An Euler-Poisson Scheme for Lévy driven SDEs

A. Ferreiro-Castilla* A.E. Kyprianou† R. Scheichl†

January 27, 2015

Abstract

We describe an Euler scheme to approximate solutions of Lévy driven Stochastic Differential Equations (SDE) where the grid points are given by the arrival times of a Poisson process and thus are random. This result extends previous work of the authors in Ferreiro-Castilla et al. [11]. We provide a complete numerical analysis of the algorithm to approximate the terminal value of the SDE and proof that the mean square error converges with rate $\mathcal{O}(n^{-1/2})$. The only requirement of the methodology is to have exact samples from the resolvent of the Lévy process driving the SDE. Classical examples, such as stable processes, subclasses of spectrally one–sided Lévy processes and new families, such as meromorphic Lévy processes (Kuznetsov et al. [20]), are examples for which our algorithm provides an interesting alternative to existing methods, due to its straightforward implementation and its robustness with respect to the jump structure of the driving Lévy process.

Key words and phrases: Lévy processes, meromorphic Lévy processes, stochastic differential equations, Euler schemes.

MSC 2000 subject classifications: 60H10, 65C05

1 Introduction

Let $Y := \{Y_t\}_{t\in[0,T]}$ be the solution of the stochastic differential equation (SDE)

$$Y_t = y_0 + \int_0^t a(Y_{s-})\mathrm{d}X_s \qquad t \in [0,T] \ , \tag{1}$$

where a is smooth enough so that (1) has a strong solution. There is a great need from applications in mathematical finance, insurance mathematics, mathematical biology, physics and engineering to solve such SDEs numerically (see e.g. [6, 14, 28, 29]). Most studies deal with the case that $X := \{X_t\}_{t\in[0,T]}$ is a Wiener process. The complete path of X is numerically intractable and, ultimately, any numerical scheme can only be based on simulating the increments of the driving process. Therefore, typical approximation schemes rely on Taylor type approximations of the integral. For Itô integrals with respect to Wiener processes, Taylor expansions of arbitrary order are available and therefore approximations of arbitrary convergence rate (cf. Kloeden and Platen [17]).

Several problems arise when X in (1) is replaced by a Lévy process. For instance, increments of X are not available in general and approximations of the driving process are required. Moreover, multiple stochastic integrals with respect to Poisson measures are more difficult to handle and most numerical schemes are based on modifications of a first order Taylor approximation or an Euler scheme, although higher order schemes can be described as in Baran [3]. The basic Euler scheme for (1) is then

$$\widehat{Y}_0 = y_0 \ , \qquad \widehat{Y}_{t_{i+1}} = \widehat{Y}_{t_i} + a(\widehat{Y}_{t_i})(X_{t_{i+1}} - X_{t_i}), \qquad \text{for} \ \ 0 \le i \le n-1, \tag{2}$$

where $\{t_i\}_{0\le i\le n}$ (typically $t_i = \frac{iT}{n}$) is a deterministic partition of $[0,T]$ and $n \in \mathbb{N}$. For the exact Euler scheme, where the increments of the Lévy process X are available, convergence rates are explicit for the

*Corresponding author. Supported by a Royal Society Newton International Fellowship.

†Department of Mathematical Sciences, University of Bath, Claverton Down, Bath, BA2 7AY, U.K.

weak and the strong error. The weak error refers to the convergence rate of $|\mathbb{E}[f(Y_T)] - \mathbb{E}[f(\widehat{Y}_T)]|$ for a function f in a suitable class. Protter and Talay [25] require $f \in \mathcal{C}^4(\mathbb{R})$ in addition to some condition on the first moments of X to show $|\mathbb{E}[f(Y_T)] - \mathbb{E}[f(\widehat{Y}_T)]| = \mathcal{O}(n^{-1})$. The literature on the strong error estimates is less extensive. The strong error refers to the p-th moment, for $p \geq 1$, of the pathwise convergence, i.e. $\mathbb{E}[\sup_{t\in[0,T]} |Y_t - \widehat{Y}_t|^p]$. It can be inferred from Dereich and Heidenreich [9] that under the assumption that finite second moments of X exist, we also have $\mathbb{E}[\sup_{t\in[0,T]} |Y_t - \widehat{Y}_t|^2] = \mathcal{O}(n^{-1})$.

However, the above convergence rates are theoretical, since the exact distributions of the increments of Lévy processes are in general not available and an extra approximation error needs to be incorporated. See for instance Jacod et al. [15] for a weak error estimate with fairly general assumptions on the approximation of the increments of X, or Dereich and Heidenreich [9] for a strong error estimate where the jump component of X is truncated below a certain threshold. Indeed, the most common approach relies on the Lévy-Itô decomposition and removes the jumps below a given threshold, transforming the original Lévy process into a jump diffusion process. Therefore, the final convergence rates depend in general on the structure of the small jumps. Compound Poisson processes are piecewise constant processes with jumps happening at the arrival times of a Poisson point process. Hence, a more promising modification is to move away from the deterministic equally-spaced grid points in (2). A jump-adapted discretization scheme consists in interlacing an equally-spaced grid for the approximation of the continuous component of the driving process, with a random grid given by the jump times of the purely discontinuous part, as described in Rubenthaler [26]. In its simplest form, the approximation can perform very poorly when the jump component has paths of infinite p-variation, with p close to 2, as shown in Dereich and Heidenreich [9] (recall that all Lévy processes have finite 2-variation paths). A more sensible approach is to substitute the small jumps by a Gaussian correction as performed in Dereich [8], but this method has its limitations as discussed in Asmussen and Rosiński [2]. A novel approach described in Kohatsu-Higa et al. [18] is to approximate the small jumps with an extra compound Poisson process matching a given number of moments of the original driving process, provided these moments exist. Convergence rates for weak errors are derived under further assumptions on the smoothness of the function f. Under the assumption that the Lévy measure is a regularly varying function, the authors in [18] combine the above approach with a high order scheme for the continuous part, obtaining arbitrary convergence rates for the weak error.

The aim of this paper is to describe an Euler scheme defined entirely on a random grid, built from the arrival times of a Poisson process. In all the methodologies mentioned above, the largest time step in the Euler approximation is bounded above by a constant. In our scheme this feature can no longer be assumed, as the inter-arrival times of a Poisson process are exponentially distributed. The origin of this scheme is based on recent developments for Wiener-Hopf factorizations of Lévy processes by Kuznetsov [19] and Kuznetsov et al. [20, 21]. The Wiener-Hopf factorization is a distributional decomposition of the path of a Lévy process in terms of the running supremum and the running infimum. In Ferreiro-Castilla et al. [11] this factorization is used to sample from the bivariate distribution of $(X_t, \sup_{s<t} X_s)$ by constructing a random walk approximation with time steps chosen according to a an exponential distribution, i.e. the arrival times of a Poisson process. This scheme effectively constructs a skeleton of the path of X and therefore it is natural to investigate also how this skeleton would perform to obtain approximations of (1).

Although the skeleton constructs a random walk approximation of the path which captures not only the end point but the supremum over each exponential time step, in the present paper we will consider an Euler scheme for the solution Y_T of (1) at the end point only. Therefore, the proposed algorithm is a modification of the Euler scheme where we assume that we can sample from the distribution of $X_{\mathbf{e}(n/T)}$ for exponentially distributed time steps $\mathbf{e}(n/T)$ with mean T/n independent of X. In other words, the grid points in our Euler scheme are given by a Poisson point process with rate n/T denoted by $\mathbf{N}(n/T)$, where the mean T/n plays the role of the grid size. We will call our scheme the Euler-Poisson scheme. Our analysis does not assume any way of obtaining the distribution of $X_{\mathbf{e}(n/T)}$ and there is no reason why the latter should be easier than the distribution of X_1, for a general Lévy process. Nevertheless, for a large class of processes called meromorphic Lévy processes, see Kuznetsov et al. [21], the distribution of $X_{\mathbf{e}(n/T)}$ is explicit and samples from it are numerically easy to obtain. However, the most important and significant advantage is that in contrast to the more classical methods mentioned above, our numerical performance does not depend on the jump structure of X. The main result of the paper derives the convergence rate in mean square error for the approximation $\widetilde{Y}_n$ of Y_T obtained via the Euler-Poisson scheme, showing that $\mathbb{E}[|Y_T - \widetilde{Y}_n|^2] = \mathcal{O}(n^{-1/2})$. We will also show that our methodology is closely related to classical discretization schemes for the Partial Integro-Differential Equation associated with computing $\mathbb{E}[f(Y_T)]$, for a given function f.

