A MONTE CARLO ALGORITHM FOR THE EXTREMA OF TEMPERED STABLE PROCESSES

JORGE IGNACIO GONZÁLEZ CÁZARES 🆔* ** AND
ALEKSANDAR MIJATOVIĆ, 🆔* *** *University of Warwick and The Alan Turing Institute*

Abstract

We develop a novel Monte Carlo algorithm for the vector consisting of the supremum, the time at which the supremum is attained, and the position at a given (constant) time of an exponentially tempered Lévy process. The algorithm, based on the increments of the process without tempering, converges geometrically fast (as a function of the computational cost) for discontinuous and locally Lipschitz functions of the vector. We prove that the corresponding multilevel Monte Carlo estimator has optimal computational complexity (i.e. of order ε^{-2} if the mean squared error is at most ε^2) and provide its central limit theorem (CLT). Using the CLT we construct confidence intervals for barrier option prices and various risk measures based on drawdown under the tempered stable (CGMY) model calibrated/estimated on real-world data. We provide non-asymptotic and asymptotic comparisons of our algorithm with existing approximations, leading to rule-of-thumb principles guiding users to the best method for a given set of parameters. We illustrate the performance of the algorithm with numerical examples.

Keywords: Supremum; tempered; Lévy process; simulation; Monte Carlo estimation; geometric convergence; multilevel Monte Carlo

2020 Mathematics Subject Classification: Primary 60G51
Secondary 65C05

1. Introduction

1.1. Setting and motivation

The class of tempered stable processes is very popular in the financial modelling of asset prices of risky assets; see e.g. [10]. A tempered stable process $X = (X_t)_{t \geq 0}$ naturally addresses the shortcomings of diffusion models by allowing for the large (often heavy-tailed and asymmetric) sudden movements of the asset price observed in the markets, while preserving the exponential moments required in exponential Lévy models $S_0 e^X$ of asset prices [7, 10, 26, 36]. Of particular interest in this context are the expected drawdown (the current decline from a historical peak) and its duration (the elapsed time since the historical peak) (see e.g. [3, 8, 28, 38, 39]), as well as barrier option prices [2, 14, 27, 37] and the estimation of ruin probabilities in insurance [25, 29, 31]. In these application areas, the key quantities are the historic maximum

$\overline{X}_T$ at time T, the time $\tau_T(X)$ at which this maximum was attained during the time interval $[0,T]$, and the value X_T of the process X at time T.

In this paper we focus on the Monte Carlo (MC) estimation of $\mathbb{E}[g(X_T, \overline{X}_T, \tau_T(X))]$, where the payoff g is (locally) Lipschitz or of barrier type (cf. Proposition 3.1 below), covering the aforementioned applications. We construct a novel *tempered stick-breaking algorithm* (TSB-Alg), applicable to a tempered Lévy process, if the increments of the process without tempering can be simulated, which clearly holds if X is a tempered stable process. We show that the bias of TSB-Alg decays geometrically fast in its computational cost for functions g that are either locally Lipschitz or of barrier type (see Subsection 3 for details). We prove that the corresponding multilevel Monte Carlo (MLMC) estimator has optimal computational complexity (i.e. of order ε^{-2} if the mean squared error is at most ε^2) and establish the central limit theorem (CLT) for the MLMC estimator. Using the CLT we construct confidence intervals for barrier option prices and various risk measures based on drawdown under the tempered stable (CGMY) model. TSB-Alg combines the stick-breaking algorithm in [18] with the exponential change of measure for Lévy processes, also applied in [33] for the MC pricing of European options. A short YouTube video [17] describes TSB-Alg and the results of this paper.

1.2. Comparison with the literature

Exact simulation of the drawdown is currently out of reach. Under the assumption that the increments of the Lévy process X can be simulated (an assumption *not* satisfied by tempered stable models of infinite variation), the algorithm SB-Alg in [18] has a geometrically small bias, significantly outperforming other algorithms for which the computational complexity analysis has been carried out. For instance, the computational complexity analysis for the procedures presented in [12, 24], applicable to tempered stable processes of finite variation, has not been carried out, so that a direct quantitative comparison with SB-Alg [18] is not possible at present. If the increments cannot be sampled, a general approach utilises the Gaussian approximation of small jumps, in which case the algorithm SBG-Alg [16] outperforms its competitors (e.g. random walk approximation; see [16] for details), while retaining polynomially small bias. Thus it suffices to compare TSB-Alg below with SB-Alg [18] and SBG-Alg [16]. Table 1 below provides a summary of the properties of TSB-Alg, SB-Alg, and SBG-Alg as well as the asymptotic computational complexities of the corresponding MC and MLMC estimators based on these algorithms (see also Subsection 3.4 below for a detailed comparison).

The stick-breaking (SB) representation in (2) plays a central role in the algorithms TSB-Alg, SB-Alg, and SBG-Alg. The SB representation was used in [18] to obtain an approximation χ_n of $\chi := (X_T, \overline{X}_T, \tau_T(X))$ that converges geometrically fast in the computational cost when the increments of X can be simulated exactly. In [16], the same representation was used in conjunction with a small-jump Gaussian approximation for arbitrary Lévy processes. In the present work we address a situation in between those of the two aforementioned papers, using TSB-Alg below. TSB-Alg preserves the geometric convergence in the computational cost of SB-Alg, while being applicable to general tempered stable processes (unlike SB-Alg [18] in the infinite-variation case) and asymptotically outperforming SBG-Alg [16]; see Table 1 for an overview.

1.3. Organisation

The remainder of the paper is structured as follows. In Section 2 we recall the SB representation and construct TSB-Alg. In Section 3 we describe the geometric decay of the bias and the strong error in L^p and explain what the computational complexities of the MC and MLMC estimators are. We discuss briefly in Subsection 3.3 the construction and properties of

TABLE 1. Summary of the properties of TSB-Alg, SB-Alg [18], and SBG-Alg [16]. The index β_*, defined in (21) below, is *slightly* larger than the Blumenthal–Getoor index β; see Section 5 below for details. The bias and level variance are parametrised by computational effort n as $n \to \infty$, while the MC and MLMC complexities are parametrised by the accuracy ε (i.e., the mean squared error is at most ε^2) as $\varepsilon \to 0$.

Algorithm	TSB-Alg	SB-Alg [18]	SBG-Alg [16]
Class of Lévy processes to which algorithm applies	Tempered Lévy process for which the increments of the process without tempering can be simulated (includes *all* tempered stable processes!)	Lévy process whose increments can be simulated (among tempered stable, includes *only* finite-variation processes)	Lévy process whose jumps of magnitude greater than any $\delta > 0$ can be simulated (includes *all* tempered stable processes)
Bias and level variance	Both decay exponentially as $\mathcal{O}(e^{-\vartheta_g n})$ where $\vartheta > 0$ depends on g and β (see Proposition 3.1 below)	Both decay exponentially as $\mathcal{O}(e^{-\vartheta_g n})$ where $\vartheta > 0$ depends on g and β (see [18, Propositions 1–3])	Bias and level variance decay polynomially as $\mathcal{O}(n^{-p})$ and $\mathcal{O}(n^{-q})$, respectively, where $p \geq q > 0$ depend on g and β (see [16, Section 3.2] for bias and [16, Section 6.5.2] for level variance)
MC complexity	$\mathcal{O}(\varepsilon^{-2} \log(1/\varepsilon))$ for locally Lipschitz or barrier-type g (see Section 3 below)	$\mathcal{O}(\varepsilon^{-2} \log(1/\varepsilon))$ for locally Lipschitz or barrier-type g (see [18, Section 2.4])	$\mathcal{O}(\varepsilon^{-2-\beta_*})$ for locally Lipschitz g; otherwise, complexity is larger (see [16, Table 2])
MLMC complexity	$\mathcal{O}(\varepsilon^{-2})$ for locally Lipschitz or barrier-type g (see Section 3 below)	$\mathcal{O}(\varepsilon^{-2})$ for locally Lipschitz or barrier-type g (see [18, Section 2.4])	$\mathcal{O}(\varepsilon^{-2\min\{\beta_*,1\}})$ for locally Lipschitz g; otherwise, complexity is larger (see [16, Table 3])

unbiased estimators based on TSB-Alg. In Subsection 3.4 we provide an in-depth comparison of TSB-Alg with the SB and SBG algorithms, identifying where each algorithm outperforms the others. In Section 4 we consider the case of tempered stable processes and illustrate the previously described results with numerical examples. The proofs of all the results except Theorem 2.1, which is stated and proved in Section 2, are given in Section 5 below.

2. The tempered stick-breaking algorithm

Let $T > 0$ be a time horizon and $\boldsymbol{\lambda} = (\lambda_+, \lambda_-) \in \mathbb{R}_+^2$ a vector with non-negative coordinates, different from the origin $\mathbf{0} = (0, 0)$. A stochastic process $X = (X_t)_{t \in [0,T]}$ is said to be a *tempered Lévy process* under the probability measure $\mathbb{P}_{\boldsymbol{\lambda}}$ if its characteristic function satisfies

$$t^{-1} \log \mathbb{E}_{\boldsymbol{\lambda}}\big[e^{iuX_t}\big] = iub_{\boldsymbol{\lambda}} - \frac{1}{2}\sigma^2 u^2 + \int_{\mathbb{R}\setminus\{0\}} \big(e^{iux} - 1 - iux \cdot \mathbb{1}_{(-1,1)}(x)\big) \nu_{\boldsymbol{\lambda}}(\mathrm{d}x),$$

$$\text{for all } u \in \mathbb{R}, \ t > 0,$$

where $\mathbb{E}_\lambda$ denotes the expectation under the probability measure $\mathbb{P}_\lambda$, and the generating (or Lévy) triplet $(\sigma^2, \nu_\lambda, b_\lambda)$ is given by

$$\nu_\lambda(\mathrm{d}x) := e^{-\lambda_{\mathrm{sgn}(x)}|x|}\nu(\mathrm{d}x) \qquad \text{and} \qquad b_\lambda := b + \int_{(-1,1)} x\big(e^{-\lambda_{\mathrm{sgn}(x)}|x|} - 1\big)\nu(\mathrm{d}x), \quad (1)$$

where $\sigma^2 \in \mathbb{R}_+$, $b \in \mathbb{R}$, and ν is a Lévy measure on $\mathbb{R} \setminus \{0\}$, i.e. ν satisfies $\int_{(-1,1)} x^2 \nu(\mathrm{d}x) < \infty$ and $\nu(\mathbb{R} \setminus (-1, 1)) < \infty$. (All Lévy triplets in this paper are given with respect to the cutoff function $x \mapsto \mathbb{1}_{(-1,1)}(x)$, and the sign function in (1) is defined as $\mathrm{sgn}(x) := \mathbb{1}_{\{x>0\}} - \mathbb{1}_{\{x<0\}}$.) The triplet $(\sigma^2, \nu_\lambda, b_\lambda)$ uniquely determines the law of X via the Lévy–Khintchine formula for the characteristic function of X_t for $t > 0$ given in the displays above (for details, see [35, Theorems 7.10 and 8.1, Definition 8.2]).

