q^2 \int_0^T \left(x_s \right)^{2q-2} ds, \quad (81)
$$

and thus

$$
\mathbb{E} \left(e^{\epsilon \gamma_T} \right) \le \mathbb{E} \left(e^{\epsilon 128(k_3)^2 q^2 \int_0^T (x_s)^{2q-2} ds} \right), \quad (82)
$$

where $-1 < 2q - 2 < 0$. It remains to bound the exponential inverse moments of (x_t) defined through the stochastic integral equation (2). Exponential inverse moments for the CIR process are known [10, Theorem 3.1] and are given by

$$
\mathbb{E} e^{\delta \int_0^t (\zeta_s(\lambda))^{-1} ds} \le C_{\mathrm{HK}} \left(\zeta_0 \right)^{-(1/2)(\nu(\lambda) - \sqrt{\nu(\lambda)^2 + 8(\delta/(K_3)^2)})}, \quad (83)
$$

for $0 \le \delta \le (2\lambda/(K_3)^2 - 1)^2 ((K_3)^2/8) =: \nu(\lambda)^2 ((K_3)^2/8)$ where the positive constant C_{HK} is explicitly given in [10, (10)] and depends on the parameters k_2, k_3, T, q, but is independent of ζ_0. Thus the other condition that we require for parameter λ is

$$
\lambda \ge 2 \left(1 - q \right) \sqrt{2\delta} \left(k_3 \right) + 2 \left(1 - q \right)^2 \left(k_3 \right)^2. \quad (84)
$$

When (84) is satisfied then (75) is satisfied too; thus there is actually no restriction on the coefficient δ in (83) since we can always choose appropriately a λ such that (84) holds. Relation (82) becomes

$$
\mathbb{E} \left(e^{\epsilon \gamma_T} \right) \le \mathbb{E} \left(e^{\epsilon 128(k_3)^2 q^2 \int_0^T (v_s)^{-1} ds} \right)
$$

$$
\le \mathbb{E} \left(e^{\epsilon 128(k_3)^2 q^2 \int_0^T (\zeta_s(\lambda))^{-1} ds} \right). \quad (85)
$$

We therefore require that

$$
128 \left(k_3 \right)^2 q^2 \epsilon \le \left(\nu \left(\lambda \right) \right)^2 \frac{\left(K_3 \right)^2}{8} \quad (86)
$$

and can always find a $\epsilon > 1$, such that the above relation holds by choosing appropriately λ as discussed before.

Relation (85) becomes

$$
\mathbb{E} \left(e^{\epsilon \gamma_T} \right) \le C_{\mathrm{HK}} \left(\zeta_0 \right)^{-\nu(\lambda)/2}, \quad (87)
$$

and therefore

$$
\mathbb{P} \left(\gamma_T \ge u \right) \le C_{\mathrm{HK}} \left(x_0 \right)^{(1-q)\nu(\lambda)} e^{-\epsilon u}, \quad (88)
$$

where λ is chosen such that (86) holds with $\epsilon > 1$. We conclude

$$
\mathbb{E} \sup_{0 \le t \le T} \left(\mathscr{E}_t \right)^2
$$

$$
\le 16 J_6 e^{6(k_2)^2 T^2} \Delta^{q-1/2}
$$

$$
+ 8 J_6 C_{\mathrm{HK}} \left(x_0 \right)^{(1-q)\nu(\lambda)} \Delta^{q-1/2} \int_0^\infty e^{(1-\epsilon)u} du \quad (89)
$$

$$
\le C \cdot \Delta^{q-1/2},
$$

by choosing $\epsilon > 1$ where $C = C(k_1, k_2, k_3, T, q, \epsilon) := 8 J_6 (2 e^{6(k_2)^2 T^2} + (C_{\mathrm{HK}}/(\epsilon - 1))(x_0)^{(1-q)\nu(\lambda)})$, is as given in statement of Theorem 2.

5. Alternative Approach with Reduced Rate of Convergence

In this section we briefly discuss the case where instead of (12) we use directly (6). Then, Lemmas 3, 4, and 5 still hold; that is, the moment bounds and error bounds of (y_t^{SD}), as well as the moment bounds involving the auxiliary process (h_t), are true. The proof of the convergence results follows the same lines as in Sections 3 and 4. The main difference is in the estimation (47) that now becomes

$$\left| \operatorname{sgn}(z_s) \, g(y_{\hat{s}}, y_s) - g(x_s, x_s) \right|^2 \le 3 (k_3)^2$$

$$\cdot \left((y_{\hat{s}})^{2q-1} \, y_s \, (\operatorname{sgn}(z_s) - 1)^2 \right. \tag{90}$$

$$\left. + y_s \, |y_{\hat{s}} - y_s|^{2q-1} + |h_s - y_s| + |h_s - x_s| \right).$$

The first term on the right-hand side of the above inequality containing the $\operatorname{sgn}(z_s)$ will contribute in a negative way to the rate of convergence. We do not give all the details but just mention that in order to bound the expectation of that term, which can be done in the following way,

$$\mathbb{E} (y_{\hat{s}})^{2q-1} \, y_s \, |\operatorname{sgn}(z_s) - 1|^2 = \mathbb{E} \left(4 (y_{t_n})^{2q-1} \, y_s \mathbb{1}_{\{z_s \le 0\}} \right)$$

$$\le 4 \mathbb{E} \left| (y_{t_n})^{2q-1} \, y_s - (y_{t_n})^{2q} \right|$$

$$+ 4 \mathbb{E} \left((y_{t_n})^{2q} \mathbb{1}_{\{z_s \le 0\}} \mathbb{1}_{\{y_{t_n} \le \Delta^{1-2\xi}\}} \right)$$

$$+ 4 \mathbb{E} \left((y_{t_n})^{2q} \mathbb{1}_{\{z_s \le 0\}} \mathbb{1}_{\{y_{t_n} > \Delta^{1-2\xi}\}} \right) \tag{91}$$

$$\le 4 \mathbb{E} \left| (y_{t_n})^{2q-1} \, (y_s - y_{t_n}) \right| + 4\Delta^{2q-4q\xi}$$

$$+ 4 \sqrt{\mathbb{E} (y_{t_n})^{4q}} \sqrt{\mathbb{P} \left(\{z_s \le 0\} \cap \{y_{t_n} > \Delta^{1-2\xi}\} \right)},$$

we need to estimate the probability of z_t being negative when at the same time $y_{t_n} > \Delta^{1-2\xi}$, for $0 < \xi < 1/2$.

Lemma 8. *For every $t \in [t_n, t_{n+1}]$ it holds that*

$$\mathbb{P} \left(\{z_t \le 0\} \cap \{y_{t_n} > \Delta^{1-2\xi}\} \right) \le C_{k_2, k_3, \theta, \Delta} \sqrt{\Delta}, \tag{92}$$

where $C_{k_2, k_3, \theta, \Delta} := k_3 / \sqrt{1 - k_2(2 - \theta)\Delta}$ and $\Delta(2 - \theta) < 1/k_2$ and $(k_3)^2/(1 + k_2\theta\Delta) \le 4k_2$. Relation (92) implies that $\mathbb{P}(\{z_t \le 0\} \cap \{y_{t_n} > \Delta^{1-2\xi}\}) = O(\sqrt{\Delta})$ as $\Delta \downarrow 0$.

