ESAIM: PS 15 (2011) 233–248
DOI: 10.1051/ps/2009017

ESAIM: Probability and Statistics
www.esaim-ps.org

SIMULATION AND APPROXIMATION OF LÉVY-DRIVEN STOCHASTIC DIFFERENTIAL EQUATIONS

Nicolas Fournier[1]

Abstract. We consider the approximate Euler scheme for Lévy-driven stochastic differential equations. We study the rate of convergence in law of the paths. We show that when approximating the small jumps by Gaussian variables, the convergence is much faster than when simply neglecting them. For example, when the Lévy measure of the driving process behaves like $|z|^{-1-\alpha}\mathrm{d}z$ near 0, for some $\alpha \in (1,2)$, we obtain an error of order $1/\sqrt{n}$ with a computational cost of order n^α. For a similar error when neglecting the small jumps, see [S. Rubenthaler, Numerical simulation of the solution of a stochastic differential equation driven by a Lévy process. *Stochastic Process. Appl.* **103** (2003) 311–349], the computational cost is of order $n^{\alpha/(2-\alpha)}$, which is huge when α is close to 2. In the same spirit, we study the problem of the approximation of a Lévy-driven S.D.E. by a Brownian S.D.E. when the Lévy process has no large jumps. Our results rely on some results of [E. Rio, Upper bounds for minimal distances in the central limit theorem. *Ann. Inst. Henri Poincaré Probab. Stat.* **45** (2009) 802–817] about the central limit theorem, in the spirit of the famous paper by Komlós-Major-Tsunády [J. Komlós, P. Major and G. Tusnády, An approximation of partial sums of independent rvs and the sample df I. *Z. Wahrsch. verw. Gebiete* **32** (1975) 111–131].

Mathematics Subject Classification. 60H35, 60H10, 60J75.

Received April 10, 2009. Revised September 11, 2009.

1. Introduction

Let $(Z_t)_{t\geq 0}$ be a one-dimensional square integrable Lévy process. Then for some $a \in \mathbb{R}$, $b \in \mathbb{R}_+$ and some measure ν on $\mathbb{R}_* := \mathbb{R}\backslash\{0\}$ satisfying $\int_{\mathbb{R}_*} z^2\nu(\mathrm{d}z) < \infty$,

$$Z_t = at + bB_t + \int_0^t \int_{\mathbb{R}_*} z\tilde{N}(\mathrm{d}s, \mathrm{d}z), \tag{1.1}$$

where $(B_t)_{t\geq 0}$ is a standard Brownian motion, independent of a Poisson measure $N(\mathrm{d}s, \mathrm{d}z)$ on $[0,\infty)\times\mathbb{R}_*$ with intensity measure $\mathrm{d}s\nu(\mathrm{d}z)$ and where $\tilde{N}$ is its compensated Poisson measure, see Jacod-Shiryaev [10].

We consider, for some $x \in \mathbb{R}$ and some function $\sigma : \mathbb{R} \mapsto \mathbb{R}$, the S.D.E.

$$X_t = x + \int_0^t \sigma(X_{s-})\mathrm{d}Z_s. \tag{1.2}$$

Keywords and phrases. Lévy processes, stochastic differential equations, Monte-Carlo methods, simulation, Wasserstein distance.

[1] LAMA UMR 8050, Faculté de Sciences et Technologies, Université Paris Est, 61 avenue du Général de Gaulle, 94010 Créteil Cedex, France; nicolas.fournier@univ-paris12.fr

Article published by EDP Sciences

© EDP Sciences, SMAI 2011

https://doi.org/10.1051/ps/2009017 Published online by Cambridge University Press

Using some classical results (see *e.g.* Ikeda-Watanabe [5]), there is strong existence and uniqueness for (1.2) as soon as σ is Lipschitz continuous: for any given couple (B,N), there exists an unique càdlàg adapted solution $(X_t)_{t\geq 0}$ to (1.2). By *adapted*, we mean adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$ generated by (B,N).

We consider two related problems in this paper. The first one deals with the numerical approximation of the solution $(X_t)_{t\geq 0}$. It is discussed in the next section in the case where σ is bounded and Lipschitz continuous. We extend our study in Section 7 to the case where σ is locally Lipschitz continuous with at most linear growth, and where the Lévy measure ν satisfies only $\int_{\mathbb{R}_*} \min(z^2,1)\nu(\mathrm{d}z) < \infty$.

The second problem concerns the approximation of $(X_t)_{t\geq 0}$ by the solution to a Brownian S.D.E., when Z has only small jumps, and is discussed in Section 3.

Our results are based on a recent work of Rio [13] that concerns the rate of convergence in the central limit theorem, when using the quadratic Wasserstein distance. This result and its application to Lévy processes are discussed in Section 4.

The proofs are handled in Sections 5 and 6. We give some numerical illustrations in Section 8.

2. Numerical simulation

The first goal of this paper is to study a numerical scheme to solve (1.2). The first idea is to perform an Euler scheme $(X^n_{i/n})_{n\geq 0}$ with time-step $1/n$, see Jacod [6], Jacod-Protter [9], Protter-Talay [12] for rates of convergence. However, this is generally not a good scheme in practise, unless one knows how to simulate the increments of the underlying Lévy process, which is the case *e.g.* when Z is a stable process.

We assume here that the Lévy measure ν is known explicitly: one can thus simulate random variables with law $\nu(\mathrm{d}z)\mathbf{1}_A(z)/\nu(A)$, for any A such that $\nu(A)<\infty$.

The first idea is to approximate the increments of Z by $\widehat{\Delta}^{n,\epsilon}_i = Z^\epsilon_{i/n} - Z^\epsilon_{(i-1)/n}$, where Z^ϵ_t is the same Lévy process as Z without its (compensated) jumps smaller than ϵ. However, Asmussen-Rosinski [1] have shown that for a Lévy process with many small jumps, it is more convenient to approximate small jumps by some Gaussian variables than to neglect them. We thus introduce $\Delta^{n,\epsilon}_i = \widehat{\Delta}^{n,\epsilon}_i + U^{n,\epsilon}_i$, where $U^{n,\epsilon}_i$ is Gaussian with same mean and variance as the neglected jumps and is independent of $\widehat{\Delta}^{n,\epsilon}_i$. The arguments of [1] concern only Lévy processes and rely on explicit computations.

Let us write $(\widehat{X}^{n,\epsilon}_{[nt]/n})_{t\geq 0}$ (resp. $(X^{n,\epsilon}_{[nt]/n})_{t\geq 0}$) for the Euler scheme using the approximate increments $(\widehat{\Delta}^{n,\epsilon}_i)_{i\geq 1}$ (resp. $(\Delta^{n,\epsilon}_i)_{i\geq 1}$). They of course have a similar computational cost.

Jacod-Kurtz-Méléard-Protter [8] have computed systematically the *weak* error for the *approximate* Euler scheme. In particular, they prove some very precise estimates of $\mathbb{E}[g(X^{n,\epsilon}_{[nt]/n})] - \mathbb{E}[g(X_t)]$ for g smooth enough. The obtained rate of convergence is very satisfying.

Assume now that the goal is to approximate some functional of the path of the solution (*e.g.* $\sup_{t\in[0,T]}|X_t|$). Then we have to estimate the error between the laws of the paths of the processes (not only between the laws of the time marginals). A common way to perform such an analysis is to introduce a suitable coupling between the numerical scheme $(X^{n,\epsilon}_{[nt]/n})_{t\geq 0}$ and the true solution $(X_t)_{t\geq 0}$ and to estimate the (discretized) *strong* error $\mathbb{E}[\sup_{t\in[0,T]}|X^{n,\epsilon}_{[nt]/n} - X_{[nt]/n}|^2]$. We refer to Jacod-Jakubowski-Mémin [7] for the speed of convergence of the discretized process $(X_{[nt]/n})_{t\geq 0}$ to the whole process $(X_t)_{t\geq 0}$.

Rubenthaler [14] has studied the strong error when neglecting small jumps. He obtains roughly $\mathbb{E}[\sup_{t\in[0,T]}|\widehat{X}^{n,\epsilon}_{[nt]/n} - X_{[nt]/n}|^2] \simeq C_T(n^{-1} + \int_{|z|\leq\epsilon} z^2\nu(\mathrm{d}z))$ (if $b\neq 0$). For ν very singular near 0, the obtained precision is very low. Let us mention that in some particular cases, a more efficient method was introduced in Rubenthaler-Wiktorsson [15].

Our aim here is to prove that the strong error is much lower when using $X^{n,\epsilon}_{[nt]/n}$, see Sections 2.4 and 8 below.

The main difficulty is to find a suitable coupling between the true increments $(Z_{i/n} - Z_{(i-1)/n})_{i\geq 1}$ and the

https://doi.org/10.1051/ps/2009017 Published online by Cambridge University Press

approximate increments $(\Delta_i^{n,\epsilon})_{i\geq 1}$: clearly, one considers Z, then one erases its jumps smaller than ϵ, but how to build the additional Gaussian variable in such a way that it is a.s. close to the erased jumps? We will use a recent result of Rio [13], which gives some very precise rate of convergence for the standard central limit theorem in Wasserstein distance, in the spirit of Komlós-Major-Tsunády [11].

2.1. Notation

We introduce, for $\epsilon \in (0,1)$, $k \in \mathbb{N}$,

$$F_\epsilon(\nu) = \int_{|z|>\epsilon} \nu(\mathrm{d}z), \quad m_k(\nu) = \int_{\mathbb{R}_*} |z|^k \nu(\mathrm{d}z), \tag{2.1}$$
$$m_{k,\epsilon}(\nu) = \int_{|z|\leq\epsilon} |z|^k \nu(\mathrm{d}z), \quad \beta_\epsilon(\nu) = \frac{m_{4,\epsilon}(\nu)}{m_{2,\epsilon}(\nu)}.$$

Observe that we always have $\beta_\epsilon(\nu) \leq \epsilon^2$ and $F_\epsilon(\nu) \leq \epsilon^{-2} m_2(\nu)$.