Our aim is to sample from the law of the statistic $(X_T, \overline{X}_T, \tau_T)$ consisting of the position X_T of the process X at T, the supremum $\overline{X}_T := \sup\{X_s : s \in [0, T]\}$ of X on the time interval $[0,T]$, and the time $\tau_T := \inf\{s \in [0, T] : \overline{X}_s = \overline{X}_T\}$ at which the supremum was attained in $[0,T]$. By [18, Theorem 1] there exists a coupling (X, Y, ℓ) under a probability measure $\mathbb{P}_\lambda$, such that $\ell = (\ell_n)_{n \in \mathbb{N}}$ is a stick-breaking process on $[0,T]$ based on the uniform law $U(0, 1)$ (i.e. $L_0 := T$, $L_n := L_{n-1}U_n$, and $\ell_n := L_{n-1} - L_n$ for $n \in \mathbb{N}$, where the U_n are independent and identically distributed (i.i.d.) as $U(0, 1)$), independent of the Lévy process Y with law equal to that of X, and the SB representation holds $\mathbb{P}_\lambda$-almost surely (a.s.):

$$\chi := (X_T, \overline{X}_T, \tau_T) = \sum_{n=1}^{\infty} \big(\xi_n, \max\{\xi_n, 0\}, \ell_n \cdot \mathbb{1}_{\{\xi_n > 0\}}\big), \quad \text{where} \quad \xi_n := Y_{L_{n-1}} - Y_{L_n}, \ \ n \in \mathbb{N}. \tag{2}$$

We stress that ℓ is *not* independent of X. In fact (ℓ, Y) can be expressed naturally through the geometry of the path of X (see [32, Theorem 1] and the coupling in [18]), but further details of the coupling are not important for our purposes. The key step in the construction of our algorithm is given by the following theorem. Its proof is based on the coupling described above and the change-of-measure theorem for Lévy processes [35, Theorems 33.1 and 33.2].

Theorem 2.1. *Denote by σB, $Y^{(+)}$, $Y^{(-)}$ the independent Lévy processes with generating triplets $(\sigma^2, 0, 0)$, $(0, \nu_\lambda|_{(0,\infty)}, 0)$, $(0, \nu_\lambda|_{(-\infty,0)}, 0)$, respectively, satisfying $Y_t = \sigma B_t + Y_t^{(+)} + Y_t^{(-)} + b_\lambda t$ for all $t \in [0, T]$, $\mathbb{P}_\lambda$-a.s. Let $\mathbb{E}_\lambda$ (resp. $\mathbb{E}_0$) be the expectation under $\mathbb{P}_\lambda$ (resp. $\mathbb{P}_0$) and define*

$$\Upsilon_\lambda := \exp\big(-\lambda_+ Y_T^{(+)} + \lambda_- Y_T^{(-)} - \mu_\lambda T\big), \qquad \text{where} \tag{3}$$

$$\mu_\lambda := \int_{\mathbb{R}} \big(e^{-\lambda_{\mathrm{sgn}(x)}|x|} - 1 + \lambda_{\mathrm{sgn}(x)}|x| \cdot \mathbb{1}_{(-1,1)}(x)\big)\nu(\mathrm{d}x). \tag{4}$$

Then for any $\sigma(\ell, \xi)$-measurable random variable ζ with $\mathbb{E}_\lambda|\zeta| < \infty$ we have $\mathbb{E}_\lambda[\zeta] = \mathbb{E}_0[\zeta \Upsilon_\lambda]$.

Proof. The exponential change-of-measure theorem for Lévy processes (see [35, Theorems 33.1 and 33.2]) implies that for any measurable function F with $\mathbb{E}_\lambda|F((Y_t)_{t \in [0,T]})| < \infty$, we have the identity $\mathbb{E}_\lambda[F((Y_t)_{t \in [0,T]})] = \mathbb{E}_0[F((Y_t)_{t \in [0,T]})\Upsilon_\lambda]$, where Υ_λ is defined in (3) in the statement of Theorem 2.1. Since the stick-breaking process ℓ is independent of Y under both $\mathbb{P}_\lambda$

and $\mathbb{P}_{\mathbf{0}}$, this property extends to measurable functions of $(\ell, (Y_t)_{t \in [0,T]})$ and thus to measurable functions of (ℓ, ξ), as claimed. □

By the equality in (2), the measurable function ζ of the sequences ℓ and ξ in Theorem 2.1 may be equal either to $g(\chi)$ (for any integrable function g of the statistic χ) or to its approximation $g(\chi_n)$, where χ_n is as introduced in [18]:

$$\chi_n := \left(Y_{L_n}, \max\{Y_{L_n}, 0\}, L_n \cdot \mathbb{1}_{\{Y_{L_n} > 0\}}\right) + \sum_{k=1}^{n} \left(\xi_k, \max\{\xi_k, 0\}, \ell_k \cdot \mathbb{1}_{\{\xi_k > 0\}}\right). \tag{5}$$

Theorem 2.1 enables us to sample χ_n under the probability measure $\mathbb{P}_{\mathbf{0}}$, which for any tempered stable process X makes the increments of Y stable and thus easy to simulate. Under $\mathbb{P}_{\mathbf{0}}$, the law of Y_t equals that of $Y_t^{(+)} + Y_t^{(-)} + \sigma B_t + bt$, where $\sigma B_t + bt$ is normal $N(bt, \sigma^2 t)$ with mean bt and variance $\sigma^2 t$ and the Lévy processes $Y^{(+)}$ and $Y^{(-)}$ have triplets $(0, \nu|_{(0,\infty)}, 0)$ and $(0, \nu|_{(-\infty,0)}, 0)$, respectively. Denote their distribution functions by $F^{(\pm)}(t, x) := \mathbb{P}_{\mathbf{0}}\left(Y_t^{(\pm)} \leq x\right)$, $x \in \mathbb{R}$, $t > 0$.

TSB-Alg

Unbiased simulation of $g(\chi_n)$ under $\mathbb{P}_{\lambda}$

Require: Tempering parameter $\lambda \in \mathbb{R}_+^2 \setminus \{\mathbf{0}\}$, generating triplet (σ^2, ν, b), time horizon $T > 0$, test function g, approximation level $n \in \mathbb{N}$

1: Set $L_0 = T$ and compute μ_{λ} in (4)
2: **for** $k = 1, \ldots, n$ **do**
3:      Sample $U_k \sim U(0, 1)$ and put $L_k = U_k L_{k-1}$ and $\ell_k = L_{k-1} - L_k$
4:      Sample $\xi_k^{(\pm)} \sim F^{(\pm)}(\ell_k, \cdot)$, $G_k \sim N(\ell_k b, \sigma^2 \ell_k)$ and put $\xi_k = \xi_k^{(+)} + \xi_k^{(-)} + G_k$
5: Sample $\zeta_n^{(\pm)} \sim F^{(\pm)}(L_n, \cdot)$, $H_n \sim N(L_n b, \sigma^2 L_n)$ and put $\zeta_n = \zeta_n^{(+)} + \zeta_n^{(-)} + H_n$
6: **end for**
7: Set $\chi_n = \sum_{k=1}^{n} (\xi_k, \max\{\xi_k, 0\}, \ell_k \mathbb{1}_{\{\xi_k > 0\}}) + (\zeta_n, \max\{\zeta_n, 0\}, L_n \mathbb{1}_{\{\zeta_n > 0\}})$ and $Y_T^{(\pm)} = \zeta_n^{(\pm)} + \sum_{k=1}^{n} \xi_k^{(\pm)}$
8: **return** $g(\chi_n) \exp\left(-\lambda_+ Y_T^{(+)} + \lambda_- Y_T^{(-)} - \mu_{\lambda} T\right)$

Note that the output $g(\chi_n) \Upsilon_{\lambda}$ of TSB-Alg is sampled under $\mathbb{P}_{\mathbf{0}}$ and, by Theorem 2.1 above, is unbiased since $\mathbb{E}_{\mathbf{0}}[g(\chi_n) \Upsilon_{\lambda}] = \mathbb{E}_{\lambda}[g(\chi_n)]$. As our aim is to obtain MC and MLMC estimators for $\mathbb{E}_{\lambda}[g(\chi)]$, our next task is to understand the expected error of TSB-Alg; see (6) in Subsection 3.1 below. In [18] it was proved that the approximation χ_n converges geometrically fast in computational effort (or equivalently as $n \to \infty$) to χ if the increments of Y can be sampled under $\mathbb{P}_{\lambda}$ (see [18] for more details and a discussion of the benefits of the 'correction term' $(Y_{L_n}, \max\{Y_{L_n}, 0\}, L_n \cdot \mathbb{1}_{\{Y_{L_n} > 0\}})$ in (5)). Theorem 2.1 allows us to weaken this requirement in the context of tempered Lévy processes, by requiring that we be able to sample the increments of Y under $\mathbb{P}_{\mathbf{0}}$. The main application of TSB-Alg is to general tempered stable processes, as the simulation of their increments is currently out of reach for many cases of interest (see Section 3.4 below for comparison with existing methods when it is not), making the main algorithm in [18] not applicable. Moreover, Theorem 2.1 allows us to retain the geometric convergence of χ_n established in [18]; see Section 3 below for details.

