

## CONVERGENCE, NON-NEGATIVITY AND STABILITY OF A NEW MILSTEIN SCHEME WITH APPLICATIONS TO FINANCE

DESMOND J. HIGHAM AND XUERONG MAO

Department of Mathematics and Statistics  
University of Strathclyde  
Glasgow, G1 1XH, UK

LUKASZ SZPRUCH  
School of Mathematics  
The University of Edinburgh  
Edinburgh, EH9 3JZ, UK

(Communicated by Rachel Kuske)

**ABSTRACT.** We propose and analyse a new Milstein type scheme for simulating stochastic differential equations (SDEs) with highly nonlinear coefficients. Our work is motivated by the need to justify multi-level Monte Carlo simulations for mean-reverting financial models with polynomial growth in the diffusion term. We introduce a double implicit Milstein scheme and show that it possesses desirable properties. It converges strongly and preserves non-negativity for a rich family of financial models and can reproduce linear and nonlinear stability behaviour of the underlying SDE without severe restriction on the time step. Although the scheme is implicit, we point out examples of financial models where an explicit formula for the solution to the scheme can be found.

**1. Introduction.** The majority of research on numerical methods for stochastic differential equations (SDEs) imposes restrictive global Lipschitz conditions on the drift and diffusion coefficients [20]. By contrast, many SDEs encountered in practical applications do not fall into this category. This motivates the development of new approximation methods for highly nonlinear SDEs. In addition to standard convergence analysis, it is important to study stability properties of numerical approximations (which are closely related to error propagation in time), and their ability to preserve qualitative properties of continuous time SDEs. For example, when modelling asset prices we are typically restricted to SDEs that have non-negative solutions. It is then desirable for numerical approximations to share this property. In this paper we propose a very simple strategy for generating convergent numerical approximations with good stability and positivity preservation. The key idea is to rewrite the original Itô SDE in Stratonovich form and then apply drift-implicit Milstein scheme.

In what follows, we study numerical approximation of the scalar Ito stochastic differential equation (SDE)

$$dx(t) = f(x(t))dt + g(x(t))dw(t). \quad (1)$$

---

2010 *Mathematics Subject Classification.* Primary: 60H10, 60H35; Secondary: 65J15.

*Key words and phrases.* Milstein scheme, implicit schemes stochastic differential equation, stability, strong convergence, non-negativity.

Here  $x(t) \in \mathbb{R}$  for each  $t \geq 0$ , and, for simplicity,  $x(0)$  is taken to be constant. We assume that  $f \in C^1(\mathbb{R}, \mathbb{R})$  and  $g \in C^2(\mathbb{R}, \mathbb{R})$ . Throughout we let  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$  be a complete probability space with a filtration  $\{\mathcal{F}_t\}_{t \geq 0}$  and by  $w(t)$  we denote Brownian motion defined on the probability space.

An important example of a SDE with non-globally Lipschitz continuous coefficients is the Heston stochastic volatility 3/2-model [11, 21]:

$$dx(t) = x(t)(\mu - \alpha x(t))dt + \beta x(t)^{3/2}dw(t), \quad \mu, \alpha, \beta > 0. \quad (2)$$

This equation is also used for modelling term structure dynamics [2], and can be viewed as a stochastic extension to the logistic equation [6]. Its super-linearly growing coefficients, however, raise many new questions. Recently [14] demonstrated that the standard Euler-Maruyama (EM) discretization scheme can diverge in strong and weak senses for SDEs with super-linear coefficients. However, in [26, 27] it is shown that an implicit Euler-type method strongly converges for non-linearities similar to (2). These positive results rely heavily on a one-sided Lipschitz condition satisfied by the drift coefficients of the SDE (1), that is, for some constant  $K$ ,

$$(x - y)(f(x) - f(y)) \leq K|x - y|^2, \quad \text{for } x, y \in \mathbb{R}. \quad (3)$$

The solution of (2) is non-negative and condition (3) holds in the relevant region  $x, y \geq 0$ . However, in general, numerical approximations do not preserve non-negativity and hence convergence theorems developed in [26, 27] cannot be applied in this situation. Similarly, convergence of an alternative Euler scheme that was considered in [15] also was established under one-sided Lipschitz conditions.

Preservation of positivity is a desirable modeling property, and, in many cases, non-negativity of the numerical approximation is needed in order for the scheme to be well defined. For example, evaluating the diffusion coefficient in (2) for a negative argument does not make sense. Many fixes have been proposed in the literature, but these can lead to substantial bias [23]. For more information about positivity preservation we refer the reader to [4, 17, 30, 32, 35]. It was shown in [18] that a class of balanced methods can preserve positivity under an appropriate choice of the weight functions, but strong convergence was proved only under a global Lipschitz condition [28]. Kahl et al. [17] proved that the EM scheme does not preserve positivity for any scalar SDE. In the case of the drift implicit EM scheme positivity can be preserved, [35], if the drift coefficient has a very special form, for example as in the Ait-Sahalia interest rate model [3]. It was also shown in [17] that the drift implicit Milstein scheme applied to

$$dx(t) = \alpha(\mu - x(t))dt + \beta x(t)^p dw(t) \quad \alpha, \mu, \beta > 0, \text{ and } p \in [0.5, 1], \quad (4)$$

preserves positivity with no restriction on the time step if  $p = 0.5$ . For  $p \in (0.5, 1]$  we need to restrict the step size to lie below  $\beta^{-2}$ .

Higher order approximation carries some pitfalls. It was demonstrated in [12] that the Milstein scheme applied to a linear scalar SDE has worse stability properties than EM, even once we allow for implicitness in the drift coefficient. This is undesirable, particularly in the multi-level Monte Carlo (MLMC) setting [1, 7, 8], where we are required to use many simulations with large discretization time step. It is therefore natural to look for more advanced numerical techniques that automatically capture such a property. Some attempts to address this issue were already suggested in literature. Stability analysis of the implicit Milstein methods for the linear SDEs was presented [5] and for SDEs with Lipschitz continuous coefficients

was presented in [36]. Our analysis covers a much richer family of SDEs and it seems (at least in the scalar case) that our scheme is more tractable than the one developed in [36].

In order to address the issues mentioned above, we introduce a new  $(\theta, \sigma)$ -Milstein scheme for a general scalar SDE. Given any step size  $\Delta t$ , we define the partition  $\mathcal{P}_{\Delta t} := \{t_k = k\Delta t : k = 0, 1, 2, \dots\}$  of the half line  $[0, \infty)$ . Letting  $X_{t_k}$  denote the approximation to  $x(t_k)$ , with  $X_{t_0} = x(0)$  and  $\Delta w_{t_k} = w(t_{k+1}) - w(t_k)$ , the  $(\theta, \sigma)$ -Milstein-scheme then has the following form

$$\begin{aligned}
 X_{t_{k+1}} = & X_{t_k} + \theta f(X_{t_{k+1}})\Delta t + (1 - \theta)f(X_{t_k})\Delta t + g(X_{t_k})\Delta w_{t_k} + \frac{1}{2}L^1g(X_{t_k})\Delta w_{t_k}^2 \\
 & - \frac{(1 - \sigma)}{2}L^1g(X_{t_k})\Delta t - \frac{\sigma}{2}L^1g(X_{t_{k+1}})\Delta t,
 \end{aligned} \tag{5}$$

where  $0 \leq \theta, \sigma \leq 1$  are free parameters and  $L^1 = g \frac{\partial}{\partial x}$ . We note that the  $(0, 0)$ -Milstein scheme reduces to classical Milstein approximation [29]. We will sometimes refer to  $(1, 1)$ -Milstein as the double implicit scheme. The advantage of the extra degree of implicitness offered by  $\sigma$  will become clear as we analyse the method. We note that the  $(\theta, \sigma)$ -Milstein scheme can be naturally derived from the Stratonovich-Taylor expansion. Indeed, we can rewrite SDE (1) into its Stratonovich form

$$dx(t) = \underline{f}(x(t))dt + g(x(t)) \circ dw(t),$$

where  $\underline{f}(x) = f(x) - L^1g(x)$ . In the scalar case the drift-implicit Milstein scheme for the Stratonovich SDE is given by (see [20] p. 345)

$$\bar{X}_{t_{k+1}} = \bar{X}_{t_k} + \underline{f}(\bar{X}_{t_{k+1}})\Delta t + g(\bar{X}_{t_k})\Delta w_{t_k} + \frac{1}{2}L^1g(\bar{X}_{t_k})\Delta w_{t_k}^2.$$

Hence, we note that  $(\theta, \sigma)$ -Milstein scheme may be obtained from the implicit Milstein scheme for a Stratonovich SDE.

