Finance Stoch (2017) 21:995–1026
DOI 10.1007/s00780-017-0341-7

Multilevel Monte Carlo for exponential Lévy models

Michael B. Giles[1] **· Yuan Xia**[1]

Received: 20 March 2014 / Accepted: 9 May 2017 / Published online: 6 September 2017

Abstract We apply the multilevel Monte Carlo method for option pricing problems using exponential Lévy models with a uniform timestep discretisation. For lookback and barrier options, we derive estimates of the convergence rate of the error introduced by the discrete monitoring of the running supremum of a broad class of Lévy processes. We then use these to obtain upper bounds on the multilevel Monte Carlo variance convergence rate for the variance gamma, NIG and α-stable processes. We also provide an analysis of a trapezoidal approximation for Asian options. Our method is illustrated by numerical experiments.

Keywords Multilevel Monte Carlo · Exponential Lévy models · Asian options · Lookback options · Barrier options

Mathematics Subject Classification (2010) 65C05 · 91G60

JEL Classification C15 · C63

1 Introduction

Exponential Lévy models are based on the assumption that asset returns follow a Lévy process [25, 10]. The asset price follows

$$S_t = S_0 \exp(X_t)$$

✉ Y. Xia
yuan.xia.cn@gmail.com

M.B. Giles
mike.giles@maths.ox.ac.uk

1 Mathematical Institute, Oxford University, Oxford, UK

where X is an (m, σ, ν)-Lévy process, i.e.,

$$X_t = mt + \sigma B_t + \int_0^t \int_{\{|z|\geq 1\}} z\, J(\mathrm{d}z, \mathrm{d}s) + \int_0^t \int_{\{|z|<1\}} z\big(J(\mathrm{d}z, \mathrm{d}s) - \nu(\mathrm{d}z)\mathrm{d}s\big),$$

where m is a constant, B is a Brownian motion, J is the jump measure and ν is the Lévy measure (cf. [24, Theorem I.42]).

Models with jumps give an intuitive explanation of implied volatility skews and smiles in the index option market and foreign exchange market ([10, Chap. 11]). The jump fear is mainly on the downside in the equity market which produces a premium for low-strike options; the jump risk is symmetric in the foreign exchange market so that the implied volatility has a smile shape. Chapter 7 in [10] shows that models building on pure jump processes can reproduce the stylised facts of asset returns, like heavy tails and the asymmetric distribution of increments. Since pure jump processes of finite activity without a diffusion component cannot generate a realistic path, it is natural to allow the jump activity to be infinite. In this work, we deal with infinite-activity pure jump exponential Lévy models, in particular models driven by variance gamma (VG), normal inverse Gaussian (NIG) and α-stable processes which allow direct simulation of increments.

We are interested in estimating the expected payoff value $\mathbb{E}[f(S)]$ in option pricing problems. In the case of European options, it is possible to directly sample the final value of the underlying Lévy process, but for Asian, lookback and barrier options, the option value depends on functionals of the Lévy process and so it is necessary to approximate those. In the case of a VG model with a lookback option, the convergence results in [13] show that to achieve an $\mathcal{O}(\epsilon)$ root mean square (RMS) error using a standard Monte Carlo method with a uniform timestep discretisation requires $\mathcal{O}(\epsilon^{-2})$ paths, each with $\mathcal{O}(\epsilon^{-1})$ timesteps, leading to a computational complexity of $\mathcal{O}(\epsilon^{-3})$.

In the case of a simple Brownian diffusion, Giles [16, 17] introduced a multilevel Monte Carlo (MLMC) method, reducing the computational complexity from $\mathcal{O}(\epsilon^{-3})$ to $\mathcal{O}(\epsilon^{-2})$ for a variety of payoffs. The objective of this paper is to investigate whether similar benefits can be obtained for exponential Lévy processes.

Various researchers have investigated simulation methods for the running maximum of Lévy processes. Reference [15] develops an adaptive Monte Carlo method for functionals of killed Lévy processes with a controlled bias. Small-time asymptotic expansions of the exit probability are given with computable error bounds. For evaluating the exit probability when the barrier is close to the starting point of the process, this algorithm outperforms a uniform discretisation significantly. Reference [21] develops a novel Wiener–Hopf Monte Carlo method to generate the joint distribution of $(X_T, \sup_{0\leq t\leq T} X_t)$ which is further extended to MLMC in [14], obtaining an RMS error ϵ with a computational complexity of $\mathcal{O}(\epsilon^{-3})$ for Lévy processes with bounded variation, and $\mathcal{O}\left(\epsilon^{-4}\right)$ for processes with infinite variation. The method currently cannot be directly applied to VG, NIG and α-stable processes. References [12, 11] adapt MLMC to Lévy-driven SDEs with payoffs which are Lipschitz with respect to the supremum norm. If the Lévy process does not incorporate a Brownian process, reference [11] obtains an $\mathcal{O}(\epsilon^{-(6\beta)/(4-\beta)})$ upper bound on the worst case computational complexity, where β is the BG index which will be defined later.

In contrast to those advanced techniques, we take the discretely monitored maximum based on a uniform timestep discretisation of the Lévy process as the approximation. The outline of the work is as follows. First we review the multilevel Monte Carlo method and present the three Lévy processes we consider in our numerical experiments. To prepare for the analysis of the multilevel variance of lookback and barrier, we bound the convergence rate of the discretely monitored running maximum for a large class of Lévy processes whose Lévy measures have a power law behaviour for small jumps, and have exponential tails. Based on this, we conclude by bounding the variance of the multilevel estimators. Numerical results are then presented for the multilevel Monte Carlo method applied to Asian, lookback and barrier options, using the three different exponential Lévy models.

2 Multilevel Monte Carlo (MLMC) method

For a path-dependent payoff P based on an exponential Lévy model on the time interval $[0,T]$, let $\widehat{P}_\ell$ denote its approximation using a discretisation with M^ℓ uniform timesteps of size $h_\ell = M^{-\ell}T$ on level ℓ; in the numerical results reported later, we use $M=2$. Due to the linearity of the expectation operator, we have the identity

$$\mathbb{E}[\widehat{P}_L] = \mathbb{E}[\widehat{P}_0] + \sum_{\ell=1}^{L} \mathbb{E}[\widehat{P}_\ell - \widehat{P}_{\ell-1}].$$

Let $\widehat{Y}_0$ denote the standard Monte Carlo estimate for $\mathbb{E}[\widehat{P}_0]$ using N_0 paths, and for $\ell > 0$, we use N_ℓ independent paths to estimate $\mathbb{E}[\widehat{P}_\ell - \widehat{P}_{\ell-1}]$ using

$$\widehat{Y}_\ell = N_\ell^{-1} \sum_{i=1}^{N_\ell} \big(\widehat{P}_\ell^{(i)} - \widehat{P}_{\ell-1}^{(i)}\big).$$

For a given path generated for $\widehat{P}_\ell^{(i)}$, we can calculate $\widehat{P}_{\ell-1}^{(i)}$ using the same underlying Lévy path. The multilevel method exploits the fact that $V_\ell := \mathbb{V}[\widehat{P}_\ell - \widehat{P}_{\ell-1}]$ decreases with ℓ, and adaptively chooses N_ℓ to minimise the computational cost to achieve a desired RMS error. This is summarised in the following theorem in [18, Theorem 1] and [19, Theorem 1].

Theorem 2.1 *Let P denote a functional of (S_t), and let $\widehat{P}_\ell$ denote the corresponding approximation using a discretisation with uniform timestep $h_\ell = M^{-\ell}T$. If there exist independent estimators $\widehat{Y}_\ell$ based on N_ℓ Monte Carlo samples, each with complexity C_ℓ, and positive constants $\alpha, \beta, c_1, c_2, c_3$ such that $\alpha \geq \frac{1}{2}\min(1,\beta)$ and*

i) $|\mathbb{E}[\widehat{P}_\ell - P]| \leq c_1 h_\ell^\alpha$,

ii) $\mathbb{E}[\widehat{Y}_\ell] = \begin{cases} \mathbb{E}[\widehat{P}_0], & \ell = 0, \\ \mathbb{E}[\widehat{P}_\ell - \widehat{P}_{\ell-1}], & \ell > 0, \end{cases}$

iii) $\mathbb{V}[\widehat{Y}_\ell] \leq c_2 N_\ell^{-1} h_\ell^\beta$,

iv) $C_\ell \leq c_3 N_\ell h_\ell^{-1}$, *then there exists a positive constant c_4 such that for any $\epsilon < e^{-1}$, there are values L and N_ℓ for which the multilevel estimator*

$$\widehat{Y} = \sum_{\ell=0}^{L} \widehat{Y}_\ell$$

has a mean square error with bound

$$MSE := \mathbb{E}\big[(\widehat{Y} - \mathbb{E}[P])^2\big] < \epsilon^2$$

with a computational complexity C with bound

$$C \leq \begin{cases} c_4\,\epsilon^{-2}, & \beta > 1, \\ c_4\,\epsilon^{-2}(\log\epsilon)^2, & \beta = 1, \\ c_4\,\epsilon^{-2-(1-\beta)/\alpha}, & 0 < \beta < 1. \end{cases}$$

We focus on the multilevel variance convergence rate β in the following numerical results and analysis since it is crucial in determining the computational complexity.

3 Lévy models

The numerical results to be presented later use the following three models.

3.1 Variance gamma (VG)

The VG process with parameter set (σ, θ, κ) is the Lévy process X with characteristic function $\mathbb{E}[\exp(iuX_t)] = (1 - iu\theta\kappa + \frac{1}{2}\sigma^2u^2\kappa)^{-t/\kappa}$. The Lévy measure of the VG process is ([10, Table 4.5])

$$\nu(x) = \frac{1}{\kappa\,|x|} e^{A-B|x|} \quad \text{with } A = \frac{\theta}{\sigma^2} \text{ and } B = \frac{\sqrt{\theta^2 + 2\sigma^2/\kappa}}{\sigma^2}.$$

One advantage of the VG process is that its additional parameters make it possible to fit the skewness and kurtosis of the stock returns ([10, Sect. 7.3]). Another is that it is easily simulated as we have a subordinator representation $X_t = \theta G_t + \sigma B_{G_t}$, in which B is a Brownian process and the subordinator G is a gamma process with parameters $(1/\kappa, 1/\kappa)$.

For ease of computation, we use the mean-correcting pricing measure in [25, Sect. 6.2.2], with risk-free interest rate $r = 0.05$. Let $(\exp(-rt)S_t)$ be a martingale. This results in the drift being

$$m = r + \kappa^{-1} \log\left(1 + \theta\kappa - \frac{1}{2}\sigma^2\kappa\right).$$

After transforming the parameter representation to the definition we use, the calibration in [25, Table 6.3] gives $\sigma = 0.1213$, $\theta = -0.1436$, $\kappa = 0.1686$.

3.2 Normal inverse Gaussian (NIG)

The NIG process with parameter set (σ, θ, κ) is the Lévy process X with characteristic function $\mathbb{E}[\exp(iuX_t)] = \exp(\frac{t}{\kappa} - \frac{t}{\kappa}\sqrt{1 - 2iu\theta\kappa + \kappa\sigma^2 u^2})$ and Lévy measure

$$\nu(x) = \frac{C}{\kappa|x|} e^{Ax} K_1(B|x|)$$

with

$$A = \frac{\theta}{\sigma^2}, \quad B = \frac{\sqrt{\theta^2 + \sigma^2/\kappa}}{\sigma^2}, \quad C = \frac{\sqrt{\theta^2 + 2\sigma^2/\kappa}}{2\pi\sigma\sqrt{\kappa}},$$

where $K_1(x)$ is the modified Bessel function of the second kind (see [10, Sect. 4.4.3]). As $x \to 0$, $K_1(x) \sim \frac{1}{x} + \mathcal{O}(1)$, while as $x \to \infty$, $K_1(x) \sim e^{-x}\sqrt{\frac{\pi}{2|x|}}(1 + \mathcal{O}(\frac{1}{|x|}))$.

In terms of simulation, the NIG process can be represented as $X_t = \theta I_t + \sigma B_{I_t}$, where the subordinator I is an inverse Gaussian process with parameters $(\frac{1}{\kappa}, 1)$. Algorithm 6.9 in [10] can be used to generate inverse Gaussian samples.

Using the mean-correcting pricing measure leads to

$$m = r - \kappa^{-1} + \pi C B \kappa^{-1}\sqrt{B^2 - (A+1)^2}.$$

Following the calibration in [25], we use the parameters $\sigma = 0.1836$, $\theta = -0.1313$, $\kappa = 1.2819$, and again use the risk-free interest rate $r = 0.05$.

3.3 Spectrally negative α-stable process

The scalar spectrally negative α-stable process has a Lévy measure of the form ([23, Sect. 1.2.6])

$$\nu(x) = \frac{B}{|x|^{\alpha+1}} \mathbf{1}_{\{x<0\}}$$

for $0 < \alpha < 2$ and some nonnegative B. We follow [23] to discuss another parametrisation of the α-stable process with characteristic function

$$\mathbb{E}[\exp(iuX_t)] = \exp\left(-tB^\alpha |u|^\alpha \left(1 + i\,\mathrm{sgn}(u)\tan\frac{\pi\alpha}{2}\right)\right) \quad \text{if } \alpha \neq 1,$$

$$\mathbb{E}[\exp(iuX_t)] = \exp\left(-tB|u|\left(1 + i\frac{2}{\pi}\mathrm{sgn}(u)\log|u|\right)\right) \quad \text{if } \alpha = 1,$$

where $\mathrm{sgn}(u) = |u|/u$ if $u \neq 0$ and $\mathrm{sgn}(0) = 0$. There are no positive jumps for the spectrally negative process, which has a finite exponential moment $\mathbb{E}[\exp(uX_t)]$ [4].