3. Monte Carlo and multilevel Monte Carlo estimators based on TSB-Alg

3.1. Bias of TSB-Alg

An application of Theorem 2.1 implies that the bias of TSB-Alg equals

$$\mathbb{E}_{\boldsymbol{\lambda}}[g(\chi)] - \mathbb{E}_{\mathbf{0}}[g(\chi_n)\Upsilon_\lambda] = \mathbb{E}_{\mathbf{0}}\big[\Delta_n^g\big], \quad \text{where } \Delta_n^g := (g(\chi) - g(\chi_n))\Upsilon_\lambda. \tag{6}$$

The natural coupling $\big(\chi, \chi_n, Y_T^{(+)}, Y_T^{(-)}\big)$ in (6) is defined by $Y_T^{(\pm)} := \sum_{k=1}^\infty \xi_k^{(\pm)}$, $\xi_k := \xi_k^{(+)} + \xi_k^{(-)} + \eta_k$ for all $k \in \mathbb{N}$, χ in (2), and χ_n in (5) with $Y_{L_n} := \sum_{k=n+1}^\infty \xi_k$, where, conditional on the stick-breaking process $\ell = (\ell_k)_{k \in \mathbb{N}}$, the random variables $\{\xi_k^{(\pm)}, \eta_k : k \in \mathbb{N}\}$ are independent and distributed as $\xi_k^{(\pm)} \sim F^{(\pm)}(\ell_k, \cdot)$ and $\eta_k \sim N\big(\ell_k b, \sigma^2 \ell_k\big)$ for $k \in \mathbb{N}$.

The following result presents the decay of the strong error Δ_n^g for Lipschitz, locally Lipschitz, and two classes of barrier-type discontinuous payoffs that arise frequently in applications. Observe that, in all cases and under the corresponding mild assumptions, the pth moment of the strong error Δ_n^g decays exponentially fast in n with a rate $\vartheta > 0$ that depends on the payoff g, the index β_* defined in (21) below, and the degree p of the considered moment. In Proposition 3.1 and throughout the paper, the notation $f(n) = \mathcal{O}(g(n))$ as $n \to \infty$ for functions $f, g \colon \mathbb{N} \to (0, \infty)$ means $\limsup_{n \to \infty} f(n)/g(n) < \infty$. Put differently, $f(n) = \mathcal{O}(g(n))$ is equivalent to $f(n)$ being bounded above by $C_0 g(n)$ for some constant $C_0 > 0$ and all $n \in \mathbb{N}$.

Proposition 3.1. *Let $\boldsymbol{\lambda} = (\lambda_+, \lambda_-)$, ν, and σ^2 be as in Section 2, and fix $p \geq 1$. Then, for the classes of payoffs $g(\chi) = g(X_T, \overline{X}_T, \tau_T)$ below, the strong error of TSB-Alg decays as follows (as $n \to \infty$).*

(Lipschitz.) *Suppose $|g(x, y, t) - g(x, y', t')| \leq K(|y - y'| + |t - t'|)$ for some K and all $(x, y, y', t, t') \in \mathbb{R} \times \mathbb{R}_+^2 \times [0, T]^2$. Then $\mathbb{E}_{\mathbf{0}}\big[\big|\Delta_n^g\big|^p\big] = \mathcal{O}\big(e^{-\vartheta_p n}\big)$, where $\vartheta_p \in [\log (3/2), \log 2]$ is in (23) below.*

(Locally Lipschitz.) *Let $|g(x, y, t) - g(x, y', t')| \leq K(|y - y'| + |t - t'|)e^{\max\{y, y'\}}$ for some constant $K > 0$ and all $(x, y, y', t, t') \in \mathbb{R} \times \mathbb{R}_+^2 \times [0, T]^2$. If $\lambda_+ \geq q > 1$ and $\int_{[1, \infty)} e^{p(q - \lambda_+)x} \nu(\mathrm{d}x) < \infty$, then $\mathbb{E}_{\mathbf{0}}\big[\big|\Delta_n^g\big|^p\big] = \mathcal{O}\big(e^{-(\vartheta_{pr}/r)n}\big)$, where $r := (1 - 1/q)^{-1} > 1$ and $\vartheta_{pr} \in [\log (3/2), \log 2]$ is as in (23).*

(Barrier type 1.) *Suppose $g(\chi) = h(X_T) \cdot \mathbb{1}\{\overline{X}_T \leq M\}$ for some $M > 0$ and a measurable bounded function $h \colon \mathbb{R} \to \mathbb{R}$. If $\mathbb{P}_{\mathbf{0}}(M < \overline{X}_T \leq M + x) \leq Kx$ for some $K > 0$ and all $x \geq 0$, then for $\alpha_* \in (1, 2]$ in (22) and any $\gamma \in (0, 1)$ we have $\mathbb{E}_{\mathbf{0}}\big[\big|\Delta_n^g\big|^p\big] = \mathcal{O}\big(e^{-[\gamma \log (2)/(\gamma + \alpha_*)]n}\big)$. Moreover, we may take $\gamma = 1$ if any of the following hold: $\sigma^2 > 0$ or $\int_{(-1,1)} |x| \nu(\mathrm{d}x) < \infty$ or Assumption 5.1 below.*

(Barrier type 2.) *Suppose $g(\chi) = h(X_T, \overline{X}_T) \cdot \mathbb{1}\{\tau_T \leq s\}$, where $s \in (0, T)$, and h is measurable and bounded with $|h(x, y) - h(x, y')| \leq K|y - y'|$ for some $K > 0$ and all $(x, y, y') \in \mathbb{R} \times \mathbb{R}_+^2$. If $\sigma^2 > 0$ or $\nu(\mathbb{R} \setminus \{0\}) = \infty$, then $\mathbb{E}_{\mathbf{0}}\big[\big|\Delta_n^g\big|^p\big] = \mathcal{O}\big(e^{-n/e}\big)$.*

Remark 3.1. (i) The proof of Proposition 3.1, given in Section 5 below, is based on Theorem 2.1 and analogous bounds in [18] (for Lipschitz, locally Lipschitz, and barrier-type-1 payoffs) and [16] (for barrier-type-2 payoffs). In particular, in the proof of Proposition 3.1 below, we need not assume $\lambda_+ > 0$ to apply [18, Proposition 1], which works under our standing assumption $\boldsymbol{\lambda} \neq \mathbf{0}$.

(ii) For barrier option payoffs under a tempered stable process X (i.e. the barrier-type-1 class in Proposition 3.1), we may take $\gamma = 1$ since X satisfies either $\int_{(-1,1)} |x| \nu(\mathrm{d}x) < \infty$ or Assumption 5.1.

(iii) The restriction $p \geq 1$ is not essential, as we may consider any $p > 0$ at the cost of a smaller (but still geometric) convergence rate. In particular, our standing assumption $\boldsymbol{\lambda} \neq \mathbf{0}$ (and $\lambda_+ > 1$ in the locally Lipschitz case) guarantees the finiteness of the p-moment of the strong error Δ_n^g for any $p > 0$. However, the restriction $p \geq 1$ covers the cases $p \in \{1, 2\}$ required for the MC and MLMC complexity analyses and ensures that the corresponding rate ϑ_s in (23) lies in $[\log(3/2), \log 2]$. In fact, for any payoff g in Proposition 3.1 we have $\mathbb{E}_{\mathbf{0}}\big[\big|\Delta_k^g\big|^p\big] = \mathcal{O}\big(e^{-\vartheta_g k}\big)$ for $p \in \{1, 2\}$ and a positive rate $\vartheta_g > 0$ bounded away from zero: $\vartheta_g \geq 0.23$ (resp. $\log(3/2)$, $(1 - 1/\lambda_+)\log(3/2)$) for barrier-type-1 and barrier-type-2 (resp. Lipschitz, locally Lipschitz) payoffs.

3.2. Computational complexity and the central limit theorem for the Monte Carlo and multilevel Monte Carlo estimators

Consider the MC estimator

$$\hat{\theta}_{\mathrm{MC}}^{g,n} := \frac{1}{N} \sum_{i=1}^{N} \theta_i^{g,n}, \tag{7}$$

where $\{\theta_i^{g,n}\}_{i\in\mathbb{N}}$ is the i.i.d. output of TSB-Alg with $\theta_1^{g,n} \overset{d}{=} g(\chi_n)\Upsilon_{\boldsymbol{\lambda}}$ (under $\mathbb{P}_{\mathbf{0}}$) and $n, N \in \mathbb{N}$. The corresponding MLMC estimator is given by

$$\hat{\theta}_{\mathrm{ML}}^{g,n} := \sum_{k=1}^{n} \frac{1}{N_k} \sum_{i=1}^{N_k} D_{k,i}^g, \tag{8}$$

where $\{D_{k,i}^g\}_{k,i\in\mathbb{N}}$ is an array of independent variables satisfying $D_{k,i}^g \overset{d}{=} (g(\chi_k) - g(\chi_{k-1}))\Upsilon_{\boldsymbol{\lambda}}$ and $D_{1,i}^g \overset{d}{=} g(\chi_1)\Upsilon_{\boldsymbol{\lambda}}$ (under $\mathbb{P}_{\mathbf{0}}$), for $i \in \mathbb{N}$, $k \geq 2$, and $n, N_1, \ldots, N_n \in \mathbb{N}$. Note that TSB-Alg can easily be adapted to sample the variable $D_{k,i}^g$ by drawing the ingredients for $(\chi_k, \Upsilon_{\boldsymbol{\lambda}})$ and computing $(\chi_{k-1}, \chi_k, \Upsilon_{\boldsymbol{\lambda}})$ deterministically from the output; see [18, Subsection 2.4] for further details. In the following, we refer to $\mathbb{V}_{\mathbf{0}}\big[D_{k,1}^g\big]$ as the level variance of the MLMC estimator.