The paper is organised as follows. In the next section, we will introduce the basic notation, describe the Euler-Poisson scheme and state our main result. The numerical analysis of our methodology is given in Section 3. Finally, we collect several remarks and observations regarding feasibility, extensions and its relation with PIDEs about our scheme in Section 4.

2 The Euler-Poisson scheme

2.1 Preliminaries

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$ be a filtered probability space and let $Y := \{Y_t\}_{t\in[0,T]}$ be a $\mathbb{R}^{d_Y}$-valued, adapted stochastic process which is the strong solution of the stochastic differential equation

$$Y_t = y_0 + \int_0^t a(Y_{s-})\mathrm{d}X_s \qquad t \in [0,T] \ , \tag{3}$$

where $a := \mathbb{R}^{d_Y} \to \mathbb{R}^{d_Y} \otimes \mathbb{R}^{d_X}$ is a coefficient with smoothness to be specified, $X := \{X_t\}_{t\in[0,T]}$ is a d_X-dimensional square-integrable Lévy process, $y_0 \in \mathbb{R}^{d_Y}$ and $T < \infty$. Recall that a Lévy process is a stochastic process issued from the origin which enjoys the properties of having stationary and independent increments with paths that are almost surely right-continuous with left limits. It is a well understood fact that, as a consequence, the law of every Lévy process is characterised through a triplet (b, Σ, Π), where $b \in \mathbb{R}^{d_X}$, $\Sigma \in \mathbb{R}^{d_X\times d_X}$ and Π is a measure concentrated on $\mathbb{R}^{d_X}\backslash\{0\}$ such that $\int_{\mathbb{R}^{d_X}}(1\wedge |x|^2)\Pi(\mathrm{d}x) < \infty$. For square-integrable Lévy processes we have, for all $t \geq 0$ and $\theta \in \mathbb{R}^{d_X}$,

$$\mathbb{E}[\mathrm{e}^{\mathrm{i}\langle\theta, X_t\rangle}] = \mathrm{e}^{-t\Psi(\theta)}, \quad \text{where} \quad \Psi(\theta) \ = \ \mathrm{i}\langle b,\theta\rangle + \frac{1}{2}\langle\theta, \Sigma\Sigma^{\mathrm{T}}\theta\rangle + \int_{\mathbb{R}^{d_X}}(1-\mathrm{e}^{\mathrm{i}\langle\theta,x\rangle} + \mathrm{i}\langle\theta,x\rangle)\Pi(\mathrm{d}x)$$

is the so-called characteristic exponent of the process and $\langle\cdot,\cdot\rangle$ is the usual inner product. Furthermore, the Lévy-Itô decomposition guaranties that we can decompose X as

$$X_t = \Sigma W_t + L_t + bt \qquad t \geq 0 \ , \tag{4}$$

where $W := \{W_t\}_{t\in[0,T]}$ is a d_X-dimensional Wiener process and $L := \{L_t\}_{t\in[0,T]}$ is a d_X-dimensional $L^2(\Omega, \mathcal{F}, P)$ martingale representing the compensated jumps of X. For ease of notation, we will assume in the following derivations, without loss of generality, that there exists a constant $k \in \mathbb{R}^+$ such that

$$\int_{\mathbb{R}^{d_X}} |x|^2\Pi(\mathrm{d}x) \leq k^2 \ , \qquad |\Sigma| \leq k \ , \qquad |b| \leq k \qquad \text{and} \qquad |y_0| \leq k \ .$$

We use indistinctly $|\cdot|$ to denote the Euclidean norm for vectors or the Frobenius norm for matrices.

Theorem 2.1 (Situ [28, Section 3.1]) *Consider the SDE driven by a square integrable Lévy process given in (3). Let $a := \mathbb{R}^{d_Y} \to \mathbb{R}^{d_Y} \otimes \mathbb{R}^{d_X}$ be a measurable funcion such that*

$$|a(x) - a(x')| \leq k'|x - x'| \qquad \text{and} \qquad |a(y_0)| \leq k'$$

for $x, x' \in \mathbb{R}^{d_Y}$ and $k' \in \mathbb{R}^+$. Then, equation (3) has a unique strong solution adapted to the filtration generated by X, $\mathcal{F}^X$, and there exists a positive constant K_1 depending only on k' and T such that

$$\mathbb{E}\left[\sup_{t\in[0,T]}|Y_t|^2\right] \leq K_1 \ .$$

Without loss of generality we set $k' = k$ in Theorem 2.1. In the following, all constants denoted by K_i and κ_i only depend on k and T and may be renamed without further notice in consecutive equations.

2.2 The discretization scheme

As mentioned in the introduction, this paper is concerned with a modification of the standard Euler scheme, replacing equally-spaced time steps by exponentially distributed ones, so that the grid points in our scheme are arrival times of a Poisson process. For $n \geq 1$, let $\{\mathbf{e}_i(n/T)\}_{i\geq 1}$ be an i.i.d. sequence of random variables in $(\Omega, \mathcal{F}, P)$ where $\mathbf{e}(q)$ denotes an exponential random variable such that $\mathbb{E}[\mathbf{e}(q)] = q^{-1}$ and denote by $\mathcal{G}$ the σ-algebra generated by $\{\mathbf{e}_i(n/T)\}_{i\geq 1}$, assumed to be independent of X; we set $\mathbf{e}_0 = 0$ for convenience. We will also denote by $\mathbf{N}(n/T) := \{\mathbf{N}_t(n/T)\}_{t\geq 0}$ the Poisson process with arrival times $\{t_i\}_{i\geq 0}$. In the above description the mean T/n is the analog of the grid size for deterministic spaced Euler schemes. The Euler-Poisson scheme is then given by the discrete Markov chain $\widetilde{Y} := \{\widetilde{Y}_{t_i}\}_{i\geq 0}$ defined recursively by

$$\widetilde{Y}_{t_i} := \widetilde{Y}_{t_{i-1}} + a(\widetilde{Y}_{t_{i-1}})\Delta X_{\mathbf{e}_i(n/T)} \qquad \text{for} \quad i \geq 1 \quad \text{and} \quad \widetilde{Y}_0 := y_0 \ , \tag{5}$$

where $\Delta X_{\mathbf{e}_i} := X_{\mathbf{e}_i(n/T)} - X_{\mathbf{e}_{i-1}(n/T)} \stackrel{d}{=} X_{\mathbf{e}(n/T)}$ and $t_i := \sum_{j=0}^{i} \mathbf{e}_j(n/T)$. Note that $t_i \stackrel{d}{=} \mathbf{g}(i, n/T)$, where $\mathbf{g}(\alpha, \beta)$ denotes a Gamma distribution with shape parameter α and rate parameter β. We claim that $\widetilde{Y}_{t_n}$ is an approximation of Y_T and the task of this paper is to derive the asymptotic behaviour of

$$\lim_{n\to\infty} \mathbb{E}[|Y_T - \widetilde{Y}_{t_n}|^2] \ . \tag{6}$$

Before we proceed, let us introduce a new process which stochastically interpolates the Euler-Poisson scheme. Denote by $\iota(t)$ the largest grid point before t, i.e. $\iota(t) := \sup[0, t] \cap \{t_i\}_{i\geq 0}$, and define

$$\widehat{Y}_t := y_0 + \int_0^t a(\widehat{Y}_{\iota(s-)})\mathrm{d}X_s = \widehat{Y}_{\iota(t)} + a(\widehat{Y}_{\iota(t)})(X_t - X_{\iota(t)}) \tag{7}$$

for $t \in [0, t_n \vee T]$. Notice that for $t \in [t_i, t_{i+1})$ we have $\widetilde{Y}_{t_i} = \widehat{Y}_{t_i} = \widehat{Y}_{\iota(t)}$ and hence $\widehat{Y} := \{\widehat{Y}_t\}_{t\in[0,t_n\vee T]}$ interpolates, in a random way, the chain $\widetilde{Y}$. Yet another important random variable which is going to play a crucial role in the following derivations is the largest gap of the random grid $\{t_i\}_{i\geq 0}$ restricted to $[0, T]$. Let us denote this $\mathcal{G}$-measurable random variable by

$$\tau := \sup\nolimits_{s\in[0,T]}(s - \iota(s)) \ . \tag{8}$$

2.3 Main result and feasibility of the Euler-Poisson scheme

With the above notation we can now formally state the main result of the paper, proved in Section 3.

