Weak Approximations for SDE's Driven by Lévy Processes

Arturo Kohatsu-Higa and Hoang-Long Ngo

Abstract. In this article we briefly survey recent advances in some simulation methods for Lévy driven stochastic differential equations. We give a brief description of each method and extend the one jump scheme method for some subordinated models like the NIG process. Simulations of all the presented methods are performed and compared.

Mathematics Subject Classification (2010). Primary 65Cxx, 60J05, 65xx; Secondary 65C30, 65C05.

Keywords. SDE's, Lévy process, numerical approximation, simulation.

1. Introduction

Let X be an $\mathbb{R}^N$-valued adapted stochastic process, unique solution of the stochastic differential equation (SDE) with jumps

$$X_t(x) = x + \int_0^t \tilde{V}_0(X_s(x))ds + \int_0^t V(X_s(x))dB_s + \int_0^t h(X_{s-}(x))dZ_s,\ t \in [0,1], \tag{1.1}$$

with smooth coefficients $\tilde{V}_0 : \mathbb{R}^N \to \mathbb{R}^N$, $V = (V_i^{(j)})_{j=1,\dots,N}^{i=1,\dots,d} : \mathbb{R}^N \to \mathbb{R}^N \otimes \mathbb{R}^d$, $h : \mathbb{R}^N \to \mathbb{R}^N \otimes \mathbb{R}^d$ whose derivatives of any order (≥ 1) are bounded. Here B denotes an d-dimensional standard Brownian motion and Z denotes an d-dimensional Lévy process with Lévy triplet $(\gamma, 0, \nu)$ such that all of its moments are finite unless stated otherwise.

In this report, we numerically compare and evaluate two types of discrete approximation schemes for X in order to estimate $\mathbb{E}[f(X_1)]$ for smooth functions f. More precisely, we find a discretization scheme $(X_{t_j}^{(n)}(x))_{j=0}^n$ for a partition $0 = t_0 < t_1 < \ldots < t_n = 1$ such that

$$|\mathbb{E}[f(X_1(x))] - \mathbb{E}[f(X_1^{(n)}(x))]| \leq \frac{C(f,x)}{n^m}, \tag{1.2}$$

This research was supported by a grant of JST.

for some $m \in \mathbb{N}$ and a positive constant $C(f,x)$.

In such a case, we say that such a scheme is an m-th order discretization scheme for X. The actual simulation to estimate $\mathbb{E}[f(X_1(x))]$ is carried out using Monte Carlo methods. That is, one computes $\frac{1}{N_{MC}} \sum_{i=1}^{N_{MC}} f(X_1^{(n),i}(x))$ where $X_1^{(n),i}(x), i = 1, \ldots, N_{MC}$ denotes N_{MC} i.i.d copies of $X_1^{(n)}(x)$. Therefore, using the law of large numbers, the final error of the estimate is of the order $O(\frac{1}{\sqrt{N_{MC}}} + \frac{1}{n^m})$. Then the optimal choice of n is $O(n^m) = O(\sqrt{N_{MC}})$.

From this result, we can see that there is a reduction in computation time if one can obtain a scheme with a high value of m even if the computational cost increases linearly with m. In this light, we want to address in this article, the performance of some competing approximation schemes for jump driven sde's of the type (1.1) in the infinite activity case (i.e. $\nu(\mathbb{R}^d) = \infty$) as it is in the case in many financial models.

The first simulation proposal is to simulate all the jump times and their corresponding jump sizes in the case that Z is a finite activity process (i.e. $\nu(\mathbb{R}^d) < \infty$) (see e.g. [3]). This becomes impossible in the infinite activity case as the number of jumps in any interval is infinite a.s. and therefore instead one may simulate all jump times for jump sizes bigger than a fixed small parameter ε. Following a proposal by Asmussen and Rosiński [2], the small jumps are replaced by an independent Brownian motion with variance given by $\int_{|y| \le \varepsilon} |y|^2 \nu(dy)$. It has been shown in [2] and [6] that when approximating the small jumps by Gaussian variables, the convergence rates, which are measured by either the Kolmogorov distance between laws of processes at a fixed time or the mean square of the supremum of the error during a finite and fixed interval, are significantly improved (see also [12]). In our context of weak approximation (1.2), we would also like that the approximation of small jumps should be accurate. When drift and/or continuous diffusion components appear in the stochastic differential equation then one naturally faces an optimality problem. That is, how to match the computational effort done on the jump part with efficient approximation schemes for the drift and the Brownian part of the equation between jump times.

This issue was addressed in the article [8]. The method introduced in that article will be one of the methods that we will use in our comparison. In that method one uses all the jumps of size bigger than ε, the jump times become time partition points and one approximates the effect of the drift up to a high order of accuracy and there is no continuous diffusion part. In [8] it is proven that the rate of convergence is fast but the calculation time may be long.

On another article [14], the authors take a different point of view. Instead of using a random time partition points given by the jumps times corresponding to jump sizes bigger than ε, a fixed time partition is used and an approximation for the increments of the Lévy process is used. In this approximation, one uses an approximation with at most a finite number of jumps per interval. The maximum number of jumps is set by the user and therefore this becomes a limitation on the computation time by the part of the

user. This approach which, in some sense, goes in the reverse direction of the scheme in [8] assumes that an approximation for the drift and the continuous diffusion part have been set and one tries to find a simple approximation of the increment of the Lévy process so as to match the computational effort invested in the Brownian and drift part of the equation. In order to introduce this method, one needs to explain the framework of the operator splitting method in its stochastic form.

This method is well known as a numerical method for partial differential equations. The idea is to use it, finding stochastic representations for the approximating splitting therefore providing new simulations methods based on composition of flows which parallel the composition of semigroups. This idea has been successfully used for stochastic differential equations driven by Brownian motion (see [9] and [10]).

Nevertheless, it should be noted that the performance of every estimation scheme depends on the activity level of small jump of the driving Lévy process Z, which is measured by the Blumenthal-Getoor index

$$\varrho = \inf\{p \geq 0 : \int_{\|x\| \leq 1} \|x\|^p \nu(dx) < \infty\}.$$

Since the Lévy measure ν satisfying $\int_{\|x\| \leq 1} \|x\|^2 \nu(dx) < \infty$, the index $\varrho \in [0, 2]$.

The goal of the present article is to give a non-technical introduction to these schemes and to present a throughout simulation study in order to assess the properties of the approximation schemes described above. Therefore we refer the reader for the proofs to the corresponding articles and we only give here the intuition behind the schemes.

In order to give the reader an idea of what are the technical conditions that need to be satisfied to obtain a new scheme, in Section 2.4.5, we deal with one case that was not treated in [14]. The case we study corresponds to a normal subordinated model. In this case, we have that $\int |y| \nu(dy) = \infty$. We will verify the main two conditions needed in order to establish the weak rate for the approximation method which follows from the main theorems 4.1, 4.3 and 5.1 in [14].

2. Approximation Schemes

In this section, we define the approximation schemes for equation (1.1) which we will compare in this paper. For proofs we refer the reader to the corresponding theoretical articles. We strive here for understanding and intuition of these schemes.

2.1. Euler's Scheme

The Euler scheme is the most natural approximation scheme. Its programming flow is as follows. We denote $t_i^n = i/n, i = 1, \ldots, n$.

1. Generate a sequence of independent random variables ΔZ_i^n, $i = 0, \dots, n-1$, which have the same distribution as $Z_{1/n}$.
2. Generate a sequence of independent random variables ΔB_i^n, $i = 0, \dots, n-1$, which have the same distribution as $B_{1/n}$.
3. $\bar{X}_0 = x$ and for $i = 0, \dots, n-1$,

$$\bar{X}_{(i+1)/n} = \bar{X}_{i/n} + \frac{1}{n}\tilde{V}_0(\bar{X}_{i/n}(x)) + V(\bar{X}_{i/n}(x))\Delta B_i^n + h(\bar{X}_{i/n}(x))\Delta Z_i^n.$$

Various articles and results have been written on this scheme. The main problem with this scheme is that it assumes that one can simulate the Brownian increment and the Lévy increment with the same computational effort. This is hardly the case in general, as the law of Lévy processes is generally given through their characteristic function. Therefore in general, an inversion procedure is needed. For more on this direction, see [7].

This simulation scheme is an approximation scheme of order 1 under sufficient conditions on the Lévy measure and it has been proven in e.g. [11].

2.2. Jump-Size Adapted Discretization Schemes

The purpose of this section is to introduce a simulation method which uses all the jumps associated with the Lévy process whose norm are bigger than a certain fixed value ε. As the number of this type of jumps is finite on finite intervals then this approximation process defines a compound Poisson process. Therefore its simulation may be possible if we assume that the jump distribution can be simulated. The main drawback of the method is that it may take long time to compute as ε becomes small. On the other hand, it is a very accurate method. For further details, we refer the reader to [8].

To introduce the method, suppose that $V = \tilde{V}_0 = 0$ and Z is a d-dimensional Lévy process without diffusion component. That is,

$$Z_t = \gamma t + \int_0^t \int_{|y|\le 1} y\hat{N}(dy, ds) + \int_0^t \int_{|y|>1} yN(dy, ds), \quad t \in [0, 1].$$

Here, $\gamma \in \mathbb{R}^d$ and N is a Poisson random measure on $\mathbb{R}^d \times [0, \infty]$ with intensity ν satisfying $\int (1 \wedge |y|^2)\nu(dy) < \infty$. $\hat{N}(dy, ds) = N(dy, ds) - \nu(dy)ds$ denotes the compensated version of N.

Consider a family of measurable functions $(\chi_\varepsilon)_{\varepsilon>0} : \mathbb{R}^d \to [0, 1]$ such that $\int_{\mathbb{R}^d} \chi_\varepsilon(y)\nu(dy) < \infty$ for all $\varepsilon > 0$, and $\lim_{\varepsilon\to 0} \chi_\varepsilon(y) = 0$, for all $y \neq 0$. This function will serve as the localization function for the jumps which will be simulated. Therefore, unless explicitly mentioned otherwise, we will usually take $\chi_\varepsilon(y) = 1(|y| > \varepsilon)$.

We assume that the associated Lévy measure ν satisfies that

$$\nu(\mathbb{R}^d) = \infty, \quad \int_{\mathbb{R}^d} |y|^2 \nu(dy) < \infty.$$

Let N_ε be a Poisson random measure with intensity $\chi_\varepsilon \nu \times ds$ and $\hat{N}_\varepsilon$ its compensated Poisson random measure. Denote $\overline{\hat{N}}_\varepsilon$ a compensated Poisson random measure with intensity $\overline{\chi}_\varepsilon \nu \times ds$, where $\overline{\chi}_\varepsilon = 1 - \chi_\varepsilon$.

The processes Z can then be represented in law as follows

$$\begin{aligned}
Z_t &\stackrel{d}{=} \gamma_\varepsilon t + Z_t^\varepsilon + R_t^\varepsilon,\\
\gamma_\varepsilon &= \gamma - \int_{|y|\le 1} y\chi_\varepsilon \nu(dy) + \int_{|y|>1} y\overline{\chi}_\varepsilon \nu(dy),\\
Z_t^\varepsilon &= \int_0^t \int_{\mathbb{R}^d} y N_\varepsilon(dy, ds),\\
R_t^\varepsilon &= \int_0^t \int_{\mathbb{R}^d} y \hat{\overline{N}}_\varepsilon(dy, ds).
\end{aligned}$$

We denote by $\lambda_\varepsilon = \int_{\mathbb{R}^d} \chi_\varepsilon(y)\nu(dy)$ the intensity of Z^ε, by $T_i^\varepsilon, i \in \mathbb{N}$ the i-th jump time of Z^ε with $T_0^\varepsilon = 0$, and by $\Sigma_\varepsilon = \Big(\int_{\mathbb{R}^d} y_i y_j \overline{\chi}_\varepsilon(y)\nu(dy)\Big)_{1\le i,j\le d}$ the covariance matrix of R_1^ε. In the one-dimensional case, $d = 1$, we set $\sigma_\varepsilon^2 = (\Sigma_\varepsilon)_{11}$. Given $\varepsilon > 0$ and Lévy measure ν, one can compute $\lambda_\varepsilon, \Sigma_\varepsilon$ and generate the sequence (T_i^ε). The random variable R_1^ε will be approximated using a Gaussian random variable with mean zero and variance Σ_ε. This is the so-called Asmussen-Rosiński approximation.

2.2.1. Kohatsu-Tankov Scheme in Dimension One. The following one dimensional scheme ($d = N = 1$) uses an explicit transformation between jump times in order to solve explicitly the ODE.

$$dX_t = h(X_t)dt, \quad X_0 = x.$$

Suppose that $1/h$ is locally integrable, this equation has a solution

$$X_t = \theta(t; x) = F^{-1}(t + F(x)),$$

where F is a primitive of $1/h$.

We define inductively $\hat{X}(0) = X_0$ and for $i \ge 0$,

$$\begin{aligned}
\hat{X}(T_{i+1}^\varepsilon -) = \theta\Big(&\gamma_\varepsilon(T_{i+1}^\varepsilon - T_i^\varepsilon) + \sigma_\varepsilon(W(T_{i+1}^\varepsilon) - W(T_i^\varepsilon))\\
&- \frac{1}{2}h'(\hat{X}(T_i^\varepsilon))\sigma_\varepsilon^2(T_{i+1}^\varepsilon - T_i^\varepsilon); \hat{X}(T_i^\varepsilon)\Big), \qquad (2.1)\\
\hat{X}(T_{i+1}^\varepsilon) = \hat{X}(T_{i+1}^\varepsilon -) &+ h(\hat{X}(T_{i+1}^\varepsilon -))\Delta Z(T_{i+1}^\varepsilon). \qquad (2.2)
\end{aligned}$$

For an arbitrary point t, we define

$$\hat{X}(t) = \theta\Big(\gamma_\varepsilon(t - \eta_t) + \sigma_\varepsilon(W(t) - W(\eta_t)) - \frac{1}{2}h'(\hat{X}(\eta_t))\sigma_\varepsilon^2(t - \eta_t); \hat{X}(\eta_t)\Big), \qquad (2.3)$$

where $\eta_t = \sup\{T_i^\varepsilon : T_i^\varepsilon \le t\}$.

The logic behind the above scheme should be clear. Between jumps we use a high order approximation to the solution of the stochastic differential equation driven by the drift coefficient γ_ε and the Wiener process W which replaces the small jumps (i.e. Asmussen-Rosiński approximation). When a jump happens the corresponding jump is added to the system.

The rate of convergence of this scheme, under the condition $\int |y|^6 \nu(dy) < \infty$ is given by

$$\left|\mathbb{E}(f(\hat{X}_1) - f(X_1))\right| \leq C\Big(\frac{\sigma_\varepsilon^2}{\lambda_\varepsilon}(\sigma_\varepsilon^2 + |\gamma_\varepsilon|) + \int_{\mathbb{R}} |y|^3 \bar{\chi}_\varepsilon \nu(dy)\Big),$$

for some constant $C > 0$ not depend on ε (see [8, Theorem 2]).

The above scheme can be applied when F can be computed explicitly. Otherwise, one has to resort to approximations for F and then the order of the approximation becomes an important issue. See [8] for more comments on this matter.