Proof of Lemma 8. By the definition (7) of (z_t) for $t \in [t_n, t_{n+1}]$ and for $0 < \xi < 1/2$, we have that

$$A := \{z_t \le 0\} \cap \{y_{t_n} > \Delta^{1-2\xi}\}$$

$$= \left\{ (y_{t_n})^{q-1/2} \left(W_t - W_{t_n} \right) \le -\frac{2(1 + k_2\theta\Delta)}{k_3} \sqrt{y_n} \right\} \tag{93}$$

$$\cap \{y_{t_n} > \Delta^{1-2\xi}\} \subseteq A_1 \cup A_2,$$

where

$$A_1$$

$$:= \left\{ W_t - W_{t_n} \le -\frac{2(1 + k_2\theta\Delta)}{k_3} \sqrt{(y_n)} \, (y_{t_n})^{-q+1/2} \right\}$$

$$\cap \{y_{t_n} \ge 1\}, \tag{94}$$

$$A_2$$

$$:= \left\{ W_t - W_{t_n} \le -\frac{2(1 + k_2\theta\Delta)}{k_3} \sqrt{(y_n)} \, (y_{t_n})^{-q+1/2} \right\}$$

$$\cap \{1 > y_{t_n} > \Delta^{1-2\xi}\}.$$

The following inclusion relations hold for the event A_1:

$$A_1 \subseteq \left\{ \Delta W_n \le -\frac{2(1 + k_2\theta\Delta)}{k_3} \sqrt{y_{t_n} \left(1 - \frac{k_2\Delta}{1 + k_2\theta\Delta} \right) - \frac{(k_3)^2 \Delta}{4(1 + k_2\theta\Delta)^2} \, (y_{t_n})^{2q-1} \, (y_{t_n})^{-q+1/2}} \right\} \cap \{y_{t_n} \ge 1\}$$

$$\subseteq \left\{ \Delta W_n \le -\frac{2(1 + k_2\theta\Delta)}{k_3} \sqrt{\frac{1 - k_2(2 - \theta)\Delta}{1 + k_2\theta\Delta} - \frac{(k_3)^2 \Delta}{4(1 + k_2\theta\Delta)^2}} \right\} \tag{95}$$

$$\subseteq \left\{ \frac{\Delta W_n}{\sqrt{t - t_n}} \le -\frac{2}{k_3} \frac{\sqrt{(1 - k_2(2 - \theta)\Delta)(1 + k_2\theta\Delta)}}{\sqrt{t - t_n}} \right\}$$

when $\Delta(2 - \theta) < 1/k_2$ and $(k_3)^2/(1 + k_2\theta\Delta) \leq 4k_2$, where $\Delta W_n := W_t - W_{t_n}$. We obtain

$$\mathbb{P}\left(G \leq -\beta\right) = \int_{-\infty}^{-\beta} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du \leq \int_{-\infty}^{-\beta} e^{-u^2/2} du$$
$$= \int_{\beta}^{\infty} e^{-u^2/2} du \leq \frac{1}{\beta} e^{-(\beta)^2/2}, \tag{96}$$

for every standard normal random variable G, where in the last step we have used [17, (9.20), page 112] valid for $\beta > 0$.

Using the fact that $\Delta W_n/\sqrt{t - t_n}$ is a standard normal r.v. and ignoring the exponential term in (96), since its exponent is negative, we get that

$$\mathbb{P}\left(A_1\right) \leq \frac{k_3}{2\sqrt{\left(1 - k_2\left(2 - \theta\right)\Delta\right)}} \sqrt{t - t_n}$$
$$\leq C_{k_2,k_3,\theta,\Delta} \sqrt{\Delta}. \tag{97}$$

The following inclusion relations hold for the event A_2:

$$A_2 \subseteq \left\{ \Delta W_n \leq -\frac{2\left(1 + k_2\theta\Delta\right)}{k_3} \sqrt{y_{t_n} \frac{1 - k_2\left(1 - \theta\right)\Delta}{1 + k_2\theta\Delta} + \frac{k_1\Delta}{1 + k_2\theta\Delta} - \frac{(k_3)^2 \Delta}{4\left(1 + k_2\theta\Delta\right)^2} \left(y_{t_n}\right)^{2q-1}} \left(y_{t_n}\right)^{-q+1/2} \right\}$$

$$\cap \left\{1 > y_{t_n} > \Delta^{1-2\xi}\right\} \subseteq \left\{ \Delta W_n \leq -\frac{2\left(1 + k_2\theta\Delta\right)}{k_3} \sqrt{\Delta^{1-2\xi} \frac{1 - k_2\left(1 - \theta\right)\Delta}{1 + k_2\theta\Delta} + \left(k_1 - \frac{(k_3)^2}{4\left(1 + k_2\theta\Delta\right)}\right) \frac{\Delta}{1 + k_2\theta\Delta}} \right\} \tag{98}$$

$$\subseteq \left\{ \frac{\Delta W_n}{\sqrt{t - t_n}} \leq -\frac{2}{k_3} \frac{\sqrt{\left(1 - k_2\left(1 - \theta\right)\Delta\right)\left(1 + k_2\theta\Delta\right)}}{\sqrt{t - t_n}} \Delta^{1/2-\xi} \right\}$$

when $\Delta(1-\theta) < 1/k_2$ and $(k_3)^2/(1+k_2\theta\Delta) \leq 4k_1$. Using again (96) we have that

$$\mathbb{P}\left(A_2\right) \leq \frac{k_3}{2\sqrt{\left(1 - k_2\left(1 - \theta\right)\Delta\right)}}$$
$$\cdot \sqrt{t - t_n} \Delta^{\xi-1/2} e^{-(2/(k_3)^2)((1-k_2(1-\theta)\Delta)(1+k_2\theta\Delta)/\sqrt{t-t_n})\Delta^{1-2\xi}}$$
$$\leq \frac{k_3}{2\sqrt{\left(1 - k_2\left(1 - \theta\right)\Delta\right)}}$$
$$\cdot \Delta^{\xi} e^{-(2/(k_3)^2)(1-k_2(1-\theta)\Delta)(1+k_2\theta\Delta)\Delta^{-2\xi}}. \tag{99}$$

Taking probabilities in the inclusion relation (93) and using (97) and (99) we get

$$\mathbb{P}\left(A\right) \leq \mathbb{P}\left(A_1\right) + \mathbb{P}\left(A_2\right) \leq C_{k_2,k_3,\theta,\Delta} \sqrt{\Delta}$$
$$+ \frac{k_3}{2\sqrt{\left(1 - k_2\left(1 - \theta\right)\Delta\right)}} \tag{100}$$
$$\cdot \Delta^{\xi} e^{-(2/(k_3)^2)(1-k_2(1-\theta)\Delta)(1+k_2\theta\Delta)\Delta^{-2\xi}} \leq C_{k_2,k_3,\theta,\Delta} \sqrt{\Delta},$$

since $\Delta^{\xi} e^{-\Delta^{-2\xi}} = o(\sqrt{\Delta})$ as $\Delta \downarrow 0$. Finally, note that $C_{k_2,k_3,\theta,\Delta} = k_3/\sqrt{1 - k_2(2 - \theta)\Delta} \to k_3$ as $\Delta \downarrow 0$ which justifies the $O(\sqrt{\Delta})$ notation (see, e.g., [24]). □

6. Numerical Experiments

We discretize the interval $[0, T]$ with a number of steps in power of 2. The semidiscrete (SD) scheme is given by

$$y_{t_{n+1}}^{\text{SD}} = \left(\sqrt{y_{t_n}\left(1 - \frac{k_2\Delta}{1 + k_2\theta\Delta}\right) + \frac{k_1\Delta}{1 + k_2\theta\Delta} - \frac{(k_3)^2 \Delta}{\left(4\left(1 + k_2\theta\Delta\right)^2\right)} \left(y_{t_n}\right)^{2q-1}} + \frac{k_3}{\left(2\left(1 + k_2\theta\Delta\right)\right)} \left(y_{t_n}\right)^{q-1/2} \Delta W_n \right)^2, \tag{101}$$

for $n = 0, \ldots, N - 1$, where $\Delta W_n := W_{t_{n+1}} - W_{t_n}$ are the increments of the Brownian motion which are Gaussian random variables with $\Delta W_n \sim \mathcal{N}(0, \Delta)$.