For $n \in \mathbb{N}$ and $t \geq 0$, we set $\rho_n(t) = [nt]/n$, where $[x]$ is the integer part of x.

2.2. Numerical scheme

Let $n \in \mathbb{N}$ and $\epsilon \in (0,1)$ be fixed. We introduce an i.i.d. sequence $(\Delta_i^{n,\epsilon})_{i\geq 1}$ of random variables, with

$$\Delta_1^{n,\epsilon} = a_{n,\epsilon} + b_{n,\epsilon} G + \sum_{i=1}^{N_{n,\epsilon}} Y_i^\epsilon, \tag{2.2}$$

where $a_{n,\epsilon} = (a - \int_{|z|>\epsilon} z\nu(\mathrm{d}z))/n$, where $b_{n,\epsilon}^2 = (b^2 + m_{2,\epsilon}(\nu))/n$, where G is Gaussian with mean 0 and variance 1, where $N_{n,\epsilon}$ is Poisson distributed with mean $F_\epsilon(\nu)/n$ and where $Y_1^\epsilon, Y_2^\epsilon, \ldots$ are i.i.d. with law $\nu(\mathrm{d}z)\mathbf{1}_{|z|>\epsilon}/F_\epsilon(\nu)$. All these random variables are assumed to be independent. Then we introduce the scheme

$$X_0^{n,\epsilon} = x, \quad X_{(i+1)/n}^{n,\epsilon} = X_{i/n}^{n,\epsilon} + \sigma(X_{i/n}^{n,\epsilon})\Delta_{i+1}^{n,\epsilon} \quad (i \geq 0). \tag{2.3}$$

Remark 2.1.

(i) The cost of simulation of $\Delta_1^{n,\epsilon}$ is of order $1 + \mathbb{E}[N_{n,\epsilon}] = 1 + F_\epsilon(\nu)/n$, whence that of $(X_{\rho_n(t)}^{n,\epsilon})_{t\in[0,T]}$ is of order $Tn(1 + F_\epsilon(\nu)/n) = T(n + F_\epsilon(\nu))$, as in [14].

(ii) $\Delta_{i+1}^{n,\epsilon}$ has the same law as $Z_{(i+1)/n}^\epsilon - Z_{i/n}^\epsilon + U_i^{n,\epsilon}$, where $U_i^{n,\epsilon}$ is Gaussian with same mean and variance as $\int_{i/n}^{(i+1)/n} \int_{|z|\leq\epsilon} z\tilde{N}(\mathrm{d}s, \mathrm{d}z)$ and where $Z_t^\epsilon = at + bB_t + \int_0^t \int_{|z|>\epsilon} z\tilde{N}(\mathrm{d}s, \mathrm{d}z)$.

(iii) The key argument of the paper is to use a suitable coupling between $U_i^{n,\epsilon}$ and $\int_{i/n}^{(i+1)/n} \int_{|z|\leq\epsilon} z\tilde{N}(\mathrm{d}s, \mathrm{d}z)$. As shown in Lemma 5.2, there exists such a coupling satisfying $E[|U_i^{n,\epsilon} - \int_{i/n}^{(i+1)/n} \int_{|z|\leq\epsilon} z\tilde{N}(\mathrm{d}s, \mathrm{d}z)|^2] \leq C\beta_\epsilon(\nu)$. Then we will choose $\Delta_{i+1}^{n,\epsilon} = Z_{(i+1)/n}^\epsilon - Z_{i/n}^\epsilon + U_i^{n,\epsilon}$, which thus satisfies $E[|\Delta_{i+1}^{n,\epsilon} - (Z_{(i+1)/n} - Z_{i/n})|^2] \leq C\beta_\epsilon(\nu)$.

2.3. Main result

We may now state our main result.

https://doi.org/10.1051/ps/2009017 Published online by Cambridge University Press

Theorem 2.2. *Assume that $\sigma : \mathbb{R} \mapsto \mathbb{R}$ is bounded and Lipschitz continuous. Let $\epsilon \in (0,1)$ and $n \in \mathbb{N}$. There is a coupling between a solution $(X_t)_{t\geq 0}$ to (1.2) and an approximated solution $(X^{n,\epsilon}_{\rho_n(t)})_{t\geq 0}$ as in Subsection 2.2 such that for all $T > 0$,*

$$\mathbb{E}\left[\sup_{t\in[0,T]} |X_{\rho_n(t)} - X^{n,\epsilon}_{\rho_n(t)}|^2\right] \leq C_T\left(n^{-1} + n\beta_\epsilon(\nu)\right),$$

where the constant C_T depends only on $T, \sigma, a, b, m_2(\nu)$.

The first bound n^{-1} is due to the time discretization (Euler scheme) and the second bound $n\beta_\epsilon(\nu)$ is due to the approximation of the increments of the Lévy process. As noted by Jacod [6], the first bound may be improved if there is no Brownian motion $b = 0$ (but we have to work with some weaker norm).

2.4. **Optimization**

Choose $\epsilon = 1/n$. Then recalling that $\beta_\epsilon(\nu) \leq \epsilon^2$, we get

$$\mathbb{E}\left[\sup_{t\in[0,T]} |X_{\rho_n(t)} - X^{n,1/n}_{\rho_n(t)}|^2\right] \leq C_T/n,$$

for a mean cost to simulate $(X^{n,1/n}_{\rho_n(t)})_{t\in[0,T]}$ of order $T(n + F_{1/n}(\nu))$.

- We always have $F_\epsilon(\nu) \leq m_2(\nu)\epsilon^{-2}$, so that the cost is always smaller than CTn^2.
- If $\nu(\mathrm{d}z) \overset{z\to 0}{\simeq} |z|^{-1-\alpha}\mathrm{d}z$ for some $\alpha \in (0,2)$, then $F_\epsilon(\nu) \simeq \epsilon^{-\alpha}$, so that the cost is of order $T(n + n^\alpha)$.

When neglecting the small jumps, one gets, for a mean cost of order $T(n + F_\epsilon(\nu))$,

$$\mathbb{E}\left[\sup_{t\in[0,T]} |X_{\rho_n(t)} - \hat{X}^{n,\epsilon}_{\rho_n(t)}|^2\right] \leq C_T(1/n + m_2(\epsilon)),$$

see [14]. In the case where $\nu(\mathrm{d}z) \overset{z\to 0}{\simeq} |z|^{-1-\alpha}\mathrm{d}z$ for some $\alpha \in (0,2)$, we have $m_2(\epsilon) \simeq \epsilon^{2-\alpha}$ and $F_\epsilon(\nu) \simeq \epsilon^{-\alpha}$. Thus to get an mean squared error of order $1/n$, one has to choose $\epsilon = n^{-1/(2-\alpha)}$, which yields a cost of order $T(n + n^{\alpha/(2-\alpha)})$. This is huge when α is close to 2.

2.5. **Discussion**

The computational cost to get a given precision does not explode when the Lévy measure becomes very singular near 0. The more ν is singular at 0, the more there are jumps greater than ϵ, which costs many simulations. But the more it is singular, the more the jumps smaller than ϵ are well-approximated by Gaussian random variables. These two phenomena are in competition and we prove that the second one compensates (partly) the first one.

Our result involves a suitable coupling between the solution $(X_t)_{t\geq 0}$ and its approximation $(X^{n,\epsilon}_t)_{t\geq 0}$. This might seem uninteresting in practise, since by assumption, $(X_t)_{t\geq 0}$ is completely unknown. But this is always the case when dealing with *strong* errors. In our opinion, this is just an artificial way to estimate the rate of convergence of the paths *in law*, using a Wasserstein type distance. For example, our result allows us to estimate the error when approximating $\mathbb{E}[F((X_{\rho_n(t)})_{t\in[0,T]})]$ by $\mathbb{E}[F((X^{n,\epsilon}_{\rho_n(t)})_{t\in[0,T]})]$, for any Lipschitz functional F.

Recall that Theorem 2.2 is extended in Section 7 to the case where σ is locally Lipschitz with at most linear growth and where $m_2(\nu)$ might be infinite.

Let us finally mention that the simulation algorithm can easily be adapted to the case of dimension $d \geq 2$. We believe that the result still holds. However, the result of Rio [13] is not known in the multidimensional setting. We could use instead the results of Einmahl [3] or Zaitsev [19]. This would be much more technical.

https://doi.org/10.1051/ps/2009017 Published online by Cambridge University Press

3. Brownian approximation

It is classical in applied sciences to approximate discontinuous phenomena by continuous ones. Here, we examine how far the solution to (1.2) is from the solution of a continuous Brownian SDE. Let us mention some more complicated models where such a problem occurs.

(a) The Boltzmann equation is a P.D.E. that can be related to a Poisson-driven S.D.E. (see Tanaka [16]), while the Landau equation can be related to a Brownian S.D.E. (see Guérin [4]). In the grazing collision limit, the Boltzmann equation is known to converge to the Landau equation (see Villani [17]). However, no convergence rate is known.

(b) In [18], Walsh approximates a Poisson-driven stochastic heat equation by a white-noise driven S.P.D.E. He proves some convergence results, without rate.

Here, we consider only the simple case of one-dimensional Lévy-driven S.D.E.s, but we hope that the main ideas of our proof might apply to more complicated models.