In this work, we allow for a nonlinear drift coefficient and show that once  $p > 0.5$  in (4) the step size restriction for non-negativity can be eliminated by the  $(\theta, \sigma)$ -Milstein method. We also present fairly general conditions for a family of Milstein schemes to preserve positivity. Due to that property the one-sided Lipschitz structure (3) is preserved. Hence, the new scheme can be shown to converge strongly to the solution of the SDE (2). Numerically we observe that the rate of strong convergence is 1, which Giles [7, 8] has shown to be the optimal rate from the MLMC perspective.

The material is structured as follows. Section 2 contains proofs of the existence of positive solutions to (5) under appropriate conditions. In Section 3 we consider stability properties of the double implicit scheme. As motivation we demonstrate that for linear test SDEs we can significantly improve stability properties of the Milstein scheme. We then extend this result to a more general nonlinear setting. In Section 4 we develop the convergence results. We give conclusions in Section 5.

**2. Existence of a solution for the implicit schemes.** We begin with conditions that guarantee the existence of a unique solution to (5). These will motivate the assumptions that we use to force positivity.

**Lemma 2.1.** *Let  $F$  be a function defined on  $\mathbb{R}$  and consider the equation*

$$F(x) = b, \tag{6}$$

for a given  $b \in \mathbb{R}$ . If  $F$  is strictly monotone, i.e.,

$$(x - y)(F(x) - F(y)) > 0, \quad (7)$$

for all  $x, y \in \mathbb{R}$ ,  $x \neq y$ , then equation (6) has at most one solution. If  $F$  is continuous and coercive, i.e.,

$$\lim_{|x| \rightarrow \infty} \frac{x F(x)}{|x|} = \infty, \quad (8)$$

then for every  $b \in \mathbb{R}$ , equation (6) has a solution  $x \in \mathbb{R}$ . Moreover, the inverse operator  $F^{-1}$  exists.

A proof follows directly from Theorem 26.A in [37]. In order to prove that the  $(\theta, \sigma)$ -Milstein (5) scheme is well defined we impose two conditions.

**Assumption 2.2.** Coefficients  $f$  and  $g$  in (1) are locally Lipschitz continuous and satisfy the following two conditions:

One-sided Lipschitz condition. There exists a constant  $K > 0$  such that

$$(x - y)(f(x) - f(y)) \leq K|x - y|^2 \quad \text{for all } x, y \in \mathbb{R}. \quad (9)$$

Monotone condition. Operator  $L^1$  acting on  $g$  satisfies

$$(x - y)(L^1 g(x) - L^1 g(y)) \geq 0 \quad \text{for all } x, y \in \mathbb{R}. \quad (10)$$

**Remark 1.** From Assumption 2.2 and Young's inequality we may show that the drift coefficient  $f$  satisfies a one-sided Lipschitz-type condition

$$x f(x) \leq K|x|^2 + x f(0) \leq a + b|x|^2 \quad \text{for all } x \in \mathbb{R},$$

where  $a = 0.5|f(0)|^2$  and  $b = (2K + 1)/2$ . Also from Assumption 2.2 we can show that  $xL^1 g(x)$  is bounded below by a linear function

$$xL^1 g(x) \geq xL^1 g(0) \quad \text{for all } x, y \in \mathbb{R}. \quad (11)$$

**Lemma 2.3.** Define, for any given  $\Delta t < (\theta(K + 1))^{-1}$ ,

$$F(x) = x - \theta f(x)\Delta t + \frac{\sigma}{2}L^1 g(x), \quad x \in \mathbb{R}. \quad (12)$$

Then under Assumption 2.2, for any  $b \in \mathbb{R}$  there exists a unique  $x \in \mathbb{R}$  such that  $F(x) = b$  and hence the method (5) is well defined.

*Proof.* In view of Lemma 2.1 it is enough to show that the function  $F$  is continuous, coercive and strictly monotone. Clearly,  $F(x)$  is continuous on  $\mathbb{R}$ . By Assumption 2.2,

$$(x - y)(F(x) - F(y)) \geq |x - y|^2 - \theta K \Delta t |x - y|^2 = (1 - \theta K \Delta t) |x - y|^2 > 0,$$

for  $\Delta t < (\theta(K + 1))^{-1}$ . Also by Assumption 2.2 and Remark 1

$$\begin{aligned} xF(x) &= x(x - \theta f(x)\Delta t) + \frac{\sigma}{2}L^1 g(x)\Delta t \\ &\geq |x|^2 \left(1 - \theta \frac{2K + 1}{2} \Delta t\right) - \frac{\theta}{2}x|f(0)|^2 \Delta t + \frac{\sigma}{2}xL^1 g(0)\Delta t, \end{aligned} \quad (13)$$

so  $F$  is coercive.  $\square$

From now on we assume that

$$\Delta t < (\theta(K + 1))^{-1} \quad (14)$$

**2.1. Existence of a positive solution for the  $(\theta, \sigma)$ -Milstein scheme.** In this subsection we introduce assumptions on coefficients  $f$  and  $g$  of equation (1) that allow us to prove the existence of a positive solution to (5).

**Definition 2.4.** Given  $x(0) > 0$ , if the solution of (1) satisfies  $\mathbb{P}(\{x(t) > 0 : t > 0\}) = 1$  (or  $\mathbb{P}(\{x(t) \geq 0 : t > 0\}) = 1$ ), then a stochastic one-step integration scheme computing approximations  $X_{t_k} \approx x(t_k)$  preserves positivity (non-negativity) if

$$\mathbb{P}(\{X_{t_{k+1}} > 0 | X_{t_k} > 0\}) = 1 \quad (\mathbb{P}(\{X_{t_{k+1}} \geq 0 | X_{t_k} \geq 0\}) = 1).$$

Let us note that to use the ideas from Lemma 2.3 to prove the existence of a positive solution to the implicit scheme we need to assume that a one-sided Lipschitz condition on  $f$  and monotone condition on  $L^1g$  hold only on the positive domain. This significantly relaxes the conditions required for the existence and uniqueness of a solution to the implicit scheme (5).

**Assumption 2.5.** *Coefficients  $f$  and  $g$  of the equation (1) are locally Lipschitz continuous and satisfy the following two conditions:*

One-sided Lipschitz condition on  $\mathbb{R}_+$ . *There exists a constant  $K > 0$  such that*

$$(x - y)(f(x) - f(y)) \leq K|x - y|^2 \quad \text{for all } x, y \in \mathbb{R}_+. \tag{15}$$

Monotone condition on  $\mathbb{R}_+$ . *Operator  $L^1$  acting on  $g$  satisfies*

$$(x - y)(L^1g(x) - L^1g(y)) \geq 0 \quad \text{for all } x, y \in \mathbb{R}_+. \tag{16}$$

Many mean-reverting models with super- and sub-linear diffusion coefficients satisfy Assumption 2.5; for example, the mean-reverting SDEs

$$dx(t) = (\mu - x(t)^q)dt + x(t)^p dw(t) \quad \text{for } x \in \mathbb{R},$$

with  $\mu, q > 0$  and  $p \geq 0.5$ .

In general, boundary behavior of one-dimensional SDEs can be fully characterized by the Feller test [19]. Let us consider the interval  $(0, \infty)$ . We assume that  $f$  and  $g$  are locally Lipschitz continuous in  $(0, \infty)$  and that  $g^2(x) > 0$ , for  $x \in (0, \infty)$ . Let us also define the scale function

$$p(x) = \int_c^x \exp \left[ -2 \int_c^s \frac{f(z)}{g^2(z)} dz \right] ds,$$

where  $c \in \mathbb{R}$ . Since we analyse the behaviour of the above function at 0, we assume that  $c > x$ . By Proposition 5.22 in [19] we have that if  $p(0+) = -\infty$  then  $\mathbb{P}[\inf_{0 \leq t < \infty} x(t) = 0] = 1$ . Therefore, in order to show that the solution to (1) is non-negative it is enough to show that  $p(0+) = -\infty$ .