The ALF (Alfonsi) scheme [4, Section 3] is an implicit scheme which requires solving the nonlinear equation

$$Y_{n+1} = y_{t_n} + (1 - q)\left(k_1\left(Y_{n+1}\right)^{-q/(1-q)} - k_2 Y_{n+1}\right.$$

$$\left. - \frac{q\left(k_3\right)^2}{2} \left(Y_{n+1}\right)^{-1} \right)\Delta + k_3\left(1 - q\right)\Delta W_n, \tag{102}$$

and then computing $y_{t_{n+1}}^{\text{ALF}} = \left(Y_{n+1}\right)^{1/(1-q)}$. The estimation of Y_{n+1} in (102) can be done, for example, with Newton's method but requires a small enough Δ (in the CIR case, that is, when

$q = 1/2$, (102) simplifies to a solution of a quadratic equation). We also consider a scheme recently proposed in [16] using again the SD method, but in a different way,

$$
y_{t_{n+1}}^{\text{HAL}}(q) = \left| \left(y_{t_n}(1 - k_2\Delta) + k_1\Delta \right. \right.
$$
$$
\left. \left. - \frac{q(k_3)^2 \Delta}{2} \left(y_{t_n} \right)^{2q-1} \right)^{1-q} + k_3(1-q) \right. \tag{103}
$$
$$
\left. \cdot \Delta W_n \right|^{1/(1-q)},
$$

for $n = 0, \ldots, N - 1$. Note the similarity in the expressions of (103) and the SD scheme (101) proposed here. This is not strange, because they both rely on the same way of splitting the drift coefficient. In particular, in the explicit HAL scheme, the following process is considered:

$$
y_t^{\text{HAL}}(q) = y_{t_n} + \tilde{f}_1\left(y_{t_n} \right) \cdot \Delta + \int_{t_n}^t \tilde{f}_2(y_s)\, ds
$$
$$
+ \int_{t_n}^t \text{sgn}(z_s)\, \tilde{g}(y_s)\, dW_s, \tag{104}
$$

for $t \in (t_n, t_{n+1}]$ with $y_0 = x_0$ a.s. where now

$$
f(x)
$$
$$
= \underbrace{k_1 - k_2 x - \left(q(k_3)^2/2 \right) x^{2q-1}}_{\tilde{f}_1(x)}
$$
$$
+ \underbrace{\left(q(k_3)^2/2 \right) x^{2q-1}}_{\tilde{f}_2(x)},
$$
$$
\tilde{g}(x) = k_3 x^q, \tag{105}
$$
$$
z_t
$$
$$
= \left(y_{t_n}(1 - k_2\Delta) + k_1\Delta - \frac{q(k_3)^2 \Delta}{2} \left(y_{t_n} \right)^{2q-1} \right)^{1-q}
$$
$$
+ k_3(1-q)\left(W_t - W_{t_n} \right).
$$

A comparison with (3) and (4) shows that $\tilde{f}_2(x) = 2qf_2(x)$ and $\tilde{g}(x) = g(x, x)$, for $\theta = 0$. We write (104) again as

$$
y_t^{\text{HAL}}(q) = y_{t_n} + \left(k_1 - k_2 y_{t_n} - \frac{q(k_3)^2}{2} \left(y_{t_n} \right)^{2q-1} \right)
$$
$$
\cdot \Delta + \int_{t_n}^t \frac{q(k_3)^2}{2} (y_s)^{2q-1}\, ds \tag{106}
$$
$$
+ k_3 \int_{t_n}^t \text{sgn}(z_s)(y_s)^q\, dW_s,
$$

and the process (106) is well defined when

$$
(k_3)^2 \le \frac{2}{q} k_1,
$$
$$
\Delta \le \frac{2}{2k_2 + q(k_3)^2}. \tag{107}
$$

The reader can compare again with (6) for $\theta = 0$. Solving for y_t, we end up with $y_t^{\text{HAL}}(q) = |z_t|^{1/(1-q)}$. The main result in [16] is

$$
\mathbb{E}\left| y_t^{\text{HAL}} - x_t \right|^2 \le C \cdot \Delta^{2q(q-1/2)}, \tag{108}
$$

when (107) holds, implying a rate of convergence at least $q(q - 1/2)$ which is bigger than the rate of convergence of the SD scheme proposed here which is at least $(1/2)(q - 1/2)$ (see Theorem 2).

We also consider two more linear-implicit schemes that were stated in the Introduction and discussed in Appendix B. Namely, we compare with the balanced implicit method (BIM) with appropriate weight functions to guarantee positivity ([25, Theorem 5.9]), which reads

$$
y_{t_{n+1}}^{\text{BIM}}(q) = \frac{y_{t_n} + k_1\Delta + k_3\left(y_{t_n} \right)^q (\Delta W_n + |\Delta W_n|)}{1 + k_2\Delta + k_3\left(y_{t_n} \right)^{q-1} |\Delta W_n|}, \tag{109}
$$

and the balanced Milstein method (BMM) with the suggested weight functions [25, Theorem 5.9] that is given by

$$
y_{t_{n+1}}^{\text{BMM}}(q) = \frac{y_{t_n} + \left(k_1 + (\Theta - 1) k_2 y_{t_n} \right)\Delta + k_3\left(y_{t_n} \right)^q \Delta W_n + \left(q(k_3)^2/2 \right)\left(y_{t_n} \right)^{2q-1}(\Delta W_n)^2}{1 + \Theta k_2\Delta + \left(q(k_3)^2/2 \right)\left| y_{t_n} \right|^{2q-2}\Delta}. \tag{110}
$$

We take the relaxation parameter Θ to be $1/2$ as recommended in [25, (5.10)].

We aim to show experimentally the order of convergence for the above positivity preserving methods for the estimation of the true solution of the CEV model (2), that is, the semidiscrete methods SD (101) and the HAL scheme (103), as well as the implicit ALF scheme (102) and the linear-implicit schemes BIM and BMM. The choice of the parameters is the same as in [15, Figure 6] with $k_3 = 0.4$. In particular $(x_0, k_1, k_2, k_3, q, T) = (1/16, 1/16, 1, 0.4, 3/4, 1)$.

Furthermore, we would also like to reveal the dependence of the order of the semidiscrete methods on q; that is, we want to verify our theoretical results and in particular the order shown in Theorem 2. We take the level of implicitness of the SD method (101) to be $\theta = 1$; that is, we consider the fully implicit scheme. We also discuss the fully explicit scheme, that is, when $\theta = 0$, but also an intermediate scheme $\theta = 1/2$, in Section 7.

