We present new algorithms for weak approximation of stochastic differential equations driven by pure jump Lévy processes. The method is built upon adaptive non-uniform discretization based on the jump times of the driving process coupled with suitable approximations of the solutions between these jump times. Our technique avoids the costly simulation of the increments of the Lévy process and in many cases achieves better convergence rates than the traditional schemes with equal time steps. To illustrate the method, we consider applications to simulation of portfolio strategies and option pricing in the Libor market model with jumps.

Key words: Lévy-driven stochastic differential equation, Euler scheme, jump-adapted discretization, weak approximation, Libor market model with jumps.

2000 Mathematics Subject Classification: Primary 60H35, Secondary 65C05, 60G51.

1 Introduction

Let $Z$ be a $d$-dimensional Lévy process without diffusion component, that is,

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

*Corresponding author
Here $\gamma \in \mathbb{R}^d$, $N$ is a Poisson random measure on $\mathbb{R}^d \times [0, \infty)$ with intensity $\nu$ satisfying $\int 1 \wedge |y|^2 \nu(dy) < \infty$ and $\tilde{N}(dy, ds) = N(dy, ds) - \nu(dy)ds$ denotes the compensated version of $N$. Further, let $X$ be an $\mathbb{R}^n$-valued adapted stochastic process, unique solution of the stochastic differential equation

$$X_t = X_0 + \int_0^t h(X_s^-)dZ_s, \quad t \in [0, 1],$$

(1)

where $h$ is an $n \times d$ matrix.

The traditional uniformly-spaced discretization schemes for (1) [15, 10] suffer from two difficulties: first, in general, we do not know how to simulate the driving Lévy process and second, a large jump of $Z$ occurring between two discretization points can lead to an important discretization error. On the other hand, in many cases of interest, the Lévy measure of $Z$ is known and one can easily simulate the jumps of $Z$ bigger in absolute value than a certain $\varepsilon > 0$.

As our benchmark example, consider the one-dimensional tempered stable Lévy process [5] which has the Lévy density

$$\nu(x) = \frac{C_-e^{-\lambda_-|x|}}{|x|^{1+\alpha_-}}1_{x<0} + \frac{C_+e^{-\lambda_+|x|}}{|x|^{1+\alpha_+}}1_{x>0}.$$  

(2)

This process, also known as the CGMY model [3] when $C_- = C_+$ and $\alpha_- = \alpha_+$, is a popular representation for the logarithm of stock price in financial markets. No easy algorithm is available for simulating the increments of this process (cf [12, 14]), however, it is straightforward to simulate random variables with density

$$\nu(x)1_{|x|>\varepsilon} \int_{|x|>\varepsilon} \nu(x)dx$$  

(3)

using the rejection method [5, example 6.9]. Sometimes, other approximations of $Z$ by a compound Poisson process, such as the series representations [16] lead to simpler simulation algorithms.

A natural idea due to Rubenthaler [17] (in the context of finite-intensity jump processes, this idea appears also in [2, 13]), is to replace the process $Z$ with a suitable compound Poisson approximation and place the discretization points at the jump times of the compound Poisson process. When the jumps of $Z$ are highly concentrated around zero, however, this approximation is too rough and can lead to convergence rates which are arbitrarily slow.

In this paper we are interested in approximating the solution of (1) in the weak sense. We use the idea of Asmussen and Rosiński [1] (see also [4]) and replace the small jumps of $Z$ with a Brownian motion between the jump times of the compound Poisson approximation. We then replace the solution of the continuous SDE between the jump times by a suitable approximation, which is explicitly computable. Combined, these two steps lead to a much lower discretization error than in [17]. If $N$ is the average number of operations required to simulate a single trajectory, the discretization error of our method converges
to zero faster than $\frac{1}{\sqrt{N}}$ in all cases and faster than $\frac{1}{N}$ if the Lévy measure of $Z$ is locally symmetric near zero\(^1\). In practice, the convergence rates are always better than these worst-case bounds: for example, if $\alpha_- = \alpha_+ = 1.8$ in (2), our method has the theoretical convergence rate of $N^{-1.22}$ (compared to only $N^{-0.11}$ without the Brownian approximation) and when $\alpha_- = \alpha_+ = 0.5$, the rate of our method is $N^{-7}$.

Consider a family of measurable functions $(\chi^\varepsilon)_{\varepsilon > 0} : \mathbb{R}^d \to [0,1]$ such that $\int_{\mathbb{R}^d} \chi^\varepsilon(y) \nu(dy) < \infty$ and $\int_{|y| > 1} |y|^2(1 - \chi^\varepsilon)(y)\nu(dy) < \infty$ for all $\varepsilon > 0$ and $\lim_{\varepsilon \downarrow 0} \chi^\varepsilon(y) = 1$ for all $y \neq 0$. The Lévy measure $\nu$ will be approximated by finite measures $\chi^\varepsilon \nu$. The most simple such approximation is $\chi^\varepsilon(y) := 1_{|y| > \varepsilon}$, but others can also be useful (see Example 7 below). We denote by $N^\varepsilon$ a Poisson random measure with intensity $\chi^\varepsilon \nu \times ds$ and by $\hat{N}^\varepsilon$ its compensated Poisson random measure. Similarly, we denote by $\bar{N}^\varepsilon$ the compensated Poisson random measure with intensity $\bar{\chi}^\varepsilon \nu \times ds$, where $\bar{\chi}^\varepsilon := 1 - \chi^\varepsilon$. The process $Z$ can then be represented in law as follows:

$$Z_t = \gamma_t^\varepsilon + Z^\varepsilon_t + R^\varepsilon_t,$$

$$\gamma_t^\varepsilon = \gamma - \int_{|y| \leq 1} y\chi^\varepsilon \nu(dy) + \int_{|y| > 1} y\bar{\chi}^\varepsilon \nu(dy),$$

$$Z^\varepsilon_t = \int_0^t \int_{\mathbb{R}^d} yN^\varepsilon(dy,ds),$$

$$R^\varepsilon_t = \int_0^t \int_{\mathbb{R}^d} y\hat{N}^\varepsilon(dy,ds).$$

We denote by $\lambda^\varepsilon = \int_{\mathbb{R}^d} \chi^\varepsilon(y) \nu(dy)$ the intensity of $Z^\varepsilon$, by $(N^\varepsilon_t)$ the Poisson process which has the same jump times as $Z^\varepsilon$, and by $T^\varepsilon_i$, $i \in \mathbb{N}$, the jump times of $Z^\varepsilon$. Furthermore, we denote by $\Sigma^\varepsilon$ the variance-covariance matrix of $R^\varepsilon_t$:

$$\Sigma^\varepsilon_{ij} = \int_{\mathbb{R}^d} y_i y_j \bar{\chi}^\varepsilon \nu(dy),$$

and in the one-dimensional case we set $\sigma^2 := \Sigma^\varepsilon_{11}$. Sometimes we shall use the following technical assumption on $(\chi^\varepsilon)$:

(A) $\forall n \geq 2$, $\exists C$ such that

$$\int_{\mathbb{R}^d} |z|^{n+1} \bar{\chi}^\varepsilon \nu(dz) \leq C \int_{\mathbb{R}^d} |z|^n \bar{\chi}^\varepsilon \nu(dz)$$

for $\varepsilon$ sufficiently small.

This assumption, roughly, means that the approximation $(\chi^\varepsilon)$ does not remove small jumps faster than big jumps, and it is clearly satisfied by $\chi^\varepsilon(y) = 1_{|y| > \varepsilon}$.

\(^1\)For the process (3) this corresponds to having $\alpha_+ = \alpha_- = 1$ and $C_+ = C_-$. 

Supposing that no further discretization is done between the jump times of $Z^\varepsilon$, the computational complexity of simulating a single approximate trajectory is proportional to the random number $N^\varepsilon_1$. To compare our method to the traditional equally-spaced discretizations, we measure instead the computational complexity by the average number of jumps of $Z^\varepsilon$ on $[0, 1]$, equal to $\lambda^\varepsilon$. We say that the approximation $\hat{X}^\varepsilon$ converges weakly for some order $\alpha > 0$ if, for a sufficiently smooth function $f$,

$$|E[f(X_1)] - E[f(\hat{X}_1)]| \leq K\lambda^\varepsilon - \alpha$$

for some $K > 0$ and all $\varepsilon$ sufficiently small.