**Assumption 2.6.** *The coefficients  $f$  and  $g$  in (1) satisfy the following conditions:*

$$f(0) \geq 0, \quad g^2(x) > 0 \text{ for } x \in (0, \infty), \text{ and } g(0) = 0 \text{ for } x = 0.$$

We assume that  $g(0) = 0$  and  $f(0) \geq 0$  in the assumption above in order for  $x(t)$  to stay positive if  $x(0) > 0$ . To see this, we first consider the auxiliary SDE

$$dy(t) = [f(y(t)) - f(0)]dt + g(y(t))dw(t)$$

with initial value  $y(0) = x(0) > 0$ . Recalling that  $f \in C^1(\mathbb{R}, \mathbb{R})$  and  $g \in C^2(\mathbb{R}, \mathbb{R})$ , we can apply [25, Lemma 3.2 on page 120] to obtain that  $y(t) > 0$  almost surely. Observing that  $f(x) \geq [f(x) - f(0)]$  (as we assume  $f(0) \geq 0$ ), the classical stochastic comparison theorem (see e.g. [24]) shows that  $x(t) \geq y(t) > 0$  almost surely.

**Theorem 2.7.** *Let Assumptions 2.5 and 2.6 hold. In addition we require that*

$$L^1g(x) > 0 \quad \text{for } x > 0.$$

*Then there exists a unique positive solution to the  $(\theta, \sigma)$ -Milstein scheme (5) if*

$$x - \frac{g^2(x)}{2L^1g(x)} + (1 - \theta)f(x)\Delta t - \frac{(1 - \sigma)}{2}L^1g(x)\Delta t > -\theta f(0)\Delta t, \quad x > 0. \quad (17)$$

*Similarly, a unique non-negative solution exists if*

$$x - \frac{g^2(x)}{2L^1g(x)} + (1 - \theta)f(x)\Delta t - \frac{(1 - \sigma)}{2}L^1g(x)\Delta t \geq -\theta f(0)\Delta t, \quad x > 0. \quad (18)$$

*Proof.* In view of Lemma 2.1 and Definition 2.4 in order to prove the lemma we analyse the following equation

$$F(X_{t_{k+1}}) = b(X_{t_k}), \quad \forall k,$$

with  $F(x)$  defined in (12) and

$$b(x) = x + (1 - \theta)f(x)\Delta t + g(x)\Delta w_{t_{k+1}} + \frac{1}{2}L^1g(x)\Delta w_{t_{k+1}}^2 - \frac{(1 - \sigma)}{2}L^1g(x)\Delta t.$$

First we prove that  $\mathbb{P}\{X_{t_{k+1}} > 0 \mid X_{t_k} > 0\} = 1$ . By Assumptions 2.5 operator  $F$  in (12) is monotone on  $(0, \infty)$  and we have

$$\lim_{x \rightarrow \infty} \frac{x F(x)}{|x|} = \infty.$$

By Assumption 2.6 we arrive at

$$\lim_{x \rightarrow 0^+} \frac{x F(x)}{|x|} = -\theta f(0)\Delta t.$$

Hence the operator  $F$  is coercive on  $(0, \infty)$ . Due to Lemma 2.1, we may complete the proof by showing

$$b(x) > -\theta f(0)\Delta t, \quad \text{for } x > 0,$$

from which it follows that there exists a positive solution to  $F(X_{t_{k+1}}) = b(X_{t_k})$ . First, for any given  $x > 0$  we find the minimum of the function

$$H(y) = g(x)y + \frac{1}{2}L^1g(x)y^2.$$

Under the assumption  $L^1g(x) > 0$ , for  $x > 0$ , this function possesses a global minimum

$$\min_y H(y) = -\frac{g^2(x)}{2L^1g(x)}.$$

Hence

$$b(x) \geq x + (1 - \theta)f(x)\Delta t - \frac{(1 - \sigma)}{2}L^1g(x)\Delta t - \frac{g^2(x)}{2L^1g(x)} > -\theta f(0)\Delta t, \quad x > 0,$$

as required. For the non-negative case we have  $b(x) \geq -\theta f(0)\Delta t$ ,  $x > 0$ . In that case we also need to check what happens if for some  $k$  we have the following event  $\{X_{t_{k+1}} = 0 \mid X_{t_k} > 0\}$  (that corresponds to the case where  $b(x) = -\theta f(0)\Delta t$ ). Then by Assumption 2.6,  $b(0) = (1 - \theta)f(0)\Delta t$  and so  $b(0) \geq -\theta f(0)\Delta t$ .  $\square$

For the fully implicit (1, 1)-Milstein scheme we see from (17) that a condition guaranteeing non-negativity independently of  $\Delta t$  is

$$x - \frac{g^2(x)}{2L^1g(x)} \geq 0, \quad x > 0.$$

**2.2. Example: Heston volatility model.** Now we demonstrate that approximation of the 3/2-Heston volatility model (2) with the double implicit Milstein scheme preserves non-negativity. We point out that implicitness in the numerical approximation does not increase computational cost in this case, since we are able to find an explicit solution. This often will be the case in mathematical finance, where typical models have drift and diffusion coefficients of a polynomial type.

The (1, 1)-Milstein scheme has the form

$$X_{t_{k+1}} = X_{t_k} + f(X_{t_{k+1}})\Delta t + g(X_{t_k})\Delta w_{t_k} + \frac{1}{2}L^1g(X_{t_k})\Delta w_{t_k}^2 - \frac{1}{2}L^1g(X_{t_{k+1}})\Delta t, \tag{19}$$

where now  $f(x) = \mu x - \alpha x^2$ ,  $g(x) = \beta x^{3/2}$  and  $L^1g(x) = \frac{3}{2}\beta^2 x^2$ . Clearly, the coefficients of equation (2) satisfy Assumptions 2.5 and 2.6. Hence, we may show that (19) has a unique non-negative solution by verifying condition (18) in Theorem 2.7. This reduces to  $x - x/3 \geq 0$  for  $x \geq 0$  and the result follows.

An explicit formula for  $X_{t_{k+1}}$  can be found by solving the relevant quadratic equation and choosing the positive solution, to give

$$X_{t_{k+1}} = (2(\alpha + \frac{3}{4}\beta^2\Delta t))^{-1} \times \left( \sqrt{(1 - \mu\Delta t)^2 + 4(\alpha + \frac{3}{4}\beta^2)\Delta t(X_{t_k} + \beta X_{t_k}^{3/2}\Delta w_{t_k} + \frac{3}{4}\beta^2 X_{t_k}^2 w_{t_k}^2)} - (1 - \mu\Delta t) \right).$$

**3. Stability analysis.** In this section we examine the global stability of the  $(\sigma, \theta)$ -Milstein scheme (5). The stability conditions we derive are related to mean-square stability, and we are interested in results that do not put severe restrictions on the time step. We begin with linear test equations where we can derive sharp results and represent stability regions graphically.