2.2.2. Kohatsu-Tankov Scheme in Higher Dimension. This scheme uses instead a Taylor expansion between jumps as the respective stochastic differential equation between jumps can not be solved explicitly. We denote

$$\begin{aligned}
&\tilde{X}(t) = Y^0(t) + Y_1(t), \quad t > \eta_t,\\
&\tilde{X}(T_i^\varepsilon) = \tilde{X}(T_i^\varepsilon -) + h(\tilde{X}(T_i^\varepsilon -))\Delta Z(T_i^\varepsilon),\\
&Y^0(t) = \tilde{X}(\eta_t) + \int_{\eta_t}^t h(Y^0(t))\gamma_\varepsilon ds,\\
&Y_1(t) = \sum_{i=1}^N \int_{\eta_t}^t \frac{\partial h}{\partial x_i}(Y^0(s))Y_1^i(s)\gamma_\varepsilon ds + \int_{\eta_t}^t h(Y^0(s))dW^\varepsilon(s),
\end{aligned}$$

where

- W^ε is a d-dimensional Brownian motion with covariance matrix Σ_ε independent of Z;
- the random vector $Y_1(t)$ conditioned on $T_i^\varepsilon,\ i \in \mathbb{N},\ t \in (T_j^\varepsilon, T_{j+1}^\varepsilon)$ and $\tilde{X}(T_j^\varepsilon)$ is a Gaussian random vector with conditional covariance matrix $\Omega(t)$ which satisfies the (matrix) linear equation

$$\Omega(t) = \int_{\eta_t}^t (\Omega(s)M(s) + M^\perp(s)\Omega^\perp(s) + N(s))ds,$$

where $M^\perp$ denotes the transpose of the matrix M and

$$M_{ij}(t) = \frac{\partial h_{jk}(Y^0(t))}{\partial x_i}\gamma_\varepsilon^k \quad \text{and} \quad N(t) = h(Y^0(t))\Sigma_\varepsilon h^\perp(Y^0(t)).$$

The rate of convergence of the above scheme under the condition $\int |y|^6 \nu(dy) < \infty$ is given by

$$\left|\mathbb{E}(f(\tilde{X}_1) - f(X_1))\right| \leq C\Big(\frac{\|\Sigma_\varepsilon\|}{\lambda_\varepsilon}(\|\Sigma_\varepsilon\| + |\gamma_\varepsilon|) + \int_{\mathbb{R}} |y|^3 \bar{\chi}_\varepsilon \nu(dy)\Big),$$

for some constant $C > 0$ which does not depend on ε (see Theorem 16 [8]).

2.3. Operator Splitting Schemes

We define $V_0 = \tilde{V}_0 - \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{N}\frac{\partial V_i}{\partial x_j}V_i^{(j)}$. Then Equation (1.1) can be rewritten in the following Stratonovich form

$$X_t(x) = x + \sum_{i=0}^{d}\int_0^t V_i(X_{s-}(x)) \circ dB_s^i + \int_0^t h(X_{s-}(x))dZ_s, \tag{2.4}$$

where $B_t^0 = t$. We define the semigroup P_t by

$$P_t f(x) = \mathbb{E}[f(X_t(x))],$$

where $f : \mathbb{R}^N \to \mathbb{R}$ is a continuous smooth function with polynomial growth at infinity.

We will approximate $P_t f(x) = \mathbb{E}[f(X_t(x))]$ by using its Taylor expansion for small $t > 0$. We will first compute, using Itô's formula

$$\frac{P_h f(x) - f(x)}{h}.$$

For this, note that Itô's formula gives

$$\begin{aligned} f(X_h(x)) - f(x) &= \int_0^h \nabla f(X_s(x))dX_s^c(x) \\ &+ \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{N}\int_0^h D_{ij}^2 f(X_s(x))d\langle X^i(x), X^j(x)\rangle_s \\ &+ \sum_{s\le h}\{f(X_{s-}(x) + h(X_{s-}(x))\Delta Z_s) \\ &\qquad - f(X_{s-}(x)) - \nabla f(X_{s-}(x))h(X_{s-}(x))\Delta Z_s\}. \end{aligned}$$

After taking expectations and limits we obtain:

$$\lim_{h\to 0}\frac{P_h f(x) - f(x)}{h} = Lf(x) = \sum_{k=0}^{d+1} L_i f(x), \tag{2.5}$$

where

$$\begin{aligned} L_0 f(x) &\equiv \tilde{V}_0 f(x) = \sum_{k=1}^{N}\frac{\partial f}{\partial x_k}(x)\tilde{V}_0^{(k)}(x), \\ L_i f(x) &= \frac{1}{2}V_i^2 f(x) = \frac{1}{2}\sum_{j,k=1}^{N}\frac{\partial^2 f(x)}{\partial x_j \partial x_k}V_i^{(j)}V_i^{(k)}(x), \quad i = 1,\dots,d, \\ L_{d+1} f(x) &= \nabla f(x)h(x)\gamma + \int (f(x + h(x)y) - f(x) - \nabla f(x)h(x)y)\nu(dy). \end{aligned} \tag{2.6}$$

From the above calculation one clearly understands that the operator L_0 is associated to the drift of Equation (2.4), L_i for $i = 1, ..., d$ is associated to

the i-th Brownian motion and L_{d+1} is associated to the Lévy process. Note also that

$$\begin{aligned}\frac{\partial P_t f(x)}{\partial t} &= LP_t f(x)\\ \frac{\partial^k P_t f(x)}{\partial t^k} &= L^k P_t f(x).\end{aligned}$$

Operator L is called the generator of P and this fact is usually written as $P_t = e^{tL}$. Due to the semigroup property of P, say $P_{t_1+t_2} = P_{t_1}P_{t_2}$, one understands that in order to approximate X one needs only to approximate P_t for small values of t and then use the following composition property

IMPORTANT PROPERTY: Let Y^1 and Y^2 be two independent stochastic processes generating semigroups R^1 and R^2 and with generators K^1 and K^2 respectively, then

$$\mathbb{E}[f(Y_t^1(Y_t^2(x)))] = R_t^2 R_t^1 f(x) = e^{tK^2} e^{tK^1} f(x).$$

Note that the operators above are not in general commutative.

In fact, if we iterate the above arguments we have that for a smooth function f,

$$P_t f(x) = f(x) + tLf(x) + \frac{t^2}{2}L^2 f(x) + ... = e^{tL} f(x).$$

Example. In this example, we retake the case of the Euler scheme in Section 2.1 and analyze it in the light of the previous argument.

Now let Q be the "semigroup" associated to the Euler scheme. That is, define $Q_t f(x) = E[f(\bar{X}_t)]$ for $t \leq \frac{1}{n}$. Then one can obtain the following expansion

$$Q_t f(x) = f(x) + t\bar{L}_1 f(x) + \frac{t^2}{2}\bar{L}_2 f(x) + ...$$

In fact, let $h \leq 1/n$ then

$$\begin{aligned}f(X_h^n(x)) - f(x) = &\int_0^h \nabla f(X_s^n(x)) dX_s^{n,c}(x)\\ &+ \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{N}\int_0^h D_{ij}^2 f(X_s^n(x)) d\langle X^{n,i}(x), X^{n,j}(x)\rangle_s\\ &+ \sum_{s\leq h}\left\{f(X_{s-}^n(x) + h(X_{s-}^n(x))\Delta Y_s) - f(X_{s-}^n(x))\right.\\ &\qquad\left. - \nabla f(X_{s-}^n(x)) h(X_{s-}^n(x))\Delta Y_s\right\}.\end{aligned}$$

After some calculation one obtains that $\bar{L}_1 = L$ and that $\bar{L}_2 \neq L_2$. Therefore one has that the local error $P_t f(x) - Q_t f(x) = O(t^2)$. The proof

finishes by using the following telescoping decomposition

$$\begin{aligned} E[f(X_1(x))] - E[f(\bar{X}_1(x))] &= -\sum_{i=1}^{n} \left\{ Q_{1/n}^{i} P_{1-t_i} f(x) - Q_{1/n}^{i-1} P_{1-t_{i-1}} f(x) \right\} \\ &= \sum_{i=1}^{n} \left\{ Q_{1/n}^{i-1} Q_{1/n} P_{1-t_i} f(x) - Q_{1/n}^{i-1} P_{1/n} P_{1-t_i} f(x) \right\} \\ &= \sum_{i=1}^{n} \left\{ Q_{1/n}^{i-1} \Big(Q_{1/n} - P_{1/n} \Big) P_{1-t_i} f(x) \right\}. \end{aligned}$$

Remark 2.1. One also needs the "stability" property of the operator Q in order to finish the argument above. That is, we need two properties: (1) the different $Q_{1/n} f - P_{1/n} f$ is of order $O(n^{-2})$ under certain regularity conditions on f (e.g. $f \in C_p^3$); and (2) the iteration $Q_{1/n}^{i-1}$ preserves the error rate of $Q_{1/n} - P_{1/n}$ without demanding any further regularity of $(Q_{1/n} - P_{1/n}) P_{1-t_i} f(x)$.

Next, we define the following stochastic processes $X_{i,t}(x),\ i = 0, \ldots, d+1$, usually called coordinate processes, which will correspond to the operator decomposition in (2.5) and which are the unique solutions of

$$\begin{aligned} X_{0,t}(x) &= x + \int_0^t V_0(X_{0,s}(x)) ds, \\ X_{i,t}(x) &= x + \int_0^t V_i(X_{i,s}(x)) \circ dB_s^i, \quad 1 \le i \le d, \\ X_{d+1,t}(x) &= x + \int_0^t h(X_{d+1,s-}(x)) dZ_s. \end{aligned}$$

Then we define

$$Q_{i,t} f(x) = \mathbb{E}[f(X_{i,t}(x))],$$

for a continuous function $f : \mathbb{R}^N \to \mathbb{R}$ with polynomial growth at infinity.

$$P_t = e^{tL} = \sum_{k=0}^{m} \frac{t^k}{k!} L^k + \mathcal{O}(t^{m+1}).$$

Note that $L = \sum_{i=0}^{d+1} L_i$ and we also let

$$P_t^i = e^{tL_i} = \sum_{k=0}^{m} \frac{t^k}{k!} L_i^k + \mathcal{O}(t^{m+1}).$$

Our next goal is to approximate e^{tL}, through a combination of the "coordinate" semigroups e^{sL_i}'s such that

$$e^{tL} - \sum_{j=1}^{k} \xi_j e^{t_{1,j} A_{1,j}} \cdots e^{t_{\ell_j,j} A_{\ell_j,j}} = \mathcal{O}(t^{m+1}),$$

with $t_{i,j} = t_{i,j}(t) > 0$, $A_{i,j} \in \{L_0, L_1, \ldots, L_{d+1}\}$ and weights $\{\xi_j\} \subset [0,1]$ with $\sum_{j=1}^k \xi_j = 1$. In this case, one can define

$$Q_t = \sum_{j=1}^{k} \xi_j e^{t_{1,j}A_{1,j}} \cdots e^{t_{\ell_j,j}A_{\ell_j,j}}. \tag{2.7}$$

If needed one may further approximate each $e^{t_{1,j}A_{1,j}}$ (m-th order scheme) and in that case the definition of Q has to be further modified.

For simplicity let $d+1 = 2$ then

$$\begin{aligned} e^{tL} &= I + tL + \frac{t^2}{2}L^2 + O(t^3), \\ e^{tL_1}e^{tL_2} &= (I + tL_1 + \frac{t^2}{2}L_1^2 + ...)(I + tL_2 + \frac{t^2}{2}L_2^2 + ...) \\ &= I + tL + \frac{t^2}{2}\left(L_2^2 + L_1^2 + 2L_1L_2\right) + O(t^3), \end{aligned}$$

then

$$e^{tL} - e^{tL_1}e^{tL_2} = O(t^2).$$

Therefore the composition of the semigroups in the above order will lead to an approximation with local error of order $O(t^2)$. This approximation can be improved by randomizing it as follows

$$e^{tL} - \frac{1}{2}e^{tL_1}e^{tL_2} - \frac{1}{2}e^{tL_2}e^{tL_1} = O(t^3),$$

since $L^2 = L_1^2 + L_2^2 + L_1L_2 + L_2L_1$. Finally one needs to obtain a stochastic representation for $\frac{1}{2}e^{tL_1}e^{tL_2} + \frac{1}{2}e^{tL_2}e^{tL_1}$ and possibly approximate each coordinate process. These approximation methods for semigroups can be generalized in higher dimension as follows:

Example. Examples of schemes of order $2 = O(t^3)$:
Ninomiya-Ninomiya (see [10]):

$$Q_t = \frac{1}{2}e^{\frac{t}{2}L_0}e^{tL_1} \cdots e^{tL_{d+1}}e^{\frac{t}{2}L_0} + \frac{1}{2}e^{\frac{t}{2}L_0}e^{tL_{d+1}} \cdots e^{tL_1}e^{\frac{t}{2}L_0}.$$

Ninomiya-Victoir (see [9]):

$$Q_t = \frac{1}{2}e^{tL_0}e^{tL_1} \cdots e^{tL_{d+1}} + \frac{1}{2}e^{tL_{d+1}} \cdots e^{tL_1}e^{tL_0}.$$

Splitting (Strang) method:

$$Q_t = e^{\frac{t}{2}L_0} \cdots e^{\frac{t}{2}L_d}e^{tL_{d+1}}e^{\frac{t}{2}L_d} \cdots e^{\frac{t}{2}L_0}. \tag{2.8}$$

It is easy to see that all the approximation operators Q mentioned above are special case of (2.7). For example, the operator Q in Splitting (Strang) method is deduced from (2.7) by putting $k = 1; \xi = 1; t_i = \frac{t}{2}$, for $1 \le i \le 2d+1, i \ne d+1$, $t_{d+1} = t$; $A_i = A_{2d+2-i} = L_{i-1}$, for $i = 1, \ldots, d+1$.

Splitting is a classical idea that is used in approximations for partial differential equations. The only new feature in the present situation is that we make use of stochastic representations in order to obtain the associated

Monte Carlo method to (2.7). So the idea of this approximation method is to combine the above algebraic approach with its stochastic representation and if necessary the associated approximation of the stochastic representation in order to obtain the definition of Q.

The first approximation is obtained through the algebraic semigroup methods described above. The second approximation corresponds to an approximation to the corresponding semigroup $e^{t\ell_j,jA_{\ell_j,j}}$ which is amenable to a stochastic representation and that can be easily simulated or easily approximated and then simulated. In the remainder of the paper, we will concentrate on this second aspect of the approximations.

2.4. Stochastic Representations and Their Approximations

In this section we will show various cases where we approximate or simulate directly the stochastic representation of Q.

2.4.1. Diffusion Process With a Finite Number of Jumps Per Interval. In this section we will consider a full example by considering equation (2.4) in the particular case that Z is a compound Poisson process. First, we need to approximate the semigroup associated to the coordinate processes defined by

$$Q_{i,t}f(x) := E[f(X_{i,t}(x))].$$

In the case of $i = 1, ..., d$ we can approximate Q using the following result. Before that we need to introduce the exponential mapping. For given $\alpha : \mathbb{R}^N \to \mathbb{R}^N$, denote by $z_t(\alpha, x)$ the solution of

$$\frac{dz_s(\alpha,x)}{ds} = \alpha(z_s(\alpha,x)),\ z_0(\alpha,x) = x,\ s \in [0,1].$$

Theorem 2.2. *Let $V_i : \mathbb{R}^N \to \mathbb{R}^N$ be a smooth function satisfying the linear growth condition: $|V_i(x)| \le C(1 + |x|)$. Let $z_s(B_t^i V_i, x)$, $s \in [0,1]$ be the exponential map defined as above for fixed $t \in [0,1]$.*

For $i = 0, 1, ..., d$, the sde

$$X_{i,t}(x) = x + \int_0^t V_i(X_{i,s}(x)) \circ dB_s^i$$

has a unique solution given by

$$X_{i,t}(x) = z_1(B_t^i V_i, x).$$

Idea of the proof: Differentiating, we obtain

$$\frac{dz_s(\alpha V_i,x)}{d\alpha} = \int_0^s V_i(z_u(\alpha V_i,x))du + \alpha \int_0^s \nabla V_i(z_u^i(\alpha,x))\frac{dz_u(\alpha V_i,x)}{d\alpha}du.$$

This gives by Itô's formula that

$$dz_1(B_t^i V_i, x) = \frac{dz_1(B_t^i V_i,x)}{d\alpha} \circ dB_t^i.$$

Therefore the result follows if one proves that (exercise)

$$\frac{dz_1(\alpha V_i,x)}{d\alpha} = V_i(z_1(\alpha V_i,x)).$$

Now, the process X can be simulated as follows: First, we can solve for each time interval $(t_i, t_{i+1}]$ the coordinate processes equations

1. Solve (say exactly) the $d+1$ ODE's

$$\begin{aligned} X_{0,t/2}(x) &= x + \int_0^{t/2} V_0(X_{0,s}(x))ds \\ \frac{dz_s(B^i_{t/2}V_i, x)}{ds} &= B^i_{t/2}V_i(z_s(B^i_{t/2}V_i, x)),\ z_0(B^i_{t/2}V_i, x) = x, \end{aligned}$$

for $s \in [0,1]$ and $i = 1, ..., d$. Denote $X_{i,t}(x) = z_1(B^i_{t/2}V_i, x)$.