We want to estimate the endpoint $\mathscr{L}^2$-norm $\epsilon = \sqrt{\mathbb{E}|y^{(\Delta)}(T) - x_T|^2}$ of the difference between the numerical scheme evaluated at step size Δ and the exact solution of (2). For that purpose, we compute M batches of L simulation paths, where each batch is estimated by $\widehat{\epsilon}_j = (1/L) \sum_{i=1}^{L} |y_{i,j}^{(\Delta)}(T) - y_{i,j}^{(\mathrm{ref})}(T)|^2$ and the Monte Carlo estimator of the error is

$$\widehat{\epsilon} = \sqrt{\frac{1}{ML} \sum_{j=1}^{M} \sum_{i=1}^{L} \left| y_{i,j}^{(\Delta)}(T) - y_{i,j}^{(\mathrm{ref})}(T) \right|^2}, \qquad (111)$$

and requires $M \cdot L$ Monte Carlo sample paths. The reference solution is evaluated at step size 2^{-14} of the numerical scheme. For the SD case, we have shown in Theorems 1 and 2 that it strongly converges to the exact solution. We simulate $100 \cdot 100 = 10000$ paths, where the choice for $L = 100$ is as in [28, page 118]. The choice of the number of trajectories $M \cdot L = 10^4$ is also considered in [26, Section 5] where a fundamental mean-square theorem is proved for SDEs with superlinear growing coefficients satisfying a one-side Lipschitz condition, but unfortunately it is not positivity preserving. Of course, the number of Monte Carlo paths has to be sufficiently large, so as not to significantly hinder the mean-square errors.

We plot in a $\log_2$-$\log_2$ scale and error bars represent 98% confidence intervals. The results are shown in Table 1 and Figure 1. Table 1 does not present the computed Monte Carlo errors with 98% confidence, since they were at least 9 times smaller that the mean-square errors.

In Table 2 we present the computational times (we simulate with 3.06 GHz Intel Pentium, 1.49 GB of RAM in MATLAB *R2014b* Software. The random number generator is Mersenne Twister. The evaluated times do not include the random number generation time, since all the methods we compare involve the same amount of random numbers) of fully implicit SD, HAL, ALF, BIM, and BMM, for the same problem. Figure 2 shows the relation between the error and computer time consumption. As one can see from Table 2 the CPU times for ALF are at least 1000 times bigger than the other schemes; thus we choose in Figure 2 to restrict our attention to the rest of the methods.

We show, in Table 3, the $\mathscr{L}^2$-distance between our proposed method and the other methods for the numerical approximation of (2). We work as before and estimate the distance,

$$d(G, H) = \sqrt{\frac{1}{ML} \sum_{j=1}^{M} \sum_{i=1}^{L} \left| y_{i,j}^{(\Delta, G)}(T) - y_{i,j}^{(\Delta, H)}(T) \right|^2}, \qquad (112)$$

between method G and H, by considering sufficient small Δ, and in particular for $\Delta = 10^{-2}, 10^{-3}, 10^{-4}$.
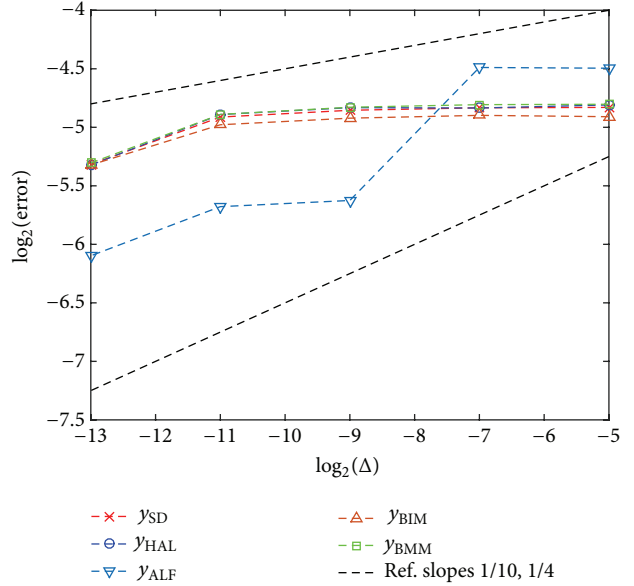

FIGURE 1: Convergence of fully implicit SD, HAL, ALF, BIM, and BMM schemes applied to SDE (2) with parameters $(x_0, k_1, k_2, k_3, q, T) = (1/16, 1/16, 1, 0.4, 3/4, 1)$.
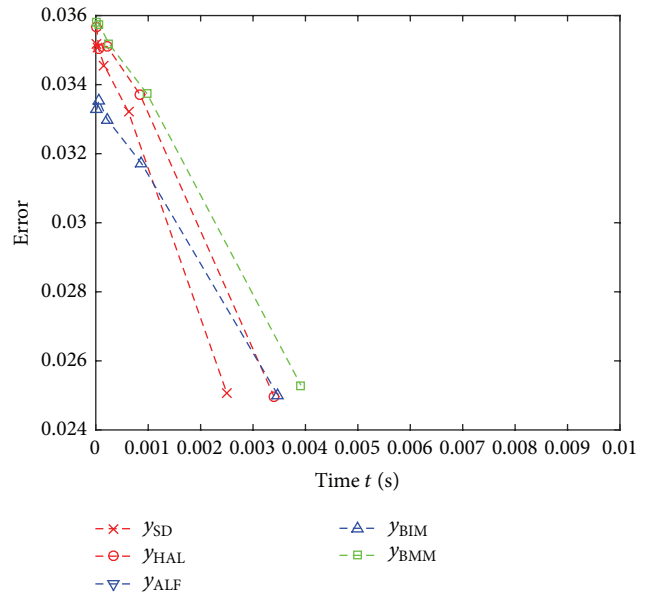

FIGURE 2: Strong convergence error of the mean-reverting CEV process (2) as a function of CPU time (in sec) using positivity preserving schemes SD, HAL, ALF, BIM, and BMM with $(x_0, k_1, k_2, k_3, q, T) = (1/16, 100, 0.05, 1/16, 1, 0.4, 3/4, 1)$.

Finally, we examine the behavior of all the methods for a value of the parameter q close to $1/2$. The results are shown in Table 4.

The following points of discussion are worth mentioning:

(i) The performance of all methods, as shown in Table 1 and Figure 1, implies, in terms of error estimates, that the implicit ALF scheme performs better, for values of discretization steps $\Delta \leq 2^{-9}$. Actually for these

TABLE 1: 98% error and step size of fully implicit SD, HAL, ALF, BIM, and BMM scheme for (2) with $(x_0, k_1, k_2, k_3, q, T) = (1/16, 1/16, 1, 0.4, 3/4, 1)$.

Step Δ	SD ($\theta = 1$)	Rate	HAL	Rate	ALF	Rate	BIM	Rate	BMM	Rate
2^{-5}	0.035161	—	0.03568	—	0.044343	—	**0.033275**	—	0.035819	–
2^{-7}	0.035082	0.0016	0.035034	0.0131	0.044539	−0.0032	**0.033541**	−0.0057	0.035737	0.0017
2^{-9}	0.034565	0.1071	0.035109	−0.0015	**0.020261**	0.5682	0.032988	0.012	0.035192	0.0111
2^{-11}	0.033205	0.029	0.033709	0.029	**0.019532**	0.0264	0.031724	0.0282	0.033731	0.0306
2^{-13}	0.025078	0.2025	0.024978	0.2162	**0.014600**	0.2099	0.024998	0.1719	0.025292	0.2077

TABLE 2: Average computational time (in seconds) for a path, for different discretizations, for all considered positivity preserving methods for the mean-reverting CEV process (2) with $q = 3/4$.

