Available at: http://publications.ictp.it IC/2010/093

United Nations Educational, Scientific and Cultural Organization
and
International Atomic Energy Agency
THE ABDUS SALAM INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS

STRONG LOCAL LINEARIZATION METHODS FOR THE NUMERICAL INTEGRATION OF STOCHASTIC DIFFERENTIAL EQUATIONS WITH ADDITIVE NOISE: AN OVERVIEW

Juan Carlos Jimenez
Instituto de Cibernética, Matemática y Física, Departamento de Matemática Interdisciplinaria, Calle 15, No. 551, e/ C y D, Vedado, La Habana 4, C.P. 10400, Cuba
and
The Abdus Salam International Centre for Theoretical Physics, Trieste, Italy.

Abstract

Strong Local Linearization (LL) methods conform a class of one-step explicit integrators for SDEs with additive noise derived from the following primary and common strategy: the drift coefficient of the differential equation is locally (piecewise) approximated through a linear Ito-Taylor expansion at each time step, thus obtaining successive linear equations that are explicitly integrated. Hereafter, the LL approach may include some additional strategies to improve that basic affine approximation. Theoretical and practical results have shown that the LL integrators have a number of convenient properties. These include arbitrary order of convergence, A-stability, preservation of the dynamic properties of the linear systems, low computational cost, and others. Remarkably, for nonlinear equations in general, these integrators show a stability similar to that of implicit schemes, but with much lower computational cost (comparable to conventional explicit schemes). In this paper, a review of the LL methods and their properties is presented.

MIRAMARE – TRIESTE
December 2010

1 Introduction

Stochastic Differential Equations (SDEs) with additive noise arise as natural mathematical models for describing random processes in a variety of application areas. For example, models for the blood clotting systems [15], cellular energetics [53], stochastic annealing [16], electrical activity of neural masses [52, 43] and noisy oscillators in a diversity of physical systems [17].

In several practical situations, it is important that the trajectories (the sample paths) of the numerical approximations be close to the strong solution of an SDE. These direct simulations of the trajectories can provide considerable insight into the qualitative behavior and dynamics of the SDE [1, 9]. Indeed, in many cases, these numerical trajectories can be interpreted as stochastic flows. There are many examples of stochastic flows on manifolds that come up when addressing stability and bifurcation issues. In general, the addition of noise to a deterministic system can drastically modified its deterministic dynamics. For instance, small noisy perturbation of a deterministic system can make bifurcation ill defined in a region near the critical value, can make trajectories flip backwards and forwards between different steady-states or can cause a crossing of attraction domains of stable equilibrium points. Therefore, the numerical integration of such stochastic systems should be done with care.

Currently, there is a variety of strong numerical integrator for SDEs (see, e.g., extensive surveys in [33, 49] and comparative studies by simulations in [34, 45]). Although the majority of them have been developed for equations with multiplicative noise, they may be applied to SDEs with additive noise too. It is well-known that an appropriate trade-off among the basic requirements of stability, high order of convergence, and low computational effort of a numerical integrator is difficult to achieve in general. From the stability viewpoint, the common approach to construct integrators with large stability regions is through implicit methods. In the better case, some of them satisfy the elemental $A-$stability criterion, but at expense of a high computational cost. Indeed, they usually involve the numerical solution of a system of nonlinear algebraic equations at each integration step. This algorithmic drawback is lessened in predictor-corrector methods (e.g., those based on explicit Runge-Kutta or Theta methods [49]) and linear-implicit methods (e.g., Rosenbrock methods [2]). However, the former ones are less stable than fully implicit and linear-implicit methods, while the latter ones still involve much more computational effort than explicit integrators, specially for high dimensional SDEs. From the viewpoint of order of convergence, increasing the order entails calculation of costly multiple stochastic integrals and, in some cases, of high order derivatives. This is so for both implicit and explicit methods, being much more critical for the implicit ones. Hence, explicit methods are preferable for providing high order of convergence with lower algorithmic complexity. But, unfortunately, standard explicit schemes are not $A-$stable. In practice, there are many examples of SDEs with bounded trajectories in which, for any fixed step-size of the time discretization, the numerical solution becomes explosive when the initial value is in a certain region of the phase space [41, 3].

The above scenario analogously holds for the majority of numerical methods specially designed for SDEs with additive noise. For example, Adams-type methods proposed in [14, 5] are explicit linear multi-step methods that are not $A-$stable when their order of convergence is high [18]. Euler-type methods [24], as well as Runge-Kutta methods [11, 22], are $A-$stable just in their implicit variants. A relevant exception are the strong Local Linearization integrators proposed in [41, 3, 50, 31, 26, 13], which are $A-$stable explicit schemes that can reach high order of convergence.

The Local Linearization (LL) approach was early introduced by Ozaki in [41] as an alternative integration method that attempts to overcome the instabilities of other methods for SDEs with additive noise. Initially, the strong LL method was based on an heuristics directed to obtain a explicit scheme in term of a multivariate autoregressive model with state-dependent coefficients and random term driven by the Wiener process of the SDE. From dynamical viewpoint, the proposed LL scheme provides satisfactory results in the integration of scalar SDEs. However, such formulation of the LL method is ambiguous in regard to the definition of the random term that appears in the numerical scheme and it is not well defined for multidimensional SDE. Later, the method was reconsidered and reformulated in [3], [50] and [31] overcoming the previously mentioned shortcomings of the original approach and clarifying its derivation. In these papers, the strong LL discretizations are obtained from a piecewise linearization of the given differential equation, from which results a sequence of linear SDEs whose explicit strong solutions are recursively computed at each time step. The above procedure involves the use of truncated Ito-Taylor expansions of the drift and diffusion coefficients. This makes a major difference with other numerical methods, which are in general obtained from truncated Ito-Taylor expansions of the unknown solution of the SDE. However, a limitation of theses LL integrators is their moderate order of convergence; namely, order 1 or 1.5. To overcome this, high order local linear discretizations were introduced in [13]. In addition, extensions of the LL methods for SDEs with jumps and for SDEs driven by semimartingales have been considered in [29] and [47], respectively.

In this paper, a review of the strong Local Linearization methods and their properties is presented. The paper is organized as follows. In sections 2 and 3, the formulation of the Local Linearization methods, as well as, the convergence rate of their associated discretizations is presented. Section 4 deals with the stability and dynamics of these discretizations for linear systems. Section 5 presents an extension of the Local Linearization methods for SDEs with jumps. Section 6 focuses on the numerical implementation of the Local Linear discretizations, that is, on the so-called Local Linearization schemes. Finally, a brief comment on the performance of the these schemes in simulations is presented in the last section.

2 Local Linear approximation

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space, and $\{\mathcal{F}_t,\, t \geq t_0\}$ be an increasing right continuous family of complete sub σ-algebras of $\mathcal{F}$. Consider a d-dimensional diffusion process $\mathbf{x}$ defined by the following stochastic differential equation with additive noise

$$d\mathbf{x}(t) = \mathbf{f}(t, \mathbf{x}(t))dt + \sum_{i=1}^{m} \mathbf{g}_i(t) d\mathbf{w}^i(t) \tag{1}$$

$$\mathbf{x}(t_0) = \mathbf{x}_0, \tag{2}$$

where the drift coefficient $\mathbf{f} : [t_0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ and the diffusion coefficient $\mathbf{g}_i : [t_0, T] \to \mathbb{R}^d$ are differentiable functions, $\mathbf{w} = (\mathbf{w}^1, \ldots, \mathbf{w}^m)$ is an m-dimensional $\mathcal{F}_t$-adapted standard Wiener process, and $\mathbf{x}_0$ is an $\mathcal{F}_{t_0}$-measurable random vector. The standard conditions for the existence and uniqueness of a strong solution are assumed.

According to the formulation of the Local Linearization approach presented in the previous ICTP preprint [27] for deterministic equations, the Local Linear discretization of the equation (1) shall be obtained from the linearization of the function $\mathbf{f}$. But, unlike the deterministic case, two form of linearizations are now possible: the classical one obtained from a truncated Taylor expansion, and other derived from a truncated Ito-Taylor expansion. Let us see this in detail.

Consider the time discretization $(t)_h = \{t_n : n = 0, 1, \ldots, N\}$, with maximum step-size $h \in (0, 1)$, defined as a sequence of $\mathcal{F}$-stopping times that satisfy

$$t_0 < t_1 < \cdots < t_N = T$$

and

$$\sup_n (h_n) \leq h$$

w.p.1, where t_n is $\mathcal{F}_{t_n}$-measurable for each $n = 0, 1, \ldots, N-1$, and $h_n = t_{n+1} - t_n$. In addition, let us denote

$$n_t = \max\{n = 0, 1, 2, \ldots : \ t_n \leq t \text{ and } t_n \in (t)_h\}$$

for all $t \in [t_0, T]$.

Suppose that $\mathbf{y}_n \in \mathbb{R}^d$ is a point close to $\mathbf{x}(t_n)$ for all $t_n \in (t)_h$. Hence, according to the linear Taylor and Ito-Taylor expansions, the drift coefficient $\mathbf{f}$ can be linearly approximated around the point $(t_n, \mathbf{y}_n)$ as

$$\mathbf{f}(s, \mathbf{u}) \approx \mathbf{A}_n \mathbf{u} + \mathbf{a}_n^{\gamma}(s), \tag{3}$$

for all $\mathbf{u} \in \mathbb{R}^m$ and $s \in \mathbb{R}$, where $\mathbf{A}_n = \mathbf{f}_{\mathbf{x}}(t_n, \mathbf{y}_n)$ is a constant matrix,

$$\mathbf{a}_n^{\gamma}(s) = \begin{cases} \mathbf{f}(t_n, \mathbf{y}_n) - \mathbf{f}_{\mathbf{x}}(t_n, \mathbf{y}_n)\mathbf{y}_n + \mathbf{f}_t(t_n, \mathbf{y}_n)(s - t_n) & \text{for } \gamma = 1 \\ \mathbf{a}_n^1(s) + \frac{1}{2} \sum\limits_{k,l=1}^{d} \sum\limits_{j=1}^{m} \mathbf{g}_j^k(t_n) \mathbf{g}_j^l(t_n) \frac{\partial^2 \mathbf{f}(t_n, \mathbf{y}_n)}{\partial \mathbf{x}^k \partial \mathbf{x}^l}(s - t_n) & \text{for } \gamma = 1.5 \end{cases}$$

is a vector function, and the symbol γ indicates the order of strong convergence of the expansion employed for approximating $\mathbf{f}$: $\gamma = 1$ for the Taylor expansion and $\gamma = 1.5$ for the

Ito-Taylor. Here, $\mathbf{f_x}$ and $\mathbf{f}_t$ denote the partial derivatives of $\mathbf{f}$ with respect to the variables $\mathbf{x}$ and t, respectively.

Thus, by using the approximation (3), the solution of (1)-(2) can be locally approximated by the solution of the linear SDE

$$d\mathbf{y}\left(t\right) = (\mathbf{A}_n\mathbf{y}(t) + \mathbf{a}_n^{\gamma}\left(t\right))dt + \sum_{i=1}^{m}\mathbf{g}_i(t)d\mathbf{w}^i(t), \quad t \in (t_n, t_{n+1}], \tag{4}$$

$$\mathbf{y}\left(t_n\right) = \mathbf{y}_n \tag{5}$$

i.e., by the expression

$$\mathbf{y}\left(t\right) = e^{\mathbf{A}_n(t-t_n)}\left\{\mathbf{y}_n + \int_{t_n}^{t} e^{-\mathbf{A}_n(u-t_n)}\mathbf{a}_n^{\gamma}\left(u\right)du + \sum_{i=1}^{m}\int_{t_n}^{t} e^{-\mathbf{A}_n(u-t_n)}\mathbf{g}_i(u)d\mathbf{w}^i(u)\right\} \tag{6}$$

for all $t \in [t_n, t_{n+1}]$. By using the integral identity

$$\int_0^{\Delta} e^{-\mathbf{A}_n u}du\ \mathbf{A}_n = -(e^{-\mathbf{A}_n\Delta} - \mathbf{I}), \qquad \Delta \geq 0$$

the expression (6) can be rewritten as

$$\mathbf{y}\left(t\right) = \mathbf{y}_n + \phi_{\gamma}\left(t_n, \mathbf{y}_n; t - t_n\right) + \xi\left(t_n, \mathbf{y}_n; t - t_n\right), \tag{7}$$

where

$$\phi_{\gamma}\left(t_n, \mathbf{y}_n; \delta\right) = \int_0^{\delta} e^{\mathbf{A}_n(\delta-u)}(\mathbf{A}_n\mathbf{y}_n + \mathbf{a}_n^{\gamma}\left(t_n + u\right))du \tag{8}$$

is a deterministic function, and

$$\xi\left(t_n, \mathbf{y}_n; \delta\right) = \sum_{i=1}^{m}\int_{t_n}^{t_n+\delta} e^{\mathbf{A}_n(t_n+\delta-u)}\mathbf{g}_i(u)d\mathbf{w}^i(u) \tag{9}$$

is a Gaussian random variable with zero mean and variance

$$\mathbf{\Sigma}_{\xi}(t_n, \mathbf{y}_n; \delta) = \int_0^{\delta} e^{\mathbf{A}_n(\delta-u)}\mathbf{G}(t_n + u)\mathbf{G}^{\intercal}(t_n + u)e^{\mathbf{A}_n^{\intercal}(\delta-u)}du,$$

being $\mathbf{G} = [\mathbf{g}_1, \ldots, \mathbf{g}_m]$ a $d \times m$ matrix.

In this way, by setting $\mathbf{y}_0 = \mathbf{x}(t_0)$ and iteratively evaluating the expression (7) at each point $t_{n+1} \in (t)_h$, a sequence of points $\mathbf{y}_{n+1}$ can be obtained as an approximation to the solution $\mathbf{x}$ of (1)-(2). More precisely, this can be defined as follows.

Definition 1 *For a given time discretization $(t)_h$, the order-γ $(= 1, 1.5)$ strong Local Linear discretization of the solution of (1)-(2) is defined by the recurrent relation*

$$\mathbf{y}_{n+1} = \mathbf{y}_n + \phi_{\gamma}(t_n, \mathbf{y}_n; h_n) + \xi(t_n, \mathbf{y}_n; h_n), \tag{10}$$

where the initial point $\mathbf{y}_0$ is an A_0-measurable random vector, and ϕ_{γ} and ξ are given by (8) and (9), respectively.

Moreover, an approximation for $\mathbf{x}$ in the whole interval $[t_0, T]$ is stated in the next definition.

Definition 2 *For a given time discretization $(t)_h$, the stochastic process $\mathbf{y} = \{\mathbf{y}(t),\ t \in [t_0, T]\}$ is called order-γ strong Local Linear approximation of the solution of (1)-(2) if*

$$\mathbf{y}(t) = \mathbf{y}_{n_t} + \phi_\gamma(t_{n_t}, \mathbf{y}_{n_t}; t - t_{n_t}) + \xi(t_{n_t}, \mathbf{y}_{n_t}; t - t_{n_t}), \tag{11}$$

where $\mathbf{y}_{n_t}$ is the Local Linear discretization (10).

It is obvious that the Local Linear approximation (11) is a continuous time stochastic process that coincides with the Local Linear discretization (10) at each point of the time discretization $(t)_h$.

Note that, for all t, the random variable $\xi(t_{n_t}, \mathbf{y}_{n_t}; t - t_{n_t})$ defines the stochastic processes $\xi = \{\xi(t_{n_t}, \mathbf{y}_{n_t}; t - t_{n_t}),\ t \in [t_0, T]\}$, whose realizations depend of the realizations of the $m-$dimensional Wiener process $\mathbf{w} = (\mathbf{w}^1, \ldots, \mathbf{w}^m)$. This justify the use of terminology "strong" in the above definitions. In addition, note that for $G \equiv 0$ the LL approximation (11) for SDEs reduces to the LL approximation for ordinary differential equations (ODEs), which was subject of review in the previous ICTP preprint [27].

2.1 Convergence analysis

In this section, conditions for the strong convergence of the Local Linear approximation (11) are presented. Although this subject was early studied in [50], here we follows the results of [28], which provides more accurate order of convergence.

The first result concerns with a bound for the variance of the LL approximation paths.

Denote by $C^{1,2}$ the space of functions from $\mathbb{R} \times \mathbb{R}^d$ to $\mathbb{R}^d$ that are once and twice continuously differentiable in their first and second argument, respectively. Further, denote by $[\cdot]$ the integer part of a real number and by δ the delta de Kronecker.

Theorem 3 *Suppose that the drift and diffusion coefficient of (1) have components*

$$\mathbf{f}^k \in \mathcal{C}^{1,[\gamma+0.5]}([t_0, T] \times \mathbb{R}^d, \mathbb{R}) \quad \text{and} \quad \mathbf{g}_i^k \in \mathcal{C}^1([t_0, T], \mathbb{R}), \tag{12}$$

and satisfy the conditions

$$|\mathbf{f}(s, \mathbf{u})| + \sum_{i=1}^{m} |\mathbf{g}_i(s)| \leq K(1 + |\mathbf{u}|)$$

and

$$\left|\frac{\partial \mathbf{f}(s, \mathbf{u})}{\partial t}\right| + \left|\frac{\partial \mathbf{f}(s, \mathbf{u})}{\partial \mathbf{x}}\right| + \left|\frac{\partial^2 \mathbf{f}(s, \mathbf{u})}{\partial \mathbf{x}^2}\right| \delta_\gamma^{1.5} \leq K \tag{13}$$

for all $s \in [t_0, T]$ and $\mathbf{u} \in \mathbb{R}^d$, where K is a positive constant. Then, if $E(|\mathbf{y}_0|^2) < \infty$, the order-$\gamma$ LL approximation $\mathbf{y}$ defined in (11) satisfies

$$E\left(\sup_{t_0 \leq t \leq T} |\mathbf{y}(t)|^2 \,\middle|\, \mathcal{F}_{t_0}\right) \leq C(1 + |\mathbf{y}_0|^2), \tag{14}$$

where C is a positive constant.

Proof. As in [28] ■

Note that, in the case of $\gamma = 1$, the bound (14) can be derived from the standard conditions that ensure existence and uniqueness of a strong solution for (1)-(2) plus the continuity condition (12). Indeed, by using the Finite Increments inequality, (12) and the Lipschitsz condition for $\mathbf{f}$ and $\mathbf{g}_i$ imply (13). Further, note that a bound for higher order moments of the LL approximation, i.e., for $E(\sup\limits_{t_0 \leq t \leq T} |\mathbf{y}(t)|^{2q} |\mathcal{F}_{t_0})$, can be obtained too. See, for instance, Lemma 6 in [10].