For this case, the mean-correcting drift is

$$m = r + B^\alpha \sec\frac{\alpha\pi}{2}.$$

Sample paths of α-stable processes can be generated by the algorithm in [5]. Following [4], we use the parameters $\alpha = 1.5597$ and $B = 0.1486$.

4 Key numerical analysis results

The variance $V_\ell = \mathbb{V}[\widehat{P}_\ell - \widehat{P}_{\ell-1}]$ of the multilevel correction depends on the behaviour of the difference between the continuously and discretely monitored suprema of X, defined for a unit time interval as

$$D_n = \sup_{0\le t\le 1} X_t - \max_{i=0,1,\dots,n} X_{i/n}.$$

To derive the order of weak convergence for lookback-type payoffs, we are concerned with $\mathbb{E}[D_n]$, which is extensively studied in the literature. For example, [13, 8, 9] derive asymptotic expansions for jump-diffusion, VG, NIG processes, as well as estimates for general Lévy processes, by using Spitzer's identity [26].

A key result due to Chen [8, Theorem 5.2.1] is the following:

Theorem 4.1 *Suppose X is a scalar Lévy process with triple (m, σ, ν), with finite first moment, i.e.,*

$$\int_{\{|x|>1\}} |x|\,\nu(\mathrm{d}x) < \infty.$$

Then $D_n = \sup_{0\le t\le 1} X_t - \max_{i=0,1,\dots,n} X_{i/n}$ satisfies the following:

1. *If $\sigma > 0$, then*

$$\mathbb{E}[D_n] = \mathcal{O}(1/\sqrt{n}).$$

2. *If $\sigma = 0$ and X is of finite variation, i.e., $\int_{\{|x|<1\}} |x|\,\nu(\mathrm{d}x) < \infty$, then*

$$\mathbb{E}[D_n] = \mathcal{O}(\log n/n).$$

3. *If $\sigma = 0$ and X is of infinite variation, then*

$$\mathbb{E}[D_n] = o(n^{-1/\beta+\delta}),$$

where

$$\beta = \inf\left\{\alpha > 0 : \int_{\{|x|<1\}} |x|^\alpha\,\nu(\mathrm{d}x) < \infty\right\}$$

is the Blumenthal–Getoor index of X, and $\delta > 0$ is an arbitrarily small strictly positive constant.

The VG process has finite variation with Blumenthal–Getoor index 0; the NIG process has infinite variation with Blumenthal–Getoor index 1. They correspond to the second and third cases of Theorem 4.1, respectively.

For the multilevel variance analysis, we require higher moments of D_n. In the pure Brownian case, Asmussen et al. [1] obtain the asymptotic distribution of D_n, which in turn gives the asymptotic behaviour of $\mathbb{E}[D_n^2]$. [13] extends their result to finite activity jump processes with non-zero diffusion.

However, in this paper we are looking at infinite activity jump processes. Our main new result is therefore concerned with the L^p convergence rate of D_n for pure jump Lévy processes. This will be used later to bound the variance of the multilevel Monte Carlo correction term V_ℓ for both lookback and barrier options.

Theorem 4.2 *Let X be a scalar pure jump Lévy process, and suppose its Lévy measure $\nu(x)$ satisfies*

$$C_2\,|x|^{-1-\alpha} \le \nu(x) \le C_1\,|x|^{-1-\alpha} \quad \text{for } |x| \le 1,$$
$$\nu(x) \le \exp\left(-C_3\,|x|\right) \quad \text{for } |x| > 1,$$

where $C_1, C_2, C_3 > 0,\ 0 \le \alpha < 2$ are constants. Then for $p \ge 1$,

$$D_n = \sup_{0\le t\le 1} X_t \; - \max_{i=0,1,\ldots,n} X_{i/n}$$

satisfies

$$\mathbb{E}[D_n^p] = \begin{cases} \mathcal{O}\left(1/n\right), & p > 2\alpha, \\ \mathcal{O}((\log n/n)^{\frac{p}{2\alpha}}), & p \le 2\alpha. \end{cases}$$

If in addition X is spectrally negative, i.e., $\nu(x) = 0$ for $x > 0$, then

$$\mathbb{E}[D_n^p] = \begin{cases} \mathcal{O}(n^{-p}), & 0 \le \alpha < 1, \\ o(n^{-p/\alpha+\delta}), & 1 \le \alpha < 2, \end{cases}$$

for any $\delta > 0$.

We give the proof of this result later in Sect. 7.6. Note that for $p = 1$, the general bound in Theorem 4.2 is slightly sharper than Chen's result for $\alpha < \frac{1}{2}$, is the same for $\alpha = \frac{1}{2}$, and is not as tight as Chen's result for $\frac{1}{2} < \alpha < 2$; the spectrally negative bound is slightly sharper than Chen's result for $\alpha < 1$, and the bound is the same for $1 \le \alpha < 2$.

5 MLMC analysis

5.1 Asian options

We consider the analysis for a Lipschitz arithmetic Asian payoff $P = P(\overline{S})$, where

$$\overline{S} = S_0\; T^{-1} \int_0^T \exp\left(X_t\right) \mathrm{d}t$$

and P is Lipschitz such that $|P(S_1) - P(S_2)| \le L_K |S_1 - S_2|$. We approximate the integral using a trapezoidal approximation

$$\overline{\overline{S}} := S_0\; T^{-1} \sum_{j=0}^{n-1} \frac{1}{2} h\big(\exp(X_{jh}) + \exp(X_{(j+1)h})\big),$$

and the approximated payoff is then $\widehat{P} = P(\overline{\overline{S}})$.

Proposition 5.1 *Let X be a scalar Lévy process underlying an exponential Lévy model. If $\overline{\overline{S}}$, $\overline{S}$ are as defined above and $\int_{\{|z|>1\}} e^{2z}\,\nu(\mathrm{d}z) < \infty$, then*

$$\mathbb{E}\big[(\overline{\overline{S}} - \overline{S})^2\big] = \mathcal{O}(h^2).$$

The proof will be given later in Sect. 7.1. Using the Lipschitz property, the weak convergence for the numerical approximation is given by

$$|\mathbb{E}[\widehat{P}_\ell - P]| \le L_K \mathbb{E}[|\overline{\overline{S}}_\ell - \overline{S}|] \le L_K \big(\mathbb{E}[(\overline{\overline{S}} - \overline{S})^2]\big)^{1/2},$$

while the convergence of the MLMC variance follows from

$$\begin{aligned} V_\ell &\le \mathbb{E}[(\widehat{P}_\ell - \widehat{P}_{\ell-1})^2] \\ &\le 2\mathbb{E}[(\widehat{P}_\ell - P)^2] + 2\,\mathbb{E}[(\widehat{P}_{\ell-1} - P)^2] \\ &\le 2L_K^2\ \mathbb{E}[(\overline{\overline{S}}_\ell - \overline{S})^2] + 2L_K^2\ \mathbb{E}[(\overline{\overline{S}}_{\ell-1} - \overline{S})^2]. \end{aligned}$$

5.2 Lookback options

In exponential Lévy models, the moment generating function $\mathbb{E}[\exp(q \sup_{0\le t\le T} X_t)]$ can be infinite for a large value of q. To avoid problems due to this, we consider a lookback put option which has a bounded payoff

$$P = \exp(-rT)\big(K - S_0 \exp(m)\big)^+, \tag{5.1}$$

where $m = \sup_{0\le t\le T} X_t$. Note that P is a Lipschitz function of m, since we have $|P'(x)| \le K$. Without loss of generality, we assume $T = 1$ in the following.

Because of the Lipschitz property, we have the estimate $|\mathbb{E}[P - \widehat{P}_\ell]| \le K\,\mathbb{E}[D_n]$, where $n = M^\ell = h_\ell^{-1}$. Therefore we obtain weak convergence for the processes covered by Theorem 4.1, with the convergence rate given by that theorem.

To analyse the variance, $V_\ell = \mathbb{V}[\widehat{P}_\ell - \widehat{P}_{\ell-1}]$, we first note that

$$0 \le \max_{0\le i\le M^\ell} X_{i/M^\ell} - \max_{0\le i\le M^{\ell-1}} X_{i/M^{\ell-1}} \le \sup_{0\le t\le 1} X_t - \max_{0\le i\le M^{\ell-1}} X_{i/M^{\ell-1}} = D_n,$$

where $n = M^{\ell-1}$. Hence, we have

$$V_\ell \le \mathbb{E}[(\widehat{P}_\ell - \widehat{P}_{\ell-1})^2] \le K^2\,\mathbb{E}[D_n^2].$$

Theorem 4.2 then provides the following bounds on the variance for the VG, NIG and spectrally negative α-stable processes.

Proposition 5.2 *Let X be a scalar Lévy process underlying an exponential Lévy model. For the Lipschitz lookback put payoff* (5.1), *we have the following multilevel variance convergence rate results*:

1. *If X is a variance gamma* (*VG*) *process*, *then $V_\ell = \mathcal{O}(h_\ell)$.*
2. *If X is a normal inverse Gaussian* (*NIG*) *process*, *then $V_\ell = \mathcal{O}(h_\ell |\log h_\ell|)$.*
3. *If X is a spectrally negative α-stable process with $\alpha > 1$*, *then $V_\ell = o(h_\ell^{2/\alpha-\delta})$*, *for any small $\delta > 0$.*

5.3 Barrier options

We consider a bounded up-and-out barrier option with discounted payoff

$$P = \exp(-rT)\, f(S_T)\, \mathbf{1}_{\{\sup_{0<t<T} S_t < B\}} = \exp(-rT)\, f(S_T)\, \mathbf{1}_{\{m<\log(B/S_0)\}}, \tag{5.2}$$

where again $m = \sup_{0<t<T} X_t$, and $|f(x)| \le F$ is bounded. On level ℓ, the numerical approximation is

$$\widehat{P}_\ell = \exp(-rT)\; f(S_T)\; \mathbf{1}_{\{\widehat{m}_\ell<\log(B/S_0)\}}, \tag{5.3}$$

where $\widehat{m}_\ell = \max_{0\le i\le M^\ell} X_{ih_\ell}$.

Our analysis for NIG and the spectrally negative α-stable processes requires the following quite general result.

Proposition 5.3 *If m is a random variable with a locally bounded density in a neighbourhood of B and $\widehat{m}$ is a numerical approximation to m, then for any $p > 0$, there exists a constant $C_p(B)$ such that*

$$\mathbb{E}[|\mathbf{1}_{\{m<B\}} - \mathbf{1}_{\{\widehat{m}<B\}}|] < C_p(B) \|m - \widehat{m}\|_p^{p/(p+1)}.$$

Proof This result was first proved by Avikainen (Lemma 3.4 in [2]), but we give here a simpler proof. If, for some fixed $X > 0$, we have $|m - B| > X$ and $|m - \widehat{m}| < X$, then $\mathbf{1}_{\{m<B\}} - \mathbf{1}_{\{\widehat{m}<B\}} = 0$. Hence,

$$\begin{aligned}\mathbb{E}[|\mathbf{1}_{\{m<B\}} - \mathbf{1}_{\{\widehat{m}<B\}}|] &\le \mathbb{P}[|m - B| \le X] + \mathbb{P}[|m - \widehat{m}| \ge X]\\ &\le 2\, \rho_{\sup}(B)\, X + X^{-p}\; \|m - \widehat{m}\|_p^p,\end{aligned}$$

with the first term being due to the local bound $\rho_{\sup}(B)$ of the density of m and the second to the Markov inequality. Differentiating the upper bound with respect to X, we find that it is minimised by choosing $X^{p+1} = \frac{p}{2\,\rho_{\sup}(B)}\; \|m - \widehat{m}\|_p^p$, and we then get the desired bound. □

Our analysis for the variance gamma process requires a sharper result customised to the properties of Lévy processes.

Proposition 5.4 *If X is a scalar pure jump Lévy process satisfying the conditions of Theorem 4.2 with $0 \le \alpha \le \frac{1}{2}$ and m and $\widehat{m}_n$ are the continuously and discretely monitored suprema of X and m has a locally bounded density in a neighbourhood of B, then*

$$\mathbb{E}[|\mathbf{1}_{\{m<B\}} - \mathbf{1}_{\{\widehat{m}<B\}}|] = o\big(n^{-1/(1+2\alpha)+\delta}\big),$$

for any $\delta > 0$.

The proof is given later in Sect. 7.7.

Both of the above propositions require the condition that the supremum m has a locally bounded density for all strictly positive values. There is considerable current research on the supremum of Lévy processes [6, 7, 20, 22]. In particular, the comments following [7, Proposition 2] indicate that the condition is satisfied by stable processes, and by a wide class of symmetric subordinated Brownian motions. Unfortunately, the VG and NIG processes in the current paper are not symmetric, so at present they lie outside the range of current theory, but new theory under development [3] will extend the property to a larger class of Lévy processes including both VG and NIG.