In this paper, we provide two versions of the jump-adapted discretization scheme for pure jump Lévy processes. The scheme presented in Section 2 is easier to use and implement but works in the case $d = n = 1$. A fully general scheme is then presented in Section 3. Section 4 presents two detailed examples where our method is applied to the problems of long term portfolio simulation and of option pricing in the Libor market model with jumps.

## 2 One-dimensional SDE

For our first scheme we take $d = n = 1$ and consider the ordinary differential equation

$$dX_t = h(X_t)dt, \quad X_0 = x. \quad (4)$$

In the one-dimensional case, the solution to this equation can always be written:

$$X_t := \theta(t; x) = F^{-1}(F(x) + \sigma \varepsilon (W(t) - W(T^\varepsilon_i)) - \frac{1}{2} h'(\hat{X}(T^\varepsilon_i))\sigma^2\varepsilon (T^\varepsilon_{i+1} - T^\varepsilon_i); \hat{X}(T^\varepsilon_i)). \quad (5)$$

Similarly, for any other point $t \in (T^\varepsilon_i, T^\varepsilon_{i+1})$, we define

$$\hat{X}(t) = \theta(\gamma \varepsilon (t - \eta_t) + \sigma \varepsilon (W(t) - W(\eta_t)) - \frac{1}{2} h'(\hat{X}(\eta_t))\sigma^2\varepsilon (t - \eta_t); \hat{X}(\eta_t)) \quad (7)$$

where we set $\eta_t := \sup\{T^\varepsilon_i : T^\varepsilon_i \leq t\}$. Here $W$ denotes a one dimensional Brownian motion.

Therefore, the idea is to replace the original equation with an Asmussen-Rosiński type approximation which is explicitly solvable between the times of
big jumps and is exact for all \( h \) if the driving process is deterministic, and for affine \( h \) in all cases. The purpose of this construction becomes clear from the following lemma.

**Lemma 1.** Let \( h \in C^1 \). The process \((\hat{X}(t))\) defined by (5)-(7) is the solution of the stochastic differential equation

\[
d\hat{X}_t = h(\hat{X}_t) \left\{ dZ_t + \sigma dW_t + \gamma dt + \frac{1}{2} (h'(\hat{X}_t) - h'(\hat{X}_{\eta(t)})) \right\}.
\]

**Proof.** It is enough to show that the process \( Y_t := \theta(\gamma t + \sigma W_t - \frac{1}{2} h'(x) \sigma^2 t; x) \) (8) is the solution of the continuous SDE

\[
dY_t = h(Y_t) \left\{ \sigma dW_t + \gamma dt + \frac{1}{2} (h'(Y_t) - h'(x)) \sigma^2 dt \right\}.
\]

This follows by an application of Itô formula to (8) using

\[
\frac{\partial \theta(t, z)}{\partial t} = h(\theta(t, z))
\]

\[
\frac{\partial^2 \theta(t, z)}{\partial t^2} = h'(\theta(t, z)) h(\theta(t, z)).
\]

\( \square \)

For the convergence analysis, we introduce two sets of conditions (parameterized by an integer number \( n \)):

\( (H_n) \) \( f \in C^n, \ h \in C^n, \ f^{(k)} \) and \( h^{(k)} \) are bounded for \( 1 \leq k \leq n \) and \( \int |z|^{2n} \nu(dz) < \infty \).

\( (H_n') \) \( f \in C^n, \ h \in C^n, \ h^{(k)} \) are bounded for \( 1 \leq k \leq n \), \( f^{(k)} \) have at most polynomial growth for \( 1 \leq k \leq n \) and \( \int |z|^k \nu(dz) < \infty \) for all \( k \geq 1 \).

**Theorem 2.**

(i) Assume \( (H_3) \) or \( (H'_3) + (A) \). Then

\[
|E[f(\hat{X}_1) - f(X_1)]| \leq C \left( \frac{\sigma^2}{\lambda \epsilon} (\sigma^2 + |\gamma|) + \int \frac{y^3}{\lambda \epsilon} \nu(dy) \right).
\]

(ii) Assume \( (H_4) \) or \( (H'_4) + (A) \), let \( \gamma_\epsilon \) be bounded and let the measure \( \nu \) satisfy

\[
\int y^4 \lambda \epsilon \nu(dy) \leq \int |y|^4 \lambda \epsilon \nu_0(dy) (9)
\]
for some measure \( \nu_0 \). Then
\[
|E[f(\hat{X}_1) - f(X_1)]| \leq C \left( \frac{\sigma^2}{\lambda \varepsilon} (\sigma^2 + |\gamma_\varepsilon|) + \int_\mathbb{R} |y|^4 \chi_\varepsilon (\nu_0 + \nu)(dy) \right)
\]
and in particular,
\[
|E[f(\hat{X}_1) - f(X_1)]| \leq C \left( \frac{\sigma^2}{\lambda \varepsilon} + \int_\mathbb{R} |y|^4 \chi_\varepsilon (\nu_0 + \nu)(dy) \right).
\]

Remark 3. (i) Under the polynomial growth assumptions (H_3') or (H_4'), the constants in the above theorem may depend on the initial value \( x \).

(ii) Condition (9) is satisfied, for example, if \( \chi_\varepsilon(y) = \chi_\varepsilon(-y) \) for all \( y \) and \( \varepsilon \) and if \( \nu \) is locally symmetric near zero. That is, \( \nu(dy) = (1 + \xi(y)) \nu_0(dy) \), where \( \nu_0 \) is a symmetric measure and \( \xi(y) = O(y) \) for \( y \to 0 \).

Remark 4. The convergence rates given in Theorem 2 are obtained under the assumption that the equation (4) is solved explicitly. If a discretization scheme is used for this equation as well, the total error will be bounded from below by the error of this scheme. However, usually this constraint is not very important since it is easy to construct a discretization scheme for (4) with a high convergence order, for example, the classical Runge-Kutta scheme has convergence order 4.

Corollary 5 (Worst-case convergence rates). Let \( \chi_\varepsilon(x) = 1_{|x| > \varepsilon} \). Then, under the conditions of part (i) of Theorem 2,
\[
|E[f(\hat{X}_1) - f(X_1)]| \leq o(\lambda^{-\frac{1}{2}}),
\]
and under the conditions of part (ii) of Theorem 2,
\[
|E[f(\hat{X}_1) - f(X_1)]| \leq o(\lambda^{-1}).
\]

Proof. These worst-case bounds follow from the following estimates. First, for every Lévy process, \( \sigma_\varepsilon \to 0 \) as \( \varepsilon \to 0 \). By the dominated convergence theorem, \( \varepsilon^2 \lambda_\varepsilon \to 0 \) as \( \varepsilon \to 0 \). This implies
\[
\sqrt{\lambda_\varepsilon} \int_{|y| \leq \varepsilon} |y|^3 \nu(dy) \leq \sqrt{\varepsilon^2 \lambda_\varepsilon \sigma^2_\varepsilon} \xrightarrow[\varepsilon \to 0]{} 0
\]
and similarly
\[
\lambda_\varepsilon \int_{|y| \leq \varepsilon} |y|^4 \nu(dy) \leq \varepsilon^2 \lambda_\varepsilon \sigma^2_\varepsilon \xrightarrow[\varepsilon \to 0]{} 0.
\]

