Jump-adapted discretization schemes for Lévy-driven SDEs

Arturo Kohatsu-Higa
Osaka University and Japan Science and Technology Agency
Graduate School of Engineering Sciences
Machikaneyama cho 1-3, Osaka 560-8531, Japan
E-mail: kohatsu@sigmath.es.osaka-u.ac.jp

Peter Tankov*
Ecole Polytechnique
Centre de Mathématiques Appliquées
91128 Palaiseau Cedex France
Email: peter.tankov@polytechnique.org

Abstract

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 times of large jumps of the driving process. To approximate the solution between these times we replace the small jumps with a Brownian motion, and construct and approximate solution of the resulting continuous SDE. Our technique avoids the simulation of the increments of the Lévy process, for which algorithms are not available in general, and in many cases achieves better convergence rates than the traditional Euler scheme with equal time steps. To illustrate the method, we discuss an application to 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,

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

*Corresponding author

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

$$X_t = X_0 + \int_0^t h(X_{s-})dZ_s, \quad t \in [0,1], \tag{1}$$

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

In this article, we propose a simulation method for X and study its rate of convergence. In particular, we are interested in the case when $\nu(\mathbb{R}^d) = \infty$, that is, there is an infinite number of jumps in every interval of nonzero length a.s. The traditional method to simulate X is to use the Euler scheme which is a uniformly-spaced discretization scheme for (1) [19, 12]. This method suffers from two difficulties: first, for a general Lévy measure ν, there is no available algorithm to simulate the increments of the driving Lévy process and second, a large jump of Z occurring between two discretization points can lead to an important discretization error.

A natural idea due to Rubenthaler [21] (in the context of finite-intensity jump processes, this idea appears also in [3, 16]), 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.

On the other hand, Mordecki et. al [16], propose an adaptive method for a stochastic differerential equation driven by a compound Poisson process and a Brownian motion. In that setting, in order to control the errors that may come from large jumps they propose the simulation of all the jumps of the compounded Poisson process and therefore their analysis cannot be extended to the case $\nu(\mathbb{R}^d) = \infty$.

We assume in this article that the Lévy measure of Z is known as it is the case in many applications. In such situation, one can easily simulate the jumps of Z larger in absolute value than a certain $\varepsilon > 0$. Sometimes, other approximations of Z by a compound Poisson process, such as the series representations [20] lead to simpler simulation algorithms (see Example 6).

As we are interested in the case $\nu(\mathbb{R}^d) = \infty$, this means that there are many small jumps and therefore an approximation of those jumps should improve the approximation scheme. In order to take into account the jumps smaller than ε, we use the idea of Asmussen and Rosiński [1] (see also [5]) and replace all the jumps of Z smaller than ε between the jump times of the compound Poisson approximation with $\sigma_\varepsilon W$ where W is a Brownian motion and σ_ε is a coefficient which equals the second moment associated to the jumps smaller than ε.

Combining these two ideas we propose the following approximating scheme. At any time when there is a jump larger than ε, we compute the jump size and the change in the approximating system. Between two jumps larger than ε, we use an approximate solution of the continuous SDE driven by a $\sigma_\varepsilon W$.

The approximation of the solution to the continuous SDE between large jumps is constructed as a perturbation around the solution of the deterministic ODE, obtained by removing the small jumps completely. This ODE can be solved either explicitly or using a Runge-Kutta type approach.

We separate our study into two cases: the one-dimensional and the multi-dimensional case. This is done because the approximate solution of an one dimensional stochastic differential equation driven by a Brownian motion can be written using solutions for ordinary differential equations and therefore the approximation between two jumps larger than ε can be achieved easily. In the multidimensional case, we use a Taylor approximation in order to obtain a scheme between two jumps larger than ε.

We denote the approximation obtained with our method by $\hat{X}^\varepsilon$ for the one-dimensional scheme and $\tilde{X}^\varepsilon$ for the multidimensional scheme. The theoretical goal in this article is to study the behavior of the weak approximation error defined by

$$|E\left[f(X_1)\right] - E\left[f(Y_1^\varepsilon)\right]|$$

for $Y^\varepsilon = \hat{X}^\varepsilon,\ \tilde{X}^\varepsilon$.

Supposing that no further discretization is done between the times of jumps larger than ε, the computational complexity of simulating a single approximate trajectory on the time interval $[0,1]$ is proportional to the number of such jumps, which is a random variable. To compare our method to the traditional equally-spaced discretizations, we measure instead the computational complexity by the *average* number of jumps larger than ε on $[0,1]$, denoted by λ_ε. We say that the approximation Y^ε converges weakly for some order $p > 0$ if, for a sufficiently smooth function f,

$$|E[f(X_1)] - E[f(Y_1^\varepsilon)]| \leq K\lambda_\varepsilon^{-p}$$

for some $K > 0$ and all ε sufficiently small.

The exact order of convergence depends on the characteristics of the Lévy process. In particular, if the Lévy measure has a singularity of the form $\frac{1}{|x|^{1+\alpha}}$ near zero, the order of our schemes is $(\frac{3}{\alpha} - 1) \wedge (\frac{2}{\alpha})$. In the same setting, [21] obtains a strictly lower order of $(\frac{1}{\alpha} - \frac{1}{2}) \wedge 1$. It is therefore clear that the introduction of the Asmussen-Rosiński approach leads to an improvement in the rate of convergence.

We can use our theoretical results in order to choose ε in the above method in an optimal way. This means that the study of the error of approximation gives exactly the order of ε in order to achieve a certain order of error. The fact that $\nu(\mathbb{R}^d) = \infty$, which implies that the weak error becomes arbitrarily small as ε decreases, plays an important role in this conclusion.

We start with a section of Preliminaries in order to introduce the notation and the general truncation method for removing small jumps. This is done in some generality in order to later introduce examples using series approximations (see Example 6 and the first example in Section 5) where the truncation functions are non-trivial.

Next, we provide two versions of the jump-adapted discretization scheme for pure jump Lévy processes. The scheme presented in Section 3 is easier to use and implement but works in the case $d = n = 1$. A fully general scheme is then presented in Section 4.

We close the article with some simulation experiments. In the first one, we check that the theoretical rates obtained in the article coincide with the observed rates in the simulation of an example using the Normal inverse Gaussian model. In this model there exists an explicit algorithm for simulating the increments, which makes it possible to compare our method with the Euler scheme.

In our second example, we apply our methodology to a financial problem of option valuation in the Libor market models with jumps. Although this second problem does not satisfy all the assumptions necessary to establish the theoretical convergence rates, we observe very fast convergence, which shows that the results proved in the previous sections may be satisfied in greater generality. In this article, we require f to have some smoothness properties while in the financial applications f usually does not have these properties. The generalization of the theoretical results to nonsmooth f may be carried out using density estimation techniques (see [17], [14] among others). As these estimates come at the cost of a much greater technicality, and impose restrictions on the choice of the Lévy measure, we do not discuss them here.

Throughout the article we use $X(t)$ or X_t to denote the value at time t of the stochastic process X. Positive constants will be denoted by C and they may change from one line to the next. $C_b^n(A)$ denotes the set of n times differentiable functions on A whose n derivatives are bounded.

2 Preliminaries

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 ν will be assumed to satisfy $\nu(\mathbb{R}^d) = \infty$ and $\int_{\mathbb{R}^d} |y|^2\nu(dy) < \infty$. This measure 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 6 below and Section 5). We denote by N_ε a Poisson random measure with intensity $\chi_\varepsilon\nu \times ds$ and by $\widehat{N_\varepsilon}$ its compensated Poisson random measure. Similarly, we denote by $\widehat{\bar{N}}_\varepsilon$ 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:

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

We denote by $\lambda_\varepsilon = \int_{\mathbb{R}^d} \chi_\varepsilon \nu(dy)$ the intensity of Z^ε, by (N_t^ε) the Poisson process which has the same jump times as Z^ε, and by T_i^ε, $i \in \mathbb{N}$, the jump times of Z^ε with $T_0^\varepsilon = 0$. $\hat{Z}^\varepsilon$ denotes the compensated version of the process Z^ε. $\mathcal{F}$ will denote the filtration generated by N and an independent Brownian motion W which will be used for the approximation of small jumps.

Furthermore, we denote by Σ^ε the variance-covariance matrix of R_1^ε:

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

and in the one-dimensional case ($d = 1$) we set $\sigma_\varepsilon^2 := \Sigma_{11}^\varepsilon$. Note that due to the previous assumptions on χ_ε, we have that $\sup_{\varepsilon\in(0,1)} \|\Sigma^\varepsilon\| < \infty$.

Sometimes we shall use the following technical assumption on (χ_ε):

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

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

for ε sufficiently small.

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

3 One-dimensional SDE

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

$$dX_t = h(X_t)dt, \quad X_0 = x. \tag{2}$$

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

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

where F is the primitive of $\frac{1}{h(x)}$. Therefore, we assume that $\frac{1}{h(x)}$ is a locally integrable function. Alternatively, a high-order discretization scheme (such as Runge-Kutta) can be used and is easy to construct (see Proposition 7).

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

$$\hat{X}(T^\varepsilon_{i+1}-) = \theta\big(\gamma_\varepsilon(T^\varepsilon_{i+1} - T^\varepsilon_i) + \sigma_\varepsilon(W(T^\varepsilon_{i+1}) - W(T^\varepsilon_i)) - \frac{1}{2}h'^\varepsilon(X^\varepsilon_{T_i})\sigma^2_\varepsilon(T^\varepsilon_{i+1} - T^\varepsilon_i); \hat{X}(T^\varepsilon_i)\big) \tag{3}$$

$$\hat{X}(T^\varepsilon_{i+1}) = \hat{X}(T^\varepsilon_{i+1}-) + h(\hat{X}(T^\varepsilon_{i+1}-))\Delta Z(T^\varepsilon_{i+1}). \tag{4}$$

Similarly, for an arbitrary point t, we define

$$\hat{X}(t) = \theta\big(\gamma_\varepsilon(t - \eta_t) + \sigma_\varepsilon(W(t) - W(\eta_t)) - \frac{1}{2}h'(\hat{X}(\eta_t))\sigma^2_\varepsilon(t - \eta_t); \hat{X}(\eta_t)\big), \tag{5}$$

where we η_t is the discretization point immediately preceeding t: $\eta_t := \sup\{T^\varepsilon_i : T^\varepsilon_i \leq t\}$. 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 large 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 $h \in C^1(\mathbb{R})$. The process $\hat{X}$ defined by (3)–(5) is the solution of the stochastic differential equation*

$$d\hat{X}_t = h(\hat{X}_{t-})\left\{dZ^\varepsilon_t + \sigma_\varepsilon dW_t + \gamma_\varepsilon dt + \frac{1}{2}(h'(\hat{X}_t) - h'(\hat{X}_{\eta(t)}))\sigma^2_\varepsilon dt\right\}.$$

Proof. It is enough to show that the process

$$Y_t := \theta(\gamma_\varepsilon t + \sigma_\varepsilon W_t - \frac{1}{2}h'(x)\sigma^2_\varepsilon t; x) \tag{6}$$

is the solution of the continuous SDE

$$dY_t = h(Y_t)\left\{\sigma_\varepsilon dW_t + \gamma_\varepsilon dt + \frac{1}{2}(h'(Y_t) - h'(x))\sigma^2_\varepsilon dt\right\}.$$

This follows by an application of Itô formula to (6) using

$$\frac{\partial\theta(t;z)}{\partial t} = h(\theta(t;z))$$
$$\frac{\partial^2\theta(t;z)}{\partial t^2} = h'(\theta(t;z))h(\theta(t;z)).$$

□

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

$(\mathbf{H_n})$ $f \in C_b^n$, $h \in C_b^n$ and $\int |z|^{2n}\nu(dz) < \infty$.