In what follow, the following conventional definitions and notations from [32] are needed. We briefly recall that $\mathcal{M}$ denotes the set of all the multi-indexes $\alpha = (j_1, \ldots, j_{l(\alpha)})$ with $j_i \in \{0, 1, \ldots, m\}$ and $i = 1, \ldots, l(\alpha)$, where m is the dimension of $\mathbf{w}$ in (1). $l(\alpha)$ denotes the length of the multi-index α and $n(\alpha)$ the number of its zero components. $-\alpha$ and $\alpha-$ are the multi-indexes in $\mathcal{M}$ obtained by deleting the first and the last component of α, respectively. The multi-index of length zero will be denoted by v. Denote by $I_\alpha [\cdot]_{t_n, t_n+h_n}$ the multiple Ito integrals for all $\alpha \in \mathcal{M}$, and by H_α the set of adapted right continuous processes $h = \{h(t),\, t \geq t_0\}$ with left hand limits such that $\{I_{\alpha-}[h(\cdot)]_{\rho,t},\, t \geq t_0\} \in H_{(j_{l(\alpha)})}$. Further,

$$L^0 = \frac{\partial}{\partial t} + \sum_{k=1}^{d} \mathbf{f}^k \frac{\partial}{\partial \mathbf{x}^k} + \frac{1}{2} \sum_{k,l=1}^{d} \sum_{j=1}^{m} \mathbf{g}_j^k \mathbf{g}_j^l \, \frac{\partial^2}{\partial \mathbf{x}^k \partial \mathbf{x}^l}$$

denotes the diffusion operator for the SDE (1), and

$$L^j = \sum_{k=1}^{d} \mathbf{g}_j^k \frac{\partial}{\partial \mathbf{x}^k},$$

for $j = 1, \ldots, m$. Thus, the Ito coefficient function λ_α corresponding to the SDE (1) is defined as

$$\lambda_\alpha = \begin{cases} L^{j_1} \ldots L^{j_{l(\alpha)-1}} \mathbf{f}, & \text{for } j_{l(\alpha)} = 0 \\ L^{j_1} \ldots L^{j_{l(\alpha)-1}} \mathbf{g}^{j_{l(\alpha)}}, & \text{for } j_{l(\alpha)} \neq 0 \end{cases}$$

for all $\alpha \in \mathcal{M}$.

Lemma 4 *Let*

$$\mathcal{A}_\gamma = \left\{ \alpha \in \mathcal{M} : l(\alpha) + n(\alpha) \leq 2\gamma \text{ or } l(\alpha) = n(\alpha) = \gamma + \frac{1}{2} \right\} \tag{15}$$

be a hierarchical set, and $\mathcal{B}(\mathcal{A}_\gamma) = \{\alpha \in \mathcal{M} \backslash \mathcal{A}_\gamma : -\alpha \in \mathcal{A}_\gamma\}$ *the remainder set of* $\mathcal{A}_\gamma$. *Further, let* $\mathbf{y}$ *be the LL approximation (11), and* $\mathbf{z} = \{\mathbf{z}(t),\, t \in [t_0, T]\}$ *the stochastic process defined by*

$$\begin{aligned} \mathbf{z}(t) = \mathbf{y}_{n_t} &+ \sum_{\alpha \in \mathcal{A}_\gamma \backslash \{\nu\}} I_\alpha[\Lambda_\alpha(t_{n_t}, \mathbf{y}_{n_t}; t_{n_t}, \mathbf{y}_{n_t})]_{t_{n_t}, t} \\ &+ \sum_{\alpha \in \mathcal{B}(\mathcal{A}_\gamma)} I_\alpha[\Lambda_\alpha(\cdot, \mathbf{y}_\cdot; t_{n_t}, \mathbf{y}_{n_t})]_{t_{n_t}, t}, \end{aligned} \tag{16}$$

where, for any given $(t_{n_t}, \mathbf{y}_{n_t})$,

$$\Lambda_\alpha(s, \mathbf{v}; t_{n_t}, \mathbf{y}_{n_t}) = \begin{cases} L^{j_1} \ldots L^{j_{l(\alpha)-1}} \mathbf{p}_\gamma(s, \mathbf{v}; t_{n_t}, \mathbf{y}_{n_t}) & \text{if } j_{l(\alpha)} = 0 \\ L^{j_1} \ldots L^{j_{l(\alpha)-1}} \mathbf{G}^{j_l}(s) & \text{if } j_{l(\alpha)} \neq 0 \end{cases}$$

is a function of s *and* $\mathbf{v}$*, and*

$$\mathbf{p}_{\gamma}(s,\mathbf{v};r,\mathbf{u})=\begin{cases}\mathbf{f}(r,\mathbf{u})+\mathbf{f}_{\mathbf{x}}(r,\mathbf{u})(\mathbf{v}-\mathbf{u})+\mathbf{f}_{t}(r,\mathbf{u})(s-r) & \text{for } \gamma=1\\ \mathbf{f}(r,\mathbf{u})+\mathbf{f}_{\mathbf{x}}(r,\mathbf{u})(\mathbf{v}-\mathbf{u})+(\mathbf{f}_{t}(r,\mathbf{u})\ +\frac{1}{2}\sum\limits_{k,l=1}^{d}[\mathbf{G}(r)\mathbf{G}^{\intercal}(r)]^{k,l}\frac{\partial^{2}\mathbf{f}(r,\mathbf{u})}{\partial\mathbf{x}^{k}\partial\mathbf{x}^{l}})\,(s-r) & \text{for } \gamma=1.5\end{cases}$$

for all $r,s\in[t_0,T]$*, and* $\mathbf{u},\mathbf{v}\in\mathbb{R}^{d}$*. Then*

$$\mathbf{y}\equiv\mathbf{z}.$$

Moreover,

$$I_{\alpha}[\Lambda_{\alpha}(t_{n_s},\mathbf{y}_{n_s};t_{n_s},\mathbf{y}_{n_s})]_{t_{n_s},s}=I_{\alpha}[\lambda_{\alpha}(t_{n_s},\mathbf{y}_{n_s})]_{t_{n_s},s}$$

for all $\alpha\in\mathcal{A}_{\gamma}$ *and* $s\in[t_0,T]$*;*

$$E\left(\sup_{t_0\leq s\leq T}|\Lambda_{\alpha}(s,\mathbf{y}(s);t_{n_s},\mathbf{y}_{n_s})|^{2}\,\Big|\mathcal{F}_{t_0}\right)\leq C(1+|\mathbf{y}_0|^{2})$$

for all $\alpha\in\mathcal{B}(\mathcal{A}_{\gamma})$*, where* λ_{α} *denotes the Ito coefficient function corresponding to the SDE (1), and* C *is a positive constant.*

Proof. As in [28]. ■

Note that, the stochastic process $\mathbf{z}$ defined in the previous lemma is solution of the piecewise linear SDE (4)-(5) and $\Lambda_{\alpha}(\cdot;t_{n_t},\mathbf{y}_{n_t})$ denotes the Ito coefficient functions corresponding to that equation. Therefore, (16) is the Ito-Taylor expansion of the LL approximation (11).

Next lemma provides the rate of convergence of the LL approximations under conventional assumptions.

Lemma 5 *Let* $\mathcal{A}_{\gamma}$ *and* $\mathcal{B}(\mathcal{A}_{\gamma})$ *be the hierarchical and remainder sets defined as in Lemma 4. Suppose that the Ito coefficient functions* λ_{α} *corresponding to the SDE (1) satisfy:*

$$|\lambda_{\alpha}(t,\mathbf{u})-\lambda_{\alpha}(t,\mathbf{v})|\leq K_1\,|\mathbf{u}-\mathbf{v}| \tag{17}$$

for all $\alpha\in\mathcal{A}_{\gamma}$*,* $t\in[t_0,T]$ *and* $\mathbf{u},\mathbf{v}\in\mathbb{R}^{d}$*;*

$$\lambda_{-\alpha}\in C^{1,2} \text{ and } \lambda_{\alpha}\in H_{\alpha} \tag{18}$$

for all $\alpha\in\mathcal{A}_{\gamma}\cup\mathcal{B}(\mathcal{A}_{\gamma})$*; and*

$$|\lambda_{\alpha}(t,\mathbf{u})|\leq K_2(1+|\mathbf{u}|) \tag{19}$$

for all $\alpha\in\mathcal{A}_{\gamma}\cup\mathcal{B}(\mathcal{A}_{\gamma})$*,* $t\in[t_0,T]$ *and* $\mathbf{u}\in\mathbb{R}^{d}$*. Then*

$$E\left(\sup_{t_0\leq t\leq T}|\mathbf{x}(t)-\mathbf{y}(t)|^{2}\,\Big|\mathcal{F}_{t_0}\right)\leq K_3\,|\mathbf{x}_0-\mathbf{y}_0|^{2}+K_4(T-t_0)(1+|\mathbf{x}_0|^{2})h^{2\gamma}$$
$$+(T-t_0)(K_5(1+|\mathbf{x}_0|^{2})+K_6(1+|\mathbf{y}_0|^{2}))h^{2\gamma+1}$$

for $\gamma = 1.5$;

$$E\left(\sup_{t_0 \le t \le T} |\mathbf{x}(t) - \mathbf{y}(t)|^2 \Big| \mathcal{F}_{t_0}\right) \le K_3 |\mathbf{x}_0 - \mathbf{y}_0|^2$$
$$+ (T - t_0)(K_7(1 + |\mathbf{x}_0|^2) + K_8(1 + |\mathbf{y}_0|^2))h^{2\gamma}$$

for $\gamma = 1$; *and*

$$E\left(\sup_{t_0 \le t \le T} |\mathbf{x}(t) - \mathbf{y}(t)|^2 \Big| \mathcal{F}_{t_0}\right) \le K_3 |\mathbf{x}_0 - \mathbf{y}_0|^2$$
$$+ (T - t_0)(K_5(1 + |\mathbf{x}_0|^2) + K_6(1 + |\mathbf{y}_0|^2))h^4$$

for $\gamma = 1, 1.5$ *and* $\mathbf{G} \equiv \mathbf{0}$. *Here,* $\mathbf{x}$ *and* $\mathbf{y}$ *denote the solution of (1) and the LL approximation (11), respectively, and each* K_i *is a positive constant.*

Proof. As in [28]. ■

Next lemma provides a relation between the conventional convergence conditions for discretizations of SDEs with the continuity condition for $\mathbf{f}$ and $\mathbf{g}_i$, which is easier to verify in many practical cases.

Denote by $\mathcal{C}_B^l$ the space of l time continuously differentiable functions with uniformly bounded partial derivatives up to order l.

Lemma 6 *Suppose that the drift and diffusion coefficient of (1) have components*

$$\mathbf{f}^k \in \mathcal{C}_B^{2\gamma,2\gamma+1}([t_0, T] \times \mathbb{R}^d, \mathbb{R}) \quad \text{and} \quad \mathbf{g}_i^k \in \mathcal{C}_B^{2\gamma}([t_0, T], \mathbb{R}).$$

Then the Ito coefficient function λ_α *corresponding to the SDE (1) satisfies the conditions (17)-(19).*

Proof. It follows, after some easy calculations, by combining the uniformly bounded property for the derivatives of $\mathbf{f}$ and $\mathbf{g}$ with the Taylor formula with Lagrange remainder. ■

The main convergence result for the LL approximation is the following.

Theorem 7 *Let* $\mathbf{x}$ *be the solution of the SDE (1)-(2), and* $\mathbf{y}$ *the order-*γ *strong Local Linear approximation of* $\mathbf{x}$ *defined by (11). Suppose that the drift and diffusion coefficient of (1) have components*

$$\mathbf{f}^k \in \mathcal{C}_B^{2\gamma,2\gamma+1}([t_0, T] \times \mathbb{R}^d, \mathbb{R}) \quad \text{and} \quad \mathbf{g}_i^k \in \mathcal{C}_B^{2\gamma}([t_0, T], \mathbb{R});$$

and the initial conditions of $\mathbf{x}$ *and* $\mathbf{y}$ *satisfying*

$$E\left(|\mathbf{x}_0|^2\right) < \infty, \quad E\left(|\mathbf{y}_0|^2\right) < \infty, \quad \text{and} \quad E\left(|\mathbf{x}_0 - \mathbf{y}_0|^2\right) \le K h^{2\gamma},$$

for some positive constant K. *Then*

$$E\left(\sup_{t_0 \le t \le T} |\mathbf{x}(t) - \mathbf{y}(t)|^q\right) = O(h^{q\gamma})$$

with $q = 1, 2$.

Proof. It follows from Lemmas 5 and 6, and the Lyapunov inequality. ■

The Theorem 7 states that the order-γ strong Local Linear approximation $\mathbf{y}$ converges strong and uniformly to $\mathbf{x}$ on $[t_0, T]$ with order γ.

In addition, note that in the particular case of the nonautonomous ODEs (i.e., $\mathbf{G} \equiv 0$), the order-γ LL approximation (11) reduces to the expression

$$\mathbf{y}(t) = \mathbf{y}_{n_t} + \phi_1(t_{n_t}, \mathbf{y}_{n_t}; t - t_{n_t})$$

for both values of γ, that is, to the LL approximation for ODEs. In this case, from Lemma 5, the Lyapunov inequality and the condition $E(|\mathbf{x}_0 - \mathbf{y}_0|^2) \leq K_3 h^4$ follows that

$$E\left(\sup_{t_0 \leq t \leq T} |\mathbf{x}(t) - \mathbf{y}(t)| \right) = O(h^2),$$

which agrees with the convergence order of the LL approximation for ODEs when $\mathbf{y}_0$ is non random variable.

3 High Order Local Linear approximations

In this section a modification of the LL method is presented in order to improve its order of convergence. The main idea consists in obtaining a convenient approximation to the local (stochastic) remainder term $\mathbf{r}$ resulting from the approximation of the exact solution $\mathbf{x}$ of (1)-(2) by local deterministic term

$$\mathbf{l}(t; t_n, \mathbf{y}_n) = \mathbf{y}_n + \phi_\gamma(t_n, \mathbf{y}_n; t - t_n)$$

of LL approximation (11) for all $t \in [t_n, t_{n+1}]$. Notice that, given $t_n \in (t)_h$, the LL approximation (11) provides an approximation to the solution $\mathbf{u}$ of the local nonlinear equation

$$d\mathbf{u}(t) = \mathbf{f}\left(t, \mathbf{u}(t)\right) dt + \sum_{j=1}^{m} \mathbf{g}_j\left(t\right) d\mathbf{w}^j(t), \qquad t \in (t_n, t_{n+1}]$$
$$\mathbf{u}(t_n) = \mathbf{y}_n.$$

Since $\mathbf{l}(\cdot; t_n, \mathbf{y}_n)$ satisfies the linear ODE

$$d\mathbf{l}\left(t\right) = (\mathbf{A}_n \mathbf{l}(t) + \mathbf{a}_n^\gamma\left(t\right))dt, \quad t \in (t_n, t_{n+1}],$$
$$\mathbf{l}\left(t_n\right) = \mathbf{y}_n,$$

the remainder term $\mathbf{r} = \mathbf{u} - \mathbf{l}$ satisfies the SDE

$$d\mathbf{r}\left(t\right) = \mathbf{q}_\gamma\left(t_n, \mathbf{y}_n; t, \mathbf{r}(t)\right) dt + \sum_{j=1}^{m} \mathbf{g}_j\left(t\right) d\mathbf{w}^j(t), \qquad t \in (t_n, t_{n+1}], \tag{20}$$
$$\mathbf{r}\left(t_n\right) = \mathbf{0}, \tag{21}$$

where $\mathbf{A}_n$ and $\mathbf{a}_n^{\gamma}$ are defined as in (4), and

$$\mathbf{q}_{\gamma}\left(t_n, \mathbf{y}_n; s, \mathbf{r}\right) = \mathbf{f}(s, \mathbf{l}\left(s; t_n, \mathbf{y}_n\right) + \mathbf{r}) - \mathbf{f}_{\mathbf{x}}(t_n, \mathbf{y}_n)\mathbf{l}(s; t_n, \mathbf{y}_n) - \mathbf{a}_n^{\gamma}\left(s\right)$$

for all $s \in [t_0, T]$ and $\mathbf{r} \in \mathbb{R}^d$. Obviously, if $\mathbf{e}\left(.; t_n, \mathbf{y}_n\right)$ is a high order (≥ 2) approximation to $\mathbf{r}$ on $[t_n, t_{n+1}]$ then

$$\mathbf{y}_{n+1} = \mathbf{l}\left(t_n + h_n; t_n, \mathbf{y}_n\right) + \mathbf{e}\left(t_n + h_n; t_n, \mathbf{y}_n\right) \tag{22}$$

provides a better approximation to $\mathbf{u}(t_{n+1})$ than the LL approximation (11), and so a better approximation to $\mathbf{x}(t_{n+1})$. This motivates the definition of the following High Order Local Linear (HOLL) discretizations.

Definition 8 *For a given time discretization $(t)_h$, the order-κ (≥ 2) strong Local Linear discretization of the SDE (1)-(2) is defined by the recursive expression (22) starting at $\mathbf{y}_0 = \mathbf{x}_0$, where $\mathbf{e}\left(\cdot; t_n, \mathbf{y}_n\right)$ is an approximation to the solution of the auxiliary equation (20)-(21) such that*

$$E\left(\left|\mathbf{x}\left(t_n\right) - \mathbf{y}_n\right|^2\right) = O(h^{\kappa})$$

holds for all $t_n \in (t)_h$.

Moreover, an approximation for $\mathbf{x}$ in the whole interval $[t_0, T]$ is stated in the next definition.

Definition 9 *For a given time discretization $(t)_h$, the order-κ (≥ 2) strong Local Linear approximation of the SDE (1)-(2) is defined by the recursive expression*

$$\mathbf{y}(t) = \mathbf{l}\left(t; t_{n_t}, \mathbf{y}_{n_t}\right) + \mathbf{e}\left(t; t_{n_t}, \mathbf{y}_{n_t}\right),$$

where $\mathbf{y}_{n_t}$ is the order-κ (≥ 2) strong Local Linear discretization (22).