Step Δ	Time/path (in sec): Implicit SD	HAL	ALF	BIM	BMM
2^{-5}	**0.000013**	0.0000164	0.0221883	0.0000174	0.0000196
2^{-7}	**0.0000422**	0.0000558	0.0841705	0.0000584	0.0000657
2^{-9}	**0.0001586**	0.0002137	0.2453943	0.0002207	0.0002482
2^{-11}	**0.0006243**	0.0008437	0.9768619	0.0008703	0.0009795
2^{-13}	**0.0024975**	0.0033977	3.9096332	0.0034785	0.0039143

TABLE 3: The $\mathcal{L}^2$-distance between all the considered numerical schemes applied to SDE (2) with parameter set $(x_0, k_1, k_2, k_3, q, T) = (1/16, 1/16, 1, 0.4, 3/4, 1)$.

Step Δ	98%-d (SD, HAL)	98%-d (SD, ALF)	98%-d (SD, BIM)	98%-d (SD, BMM)
10^{-2}	0.0005727	0.0716140	0.0038373	0.0005312
10^{-3}	0.0001577	0.0286630	0.0013460	0.0001564
10^{-4}	0.0000498	0.0283117	0.0004448	0.0000498

step sizes the ALF method starts to converge, and the same is true for the HAL and BIM methods. All the methods except ALF, that is, the semidiscrete SD and HAL, the BIM, and the BMM have a similar behavior for all values of Δ in the sense of error estimation as Figure 1 shows. The similarity of SD, HAL, BIM, BMM, and especially between SD and BMM is also indicated in Table 3, where we see how close they are with respect to the $\mathcal{L}^2$-norm. Nevertheless, Table 3 also shows that in order to get an accuracy to at least two decimal digits, which in practice may be adequate concerning that we want, for example, to evaluate an option and thus our results are in euros, we are free to use any of the above available methods. We may then choose the fastest one, as will be discussed later on.

(ii) The experimental strong order of convergence of implicit SD for problem (2) is 1/5 (at least $1/2(q − 1/2) = 1/8$ as shown theoretically and presented in Table 1). We also see that all methods converge with similar orders and the theoretically rate 1 of the ALF method [4] does not hold for these values of Δ. Thus, again we see that the rate in practical situations does not necessarily matter, if one has to consider very small values of Δ to achieve it. Moreover, we present

in Table 5 the performance of the explicit SD method and see that it is very close to the implicit, which is of course natural to happen.

(iii) Table 4 concerns the case where the parameter q is 0.55. We do not present the ALF method since it required smaller values of Δ. All the methods again behave quite the same, with the BIM performing better with respect to error estimation.

(iv) In practice, the computer time consumed to provide a desired level of accuracy is of great importance. In particular, in financial applications, a scheme is considered better when, except of its accuracy, it is implemented faster. As mentioned before, the SD method as well as the HAL method performs well in that aspect, compared to the implicit ALF method, which requires the estimation of a root of a nonlinear equation in each step and is therefore time consuming. This is presented in Table 2 and Figure 2 which illustrates the advantage of the semidiscrete method SD, performing slightly better than HAL and BMM, better than BIM, and of course a lot better compared with ALF (over 1000 times quicker to achieve an accuracy of almost two decimal digits). Moreover, the explicit SD performs slightly better in that aspect.

(v) A negative step of a numerical method appears when the computer-generated random variable exceeds a certain threshold, which tends to increase as the step size Δ decreases. Thus, the undesirable effect of negative values that are produced by some numerical schemes (such as the explicit Euler (EM) and standard Milstein (M)) tends to disappear, since, after a certain small step size, the threshold exceeds the maximum standard normal random number attainable by the computer system.

TABLE 4: 98% error and step size of fully implicit SD, HAL, ALF, BIM, and BMM scheme for (2) with $(x_0, k_1, k_2, k_3, q, T) = (1/16, 1/16, 1, 0.4, 0.55, 1)$.

Step Δ	SD-error ($\theta = 1$)	Rate	HAL-error	Rate	BIM-error	Rate	BMM-error	Rate
2^{-10}	0.059129	—	0.059666	—	**0.055819**	—	0.059661	—
2^{-11}	0.058178	0.0234	0.058669	0.0243	**0.055033**	0.0205	0.058667	0.0242
2^{-12}	0.054582	0.092	0.054989	0.0935	**0.051796**	0.0875	0.054987	0.0935
2^{-13}	0.044065	0.3088	0.044288	0.3122	**0.042444**	0.2873	0.044287	0.3122

TABLE 5: The performance of fully explicit SD scheme (101) applied to SDE (2) with parameter set $(x_0, k_1, k_2, k_3, q, T) = (1/16, 1/16, 1, 0.4, 3/4, 1)$.

Step Δ	98% SD-error ($\theta = 0$)	Rate
2^{-7}	0.0344244	—
2^{-9}	0.0342415	0.0038
2^{-11}	0.0331273	0.0239
2^{-13}	0.0250195	0.2025

7. Approximation of Stochastic Model (1)

So far we have focused on the process (V_t), which is one part of system (1). Nevertheless, it can be treated independently, since the only way that it interacts with the process (S_t) is through the correlation ρ of the Wiener processes. First we apply Itô's formula on $\ln(S_t)$ to get

$$
\ln S_t = \ln S_0 + \int_0^t \mu \, du - \frac{1}{2} \int_0^t \left(V_u\right)^{2p} du \\
+ \int_0^t \left(V_u\right)^p dW_u, \quad t \in [0, T].
\tag{113}
$$

Then, we consider two different schemes for the integration of (113) (the reason for not considering other schemes such as the two-dimensional Milstein is that they generally are time consuming, since they involve additional random number generation for the approximation of double Wiener integrals). The first is the EM scheme which reads

$$
\ln S_{t_{n+1}}^{\text{EM}} = \ln S_{t_n} + \mu \Delta - \frac{1}{2} \left(V_{t_n}\right)^{2p} \Delta + \left(V_{t_n}\right)^p \Delta W_n,
\tag{114}
$$

which has strong convergence order 1/2 and is easy to implement. The second scheme, which is based on an interpolation of the drift term and an interpolation of the diffusion term, considering decorrelation of the diffusion term, including a higher order Milstein term [15, Section 4.2], is denoted IJK and is given by [15, (137)]

$$
\ln S_{t_{n+1}}^{\text{IJK}} = \ln S_{t_n} + \mu \Delta - \frac{1}{4} \left(\left(V_{t_n}\right)^{2p} + \left(V_{t_{n+1}}\right)^{2p} \right) \Delta \\
+ \rho \left(V_{t_n}\right)^p \Delta \widetilde{W}_n \\
+ \frac{1}{2} \left(\left(V_{t_n}\right)^p + \left(V_{t_{n+1}}\right)^p \right) \left(\Delta W_n - \rho \Delta \widetilde{W}_n \right) \\
+ \frac{1}{2} \rho p k_3 \left(V_{t_n}\right)^{q+p-1} \left(\left(\Delta \widetilde{W}_n\right)^2 - \Delta \right).
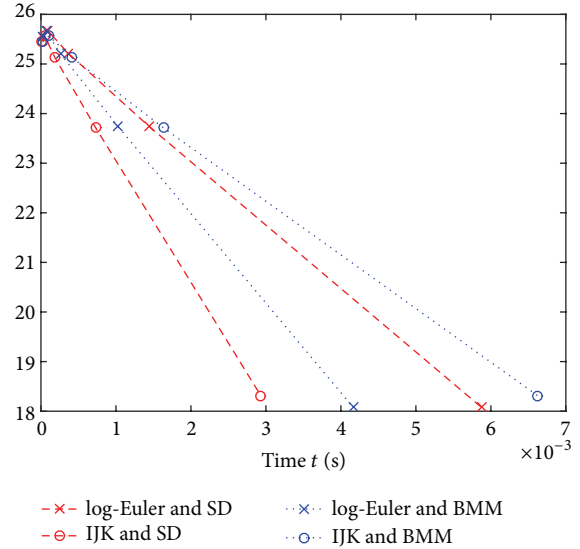\tag{115}
$$