$(\mathbf{H'_n})$ $f \in C^n$, $h \in C_b^n$, $f^{(k)}$ have at most polynomial growth for $1 \le k \le n$ and $\int |z|^k \nu(dz) < \infty$ for all $k \ge 1$.

Theorem 2.

(i) Assume $(\mathbf{H_3})$ *or* $(\mathbf{H'_3}) + (\mathbf{A})$. *Then*

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

(ii) Assume $(\mathbf{H_4})$ *or* $(\mathbf{H'_4}) + (\mathbf{A})$, *let* γ_ε *be bounded and let the measure* ν *satisfy*

$$\left|\int y^3 \bar{\chi}_\varepsilon \nu(dy)\right| \le C \int |y|^4 \bar{\chi}_\varepsilon \nu_0(dy) \tag{7}$$

for some measure ν_0 *and some posituve constant* C *independent of* ε . *Then*

$$|E[f(\hat{X}_1) - f(X_1)]| \le C\left(\frac{\sigma_\varepsilon^2}{\lambda_\varepsilon}(\sigma_\varepsilon^2 + |\gamma_\varepsilon|) + \int_{\mathbb{R}} |y|^4 \bar{\chi}_\varepsilon (\nu_0 + \nu)(dy)\right)$$

and in particular,

$$|E[f(\hat{X}_1) - f(X_1)]| \le C\left(\frac{\sigma_\varepsilon^2}{\lambda_\varepsilon} + \int_{\mathbb{R}} |y|^4 \bar{\chi}_\varepsilon (\nu_0 + \nu)(dy)\right).$$

Remark 3. *(i)Under the polynomial growth assumptions* $(\mathbf{H'_3})$ *or* $(\mathbf{H'_4})$, *the constants in the above theorem may depend on the initial value* x.
(ii) Condition (7) is satisfied, for example, if $\chi_\varepsilon(y) = \chi_\varepsilon(-y)$ *for all* y *and* ε *and if* ν *is locally symmetric near zero. That is,* $\nu(dy) = (1 + \xi(y))\nu_0(dy)$, *where* ν_0 *is a symmetric measure satuisfying suitable integrability conditions and* $\xi(y) = O(y)$ *for* $y \to 0$.

Corollary 4 (Worst-case convergence rates)**.** *Let* $\chi_\varepsilon(x) = 1_{|x|>\varepsilon}$. *Then, under the conditions* $(\mathbf{H_3})$ *or* $(\mathbf{H'_3}) + (\mathbf{A})$,

$$|E[f(\hat{X}_1) - f(X_1)]| \le o(\lambda_\varepsilon^{-\frac{1}{2}}),$$

and under the conditions of part (ii) of Theorem 2,

$$|E[f(\hat{X}_1) - f(X_1)]| \le o(\lambda_\varepsilon^{-1}).$$

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

$$\sqrt{\lambda_\varepsilon} \int_{|y| \le \varepsilon} |y|^3 \nu(dy) \le \sqrt{\varepsilon^2 \lambda_\varepsilon}\, \sigma_\varepsilon^2 \xrightarrow{\varepsilon \to 0} 0$$

and similarly

$$\lambda_\varepsilon \int_{|y|\leq\varepsilon} |y|^4 \nu(dy) \leq \varepsilon^2 \lambda_\varepsilon \sigma_\varepsilon^2 \xrightarrow{\varepsilon\to 0} 0.$$

□

Example 5 (Stable-like behavior). Once again, let $\chi_\varepsilon(x) = 1_{|x|>\varepsilon}$, and assume that the Lévy measure has an α-stable-like behavior near zero ($\alpha \in (0,2)$), meaning that ν has a density $\nu(y)$ satisfying

$$\nu(y) = \frac{g(y)}{|y|^{1+\alpha}}, \tag{8}$$

where $\alpha \in (0,2)$ and g has finite nonzero right and left limits at zero. This is the case, for example, for the tempered stable process (CGMY) [6]. Then in this case, we have that for $k \neq \alpha$

$$\int_{|y|\leq\varepsilon} |y|^k \nu(dy) = O(\varepsilon^{k-\alpha}) \quad \text{and} \int_{|y|\leq\varepsilon} |y|^\alpha \nu(dy) = O(\log(\varepsilon)).$$

Therefore in particular, we have that $\lambda_\varepsilon = O(\varepsilon^{-\alpha})$, $\sigma_\varepsilon^2 = O(\lambda_\varepsilon^{1-\frac{2}{\alpha}})$ and $\gamma_\varepsilon = \gamma + \int_{|y|>1} \nu(dy) + O(\lambda_\varepsilon^{1-\frac{1}{\alpha}})$

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

$$|E[f(\hat{X}_1) - f(X_1)]| \leq O(\lambda_\varepsilon^{(1-\frac{3}{\alpha})\vee(-\frac{2}{\alpha})}).$$

By comparaison, in in the same setting, [21] obtains a convergence rate of only $O(\lambda_\varepsilon^{(\frac{1}{2}-\frac{1}{\alpha})\vee(-1)})$. If the Lévy measure is locally symmetric near zero, our scheme has the improved convergence rate

$$|E[f(\hat{X}_1) - f(X_1)]| \leq O(\lambda_\varepsilon^{-\frac{2}{\alpha}})$$

for all $\alpha \in (0,2)$.

Example 6 (Simulation using series representation). In this example we explain why it can be useful to define truncation functions other than $\chi_\varepsilon(x) = 1_{|x|>\varepsilon}$. The gamma process has Lévy density $\nu(z) = \frac{ce^{-\lambda z}}{z} 1_{z>0}$. If one uses the truncation function $\chi_\varepsilon(x) = 1_{|x|>\varepsilon}$, one will need to simulate random variables with law $\frac{ce^{-\lambda z}}{z\nu((\varepsilon,\infty))} 1_{z>\varepsilon}$, which may be costly. Instead, one can use one of the many series representations for the gamma process [20]. Maybe the most convenient one is

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

where (Γ_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

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

defines a compound Poisson process with Lévy density

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

and therefore this series representation corresponds to

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

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

Discretization of the ODE The convergence rates given in Theorem 2 are obtained under the assumption that the equation (2) is solved explicitly. In this remark, we consider the situation when a discretization scheme (e.g. Runge-Kutta) is used for (2) as well. Let $\theta^d(t;x)$ be the approximate solution of (2) at time t with initial condition $X_0 = x$ obtained with one step of the ODE approximation scheme of our choice. We assume that there exists $q \ge 1$ and $C < \infty$, such that

$$|\theta^d(t;x) - \theta(t;x)| \le C|t|^{q+1}, \quad \forall x, \forall t. \tag{9}$$

For example, the classical Runge-Kutta scheme of order q satisfies the condition (9) provided that *the function h and its derivatives of orders up to p are bounded*. We refer the reader to [11] for details on Runge-Kutta schemes and the corresponding error estimates.

We introduce a discretization scheme using approximate solution of the ODE by defining inductively $\hat{X}^d(0) = X_0$ and for $i \ge 1$,

$$\begin{aligned}\hat{X}^d(T_{i+1}^\varepsilon -) = \theta^d\big(&\gamma_\varepsilon(T_{i+1}^\varepsilon - T_i^\varepsilon) + \sigma_\varepsilon(W(T_{i+1}^\varepsilon) - W(T_i^\varepsilon)) \\ &- \frac{1}{2}h'(\hat{X}^d(T_i^\varepsilon))\sigma_\varepsilon^2(T_{i+1}^\varepsilon - T_i^\varepsilon); \hat{X}^d(T_i^\varepsilon)\big)\end{aligned} \tag{10}$$

$$\hat{X}^d(T_i^\varepsilon) = \hat{X}^d(T_i^\varepsilon -) + h(\hat{X}^d(T_i^\varepsilon -))\Delta Z(T_i^\varepsilon). \tag{11}$$

This means, that although we use an approximation scheme for the solution of the ODE, we do not introduce additional discretization points between consecutive jump times of Z^ε. We shall see from the subsequent analysis that this is not necessary if the ODE approximation scheme has sufficiently high order.

Similarly, for an arbitrary point t, we define

$$\hat{X}^d(t) = \theta^d\big(\gamma_\varepsilon(t - \eta_t) + \sigma_\varepsilon(W(t) - W(\eta_t)) - \frac{1}{2}h'(\hat{X}^d(\eta_t))\sigma_\varepsilon^2(t - \eta_t); \hat{X}^d(\eta_t)\big). \tag{12}$$

Finally, for technical reasons, we define an auxiliary discretization scheme $\hat{X}^{dc}_t$, which is defined in each interval $[T^\varepsilon_i, T^\varepsilon_{i+1}]$ as follows. First, $X^{dc}(T^\varepsilon_i) = \hat{X}^d(T^\varepsilon_i)$. Next for $t \in (T^\varepsilon_i, T^\varepsilon_{i+1})$ we define X^{dc} by

$$\hat{X}^{dc}(t) = \theta\big(\gamma_\varepsilon(t-\eta_t) + \sigma_\varepsilon(W(t)-W(\eta_t)) - \frac{1}{2}h'(\hat{X}^d(\eta_t))\sigma^2_\varepsilon(t-\eta_t); \hat{X}^d(\eta_t)\big). \tag{13}$$

Finally X^{dc} jumps at T^ε_{i+1} so that $X^{dc}(T^\varepsilon_{i+1}) = \hat{X}^d(T^\varepsilon_{i+1})$.

If the discretization scheme (10)–(11) is used instead of the exact ODE solution, the analog of Theorem 2 takes the following form.