Consider the Lévy process introduced in (1.1), consider $x \in \mathbb{R}$, $\sigma : \mathbb{R} \mapsto \mathbb{R}$ Lipschitz continuous, and the unique solution $(X_t)_{t\geq 0}$ to (1.2). Recall (2.1), consider a Brownian motion $(W_t)_{t\geq 0}$ and set

$$\tilde{Z}_t = at + \sqrt{b^2 + m_2(\nu)}W_t, \tag{3.1}$$

which has the same mean and variance as Z_t. Let $(\tilde{X}_t)_{t\geq 0}$ be the unique solution to

$$\tilde{X}_t = x + \int_0^t \sigma(\tilde{X}_{s-})\mathrm{d}\tilde{Z}_s. \tag{3.2}$$

Theorem 3.1. *Assume that σ is Lipschitz continuous and bounded. Then it is possible to couple the solutions $(X_t)_{t\geq 0}$ to (1.2) and $(\tilde{X}_t)_{t\geq 0}$ to (3.2) in such a way that for all $p \geq 4$, all $T > 0$, all $n \geq 1$,*

$$\mathbb{E}\left[\sup_{t\in[0,T]} |X_t - \tilde{X}_t|^2\right] \leq C_{T,p}\left(n^{2/p-1} + m_p(\nu)^{2/p} + nm_4(\nu)\right),$$

where $C_{T,p}$ depends only on $p, T, \sigma, a, b, m_2(\nu)$.

If we only know that $m_4(\nu) < \infty$, we choose $p = 4$ and $n = [m_4(\nu)^{-2/3}] + 1$ to get $\mathbb{E}\left[\sup_{t\in[0,T]} |X_t - \tilde{X}_t|^2\right] \leq C_T(m_4(\nu)^{1/3} + m_4(\nu))$.

Consider a sequence of Lévy processes $(Z^\epsilon_t)_{t\geq 0}$ with drift a, diffusion coefficient b and Lévy measure ν_ϵ, such that $z^2\nu_\epsilon(\mathrm{d}z)$ tends weakly to the Dirac mass $\delta_0(\mathrm{d}z)$. Then $\lim_{\epsilon\to 0} m_2(\nu_\epsilon) = 1$, while in almost all cases, $\lim_{\epsilon\to 0} m_p(\nu_\epsilon) = 0$ for some (or all) $p > 2$.

Consider the solution to $X^\epsilon_t = x + \int_0^t \sigma(X^\epsilon_{s-})\mathrm{d}Z^\epsilon_s$. Then it is well-known and easy to show that $(X^\epsilon_t)_{t\geq 0}$ tends in law to the solution of a Brownian S.D.E. Theorem 3.1 provides a rate of convergence (for some Wasserstein distance). To our knowledge, it is the first result in that direction. For example, we will immediately deduce the following corollary.

Corollary 3.2. *Assume that σ is Lipschitz continuous and bounded. Assume that $\nu(\{|z| > \epsilon\}) = 0$ for some $\epsilon \in (0, 1]$. Then it is possible to couple the solutions $(X_t)_{t\geq 0}$ to (1.2) and $(\tilde{X}_t)_{t\geq 0}$ to (3.2) in such a way that for all $\eta \in (0, 1)$, all $T > 0$,*

$$\mathbb{E}\left[\sup_{t\in[0,T]} |X_t - \tilde{X}_t|^2\right] \leq C_{T,\eta}\epsilon^{1-\eta}$$

where $C_{T,\eta}$ depends only on $\eta, T, \sigma, a, b, m_2(\nu)$.

https://doi.org/10.1051/ps/2009017 Published online by Cambridge University Press

4. Coupling results

Consider two laws P, Q on $\mathbb{R}$ with finite variance. The Wasserstein distance $\mathcal{W}_2$ is defined by

$$\mathcal{W}_2^2(P,Q) = \inf\left\{\mathbb{E}\left[|X-Y|^2\right],\ \mathcal{L}(X)=P, \mathcal{L}(Y)=Q\right\}.$$

With an abuse of notation, we also write $\mathcal{W}_2(X,Y) = \mathcal{W}_2(X,Q) = \mathcal{W}_2(P,Q)$ if $\mathcal{L}(X)=P$ and $\mathcal{L}(Y)=Q$. We recall the following result of Rio [13], Theorem 4.1.

Theorem 4.1. *There is an universal constant C such that for any sequence of i.i.d. random variables $(Y_i)_{i\geq 1}$ with mean 0 and variance θ^2, for any $n \geq 1$,*

$$\mathcal{W}_2^2\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n Y_i, \mathcal{N}(0,\theta^2)\right) \leq C\frac{\mathbb{E}[Y_1^4]}{n\theta^2}.$$

Here $\mathcal{N}(0,\theta^2)$ is the Gaussian distribution with mean 0 and variance θ^2. Recall now (2.1).

Corollary 4.2. *Consider a pure jump centered Lévy process $(Y_t)_{t\geq 0}$ with Lévy measure μ. In other words $Y_t = \int_0^t \int_{\mathbb{R}_*} z\tilde{M}(\mathrm{d}s,\mathrm{d}z)$, where $\tilde{M}$ is a compensated Poisson measure with intensity $\mathrm{d}s\mu(\mathrm{d}z)$. There is an universal constant C such that*

$$\forall\, t\geq 0, \quad \mathcal{W}_2^2\left(Y_t, \mathcal{N}(0, tm_2(\mu))\right) \leq C\frac{m_4(\mu)}{m_2(\mu)}.$$

Proof. Let $t>0$. For $n\geq 1$, $i\geq 1$, write $Y_i^n = n^{1/2}\int_{(i-1)t/n}^{it/n}\int_{\mathbb{R}_*} z\tilde{M}(\mathrm{d}s,\mathrm{d}z)$, whence $Y_t = n^{-1/2}\sum_{i=1}^n Y_i^n$. The Y_i^n are i.i.d., centered, $\mathbb{E}[(Y_1^n)^2] = tm_2(\mu)$, and

$$\begin{aligned}\mathbb{E}[(Y_1^n)^4] &= n^2\mathbb{E}\left[\left(\int_0^{t/n}\int_{\mathbb{R}_*} z^2 M(\mathrm{d}s,\mathrm{d}z)\right)^2\right]\\ &= n^2\mathbb{E}\left[\left(\int_0^{t/n}\int_{\mathbb{R}_*} z^2\tilde{M}(\mathrm{d}s,\mathrm{d}z) + (t/n)m_2(\mu)\right)^2\right]\\ &= n^2\left[tm_4(\mu)/n + (tm_2(\mu)/n)^2\right] = ntm_4(\mu) + t^2m_2(\mu).\end{aligned}$$

Using Theorem 4.1, we get

$$\mathcal{W}_2^2\left(Y_t, \mathcal{N}(0,tm_2(\mu))\right) \leq C\frac{ntm_4(\mu)+t^2m_2(\mu)}{ntm_2(\mu)} \overset{n\to\infty}{\longrightarrow} C\frac{m_4(\mu)}{m_2(\mu)},$$

which concludes the proof. □

This result is quite surprising at first glance: since the variances of the involved variables are $tm_2(\mu)$, it would be natural to get a bound that decreases to 0 as t decreases to 0 (and that explodes for large t). Of course, we deduce the bound $\mathcal{W}_2^2(Y_t,\mathcal{N}(0,tm_2(\mu))) \leq C\min(m_4(\mu)/m_2(\mu), tm_2(\mu))$, but this is now optimal, as shown in the following example.

Example. Consider, for $\epsilon>0$, $\mu_\epsilon = (2\epsilon^2)^{-1}(\delta_\epsilon+\delta_{-\epsilon})$ and the corresponding pure jump (centered) Lévy process $(Y_t^\epsilon)_{t\geq 0}$. It takes its values in $\epsilon\mathbb{Z}$. Observe that $m_2(\mu_\epsilon)=1$ and $m_4(\mu_\epsilon)=\epsilon^2$. There is $c>0$ such that for all $t\geq 0$, all $\epsilon>0$, $\mathcal{W}_2^2(Y_t^\epsilon,\mathcal{N}(0,t)) \geq c\min(\epsilon^2,t) = c\min(m_4(\mu_\epsilon)/m_2(\mu_\epsilon), tm_2(\mu_\epsilon))$. Indeed,

• if $t\leq\epsilon^2$, then $\mathbb{P}(Y_t^\epsilon=0) \geq \mathrm{e}^{-t\mu_\epsilon(\mathbb{R})} = \mathrm{e}^{-t/\epsilon^2} \geq 1/\mathrm{e}$, from which the lower-bound $\mathcal{W}_2^2(Y_t^\epsilon,\mathcal{N}(0,t)) \geq ct = c\min(t,\epsilon^2)$ is easily deduced;

https://doi.org/10.1051/ps/2009017 Published online by Cambridge University Press

• if $t \geq \epsilon^2$, use that $\mathcal{W}_2^2(Y_t^\epsilon, \mathcal{N}(0,t)) \geq \mathbb{E}[\min_{n\in\mathbb{Z}} |t^{1/2}G - n\epsilon|^2] = t\mathbb{E}[\min_{n\in\mathbb{Z}} |G - n\epsilon t^{-1/2}|^2]$, where G is Gaussian with mean 0 and variance 1. Tedious computations show that there is $c > 0$ such that for any $a \in (0,1]$, $\mathbb{E}[\min_{n\in\mathbb{Z}} |G - na|^2] \geq (a/4)^2\mathbb{P}(G \in \cup_{n\in\mathbb{Z}}[(n+1/4)a, (n+3/4)a]) \geq ca^2$. Hence $\mathcal{W}_2^2(Y_t^\epsilon, \mathcal{N}(0,t)) \geq ct(\epsilon t^{-1/2})^2 = c\epsilon^2 = c\min(t, \epsilon^2)$.

5. Proof of Theorem 2.2

We recall elementary results about the Euler scheme for (1.2) in Section 5.1. We introduce our coupling in Section 5.2, which allows us to compare our scheme with the Euler scheme in Section 5.3. We conclude in Section 5.4. We assume in the whole section that σ is bounded and Lipschitz continuous.

5.1. Euler scheme

We introduce the Euler scheme with step $1/n$ associated to (1.2). Let

$$\Delta_i^n = Z_{i/n} - Z_{(i-1)/n} \quad (i \geq 1), \tag{5.1}$$

$$X_0^n = x, \quad X_{(i+1)/n}^n = X_{i/n}^n + \sigma\left(X_{i/n}^n\right)\Delta_{i+1}^n \quad (i \geq 0). \tag{5.2}$$

The following result is classical.