Example 6 (Stable-like behavior). Once again, let \( \chi_\varepsilon(x) = 1_{|x| > \varepsilon} \), and assume that the Lévy measure has an \( \alpha \)-stable-like behavior near zero, meaning that \( \nu \) has a density \( \nu(z) \) satisfying
\[
\nu(z) = \frac{g(z)}{|z|^{1+\alpha}},
\]
where $g$ has finite nonzero right and left limits at zero as, for example, for the tempered stable process (CGMY) [5]. Then $\sigma^2 = O(\varepsilon^{2-\alpha})$, $\lambda_\varepsilon = O(\varepsilon^{-\alpha})$,

$$\int_{|y|\leq \varepsilon} |y|^3 \nu(dy) = O(\varepsilon^{3-\alpha}) \quad \text{and} \quad \int_{|y|\leq \varepsilon} |y|^4 \nu(dy) = O(\varepsilon^{4-\alpha}),$$

so that in general for $\alpha \in (0, 2)$

$$|E[f(\hat{X}_1) - f(X_1)]| \leq O(\lambda_\varepsilon^{1-\frac{3}{4}}),$$

and if the Lévy measure is locally symmetric near zero and $\alpha \in (0, 1),$

$$|E[f(\hat{X}_1) - f(X_1)]| \leq O(\lambda_\varepsilon^{1-\frac{4}{5}}).$$

**Example 7 (Simulation using series representation).** In this example we explain why it can be useful to make other choices of truncation than $\chi_\varepsilon(x) = 1_{|x|>\varepsilon}$. The gamma process has Lévy density $\nu(z) = ce^{-\lambda z}z^{-1}1_{z>0}$. If one uses the truncation function $\chi_\varepsilon(x) = 1_{|x|>\varepsilon}$, one will need to simulate random variables with law $\frac{ce^{-\lambda x}z^{-1}1_{z>\varepsilon}}{ce^{-\lambda x}1_{z>\varepsilon}}$, which may be costly. Instead, one can use one of the many series representations for the gamma process [16]. Maybe the most convenient one is

$$X_t = \sum_{i=1}^{\infty} \lambda^{-1} e^{-\Gamma_i/c} V_i U_i, \quad 1_{U_i \leq t},$$

where $(\Gamma_i)$ is a sequence of jump times of a standard Poisson process, $(U_i)$ is an independent sequence of independent random variables uniformly distributed on $[0, 1]$ and $(V_i)$ is an independent sequence of independent standard exponential random variables. For every $\tau > 0$, the truncated sum

$$X_t = \sum_{i: \Gamma_i < \tau} \lambda^{-1} e^{-\Gamma_i/c} V_i 1_{U_i \leq t},$$

defines a compound Poisson process with Lévy density

$$\nu_\tau(x) = \frac{e^{-\lambda x}}{x} \left[ e^{-\lambda x} - e^{-\lambda x e^{\tau/c}} \right] 1_{x>0},$$

and therefore this series representation corresponds to

$$\chi_\varepsilon(x) = 1 - e^{\frac{c}{\lambda}x(e^{\tau/c} - 1)},$$

where $\varepsilon$ can be linked to $\tau$, for example, by setting $\tau = \frac{1}{\varepsilon}$.

Theorem 2 will be proved after a series of lemmas.

**Lemma 8 (Bounds on moments of $X$ and $\hat{X}$).** Assume

$$\int_{\mathbb{R}} |z|^p \nu(dz) < \infty \quad \text{for some } p \geq 2,$$

7
\( h \in C^1(\mathbb{R}) \) and

\[ |h(x)| \leq K(1 + |x|) \quad \text{and} \quad |h'(x)| \leq K \quad \text{for some } K < \infty. \]

Then there exists a constant \( C > 0 \) (which may depend on \( p \) but not on \( \varepsilon \)) such that

\[
E[ \sup_{0 \leq s \leq t} |X_s|^p ] \leq C(1 + |x|^p), \tag{10}
\]

\[
E[ \sup_{0 \leq s \leq t} |\hat{X}_s|^p ] \leq C(1 + |x|^p). \tag{11}
\]

Proof. We shall concentrate on the bound (11), the other one will follow by making \( \varepsilon \) go to zero. We have

\[
\hat{X}_t = x + \int_0^t h(\hat{X}_s)\tilde{\gamma}_s ds + \int_0^t h(\hat{X}_s) d\tilde{Z}_s
\]

with

\[
\tilde{\gamma}_t = \gamma + \int_{|z| > 1} z\nu(dz) + \frac{1}{2} \sigma^2 \{ h'(\hat{X}_t) - h'(\hat{X}_{\eta(t)}) \},
\]

\[
\tilde{Z}_t = \sigma\varepsilon W_t + \int_0^t \int_{\mathbb{R}} z\hat{N}_\varepsilon(dz, ds).
\]

Observe that \( \tilde{\gamma}_t \) is bounded uniformly on \( t \) and on \( \varepsilon \). Using a predictable version of the Burkholder-Davis-Gundy inequality [6, lemma 2.1], we then obtain for \( t \leq 1 \) and for some constant \( C < \infty \) which may change from line to line

\[
E[ \sup_{0 \leq s \leq t} |\hat{X}_s|^p ] \leq CE \left[ |x|^p + \int_0^t |h(\hat{X}_s)|^p |\tilde{\gamma}_s|^p ds \right. \\
+ \left. \left( \int_0^t h^2(\hat{X}_s)(\sigma^2 + \int_{\mathbb{R}} z^2\hat{\nu}(dz)) ds \right)^{p/2} + \int_0^t |h(\hat{X}_s)|^p ds \int_{\mathbb{R}} |z|^p\hat{\nu}(dz) \right]^{p/2}
\]

\[
\leq CE \left[ |x|^p + \int_0^t |h(\hat{X}_s)|^p ds \right] \leq CE \left[ 1 + |x|^p + \int_0^t |\hat{X}_s|^p ds \right].
\]

The bound (11) now follows from Gronwall’s inequality. \( \Box \)

Lemma 9 (Derivatives of the flow). Let \( p \geq 2 \) and for an integer \( n \geq 1 \) assume

\[
\int_{\mathbb{R}} |z|^n \hat{\nu}(dz) < \infty,
\]

\( h \in C^n(\mathbb{R}) \) and

\[ |h(x)| \leq K(1 + |x|) \quad \text{and} \quad |h^{(k)}(x)| \leq K \]

for some \( K < \infty \) and all \( k \) with \( 1 \leq k \leq n \). Then

\[
E \left[ \left| \frac{\partial^k}{\partial x^k} X_T^{(t,x)} \right|^p \right] < \infty
\]

for all \( k \) with \( 1 \leq k \leq n \).
Proof. See the proof of lemma 4.2 in [15].

**Lemma 10.** Let $u(t,x) := E^{(t,x)}[f(X_T)]$.

(i) Assume $(H_n)$ with $n \geq 2$. Then $u \in C^{1,n}([0,1] \times \mathbb{R})$, $\frac{\partial u}{\partial x}$ are uniformly bounded for $1 \leq k \leq n$ and $u$ is a solution of the equation

$$
\frac{\partial u}{\partial t}(t,x) + \gamma \frac{\partial u}{\partial x}(t,x)h(x) + \int_{|y| \leq 1} \left( u(t,x+hy) - u(t,x) - \frac{\partial u}{\partial x}(t,x)hy \right) \nu(dy)
+ \int_{|y| > 1} (u(t,x+hy) - u(t,x)) \nu(dy) = 0
$$

(12)

and $u(1,x) = f(x)$.

(ii) Assume $(H'_n)$ with $n \geq 2$. Then $u \in C^{1,n}([0,1] \times \mathbb{R})$, $u$ is a solution of equation (12) and there exist $C < \infty$ and $p > 0$ with

$$
\left| \frac{\partial ^k u(t,x)}{\partial x^k} \right| \leq C(1 + |x|^p)
$$

for all $t \in [0,1]$, $x \in \mathbb{R}$ and $1 \leq k \leq n$.

**Proof.** The derivative $\frac{\partial u}{\partial x}$ satisfies

$$
\frac{\partial u(t,x)}{\partial x} = E^{(t,x)} \left[ f'(X^t(x)) \frac{\partial}{\partial x} X^t(x) \right].
$$

The interchange of the derivative and the expectation is justified using lemma 9. The boundedness under $(H_n)$ or the polynomial growth under $(H'_n)$ then follow from lemmas 8 and 9. The other derivatives with respect to $x$ are obtained by successive differentiations under the expectation. The derivative with respect to $t$ is obtained from Itô’s formula applied to $f(X^t(x))$. In fact, note that since $Z$ is a Lévy process and $h$ does not depend on $t$, $E[f(X^t(x))] = E[f(X^0(x))]$ and hence it is sufficient to study the derivative $\frac{\partial}{\partial x} E[(X^t(x))].$ The Itô formula yields

$$
E[f(X_t)] = f(x) + \int_0^t f'(X_{s-})h(X_{s-})dZ_s
+ E \int_0^t \int_{\mathbb{R}} \{ f(X_{s-} + h(X_{s-})z) - f(X_{s-}) - f'(X_{s-})h(X_{s-})z \} \nu(dz,ds).
$$

Denoting by $\tilde{Z}$ the martingale part of $Z$ and by $\tilde{\gamma} := \gamma + \int_{|z| > 1} z \nu(dz)$ the residual drift, and using lemma 8, we get

$$
E[f(X_t)] = f(x) + \int_0^t ds \tilde{\gamma} E[f'(X_s)h(X_s)]
+ \int_0^t ds E \left[ \int_{\mathbb{R}} \{ f(X_s + h(X_s)z) - f(X_s) - f'(X_s)h(X_s)z \} \nu(dz) \right],
$$
and therefore
\[
\frac{\partial E[f(X_t)]}{\partial t} = \gamma E[f'(X_t)h(X_t)] + E\left[ \int_{\mathbb{R}} \left( f(X_t + h(X_t)z) - f(X_t) - f'(X_t)h(X_t)z \right) \nu(dz) \right] \\
= \gamma E[f'(X_t)h(X_t)] + \int_0^1 d\theta E\left[ \int_{\mathbb{R}} (1 - \theta)h^2(X_t)z^2 \frac{\partial^2 f(X_t + \theta z h(X_t))}{\partial x^2} \nu(dz) \right].
\]

Now, once again, lemma 8 allows to prove the finiteness of this expression. Finally, equation (12) is a consequence of Itô’s formula applied, this time, to \( u(t, X_t) \).