2. Solve (say exactly) the $d+1$ ODE's

$$\begin{aligned} \bar{X}_{0,t/2}(x) &= x + \int_0^{t/2} V_0(\bar{X}_{0,s}(x))ds \\ \frac{dz_s(\bar{B}^i_{t/2}V_i, x)}{ds} &= \bar{B}^i_{t/2}V_i(z_s(\bar{B}^i_{t/2}V_i, x)),\ z_0(\bar{B}^i_{t/2}V_i, x) = x, \end{aligned}$$

for $s \in [0,1]$, $i = 1, ..., d$ and $\bar{B}$ is an independent copy of B. Denote $\bar{X}_{i,t}(x) = z_1(\bar{B}^i_{t/2}V_i, x)$.

3. Solve (say exactly) the difference equation

$$X_{d+1,t}(x) = x + \int_0^t h(X_{d+1,s-}(x))dZ_s.$$

The global idea is that $X_{i,t}(x)$ represents the process that has as generator L^i. Hence, we use (say) the splitting (Strang) formula (2.8) to write

$$\hat{X}_t(x) = X_{0,t/2} \circ ... \circ X_{d,t/2} \circ X_{d+1,t} \circ \bar{X}_{d,t/2} \circ \ldots \circ \bar{X}_{1,t/2} \circ \bar{X}_{0,t/2}(x).$$

This gives a scheme of order 2.

On the other hand, if we approximate processes $X_{i,t}$ and $\bar{X}_{i,t}$ with a good high order approximation, say Y^i_t and $\bar{Y}^i_t$, respectively, then we can obtain an approximate of X by

$$\tilde{X}_t(x) = Y^0_{t/2} \circ ... \circ Y^d_{t/2} \circ Y^{d+1}_t \circ \bar{Y}^d_{t/2} .. \circ \bar{Y}^0_{t/2}(x).$$

The semigroup associated with the process X_t is Q_t in (2.7). Finally the Monte Carlo method is given by

$$\frac{1}{M}\sum_{j=1}^{M} f((\tilde{X}_{1/n} \circ ... \circ \tilde{X}_{1/n}(x))^{(j)}).$$

In the particular case that $X_{i,\cdot}$ can be solved exactly one can always take $Y^i_s = X_{i,s},\ s = t/2,\ t$.

Remark 2.3. 1. If one can solve the above ODE and difference equations 1, 2 and 3 without much effort then the scheme can be implemented. But if so, there is no reason to use the splitting method of order 2. One can use a a higher order method that will lead to better accuracy just by using compositions.

2. It is very rarely the case when one can in fact solve explicitly 1, 2 amd 3 above. Usually one has to approximate the solutions of the ODE's. Then the order of approximation has to match the order of the semigroup approximation method used. For example, in the above case if we use methods of order 2 to approach the ODE's then the order of the whole scheme will be order 2. In this case also the definition of Q has to be changed into the semigroup associated with the approximation process.

3. For the compounded Poisson case, we have that if λ is large, many jumps will appear in any interval and so the calculation time will be long. Later, we will see that we do not need to consider all the jumps in order to obtain an approximation of order 2. This can be intuitively understood because the probability of having two or more jumps in an interval of size t is $O(t^2)$.

2.4.2. On the Two Basic Properties in Order to Prove the Error of Approximation. In order to find an approximation of order n, one needs to check the two conditions mentioned in Remark 2.1. That is,

1. $Q_t f$ preserves the regularity properties of the function f.
2. $(Q_t - P_t)(f) = O(t^{n+1})$.

According to the operator splitting scheme explained in the previous section, one may even verify these conditions for each of the operators used in the decomposition. This first property when written mathematically becomes:

(H1). For $f_p(x) := |x|^{2p}$ $(p \in \mathbf{N})$,

$$Q_t f_p(x) \leq (1 + Kt) f_p(x) + K't$$

for $K = K(T, p)$, $K' = K'(T, p) > 0$.

This condition expresses the fact that Q does not alter the smoothness properties of the function f_p. The following condition expresses the fact that $P_t - Q_t = O(t^{n+1})$ and therefore the resulting scheme will be of order n. To be precise, we need to recall the definition of the functional space C_p^m for each $m \in \mathbb{N}$ and $p > 0$. For each function $f : \mathbb{R}^n \to \mathbb{R}$ in C^m, denote

$$\|f\|_{C_p^m} := \inf\{C \geq 0 : |\partial_x^\alpha f(x)| \leq C(1 + |x|^p),\ 0 \leq |\alpha| \leq m, x \in \mathbb{R}^n\}.$$

Then, denote

$$C_p^m = \{f \in C^m : \|f\|_{C_p^m} < \infty\}.$$

The Property 2 above when written mathematically becomes:

(H2). $\left|E[f(\bar{X}_t)] - E[f(X_t)]\right| \leq \|f\|_{C_p^{2n}}(1 + |x|^{p+n})t^{n+1}$.

Or in a more generalized form for $q \equiv q(n, p)$ and $m \equiv m(n)$

$$|E[f(\bar{X}_t)] - E[f(X_t)]| \leq \|f\|_{C_p^m}(1 + |x|^q)t^{n+1}.$$

In Section 2.4.5, we propose an scheme and verify that conditions (H1) and (H2) are valid in the case that $\int |y|\nu(dy) = \infty$.

2.4.3. The Study of the Jump-Size Adapted Scheme Using the Operator Splitting Method. Here we only discuss the approximation of the $(d+1)$th coordinate which corresponds to the jump process. Define for $\varepsilon > 0$ the finite activity Lévy process (Z_t^ε) with the Lévy triple $(\gamma, 0, \nu^\varepsilon)$ where the Lévy measure ν^ε is defined by

$$\nu^\varepsilon(E) = \nu(E \cap \{y : |y| > \varepsilon\}),\ E \in \mathcal{B}(\mathbb{R}_0^d).$$

We consider the approximate SDE

$$Y_t^{d+1,\varepsilon}(x) = x + \int_0^t h(Y_{s-}^{d+1,\varepsilon}(x))(dZ_s^\varepsilon + \gamma_\varepsilon ds).$$

In this case it is clear that the order to approximation on the jumps components is given by

$$\begin{aligned} &E[f(X_{d+1,t}(x))] - E[f(Y_t^{d+1,\varepsilon}(x))] \\ &\qquad = t\int_{|y|\le\varepsilon} (f(x+h(x)y) - f(x) - \nabla f(x)h(x)y)\nu(dy) + O(t^2). \end{aligned}$$

By a further Taylor expansion one obtains that

$$\begin{aligned} &\int_{|y|\le\varepsilon} (f(x+h(x)y) - f(x) - \nabla f(x)h(x)y)\nu(dy) \\ &\approx D^2 f(x)h(x)^{\otimes 2}\int_{|y|\le\varepsilon} |y|^2\nu(dy) + R\int_{|y|\le\varepsilon} |y|^3\nu(dy). \end{aligned}$$

Therefore one sees that if $\varepsilon > 0$ is chosen so that $\int_{|y|\le\varepsilon} |y|^2\nu(dy) = Ct$ then (H2) is satisfied with $n = 1$. Furthermore, the Asmussen-Rosiński approach [2] corresponds to the first term in the above expansion.

The verification of (H1) in this case is done in [14].

2.4.4. Approximate Small Jumps Scheme. In this section, we give an approximation scheme which uses a limited number of jumps per interval. We assume that $\int_{|y|<1} |y|\nu(dy) < \infty$. Then we further decompose the operator L_{d+1} defined in (2.6) as follows

$$\begin{aligned} L_{d+1} &= L_{d+1}^{1,\varepsilon} + L_{d+1}^{2,\varepsilon} + L_{d+1}^{3,\varepsilon}, \\ L_{d+1}^{1,\varepsilon} f(x) &:= \nabla f(x)h(x)\left(\gamma - \int_{\varepsilon<|y|\le 1} y\nu(dy)\right), \\ L_{d+1}^{2,\varepsilon} f(x) &:= \int_{|y|\le\varepsilon} (f(x+h(x)y) - f(x) - \nabla f(x)h(x)y)\nu(dy), \\ L_{d+1}^{3,\varepsilon} f(x) &:= \int_{\varepsilon<|y|} f(x+h(x)y) - f(x)\nu(dy). \end{aligned}$$

The operator $L_{d+1}^{1,\varepsilon}$ can be exactly generated using

$$\bar{X}_{d+1,t}^{1,\varepsilon} = x + \left(\gamma - \int_{\varepsilon<|y|\le 1} y\nu(dy)\right)\int_0^t h\left(\bar{X}_{d+1,s}^{1,\varepsilon}\right) ds.$$

For $L^{2,\varepsilon}_{d+1}$ one can use Asmussen-Rosinski (see [14]). We discuss $L^{3,\varepsilon}_{d+1}$.

The approximation for $L^{3,\varepsilon}_{d+1}$ is defined as follows. Let $\lambda_\varepsilon = \int_{|y|>\varepsilon} \nu(dy)$, $G_\varepsilon(dy) = \lambda_\varepsilon^{-1} 1_{|y|>\varepsilon}\nu(dy)$, and let $Z^\varepsilon \sim G_\varepsilon$ and let S^ε be a Bernoulli random variable independent of Z^ε.

$$\bar{X}^{3,\varepsilon}_{d+1,t}(x) = \begin{cases} x & \text{if } S^\varepsilon = 0, \\ x + h(x)Z^\varepsilon & \text{if } S^\varepsilon = 1. \end{cases}$$

Lemma 2.4. *Assume that* $\left|\lambda_\varepsilon^{-1}\mathbb{P}\left[S^\varepsilon = 1\right] - t\right| \le Ct^2$ *then*

$$\left|\mathbb{E}\left[f(\bar{X}^{3,\varepsilon}_{d+1,t})\right] - f(x) - tL^{3,\varepsilon}_{d+1}f(x)\right| \le Ct^2 \left\|f\right\|_{C^1_p}(1+|x|^{p+1}) \int_{|y|>\varepsilon} |y|\nu(dy).$$

In [14] an approximation for $L^{3,\varepsilon}_{d+1}$ with importance sampling and restriction on the number of jumps is proposed.

Remark 2.5. This approximate small jumps scheme has some advantages in comparison with the Jump-size adapted discretization schemes presented in Section 2.2. The first advantage is that in the former scheme we can control the number of jumps needed to be simulated. This fact is important especially in the case that it takes time to generate jump sizes. The second advantage is that the former scheme can be applied for SDE driven by both Brownian motion and jump processes while the latter scheme can be applied only for SDE driven by pure jump processes.

An extension with at most two jumps per interval: Considering more jumps per interval will give higher order approximations.

For $L^{3,\varepsilon}_{d+1}$ one can do the following: Let $G_\varepsilon(dy) = \lambda_\varepsilon^{-1} 1_{|y|>\varepsilon}\nu(dy)$, $\lambda_\varepsilon = \int_{|y|>\varepsilon}\nu(dy)$ and let Z^ε_1, $Z^\varepsilon_2 \sim G_\varepsilon$ independent between themselves and let S^ε_1 and S^ε_2 be two independent Bernoulli random variables independent of Z^ε_1, Z^ε_2.

$$\bar{X}^{3,\varepsilon}_{d+1,t}(x) := \begin{cases} x & \text{if } S^\varepsilon_1 = 0, \\ x + h(x)Z^\varepsilon_1 & \text{if } S^\varepsilon_1 = 1 \text{ and } S^\varepsilon_2 = 0 \\ x + h(x)Z^\varepsilon_1 + h(x + h(x)Z^\varepsilon_1)Z^\varepsilon_2 & \text{if } S^\varepsilon_1 = 1 \text{ and } S^\varepsilon_2 = 1. \end{cases}$$

Denote

$$p_\varepsilon := \mathbb{P}\left[S^\varepsilon_1 = 1\right]\left(1 + \mathbb{P}\left[S^\varepsilon_2 = 1\right]\right),$$
$$q_\varepsilon := \mathbb{P}\left[S^\varepsilon_1 = 1\right]\mathbb{P}\left[S^\varepsilon_2 = 1\right].$$

Lemma 2.6. *Assume that* $\left|\lambda_\varepsilon^{-1}p_\varepsilon - t\right| \le Ct^3$ *and* $\left|2\lambda_\varepsilon^{-2}q_\varepsilon - t^2\right| \le Ct^3$ *then*

$$\left|\mathbb{E}\left[f(\hat{X}^{3,\varepsilon}_{d+1,t})\right] - f(x) - tL^3_{d+1}f(x) - \frac{t^2}{2}\left(L^3_{d+1}\right)^2 f(x)\right|$$
$$\le Ct^3 \left\|f\right\|_{C^2_p}(1+|x|^{p+2})\left(1 + \left(\int_{|y|>\varepsilon}|y|\nu(dy)\right)^2\right).$$

2.4.5. A Case Study. In some cases it is possible to introduce the limited number of jumps scheme even when $\int_{|y|\leq 1}|y|\nu(dy)=\infty$. We suppose the one dimensional case for simplicity. Let S be a subordinator (an increasing Lévy process on $\mathbb{R}$) with Lévy density ρ and drift γ_S. That is,

$$S_t=\beta_0 t+\int_0^t\int_0^\infty zN_S(dz,dt),\quad \beta_0=\gamma_S-\int_0^\infty z\rho(z)dz,$$

where N_S is a Poisson random measure on $[0,\infty)\times[0,\infty)$ with intensity $\rho(z)dz$ and

$$\beta_0\geq 0,\quad \int_0^\infty (1\wedge z)\rho(z)dz<\infty. \tag{2.9}$$

Let $Z_t=\theta S_t+\sigma W_{S_t}$ where W is a standard Brownian motion independent of S. This is the setup in Section 2.2 in the particular case that the Lévy process Z is a subordinate to a Brownian motion with drift. It follows from Theorem 30.1 in [13] that Z is a Lévy process with the generating triplet (γ,A,ν) defined as follows

$$\begin{aligned} A&=\sigma\beta_0,\\ \nu(dy)&=\int_0^\infty \frac{1}{\sqrt{2\pi\sigma^2 t}}\exp\Big(-\frac{(y-\theta t)^2}{2\sigma^2 t}\Big)\rho(t)dtdy,\\ \gamma&=\theta\beta_0+\int_0^\infty\int_{|y|\leq 1}\frac{y}{\sqrt{2\pi\sigma^2 t}}\exp\Big(-\frac{(y-\theta t)^2}{2\sigma^2 t}\Big)\rho(t)dy\,dt. \end{aligned} \tag{2.10}$$

Let ϱ denote the Blumenthal-Getoor index of S. That is,

$$\varrho=\inf\{p>0:\int_0^1 z^p\rho(z)dz<\infty\}.$$

It follows from (2.9) that $\varrho\in[0,1]$.

The Blumenthal-Getoor index plays an essential role in our approximation. The following result relates the Blumenthal-Getoor indices of S and Z.