Theorem 2.2 *Under the assumptions of Theorem 2.1, there exists a positive constant K_2 depending only on k and T such that*

$$\mathbb{E}[|Y_T - \widetilde{Y}_{t_n}|^2] \leq K_2 n^{-1/2} \ .$$

It is clear from the preceding section that the Euler-Poisson method is of practical interest only if samples from the distribution of $X_{\mathbf{e}(q)}$ are available. In general, there is no reason why the latter distribution is easier to handle than the distribution of X_1 itself. Nevertheless, recent developments in Wiener-Hopf theory for 1-dimensional Lévy processes have provided a rich enough variety of examples for which the necessary distributional sampling can be performed and thus the Euler-Poisson scheme may lead to simpler numerical techniques for (3). This family of processes are called meromorphic Lévy processes, see Kuznetsov et al. [20, 21]. For the class of meromorphic Lévy processes, the Wiener-Hopf factors are explicit and hence we can efficiently sample from the distribution of $X_{\mathbf{e}(q)}$ through its factorization. Indeed, numerical algorithms involving the computation of $X_{\mathbf{e}(q)}$ for meromorphic Lévy processes are very easy to implement and robust with respect to the jump structure, see for example Ferreiro-Castilla et al. [11]. One large subfamily of such processes is the β-class of Lévy processes, which also conveniently offers all the desirable properties of better known Lévy processes that are used in mathematical finance, such as CGMY processes, VG processes or Meixner processes; see for example the discussions in Ferreiro-Castilla and Schoutens [12] and Schoutens and van Damme [27]. This brings the possibility to study new processes associated to the SDE (3). For instance, the results in Ferreiro-Castilla et al. [11] and the ones presented here suggest that we can sample and numerically analyze approximate solutions for SDEs like

$$Y_t = y_0 + \int_0^t a(Y_{s-}, \overline{X}_{s-})\mathrm{d}X_s \qquad \text{or} \qquad Y_t = y_0 + \int_0^t a(Y_{s-}, \underline{X}_{s-})\mathrm{d}X_s \ ,$$

where $\overline{X}_t = \sup_{s\leq t} X_s$ and $\underline{X}_t := \inf_{s\leq t} X_s$. To our knowledge, such SDEs have not yet been numerically considered in the literature, but it is not difficult to imagine applications of such processes. For instance, models that appear in stochastic dynamics for population or chemical reactions might be modeled by the above SDEs where the knowledge of $\overline{X}$ can replace the artificial barrier restrictions that are usually imposed on the driving processes due to physical constrains (e.g. Situ [28, Chapter 11]). In financial mathematics it might be used to model drawdown or barrier constraints on credit derivatives.

3 Numerical Analysis

The construction of the Euler-Poisson scheme uses a random grid that is supported on an interval that can be smaller or bigger than $[0,T]$. We will split the mean square error described in (6) between what we denote by the discretization error and the hitting error. To fix ideas, let us write

$$|Y_T - \widetilde{Y}_{t_n}| = |Y_T - \widehat{Y}_{t_n}| \leq |Y_T - \widehat{Y}_T| + |\widehat{Y}_T - \widehat{Y}_{t_n}| \ , \tag{9}$$

where the first term on the right hand side of the above inequality corresponds to the discretization error and the second term to the hitting error.

3.1 The discretization error

Heuristically, the discretization error should behave as in the classical Euler scheme for deterministic equally-spaced grid points. In order to see this, we first derive a technical lemma which obtains the analogous result for $\widehat{Y}$ to the one described in Theorem 2.1 for Y.

Lemma 3.1 *Under the assumptions of Theorem 2.1, the process $\widehat{Y}$ defined in (7) is adapted to $\mathcal{G}\bigvee\mathcal{F}^X$ and there exists a constant $K_3 > 0$ such that*

$$\text{(i)}\ \ \mathbb{E}\left[\sup\nolimits_{t\in[0,T]}|\widehat{Y}_t|^2\right] \leq K_3 \qquad \text{and} \qquad \text{(ii)}\ \ \mathbb{E}\left[\sup\nolimits_{t\in[0,T]}|\widehat{Y}_t|^2 \;\middle|\; \mathcal{G}\right] \leq K_3 \ .$$

Proof. The adaptivity property is clear from the right hand side of (7). The square integrability of the first part (i) follows similarly as in the proof of Theorem 2.1, which we briefly review here for the sake of completeness. Let $\sigma_N := \inf\{t > 0 \mid |\widehat{Y}_t| > N\}$, $t \in [0,T]$. Then, using the definition of $\widehat{Y}$ and the Cauchy-Schwarz inequality for the random Lebesgue integral, we have

$$\begin{aligned}\frac{1}{3}|\widehat{Y}_{t\wedge\sigma_N}|^2 &\leq |y_0|^2 + \left|\int_0^{t\wedge\sigma_N} a(\widehat{Y}_{\iota(s)})b\mathrm{d}s\right|^2 + \left|\int_0^{t\wedge\sigma_N} a(\widehat{Y}_{\iota(s-)})\mathrm{d}(\Sigma W_s + L_s)\right|^2 \\ &\leq |y_0|^2 + (t\wedge\sigma_N)k^2\int_0^{t\wedge\sigma_N}|a(\widehat{Y}_{\iota(s)})|^2\mathrm{d}s + \left|\int_0^{t\wedge\sigma_N} a(\widehat{Y}_{\iota(s-)})\mathrm{d}(\Sigma W_s + L_s)\right|^2 \ .\end{aligned} \tag{10}$$

Using the Lipschitz condition of a, we further derive the growth condition

$$|a(x)|^2 = |a(x) - a(y_0) + a(y_0)|^2 \leq 4k^2|x|^2 + 2k^2(2k^2+1) \leq K_0|x|^2 + K_0 \ , \tag{11}$$

for a constant K_0 depending on k only. Hence, using the definition of the stopping time σ_N, we conclude that the stochastic integral in (10) is a square-integrable martingale, to which we apply Doob's inequality and the Itô isometry to obtain

$$\begin{aligned}\frac{1}{3}\mathbb{E}\left[\sup_{r\le t\wedge\sigma_N}|\widehat{Y}_r|^2\right] &\le k^2+tk^2\mathbb{E}\left[\int_0^{t\wedge\sigma_N}|a(\widehat{Y}_{\iota(s)})|^2\mathrm{d}s\right]+8k^2\mathbb{E}\left[\int_0^{t\wedge\sigma_N}|a(\widehat{Y}_{\iota(s)})|^2\mathrm{d}s\right]\\ &\le k^2+(tk^2+8k^2)\left(K_0\mathbb{E}\left[\int_0^{t\wedge\sigma_N}|\widehat{Y}_{\iota(s)}|^2\mathrm{d}s\right]+K_0t\right)\\ &\le \kappa_1+\kappa_1\int_0^t\mathbb{E}\left[\sup_{r\le s\wedge\sigma_N}|\widehat{Y}_r|^2\right]\mathrm{d}s\ ,\end{aligned}$$

where κ_1 is a constant only depending on k and T. Finally, Gronwall's lemma gives

$$\mathbb{E}\left[\sup_{r\le t\wedge\sigma_N}|\widehat{Y}_r|^2\right]\le 3\kappa_1 e^{3\kappa_1 t}\le 3\kappa_1 e^{3\kappa_1 T}=K_3$$

and (i) follows by letting $N\uparrow\infty$. The second part of the claim follows analogously by noting that X is independent of $\mathcal{G}$; therefore, conditioned on $\mathcal{G}$, the stochastic integral

$$\int_0^{t\wedge\sigma_N}a(\widehat{Y}_{\iota(s-)})\mathrm{d}(\Sigma W_s+L_s)$$

is a martingale with respect to $\mathcal{F}^X$, allowing us to use conditioned versions of Doob's inequality and of Itô isometry. The bound in (ii) then follows in the same way as above for (i). □

The following theorem derives the asymptotic behavior for the discretization error which ultimately depends on the random grid size τ defined in (8). The necessary results to obtain bounds for the moments of τ are derived in Appendix A.