**Lemma 11.** Let \( f : [0, \infty) \to [0, \infty) \) be an increasing function. Then
\[
E\left[ \int_0^1 f(t - \eta(t))dt \right] \leq E[f(\tau)],
\]
where \( \tau \) is an exponential random variable with intensity \( \lambda_\epsilon \). In particular,
\[
E\left[ \int_0^1 (t - \eta(t))^{p+1}dt \right] \leq \frac{(p + 1)!}{\lambda_\epsilon^{p+1}}.
\]

**Proof.** Let \( k_\epsilon = \sup\{k : T_k^\epsilon < 1\} \). Then
\[
E\left[ \int_0^1 f(t - \eta(t))dt \right] = E\left[ \sum_{l=1}^{k_\epsilon} \int_{T_{l-1}^\epsilon}^{T_l^\epsilon} f(t - T_{l-1}^\epsilon)dt + \int_{T_{k_\epsilon}}^1 f(t - T_{k_\epsilon})dt \right] \\
= E\left[ \sum_{l=1}^{k_\epsilon} \int_{T_{l-1}^\epsilon}^{T_l^\epsilon} f(T_l^\epsilon - t)dt + \int_{T_{k_\epsilon}}^1 f(1 - t)dt \right] \\
\leq E\left[ \sum_{l=1}^{k_\epsilon} \int_{T_{l-1}^\epsilon}^{T_l^\epsilon} f(T_l^\epsilon - t)dt + \int_{T_{k_\epsilon}}^1 f(T_{k_\epsilon+1} - t)dt \right] \\
= E\left[ \int_0^1 f(\inf\{T_l^\epsilon : T_l^\epsilon > t\} - t)dt \right] = E[f(\tau)].
\]

The second statement of the lemma is a direct consequence of the first one. \( \square \)

**Proof of theorem 2.** From Itô’s formula and lemmas 8 and 10,
\[
E[f(\hat{X}_t) - f(X_t)] = E[u(1, \hat{X}_1) - u(0, X_0)] \\
= \int_0^1 dt E\left[ \frac{1}{2} \frac{\partial^2 u(t, \hat{X}_t)}{\partial x^2} \sigma^2(\hat{X}_t) \\
- \int_{\mathbb{R}} \hat{X}_t \nu(dy)u(t, \hat{X}_t + h(\hat{X}_t)y) - u(t, \hat{X}_t) - \frac{\partial u(t, \hat{X}_t)}{\partial x} h(\hat{X}_t)y \right] \\
+ \int_0^1 dt E\left[ \frac{1}{2} \sigma^2 \frac{\partial u(t, \hat{X}_t)}{\partial x} h(\hat{X}_t)(h'(\hat{X}_t) - h'(\hat{X}_{\eta(t)})) \right]. (13)
\]
Denote the expectation in (13) by $A_t$ and the one in (14) by $B_t$. From lemmas 10 and 8, we then have, under the hypothesis $(H_3)$,
\[
|A_t| \leq \left| E \left[ \int_{\mathbb{R}} \frac{1}{2} y^3 h^3(\hat{X}_t) \bar{\chi}_\varepsilon \int_0^1 (1 - \theta)^2 \frac{\partial^4 u(t, \hat{X}_t + \theta y h(\hat{X}_t))}{\partial x^4} d\theta \nu(dy) \right] \right| \\
\leq CE[1 + |h|^3(\hat{X}_t)] \int_{\mathbb{R}} |y|^3 \bar{\chi}_\varepsilon \nu(dy) \leq C(1 + |x|^3) \int_{\mathbb{R}} |y|^3 \bar{\chi}_\varepsilon \nu(dy).
\]

If, instead, the polynomial growth condition $(H'_3)$ is satisfied, then for some $p > 0$,
\[
A_t \leq CE \left[ \int_{\mathbb{R}} |y|^3 (1 + |h|^3(\hat{X}_t))(1 + |\hat{X}_t|^p + |y h(\hat{X}_t)|^p) \bar{\chi}_\varepsilon \nu(dy) \right]
\]
and once again, we use lemma 8 together with the assumption $(A)$, because the Lévy measure is now supposed integrable with any power of $y$. Under the condition of part (ii) and the hypothesis $(H_4)$,
\[
|A_t| \leq \left| E \left[ \frac{\partial^3 u(t, \hat{X}_t)}{\partial x^3} h^3(\hat{X}_t) \int_{\mathbb{R}} y^3 \bar{\chi}_\varepsilon \nu(dy) \right] \right| \\
+ \left| E \left[ \int_{\mathbb{R}} \frac{1}{24} y^4 h^4(\hat{X}_t) \int_0^1 (1 - \theta)^3 \frac{\partial^4 u(t, \hat{X}_t + \theta y h(\hat{X}_t))}{\partial x^4} d\theta \bar{\chi}_\varepsilon \nu(dy) \right] \right| \\
\leq C(1 + |x|^4) \int_{\mathbb{R}} |y|^4 \bar{\chi}_\varepsilon (\nu_0 + \nu)(dy).
\]
The case of $(H'_4)$ is treated as above. To analyze the term $B_t$, define
\[
H(t, x) = \frac{\partial u(t, x)}{\partial x} h(x)
\]
and assume, to fix the notation, that $(H_3)$ is satisfied. Then, once again by
Note that lemmas 10 and 8, Taylor formula and the Cauchy-Schwartz inequality,

\[
|B_t| \leq \frac{1}{2} \sigma^2 \left| E\left[ (H(t, \tilde{X}_t) - H(t, \tilde{X}_{\eta(t)}))(h'(\tilde{X}_t) - h'(\tilde{X}_{\eta(t)})) \right] \right|
+ \frac{1}{2} \sigma^2 \left| E \left[ H(t, \tilde{X}_{\eta(t)})h'(\tilde{X}_t) - h'(\tilde{X}_{\eta(t)}) \right] \right|
\leq \frac{1}{2} \sigma^2 \left| E \left[ (\tilde{X}_t - \tilde{X}_{\eta(t)})^2 \int_0^1 d\theta \frac{\partial H}{\partial x}(\tilde{X}_{\eta(t)} + \theta(\tilde{X}_t - \tilde{X}_{\eta(t)})) \right] \right|
\times \int_0^1 d\theta h''(\tilde{X}_{\eta(t)} + \theta(\tilde{X}_t - \tilde{X}_{\eta(t)})) \right|
+ \frac{1}{2} \sigma^2 \left| E \left[ (\tilde{X}_t - \tilde{X}_{\eta(t)})^2 \int_0^1 (1 - \theta) h''(1) (\tilde{X}_{\eta(t)} + \theta(\tilde{X}_t - \tilde{X}_{\eta(t)})) \right] \right|
\leq C \sigma^2 |E[|X_t - X_{\eta(t)}|^4]^{1/2} + \frac{1}{2} \sigma^2 \left| E[H(t, \tilde{X}_{\eta(t)})h''(\tilde{X}_{\eta(t)})E[\tilde{X}_t - \tilde{X}_{\eta(t)}|F_{\eta(t)}]|] \right|
\right).
\]

(15)