Clearly, a simple way to construct a high order approximation $\mathbf{e}$ to the correction term $\mathbf{r}$ is by applying some one-step explicit numerical integrator to the auxiliary problem (20)-(21). Indeed, if $\mathbf{u}_{n+1} = \mathbf{u}_n + \mathbf{\Lambda}^{\mathbf{y}_n}(t_n, \mathbf{u}_n; h_n)$ is some one-step explicit numerical integrator for this equation, then $\mathbf{e}(t_n + h_n; t_n, \mathbf{y}_n) = \mathbf{\Lambda}^{\mathbf{y}_n}(t_n, \mathbf{0}; h_n)$ and the HOLL discretization can be written as

$$\mathbf{y}_{n+1} = \mathbf{y}_n + \phi_{\gamma}(t_n, \mathbf{y}_n; h_n) + \mathbf{\Lambda}^{\mathbf{y}_n}(t_n, \mathbf{0}; h_n),$$

where the map ϕ_{γ} is defined in (8). In this way, a variety of HOLL discretizations can be constructed.

In next, two classes of HOLL discretization will be consider: the Local Linear - Taylor (LLT) and Local Linear - Runge Kutta (LLRK) discretizations. In the former, the solution of the auxiliary equation (20)-(21) is approximated by means of some explicit Ito-Taylor scheme, whereas in the later, explicit Runge Kutta schemes are used for that purpose.

3.1 Local Linear - Taylor discretizations

Let us consider the order-κ strong Ito-Taylor discretization $\mathbf{u}_{n+1} = \mathbf{u}_n + \mathbf{\Lambda}^{\mathbf{y}_n}(t_n, \mathbf{u}_n; h_n)$ defined by

$$\mathbf{\Lambda}^{\mathbf{y}_n}(t_n, \mathbf{u}_n; h_n) = \sum_{\alpha\in\mathcal{A}_\gamma} \rho_\alpha^{t_n,\mathbf{y}_n}(t_n, \mathbf{u}_n)\, I_\alpha[1]_{t_n,t_n+h_n}, \tag{23}$$

where $\mathcal{A}_\gamma$ is the hierarchical set (15) and

$$\rho_\alpha^{t_n,\mathbf{y}_n}(\cdot) = \begin{cases} L^{j_1}\dots L^{j_{l-1}}\mathbf{q}_\gamma(t_n, \mathbf{y}_n; \cdot), & \text{for } j_l = 0 \\ L^{j_1}\dots L^{j_{l-1}}\mathbf{g}^{j_l}, & \text{for } j_l \neq 0 \end{cases}$$

denotes the Ito coefficient functions corresponding to the auxiliary equation (20)-(21).

Definition 10 *An order-κ strong Local Linear - Taylor discretization is an order-κ strong Local Linear discretization of the form*

$$\mathbf{y}_{n+1} = \mathbf{y}_n + \phi_\gamma(t_n, \mathbf{y}_n; h_n) + \mathbf{\Lambda}^{\mathbf{y}_n}(t_n, \mathbf{0}; h_n),$$

where ϕ_γ is the map defined in (8) and $\mathbf{\Lambda}^{\mathbf{y}_n}(t_n, \mathbf{0}; h_n)$, defined by (23), is the map of the strong Ito-Taylor discretization used to approximate the solution of the auxiliary equation (20)-(21).

As an example of this class discretizations, let us consider one based on the order-2 strong Taylor scheme described in [33], Section 10.5. For simplicity, denote $\mathbf{q}_n(\cdot) \equiv \mathbf{q}_\gamma(t_n, \mathbf{y}_n; \cdot)$. Thus, we have

$$\begin{aligned}\mathbf{\Lambda}^{\mathbf{y}_n}(t_n, \mathbf{0}; h_n) &= \mathbf{q}_n(t_n, \mathbf{0})h_n + \frac{h_n^2}{2}\underline{L}^0\mathbf{q}_n(t_n, \mathbf{0}) + \sum_{j=1}^m \mathbf{g}_j(t_n)\Delta\mathbf{w}_n^j + \sum_{j=1}^m L^j\mathbf{q}_n(t_n, \mathbf{0})J_{(j,0)} \\ &+ \sum_{j=1}^m L^0\mathbf{g}^j(t_n)J_{(0,j)} + \sum_{j_1,j_2=1}^m L^{j_1}L^{j_2}\mathbf{q}_n(t_n, \mathbf{0})J_{(j_1,j_2,0)},\end{aligned}$$

where

$$\underline{L}^0 = \frac{\partial}{\partial t} + \sum_{i=1}^d \mathbf{q}_n^i \frac{\partial}{\partial x^i} \qquad \text{and} \qquad L^j = \sum_{i=1}^d \mathbf{g}_j^i \frac{\partial}{\partial x^i} \qquad j = 1, \dots, m,$$

$\Delta\mathbf{w}_n^j = \mathbf{w}_{n+1}^j - \mathbf{w}_n^j$ and $J_{(j,0)}$, $J_{(0,j)}$, $J_{(j_1,j_2,0)}$ denote Stratonovich multiple integrals over the time interval $[t_n, t_{n+1}]$.

From the rules of matrix differential calculus it is follows that

$$\begin{aligned}\mathbf{q}_n(t_n, \mathbf{0}) &= \mathbf{0}, \\ \underline{L}^0\mathbf{q}_n(t_n, \mathbf{0}) &= \delta_\gamma^{1.5}\frac{1}{2}\sum_{j=1}^m(\mathbf{I}_{d\times d}\otimes \mathbf{g}_j^\intercal(t_n))\mathbf{f}_{\mathbf{xx}}(t_n, \mathbf{y}_n)\mathbf{g}_j(t_n), \\ L^j\mathbf{q}_n(t_n, \mathbf{0}) &= \mathbf{f}_\mathbf{x}(t_n, \mathbf{y}_n)\mathbf{g}_j(t_n), \\ L^0\mathbf{g}_j(t_n) &= \frac{d\mathbf{g}_j(t_n)}{dt}, \\ L^{j_1}L^{j_2}\mathbf{q}_n(t_n, \mathbf{0}) &= (\mathbf{I}_{d\times d}\otimes \mathbf{g}_{j_2}^\intercal(t_n))\mathbf{f}_{\mathbf{xx}}(t_n, \mathbf{y}_n)\mathbf{g}_{j_1}(t_n),\end{aligned}$$

where the $d^2 \times d$ matrix $\mathbf{f_{xx}}$ denotes the Hessian of $\mathbf{f}$. Thus, the corresponding order-2 LL-Taylor discretization reduces to

$$
\begin{aligned}
\mathbf{y}_{n+1} &= \mathbf{y}_n + \phi_\gamma\left(t_n, \mathbf{y}_n; h_n\right) + \sum_{j=1}^{m} \mathbf{g}_j\left(t_n\right) \Delta \mathbf{w}_n^j + \sum_{j=1}^{m} \mathbf{f_x}(t_n, \mathbf{y}_n) \mathbf{g}_j\left(t_n\right) J_{(j,0)} \\
&+ \sum_{j=1}^{m} \frac{d\mathbf{g}_j\left(t_n\right)}{dt} J_{(0,j)} + \sum_{j_1,j_2=1}^{m} (\mathbf{I}_{d\times d} \otimes \mathbf{g}_{j_2}^{\intercal}\left(t_n\right)) \mathbf{f_{xx}}(t_n, \mathbf{y}_n) \mathbf{g}_{j_1}\left(t_n\right) J_{(j_1,j_2,0)} \\
&+ \delta_\gamma^{1.5} \frac{h_n^2}{4} \sum_{j=1}^{m} (\mathbf{I}_{d\times d} \otimes \mathbf{g}_j^{\intercal}\left(t_n\right)) \mathbf{f_{xx}}(t_n, \mathbf{y}_n) \mathbf{g}_j\left(t_n\right). \qquad (24)
\end{aligned}
$$

A drawback of this class of discretizations is that they require the computation of high order derivatives. This can be overcome by Local Linear - Runge Kutta discretizations as will be show next.

3.2 Local Linear - Runge Kutta discretizations

In order of consider the general formulation of strong Runge-Kuta (RK) methods, its is convenient to rewrite the nonautonomous auxiliary equation (20)-(21) in its equivalent autonomous Stratonovich form

$$
\begin{aligned}
d\mathbf{u}\left(t\right) &= \mathbf{f}_0\left(\mathbf{u}(t)\right) dt + \sum_{j=1}^{m} \mathbf{f}_j\left(\mathbf{u}(t)\right) \circ d\mathbf{w}^j(t), \qquad t \in (t_n, t_{n+1}], \\
\mathbf{u}\left(t_n\right) &= \mathbf{u}_n,
\end{aligned}
$$

where $\mathbf{u} = [\mathbf{r}\ t]^\intercal$, $\mathbf{u}_n = [\mathbf{0}\ t_n]^\intercal$, $\mathbf{f}_0 = [\mathbf{q}_\gamma(t_n, \mathbf{y}_n)\ 1]^\intercal$ and $\mathbf{f}_j = [\mathbf{g}_j\ 0]^\intercal$ for $j = 1, \ldots, m$. For this equation, the order-κ, s-stages explicit strong RK scheme equation has the general form [8]

$$
\begin{aligned}
\mathbf{v}_i &= \mathbf{u}_n + \sum_{k=0}^{m} \sum_{j=1}^{i-1} \mathcal{Z}_k^{i,j} \mathbf{f}_k(\mathbf{v}_j), \qquad i = 1, \ldots, s \\
\mathbf{u}_{n+1} &= \mathbf{u}_n + \sum_{k=0}^{m} \sum_{j=1}^{s} z_k^j \mathbf{f}_k(\mathbf{v}_j),
\end{aligned}
$$

where the matrices $\mathcal{Z}_k = (\mathcal{Z}_k^{i,j})$ and the vectors $z_k = (z_k^j)$ have certain random variables as their elements. In this way, the LK scheme for the auxiliary equation (20)-(21) is defined by the map

$$
\mathbf{\Lambda}^{\mathbf{y}_n}\left(t_n, \mathbf{0}; h_n\right) = (\mathbf{u}_{n+1}^1, \ldots, \mathbf{u}_{n+1}^d), \qquad (25)
$$

where $\mathbf{u}_{n+1} = (\mathbf{u}_{n+1}^1, \ldots, \mathbf{u}_{n+1}^d, \mathbf{u}_{n+1}^{d+1})$.

Definition 11 *An order-κ strong Local Linear - Runge Kutta discretization is an order-κ strong Local Linear discretization of the form*

$$
\mathbf{y}_{n+1} = \mathbf{y}_n + \phi_\gamma\left(t_n, \mathbf{y}_n; h_n\right) + \mathbf{\Lambda}^{\mathbf{y}_n}\left(t_n, \mathbf{0}; h_n\right),
$$

where ϕ_γ is the map defined in (8) and $\mathbf{\Lambda}^{\mathbf{y}_n}\left(t_n, \mathbf{0}; h_n\right)$, defined by (25), is the map of the strong Runge Kutta scheme used to approximate the solution of the auxiliary equation (20)-(21).

As examples of LLRK discretizations, we will focus on those based on two well-known explicit RK schemes with order 2 described in [7] and [33], respectively. For simplicity, in what follows we set $m=1$, $\gamma=1$ and denote $\mathbf{q}_n(\cdot)\equiv\mathbf{q}_\gamma(t_n,\mathbf{y}_n;\cdot)$.

The RK scheme introduced in [7] yields the following approximation to the solution of the auxiliary equation (20)-(21):

$$\mathbf{\Lambda}^{\mathbf{y}_n}(t_n,\mathbf{0};h_n)=h_n\sum_{j=1}^{4}\alpha_j\mathbf{q}_n(t_n+c_jh_n,\mathbf{k}_j)+\sum_{j=1}^{4}(\gamma_j^{(1)}\Delta\mathbf{w}_n+\gamma_j^{(2)}\frac{J_{(1,0)}}{h_n})\mathbf{g}(t_n+c_jh_n),$$

where $\mathbf{q}_n\equiv\mathbf{q}_\gamma^{t_n,\mathbf{y}_n}$,

$$\mathbf{k}_i=h_n\sum_{j=1}^{i-1}a_{ij}\mathbf{q}_n(t_n+c_jh_n,\mathbf{k}_j)+\sum_{j=1}^{i-1}(b_{ij}^{(1)}\Delta\mathbf{w}_n+b_{ij}^{(2)}\frac{J_{(1,0)}}{h_n})\mathbf{g}(t_n+c_jh_n),\quad i=1,\ldots,4,$$

and the coefficients a_{ij}, α_j, c_j, $b_{ij}^{(1)}$, $b_{ij}^{(2)}$, $\gamma_j^{(1)}$ and $\gamma_j^{(2)}$ are defined as

$$a_{ij}=\begin{bmatrix}0&&&\\\frac{1}{2}&0&&\\0&\frac{1}{2}&0&\\0&0&1&0\end{bmatrix},\qquad\alpha_j=\begin{bmatrix}\frac{1}{6}&\frac{1}{3}&\frac{1}{3}&\frac{1}{6}\end{bmatrix},\qquad c_j=\begin{bmatrix}0&\frac{1}{2}&\frac{1}{2}&1\end{bmatrix},$$

$$b_{ij}^{(1)}=\begin{bmatrix}0&&&\\-0.7242916356&0&&\\0.42373534060&-0.1994437050&0&\\-1.5784755060&0.8401003430&1.738375163&0\end{bmatrix},\quad b_{ij}^{(2)}=\begin{bmatrix}0&&&\\2.702000410&0&&\\1.757261649&0&0&\\-2.918524118&0&0&0\end{bmatrix},$$

$$\gamma_j^{(1)}=\begin{bmatrix}-0.7800788474&0.07363768240&1.486520013&0.2199211524\end{bmatrix},$$

$$\gamma_j^{(2)}=\begin{bmatrix}1.693950844&1.636107882&-3.024009558&-0.3060491602\end{bmatrix}.$$

The corresponding LLRK discretization is then given by

$$\mathbf{y}_{n+1}=\mathbf{y}_n+\phi_\gamma(t_n,\mathbf{y}_n;h_n)+h_n\sum_{j=1}^{4}\alpha_j\mathbf{q}_n(t_n+c_jh_n,\mathbf{k}_j)+\sum_{j=1}^{4}(\gamma_j^{(1)}\Delta\mathbf{w}_n+\gamma_j^{(2)}\frac{J_{(1,0)}}{h_n})\mathbf{g}(t_n+c_jh_n). \tag{26}$$

On the other hand, the RK scheme discussed in [33], pp. 384, involves more stochastic multiple integrals than the RK method just described. For $m=\gamma=1$, it leads to the following approximation to the solution of (20)-(21):

$$\mathbf{\Lambda}^{\mathbf{y}_n}(t_n,\mathbf{0};h_n)=\frac{h_n}{2}\{\mathbf{k}_1+\mathbf{k}_2\}+\mathbf{g}(t_n)\Delta\mathbf{w}_n+\frac{1}{h_n}\{\mathbf{g}(t_{n+1})-\mathbf{g}(t_n)\}J_{(0,1)},$$

where

$$\begin{aligned}\mathbf{k}_1&=\mathbf{q}_n(t_n+\frac{1}{2}h_n,\alpha_+)\\&=\mathbf{f}(t_n+\frac{h_n}{2},\mathbf{y}_n+\phi_\gamma(t_n,\mathbf{y}_n;\frac{h_n}{2})+\alpha_+)-\mathbf{f}_\mathbf{x}(t_n,\mathbf{y}_n)\phi_\gamma(t_n,\mathbf{y}_n;\frac{h_n}{2})-\mathbf{f}(t_n,\mathbf{y}_n)-\mathbf{f}_t(t_n,\mathbf{y}_n)\frac{h_n}{2},\\\mathbf{k}_2&=\mathbf{q}_n(t_n+\frac{1}{2}h_n,\alpha_-)\\&=\mathbf{f}(t_n+\frac{h_n}{2},\mathbf{y}_n+\phi_\gamma(t_n,\mathbf{y}_n;\frac{h_n}{2}))+\alpha_--\mathbf{f}_\mathbf{x}(t_n,\mathbf{y}_n)\phi_\gamma(t_n,\mathbf{y}_n;\frac{h_n}{2})-\mathbf{f}(t_n,\mathbf{y}_n)-\mathbf{f}_t(t_n,\mathbf{y}_n)\frac{h_n}{2},\end{aligned}$$

and

$$\alpha_{\substack{+\\-}} = \frac{1}{2}\mathbf{q}_n(t_n,\mathbf{0})\,h_n + \frac{1}{h_n}\mathbf{g}(t_n)\left\{J_{(1,0)} \substack{+\\-} \sqrt{2J_{(1,1,0)}h_n - (J_{(1,0)})^2}\right\}$$
$$= \frac{1}{h_n}\mathbf{g}(t_n)\left\{J_{(1,0)} \substack{+\\-} \sqrt{2J_{(1,1,0)}h_n - \left(J_{(1,0)}\right)^2}\right\}.$$

The corresponding LLRK discretization can then be written as

$$\mathbf{y}_{n+1} = \mathbf{y}_n + \phi_\gamma(t_n,\mathbf{y}_n;h_n) + \frac{h_n}{2}\{\mathbf{k}_1+\mathbf{k}_2\} + \mathbf{g}(t_n)\,\Delta\mathbf{w}_n + \frac{1}{h_n}\{\mathbf{g}(t_{n+1})-\mathbf{g}(t_n)\}\,J_{(0,1)}. \quad (27)$$

3.3 Convergence analysis

In this section, the convergence of the HOLL discretizations is study. The first main result is the following.