The computational complexity analysis of the MC and MLMC estimators is given in the next result (the usual notation $\lceil x \rceil := \inf\{k \in \mathbb{N} : k \geq x\}$, $x \in \mathbb{R}_+$, is used for the ceiling function). In Proposition 3.2 and throughout the paper, the computational cost of an algorithm is measured as the total number of operations carried out by the algorithm. In particular, we assume that the following operations have computational costs uniformly bounded by some constant (measured, for instance, in units of time): simulation from the uniform law; simulation from the laws $F^{(\pm)}(t, \cdot)$, $t > 0$; evaluation of elementary mathematical operations such as addition, subtraction, multiplication, and division; and evaluation of elementary functions such as exp, log, sin, cos, tan, and arctan.

Proposition 3.2. *Let the payoff g from Proposition* 3.1 *also satisfy* $\mathbb{E}_{\mathbf{0}}\big[g(\chi)^2\Upsilon_{\boldsymbol{\lambda}}^2\big] < \infty$. *For any* $\varepsilon > 0$, *let* $n(\varepsilon) := \inf\{k \in \mathbb{N} : |\mathbb{E}_{\mathbf{0}}[g(\chi_k)\Upsilon_{\boldsymbol{\lambda}}] - \mathbb{E}_{\boldsymbol{\lambda}}[g(\chi)]| \leq \varepsilon/\sqrt{2}\}$. *Let c be an upper bound on the expected computational cost of line* 2 *in TSB-Alg for a time horizon bounded by T, and let* $\mathbb{V}_{\mathbf{0}}[\,\cdot\,]$ *denote the variance under the probability measure* $\mathbb{P}_{\mathbf{0}}$.

(MC.) Suppose $n = n(\varepsilon)$ and $N = \lceil 2\varepsilon^{-2}\mathbb{V}_\mathbf{0}[g(\chi_n)\Upsilon_\lambda]\rceil$. Then the MC estimator $\hat{\theta}_{\mathrm{MC}}^{g,n}$ is ε-accurate, i.e. $\mathbb{E}_\mathbf{0}\big[\big|\hat{\theta}_{\mathrm{MC}}^{g,n} - \mathbb{E}_\lambda[g(\chi)]\big|^2\big] \le \varepsilon^2$, with expected cost $\mathcal{C}_{\mathrm{MC}}(\varepsilon) := c(n+1)N = \mathcal{O}\big(\varepsilon^{-2}\log(1/\varepsilon)\big)$ as $\varepsilon \to 0$.

(MLMC.) Suppose $n = n(\varepsilon)$ and set

$$N_k := \left\lceil 2\varepsilon^{-2}\sqrt{\mathbb{V}_\mathbf{0}\big[D_{k,1}^g\big]/k}\left(\sum_{j=1}^n \sqrt{j\mathbb{V}_\mathbf{0}\big[D_{j,1}^g\big]}\right)\right\rceil, \quad k \in \{1, \dots, n\}. \tag{9}$$

Then the MLMC estimator $\hat{\theta}_{\mathrm{ML}}^{g,n}$ is ε-accurate, and the corresponding expected cost equals

$$\mathcal{C}_{\mathrm{ML}}(\varepsilon) := 2c\varepsilon^{-2}\left(\sum_{k=1}^n \sqrt{k\mathbb{V}_\mathbf{0}\big[D_{k,1}^g\big]}\right)^2 = \mathcal{O}\big(\varepsilon^{-2}\big) \quad \text{as } \varepsilon \to 0. \tag{10}$$

Proposition 3.1 (with $p = 1$) implies that the bias in (6) equals $\mathbb{E}_\mathbf{0}\big[\Delta_n^g\big] = \mathcal{O}\big(e^{-\vartheta_g n}\big)$ as $n \to \infty$ for some $\vartheta_g > 0$. Thus, the integer $n(\varepsilon)$ in Proposition 3.2 is finite for all payoffs g considered in Proposition 3.1, and moreover, $n(\varepsilon) = \mathcal{O}(\log(1/\varepsilon))$ as $\varepsilon \to 0$ in all cases. In addition, by Remark 3.1(i) above, the variance of $\theta_1^{g,k}$ is bounded in $k \in \mathbb{N}$:

$$\mathbb{V}_\mathbf{0}\big[\theta_1^{g,k}\big] \le \mathbb{E}_\mathbf{0}\big[g(\chi_k)^2\Upsilon_\lambda^2\big] \le 2\mathbb{E}_\mathbf{0}\big[g(\chi)^2\Upsilon_\lambda^2\big] + 2\mathbb{E}_\mathbf{0}\big[\big(\Delta_k^g\big)^2\big] \to 2\mathbb{E}_\mathbf{0}\big[g(\chi)^2\Upsilon_\lambda^2\big] < \infty \quad \text{as } k \to \infty.$$

Note that barrier-type payoffs g considered in Proposition 3.1 satisfy the second moment assumption, while in the Lipschitz or locally Lipschitz cases it is sufficient to assume additionally that λ_+ is either positive or strictly greater than one, respectively. Moreover, $\mathbb{V}_\mathbf{0}\big[D_{k,1}^g\big] \le 2\mathbb{E}_\mathbf{0}\big[\big(\Delta_k^g\big)^2 + \big(\Delta_{k-1}^g\big)^2\big] = \mathcal{O}\big(e^{-\vartheta_g k}\big)$ for $\vartheta_g > 0$ bounded away from zero (see Remark 3.1(iii) above). These facts and the standard complexity analysis for MLMC (see e.g. [16, Appendix A] and the references therein) imply that the estimators $\hat{\theta}_{\mathrm{MC}}^{g,n}$ and $\hat{\theta}_{\mathrm{ML}}^{g,n}$ are ε-accurate with the stated computational costs, proving Proposition 3.2.

We stress that the payoffs g in Proposition 3.2 include discontinuous payoffs in the supremum $\overline{X}_T$ (barrier type 1) or the time τ_T when this supremum is attained (barrier type 2), with the corresponding computational complexities of the MC and MLMC estimators given by $\mathcal{O}\big(\varepsilon^{-2}\log(1/\varepsilon)\big)$ and $\mathcal{O}\big(\varepsilon^{-2}\big)$, respectively. This theoretical prediction matches the numerical performance of TSB-Alg for barrier options and the modified ulcer index; see Section 4.2 below.

For obtaining confidence intervals (CIs) for the MC and MLMC estimators $\hat{\theta}_{\mathrm{MC}}^{g,n}$ and $\hat{\theta}_{\mathrm{ML}}^{g,n}$, a CLT can be very helpful. (The CIs derived in this paper do not account for model uncertainty or the uncertainty in the estimation or calibration of the parameters.) In fact, the CLT is necessary to construct a CI if the constants in the bounds on the bias in Proposition 3.1 are not explicitly known (e.g. for barrier-type-1 payoffs, the constant depends on the unknown value of the density of the supremum $\overline{X}_T$ at the barrier); see the discussion in [18, Section 2.3]. Moreover, even if the bias can be controlled explicitly, the concentration inequalities typically lead to wider CIs than those based on the CLT; see [4, 20]. The following result establishes the CLT for the MC and MLMC estimators valid for payoffs considered in Proposition 3.1.

Theorem 3.1. *[CLT for TSB-Alg] Let g be any of the payoffs in Proposition 3.1, satisfying in addition $\mathbb{E}_\mathbf{0}\big[g(\chi)^2\Upsilon_\lambda^2\big] < \infty$. Let $\vartheta_g \in (0, \log 2)$ be the rate satisfying $\mathbb{E}_\mathbf{0}\big[\big|\Delta_n^g\big|\big] = \mathcal{O}\big(e^{-\vartheta_g n}\big)$, given in Proposition 3.1 and Remark 3.1(iii) above (with $p = 1$). Fix $c_0 > 1/\vartheta_g$, let $n = n(\varepsilon) := \lceil c_0 \log(1/\varepsilon)\rceil$, and suppose N and $N_1, \dots, N_n$ are as in Proposition 3.2. Then the MC and*

MLMC estimators $\hat{\theta}_{MC}^{g,n}$ and $\hat{\theta}_{ML}^{g,n}$ respectively satisfy the following CLTs (Z is a standard normal random variable):

$$\sqrt{2}\varepsilon^{-1}\big(\hat{\theta}_{MC}^{g,n(\varepsilon)} - \mathbb{E}_{\boldsymbol{\lambda}}[g(\chi)]\big) \xrightarrow{d} Z, \qquad \text{and} \qquad \sqrt{2}\varepsilon^{-1}\big(\hat{\theta}_{ML}^{g,n(\varepsilon)} - \mathbb{E}_{\boldsymbol{\lambda}}[g(\chi)]\big) \xrightarrow{d} Z, \quad \text{as } \varepsilon \to 0. \tag{11}$$

Theorem 3.1 works well in practice: in Figure 7 of Section 4.2 below we construct CIs (as a function of decreasing accuracy ε) for an MLMC estimator of a barrier option price under a tempered stable model. The rate c_0 can be taken arbitrarily close to $1/\vartheta_g$, where ϑ_g is the corresponding geometric rate of decay of the error for the payoff g in Proposition 3.1 (ϑ_g is bounded away from zero by Remark 3.1(iii) above), ensuring that the bias of the estimators vanishes in the limit.