Theorem 3.2 *Under the assumptions of Theorem 2.1, there exists a constant $K_4>0$ such that*

$$\mathbb{E}\left[\sup_{t\in[0,T]}|Y_t-\widehat{Y}_t|^2\right]\le K_4\,n^{-1}\log(n)\,.$$

Proof. Let $t\in[0,T]$ and define

$$Z_t:=Y_t-\widehat{Y}_t=\int_0^t(a(Y_s)-a(\widehat{Y}_{\iota(s)}))b\mathrm{d}s+\int_0^t(a(Y_{s-})-a(\widehat{Y}_{\iota(s-)}))\mathrm{d}(\Sigma W_s+L_s)\ . \tag{12}$$

From Theorem 2.1 and Lemma 3.1, we deduce that the stochastic integral on the right hand side of (12) is a square integrable martingale with respect to the filtration $\mathcal{G}\bigvee\mathcal{F}^X$. We apply Cauchy-Schwarz inequality to the random Lebesgue integral and Doob's martingale inequality plus the Itô isometry to the stochastic integral in (12) to end up with

$$\begin{aligned}\frac{1}{2}\mathbb{E}[\sup_{r<t}|Z_r|^2] &\le \mathbb{E}\left[\sup_{r<t}\left(\int_0^r(a(Y_s)-a(\widehat{Y}_{\iota(s)}))b\mathrm{d}s\right)^2+\left(\int_0^r(a(Y_{s-})-a(\widehat{Y}_{\iota(s-)}))\mathrm{d}(\Sigma W_s+L_s)\right)^2\right]\\ &\le k^2\mathbb{E}\left[t\int_0^t|Y_s-\widehat{Y}_{\iota(s)}|^2\mathrm{d}s\right]+8k^2\mathbb{E}\left[\int_0^t|Y_s-\widehat{Y}_{\iota(s)}|^2\mathrm{d}s\right]\\ &\le \kappa_2\int_0^t\mathbb{E}[|Z_s|^2]+\mathbb{E}[|\widehat{Y}_s-\widehat{Y}_{\iota(s)}|^2]\mathrm{d}s\ \le\ \kappa_2\int_0^t\mathbb{E}[\sup_{r<s}|Z_r|^2]+\mathbb{E}[|\widehat{Y}_s-\widehat{Y}_{\iota(s)}|^2]\mathrm{d}s\,,\end{aligned} \tag{13}$$

where κ_2 is a positive constant depending on k and T only. The next objective is to use Gronwall's lemma in inequality (13). This will rely on controlling $|\widehat{Y}_s-\widehat{Y}_{\iota(s)}|$. Since X has independent increments and due to the growth condition of $a(x)$ in (11), we can write

$$\begin{aligned}\mathbb{E}[|\widehat{Y}_s-\widehat{Y}_{\iota(s)}|^2] &= \mathbb{E}[|a(\widehat{Y}_{\iota(s)})|^2]\mathbb{E}[|X_s-X_{\iota(s)}|^2]\le\left(K_0\mathbb{E}[|\widehat{Y}_{\iota(s)}|^2]+K_0\right)\mathbb{E}[|X_s-X_{\iota(s)}|^2]\\ &\le \left(2K_0\mathbb{E}[|Z_{\iota(s)}|^2]+2K_0\mathbb{E}[|Y_{\iota(s)}|^2]+K_0\right)\mathbb{E}[|X_s-X_{\iota(s)}|^2]\ .\end{aligned} \tag{14}$$

Now,

$$\mathbb{E}[|X_s - X_{\iota(s)}|^2] \le k^2 \mathbb{E}[2\tau + \tau^2] \le k^2(2T + T^2)$$

and so together with (13) and (14), as well as Theorem 2.1, we obtain

$$\mathbb{E}\left[\sup_{r<t} |Z_r|^2\right] \le \kappa_2 \mathbb{E}[2\tau + \tau^2] + \kappa_2 \int_0^t \mathbb{E}\left[\sup_{r<s} |Z_r|^2\right] \mathrm{d}s\,,$$

where we renamed the constant κ_2. It follows from Gronwall's inequality that

$$\mathbb{E}\left[\sup_{t\in[0,T]} |Y_t - \widehat{Y}_t|^2\right] \le \mathbb{E}[2\tau + \tau^2]\kappa_2 e^{T\kappa_2} = K_4 \mathbb{E}[2\tau + \tau^2].$$

This completes the proof of the theorem up to bounding $\mathbb{E}[2\tau + \tau^2]$. This bound follows from Proposition A.1 in Appendix A. $\square$

3.2 The hitting error

The next result derives the asymptotic behaviour for the hitting error, which boils down to measuring how fast the random time t_n converges to T. This, in turn, is controlled by the variance of a Gamma distribution. Before we proceed, let us first derive two technical lemmas in the spirit of Lemma 3.1.

Lemma 3.3 *Under the assumptions of Theorem 2.1, the process $\widehat{Y}$ defined in (7) is adapted to $\mathcal{G} \bigvee \mathcal{F}^X$ and there exists a constant $K_5 > 0$ such that*

$$\max_{0\le i\le n} \mathbb{E}[|\widehat{Y}_{t_i}|^2] \le K_5\,.$$

Proof. Fix $i > 0$ and recall the definition of $\widehat{Y}_{t_i}$ in (7) to write

$$\begin{aligned}\mathbb{E}[|\widehat{Y}_{t_i}|^2] &= \mathbb{E}[|\widehat{Y}_{t_{i-1}}|^2] + \mathbb{E}[|a(\widehat{Y}_{t_{i-1}})|^2]\mathbb{E}[|X_{t_i} - X_{t_{i-1}}|^2] + 2\mathbb{E}[\widehat{Y}_{t_{i-1}}^{\mathrm{T}} a(\widehat{Y}_{t_{i-1}})]\mathbb{E}[X_{t_i} - X_{t_{i-1}}] \\ &\le \mathbb{E}[|\widehat{Y}_{t_{i-1}}|^2]\left(1 + K_0 2k^2 \frac{T}{n}\left(1 + \frac{T}{n}\right) + 2\sqrt{K_0}k\frac{T}{n}\right) + K_0 2k^2 \frac{T}{n}\left(1 + \frac{T}{n}\right) + 2\sqrt{K_0}k\frac{T}{n}\,, \end{aligned} \qquad (15)$$

where we used that $t_i - t_{i-1} \overset{d}{=} \mathbf{e}(n/T)$ and the orthogonal decomposition of X in (4), as well as the growth condition (11) and the following inequality, which follows from the assumptions on $a(x)$:

$$|x^{\mathrm{T}} a(x)| \le \sqrt{K_0}|x|^2 + \sqrt{K_0}\,.$$

It is then clear from (15) that there exists a constant κ_3, depending on k and T only, such that

$$\mathbb{E}[|\widehat{Y}_{t_i}|^2] \le \mathbb{E}[|\widehat{Y}_{t_{i-1}}|^2]\left(1 + \frac{\kappa_3}{n}\right) + \frac{\kappa_3}{n} \le |y_0|^2\left(1 + \frac{\kappa_3}{n}\right)^i + i\exp\left(i\frac{\kappa_3}{n}\right)\frac{\kappa_3}{n}\,,$$

which follows from the argument that if $x_{m+1} \le \alpha x_m + \beta$ and $\alpha \ge 1$, then $x_m \le \alpha^m x_0 + m e^{m(\alpha-1)}\beta$. Finally,

$$\max_{0\le i\le n} \mathbb{E}[|\widehat{Y}_{t_i}|^2] \le |y_0|^2\left(1 + \frac{\kappa_3}{n}\right)^n + e^{\kappa_3}\kappa_3 \le e^{\kappa_3}\left(k^2 + \kappa_3\right)$$

which concludes the proof. $\square$

Lemma 3.4 *Under the assumptions of Theorem 2.1, the process $\widehat{Y}$ defined in (7) is adapted to $\mathcal{G} \bigvee \mathcal{F}^X$ and there exists a constant $K_6 > 0$ such that*

$$\text{(i)}\ \ \mathbb{E}\left[\max_{0\le i\le n} |\widehat{Y}_{t_i}|^2\right] \le K_6 \quad \text{and} \quad \text{(ii)}\ \ \mathbb{E}\left[\left(\mathbb{E}\left[\max_{0\le i\le n} |\widehat{Y}_{t_i}|^2 | \mathcal{G}\right]\right)^2\right] \le K_6\,.$$

Proof. We define $\Delta\widehat{Y}_i := \widehat{Y}_{t_{i+1}} - \widehat{Y}_{t_i}$ and use the same principles as in (15) and Lemma 3.3 to derive

$$\mathbb{E}[|\Delta\widehat{Y}_{t_i}|^2] = \mathbb{E}[|a(\widehat{Y}_{t_i})|^2]\mathbb{E}[|X_{t_{i+1}} - X_{t_i}|^2] \le (K_0K_5 + K_0)2k^2\frac{T}{n}\left(1+\frac{T}{n}\right), \quad \text{for } 0 \le i \le n-1,$$

and hence there exists a constant κ_4 depending only on k and T such that

$$\max\nolimits_{0\le i\le n-1} \mathbb{E}[|\Delta\widehat{Y}_{t_i}|^2] \le \kappa_4\, n^{-1}\,. \tag{16}$$