**3.1. Linear mean-square stability.** For the linear test SDE

$$dx(t) = \alpha x(t)dt + \mu x(t)dw(t), \tag{20}$$

the property of mean-square stability,

$$\lim_{t \rightarrow \infty} \mathbb{E} |x(t)|^2 = 0,$$

is characterized by

$$2\alpha + \mu^2 < 0. \tag{21}$$

For the  $\theta$ -Milstein scheme on (20),

$$X_{t_{k+1}} = X_{t_k} + \theta\alpha X_{t_{k+1}}\Delta t + (1 - \theta)\alpha X_{t_k}\Delta t + \mu X_{t_k}\Delta w_{t_{k+1}} + \frac{1}{2}\mu^2 X_{t_k}[\Delta w_{t_{k+1}}^2 - \Delta t],$$

the analogous property

$$\lim_{k \rightarrow \infty} \mathbb{E} |X_{t_k}|^2 = 0, \tag{22}$$

was studied in [12]. In particular, the linear stability region

$$R_{MS} := \{\Delta t\alpha, \Delta t\mu^2 \in \mathbb{R} : \text{method mean-square stable on (20)}\} \tag{23}$$

was shown to be significantly smaller than that for the corresponding Euler-based scheme. We now examine the new Milstein scheme (5) in this setting, which reduces to

$$\begin{aligned} X_{t_{k+1}} &= X_{t_k} + \theta\alpha X_{t_{k+1}}\Delta t + (1-\theta)\alpha X_{t_k}\Delta t + \mu X_{t_k}\Delta w_{t_{k+1}} \\ &\quad + \frac{1}{2}\mu^2 X_{t_k}\Delta w_{t_{k+1}}^2 - \frac{(1-\sigma)}{2}\mu^2 X_{t_k}\Delta t - \frac{\sigma}{2}\mu^2 X_{t_{k+1}}\Delta t. \end{aligned} \quad (24)$$

**Theorem 3.1.** *The  $(\theta, \sigma)$ -Milstein scheme (24) is linearly mean-square stable, (22), if and only if*

$$2\alpha + \mu^2 + \Delta t\alpha^2(1-2\theta) + \frac{\Delta t\mu^2}{2}(2\sigma\alpha + \mu^2) < 0. \quad (25)$$

*Proof.* We rewrite (24) as a recurrence of the form

$$X_{t_{k+1}} = X_{t_k} \left( p + q\xi_{t_{k+1}} + r\xi_{t_{k+1}}^2 \right),$$

where  $\xi \sim \mathbb{N}(0, 1)$  and

$$\begin{aligned} p &= \frac{1 + (1-\theta)\alpha\Delta t - \frac{(1-\sigma)}{2}\mu^2\Delta t}{1 - \theta\alpha\Delta t + \frac{\sigma}{2}\mu^2\Delta t}, \\ q &= \frac{\mu\sqrt{\Delta t}}{1 - \theta\alpha\Delta t + \frac{\sigma}{2}\mu^2\Delta t}, \\ r &= \frac{\frac{1}{2}\mu^2\Delta t}{1 - \theta\alpha\Delta t + \frac{\sigma}{2}\mu^2\Delta t}. \end{aligned}$$

Then

$$|X_{t_{k+1}}|^2 = |X_{t_k}|^2 \left( p^2 + q^2\xi_{t_{k+1}}^2 + r^2\xi_{t_{k+1}}^4 + 2pq\xi_{t_{k+1}} + 2pr\xi_{t_{k+1}}^2 + 2qr\xi_{t_{k+1}}^3 \right).$$

Taking conditional expectation of both sides lead us to

$$\mathbb{E}[|X_{t_{k+1}}|^2 | \mathcal{F}_{t_k}] = |X_{t_k}|^2 (p^2 + q^2 + 3r^2 + 2pr).$$

Taking conditional expectation of both sides again we obtain

$$\mathbb{E}|X_{t_{k+1}}|^2 = \mathbb{E}|X_{t_k}|^2 (p^2 + q^2 + 3r^2 + 2pr). \quad (26)$$

Therefore stability is characterized by  $(p+r)^2 + q^2 + 2r^2 < 1$ . This is equivalent to (25), as required.  $\square$

**Remark 2.** Let us observe that for  $\theta = 0.5$  and  $\sigma = 1$  we have recovered precisely the condition (21) for the underlying SDE, so the method perfectly reproduces stability for any step-size. More generally, for  $\theta \geq 0.5$  and  $\sigma = 1$  we have the property that “problem stable implies method stable for all  $\Delta t$ ”, which is referred to as A-stability in the deterministic literature.

Motivated by [12, 13] we will draw stability regions for (24) in the  $x$ - $y$  plane, where  $x = \alpha\Delta t$  and  $y = \mu^2\Delta t$ . In Figure 1 the stability region of the underlying SDE (20) is shaded light grey. The upper pictures in Figure 1 superimpose the stability region of the  $(\theta, 0)$ -Milstein scheme with  $\theta = 0, 0.5, 1$ , respectively, using darker shading. We see that even in the case of a linear scalar equation we are not able to reproduce the stability region of the underlying test equation (20). However, by introducing additional implicitness we overcome this poor performance. The lower pictures in Figure 1 superimpose the stability region of the  $(\theta, \sigma)$ -Milstein



scheme with  $(0, 1)$ ,  $(0.5, 1)$ ,  $(1, 1)$ , respectively. As stated in Remark 2, we recover exactly the stability region of underlying test SDE (20) for  $\theta = 0.5$  and  $\sigma = 1$ .

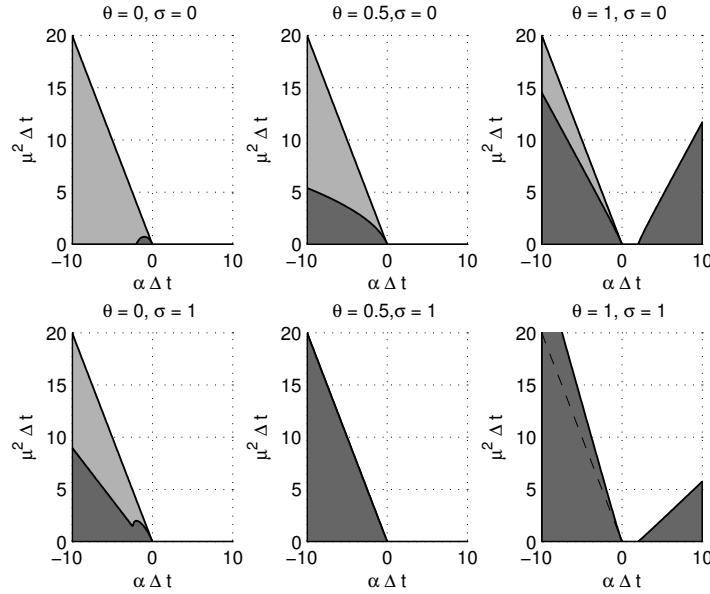


FIGURE 1. Mean square stability for linear SDE (20). Light shading: linear mean-square stability of the SDE. Darker shading: linear mean-square stability of Implicit Milstein scheme (upper) and double-implicit Milstein scheme (lower).

**3.2. Almost sure stability.** In this section we present an almost sure stability analysis for nonlinear SDEs in the spirit of the LaSalle invariance principle that in the stochastic setting was presented in [33]. We begin by stating a result that combines Doob’s Decomposition and Martingale Convergence Theorems.

**Theorem 3.2** (Lipster and Shiryaev [22]). *Let  $Z = \{Z_n\}_{n \in \mathbb{N}}$  be a nonnegative decomposable stochastic process with Doob-Meyer decomposition  $Z_n = Z_0 + A_n^1 - A_n^2 + M_n$ , where  $A^1 = \{A_n^1\}_{n \in \mathbb{N}}$  and  $A^2 = \{A_n^2\}_{n \in \mathbb{N}}$  are a.s. nondecreasing, predictable processes with  $A_0^1 = A_0^2 = 0$ , and  $M = \{M_n\}_{n \in \mathbb{N}}$  is local  $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ -martingale with  $M_0 = 0$ . Then*

$$\left\{ \omega : \lim_{n \rightarrow \infty} A^1(n) < \infty \right\} \subseteq \left\{ \omega : \lim_{n \rightarrow \infty} A^2(n) < \infty \right\} \cap \left\{ \lim_{n \rightarrow \infty} Z_n < \infty \right\} \quad a.s.$$

In [33], the authors proved a very general Stochastic LaSalle Theorem in which stability is described by the kernel of some function  $z \in C(\mathbb{R}; \mathbb{R}_+)$ . Here we present a simplified version of their theorem, with fixed Lyapunov function  $V(x) = |x|^2$ .

**Theorem 3.3** (Shen et al. [33]). *Let local Lipschitz conditions hold for  $f$  and  $g$ . Assume further that there exists a function  $z \in C(\mathbb{R}; \mathbb{R}_+)$  such that*

$$xf(x) + \frac{1}{2}|g(x)|^2 \leq -z(x), \tag{27}$$

for all  $x \in \mathbb{R}$ . For any  $x_0 \in \mathbb{R}$ , the solution  $(x(t))_{t \geq 0}$  of (1) then has the properties that

$$\limsup_{t \rightarrow \infty} |x(t)|^2 < \infty \quad \text{a.s.} \quad \text{and} \quad \lim_{t \rightarrow \infty} z(x(t)) = 0 \quad \text{a.s.}$$

Further if  $z(x) = 0$  if and only if  $x = 0$ , then

$$\lim_{t \rightarrow \infty} x(t) = 0 \quad \text{a.s.} \quad \forall x \in \mathbb{R}.$$

Now we present a counterpart of this Stochastic LaSalle Theorem for the new Milstein scheme.