Proposition 5.1. *Consider a Lévy process $(Z_t)_{t\geq 0}$ as in (1.1). For $(X_t)_{t\geq 0}$ the solution to (1.2) and for $(X_{i/n}^n)_{i\geq 0}$ defined in (5.1)–(5.2),*

$$\mathbb{E}\left[\sup_{t\in[0,T]} |X_{\rho_n(t)} - X_{\rho_n(t)}^n|^2\right] \leq C_T/n,$$

where C_T depends only on $T, a, b, m_2(\nu)$ and σ.

We give a proof for the sake of completeness.

Proof. Using the Doob and Cauchy-Schwarz inequalities, we get, for $0 \leq s \leq t \leq T$,

$$\begin{aligned}\mathbb{E}\left[\sup_{u\in[s,t]} |X_u - X_s|^2\right] &\leq C\mathbb{E}\left[\left(a\int_s^t |\sigma(X_u)|\mathrm{d}u\right)^2 + \sup_{u\in[s,t]}\left(b\int_s^u \sigma(X_v)\mathrm{d}B_v\right)^2\right. \\ &\qquad \left. + \sup_{u\in[s,t]}\left(\int_s^u\int_{\mathbb{R}_*} \sigma(X_{v-})z\tilde{N}(\mathrm{d}s,\mathrm{d}v)\right)^2\right] \\ &\leq C_T\int_s^t (a^2 + b^2 + m_2(\nu))||\sigma||_\infty^2\mathrm{d}v \leq C_T(t-s).\end{aligned} \tag{5.3}$$

Observe now that $X_{\rho_n(t)}^n = x + \int_0^{\rho_n(t)} \sigma(X_{\rho_n(s)-}^n)\mathrm{d}Z_s$. Setting $A_t^n = \sup_{[0,t]} |X_{\rho_n(s)} - X_{\rho_n(s)}^n|^2$, we thus get $A_t^n = \sup_{[0,t]} |\int_0^{\rho_n(s)} (\sigma(X_{u-}) - \sigma(X_{\rho_n(u)-}^n))\mathrm{d}Z_u|^2$. Using the same arguments as in (5.3), then the Lipschitz property of σ and (5.3), we get

$$\begin{aligned}\mathbb{E}[A_t^n] &\leq C_T\int_0^{\rho_n(t)} (a^2 + b^2 + m_2(\nu))\mathbb{E}[(\sigma(X_s) - \sigma(X_{\rho_n(s)}^n))^2]\mathrm{d}s \\ &\leq C_T\int_0^t \mathbb{E}[(X_s - X_{\rho_n(s)})^2 + (X_{\rho_n(s)} - X_{\rho_n(s)}^n)^2]\mathrm{d}s \\ &\leq C_T\int_0^t (|s - \rho_n(s)| + \mathbb{E}[A_s^n])\,\mathrm{d}s.\end{aligned}$$

https://doi.org/10.1051/ps/2009017 Published online by Cambridge University Press

We conclude using that $|s-\rho_n(s)| \le 1/n$ and the Gronwall lemma. □

5.2. Coupling

We now introduce a suitable coupling between the Euler scheme (see Sect. 5.1) and our numerical scheme (see Sect. 2.2). Recall (2.1).

Lemma 5.2. *Let $n \in \mathbb{N}$ and $\epsilon > 0$. There exist two coupled families of i.i.d. random variables $(\Delta_i^n)_{i\ge1}$ and $(\Delta_i^{n,\epsilon})_{i\ge1}$, distributed respectively as in (5.1) and (2.2) in such a way that for each $i \ge 1$,*

$$\mathbb{E}[(\Delta_i^n - \Delta_i^{n,\epsilon})^2] \le C\beta_\epsilon(\nu),$$

where C is an universal constant. Furthermore, for all $\epsilon > 0$, all $n \in \mathbb{N}$, all $i \ge 1$,

$$\mathbb{E}[\Delta_i^n] = \mathbb{E}[\Delta_i^{n,\epsilon}] = \frac{a}{n}, \quad \mathbb{V}\mathrm{ar}[\Delta_i^n] = \mathbb{V}\mathrm{ar}[\Delta_i^{n,\epsilon}] = \frac{b^2 + m_2(\nu)}{n}\cdot$$

Proof. It of course suffices to build $(\Delta_1^n, \Delta_1^{n,\epsilon})$ and then to take independent copies. Consider a Poisson measure $N(\mathrm{d}s,\mathrm{d}z)$ with intensity measure $\mathrm{d}s\nu(\mathrm{d}z)\mathbf{1}_{\{|z|\le\epsilon\}}$ on $[0,\infty)\times\{|z|\le\epsilon\}$. Observe that $\int_0^t\int_{|z|\le\epsilon} z\tilde N(\mathrm{d}s,\mathrm{d}z)$ is a centered pure jump Lévy process with Lévy measure $\nu_\epsilon(\mathrm{d}z) = \mathbf{1}_{|z|\le\epsilon}\nu(\mathrm{d}z)$. Then we use Corollary 4.2 and enlarge the underlying probability space if necessary: there is a Gaussian random variable $G_1^{n,\epsilon}$ with mean 0 and variance $m_2(\nu_\epsilon)/n = m_{2,\epsilon}(\nu)/n$ such that $\mathbb{E}\left[|\int_0^{1/n}\int_{|z|\le\epsilon} z\tilde N(\mathrm{d}s,\mathrm{d}z) - G_1^{n,\epsilon}|^2\right] \le Cm_4(\nu_\epsilon)/m_2(\nu_\epsilon) = C\beta_\epsilon(\nu)$.

We consider a Brownian motion $(B_t)_{t\ge0}$ and a Poisson measure N with intensity measure $\mathrm{d}s\nu(\mathrm{d}z)\mathbf{1}_{\{|z|>\epsilon\}}$ on $[0,\infty)\times\{|z|>\epsilon\}$, independent of the couple $(G_1^{n,\epsilon}, \int_0^{1/n}\int_{|z|\le\epsilon} z\tilde N(\mathrm{d}s,\mathrm{d}z))$ and we set

- $\Delta_1^n := a/n + bB_{1/n} + \int_0^{1/n}\int_{|z|\le\epsilon} z\tilde N(\mathrm{d}s,\mathrm{d}z) + \int_0^{1/n}\int_{|z|>\epsilon} z\tilde N(\mathrm{d}s,\mathrm{d}z)$;
- $\Delta_1^{n,\epsilon} := a/n + bB_{1/n} + G_1^{n,\epsilon} + \int_0^{1/n}\int_{|z|>\epsilon} z\tilde N(\mathrm{d}s,\mathrm{d}z)$.

Then Δ_1^n has obviously the same law as $Z_{1/n} - Z_0$ (see (1.1) and (5.1)), while $\Delta_1^{n,\epsilon}$ has also the desired law (see (2.2)). Indeed, $bB_{1/n} + G_1^{n,\epsilon}$ has a centered Gaussian law with variance $b^2/n + m_{2,\epsilon}(\nu)/n = b_{n,\epsilon}^2$ and $a/n + \int_0^{1/n}\int_{|z|>\epsilon} z\tilde N(\mathrm{d}s,\mathrm{d}z) = a_{n,\epsilon} + \int_0^{1/n}\int_{|z|>\epsilon} zN(\mathrm{d}s,\mathrm{d}z)$. This last integral can be represented as in (2.2). Finally $\mathbb{E}[(\Delta_1^n - \Delta_1^{n,\epsilon})^2] \le \mathbb{E}\left[|\int_0^{1/n}\int_{|z|\le\epsilon} z\tilde N(\mathrm{d}s,\mathrm{d}z) - G_1^{n,\epsilon}|^2\right] \le C\beta_\epsilon(\nu)$ and the mean and variance estimates are obvious. □

5.3. Estimates

We now compare our scheme with the Euler scheme. To this end, we introduce some notation. First, we consider the sequence $(\Delta_i^n, \Delta_i^{n,\epsilon})_{i\ge1}$ introduced in Lemma 5.2. Then we consider $(X_{i/n}^n)_{i\ge0}$ and $(X_{i/n}^{n,\epsilon})_{i\ge0}$ defined in (5.2) and (2.3). We introduce the filtration $\mathcal{F}_i^{n,\epsilon} = \sigma(\Delta_k^n, \Delta_k^{n,\epsilon}, k \le i)$ and the processes, for $i \ge 0$ (with $V_0^{n,\epsilon} = 0$)

$$Y_i^{n,\epsilon} = X_{i/n}^n - X_{i/n}^{n,\epsilon}, \quad V_i^{n,\epsilon} = \frac{a}{n}\sum_{k=0}^{i-1}[\sigma(X_{k/n}^n) - \sigma(X_{k/n}^{n,\epsilon})], \quad M_i^{n,\epsilon} = Y_i^{n,\epsilon} - V_i^{n,\epsilon}.$$

Lemma 5.3. *There is a constant C, depending only on $\sigma, a, b, m_2(\nu)$ such that for all $N \ge 1$,*

$$\mathbb{E}\left[\sup_{i=0,\dots,N}|Y_i^{n,\epsilon}|^2\right] \le Cn\beta_\epsilon(\nu)(1+C/n)^N(1+N^2/n^2).$$

Proof. We divide the proof into four steps.

https://doi.org/10.1051/ps/2009017 Published online by Cambridge University Press