Consider now the filtration $\mathcal{H}_i := \sigma\langle \widehat{Y}_{t_j},\ 0 \le j \le i\rangle$ and the auxiliary random variables

$$Z_i := \Delta\widehat{Y}_{t_i} - \mathbb{E}[\Delta\widehat{Y}_{t_i}|\mathcal{H}_i],$$

for $0 \le i \le n-1$. It is clear that Z_i is $\mathcal{H}_{i+1}$-measurable and it is not difficult to check that $\sum_{j=0}^{i} Z_i$ is a martingale such that $\mathbb{E}[Z_iZ_j] = 0$ if $i \ne j$. Therefore we can write

$$\begin{aligned}
\max_{0\le i\le n} |\widehat{Y}_{t_i}|^2 &\le 2\left(|y_0|^2 + \max_{0\le i\le n-1}\left|\sum_{j=0}^{i}\Delta\widehat{Y}_{t_j}\right|^2\right) = 2\left(|y_0|^2 + \max_{0\le i\le n-1}\left|\sum_{j=0}^{i} Z_j + \mathbb{E}[\Delta\widehat{Y}_{t_j}|\mathcal{H}_j]\right|^2\right) \\
&\le 2\left(|y_0|^2 + 2\underbrace{\max_{0\le i\le n-1}\left|\sum_{j=0}^{i} Z_j\right|^2}_{(*)} + 2\underbrace{\max_{0\le i\le n-1}\left|\sum_{j=0}^{i}\mathbb{E}[\Delta\widehat{Y}_{t_j}|\mathcal{H}_j]\right|^2}_{(**)}\right).
\end{aligned} \tag{17}$$

We now use Doob's martingale inequality and the orthogonality of $\{Z_i\}_{i=0}^{n-1}$ to bound $(*)$. Combining this with Jensen's inequality and (16) we find that

$$\mathbb{E}[(*)] \le \mathbb{E}\left[\sum_{j=0}^{n-1}|Z_j|^2\right] \le 2\mathbb{E}\left[\sum_{j=0}^{n-1}|\Delta\widehat{Y}_{t_j}|^2 + |\mathbb{E}[\Delta\widehat{Y}_{t_j}|\mathcal{H}_j]|^2\right] \le 4\sum_{j=0}^{n-1}\mathbb{E}\left[|\Delta\widehat{Y}_{t_j}|^2\right] \le 4\kappa_4\,.$$

Similarly, using Lemma 3.3, one obtains

$$\mathbb{E}[(**)] \le \mathbb{E}\left[\left(\sum_{j=0}^{n-1}|\mathbb{E}[\Delta\widehat{Y}_{t_j}|\mathcal{H}_j]|\right)^2\right] = \mathbb{E}\left[\left(\sum_{j=0}^{n-1}|a(\widehat{Y}_{t_j})|k\frac{T}{n}\right)^2\right] \le k^2T^2(K_0K_5 + K_0)\,.$$

The first part (i) follows substituting the upper bounds for $\mathbb{E}[(*)]$ and $\mathbb{E}[(**)]$ into (17).

For the second part (ii) we consider $\mathcal{H}_i := \mathcal{G}\bigvee\sigma\langle\widehat{Y}_{t_j},\ 0 \le j \le i\rangle$ and reproduce the above derivations up to (17). By the definition of $\Delta\widehat{Y}_{t_i}$, we have

$$\mathbb{E}[(*)|\mathcal{G}] \le 4\sum_{j=0}^{n-1}\mathbb{E}\left[|\Delta\widehat{Y}_{t_j}|^2\,\middle|\,\mathcal{G}\right] \le 8k^2\max_{0\le i\le n-1}\mathbb{E}\left[|a(\widehat{Y}_{t_i})|^2\,\middle|\,\mathcal{G}\right]\sum_{j=0}^{n}\mathbf{e}_j\,(1+\mathbf{e}_j) \tag{18}$$

$$\mathbb{E}[(**)|\mathcal{G}] \le \mathbb{E}\left[\left(\sum_{j=0}^{n-1}|a(\widehat{Y}_{t_j})|k\mathbf{e}_j\right)^2\,\middle|\,\mathcal{G}\right] \le k^2\max_{0\le i\le n-1}\mathbb{E}\left[|a(\widehat{Y}_{t_i})|^2\,\middle|\,\mathcal{G}\right] n\sum_{j=0}^{n}\mathbf{e}_j^2\,. \tag{19}$$

Therefore, to prove (ii) it is enough to recall (17) and to show that $\mathbb{E}\left[(\mathbb{E}[(*)|\mathcal{G}])^2 + (\mathbb{E}[(**)|\mathcal{G}])^2\right] \le \kappa_5$, for some constant κ_5 depending on k and T only. Using Cauchy-Schwarz inequality and renaming κ_5, a sufficient condition for this claim to be true is

$$\mathbb{E}\left[\left(\max_{0\le i\le n-1}\mathbb{E}\left[|a(\widehat{Y}_{t_i})|^2\,\middle|\,\mathcal{G}\right]\right)^4\right] + \mathbb{E}\left[\left(n\sum_{j=0}^{n}\mathbf{e}_j^2\right)^4 + \left(\sum_{j=0}^{n}\mathbf{e}_j\,(1+\mathbf{e}_j)\right)^4\right] \le \kappa_5\,. \tag{20}$$

Since $\mathbb{E}[\mathbf{e}(n/T)^i] = i!T^i/n^i$ for $i \ge 1$, one can check that

$$\mathbb{E}\left[\left(n\sum_{j=0}^{n}\mathbf{e}_j^2\right)^4 + \left(\sum_{j=0}^{n}\mathbf{e}_j\,(1+\mathbf{e}_j)\right)^4\right] \le 8!T^8 + 8\left(4!T^4 + \frac{8!T^8}{n^4}\right) \tag{21}$$

and the second term in (20) is bounded. Adapting the left hand side of (15) to incorporate the conditional expectation we write

$$\mathbb{E}\left[|\widehat{Y}_{t_i}|^2 \,\middle|\, \mathcal{G}\right] \leq \mathbb{E}\left[|\widehat{Y}_{t_{i-1}}|^2 \,\middle|\, \mathcal{G}\right]\left(1 + K_0 2k^2(\mathbf{e}_i + \mathbf{e}_i^2) + 2\sqrt{K_0}k\mathbf{e}_i\right) + K_0 2k^2(\mathbf{e}_i + \mathbf{e}_i^2) + 2\sqrt{K_0}k\mathbf{e}_i \ ,$$

where κ_6 is a constant that only depends on k. Using again a recurrence argument, we easily see that

$$\max_{0\leq i\leq n} \mathbb{E}\left[|\widehat{Y}_{t_i}|^2 \,\middle|\, \mathcal{G}\right] \leq |y_0|^2 \prod_{i=1}^n \left(1 + \kappa_6(\mathbf{e}_i + \mathbf{e}_i^2)\right) + \sum_{i=1}^n \kappa_6(\mathbf{e}_i + \mathbf{e}_i^2) \prod_{j=i+1}^n \left(1 + \kappa_6(\mathbf{e}_j + \mathbf{e}_j^2)\right) .$$

Finally, the bound on the first term in (20) follows from this inequality via the same sort of manipulations that were used in (21). Since (20) holds, so do (18) and (19), which proves (ii). □

Proposition 3.5 *Under the assumptions of Theorem 2.1, there exists a constant $K_7 > 0$ such that*

$$\mathbb{E}[|\widehat{Y}_T - \widehat{Y}_{t_n}|^2] \leq K_7 n^{-1/2} .$$

Proof. Let us write

$$\frac{1}{2}|\widehat{Y}_T - \widehat{Y}_{t_n}|^2 \leq \left|\int_{t_n}^T a(\widehat{Y}_{\iota(s)})b\mathrm{d}s\right|^2 + \left|\int_{t_n}^T a(\widehat{Y}_{\iota(s-)})\mathrm{d}(\Sigma W_s + L_s)\right|^2 .$$