By Lemma 5.2 below, the definition of the sample sizes N and $N_1, \ldots, N_n$ in Proposition 3.2 implies that the variances of the estimators $\hat{\theta}_{MC}^{g,n(\varepsilon)}$ and $\hat{\theta}_{ML}^{g,n(\varepsilon)}$ (under $\mathbb{P}_{\boldsymbol{0}}$) satisfy

$$\frac{\mathbb{V}_{\boldsymbol{0}}\big[\hat{\theta}_{MC}^{g,n(\varepsilon)}\big]}{\varepsilon^2/2} = \frac{\mathbb{V}_{\boldsymbol{0}}\big[\theta_1^{g,n(\varepsilon)}\big]}{\varepsilon^2 N/2} \to 1 \quad \text{and} \quad \frac{\mathbb{V}_{\boldsymbol{0}}\big[\hat{\theta}_{ML}^{g,n(\varepsilon)}\big]}{\varepsilon^2/2} = \sum_{k=1}^{n(\varepsilon)} \frac{\mathbb{V}_{\boldsymbol{0}}\big[D_{k,1}^g\big]}{\varepsilon^2 N_k/2} \to 1 \quad \text{as } \varepsilon \to 0.$$

Hence, the CLT in (11) can also be expressed as follows:

$$\big(\hat{\theta}_{MC}^{g,n(\varepsilon)} - \mathbb{E}_{\boldsymbol{\lambda}}[g(\chi)]\big)/\mathbb{V}_{\boldsymbol{0}}\big[\hat{\theta}_{MC}^{g,n(\varepsilon)}\big]^{1/2} \xrightarrow{d} Z$$

$$\text{and} \quad \big(\hat{\theta}_{ML}^{g,n(\varepsilon)} - \mathbb{E}_{\boldsymbol{\lambda}}[g(\chi)]\big)/\mathbb{V}_{\boldsymbol{0}}\big[\hat{\theta}_{ML}^{g,n(\varepsilon)}\big]^{1/2} \xrightarrow{d} Z, \quad \text{as } \varepsilon \to 0.$$

Since the variances $\mathbb{V}_{\boldsymbol{0}}\big[\hat{\theta}_{MC}^{g,n(\varepsilon)}\big]$ and $\mathbb{V}_{\boldsymbol{0}}\big[\hat{\theta}_{ML}^{g,n(\varepsilon)}\big]$ can be estimated from the sample, this is in fact how the CLT is often applied in practice. The proof of Theorem 3.1 is based on the CLT for triangular arrays and amounts to verifying Lindeberg's condition for the estimators $\hat{\theta}_{MC}^{g,n}$ and $\hat{\theta}_{ML}^{g,n}$; see Section 5 below.

3.3. Unbiased estimators

It is known that when the MLMC complexity is optimal, it is possible to construct unbiased estimators without altering the optimal computational complexity $\mathcal{O}\big(\varepsilon^{-2}\big)$ as $\varepsilon \to 0$. Such debiasing techniques are based around randomising the number of levels $\sup\{k \in \mathbb{N}: N_k > 0\}$ and number of samples $(N_k)_{k\in\mathbb{N}}$ at each level of the variables $\{D_{k,i}^g\}_{k,i\in\mathbb{N}}$ in the MLMC estimator in (8); see e.g. [30, 34, 40]. More precisely, following [40, Theorem 7], for any g in Proposition 3.1 and a parameter $N \in \mathbb{N}$, we may construct an almost surely finite sequence of random integers $(N_k)_{k\in\mathbb{N}}$ (i.e. $\sup\{k \in \mathbb{N}: N_k > 0\} < \infty$ a.s.) with explicit means $\mathbb{E}_{\boldsymbol{0}}[N_k] > 0$, such that the estimator

$$\hat{P}_N := \sum_{k=1}^{\infty} \frac{1}{\mathbb{E}_{\boldsymbol{0}}[N_k]} \sum_{i=1}^{N_k} D_{k,i}^g \tag{12}$$

is unbiased, $\mathbb{E}_{\boldsymbol{0}}[\hat{P}_N] = \mathbb{E}_{\boldsymbol{\lambda}}[g(\chi)]$, and its variance $\mathbb{V}_{\boldsymbol{0}}[\hat{P}_N]$ and expected computational cost (under $\mathbb{P}_{\boldsymbol{0}}$) are proportional to $1/N$ and N, respectively, as $N \to \infty$. The MC complexity analysis of the estimator $\hat{P}_N$ is then nearly identical to that of the classical MC estimator for exact simulation algorithms.

There are several parametric ways of constructing the random sequence $(N_k)_{k\in\mathbb{N}}$ (see e.g. [40]), and it is also possible to optimise for the parameters appearing in these constructions

as a function of the considered payoff g and other characteristics of X (see e.g. [18, Section 2.5]). The details of such optimisations have been omitted in the present work in the interest of brevity, since they coincide with those found in [18, Section 2.5 and Appendix A.3]. We comment that, as pointed out by the referee, in addition to achieving the optimal complexity, unbiased estimators (referred to as 'single-term estimators' in [34, 40]) also achieve the same asymptotic variance as the MLMC estimators (see [40, Theorem 19]). Moreover, there are cases where unbiased estimators of a different form (referred to as 'coupled-sum estimators' in [34, 40]) achieve much smaller asymptotic variance than the MLMC (see Table 17 in the online supplement of [34]). We are grateful to the referee for this observation.

3.4. Comparisons

This subsection provides non-asymptotic and asymptotic performance comparisons of estimators based on TSB-Alg. The main aim is to develop rule-of-thumb principles guiding the user to the most effective estimator. In Subsection 3.4.1, for a given value of the accuracy ε, we compare the computational complexity of the MC and MLMC estimators based on TSB-Alg. The MLMC estimator based on TSB-Alg is compared with the ones based on SB-Alg [18] with rejection sampling (available only when the jumps of X are of finite variation) and SBG-Alg [16] in Subsections 3.4.2 and 3.4.3, respectively. In both cases we analyse the behaviour of the computational complexity in two regimes: (I) ε tending to 0 and fixed time horizon T; (II) fixed ε and time horizon T tending to 0 or ∞.

Regime (II) is useful when there is a limited benefit to arbitrary accuracy in ε but the constants may vary greatly with the time horizon T. For example, in option pricing, estimators with accuracy smaller than a basis point are of limited interest. For simplicity, in the remainder of this subsection the payoff g is assumed to be Lipschitz. However, analogous comparisons can be made for other payoffs under appropriate assumptions.

3.4.1. *Comparison between the Monte Carlo and multilevel Monte Carlo estimators based on TSB-Alg* Recall first that both MC and MLMC estimators have the same bias, since the latter estimator is a telescoping sum of a sequence of the former estimators, controlled by $n(\varepsilon)$ given in Theorem 3.1 above.

Regime (I). Propositions 3.1 and 3.2 imply that the MLMC estimator outperforms the MC estimator as $\varepsilon \to 0$. Moreover, since $\mathbb{V}_{\mathbf{0}}[g(\chi_n)\Upsilon_\lambda] \to \mathbb{V}_{\mathbf{0}}[g(\chi)\Upsilon_\lambda]$ and $\varepsilon^2 \mathcal{C}_{\mathrm{ML}}(\varepsilon) \to 2c\big(\sum_{k=1}^{\infty} (k\mathbb{V}_{\mathbf{0}}[D_{k,1}^g])^{1/2}\big)^2 < \infty$ as $\varepsilon \to 0$, the MLMC estimator is preferable to the MC estimator for $\varepsilon > 0$ satisfying

$$ n(\varepsilon) > \left(\sum_{k=1}^{\infty} \sqrt{k\mathbb{V}_{\mathbf{0}}\big[D_{k,1}^g\big]} \right)^2 / \mathbb{V}_{\mathbf{0}}\big[g(\chi)\Upsilon_\lambda\big]. $$

Since the payoff g is Lipschitz, Proposition 3.1 implies that this condition is close to

$$ \log(1/\varepsilon) > \vartheta_1 \left(\sum_{k=1}^{\infty} \sqrt{k\big(2^{-k} - e^{-2\vartheta_1 k}\big)} \right)^2, $$

where we recall that $\vartheta_1 \in [\log(3/2), \log 2]$ is defined in (23) below. In particular, the latter condition is equivalent to $\varepsilon < 0.0915$ if $\vartheta_1 = \log(3/2)$, or $\varepsilon < 5.06 \times 10^{-5}$ if $\vartheta_1 = \log 2$.

Regime (II). Assume $\varepsilon > 0$ is fixed. In this case, the MC and MLMC estimators share the value of $n = n(\varepsilon)$, which is $\mathcal{O}(\max\{1, \log T\})$ as either $T \to 0$ or $T \to \infty$. Moreover, the variances $\mathbb{V}_{\mathbf{0}}[g(\chi_n)\Upsilon_\lambda]$ (appearing in $\mathcal{C}_{\mathrm{MC}}$) and $\mathbb{V}_{\mathbf{0}}[D_{k,1}^g]$, $k \in \mathbb{N}$ (appearing in $\mathcal{C}_{\mathrm{ML}}$; see

Proposition 3.2 above) are all proportional to $\mathcal{O}\big((T + T^2)e^{(\mu_{2\lambda} - 2\mu_\lambda)T}\big)$ as either $T \to 0$ or $T \to \infty$. Therefore, by Proposition 3.2, the quotient $\mathcal{C}_{\mathrm{MC}}/\mathcal{C}_{\mathrm{ML}}$ is proportional to a constant as $T \to 0$ and a multiple of $\log T$ as $T \to \infty$.

In conclusion, the MLMC estimator is preferable to the MC estimator when either ε is small or T is large. Otherwise, when ε is not small and T is small, the two estimators have similar performance.