Theorem 12 *Let*

$$\mathbf{y}_{n+1} = \mathbf{y}_n + \phi_\gamma(t_n,\mathbf{y}_n;h_n) + \mathbf{\Lambda}^{\mathbf{y}_n}(t_n,\mathbf{0};h_n)$$

be a HOLL discretization of the SDE (1)-(2), where ϕ_γ *is defined by (8) and* $\mathbf{u}_{n+1} = \mathbf{u}_n + \mathbf{\Lambda}^{\mathbf{y}_n}(t_n,\mathbf{u}_n;h_n)$ *is an strong numerical integrator for the auxiliary equation (20)-(21). If there exist constants* $\gamma_2 \geq 1/2$, $\gamma_1 \geq \gamma_2 + 1/2$ *and* $C_1 > 0$ *such that conditions*

$$\left|E\left(\mathbf{r}(t_n+h_n;\,t_n,\mathbf{y}_n) - \mathbf{\Lambda}^{\mathbf{y}_n}(t_n,\mathbf{0};h_n)\middle|\mathcal{F}_{t_n}\right)\right| \leq C_1(1+|\mathbf{y}_n|^2)^{1/2}h^{\gamma_1}, \quad (28)$$

and

$$\left(E\left(|\mathbf{r}(t_n+h_n;\,t_n,\mathbf{y}_n) - \mathbf{\Lambda}^{\mathbf{y}_n}(t_n,\mathbf{0};h_n)|^2\middle|\mathcal{F}_{t_n}\right)\right)^{1/2} \leq C_1(1+|\mathbf{y}_n|^2)^{1/2}h^{\gamma_2}, \quad (29)$$

hold, then

$$\left(E\left(|\mathbf{y}_n|^2\middle|\mathcal{F}_{t_0}\right)\right)^{1/2} \leq C_2(1+|\mathbf{y}_0|^2)^{1/2}$$
$$\left(E\left(|\mathbf{x}(t_n)-\mathbf{y}_n|^2\middle|\mathcal{F}_{t_0}\right)\right)^{1/2} \leq C_2(1+|\mathbf{x}_0|^2)^{1/2}h^{\gamma_2-1/2}$$

for all $t_n \in (t)_h$ *and some constant* $C_2 > 0$.

Proof. As in [13]. ■

Theorem 12 states that the HOLL discretization $\mathbf{y}_n$ has global order of convergence $\kappa = \gamma_2 - 1/2$. This result has been derived as a straightforward consequence of the well-known Theorem 1.1 in [37], which is stated under quite general conditions on $\mathbb{R}^d$.

In practice, in order to study the global order of convergence of a particular HOLL discretization, it is necessary to verify inequalities (28) and (29) of Theorem 12. As is standard in convergence analysis of discretizations for SDEs [33, 9], this can be carried out quite straightforwardly by matching the Ito-Taylor expansions of the approximate and exact solutions.

The next Lemma shows that, if the auxiliary equation (20) is approximated by an order-κ Ito-Taylor expansion, then the corresponding HOLL discretization has global order of convergence κ.

Lemma 13 *Let*

$$\mathbf{y}_{n+1} = \mathbf{y}_n + \phi_\gamma(t_n, \mathbf{y}_n; h_n) + \mathbf{\Lambda}^{\mathbf{y}_n}(t_n, \mathbf{0}; h_n)$$

be the HOLL discretization based on the order-κ strong Ito-Taylor discretization $\mathbf{u}_{n+1} = \mathbf{u}_n + \mathbf{\Lambda}^{\mathbf{y}_n}(t_n, \mathbf{u}_n; h_n)$ *defined by*

$$\mathbf{\Lambda}^{\mathbf{y}_n}(t_n, \mathbf{u}_n; h_n) = \sum_{\alpha \in \mathcal{A}_\kappa} \rho_\alpha^{t_n, \mathbf{y}_n}(t_n, \mathbf{u}_n) I_\alpha [1]_{t_n, t_n+h_n}, \tag{30}$$

where $\rho_\alpha^{t_n,\mathbf{y}_n}$ *denotes the Ito Coefficient functions corresponding to the auxiliary equation (20)-(21). Suppose that the Ito Coefficient functions* λ_α *corresponding to the SDE (1) satisfy the conditions of Lemma 5 for the hierarchical set* $\mathcal{A}_\kappa$*. Then*

$$\left(E\left|\mathbf{x}(t_n) - \mathbf{y}_n\right|^2 \middle| \mathcal{F}_{t_0}\right)^{1/2} \leq C(1 + |\mathbf{x}_0|^2)^{1/2} h^\kappa$$

for all $t_n \in (t)_h$.

Proof. As in [13]. ∎

The next result extends Lemma 13 to the practical case in which the integration of the auxiliary equation (20) is carried out by means of some strong Ito scheme.

Theorem 14 *Let* $\mathbf{x}$ *be the solution of the SDE (1)-(2), and*

$$\mathbf{y}_{n+1} = \mathbf{y}_n + \phi_\gamma(t_n, \mathbf{y}_n; h_n) + \mathbf{\Lambda}^{\mathbf{y}_n}(t_n, \mathbf{0}; h_n)$$

be an HOLL discretization based on the order-κ general strong Ito scheme $\mathbf{u}_{n+1} = \mathbf{u}_n + \mathbf{\Lambda}^{\mathbf{y}_n}(t_n, \mathbf{u}_n; h_n)$ *defined by*

$$\mathbf{\Lambda}^{\mathbf{y}_n}(t_n, \mathbf{u}_n; h_n) = \sum_{\alpha \in \mathcal{A}_\kappa} \widetilde{\rho}_\alpha^{t_n, \mathbf{y}_n}(t_n, \mathbf{u}_n) \widetilde{I}_\alpha [1]_{t_n, t_n+h_n} + \varepsilon^{t_n, \mathbf{y}_n}(t_n + h_n), \tag{31}$$

where the first right hand term is an approximation to Ito-Taylor expansion (30) and $\varepsilon^{t_n,\mathbf{y}_n}$ *is the remainder term. Suppose that the drift and diffusion coefficient of (1) have components*

$$\mathbf{f}^j \in \mathcal{C}_B^{2\kappa, 2\kappa+1}([t_0, T] \times \mathbb{R}^d, \mathbb{R}) \quad \text{and} \quad \mathbf{g}_i^j \in \mathcal{C}_B^{2\kappa}([t_0, T], \mathbb{R});$$

and the initial conditions of $\mathbf{x}$ *and* $\mathbf{y}$ *satisfying*

$$E\left(|\mathbf{x}_0|^2\right) < \infty, \quad E\left(|\mathbf{y}_0|^2\right) < \infty, \quad \text{and} \quad E\left(|\mathbf{x}_0 - \mathbf{y}_0|^2\right) \leq K h^{2\kappa},$$

for some positive constant K. Further, suppose that the approximations $\widetilde{\rho}_\alpha^{t_n,\mathbf{y}_n}$ *and* $\widetilde{I}_\alpha$ *to* $\rho_\alpha^{t_n,\mathbf{y}_n}$ *and* I_α*, and the remainder term* $\varepsilon^{t_n,\mathbf{y}_n}$ *satisfy the conditions*

$$E\left(\left|\rho_\alpha^{t_n,\mathbf{y}_n}(t_n, \mathbf{0}) - \widetilde{\rho}_\alpha^{t_n,\mathbf{y}_n}(t_n, \mathbf{0})\right|^2\right) \leq K_1 h^{2\kappa},$$

$$E\left(\left|\textstyle\sum_{0 \leq i \leq n-1} \varepsilon^{t_i, \mathbf{y}_i}(t_{i+1})\right|^2\right) \leq K_1 h^{2\kappa}$$

and

$$E\left(\left|I_{\alpha}-\widetilde{I}_{\alpha}\right|^{2}\right)\leq K_{1}h^{2\kappa+1}$$

for all $t_n \in (t)_h$, $\alpha \in \mathcal{A}_\kappa$, *and some positive constant* K_1. *Then*

$$E\left(\max_{n}\left|\mathbf{x}\left(t_{n}\right)-\mathbf{y}_{n}\right|^{2}\right)\leq Ch^{2\kappa},$$

where C *is a positive constant.*

Proof. As in [13]. ■

The class of HOLL discretizations covered by Theorem 14 is quite wide. Indeed, standard schemes that may be used to solve the auxiliary equation (20)-(21) typically can be written in the form of a general strong Ito approximation (31) whose coefficient functions $\widetilde{\rho}_{\alpha}^{t_n,\mathbf{y}_n}$ and the remainders $\varepsilon^{t_n,\mathbf{y}_n}$ fulfill the conditions of this theorem. It is so, in particular, for standard schemes derived from the Ito-Taylor expansion (such as strong Taylor and multi-step schemes, see chapters 10 and 11 in [33]) and Runge-Kutta schemes [8]. Thus, in general one can conclude that a HOLL discretization inherits the (high) order of convergence of the numerical scheme taken up to integrate the auxiliary equation.

It is work of noting that, for some simple schemes, the order of strong global convergence as provided by Theorem 14 can be obtained under weaker conditions [20]. Furthermore, the thesis of Theorem 12 can be strengthened to uniform convergence if more restrictive conditions about the coefficients of the SDE are assumed, such as those in Theorem 5.6.3 in [49]. In addition, in some practical cases such as models for prices in finances or for concentrations of substance in chemical reactions, it is necessary to state the convergence of the numerical integrators into certain domain $\mathcal{D} \subset \mathbb{R}^d$. With that purpose, modifications to HOLL discretizations can be carry out by following the procedure introduced in [25] based on modified Ito-Taylor expansions for auxiliary drift and diffusion coefficients. In this situation, no global Lipschitz conditions for the coefficients of the SDE and their derivatives are required.

4 Dynamics of the Local Linear discretizations for linear equations

Consider the scalar linear system (test equation)

$$dx\left(t\right)=\lambda x\left(t\right)dt+dw(t),\quad t\in\mathbb{R},\tag{32}$$

where λ is a complex number and w is a (two-sided) standard Wiener process [1]. Denote by μ_t the (marginal) distribution of the solution $x(t)$ of (32) at time t. If $\mathrm{Re}(\lambda) < 0$ then the dynamical system generated by this equation has a (unique) stationary distribution μ and a stationary solution x^∞. From a random dynamical viewpoint, these two aspects characterize the asymptotic behavior of the system in the phenomenological and pathwise sense, respectively (see [1]). More precisely, μ is the limit of the flow of distributions $(\mu_t : t \geq 0)$ while the stationary solution determines the random attractor of the flow of solutions. It is desirable that a numerical integrator reproduces both dynamical features. Some criteria for this purpose were proposed in [13]. This is discussed next, as well as its application to HOLL discretizations.

4.1 Ergodicity and $A-$stability

The standard concept of $A-$stability for numerical integrators of SDEs with additive noise is based on the test equation (32) (see, e.g., [22, 33, 2, 38]). For a uniform time partition $(t)_h$ with $t_n = nh$, $n \in \mathbb{Z}$, let

$$y_{n+1} = \phi(\lambda h)\, y_n + \xi_n \tag{33}$$

be the recursion resulting from a given numerical integrator when applied to equation (32). It is assumed that the random variable ξ_n does not dependent on y_k for all $k \in \mathbb{Z}$. The region of absolute stability of the integrator is defined as

$$S = \{\lambda h \in \mathbb{C} : h > 0, \quad \mathrm{Re}(\lambda) < 0, \quad |\phi(\lambda h)| < 1\}.$$

A numerical integrator is said to be $A-$stable if S contains the whole left half of the complex plane. It can be shown [4] that a discretization is $A-$stable if and only if it is numerically stable in the mean-square sense for the test equation (32) (this should not be confused with the standard concept of mean-square stability [46, 49, 21], which is not applicable in case of equations with additive noise).

Theorem 15 *LL discretizations of any order are $A-$stable.*

Proof. As in [13]. ■

Note that a critical drawback of the concept of $A-$stability is that it does not directly ask for the numerical integrator to reproduce either phenomenological or pathwise dynamical aspects of the stochastic system. Notwithstanding, there is an immediate link between $A-$stability and conservation of ergodicity of linear test equations, as the following lemma shows.

Lemma 16 *For the test equation (32) with $\mathrm{Re}(\lambda) < 0$, assume that a discretization (33) is $A-$stable and satisfies the following condition:*

$$\xi_n,\ n \in \mathbb{Z},\ \text{are independent and identically distributed with } E\left(\xi_1^2\right) < \infty. \tag{34}$$

Then, (33) defines an ergodic discrete-time process.

Proof. As in [13]. ■

Thus, in general $A-$stability ensures preservation of ergodicity for linear equations with additive noise; and so, because Theorem 15, all HOLL discretization retains the ergodicity of these equations.

However, as it has been pointed out for a number of numerical integrators (see e.g. [51], [48], [36] and references therein), $A-$stability is not a sufficient condition for the ergodicity preservation in cases of more complex SDEs. In particular, preservation of ergodicity by LL discretizations in case of nonlinear SDEs with additive noise has been proved in [19]. Nevertheless, it should be emphasized that ergodicity of a numerical scheme is a stability property in the phenomenological sense (i.e., in regards to the evolution of its probability distribution): it guarantees that the numerical system has an stationary distribution υ and the flow of marginal distributions υ_{t_n} of the discretization y_n converges to υ as n increases. This does not characterize the asymptotic pathwise behavior of the approximate solution, i.e. stability in the pathwise or random dynamical sense, which is the most meaningful one for strong numerical integrators. This will be dealt with next.

4.2 Random attractor and $RA-$stability

The concept of attractor plays a fundamental role in understanding the asymptotic behavior of deterministic dynamical systems. Likewise, pathwise asymptotic behavior of a random dynamical system (RDS) can be characterized through random attractors ([1, 32, 40]). It is then natural to ask for a strong numerical integrator to (approximately) reproduce the random attractors of the original system (see, e.g., [32, 44] for works on this viewpoint in a general setting). For the particular test equation (32), this requirement can be explicitly elaborated. For this, we first recall some basic definitions from the theory of RDS (for more details, see the references just mentioned).

A RDS over a measure preserving flow of transformations $(\Omega, \mathcal{F}, P, (\theta_t)_{t\in\mathbb{T}})$ with time $\mathbb{T}$ $(0 \in \mathbb{T})$ and phase space $(\mathcal{X}, \mathcal{B}) = \left(\mathbb{R}^d, \mathcal{B}\right)$ is a mapping $\varphi : \mathbb{T} \times \Omega \times \mathcal{X} \to \mathcal{X}$ with the following properties:

i) φ is $\mathcal{B}(\mathbb{T}) \times \mathcal{F} \times \mathcal{B}$, $\mathcal{B}-$measurable;

ii) The mappings $\varphi(t, \omega) := \varphi(t, \omega, .) : \mathcal{X} \to \mathcal{X}$ form a cocycle over θ, i.e., $\varphi(0, \omega) = id_{\mathcal{X}}$ and $\varphi(t + s, \omega) = \varphi(t, \theta_s\omega) \circ \varphi(s, \omega)$ for $s, t \in \mathbb{T}$ and $\omega \in \Omega$.

A random set B is a family $(B\,(\omega) : \omega \in \Omega)$ of subsets of $\mathcal{X}$ such that its graph $\{(\omega, x) : x \in B\,(\omega)\}$ is $\mathcal{F}$, $\mathcal{B}-$measurable. A random compact set is a random set A such that $A(\omega)$ is compact for all $\omega \in \Omega$ and $\omega \longmapsto d\,(x, A(\omega)) = \inf\{d\,(x, y) : y \in A(\omega)\}$ is measurable for each $x \in \mathcal{X}$.

Following [40] (see also [1]), for clearness, we will use the definition of random attractor in terms of forward attraction instead of pullback convergence. A random compact set A is called a global attractor of the RDS φ over $(\Omega, \mathcal{F}, P, (\theta_t)_{t\in\mathbb{R}})$ if it satisfies the following properties:

i) forward invariance: $\varphi(t, \omega)A(\omega) = A(\theta_t\omega)$ for all $t \in \mathbb{T}$, $\omega \in \Omega$;

ii) forward attraction: for any random set C, $d_H(\varphi(t, \omega)C(\omega), A(\theta_t\omega)) \to 0$ in probability as $t \to \infty$, where d_H denotes the Hausdorff semi-metric, i.e., $d_H\,(D_1, D_2) = \sup_{x\in D_1} \inf_{y\in D_2} d\,(x, y)$.

The test equation (32) with $\operatorname{Re}(\lambda) < 0$ generates a RDS φ over $(\Omega, \mathcal{F}, P, (\theta_t)_{t\in\mathbb{R}})$, where $\Omega = C\,(\mathbb{R}, \mathbb{R})$, P is the Wiener measure, θ is the flow of shifts $\theta_t : (\omega_.) \longmapsto (\omega_{.+t} - \omega_t)$, and φ

is defined by the flow of orbits $\varphi(t,\omega)x_0 = x\,(t;x_0)$. Here, $x\,(t;x_0)$ denotes the solution of (32) with initial value x_0 at $t=0$. The global random attractor of this RDS is $A=\{x^\infty\,(0)\}$, where x^∞ denotes the stationary solution

$$x^\infty\,(t) = e^{\lambda t}\int_{-\infty}^{t} e^{-\lambda s}dW_s, \qquad t\in\mathbb{R}.$$

Likewise, if the recursion (33) associated with a discretization method has independent and identically distributed errors ξ_n, $n\in\mathbb{Z}$, then it generates a discrete-time RDS $\widetilde{\varphi}$ over $(\widetilde{\Omega},\widetilde{\mathcal{F}},\widetilde{P},(\widetilde{\theta}_n)_{n\in\mathbb{Z}})$, where $\widetilde{\Omega}=\mathbb{R}^{\mathbb{Z}}$, $\widetilde{P}$ is the distribution on $\widetilde{\Omega}$ induced by P through the map $\omega\longmapsto(\xi_{t_n}(\omega))_{n\in\mathbb{Z}}$, $\widetilde{\theta}$ is the flow of discrete time shifts $\widetilde{\theta}_n:(u_\cdot)\longmapsto(u_{\cdot+n})$, and $\widetilde{\varphi}$ is defined by the flow of orbits $\widetilde{\varphi}(t_n,\omega)y_0=y_n$. If the discretization is also $A-$stable then this RDS has a global random attractor under quite general assumptions. For later reference, this property is stated in the following.

Lemma 17 *For the test equation (32) with* $\mathrm{Re}\,(\lambda)<0$*, assume that a discretization method (33) is* $A-$*stable and satisfies the condition (34). Then, its corresponding discrete time RDS has the stationary solution* $y^\infty=(y_n^\infty)_{n\in\mathbb{Z}}$ *given by*

$$y_n^\infty=\sum_{k=0}^{\infty}\phi^k\,(\lambda h)\,\xi_{n-k},$$

and the global random attractor $\widetilde{A}=\{y_0^\infty\}$.

Proof. As in [13]. ■

Thus, the requirement of closeness between the random attractor of the original system and that of the discretization method leads, in case of the linear test system (32), to compare the random point sets A and $\widetilde{A}$. This motivates the introduction of the following concept.