According to part (i) of Lemmas 3.1 and 3.4, the stochastic integral in the above decomposition is a square integrable martingale with respect to $\mathcal{G}\bigvee\mathcal{F}^X$. Hence, we can use again Cauchy-Schwarz inequality to the random Lebesgue integral and the Itô isometry for the stochastic integral to obtain

$$\begin{aligned}
\frac{1}{2}\mathbb{E}[|\widehat{Y}_T - \widehat{Y}_{t_n}|^2] \;&\leq\; \mathbb{E}\left[\left(k^2|T - t_n| + 2k^2\right)\int_{t_n}^T |a(\widehat{Y}_{\iota(s)})|^2\mathrm{d}s\right] \\
&= k^2\mathbb{E}\left[(|T - t_n| + 2)\int_{t_n}^T \mathbb{E}\left[|a(\widehat{Y}_{\iota(s)})|^2 \,\middle|\, \mathcal{G}\right]\mathrm{d}s\right] \\
&\leq k^2\mathbb{E}\left[\left(|T - t_n|^2 + 2|T - t_n|\right)\left(\sup_{t\in[0,T\vee t_n]}\mathbb{E}\left[|a(\widehat{Y}_{\iota(s)})|^2 \,\middle|\, \mathcal{G}\right]\right)\right] \\
&\leq k^2\Bigg(\underbrace{\mathbb{E}\left[\left(|T - t_n|^2 + 2|T - t_n|\right)^2\right]}_{(\dagger)}\ \underbrace{\mathbb{E}\left[\left(\sup_{t\in[0,T\vee t_n]}\mathbb{E}\left[|a(\widehat{Y}_{\iota(t)})|^2|\mathcal{G}\right]\right)^2\right]}_{(\dagger\dagger)}\Bigg)^{\frac{1}{2}} . \qquad (22)
\end{aligned}$$

Note that we have used that $\{t_i\}_{i\geq 0}$ are measurable with respect to $\mathcal{G}$. Thanks to part (ii) of Lemmas 3.1 and 3.4 we can bound ($\dagger\dagger$) by some constant κ_7 depending on k and T only:

$$(\dagger\dagger) \leq \mathbb{E}\left[\left(\sup_{t\in[0,T]}\mathbb{E}\left[|a(\widehat{Y}_t)|^2|\mathcal{G}\right] + \max_{0\leq i\leq n}\mathbb{E}\left[|a(\widehat{Y}_{t_i})|^2|\mathcal{G}\right]\right)^2\right] \leq \kappa_7 .$$

To compute the expression in ($\dagger$) we recall that $t_n \stackrel{d}{=} \mathbf{g}(n, n/T)$. Hence, we can apply Jensen's inequality to bound the first three moments of the difference $|T - t_n|$ from above by powers of the fourth moment $\mathbb{E}[|T - \mathbf{g}(n, n/T)|^4] = 3T^4(2 + n)n^{-3}$, i.e.

$$\begin{aligned}
(\dagger) \;&=\; \mathbb{E}[|T - \mathbf{g}(n, n/T)|^4] + 4\mathbb{E}[|T - \mathbf{g}(n, n/T)|^3] + 4\mathbb{E}[|T - \mathbf{g}(n, n/T)|^2] \\
&\leq \left(\frac{3T^4(2+n)}{n^3}\right) + 4\left(\frac{3T^4(2+n)}{n^3}\right)^{3/4} + 4\left(\frac{3T^4(2+n)}{n^3}\right)^{1/2} .
\end{aligned}$$

Recall now (22) and the upper bounds for ($\dagger$) and ($\dagger\dagger$) to conclude the proof. □

Proof of Theorem 2.2. Using the decomposition of the mean square error in (9), the proof of the main result of the paper is now a mere corollary of Theorem 3.2 and Proposition 3.5. □

4 Remarks on the Euler-Poisson scheme

4.1 Enhanced Euler-Poisson scheme

The Euler-Poisson scheme has a deterministic number of iterations, but since it is supported on a random grid, it is natural to investigate if there is a more efficient way to stop the algorithm.

Recall the Poisson process $\mathbf{N}(n/T)$ defined in Section 2.2 and define $\mathbf{T}(n,T) := t_{\mathbf{N}_T+1}$, where we drop the dependence on n/T for ease of notation. Consider the Euler-Poisson scheme now stopped at the random iteration dictated by $\mathbf{N}_T + 1$, i.e. $\mathbf{T}(n,T)$ is the grid point closest to and bigger than T. In other words, this enhanced Euler-Poisson scheme considers $\widehat{Y}_{\mathbf{T}(n,T)}$ as the approximation of Y_T.

Corollary 4.1 *Under the assumptions of Theorem 2.1, there exists a constant $K_8 > 0$ such that*

$$\mathbb{E}[|Y_T - \widetilde{Y}_{\mathbf{T}(n,T)}|^2] \leq K_8\, n^{-1} \log(n)\,.$$

Proof. We first proof an analogous result to Proposition 3.5 for the random iteration $\mathbf{N}_T + 1$. From the construction of $\widehat{Y}$, recalling that $\widetilde{Y}_{\mathbf{T}(n,T)} = \widehat{Y}_{\mathbf{T}(n,T)}$, we write

$$\begin{aligned}
\mathbb{E}[|\widehat{Y}_T - \widehat{Y}_{\mathbf{T}(n,T)}|^2] &= \mathbb{E}[|a(\widehat{Y}_{\iota(T)})|^2]\mathbb{E}[|X_{\mathbf{T}(n,T)} - X_T|^2] \\
&\leq (K_0K_3 + K_0)\Big(k^2\mathbb{E}[|\mathbf{T}(n,T) - T|^2] + 2k^2\mathbb{E}[|\mathbf{T}(n,T) - T|]\Big) \\
&= (K_0K_3 + K_0)\left(k^2\frac{T^2}{n^2} + 2k^2\frac{T}{n}\right)\,, \qquad (23)
\end{aligned}$$

where the only difference with the proof of Proposition 3.5 is the fact that due to the lack of memory property $\mathbf{T}(n,T) - T \stackrel{d}{=} \mathbf{e}(n/T)$ and that we have used (11) and Lemmas 3.1 and 3.4 to bound $a(\widehat{Y}_{\iota(T)})$. To prove the claim of the result we just need to split the error $|Y_T - \widetilde{Y}_{\mathbf{T}(n,T)}|$ into a discretization error and a hitting error, as done in (9), and then use Theorem 3.2 together with (23). $\square$

Thus, this enhanced Euler-Poisson scheme is quasi-optimal. Another equivalent modification would be to use as the final point $\widetilde{\mathbf{T}}(n,T) := t_{\mathbf{N}_T}$, i.e. the closest point in the Poisson grid that is smaller than T. This modification also leads to a quasi-optimal convergence. However, unfortunately to construct either $\widehat{Y}_{\mathbf{T}(n,T)}$ or $\widehat{Y}_{\widetilde{\mathbf{T}}(n,T)}$ we need to be able to sample from the bivariate $(\Delta X_{\mathbf{e}_i}, \mathbf{e}_i)$ and not just from the resolvent of X, and thus the univariate $\Delta X_{\mathbf{e}_i}$. The Wiener-Hopf factorisation does not provide the pair $(\Delta X_{\mathbf{e}_i}, \mathbf{e}_i)$ and so far there is also no other approach. Therefore, the enhancement is of little practical relevance. Moreover, if the distribution of $(\Delta X_{\mathbf{e}_i}, \mathbf{e}_i)$ is available then the distribution of X_t is given by

$$\mathbb{P}(\Delta X_{\mathbf{e}(q)} \in \mathrm{d}x, \mathbf{e}(q) \in \mathrm{d}t) = \mathbb{P}(X_t \in \mathrm{d}x)\; qe^{-qt}\mathrm{d}t\,. \qquad (24)$$

and one might as well use the classical Euler scheme for SDEs (also known as Euler-Maruyama). The only advantage of the enhanced Euler-Poisson algorithm over Euler-Maruyama would be to avoid the Laplace transformation in (24).

4.2 Heuristics behind the Euler-Poisson scheme

The Feynman-Kac representation identifies conditional expectations of functionals of the solution of a SDE as solutions of a certain Partial Integro Differential Equation (PIDE). This section aims to formalize the relationship between the discretization procedure given by the Euler-Poisson scheme in (5) and its counterpart in the PIDE representation. We claim that, in some sense, the solution Y of (3) sampled over a random grid generated by the arrival times of a Poisson process is more natural, since it is equivalent to perform a discretization in time by the method of lines to the associated Feynman-Kac equation. We are not the first to point out this relationship. It was also the basis of Carr [4], where an approximation for American options of finite maturity is obtained by randomizing the time horizon by an Erlang distribution. Matache et al. [22] also point out informally the relation between a deterministic discretization in time of a Feynman-Kac PIDE and its probabilistic counterpart.