Proposition 7. *(i) Assume* $(\mathbf{H_3})$ *and* (9)*. Then*

$$|E[f(\hat{X}^d_1) - f(X_1)]| \le C\left(\frac{\sigma^2_\varepsilon}{\lambda_\varepsilon}(\sigma^2_\varepsilon + |\gamma_\varepsilon|) + \int_{\mathbb{R}} |y|^3 \bar{\chi}_\varepsilon \nu(dy) + \frac{|\gamma_\varepsilon|^{q+1}+1}{\lambda^q_\varepsilon} + \frac{\sigma^{q+1}_\varepsilon}{\lambda_\varepsilon^{\frac{q-1}{2}}}\right).$$

(ii) Assume $(\mathbf{H_4})$ *and* (9)*, let* γ_ε *be bounded and let the measure* ν *satisfy* (7)*. Then*

$$|E[f(\hat{X}^d_1) - f(X_1)]| \le C\left(\frac{\sigma^2_\varepsilon}{\lambda_\varepsilon}(\sigma^2_\varepsilon + |\gamma_\varepsilon|) + \int_{\mathbb{R}} |y|^4 \bar{\chi}_\varepsilon (\nu_0+\nu)(dy) + \frac{1}{\lambda^q_\varepsilon} + \frac{\sigma^{q+1}_\varepsilon}{\lambda_\varepsilon^{\frac{q-1}{2}}}\right)$$

and in particular,

$$|E[f(\hat{X}^d_1) - f(X_1)]| \le C\left(\frac{\sigma^2_\varepsilon}{\lambda_\varepsilon} + \int_{\mathbb{R}} |y|^4 \bar{\chi}_\varepsilon (\nu_0+\nu)(dy) + \frac{1}{\lambda^q_\varepsilon} + \frac{\sigma^{q+1}_\varepsilon}{\lambda_\varepsilon^{\frac{q-1}{2}}}\right).$$

Remark 8. *The same result can be shown to be true under the conditions* $(\mathbf{H'_3}) + (\mathbf{A})$ *instead of* $(\mathbf{H_3})$ *or* $(\mathbf{H'_4}) + (\mathbf{A})$ *instead of* $(\mathbf{H_4})$ *but we omit this discussion to save space.*

Remark 9. *[Choice of the order of ODE discretization scheme] To understand the effect of the order* q *of the discretization scheme used for the deterministic ODE on the convergence rates in proposition 7, let us compute the rates for stable-like Lévy measure* (8)*. We denote by* ε^c *the upper bound on the error given by Theorem 2 (part i. or ii., depending on the context), by* ε^d *the upper bound on the additional error introduced by ODE discretization:*

$$\varepsilon^d = \frac{|\gamma_\varepsilon|^{q+1}+1}{\lambda^q_\varepsilon} + \frac{\sigma^{q+1}_\varepsilon}{\lambda_\varepsilon^{\frac{q-1}{2}}}.$$

A simple computation then yields: $\varepsilon^d = O(\lambda_\varepsilon^{-q})$ *for* $0 < \alpha < 1$ *and* $\varepsilon^d = O(\lambda_\varepsilon^{1-\frac{q+1}{\alpha}})$ *for* $1 < \alpha < 2$*. In particular, to have the same worst-case convergence rates for* ε^d *as for* ε^c*, one must take* $q \ge 2$ *in the general case and* $q \ge 3$ *in the locally symmetric case.*

The proofs Theorem 2 and Proposition 7 will be proved after a series of lemmas.

Lemma 10. *Let $f : [0,\infty) \to [0,\infty)$ be an increasing function. Then*

$$E\left[\int_0^1 f(t-\eta(t))dt\right] \le E[f(\tau)],$$

where τ is an exponential random variable with intensity λ_ε. In particular, for $p > 0$, we have

$$E\left[\int_0^1 (t-\eta(t))^p dt\right] \le \frac{p!}{\lambda_\varepsilon^p}.$$

Proof. Let $k_\varepsilon = \sup\{k : T_k^\varepsilon < 1\}$. Then

$$\begin{aligned}
E\left[\int_0^1 f(t-\eta(t))dt\right] &= E\left[\sum_{i=1}^{k_\varepsilon}\int_{T_{i-1}^\varepsilon}^{T_i^\varepsilon} f(t-T_{i-1}^\varepsilon)dt + \int_{T_{k_\varepsilon}^\varepsilon}^1 f(t-T_{k_\varepsilon}^\varepsilon)dt\right] \\
&= E\left[\sum_{i=1}^{k_\varepsilon}\int_{T_{i-1}^\varepsilon}^{T_i^\varepsilon} f(T_i^\varepsilon - t)dt + \int_{T_{k_\varepsilon}^\varepsilon}^1 f(1-t)dt\right] \\
&\le E\left[\sum_{i=1}^{k_\varepsilon}\int_{T_{i-1}^\varepsilon}^{T_i^\varepsilon} f(T_i^\varepsilon - t)dt + \int_{T_{k_\varepsilon}^\varepsilon}^1 f(T_{k_\varepsilon+1}^\varepsilon - t)dt\right] \\
&= E\left[\int_0^1 f(\inf\{T_i^\varepsilon : T_i^\varepsilon > t\} - t)dt\right] = E[f(\tau)].
\end{aligned}$$

The second statement of the lemma is a direct consequence of the first one. □

Lemma 11 (Bounds on moments of X, $\hat{X}$ and $\hat{X}^{dc}$)**.** *Assume*

$$\int_{\mathbb{R}} |z|^p \nu(dz) < \infty \quad \text{for some } p \ge 2, \tag{14}$$

and $h \in C_b^1(\mathbb{R})$. Then there exists a constant $C > 0$ (which may depend on p but not on ε) such that

$$E[\sup_{0\le s\le 1} |X_s|^p] \le C(1+|x|^p), \tag{15}$$

$$E[\sup_{0\le s\le 1} |\hat{X}_s|^p] \le C(1+|x|^p). \tag{16}$$

Assume in addition that the discretization scheme used for defining $\hat{X}^d$ and $\hat{X}^{dc}$ satisfies the condition (9)*. Then also*

$$E[\sup_{0\le s\le 1} |\hat{X}_s^{dc}|^p] \le C(1+|x|^p). \tag{17}$$

Proof. We shall concentrate on the bound (17). The bound (16) will then follow by taking $\theta^d \equiv \theta$ and the bound (15) will follow from (16) by making ε go to zero. The process $\hat{X}^{dc}$ satisfies the stochastic differential equation (cf (6)):

$$\begin{aligned}
d\hat{X}_t^{dc} &= h(\hat{X}_t^{dc})\{\gamma_\varepsilon dt + \frac{1}{2}(h'(\hat{X}_t^{dc}) - h'(\hat{X}_{\eta_t}^{d}))\sigma_\varepsilon^2 dt + \sigma_\varepsilon dW_t\} \\
&\quad + h(\hat{X}_{t-}^{d})dZ_t^\varepsilon + (\hat{X}_{t-}^{d} - \hat{X}_{t-}^{dc})dN_t^\varepsilon \\
&= \{h(\hat{X}_t^{dc}) - h(\hat{X}_t^{d})\}\gamma_\varepsilon dt + h(\hat{X}_t^{d})\{\gamma + \int_{|z|>1} z\nu(dz)\}dt \\
&\quad + \frac{1}{2}h(\hat{X}_t^{dc})(h'(\hat{X}_t^{dc}) - h'(\hat{X}_{\eta_t}^{d}))\sigma_\varepsilon^2 dt + h(\hat{X}_t^{dc})\sigma_\varepsilon dW_t \\
&\quad + h(\hat{X}_{t-}^{d})d\hat{Z}_t^\varepsilon + (\hat{X}_{t-}^{d} - \hat{X}_{t-}^{dc})d\tilde{N}_t^\varepsilon + \lambda_\varepsilon(\hat{X}_t^{d} - \hat{X}_t^{dc})dt,
\end{aligned}$$

where $\hat{Z}^\varepsilon$ and $\hat{N}^\varepsilon$ are compensated versions of Z^ε and N^ε. Using a predictable version of the Burkholder-Davis-Gundy inequality [7, lemma 2.1] and the fact that h' is bounded, we then obtain for $t \leq 1$ and for some constant $C < \infty$, independent of ε and which may change from inequality to inequality

$$\begin{aligned}
&E[\sup_{0\leq s\leq t} |\hat{X}_s^{dc}|^p] \leq CE\Bigg[|x|^p + |\gamma_\varepsilon|^p \int_0^t |h(\hat{X}_s^{d}) - h(\hat{X}_s^{dc})|^p ds \\
&+ \int_0^t (|h(\hat{X}_s^{d})|^p + |h(\hat{X}_s^{dc})|^p)ds + \lambda_\varepsilon^p \int_0^t |\hat{X}_s^{d} - \hat{X}_s^{dc}|^p ds \\
&+ \int_0^t |\hat{X}_s^{d} - \hat{X}_s^{dc}|^p \lambda_\varepsilon ds + \left(\int_0^t |h(\hat{X}_s^{d})|^2 \int_{\mathbb{R}} |z|^2 \chi_\varepsilon(z)\nu(dz)ds\right)^{p/2} \\
&+ \int_0^t |h(\hat{X}_s^{d})|^p \int_{\mathbb{R}} |z|^p \chi_\varepsilon(z)\nu(dz)ds + \left(\int_0^t (\hat{X}_s^{d} - \hat{X}_s^{dc})^2 \lambda_\varepsilon ds\right)^{p/2} \\
&+ \left(\int_0^t |h(\hat{X}_s^{dc})|^2 \sigma_\varepsilon^2 ds\right)^{p/2}\Bigg]
\end{aligned}$$

Using Jensen's inequality, the assumption (14) the Lipschitz property of h, and the bound

$$|\gamma_\varepsilon|^2 \leq C\left(\lambda_\varepsilon \int_{\mathbb{R}} z^2 \chi_\varepsilon(z)\nu(dz) + 1\right), \tag{18}$$

the above inequality simplifies to

$$\begin{aligned}
&E[\sup_{0\leq s\leq t} |\hat{X}_s^{dc}|^p] \leq CE\left[|x|^p + \int_0^t |h(\hat{X}_s^{dc})|^p ds + (1+\lambda_\varepsilon^p)\int_0^t |\hat{X}_s^{d} - \hat{X}_s^{dc}|^p ds\right] \\
&\leq CE\left[1 + |x|^p + \int_0^t |\hat{X}_s^{dc}|^p ds + (1+\lambda_\varepsilon^p)\int_0^t |\hat{X}_s^{d} - \hat{X}_s^{dc}|^p ds\right].
\end{aligned}$$

Let $\mathcal{G}^\varepsilon$ be the sigma algebra generated by Z^ε. Using the bound (9), we get

$$|\hat{X}_s^{d} - \hat{X}_s^{dc}| \leq |\gamma_\varepsilon(s-\eta_s) + \sigma_\varepsilon(W(s) - W(\eta_s)) - \frac{1}{2}h'(\hat{X}^d(\eta_s))\sigma_\varepsilon^2(s-\eta_s)|^{q+1}$$

and therefore

$$E\left[|\hat{X}_s^d - \hat{X}_s^{dc}|^p\right] = E\left[E\left[|\hat{X}_s^d - \hat{X}_s^{dc}|^p \middle| \mathcal{G}^\varepsilon\right]\right]$$
$$\leq CE\left[\left(|\gamma_\varepsilon|^{p(q+1)} + 1\right)(s-\eta_s)^{p(q+1)} + (s-\eta_s)^{\frac{p(q+1)}{2}}\right].$$

Using Lemma 10,

$$E\left[(1+\lambda_\varepsilon^p)\int_0^t |\hat{X}_s^d - \hat{X}_s^{dc}|^p ds\right] \leq \frac{|\gamma_\varepsilon|^{p(q+1)}+1}{\lambda_\varepsilon^{qp}} + \frac{1}{\lambda_\varepsilon^{p(q-1)/2}},$$

which is bounded uniformly on ε by the inequality (18). The bound (17) now follows from Gronwall's inequality. □

Lemma 12 (Derivatives of the flow)**.** *Let $p \geq 2$ and for an integer $n \geq 1$ assume*

$$\int_{\mathbb{R}} |z|^{np} \nu(dz) < \infty,$$

$h \in C_b^n(\mathbb{R})$. Then

$$E\left[\left|\frac{\partial^k}{\partial x^k} X_1^{(t,x)}\right|^p\right] < \infty$$

for all k with $1 \leq k \leq n$.