Definition 18 *A numerical integrator is called Random* $A-$*stable (*$RA-$*stable) if, when applied to the test equation (32) with* $\mathrm{Re}\,(\lambda)<0$,

i) it has a global random attractor y_0^∞*, for all* $h>0$*; and*

ii) the global random attractor y_0^∞ *converges in probability to the random attractor* $x^\infty\,(0)$ *of the linear system (32) as* $h\to0$.

The following results state sufficient conditions for $RA-$stability.

Lemma 19 *For the test equation (32) with* $\mathrm{Re}\,(\lambda)<0$*, assume that a discretization (33) is* $A-$*stable and satisfies the condition (34). In addition, assume that there exist*

i) constants $\delta>1$ *and* $C>0$ *such that*

$$\left|\phi\,(\lambda h)-e^{\lambda h}\right|\le Ch^\delta;$$

ii) constants $\gamma_1 \geq \gamma_2 + 1/2$ *and* $C_1 > 0$ *such that*

$$|E(\xi_1) - E(\eta_1)| \leq C_1 h^{\gamma_1};$$

iii) constants $\gamma_2 > 1/2$ *and* $C_2 > 0$ *such that*

$$\left(E|\xi_1 - \eta_1|^2\right)^{1/2} \leq C_2 h^{\gamma_2},$$

where

$$\eta_n = e^{\lambda t_n} \int_{t_{n-1}}^{t_n} e^{-\lambda s} dW_s.$$

Then, the numerical integrator (33) is $RA-$*stable. Specifically, there exist constants* A *and* B *such that*

$$\left(E|y_0^\infty - x^\infty(0)|^2\right)^{1/2} \leq A h^{\gamma_2 - 1/2} + B h^{\delta - 1}.$$

Proof. As in [13]. ■

As a consequence of the previous lemma, the following central criterion for the $RA-$stability of a numerical integrator is obtained.

Theorem 20 *Let* $x(s;t,z)$ *be the solution of the test equation (32) for* $s \geq t$ *with initial value* $x(t;t,z) = z$*, and let* $y_{n+1} = y_n + \digamma(t_n, y_n; h)$ *be a numerical integrator satisfying the following three conditions:*

i) for all $z \in \mathbb{R}$ *and* $t \in [t_0, T-h]$*,*

$$|E(x(t+h;t,z) - z - \digamma(t,z;h))| \leq C\left(1+z^2\right)^{1/2} h^{\gamma_1}, \tag{35}$$

$$\left(E\left(|x(t+h;t,z) - z - \digamma(t,z;h)|^2\right)\right)^{1/2} \leq C\left(1+z^2\right)^{1/2} h^{\gamma_2}, \tag{36}$$

where $\gamma_2 > 1/2$*,* $\gamma_1 \geq \gamma_2 + 1/2$ *and* $C > 0$*;*

ii) when applied to the test equation (32), y *reduces to the recursion (33) fulfilling the condition (34); and*

iii) y *is* $A-$*stable.*

Then, the numerical integrator y *is* $RA-$*stable. Moreover, if* $x^\infty(0)$ *is the random attractor of (32) and* y_0^∞ *is the random attractor of* y *when applied to (32), then there exists a constant* $A > 0$ *such that*

$$\left(E\left(|y_0^\infty - x^\infty(0)|^2\right)\right)^{1/2} \leq A h^\gamma,$$

where $\gamma = \gamma_2 - 1/2$*.*

Proof. As in [13]. ■

Notice that the conditions (35)-(36) of Theorem 20 are just the standard (local) conditions for guaranteeing (global) order of convergence γ of a numerical integrator. In this sense, Theorem 20 means that a consistent discretization method (under said standard conditions) is $RA-$stable if it is $A-$stable. Moreover, it states that the random attractor of a linear system is approached by the attractor of an $RA-$stable numerical scheme with a rate that is just its order of convergence γ.

In general, standard explicit integrators do not fulfill the $A-$stability assumption of Theorem 20. A number of implicit methods also fail to satisfy this condition. Other implicit methods are $A-$stable and therefore $RA-$stable, but their order of convergence is low, so providing slow approximation to the exact attractors. There are also classes of high order implicit integrators that are $RA-$stable, but they involve in general considerably computational cost.

In regards to LL and HOLL discretizations, the following result holds as a straightforward consequence of the theorems 15 and 20.

Theorem 21 *LL discretizations of any order are $RA-$stable.*

According to Theorem 21, LL and HOLL discretizations are $RA-$stable, though the latter achieves more precise approximation to random attractors of linear systems than the former due to higher order of convergence.

5 Local Linear approximation for equations with jumps

Let $(\Omega, \mathcal{F}, P)$ be the underlying complete probability space, and $\{\mathcal{F}_t,\, t \geq t_0\}$ an increasing right continuous family of complete sub σ-algebras of $\mathcal{F}$. Consider a d-dimensional jump diffusion process $\mathbf{x}$ defined by the following stochastic differential equation

$$d\mathbf{x}(t) = \mathbf{f}(t, \mathbf{x}(t))dt + \sum_{i=1}^{m} \mathbf{g}_i(t)d\mathbf{w}^i(t) + \sum_{i=1}^{p} \mathbf{h}_i(t, \mathbf{x}(t))d\mathbf{q}^i(t) \tag{37}$$

$$\mathbf{x}(t_0) = \mathbf{x}_0, \tag{38}$$

where $\mathbf{f}$, $\mathbf{g}_i$ and $\mathbf{h}_i$ are differentiable functions on $[t_0, T] \times \mathbb{R}^d$, and $\mathbf{w} = (\mathbf{w}^1, \ldots, \mathbf{w}^m)$ is an m-dimensional $\mathcal{F}_\tau$-adapted standard Wiener process. Here, each $\mathbf{q}^i(t)$ could be either an $\mathcal{F}_t$-adapted Poisson counting process $\mathbf{n}^i(t)$ with intensity μ^i, or an $\mathcal{F}_t$-adapted compensated Poisson processes $\mathbf{n}^i(t) - \mu^i t$. It is assumed that $\mathbf{w}^i(t)$ and $\mathbf{q}^j(t)$ are all independent with zero probability of simultaneous jumps. Linear growth restriction, uniform Lipschitz and smoothness conditions for the functions $\mathbf{f}$, $\mathbf{g}_i$ and $\mathbf{h}_i$ are also assumed in order to ensure a unique strong solution of the equation (37)-(38).

Consider the time discretization $(t)_h = \{t_n : n = 0, 1, \ldots, N\}$, with maximum step-size $h \in (0, 1)$, defined as a sequence of $\mathcal{F}$-stopping times that satisfy

$$t_0 < t_1 < \cdots < t_N = T$$

and

$$\sup_n (h_n) \leq h$$

w.p.1, where t_n is $\mathcal{F}_{t_n}$-measurable for each $n = 0, 1, \ldots, N-1$, and $h_n = t_{n+1} - t_n$. In addition, let us denote

$$n_t = \max\{n = 0, 1, 2, \ldots : \ t_n \leq t \text{ and } t_n \in (t)_h\}$$

for all $t \in [t_0, T]$.

Further, consider the sequence of jump times $\{\sigma\}_{\mu^i} = \{\sigma_{i,n} : n = 0, 1, 2, \ldots\}$ associated to $\mathbf{q}^i$, which is defined as an increasing sequence of random variables such that $\sigma_{i,n+1} - \sigma_{i,n}$ is exponentially distributed with parameter μ^i for all n and i. Without loss of generality, it is assumed that $\{\sigma\}_{\mu^i} \subset (t)_h$, for all $i = 1, \ldots, p$. In addition, let us assume that only the first r Poisson processes $\mathbf{q}^i$ are compensated.

It is well known (see, e.g., [42]) that the solution of (37) is given by

$$\mathbf{x}(t) = \mathbf{x}(t-) + \sum_{i=1}^{p} \mathbf{h}_i(t, \mathbf{x}(t-))\Delta\mathbf{n}_t^i \tag{39}$$

where $\Delta\mathbf{n}_t^i$ is the increment of the process $\mathbf{n}^i$ at the time instant t, and $\mathbf{x}(t-)$ denotes the solution of the SDE

$$d\mathbf{z}(t) = \overline{\mathbf{f}}(t, \mathbf{z}(t))dt + \sum_{i=1}^{m} \mathbf{g}_i(t)d\mathbf{w}^i(t) \tag{40}$$

$$\mathbf{z}(\sigma_{i,n}) = \mathbf{x}(\sigma_{i,n}) \tag{41}$$

with drift coefficient

$$\overline{\mathbf{f}}(t, \mathbf{z}(t)) = \mathbf{f}(t, \mathbf{z}(t)) - \sum_{j=1}^{r} \mathbf{h}_j(t, \mathbf{z}(t))\mu^j \tag{42}$$

for all t between two consecutive jump times $\sigma_{i,n}$ and $\sigma_{j,l}$. Thus, an strong approximation to $\mathbf{x}$ could be easily obtained by combining a strong numerical integrator for (40)-(41) with the expression (39).

In case that strong LL discretizations as defined in Sections 2 and 3 are consider, the above leads to the following two definitions.

Definition 22 *For a given time discretization $(t)_h$, the order-γ strong Local Linear discretization of the jump diffusion process $\mathbf{x}$ is defined by the recursive relation*

$$\mathbf{y}_{n+1} = \mathbf{y}_{n+1^-} + \sum_{i=1}^{p} \mathbf{h}_i(t_{n+1}, \mathbf{y}_{n+1^-})\Delta\mathbf{n}_{t_{n+1}}^i, \tag{43}$$

where $\mathbf{y}_{n+1^-}$ denotes an order-γ strong Local Linear discretization of the SDE (40)-(41) at t_{n+1}.

Definition 23 *For a given time discretization $(t)_h$, the stochastic process $\mathbf{y} = \{\mathbf{y}(t),\, t \in [t_0, T]\}$ is called the order-γ strong Local Linear approximation of the jump diffusion process $\mathbf{x}$ if*

$$\mathbf{y}(t) = \mathbf{y}(t-) + \sum_{i=1}^{p} \mathbf{h}_i(t, \mathbf{y}(t-))\Delta\mathbf{n}_t^i, \tag{44}$$

where $\mathbf{y}(t-)$ denotes an order-γ strong Local Linear approximation of the SDE (40)-(41) at t.

Note that the above definitions are valid for SDEs driven for nonhomogeneous Poisson processes too, i.e., when the intensity of such a processes is not a constant on $[t_0, T]$. In this case, μ^j must be replaced by $\mu^j(t)$ in (42). Further note that the LL discretization (43) belong to the class of the so called *jump-adapted* discretizations because it is defined on a time discretization $(t)_h$ that includes all the jump times of the Poisson measure. Alternatively, *regular* LL discretizations (that do not include these jump times) can be constructed as well, but they have been no consider so far. See [6] for an exhaustive presentation of this last subject.

5.1 Convergence analysis

In order to study the convergence of the LL approximations (44), the following general result is useful.

Lemma 24 *Let* $\mathbf{z}$ *be the solution of the SDE (40), and* $\widetilde{\mathbf{z}}(t) = \widetilde{\mathbf{z}}_{n_t} + \digamma(t_{n_t}, \widetilde{\mathbf{z}}_{n_t}; t - t_{n_t})$ *the approximation to* $\mathbf{z}(t)$ *defined by the numerical integrator* $\widetilde{\mathbf{z}}_n = \widetilde{\mathbf{z}}_{n-1} + \digamma(t_{n-1}, \widetilde{\mathbf{z}}_{n-1}; h_{n-1})$ *such that*

$$E\left(\sup_{t_0 \le s \le T} |\widetilde{\mathbf{z}}(s)|^2 \,|\mathcal{F}_{t_0}\right) \le K(1 + |\widetilde{\mathbf{z}}(t_0)|^2)$$

and

$$E\left(\sup_{t_0 \le s \le T} |\mathbf{z}(s) - \widetilde{\mathbf{z}}(s)|^2 \,|\mathcal{F}_{t_0}\right) \le K h^{2\kappa},$$

where K *is a positive constant independent of* $\widetilde{\mathbf{z}}(t_0)$*. Let* $\mathbf{y} = \{\mathbf{y}(t),\ t \in [t_0, T],\ \mathbf{y}(t_0) = \mathbf{x}_0\}$ *be the stochastic process defined as*

$$\mathbf{y}(t) = \mathbf{y}(t-) + \sum_{i=1}^{p} \mathbf{h}_i(t, \mathbf{y}(t-)) \Delta \mathbf{n}_t^i,$$

where $\mathbf{y}(t-)$ *denotes the value of the approximation* $\widetilde{\mathbf{z}}(t)$ *to the solution of (40) on the interval* $[t_{n-1}, t_n]$ *with initial condition* $\mathbf{z}(t_{n_t}) = \mathbf{y}(t_{n_t})$*, and* $\Delta \mathbf{n}_t^i$ *the increment of the Poisson process* $\mathbf{n}^i$ *at the time instant* t*. If the functions* $\mathbf{h}_i$ *defined in (37) satisfy the conditions*

$$|\mathbf{h}_i(t, \mathbf{u}) - \mathbf{h}_i(t, \mathbf{v})| \le K\,|\mathbf{u} - \mathbf{v}|$$

and

$$|\mathbf{h}_i(t, \mathbf{u})| \le K(1 + |\mathbf{u}|)$$

for $t \in [t_0, T]$ *and* $\mathbf{u}, \mathbf{v} \in \mathbb{R}^d$*, then there exist positive constants* D_1, D_3, D_4, D_5 *such that*

$$E\left(\sup_{t_0 \le s \le T} |\mathbf{y}(s)|^2 \,|\mathcal{F}_{t_0}\right) \le D_1(1 + |\mathbf{y}(t_0)|^2)$$

and

$$E\left(\sup_{t_0 \le t \le T} |\mathbf{x}(t) - \mathbf{y}(t)|^2 \,|\mathcal{F}_{t_0}\right) \le D_4\,|\mathbf{x}_0 - \mathbf{y}(t_0)|^2 + \left(D_5(1 + |\mathbf{x}_0|^2) + D_3(1 + |\mathbf{y}(t_0)|^2)\right) h^{2\kappa},$$

where $\mathbf{x}$ *is the solution of the SDE with jumps (37)-(38).*

Proof. As in [29]. ■

The following theorem states the convergence rate of the LL approximations for the jump diffusion processes.

Theorem 25 *Suppose that the functions* $\mathbf{h}_i$ *defined in (37) satisfy the inequalities*

$$|\mathbf{h}_i(t,\mathbf{u}) - \mathbf{h}_i(t,\mathbf{v})| \leq K\,|\mathbf{u}-\mathbf{v}|$$

and

$$|\mathbf{h}_i(t,\mathbf{u})| \leq K(1+|\mathbf{u}|)$$

for some positive constant K*,* $t \in [t_0,T]$ *and* $\mathbf{u},\mathbf{v} \in \mathbb{R}^d$*. Further, suppose that the functions* $\overline{\mathbf{f}}$ *and* $\mathbf{g}_i$ *defined in (40) satisfies the conditions under which a Local Linear approximation for SDEs converges to the solutions of (40) with order* κ*. Then the corresponding Local Linear approximation (44) to the solution* $\mathbf{x}$ *of the SDE with jump (37)-(38) satisfies*

$$E\left(\sup_{t_0\leq s\leq T} |\mathbf{y}(s)|^2\,|\mathcal{F}_{t_0}\right) \leq C_1\ (1+|\mathbf{y}(t_0)|^2),$$

where C_1 *is a positive constant. Moreover, if the initial conditions satisfy*

$$E\left(|\mathbf{x}_0|^2\right) < \infty, \qquad E\left(|\mathbf{y}_0|^2\right) < \infty, \qquad \text{and} \quad E\left(|\mathbf{x}_0-\mathbf{y}_0|^2\right) \leq K_1 h^{2\kappa},$$

where K_1 *is a positive constant, then*

$$E\left(\sup_{t_0\leq s\leq T} |\mathbf{x}(s)-\mathbf{y}(s)|^2\,|\mathcal{F}_{t_0}\right) \leq C_2 h^{2\kappa},$$

where C_2 *is a positive constant.*

Proof. As in [29]. ■

6 Local Linearization schemes

This section deals with the practical issues of the LL method, that is, with the so called Local Linearization schemes for SDEs.

Roughly speaking, every numerical implementation of a Local Linear discretization of any order is generically called Local Linearization scheme. More precisely, this can be defined as follows.

Definition 26 *For an order-*κ *strong Local Linear discretization* $\mathbf{y}_{n+1} = \mathbf{y}_n + \digamma_\kappa(t_n,\mathbf{y}_n;h_n)$ *of a SDE, all recursion of the form*

$$\widetilde{\mathbf{y}}_{n+1} = \widetilde{\mathbf{y}}_n + \widetilde{\digamma}_\kappa(t_n,\widetilde{\mathbf{y}}_n;h_n), \qquad \text{with } \widetilde{\mathbf{y}}_0 = \mathbf{y}_0,$$

where $\widetilde{\digamma}_\kappa$ *denotes a numerical algorithm to compute* $\digamma_\kappa$*, is generically called strong Local Linearization scheme.*

Regularly, the numerical implementation of an order-κ Local Linear discretization with $\kappa = 1, 1.5$ is simply called Local Linearization (LL) scheme. Otherwise, that is if $\kappa > 1.5$, the numerical implementation is called High Order Local Linearization (HOLL) scheme.

6.1 LL schemes

There is a large variety of LL schemes which differ with respect to the algorithm that is used in the numerical implementation of the Local Linear discretization

$$\mathbf{y}_{n+1} = \mathbf{y}_n + \phi_\gamma(t_n, \mathbf{y}_n; h_n) + \xi(t_n, \mathbf{y}_n; h_n)$$

defined in (10). Depending of the way of computing ϕ_γ and ξ**,** different LL scheme can be obtained. We refer to the ICTP preprint [27] for a detailed presentation of a number of methods for computing ϕ_γ. Thus, here, we mostly focus on the computation of ξ.