Note that

\[
\tilde{X}_t - \tilde{X}_{\eta(t)} = \int_{\eta(t)}^t h(\tilde{X}_s) \left\{ \sigma \varepsilon dW_s + \gamma \varepsilon d\sigma + \frac{1}{2} (h'(\tilde{X}_s) - h'(\tilde{X}_{\eta(s)})) \sigma^2 ds \right\},
\]

\[
E[\tilde{X}_t - \tilde{X}_{\eta(t)}|F_{\eta(t)}] = E \left[ \int_{\eta(t)}^t h(\tilde{X}_s) \left\{ \gamma \varepsilon d\sigma + \frac{1}{2} (h'(\tilde{X}_s) - h'(\tilde{X}_{\eta(s)})) \sigma^2 ds \right\} |F_{\eta(t)} \right]
\]

The Burkholder-Davis-Gundy inequality and the Cauchy-Schwartz inequality and lemma 8 then give, under the hypothesis (H3):

\[
E[|\tilde{X}_t - \tilde{X}_{\eta(t)}|^4] \leq C E \left[ \left( \int_{\eta(t)}^t h'(\tilde{X}_s) \right)^4 \right]
+ C \sigma^2 [E \left( \int_{\eta(t)}^t h'(\tilde{X}_s) \right)^2]^{1/2}
\leq C([\gamma \varepsilon] + \sigma^2)^4 E[(1 + \sup \tilde{X})^4] + C \sigma^2 E[(1 + \sup \tilde{X})^2]
\leq C([\gamma \varepsilon] + \sigma^2)^4 E[(1 + \sup \tilde{X})^4]^{1/2} + C \sigma^2 E[(1 + \sup \tilde{X})^{1/2}]
\]

Similarly, for the second term in (15), we get

\[
\frac{1}{2} \sigma^2 \left| E[H(t, \tilde{X}_{\eta(t)})h''(\tilde{X}_{\eta(t)})E[\tilde{X}_t - \tilde{X}_{\eta(t)}|F_{\eta(t)}]|] \right| \leq C \sigma^2 [E[(1 + \sigma^2)]^2 E[(1 + \gamma \varepsilon)]^{1/2}]
\]

12
Assembling together the estimates for the two terms in (15),
\[ |B_t| \leq C(\gamma_c + \sigma^2)E[(t-\eta(t))^2]^{\frac{1}{2}} + C\sigma^2 E[(t-\eta(t))^4]^{\frac{1}{2}} + (\gamma_c + \sigma^2)E[(t-\eta(t))^2]^{\frac{1}{2}} \]
Using the Jensen inequality and lemma 11, this is further reduced to
\[ \int_0^1 |B_t| dt \leq C\sigma^2 \left( \frac{\gamma_c^2}{\lambda^2} + \frac{\gamma_c^2}{\lambda^2} \right). \]
From the Cauchy-Schwartz inequality we get
\[ \left( \int_{|y| \leq 1} |y| \chi_{C^2} \nu(dy) \right)^2 \leq \lambda \int_{|y| \leq 1} y^2 \chi_{C^2} \nu(dy) \leq C\lambda^2, \]
which implies that \( |\gamma_c| \leq C\sqrt{\lambda} \) and finally \( |B| \leq C\sigma^2 \frac{\gamma_c + \sigma^2}{\lambda^2} \). Assembling these estimates with the ones for \( A \), we complete the proof under the assumptions \((H_3)\) (or \((H_4)\)). Under the assumptions \((H_3')\) or \((H_4')\) the proof is done in a similar fashion. \( \square \)

3 Approximating multidimensional SDE using expansions

In this section we propose an alternative approximation scheme, which yields similar rates to the ones obtained in section 2 but has the advantage of being applicable in the multidimensional case. On the other hand, it is a little more difficult to implement. As before, we start by replacing the small jumps of \( Z \) with a suitable Brownian motion, yielding the SDE
\[ d\bar{X}_t = h(\bar{X}_t)\{\gamma_c dt + dW_t + dZ_t\}, \tag{16} \]
where \( W^c \) is a \( d \)-dimensional Brownian motion with covariance matrix \( \Sigma^c \). This process can also be written as
\[ \bar{X}(t) = \bar{X}(\eta_t) + \int_{\eta_t}^t h(\bar{X}(s)) dW^c(s) + \int_{\eta_t}^t h(\bar{X}(s)) \gamma_c ds, \]
\[ \bar{X}(T^c_i) = \bar{X}(T^c_i-) + h(\bar{X}(T^c_i-))\Delta Z(T^c_i). \]
The idea now is to expand the solution of (16) between the jumps of \( Z^c \) around the solution of the deterministic dynamical system (4), treating the stochastic term as a small random perturbation (see [8, Chapter 2]).

Assume that the coefficient \( h \) is Lipschitz and consider a family of processes \((Y^\alpha)_{0 \leq \alpha \leq 1}\) defined by
\[ Y^\alpha(t) = \bar{X}(\eta_t) + \alpha \int_{\eta_t}^t h(Y^\alpha(s)) dW^c(s) + \int_{\eta_t}^t h(Y^\alpha(s)) \gamma_c ds. \]
Our idea is to replace the process $X := Y^1$ with its first-order Taylor approximation:

$$
\tilde{X}(t) \approx Y^0(t) + \frac{\partial}{\partial \alpha} Y^\alpha(t)|_{\alpha=0}.
$$

Therefore, the new approximation $\tilde{X}$ is defined by

$$
\tilde{X}(t) = Y_0(t) + Y_1(t), \quad t > \eta,
$$

$$
\tilde{X}(T^i_\varepsilon) = \tilde{X}(T^i_\varepsilon -) + h(\tilde{X}(T^i_\varepsilon -))\Delta Z(T^i_\varepsilon),
$$

$$
Y_0(t) = \tilde{X}(\eta_t) + \int_{\eta_t}^t h(Y_0(s)) \gamma_\varepsilon ds
$$

$$
Y_1(t) = \int_{\eta_t}^t \frac{\partial h}{\partial x_i}(Y_0(s)) Y^r_i(s) \gamma_\varepsilon ds + \int_{\eta_t}^t h(Y_0(s)) dW^\varepsilon(s)
$$

where we used the Einstein convention for summation over repeated indices.

Conditionally on the jump times $(T^i_\varepsilon)_{i \geq 1}$, the random vector $Y_1(t)$ is Gaussian with mean zero. Applying the integration by parts formula to $Y^r_i(t)Y^r_j(t)$, it can be shown that its covariance matrix $\Omega(t)$ 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

$$
M_{ij}(t) = \frac{\partial h_{ik}(Y^0(t))}{\partial x_j} \gamma^k_\varepsilon \quad \text{and} \quad N(t) = h(Y_0(t)) \Sigma^\varepsilon h(Y_0(t)).
$$

Successive applications of Gronwall’s inequality yield the following bounds for $Y_0$ and $\Omega$:

$$
\|Y_0(t)\| \leq (\|Y_0(\eta_t)\| + C\|\gamma_\varepsilon\|(t-\eta_t)) e^{C\|\gamma_\varepsilon\|(t-\eta_t)}
$$

$$
\|\Omega(t)\| \leq C\|\Sigma^\varepsilon\|(t-\eta_t)(1 + \|Y_0(\eta_t)\|^2) e^{C\|\gamma_\varepsilon\|(1+\|\gamma_\varepsilon\|)(t-\eta_t)}.
$$

**Lemma 12.** Assume $\nu(\mathbb{R}^d) = \infty$. Then

$$
\lim_{\varepsilon \to 0} \frac{|\gamma_\varepsilon|^2}{\lambda_\varepsilon} = 0.
$$

**Proof.** Left to the reader as an exercise.

To prove Theorem 14, we need to generalize lemmas 8, 9 and 10 to the multi-dimensional setting. While the generalization of the last two lemmas is straightforward, the first one requires a little work, because it needs to be adapted to the new discretization scheme.

**Lemma 13 (Bounds on moments of $X$).** Assume $\nu(\mathbb{R}^d) = \infty$, $h \in C^1(\mathbb{R}^n)$ and

$$
\int_{z \in \mathbb{R}^d} |z|^p \nu(dz) < \infty \quad \text{for some} \ p \geq 2.
$$
Then there exists a constant $C > 0$ (which may depend on $p$ but not on $\varepsilon$) such that for all $\varepsilon$ sufficiently small,

$$E \sup_{0 \leq s \leq t} |\tilde{X}_s|^p \leq C(1 + |x|^p). \tag{20}$$

**Proof.** Denote $h_t := h(\tilde{X}_t)$ and $\tilde{h}_t := h(Y_0(t))$. The SDE for $\tilde{X}$ can be rewritten as

$$d\tilde{X}_t = h_t \tilde{Z}_t + \tilde{h}_t dW_t + h_t \tilde{\gamma}_t dt + \left(\tilde{h}_t - h_t\right)\gamma_t dt + \frac{\partial h}{\partial x_t} (Y_0(t)) Y'_t(t) \gamma_t dt,$$

where $\tilde{Z}_t = \int_0^t \int_{|y| > \varepsilon} g\tilde{N}_t(dy, ds)$, $\tilde{\gamma}_t = \gamma + \int_{|y| > \varepsilon} y\nu(dy)$.