Lemma 2.7. *The Blumenthal-Getoor index of S is ϱ if and only if the Blumenthal-Getoor index of Z is 2ϱ.*

Proof. Since the integral on $[-1,0]$ can be converted into an integral on $[0,1]$ by doing a change of variable $w=-y$, for any $\alpha\in(0,1)$, we have

$$\begin{aligned} &\int_{|y|\leq 1}|y|^{2\alpha}\nu(dy)\\ &=\int_0^1 dy\int_0^\infty y^{2\alpha}\frac{1}{\sqrt{2\pi\sigma^2 t}}\Big[\exp\Big(-\frac{(y-\theta t)^2}{2\sigma^2 t}\Big)+\exp\Big(-\frac{(y+\theta t)^2}{2\sigma^2 t}\Big)\Big]\rho(t)dt\\ &=\int_0^1 dy\int_0^\infty \frac{y^{2\alpha}}{\sqrt{2\pi\sigma^2 t}}\exp\Big(-\frac{y^2+\theta^2t^2}{2\sigma^2 t}\Big)\Big[\exp\Big(\frac{\theta y}{\sigma^2}\Big)+\exp\Big(-\frac{\theta y}{\sigma^2}\Big)\Big]\rho(t)dt. \end{aligned}$$

Since $\exp\left(\frac{\theta y}{\sigma^2}\right) + \exp\left(-\frac{\theta y}{\sigma^2}\right) \geq 2$, we have

$$\int_{|y|\leq 1} |y|^{2\alpha}\nu(dy) \geq 2\int_0^1 dy \int_0^\infty \frac{y^{2\alpha}}{\sqrt{2\pi\sigma^2 t}} \exp\Big(-\frac{y^2+\theta^2t^2}{2\sigma^2 t}\Big)\rho(t)dt.$$

Since the integrand is non-negative, by using Fubini theorem, we have using the change of variables $z = \frac{y}{\sigma\sqrt{t}}$,

$$\begin{aligned}&\int_{|y|\leq 1} |y|^{2\alpha}\nu(dy)\\ &\geq C\sigma^{2\alpha}\int_0^1 dt \int_0^{1/(\sigma\sqrt{t})} z^{2\alpha}e^{-z^2/2}t^\alpha \exp\Big(-\frac{\theta^2 t}{2\sigma^2}\Big)\rho(t)dz.\end{aligned}$$

For each $t \in (0,1)$, one has

$$\int_0^{1/(\sigma\sqrt{t})} z^{2\alpha}e^{-z^2/2}dz \geq \int_0^{1/\sigma} z^{2\alpha}e^{-z^2/2}dz > 0,$$

and $\exp\big(-\frac{\theta^2 t}{2\sigma^2}\big) \geq \exp\big(-\frac{\theta^2}{2\sigma^2}\big)$. Hence,

$$\int_{|y|\leq 1} |y|^{2\alpha}\nu(dy) \geq C\int_0^1 t^\alpha \rho(t)dt. \tag{2.11}$$

On the other hand, one has

$$\begin{aligned}&\int_{|y|\leq 1} |y|^{2\alpha}\nu(dy)\\ &\leq 2\int_0^1 dy \int_0^\infty \frac{y^{2\alpha}}{\sqrt{2\pi\sigma^2 t}} \exp\Big(-\frac{y^2+\theta^2t^2}{2\sigma^2 t}\Big)\exp\Big(\frac{\theta y}{\sigma^2}\Big)\rho(t)dt.\\ &= 2\int_0^1 dy \left(\int_0^1 + \int_1^\infty\right) \frac{y^{2\alpha}}{\sqrt{2\pi\sigma^2 t}} \exp\Big(-\frac{y^2+\theta^2t^2}{2\sigma^2 t}\Big)\exp\Big(\frac{\theta y}{\sigma^2}\Big)\rho(t)dt.\end{aligned}$$

The second term above is less than $\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(\frac{\theta}{\sigma^2}\right)\int_1^\infty \rho(t)dt < \infty$ while the first term is bounded by (using again $z = \frac{y}{\sigma\sqrt{t}}$),

$$C\int_0^\infty z^{2\alpha}e^{-z^2/2}dz \int_0^1 t^\alpha\rho(t)dt \leq C\int_0^1 t^\alpha\rho(t)dt.$$

Hence

$$\int_{|y|\leq 1} |y|^{2\alpha}\nu(dy) \leq C\big(1+\int_0^1 t^\alpha\rho(t)dt\big).$$

This fact together with (2.11) implies, for any $\alpha \in (0,1)$,

$$\int_{|y|\leq 1} |y|^{2\alpha}\nu(dy) < \infty \Leftrightarrow \int_0^1 t^\alpha\rho(t)dt < \infty.$$

This yields the desired result. □

Remark 2.8. It follows from Lemma 2.7 that if the Blumenthal-Getoor index ρ of subordinator S is bigger than $1/2$, then

$$\int_{|y|\le 1} |y|\nu(dy) = \infty.$$

The following simple observation plays an important role in the next discussion.

Lemma 2.9. *Suppose that* $\varrho \in (0,1)$ *and* $\int_1^\infty t\rho(t)dt < \infty$, *then*

$$\Big| \int_{|y|\le 1} y\nu(dy) \Big| < \infty.$$

Proof. We have

$$\begin{aligned}
&\int_{|y|\le 1} y\nu(dy) \\
&= \int_0^1 dy \int_0^\infty \frac{y}{\sqrt{2\pi\sigma^2 t}} \exp\Big(-\frac{y^2+\theta^2t^2}{2\sigma^2 t}\Big)\Big[\exp\Big(\frac{\theta y}{\sigma^2}\Big) - \exp\Big(-\frac{\theta y}{\sigma^2}\Big)\Big]\rho(t)dt,
\end{aligned}$$

Hence as $e^x - e^{-x} \le 2xe^x$ for $x \in [0,1]$ and for each $\beta > 0$, $\sup_{x>0} x^\beta e^{-x} < \infty$, we obtain

$$\begin{aligned}
&\Big| \int_{|y|\le 1} y\nu(dy) \Big| \\
&\le 2\int_0^1 dy \int_0^\infty \frac{y}{\sqrt{2\pi\sigma^2 t}} \exp\Big(-\frac{y^2+\theta^2t^2}{2\sigma^2 t}\Big)\frac{|\theta| y}{\sigma^2}\exp\Big(\frac{|\theta|y}{\sigma^2}\Big)\rho(t)dt \\
&\le C\int_0^1 dy \int_0^\infty \Big(\frac{y^2}{2\sigma^2 t}\Big)^{1+\delta/2} \exp\Big(-\frac{y^2}{2\sigma^2 t}\Big) y^{-\delta} t^{(\delta+1)/2}\rho(t)dt \\
&\le C\int_0^1 y^{-\delta}dy \int_0^\infty t^{(\delta+1)/2}\rho(t)dt < \infty,
\end{aligned}$$

for some constant $\delta \in (2\varrho - 1, 1)$ where C is a positive constant that depends on σ^2. □

Throughout the rest of this section, we suppose that $\varrho < 1$. Then we can rewrite $\gamma = \theta\beta_0 + \int_{|x|\le 1} x\nu(dx)$. We decompose the operator L_{d+1} defined in (2.6) by $L_{d+1} = L^1_{d+1} + L^2_{d+1}$, where

$$\begin{aligned}
L^1_{d+1}f(x) &:= \theta\beta_0 h(x) f'(x), \\
L^2_{d+1}f(x) &:= \int_{\mathbb{R}} \big(f(x + h(x)y) - f(x)\big)\nu(dy).
\end{aligned} \tag{2.12}$$

The operator L^1_{d+1} can be exactly generated using

$$\overline{X}^1_{d+1,t}(x) = x + \theta\beta_0 \int_0^t h(\overline{X}^1_{d+1,s}(x))ds,$$

as before. The approximation for L^2_{d+1} is defined as follows: For some $\varepsilon \in (0,1)$ which will be specified later, let $H_\varepsilon(x) = C_\varepsilon^{-1} 1_{x>\varepsilon}\rho(x)$, $C_\varepsilon =$

$\int_\varepsilon^\infty \rho(x)dx$ and $\zeta_\varepsilon = \int_0^\varepsilon x\rho(x)dx$. Furthermore let S^ε be a Bernoulli random variable with $p_i = \mathbb{P}[S^\varepsilon = i], i = 0, 1$. We define

$$\overline{X}_t^{2,\varepsilon}(x) = \begin{cases} x + h(x)\big(\zeta_\varepsilon t\theta + \sigma\sqrt{\zeta_\varepsilon t}Z\big) & \text{if } S^\varepsilon = 0, \\ x + h(x)\big(\zeta_\varepsilon t\theta + \theta U^\varepsilon + \sigma\sqrt{\zeta_\varepsilon t + U^\varepsilon}Z\big) & \text{if } S^\varepsilon = 1, \end{cases}$$

where Z is a standard normal random variable and U^ε is a random variable with density function H_ε. We suppose that Z, U^ε, S^ε and W are mutually independent. Throughout this section we assume without loss of generality that $t \le 1$.

We need the following auxiliary estimate.

Lemma 2.10. *For any $\varrho_0 \in (\varrho, 1)$, there exists a positive constant $C(\varrho_0)$ which does not depend on ε such that for any $\varepsilon \in (0, 1)$,*

$$\zeta_\varepsilon \le C(\varrho_0)\varepsilon^{1-\varrho_0}, \tag{2.13}$$

$$C_\varepsilon = \int_\varepsilon^\infty \rho(z)dz \le C(\varrho_0)\varepsilon^{-\varrho_0}, \tag{2.14}$$

$$\int_\varepsilon^\infty \sqrt{z}\rho(z)dz \le C(\varrho_0)(1 + \varepsilon^{1/2-\rho_0}), \tag{2.15}$$

$$\int_0^\varepsilon z^{3/2}\rho(z)dz \le C(\varrho_0)\varepsilon^{3/2-\varrho_0}, \tag{2.16}$$

Proof. Because $\varrho_0 \in (\varrho, 1)$, there exists a positive constant $C_0 = C(\varrho_0)$ such that

$$\int_0^1 z^{\varrho_0}\rho(z)dz < C_0.$$

Hence one has the following estimate for ζ_ε,

$$\zeta_\varepsilon = \int_0^\varepsilon z\rho(z)dz = \int_0^\varepsilon z^{1-\varrho_0}z^{\varrho_0}\rho(z)dz \le \varepsilon^{1-\varrho_0}\int_0^\varepsilon z^{\varrho_0}\rho(z)ds \le C_0\varepsilon^{1-\varrho_0}.$$

Next one has,

$$\begin{aligned}\int_\varepsilon^\infty \rho(z)dz &= \int_\varepsilon^1 \rho(z)dz + \int_1^\infty \rho(z)dz \le \int_\varepsilon^1 \Big(\frac{z}{\varepsilon}\Big)^{\varrho_0}\rho(z)dz + \int_1^\infty \rho(z)dz \\ &\le C_0\varepsilon^{-\varrho_0} + \int_1^\infty \rho(z)dz \le C_1\varepsilon^{-\varrho_0},\end{aligned}$$

where $C_1 = C_0 + \int_1^\infty \rho(z)dz < \infty$ since $\varepsilon < 1$ and $\int_1^\infty \rho(z)dz < \infty$. A similar calculation gives (2.15).

Finally, one has that as $\varrho_0 < \frac{3}{2}$,

$$\int_0^\varepsilon z^{3/2}\rho(z)dz = \int_0^\varepsilon z^{-\varrho_0+3/2}z^{\varrho_0}\rho(z)dz \le C_0\varepsilon^{-\varrho_0+3/2}.$$

□

The following lemma will be used to justify condition (H2).

Lemma 2.11. *Assume that $f \in C_p^4$ for some $p > 1$, $\mathbb{P}[S^\varepsilon = 1] = C_\varepsilon t < 1$ and $\int_1^\infty z^{p+2}\rho(z)dz < \infty$, then for each $\varrho_0 \in (\varrho, 1)$, there exists a positive constant $C(\varrho_0)$ which does not depend on ε and t such that*

$$\begin{aligned}&\left|\mathbb{E}\left[f(\overline{X}_t^{2,\varepsilon}(x)\right] - f(x) - tL_{d+1}^2 f(x)\right| \\ &\leq C(\varrho_0)(1+|x|^{p+4})\|f\|_{C_p^4}\left(t^{3/2}\varepsilon^{3(1-\varrho_0)/2} + t^2\varepsilon^{3/2-2\varrho_0} + t^2\varepsilon^{1-\varrho_0} + t\varepsilon^{3/2-\varrho_0}\right)\end{aligned} \tag{2.17}$$

Proof. Before, we start the proof, we remind the reader the properties that will be used repeatedly without further mention. These are: 1. $\sup_{0<\varepsilon<1}\zeta_\varepsilon < \infty$ and $\zeta_\varepsilon \to 0$ as $\varepsilon \to 0$. 2. $|h(x)|^q \leq C(1+|x|^q)$ and 3. $\sup_{0<\varepsilon<1} C_\varepsilon t \leq 1$.

1) First we expand $\mathbb{E}\left[f(\overline{X}_t^{2,\varepsilon}(x))\right] - f(x)$. Set $p_i = \mathbb{P}[S^\varepsilon = i],\ i = 0, 1$. Using Taylor's expansion, one has

$$\begin{aligned}&\mathbb{E}\left[f(\overline{X}_t^{2,\varepsilon}(x))\right] - f(x) \\ =&p_0\int_{\mathbb{R}}\frac{e^{-y^2/2}}{\sqrt{2\pi}}h(x)(\zeta_\varepsilon t\theta + \sigma\sqrt{\zeta_\varepsilon t}y)\int_0^1 f'(x+uh(x)(\zeta_\varepsilon t\theta + \sigma\sqrt{\zeta_\varepsilon t}y))du\,dy \\ &+ C_\varepsilon^{-1}p_1\int_{\mathbb{R}}\frac{e^{-y^2/2}}{\sqrt{2\pi}}\int_\varepsilon^\infty h(x)(\zeta_\varepsilon t\theta + \theta z + \sigma\sqrt{\zeta_\varepsilon t + z}y)\rho(z) \\ &\quad\times\int_0^1 f'(x+uh(x)(\zeta_\varepsilon t\theta + \theta z + \sigma\sqrt{\zeta_\varepsilon t + z}y))du\,dz\,dy \\ =&I_1 + I_2.\end{aligned}$$

Using Taylor's expansion again, I_1 becomes

$$\begin{aligned}I_1 =&p_0\int_{\mathbb{R}}\frac{e^{-y^2/2}}{\sqrt{2\pi}}h(x)(\zeta_\varepsilon t\theta + \sigma\sqrt{\zeta_\varepsilon t}y)f'(x)dy \\ &+ p_0\int_{\mathbb{R}}\frac{e^{-y^2/2}}{\sqrt{2\pi}}h(x)^2(\zeta_\varepsilon t\theta + \sigma\sqrt{\zeta_\varepsilon t}y)^2\int_0^1 du \\ &\quad\times\int_0^1 uf''(x+uvh(x)(\zeta_\varepsilon t\theta + \sigma\sqrt{\zeta_\varepsilon t}y))dv\,dy \\ =&p_0h(x)f'(x)\zeta_\varepsilon t\theta \\ &+ \frac{p_0}{2}h(x)^2\int_{\mathbb{R}}\frac{e^{-y^2/2}}{\sqrt{2\pi}}\zeta_\varepsilon t\sigma^2y^2f''(x)\,dy \\ &+ p_0h(x)^3\int_{\mathbb{R}}\frac{e^{-y^2/2}}{\sqrt{2\pi}}\zeta_\varepsilon t\sigma^2y^2(\zeta_\varepsilon t\theta + \sigma\sqrt{\zeta_\varepsilon t}y)\int_0^1 du\int_0^1 dv \\ &\quad\int_0^1 u^2vf'''(x+uvwh(x)(\zeta_\varepsilon t\theta + \sigma\sqrt{\zeta_\varepsilon t}y))dw\,dy\end{aligned}$$

$$
\begin{aligned}
&+ p_0 h(x)^2 \int_{\mathbb{R}} \frac{e^{-y^2/2}}{\sqrt{2\pi}} ((\zeta_\varepsilon t\theta)^2 + 2\sigma\theta y(\zeta_\varepsilon t)^{3/2}) \int_0^1 du \\
&\qquad \times \int_0^1 u f''(x + uvh(x)(\zeta_\varepsilon t\theta + \sigma\sqrt{\zeta_\varepsilon t}y))dv\, dy \\
&= I_{11} + I_{12} + I_{13} + I_{14}, \qquad (2.18)
\end{aligned}
$$

where the second equation is obtained by writing $(\zeta_\varepsilon t\theta + \sigma\sqrt{\zeta_\varepsilon t}y)^2 = \sigma^2 y^2 \zeta_\varepsilon t + ((\zeta_\varepsilon t\theta)^2 + 2\sigma\theta y(\zeta_\varepsilon t)^{3/2})$ and using the fact that $\int_{\mathbb{R}} y e^{-y^2/2} dy = 0$.