A directly way to compute ξ is by means of the stochastic cuadrature rule or by its composite version [54], which yields to the approximations

$$\widetilde{\xi}\,(t_n, \widetilde{\mathbf{y}}_n; h_n) = \sum_{i=1}^{m} e^{\widetilde{\mathbf{A}}_n h_n} \mathbf{g}_i(t_n)(\mathbf{w}^i(t_{n+1}) - \mathbf{w}^i(t_n)) \tag{45}$$

and

$$\widetilde{\xi}\,(t_n, \widetilde{\mathbf{y}}_n; h_n) = \sum_{j=0}^{r_n-1} e^{\frac{1}{r_n}\widetilde{\mathbf{A}}_n h_n (r_n - j)} \sum_{i=1}^{m} \mathbf{g}_i(\tau_j)(\mathbf{w}^i(\tau_{j+1}) - \mathbf{w}^i(\tau_j)), \tag{46}$$

respectively, where $\widetilde{\mathbf{A}}_n = \mathbf{f_x}(t_n, \widetilde{\mathbf{y}}_n)$, r_n is a integer number such that $r_n \sim \frac{1}{h_n^\alpha}$ with $\alpha \geq 0$ and $\tau_j = t_n + jh_n/r_n$.

Alternatively, since the integral involved in the definition of ξ does not dependent of the Winner process $\mathbf{w}$, ξ can be rewritten as

$$\begin{aligned}\xi\,(t_n, \mathbf{y}_n; h_n) = &\sum_{i=1}^{m} \mathbf{g}_i(t_{n+1})\mathbf{w}^i(t_{n+1}) - e^{\mathbf{A}_n h_n} \mathbf{g}_i(t_n)\mathbf{w}^i(t_n) \\ &- \int_{t_n}^{t_{n+1}} e^{\mathbf{A}_n (t_{n+1} - u)} (\frac{d}{du}\mathbf{g}_i(u) - \mathbf{A}_n \mathbf{g}_i(u))\mathbf{w}^i(u) du.\end{aligned}$$

By applying the deterministic composite Trapezoidal rule (see, e.g., [12]) to the integral in the last expression, the following approximation to ξ is obtained

$$\widetilde{\xi}\,(t_n, \widetilde{\mathbf{y}}_n; h_n) = \sum_{i=1}^{m} \mathbf{g}_i(t_{n+1})\mathbf{w}^i(t_{n+1}) - e^{\widetilde{\mathbf{A}}_n h_n} \mathbf{g}_i(t_n)\mathbf{w}^i(t_n) - \frac{h_n}{2r_n} \sum_{j=0}^{r_n-1} \mathbf{q}_i(j+1) + \mathbf{q}_i(j), \tag{47}$$

where $\mathbf{q}_i(j) = e^{\frac{1}{r_n}\widetilde{\mathbf{A}}_n h_n (r_n - j)} (\frac{d}{du}\mathbf{g}_i(\tau_j) - \widetilde{\mathbf{A}}_n \mathbf{g}_i(\tau_j))\mathbf{w}^i(\tau_j)$, and r_n and τ_j defined as before.

Approximations of the form (46) and (47) have been consider in [50] and in [3, 31], respectively.

On the other hand, conventional numerical integrators can be applied to compute ξ. Indeed, since

$$\xi\,(t_n, \mathbf{y}_n; t - t_n) = e^{\mathbf{A}_n t} \eta(t), \tag{48}$$

where

$$\eta(t) = \sum_{i=1}^{m} \int_{t_n}^{t} e^{-\mathbf{A}_n u} \mathbf{g}_i(u) d\mathbf{w}^i(u)$$

is the solution of the SDE

$$d\eta(t) = \sum_{i=1}^{m} e^{-\mathbf{A}_n u} \mathbf{g}_i(u) d\mathbf{w}^i(u)$$
$$\eta(t_n) = \mathbf{0}$$

for $t \in [t_n, t_{n+1}]$, the application of the Ito formula to (48) yields

$$d\rho(t) = \mathbf{A}_n \rho(t) dt + \sum_{i=1}^{m} \mathbf{g}_i(u) d\mathbf{w}^i(u)$$
$$\rho(t_n) = \mathbf{0}$$

for $t \in [t_n, t_{n+1}]$. Since $\rho(t) \equiv \xi(t_n, \mathbf{y}_n; t - t_n)$, any approximation to ρ given by a numerical integrator is also an approximation to ξ. When the Euler or Milstein scheme is used for that propose the approximation

$$\widetilde{\xi}(t_n, \widetilde{\mathbf{y}}_n; h_n) = \sum_{i=1}^{m} \mathbf{g}_i(t_n) \left(\mathbf{w}^i(t_{n+1}) - \mathbf{w}^i(t_n) \right) \tag{49}$$

is obtained. A different approximation is provided, for instance, by the order-1.5 strong Taylor scheme defined in [32]. This yields to the approximation

$$\widetilde{\xi}(t_n, \widetilde{\mathbf{y}}_n; h_n) = \sum_{i=1}^{m} \left(\mathbf{g}_i(t_n) \Delta \mathbf{w}_n^i + \widetilde{\mathbf{A}}_n \mathbf{g}_i(t_n) \Delta \mathbf{z}_n^i + \frac{d\mathbf{g}_i(t_n)}{dt} (\Delta \mathbf{w}_n^i h_n - \Delta \mathbf{z}_n^i) \right), \tag{50}$$

where $\Delta \mathbf{w}_n^i = \mathbf{w}^i(t_{n+1}) - \mathbf{w}^i(t_n)$ and $\Delta \mathbf{z}_n^i$ represents the double Ito integral $\Delta \mathbf{z}_n^i = \int_{t_n}^{t_{n+1}} \int_{t_n}^{s_2} d\mathbf{w}^i(s_1) ds_2$. The random variable $\Delta \mathbf{w}_n^i$ is normally distributed with mean $E(\Delta \mathbf{w}_n^i) = 0$ and variance $E\left((\Delta \mathbf{w}_n^i)^2\right) = h_n$, whereas $\Delta \mathbf{z}_n^i$ is a Gaussian random variable with mean $E(\Delta \mathbf{z}_n^i) = 0$, variance $E\left((\Delta \mathbf{z}_n^i)^2\right) = \frac{1}{3} h_n^3$ and covariance $E(\Delta \mathbf{w}_n^i \Delta \mathbf{z}_n^i) = \frac{1}{2} h_n^2$.

Approximations (47) with $\alpha = 0$, and (45) satisfy the inequality

$$E\left(\left| \xi(t_n, \widetilde{\mathbf{y}}_n; h_n) - \widetilde{\xi}(t_n, \widetilde{\mathbf{y}}_n; h_n) \right|^2 \middle| \mathcal{F}_{t_n} \right) \leq C h_n^{2p+1}$$

with $p = 1$ due to theorems of Section 4.8 in [12] and Theorem 2 in [54], respectively. Whereas, due to the mentioned theorems, approximations (47) with $\alpha = 1$, and (46) with $\alpha = 0.5$ satisfy the same inequality but with $p = 1.5$. Clearly, among them, approximations (49) and (50) are computationally less demanding than that the others.

Alternatively, conventional numerical integrators can be also directly applied to approximate the piecewise linear equation (4)-(5) for each interval $[t_n, t_{n+1}]$ for obtaining an approximation to the LL discretization $\mathbf{y}_{n+1}$ at each point $t_{n+1} \in (t)_h$. For instance, the approximations provided by the numerical schemes proposed in [35], which are specifically designed of linear SDEs.

Clearly, the error of an LL scheme depends of both, the discretization error consider in Section 2.1 and the error derived from the numerical implementation of the LL discretization. This is stated in the following theorems [30].

Theorem 27 *Let* $\mathbf{x}$ *be the solution of the SDE (1)-(2),*

$$\mathbf{y}_{n+1} = \mathbf{y}_n + \phi_\gamma(t_n, \mathbf{y}_n; h_n) + \xi(t_n, \mathbf{y}_n; h_n)$$

the Local Linear discretization defined in (10), and

$$\widetilde{\mathbf{y}}_{n+1} = \widetilde{\mathbf{y}}_n + \widetilde{\phi}_\gamma(t_n, \widetilde{\mathbf{y}}_n; h_n) + \widetilde{\xi}(t_n, \widetilde{\mathbf{y}}_n; h_n)$$

a numerical implementation of $\mathbf{y}_{n+1}$*, where* $\widetilde{\phi}_\gamma$ *and* $\widetilde{\xi}$ *denote numerical algorithms to compute* ϕ_γ *and* ξ*. Suppose that* $\widetilde{\phi}_\gamma$ *and* $\widetilde{\xi}$ *fulfill the local conditions*

$$\left| E\left(\phi_\gamma(t_n, \widetilde{\mathbf{y}}_n; h_n) - \widetilde{\phi}_\gamma(t_n, \widetilde{\mathbf{y}}_n; h_n) \middle| \mathcal{F}_{t_n} \right) \right| \leq K_1(1 + |\widetilde{\mathbf{y}}_n|^2)^{1/2} \ h_n^{k_1} \tag{51}$$

$$(E\left(\left| \phi_\gamma(t_n, \widetilde{\mathbf{y}}_n; h_n) - \widetilde{\phi}_\gamma(t_n, \widetilde{\mathbf{y}}_n; h_n) \right|^2 \middle| \mathcal{F}_{t_n} \right))^{1/2} \leq K_1(1 + |\widetilde{\mathbf{y}}_n|^2)^{1/2} \ h_n^{k_2+1/2} \tag{52}$$

and

$$\left| E\left(\xi(t_n, \widetilde{\mathbf{y}}_n; h_n) - \widetilde{\xi}(t_n, \widetilde{\mathbf{y}}_n; h_n) \middle| \mathcal{F}_{t_n} \right) \right| \leq K_2(1 + |\widetilde{\mathbf{y}}_n|^2)^{1/2} \ h_n^{p_1}$$

$$(E\left(\left| \xi(t_n, \widetilde{\mathbf{y}}_n; h_n) - \widetilde{\xi}(t_n, \widetilde{\mathbf{y}}_n; h_n) \right|^2 \middle| \mathcal{F}_{t_n} \right))^{1/2} \leq K_2(1 + |\widetilde{\mathbf{y}}_n|^2)^{1/2} \ h_n^{p_2+1/2}$$

for some positive numbers $k_2, p_2 \geq 1/2$*,* $k_1 \geq k_2 + 1$*,* $p_1 \geq p_2 + 1$ *and all* $t_n \in (t)_h$*, where* K_1 *and* K_2 *are positive constants. Then there exist positive constants* M_1 *and* M_2 *such that*

$$E\left(|\widetilde{\mathbf{y}}_n|^2 \middle| \mathcal{F}_{t_0} \right) \leq M_1(1 + E\left(|\widetilde{\mathbf{y}}_0|^2 \right))$$

and

$$E\left(|\mathbf{x}(t_n) - \widetilde{\mathbf{y}}_n|^2 \middle| \mathcal{F}_{t_0} \right) \leq M_2 h^{\min\{2\gamma, 2k_2, 2p_2\}}$$

for all $t_n \in (t)_h$*.*

Proof. As in [30]. ■

Similarly, the following result can be derived.

Theorem 28 *Let* $\mathbf{x}$ *be the solution of the SDE (1)-(2), and let* $\widetilde{\mathbf{y}}_{n+1}$ *be the solution provided at* t_{n+1} *by an one-step numerical integrator when applies to piecewise linear SDE*

$$d\mathbf{z}(t) = (\mathbf{B}_n \mathbf{z}(t) + \mathbf{b}_n^\gamma(t))dt + \sum_{i=1}^{m} \mathbf{g}_i(t) d\mathbf{w}^i(t), \quad t \in (t_n, t_{n+1}],$$

$$\mathbf{z}(t_n) = \widetilde{\mathbf{y}}_n \ ,$$

where $\mathbf{B}_n = \mathbf{f}_{\mathbf{x}}(t_n, \widetilde{\mathbf{y}}_n)$ *is a constant matrix,*

$$\mathbf{b}_n^{\gamma}(s) = \begin{cases} \mathbf{f}(t_n, \widetilde{\mathbf{y}}_n) - \mathbf{f}_{\mathbf{x}}(t_n, \widetilde{\mathbf{y}}_n)\widetilde{\mathbf{y}}_n + \mathbf{f}_t(t_n, \widetilde{\mathbf{y}}_n)(s - t_n) & \text{for } \gamma = 1 \\ \mathbf{b}_n^1(s) + \frac{1}{2}\sum\limits_{k,l=1}^{d}\sum\limits_{j=1}^{m} \mathbf{g}_j^k(t_n)\mathbf{g}_j^l(t_n)\frac{\partial^2 \mathbf{f}(t_n,\widetilde{\mathbf{y}}_n)}{\partial \mathbf{x}^k \partial \mathbf{x}^l}(s - t_n) & \text{for } \gamma = 1.5 \end{cases}.$$

Suppose that $\widetilde{\mathbf{y}}_{n+1}$ *fulfills the local conditions*

$$\left| E\left(\mathbf{z}(t_{n+1}) - \widetilde{\mathbf{y}}_{n+1} \middle| \mathcal{F}_{t_n} \right) \right| \leq K(1 + |\mathbf{z}(t_n)|^2)^{1/2} \; h_n^{p_1}$$

$$\left(E\left(|\mathbf{z}(t_{n+1}) - \widetilde{\mathbf{y}}_{n+1}|^2 \middle| \mathcal{F}_{t_n} \right) \right)^{1/2} \leq K(1 + |\mathbf{z}(t_n)|^2)^{1/2} \; h_n^{p_2+1/2}$$

for some positive numbers $p_2 \geq 1/2$, $p_1 \geq p_2 + 1$ *and all* $t_n \in (t)_h$, *where* K *is a positive constant. Then there exist positive constants* M_1 *and* M_2 *such that*

$$E\left(|\widetilde{\mathbf{y}}_n|^2 \middle| \mathcal{F}_{t_0} \right) \leq M_1(1 + |\widetilde{\mathbf{y}}_0|^2)$$

and

$$E\left(|\mathbf{x}(t_n) - \widetilde{\mathbf{y}}_n|^2 \middle| \mathcal{F}_{t_0} \right) \leq M_2 h^{\min\{2\gamma, 2p_2\}}$$

for all $t_n \in (t)_h$.

In order to illustrate the application of previous theorems, let us consider the LL schemes obtained from the Padé approximation for the computation of ϕ_γ and by numerical schemes for the integration of ξ.

We recall from [26] that ϕ_γ can be rewritten in the general form

$$\phi_\gamma(t_n, \widetilde{\mathbf{y}}_n; h_n) = \mathbf{L}e^{\widetilde{\mathbf{D}}_n h_n}\mathbf{r},$$

where the matrices $\widetilde{\mathbf{D}}_n$, $\mathbf{L}$ and $\mathbf{r}$ have different form depending of the type of equation (1) and value of γ. That is,

$$\widetilde{\mathbf{D}}_n = \begin{bmatrix} \mathbf{f}_{\mathbf{x}}(t_n, \widetilde{\mathbf{y}}_n) & \mathbf{f}_t(t_n, \widetilde{\mathbf{y}}_n) + \frac{1}{2}\sum\limits_{k,l=1}^{d}\sum\limits_{j=1}^{m} \mathbf{g}_j^k(t_n)\mathbf{g}_j^l(t_n)\frac{\partial^2 \mathbf{f}(t_n,\widetilde{\mathbf{y}}_n)}{\partial \mathbf{x}^k \partial \mathbf{x}^l} & \mathbf{f}(t_n, \widetilde{\mathbf{y}}_n) \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \in \mathbb{R}^{(d+2)\times(d+2)},$$

$\mathbf{L} = \begin{bmatrix} \mathbf{I}_d & \mathbf{0}_{d\times 2} \end{bmatrix}$ and $\mathbf{r}^\intercal = \begin{bmatrix} \mathbf{0}_{1\times(d+1)} & 1 \end{bmatrix}$ for nonautonomous SDEs and $\gamma = 1.5$;

$$\widetilde{\mathbf{D}}_n = \begin{bmatrix} \mathbf{f}_{\mathbf{x}}(t_n, \widetilde{\mathbf{y}}_n) & \mathbf{f}_t(t_n, \widetilde{\mathbf{y}}_n) & \mathbf{f}(t_n, \widetilde{\mathbf{y}}_n) \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \in \mathbb{R}^{(d+2)\times(d+2)},$$

$\mathbf{L} = \begin{bmatrix} \mathbf{I}_d & \mathbf{0}_{d\times 2} \end{bmatrix}$ and $\mathbf{r}^\intercal = \begin{bmatrix} \mathbf{0}_{1\times(d+1)} & 1 \end{bmatrix}$ for nonautonomous SDEs and $\gamma = 1$; and

$$\widetilde{\mathbf{D}}_n = \begin{bmatrix} \mathbf{f}_{\mathbf{x}}(\widetilde{\mathbf{y}}_n) & \mathbf{f}(\widetilde{\mathbf{y}}_n) \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{(d+1)\times(d+1)},$$

$\mathbf{L} = \left[\begin{array}{cc} \mathbf{I}_d & \mathbf{0}_{d\times 1} \end{array} \right]$ and $\mathbf{r}^\intercal = \left[\begin{array}{cc} \mathbf{0}_{1\times d} & 1 \end{array} \right]$ for autonomous SDEs and $\gamma = 1$. If the computation of $e^{h_n \widetilde{\mathbf{D}}_n}$ is carried out by means of the rational Padé approximation with the 'scaling and squaring' procedure [39], the following approximation to ϕ_γ is obtained

$$\widetilde{\phi}_\gamma\left(t_n, \widetilde{\mathbf{y}}_n; h_n\right) = \mathbf{L}(\mathbf{P}_{p,q}(2^{-k_n}\widetilde{\mathbf{D}}_n h_n))^{2^{k_n}}\mathbf{r}, \tag{53}$$

where $\mathbf{P}_{p,q}(2^{-k_n}\widetilde{\mathbf{D}}_n h_n)$ denotes the (p,q)-Padé approximation of $e^{2_n^{-k_n}\widetilde{\mathbf{D}}_n h_n}$ and k_n is the smallest integer number such that $\left|2^{-k_n}\widetilde{\mathbf{D}}_n h_n\right| \leq \frac{1}{2}$.