We now bound the weak convergence of the estimator and the multilevel variance convergence.

Proposition 5.5 *Let X be a scalar Lévy process underlying an exponential Lévy model. For the up-and-out barrier option payoff* (5.2) *with the numerical approximation* (5.3), *we have the following rates of convergence for the multilevel correction variance and the weak error, assuming that m has a bounded density*:

1. *If X is a variance gamma* (*VG*) *process*, *then*

$$V_\ell = o(h_\ell^{1-\delta}),$$
$$|\mathbb{E}[\widehat{P} - P]| = o(h_\ell^{1-\delta}),$$

   where δ is an arbitrary positive number.
2. *If X is a NIG process*, *then*

$$V_\ell = o(h_\ell^{1/2-\delta}),$$
$$|\mathbb{E}[\widehat{P} - P]| = o(h_\ell^{1/2-\delta}),$$

   where δ is an arbitrary positive number.
3. *If X is a spectrally negative α-stable process with $\alpha > 1$*, *then*

$$V_\ell = o(h_\ell^{\frac{1}{\alpha}-\delta}),$$
$$|\mathbb{E}[\widehat{P} - P]| = o(h_\ell^{\frac{1}{\alpha}-\delta}),$$

   where δ is an arbitrary positive number.

Proof The variance of the multilevel correction term is bounded by

$$V_\ell \le \mathbb{E}[(\widehat{P}_\ell - \widehat{P}_{\ell-1})^2] \le 2\,\mathbb{E}[(\widehat{P}_\ell - P)^2] + 2\,\mathbb{E}[(\widehat{P}_{\ell-1} - P)^2].$$

For an up-and-out barrier option, since the payoff is bounded, we have

$$\mathbb{E}[(\widehat{P}_\ell - P)^2] \le F^2\,\mathbb{E}[\mathbf{1}_{\{\widehat{m}_n < \log(B/S_0)\}} - \mathbf{1}_{\{m < \log(B/S_0)\}}],$$
$$|\mathbb{E}[\widehat{P}_\ell - P]| \le F\,\mathbb{E}[\mathbf{1}_{\{\widehat{m}_n < \log(B/S_0)\}} - \mathbf{1}_{\{m < \log(B/S_0)\}}],$$

where $n = M^\ell$.

The bounds for the VG process come from Proposition 5.4 together with the results from Theorem 4.2.

The bounds for the NIG come from taking $p=1$ in Proposition 5.3 together with Chen's result in Theorem 4.1.

The bounds for the spectrally negative α-stable process come from Proposition 5.3 together with the results from Theorem 4.2. The latter gives

$$\|m-\widehat{m}\|_p^{p/(p+1)} = (\mathbb{E}[|m-\widehat{m}|^p])^{1/(p+1)} = o(h^{\frac{p}{(p+1)\alpha}-\frac{\delta}{p+1}}).$$

We then obtain the desired bound by taking p to be sufficiently large. □

6 Numerical results

We have numerical results for three different Lévy models: variance gamma, normal inverse Gaussian and α-stable processes, and three different options: Asian, lookback and barrier.

The current code is based on Giles' MATLAB code [17], using which we generate standardised numerical results and a set of four figures. The top two plots correspond to a set of experiments to investigate how the variance and mean for both $\widehat{P}_\ell$ and $\widehat{P}_\ell - \widehat{P}_{\ell-1}$ vary with level ℓ. The top left plot shows the values for $\log_2$ variance, so that the absolute value of the slope of the line for $\log_2 \mathbb{V}[\widehat{P}_\ell - \widehat{P}_{\ell-1}]$ indicates the convergence rate β of V_ℓ in condition iii) of Theorem 2.1. Similarly, the absolute value of the slope of the line for $\log_2 |\mathbb{E}[\widehat{P}_\ell - \widehat{P}_{\ell-1}]|$ in the top right plot indicates the weak convergence rate α in the condition i) of Theorem 2.1.

The bottom two plots correspond to a set of MLMC calculations for different values of the desired accuracy ϵ. Each line in the bottom left plot corresponds to one multilevel calculation and displays the number of samples N_ℓ on each level. Note that as ϵ is varied, the MLMC algorithm automatically decides how many levels are required to reduce the weak error appropriately. The optimal number of samples on each level is based on an empirical estimation of the multilevel correction variance V_ℓ, together with the use of a Lagrange multiplier to determine how best to minimise the overall computational cost for a given target accuracy. A complete description of the algorithm is given in [19, Sect. 3.1, Algorithm 1]. The bottom right plots show the variation of the computational complexity C with the desired accuracy ϵ. In the best cases, the MLMC complexity is $\mathcal{O}(\epsilon^{-2})$, and therefore the plot is of $\epsilon^2 C$ versus ϵ so that we can see whether this is achieved, and compare the complexity to that of the standard Monte Carlo method.

6.1 Asian option

The Asian option we consider is an arithmetic Asian call option with discounted payoff

$$P = \exp(-rT)\ \max(0, \overline{S}-K),$$

where $T = 1$, $r = 0.05$, $S_0 = 100$, $K = 100$ and

$$\overline{S} = S_0\, T^{-1} \int_0^T \exp(X_t)\, \mathrm{d}t.$$

For a general Lévy process, it is not easy to directly sample the integral process. We use the trapezoidal approximation

$$\overline{\overline{S}} := S_0\, T^{-1} \sum_{j=0}^{n-1} \frac{1}{2} h \big(\exp(X_{jh}) + \exp(X_{(j+1)h}) \big),$$

where $n = T/h$ is the number of timesteps. The payoff approximation is then

$$\widehat{P} = \exp(-rT)\ \max(0, \overline{\overline{S}} - K).$$

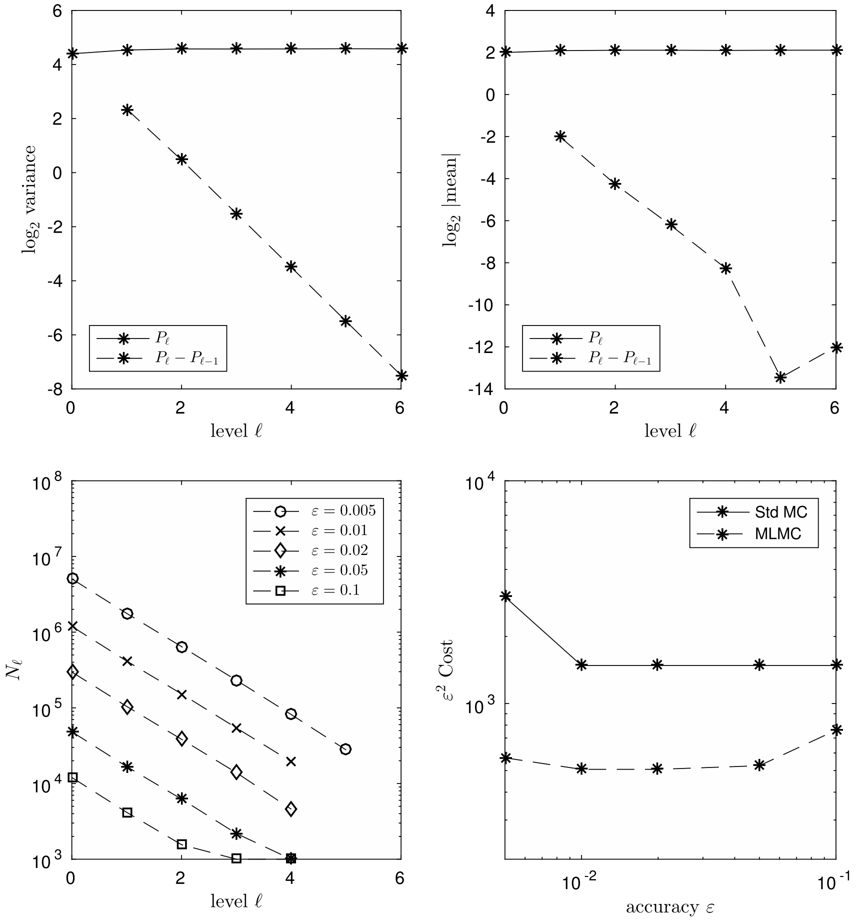

Fig. 1 Asian option in variance gamma model

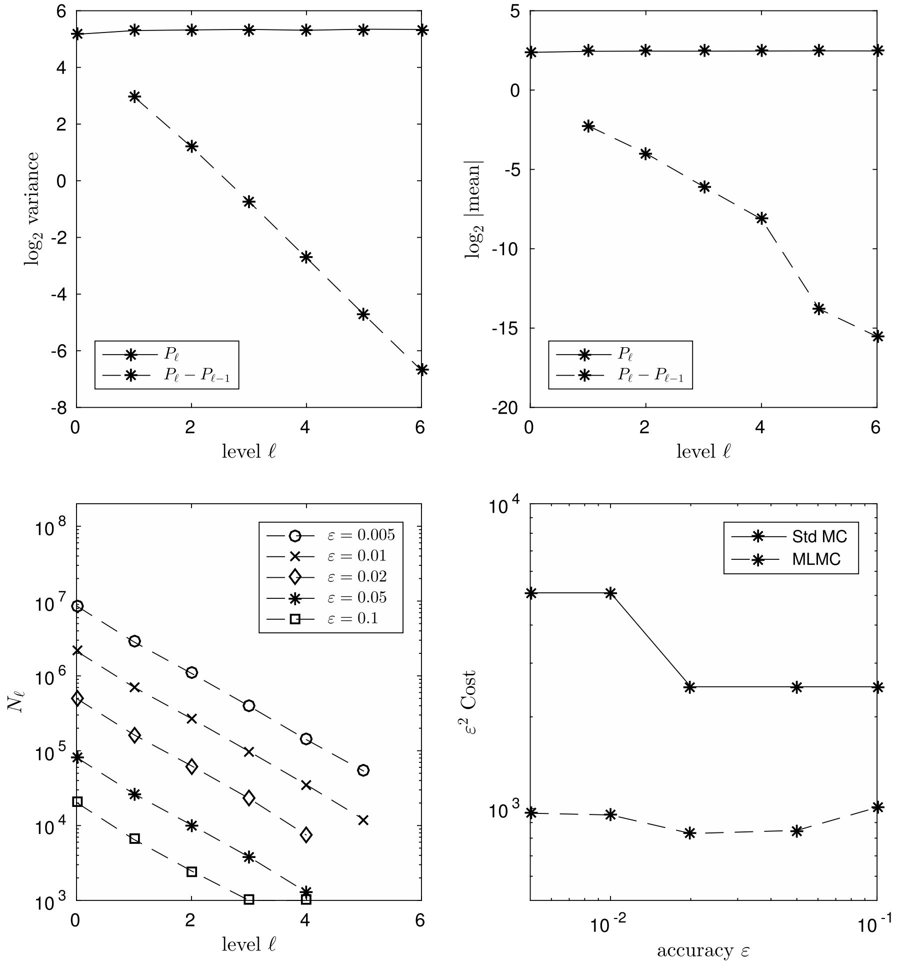

Fig. 2 Asian option in normal inverse Gaussian model

In the multilevel estimator, the approximation $\widehat{P}_\ell$ on level ℓ is obtained using $n_\ell := 2^\ell$ timesteps.

Figures 1, 2, 3 are for the VG, NIG and α-stable models, respectively. The numerical results in the top right plots indicate approximately second order weak convergence. With the standard Monte Carlo method, the top left plots show that the variance is approximately independent of the level, or equivalently the timestep, and therefore, the standard Monte Carlo calculation has computational cost $\mathcal{O}(\epsilon^{-2} n_\ell) = \mathcal{O}(\epsilon^{-2.5})$. Multiplying this cost by ϵ^2 to create the bottom right complexity plots, the scaled cost is $\mathcal{O}(n_\ell)$ and therefore goes up in steps as ϵ is reduced, when decreasing ϵ requires an increase in the value of the finest level L. On the other hand, the convergence rate of the variance of the MLMC estimator is approximately 1.2 for VG, 2.0 for NIG and 2 for the α-stable model. Since in all three cases we have $\beta > 1$, the MLMC theorem gives a complexity which is $\mathcal{O}(\epsilon^{-2})$ which is consistent

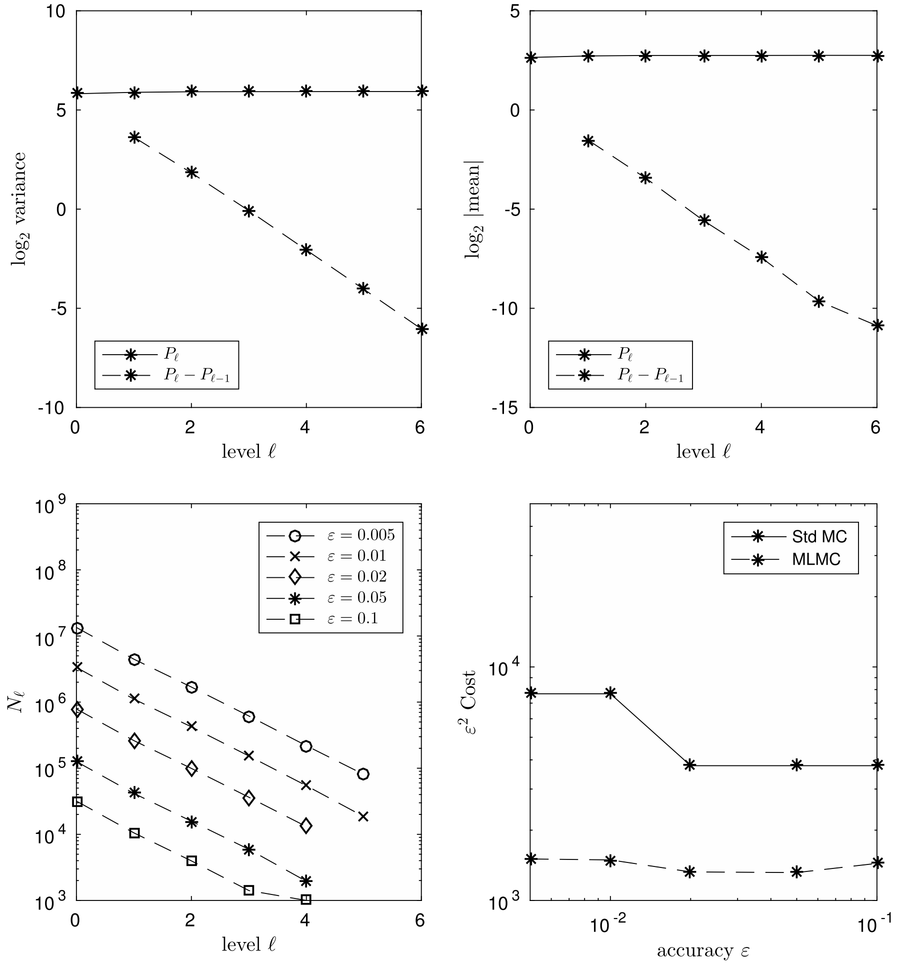

Fig. 3 Asian option in spectrally negative α-stable model

with the results in the bottom right plots which show little variation in $\epsilon^2 C$ for the MLMC estimator.