FIGURE 3: Strong convergence error of the financial underlying process (S_t), as a function of CPU time (in sec) using log-Euler or IJK method with SD or BMM scheme for (1) with $(x_0, S_0, \mu, k_1, k_2, k_3, q, T) = (1/16, 100, 0.05, 1/16, 1, 0.4, 3/4, 1)$, correlation $\rho = -0.8$.

We therefore consider the EM scheme (114) combined with SD (101) and the IJK scheme (115) combined with SD (101) and compare with the case where the stochastic variance ($p = 1/2$) is integrated with BMM scheme (110), for three different correlation parameters, $\rho = 0, \rho = -0.4$, and $\rho = -0.8$ with $S_0 = 100, \mu = 0.05$, as in [15, Section 5]. We present in Tables 6, 7, and 8 the errors, in the sense of distance (112), for all the above considered ways of numerical integration of process (S_t), for different step sizes, as well as the average computational time (in seconds) consumed for each discretization. We also give an illustrative representation just for one case ($\rho = -0.8$) in Figure 3.

Tables 6, 7, and 8 indicate that in all cases the favorable choice is to integrate (S_t) using IJK method combined with the SD scheme for (V_t) in model (1). The IJK-SD approximation of system (1) seems to be the better one, with respect to CPU time, for every correlation coefficient considered.

8. Conclusion

In this paper, we exploit further the semidiscrete method (SD), which originally appeared in [8], to numerically approximate stochastic processes that appear in financial

TABLE 6: 98% error, step size, and average computational time of numerical integration of process (S_t) using log-Euler or IJK method with SD or BMM scheme for (1) with $(x_0, S_0, \mu, k_1, k_2, k_3, q, T) = (1/16, 100, 0.05, 1/16, 1, 0.4, 3/4, 1)$ and correlation $\rho = 0$.

Step Δ	EM and SD-error ($\theta = 0.5$)	IJK and SD-error ($\theta = 0.5$)	EM and BMM-error ($\Theta = 0.5$)	IJK and BMM-error ($\Theta = 0.5$)
2^{-5}	26.901 (0.0000261)	26.901 (**0.0000159**)	26.891 (0.00002)	26.890 (0.0000294)
2^{-7}	27.288 (0.0000919)	27.288 (**0.0000492**)	27.277 (0.0000676)	27.277 (0.0001043)
2^{-9}	27.298 (0.0003595)	27.297 (**0.0001843**)	27.289 (0.0002610)	27.288 (0.0004081)
2^{-11}	25.057 (0.0014255)	25.058 (**0.0007309**)	25.051 (0.0010309)	25.051 (0.0016191)
2^{-13}	19.441 (0.0057322)	19.441 (**0.0028928**)	19.442 (0.0041177)	19.442 (0.0064721)

TABLE 7: 98% error, step size, and average computational time of numerical integration of process (S_t) using log-Euler or IJK method with SD or BMM scheme for (1) with $(x_0, S_0, \mu, k_1, k_2, k_3, q, T) = (1/16, 100, 0.05, 1/16, 1, 0.4, 3/4, 1)$, correlation $\rho = -0.4$.

Step Δ	EM and SD-error ($\theta = 0.5$)	IJK and SD-error ($\theta = 0.5$)	EM and BMM-error ($\Theta = 0.5$)	IJK and BMM-error ($\Theta = 0.5$)
2^{-5}	26.382 (0.0000266)	26.331 (**0.0000161**)	26.372 (0.0000202)	26.324 (0.00003)
2^{-7}	26.448 (0.0000951)	26.396 (**0.000005**)	26.439 (0.0000691)	26.389 (0.0001081)
2^{-9}	25.951 (0.0003631)	25.909 (**0.000184**)	25.944 (0.0002606)	25.904 (0.0004131)
2^{-11}	24.540 (0.0014506)	24.494 (**0.0007355**)	24.531 (0.0010378)	24.486 (0.0016495)
2^{-13}	18.738 (0.0060748)	18.749 (**0.0030185**)	18.735 (0.0042868)	18.747 (0.0068395)

TABLE 8: 98% error, step size, and average computational time of numerical integration of process (S_t) using log-Euler or IJK method with SD or BMM scheme for (1) with $(x_0, S_0, \mu, k_1, k_2, k_3, q, T) = (1/16, 100, 0.05, 1/16, 1, 0.4, 3/4, 1)$, correlation $\rho = -0.8$.

Step Δ	EM and SD-error ($\theta = 0.5$)	IJK and SD-error ($\theta = 0.5$)	EM and BMM-error ($\Theta = 0.5$)	IJK and BMM-error ($\Theta = 0.5$)
2^{-5}	25.552 (0.0000263)	25.455 (**0.0000159**)	25.541 (0.0000199)	25.449 (0.0000296)
2^{-7}	25.670 (0.0000932)	25.569 (**0.0000494**)	25.659 (0.0000678)	25.564 (0.0001059)
2^{-9}	25.217 (0.0003622)	25.137 (**0.0001835**)	25.208 (0.0002595)	25.132 (0.0004111)
2^{-11}	23.743 (0.0014407)	23.711 (**0.0007306**)	23.734 (0.0010307)	23.707 (0.0016376)
2^{-13}	18.082 (0.005871)	18.316 (**0.0029312**)	18.078 (0.0041637)	18.312 (0.0066239)

mathematics and are meant to be nonnegative. In [23] we examined the Heston 3/2-model, that is, a mean-reverting process with superlinear diffusion, described by a SDE of the form (2) with $q = 3/2$. Now, we deal with SDEs with sublinear diffusion coefficients of the type $(x_t)^q$ with $1/2 < q < 1$. These kinds of SDEs, called mean-reverting CEV processes, appear in stochastic models, where they represent the instantaneous volatility-variance of an underlying financially observable variable. We prove theoretically the strong convergence of our proposed SD scheme, revealing the order of convergence. The resulting polynomial rate is shown in Theorem 1. We present a comparative study between various positivity preserving schemes and the SD method seems to be the best with respect to CPU time consumption. The advantage of the SD method here is that although implicit, it has an explicit formula and thus requires fewer arithmetic operations and consequently less computational time. Moreover, our method can cover cases where (2) has time varying coefficients, that is, $k_1(t)$, $k_2(t)$, $k_3(t)$.