By predictable Burkholder-Davis-Gundy inequality [6, lemma 2.1], we then have

$$E \left[ \sup_{0 \leq s \leq t} \|\tilde{X}_s\|^p \right] \leq CE \left[ \|x|^p + \left( \int_0^t \|h_s\|^2 \int |z| \chi_{\varepsilon}(d\nu)(ds) \right)^{p/2} 
+ \int_0^t \|h_s\|^p ds \int |z|^p \chi_{\varepsilon}(d\nu) + \left( \int_0^t \|h_s\|^2 \Sigma_\varepsilon \right)^{p/2} 
+ \int_0^t \|h_s\|^p |\tilde{\gamma}|^p ds + |\gamma_\varepsilon|^p \int_0^t \|\tilde{h}_s - h_s\|^p ds + |\gamma_\varepsilon|^p \int_0^t \|Y_1(s)\|^p ds \right]$$

$$\leq CE \left[ \|x|^p + \int_0^t \|h_s\|^p ds + (1 + |\gamma_\varepsilon|^p) \int_0^t \|\tilde{h}_s - h_s\|^p ds + |\gamma_\varepsilon|^p \int_0^t \|Y_1(s)\|^p ds \right],$$

where the constant $C$ does not depend on $\varepsilon$ and may change from line to line. Since $h'$ is bounded, we have

$$E \left[ \sup_{0 \leq s \leq t} \|\tilde{X}_s\|^p \right] \leq CE \left[ \|x|^p + \int_0^t \|h_s\|^p ds + (1 + |\gamma_\varepsilon|^p) \int_0^t \|Y_1(s)\|^p ds \right]. \tag{21}$$

Using (19), the last term can be estimated as

$$E \left[ (1 + |\gamma_\varepsilon|^p) \int_0^t \|Y_1(s)\|^p ds \right] = CE \left[ (1 + |\gamma_\varepsilon|^p) \int_0^t E \left[ \|Y_1(s)\|^p |F_{\eta(s)} \right] ds \right]$$

$$\leq CE \left[ (1 + |\gamma_\varepsilon|^p) \int_0^t \|\Omega(s)\|^p/2 ds \right]$$

$$\leq CE \left[ (1 + |\gamma_\varepsilon|^p) \int_0^t (1 + \|\hat{X}_{\eta(s)}\|^p) \|\Sigma_\varepsilon\|^{p/2} (s - \eta_s)^{p/2} e^{C|\gamma_\varepsilon|(1 + |\gamma_\varepsilon|)(s - \eta_s) ds} \right]$$

$$\leq CE \left[ (1 + |\gamma_\varepsilon|^p) \int_0^t (1 + \|\hat{X}_{\eta(s)}\|^p) \|\Sigma_\varepsilon\|^{p/2} e^{2C|\gamma_\varepsilon|(1 + |\gamma_\varepsilon|) r} ds \right],$$
where $\tau$ is an independent exponential random variable with parameter $\lambda_\varepsilon$. Due to Lemma 12, for $\varepsilon$ sufficiently small, the expectation with respect to $\tau$ exists, and computing it explicitly we obtain

$$E \left[ (1 + |\gamma_\varepsilon|^p) \int_0^t \|Y_1(s)\|^p ds \right] \leq C_\varepsilon E \left[ \int_0^t (1 + \|\bar{X}_{\eta(s)}\|^p) ds \right],$$

where $C_\varepsilon \to 0$ as $\varepsilon \to 0$. Inequality (21) therefore becomes

$$E \left[ \sup_{0 \leq s \leq t} \|\bar{X}_s\|^p \right] \leq CE \left[ \|x\|^p + \int_0^t \|h_s\|^p ds + C_\varepsilon \int_0^t (1 + \|\bar{X}_{\eta(s)}\|^p) ds \right]$$

$$\leq CE \left[ \|x\|^p + \int_0^t \|h_s\|^p ds + C_\varepsilon t(1 + \sup_{0 \leq s \leq t} \|\bar{X}_s\|^p) \right],$$

which implies that for $\varepsilon$ sufficiently small,

$$E \left[ \sup_{0 \leq s \leq t} \|\bar{X}_s\|^p \right] \leq CE \left[ 1 + \|x\|^p + \int_0^t \|h_s\|^p ds \right],$$

and we get (20) by Gronwall’s lemma.

\[\square\]

**Theorem 14.**

(i) Assume (H$_3$) or (H'$_3$) + (A). Then

$$|E[f(\bar{X}_1) - f(X_1)]| \leq C \left( \frac{\|\Sigma_\varepsilon\|}{\lambda_\varepsilon} (\|\Sigma_\varepsilon\| + |\gamma_\varepsilon|) + \int_{\mathbb{R}^d} |y|^4 \bar{\chi}_\varepsilon \nu(dy) \right).$$

(ii) Assume (H$_4$) or (H'$_4$) + (A), let $\gamma_\varepsilon$ be bounded and suppose that for some measure $\nu_0$

$$\int_{\mathbb{R}^d} y_i y_j y_k \bar{\chi}_\varepsilon \nu(dy) \leq \int_{\mathbb{R}^d} |y|^4 \bar{\chi}_\varepsilon \nu_0(dy)$$

for all $i,j,k$ and all $\varepsilon$ sufficiently small. Then

$$|E[f(\bar{X}_1) - f(X_1)]| \leq C \left( \frac{\|\Sigma_\varepsilon\|}{\lambda_\varepsilon} + \int_{\mathbb{R}^d} |y|^4 \bar{\chi}_\varepsilon (\nu_0 + \nu)(dy) \right).$$

**Proof.** By Itô formula we have

$$E[f(\bar{X}_1) - f(X_1)] = E[u(1, \bar{X}_1) - u(0, X_0)]$$

$$= \int_0^1 dt E \left[ \frac{\partial u(t, \bar{X}(t))}{\partial x_i} \left\{ h_{ij}(Y_0(t)) + \frac{\partial h_{ij}(Y_0(t))}{\partial x_k} Y_{ik}(t) - h_{ij}(\bar{X}(t)) \right\} \gamma_\varepsilon^j \right]$$

$$+ \frac{1}{2} \frac{\partial^2 u(t, \bar{X}(t))}{\partial x_i \partial x_j} h_{ik}(Y_0(t)) \Sigma_{\varepsilon k} h_{jl}(Y_0(t))$$

$$- \int_{\mathbb{R}^d} \left\{ u(t, \bar{X}(t) + h(\bar{X}(t))y) - u(t, \bar{X}(t)) - \frac{\partial u(t, \bar{X}_1)}{\partial x_i} h_{ij}(\bar{X}(t))y_j \right\} \bar{\chi}_\varepsilon \nu(dy).$$

**16**
Denote the expectation term in (22) by $A_t$ and the sum of the terms in (23) and (24) by $B_t$. The term $A_t$ satisfies

$$|A_t| \leq CE \left[ \left| \frac{\partial u(t, X(t))}{\partial x} \right| |Y_1(t)|^2 |\gamma_\varepsilon| \right]$$

Under the assumption (H$_4$) or (H$_3$), using (19), we have

$$|A| \leq CE \left[ |Y_1(t)|^2 |\gamma_\varepsilon| \leq CE \left[ \|\Omega(t)\| |\gamma_\varepsilon| \right]
\leq CE [1 + |Y_0(\eta_0)|^4 |\Sigma_\varepsilon| |\gamma_\varepsilon| |E[(t - \eta_0) e^{C|\gamma_\varepsilon| (1 + |\gamma_\varepsilon|)(t - \eta_0)}]].$$

Using Lemmas 11 and 13, we then get

$$\int_0^1 |A_t| dt \leq C(1 + |x|^2) \frac{\|\Sigma_\varepsilon\| |\gamma_\varepsilon|}{\lambda_\varepsilon}.$$  

Under (H$_4$) or (H$_3$) this result can be obtained along the same lines.