Step 1. We prove that for all $i \geq 0$, $\mathbb{E}\left[|Y_i^{n,\epsilon}|^2\right] \leq Cn\beta_\epsilon(\nu)(1+C/n)^i$. First,

$$\begin{aligned}\mathbb{E}[|Y_{i+1}^{n,\epsilon}|^2] &= \mathbb{E}[|Y_i^{n,\epsilon}|^2] + \mathbb{E}[(\sigma(X_{i/n}^n)\Delta_{i+1}^n - \sigma(X_{i/n}^{n,\epsilon})\Delta_{i+1}^{n,\epsilon})^2] \\ &\quad + 2\mathbb{E}\left[Y_i^{n,\epsilon}(\sigma(X_{i/n}^n)\Delta_{i+1}^n - \sigma(X_{i/n}^{n,\epsilon})\Delta_{i+1}^{n,\epsilon})\right] = \mathbb{E}[|Y_i^{n,\epsilon}|^2] + I_i^{n,\epsilon} + J_i^{n,\epsilon}.\end{aligned}$$

Now, using Lemma 5.2 and that $(\Delta_{i+1}^n, \Delta_{i+1}^{n,\epsilon})$ is independent of $\mathcal{F}_i^{n,\epsilon}$, we deduce that

$$J_i^{n,\epsilon} = \frac{2a}{n}\mathbb{E}\left[Y_i^{n,\epsilon}(\sigma(X_{i/n}^n) - \sigma(X_{i/n}^{n,\epsilon}))\right] \leq \frac{C}{n}E[|Y_i^{n,\epsilon}|^2],$$

since σ is Lipschitz continuous. Using now the Lipschitz continuity and the boundedness of σ, together with Lemma 5.2 and the independence of $(\Delta_{i+1}^n, \Delta_{i+1}^{n,\epsilon})$ with respect to $\mathcal{F}_i^{n,\epsilon}$, we get

$$I_i^{n,\epsilon} \leq C\mathbb{E}[|Y_i^{n,\epsilon}|^2(\Delta_{i+1}^{n,\epsilon})^2] + C\mathbb{E}[(\Delta_{i+1}^{n,\epsilon} - \Delta_{i+1}^n)^2] \leq \frac{C}{n}E[|Y_i^{n,\epsilon}|^2] + C\beta_\epsilon(\nu).$$

Finally, we get

$$\mathbb{E}[|Y_{i+1}^{n,\epsilon}|^2] \leq (1+C/n)E[|Y_i^{n,\epsilon}|^2] + C\beta_\epsilon(\nu).$$

Since $Y_0^{n,\epsilon} = 0$, this entails that $E[|Y_i^{n,\epsilon}|^2] \leq C\beta_\epsilon(\nu)[1 + (1+C/n) + ... + (1+C/n)^{i-1}] \leq Cn\beta_\epsilon(\nu)(1+C/n)^i$.

Step 2. We check that for $N \geq 1$, $\mathbb{E}[\sup_{0,\ldots,N}|V_i^{n,\epsilon}|^2] \leq Cn\beta_\epsilon(\nu)(1+C/n)^N N^2/n^2$. It suffices to use the Lipschitz property of σ, the Cauchy-Schwarz inequality and then Step 1:

$$\begin{aligned}\mathbb{E}\left[\sup_{1,\ldots,N}|V_i^{n,\epsilon}|^2\right] &\leq C\mathbb{E}\left[\left(\frac{1}{n}\sum_{i=0}^{N-1}|Y_i^{n,\epsilon}|\right)^2\right] \leq C\frac{N}{n^2}\sum_{i=0}^{N-1}\mathbb{E}[|Y_i^{n,\epsilon}|^2] \\ &\leq C\frac{N^2}{n^2}n\beta_\epsilon(\nu)(1+C/n)^N.\end{aligned}$$

Step 3. We now verify that $(M_i^{n,\epsilon})_{i\geq 0}$ is a $(\mathcal{F}_i^{n,\epsilon})_{i\geq 0}$-martingale. We have $M_{i+1}^{n,\epsilon} - M_i^{n,\epsilon} = \sigma(X_{i/n}^n)[\Delta_{i+1}^n - a/n] - \sigma(X_{i/n}^{n,\epsilon})[\Delta_{i+1}^{n,\epsilon} - a/n]$. The step is finished, since the variables $\Delta_{i+1}^n - a/n$ and $\Delta_{i+1}^{n,\epsilon} - a/n$ are centered and independent of $\mathcal{F}_i^{n,\epsilon}$.

Step 4. Using the Doob inequality and then Steps 1 and 2, we get

$$\begin{aligned}\mathbb{E}\left[\sup_{i=0,\ldots,N}|M_i^{n,\epsilon}|^2\right] &\leq C\sup_{i=0,\ldots,N}\mathbb{E}\left[|M_i^{n,\epsilon}|^2\right] \\ &\leq C\sup_{i=0,\ldots,N}\mathbb{E}\left[|Y_i^{n,\epsilon}|^2\right] + C\sup_{i=0,\ldots,N}\mathbb{E}\left[|V_i^{n,\epsilon}|^2\right] \\ &\leq Cn\beta_\epsilon(\nu)(1+C/n)^N(1+N^2/n^2).\end{aligned}$$

But now, since $Y_i^{n,\epsilon} = M_i^{n,\epsilon} + V_i^{n,\epsilon}$,

$$\mathbb{E}\left[\sup_{i=0,\ldots,N}|Y_i^{n,\epsilon}|^2\right] \leq C\mathbb{E}\left[\sup_{i=0,\ldots,N}|M_i^{n,\epsilon}|^2\right] + C\mathbb{E}\left[\sup_{i=0,\ldots,N}|V_i^{n,\epsilon}|^2\right],$$

which allows us to conclude. □

Let us rewrite these estimates in terms of X^n and $X^{n,\epsilon}$.

https://doi.org/10.1051/ps/2009017 Published online by Cambridge University Press

Lemma 5.4. *Consider the sequence $(\Delta_i^n, \Delta_i^{n,\epsilon})_{i\geq 1}$ introduced in Lemma 5.2 and then $(X^n_{i/n})_{i\geq 0}$ and $(X^{n,\epsilon}_{i/n})_{i\geq 0}$ defined in (5.2) and (2.3). For all $T \geq 0$,*

$$\mathbb{E}\left[\sup_{t\in[0,T]} |X^n_{\rho_n(t)} - X^{n,\epsilon}_{\rho_n(t)}|^2\right] \leq C_T n\beta_\epsilon(\nu),$$

where C_T depends only on $T, a, b, m_2(\nu), \sigma$.

Proof. With the previous notation, $\sup_{t\in[0,T]} |X^n_{\rho_n(t)} - X^{n,\epsilon}_{\rho_n(t)}| = \sup_{i=0,\dots,[nT]} |Y_i^{n,\epsilon}|$. Thus using Lemma 5.3, we get the bound $Cn\beta_\epsilon(\nu)(1+C/n)^{[nT]}(1+[nT]^2/n^2) \leq Cn\beta_\epsilon(\nu)\mathrm{e}^{CT}(1+T^2)$, which ends the proof. □

5.4. **Conclusion**

We finally give the

Proof of Theorem 2.2. Fix $n \in \mathbb{N}$ and $\epsilon > 0$. Denote by $Q(\mathrm{d}u, \mathrm{d}v)$ the joint law of $(\Delta_1^n, \Delta_1^{n,\epsilon})$ built in Lemma 5.2 and write $Q(\mathrm{d}u, \mathrm{d}v) = Q_1(\mathrm{d}u)R(u, \mathrm{d}v)$, where $Q_1(\mathrm{d}u)$ is the law of Δ_1^n and where $R(u, \mathrm{d}v)$ is the law of $\Delta_1^{n,\epsilon}$ conditionally to $\Delta_1^n = u$.

Consider a Lévy process $(Z_t)_{t\geq 0}$ as in (1.1) and $(X_t)_{t\geq 0}$ the corresponding solution to (1.2). Set, for $i \geq 0$, $\Delta_i^n = Z_{i/n} - Z_{(i-1)/n}$ and consider the Euler scheme $(X^n_{i/n})_{i\geq 0}$ as in (5.2). For each $i \geq 1$, let $\Delta_i^{n,\epsilon}$ be distributed according to $R(\Delta_i^n, \mathrm{d}v)$, in such a way that $(\Delta_i^{n,\epsilon})_{i\geq 1}$ is an i.i.d. sequence. Finally, let $(X^{n,\epsilon}_{i/n})_{i\geq 0}$ as in (2.3).

By this way, the processes $(X_t)_{t\geq 0}$, $(X^n_{i/n})_{i\geq 0}$ and $(X^{n,\epsilon}_{i/n})_{i\geq 0}$ are coupled in such a way that we may apply Proposition 5.1 and Lemma 5.4. We get

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in[0,T]} |X_{\rho_n(t)} - X^{n,\epsilon}_{\rho_n(t)}|^2\right] &\leq 2\mathbb{E}\left[\sup_{t\in[0,T]} |X_{\rho_n(t)} - X^n_{\rho_n(t)}|^2 + \sup_{t\in[0,T]} |X^n_{\rho_n(t)} - X^{n,\epsilon}_{\rho_n(t)}|^2\right] \\ &\leq C_T[n^{-1} + n\beta_\epsilon(\nu)].\end{aligned}$$

This concludes the proof. □

6. Proofs of Theorem 3.1 and Corollary 3.2

We assume in the whole section that σ is bounded and Lipschtiz continuous. We start with a technical lemma.

Lemma 6.1. *Let $(X_t)_{t\geq 0}$ and $(\tilde{X}_t)_{t\geq 0}$ be solutions to (1.2) and (3.2). Then for $p \geq 2$, for all $t_0 \geq 0$, all $h \in (0, 1]$,*

$$\mathbb{E}\left[\sup_{t\in[t_0,t_0+h]} |X_t - X_{t_0}|^p\right] \leq C_p(h^{p/2} + hm_p(\nu)),$$

$$\mathbb{E}\left[\sup_{t\in[t_0,t_0+h]} |\tilde{X}_t - \tilde{X}_{t_0}|^p\right] \leq C_p h^{p/2},$$

where C_p depends only on $p, \sigma, a, b, m_2(\nu)$.