Proof. See the proof of lemma 4.2 in [19]. □

Lemma 13. *Let $u(t,x) := E^{(t,x)}[f(X_1)]$.*

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

$$\begin{aligned}\frac{\partial u}{\partial t}(t,x) + \gamma\frac{\partial u}{\partial x}(t,x)h(x) &\\ + \int_{|y|\leq 1}\left(u(t,x+h(x)y) - u(t,x) - \frac{\partial u}{\partial x}(t,x)h(x)y\right)\nu(dy) &\\ + \int_{|y|>1}\left(u(t,x+h(x)y) - u(t,x)\right)\nu(dy) = 0 \qquad (19)&\\ u(1,x) = f(x).&\end{aligned}$$

(ii) Assume $(\mathbf{H_n'})$ with $n \geq 2$. Then $u \in C^{1,n}([0,1]\times\mathbb{R})$, u is a solution of equation (19) 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

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

The interchange of the derivative and the expectation is justified using lemma 12. The boundedness under $(\mathbf{H_n})$ or the polynomial growth under $(\mathbf{H'_n})$ then follow from lemmas 11 and 12. 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 $f(X_T^{(t,x)})$. In fact, note that since Z is a Lévy process and h does not depend on t, $E[f(X_T^{(t,x)})] = E[f(X_{T-t}^{(0,x)})]$ and hence it is sufficient to study the derivative $\frac{\partial E[f(X_t^{(0,x)})]}{\partial t}$. The Itô formula yields

$$\begin{aligned}E[f(X_t)] &= f(x) + E\left[\int_0^t f'(X_{s-})h(X_{s-})dZ_s\right] \\ &+ E\left[\int_0^t\int_{\mathbb{R}}\{f(X_{s-}+h(X_{s-})z) - f(X_{s-}) - f'(X_{s-})h(X_{s-})z\}N(dz,ds)\right].\end{aligned}$$

Denoting by $\hat{Z}$ the martingale part of Z and by $\tilde{\gamma} := \gamma + \int_{|z|>1} z\nu(dz)$ the residual drift, and using lemma 11, we get

$$\begin{aligned}E[f(X_t)] &= f(x) + \tilde{\gamma}\int_0^t E[f'(X_s)h(X_s)]ds \\ &+ \int_0^t E\left[\int_{\mathbb{R}}\{f(X_s+h(X_s)z) - f(X_s) - f'(X_s)h(X_s)z\}\nu(dz)\right]ds,\end{aligned}$$

and therefore

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

Now, once again, lemma 11 allows to prove the finiteness of this expression. Finally, equation (19) is a consequence of Itô's formula applied, this time, to $u(t, X_t)$. □

Proof of Theorem 2. From Itô's formula and lemmas 11 and 13,

$$
\begin{aligned}
&E[f(\hat{X}_1) - f(X_1)] = E[u(1,\hat{X}_1) - u(0,X_0)] \\
&= \int_0^1 dtE\left[\frac{1}{2}\frac{\partial^2 u(t,\hat{X}_t)}{\partial x^2}\sigma_\varepsilon^2 h^2(\hat{X}_t)\right. \\
&\qquad \left. - \int_{\mathbb{R}} \bar{\chi}_\varepsilon \nu(dy)(u(t,\hat{X}_t + h(\hat{X}_t)y) - u(t,\hat{X}_t) - \frac{\partial u(t,\hat{X}_t)}{\partial x}h(\hat{X}_t)y)\right] \qquad (20) \\
&+ \int_0^1 dtE\left[\frac{1}{2}\sigma_\varepsilon^2 \frac{\partial u(t,\hat{X}_t)}{\partial x}h(\hat{X}_t)(h'(\hat{X}_t) - h'(\hat{X}_{\eta(t)}))\right]. \qquad (21)
\end{aligned}
$$

Denote the expectation in (20) by A_t and the one in (21) by B_t. From lemmas 13 and 11, we then have, under the hypothesis $(\mathbf{H_3})$,

$$
\begin{aligned}
|A_t| &\leq \left| E\left[\int_{\mathbb{R}} \frac{1}{2}y^3 h^3(\hat{X}_t)\bar{\chi}_\varepsilon \int_0^1 (1-\theta)^2 \frac{\partial^3 u(t,\hat{X}_t + \theta y h(\hat{X}_t))}{\partial x^3} d\theta \nu(dy)\right]\right| \\
&\leq CE[|h|^3(\hat{X}_t)] \int_{\mathbb{R}} |y|^3 \bar{\chi}_\varepsilon \nu(dy) \leq C(1+|x|^3)\int_{\mathbb{R}} |y|^3 \bar{\chi}_\varepsilon \nu(dy).
\end{aligned}
$$

If, instead, the polynomial growth condition $(\mathbf{H_3'})$ is satisfied, then for some $p > 0$,

$$
A_t \leq CE\left[\int_{\mathbb{R}} |y|^3(1+|h|^3(\hat{X}_t))(1+|\hat{X}_t|^p + |yh(\hat{X}_t)|^p)\bar{\chi}_\varepsilon \nu(dy)\right]
$$

and once again, we use lemma 11 together with the assumption $(\mathbf{A})$, because the Lévy measure integrates any polynomial in y. Under the condition of part (ii) and the hypothesis $(\mathbf{H_4})$,

$$
\begin{aligned}
|A_t| &\leq \left| E\left[\frac{\partial^3 u(t,\hat{X}_t)}{\partial x^3} h^3(\hat{X}_t)\int_{\mathbb{R}} y^3 \bar{\chi}_\varepsilon \nu(dy)\right]\right| \\
&\quad + \left| E\left[\int_{\mathbb{R}} \frac{1}{24}y^4 h^4(\hat{X}_t)\int_0^1 (1-\theta)^3 \frac{\partial^4 u(t,\hat{X}_t + \theta y h(\hat{X}_t))}{\partial x^4} d\theta \bar{\chi}_\varepsilon \nu(dy)\right]\right| \\
&\leq C(1+|x|^4)\int_{\mathbb{R}} |y|^4 \bar{\chi}_\varepsilon(\nu_0 + \nu)(dy).
\end{aligned}
$$

The case of $(\mathbf{H_4'})$ 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 $(\mathbf{H_3})$ is satisfied. Note that $\eta(t)$ is a stopping time and then, once again by lemmas 13 and 11, Taylor formula and

the Cauchy-Schwartz inequality,

$$
\begin{aligned}
|B_t| &\leq \frac{1}{2}\sigma_\varepsilon^2 \left| E[(H(t,\hat{X}_t) - H(t,\hat{X}_{\eta(t)}))(h'(\hat{X}_t) - h'(\hat{X}_{\eta(t)}))] \right| \\
&+ \frac{1}{2}\sigma_\varepsilon^2 \left| E\left[H(t,\hat{X}_{\eta(t)})(h'(\hat{X}_t) - h'(\hat{X}_{\eta(t)})) \right] \right| \\
&\leq \frac{1}{2}\sigma_\varepsilon^2 \left| E\left[(\hat{X}_t - \hat{X}_{\eta(t)})^2 \int_0^1 d\theta \frac{\partial H}{\partial x}(\hat{X}_{\eta(t)} + \theta(\hat{X}_t - \hat{X}_{\eta(t)})) \right.\right. \\
&\qquad \left.\left. \times \int_0^1 d\theta' h''(\hat{X}_{\eta(t)} + \theta'(\hat{X}_t - \hat{X}_{\eta(t)})) \right] \right| \\
&+ \frac{1}{2}\sigma_\varepsilon^2 \left| E\left[H(t,\hat{X}_{\eta(t)}) E\left[(h'(\hat{X}_t) - h'(\hat{X}_{\eta(t)}))|\mathcal{F}_{\eta(t)} \right] \right] \right| \\
&\leq C\sigma_\varepsilon^2 E[(\hat{X}_t - \hat{X}_{\eta(t)})^2 (1 + |\hat{X}_t| + |\hat{X}_{\eta(t)}|)] \\
&+ \frac{1}{2}\sigma_\varepsilon^2 \left| E[H(t,\hat{X}_{\eta(t)}) h''(\hat{X}_{\eta(t)}) E[\hat{X}_t - \hat{X}_{\eta(t)}|\mathcal{F}_{\eta(t)}]] \right| \\
&+ \frac{1}{2}\sigma_\varepsilon^2 E\left[H(t,\hat{X}_{\eta(t)})(\hat{X}_t - \hat{X}_{\eta(t)})^2 \int_0^1 d\theta (1-\theta) h^{(3)}(\hat{X}_{\eta(t)} + \theta(\hat{X}_t - \hat{X}_{\eta(t)})) \right] \\
&\leq C\sigma_\varepsilon^2 E[(\hat{X}_t - \hat{X}_{\eta(t)})^4]^{1/2} \\
&+ \frac{1}{2}\sigma_\varepsilon^2 \left| E[H(t,\hat{X}_{\eta(t)}) h''(\hat{X}_{\eta(t)}) E[\hat{X}_t - \hat{X}_{\eta(t)}|\mathcal{F}_{\eta(t)}]] \right|. \qquad (22)
\end{aligned}
$$

Note that

$$
\hat{X}_t - \hat{X}_{\eta(t)} = \int_{\eta(t)}^t h(\hat{X}_s) \left\{ \sigma_\varepsilon dW_s + \gamma_\varepsilon ds + \frac{1}{2}(h'(\hat{X}_s) - h'(\hat{X}_{\eta_s}))\sigma_\varepsilon^2 ds \right\},
$$

$$
E[\hat{X}_t - \hat{X}_{\eta(t)}|\mathcal{F}_{\eta(t)}] = E\left[\int_{\eta(t)}^t h(\hat{X}_s) \left\{ \gamma_\varepsilon ds + \frac{1}{2}(h'(\hat{X}_s) - h'(\hat{X}_{\eta_s}))\sigma_\varepsilon^2 ds \right\} |\mathcal{F}_{\eta(t)} \right].
$$

The Burkholder-Davis-Gundy inequality and, the Cauchy-Schwartz inequality and lemma 11 then give, under the hypothesis ($\mathbf{H_3}$):

$$
\begin{aligned}
E[(\hat{X}_t - \hat{X}_{\eta(t)})^4] &\leq CE\left[\left| \int_{\eta(t)}^t h(\hat{X}_s) \left\{ \gamma_\varepsilon ds + \frac{1}{2}(h'(\hat{X}_s) - h'(\hat{X}_{\eta_s}))\sigma_\varepsilon^2 ds \right\} \right|^4 \right] \\
&+ C\sigma_\varepsilon^4 E\left[\left(\int_{\eta(t)}^t h^2(\hat{X}_s) ds \right)^2 \right] \\
&\leq C(|\gamma_\varepsilon| + \sigma_\varepsilon^2)^4 E[(t - \eta(t))^4 (1 + \sup_{s \leq t} |\hat{X}_s|)^4] \\
&+ C\sigma_\varepsilon^4 E[(t - \eta(t))^2 (1 + \sup_{s \leq t} |\hat{X}_s|)^4] \\
&\leq C(|\gamma_\varepsilon| + \sigma_\varepsilon^2)^4 E[(t - \eta(t))^8]^{\frac{1}{2}} + C\sigma_\varepsilon^4 E[(t - \eta(t))^4]^{\frac{1}{2}}.
\end{aligned}
$$

Similarly, for the second term in (22), we get

$$\frac{1}{2}\sigma_\varepsilon^2 \left| E[H(t,\hat{X}_{\eta(t)})h''(\hat{X}_{\eta(t)})E[\hat{X}_t - \hat{X}_{\eta(t)}|\mathcal{F}_{\eta(t)}]] \right| \\ \leq C\sigma_\varepsilon^2(|\gamma_\varepsilon| + \sigma_\varepsilon^2)E[(t-\eta(t))^2]^{\frac{1}{2}}.$$