Let us now turn to the term $B_t$. It is rewritten via

$$B_t = E \left[ - \int_{R^4} \left\{ u(t, \tilde{X}(t) + h(\tilde{X}(t))y) - u(t, \tilde{X}(t)) - \frac{\partial u(t, \tilde{X}(t))}{\partial x_i} h_{ij}(\tilde{X}(t))y_j \right\} \tilde{x}_\varepsilon \nu(dy) 
- \frac{\partial^2 u(t, \tilde{X}(t))}{\partial x_i \partial x_j} h_{ik}(\tilde{X}(t))h_{jl}(\tilde{X}(t))y_k y_l \right\} \tilde{x}_\varepsilon \nu(dy) 
+ \frac{\partial^2 u(t, Y_0(t))}{\partial x_i \partial x_j} \Sigma_{\varepsilon ij} \left\{ h_{ik}(Y_0(t))h_{jl}(Y_0(t)) - h_{ik}(\tilde{X}(t))h_{jl}(\tilde{X}(t)) \right\} 
+ \left( \frac{\partial^2 u(t, \tilde{X}(t))}{\partial x_i \partial x_j} - \frac{\partial^2 u(t, \tilde{X}(t))}{\partial x_i \partial x_j} \right) \Sigma_{\varepsilon ij} \left\{ h_{ik}(Y_0(t))h_{jl}(Y_0(t)) - h_{ik}(\tilde{X}(t))h_{jl}(\tilde{X}(t)) \right\} \right]$$

Denote the first two lines by $B_1(t)$, the third line by $B_2(t)$ and the fourth by $B_3(t)$. Under the hypotheses (H$_4$) or (H$_3$), we get

$$|B_1(t)| \leq E \left| \int_{R^4} \frac{\partial^2 u(t, \tilde{X}(t) + \theta h(\tilde{X}(t))y)}{\partial x_i \partial x_j \partial k} h_{il}(\tilde{X}(t))h_{jm}(\tilde{X}(t))y_i y_j y_m y_n \tilde{x}_\varepsilon \nu(dy) \right|$$

$$\leq CE \|h(\tilde{X}_t)\|^2 \int_{R^4} |y|^3 \tilde{x}_\varepsilon \nu(dy) \leq C(1 + |x|^3) \int_{R^4} |y|^3 \tilde{x}_\varepsilon \nu(dy).$$

Let $F_{ikj}(x) := h_{ik}(x)h_{jl}(x)$. Using the fact that conditionally on $\mathcal{F}_{\eta_0}$, $Y_1$ is a centered Gaussian process,

$$|B_2(t)| \leq E \left| \frac{\partial^2 u(t, Y_0(t))}{\partial x_i \partial x_j} \Sigma_{\varepsilon ij} \left\{ F_{ikj}(Y_0(t)) - F_{ikj}(\tilde{X}(t)) \right\} \mathcal{F}_{\eta_0} \right|$$

$$\leq E \left| \frac{\partial^2 u(t, Y_0(t))}{\partial x_i \partial x_j} \Sigma_{\varepsilon ij} Y_1^m(t)Y_1^n(t) \int_0^1 (1 - \theta) \frac{\partial^2 F_{ikj}(Y_0(t) + \theta Y_1(t))}{\partial x_p \partial x_q} d\theta \right|$$

$$\leq CE \left[ \|\Sigma_\varepsilon\| |Y_1(t)|^2 (1 + \sup_x |\tilde{X}_s|) \right].$$
Using once again Lemmas 11 and 13, we obtain
\[ \int_0^1 |B_2| dt \leq C(1 + |x|) \frac{\|\Sigma_x\|^2}{\lambda_x}. \]

With a similar reasoning we obtain that \( B_3 \) also satisfies
\[ \int_0^1 |B_3| dt \leq C(1 + |x|) \frac{\|\Sigma_x\|^2}{\lambda_x}. \]

The proof of the first part of the theorem under assumptions (\( H'_4 \)) or (\( H'_3 \)) is done along the same lines.

\[ \square \]

4 Applications

Stochastic models driven by Lévy processes are gaining increasing popularity in financial mathematics, where they offer a much more realistic description of the underlying risks than the traditional diffusion-based models. In this context, many quantities of interest are given by solutions of stochastic differential equations which cannot be solved explicitly. In this paper we consider two such applications: portfolio management with constraints and the Libor market model with jumps.

**One-dimensional SDE: portfolio management with constraints** The value \( V_t \) of a portfolio in various common portfolio management strategies can often be expressed as solution to SDE. One example is the so-called Constant Proportion Portfolio Insurance (CPPI) strategy which aims at maintaining the value of the portfolio above some prespecified guaranteed level \( B_t \) and consists in investing \( m(V_t - B_t) \) into the risky asset \( S_t \) and the remaining amount into the risk-free bond, where \( B_t \) is the time-\( t \) value of the guarantee:

\[
dV_t = m(V_t - B_t) \frac{dS_t}{S_t} + (V_t - m(V_t - B_t)) r dt.
\]

For simplicity we shall suppose that the guarantee is the value of a zero-coupon bond expiring at time \( t \) with nominal amount \( N = 1 \): \( B_t = e^{-r(T-t)}N \). This equation can then easily be solved analytically, but often additional constraints imposed on the portfolio make the analytic solution impossible. For instance, often leverage is not allowed, and the portfolio dynamics becomes

\[
dV_t = \min(V_t - m(V_t - B_t)) \frac{dS_t}{S_t} + (V_t - \min(V_t - m(V_t - B_t))) r dt.
\]

(25)

Suppose that the stock \( S \) follows an exponential Lévy model:

\[
dZ^*_t = \frac{dS_t}{S_t} - r dt,
\]
where $Z^*$ is a pure-jump Lévy process, and let $V^*_t = \frac{V_t}{B_t}$ be the value of the portfolio computed using $B_t$ as numéraire. Equation (25) can be simplified to

$$dV^*_t = \min(V^*_{t-}, m(V^*_t - 1))dZ^*_t.$$  

(26)

This equation cannot be solved explicitly, but the corresponding deterministic ODE

$$dX_t = \min(X_t, m(X_t - 1))dt$$

admits a simple explicit solution for $m > 1$: if $x \geq \frac{m}{m-1}$, then

$$X_t = xe^{t}.$$  

Otherwise, for $1 < x < \frac{m}{m-1}$

$$X_t = 1 + (x-1)e^{mt}, \quad t \leq t^*$$

$$X_t = \frac{m}{m-1} [(x-1)(m-1)]^{1/m} e^{t}, \quad t \geq t^*$$

with $t^* = -\frac{1}{m} \log [(x-1)(m-1)]$.

Figure 1 shows the trajectories of the solution of (26) simulated in the variance gamma model. The variance gamma model is a pure-jump Lévy process with Lévy density

$$\nu(x) = \frac{Ce^{-\lambda_+|x|}}{|x|}1_{x<0} + \frac{Ce^{-\lambda_-x}}{x}1_{x>0}.$$  

The variance gamma process is thus a difference of two independent gamma processes and we simulate it using the approximation of Example 7. The “near-exact” solution was produced using the standard Euler scheme with 10000 discretization steps. With 100 jumps, the jump-adapted scheme has a greater error than the Euler scheme with 100 points, because the Euler scheme uses exact simulation of increments. However for 200 jumps, the error of the jump-adapted scheme virtually disappears, whereas the Euler scheme still has a significant error even with 500 steps. This illustrates the exponential convergence of the jump-adapted scheme for the variance gamma process. Note that these graphs show the trajectories of the solution and hence illustrate the strong convergence. The effect of the weak approximation of small jumps by Brownian motion is therefore not visible, but even without this additional improvement, the convergence for the variance gamma process is very fast.

**Multidimensional SDE: Libor market model with jumps**  
One important example of a non-trivial multidimensional SDE arises in the Libor market model, which describes joint arbitrage-free dynamics of a set of forward interest rates. Libor market models with jumps were considered among others by Jamshidian [11], Glasserman and Kou [9] and Eberlein and Özkan [7]. Let $T_i = T_1 + (i-1)\delta$, $i = 1, \ldots, n+1$ be a set of dates called tenor dates. The Libor rate $L^*_t$ is the forward interest rate, defined at date $t$ for the period $[T_i, T_{i+1}]$. If
Figure 1: CPPI portfolio simulation with $V_0^* = 1.5$ and $m = 2$ in the variance gamma model (no diffusion component added). Left: jump-adapted scheme. Right: standard Euler scheme.