3.4.2. *Comparison with SB-Alg* In the special case when the jumps of X have finite variation (or equivalently, $\int_{(-1,1)} |x|\nu(\mathrm{d}x) < \infty$), the increments X_t can be simulated under $\mathbb{P}_\lambda$ using rejection sampling (see [19, 23]), making SB-Alg [18] applicable to sample χ_n (see (5) for its definition) under $\mathbb{P}_\lambda$. The rejection sampling is performed for each of the increments of the subordinators

$$\widetilde{Y}_t^{(\pm)} := \pm Y_t^{(\pm)} + d_\pm t, \quad \text{where} \quad d_+ := \int_{(0,1)} x\nu(\mathrm{d}x) \quad \text{and} \quad d_- := \int_{(-1,0)} |x|\nu(\mathrm{d}x),$$

and the processes $Y^{(\pm)}$ are as in Theorem 2.1. The algorithm proposes samples under $\mathbb{P}_0$ and rejects independently with probability $\exp\big(-\lambda_\pm \widetilde{Y}_t^{(\pm)}\big)$. Let $\boldsymbol{\lambda}_+ := (\lambda_+, 0)$ and $\boldsymbol{\lambda}_- := (0, \lambda_-)$; then the expected number of proposals required for each sample equals $\exp\big(\gamma_{\boldsymbol{\lambda}}^{(\pm)}t\big) = 1/\mathbb{E}_0\big[\exp\big(-\lambda_\pm \widetilde{Y}_t^{(\pm)}\big)\big]$, where we define

$$\gamma_{\boldsymbol{\lambda}}^{(\pm)} := \lambda_\pm d_\pm - \mu_{\boldsymbol{\lambda}_\pm} = \int_{\mathbb{R}_\pm} \big(1 - e^{-\lambda_\pm |x|}\big)\nu(\mathrm{d}x) \in [0, \infty). \tag{13}$$

(Note that $\mu_{p\boldsymbol{\lambda}} - p\mu_{\boldsymbol{\lambda}} = p\big(\gamma_{\boldsymbol{\lambda}}^{(+)} + \gamma_{\boldsymbol{\lambda}}^{(-)}\big) - \big(\gamma_{p\boldsymbol{\lambda}}^{(+)} + \gamma_{p\boldsymbol{\lambda}}^{(-)}\big)$; see (4).)

We need the following elementary lemma to analyse the computational complexity of SB-Alg with rejection sampling.

Lemma 3.1. *(a) Let ℓ be a stick-breaking process on* [0,1]*; then for any $n \in \mathbb{N}$ we have*

$$0 \le n + \int_0^1 \frac{1}{x}\big(e^{cx} - 1\big)\mathrm{d}x - \sum_{k=1}^n \mathbb{E}\big[e^{c\ell_k}\big] \le 2^{-n} \int_0^1 \frac{1}{x}\big(e^{cx} - 1\big)\mathrm{d}x. \tag{14}$$

(b) We have $c^{-1}e^{-c}\big(1 + c^2\big) \int_0^1 x^{-1}\big(e^{cx} - 1\big)\mathrm{d}x \to 1$ as either $c \to 0$ or $c \to \infty$.

Assume that the simulation of the increments $Y_t^{(\pm)}$ under $\mathbb{P}_0$ has constant cost independent of the time horizon t (we also assume without loss of generality that the simulation of uniform random variables and the evaluation of operators such as sums and products and of the exponential functions have constant cost). Since SB-Alg requires the rejection sampling to be carried out over random stick-breaking lengths, the expected cost of SB-Alg with rejection sampling is, by (14) in Lemma 3.1, asymptotically proportional to

$$\sum_{k=1}^n \mathbb{E}\big[e^{\gamma_{\boldsymbol{\lambda}}^{(+)}\ell_k} + e^{\gamma_{\boldsymbol{\lambda}}^{(-)}\ell_k}\big] = 2n + \big(1 + \mathcal{O}(2^{-n})\big) \int_0^1 \frac{1}{x}\big(e^{\gamma_{\boldsymbol{\lambda}}^{(+)}Tx} + e^{\gamma_{\boldsymbol{\lambda}}^{(-)}Tx} - 2\big)\mathrm{d}x, \quad \text{as } n \to \infty. \tag{15}$$

In fact, by Lemma 3.1(a), the absolute value of the term $\mathcal{O}(2^{-n})$ is bounded by 2^{-n}. Moreover, by Lemma 3.1(b), the integral in (15) may be replaced with an explicit expression

$$\Gamma_{\boldsymbol{\lambda}}(T) := \frac{\gamma_{\boldsymbol{\lambda}}^{(+)}T}{1 + \big(\gamma_{\boldsymbol{\lambda}}^{(+)}T\big)^2} e^{\gamma_{\boldsymbol{\lambda}}^{(+)}T} + \frac{\gamma_{\boldsymbol{\lambda}}^{(-)}T}{1 + \big(\gamma_{\boldsymbol{\lambda}}^{(-)}T\big)^2} e^{\gamma_{\boldsymbol{\lambda}}^{(-)}T}, \tag{16}$$

as T tends to either 0 or ∞.

TABLE 2. Level variance and expected cost of SB-Alg and TSB-Alg, up to multiplicative constants that do not depend on the time horizon T. The bounds on the level variances (defined in the first paragraph of Subsection 3.2 above) follow from [18, Theorem 2] for SB-Alg and Proposition 3.1 for TSB-Alg.

Algorithm	Bias	Level variance	Cost
SB-Alg	$e^{-\vartheta_1 n}\big(\sqrt{T}+T\big)$	$e^{-\log(2)n}\big(T+T^2\big)$	$2n + \Gamma_\lambda(T)$
TSB-Alg	$e^{-\vartheta_1 n}\big(\sqrt{T}+T\big)$	$e^{-\log(2)n}\big(T+T^2\big)e^{(\mu_{2\lambda}-2\mu_\lambda)T}$	$2n$

Table 2 shows how SB-Alg with rejection sampling compares to TSB-Alg above.

Regime (I). By Table 2, we can deduce that the MC and MLMC estimators of both algorithms have the complexities $\mathcal{O}\big(\varepsilon^{-2}\log(1/\varepsilon)\big)$ and $\mathcal{O}\big(\varepsilon^{-2}\big)$ as $\varepsilon \to 0$, respectively, for all the payoffs considered in Proposition 3.1.

Regime (II). Assume ε is fixed. The biases of the two algorithms agree and equal $\mathcal{O}\big(e^{-\vartheta_1 n}\big(\sqrt{T}+T\big)\big)$, implying $n = \log\big((\sqrt{T}+T)/\varepsilon\big)/\vartheta_1 + \mathcal{O}(1)$ (with ϑ_1 defined in (23)). The level variances of SB-Alg and TSB-Alg are $\mathcal{O}\big(e^{-\log(2)n}\big(T+T^2\big)\big)$ and $\mathcal{O}\big(e^{-\log(2)n}\big(T+T^2\big)e^{(\mu_{2\lambda}-2\mu_\lambda)T}\big)$, with costs $\mathcal{O}(n+\Gamma_\lambda(T))$ and $\mathcal{O}(n)$, respectively. Thus, by (16), the ratios of the level variance and cost converge to 1 as $T \to 0$, so the ratio of the complexities of both algorithms converges to 1, implying that one should use TSB-Alg, as its performance in this regime is no worse than that of the other algorithm, but its implementation is simpler. For moderate or large values of T, by [16, Equation (A.3)], the computational complexity of the MLMC estimator based on TSB-Alg is proportional to $\varepsilon^{-2}e^{(\mu_{2\lambda}-2\mu_\lambda)T}(T+T^2)$ and that of SB-Alg is proportional to $\varepsilon^{-2}(1+\Gamma_\lambda(T))(T+T^2)$, where, in both cases, the proportionality constants do not depend on the time horizon T. Since both constants are exponential in T, for large T we need only compare $\max\big\{\gamma_\lambda^{(+)}, \gamma_\lambda^{(-)}\big\}$ with $\mu_{2\lambda}-2\mu_\lambda$. Indeed, TSB-Alg has a smaller complexity than SB-Alg for all sufficiently large T if and only if $\mu_{2\lambda}-2\mu_\lambda < \max\big\{\gamma_\lambda^{(+)}, \gamma_\lambda^{(-)}\big\}$. In Subsection 4.3.1 below, we provide an explicit criterion for the tempered stable process in terms of the parameters; see Figure 8.

In conclusion, when X has jumps of finite variation, it is preferable to use TSB-Alg if T is small or $e^{(\mu_{2\lambda}-2\mu_\lambda)T} < 1 + \Gamma_\lambda(T)$. Moreover, this typically holds if the Blumenthal–Getoor index of X is larger than $\log_2(3/2) < 0.6$; see Subsection 4.3.1 and Figure 8 below for details.

3.4.3. *Comparison with SBG-Alg.* Given any cutoff level $\kappa \in (0, 1]$, the algorithm SBG-Alg approximates the Lévy process X (under $\mathbb{P}_\lambda$ with the generating triplet $\big(\sigma^2, \nu_\lambda, b_\lambda\big)$) with a jump-diffusion $X^{(\kappa)}$ and samples exactly the vector $\big(X_T^{(\kappa)}, \overline{X}_T^{(\kappa)}, \tau_T(X^{(\kappa)})\big)$; see [16] for details. The bias of SBG-Alg is $\mathcal{O}\big(\min\{\sqrt{T}\kappa^{1-\beta_*/2}, \kappa\}\big(1+\log^+(T\kappa^{-\beta_*})\big)\big)$ with corresponding expected cost $\mathcal{O}\big(T\kappa^{-\beta_*}\big)$. Thus, when parametrised by computational effort n, the bias of SBG-Alg is $\mathcal{O}\big(T^{1/\beta_*}n^{-1/\beta_*}\log n\big)$ while the bias of TSB-Alg is $\mathcal{O}\big(e^{-\vartheta_1 n}\big(\sqrt{T}+T\big)\big)$.