Theorem 4.2 (Situ [28, Section 8.17]) *Consider the following integro-differential operator*

$$\begin{aligned}\mathcal{A}_Y g(x) \;&:=\; \langle a(x)b, \nabla\rangle g(x) + \frac{1}{2}\langle a(x)\Sigma\Sigma^T a^T(x)\nabla, \nabla\rangle g(x) \\ &\quad + \int_{\mathbb{R}^{d_X}} \Big(g(x + a(x)z) - g(x) - \langle a(x)z, \nabla\rangle g(x)\Big)\Pi(\mathrm{d}z) \ ,\end{aligned}$$

taking values in $\mathcal{C}^{1,2}([0,T]\times\mathbb{R}^{d_Y},\mathbb{R})$. Let us assume that the assumptions of Theorem 2.1 hold, that

(i) $a := \mathbb{R}^{d_Y} \to \mathbb{R}^{d_Y}\otimes\mathbb{R}^{d_X}$ is bounded, and that

(ii) there exists $\delta_1, \delta_2 > 0$ such that $\delta_1|\lambda|^2 \le \langle a(x)\Sigma\Sigma^T a^T(x)\lambda, \lambda\rangle \le \delta_2|\lambda|^2$ for all $x, \lambda \in \mathbb{R}^{d_Y}$.

Let $u(t,x) \in \mathcal{C}^{1,2}([0,T]\times\mathbb{R}^{d_Y},\mathbb{R})$ be a classical solution of the PIDE

$$\frac{\partial}{\partial t}u(t,x) = \mathcal{A}_Y u(t,x) \tag{25}$$

with initial condition $u(0,x) = f(x)$, for some bounded continuous function $f : \mathbb{R}^{d_Y} \to \mathbb{R}$, i.e. $f \in \mathcal{C}_0$. Then

$$u(T-t,x) = \mathbb{E}\big[f(Y_T)\,\big|\,Y_t = x\big] = \mathbb{E}\big[f(Y_{T-t})\,\big|\,Y_0 = x\big] := \mathbb{E}_x[f(Y_{T-t})] \ , \tag{26}$$

where Y is the unique strong solution of (3) and $0 \le t \le T$.

The converse of the preceding statement also holds with appropriate assumptions. It can be written under much more general assumptions and in terms of weak solutions of the PIDE, but the simpler statement above is enough to make the point in this section. A typical setting where the above relation is exploited happens when (26) represents the price of an option under the risky asset Y which is computed by numerically solving the associated PIDE. The celebrated Black-Scholes formula is an example of this approach when the underlying process follows a geometric Brownian motion; for incomplete markets generated by Lévy processes similar formulas hold (cf. Chan [5]).

Recall the random times $\{t_i\}_{i\ge 0}$ defined in Section 2.2 as the arrival times of a Poisson process $\mathbf{N}$, and consider the Laplace-Carlson transform, $\mathcal{L}$, of $u(t,x)$, that is

$$\mathcal{L}[u](x) = \int_0^\infty \frac{n}{T}\exp\Big(-\frac{nt}{T}\Big)u(t,x)\mathrm{d}t = \int_0^\infty \frac{n}{T}\exp\Big(-\frac{nt}{T}\Big)\mathbb{E}_x[f(Y_t)]\mathrm{d}t = \mathbb{E}_x\Big[f\big(Y_{\mathbf{e}(\frac{n}{T})}\big)\Big] = \mathbb{E}_x[f(Y_{t_1})], \tag{27}$$

where we have used the boundedness of $f \in \mathcal{C}_0$ to apply Fubini's theorem. Note that the last term in the above equation corresponds to the expectation of the solution in (3) at the first arrival time of the Poisson process $\mathbf{N}$. Moreover, due to the boundedness of f we can also interchange the differential operator $\mathcal{A}_Y$ and the transform $\mathcal{L}$ to obtain the integro-differential equation satisfied by the Laplace-Carlson transform:

$$\frac{\mathcal{L}[u](x) - f(x)}{T/n} = \mathcal{A}_Y\mathcal{L}[u](x)\,, \tag{28}$$

which contains a difference instead of the differential $\frac{\partial}{\partial t}$ in (25). Due to the homogeneity of $\mathcal{A}_Y$, this turns out to be of the same form as the first order finite difference approximation in time of (25) with respect to $\mathcal{L}[u]$ instead of u. To fix ideas, the following proposition explicitly relates the solution Y at the arrival times of $\mathbf{N}$ with the iterates of what is known in the literature as the method of lines or Rothe's method for PIDEs.

Proposition 4.3 *Under the assumptions of Theorems 2.1 and 4.2, consider Rothe's method for (25), given by*

$$\frac{u_i(x) - u_{i-1}(x)}{T/n} = \mathcal{A}_Y u_i(x) \ , \tag{29}$$

for $i = 1,\dots,n$ with $u_0(x) = f(x)$. Then, for all $i = 1,\dots,n$,

$$u_i(x) = \mathbb{E}_x[f(Y_{t_i})] \ .$$

Proof. It is clear that the solution of (3) given by Theorem 2.1 has the strong Markov property (cf. Protter [24, Theorem 32 p. 294]). Therefore, we write

$$\mathbb{E}_x[f(Y_{t_i})] = \mathbb{E}_x[\mathbb{E}_{Y_{t_1}}[\mathbb{E}_{Y_{t_2}}[\cdots \mathbb{E}_{Y_{t_{i-1}}}[f(Y_{t_i})]\cdots]]] \ ,$$

and apply recursively the arguments derived from (27) and (28) in the above nested expectations to obtain the recursive solutions that solve the system of differential equations in (29). $\square$

4.3 Pathwise convergence

The Euler-Poisson scheme is supported on a random grid and there is no straightforward way to perform a pathwise numerical analysis of the algorithm. Nevertheless the above analogy with Rothe's method suggests that one may try to study the behavior of

$$\mathbb{E}\left[\max_{1\le i\le n}|Y_{iT/n}-\widetilde{Y}_{t_i}|^2\right] = \mathbb{E}\left[\max_{1\le i\le n}|Y_{iT/n}-\widehat{Y}_{t_i}|^2\right] \ .$$

Indeed, Theorem 3.2 states a pathwise result for the discretization error and hence, using the decomposition in (9), one would only need to obtain a pathwise analogue of the hitting error in order to study the above quantity, i.e. a pathwise generalization of Proposition 3.5. Unfortunately the latter is not true. A weaker statement that can be proved and involves the entire path of the Euler-Poisson scheme is

$$\max_{1\le i\le n}\mathbb{E}\left[|Y_{iT/n}-\widetilde{Y}_{t_i}|^2\right] \le K_9\, n^{-1/2}\,, \tag{30}$$

where $K_9 > 0$ depends only on k and T. The result in (30) is a direct consequence of Theorem 3.2 and Proposition 3.5 together with the following observation (derivable from Doob [10, Theorem 5.1])

$$\mathbb{E}\left[\max_{1\le i\le n}\left|t_i-\frac{Ti}{n}\right|^p\right] \le 8\,\mathbb{E}\Big[\,|t_n-T|^p\,\Big], \qquad \text{for} \quad p\ge 1 \ .$$

A Appendix – Moments of τ

Let $m \in \mathbb{N}$. If a Poisson process $\mathbf{N}$ has m arrivals up to time T, then those m arrival times have the same distribution as m ordered independent uniform random variables on $[0, T]$. Therefore, in order to study the random variable τ defined in (8), we can start by studying the largest partition on the interval $[0, 1]$ defined by m independent uniform random variables in $[0, 1]$.