Proof. It clearly suffices to treat the case of $(X_t)_{t\geq 0}$, because $(\tilde{X}_t)_{t\geq 0}$ solves the same equation (with ν replaced by 0 and b replaced by $b + m_2(\nu)$). Let thus $p \geq 2$.

https://doi.org/10.1051/ps/2009017 Published online by Cambridge University Press

Using the Burkholder-Davies-Gundy inequality and the boundedness of σ, we get

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in[t_0,t_0+h]}|X_t-X_{t_0}|^p\right] &\le C_p\mathbb{E}\left[\left(\int_{t_0}^{t_0+h}|a\sigma(X_s)|\mathrm{d}s\right)^p\right]+C_p\mathbb{E}\left[\left(\int_{t_0}^{t_0+h}b^2\sigma^2(X_s)\mathrm{d}s\right)^{p/2}\right]\\ &\quad+C_p\mathbb{E}\left[\left(\int_{t_0}^{t_0+h}\int_{\mathbb{R}_*}\sigma^2(X_s)z^2N(\mathrm{d}s,\mathrm{d}z)\right)^{p/2}\right]\\ &\le C_ph^p+C_ph^{p/2}+C_p\mathbb{E}\left[\left(\int_{t_0}^{t_0+h}\int_{\mathbb{R}_*}z^2N(\mathrm{d}s,\mathrm{d}z)\right)^{p/2}\right]\\ &\le C_ph^{p/2}+C_p\mathbb{E}[U_h^{p/2}],\end{aligned}$$

where $U_t=\int_0^t\int_{\mathbb{R}_*}z^2N(\mathrm{d}s,\mathrm{d}z)$. It remains to check that for $t\ge 0$, $\mathbb{E}[U_t^{p/2}]\le C_p(t^{p/2}+tm_p(\nu))$. But, with C_p depending on $m_2(\nu)$,

$$\begin{aligned}\mathbb{E}[U_t^{p/2}] &= \int_0^t\mathrm{d}s\int_{\mathbb{R}_*}\nu(\mathrm{d}z)\mathbb{E}[(U_s+z^2)^{p/2}-U_s^{p/2}]\\ &\le C_p\int_0^t\mathrm{d}s\int_{\mathbb{R}_*}\nu(\mathrm{d}z)\mathbb{E}[z^2U_s^{p/2-1}+|z|^p]\le C_p\int_0^t\mathbb{E}[U_s^{p/2-1}]\mathrm{d}s+C_pm_p(\nu)t\\ &\le C_p\int_0^t\mathbb{E}[U_s^{p/2}]\epsilon^{-1}\mathrm{d}s+C_p(\epsilon^{p/2-1}+m_p(\nu))t,\end{aligned}$$

for any $\epsilon>0$. Hence $\mathbb{E}[U_t^{p/2}]\le C_p(\epsilon^{p/2-1}t+m_p(\nu)t)\mathrm{e}^{C_pt/\epsilon}$ by the Gronwall lemma. Choosing $\epsilon=t$, we conclude that $\mathbb{E}[U_t^{p/2}]\le C_p(t^{p/2}+m_p(\nu)t)$. □

Proof of Theorem 3.1. We fix $n\ge 1$, $T>0$ and $p\ge 4$.

Step 1. Using Lemma 4.2 (see also Lem. 5.2) we deduce that we may couple two i.i.d. families $(\Delta_i^n)_{i\ge1}$ and $(\tilde\Delta_i^n)_{i\ge1}$, in such a way that:

- $(\Delta_i^n)_{i\ge1}$ has the same law as the increments $(Z_{i/n}-Z_{(i-1)/n})_{i\ge1}$ of the Lévy process (1.1);
- $(\tilde\Delta_i^n)_{i\ge1}$ has the same law as the increments $(\tilde Z_{i/n}-\tilde Z_{(i-1)/n})_{i\ge1}$ of the Lévy process (3.1);
- for all $i\ge1$, $\mathbb{E}[(\Delta_i^n-\tilde\Delta_i^n)^2]\le Cm_4(\nu)$ (we allow constants to depend on $m_2(\nu)$).

Step 2. We then set $X_0^n=\tilde X_0^n=x$ and for $i\ge1$, $X_{i/n}^n=X_{(i-1)/n}^n+\sigma(X_{(i-1)/n}^n)\Delta_i^n$ and $\tilde X_{i/n}^n=\tilde X_{(i-1)/n}^n+\sigma(\tilde X_{(i-1)/n}^n)\tilde\Delta_i^n$. Using exactly the same arguments as in Lemmas 5.3 and 5.4, we deduce that $\mathbb{E}\left[\sup_{t\in[0,T]}|X_{\rho_n(t)}^n-\tilde X_{\rho_n(t)}^n|^2\right]\le C_Tnm_4(\nu)$, where C_T depends only on $T,\sigma,a,b,m_2(\nu)$.

Step 3. But $(X_{\rho_n(t)}^n)_{t\ge0}$ is the Euler discretization of (1.2), while $(\tilde X_{\rho_n(t)}^n)_{t\ge0}$ is the Euler discretization of (3.2). Hence using Step 2, Proposition 5.1 and a suitable coupling as in the final proof of Theorem 2.2, $\mathbb{E}\left[\sup_{t\in[0,T]}|X_{\rho_n(t)}-\tilde X_{\rho_n(t)}|^2\right]\le C_T(1/n+nm_4(\nu))$.

Step 4. We now prove that $\mathbb{E}\left[\sup_{t\in[0,T]}|X_t-X_{\rho_n(t)}|^2\right]\le C_{T,p}(n^{2/p-1}+m_p(\nu)^{2/p})$.

https://doi.org/10.1051/ps/2009017 Published online by Cambridge University Press

We set $\Gamma_i = \sup_{t\in[i/n,(i+1)/n]} |X_t - X_{\rho_n(t)}| = \sup_{t\in[i/n,(i+1)/n]} |X_t - X_{i/n}|$. By Lemma 6.1, $\mathbb{E}[\Gamma_i^p] \le C_p[(1/n)^{p/2} + m_p(\nu)/n]$. Thus, since $p \ge 2$,

$$\mathbb{E}\left[\sup_{t\in[0,T]} |X_t - X_{\rho_n(t)}|^2\right] \le \mathbb{E}\left[\sup_{1,\dots,[nT]} \Gamma_i^2\right] \le \mathbb{E}\left[\sup_{1,\dots,[nT]} \Gamma_i^p\right]^{2/p} \le \mathbb{E}\left[\sum_{i=1}^{[nT]} \Gamma_i^p\right]^{2/p}$$
$$\le C_{T,p} n^{2/p} \left[(1/n)^{p/2} + m_p(\nu)/n\right]^{2/p},$$

which ends the step.

Step 5. Exactly as in Step 4, we get $\mathbb{E}\left[\sup_{t\in[0,T]} |\tilde{X}_t - \tilde{X}_{\rho_n(t)}|^2\right] \le C_{T,p} n^{2/p-1}$.

Step 6. Using Steps 3, 4 and 5, we deduce that with a suitable coupling, we have $\mathbb{E}[\sup_{t\in[0,T]} |X_t - \tilde{X}_t|^2] \le C_{T,p}(n^{2/p-1} + m_p(\nu)^{2/p} + n^{-1} + n m_4(\nu))$. □

We conclude this section with the

Proof of Corollary 3.2. Since $\nu(\{|z| > \epsilon\}) = 0$, we deduce that $m_p(\nu) \le m_2(\nu)\epsilon^{p-2}$, for any $p \ge 2$. Applying Theorem 3.1 and choosing $n = [\epsilon^{-p/(p-1)}]$, we get the bound

$$C_{T,p}\left(\epsilon^{(1-2/p)(p/(p-1))} + \epsilon^{(p-2)(2/p)} + \epsilon^{2-p/(p-1)}\right) \le C_{T,p}(\epsilon^{1-1/(p-1)} + \epsilon^{2-4/p}).$$

Hence for $\eta \in (0,1)$, it is possible to get the bound $C_{T,\eta}\epsilon^{1-\eta}$, choosing p large enough. □

7. Generalization

The goal of this section is to generalize Theorem 2.2 by using a standard localization argument. In the important case where the driving Lévy process behaves like a stable process, it does not hold that $m_2(\nu) < \infty$. We want here to treat the most general case, that is $\int_{\mathbb{R}_*} \min(z^2, 1)\nu(\mathrm{d}z) < \infty$. We will also show how to deal with the case where σ is only locally Lipschitz continuous with at most linear growth. Of course, we can not work in L^2 any more, so that we will only prove a tightness result. This will also determine the rate of convergence in law of the paths, but in a weaker sense.

We thus consider a general one-dimensional Lévy process with Lévy measure ν:

$$Z_t = at + bB_t + \int_0^t \int_{[-1,1]\setminus\{0\}} z\tilde{N}(\mathrm{d}s, \mathrm{d}z) + \int_0^t \int_{\mathbb{R}\setminus[-1,1]} zN(\mathrm{d}s, \mathrm{d}z), \tag{7.1}$$

where $(B_t)_{t\ge0}$ is a standard Brownian motion, independent of a Poisson measure $N(\mathrm{d}s, \mathrm{d}z)$ on $[0,\infty)\times\mathbb{R}_*$ with intensity measure $\mathrm{d}s\nu(\mathrm{d}z)$ and where $\tilde{N}$ is its compensated Poisson measure. Assuming that σ is locally Lipschitz continuous with at most linear growth, it is well-known that (1.2) has a unique càdlàg adapted strong solution $(X_t)_{t\ge0}$.