The second term I_{12} can be rewritten as

$$
I_{12} = \frac{p_0}{2} \zeta_\varepsilon t \sigma^2 h(x)^2 f''(x). \qquad (2.19)
$$

Since $f \in C_p^4$, one can show that

$$
\begin{aligned}
|I_{13}| &\le C(\zeta_\varepsilon t)^{3/2} \|f\|_{C_p^4} \int_{\mathbb{R}} e^{-y^2/2} y^2 (1 + |x|^3)(\sqrt{\zeta_\varepsilon t} + y) \int_0^1 du \int_0^1 dv \\
&\qquad \times \int_0^1 u^2 v (1 + |x|^p) \Big(1 + u^p v^p \big((\zeta_\varepsilon t)^p + (\zeta_\varepsilon t)^{p/2} y^p \big) \Big) dw dy \\
&\le C(1 + |x|^{p+3}) \|f\|_{C_p^4} (\zeta_\varepsilon t)^{3/2}.
\end{aligned}
$$

It follows from (2.13) that

$$
|I_{13}| \le C(\varrho_0)(1 + |x|^{p+4}) \|f\|_{C_p^4} t^{3/2} \varepsilon^{3(1-\varrho_0)/2}.
$$

After using a similar argument for I_{14}, we finally get

$$
|I_{13}| + |I_{14}| \le C(\varrho_0)(1 + |x|^{p+4}) \|f\|_{C_p^4} t^{3/2} \varepsilon^{3(1-\varrho_0)/2}. \qquad (2.20)
$$

Furthermore, one has

$$
\begin{aligned}
I_2 =& th(x)^2 \int_{\mathbb{R}} dy \frac{e^{-y^2/2}}{\sqrt{2\pi}} \int_\varepsilon^\infty (\zeta_\varepsilon t\theta + \theta z + \sigma\sqrt{\zeta_\varepsilon t + z}y)^2 \rho(z) \int_0^1 du \int_0^1 u \\
&\times f''(x + uvh(x)(\zeta_\varepsilon t\theta + \theta z + \sigma\sqrt{\zeta_\varepsilon t + z}y))dv\, dz \\
&+ th(x) f'(x) \int_{\mathbb{R}} dy \frac{e^{-y^2/2}}{\sqrt{2\pi}} \int_\varepsilon^\infty (\zeta_\varepsilon t\theta + \theta z + \sigma\sqrt{\zeta_\varepsilon t + z}y)\rho(z) dz.
\end{aligned}
$$

In the first integral, we decompose $(\zeta_\varepsilon t\theta + \theta z + \sigma\sqrt{\zeta_\varepsilon t + z}y)^2 = (\zeta_\varepsilon t\theta)^2 + 2\zeta_\varepsilon t\theta(\theta z + \sigma\sqrt{\zeta_\varepsilon t + z}y) + (\theta z + \sigma\sqrt{\zeta_\varepsilon t + z}y)^2$ to define I_{21}, I_{22} and I_{23}; and in the second integral, θz together with $\zeta_\varepsilon t\theta$ define I_{24} and I_{25}, respectively. Note that, the last integral corresponding to $\sigma\sqrt{\zeta_\varepsilon t + z}$ is zero. In detail, we write

$$
I_2 = \sum_{i=1}^5 I_{2i},
$$

where

$$
\begin{aligned}
I_{21} &= th(x)^2 \int_{\mathbb{R}} dy \frac{e^{-y^2/2}}{\sqrt{2\pi}} \int_\varepsilon^\infty (\zeta_\varepsilon t\theta)^2 \rho(z) \int_0^1 du \\
&\quad \times \int_0^1 u f''(x + uvh(x)(\zeta_\varepsilon t\theta + \theta z + \sigma\sqrt{\zeta_\varepsilon t + z}y))dv\, dz, \\
I_{22} &= th(x)^2 \int_{\mathbb{R}} dy \frac{e^{-y^2/2}}{\sqrt{2\pi}} \int_\varepsilon^\infty 2\zeta_\varepsilon t\theta(\theta z + \sigma\sqrt{\zeta_\varepsilon t + z}y)\rho(z) \int_0^1 du \int_0^1 u \\
&\quad \times f''(x + uvh(x)(\zeta_\varepsilon t\theta + \theta z + \sigma\sqrt{\zeta_\varepsilon t + z}y))dv\, dz, \\
I_{23} &= th(x)^2 \int_{\mathbb{R}} dy \frac{e^{-y^2/2}}{\sqrt{2\pi}} \int_\varepsilon^\infty (\theta z + \sigma\sqrt{\zeta_\varepsilon t + z}y)^2 \rho(z) \int_0^1 du \int_0^1 u \\
&\quad \times f''(x + uvh(x)(\zeta_\varepsilon t\theta + \theta z + \sigma\sqrt{\zeta_\varepsilon t + z}y))dv\, dz, \\
I_{24} &= th(x) f'(x)\theta \int_\varepsilon^\infty z\rho(z)dz,
\end{aligned}
$$

and

$$
I_{25} = t^2 \zeta_\varepsilon \theta h(x) f'(x) \int_\varepsilon^\infty \rho(z)dz = p_1 t \zeta_\varepsilon \theta h(x) f'(x).
$$

We have

$$
I_{11} + I_{25} + I_{24} = th(x) f'(x)\theta \int_0^\infty z\rho(z)dz. \tag{2.21}
$$

Since $f \in C_p^4$ and $\varrho < 1$, one gets

$$
\int_\varepsilon^\infty z^q \rho(z)dz \le \int_0^1 z\rho(z)dz + \int_1^\infty z^{p+2}\rho(z)dz < \infty,
$$

for all $1 \le q \le p+2$. Furthermore, one has

$$
\begin{aligned}
|I_{21}| &\le Ct^3 \zeta_\varepsilon^2 (1+x^2)\|f\|_{C_p^4} \int_{\mathbb{R}} dy e^{-y^2/2} \int_\varepsilon^\infty \rho(z) \int_0^1 du \\
&\quad \times \int_0^1 u\Big(1 + |x|^p + u^p v^p (1+|x|^p)\big(t^p + z^p + (z^{p/2} + (\zeta_\varepsilon t)^{p/2})y^p\big)\Big)dv\, dz \\
&\le Ct^3 \zeta_\varepsilon^2 (1+|x|^{p+2})\|f\|_{C_p^4} \int_\varepsilon^\infty (1+z^p)\rho(z)dz. \\
&\le C(1+|x|^{p+2})\|f\|_{C_p^4} C_\varepsilon t^3 \zeta_\varepsilon^2 \\
&\le C(1+|x|^{p+2})\|f\|_{C_p^4} t^2 \zeta_\varepsilon^2.
\end{aligned}
$$

The last inequality follows from the fact that $p_1 = C_\varepsilon t \le 1$. It then follows from (2.13) that

$$
|I_{21}| \le C(\varrho_0)(1+|x|^{p+4})\|f\|_{C_p^4} t^2 \varepsilon^{1-\varrho_0}.
$$

By a similar argument, one has

$$
\begin{aligned}
|I_{22}| &\le Ct^2\zeta_\varepsilon(1+|x|^{p+2})\|f\|_{C_p^4}\Big(\int_\varepsilon^\infty(\sqrt{z}+z^{p+1})\rho(z)dz+\sqrt{t\zeta_\varepsilon}\int_\varepsilon^\infty\rho(z)dz\Big)\\
&\le Ct^2\zeta_\varepsilon(1+|x|^{p+2})\|f\|_{C_p^4}\Big(1+\int_\varepsilon^\infty\sqrt{z}\rho(z)dz+C_\varepsilon\sqrt{t\zeta_\varepsilon}\Big)\\
&\le C(1+|x|^{p+2})\|f\|_{C_p^4}\Big(t^2\zeta_\varepsilon\int_\varepsilon^\infty\sqrt{z}\rho(z)dz+(t\zeta_\varepsilon)^{3/2}+t^2\zeta_\varepsilon\Big).
\end{aligned}
$$

It follows from (2.13) and (2.15) that

$$
|I_{22}|\le C(1+|x|^{p+2})\|f\|_{C_p^4}\big(t^2\varepsilon^{1-\varrho_0}+t^2\varepsilon^{3/2-2\varrho_0}+t^{3/2}\varepsilon^{3(1-\varrho_0)/2}\big). \quad (2.22)
$$

Next, by applying Taylor's expansion for f'', one gets $I_{23}=I_{23a}+I_{23b}$, where

$$
I_{23a}=th(x)^2f''(x)\int_{\mathbb{R}}dy\frac{e^{-y^2/2}}{2\sqrt{2\pi}}\int_\varepsilon^\infty(\theta z+\sigma\sqrt{\zeta_\varepsilon t+z}y)^2\rho(z)dz, \quad (2.23)
$$

and

$$
\begin{aligned}
I_{23b}&=th(x)^3\int_{\mathbb{R}}\frac{e^{-y^2/2}}{\sqrt{2\pi}}\int_\varepsilon^\infty(\theta z+\sigma\sqrt{\zeta_\varepsilon t+z}y)^2(\zeta_\varepsilon t\theta+\theta z+\sigma\sqrt{\zeta_\varepsilon t+z}y)\rho(z)\\
&\times\int_0^1du\int_0^1u^2v\int_0^1f'''(x+uvwh(x)(\zeta_\varepsilon t\theta+\theta z+\sigma\sqrt{\zeta_\varepsilon t+z}y))dw\,dv\,dz\,dy.
\end{aligned} \quad (2.24)
$$

2) Next, we expand $L_{d+1}^2f(x)$. It follows from (2.10) and (2.12) and Taylor's expansion for f that

$$
\begin{aligned}
&L_{d+1}^2f(x)\\
&=h(x)\int_{\mathbb{R}}y\int_0^1f'(x+uh(x)y)du\int_0^\infty\frac{1}{\sqrt{2\pi\sigma^2z}}\exp\Big(-\frac{(y-\theta z)^2}{2\sigma^2z}\Big)\rho(z)dz\,dy\\
&=h(x)\int_{\mathbb{R}}\int_0^\infty(\theta z+\sigma\sqrt{z}y)\int_0^1f'(x+uh(x)(\theta z+\sigma\sqrt{z}y))du\frac{e^{-y^2/2}}{\sqrt{2\pi}}\rho(z)dzdy\\
&=h(x)^2\int_{\mathbb{R}}dy\frac{e^{-y^2/2}}{\sqrt{2\pi}}\Big(\int_0^\varepsilon+\int_\varepsilon^\infty\Big)(\theta z+\sigma\sqrt{z}y)^2\int_0^1du\\
&\quad\times\int_0^1uf''(x+uvh(x)(\theta z+\sigma\sqrt{z}y))dv\rho(z)dz\\
&\quad+h(x)f'(x)\theta\int_0^\infty z\rho(z)dz\\
&=J_1+J_2+J_3,
\end{aligned}
$$

where the second equation follows by using an appropriate change of variables. By applying Taylor's expansion for f'', one gets $J_2 = J_{21} + J_{22}$, where

$$\begin{aligned} J_{21} &= h(x)^2 \int_{\mathbb{R}} dy \frac{e^{-y^2/2}}{\sqrt{2\pi}} \int_\varepsilon^\infty (\theta z + \sigma\sqrt{z}y)^2 \int_0^1 du \int_0^1 u f''(x) dv \rho(z) dz \\ &= h(x)^2 f''(x) \int_{\mathbb{R}} dy \frac{e^{-y^2/2}}{2\sqrt{2\pi}} \int_\varepsilon^\infty (\theta z + \sigma\sqrt{z}y)^2 \rho(z) dz, \end{aligned} \tag{2.25}$$

and

$$\begin{aligned} J_{22} = h(x)^3 \int_{\mathbb{R}} dy \frac{e^{-y^2/2}}{\sqrt{2\pi}} \int_\varepsilon^\infty (\theta z + \sigma\sqrt{z}y)^3 \int_0^1 du \int_0^1 u^2 v \\ \times \int_0^1 f'''(x + uvwh(x)(\theta z + \sigma\sqrt{z}y)) dw dv \rho(z) dz. \end{aligned} \tag{2.26}$$

Next, we write $(\theta z + \sigma\sqrt{z}y)^2 = \sigma^2 y^2 z + (\theta^2 z^2 + 2\sigma\theta y z^{3/2})$ and after applying Taylor's expansion for f'', we get

$$\begin{aligned} J_1 =& \sigma^2 h(x)^2 f''(x) \int_{\mathbb{R}} y^2 \frac{e^{-y^2/2}}{2\sqrt{2\pi}} dy \int_0^\varepsilon z\rho(z) dz \\ &+ \sigma^2 h(x)^3 \int_{\mathbb{R}} y^2 \frac{e^{-y^2/2}}{\sqrt{2\pi}} \int_0^\varepsilon \rho(z) \int_0^1 du \int_0^1 u^2 v (\theta z + \sigma\sqrt{z}y)^3 \\ &\times \int_0^1 f'''(x + uvwh(x)(\theta z + \sigma\sqrt{z}y)) dw\, dv\, dz\, dy \\ &+ h(x)^2 \int_{\mathbb{R}} dy \int_0^\varepsilon dz (\theta^2 z^2 + 2\theta\sigma y z^{3/2}) \rho(z) \frac{e^{-y^2/2}}{\sqrt{2\pi}} \int_0^1 du \\ &\times \int_0^1 u f''(x + uvh(x)(\theta z + \sigma\sqrt{z}y)) dv \\ =& J_{11} + J_{12} + J_{13}. \end{aligned}$$

Using a similar argument as before, one has

$$|J_{12}| + |J_{13}| \le C(1 + |x|^{p+3}) \|f\|_{C_p^4} \int_0^\varepsilon z^{3/2} \rho(z) dz.$$

Therefore, it follows from (2.16) that

$$|J_{12}| + |J_{13}| \le C(\varrho_0)(1 + |x|^{p+4}) \|f\|_{C_p^4} \varepsilon^{3/2 - \varrho_0}. \tag{2.27}$$

3) Now we compare the factors of $\mathbb{E}\big[f(\overline{X}_t^{2,\varepsilon}(x))\big] - f(x)$ and $L_{d+1}^2 f(x)$. First, it follows from (2.21) that

$$I_{11} + I_{25} + I_{24} = tJ_3. \tag{2.28}$$

And it follows from (2.19) that

$$tJ_{11} - I_{12} = \frac{1}{2}\sigma^2 p_1 t \zeta_\varepsilon h(x)^2 f''(x). \tag{2.29}$$