**Theorem 3.4.** *Let Assumption 2.2 hold. Assume that for the  $(\theta, \sigma)$ -Milstein scheme (5) there exists a function  $z \in C(\mathbb{R}; \mathbb{R}_+)$  such that*

$$\begin{aligned} & 2xf(x) + |g(x)|^2 + (1 - 2\theta) |f(x)|^2 \Delta t \\ & + \frac{\Delta t}{2} L^1 g(x)(2\sigma f(x) + L^1 g(x)) \leq -z(x) \quad \text{for all } x \in \mathbb{R}. \end{aligned} \tag{28}$$

Then

$$\limsup_{k \rightarrow \infty} |X_{t_k}|^2 < \infty$$

and

$$\lim_{k \rightarrow \infty} z(X_{t_k}) = 0 \quad \text{a.s.}$$

Further if  $z(x) = 0$  if and only if  $x = 0$  then

$$\lim_{k \rightarrow \infty} X_{t_k} = 0 \quad \text{a.s.}$$

*Proof.* We can rewrite  $F$  in (12) as

$$F(X_{t_{k+1}}) = F(X_{t_k}) + f(X_{t_k})\Delta t + g(X_{t_k})\Delta w_{t_{k+1}} + \frac{1}{2}L^1 g(X_{t_k})(\Delta w_{t_{k+1}}^2 - \Delta t).$$

Squaring both sides, we arrive at

$$\begin{aligned} |F(X_{t_{k+1}})|^2 &= |F(X_{t_k})|^2 + |f(X_{t_k})|^2 \Delta t^2 + |g(X_{t_k})|^2 \Delta t \\ &+ \frac{1}{2} |L^1 g(X_{t_k})|^2 \Delta t^2 + 2F(X_{t_k})f(X_{t_k})\Delta t + m_{k+1}, \end{aligned}$$

where

$$\begin{aligned} m_{k+1} &= |g(X_{t_k})|^2 (\Delta w_{t_{k+1}}^2 - \Delta t) + \frac{1}{2} |L^1 g(X_{t_k})|^2 [(\Delta w_{t_{k+1}}^2 - \Delta t)^2 - 2\Delta t^2] \\ &+ 2F(X_{t_k}) \left[ g(X_{t_k})\Delta w_{t_{k+1}} + \frac{1}{2}L^1 g(X_{t_k})(\Delta w_{t_{k+1}}^2 - \Delta t) \right] \\ &+ 2f(X_{t_k})\Delta t \left[ g(X_{t_k})\Delta w_{t_{k+1}} + \frac{1}{2}L^1 g(X_{t_k})(\Delta w_{t_{k+1}}^2 - \Delta t) \right] \\ &+ g(X_{t_k})L^1 g(X_{t_k})(\Delta w_{t_{k+1}}^2 - \Delta t)\Delta w_{t_{k+1}} \end{aligned} \tag{29}$$

is a local martingale difference. From the definition of  $F$  we arrive at

$$|F(X_{t_{k+1}})|^2 = |F(X_{t_k})|^2 + 2X_k f(X_{t_k})\Delta t + |g(X_{t_k})|^2 \Delta t \tag{30}$$

$$\begin{aligned} &+ \frac{1}{2}L^1 g(X_{t_k}) [L^1 g(X_{t_k}) + 2\sigma f(X_{t_k})] \Delta t^2 \\ &+ (1 - 2\theta) |f(X_{t_k})|^2 \Delta t^2 + m_{k+1}. \end{aligned} \tag{31}$$

Therefore

$$|F(X_{t_{k+1}})|^2 = |F(X_{t_k})|^2 - A_{t_k} \Delta t + m_{k+1},$$

where

$$A_{t_k}(x) = -\left(2X_{t_k}f(X_{t_k}) + |g(X_{t_k})|^2 + \frac{1}{2}L^1g(X_{t_k}) [L^1g(X_{t_k}) + 2\sigma f(X_{t_k})] \Delta t + (1 - 2\theta) |f(X_{t_k})|^2 \Delta t\right).$$

Hence, we have obtained a decomposition that allows us to apply Theorem 3.2, i.e.,

$$|F(X_{t_{N+1}})|^2 = |F(X_{t_0})|^2 - \sum_{k=0}^N A_{t_k} \Delta t + \sum_{k=0}^N m_{k+1}.$$

Theorem 3.2 gives  $\lim_{k \rightarrow \infty} |F(X_{t_k})|^2 < \infty$ . By condition (28) and (11)

$$\begin{aligned} |F(x)|^2 &= \left(x - \theta f(x)\Delta t + \frac{1}{2}\sigma L^1g(x)\Delta t\right)^2 \\ &= |x|^2 - 2\theta x f(x)\Delta t - \theta \sigma f(x)L^1g(x)\Delta t^2 + \theta^2(f(x))^2 \Delta t^2 \\ &\quad + \frac{1}{4}\sigma^2(L^1g(x))^2 \Delta t^2 + \sigma x L^1g(x)\Delta t \\ &\geq |x|^2 + \theta z(x)\Delta t + \sigma x L^1g(x)\Delta t \\ &\geq |x|^2 - \sigma |x| |L^1g(0)| \Delta t \\ &\geq (1 - 0.5\Delta t) |x|^2 - 0.5 \sigma^2 |L^1g(0)|^2 \Delta t. \end{aligned}$$

Hence  $\limsup_{k \rightarrow \infty} |X(t_k)|^2$  exists and is finite almost surely. Another implication of Theorem 3.2 is

$$\sum_{k=0}^{\infty} z(X_{t_k})\Delta t \leq \sum_{k=0}^{\infty} A_{t_k} \Delta t < \infty \quad \text{a.s.},$$

as required. □

In the case where (5) is non-negative it is enough if condition (28) holds on the non-negative half line.

**Theorem 3.5.** *Let conditions required for existence of non-negative solution in Theorem 2.7 hold. Assume that for the  $(\theta, \sigma)$ -Milstein scheme (5) there exists a function  $z \in C(\mathbb{R}; \mathbb{R}_+)$  such that*

$$\begin{aligned} 2xf(x) + |g(x)|^2 + (1 - 2\theta) |f(x)|^2 \Delta t \\ + \frac{\Delta t}{2} L^1g(x)(2\sigma f(x) + L^1g(x)) \leq -z(x), \quad \text{for all } x \in \mathbb{R}_+. \end{aligned}$$

Then

$$\limsup_{k \rightarrow \infty} |X(t_k)|^2 < \infty$$

and

$$\lim_{k \rightarrow \infty} z(X_{t_k}) = 0 \quad \text{a.s.}$$

Further if  $z(x) = 0$  if and only if  $x = 0$  then

$$\lim_{k \rightarrow \infty} X_{t_k} = 0 \quad \text{a.s.}$$

*Proof.* The proof is analogous to the proof of Theorem 3.4. □

**Remark 3.** Following on from Remark 2, suppose that (27) holds, so the results of Theorem 3.3 hold for the SDE. Then, to minimize restrictions on the stepsize in (28), the choice  $\theta = 0.5$  is clearly best, and the extra freedom allowed by the parameter  $\sigma$  can be used to exploit dissipativity. For example, on the SDE

$$dx(t) = -x(t)^3 dt + x(t)^2 dw(t),$$

we have

$$L^1 g(x)(2\sigma f(x) + L^1 g(x)) = L^1 g(x)(-2\sigma x^3 + 2x^3),$$

so the choice  $\sigma = 1$  makes (28) independent of  $\Delta t$  and identical to (27).

**4. Convergence result.** In this section we show that the numerical approximation (5) strongly converges to the solution of (1) under fairly general conditions. We will not establish the rate of convergence, but we perform numerical experiments that suggest a rate of 1. We note that the (1, 1) scheme was considered in [20] (chapter 12) as an alternative to the more typical (1, 0) version. In particular, those authors showed that when the coefficients  $f$ ,  $g$  and  $L^1 g(x)$  in (1) are globally Lipschitz, the (1, 1) case retains the usual first order of strong convergence. This result is easily extended to the general  $(\theta, \sigma)$  case.