For $n \ge 1$ and $\epsilon \in (0,1)$, we introduce, as in (2.2), a sequence of i.i.d. random variables $(\Delta_i^{n,\epsilon})_{i\ge1}$, with

$$\Delta_1^{n,\epsilon} = a_{n,\epsilon} + b_{n,\epsilon}G + \sum_{i=1}^{N_{n,\epsilon}} Y_i^\epsilon, \tag{7.2}$$

where $a_{n,\epsilon} = (a - \int_{\epsilon<|z|\le1} z\nu(\mathrm{d}z))/n$, where $b_{n,\epsilon}^2 = (b^2 + m_{2,\epsilon}(\nu))/n$, where G is Gaussian with mean 0 and variance 1, where $N_{n,\epsilon}$ is Poisson distributed with mean $F_\epsilon(\nu)/n$ and where $Y_1^\epsilon, Y_2^\epsilon, \dots$ are i.i.d. with law $\nu(\mathrm{d}z)\mathbf{1}_{|z|>\epsilon}/F_\epsilon(\nu)$. The only difference with (2.2) is the value of $a_{n,\epsilon}$. Then we introduce the approximated solution $(X^{n,\epsilon}_{\rho_n(t)})_{t\ge0}$ as in (2.3).

https://doi.org/10.1051/ps/2009017 Published online by Cambridge University Press

Theorem 7.1. *Assume that $\sigma : \mathbb{R} \mapsto \mathbb{R}$ is locally Lipschitz continuous with at most linear growth and that $\int_{\mathbb{R}_*} \min(z^2, 1)\nu(\mathrm{d}z) < \infty$. There exists, for each $n \in \mathbb{N}$ and each $\epsilon \in (0,1)$, a coupling between a solution $(X_t)_{t\geq 0}$ to (1.2) and an approximated solution $(X^{n,\epsilon}_{\rho_n(t)})_{t\geq 0}$, such that for all $T > 0$, recall (2.1),*

$$\lim_{A\to\infty} \sup_{n\in\mathbb{N},\epsilon\in(0,1)} \mathbb{P}\left((n^{-1} + n\beta_\epsilon(\nu))^{-1/2} \sup_{[0,T]} |X_{\rho_n(t)} - X^{n,\epsilon}_{\rho_n(t)}| \geq A \right) = 0. \tag{7.3}$$

Proof. The proof relies on a standard localization argument, and we only give the main steps. We fix the terminal time $T > 0$.

Step 1. First we define, for $K > 1$, the Lévy process

$$Z^K_t = at + bB_t + \int_0^t \int_{|z|\leq 1} z\tilde{N}(\mathrm{d}s, \mathrm{d}z) + \int_0^t \int_{1\leq|z|\leq K} zN(\mathrm{d}s, \mathrm{d}z),$$

whose Lévy measure $\nu_K(\mathrm{d}z) = \mathbf{1}_{\{|z|\leq K\}}\nu(\mathrm{d}z)$ satisfies $m_2(\nu_K) < \infty$. We also introduce the sequence of i.i.d. random variables $(\Delta^{n,\epsilon,K}_i)_{i\geq 1}$ as previously, using ν_K instead of ν. We denote by $(X^K_t)_{t\geq 0}$ the solution to (1.2) with Z^K instead of Z, and by $(X^{n,\epsilon,K}_{\rho_n(t)})_{t\geq 0}$ the approximated solution with $(\Delta^{n,\epsilon,K}_i)_{i\geq 1}$ instead of $(\Delta^{n,\epsilon}_i)_{i\geq 1}$. One easily checks that with a suitable coupling between $(\Delta^{n,\epsilon,K}_i)_{i\geq 1}$ and $(\Delta^{n,\epsilon}_i)_{i\geq 1}$,

$$\mathbb{P}\left((X^K_t)_{t\in[0,T]} = (X_t)_{t\in[0,T]}, (X^{n,\epsilon,K}_{\rho_n(t)})_{t\in[0,T]} = (X^{n,\epsilon}_{\rho_n(t)})_{t\in[0,T]} \right) \geq \mathbb{P}(N([0,T]\times\{|z|\geq K\}) = 0) = \mathrm{e}^{-T\nu(\{|z|\geq K\})}.$$

Step 2. Since σ has at most linear growth and since $m_2(\nu_K) < \infty$, one easily checks that

$$\sup_{n\in\mathbb{N},\epsilon\in(0,1)} \mathbb{E}\left[\sup_{[0,T]} \left(|X^K_t|^2 + |X^{n,\epsilon,K}_{\rho_n(t)}|^2 \right) \right] \leq C_{K,T},$$

where the constant $C_{K,T}$ possibly explodes when K increases to infinity.

Step 3. For $L > 0$, we introduce $\sigma_L(x) = \sigma((x \vee -L) \wedge L)$ and we denote by $(X^{K,L}_t)_{t\geq 0}$ and by $(X^{n,\epsilon,K,L}_{\rho_n(t)})_{t\geq 0}$ the solutions to (1.2) and (2.3) with Z^K instead of Z, $(\Delta^{n,\epsilon,K}_i)_{i\geq 1}$ instead of $(\Delta^{n,\epsilon}_i)_{i\geq 1}$, and with σ_L instead of σ. Using Step 2 and a uniqueness argument, it is classically deduced that

$$\mathbb{P}\left((X^{K,L}_t)_{t\in[0,T]} = (X^K_t)_{t\in[0,T]}, (X^{n,\epsilon,K,L}_{\rho_n(t)})_{t\in[0,T]} = (X^{n,\epsilon,K}_{\rho_n(t)})_{t\in[0,T]} \right) \geq 1 - \frac{C_{K,T}}{L^2}.$$

Step 4. Since $m_2(\nu_K) < \infty$ and since σ_L is bounded and Lipschitz continuous, we may apply Theorem 2.2: for each K, L, there exists a constant $C_{K,L,T}$ such that for all $n \in \mathbb{N}$, all $\epsilon \in (0,1)$, with a suitable coupling,

$$\mathbb{E}\left[\sup_{[0,T]} |X^{K,L}_{\rho_n(t)} - X^{n,\epsilon,K,L}_{\rho_n(t)}|^2 \right] \leq C_{K,L,T}(n^{-1} + \beta_\epsilon(\nu_L)).$$

Since $L \geq 1 > \epsilon$, we obviously have $\beta_\epsilon(\nu_L) = \beta_\epsilon(\nu)$.

Step 5. Using Steps 1, 3 and 4, we get, for all $A > 0$, $K \geq 1$, $L > 0$, $n \in \mathbb{N}, \epsilon \in (0,1)$,

$$\mathbb{P}\left[(n^{-1} + \beta_\epsilon(\nu_L))^{-1/2} \sup_{[0,T]} |X_{\rho_n(t)} - X^{n,\epsilon}_{\rho_n(t)}| \geq A \right] \leq 1 - \mathrm{e}^{-T\nu(\{|z|\geq K\})} + \frac{C_{K,T}}{L^2} + \frac{C_{K,L,T}}{A^2}.$$

https://doi.org/10.1051/ps/2009017 Published online by Cambridge University Press

Consequently, for all $K \geq 1$, all $L > 0$,

$$\limsup_{A\to\infty} \sup_{n\in\mathbb{N},\epsilon\in(0,1)} \mathbb{P}\left[(n^{-1}+\beta_\epsilon(\nu_L))^{-1/2} \sup_{[0,T]} |X_{\rho_n(t)} - X^{n,\epsilon}_{\rho_n(t)}| \geq A\right] \leq 1 - \mathrm{e}^{-T\nu(\{|z|\geq K\})} + \frac{C_{K,T}}{L^2}.$$

Taking the limit as $L \to \infty$ and then the limit as $K \to \infty$, the conclusion follows. □

8. Numerical illustration

To illustrate our result, we consider the case of a stable driving Lévy process, of which the increments may be simulated exactly, see Chambers-Mallows-Stuck [2]. We consider the Lévy process (7.1) with $a = b = 0$ and $\nu(\mathrm{d}z) = |z|^{-\alpha-1}\mathrm{d}z$, for $\alpha = 1.8$, and the stochastic differential equation (1.2) starting from $x = 0$, with $\sigma(y) = (1+y^2)/(1+y^4)$. We are interested in the law of $V = \sup_{t\in[0,1]} X_t$.

We introduce the *exact* Euler scheme $(X^n_{i/n})_{i=0,\dots,n}$ with time-step $1/n$, which can be simulated exactly in this particular case.

For $\epsilon > 0$, we introduce the scheme $(X^{n,\epsilon}_{i/n})_{i=0,\dots,n}$ studied in this paper, defined by (2.2)–(2.3), and we denote by $(\hat{X}^{n,\epsilon}_{i/n})_{i=0,\dots,n}$ the scheme where the jumps smaller than ϵ are simply neglected [14]. We also set

$$V^n = \sup_{i=0,\dots,n} X^n_{i/n}, \quad V^{n,\epsilon} = \sup_{i=0,\dots,n} X^{n,\epsilon}_{i/n}, \quad \hat{V}^{n,\epsilon} = \sup_{i=0,\dots,n} \hat{X}^{n,\epsilon}_{i/n}.$$

In all the simulations, we have chosen $n = 1000$. Using a Monte-Carlo method (with 10^6 simulations) and have estimated the density of V^{1000}. This density is drawn on all the figures (its shape is slightly surprising). Next, we have performed 10^5 simulations of $\hat{V}^{1000,\epsilon}$ and $V^{1000,\epsilon}$ for different values of ϵ, and we have drawn a histogram.

As we can see from the figures below, the law of $V^{1000,\epsilon}$ is really much more close to the law of V^{1000} than the law of $\hat{V}^{1000,\epsilon}$.