Assembling together the estimates for the two terms in (22),

$$|B_t| \leq C\sigma_\varepsilon^2((|\gamma_\varepsilon|+\sigma_\varepsilon^2)^2 E[(t-\eta(t))^8]^{\frac{1}{4}} + \sigma_\varepsilon^2 E[(t-\eta(t))^4]^{\frac{1}{4}} + (|\gamma_\varepsilon|+\sigma_\varepsilon^2)E[(t-\eta(t))^2]^{\frac{1}{2}})$$

Using the Jensen inequality and lemma 10, this is further reduced to

$$\int_0^1 |B_t|dt \leq C\sigma_\varepsilon^2 \left(\frac{(|\gamma_\varepsilon| + \sigma_\varepsilon^2)^2}{\lambda_\varepsilon^2} + \frac{|\gamma_\varepsilon| + \sigma_\varepsilon^2}{\lambda_\varepsilon} \right).$$

From the Cauchy-Schwartz inequality we get

$$\left(\int_{|y|\leq 1} |y| \chi_\varepsilon \nu(dy) \right)^2 \leq \lambda_\varepsilon \int_{|y|\leq 1} y^2 \chi_\varepsilon \nu(dy) \leq C\lambda_\varepsilon,$$

which implies that $|\gamma_\varepsilon| \leq C\sqrt{\lambda_\varepsilon}$ and finally $|B| \leq C\sigma_\varepsilon^2 \frac{|\gamma_\varepsilon|+\sigma_\varepsilon^2}{\lambda_\varepsilon}$. Assembling these estimates with the ones for A, we complete the proof under the assumptions $(\mathbf{H_3})$ (or $(\mathbf{H_4})$). Under the assumptions $(\mathbf{H_3'})$ or $(\mathbf{H_4'})$ the proof is done in a similar fashion. □

Proof of proposition 7. Recall the notation $(T_i^\varepsilon)_{i\geq 0}$ for the jump times of the process Z^ε and $(N_t^\varepsilon)_{t\geq 0}$ for the Poisson process counting the jumps of Z^ε. We

also define $T_0^\varepsilon = 0$ and, abusing the notation, we set $T^\varepsilon_{N_1^\varepsilon+1} := 1$. Then,

$$
\begin{aligned}
&|E[f(\hat{X}_1^d) - f(X_1)]| = |E[u(1,\hat{X}_1^d) - u(0,X_0)]| \\
&= \left| E\left[\sum_{i=1}^{N_1^\varepsilon+1} \{u(T_i^\varepsilon, \hat{X}^d_{T_i^\varepsilon}) - u(T_{i-1}^\varepsilon, \hat{X}^d_{T_{i-1}^\varepsilon})\} \right] \right| \\
&= \left| E\Bigg[\sum_{i=1}^{N_1^\varepsilon} \{u(T_i^\varepsilon, \hat{X}^d_{T_i^\varepsilon}) - u(T_i^\varepsilon, \hat{X}^d_{T_i^\varepsilon -}) \} \right. \\
&\qquad \left. + \sum_{i=1}^{N_1^\varepsilon+1} \{u(T_i^\varepsilon, \hat{X}^d_{T_i^\varepsilon -}) - u(T_i^\varepsilon, \hat{X}^{dc}_{T_i^\varepsilon -}) + u(T_i^\varepsilon, \hat{X}^{dc}_{T_i^\varepsilon -}) - u(T_{i-1}^\varepsilon, \hat{X}^{dc}_{T_{i-1}^\varepsilon})\} \Bigg] \right| \\
&\le CE\left[\sum_{i=1}^{N_1^\varepsilon+1} |\hat{X}^d_{T_i^\varepsilon -} - \hat{X}^{dc}_{T_i^\varepsilon -}| \right] \\
&\qquad + \left| E\Bigg[\sum_{i=1}^{N_1^\varepsilon} \{u(T_i, \hat{X}^d_{T_i^\varepsilon -} + h(\hat{X}^d_{T_i^\varepsilon -})\Delta Z_{T_i^\varepsilon}) - u(T_i, \hat{X}^d_{T_i^\varepsilon -})\} \right. \\
&\qquad + \int_0^1 dt \Big\{ \frac{\partial u(t,\hat{X}_t^{dc})}{\partial t} + \frac{1}{2}\frac{\partial^2 u(t,\hat{X}_t^{dc})}{\partial x^2}\sigma_\varepsilon^2 h^2(\hat{X}_t^{dc}) + \frac{\partial u(t,\hat{X}_t^{dc})}{\partial x}\gamma_\varepsilon h(\hat{X}_t^{dc}) \\
&\qquad \left. + \frac{\sigma_\varepsilon^2}{2}\frac{\partial u(t,\hat{X}_t^{dc})}{\partial x} h(\hat{X}_t^{dc})(h'(\hat{X}_t^{dc}) - h'(\hat{X}_{\eta_t}^{dc})) \Big\} \Bigg] \right| \\
&= CE\left[\sum_{i=1}^{N_1^\varepsilon+1} |\hat{X}^d_{T_i^\varepsilon -} - \hat{X}^{dc}_{T_i^\varepsilon -}| \right] + \left| E\left[\int_0^1 dt \int_{\mathbb{R}} \chi_\varepsilon \nu(dy) \{u(t, \hat{X}_t^d + h(\hat{X}_t^d)y) \right. \right. \\
&\qquad \left. \left. - u(t,\hat{X}_t^{dc} + h(\hat{X}_t^{dc})y) + u(t,\hat{X}_t^d) - u(t,\hat{X}_t^{dc})\} \right] \right| \\
&\qquad + \left| E\Bigg[\int_0^1 dt \Big\{ \frac{\partial u(t,\hat{X}_t^{dc})}{\partial t} + \frac{1}{2}\frac{\partial^2 u(t,\hat{X}_t^{dc})}{\partial x^2}\sigma_\varepsilon^2 h^2(\hat{X}_t^{dc}) + \frac{\partial u(t,\hat{X}_t^{dc})}{\partial x}\gamma_\varepsilon h(\hat{X}_t^{dc}) \right. \right. \\
&\qquad + \frac{\sigma_\varepsilon^2}{2}\frac{\partial u(t,\hat{X}_t^{dc})}{\partial x} h(\hat{X}_t^{dc})(h'(\hat{X}_t^{dc}) - h'(\hat{X}_{\eta_t}^{dc})) \\
&\qquad \left. \left. + \int_{\mathbb{R}} \chi_\varepsilon \nu(dy)\{u(t,\hat{X}_t^{dc} + h(\hat{X}_t^{dc})y) - u(t,\hat{X}_t^{dc})\} \Big\} \Bigg] \right| \qquad (23)
\end{aligned}
$$

The third term above is exactly the same expression as in (20) and (21), with $\hat{X}$ replaced by $\hat{X}^{dc}$. Since lemma 11 also applies to $\hat{X}^{dc}$, the third term can be estimated in the same way as in the proof of the theorem 2 and yields the same error bound. It remains then to estimate the first and the second terms.

Let $\mathcal{F}^J := \sigma(N_t^\varepsilon, 0 \le t \le 1)$ and $\mathcal{F}_t^J := \mathcal{F}_t \vee \mathcal{F}^J$. Then the first term in the

right-hand side satisfies

$$
\begin{aligned}
E\left[\sum_{i=1}^{N_1^\varepsilon+1}|\hat{X}^d_{T_i^\varepsilon-}-\hat{X}^{dc}_{T_i^\varepsilon-}|\right] &= E\left[\sum_{i=1}^{N_1^\varepsilon+1}E\left[|\hat{X}^d_{T_i^\varepsilon-}-\hat{X}^{dc}_{T_i^\varepsilon-}|\big|\mathcal{F}^J_{T^\varepsilon_{i-1}}\right]\right]\\
&\le CE\Bigg[\sum_{i=1}^{N_1^\varepsilon+1}E\Big[|\gamma_\varepsilon(T_i^\varepsilon-T_{i-1}^\varepsilon)+\sigma_\varepsilon(W_{T_i^\varepsilon}-W_{T^\varepsilon_{i-1}})\\
&\qquad -\frac{1}{2}h'(\hat{X}^d_{T^\varepsilon_{i-1}})\sigma_\varepsilon^2(T_i^\varepsilon-T_{i-1}^\varepsilon)|^{q+1}\big|\mathcal{F}^J_{T^\varepsilon_{i-1}}\Big]\Bigg]\\
&\le CE\left[\sum_{i=1}^{N_1^\varepsilon+1}\left\{(|\gamma_\varepsilon|^{q+1}+1)(T_i^\varepsilon-T_{i-1}^\varepsilon)^{q+1}+\sigma_\varepsilon^{q+1}(T_i^\varepsilon-T_{i-1}^\varepsilon)^{\frac{q+1}{2}}\right\}\right]\\
&\le C(\gamma_\varepsilon^{q+1}+1)E\left[\int_0^1(t-\eta_t)^q dt\right]+C\sigma_\varepsilon^{q+1}E\left[\int_0^1(t-\eta_t)^{\frac{q-1}{2}}dt\right]\\
&\le C\frac{\gamma_\varepsilon^{q+1}+1}{\lambda_\varepsilon^q}+C\frac{\sigma_\varepsilon^{q+1}}{\lambda_\varepsilon^{\frac{q-1}{2}}},
\end{aligned}
$$

where we used lemma 10 in the last line.

Finally, the second term in the right-hand side of (23) can be estimated from above by:

$$
CE\left[\int_0^1 dt\int_{\mathbb{R}}\chi_\varepsilon\nu(dy)(1+|y|)|\hat{X}^d_t-\hat{X}^{dc}_t|\right]\le C\lambda_\varepsilon E\left[\int_0^1 dt|\hat{X}^d_t-\hat{X}^{dc}_t|\right],
$$

and from now on one proceeds similarly to the estimation of the first term above (and obtains the same bound). □

4 Approximating multidimensional SDE using expansions

In this section we propose an alternative approximation scheme, which yields similar rates to the ones obtained in section 3 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 d-dimensional Brownian motion W^ε with covariance matrix Σ^ε independent of Z, yielding the SDE

$$
d\bar{X}_t = h(\bar{X}_{t-})\{\gamma_\varepsilon dt + dW_t^\varepsilon + dZ_t^\varepsilon\}. \tag{24}
$$

This process can also be written as

$$
\begin{aligned}
\bar{X}(t) &= \bar{X}(\eta_t)+\int_{\eta_t}^t h\left(\bar{X}(s)\right)dW^\varepsilon(s)+\int_{\eta_t}^t h\left(\bar{X}(s)\right)\gamma_\varepsilon ds,\\
\bar{X}(T_i^\varepsilon) &= \bar{X}(T_i^\varepsilon-)+h(\bar{X}(T_i^\varepsilon-))\Delta Z(T_i^\varepsilon).
\end{aligned}
$$

The idea now is to expand the solution of (24) between the jumps of Z^ε around the solution of the deterministic dynamical system (2), treating the stochastic term as a small random perturbation (see [9, Chapter 2]).