For this Asian option, MLMC is 3–8 times more efficient than standard MC. The gains are modest because the high rate of weak convergence means that only 4 levels of refinement are required in most cases, so there is only a $2^4 = 16$ difference in cost between each MC path calculation on the finest level, and each of the MLMC path calculations on the coarsest level.

6.2 Lookback option

The lookback option we consider is a put option on the floating underlying,

$$P = \exp(-rT)\left(K - \sup_{0 \le t \le T} S_t\right)^+ = \exp(-rT)\big(K - S_0 \exp(m)\big)^+,$$

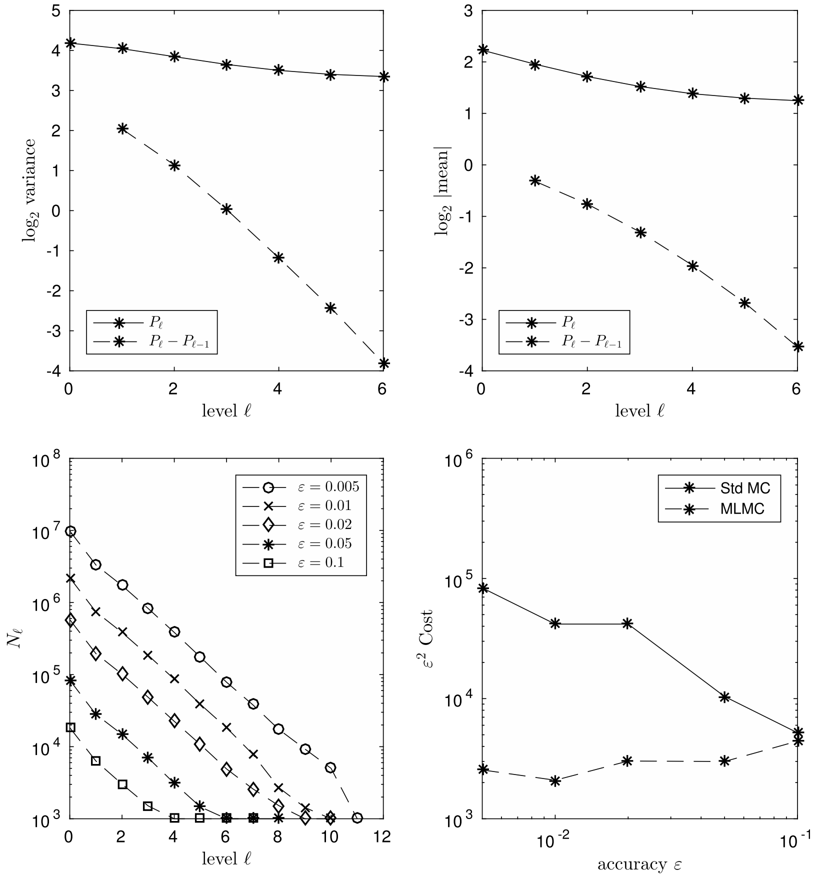

Fig. 4 Lookback option with variance gamma model

where $m = \sup_{0 \le t \le T} X_t$, with $T = 1$, $r = 0.05$, $S_0 = 100$, $K = 110$. We use the discretely monitored maximum as the approximation, so that

$$\widehat{P}_\ell = \exp(-rT)\big(K - S_0 \exp(\widehat{m}_\ell)\big)^+, \quad \widehat{m}_\ell = \max_{0 \le j \le n_\ell} X_{jh_\ell}.$$

Figures 4, 5, 6 show the numerical results for the VG, NIG and α-stable models. The most obvious difference compared to the Asian option is a greatly reduced order of weak convergence, approximately 1, 0.8 and 0.6 in the respective cases. This reduced weak convergence leads to a big increase in the finest approximation level, which in turn greatly increases the standard MC cost but does not significantly change the MLMC cost. Hence, the computational savings are much greater than for the Asian option, with savings of up to a factor of 30.

The small erratic fluctuation in N_ℓ on levels greater than 5 is due to poor estimates of the variance due to a limited number of samples. This also appears later for the barrier option.

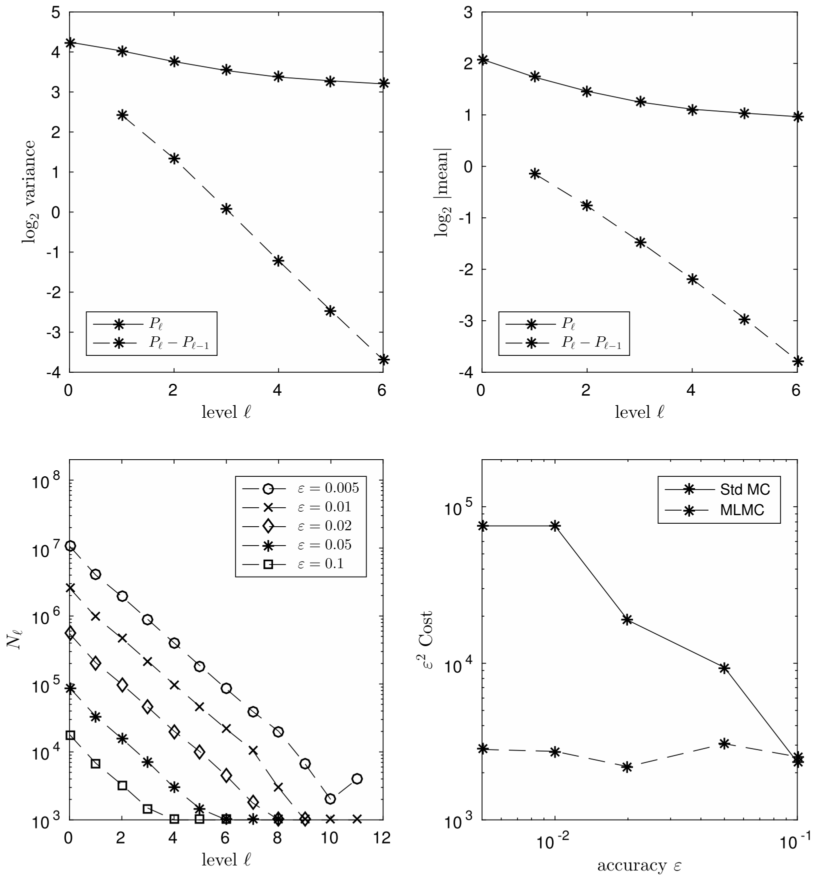

Fig. 5 Lookback option with normal inverse Gaussian model

6.3 Barrier option

The barrier option is an up-and-out call with payoff

$$P = \exp(-rT)\,(S_T - K)^+\,\mathbf{1}_{\{\sup_{0\le t\le T} S(t)<B\}} = \exp(-rT)\,(S_T - K)^+\,\mathbf{1}_{\{m<\log(B/S_0)\}},$$

with $T = 1$, $r = 0.05$, $S_0 = 100$, $K = 100$, $B = 115$. The discretely monitored approximation is

$$\widehat{P}_\ell = \exp(-rT)\,(S_T - K)^+\,\mathbf{1}_{\{\hat{m}_\ell<\log(B/S_0)\}}, \quad \widehat{m}_\ell = \max_{0\le j\le n_\ell} X_{jh_\ell}$$

With the barrier option (Figs. 7, 8, 9), the most noticeable change from the previous options is a reduction in the rate of convergence β of the MLMC variance, with $\beta \approx 0.75, 0.5, 0.6$ in the three cases. For $\beta < 1$, the MLMC theorem proves a complexity which is $\mathcal{O}(\epsilon^{-2-(1-\beta)/\alpha})$, with α here being the rate of weak convergence. The fact that the MLMC complexity is not $\mathcal{O}(\epsilon^{-2})$ is clearly visible from the bottom

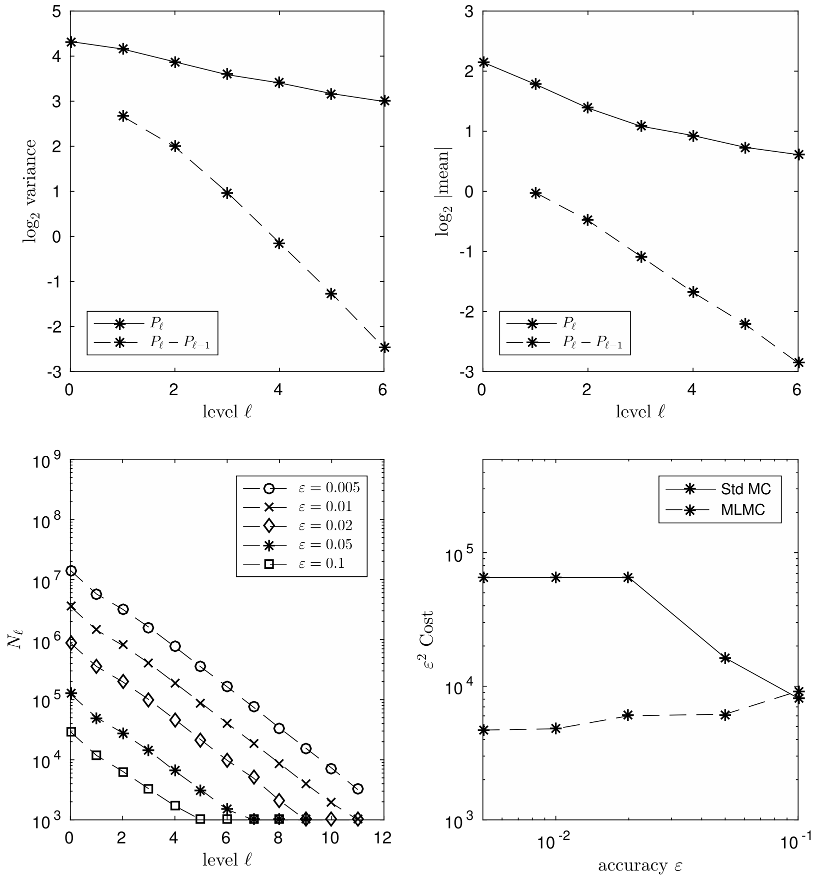

Fig. 6 Lookback option with spectrally negative α-stable model

right complexity plots, but there are still significant savings compared to the standard MC computations.

6.4 Summary and discussion

Table 1 summarises the convergence rates for the weak error $\mathbb{E}[\widehat{P}_\ell - P]$ and the MLMC variance $V_\ell = \mathbb{V}[\widehat{P}_\ell - \widehat{P}_{\ell-1}]$ given by Propositions 5.1, 5.2, 5.5, and the empirical convergence rates observed in the numerical experiments.

In general, the agreement between the analysis and the numerical rates of convergence is quite good, suggesting that in most cases the analysis may be sharp. The most obvious gap between the two is with the weak order of convergence for the Asian option with all three models; the analysis proves an $\mathcal{O}(h)$ bound, whereas the numerical results suggest it is actually $\mathcal{O}(h^2)$. The numerical results are perhaps not surprising as $\mathcal{O}(h^2)$ is the order of convergence of trapezoidal integration of a smooth function, and therefore it is the order one would expect if the payoff was simply a multiple of $\overline{S}$.