Next, it follows from (2.23) and (2.25) that

$$
\begin{aligned}
&I_{23a} - tJ_{21} \\
&= th(x)^2 f''(x) \int_{\mathbb{R}} dy \frac{e^{-y^2/2}}{2\sqrt{2\pi}} \int_{\varepsilon}^{\infty} \Big(\sigma^2 \zeta_\varepsilon t y^2 + 2\theta\sigma z(\sqrt{z+\zeta_\varepsilon t} - \sqrt{z}) y \Big) \rho(z) dz,
\end{aligned}
$$

hence as $p_1 = C_\varepsilon t$ it follows from the above and (2.29)

$$
I_{23a} + I_{12} = t(J_{11} + J_{21}). \tag{2.30}
$$

Next, we compare I_{23b} and J_{22}. It follows from (2.24) and (2.26) that $I_{23b} - tJ_{22} = K_1 + K_2$, where

$$
\begin{aligned}
K_1 = th(x)^3 &\int_{\mathbb{R}} dy \int_{\varepsilon}^{\infty} \Big((\theta z + \sigma\sqrt{\zeta_\varepsilon t + z} y)^2 (\zeta_\varepsilon t\theta + \theta z + \sigma\sqrt{\zeta_\varepsilon t + z} y) \\
&- (\theta z + \sigma\sqrt{z} y)^3 \Big) \rho(z) \int_0^1 du \int_0^1 u^2 v \\
&\times \int_0^1 \frac{e^{-y^2/2}}{\sqrt{2\pi}} f'''(x + uvwh(x)(\zeta_\varepsilon t\theta + \theta z + \sigma\sqrt{\zeta_\varepsilon t + z} y)) dw\, dv\, dz,
\end{aligned}
$$

and

$$
\begin{aligned}
K_2 = &th(x)^3 \int_{\mathbb{R}} dy \int_{\varepsilon}^{\infty} (\theta z + \sigma\sqrt{z} y)^3 \rho(z) \int_0^1 du \int_0^1 dv \int_0^1 dw\, u^2 v \frac{e^{-y^2/2}}{\sqrt{2\pi}} \\
&\times \Big(f'''(x + uvwh(x)(\zeta_\varepsilon t\theta + \theta z + \sigma\sqrt{\zeta_\varepsilon t + z} y) \\
&\quad - f'''(x + uvwh(x)(\zeta_\varepsilon t\theta + \theta z + \sigma\sqrt{z} y)) \Big) dz.
\end{aligned}
$$

We rewrite $K_1 = K_{11} + K_{12}$, where

$$
\begin{aligned}
K_{11} = t^2 \zeta_\varepsilon \theta h(x)^3 &\int_{\mathbb{R}} dy \int_{\varepsilon}^{\infty} (\theta z + \sigma\sqrt{\zeta_\varepsilon t + z} y)^2 \rho(z) \int_0^1 du \int_0^1 u^2 v \\
&\times \int_0^1 \frac{e^{-y^2/2}}{\sqrt{2\pi}} f'''(x + uvwh(x)(\zeta_\varepsilon t\theta + \theta z + \sigma\sqrt{\zeta_\varepsilon t + z} y)) dw\, dv\, dz,
\end{aligned}
$$

and

$$
\begin{aligned}
K_{12} = \ & th(x)^3 \int_{\mathbb{R}} dy \int_{\varepsilon}^{\infty} \Big((\theta z + \sigma\sqrt{\zeta_\varepsilon t + z} y)^3 - (\theta z + \sigma\sqrt{z} y)^3 \Big) \rho(z) \int_0^1 du \\
&\times \int_0^1 u^2 v \int_0^1 \frac{e^{-y^2/2}}{\sqrt{2\pi}} f'''(x + uvwh(x)(\zeta_\varepsilon t\theta + \theta z + \sigma\sqrt{\zeta_\varepsilon t + z} y)) dw\, dv\, dz.
\end{aligned}
$$

By using a similar argument as before, we have

$$
\begin{aligned}
|K_{11}| &\le C(1 + |x|^{p+3}) \|f\|_{C_p^4} \zeta_\varepsilon t^2 \int_{\varepsilon}^{\infty} (\zeta_\varepsilon t + z + z^{p+2}) \rho(z) dz \\
&\le C(1 + |x|^{p+3}) \|f\|_{C_p^4} \zeta_\varepsilon t^2 (1 + C_\varepsilon \zeta_\varepsilon t) \\
&\le C(\varrho_0)(1 + |x|^{p+4}) \|f\|_{C_p^4} t^2 \varepsilon^{1-\varrho_0},
\end{aligned}
$$

where the last inequality follows from the fact that $C_\varepsilon t \le 1$ and (2.13). Next, we will estimate K_{12}. First, we have

$$\begin{aligned}
&(\theta z + \sigma\sqrt{\zeta_\varepsilon t + z}y)^3 - (\theta z + \sigma\sqrt{z}y)^3 \\
&= \sigma y(\sqrt{\zeta_\varepsilon t + z} - \sqrt{z})\big((\theta z + \sigma\sqrt{\zeta_\varepsilon t + z}y)^2 + (\theta z + \sigma\sqrt{\zeta_\varepsilon t + z}y)(\theta z + \sigma\sqrt{z}y) \\
&\qquad + (\theta z + \sigma\sqrt{z}y)^2\big) \\
&\le Cy(\sqrt{\zeta_\varepsilon t + z} - \sqrt{z})(z^2 + y^2 z + y^2\zeta_\varepsilon t).
\end{aligned}$$

Furthermore, since

$$\begin{aligned}
\int_\varepsilon^\infty (\sqrt{\zeta_\varepsilon t + z} - \sqrt{z}) z\rho(z)dz &\le \zeta_\varepsilon t \int_\varepsilon^\infty \frac{z\rho(z)}{\sqrt{\zeta_\varepsilon t + z} + \sqrt{z}}dz \le \\
&\le \zeta_\varepsilon t \int_\varepsilon^\infty \sqrt{z}\rho(z)dz, \\
\int_\varepsilon^\infty (\sqrt{\zeta_\varepsilon t + z} - \sqrt{z})\rho(z)dz &\le \sqrt{t\zeta_\varepsilon} \int_\varepsilon^\infty \rho(z)dz = C_\varepsilon\sqrt{t\zeta_\varepsilon},
\end{aligned}$$

and for all $q \ge 3/2$,

$$\int_\varepsilon^\infty (\sqrt{\zeta_\varepsilon t + z} - \sqrt{z}) z^q \rho(z)dz \le \zeta_\varepsilon t \int_\varepsilon^\infty z^{q-1/2}\rho(z)dz \le C\zeta_\varepsilon t,$$

we have

$$\begin{aligned}
|K_{12}| &\le C(1 + |x|^{p+3})\|f\|_{C_p^4} t^2\zeta_\varepsilon \Big(1 + C_\varepsilon(\zeta_\varepsilon t)^{1/2} + \int_\varepsilon^\infty \sqrt{z}\rho(z)dz\Big) \\
&\le C(1 + |x|^{p+3})\|f\|_{C_p^4} \Big(t^2\zeta_\varepsilon + (\zeta_\varepsilon t)^{3/2} + t^2\zeta_\varepsilon \int_\varepsilon^\infty \sqrt{z}\rho(z)dz\Big) \\
&\le C(1 + |x|^{p+4})\|f\|_{C_p^4} \big(t^2\varepsilon^{1-\varrho_0} + t^{3/2}\varepsilon^{3(1-\varrho_0)/2} + t^2\varepsilon^{3/2-2\varrho_0}\big), \quad (2.31)
\end{aligned}$$

where the last inequality follows from (2.13)–(2.15). Now we estimate K_2. Since $f \in C_p^4$, for any $u, v \in \mathbb{R}$, one has $|f'''(u) - f'''(v)| \le C\|f\|_{C_p^4}|u - v|(1 + |u|^p + |v|^p)$, hence

$$\begin{aligned}
|K_2| \le& Ct|h(x)|^4\|f\|_{C_p^4} \int_{\mathbb{R}} dy \int_\varepsilon^\infty |(\theta z + \sigma\sqrt{z}y)^3 y|(\sqrt{z + \zeta_\varepsilon t} - \sqrt{z})\rho(z) \\
&\times \int_0^1 du \int_0^1 dv \int_0^1 dw u^2 v \frac{e^{-y^2/2}}{\sqrt{2\pi}} uv \\
&\times \Big(1 + |x|^p + u^p v^p w^p (1 + |x|^p)\big(\zeta_\varepsilon^p t^p + z^p + (z^{p/2} + \zeta_\varepsilon^{p/2} t^{p/2})y^p\big)\Big)dz \\
\le& Ch(x)^4\|f\|_{C_p^4} t^2\zeta_\varepsilon \int_{\mathbb{R}} dy \int_\varepsilon^\infty |yz(\sqrt{z} + y)^3|\rho(z) \\
&\times \int_0^1 du \int_0^1 dv \int_0^1 dw\, u^2 v \frac{e^{-y^2/2}}{\sqrt{2\pi}} uv \\
&\times \Big(1 + |x|^p + u^p v^p w^p (1 + |x|^p)\big(\zeta_\varepsilon^p t^p + z^p + (z^{p/2} + \zeta_\varepsilon^{p/2} t^{p/2})y^p\big)\Big)dz.
\end{aligned}$$

Therefore

$$|K_2| \le C(1+|x|^{p+4})\|f\|_{C_p^4} t^2 \zeta_\varepsilon \int_\varepsilon^\infty z\rho(z)dz$$
$$\le C(1+|x|^{p+4})\|f\|_{C_p^4} t^2 \varepsilon^{1-\varrho_0}, \tag{2.32}$$

since $\int_\varepsilon^\infty z\rho(z)dz < \infty$.

4) Finally, it follows from (2.20), (2.22), (2.27), (2.28) and (2.30)-(2.32) that for any $\varrho_0 \in (\varrho, 1)$, there exists a positive constant $C(\varrho_0)$ which does not depend on ε such that

$$\big|\mathbb{E}\big[f(\overline{X}_t^{2,\varepsilon}(x)\big] - f(x) - tL_{d+1}^2 f(x)\big|$$
$$\le C(\varrho_0)(1+|x|^{p+4})\|f\|_{C_p^4}\Big(t^{3/2}\varepsilon^{3(1-\varrho_0)/2} + t^2\varepsilon^{3/2-2\varrho_0} + t^2\varepsilon^{1-\varrho_0} + t\varepsilon^{3/2-\varrho_0}\Big),$$

this implies (2.17). □

Next, the parameter ε should be chosen in order to obtain the best bound in (2.17). After a simple calculation, we have the following result.

Lemma 2.12. *Assume that $f \in C_p^4$ and $\int_1^\infty z^{p+2}\rho(z)dz < \infty$. For each $\varrho_0 \in (\varrho, 1)$, if we choose $\varepsilon = O(t^{1/\varrho_0})$ and such that $C_\varepsilon t < 1$, then there exists a positive constant $C(\varrho_0)$ which does not depend on ε and t such that*

$$\big|\mathbb{E}\big[f(\overline{X}_t^{2,\varepsilon}(x)\big] - f(x) - tL_{d+1}^2 f(x)\big| \le C(\varrho_0)(1+|x|^{p+4})\|f\|_{C_p^4}(t^{1+1/\varrho_0} + t^{3/(2\rho_0)}).$$

This result shows that if $\varepsilon = O(t^{1/\rho_0})$ and $C_\varepsilon t < 1$ then the analog of (H2) will be satisfied ($n = 1$ if $\rho \ge 1/2$ and $n = 2$ if $\rho < 1/2$). We also remark that the fact that $\varepsilon = \kappa_0 t^{1/\rho_0}$ together with $C_\varepsilon t < 1$ results on a choice for κ_0.

The next lemma verifies condition (H1) which corresponds to the assumption $(\mathcal{M})$ in [14].

Lemma 2.13. *Assume that $\mathbb{P}[S^\varepsilon = 1] = C_\varepsilon t < 1$. Then for any $p \ge 2$ such that $\int_1^\infty z^p\rho(z)dz < \infty$, there exist constants K and K' satisfying*

$$\mathbb{E}\big[|\overline{X}_t^{2,\varepsilon}(x)|^p\big] \le (1+Kt)|x|^p + K't.$$

Proof. We first denote $f(x) = |x|^p$ and write $\mathbb{E}[f(\bar{X}_t^{2,\varepsilon}(x)] - f(x) = I_1 + I_2$ as in the first part of the proof of Lemma 2.11. We need to show that

$$|I_1| + |I_2| \le ct(1+|x|^p), \quad \forall x \in \mathbb{R}.$$

It follows from (2.18) that

$$|I_1| \le t\zeta_\varepsilon p\theta|h(x)||x|^{p-1}\int_{\mathbb{R}} \frac{e^{-y^2/2}}{\sqrt{2\pi}}dy$$
$$+ t\zeta_\varepsilon p(p-1)h(x)^2 \int_{\mathbb{R}} \frac{e^{-y^2/2}}{\sqrt{2\pi}}(\sqrt{\zeta_\varepsilon t}\theta + \sigma y)^2 \int_0^1 du$$
$$\times \int_0^1 u|x + uvh(x)(\zeta_\varepsilon t\theta + \sigma\sqrt{\zeta_\varepsilon t}y)|^{p-2}dvdy.$$

Since $|h(x)| \leq C(1+|x|)$, we get

$$\begin{aligned} |I_1| &\leq Ct\zeta_\varepsilon(1+|x|^p) \\ &\leq Ct(1+|x|^p), \end{aligned}$$

since ζ_ε is bounded. It remains to bound I_2. Since $C_\varepsilon^{-1}p_1 = t$, we have

$$\begin{aligned} I_2 \leq tp \int_{\mathbb{R}} \frac{e^{-y^2/2}}{\sqrt{2\pi}} \int_\varepsilon^\infty |h(x)|(\zeta_\varepsilon t\theta + \theta z + \sigma\sqrt{\zeta_\varepsilon t + z}|y|)\rho(z) \\ \times \int_0^1 |x + uh(x)(\zeta_\varepsilon t\theta + \theta z + \sigma\sqrt{\zeta_\varepsilon t + z}y)|^{p-1} du dz dy. \end{aligned}$$

And it follows from (2.13) that

$$\begin{aligned} |I_2| &\leq Ct(1+|x|^p)\Big(1 + \int_\varepsilon^\infty z^p \rho(z)dz\Big) \\ &\leq Ct(1+|x|^p), \end{aligned}$$

since $C_\varepsilon t < 1$ and $\int_\varepsilon^\infty z^p\rho(z)dz \leq \int_0^1 z\rho(z)dz + \int_1^\infty z^p\rho(z)dz < \infty$. □

Finally, the rate of convergence of our scheme can be established by following the guide in Section 2.4.2 under appropriate regularity conditions.

3. Numerical Study

In our numerical study, we concentrate on examples from three points of view.

- Lévy measures with different values of the Blumenthal-Getoor index.
- Different types of expectations. That is, to consider the calculation of $\mathbb{E}[f(X_1)]$ for different functions f.
- Different types of SDE's (different types of coefficients). In particular, we consider oscillating type of coefficients $\sin(ax)$ for different values of a.

In some of the cases considered, the theory provided so far does not tell us the theoretical rate of convergence. Still, by doing the simulations one can learn that the proposed methods still apply and points to further possible theoretical extensions of these methods.

3.1. Example 1: Weak Approximation for an SDE Driven by a Tempered Stable Lévy Process

Let Z be a tempered stable Lévy process with Lévy measure ν defined on $\mathbb{R}$ by

$$\nu(dy) = \frac{1}{|y|^{\alpha+1}}\big(c_+e^{-\lambda_+|y|}1_{y>0} + c_-e^{-\lambda_-|y|}1_{y<0}\big)dy.$$

3.1.1. We consider the following one dimensional SDE

$$dX_t = \sin(aX_{t-})dZ_t, \quad X_0 = 1.$$

We will estimate $\mathbb{E}[f(X_1)]$ with $f(x) = 2-2\cos(x-X_0)$ using various schemes mentioned in the last section.

We choose the parameter values $a = 1.0$, $\lambda_+ = 1.0, c_+ = 1, \gamma = 1$ and $c_- = 0$. Figures 1 and 2 present the Monte Carlo estimators and the corresponding variances obtained by using various schemes with the jump index Blumenthal-Getoor $\alpha = 0.1$ and $\alpha = 0.9$, respectively. The symbols ES, JSAS, 1JS and 2JS stand for Euler scheme (Section 2.1), Jump size adapted scheme (Section 2.2), One jump scheme (Section 2.4.4) and Two jump scheme (Section 2.4.4), respectively.