Assume that the coefficient h is Lipschitz and consider a family of processes $(Y^\alpha)_{0\le\alpha\le 1}$ defined by

$$Y^\alpha(t) = \bar{X}(\eta_t) + \alpha \int_{\eta_t}^t h\left(Y^\alpha(s)\right) dW^\varepsilon(s) + \int_{\eta_t}^t h\left(Y^\alpha(s)\right)\gamma_\varepsilon ds$$

Our idea is to replace the process $\bar{X} := Y^1$ with its first-order Taylor approximation:

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

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

$$\begin{aligned}
\tilde{X}(t) &= Y^0(t) + Y_1(t), \quad t > \eta_t,\\
\tilde{X}(T_i^\varepsilon) &= \tilde{X}(T_i^\varepsilon -) + h(\tilde{X}(T_i^\varepsilon -))\Delta Z(T_i^\varepsilon),\\
Y^0(t) &= \tilde{X}(\eta_t) + \int_{\eta_t}^t h(Y^0(t))\gamma_\varepsilon ds\\
Y_1(t) &= \int_{\eta_t}^t \frac{\partial h}{\partial x_i}\left(Y^0(s)\right) Y_1^i(s)\gamma_\varepsilon ds + \int_{\eta_t}^t h\left(Y^0(s)\right) dW^\varepsilon(s)
\end{aligned} \tag{25}$$

where we used the Einstein convention for summation over repeated indices. Note that the random vector $Y_1(t)$ conditioned on T_i^ε, $i \in \mathbb{N}$, $t \in (T_j^\varepsilon, T_{j+1}^\varepsilon)$ and $\tilde{X}(T_j^\varepsilon)$ is a Gaussian random vector with conditional covariance matrix $\Omega(t)$ which satisfies the (matrix) linear equation

$$\Omega(t) = \int_{\eta_t}^t (\Omega(s)M(s) + M^\perp(s)\Omega^\perp(s) + N(s))ds \tag{26}$$

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

$$M_{ij}(t) = \frac{\partial h_{jk}(Y_0(t))}{\partial x_i}\gamma_\varepsilon^k \quad \text{and} \quad N(t) = h(Y_0(t))\Sigma^\varepsilon h^\perp(Y_0(t)).$$

Successive applications of Gronwall's inequality yield the following bounds for Y_0 and Ω.

$$\begin{aligned}
&|Y_0(t)| \le \left(|\tilde{X}(\eta_t)| + C|\gamma_\varepsilon|(t-\eta_t)\right) e^{C|\gamma_\varepsilon|(t-\eta_t)}\\
&\|\Omega(t)\| \le C\|\Sigma_\varepsilon\|(t-\eta_t)(1 + (|\tilde{X}(\eta_t)| + C|\gamma_\varepsilon|(t-\eta_t))^2)e^{3C|\gamma_\varepsilon|(t-\eta_t)}.
\end{aligned} \tag{27}$$

Lemma 14. *Under our standing assumptions $\nu(\mathbb{R}^d) = \infty$ and $\int_{|y|>1} |y|\nu(dy) < \infty$,*

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

Proof. Left to the reader as an exercise. □

To prove the result on weak convergence (Theorem 16), we need to generalize lemmas 11, 12 and 13 to the multidimensional 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 15 (Bounds on moments of $\tilde{X}$)**.** *Assume $h \in C_b^1(\mathbb{R}^n)$ and*

$$\int_{z\in\mathbb{R}^d} |z|^p \nu(dz) < \infty \quad \textit{for some } p \geq 2.$$

Then there exists a constant $C > 0$ (which may depend on p but not on ε) such that for all ε sufficiently small,

$$E[\sup_{0\leq s\leq 1} |\tilde{X}_s|^p] \leq C(1 + |x|^p). \tag{28}$$

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

$$d\tilde{X}_t = h_{t-} d\hat{Z}_t^\varepsilon + \tilde{h}_t dW_t^\varepsilon + h_t\tilde{\gamma}dt + (\tilde{h}_t - h_t)\gamma_\varepsilon dt + \frac{\partial h}{\partial x_i}(Y_0(t))\, Y_1^i(t)\gamma_\varepsilon dt, \tag{29}$$

$$\text{where} \quad \hat{Z}_t = \int_0^t \int_{|y|>\varepsilon} y\hat{N}^\varepsilon(dy, ds), \qquad \tilde{\gamma} = \gamma + \int_{|y|>1} y\nu(dy).$$

By predictable Burkholder-Davis-Gundy inequality [7, lemma 2.1], we then have

$$\begin{aligned}
E\left[\sup_{0\leq s\leq t} \|\tilde{X}_s\|^p\right] &\leq CE\left[\|x\|^p + \left(\int_0^t \|h_s\|^2 \int |z|^2\chi_\varepsilon\nu(dz)ds\right)^{p/2}\right. \\
&\quad + \int_0^t \|h_s\|^p ds \int |z|^p\chi_\varepsilon\nu(dz) + \left(\int_0^t \|\tilde{h}_s\|^2\|\Sigma_\varepsilon\|ds\right)^{p/2} \\
&\quad \left. + \int_0^t \|h_s\|^p|\tilde{\gamma}|^p ds + |\gamma_\varepsilon|^p \int_0^t \|\tilde{h}_s - h_s\|^p ds + |\gamma_\varepsilon|^p \int_0^t \|Y_1(s)\|^p ds\right] \\
&\leq CE\left[\|x\|^p + \int_0^t \|h_s\|^p ds + (1 + |\gamma_\varepsilon|^p)\int_0^t \|\tilde{h}_s - h_s\|^p ds + |\gamma_\varepsilon|^p \int_0^t \|Y_1(s)\|^p ds\right],
\end{aligned}$$

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

$$E\left[\sup_{0\leq s\leq t} \|\tilde{X}_s\|^p\right] \leq CE\left[\|x\|^p + \int_0^t \|h_s\|^p ds + (1 + |\gamma_\varepsilon|^p)\int_0^t \|Y_1(s)\|^p ds\right]. \tag{30}$$

Using (27), the last term can be estimated as

$$E\left[(1+|\gamma_\varepsilon|^p)\int_0^t \|Y_1(s)\|^p ds\right] = CE\left[(1+|\gamma_\varepsilon|^p)\int_0^t E\left[\|Y_1(s)\|^p|\mathcal{F}_{\eta(s)}\right] ds\right]$$
$$\leq CE\left[(1+|\gamma_\varepsilon|^p)\int_0^t \|\Omega(s)\|^{p/2} ds\right]$$
$$\leq CE\left[(1+|\gamma_\varepsilon|^p)\int_0^t (1+\|\tilde{X}_{\eta(s)})\|^p)\|\Sigma_\varepsilon\|^{p/2}(s-\eta_s)^{p/2}e^{C|\gamma_\varepsilon|(1+|\gamma_\varepsilon|)(s-\eta_s)} ds\right]$$
$$\leq CE\left[(1+|\gamma_\varepsilon|^p)\int_0^t (1+\|\tilde{X}_{\eta(s)})\|^p)\|\Sigma_\varepsilon\|^{p/2}\tau^{p/2}e^{C(1+|\gamma_\varepsilon|)\tau} ds\right],$$

where τ is an independent exponential random variable with parameter λ_ε. Due to Lemma 14, for ε sufficiently small, the expectation with respect to τ exists, and computing it explicitly we obtain

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

where $C_\varepsilon \to 0$ as $\varepsilon \to 0$. Inequality (30) therefore becomes

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

which implies that for ε sufficiently small,

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

and we get (28) by Gronwall's lemma. □

Theorem 16.

(i) Assume $(\mathbf{H_3})$ *or* $(\mathbf{H_3'})+(\mathbf{A})$. *Then*

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

(ii) Assume $(\mathbf{H_4})$ *or* $(\mathbf{H_4'})+(\mathbf{A})$, *let* γ_ε *be bounded and suppose that for some measure* ν_0

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

for all i,j,k *and all* ε *sufficiently small. Then*

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

Proof. By Itô formula and (29) we have

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

$$= \int_0^1 dt E\left[\frac{\partial u(t,\tilde{X}(t))}{\partial x_i}\left\{h_{ij}(Y^0(t)) + \frac{\partial h_{ij}(Y^0(t))}{\partial x_k}Y_1^k(t) - h_{ij}(\tilde{X}(t))\right\}\gamma_\varepsilon^j\right. \tag{31}$$

$$+ \frac{1}{2}\frac{\partial^2 u(\tilde{X}(t))}{\partial x_i \partial x_j}h_{ik}(Y_0(t))\Sigma_\varepsilon^{kl}h_{jl}(Y_0(t)) \tag{32}$$

$$\left. - \int_{\mathbb{R}^d}\left\{u(t,\tilde{X}(t) + h(\tilde{X}(t))y) - u(t,\tilde{X}(t)) - \frac{\partial u(t,\tilde{X}_t)}{\partial x_i}h_{ij}(\tilde{X}(t))y_j\right\}\bar{\chi}_\varepsilon\nu(dy)\right]. \tag{33}$$

Denote the expectation term in (31) by A_t and the sum of the terms in (32) and (33) by B_t. The term A_t satisfies

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

Under the assumption ($\mathbf{H_4}$) or ($\mathbf{H_3}$), using (27), we have

$$\begin{aligned}|A_t| &\le C|\gamma_\varepsilon|E\left[|Y_1(t)|^2\right] \le C|\gamma_\varepsilon|E[\|\Omega(t)\|] \\ &\le C\|\Sigma_\varepsilon\|\left\{E[e^{3C|\gamma_\varepsilon|(t-\eta_t)}(t-\eta_t)](1+E[|\tilde{X}(\eta_t)|^2])|\gamma_\varepsilon|\right. \\ &\left. + C^2|\gamma_\varepsilon|^2E[e^{3C|\gamma_\varepsilon|(t-\eta_t)}(t-\eta_t)^3]\right\}.\end{aligned}$$

Using Lemmas 10, 14 and 15, we then get

$$\begin{aligned}\int_0^1 |A_t|dt &\le C(1+|x|^2)\frac{\|\Sigma_\varepsilon\|\lambda_\varepsilon|\gamma_\varepsilon|}{(\lambda_\varepsilon - 3C|\gamma_\varepsilon|)^2}\left\{1 + \frac{|\gamma_\varepsilon|}{(\lambda_\varepsilon - 3C|\gamma_\varepsilon|)^2}\right\} \\ &\le C(1+|x|^2)\frac{\|\Sigma_\varepsilon\|}{\lambda_\varepsilon}\left\{1 + \frac{|\gamma_\varepsilon|}{\lambda_\varepsilon^2}\right\}|\gamma_\varepsilon|\end{aligned}$$

Under ($\mathbf{H_4'}$) or ($\mathbf{H_3'}$) this result can be obtained along the same lines.