$B_t(T)$ is the price at time $t$ of a zero-coupon bond, that is, a bond which pays to its holder 1 unit at date $T$, the Libor rates can be computed via

$$L_i^t = \frac{1}{\delta} \left( \frac{B_t(T_i)}{B_t(T_{i+1})} - 1 \right)$$

A Libor market model (LMM) is a model describing an arbitrage-free dynamics of a set of Libor rates.

In this example, we shall use a simplified version of the general semimartingale LMM given in [11], supposing that all Libor rates are driven by a $d$-dimensional pure jump Lévy process. In this case, following [11], an arbitrage-free dynamics of $L_1^t, \ldots, L_n^t$ can be constructed via the multi-dimensional SDE

$$\frac{dL_i^t}{L_i^t} = \sigma^i(t) dZ_t - \int_{\mathbb{R}^d} \sigma^i(t) z \left[ \prod_{j=i+1}^n \left( 1 + \frac{\delta L_j^t \sigma^j(t) z}{1 + \delta L_j^t} \right) - 1 \right] \nu(dz)dt. \quad (27)$$

Here $Z$ is a $d$-dimensional martingale pure jump Lévy process with Lévy measure $\nu$, and $\sigma^i(t)$ are $d$-dimensional deterministic volatility functions.

Equation (27) gives the Libor dynamics under the so-called terminal measure, which corresponds to using the last zero-coupon bond $B_t(T_{n+1})$ as numéraire. This means that the price at time $t$ of an option having payoff $H = h(L_1^{T_1}, \ldots, L_n^{T_n})$
at time $T_1$ is given by

$$\pi_t(H) = B_t(T_{n+1}) E \left[ \frac{h(L_{T_1}^1, \ldots, L_{T_1}^n)}{B_t(T_{n+1})} \bigg| \mathcal{F}_t \right]$$

$$= B_t(T_{n+1}) E \left[ h(L_{T_1}^1, \ldots, L_{T_1}^n) \prod_{i=1}^n (1 + \delta L_{T_1}^i) \bigg| \mathcal{F}_t \right]$$

$$= \frac{B_t(T_1)}{\prod_{i=1}^n (1 + \delta L_{T_1}^i)} E \left[ h(L_{T_1}^1, \ldots, L_{T_1}^n) \prod_{i=1}^n (1 + \delta L_{T_1}^i) \bigg| \mathcal{F}_t \right]$$

(28)

The price of any such option can therefore be computed by Monte Carlo using equation (27).

Introducing a $n+1$-dimensional state process $X_t$ with $X_0^t \equiv t$ and $X_i^t = L_i^t$ for $1 \leq i \leq n$, a $d+1$-dimensional driving Lévy process $\tilde{Z}_t = (t, Z_t)^\top$, and a $(n+1) \times (d+1)$-dimensional function $h(x)$ defined by $h_{11} = 1$, $h_{1j} = 0$ for $j = 2, \ldots, d+1$, $h_{i1} = f_i(x)$ and $h_{ij} = \sigma_{ij}^t(x_0)$ with

$$f_i(x) := -\int_{\mathbb{R}^d} \sigma^t(x_0) z \left[ \prod_{j=i+1}^n \left( 1 + \frac{\delta x_j \sigma^t(x_0) z}{1 + \delta x_j} \right) - 1 \right] \nu(dz),$$

we see that the equation (27) takes the homogeneous form $dX_i = h(X_{t-}) d\tilde{Z}_t$, to which the discretization scheme of section 3 can be readily applied.

For the purposes of illustration, we simplify the model even further, taking $d = 1$ and supposing that the functions $\sigma^t(t)$ are constant: $\sigma^t(t) \equiv 1$. The driving Lévy process $Z$ is supposed to follow the CGMY model with two alternative sets of parameters: $\alpha = 0.5$ and $C = 1.5$ in Case 1 and $\alpha = 1.8$ and $C = 0.01$ in Case 2. In both cases, we take $\lambda_+ = 10$ and $\lambda_- = 20$. The set of tenor dates for the Libor market model is $\{5, 6, 7, 8, 9, 10\}$ years, which corresponds to a stochastic differential equation in dimension 5. The initial values of all forward Libor rates are all equal to 15%. This big a value was taken to emphasize the non-linear effects in the simulation.

For the numerical implementation of the scheme of section 3, we solved the equations (17) and (18) simultaneously using the classical 4-th order Runge-Kutta scheme, with time step equal to $\max(1, T_{n+1} - T_1^n)$. This means that most of the time there is only one step of the Runge-Kutta scheme between two consecutive jump times, and our numerical tests have shown that even in this case the error associated to the Runge-Kutta scheme is negligible compared to the stochastic approximation error.

As a sanity check, we first compute the price of a zero-coupon bond with maturity $T_1$, which corresponds to taking $h \equiv 1$ in equation (28). By construction of the model, if the SDE is solved correctly, we must recover the input price of the zero-coupon bond. Figure 2 shows the ratio of the zero coupon bond price estimated using the first-order scheme of section 3 to the input value. For comparison, we also give the value computed using the 0-order approximation.
Figure 2: Ratio of estimated to theoretical zero coupon bond price in Case 1 (left) and Case 2 (right) with 1 standard deviation bounds. $10^5$ trajectories were used for all points except the three points with the biggest intensity in the right graph, where $10^4$ simulations were made to save time.

We do not compare our results with the classical Euler scheme because this would require us to simulate the increments of the CGMY process for which no standard algorithm is available.

The graphs in Figure 2 show that already for the intensity of 1 jump per year, the true price of the zero coupon is within the Monte Carlo confidence bounds for the 1st order scheme, on the other hand, for the 0-th order scheme the convergence is very slow, especially in case 2. This is consistent with the theoretical convergence rates which are of order of $\lambda^{-3}$ in Case 1 and $\lambda^{-0.11}$ in Case 2 for 0-th order approximation and, respectively, $\lambda^{-7}$ and $\lambda^{-1.22}$ for 1-st order approximation.

Next, we use our method to compute the price of the so-called receiver swaption, which gives its holder the right but not the obligation to enter an interest rate swap with fixed rate $K$ at date $T_1$. This means that its pay-off at date $T_1$ is equal to

$$h(L_{T_1}, \ldots, L_{T_1}) = \left(1 - B_{T_1}(T_{n+1}) - K\delta \sum_{i=1}^{n} B_{T_1}(T_{i+1})\right)^+$$

$$= \left(\prod_{i=1}^{n}(1 + \delta L_{T_1}^i)^{-1} - 1 - K\delta \sum_{i=1}^{n} \prod_{j=1}^{i}(1 + \delta L_{T_1}^j)^{-1}\right)^+.$$

Figure 3 shows the price of this product with $K = 15\%$ estimated using the method of section 3, and compared once again to the 0-order scheme. The theoretical value is not known in closed form in this case, but we see that despite the fact that the pay-off function is not differentiable, the convergence of the 1-st order scheme is achieved very quickly while the 0-th order scheme takes a long time to converge.
Figure 3: Estimated price of an ATM receiver swaption with maturity 5 years in Case 1 (left) and Case 2 (right) with 1 standard deviation bounds. The Monte Carlo simulation was performed with $10^5$ trajectories.

Finally, Figure 4 shows the execution time for the 1-st order and the 0-th order scheme as a function of the jump intensity $\lambda$ (this dependence is very similar in cases 1 and 2). These times were obtained on a standard (rather old) Pentium-III PC using a very simple implementation of the scheme of section 3, without any code optimization or variance reduction which could accelerate the computation. Despite the fact that for the same intensity, the 1-st order scheme needs 5 times as much computational effort as the 0-th order scheme, the improvement of convergence is such that even in Case 1 it is advantageous to use the 1-st order scheme.

Acknowledgement

This research of Peter Tankov is supported by the Chair Financial Risks of the Risk Foundation sponsored by Société Générale, the Chair Derivatives of the Future sponsored by the Fédération Bancaire Française, and the Chair Finance and Sustainable Development sponsored by EDF and Calyon.

References


Figure 4: Execution times for $10^5$ trajectories on a Pentium-III PC.


[12] D. Madan and M. Yor, CGMY and Meixner subordinators are absolutely continuous with respect to one sided stable subordinators. Prépublication du Laboratoire de Probabilités et Modèles Aléatoires.