The time (in seconds) needed for each simulation of $\hat{V}^{1000,\epsilon}$ and $V^{1000,\epsilon}$ is as follows:

	$\epsilon = 0.1$	$\epsilon = 0.01$	$\epsilon = 0.001$
$\hat{V}^{1000,\epsilon}$	$3.6\ 10^{-4}$	$2.7\ 10^{-3}$	$9.5\ 10^{-2}$
$V^{1000,\epsilon}$	$6.6\ 10^{-4}$	$3.1\ 10^{-3}$	$9.5\ 10^{-2}$

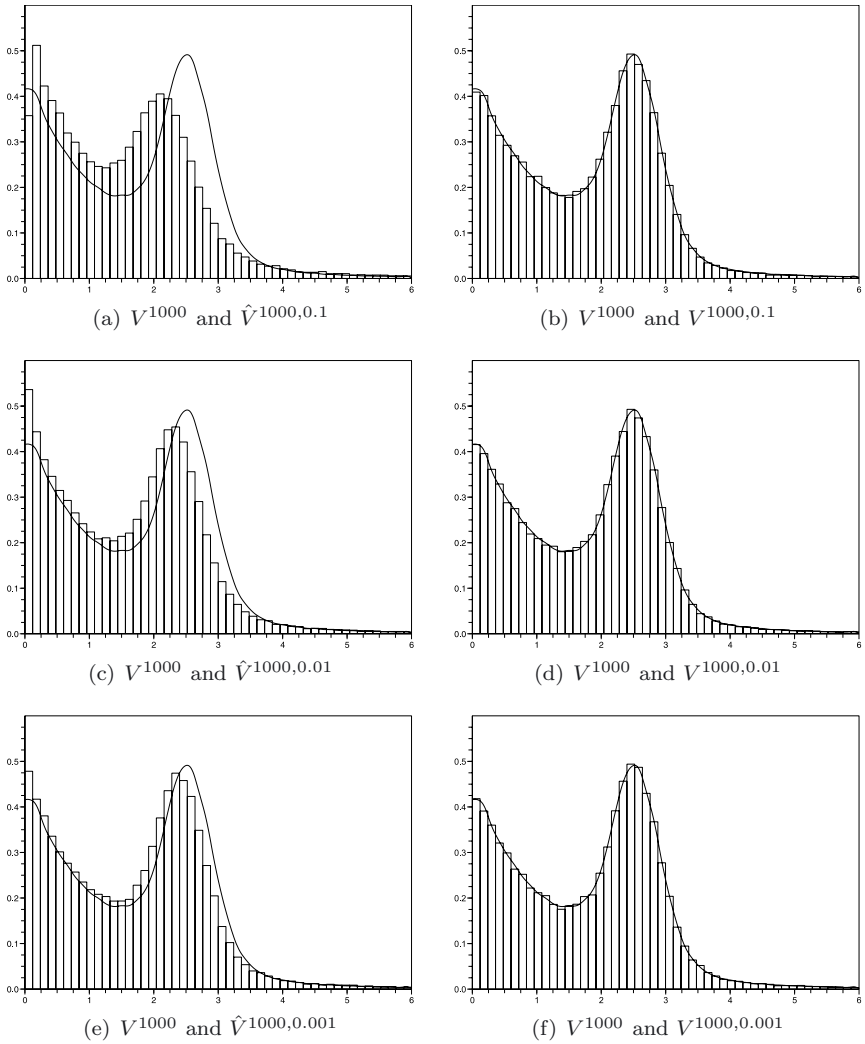
One observes that for ϵ very small, the additional cost to simulate $V^{n,\epsilon}$ is insignificant. This is natural: the simulation schemes of $\hat{V}^{n,\epsilon}$ and $V^{n,\epsilon}$ are the same, except that one needs to simulate additionally n Gaussian random variables for $V^{n,\epsilon}$. Thus the additional cost does not depend on ϵ, and becomes insignificant when ϵ is small.

It took about 66 seconds to obtain Figure 1(b), and about 9500 seconds to get Figure 1(e). Clearly, Figure 1(b) is much more convincing. As a conclusion, it is really better to approximate the small jumps by Gaussian random variables than to neglect them, both from a theoretical and numerical point of view.

The shape of the density of V is quite surprising. Numerical simulations with other values of α (or even with Z replaced by a standard Brownian motion) produce similar shapes. This seems to be due to the fact σ is not monotonous on $\mathbb{R}_+$. In Figure 2, we have drawn simulations with $\sigma(y) = \sqrt{1+y^2}$, for which the density of V has a more classical shape. But in Figure 3, we have considered the oscillating case $\sigma(y) = 1.2 + \sin(4y)$, which seems to produce infinitely many modes for the density of V.

Acknowledgements. I wish to thank Jean Jacod for fruitful discussions. I am also grateful to the anonymous referees for their constructive observations.

https://doi.org/10.1051/ps/2009017 Published online by Cambridge University Press

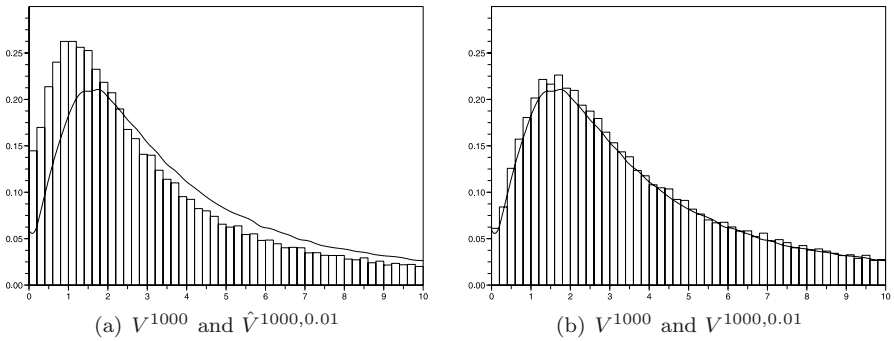


(a) V^{1000} and $\hat{V}^{1000,0.1}$ (b) V^{1000} and $V^{1000,0.1}$

(c) V^{1000} and $\hat{V}^{1000,0.01}$ (d) V^{1000} and $V^{1000,0.01}$

(e) V^{1000} and $\hat{V}^{1000,0.001}$ (f) V^{1000} and $V^{1000,0.001}$

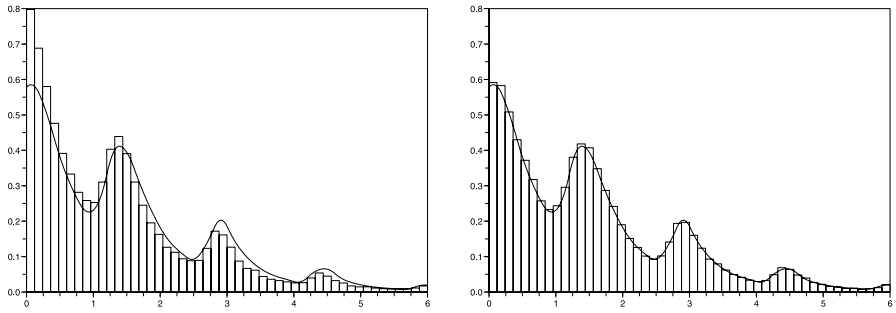
FIGURE 1. Simulations with $\sigma(y) = (1+y^2)/(1+y^4)$.



(a) V^{1000} and $\hat{V}^{1000,0.01}$ (b) V^{1000} and $V^{1000,0.01}$

FIGURE 2. Simulations with $\sigma(y) = \sqrt{1+y^2}$.

https://doi.org/10.1051/ps/2009017 Published online by Cambridge University Press



(a) V^{1000} and $\hat{V}^{1000,0.01}$ (b) V^{1000} and $V^{1000,0.01}$

FIGURE 3. Simulations with $\sigma(y) = 1.2 + \sin(4y)$.

REFERENCES

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

[2] J.M. Chambers, C.L. Mallows and B.W. Stuck, A method for simulating stable random variables. *J. Amer. Statist. Assoc.* **71** (1976) 340–344.

[3] U. Einmahl, Extensions of results of Komlos, Major, and Tusnady to the multivariate case. *J. Multivariate Anal.* **28** (1989) 20–68.

[4] H. Guérin, Solving Landau equation for some soft potentials through a probabilistic approach. *Ann. Appl. Probab.* **13** (2003) 515–539.

[5] N. Ikeda and S. Watanabe, *Stochastic differential equations and diffusion processes,* Second edition. North-Holland Mathematical Library, 24. North-Holland Publishing Co., Amsterdam; Kodansha, Ltd., Tokyo (1989).

[6] J. Jacod, The Euler scheme for Lévy driven stochastic differential equations: limit theorems. *Ann. Probab.* **32** (2004) 1830–1872.

[7] J. Jacod, A. Jakubowski and J. Mémin, On asymptotic errors in discretization of processes. *Ann. Probab.* **31** (2003) 592–608.

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

[9] J. Jacod and P. Protter, Asymptotic error distributions for the Euler method for stochastic differential equations. *Ann. Probab.* **26** (1998) 267–307.

[10] J. Jacod and A.N. Shiryaev, *Limit theorems for stochastic processes,* second edition. Springer-Verlag, Berlin (2003).

[11] J. Komlós, P. Major and G. Tusnády, An approximation of partial sums of independent rvs and the sample df I. *Z. Wahrsch. verw. Gebiete* **32** (1975) 111–131.

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

[13] E. Rio, Upper bounds for minimal distances in the central limit theorem. *Ann. Inst. Henri Poincaré Probab. Stat.* **45** (2009) 802–817.

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

[15] S. Rubenthaler and M. Wiktorsson, Improved convergence rate for the simulation of stochastic differential equations driven by subordinated Lévy processes. *Stochastic Process. Appl.* **108** (2003) 1–26.

[16] H. Tanaka, Probabilistic treatment of the Boltzmann equation of Maxwellian molecules. *Z. Wahrsch. Verw. Gebiete* **46** (1978/79) 67–105.

[17] C. Villani, On a new class of weak solutions to the spatially homogeneous Boltzmann and Landau equations. *Arch. Rat. Mech. Anal.* **143** (1998) 273–307.

[18] J.B. Walsh, A stochastic model of neural response. *Adv. Appl. Prob.* **13** (1981) 231–281.

[19] A. Yu. Zaitsev, Estimates for the strong approximation in multidimensional central limit theorem. *Proceedings of the International Congress of Mathematicians*, Vol. III (2002) 107–116.

https://doi.org/10.1051/ps/2009017 Published online by Cambridge University Press