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

$$\begin{aligned}B_t = E\Bigg[&-\int_{\mathbb{R}^d}\Big\{u(t,\tilde{X}(t)+h(\tilde{X}(t))y) - u(t,\tilde{X}(t)) - \frac{\partial u(t,\tilde{X}(t))}{\partial x_i}h_{ij}(\tilde{X}(t))y_j \\ &- \frac{\partial^2 u(t,\tilde{X}_t)}{\partial x_i\partial x_j}h_{ik}(\tilde{X}(t))h_{jl}(\tilde{X}(t))y_k y_l\Big\}\bar{\chi}_\varepsilon\nu(dy) \\ &+ \frac{\partial^2 u(t,Y^0(t))}{\partial x_i\partial x_j}\Sigma_\varepsilon^{kl}\left\{h_{ik}(Y^0(t))h_{jl}(Y^0(t)) - h_{ik}(\tilde{X}(t))h_{jl}(\tilde{X}(t))\right\} \\ &+ \left(\frac{\partial^2 u(t,\tilde{X}(t))}{\partial x_i\partial x_j} - \frac{\partial^2 u(t,Y^0(t))}{\partial x_i\partial x_j}\right)\Sigma_\varepsilon^{kl}\left\{h_{ik}(Y^0(t))h_{jl}(Y^0(t)) - h_{ik}(\tilde{X}(t))h_{jl}(\tilde{X}(t))\right\}\Bigg]\end{aligned}$$

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 $(\mathbf{H_4})$ or $(\mathbf{H_3})$, we get

$$
\begin{aligned}
&|B_1(t)| \\
&\leq E\left|\int_{\mathbb{R}^d} \frac{\partial^3 u(t,\tilde{X}_t+\theta h(\tilde{X}(t))y)}{\partial x_i \partial x_j \partial k} h_{il}(\tilde{X}(t))h_{jm}(\tilde{X}_t)h_{kn}(\tilde{X}(t))y_l y_m y_n \bar{\chi}_\varepsilon \nu(dy)\right| \\
&\leq CE[\|h(\tilde{X}_t)\|^3]\int_{\mathbb{R}^d}|y|^3\bar{\chi}_\varepsilon\nu(dy) \leq C(1+|x|^3)\int_{\mathbb{R}^d}|y|^3\bar{\chi}_\varepsilon\nu(dy).
\end{aligned}
$$

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

$$
\begin{aligned}
|B_2(t)| &\leq E\left[\left|\frac{\partial^2 u(t,Y^0(t))}{\partial x_i \partial x_j}\Sigma_\varepsilon^{kl} E\left\{F_{ikjl}(Y^0(t)) - F_{ikjl}(\tilde{X}(t))\Big/\mathcal{F}_{\eta_t}\right\}\right|\right] \\
&\leq E\left[\left|\frac{\partial^2 u(t,Y^0(t))}{\partial x_i \partial x_j}\Sigma_\varepsilon^{kl} Y_1^p(t)Y_1^q(t)\int_0^1(1-\theta)\frac{\partial^2 F_{ikjl}(Y^0(t)+\theta Y_1(t))}{\partial x_p \partial x_q}d\theta\right|\right] \\
&\leq C\|\Sigma_\varepsilon\| E\left[|Y_1(t)|^2(1+\sup_{s\leq t}|\tilde{X}_s|)\right].
\end{aligned}
$$

Using once again Lemmas 10, 14 and 15, we obtain

$$
\int_0^1 |B_2(t)|dt \leq \frac{C\|\Sigma_\varepsilon\|^2}{\lambda_\varepsilon}(1+|x|)^3\left\{1+\frac{|\gamma_\varepsilon|^4}{\lambda_\varepsilon^4}\right\}^{1/2}.
$$

With a similar reasoning we obtain that B_3 is also upper bounded as B_2

The proof of the first part of the theorem under assumptions $(\mathbf{H_4'})$ or $(\mathbf{H_3'})$ is done along the same lines. □

5 Examples and applications

5.1 Example 1: Weak convergence for an SDE driven by a NIG Lévy process

In our first example, we verify the theoretical results on a concrete example of a SDE driven by a normal inverse Gaussian (NIG) Lévy process [2], which has characteristic function

$$
\phi_t(u) := E[e^{iuZ_t}] = \exp\left\{-\delta t\left(\sqrt{\alpha^2-(\beta-iu)^2}-\sqrt{\alpha^2-\beta^2}\right)\right\},
$$

where $\alpha > 0$, $\beta \in (-\alpha, \alpha)$ 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 the modified Bessel function of the second kind. The NIG process has stable-like behavior of small jumps with $\nu(x) \sim \frac{const}{|x|^2}$, $x \to 0$, and exponential tails with $\nu(x) \sim const \times |x|^{-3/2} e^{-\alpha|x|+\beta x}$, $x \to \pm\infty$.

The NIG process can be represented as $Z_t = \theta Y_t + \sigma W_{Y_t}$, where W is a standard Brownian motion and Y is an inverse Gaussian subordinator: a pure jump Lévy process with Lévy density $\rho(x) = \frac{1}{\sqrt{2\pi k}} \frac{e^{-\frac{x}{2\kappa}}}{|x|^{3/2}}$. The parameters (σ, θ, κ) are related to (α, β, δ) via

$$\kappa = \frac{1}{\delta\sqrt{\alpha^2 - \beta^2}}, \quad \theta = \frac{\beta\delta}{\sqrt{\alpha^2 - \beta^2}}, \quad \sigma^2 = \frac{\delta}{\sqrt{\alpha^2 - \beta^2}}.$$

Thanks to this representation, increments of the NIG process can be simulated explicitly (see [6, algorithms 6.9 and 6.10]), which enables us to compare our jump-adapted algorithms with the classical Euler scheme.

Since the Lévy density ν involves a special function, simulation from the density $\frac{\nu(x)1_{|x|>\varepsilon}}{\int_{|x|>\varepsilon}\nu(x)dx}$ is rather intricate. We therefore propose another choice of the truncation function χ_ε which is based on the time-changed Brownian motion representation of the NIG process and simplifies the simulation. Let p_t be the (Gaussian) density of $\theta t + \sigma W_t$. The Lévy density of the NIG process can be represented as [22, Theorem 30.1]

$$\nu(x) = \int_0^\infty p_t(x)\rho(t)dt.$$

We first define the (finite) measure ν_ε via

$$\nu_\varepsilon(x) := \int_\varepsilon^\infty p_t(x)\rho(t)dt,$$

and then the function χ_ε as the ratio of the two densities

$$\chi_\varepsilon(x) := \frac{\nu_\varepsilon(x)}{\nu(x)}.$$

It is easy to check that this function satisfies the required conditions as well as the assumption (A). The constants λ_ε, γ_ε and σ_ε are computed as follows:

$$\begin{aligned}
\lambda_\varepsilon &= \int_{\mathbb{R}} \nu_\varepsilon(x)dx = \int_\varepsilon^\infty \rho(t)dt = \frac{\sqrt{2}}{\sqrt{\pi\kappa\varepsilon}} e^{-\frac{\varepsilon}{2\kappa}} - \frac{2}{\kappa} N\left(-\sqrt{\frac{\varepsilon}{\kappa}}\right), \\
\gamma_\varepsilon &= \theta - \int_{\mathbb{R}} x\nu_\varepsilon(x)dx = \theta - \theta\int_\varepsilon^\infty t\rho(t)dt = \theta - 2\theta N\left(-\sqrt{\frac{\varepsilon}{\kappa}}\right) \\
\sigma_\varepsilon^2 &= \int_{\mathbb{R}} x^2(\nu(x) - \nu_\varepsilon(x))dx = \int_{\mathbb{R}} x^2 \int_0^\varepsilon p_t(x)\rho(t)dxdt = \int_0^\varepsilon (\theta^2 t^2 + \sigma^2 t)\rho(t)dt \\
&= (\sigma^2 + \kappa\theta^2)\left(1 - 2N\left(-\sqrt{\frac{\varepsilon}{\kappa}}\right)\right) - \frac{\sqrt{2\kappa\varepsilon}}{\sqrt{\pi}} e^{-\frac{\varepsilon}{2\kappa}}
\end{aligned}$$

where $N(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{x^2}{2}} dx$ is the standard normal CDF.

Random variables with density $\frac{\nu_\varepsilon}{\lambda_\varepsilon}$ can be sampled using the following algorith:

- First, sample a random variable Y with probability density $\frac{\rho(x)1_{x>\varepsilon}}{\lambda_\varepsilon}$ using the rejection method (example 6.9 in [6]);
- Conditionally on Y, sample X with law p_Y.

For our numerical example we choose to solve the one-dimensional SDE

$$dX_t = \sin(aX_t)dZ_t,$$

where Z is the NIG Lévy process (with drift adjusted to have $E[Z_t] = 0$). The solution of the corresponding deterministic ODE

$$dX_t = \sin(aX_t)dt, \quad X_0 = x$$

is given explicitly by

$$X_t = \theta(t; x) = \frac{1}{a} \arccos \frac{1 + \cos(ax) - e^{2at}(1 - \cos(ax))}{1 + \cos(ax) + e^{2at}(1 - \cos(ax))}$$

We compare the performance of the jump-adapted scheme of section 3 and of the one-dimensional version of the scheme of section 4 with the classical Euler scheme. In the one-dimensional setting, equation (26) simplifies and admits an explicit solution

$$\Omega(t) = \sigma_\varepsilon^2 h^2(Y_t^0)(t - \eta_t).$$

We compute the Monte Carlo estimator of $E[f(X_1)]$ with

$$f(x) = 2 - 2\cos(x - X_0).$$

This choice was motivated by the desire to have a function similar to x^2 but with bounded derivatives.

Figure 1 presents the Monte Carlo estimator and the corresponding errors obtained using the three schemes with parameter values $\sigma = 0.5$, $\theta = 0.4$, $\kappa = 0.6$, $a = 5$ and $X_0 = 1$. The true value (obtained with a very large number of simulations) is close to 0.13045 in this case. The estimators and errors are plotted as function of the complexity parameter N which is equal to λ_ε (that is, the average number of discretization points) for the jump-adapted schemes and to the number of discretization points for the Euler scheme. For given N, the computation using any of the jump-adapted schemes takes about 2.5 times as much time as with the Euler scheme. We simulated 10^6 trajectories for each point, leading to a Monte Carlo standard deviation of about 3×10^{-4} for every value (the standard deviation is almost independent from N).

Both jump-adapted schemes appear largely superior to the Euler scheme, and the scheme of section 3 has a better performance than that of section 4. The estimator obtained using the scheme of section 3 falls within the Monte Carlo

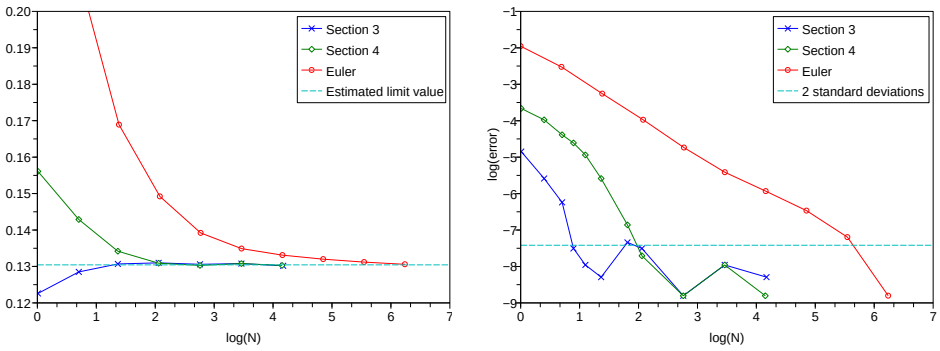

Figure 1: Numerical comparison of the jump-adapted scheme of section 3, jump-adapted scheme of section 4 and the classical Euler scheme. Left graph: the Monte Carlo estimator of $E[f(X_1)]$. The dotted line corresponds to the "true value" computed with a very large number of simulations. Right graph: log-log plot of absolute error. The dotted line corresponds to the logarithm of two standard deviations of the MC estimator.

confidence bounds already for $N = 3$, and the estimator of section 4 converges after $N = 8$, whereas the Euler scheme only converges at about $N = 300$. From the log-log plot, one can clearly identify the usual convergence rate of $\frac{1}{N}$ for the Euler scheme, whereas for the other two schemes, after a certain warm-up period, the convergence is much faster than $\frac{1}{N}$ and looks more like $\frac{1}{N^2}$. This is consistent with the theoretical result of Theorems 2 which predicts a rate of $O(\lambda_\varepsilon^{-2})$ (see example 5).