**Theorem 4.1.** *Let  $f$ ,  $g$  and  $L^1 g(x)$  be globally Lipschitz. Then the  $(\theta, \sigma)$ -Milstein scheme (5) strongly converges to the solution of the SDE (1), that is, there exists a constant  $K > 0$  such that*

$$\mathbb{E} \left[ \sup_{0 \leq t_k \leq T} |x(t_k) - X_{t_k}|^p \right] \leq K \Delta t^p, \quad \text{for } p \geq 2.$$

*Proof.* A proof follows by extending the (1, 1) case from Chapter 12 of Kloeden and Platen [20].  $\square$

Then using a localization procedure as in [10, 16] we can prove pathwise convergence without a global Lipschitz assumption. From [16] we know that scheme (5) almost surely converges to the solution of (1), that is:

**Theorem 4.2.** *Assume that the solution to SDE (1) has a strong solution. Then the  $(\theta, \sigma)$ -Milstein scheme (5) converges to the solution of the SDE (1) in the pathwise sense, that is for  $\gamma > 0$  there exists a random variable  $K = K(\omega)$ ,  $\omega \in \Omega$ , such that*

$$\sup_{0 \leq t_k \leq T} |x(t_k) - X_{t_k}| \leq K(\omega) \Delta t^{1-\gamma}, \quad \text{for } T, \gamma > 0. \quad (32)$$

Our proof is omitted as it is a straightforward adaption of the proof of Theorem 1 in [16] or Theorem 2.3 in [10].

In order to show that we also have strong convergence we need to show that the solution to (5) has bounded moments. We will prove boundedness of the moments under the following assumption.

**Assumption 4.3.** *Monotone-type condition. There exists a constant  $K$*

$$\begin{aligned} & 2xf(x) + |g(x)|^2 + (1 - 2\theta) |f(x)|^2 \Delta t \\ & + \frac{\Delta t}{2} L^1 g(x)(2\sigma f(x) + L^1 g(x)) \leq a + b|x|^2 \quad \text{for all } x \in \mathbb{R}, \end{aligned} \quad (33)$$

where  $a = 0.5 |f(0)|^2$  and  $b = (2K + 1)/2$  (see Remark 1).

The following lemma establishes a useful relation between the function  $F(x)$  defined in (12) and its argument  $x$ .

**Lemma 4.4.** *Lets Assumptions 2.2 and 4.3 hold. Then there exist constants  $c_1, c_2 > 0$  such that*

$$|F(x)|^2 \geq c_1 |x|^2 - c_2 \Delta t \quad \text{for } x \in \mathbb{R}.$$

*Proof.* By Assumptions 2.2 and 4.3 we have

$$\begin{aligned} |F(x)|^2 &\geq |x|^2 - \theta a \Delta t - \theta b |x|^2 \Delta t + \sigma x L^1 g(x) \Delta t \\ &\geq |x|^2 - \theta a \Delta t - \theta b |x|^2 \Delta t - \sigma |x| |L^1 g(0)| \Delta t \\ &\geq |x|^2 - \theta a \Delta t - \theta b |x|^2 \Delta t - \theta/2 |x|^2 \Delta t - \sigma^2/(2\theta) |L^1 g(0)|^2 \Delta t \\ &\geq (1 - (\theta b + \theta/2) \Delta t) |x|^2 - \theta a \Delta t - \sigma^2/(2\theta) |L^1 g(0)|^2 \Delta t, \end{aligned} \tag{34}$$

and we take  $c_1 = (1 - (\theta b + \theta/2) \Delta t)$  and  $c_2 = -\theta a + \sigma^2/(2\theta) |L^1 g(0)|^2$ . Due to (14)  $c_1 > 0$ . □

Our analysis uses a localization procedure. We define the stopping time  $\lambda_m$  by

$$\lambda_m = \inf\{k : |X_{t_k}| > m\}. \tag{35}$$

We observe that when  $k \in [0, \lambda_m(\omega)]$ ,  $|X_{t_{k-1}}(\omega)| \leq m$ , but we might have that  $|X_{t_k}(\omega)| > m$ , so we need the following lemma.

**Lemma 4.5.** *Let Assumptions 2.2 and 4.3 hold. Then for  $p \geq 2$  and sufficiently large integer  $m$ , there exists a constant  $K = K(p, m)$ , such that*

$$\mathbb{E} [|X_{t_k}|^p \mathbf{1}_{[0, \lambda_m]}(k)] < K \quad \text{for any } k \geq 0.$$

*Proof.* By (30) and Assumption 4.3 we obtain

$$|F(X_{t_k})|^2 \leq |F(X_{t_{k-1}})|^2 + a \Delta t + b |X_{t_{k-1}}|^2 \Delta t + \Delta m_k,$$

where  $\Delta m_{k+1}$  is defined by (29). Using the basic inequality  $(a_1 + a_2 + a_3 + a_4)^{p/2} \leq 4^{p/2-1} (a_1^p + a_2^p + a_3^p + a_4^p)$ , where  $a_i \geq 0$ , we obtain

$$|F(X_{t_k})|^p \leq 4^{p-1} \left( |F(X_{t_{k-1}})|^p + (a \Delta t)^{p/2} + b |X_{t_{k-1}}|^p \Delta t + |\Delta m_k|^{p/2} \right). \tag{36}$$

As a consequence

$$\begin{aligned} \mathbb{E} [|F(X_{t_k})|^p \mathbf{1}_{[0, \lambda_m]}(k)] &\leq 4^{p-1} \left( \mathbb{E} [|F(X_{t_{k-1}})|^p \mathbf{1}_{[0, \lambda_m]}(k)] + (a \Delta t)^{p/2} \right. \\ &\quad \left. + b m^p \Delta t + \mathbb{E} [|\Delta m_k|^{p/2} \mathbf{1}_{[0, \lambda_m]}(k)] \right). \end{aligned}$$

In order to bound  $\mathbb{E} [|\Delta m_k|^{p/2} \mathbf{1}_{[0, \lambda_m]}(k)]$  we need to consider all the terms of  $\Delta m_k$  separately. By the Cauchy-Schwarz inequality

$$\begin{aligned} &\mathbb{E} \left[ |g(X_{t_{k-1}})|^p \left| \Delta w_{t_{k+1}}^2 - \Delta t \right|^{p/2} \mathbf{1}_{[0, \lambda_m]}(k) \right] \\ &\leq \left[ \left( \mathbb{E} [|g(X_{t_{k-1}})|^{2p} \mathbf{1}_{[0, \lambda_m]}(k)] \right)^{1/2} \left( \mathbb{E} \left| \Delta w_{t_{k+1}}^2 - \Delta t \right|^p \right)^{1/2} \right]. \end{aligned}$$

Since there exists a positive constant  $C(p)$ , such that  $\mathbb{E} |\Delta w_{t_{k-1}}|^2 < C(p)$ , there exists a constant  $C(m, p)$  such that

$$\mathbb{E} \left[ |g(X_{t_{k-1}})|^p \left| \Delta w_{t_{k+1}}^2 - \Delta t \right|^{p/2} \mathbf{1}_{[0, \lambda_m]}(k) \right] \leq C(m, p).$$

In the same way we can bound all the other terms of  $\Delta m_k$ . Hence

$$\mathbb{E} [|F(X_{t_k})|^p \mathbf{1}_{[0, \lambda_m]}(k)] < C(m, p).$$

Due to Lemma 4.4 the proof is complete.  $\square$

In addition to Assumption 4.3 we require the following very mild restriction on the coefficients of the SDE.

**Assumption 4.6.** *The coefficients of equation (1) satisfy a polynomial growth condition. That is, there exists a pair of constants  $h \geq 1$  and  $H > 0$  such that*

$$|f(x)| \vee |g(x)| \leq H(1 + |x|^h), \quad \forall x. \quad (37)$$

Now we formulate the key theorem that allows us to prove a strong convergence result.