In the case that these schemes use the Asmussen-Rosiński approximation for small jumps, we append in parentheses "AR" to their symbols.

In the following, we provide some detailed information about each scheme and compute their theoretical computational costs. To be precise, we fix the error of the estimate at a level ε_0 and compute the expected computational cost, which is needed asymptotically in order to reach this level of error in the weak sense, with respect to ε_0. As usual, these calculations are exact up to constants.

First, we remark that in any scheme, if needed, we will always use the RK4 method (Runge-Kutta scheme of order 4) to solve the ode $dx_t = \sin(ax_t)dt$ (see [4]) if needed.

1. *Euler scheme without AR-correction:* We use the scheme presented in page 187 [5] to generate Z. The method consists of simulating the big jumps and replacing the small ones with their expectation.

   Fixed a jump threshold level $\varepsilon > 0$, then we approximate Z by a compound Poisson process Z^ε which has finite Lévy measure with density

$$\nu^\varepsilon(y) = \frac{1}{|y|^{\alpha+1}}\left(c_+e^{-\lambda_+|y|}1_{y>\varepsilon} + c_-e^{-\lambda_-|y|}1_{y<-\varepsilon}\right).$$

   Hence Z has intensity $\lambda_\varepsilon = \int_{\mathbb{R}} \nu^\varepsilon(y)dy$, and jump size distribution $p^\varepsilon(x) = \nu^\varepsilon(x)/\lambda_\varepsilon$. The jump size distribution can be simulated by using the acceptance-rejection method. It has been shown in [5] that the average number of loops needed to generate one random variable tends to 1 when $\varepsilon \to 0$. Hence the computational cost to generate Z^ε is proportional to $\lambda_\varepsilon = O(\varepsilon^{-\alpha})$. Therefore considering that there are t^{-1} time partition points, we obtain a total cost of $t^{-1} + \varepsilon^{-\alpha}$.

   On the other hand, the order of convergence of this scheme is

$$t + \int_{|y|\le\varepsilon} |y|^2\nu(dy) \approx t + \varepsilon^{2-\alpha}.$$

   If we choose $t = \varepsilon^{2-\alpha} = \varepsilon_0$ then the computational cost to reach to error of level ε_0 is $(\frac{1}{\varepsilon_0})^{1\vee(\alpha/(2-\alpha))}$. One should remark that this computational

cost blows up when $\alpha \to 2$ even if the error level ε_0 stays constant and sufficiently small.

2. *Euler scheme with AR-correction:* The increment of Z is generated as before with a modification: we replace the small jumps by a Brownian motion with the same local mean and variance as explained in Section 2.2.

   The order of convergence of this scheme is

$$t + \int_{|y|\le\varepsilon} |y|^3 \nu(dy) \approx t + \varepsilon^{3-\alpha}.$$

   As before, if we choose $t = \varepsilon^{3-\alpha} = \varepsilon_0$ then the computational cost to reach to error of level ε_0 is $(\frac{1}{\varepsilon})^{1\vee(\alpha/(3-\alpha))}$.

3. *JSAS without AR-correction:* The approximated solution $\hat{X}$ is defined inductively as follow: $\hat{X}(0) = X_0$ and for $i \ge 0$,

$$\hat{X}(T^\varepsilon_{i+1}-) = \theta\Big(\gamma_\varepsilon(T^\varepsilon_{i+1} - T^\varepsilon_i); \hat{X}(T^\varepsilon_i)\Big),$$
$$\hat{X}(T^\varepsilon_{i+1}) = \hat{X}(T^\varepsilon_{i+1}-) + h(\hat{X}(T^\varepsilon_{i+1}-))\Delta Z(T^\varepsilon_{i+1}).$$

   For an arbitrary point t, we define

$$\hat{X}(t) = \theta\Big(\gamma_\varepsilon(t - \eta_t); \hat{X}(\eta_t)\Big),$$

   where $\eta_t = \sup\{T^\varepsilon_i : T^\varepsilon_i \le t\}$.

   Although this scheme was not directly studied in [8]. The same ideas give that the error is of the order $\varepsilon^{2-\alpha}$. Therefore the computational cost to reach to error of level ε_0 is $(\frac{1}{\varepsilon_0})^{\frac{\alpha}{2-\alpha}}$.

   Now we introduce the cost for the schemes with limited number of jumps. Recall that for these schemes, we only consider the case that $\alpha < 1$.

4. *JSAS with AR-correction:* The computational cost is proportional to $\lambda_\varepsilon = \int_{|y|>\varepsilon} \nu(dy) = O(\varepsilon^{-\alpha})$.

   The order of convergence of this scheme is

$$\frac{\sigma^2_\varepsilon}{\lambda_\varepsilon}(\sigma^2_\varepsilon + |\gamma_\varepsilon|) + \int_{|y|\le\varepsilon} |y|^3 \bar{\chi}_\varepsilon \nu(dy) = \varepsilon^{2\wedge(3-\alpha)}.$$

   Hence, the computational cost to reach to error of level ε_0 is $\varepsilon_0^{-\alpha/(2\wedge(3-\alpha))}$.

5. *1JS without AR-correction:* The weak error of this scheme is proportional to $t + \varepsilon^{2-\alpha}$. Therefore, the optimal choice of parameters is $t = \varepsilon^{2-\alpha} = \varepsilon_0$. The computational cost is proportional to t^{-1}. On the other hand, if we choose $t = \kappa\varepsilon^{2-\alpha}$, where the positive κ is small enough such that $\mathbb{P}[S^\varepsilon = 1] = \lambda_\varepsilon t < 1$, then the computational cost to reach an error of level ε_0 is ε_0^{-1}.

6. *1JS with AR-correction:* In this case the weak error is of the order $t + \varepsilon^{3-\alpha}$. Then the calculation of cost follows as in the previous case, which gives a computational cost of ε_0^{-1}. The main difference with the previous scheme is that fitting the side condition $\mathbb{P}[S^\varepsilon = 1] = \lambda_\varepsilon t < 1$

becomes easier. In fact, we choose $\varepsilon = (\frac{t}{\kappa})^{1/\alpha}$, where the positive κ is small enough such that $\mathbb{P}[S^\varepsilon = 1] = \kappa\varepsilon^\alpha\lambda_\varepsilon = \kappa\varepsilon^\alpha \int_{|y|>\varepsilon} \nu(dy) < 1$, then the scheme is of order 1.

7. *2JS without AR-correction:* The weak error is of order $t^2 + \varepsilon^{2-\alpha}$ if the side conditions stated in Lemma 2.6 are satisfied.

   In this case, one needs to consider a non-regular choice of parameters due to these side conditions. In fact, besides the condition that the weak error has to be of order ε_0 one also needs to have that $\lambda_\varepsilon t < 2$. This raises an optimization problem which one can solve easily. The solution is to take $t = \kappa\varepsilon^{\alpha\vee(1-\frac{\alpha}{2})}$, where the positive κ is small enough such that $\mathbb{P}[S_1^\varepsilon = 1] = \lambda_\varepsilon t - \frac{\lambda_\varepsilon^2 t^2}{2} < 1$ and $\mathbb{P}[S_2^\varepsilon = 1] = \frac{\lambda_\varepsilon t}{2-\lambda_\varepsilon t} < 1$.

   This choice gives a final computational cost of $(\frac{1}{\varepsilon_0})^{\frac{1}{2}\vee\frac{\alpha}{2-\alpha}}$.

8. *2JS with AR-correction:* In this case the optimization problem that appeared in the previous case is simplified due to the higher weak rate of convergence and one may choose $t = \kappa\varepsilon^\alpha$, where the positive κ is small enough such that $\mathbb{P}[S_1^\varepsilon = 1] = \lambda_\varepsilon t - \frac{\lambda_\varepsilon^2 t^2}{2} < 1$ and $\mathbb{P}[S_2^\varepsilon = 1] = \frac{\lambda_\varepsilon t}{2-\lambda_\varepsilon t} < 1$, then the scheme is of order t^2. The computational cost to reach to error of level ε_0 is $\varepsilon_0^{-1/2}$.

Putting this information on a table for the case $\alpha < 1$, we obtain

method	ES	ES(AR)	JSAS	JSAS(AR)
cost	ε_0^{-1}	ε_0^{-1}	$(\varepsilon_0)^{-\frac{\alpha}{2-\alpha}}$	$\varepsilon_0^{-\alpha/2}$
t	$\varepsilon^{2-\alpha}$	$\varepsilon^{3-\alpha}$	–	–
ε	$\varepsilon_0^{\frac{1}{2-\alpha}}$	$\varepsilon_0^{\frac{1}{3-\alpha}}$	$\varepsilon_0^{\frac{1}{2-\alpha}}$	$\varepsilon_0^{\frac{1}{2}}$
N	t^{-1}	t^{-1}	λ_ε	λ_ε
method	1JS	1JS(AR)	2JS	2JS(AR)
cost	ε_0^{-1}	$(\varepsilon_0)^{-\frac{1}{2}\vee\frac{\alpha}{2-\alpha}}$	ε_0^{-1}	$\varepsilon_0^{-1/2}$
t	$\varepsilon^{2-\alpha}$	ε^{α}	$\varepsilon^{\alpha\vee(1-\frac{\alpha}{2})}$	ε^{α}
ε	$\varepsilon_0^{\frac{1}{2-\alpha}}$	$\varepsilon_0^{\frac{1}{3-\alpha}}$	$\varepsilon_0^{\frac{1}{2-\alpha}}$	$\varepsilon_0^{\frac{1}{3-\alpha}}$
N	t^{-1}	t^{-1}	t^{-1}	t^{-1}

From this table one can deduce that the JSAS methods have the lowest theoretical expected computational cost while the Euler scheme methods perform the worst.

This table assumes the general situation where one does not have information of how to generate the increments of the Lévy process exactly.

Let us now procceed with the experimental results. The estimator and variance of each scheme is plotted as a function of $\log(N)$ where N is the numbers of discretization points n between 0 and 1. The parameter λ_ε appearing in the JSAS method is chosen equal to N in order to allow for comparison of computational cost.

We have decided to use this as it would seem the most natural measure of computational time. The only case where this will differ with theoretical

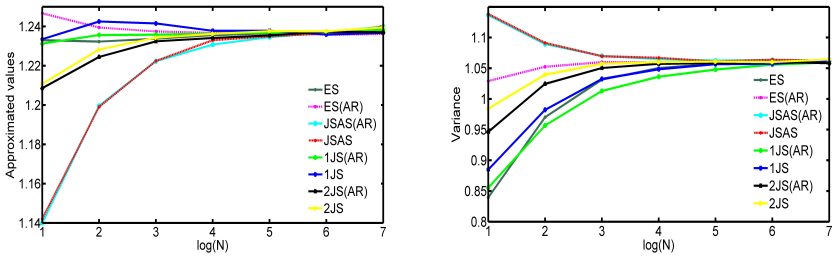

FIGURE 1. Numerical comparison of various schemes for estimating $\mathbb{E}[f(X_1)]$ with $\alpha = 0.1$. (Left: mean, right: variance)

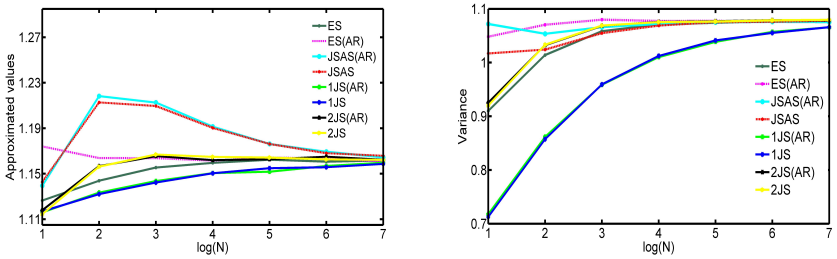

FIGURE 2. Numerical comparison of various schemes for estimating $\mathbb{E}[f(X_1)]$ with $\alpha = 0.9$. (Left: mean, right: variance)

computational time is in the case of the Euler scheme when all jumps have to be simulated.

For each point, we simulated 10^6 trajectories. In Figures 1 and 2, we see the convergence and the variance of each estimator. The computational times with respect to the case $\log(N) = 6$ are shown in Figure 3. This figure shows that the theoretical computational estimates are not necessarily accurate at this level and shows the difference of the constants in the asymptotic estimates. For example, the increase of computational time for the Euler scheme from $\alpha = 0.1$ to $\alpha = 0.9$ is due to the increase in the number of jumps. Even more, the fact that the JSAS schemes have a random number of partitions seems to play an important role in the computational time. In fact, asymptotically speaking, the number of calculations needed is a Poisson random variable with mean λ_ϵ which behaves like a Gaussian r.v. with variance proportional to λ_ϵ. From this figure, we can also see the increasing dependence of these constants with respect to the value of α.

Next, we perform the same simulation as above but with a different value of parameter a. More precisely, we choose $a = 5.0$, $\lambda_+ = 1.0$, $\alpha = 0.9, c_+ = 1, \gamma = 1$ and $c_- = 0$. The results are presented in Figures 4 and 5.

The conclusion is that in general the 1JS method is fast and gives good results for coefficients that do not oscillate too much. This contrasts with the

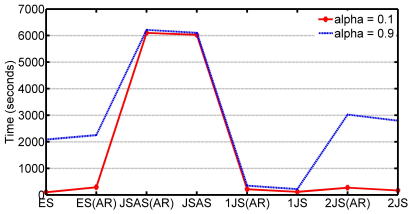

FIGURE 3. Computation time taken in the estimation of $\mathbb{E}[f(X_1)]$

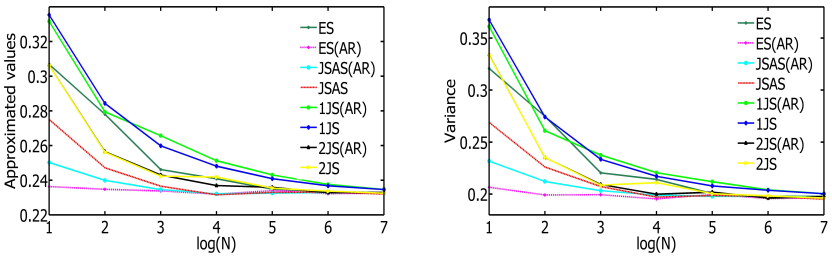

FIGURE 4. Numerical comparison of various schemes for estimating $\mathbb{E}[f(X_1)]$ with $a = 5.0$. (Left: mean, right: variance)

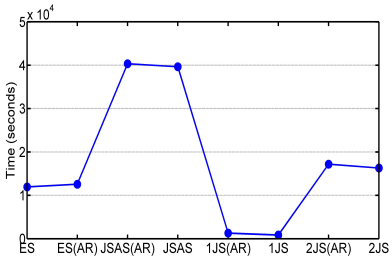

FIGURE 5. Computation time taken in the estimation of $\mathbb{E}[f(X_1)]$ with respect to $\log(N) = 6$

theoretical results shown in the previous table. This seems to be caused by the size of the constants in the error expansions.

If the coefficients have rapidly growing derivatives then the method looses accuracy and one may better use the JSAS method which may be time consuming. In between these two methods one has the 2JS method. Therefore a practical issue is how to determine before implementing the method which one should use and the range of applicability of each method.

3.1.2. The approximation schemes presented in Section 2 are only applicable for smooth functions f. However, in the next simulation, we will use this

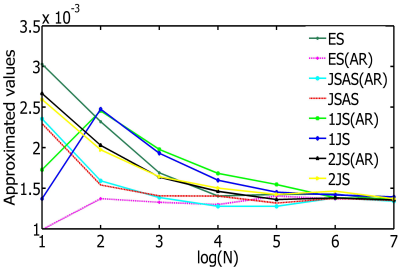

FIGURE 6. Numerical comparison of various schemes for estimating $\mathbb{P}[X_1 > 2]$

scheme to estimate the probability $\mathbb{P}(X_1 > x_0)$, or in other words, to estimate the expectation $\mathbb{E}(f(X_1))$ with $f(x) = 1_{x>x_0}$.