5.2 Example 2: A financial application: Libor market model with jumps

In the example we treat in this Section, the theoretical resuts of this paper establishing the convergence rate of the weak error cannot formally be applied due to the non-smoothness of the function f. Nevertheless, the scheme itself can be applied, and, as shown by numerical experiments, the weak error converges to zero very quickly. This shows that the methodology we have introduced can be applied with greater generality.

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. 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 [13], Glasserman and

Kou [10] and Eberlein and Özkan [8]. 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^i is the forward interest rate, defined at date t for the period $[T_i, T_{i+1}]$. If $B_t(T)$ 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

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

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 [13], supposing that all Libor rates are driven by a d-dimensional pure jump Lévy process. In this case, following [13], an arbitrage-free dynamics of $L_t^1, \ldots, L_t^n$ can be constructed via the multi-dimensional SDE

$$\frac{dL_t^i}{L_{t-}^i} = \sigma^i(t)dZ_t - \int_{\mathbb{R}^d} \sigma^i(t)z \left[\prod_{j=i+1}^{n}\left(1 + \frac{\delta L_t^j \sigma^j(t)z}{1+\delta L_t^j}\right) - 1\right]\nu(dz)dt. \qquad (34)$$

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

Equation (34) 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 $H = h(L_{T_1}^1, \ldots, L_{T_1}^n)$ at time T_1 is given by

$$\begin{aligned}
\pi_t(H) &= B_t(T_{n+1})E\left[\frac{h(L_{T_1}^1, \ldots, L_{T_1}^n)}{B_{T_1}(T_{n+1})}\middle|\mathcal{F}_t\right] \\
&= B_t(T_{n+1})E\left[h(L_{T_1}^1, \ldots, L_{T_1}^n)\prod_{i=1}^{n}(1+\delta L_{T_1}^i)\middle|\mathcal{F}_t\right] \\
&= \frac{B_t(T_1)}{\prod_{i=1}^{n}(1+\delta L_t^i)}E\left[h(L_{T_1}^1, \ldots, L_{T_1}^n)\prod_{i=1}^{n}(1+\delta L_{T_1}^i)\middle|\mathcal{F}_t\right] \qquad (35)
\end{aligned}$$

The price of any such option can therefore be computed by Monte Carlo using equation (34).

Introducing a $n+1$-dimensional state process X_t with $X_t^0 \equiv t$ and $X_t^i = L_t^i$ for $1 \le i \le n$, a $d+1$-dimensional driving Lévy process $\tilde{Z}_t = (t \quad Z_t)^\perp$, 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 $h_{ij} = \sigma_{j-1}^i(x_0)$ with

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

we see that the equation (34) takes the homogeneous form $dX_t = h(X_{t-})d\tilde{Z}_t$, to which the discretization scheme of section 4 can be readily applied.

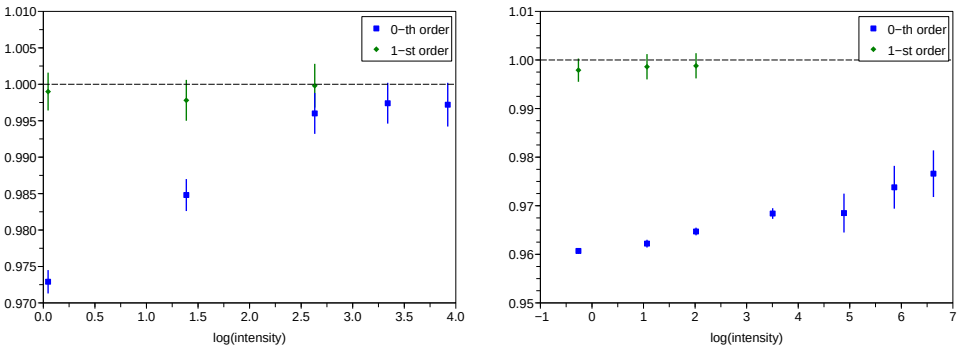

Figure 2: Ratio of estimated to theoretical zero coupon bond price in Case 1 (left) and Case 2 (right) with 1 standard deviation bounds. 10^5 trajectories were used for all points except the three points with the largest intensity in the right graph, where 10^4 simulations were made to save time.

For the purposes of illustration, we simplify the model even further, taking $d = 1$ and supposing that the functions $\sigma^i(t)$ are constant: $\sigma^i(t) \equiv 1$. The driving Lévy process Z is supposed to follow the CGMY model [6] which has Lévy density

$$\nu(x) = \frac{Ce^{-\lambda_-|x|}}{|x|^{1+\alpha}} 1_{x<0} + \frac{Ce^{-\lambda_+|x|}}{|x|^{1+\alpha}} 1_{x>0}. \tag{36}$$

No easy algorithm is available for simulating the increments of this process (cf [15, 18]), however, it is straightforward to simulate random variables with density

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

using the rejection method [6, example 6.9].

We use 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 (25) and (26) simultaneously using the classical 4-th order Runge-Kutta scheme as described in Section 3.

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 (35). 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 Y^0 only. This illustrates the impact of using the Asmussen-Rosinski type approximation as compared to neglecting the small jumps completely. 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 the approximating process equal to 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. Recall that the theoretical convergence rates of Theorem 16 and Example 5 (which formally do not apply here) are of order of λ^{-4} in Case 1 and $\lambda^{-1.11}$ in Case 2.

This is consistent with the theoretical convergence rates which are of order of λ^{-3} in Case 1 and $\lambda^{-0.11}$ in Case 2 for 0-th order approximation and, respectively, λ^{-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

$$\begin{aligned} h(L^1_{T_1},\dots,L^n_{T_1}) &= \left(1 - B_{T_1}(T_{n+1}) - K\delta\sum_{i=1}^{n} B_{T_1}(T_{i+1})\right)^+ \\ &= \left(\prod_{i=1}^{n}(1+\delta L^i_{T_1})^{-1} - 1 - K\delta\sum_{i=1}^{n}\prod_{j=1}^{i}(1+\delta L^i_{T_1})^{-1}\right)^+ . \end{aligned}$$

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.

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 λ_ε (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.

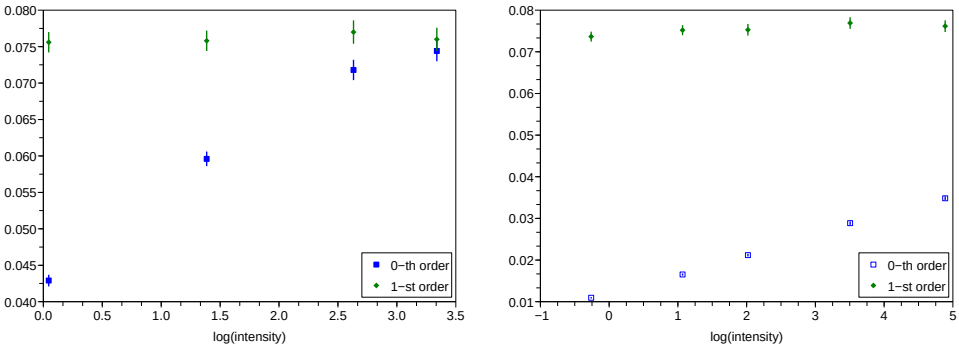

Figure 3: Estimated price of an ATM receiver swaption with maturity 5 years in Case 1 (left) and Case 2 (right) with 1 standard deviation bounds. The Monte Carlo simulation was performed with 10^5 trajectories.

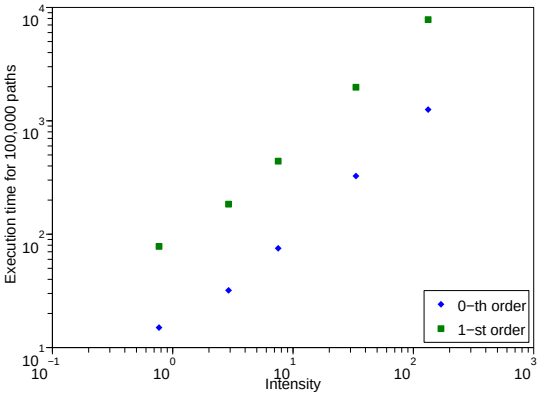

Figure 4: Execution times for 10^5 trajectories on a Pentium-III PC.

Acknowledgement

The 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.

The research of Arturo Kohatsu-Higa has been supported by grants of the Japanese Goverment.

References

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

[2] O. Barndorff-Nielsen, *Processes of normal inverse Gaussian type*, Finance Stoch., 2 (1998), pp. 41–68.

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

[4] P. Carr, H. Geman, D. Madan, and M. Yor, *The fine structure of asset returns: An empirical investigation*, Journal of Business, 75 (2002), pp. 305–332.

[5] S. Cohen and J. Rosiński, *Gaussian approximation of multivariate lévy processes with applications to simulation of tempered stable processes*, Bernoulli, 13 (2007), pp. 195–210.

[6] R. Cont and P. Tankov, *Financial Modelling with Jump Processes*, Chapman & Hall / CRC Press, 2004.

[7] K. Dzhaparidze and E. Valkeila, *On the Hellinger type distances for filtered experiments*, Probab. Theor. Relat. Fields, 85 (1990), pp. 105–117.

[8] E. Eberlein and F. Ozkan, *The Lévy LIBOR model*, Finance and Stochastics, 9 (2005), pp. 327–348.

[9] M. I. Freidlin and A. D. Wentzell, *Random perturbations of dynamical systems*, vol. 260 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag, New York, second ed., 1998. Translated from the 1979 Russian original by Joseph Szücs.

[10] P. Glasserman and S. Kou, *The term structure of forward rates with jump risk*, Math. Finance, 13 (2003), pp. 383–410.

[11] E. Hairer, S. P. Nøorsett and G. Wanner, *Solving Ordinary Differential Equations I: Nonstiff Problems*, Springer-Verlag, New York, second ed., 1993.

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

[13] F. Jamshidian, *LIBOR market models with semimartingales*, tech. report, NetAnalytics Ltd., 1999.

[14] V. Konakov, S. Menozzi, *Weak Error for stable driven SDEs: expansion of the densities*. Preprint.

[15] D. Madan and M. Yor, *CGMY and Meixner subordinators are absolutely continuous with respect to one sided stable subordinators*. Prépublication du Laboratiore de Probabilités et Modéles Aléatoires.

[16] E. Mordecki, A. Szepessy, R. Tempone, and G. E. Zouraris, *Adaptive weak approximation of diffusions with jumps*, SIAM J. Numer. Anal., 46 (2008), pp. 1732–1768.

[17] J. Picard, *On the existence of smooth densities for jump processes*, Probab. Theory Rel. Fields 105 (1996), 481-511.

[18] J. Poirot and P. Tankov, *Monte Carlo option pricing for tempered stable (CGMY) processes*, Asia-Pacific Financial Markets, 13 (2006), pp. 327–344.

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

[20] J. Rosiński, *Series representations of Lévy processes from the perspective of point processes*, in Lévy Processes — Theory and Applications, O. Barndorff-Nielsen, T. Mikosch, and S. Resnick, eds., Birkhäuser, Boston, 2001.

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

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