**Theorem 4.7.** *Let Assumptions 2.2, 4.3 and 4.6 hold. Then there exists a constant  $K = K(T)$  such that the  $(\theta, \sigma)$ -Milstein scheme (5) satisfies*

$$\sup_{0 \leq t_k \leq T} \mathbb{E} |X_{t_k}|^2 \leq K.$$

*Proof.* By (30) and Assumption 4.3 we arrive at

$$|F(X_{t_{k+1}})|^2 \leq |F(X_{t_k})|^2 + a\Delta t + b|X_{t_k}|^2 \Delta t + \Delta m_{k+1}, \quad (38)$$

where  $\Delta m_{k+1}$  is defined by (29). Let  $N$  be any non-negative integer such that  $N\Delta t \leq T$ . Summing both sides of inequality (38) from  $k = 0$  to  $N \wedge \lambda_m$ , we get

$$\begin{aligned} |F(X_{t_{N \wedge \lambda_m + 1}})|^2 &\leq |F(X_{t_0})|^2 + aT + b \sum_{k=0}^{N \wedge \lambda_m} |X_{t_k}|^2 \Delta t + \sum_{k=0}^{N \wedge \lambda_m} \Delta m_{k+1} \\ &\leq |F(X_{t_0})|^2 + aT + b \sum_{k=0}^N |X_{t_{k \wedge \lambda_m}}|^2 \Delta t + \sum_{k=0}^N \Delta m_{k+1} \mathbf{1}_{[0, \lambda_m]}(k). \end{aligned}$$

Due to Lemma 4.5  $\sum_{k=0}^N \Delta m_{k+1} \mathbf{1}_{[0, \lambda_m]}(k)$  is a martingale. Hence

$$\mathbb{E} |F(X_{t_{N \wedge \lambda_m + 1}})|^2 \leq |F(X_{t_0})|^2 + aT + b \mathbb{E} \left[ \sum_{k=1}^N |X_{t_{k \wedge \lambda_m}}|^2 \Delta t \right].$$

Due to Lemma 4.4 we have

$$\mathbb{E} |F(X_{t_{N \wedge \lambda_m + 1}})|^2 \leq |F(X_{t_0})|^2 + (a + c_2 c_1^{-1})T + b c_1^{-1} \mathbb{E} \left[ \sum_{k=0}^N |F(X_{t_{k \wedge \lambda_m}})|^2 \Delta t \right].$$

By the discrete Gronwall Lemma

$$\mathbb{E} |F(X_{t_{N \wedge \lambda_m + 1}})|^2 \leq \left[ |F(X_{t_0})|^2 + (a + c_2 c_1^{-1})T \right] \exp(b c_1^{-1} T), \quad (39)$$

where we used the fact that  $N\Delta t \leq T$ . Thus, letting  $m \rightarrow \infty$  in (39) and applying Fatou's lemma, we obtain

$$\mathbb{E} |F(X_{t_{N+1}})|^2 \leq \left[ |F(X_{t_0})|^2 + (a + c_2 c_1^{-1})T \right] \exp(b c_1^{-1} T).$$

The final bound follows from Lemma 4.4.  $\square$

We are ready to prove a strong convergence result.

**Theorem 4.8.** *Let Assumptions 2.2, 4.3 and 4.6 hold. Then the  $(\theta, \sigma)$ -Milstein scheme (5) strongly converges to the solution of the SDE (1), that is*

$$\lim_{\Delta t \rightarrow 0} \mathbb{E} |x(t_k) - X_{t_k}|^p = 0 \quad \text{for } 0 < p < 2. \tag{40}$$

*Proof.* By (32) the  $(\theta, \sigma)$ -Milstein approximation (5)  $X_{t_k}$  converges to  $x(t_k)$  in probability (Theorem 2.2 in [34]). Theorem 4.7 implies that the sequence  $\{|X_{t_k}|^{2-\varepsilon}\}_{t_k}$  is uniformly integrable (Lemma 2.3 in [34]). Therefore by the Vitali convergence theorem (Theorem 2.4 in [34]) the statement of the theorem holds.  $\square$

In the case where we can guarantee non-negativity of approximation, conditions required to prove Theorem 4.8 can be significantly relaxed.

**Assumption 4.9.** Monotone-type condition on  $\mathbb{R}_+$ . *There exists constant  $K$  such that*

$$\begin{aligned} &2xf(x) + |g(x)|^2 + (1 - 2\theta) |f(x)|^2 \Delta t \\ &+ \frac{\Delta t}{2} L^1 g(x)(2\sigma f(x) + L^1 g(x)) \leq a + b|x|^2 \quad \text{for all } x \in \mathbb{R}_+, \end{aligned} \tag{41}$$

where  $a = 0.5 |f(0)|^2$  and  $b = (2K + 1)/2$  (see Remark 1).

**Theorem 4.10.** *Let the conditions required for existence of non-negative solution in Theorem 2.7 hold. Then under Assumptions 4.6 and 4.9 the  $(\theta, \sigma)$ -Milstein scheme (5) strongly converges to the solution of the SDE (1), that is*

$$\lim_{\Delta t \rightarrow 0} \mathbb{E} |x(t_k) - X_{t_k}|^p = 0 \quad \text{for } 0 < p < 2. \tag{42}$$

*Proof.* The theorem can be proved in an analogous way to Theorem 4.8.  $\square$

It is clear that the 3/2-model (2) does not satisfy Assumption 4.3, but satisfies Assumption 4.9 as long as  $\alpha \geq \beta^2/4$ . This condition does not seem to be restrictive, as pointed out in [9].

**4.1. Numerical experiment.** In order to estimate the rate of convergence we proceed with numerical experiments for (2). We focus on the strong endpoint error,  $e_{\Delta t}^{\text{strong}} = \mathbb{E} |x(T) - X_T|$ , with  $T = 1$ . We used  $\mu = 0.1$ ,  $\alpha = 0.2$ ,  $\beta = \sqrt{0.2}$  and  $x(0) = 0.5$ . We plot  $e_{\Delta t}^{\text{strong}}$  against  $\Delta t$  on a log-log scale. Error bars representing 95% confidence intervals are shown by circles, and a reference line of slope 1 is also given. Although we do not know the explicit form of the solution, Theorem 4.8 guarantees that the (1,1)-Milstein scheme (19) strongly converges to the true solution. We therefore take the (1,1)-Milstein scheme with  $\Delta t = 2^{-14}$  as a reference solution. We compare this with the (1,1)-Milstein scheme evaluated with  $(2\Delta t, 2^3\Delta t, 2^5\Delta t, 2^7\Delta t)$  in order to estimate the rate of convergence. Since we are using a Monte Carlo method, the sampling error decays like  $1/\sqrt{M}$ , where  $M = 10000$  is the number of sample paths. From Figure 2 we see that there appears to exist a positive constant  $C$  such that

$$e_{\Delta t}^{\text{strong}} \leq C\Delta t, \quad \text{for sufficiently small } \Delta t.$$

A least squares fit for equality produced the value 1.1304 for the rate with residual of 0.2468. Hence, our results are consistent with strong order of convergence equal to one.

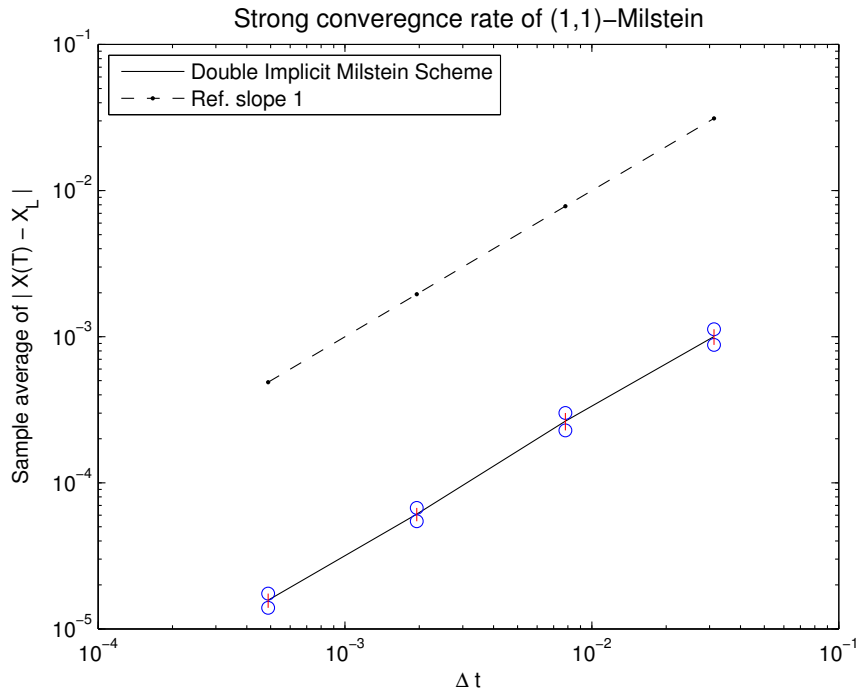


FIGURE 2. Strong error of double-implicit Milstein scheme applied to Heston 3/2 Stochastic volatility model.