Let $\{U_i\}_{i=1,\dots,m-1}$ be a sequence of i.i.d. random variables with common uniform distribution in $[0, 1]$ and consider its order statistics $U_{(i)}$, for $i = 0, \dots, m$, where $U_0 = 0$ and $U_m = 1$. Denote the largest gap by

$$\lambda_m := \max_{i=1\dots,m}\{U_{(i)}-U_{(i-1)}\} \ .$$

Recall the definition of τ in (8). The conditional distribution of τ is, up to a constant, equal to λ. Indeed, $\frac{1}{T}\tau$, conditioned on $\mathbf{N}_T$, is equal in distribution to $\lambda_{\mathbf{N}_T+1}$. In particular we have

$$\frac{1}{T}\mathbb{E}[\tau] = \mathbb{E}[\lambda_{\mathbf{N}_T+1}] \ . \tag{31}$$

Fisher [13] already studied the behaviour of λ_m and the following expression is given in Mauldon [23]:

$$\mathbb{E}[(1-\lambda_m s)^{-m}] = \frac{m!}{1-s}\prod_{j=2}^{m}\frac{1}{j-s} \qquad |s|<1/2\ , m\ge 1 \ .$$

All moments of λ_m can be expanded form the above expression and in particular, for $m \ge 1$, we have

$$m\mathbb{E}[\lambda_m] = \sum_{j=1}^{m}\frac{1}{j} = \Psi(m+1)+\gamma\,,$$

where Ψ is the digamma function (see Abramowitz and Stegun [1, Sections 6.3.2 and 6.4.10]). Recall that the function $\Psi(m+1)+\gamma$ is zero for $m=0$, positive for $m>0$ and grows asymptotically as $\log(m+1)$, i.e. $\lim_{m\to\infty}\Psi(m)/\log(m)=1$. Therefore, there is a constant $\kappa_0>0$ independent of m such that $\Psi(m+1)+\gamma\le\kappa_0\log(m+1)$ for $m\ge1$. Hence,

$$m\mathbb{E}[\lambda_m]\le\kappa_0\log(m+1),\qquad \text{for}\quad m\ge1.$$

Proposition A.1

$$\mathbb{E}[\tau]+\mathbb{E}[\tau^2]\le K_A\,n^{-1}\log(n)\,.$$

Proof. According to (31) and recalling that the arrival rate for $\mathbf{N}_T$ is n/T, we have

$$\begin{aligned}\frac{1}{T}\mathbb{E}[\tau]=\mathbb{E}[\lambda_{\mathbf{N}_T+1}] &= \sum_{k=0}^{\infty}\mathbb{E}[\lambda_{\mathbf{N}_T+1}|\mathbf{N}_T=k]\mathbb{P}(\mathbf{N}_T=k)\\ &\le \kappa_0\sum_{k=0}^{\infty}\frac{\log(k+2)}{k+1}\exp\left(-\frac{n}{T}\right)\frac{\left(\frac{n}{T}\right)^k}{k!}=\kappa_0\frac{T}{n}\sum_{k=1}^{\infty}\log(k+1)\exp\left(-\frac{n}{T}\right)\frac{\left(\frac{n}{T}\right)^k}{k!}\\ &= \frac{\kappa_0 T}{n}\mathbb{E}[\log(\mathbf{N}_T+1)]\le\frac{\kappa_0 T}{n}\log(\mathbb{E}[\mathbf{N}_T]+1)=\frac{\kappa_0 T\log(n+1)}{n}\end{aligned}$$

where the last inequality is due to Jensen's inequality and the concavity of $x\to\ln(x+1)$ for $x\ge0$. To derive the claim of the proposition we can use the crude upper bound $\lambda_m^2\le\lambda_m$, since $\lambda_m\in[0,1]$, and hence

$$\frac{1}{T^2}\mathbb{E}[\tau^2]=\mathbb{E}[\lambda^2_{\mathbf{N}_T+1}]\le\mathbb{E}[\lambda_{\mathbf{N}_T+1}]\,.$$

$\square$

References

[1] Abramowitz, M., Stegun, I. A. (1970) *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* Dover Publications, Inc., New York.

[2] Asmussen, S. and Rosiński, J. (2001) Approximations of small jumps of Lévy processes with a view towards simulation. *J. Appl. Prob.* **38**(2), 482–493.

[3] Baran, M. (2009) Approximation for solutions of Lévy-type stochastic differential equations. *Stoch. Anal. Appl.* **27**(5), 924–961.

[4] Carr, P. (1998) Randomization and the American Put. *Rev. Fin. Studies.* **11**(3), 597–626.

[5] Chan, T. (1999) Pricing contingent claims on stocks driven by Lévy processes. *Ann. Appl. Probab.* **9**(2), 504–528.

[6] Cont, R. and Tankov, P. (2004) *Financial Modeling with Jump Processes.* Chapman & Hall/CRC, Boca Raton, FL.

[7] Corcuera, J., de Spiegeleer, J., Ferreiro-Castilla, A., Kyprianou, A.E., Madan, D. and Schoutens, W. (2013) Efficient Pricing of Contingent Convertibles under Smile Conform Models. *Journal of Credit Risk* **9**(3), 121–140.

[8] Dereich, S. (2011) Multilevel Monte Carlo Algorithms for Lévy–driven SDEs with Gaussian correction. *Ann. Appl. Probab.* **21**(1), 283–311.

[9] Dereich, S. and Heidenreich, F. (2011) A multilevel Monte Carlo algorithm for Lévy–driven stochastic differential equations. *Stochastic Process. Appl.* **121**(7), 1565–1587.

[10] Doob, J.L. (1953) *Stochastic Processes.* John Wiley & Sons, Inc., New York.

[11] Ferreiro-Castilla, A., Kyprianou, A.E., Scheichl, R. and Suryanarayana, G. (2014) Multilevel Monte Carlo simulation for Lévy processes based on the Wiener-Hopf factorization. *Stochastic Process. Appl.* **124**(2), 985–1010.

[12] Ferreiro-Castilla, A. and Schoutens, W. (2012) The $\boldsymbol{\beta}$-Meixner model. *J. Comput. Appl. Math.* **236**(9), 2466–2476.

[13] Fisher, R. A. (1929) Tests of Significance in Harmonic Analysis. *Proceedings of the Royal Society of London. Series A.* **125**(796), 54–59.

[14] Gīhman, Ĭ. Ī. and Skorohod, A. V. (1972) *Stochastic differential equations.* Springer, New York.

[15] Jacod, J., Kurtz, T.G., Méléard, S. and Protter, P. (2005) The approximate Euler method for Lévy driven stochastic differential equations. *Ann. Inst. H. Poincaré Probab. Statist.* **41**, 523–558.

[16] Kallsen, J. and Tankov, P. (2006) Characterization of dependence of multidimensional Lévy processes using Lévy copulas. *J. Multivariate Anal.* **97**(7), 1551–1572.

[17] Kloeden, P.E. and Platen, E. (1992) *Numerical solution of stochastic differential equations.* Springer, Berlin.

[18] Kohatsu-Higa, A., Ortiz-Latorre, S. and Tankov, P. (2014) Optimal simulation scheme for Lévy driven stochastic differential equations. *Math. Comput.* **83**, 2293–2324.

[19] Kuznetsov (2010) Wiener-Hopf factorization and distribution of extrema for a family of Levy processes *Ann. Appl. Probab.* **20**, 1801–1830.

[20] Kuznetsov, A., Kyprianou, A. E. and Pardo, J. C. (2012) Meromorphic Lévy processes and their fluctuation identities. *Ann. Appl. Probab.* **22**(3), 1101–1135.

[21] Kuznetsov, A., Kyprianou, A.E., Pardo, J.C. and van Schaik, K. (2011) A Wiener-Hopf Monte Carlo simulation technique for Lévy process. *Ann. Appl. Probab.* **21**(6), 2171–2190.

[22] Matache, A.M., Nitsche, P.A. and Schwab, C. (2005) Wavelet Galerkin pricing of American options on Lévy driven assets. *Quantitative Finance*, **5**(4), 403–424.

[23] Mauldon, J. G. (1951) Random division of an interval. *Proc. Cambr. Philos. Soc.*, **47**(2), 331–336.

[24] Protter, P.E. (2005) *Stochastic integration and differential equations.* Springer-Verlag, Berlin.

[25] Protter, P.E. and Talay (1997) The Euler scheme for Lévy driven stochastic differential equations *Ann. Probab.*, **25**(1), 393–423.

[26] Rubenthaler, S. (2003) Numerical simulation of the solution of a stochastic differential equation driven by a Lévy process. *Stochastic Process. Appl.*, **103**(2), 311–349.

[27] Schoutens, W. and van Damme, G. (2010) The $\boldsymbol{\beta}$-variance gamma model. *Rev. Deriv. Res.* **14**(3), 263–282.

[28] Situ, R. (2005) *Theory of stochastic differential equations with jumps and applications. Mathematical and analytical techniques with applications to engineering.* Springer, New York.

[29] Sobczyk, K. (1991) *Stochastic differential equations: With applications to physics and engineering.* Kluwer Academic Publishers Group, Dordrecht.

[30] Talay, D. and Tubaro, L. (1990) Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Anal. Appl.* **8**(4), 483–509.