We choose $a = 5.0,\ \lambda_+ = 0.5,\ \alpha_+ = 0.9, c_+ = 1, \gamma = 1$ and $c_- = 0$. The results are presented in Figure 6. This study reveals that one may need to use an importance sampling method in order to improve the accuracy of the proposed method.

3.2. Example 2: Weak Approximation for an SDE Driven by a NIG Lévy Process

Let Z be a normal inverse Gaussian Lévy process whose characteristic function is defined by

$$\phi_t(u) = \mathbb{E}(e^{iuZ_t}) = \exp\Big\{ -\delta t\Big(\sqrt{\alpha^2 - (\beta - iu)^2} - \sqrt{\alpha^2 - \beta^2}\Big)\Big\}$$

where $\alpha > 0$ and $\beta \in (-\alpha, \alpha)$ and $\delta > 0$ are parameters. The Lévy density is given by

$$\nu(x) = \frac{\delta\alpha}{\pi}\frac{e^{\beta x}K_1(\alpha|x|)}{|x|},$$

where K is a modified Bessel function of the second kind. The NIG process can be defined as

$$Z_t = \theta Y_t + \sigma W_{Y_t}, \tag{3.1}$$

where W is a standard Brownian motion and Y is a inverse Gaussian subordinator: a pure jump Lévy process with Lévy density $\rho(x) = \frac{2}{\sqrt{2\kappa\pi}}\frac{e^{-\frac{x}{2\kappa}}}{|x|^{3/2}}$ and therefore $\rho_0 = 0.5$ in this case. The parameters (σ, θ, κ) and (α, β, δ) are related by

$$\begin{cases} \kappa = \frac{1}{\delta\sqrt{\alpha^2-\beta^2}} \\ \theta = \frac{\beta\delta}{\sqrt{\alpha^2-\beta^2}} \\ \sigma^2 = \frac{\delta}{\sqrt{\alpha^2-\beta^2}} \end{cases} \Leftrightarrow \begin{cases} \alpha = \frac{\sqrt{\theta^2+\sigma^2\kappa^{-1}}}{\sigma^2} \\ \beta = \theta\sigma^{-2} \\ \delta = \sigma\kappa^{-1/2}. \end{cases}$$

The representation (3.1) allows to simulate exact increments of NIG process in order to perform an Euler approximation scheme.

Let p_t be the density of $\theta t + \sigma W_t$, the density of the NIG process can be presented as $\nu(x) = \int_0^\infty p_t(x)\rho(t)dt$. We define the finite measure ν_ε via $\nu_\varepsilon(x) = \int_\varepsilon^\infty p_t(x)\rho(t)dt$, and then $\chi_\varepsilon(x) = \frac{\nu_\varepsilon(x)}{\nu(x)}$.

The constants $\lambda_\varepsilon, \gamma_\varepsilon, \sigma_\varepsilon$ are computed as follows

$$\lambda_\varepsilon = \sqrt{\frac{2}{\pi\kappa\varepsilon}}\exp\Big(-\frac{\varepsilon}{2\kappa}\Big) - \frac{2}{\kappa}N\Big(-\sqrt{\frac{\varepsilon}{\kappa}}\Big),$$

$$\gamma_\varepsilon = \theta - 2\theta N\Big(-\sqrt{\frac{\varepsilon}{\kappa}}\Big),$$

$$\sigma_\varepsilon^2 = (\sigma^2 + \kappa\theta^2)\Big(1 - 2N\Big(-\sqrt{\frac{\varepsilon}{\kappa}}\Big)\Big) - \sqrt{\frac{2\kappa\varepsilon}{\pi}}\exp\Big(-\frac{\varepsilon}{2\kappa}\Big)\theta^2,$$

where $N(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^x \exp(-\frac{x^2}{2})dx$.

We choose $\sigma = 0.5$; $\theta = 0.4$; $\kappa = 0.6$, and solve the one-dimensional SDE

$$dX_t = \sin(aX_t)d\tilde{Z}_t,$$

where $\tilde{Z}$ is the NIG process with drift adjusted to have $\mathbb{E}(\tilde{Z}_t) = 0$. In other words, $\tilde{Z}_t = Z_t - \theta t$. In this case, the values of λ_ε and σ_ε are the same as before but $\gamma_\varepsilon = -2\theta N\Big(-\sqrt{\frac{\varepsilon}{\kappa}}\Big)$.

We first use the representation (3.1) to define the Euler's scheme for X_1. Besides, we also use JSAS method and 1JS method to simulate X_1 as introduced in Section 2.4.5.

We now explain how to define the JSAS1 scheme (2.1)–(2.3) to simulate X_1:

1. $(T_i^\varepsilon)_{i\in\mathbb{N}}$ denotes jump times of a Poisson process whose intensity is λ_ε, $T_0^\varepsilon = 0$.
2. $(\Delta Z(T_i^\varepsilon))_{i\in\mathbb{N}}$ denotes a sequence of independent random variables whose density is $\frac{\nu_\varepsilon}{\lambda_\varepsilon}$.
3. The solution of the ODE $dX_t = \sin(aX_t)dt$ is approximated using the RK4 method.

A random variable with density $\frac{\nu_\varepsilon}{\lambda_\varepsilon}$ can be sampled using the following algorithm:

1. Sample a random variable Y with probability density $\frac{\rho(x)I_{x>\varepsilon}}{\lambda_\varepsilon}$ using the acceptance-rejection method (see [5], Example 6.9).
2. Conditional on Y, sample X with law p_Y. It means that X is sampled by $X = \theta Y + \sigma\sqrt{Y}W'$, where W' is a standard normal distributed random variable.

Next, we use JSAS2 method introduced in Section 2.2.2 to simulate X_1. In this setting, $\Omega(t) = \sigma_\varepsilon^2 h^2(Y_t^0)(t - \eta_t)$.

Finally, we use 1JS(AR) method defined in Section 2.4.5 to simulate X_1. We remark that this method already incorporates the Asmussen-Rosiński approximation in its definition. One can also do a similar computational cost study for this case. We do not give details but only the following table.

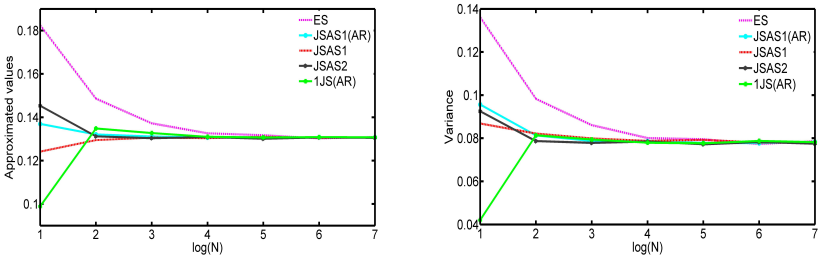

FIGURE 7. $a = 5,\ f(x) = 2 - 2\cos(x - X_0)$. Left: Mean. Right: Variance

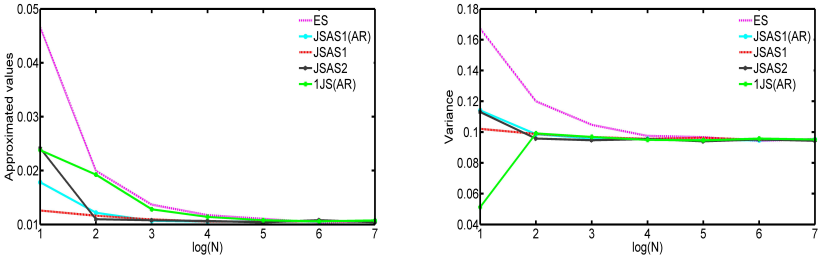

FIGURE 8. $a = 5,\ f(x) = 1_{\{x>3/2\}}$. Left: Mean. Right: Variance

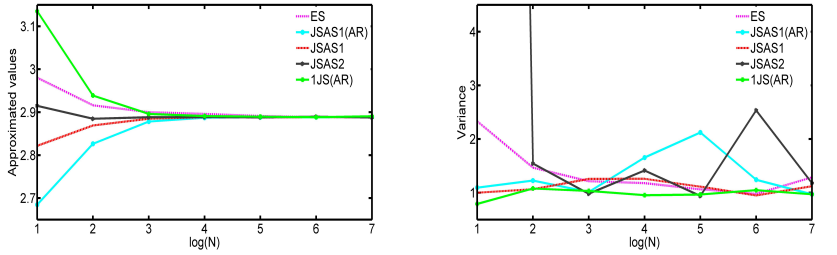

FIGURE 9. $a = 5,\ f(x) = e^x$. Left: Mean. Right: Variance

method	ES	JSAS1	JSAS1(AR)	JSAS2	1JS(AR)
cost	ε_0^{-1}	ε_0^{-1}	$\varepsilon_0^{-1/2}$	$\varepsilon_0^{-1/2}$	$\varepsilon_0^{-1/2}$

Figures 7, 8 and 9 show the Monte Carlo estimation for $\mathbb{E}[f(X_1)]$ with $f(x) = 2 - 2\cos(x - X_0)$, $f(x) = 1_{\{x>3/2\}}$ and $f(x) = e^x$, and the corresponding variances, respectively, with $a = 5.0$.

Figures 10, 11 and 12 show the Monte Carlo estimation for $\mathbb{E}[f(X_1)]$ with $f(x) = 2 - 2\cos(x - X_0)$, $f(x) = 1_{\{x>3/2\}}$ and $f(x) = e^x$, and the corresponding variances, respectively, with $a = 10.0$. The computational time

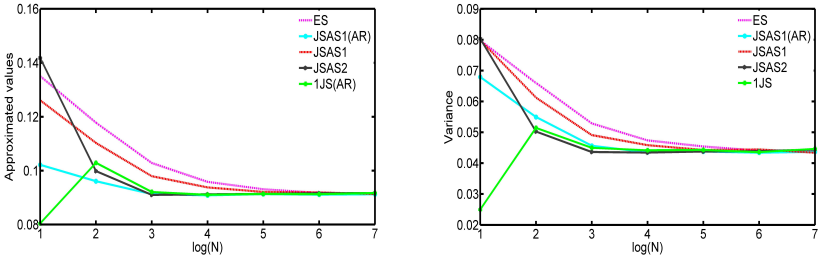

FIGURE 10. $a = 10,\ f(x) = 2 - 2\cos(x - X_0)$. Left: Mean. Right: Variance

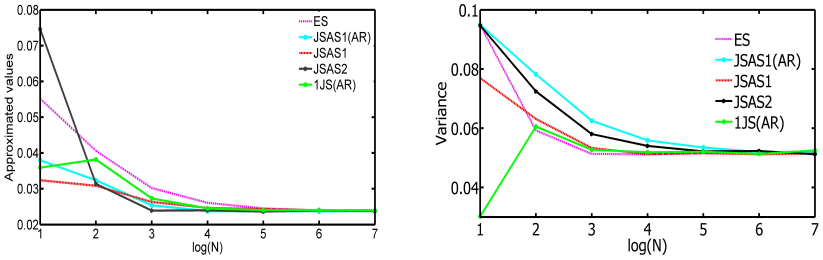

FIGURE 11. $a = 10,\ f(x) = 1_{\{x>3/2\}}$. Left: Mean. Right: Variance

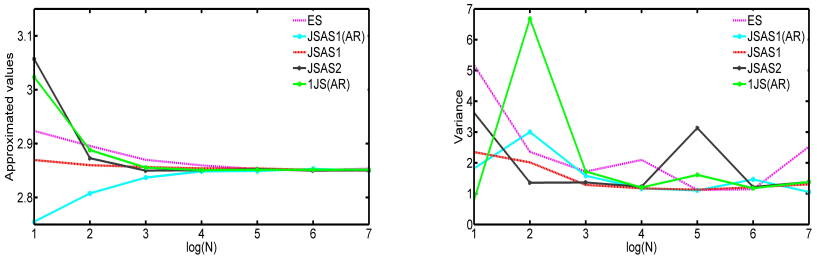

FIGURE 12. $a = 10,\ f(x) = e^x$. Left: Mean. Right: Variance

of each method with respect to the case $\log(N) = 7$ and $N_{MC} = 10^6$ are shown in Figure 13.

The conclusion is that, on one hand, JSAS methods have higher rates of convergence than the other methods. On the other hand, 1JS(AR) method defined in Section 2.4.5 is very fast and gives good results even when coefficients oscillate a lot.

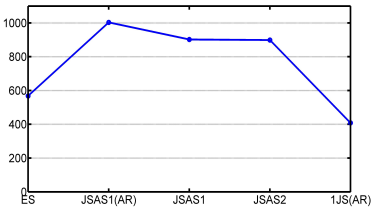

FIGURE 13. Computation time for estimating $\mathbb{E}f(X_1)$

3.3. Some Conclusions

Throughout the experiments, we see that there is a big gap between theoretical asymptotic values and the actual computational results. So far, one can see that the 1JS is a fairly efficient scheme in most situations considering its accuracy and computational time. If high accuracy is required then the JSAS or 2JS can be used. Further studies are needed which may also incorporate new schemes. We have striven here for generality and therefore many faster schemes may be provided for particular situations.

References

[1] D. Applebaum, *Lévy Processes and Stochastic Calculus*. Cambridge University Press, 2009.

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

[3] N. Bruti-Liberati and E. Platen, *Strong approximations of stochastic differential equations with jumps*. J. Comput. Appl. Math., **205 (2)** (2007), 982–1001.

[4] J.C. Butcher, *The Numerical Analysis of Ordinary Differential Equations: Runge-Kutta and General Linear Methods*. John Wiley & Sons, 1986.

[5] R. Cont and P. Tankov, *Financial Modelling with Jump Processes*. Chapman Hall CRC, 2004.

[6] N. Fournier, *Simulation and approximation of Lévy-driven stochastic differential equations*. ESAIM:PS, **15** (2011), 233–248.

[7] P. Glasserman and Z. Liu, *Sensitivity estimates from characteristic functions*. Operations Research November/December, **58 (6)** (2010), 1611–1623.

[8] A. Kohatsu-Higa and P. Tankov, *Jump-adapted discretization schemes for Lévy-driven SDEs*. Stochastic Process. Appl., **120 (11)** (2010), 2258–2285.

[9] M. Ninomiya and S. Ninomiya, *A new higher-order weak approximation scheme of stochastic differential equations and the Runge-Kutta method*. Finance and Stochastics, **13** (2009), 415–443.

[10] S. Ninomiya and N. Victoir, *Weak approximation of stochastic differential equations and application to derivative pricing*. Appl. Math. Finance, **15** (2008), 107–121.

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

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

[13] K. Sato, *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, 1999.

[14] H. Tanaka and A. Kohatsu-Higa, *An operator approach for Markov chain weak approximations with an application to infinite activity Lévy driven SDEs*. Ann. of Appl. Probab., **19 (3)** (2009), 1026–1062. (For an updated/corrected version see `http://www.math.ritsumei.ac.jp/~khts00`.)

Arturo Kohatsu-Higa
Ritsumeikan University and Japan Science and Technology Agency
Department of Mathematical Sciences
1-1-1 Nojihigashi, Kusatsu, Shiga, 525-8577, Japan
e-mail: `arturokohatsu@gmail.com`

Hoang-Long Ngo
Ritsumeikan University and Japan Science and Technology Agency
Department of Mathematical Sciences
1-1-1 Nojihigashi, Kusatsu, Shiga, 525-8577, Japan
and
Hanoi National University of Education
136 Xuan Thuy, Cau Giay, Ha Noi, Vietnam
e-mail: `lenhholong@yahoo.com`