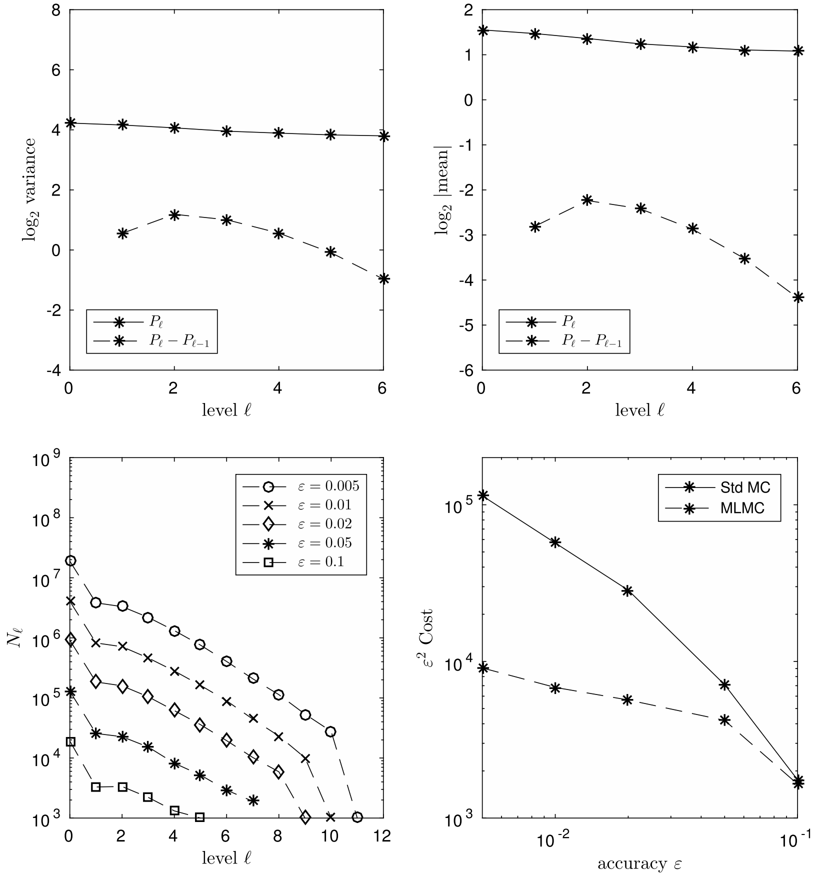

Fig. 7 Barrier option in variance gamma model

7 Proofs

7.1 Proof of Proposition 5.1

We decompose the difference between the true value and approximation into parts which we can bound separately by writing

$$
\begin{aligned}
&|\overline{S} - \widehat{\overline{S}}| \\
&= S_0\, T^{-1} \left| \int_0^T \exp(X_t)\mathrm{d}t - \sum_{j=0}^{n-1} \frac{1}{2} h \big(\exp(X_{jh}) + \exp(X_{(j+1)h})\big) \right| \\
&= S_0\, T^{-1} \left| \sum_{j=0}^{n-1} \exp(X_{jh}) \int_{jh}^{(j+1)h} \big(\exp(X_t - X_{jh}) - 1\big)\mathrm{d}t - \frac{1}{2} h \exp(X_T) + \frac{1}{2} h \right|.
\end{aligned}
$$

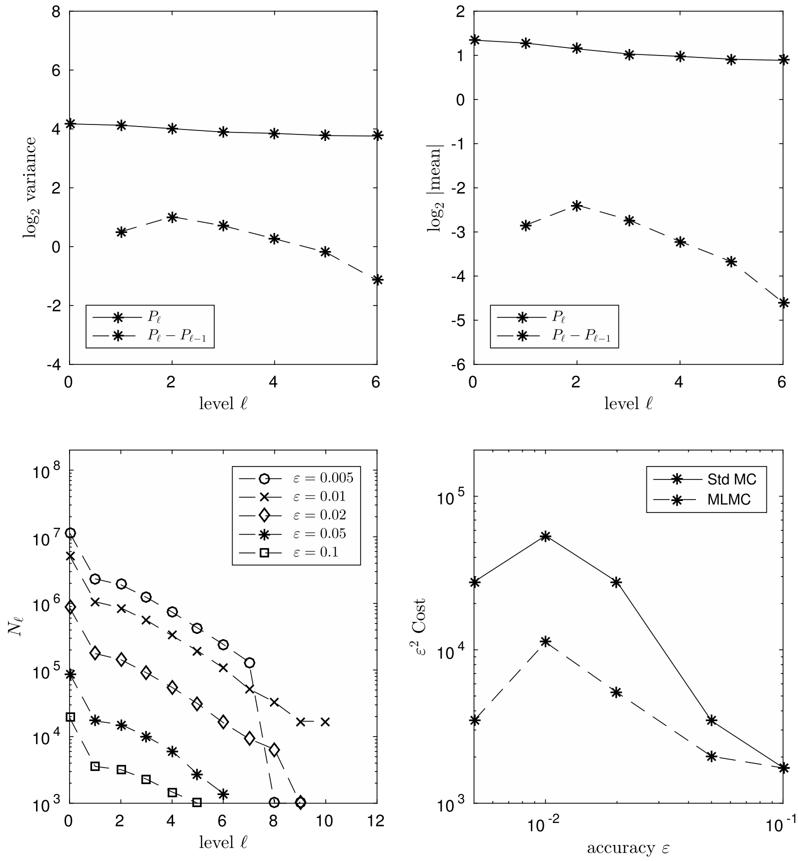

Fig. 8 Barrier option in normal inverse Gaussian model

If we define

$$b_j = \exp(X_{jh}),$$

$$I_j = \int_{jh}^{(j+1)h} \big(\exp(X_t - X_{jh}) - 1\big)\,\mathrm{d}t,$$

$$R_A = -\frac{1}{2} h \exp(X_T) + \frac{1}{2} h,$$

then

$$\mathbb{E}\big[(\widehat{\overline{S}} - \overline{S})^2\big] = T^{-2} S_0^2\, \mathbb{E}\Bigg[\Bigg|\sum_{j=0}^{n-1} b_j I_j + R_A\Bigg|^2\Bigg]$$

$$\leq 2T^{-2} S_0^2 \Bigg(\mathbb{E}\Bigg[\Bigg|\sum_{j=0}^{n-1} b_j I_j\Bigg|^2\Bigg] + \mathbb{E}[R_A^2]\Bigg).$$

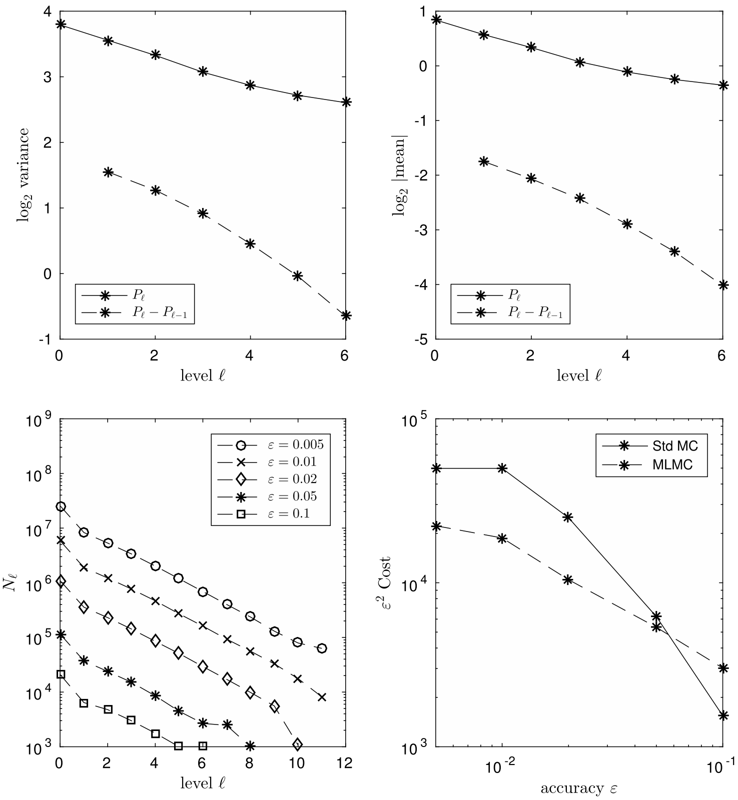

Fig. 9 Barrier option in spectrally negative α-stable model

We have $\mathbb{E}[R_A^2] = \mathcal{O}(h^2)$, and due to the independence of b_j and I_j, we obtain

$$\begin{aligned}\mathbb{E}\left[\left|\sum_{j=0}^{n-1} b_j I_j\right|^2\right] &= \mathbb{E}\left[\sum_{j=0}^{n-1} b_j^2 I_j^2 + 2\sum_{m=1}^{n-1}\sum_{j=0}^{m-1} b_m I_m b_j I_j\right] \\ &= \sum_{j=0}^{n-1} \mathbb{E}[b_j^2]\mathbb{E}[I_j^2] + 2\sum_{m=1}^{n-1}\sum_{j=0}^{m-1} \mathbb{E}[b_m I_m b_j I_j]. \qquad (7.1)\end{aligned}$$

Defining $A = 2m + \int (e^{2z} - 1 - 2z\mathbf{1}_{\{|z|<1\}})\nu(\mathrm{d}z)$, we have $\mathbb{E}[b_j^2] = e^{Ajh}$. Furthermore, by the Cauchy–Schwarz inequality,

$$\begin{aligned}\mathbb{E}[I_j^2] &\le h\,\mathbb{E}\left[\int_{jh}^{(j+1)h} \big(\exp(X_t - X_{jh}) - 1\big)^2 \mathrm{d}t\right] \\ &= h\int_0^h \mathbb{E}\left[\big(\exp(X_t) - 1\big)^2\right]\mathrm{d}t \\ &= h\left(\frac{1}{A}(e^{Ah} - 1 - Ah) - 2\frac{1}{r}(e^{rh} - 1 - rh)\right).\end{aligned}$$

Table 1 Convergence rates of weak error and variance V_ℓ for VG, NIG and α-stable processes; δ can be any small positive constant. The numerical values are estimates based on the numerical experiments

Option	VG			
	numerical		analysis	
	weak	var	weak	var
Asian	$\mathcal{O}(h^2)$	$\mathcal{O}(h^2)$	$\mathcal{O}(h)$	$\mathcal{O}(h^2)$
lookback	$\mathcal{O}(h)$	$\mathcal{O}(h^{1.2})$	$\mathcal{O}(h\|\log h\|)$	$\mathcal{O}(h)$
barrier	$\mathcal{O}(h^{0.8})$	$\mathcal{O}(h^{0.9})$	$o(h^{1-\delta})$	$o(h^{1-\delta})$

Option	NIG			
	numerical		analysis	
	weak	var	weak	var
Asian	$\mathcal{O}(h^2)$	$\mathcal{O}(h^2)$	$\mathcal{O}(h)$	$\mathcal{O}(h^2)$
lookback	$\mathcal{O}(h^{0.8})$	$\mathcal{O}(h^{1.2})$	$o(h^{1-\delta})$	$\mathcal{O}(h\|\log h\|)$
barrier	$\mathcal{O}(h^{0.4})$	$\mathcal{O}(h^{0.5})$	$o(h^{0.5-\delta})$	$o(h^{0.5-\delta})$

Option	spectrally negative α-stable with $\alpha > 1$			
	numerical for $\alpha = 1.5597$		analysis	
	weak	var	weak	var
Asian	$\mathcal{O}(h^2)$	$\mathcal{O}(h^2)$	$\mathcal{O}(h)$	$\mathcal{O}(h^2)$
lookback	$\mathcal{O}(h^{0.6})$	$\mathcal{O}(h^{1.6})$	$o(h^{1/\alpha-\delta})$	$o(h^{2/\alpha-\delta})$
barrier	$\mathcal{O}(h^{0.5})$	$\mathcal{O}(h^{0.6})$	$o(h^{1/\alpha-\delta})$	$o(h^{1/\alpha-\delta})$

Note that $1+x < e^x < 1+x+x^2$ for $0<x<1$, and therefore for $h<1/A$, we have $\mathbb{E}[I_j^2] < Ah^3$ and hence

$$\sum_{j=0}^{n-1} \mathbb{E}[b_j^2]\mathbb{E}[I_j^2] < A\,h^3 \sum_{j=0}^{n-1} e^{Ajh} = A\,\frac{e^{AT}-1}{e^{Ah}-1}\,h^3 < (e^{AT}-1)\,h^2.$$

Now we calculate the second term in (7.1). Note that for $m>j$, I_m is independent of $b_m b_j I_j$, and b_m/b_{j+1} is independent of $b_{j+1}b_j I_j$, so

$$\sum_{m=1}^{n-1}\sum_{j=0}^{m-1} \mathbb{E}[b_m I_m b_j I_j] = \sum_{m=1}^{n-1} \mathbb{E}[I_m] \sum_{j=0}^{m-1} \mathbb{E}[b_m/b_{j+1}]\mathbb{E}[b_{j+1} b_j I_j].$$

Firstly, for $h<1/r$,

$$\mathbb{E}[I_m] = \int_0^h (e^{rt}-1)\,\mathrm{d}t = r^{-1}(e^{rh}-1-rh)\; < \; r\,h^2.$$