We also treat the two-dimensional stochastic volatility model (1). In order to do that, we actually integrate the process $\ln(S_t)$ which satisfies a SDE of the form (113) and in the end transform back for (S_t). We only consider two different schemes for the integration of $\ln(S_t)$, namely, the Euler Maruyama (EM) scheme, which is easy to implement and the IJK scheme [15, (137)] which is shown to be the most efficient method, robust and simple as EM [15]. We

do not apply other two-dimensional schemes, such as, for example, the Milstein scheme, since they are in general time consuming, as they involve approximations of double Wiener integrals which require additional random number generation. We therefore combine the EM scheme with SD ((114) and (101)) and the IJK scheme with SD ((115) and (101)) and compare with the case where the stochastic variance ($p = 1/2$) is integrated with BMM scheme (110), for three different correlation parameters, $\rho = 0$, $\rho = -0.4$, and $\rho = -0.8$ with $S_0 = 100$, $\mu = 0.05$, as in [15, Section 5]. The combination IJK with SD seems to be the most favorable with respect to CPU time, for all the cases.

Appendices

A. Boundary Classification of One-Dimensional Time-Homogeneous SDEs

Let us now recall some results [17, Section 5.5] concerning the boundary behavior of SDEs of the form

$$dX_t = a(X_t) dt + b(X_t) dW_t. \tag{A.1}$$

Let $I = (l, r)$ be an interval with $-\infty \le l < r \le \infty$ and define the exit time from I to be

$$S := \inf\{t \ge 0 : X_t \notin (l, r)\}. \tag{A.2}$$

Let also the coefficients of (A.1) satisfy the following conditions:

Nondegeneracy (ND)

$$b^2(x) > 0, \quad \forall x \in I. \tag{A.3}$$

Local Integrability (LI)

$$\int_{x-\epsilon}^{x+\epsilon} \frac{1 + |a(y)|}{b^2(y)} dy < \infty, \quad \forall x \in I, \ \exists \epsilon > 0. \tag{A.4}$$

Then for $c \in I$, we can define the scale function

$$s(x) := \int_c^x e^{-2\int_c^y (a(z)/b^2(z))dz} dy, \tag{A.5}$$

whose behavior at the endpoints of I determines the boundary behavior of (X_t) [17, Proposition 5.22]. In particular, we get for the dynamics of the mean-reverting CEV process (V_t) of (1) a boundary behavior which is determined by the scale function

$$s(x) = \int_c^x \exp\left\{ -2\int_c^y \frac{k_1 - k_2 z}{(k_3)^2 z^{2q}} dz \right\} dy$$

$$= C \int_c^x \exp\left\{ -\frac{2k_1}{(k_3)^2 (1 - 2q)} y^{1-2q} \right. \tag{A.6}$$

$$\left. + \frac{2k_2}{(k_3)^2 (2 - 2q)} y^{2-2q} \right\} dy,$$

for any $x \in I$, where $C > 0$. Let $I = (0, \infty)$ and take $c = 1$. We compute

$$s(0+) = -C \int_0^1 \exp\left\{ -\frac{2k_1}{(k_3)^2 (1 - 2q)} y^{1-2q} \right.$$

$$\left. + \frac{2k_2}{(k_3)^2 (2 - 2q)} y^{2-2q} \right\} dy = -\infty, \tag{A.7}$$

when $2q > 1$; thus by [17, Proposition 5.22c] we have that $\mathbb{P}(\inf_{0 \le t} V_t > 0) = 1$.

B. Some Numerical Schemes for the Integration of the Variance-Volatility Process (V_t)

We consider a partition of the time interval $[0, T]$ with $0 = t_0 < t_1 < \cdots < t_N = T$ and discretization steps $\Delta_n := t_{n+1} - t_n$ for $n = 0, \dots, N-1$. Moreover, we denote by $\Delta W_n := W_{t_{n+1}} - W_{t_n}$ the increments of the Brownian motion. We show in the following subsections some numerical schemes for the approximation of

$$V_t = V_0 + \int_0^t (k_1 - k_2 V_s) ds + \int_0^t k_3 (V_s)^q dW_s, \tag{B.1}$$

$$t \in [0, T]$$

and make some brief comments on them. We also denote $V_n := V_{t_n}$.

B.1. Standard Euler-Maruyama Scheme. The Euler method, applied to the SDE setting, already appeared in the 50s through Maruyama [27] and thereafter there has been an extensive study on numerical approximations of solutions of SDEs (we just mention [12] for a recent review on numerical methods for SDEs with applications in finance and references therein).

The explicit Euler-Maruyama (EM) scheme for the process (V_t) is given by

$$V_{n+1}^{\text{EM}} = V_n + (k_1 - k_2 V_n) \Delta_n + k_3 (V_n)^q \Delta W_n, \tag{B.2}$$

for $n = 0, \dots, N-1$. Clearly $\mathbb{P}(V_{n+1} < 0 \mid V_n > 0) > 0$; thus the EM scheme can produce negative values with positive probability, or in the notion of [29] we say that (B.2) has a finite life time.

B.2. Standard Milstein Scheme. The standard one-dimensional Milstein (M) scheme contains some extra terms derived by Itô-Taylor expansion [1, Section 5] and applied to (V_t) reads

$$V_{n+1}^M = V_n + (k_1 - k_2 V_n) \Delta_n + k_3 (V_n)^q \Delta W_n$$

$$+ \frac{1}{2} (k_3)^2 q (V_n)^{2q-1} \left((\Delta W_n)^2 - \Delta_n \right), \tag{B.3}$$

for $n = 0, \dots, N-1$ where we have retained terms of order (Δ_n). Again (M) scheme has a finite life time.

B.3. Balanced Implicit Method. The balanced implicit method (BIM) [11, (3.2)] was the first attempt to treat the problem of invariance-preserving of specific domains of the underlying process and reads

$$V_{n+1}^{\text{BIM}} = V_n + (k_1 - k_2 V_n) \Delta_n + k_3 (V_n)^q \Delta W_n$$

$$+ \left(c^0 (V_n) \Delta_n + c^1 (V_n) |\Delta W_n| \right) (V_n - V_{n+1}), \tag{B.4}$$

for $n = 0, \dots, N-1$ where c^0 and c^1 are appropriate weight functions. The choice $c^0(x) = k_2$ and $c^1(x) = k_3 x^{q-1}$ preserves positivity [25, Section 5]. Rearranging the above equation, we get the expression

$$V_{n+1}^{\text{BIM}} = \frac{V_n + k_1 \Delta_n + k_3 (V_n)^q (\Delta W_n + |\Delta W_n|)}{1 + k_2 \Delta_n + k_3 (V_n)^{q-1} |\Delta W_n|}. \tag{B.5}$$

B.4. Balanced Milstein Method. The balanced Milstein method (BMM) was proposed in [25], for an improvement of the BIM in the stability behavior as well as in the rate of convergence. It is given by the following linear-implicit relation:

$$V_{n+1}^{\text{BMM}} = V_n + (k_1 - k_2 V_n) \Delta_n + k_3 (V_n)^q \Delta W_n$$

$$+ \frac{1}{2} (k_3)^2 q (V_n)^{2q-1} \left((\Delta W_n)^2 - \Delta_n \right)$$

$$+ \left(d^0 (V_n) \Delta_n + d^1 (V_n) \left((\Delta W_n)^2 - \Delta_n \right) \right) \tag{B.6}$$

$$\cdot (V_n - V_{n+1}),$$

for $n = 0, \ldots, N - 1$ where d^0 and d^1 are appropriate weight functions. The choice $d^0(x) = \Theta k_2 + (1/2)(k_3)^2 q|x|^{2q-2}$, where $\Theta \in [0, 1]$ and $d^1(x) = 0$ implies an eternal life time for the scheme [25, Theorem 5.9], in the sense that $\mathbb{P}(V_{n+1} > 0 \mid V_n > 0) = 1$. The step sizes Δ_n have to be such that