Theorem 29 *Let $\widetilde{\phi}_\gamma$ defined by (53) and*

$$\widetilde{\xi}_\gamma\left(t_n, \widetilde{\mathbf{y}}_n; h_n\right) = \sum_{i=1}^{m} \mathbf{g}_i(t_n)\Delta\mathbf{w}_n^i + \delta_\gamma^{1.5}\sum_{i=1}^{m}\left(\mathbf{f}_{\mathbf{x}}(t_n,\widetilde{\mathbf{y}}_n)\mathbf{g}_i(t_n)\Delta\mathbf{z}_n^i + \frac{d\mathbf{g}_i(t_n)}{dt}(\Delta\mathbf{w}_n^i h_n - \Delta\mathbf{z}_n^i)\right)$$

be the approximations to ϕ_γ and ξ provided by the Padé approximation for exponential matrices and the numerical schemes (49)-(50), respectively. Suppose that conditions of Theorem 7 hold. Then

i) $\widetilde{\phi}_\gamma$ fulfills the conditions (51)-(52) with $k_1 = p+q+1$ and $k_2 = p+q$; and

ii) the error of the LL scheme

$$\widetilde{\mathbf{y}}_{n+1} = \widetilde{\mathbf{y}}_n + \widetilde{\phi}_\gamma\left(t_n, \widetilde{\mathbf{y}}_n; h_n\right) + \widetilde{\xi}_\gamma\left(t_n, \widetilde{\mathbf{y}}_n; h_n\right)$$

is given by

$$E\left(\left|\mathbf{x}(t_n) - \widetilde{\mathbf{y}}_n\right)\right|^2\right) \leq Mh^{2\min\{\gamma, p+q\}}$$

for all $t_n \in (t)_h$, where M is a positive constant.

Proof. As in [30]. ■

We recall from theorems 32 in [27] and 15, and Lemma 16 that LL schemes like those considered in previous theorem are $A-$stable, therefore they preserve the random attractor and ergodicity of the linear SDEs. They also are geometrically ergodic for some class of nonlinear SDEs [19]. However, because the use of Padé approximations to compute ϕ_γ such schemes are not appropriate for large dimensional systems of SDEs. In such a case, LL schemes based on Krylov approximation for ϕ_γ are recommended. Indeed, when the Krylov method [23] is used, the function ϕ_γ can be approximated by

$$\widetilde{\phi}_\gamma\left(t_n, \widetilde{\mathbf{y}}_n; h_n\right) = \mathbf{L}\ \mathbf{k}_{m_n,k_n}^{p,q}(h_n, \widetilde{\mathbf{D}}_n, \mathbf{r}), \tag{54}$$

where $\mathbf{k}_{m_n,k_n}^{p,q}(h_n, \widetilde{\mathbf{D}}_n, \mathbf{r})$ denotes the Krylov-Padé approximation of $e^{h_n\widetilde{\mathbf{D}}_n}\mathbf{r}$. More precisely,

$$\mathbf{k}_{m,k}^{p,q}(\delta, \mathbf{A}, \mathbf{v}) = \beta\mathbf{V}_m\mathbf{F}_{p,q}^k(\delta\mathbf{H}_m)\mathbf{e}_1,$$

where $\beta = |\mathbf{v}|$, $\mathbf{e}_1$ is the $m-$dimensional unitary vector, and the matrices $\mathbf{V}_m = [\mathbf{v}_1, \ldots, \mathbf{v}_m]$ and $\mathbf{H}_m$ are the orthonormal basis of the Krylov space $\mathcal{K}_m = span(\mathbf{v}, \mathbf{A}\mathbf{v}, \ldots, \mathbf{A}^{m-1}\mathbf{v})$ and the upper Hessenberg matrix, respectively, resulting both from the well known Arnoldi algorithm. In addition, $\mathbf{F}_{p,q}^{k}(\delta\mathbf{H}_m) = (\mathbf{P}_{p,q}(2^{-k}\delta\mathbf{H}_m))^{2^k}$ denotes the (p,q)-Padé approximation with ´scaling and squaring´ procedure for the computation of $e^{\delta\mathbf{H}_m}$, and k is the smallest integer number such that $\left|2^{-k}\delta\mathbf{H}_m\right| \leq \frac{1}{2}$.

The second illustrative convergence result is then the following.

Theorem 30 *Let $\widetilde{\phi}_\gamma$ defined as in (54) and*

$$\widetilde{\xi}_\gamma\left(t_n,\widetilde{\mathbf{y}}_n;h_n\right) = \sum_{i=1}^{m}\mathbf{g}_i(t_n)\Delta\mathbf{w}_n^i + \delta_\gamma^{1.5}\sum_{i=1}^{m}\left(\mathbf{f}_{\mathbf{x}}(t_n,\widetilde{\mathbf{y}}_n)\mathbf{g}_i(t_n)\Delta\mathbf{z}_n^i + \frac{d\mathbf{g}_i(t_n)}{dt}(\Delta\mathbf{w}_n^i h_n - \Delta\mathbf{z}_n^i)\right)$$

be the approximations to ϕ_γ and ξ provided by the Krylov-Padé approximation for exponential matrices and the numerical schemes (49)-(50), respectively. Suppose that conditions of Theorem 7 hold, and that $m_n \geq 2h_n\left|\widetilde{\mathbf{D}}_n\right|_2$ for all n, where $|\cdot|_2$ denotes the Euclidean norm. Then

i) $\widetilde{\phi}_\gamma$ fulfills the conditions (51)-(52) with $k_1 = \min\{m, p+q+1\}$ and $k_2 = \min\{m-1, p+q\}$, where $m = \min\{m_n\}$; and

ii) the error of the LL scheme

$$\widetilde{\mathbf{y}}_{n+1} = \widetilde{\mathbf{y}}_n + \widetilde{\phi}_\gamma\left(t_n,\widetilde{\mathbf{y}}_n;h_n\right) + \widetilde{\xi}_\gamma\left(t_n,\widetilde{\mathbf{y}}_n;h_n\right)$$

is given by

$$E\left(\left|\mathbf{x}(t_n) - \widetilde{\mathbf{y}}_n\right|_2^2\right) \leq Mh^{2\min\{\gamma,m-1,p+q\}}$$

for all $t_n \in (t)_h$, where M is a positive constant.

Proof. As in [30]. ■

6.2 HOLL schemes

In general, the numerical implementation of HOLL methods requires the calculation of ϕ_γ and the simulation of multiple stochastic integrals involved in the numerical integration of the auxiliary equation (20)-(21). Obviously, the best way to deal with these tasks strongly depends on the specific HOLL scheme under consideration. However, there are a number of common implementation issues that will be discussed here.

When implementing the HOLL discretization

$$\mathbf{y}_{n+1} = \mathbf{y}_n + \phi_\gamma(t_n,\mathbf{y}_n;h_n) + \mathbf{\Lambda}^{\mathbf{y}_n}\left(t_n,\mathbf{0};h_n\right),$$

that is, when a HOLL scheme is constructed, the required evaluations of the map $\phi_\gamma(t_n,\mathbf{y}_n;\cdot)$ at a number of points $t - t_n \leq h_n$ can be computed by different algorithms. In [27] a number of them were presented, which yield the two basic recursions:

$$\widetilde{\mathbf{y}}_{n+1} = \widetilde{\mathbf{y}}_n + \widetilde{\phi}_\gamma(t_n,\widetilde{\mathbf{y}}_n;h_n) + \widetilde{\mathbf{\Lambda}}^{\widetilde{\mathbf{y}}_n}\left(t_n,\mathbf{0};h_n\right)$$

and

$$\widetilde{\mathbf{y}}_{n+1} = \widetilde{\mathbf{z}}(t_n + h_n; t_n, \widetilde{\mathbf{y}}_n) + \widetilde{\mathbf{\Lambda}}^{\widetilde{\mathbf{y}}_n}(t_n, \mathbf{0}; h_n),$$

where $\widetilde{\phi}_\gamma$ is a numerical implementation of ϕ_γ, $\widetilde{\mathbf{z}}$ is a numerical solution of the linear ODE

$$d\mathbf{z}(t) = (\mathbf{B}_n \mathbf{z}(t) + \mathbf{b}_n^\gamma(t))\, dt, \quad t \in (t_n, t_{n+1}], \tag{55}$$

$$\mathbf{z}(t_n) = \widetilde{\mathbf{y}}_n, \tag{56}$$

and $\widetilde{\mathbf{\Lambda}}^{\widetilde{\mathbf{y}}_n}$ is the map of the numerical scheme applied to the SDE

$$d\mathbf{v}(t) = \widetilde{\mathbf{q}}_\gamma(t_n, \widetilde{\mathbf{y}}_n; t, \mathbf{v}(t)) + \sum_{i=1}^{m} \mathbf{g}_i(t) d\mathbf{w}^i(t), \quad t \in (t_n, t_{n+1}], \tag{57}$$

$$\mathbf{v}(t_n) = \mathbf{0}, \tag{58}$$

with drift coefficient

$$\widetilde{\mathbf{q}}_\gamma(t_n, \widetilde{\mathbf{y}}_n; s, \xi) = \mathbf{f}(s, \widetilde{\mathbf{y}}_n + \widetilde{\phi}_\gamma(t_n, \widetilde{\mathbf{y}}_n; s - t_n) + \xi) - \mathbf{f}_{\mathbf{x}}(t_n, \widetilde{\mathbf{y}}_n)(\widetilde{\mathbf{y}}_n + \widetilde{\phi}_\gamma(t_n, \widetilde{\mathbf{y}}_n; s - t_n)) - \mathbf{b}_n^\gamma(s),$$

for the first kind of recursion, or

$$\widetilde{\mathbf{q}}_\gamma(t_n, \widetilde{\mathbf{y}}_n; s, \xi) = \mathbf{f}(s, \widetilde{\mathbf{z}}(s; t_n, \widetilde{\mathbf{y}}_n) + \xi) - \mathbf{f}_{\mathbf{x}}(t_n, \widetilde{\mathbf{y}}_n)\widetilde{\mathbf{z}}(s; t_n, \widetilde{\mathbf{y}}_n) - \mathbf{b}_n^\gamma(s)$$

for the second one. In equation (55), $\mathbf{B}_n = \mathbf{f}_{\mathbf{x}}(t_n, \widetilde{\mathbf{y}}_n)$ is a $d \times d$ constant matrix and

$$\mathbf{b}_n^\gamma(s) = \begin{cases} \mathbf{f}(t_n, \widetilde{\mathbf{y}}_n) - \mathbf{f}_{\mathbf{x}}(t_n, \widetilde{\mathbf{y}}_n)\widetilde{\mathbf{y}}_n + \mathbf{f}_t(t_n, \widetilde{\mathbf{y}}_n)(s - t_n) & \text{for } \gamma = 1 \\ \mathbf{b}_n^1(s) + \frac{1}{2} \sum\limits_{k,l=1}^{d} \sum\limits_{j=1}^{m} \mathbf{g}_j^k(t_n) \mathbf{g}_j^l(t_n) \frac{\partial^2 \mathbf{f}(t_n, \widetilde{\mathbf{y}}_n)}{\partial \mathbf{x}^k \partial \mathbf{x}^l}(s - t_n) & \text{for } \gamma = 1.5 \end{cases}$$

is a d-dimensional linear vector function.

Clearly, a HOLL scheme will preserve the order κ of the underlaying HOLL discretization only if $\widetilde{\phi}_\gamma$ and $\widetilde{\mathbf{\Lambda}}$ are suitable approximations to ϕ_γ and $\mathbf{\Lambda}$. This requirement is considered in what follows [30].

Theorem 31 *Let* $\mathbf{x}$ *be the solution of the SDE (1)-(2). With* $t_n, t_{n+1} \in (t)_h$*, let* $\widetilde{\mathbf{z}}_{n+1} = \widetilde{\mathbf{z}}_n + \mathbf{\Lambda}_1(t_n, \widetilde{\mathbf{z}}_n; h_n)$ *and* $\widetilde{\mathbf{v}}_{n+1} = \widetilde{\mathbf{v}}_n + \mathbf{\Lambda}_2^{\widetilde{\mathbf{z}}_n}(t_n, \widetilde{\mathbf{v}}_n; h_n)$ *be one-step explicit integrators of the ODE (55)-(56) and SDE (57)-(58), respectively. Suppose that these integrators are of order* r *and* p*, respectively, where the convergence for the second one is assumed to be in the strong mean-square sense. Then, the numerical scheme*

$$\widetilde{\mathbf{y}}_{n+1} = \widetilde{\mathbf{y}}_n + \mathbf{\Lambda}_1(t_n, \widetilde{\mathbf{y}}_n; h_n) + \mathbf{\Lambda}_2^{\widetilde{\mathbf{y}}_n}(t_n, \mathbf{0}; h_n)$$

satisfies

$$E\left(|\widetilde{\mathbf{y}}_n|^2 \,\middle|\, \mathcal{F}_{t_0}\right) \leq M_1(1 + |\widetilde{\mathbf{y}}_0|^2)$$

and

$$E\left(|\mathbf{x}(t_n) - \widetilde{\mathbf{y}}_n|^2 \,\middle|\, \mathcal{F}_{t_0}\right) \leq M_2 h^{2\min\{r,p\}}$$

for all $t_n \in (t)_h$*, where* M_1 *and* M_2 *are positive constants.*

Proof. As in [30]. ■

In order to show the application of previous theorem let us consider the numerical implementation of the order-2 strong LLT discretization (24) by means of the Padé approximation for ϕ_γ, and by a truncated Fourier series for the multiple Stratonovish integrals J_α.

Theorem 32 *Let*

$$\widetilde{\phi}_\gamma\left(t_n, \widetilde{\mathbf{y}}_{t_n}; h_n\right) = \mathbf{L}(\mathbf{P}_{p,q}(2^{-k_n}\widetilde{\mathbf{D}}_n h_n))^{2^{k_n}}\mathbf{r}$$

be the approximation to ϕ_γ provided by the Padé approximation (53), and let

$$\begin{aligned}\widetilde{\mathbf{\Lambda}}_\gamma^{\widetilde{\mathbf{y}}_n}\left(t_n, \mathbf{0}; h_n\right) &= \sum_{j=1}^{m}\mathbf{g}_j\left(t_n\right)\Delta\mathbf{w}_n^j + \sum_{j=1}^{m}\mathbf{f}_{\mathbf{x}}(t_n, \widetilde{\mathbf{y}}_n)\mathbf{g}_j\left(t_n\right)\widetilde{J}_{(j,0)}^2 \\ &+ \sum_{j=1}^{m}\frac{d\mathbf{g}_j\left(t_n\right)}{dt}\widetilde{J}_{(0,j)}^2 + \sum_{j_1,j_2=1}^{m}\left(\mathbf{I}_{d\times d}\otimes\mathbf{g}_{j_2}^{\intercal}\left(t_n\right)\right)\mathbf{f}_{\mathbf{xx}}(t_n, \widetilde{\mathbf{y}}_n)\mathbf{g}_{j_1}\left(t_n\right)\widetilde{J}_{(j_1,j_2,0)}^2 \\ &+ \delta_\gamma^{1.5}\frac{h_n^2}{4}\sum_{j=1}^{m}\left(\mathbf{I}_{d\times d}\otimes\mathbf{g}_j^{\intercal}\left(t_n\right)\right)\mathbf{f}_{\mathbf{xx}}(t_n, \widetilde{\mathbf{y}}_n)\mathbf{g}_j\left(t_n\right)\end{aligned}$$

be the map derived from the application of the order-2 strong Ito-Taylor scheme to the auxiliary equation (57)-(58). Here, $\widetilde{J}_\alpha^2$ is an order-2 approximation to the stochastic multiple Stratonovish integral J_α. Then, under conditions of Theorem 14, the error of the LLT scheme

$$\widetilde{\mathbf{y}}_{n+1} = \widetilde{\mathbf{y}}_n + \widetilde{\phi}_\gamma\left(t_n, \widetilde{\mathbf{y}}_n; h_n\right) + \widetilde{\mathbf{\Lambda}}_\gamma^{\widetilde{\mathbf{y}}_n}\left(t_n, \mathbf{0}; h_n\right)$$

is given by

$$E\left(\left|\mathbf{x}(t_n) - \widetilde{\mathbf{y}}_n\right|^2\right) \leq Mh^{2\min\{2,p+q\}}$$

for all $t_n \in (t)_h$, where M is a positive constant.

Proof. As in [30]. ■

For high dimensional SDEs we have the following result.

Theorem 33 *Let*

$$\widetilde{\phi}_\gamma\left(t_n, \widetilde{\mathbf{y}}_{t_n}; h_n\right) = \mathbf{L}\ \mathbf{k}_{m_n,k_n}^{p,q}(h_n, \widetilde{\mathbf{D}}_n, \mathbf{r})$$

be the approximation to ϕ_γ provided by the Krylov-Padé approximation (54), and let

$$\begin{aligned}\widetilde{\mathbf{\Lambda}}_\gamma^{\widetilde{\mathbf{y}}_n}\left(t_n, \mathbf{0}; h_n\right) &= \sum_{j=1}^{m}\mathbf{g}_j\left(t_n\right)\Delta\mathbf{w}_n^j + \sum_{j=1}^{m}\mathbf{f}_{\mathbf{x}}(t_n, \widetilde{\mathbf{y}}_n)\mathbf{g}_j\left(t_n\right)\widetilde{J}_{(j,0)}^2 \\ &+ \sum_{j=1}^{m}\frac{d\mathbf{g}_j\left(t_n\right)}{dt}\widetilde{J}_{(0,j)}^2 + \sum_{j_1,j_2=1}^{m}\left(\mathbf{I}_{d\times d}\otimes\mathbf{g}_{j_2}^{\intercal}\left(t_n\right)\right)\mathbf{f}_{\mathbf{xx}}(t_n, \widetilde{\mathbf{y}}_n)\mathbf{g}_{j_1}\left(t_n\right)\widetilde{J}_{(j_1,j_2,0)}^2 \\ &+ \delta_\gamma^{1.5}\frac{h_n^2}{4}\sum_{j=1}^{m}\left(\mathbf{I}_{d\times d}\otimes\mathbf{g}_j^{\intercal}\left(t_n\right)\right)\mathbf{f}_{\mathbf{xx}}(t_n, \widetilde{\mathbf{y}}_n)\mathbf{g}_j\left(t_n\right)\end{aligned}$$

be the map derived from the application of the order-2 strong Ito-Taylor scheme to the auxiliary equation (57)-(58). Here, $\widetilde{J}_\alpha^2$ is an order-2 approximation to the multiple Stratonovish integral

J_α. Suppose that conditions of Theorem 14 hold, and that $m_n \geq 2h_n \left|\widetilde{\mathbf{D}}_n\right|_2$ for all n, where $|\cdot|_2$ denotes the Euclidean norm. Then the error of the LLT scheme

$$\widetilde{\mathbf{y}}_{n+1} = \widetilde{\mathbf{y}}_n + \widetilde{\phi}_\gamma\left(t_n, \widetilde{\mathbf{y}}_n; h_n\right) + \widetilde{\mathbf{\Lambda}}_\gamma^{\widetilde{\mathbf{y}}_n}\left(t_n, \mathbf{0}; h_n\right)$$

is given by

$$E\left(|\mathbf{x}(t_n) - \widetilde{\mathbf{y}}_n|_2^2\right) \leq Mh^{2\min\{2,m-1,p+q\}}$$

for all $t_n \in (t)_h$, where $m = \min\{m_n\}$ and M is a positive constant.