Moreover, we have $\mathbb{E}[b_m/b_{j+1}] = e^{r(m-j-1)h}$ and

$$\begin{aligned}
\mathbb{E}[b_{j+1}b_j I_j] &= \mathbb{E}\bigg[\exp(2X_{jh})\exp(X_{(j+1)h}-X_{jh}) \\
&\qquad \times \int_{jh}^{(j+1)h}\big(\exp(X_t-X_{jh})-1\big)\,\mathrm{d}t\bigg] \\
&= \mathbb{E}[\exp(2X_{jh})]\,\mathbb{E}\bigg[\exp(X_h)\int_0^h\big(\exp(X_t)-1\big)\,\mathrm{d}t\bigg] \\
&= e^{Ajh}\int_0^h\big(\mathbb{E}[\exp(X_h-X_t)]\mathbb{E}[\exp(2X_t)]-\mathbb{E}[\exp(X_h)]\big)\,\mathrm{d}t \\
&= e^{Ajh}\int_0^h(e^{r(h-t)}e^{At}-e^{rh})\,\mathrm{d}t \\
&= e^{Ajh}e^{rh}\,\frac{e^{(A-r)h}-1-(A-r)h}{A-r}.
\end{aligned}$$

Thus, for $h<1/(A-r)$,

$$\begin{aligned}
\sum_{m=1}^{n-1}\sum_{j=0}^{m-1}\mathbb{E}[b_m/b_{j+1}]\mathbb{E}[b_{j+1}b_jI_j] &= \frac{e^{(A-r)h}-1-(A-r)}{A-r}\sum_{m=1}^{n-1}\sum_{j=0}^{m-1}e^{r(m-j)h}e^{Ajh} \\
&= \frac{e^{(A-r)h}-1-(A-r)h}{(A-r)\,(e^{(A-r)h}-1)}\sum_{m=1}^{n-1}(e^{Amh}-e^{rmh}) \\
&< h\,\frac{e^{AT}-1}{e^{Ah}-1} \\
&< A^{-1}(e^{AT}-1).
\end{aligned}$$

Hence,

$$\mathbb{E}\bigg[\sum_{m=1}^{n-1}\sum_{j=0}^{m-1}b_mI_mb_jI_j\bigg] = \sum_{m=1}^{n-1}\mathbb{E}[I_m]\sum_{j=0}^{m-1}\mathbb{E}[b_m/b_{j+1}]\mathbb{E}[b_{j+1}b_jI_j] = \mathcal{O}(h^2),$$

and we can therefore conclude that $\mathbb{E}[(\widehat{\overline{S}}-\overline{S})^2] = \mathcal{O}(h^2)$. □

7.2 Lévy process decomposition

The proofs rely on a decomposition of the Lévy process into a combination of a finite activity pure jump part, a drift part, and a residual part consisting of very small jumps.

Let X be an $(m,0,\nu)$-Lévy process, i.e.,

$$X_t = mt + \int_0^t\int_{\{|z|\ge 1\}} z\,J(\mathrm{d}z,\mathrm{d}s) + \int_0^t\int_{\{|z|<1\}} z\big(J(\mathrm{d}z,\mathrm{d}s)-\nu(\mathrm{d}z)\mathrm{d}s\big).$$

The finite activity jump part is defined by

$$X_t^\varepsilon = \int_0^t \int_{\{\varepsilon<|z|\}} z\ J(\mathrm{d}z, \mathrm{d}s) = \sum_{i=1}^{N_t} Y_i,$$

which is the compound Poisson process truncating the jumps of X smaller than ε, which is assumed to satisfy $0 < \varepsilon < 1$. The intensity of (N_t) and the c.d.f. of Y_i are

$$\lambda_\varepsilon = \int_{\{\varepsilon<|z|\}} \nu(\mathrm{d}z),$$

$$\mathbb{P}\big[Y_i < y\big] = \lambda_\varepsilon^{-1} \int_{\{z<y\}} \mathbf{1}_{\{\varepsilon<|z|\}} \nu(\mathrm{d}z).$$

The drift rate for the drift term is defined to be

$$\mu_\varepsilon = m - \int_{\{\varepsilon<|z|<1\}} z\ \nu(\mathrm{d}z),$$

so that the residual term is then a martingale, given by

$$R_t^\varepsilon := \int_0^t \int_{\{|z|\le\varepsilon\}} z\big(J(\mathrm{d}z, \mathrm{d}s) - \nu(\mathrm{d}z)\mathrm{d}s\big).$$

We define

$$\sigma_\varepsilon^2 = \int_{\{|z|\le\varepsilon\}} z^2 \nu(\mathrm{d}z), \tag{7.2}$$

so that $\mathbb{V}[R_t^\varepsilon] = \sigma_\varepsilon^2 t$. The three quantities μ_ε, λ_ε and σ_ε all play a major role in the subsequent numerical analysis.

We bound D_n by the difference between continuous maxima and 2-point maxima over all timesteps via

$$D_n = \sup_{0\le t\le 1} X_t - \max_{i=0,1,\dots,n} X_{\frac{i}{n}} \ \le \ \max_{i=0,\dots,n-1} D_n^{(i)},$$

where the random variables

$$D_n^{(i)} = \sup_{[\frac{i}{n},\frac{i+1}{n}]} X_t - \max\big(X_{\frac{i+1}{n}}, X_{\frac{i}{n}}\big)$$

are independent and identically distributed. If we now define

$$\begin{aligned} \Delta^{(i)} X_t &= X_{\frac{i}{n}+t} - X_{\frac{i}{n}}, \ \Delta^{(i)} X_t^\varepsilon = X^\varepsilon_{\frac{i}{n}+t} - X^\varepsilon_{\frac{i}{n}}, \\ \Delta^{(i)} t &= t - \tfrac{i}{n}, \qquad \Delta^{(i)} R_t^\varepsilon = R^\varepsilon_{\frac{i}{n}+t} - R^\varepsilon_{\frac{i}{n}}, \end{aligned}$$

then

$$
\begin{aligned}
D_n^{(i)} &= \sup_{[0,\frac{1}{n}]} \Delta^{(i)} X_t - (\Delta^{(i)} X_{\frac{1}{n}})^+ \\
&= \sup_{[0,\frac{1}{n}]} (\Delta^{(i)} X_t^\varepsilon + \Delta^{(i)} R_t^\varepsilon + \mu_\varepsilon \Delta^{(i)} t) - \left(\Delta^{(i)} X_{\frac{1}{n}}^\varepsilon + \Delta^{(i)} R_{\frac{1}{n}}^\varepsilon + \mu_\varepsilon \frac{1}{n} \right)^+ \\
&\le \sup_{[0,\frac{1}{n}]} (\Delta^{(i)} X_t^\varepsilon + \Delta^{(i)} R_t^\varepsilon) - (\Delta^{(i)} X_{\frac{1}{n}}^\varepsilon + \Delta^{(i)} R_{\frac{1}{n}}^\varepsilon)^+ + \frac{|\mu_\varepsilon|}{n} \\
&\le \sup_{[0,\frac{1}{n}]} \Delta^{(i)} X_t^\varepsilon - (\Delta^{(i)} X_{\frac{1}{n}}^\varepsilon)^+ + \frac{|\mu_\varepsilon|}{n} + \sup_{[0,\frac{1}{n}]} \Delta^{(i)} R_t^\varepsilon + (-\Delta^{(i)} R_{\frac{1}{n}}^\varepsilon)^+ \\
&\le \sup_{[0,\frac{1}{n}]} \Delta^{(i)} X_t^\varepsilon - (\Delta^{(i)} X_{\frac{1}{n}}^\varepsilon)^+ + \frac{|\mu_\varepsilon|}{n} + 2 \sup_{[0,\frac{1}{n}]} |\Delta^{(i)} R_t^\varepsilon|, \qquad (7.3)
\end{aligned}
$$

where we use $(a+b)^+ \le a^+ + b^+$ with $a = \Delta^{(i)} X_{\frac{1}{n}}^\varepsilon + \Delta^{(i)} R_{\frac{1}{n}}^\varepsilon + \mu_\varepsilon \frac{1}{n}$, $b = -\mu_\varepsilon \frac{1}{n}$ in the first inequality, and $a = \Delta^{(i)} X_{\frac{1}{n}}^\varepsilon + \Delta^{(i)} R_{\frac{1}{n}}^\varepsilon$, $b = -\Delta^{(i)} R_{\frac{1}{n}}^\varepsilon$ in the second.

Let $Z_n^{(i)} := \sup_{[0,\frac{1}{n}]} \Delta^{(i)} X_t^\varepsilon - (\Delta X_{\frac{1}{n}}^\varepsilon)^+$ and $S_n^{(i)} := \sup_{[0,\frac{1}{n}]} |\Delta^{(i)} R_t^\varepsilon|$. Then for $p \ge 1$, Jensen's inequality gives

$$
\begin{aligned}
\mathbb{E}[D_n^p] &\le \mathbb{E}\left[\max_{0 \le i < n} (Z_n^{(i)} + \frac{|\mu_\varepsilon|}{n} + 2\, S_n^{(i)})^p \right] \\
&\le 3^{p-1} \mathbb{E}\left[\max_{0 \le i < n} (Z_n^{(i)})^p + (|\mu_\varepsilon|/n)^p + 2^p \max_{0 \le i < n} (S_n^{(i)})^p \right] \\
&\le 3^{p-1} n\, \mathbb{E}\left[\Big(\sup_{[0,\frac{1}{n}]} X_t^\varepsilon - (X_{\frac{1}{n}}^\varepsilon)^+ \Big)^p \right] + 3^{p-1} (|\mu_\varepsilon|/n)^p \\
&\quad + 3^{p-1} 2^p\, \mathbb{E}\left[\max_{0 \le i < n} (S_n^{(i)})^p \right], \qquad (7.4)
\end{aligned}
$$

where in the final step we have used the fact that all of the $\Delta^{(i)} X_t^\varepsilon$ have the same distribution as X_t^ε.

The task now is to bound the first and third terms in the final line of (7.4).

7.3 Bounding moments of $\sup_{[0,\frac{1}{n}]} X_t^\varepsilon - (X_{\frac{1}{n}}^\varepsilon)^+$

Theorem 7.1 *Let X be a scalar Lévy process with triple $(m, 0, \nu)$, and let X_t^ε, μ_ε, λ_ε and σ_ε be defined as in Sect. 7.2. Then provided $\lambda_\varepsilon \le n$, for any $p > 1$, there exists*

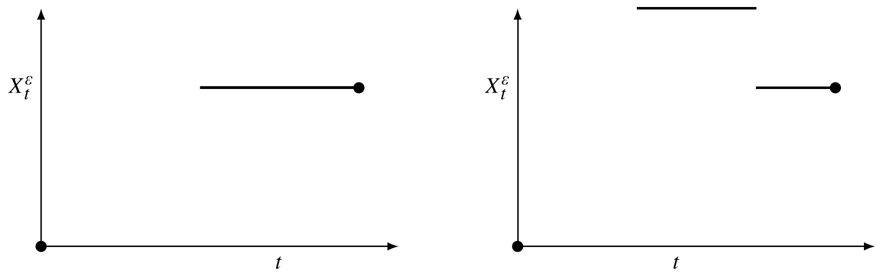

Fig. 10 Behaviour of (X_t^ε) in the case of one or two jumps in the interval $[0, \frac{1}{n}]$

a constant K_p such that

$$\mathbb{E}\Big[\Big(\sup_{[0,\frac{1}{n}]} X_t^\varepsilon - (X_{\frac{1}{n}}^\varepsilon)^+\Big)^p\Big] \leq K_p\Big(\varepsilon^p + \frac{L_\varepsilon(p)}{\lambda_\varepsilon^2}\Big)\frac{\lambda_\varepsilon^2}{n^2}, \tag{7.5}$$

where $L_\varepsilon(p) = p\int_{\{x>\varepsilon\}} x^{p-1}\lambda_x^2\,\mathrm{d}x$ is a function depending on the Lévy measure $\nu(x)$.

Proof Let

$$Z = \sup_{[0,\frac{1}{n}]} X_t^\varepsilon - (X_{\frac{1}{n}}^\varepsilon)^+.$$

We determine an upper bound on $\mathbb{E}[Z^p]$ by analysing the jump behaviour of the finite activity process (X_t^ε) in a single interval $[0, \frac{1}{n}]$.