$\Delta_n < (2q - 1)/2qk_2(1 - \Theta)$. The relaxation parameter resembles the implicitness parameter (θ in our notation). For $\Theta = 1$ there is no restriction in the step size, but it is recommended when possible [25, Remark 5.10] to take $\Theta = 1/2$. Rearranging with the above specifications leads to

$$V_{n+1}^{\text{BMM}} = \frac{V_n + (k_1 - (1 - \Theta) k_2 V_n) \Delta_n + k_3 (V_n)^q \Delta W_n + (1/2) (k_3)^2 q (V_n)^{2q-1} (\Delta W_n)^2}{1 + \Theta k_2 \Delta_n + (1/2) (k_3)^2 q |V_n|^{2q-2} \Delta_n}. \tag{B.7}$$

Finally, the proposed semidiscrete (SD) scheme reads

$$V_{n+1}^{\text{SD}} = \left(\sqrt{V_n \left(1 - \frac{k_2 \Delta}{1 + k_2 \theta \Delta} \right) + \frac{k_1 \Delta}{1 + k_2 \theta \Delta} - \frac{(k_3)^2 \Delta}{4 (1 + k_2 \theta \Delta)^2} (V_n)^{2q-1}} + \frac{k_3}{2 (1 + k_2 \theta \Delta)} (V_n)^{q-1/2} \Delta W_n \right)^2. \tag{B.8}$$

Increasing the time horizon T results in an increase of the percentage of negative paths of EM and M. On the other hand BIM, BMM, and of course SD are not affected by that, since they preserve their positivity on any interval $[0, T]$.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

References

[1] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, vol. 23, Springer, Berlin, Germany, 1995, Corrected 2nd printing.

[2] M. J. Brennan and E. S. Schwartz, "Analyzing convertible bonds," *The Journal of Financial and Quantitative Analysis*, vol. 15, no. 4, pp. 907–929, 1980.

[3] K. C. Chan, G. A. Karolyi, F. A. Longstaff, and A. B. Sanders, "An empirical comparison of alternative models of the short-term interest rate," *The Journal of Finance*, vol. 47, no. 3, pp. 1209–1227, 1992.

[4] A. Alfonsi, "Strong order one convergence of a drift implicit Euler scheme: application to the CIR process," *Statistics & Probability Letters*, vol. 83, no. 2, pp. 602–607, 2013.

[5] N. Halidias, "A new numerical scheme for the CIR process," *Monte Carlo Methods and Applications*, vol. 21, no. 3, pp. 245–253, 2015.

[6] J. C. Cox, J. E. Ingersoll, and S. A. Ross, "A theory of the term structure of interest rates," *Econometrica*, vol. 53, no. 2, pp. 385–407, 1985.

[7] L. B. G. Andersen and V. V. Piterbarg, "Moment explosions in stochastic volatility models," *Finance and Stochastics*, vol. 11, no. 1, pp. 29–50, 2007.

[8] N. Halidias, "Semi-discrete approximations for stochastic differential equations and applications," *International Journal of Computer Mathematics*, vol. 89, no. 6, pp. 780–794, 2012.

[9] N. Halidias, "A novel approach to construct numerical methods for stochastic differential equations," *Numerical Algorithms*, vol. 66, no. 1, pp. 79–87, 2014.

[10] T. R. Hurd and A. Kuznetsov, "Explicit formulas for Laplace transforms of stochastic integrals," *Markov Processes and Related Fields*, vol. 14, no. 2, pp. 277–290, 2008.

[11] G. N. Milstein, E. Platen, and H. Schurz, "Balanced implicit methods for stiff stochastic systems," *SIAM Journal on Numerical Analysis*, vol. 35, no. 3, pp. 1010–1019, 1998.

[12] P. Kloeden and A. Neuenkirch, "Convergence of numerical methods for stochastic differential equations in mathematical finance," in *Recent Developments in Computational Finance: Foundations, Algorithms and Applications*, pp. 49–80, 2013.

[13] J. Alcock and K. Burrage, "A note on the balanced method," *BIT Numerical Mathematics*, vol. 46, no. 4, pp. 689–710, 2006.

[14] J. Alcock and K. Burrage, "Stable strong order 1.0 schemes for solving stochastic ordinary differential equations," *BIT Numerical Mathematics*, vol. 52, no. 3, pp. 539–557, 2012.

[15] C. Kahl and P. Jäckel, "Fast strong approximation Monte Carlo schemes for stochastic volatility models," *Quantitative Finance*, vol. 6, no. 6, pp. 513–536, 2006.

[16] N. Halidias, "An explicit and positivity preserving numerical scheme for the mean reverting CEV model," *Japan Journal of Industrial and Applied Mathematics*, vol. 32, no. 2, pp. 545–552, 2015.

[17] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, Springer, New York, NY, USA, 1988.

[18] T. H. Gronwall, "Note on the derivatives with respect to a parameter of the solutions of a system of differential equations," *The Annals of Mathematics*, vol. 20, no. 4, pp. 292–296, 1919.

[19] T. Yamada and S. Watanabe, "On the uniqueness of solutions of stochastic differential equations," *Journal of Mathematics of Kyoto University*, vol. 11, no. 1, pp. 155–167, 1971.

[20] X. Mao, *Stochastic Differential Equations and Applications*, Horwood Publishing, Chichester, UK, 1997.

[21] A. Berkaoui, "Euler scheme for solutions of stochastic differential equations with non-Lipschitz coefficients," *Portugaliae Mathematica*, vol. 61, no. 4, pp. 461–478, 2004.

[22] S. Dereich, A. Neuenkirch, and L. Szpruch, "An Euler-type method for the strong approximation of the cox-ingersoll-ross process," *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 468, no. 2140, pp. 1105–1115, 2012.

[23] N. Halidias and I. S. Stamatiou, "On the numerical solution of some non-linear stochastic differential equations using the semi-discrete method," *Computational Methods in Applied Mathematics*, 2015.

[24] F. W. J. Olver, *Asymptotics and Special Functions*, AKP Classics, A K Peters, Wellesley, Mass, USA, 1997.

[25] C. Kahl and H. Schurz, "Balanced Milstein methods for ordinary SDEs," *Monte Carlo Methods and Applications*, vol. 12, no. 2, pp. 143–170, 2006.

[26] M. V. Tretyakov and Z. Zhang, "A fundamental mean-square convergence theorem for SDEs with locally Lipschitz coefficients and its applications," *SIAM Journal on Numerical Analysis*, vol. 51, no. 6, pp. 3135–3162, 2013.

[27] G. Maruyama, "Continuous Markov processes and stochastic equations," *Rendiconti del Circolo Matematico di Palermo*, vol. 4, no. 1, pp. 48–90, 1955.

[28] P. E. Kloeden, E. Platen, and H. Schurz, *Numerical Solution of SDE through Computer Experiments*, Springer, 2003.

[29] H. Schurz, "Numerical regularization for SDEs: construction of nonnegative solutions," *Dynamic Systems and Applications*, vol. 5, no. 3, pp. 323–352, 1996.