**5. Conclusions.** Our aim was to introduce a new discretization scheme that can be shown to work well on highly nonlinear SDEs arising in mathematical finance and to possess excellent linear and nonlinear stability properties. In order to derive the scheme we followed a very simple strategy: We rewrote the original Itô SDE in Stratonovich form and then applied a drift implicit Milstein scheme. For future research it would be very interesting to see if this simple strategy can generate other useful schemes, especially in a multidimensional setting. Also it would be interesting to compare the schemes derived here with numerical methods developed in [31], where the authors applied the Lamperti transformation to the original SDE and then approximated with drift implicit Euler scheme. Other interesting areas for follow-up work are (a) establishing a strong order of convergence for this method in a nonlinear setting, and (b) developing a theory of positivity preservation in the case of SDE systems and their numerical simulation.

## REFERENCES

- [1] Assyr Abdule and Adrian Blumenthal, *Stabilized multilevel Monte Carlo method for stiff stochastic differential equations*, EPFL-ARTICLE-183502, (2013).
- [2] D. H. Ahn and B. Gao, *A parametric nonlinear model of term structure dynamics*, Review of Financial Studies, **12** (1999), 721.
- [3] Y. Ait-Sahalia, *Testing continuous-time models of the spot interest rate*, Review of Financial Studies, **9** (1996), 385–426.



- [4] J. A. D. Appleby, M. Guzowska, C. Kelly and A. Rodkina, *Preserving positivity in solutions of discretised stochastic differential equations*, Applied Mathematics and Computation, **217** (2010), 763–774.
- [5] E. Buckwar and T. Sickenberger, *A comparative linear mean-square stability analysis of Maruyama-and Milstein-type methods*, Mathematics and Computers in Simulation, **81** (2011), 1110–1127.
- [6] T. C. Gard, “Introduction to Stochastic Differential Equations,” Marcel Dekker, New York, 1988.
- [7] M. Giles, *Improved multilevel Monte Carlo convergence using the Milstein scheme*, Monte Carlo and Quasi-Monte Carlo Methods, **2006** (2008), 343–358.
- [8] M. B. Giles, *Multilevel Monte Carlo path simulation*, Operations Research-Baltimore, **56** (2008), 607–617.
- [9] J. Goard and M. Mazur, “Stochastic Volatility Models and the Pricing of Vix Options,” Mathematical Finance, 2011.
- [10] I. Gyöngy, *A note on Euler’s approximations*, Potential Analysis, **8** (1998), 205–216.
- [11] S. L. Heston, “A Simple New Formula for Options with Stochastic Volatility,” Course Notes of Washington University in St. Louis, Missouri, 1997.
- [12] D. J. Higham, *A-stability and stochastic mean-square stability*, BIT Numerical Mathematics, **40** (2000), 404–409.
- [13] D. J. Higham, *Mean-square and asymptotic stability of the stochastic theta method*, SIAM Journal on Numerical Analysis, **38** (2000), 753–769.
- [14] M. Hutzenthaler, A. Jentzen and P. E. Kloeden, *Strong and weak divergence in finite time of Euler’s method for stochastic differential equations with non-globally Lipschitz continuous coefficients*, Proceedings of the Royal Society A, **467** (2011), 1563–1576.
- [15] M. Hutzenthaler, A. Jentzen and P. E. Kloeden, *Strong convergence of an explicit numerical method for SDEs with nonglobally Lipschitz continuous coefficients*, The Annals of Applied Probability, **22** (2012), 1611–1641.
- [16] A. Jentzen, P. E. Kloeden and A. Neuenkirch, *Pathwise approximation of stochastic differential equations on domains: Higher order convergence rates without global Lipschitz coefficients*, Numerische Mathematik, **112** (2009), 41–64.
- [17] C. Kahl, M. Gunther and T. Rosberg, *Structure preserving stochastic integration schemes in interest rate derivative modeling*, Applied Numerical Mathematics, **58** (2008), 284–295.
- [18] C. Kahl and H. Schurz, *Balanced Milstein methods for ordinary SDEs*, Monte Carlo Methods and Applications, **12** (2006), 143–170.
- [19] I. Karatzas and S. E. Shreve, “Brownian Motion and Stochastic Calculus,” Springer, 1991.
- [20] P. E. Kloeden and E. Platen, “Numerical Solution of Stochastic Differential Equations,” Springer, 1992.
- [21] A. L. Lewis, “Option Valuation Under Stochastic Volatility,” Finance Press, 2000.
- [22] R. S. Liptser and A. N. Shiriyayev, “Theory of Martingales,” Kluwer Academic Publishers, 1989.
- [23] R. Lord, R. Koekoek and D. J. C. Van Dijk, *A comparison of biased simulation schemes for stochastic volatility models*, Quantitative Finance, **10** (2010), 177–194.
- [24] X. Mao, “Stability of Stochastic Differential Equations with Respect to Semimartingales,” Longman Scientific & Technical, 1991.
- [25] X. Mao, “Stochastic Differential Equations and Their Applications,” Horwood Pub Ltd, 1997.
- [26] X. Mao and L. Szpruch, *Strong convergence rates for backward Euler–Maruyama method for non-linear dissipative-type stochastic differential equations with super-linear diffusion coefficients*, Stochastics, **85** (2012), 144–177.
- [27] X. Mao and L. Szpruch, *Strong convergence and stability of implicit numerical methods for stochastic differential equations with non-globally Lipschitz continuous coefficients*, J. Comput. Appl. Math., **238** (2013), 14–28.
- [28] G. N. Milstein, E. Platen and H. Schurz, *Balanced implicit methods for stiff stochastic systems*, SIAM Journal on Numerical Analysis, **35** (1998), 1010–1019.
- [29] G. N. Milstein and M. V. Tretyakov, “Stochastic Numerics for Mathematical Physics. Scientific Computation,” Springer-Verlag, Berlin, 2004.
- [30] Tetsuya Misawa, *A lie algebraic approach to numerical integration of stochastic differential equations*, SIAM Journal on Scientific Computing, **23** (2001), 866–890.
- [31] A. Neuenkirch and L. Szpruch, *First order strong approximations of scalar sdes with values in a domain*, Preprint. [arXiv:1209.0390](https://arxiv.org/abs/1209.0390).

- [32] H. Schurz, *Convergence and stability of balanced implicit methods for systems of SDEs*, Int. J. Numer. Anal. Model, **2** (2005), 197–220.
- [33] Y. Shen, Q. Luo and X. Mao, *The improved LaSalle-type theorems for stochastic functional differential equations*, Journal of Mathematical Analysis and Applications, **318** (2006), 134–154.
- [34] A. N. Shiryaev, “Probability,” Springer-Verlag, New York, 1996.
- [35] L. Szpruch, X. Mao, D. J. Higham and J. Pan, *Numerical simulation of a strongly nonlinear Ait-Sahalia type interest rate model*, BIT Numerical Mathematics, **51** (2011), 405–425.
- [36] X. Wang, S. Gan and D. Wang, *A family of fully implicit Milstein methods for stiff stochastic differential equations with multiplicative noise*, BIT, **52** (2012), 741–772.
- [37] E. Zeidler, “Nonlinear Functional Analysis and Its Applications,” Springer Verlag, 1990.

Received April 2012; revised February 2013.

*E-mail address:* [d.j.higham@strath.ac.uk](mailto:d.j.higham@strath.ac.uk)

*E-mail address:* [x.mao@strath.ac.uk](mailto:x.mao@strath.ac.uk)

*E-mail address:* [l.szpruch@ed.ac.uk](mailto:l.szpruch@ed.ac.uk)