Let N be the number of jumps. If $N \leq 1$, then $Z = 0$, while if $N = 2$, then $Z \leq \min(|Y_1|, |Y_2|)$. This can be seen from the behaviour of (X_t^ε) in the different scenarios illustrated in Fig. 10. More generally, if $N = k, k \geq 2$, then

$$\begin{aligned} Z > x &\Longrightarrow \exists\ 1 \leq j \leq k-1 \text{ such that } \Big|\sum_{\ell=1}^{j} Y_\ell\Big| > x,\ \Big|\sum_{\ell=j+1}^{k} Y_\ell\Big| > x \\ &\Longrightarrow \exists\ j_1, j_2 \text{ such that } |Y_{j_1}| > \frac{x}{k-1},\quad |Y_{j_2}| > \frac{x}{k-1}. \end{aligned}$$

Since

$$\begin{aligned} &\mathbb{P}\Big[\exists j_1, j_2 \text{ such that } |Y_{j_1}| > \frac{x}{k-1},\ |Y_{j_2}| > \frac{x}{k-1}\Big] \\ &\leq \sum_{(j_1,j_2)} \mathbb{P}\Big[|Y_{j_1}| > \frac{x}{k-1},\ |Y_{j_2}| > \frac{x}{k-1}\Big] \\ &= \frac{k(k-1)}{2}\mathbb{P}\Big[|Y_1| > \frac{x}{k-1}\Big]^2, \end{aligned}$$

it follows that

$$\begin{aligned}\mathbb{E}[Z^p \mid N=k] &= p\int x^{p-1}\mathbb{P}[Z>x \mid N=k]\mathrm{d}x\\ &\le \frac{k\,(k-1)}{2}\,p\int x^{p-1}\mathbb{P}\left[|Y_1|>\frac{x}{k-1}\right]^2\mathrm{d}x\\ &= \frac{k\,(k-1)}{2}\,\frac{p}{\lambda_\varepsilon^2}\int x^{p-1}\left(\int_{\{|z|>x/(k-1)\}}\mathbf{1}_{\{\varepsilon<|z|\}}\nu(\mathrm{d}z)\right)^2\mathrm{d}x\\ &= \frac{k\,(k-1)^{p+1}}{2}\,\frac{p}{\lambda_\varepsilon^2}\int x^{p-1}\left(\int_{\{|z|>x\}}\mathbf{1}_{\{\varepsilon<|z|\}}\nu(\mathrm{d}z)\right)^2\mathrm{d}x\\ &=: d_{k,p}\left(\varepsilon^p+\frac{L_\varepsilon(p)}{\lambda_\varepsilon^2}\right),\end{aligned}$$

where $d_{k,p}=\frac{1}{2}k\,(k-1)^{p+1}$. We then have

$$\begin{aligned}\mathbb{E}[Z^p] &= \sum_{k=2}^{\infty}\mathbb{E}[Z^p \mid N=k]\,\mathbb{P}[N=k]\\ &\le \left(\varepsilon^p+\frac{L_\varepsilon(p)}{\lambda_\varepsilon^2}\right)\,\exp\left(-\frac{\lambda_\varepsilon}{n}\right)\,\sum_{k=2}^{\infty}d_{k,p}\left(\frac{\lambda_\varepsilon}{n}\right)^k\frac{1}{k!}.\end{aligned}$$

For $k_p=\lceil p\rceil+2$, there exists C_p such that for any $k\ge k_p$, $d_{k,p}\le C_p\frac{k!}{(k-k_p)!}$, so

$$\begin{aligned}\sum_{k=2}^{\infty}d_{k,p}\left(\frac{\lambda_\varepsilon}{n}\right)^k\frac{1}{k!} &\le \sum_{k=2}^{k_p-1}d_{k,p}\left(\frac{\lambda_\varepsilon}{n}\right)^k\frac{1}{k!}+C_p\sum_{k=k_p}^{\infty}\left(\frac{\lambda_\varepsilon}{n}\right)^k\frac{1}{(k-k_p)!}\\ &\le \sum_{k=2}^{k_p-1}d_{k,p}\left(\frac{\lambda_\varepsilon}{n}\right)^k\frac{1}{k!}+C_p\left(\frac{\lambda_\varepsilon}{n}\right)_p^k\exp\left(\frac{\lambda_\varepsilon}{n}\right)\\ &\le K_p\left(\frac{\lambda_\varepsilon}{n}\right)^2\end{aligned}$$

for some constant K_p, where the last step uses the fact that $\lambda_\varepsilon\le n$. Therefore, we obtain the final result that

$$\mathbb{E}[Z^p]\le K_p\left(\varepsilon^p+\frac{L_\varepsilon(p)}{\lambda_\varepsilon^2}\right)\frac{\lambda_\varepsilon^2}{n^2}.$$

□

7.4 Bounding moments of $\sup_{[0,T]}|R_t^\varepsilon|$

Proposition 7.1 *Let X be a scalar Lévy process with triple $(m,0,\nu)$ and let R_t^ε, μ_ε, λ_ε and σ_ε be defined as in Sect. 7.2. Then (R_t^ε) satisfies*

$$\mathbb{E}\left[\sup_{[0,T]}|R_t^\varepsilon|^p\right]\le\begin{cases}K_p(T^{p/2}\sigma_\varepsilon^p+T\int_{\{|z|\le\varepsilon\}}|z|^p\nu(\mathrm{d}x)), & p>2,\\ K_p\,T^{p/2}\sigma_\varepsilon^p, & 1\le p\le 2,\end{cases}$$

where K_p is a constant depending on p.

Proof For any $1 \leq p \leq 2$, by Jensen's inequality and the Doob inequality,

$$\mathbb{E}\left[\sup_{0\leq t\leq T} |R_t^\varepsilon|^p\right] \leq \mathbb{E}\left[\sup_{0\leq t\leq T} |R_t^\varepsilon|^2\right]^{p/2} \leq 2^p\, \mathbb{E}[|R_T^\varepsilon|^2]^{p/2} = 2^p\, T^{p/2}\sigma_\varepsilon^p.$$

For any $p > 2$, the Burkholder–Davis–Gundy inequality gives

$$\mathbb{E}\left[\sup_{0\leq t\leq 1} |R_t^\varepsilon|^p\right] \leq \mathbb{E}\big[[R^\varepsilon]_1^{p/2}\big],$$

where $\left[R^\varepsilon\right]_t$ is the quadratic variation of R_t^ε. We can use the method in the proof of [24, Theorem V.66] to get

$$\begin{aligned}\mathbb{E}\big[[R^\varepsilon]_1^{p/2}\big] &\leq K_p\left[\left(\int_{\{|z|\leq\varepsilon\}} z^2\nu(\mathrm{d}z)\right)^{p/2} + \int_{\{|z|\leq\varepsilon\}} |z|^p\nu(\mathrm{d}z)\right] \\ &= K_p\left(\sigma_\varepsilon^p + \int_{\{|z|\leq\varepsilon\}} |z|^p\nu(\mathrm{d}z)\right),\end{aligned}$$

where K_p is a constant depending on p.

To extend this result to an arbitrary time interval $[0, T]$, we use a change of time coordinate $t' = t/T$ with associated changed Lévy measure $\nu'(\mathrm{d}z) = T\,\nu(\mathrm{d}z)$ to obtain

$$\mathbb{E}\left[\sup_{[0,T]} |R_t^\varepsilon|^p\right] \leq K_p\left[T^{p/2}\sigma_\varepsilon^{p/2} + T\int_{\{|z|\leq\varepsilon\}} |z|^p\nu(\mathrm{d}z)\right]. \qquad \square$$

7.5 Bounding moments of $\max_{0\leq i<n} S_n^{(i)}$

Proposition 7.2 *Let X be a scalar pure jump Lévy process, with Lévy measure $\nu(x)$ which satisfies*

$$C_2\,|x|^{-1-\alpha} \leq \nu(x) \leq C_1\,|x|^{-1-\alpha} \quad \textit{for}\, |x| \leq 1,$$

for constants $C_1, C_2 > 0$ and $0 \leq \alpha < 2$. If $S_n^{(i)}$ is defined as in Sect. 7.2 and $\lambda_\varepsilon \leq n$, then for $p \geq 1$ and arbitrary $\delta > 0$, there exists a constant $C_{p,\delta}$, which does not depend on n, ε, such that

$$\mathbb{E}\left[\left(\max_{0\leq i<n} S_n^{(i)}\right)^p\right] \leq C_{p,\delta}\, \varepsilon^{p-\delta}.$$

In the particular case of $\alpha = 0$, such a bound holds with $\delta = 0$.

Proof By Proposition 7.1, for $q > 2$,

$$\mathbb{E}\left[\left(\max_{0\leq i<n} S_n^{(i)}\right)^q\right] \leq n\,\mathbb{E}\left[\sup_{[0,\frac{1}{n}]} |R_t^\varepsilon|^q\right] \leq K_q\left(n^{1-q/2}\sigma_\varepsilon^q + \int_{\{|z|\leq\varepsilon\}} |z|^q\,\nu(\mathrm{d}x)\right).$$

Recalling the definition of σ_ε in (7.2), due to the assumption on $\nu(x)$, we have

$$\sigma_\varepsilon^q \le \left(\frac{2C_1}{2-\alpha}\right)^{q/2} \varepsilon^{q-q\alpha/2}, \qquad \int_{\{|z|\le\varepsilon\}} |z|^q \, \nu(\mathrm{d}x) \le \frac{2C_1}{q-\alpha}\, \varepsilon^{q-\alpha}.$$

Given $p \ge 1$, for any $q > \max(2, p)$, Jensen's inequality gives

$$\begin{aligned}\mathbb{E}\left[\left(\max_{0\le i<n} S_n^{(i)}\right)^p\right] &\le \mathbb{E}\left[\left(\max_{0\le i<n} S_n^{(i)}\right)^q\right]^{p/q} \\ &\le K_q^{p/q}\left[\left(\frac{2C_1}{2-\alpha}\right)^{q/2}\left(\frac{\varepsilon^{-\alpha}}{n}\right)^{q/2-1} + \frac{2C_1}{q-\alpha}\right]^{p/q} \varepsilon^{p-\alpha p/q}.\end{aligned}$$

If $\alpha = 0$, then the desired bound is obtained immediately. On the other hand, if $0 < \alpha < 2$, then

$$\lambda_\varepsilon \ \ge\ C_2 \int_{\{\varepsilon<|z|<1\}} \frac{1}{|z|^{\alpha+1}}\, \mathrm{d}z = \frac{2\,C_2}{\alpha}(\varepsilon^{-\alpha} - 1).$$

Since $\lambda_\varepsilon \le n$, this implies that $\varepsilon^{-\alpha} \le \frac{K\alpha}{2C_2} n + 1$, and thus $\varepsilon^{-\alpha}/n$ is bounded. Hence there exists a constant C such that

$$\mathbb{E}\left[\left(\max_{0\le i<n} S_n^{(i)}\right)^p\right] \le C\, \varepsilon^{p-\alpha p/q},$$

and by choosing q large enough so that $\alpha p/q \le \delta$, we obtain the desired bound. $\square$

7.6 Proof of Theorem 4.2

Provided $\lambda_\varepsilon \le n$, by (7.4) and (7.5) we have

$$\mathbb{E}[D_n^p] \prec \underbrace{\mathbb{E}\left[\left(\max_{0\le i<n} S_n^{(i)}\right)^p\right]}_{1)} + \underbrace{\varepsilon^p \frac{\lambda_\varepsilon^2}{n}}_{2)} + \underbrace{\frac{L_\varepsilon(p)}{n}}_{3)} + \underbrace{\left(\frac{|\mu_\varepsilon|}{n}\right)^p}_{4)}, \tag{7.6}$$

where the notation $u \prec v$ means that there exists a constant $c > 0$ independent of n such that $u < cv$.

We can now bound each term, given the specification of the Lévy measure, and if we can choose appropriately how $\varepsilon \to 0$ as $n \to \infty$ so that the RHS of (7.6) is convergent, then the convergence rate of $\mathbb{E}[D_n^p]$ can be bounded.

For $0 < x < 1$,

$$\begin{aligned}\lambda_x &\le C_1 \int_{\{x<|z|<1\}} \frac{1}{|z|^{\alpha+1}}\, \mathrm{d}z + \int_{\{1<|z|\}} \exp(-C_3\,|z|)\, \mathrm{d}z \\ &\le \begin{cases} 2C_1 \log\frac{1}{x} + \ell_1, & \alpha = 0, \\ \ell_2\, x^{-\alpha}, & 0 < \alpha < 2, \end{cases}\end{aligned} \tag{7.7}$$

where ℓ_1, ℓ_2 are constants with $\ell_2 \geq 2\, C_3^{-1}$, while for $x \geq 1$,

$$\lambda_x \leq \int_{\{x<|z|\}} \exp\left(-C_3\, |z|\right)\, \mathrm{d}z = 2\, C_3^{-1} \exp(-C_3\, x).$$

If $\alpha > 0$, then

$$\begin{aligned} L_\varepsilon(p) &= p \int_{\{x>\varepsilon\}} x^{p-1} \lambda_x^2\, \mathrm{d}x \\ &\leq \ell_2^2\, p \int_{\{x>\varepsilon\}} x^{p-1} \big(\mathbf{1}_{\{x<1\}} x^{-2\alpha} + \mathbf{1}_{\{x>1\}} \exp(-2C_3 x)\big) \mathrm{d}x \\ &\leq \begin{cases} \ell_3, & p > 2\alpha, \\ \ell_3 \log \frac{1}{\varepsilon} + \ell_4, & p = 2\alpha, \\ \ell_3 \varepsilon^{-2\alpha+p} + \ell_4, & p < 2\alpha, \end{cases} \end{aligned} \tag{7.8}$$

where ℓ_3 and ℓ_4 are additional constants. If $\alpha = 0$, it is easily verified that $L_\varepsilon(p)$ is bounded for $p \geq 1$, so (7.8) applies equally to this case.

Given $0 < \varepsilon < 1$, we have

$$|\mu_\varepsilon| = \left| m - \int_{\varepsilon<|z|<1} z\, \nu(\mathrm{d}z) \right| \leq \begin{cases} |m| + |C_1 - C_2| \, \frac{\varepsilon^{1-\alpha}-1}{\alpha-1}, & \alpha \neq 1, \\ |m| + |C_1 - C_2| \log \frac{1}{\varepsilon}, & \alpha = 1. \end{cases} \tag{7.9}$$

Subject to the condition that $\lambda_\varepsilon \leq n$, we now consider the terms in (7.6).