Regime (I). The complexities of the MC and MLMC estimators based on SBG-Alg are, by [16, Theorem 22], equal to $\mathcal{O}\big(\varepsilon^{-2-\beta_*}\big)$ and $\mathcal{O}\big(\varepsilon^{-\max\{2,2\beta_*\}}\big(1+\mathbb{1}_{\{\beta_*=1\}}\log(\varepsilon)^2\big)\big)$, respectively, where β_*, defined in (21) below, is *slightly* larger than the Blumenthal–Getoor index of X. Since, by Subsection 3.2 above, the complexities of the MC and MLMC estimators based on TSB-Alg are $\mathcal{O}\big(\varepsilon^{-2}\log(1/\varepsilon)\big)$ and $\mathcal{O}\big(\varepsilon^{-2}\big)$, respectively, the complexity of TSB-Alg is always dominated by that of SBG-Alg as $\varepsilon \to 0$.

Regime (II). Suppose $\varepsilon > 0$ is fixed. Then, as in Subsection 3.4.2 above, the computational complexity of the MLMC estimator based on TSB-Alg is $\mathcal{O}\big(\varepsilon^{-2}(T+T^2)e^{(\mu_{2\lambda}-2\mu_\lambda)T}\big)$,

where the multiplicative constant does not depend on the time horizon T. By [16, Theorems 3 and 22] and [16, Equation (A.3)], the complexity of the MLMC estimator based on SBG-Alg is $\mathcal{O}\big(\varepsilon^{-2}\big(C_1 T + C_2 T^2\big)\big)$ if $\beta_* < 1$, $\mathcal{O}\big(\varepsilon^{-2}\big(C_1 T + C_2 T^2 \log^2(1/\varepsilon)\big)\big)$ if $\beta_* = 1$, and $\mathcal{O}\big(\varepsilon^{-2}\big(C_1 T + C_2 T^2 \varepsilon^{-2(\beta_*-1)}\big)\big)$ otherwise, where

$$C_1 := e^{r\beta_*}/\big(1 - e^{r(\beta_*/2-1)}\big)^2,$$
$$C_2 := e^{r\beta_*}/\big(1 - e^{r(\beta_*-1)}\big)^2 \cdot \mathbb{1}_{\{\beta_* \neq 1\}} + (e/2)^2 \cdot \mathbb{1}_{\{\beta_* = 1\}}, \text{ and}$$
$$r := (2/|\beta_* - 1|) \log(1 + |\beta_* - 1|/\beta_*) \cdot \mathbb{1}_{\{\beta_* \neq 1\}} + 2 \cdot \mathbb{1}_{\{\beta_* = 1\}}.$$

None of the other multiplicative constants in these bounds depend on the time horizon T. Thus TSB-Alg outperforms SBG-Alg if and only if we are in one of the following cases:

- $\beta_* < 1$ and $(1 + T)e^{(\mu_{2\lambda} - 2\mu_\lambda)T} < C_1 + C_2 T$,

- $\beta_* = 1$ and $(1 + T)e^{(\mu_{2\lambda} - 2\mu_\lambda)T} < C_1 + C_2 T \log^2(1/\varepsilon)$,

- $\beta_* > 1$ and $(1 + T)e^{(\mu_{2\lambda} - 2\mu_\lambda)T} < C_1 + C_2 T \varepsilon^{-2(\beta_*-1)}$.

Note that the constant C_2 is unbounded for β_* close to 1, favouring TSB-Alg.

In conclusion, TSB-Alg is simpler than SBG-Alg [16] and outperforms it asymptotically as $\varepsilon \to 0$. Moreover, TSB-Alg performs better than SBG-Alg for a fixed accuracy $\varepsilon > 0$ if either (I) $\beta_* < 1$ and the time horizon $T \ll 1$ satisfies the inequality $T < \log(C_1)/(\mu_{2\lambda} - 2\mu_\lambda)$ or (II) $\beta_* \geq 1$ and T is not large. In Subsection 4.3.2 below, we apply this criterion to tempered stable processes; see Figure 9 for the case (I) with $\beta_* < 1$ and Figure 10 for the case (II) with $\beta_* \geq 1$.

3.5. Variance reduction via exponential martingales

It follows from Subsection 3.4 that the performance of TSB-Alg deteriorates if the expectation $\mathbb{E}_{\mathbf{0}}\big[\Upsilon_\lambda^2\big] = \exp\big((\mu_{2\lambda} - 2\mu_\lambda)T\big)$ is large, making the variance of the estimator large. This problem is mitigated by using the control variates method, a variance reduction technique, based on exponential $\mathbb{P}_{\mathbf{0}}$-martingales, at (almost) no additional computational cost.

Suppose we are interested in estimating $\mathbb{E}_\lambda[\zeta] = \mathbb{E}_{\mathbf{0}}[\zeta\Upsilon_\lambda]$, where ζ is either $g(\chi_n)$ (MC) or $g(\chi_n) - g(\chi_{n-1})$ (MLMC). We propose to draw samples of $\tilde{\zeta}$ under $\mathbb{P}_{\mathbf{0}}$, instead of $\zeta\Upsilon_\lambda$, where

$$\tilde{\zeta} := \zeta\Upsilon_\lambda - w_0(\Upsilon_\lambda - 1) - w_+(\Upsilon_{\lambda_+} - 1) - w_-(\Upsilon_{\lambda_-} - 1),$$

and $w_0, w_+, w_- \in \mathbb{R}$ are constants to be determined (recall $\lambda_+ = (\lambda_+, 0)$ and $\lambda_- = (0, \lambda_-)$). Clearly $\mathbb{E}_{\mathbf{0}}[\tilde{\zeta}] = \mathbb{E}_{\mathbf{0}}[\zeta\Upsilon_\lambda]$ since the variables Υ_λ, Υ_{λ_+}, and Υ_{λ_-} have unit mean under $\mathbb{P}_{\mathbf{0}}$. We choose w_0, w_+, and w_- to minimise the variance of $\tilde{\zeta}$, by minimising the quadratic form [15, Section 4.1.2] of $\mathbb{V}_{\mathbf{0}}[\tilde{\zeta}]$:

$$\begin{pmatrix} -1 \\ \omega_0 \\ \omega_+ \\ \omega_- \end{pmatrix}^\top \begin{pmatrix} \mathbb{V}_{\mathbf{0}}[\zeta\Upsilon_\lambda] & \mathrm{cov}_{\mathbf{0}}(\zeta\Upsilon_\lambda, \Upsilon_\lambda) & \mathrm{cov}_{\mathbf{0}}(\zeta\Upsilon_\lambda, \Upsilon_{\lambda_+}) & \mathrm{cov}_{\mathbf{0}}(\zeta\Upsilon_\lambda, \Upsilon_{\lambda_-}) \\ \mathrm{cov}_{\mathbf{0}}(\zeta\Upsilon_\lambda, \Upsilon_\lambda) & \mathbb{V}_{\mathbf{0}}[\Upsilon_\lambda] & \mathrm{cov}_{\mathbf{0}}(\Upsilon_\lambda, \Upsilon_{\lambda_+}) & \mathrm{cov}_{\mathbf{0}}(\Upsilon_\lambda, \Upsilon_{\lambda_-}) \\ \mathrm{cov}_{\mathbf{0}}(\zeta\Upsilon_\lambda, \Upsilon_{\lambda_+}) & \mathrm{cov}_{\mathbf{0}}(\Upsilon_\lambda, \Upsilon_{\lambda_+}) & \mathbb{V}_{\mathbf{0}}[\Upsilon_{\lambda_+}] & \mathrm{cov}_{\mathbf{0}}(\Upsilon_{\lambda_+}, \Upsilon_{\lambda_-}) \\ \mathrm{cov}_{\mathbf{0}}(\zeta\Upsilon_\lambda, \Upsilon_{\lambda_-}) & \mathrm{cov}_{\mathbf{0}}(\Upsilon_\lambda, \Upsilon_{\lambda_-}) & \mathrm{cov}_{\mathbf{0}}(\Upsilon_{\lambda_+}, \Upsilon_{\lambda_-}) & \mathbb{V}_{\mathbf{0}}[\Upsilon_{\lambda_-}] \end{pmatrix} \begin{pmatrix} -1 \\ \omega_0 \\ \omega_+ \\ \omega_- \end{pmatrix}.$$

The solution, in terms of the covariance matrix (under $\mathbb{P}_{\mathbf{0}}$) of the vector $(\zeta\Upsilon_\lambda, \Upsilon_\lambda, \Upsilon_{\lambda_+}, \Upsilon_{\lambda_-})$, is

$$\begin{pmatrix} w_0 \\ w_+ \\ w_- \end{pmatrix} = \begin{pmatrix} 2\mathbb{V}_{\mathbf{0}}[\Upsilon_\lambda] & \mathrm{cov}_{\mathbf{0}}(\Upsilon_\lambda, \Upsilon_{\lambda_+}) & \mathrm{cov}_{\mathbf{0}}(\Upsilon_\lambda, \Upsilon_{\lambda_-}) \\ \mathrm{cov}_{\mathbf{0}}(\Upsilon_\lambda, \Upsilon_{\lambda_+}) & 2\mathbb{V}_{\mathbf{0}}[\Upsilon_{\lambda_+}] & \mathrm{cov}_{\mathbf{0}}(\Upsilon_{\lambda_+}, \Upsilon_{\lambda_-}) \\ \mathrm{cov}_{\mathbf{0}}(\Upsilon_\lambda, \Upsilon_{\lambda_-}) & \mathrm{cov}_{\mathbf{0}}(\Upsilon_{\lambda_+}, \Upsilon_{\lambda_-}) & 2\mathbb{V}_{\mathbf{0}}[\Upsilon_{\lambda_-}] \end{pmatrix}^{-1} \begin{pmatrix} \mathrm{cov}_{\mathbf{0}}(\zeta\Upsilon_\lambda, \Upsilon_\lambda) \\ \mathrm{cov}_{\mathbf{0}}(\zeta\Upsilon_\lambda, \Upsilon_{\lambda_+}) \\ \mathrm{cov}_{\mathbf{0}}(\zeta\Upsilon_\lambda, \Upsilon_{\lambda_-}) \end{pmatrix}.$$

In practice, these covariances are estimated from the same samples that were drawn to estimate $\mathbb{E}_{\mathbf{0}}[\zeta \Upsilon_\lambda]$. The additional cost is (nearly) negligible as all the variables in the exponential martingales are by-products of TSB-Alg. It is difficult to establish theoretical guarantees for the improvement of $\tilde{\zeta}$ over $\zeta \Upsilon_\lambda$. However, since most of the variance of the estimator based on $\zeta \Upsilon_\lambda$ comes from Υ_λ, the proposal $\tilde{\zeta}$ typically performs well in applications; see e.g. the CIs in Figures 3 and 4 for a tempered stable process.