Proof. As in [30]. ■

Similar convergence orders can be obtained for the numerical implementation of the LLRK discretization (27), if the map ϕ_γ and the integrals J_α are approximated as in the previous theorems by the Padé or Krylov-Padé approximations and by an order-2 approximation, respectively. That is, for the LLRK scheme

$$\widetilde{\mathbf{y}}_{n+1} = \widetilde{\mathbf{y}}_n + \widetilde{\phi}_\gamma\left(t_n, \widetilde{\mathbf{y}}_n; h_n\right) + \frac{h_n}{2}\left\{\widetilde{\mathbf{k}}_1 + \widetilde{\mathbf{k}}_2\right\} + \mathbf{g}\left(t_n\right)\Delta\mathbf{w}_n + \frac{1}{h_n}\left\{\mathbf{g}\left(t_{n+1}\right) - \mathbf{g}\left(t_n\right)\right\}\widetilde{J}^2_{(0,1)},$$

where

$$\widetilde{\mathbf{k}}_1 = \mathbf{f}(t_n + \frac{h_n}{2}, \widetilde{\mathbf{y}}_n + \widetilde{\phi}_\gamma(t_n, \widetilde{\mathbf{y}}_n; \frac{h_n}{2}) + \widetilde{\alpha}_+) - \mathbf{f}_\mathbf{x}(t_n, \widetilde{\mathbf{y}}_n)\widetilde{\phi}_\gamma(t_n, \widetilde{\mathbf{y}}_n; \frac{h_n}{2}) - \mathbf{f}\left(t_n, \widetilde{\mathbf{y}}_n\right) - \mathbf{f}_t\left(t_n, \widetilde{\mathbf{y}}_n\right)\frac{h_n}{2},$$
$$\widetilde{\mathbf{k}}_2 = \mathbf{f}(t_n + \frac{h_n}{2}, \widetilde{\mathbf{y}}_n + \widetilde{\phi}_\gamma(t_n, \widetilde{\mathbf{y}}_n; \frac{h_n}{2}) + \widetilde{\alpha}_-) - \mathbf{f}_\mathbf{x}(t_n, \widetilde{\mathbf{y}}_n)\widetilde{\phi}_\gamma(t_n, \widetilde{\mathbf{y}}_n; \frac{h_n}{2}) - \mathbf{f}\left(t_n, \widetilde{\mathbf{y}}_n\right) - \mathbf{f}_t\left(t_n, \widetilde{\mathbf{y}}_n\right)\frac{h_n}{2}$$

and

$$\widetilde{\alpha}_{\substack{+\\-}} = \frac{1}{h_n}\mathbf{g}\left(t_n\right)\left\{\widetilde{J}^2_{(1,0)} \substack{+\\-} \sqrt{2\widetilde{J}^2_{(1,1,0)}h_n - (\widetilde{J}^2_{(1,0)})^2}\right\}.$$

We point out that the value of γ does affect neither the order of convergence nor the A-stability property of the HOLL discretizations. Typically, with the value $\gamma = 1$, the resulting HOLL schemes have less computational cost than those with $\gamma = 1.5$ (see for instance the schemes of the two previous theorems). Therefore, HOLL schemes derived from discretizations with $\gamma = 1$ are more common. Here, both cases are included by completeness.

In addition, it is worth of noting some other particularities that usually contribute to decrease the computational burden of some HOLL schemes. For instance, the initial value $\mathbf{r}\left(t_n\right) = \mathbf{0}$ of the auxiliary equation (20) typically leads to simplifications into the algebraic expressions of the schemes. Similarly, some properties of the exponential matrix usually yield to a reduction of the number of exponential matrix evaluations per integration step. With regards to these aspects, the numerical implementation of the LLRK discretization (26) is a typical example: $\mathbf{r}\left(t_n\right) = \mathbf{0}$ implies that $\mathbf{k}_0 = \mathbf{0}$, while the identity $e^{\frac{j}{2}(t-t_n)\mathbf{D}_n} = (e^{\frac{1}{2}(t-t_n)\mathbf{D}_n})^j$ allows for the computation of just one matrix exponential (namely, $e^{\frac{1}{2}(t-t_n)\mathbf{D}_n}$) and its square to obtain $\mathbf{k}_i$, $i = 1, \ldots, 4$, at each integration step.

6.3 Schemes for equations with jumps

From the results of the two previous subsections, it is easy to realize that numerical implementations of Local Linear discretizations for SDE with jumps imply the use of LL schemes for SDEs (with no jumps). Indeed, from the Definition 22, an Local Linearization scheme $\widetilde{\mathbf{y}}_n$ for the integration of the SDE with jumps (37)-(38) can be defined as

$$\widetilde{\mathbf{y}}_n = \widetilde{\mathbf{y}}_{n-} + \sum_{i=1}^{p} \mathbf{h}_i(t_n, \widetilde{\mathbf{y}}_{n-})\Delta \mathbf{n}_n^i, \tag{59}$$

where $\widetilde{\mathbf{y}}_{n-}$ denotes the value $\widetilde{\mathbf{z}}_n$ of an order-κ LL scheme for the SDE (40) on $[t_{n-1}, t_n]$ with initial condition $\mathbf{z}(t_{n-1}) = \widetilde{\mathbf{y}}_{n-1}$, and $\Delta \mathbf{n}_n^i$ is the increment of the process $\mathbf{n}_n^i$ at the time instant t_n.

In order to study the convergence of the LL scheme (59) the following general result is useful, which is a straightforward consequence of Lemma 24.

Corollary 34 *Let $\mathbf{z}_n = \mathbf{z}_{n-1} + F(t_{n-1}, \mathbf{z}_{n-1}; h_{n-1})$ be a numerical integrator for the SDE (40) such that*

$$E\left(|\mathbf{z}_n|^2 \,\middle|\, \mathcal{F}_{t_0}\right) \leq K(1 + |\mathbf{z}_0|^2)$$

$$E\left(|\mathbf{z}\,(t_n) - \mathbf{z}_n|^2 \,\middle|\, \mathcal{F}_{t_0}\right) \leq K h^{2\kappa}$$

for all $t_n \in (t)_h$, where K is a positive constant. Let $\{\mathbf{y}_n\}$ be the sequence defined by

$$\mathbf{y}_n = \mathbf{y}_{n-} + \sum_{i=1}^{p} \mathbf{h}_i(t_n, \mathbf{y}_{n-})\Delta \mathbf{n}_n^i,$$

where $\mathbf{y}_{n-}$ denotes the value $\mathbf{z}_n$ of the numerical integrator when applies to the equation (40) on the interval $[t_{n-1}, t_n]$ with initial condition $\mathbf{z}(t_{n-1}) = \mathbf{y}_{n-1}$. If the functions $\mathbf{h}_i$ defined in (37) satisfy the condition

$$|\mathbf{h}_i(t, \mathbf{u}) - \mathbf{h}_i(t, \mathbf{v})| \leq K\,|\mathbf{u} - \mathbf{v}| \tag{60}$$

and

$$|\mathbf{h}_i(t, \mathbf{u})| \leq K(1 + |\mathbf{u}|) \tag{61}$$

for $t \in [t_0, T]$ and $\mathbf{u}, \mathbf{v} \in \mathbb{R}^d$, then there exists a positive constant C such that

$$E\left(\max_n |\mathbf{y}_n|^2 \,\middle|\, \mathcal{F}_{t_0}\right) \leq C(1 + |\mathbf{y}_0|^2)$$

and

$$E\left(\max_n |\mathbf{x}\,(t_n) - \mathbf{y}_n|^2 \,\middle|\, \mathcal{F}_{t_0}\right) \leq C h^{2\kappa},$$

where $\mathbf{x}$ is the solution of the SDE with jumps (37)-(38).

The main convergence result for the LL schemes is the following.

Theorem 35 *Suppose that the functions* $\mathbf{h}_i$ *defined in (37) satisfy the inequalities (60)-(61). Further, suppose that the functions* $\overline{\mathbf{f}}$ *and* $\mathbf{g}_i$ *defined in (40) satisfies the conditions under which a LL scheme for SDEs converges to the solutions of (40) with order* κ*. Then the corresponding LL scheme (59) to the solution* $\mathbf{x}$ *of the SDE with jump (37)-(38) satisfies*

$$E\left(\max_n |\mathbf{x}(t_n) - \widetilde{\mathbf{y}}_n|^2\right) \leq Ch^{2\kappa},$$

where C *is a positive constant.*

Proof. As in [30]. ■

7 Numerical simulations

In [41, 3, 50, 31, 13] a number of numerical simulations were carried out in order to illustrate the performance of the LL schemes and compare them to other numerical integrators. With special emphasis, the simulations illustrate the high stability of the LL scheme. For nonlinear equations in general, these LL schemes show a stability similar to that of implicit schemes, but with much lower computational cost (comparable to conventional explicit schemes).

Acknowledgement. This work was concluded within the framework of the Associateship Scheme of the Abdus Salam International Centre for Theoretical Physics, Trieste, Italy.

References

[1] Arnold L., Random Dynamical Systems, Springer-Verlag, Heidelberg, 1998.

[2] Artemiev S.S. and Averina T.A., Numerical analysis of systems of ordinary and stochastic differential equations, VSP, 1997.

[3] Biscay R., Jimenez J.C., Riera J. and Valdes P., Local linearization method for the numerical solution of stochastic differential equations, Annals Inst. Statis. Math., 48 (1996) 631-644.

[4] Bokor R. H., On the stability for numerical approximations of stochastic ordinary differential equations, Electronic J. of Qualitative Theory of Differential Equations, Proc. 7th Coll. QTDE (2004) No. 15, 1-9 (http://www.math.u-szeged.hu/ejqtde/).

[5] Brugnano L., Burrage K. and Burrage P.M., Adams-type methods for the numerical solution of stochastic ordinary differential equations, BIT, 40 (2000) 451-470.

[6] Bruti-Liberati N., Numerical Solution of Stochastic Differential Equations with Jumps in Finance, PhD Thesis, School of Finance and Economics, University of Technology, Sydney, 2007.

[7] Burrage K. and Burrage P.M., High strong order explicit Runge-Kutta methods for stochastic ordinary differential equations, Applied Numerical Mathematics, 22 (1996) 81-101.

[8] Burrage K. and Burrage P.M., Order conditions of stochastic Runge–Kutta methods by B-series, SIAM J. Numer. Anal., 38 (2000) 1626–1646.

[9] Burrage K., Burrage P.M. and Tian T., Numerical methods for strong solutions of stochastic differential equations: an overview, Proceedings of the Royal Society of London, Series A, 460 (2004) 373-402.

[10] Carbonell F., Jimenez J.C., Biscay R.J., Weak local linear discretizations for stochastic differential equations: convergence and numerical schemes, J. Comput. Appl. Math., 197 (2006) 578-596.

[11] Chang C.C., Numerical solution of stochastic differential equations with constant diffusion coefficients, Math. Comp., 49 (1987) 523-542.

[12] Davis P.J. and Rabinowitz P. (1984) Methods of Numerical Integration. Academic Press, Inc., 2nd edition.

[13] de la Cruz H., Biscay R.J., Jimenez J.C., Carbonell F. and Ozaki T., High Order Local Linearization methods: an approach for constructing A-stable high order explicit schemes for stochastic differential equations with additive noise, BIT, 50 (2010) 509–539.

[14] Denk G. and Schaffler S., Adams methods for the solution of stochastic differential equations with additive noise, Computing, 59 (1997) 153-161.

[15] Fogelson A.L., A mathematical model and numerical method for studying platelet adhesion and aggregation during blood clotting, J. Comput. Phys., 50 (1984) 111-134.

[16] Geman S. and Hwang C., Diffusions for global optimization, SIAM J. Control Optim., 24 (1986) 1031-1043.

[17] Gitterman M., The noisy oscillator, World Scientific, 2005.

[18] Hairer, E. and Wanner, G. Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems, 3th ed., Springer-Verlag: Berlin, 1996.

[19] Hansen, N.R., Geometric ergodicity of discrete-time approximations to multivariate diffusions, Bernoulli, 9 (2003) 725-743.

[20] Higham D.J., Mao X., and Stuart A.M., Strong convergence of Euler type methods for nonlinear stochastic differential equations, SIAM J. on Numer. Anal., 40 (2002) 1041-1063.

[21] Higham, D.J.; Mao, X.; Yuan, Ch. Almost sure and moment exponential stability in the numerical simulation of stochastic differential equations. SIAM J. Numer.Anal., 45 (2007) 592–609.

[22] Hernandez D.B. and Spigler R., A-stability of Runge-Kutta methods for systems with additive noise, BIT, 32 (1992) 620-633.

[23] Hochbruck, M. and Lubich, C. On Krylov subspace approximations to the matrix exponential operator, SIAM Numer. Anal., 34 (1997) 1911-1925.

[24] Hofmann N., MÄuller-Gronbach T. and Ritter K., Stepsize control for the uniform approximation of systems of stochastic differential equations with additive noise, Ann. Appl. Probab., 10 (2000) 613-633.

[25] Jentzen A., Kloeden P. E. and Neuenkirch A., Pathwise approximation of stochastic differential equations on domains: higher order convergence rates without global Lipschitz coefficients, Numer. Math., 112 (2009) 41-64.

[26] Jimenez J.C., A simple algebraic expression to evaluate the Local Linearization schemes for stochastic differential equations, Appl. Math. Letters, 15 (2002) 775-780.

[27] Jimenez J.C., Local Linearization methods for the numerical integration of ordinary differential equations: An overview. ICTP preprint 035, 2009.

[28] Jimenez J.C. and Biscay R., Approximation of continuous time stochastic processes by the Local Linearization method revisited. Stochast. Anal. & Appl., (2002) 20, 105-121.

[29] Jimenez J.C. and Carbonell F., Local Linear approximations for jump diffusion processes, J. Appl. Prob., 43 (2006) 185-194.

[30] Jimenez J.C. and de la Cruz H., Convergence rate of strong Local Linearization schemes for stochastic differential equations with additive noise. Submitted.

[31] Jimenez J.C., Shoji I. and Ozaki T., Simulation of stochastic differential equations through the Local Linearization method. A comparative study, J. Statist. Physics, 94 (1999) 587-602.

[32] Kloeden, P. E., Keller, H. and Schmalfuss, B., Towards a theory of random numerical dynamics, In: Stochastic Dynamics, H. Crauel und V. M. Gundlach (Eds.) Springer-Verlag (1999) 259-282.

[33] Kloeden P.E. and Platen E., Numerical Solution of Stochastic Differential Equations, Springer-Verlag, Berlin, Second Edition, 1995.

[34] Kloeden P.E., Platen E. and Scurz H. (1994) Numerical Solution of SDE Through Computer Experiments, Springer-Verlag, Berlin.

[35] Lord G., Malham S.J.A. and Wiese A., Efficient strong integrators for linear stochastic systems. SIAM J. Numer. Anal., 46 (2008) 2892-2919.

[36] Mattingly J.C., Stuart A.M., Higham D.J., Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise, Stoch. Proces. Appl., 101 (2002) 185-232.

[37] Milshtein G.N (1988) The Numerical Integration of Stochastic Differentials Equations, Ural University Press.

[38] Milstein G.N. and Tretyakov M.V., *Stochastic Numerics for Mathematical Physics*, Springer, 2004.

[39] Moler C. and Van Loan C., Nineteen dubious ways to compute the exponential of a matrix, SIAM Review, 20 (1978) 801-836.

[40] Ochs G., Weak random attractors, Report Nr. 449, Institut fur Dynamische Systeme, 1999.

[41] Ozaki T., Nonlinear Time Series Models and Dynamical Systems. In *Handbook of Statistics*, Vol 5, edited by E.J. Hannan et al , North-Holland, 1985, 25-83.

[42] Protter P., Stochastic integration and differential equations. Springer-Verlag, 1990.

[43] Riera J.J., Wan X., Jimenez J.C. and Kawashima R., Nonlinear local electro-vascular coupling. Part I: A theoretical model, Human Brain Mapping, 27 (2006) 896-914.

[44] Robinson J.C., Stability of random attractors under perturbation and approximation, J. Diff. Equat., 186 (2002) 652–669.

[45] Saito Y. and Mitsui T., Simulation of stochastic differential equations, Ann. Inst. Statist. Math., 45 (1993) 419-432.

[46] Saito Y. and Mitsui T., Stability analysis of numerical schemes for stochastic differential equations, SIAM journal of Numerical Analysis, 33 (1996) 2254-2267.

[47] Salomón L.A. and Biscay R.J., A simulation study of the Local Linearization method for the numerical (strong) solution of stochastic differential equations driven by alpha-stable Levy motions, Revista Investigación Operacional, 28 (2007) 51-60.

[48] Schurz H. The invariance of asymptotic laws of linear stochastic systems under discretization. ZAMM Z. Angew. Math. Mech., 79 (1999) 375–382.

[49] Schurz H., Numerical analysis de stochastic differential equations without tears. In Handbook of Stochastic Analysis & Applications, D. Kannan, V. Lakahmikamtham (Eds.) Marcell Dekker, 2002, 237-358.

[50] Shoji I., Approximation of continuous time stochastic processes by a local linearization method, Math. Comput., 67 (1998) 287-298.

[51] Talay D., Second order discretization schemes of stochastic differential systems for the computation of the invariant law. Stochastic Stochas. Rep., 29 (1990) 13-36.

[52] Valdes P.A., Jimenez J.C., Riera J., Biscay R. and Ozaki T., Nonlinear EEG analysis based on a neural mass model. Biol. Cyb., 81 (1999) 415-424.

[53] Veuthey A.L. and Stucki J., The adenylate kinase reaction acts as a frequency filter towards fluctuations of ATP utilization in the cell, Biophys. Chem., 26 (1987) 19-28.

[54] Wasilowski G.W. and Zniakowski H.W. On the complexity of stochastic integration. Math. Comput., 70 (2000) 685-698.