1. By Proposition 7.2,

$$\mathbb{E}\left[\left(\max_{0\leq i<n} S_n^{(i)} \right)^p \right] \prec \varepsilon^{p-\delta}, \quad \text{for any } \delta > 0.$$

2. By (7.7),

$$\varepsilon^p \frac{\lambda_\varepsilon^2}{n} \prec n^{-1} \times \begin{cases} \varepsilon^p \log \frac{1}{\varepsilon}, & \alpha = 0, \\ \varepsilon^{p-2\alpha}, & 0 < \alpha < 2. \end{cases}$$

3. By (7.8),

$$\frac{L_\varepsilon(p)}{n} \prec n^{-1} \times \begin{cases} 1, & p > 2\alpha, \\ \log \frac{1}{\varepsilon}, & p = 2\alpha, \\ \varepsilon^{-2\alpha+p}, & p < 2\alpha. \end{cases}$$

4. By (7.9),

$$\left(\frac{|\mu_\varepsilon|}{n} \right)^p \prec n^{-p} \times \begin{cases} 1 + |C_1 - C_2|^p \varepsilon^{p(1-\alpha)}, & \alpha > 1, \\ 1 + (|C_1 - C_2| \log \frac{1}{\varepsilon})^p, & \alpha = 1, \\ 1, & \alpha < 1. \end{cases}$$

In the following, we assume $C_1 \neq C_2$.

Suppose $p \geq 2\alpha$. If we choose $\varepsilon = C n^{-2/p}$, then $\lambda_\varepsilon \prec \varepsilon^{-\alpha} \prec n^{2\alpha/p}$, and the constant C can be taken to be sufficiently small so that $\lambda_\varepsilon \leq n$ for sufficiently large n. Taking $\delta < p/2$, we find that the dominant contribution to (7.6) comes from 3), giving the desired result that

$$\mathbb{E}[D_n^p] \prec \begin{cases} n^{-1}, & p > 2\alpha, \\ \log n/n, & p = 2\alpha. \end{cases}$$

If $1 \leq p < 2\alpha$, Hölder's inequality gives $\mathbb{E}[D_n^p] \leq \mathbb{E}[D_n^{2\alpha}]^{\frac{p}{2\alpha}} \prec (\log n/n)^{\frac{p}{2\alpha}}$. For a spectrally negative process, $\sup_{[0,\frac{1}{n}]} X_t^\varepsilon - (X_{\frac{1}{n}}^\varepsilon)^+ = 0$, since X does not have positive jumps, and hence

$$\mathbb{E}[D_n^p] \leq \mathbb{E}\left[\left(\max_{0 \leq i < n} S_n^{(i)}\right)^p\right] + \left(\frac{|\mu_\varepsilon|}{n}\right)^p.$$

We can take $\varepsilon = C n^{-1/\alpha}$ with the constant C again chosen so that $\lambda_\varepsilon \leq n$ for sufficiently large n. We then obtain

$$\mathbb{E}[D_n^p] \prec \begin{cases} n^{-p/\alpha+\delta}, & \alpha \geq 1, \\ n^{-p}, & \alpha < 1, \end{cases}$$

for any $\delta > 0$. □

7.7 Proof of Proposition 5.4

We decompose the term we want to bound into parts and then balance their asymptotic orders to get the desired result.

Note that $\mathbf{1}_{\{\widehat{m}_n < B\}} - \mathbf{1}_{\{m < B\}} = 1$ only if either m is close to the barrier or the difference between discretely and continuously monitored maximum $D_n = m - \widehat{m}_n$ is large. More precisely,

$$\left\{\mathbf{1}_{\{\widehat{m}_n < B\}} - \mathbf{1}_{\{m < B\}} = 1\right\} \subseteq F \cup G,$$

where $F := \{B \leq m \leq B + n^{-r}\}$ and $G := \{D_n > n^{-r}\}$ for an $r > 0$ to be determined. Hence

$$\mathbb{E}[\mathbf{1}_{\{\widehat{m}_n < B\}} - \mathbf{1}_{\{m < B\}}] \leq \mathbb{P}[F] + \mathbb{P}[G].$$

Due to the locally bounded density for m, $\mathbb{P}[F] = \mathcal{O}(n^{-r})$. If we denote

$$Z_n^{(i)} = \sup_{[0,\frac{1}{n}]} \Delta^{(i)} X_t^\varepsilon - (\Delta^{(i)} X_{\frac{1}{n}}^\varepsilon)^+,$$

where $\Delta^{(i)} X_t$ is defined as previously in Sect. 7.2, then (7.3) gives

$$D_n \leq \max_{0 \leq i < n} Z_n^{(i)} + \frac{|\mu_\varepsilon|}{n} + \max_{0 \leq i < n} S_n^{(i)}.$$

For $\alpha < 1$, μ_ε is bounded, so $|\mu_\varepsilon| \le \frac{1}{2} n^{1-r}$ for sufficiently large n. Hence,

$$\begin{aligned}\mathbb{P}[D_n > n^{-r}] &\le \mathbb{P}\left[\max_{0\le i<n} Z_n^{(i)} + \max_{0\le i<n} S_n^{(i)} > \frac{1}{2} n^{-r}\right] \\ &\le \mathbb{P}\left[\max_{0\le i<n} Z_n^{(i)} > \frac{1}{4} n^{-r}\right] + \mathbb{P}\left[\max_{0\le i<n} S_n^{(i)} > \frac{1}{4} n^{-r}\right].\end{aligned}$$

Now, $\max_{0\le i<n} Z_n^{(i)} > 0$ requires that there are at least two jumps in one of the n intervals. The probability of two jumps in one particular interval is

$$1 - \exp\left(-\frac{\lambda_\varepsilon}{n}\right)\left(1 + \frac{\lambda_\varepsilon}{n}\right) \prec \left(\frac{\lambda_\varepsilon}{n}\right)^2$$

if $\lambda_\varepsilon \le n$, and hence

$$\mathbb{P}\left[\max_{0\le i<n} Z_n^{(i)} > \frac{1}{4} n^{-r}\right] \prec \frac{\lambda_\varepsilon^2}{n}.$$

We use the Markov inequality for the remaining term. According to Proposition 7.1, $\mathbb{E}[\max_{0\le i<n} (S_n^{(i)})^p] \prec \varepsilon^{p-\delta}$ and so

$$\mathbb{P}\left[\max_{0\le i<n} S_n^{(i)} > \frac{1}{4} n^{-r}\right] \prec \mathbb{E}\left[\max_{0\le i<n} (S_n^{(i)})^p\right] \Big/ \left(\frac{1}{4} n^{-r}\right)^p \prec \varepsilon^{p-\delta} n^{rp}.$$

Combining these elements, provided $\lambda_\varepsilon \le n$, we have

$$\mathbb{E}[\mathbf{1}_{\{\widehat{m}_n < B\}} - \mathbf{1}_{\{m<B\}}] \prec n^{-r} + \varepsilon^{p-\delta} n^{rp} + \frac{\lambda_\varepsilon^2}{n}.$$

Equating the first two terms on the right-hand side gives $\varepsilon = n^{-r(1+p)/(p-\delta)}$.

If $\alpha = 0$, then $\lambda_\varepsilon \prec \log \frac{1}{\varepsilon} \prec \log n$, so $\lambda_\varepsilon = o(n)$ is satisfied. We also have $\frac{\lambda_\varepsilon^2}{n} \prec \frac{(\log n)^2}{n}$, and therefore for any $r < 1$, we have $\mathbb{E}[\mathbf{1}_{\{\widehat{m}_n < B\}} - \mathbf{1}_{\{m<B\}}] \prec n^{-r}$.

If $0 < \alpha < 2$, then $\lambda_\varepsilon \prec \varepsilon^{-\alpha} \prec n^{r\alpha(1+p)/(p-\delta)}$, and hence we obtain that $\frac{\lambda_\varepsilon^2}{n} \prec n^{-1+2r\alpha(1+p)/(p-\delta)}$. Balancing n^{-r} and $n^{-1+2r\alpha(1+p)/(p-\delta)}$ gives $\lambda_\varepsilon = o(n)$ and

$$r = \left(1 + 2\alpha\, \frac{1+p}{p-\delta}\right)^{-1}. \tag{7.10}$$

Since $r \to \frac{1}{1+2\alpha}$ as $\delta \to 0$, and $p \to \infty$, for any fixed value of $r < \frac{1}{1+2\alpha}$, it is possible to choose appropriate values of p and δ to satisfy (7.10) and thereby conclude that $\mathbb{E}[\mathbf{1}_{\{\widehat{m}_n < B\}} - \mathbf{1}_{\{m<B\}}] \prec n^{-r}$. □

Acknowledgements This work was supported by the China Scholarship Council and the Oxford–Man Institute of Quantitative Finance. We should like to thank Ben Hambly, Andreas Kyprianou, Loïc Chaumont, Jacek Małecki and Jose Blanchet for their helpful comments.

References

1. Asmussen, S., Glynn, P., Pitman, J.: Discretization error in simulation of one-dimensional reflecting Brownian motion. Ann. Appl. Probab. **5**, 875–896 (1995)
2. Avikainen, R.: Convergence rates for approximations of functionals of SDEs. Finance Stoch. **13**, 381–401 (2009)
3. Blanchet, J.: Personal communication (2015)
4. Carr, P., Wu, L.: The finite moment log stable process and option pricing. J. Finance **58**, 753–778 (2003)
5. Chambers, J.M., Mallows, C.L., Stuck, B.: A method for simulating stable random variables. J. Am. Stat. Assoc. **71**, 340–344 (1976)
6. Chaumont, L.: On the law of the supremum of Lévy processes. Ann. Probab. **41**, 1191–1217 (2013)
7. Chaumont, L., Małecki, J.: The asymptotic behavior of the density of the supremum of Lévy processes. Ann. Inst. Henri Poincaré Probab. Stat. **52**, 1178–1195 (2016)
8. Chen, A.: Sampling Error of the Supremum of a Lévy Process. Ph.D. Thesis, University of Illinois at Urbana-Champaign (2011). Available online at https://www.ideals.illinois.edu/handle/2142/26321
9. Chen, A., Feng, L., Song, R.: On the monitoring error of the supremum of a normal jump diffusion process. J. Appl. Probab. **48**, 1021–1034 (2011)
10. Cont, R., Tankov, P.: Financial Modelling with Jump Processes. Chapman & Hall/CRC Press, London (2004)
11. Dereich, S.: Multilevel Monte Carlo algorithms for Lévy-driven SDEs with Gaussian correction. Ann. Appl. Probab. **21**, 283–311 (2011)
12. Dereich, S., Heidenreich, F.: A multilevel Monte Carlo algorithm for Lévy-driven stochastic differential equations. Stoch. Process. Appl. **121**, 1565–1587 (2011)
13. Dia, E., Lamberton, D.: Connecting discrete and continuous lookback or hindsight options in exponential Lévy models. Adv. Appl. Probab. **43**, 1136–1165 (2011)
14. Ferreiro-Castilla, A., Kyprianou, A.E., Scheichl, R., Suryanarayana, G.: Multilevel Monte Carlo simulation for Lévy processes based on the Wiener–Hopf factorisation. Stoch. Process. Appl. **124**, 985–1010 (2014)
15. Figueroa-López, J., Tankov, P.: Small-time asymptotics of stopped Lévy bridges and simulation schemes with controlled bias. Bernoulli **20**, 1126–1164 (2014)
16. Giles, M.B.: Improved multilevel Monte Carlo convergence using the Milstein scheme. In: Keller, A., et al. (eds.) Monte Carlo and Quasi-Monte Carlo Methods 2006, pp. 343–358. Springer, Berlin (2007)
17. Giles, M.B.: Multilevel Monte Carlo path simulation. Oper. Res. **56**, 607–617 (2008)
18. Giles, M.B.: Multilevel Monte Carlo methods. In: Dick, J., et al. (eds.) Monte Carlo and Quasi-Monte Carlo Methods 2012, pp. 83–103. Springer, Berlin (2012)
19. Giles, M.B.: Multilevel Monte Carlo methods. Acta Numer. **24**, 259–328 (2015)
20. Kuznetsov, A.: On extrema of stable processes. Ann. Probab. **39**, 1027–1060 (2011)
21. Kuznetsov, A., Kyprianou, A., Pardo, J., Van Schaik, K.: A Wiener–Hopf Monte Carlo simulation technique for Lévy processes. Ann. Appl. Probab. **21**, 2171–2190 (2011)
22. Kwasnicki, M., Małecki, J., Ryznar, M.: Suprema of Lévy processes. Ann. Probab. **41**, 2047–2065 (2013)
23. Kyprianou, A.: Introductory Lectures on Fluctuations of Lévy Processes with Applications. Springer, Berlin (2006)
24. Protter, P.: Stochastic Integration and Differential Equations, 2nd edn. Springer, Berlin (2004)
25. Schoutens, W.: Lévy Processes in Finance: Pricing Financial Derivatives. Wiley/Blackwell, New Jersey (2003)
26. Spitzer, F.: A combinatorial lemma and its application to probability theory. Trans. Am. Math. Soc. **82**, 323–339 (1956)