4. Extrema of tempered stable processes

4.1. Description of the model

In the present section we apply our results to the tempered stable process X. More precisely, given $\lambda \in \mathbb{R}_+^2 \setminus \{\mathbf{0}\}$, the tempered stable Lévy measure ν_λ specifies the Lévy measure ν via (1):

$$\frac{\nu(\mathrm{d}x)}{\mathrm{d}x} = \frac{c_+}{x^{\alpha_+ + 1}} \cdot \mathbb{1}_{(0,\infty)}(x) + \frac{c_-}{|x|^{\alpha_- + 1}} \cdot \mathbb{1}_{(-\infty,0)}(x), \tag{17}$$

where $c_\pm \geq 0$ and $\alpha_\pm \in (0, 2)$. The drift b is given by the tempered stable drift $b_\lambda \in \mathbb{R}$ via (1), and the constant μ_λ is given in (4). Both b and μ_λ can be computed using (18) and (19) below. TSB-Alg samples from the distribution functions $F^{(\pm)}(t, x) = \mathbb{P}_{\mathbf{0}}\big(Y_t^{(\pm)} \leq x\big)$, where $Y^{(\pm)}$ are the spectrally one-sided stable processes defined in Theorem 2.1, using the Chambers–Mallows–Stuck algorithm [9]. In the appendix we include a version of this algorithm, given explicitly in terms of the drift b and the parameters in the Lévy measure ν; see Algorithm 2 below.

Next, we provide explicit formulae for b and μ_λ in terms of special functions. We begin by expressing b in terms of b_λ (see (1) above):

$$b = b_\lambda - c_+ B_{\alpha_+, \lambda_+} + c_- B_{\alpha_-, \lambda_-}, \quad \text{where} \quad B_{a,r} := \int_0^1 \big(e^{-rx} - 1\big)x^{-a}\mathrm{d}x. \tag{18}$$

We have $B_{a,0} = 0$ for any $a \geq 0$ and, for $r > 0$,

$$B_{a,r} = \sum_{n=1}^\infty \frac{(-r)^n}{n!(n-a-1)} = \begin{cases} \big(e^{-r} - 1 + r^{a-1}\gamma(2-a, r)\big)/(1-a), & a \in (0, 2) \setminus \{1\}, \\ -\gamma - \log r - E_1(r), & a = 1, \end{cases}$$

where $\gamma(a, r) = \int_0^r e^{-x}x^{a-1}\mathrm{d}x = \sum_{n=0}^\infty (-1)^n r^{n+a}/(n!(n+a))$, $a > 0$, is the lower incomplete gamma function, $E_1(r) = \int_r^\infty e^{-x}x^{-1}\mathrm{d}x$, $r > 0$, is the exponential integral, and $\gamma = 0.57721566\ldots$ is the Euler–Mascheroni constant. Similarly, to compute μ_λ from (4), note that

$$\mu_\lambda = c_+ C_{\alpha_+, \lambda_+} + c_- C_{\alpha_-, \lambda_-}, \quad \text{where} \quad C_{a,r} := \int_0^\infty \big(e^{-rx} - 1 + rx \cdot \mathbb{1}_{(0,1)}(x)\big)x^{-a-1}\mathrm{d}x. \tag{19}$$

Clearly, $C_{a,0} = 0$ for any $a \geq 0$ and, for $r > 0$,

$$C_{a,r} = -\frac{1}{a}\big(e^{-r} + r(1 + B_{a,r})\big) + r^a \Gamma(-a, r),$$

where $\Gamma(a, r) = \int_r^\infty e^{-x}x^{a-1}dx$ is the upper incomplete gamma function. When $a < 1$, this simplifies to $C_{a,r} = r^a \Gamma(-a) + r/(1-a)$, where $\Gamma(\cdot)$ is the gamma function.

As discussed in Subsection 3.4 (see also Table 2 above), the performance of TSB-Alg deteriorates for large values of the constant $\mu_{2\lambda} - 2\mu_\lambda$. As a consequence of the formulae above,
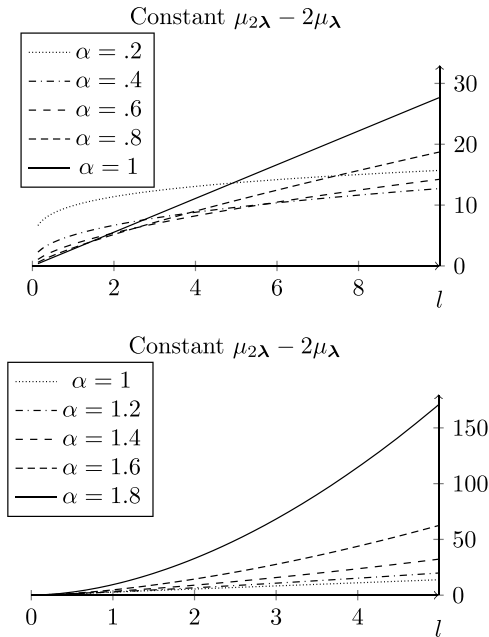
FIGURE 1. The picture shows the map $l \mapsto \mu_\lambda$, where $c_+ = c_- = 1$, $\alpha_+ = \alpha_- = \alpha$, $\lambda = (l, l)$, and $\alpha \in \{0.2, \ldots, 1.8\}$.

it is easy to see that this constant is proportional to $\lambda_+^{\alpha_+}/(2 - \alpha_+) + \lambda_-^{\alpha_-}/(2 - \alpha_-)$ as either $\lambda_\pm \to \infty$ or $\alpha_\pm \to 2$; see Figure 1.

It is natural to expect the performance of TSB-Alg to deteriorate as $\lambda_\pm \to \infty$ or $\alpha_\pm \to 2$, because the variance of the Radon–Nikodym derivative Υ_λ tends to ∞, making the variance of all the estimators very large.

4.2. Monte Carlo and multilevel Monte Carlo estimators for tempered stable processes

Let X denote the tempered stable process with generating triplet $(\sigma^2, \nu_\lambda, b_\lambda)$ given in Subsection 4.1. Then $S_t = S_0 e^{X_t}$ models the risky asset for some initial value $S_0 > 0$ and all $t \geq 0$. Observe that $\overline{S}_t = S_0 e^{\overline{X}_t}$ and that the time at which S attains its supremum on $[0,t]$ is also τ_t. Since the tails $\nu(\mathbb{R} \setminus (-x, x))$ of ν decay polynomially as $x \to \infty$ (see (17)), S_t and $\overline{S}_t$ satisfy the following moment conditions under $\mathbb{P}_\lambda$: $\mathbb{E}_\lambda[S_t^r] < \infty$ (resp. $\mathbb{E}_\lambda[\overline{S}_t^r] < \infty$) if and only if $r \in [-\lambda_-, \lambda_+]$ (resp. $r \leq \lambda_+$). In the present subsection we apply TSB-Alg to estimate $\mathbb{E}_\lambda[g(\chi)]$ using the MC and MLMC estimators in (7) and (8), respectively, for payoffs g that arise in applications and calibrated/estimated values of tempered stable model parameters.

4.2.1. *Monte Carlo estimators* Consider the estimator in (7) for the following payoff functions g: (I) the payoff of the up-and-out barrier call option $g(\chi) = \max\{S_T - K, 0\}\mathbb{1}\{\overline{S}_T \leq M\}$ and (II) the function $g(\chi) = (S_T/\overline{S}_T - 1)^2$ associated to the ulcer index (UI) given by $100\sqrt{\mathbb{E}_\lambda g(\chi)}$ (see [11]). We plot the estimated value along with a band representing the empirically constructed CIs of level 95% via Theorem 3.1; see Figures 2 and 3. We set the initial value $S_0 = 100$, assume $b_\lambda = 0$, and consider multiple time horizons T (given as fractions of a calendar year): we take $T \in \{\frac{30}{365}, \frac{90}{365}\}$ in (I) and $T \in \{\frac{14}{365}, \frac{28}{365}\}$ in (II). These time horizons are common in their respective applications, (I) option pricing and (II) real-world risk assessment
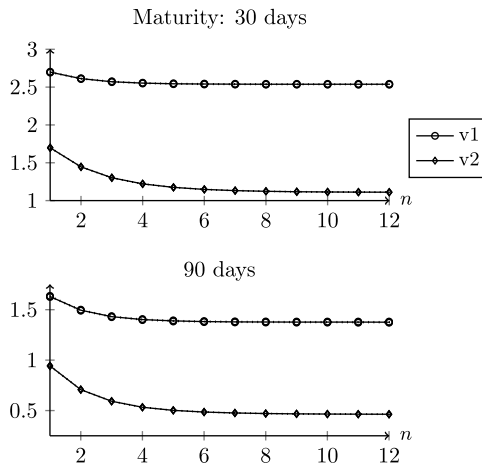
FIGURE 2. *Solid* lines indicate the estimated value of $\mathbb{E}_\lambda[g(\chi_n)]$ for the USD/JPY currency pair and the function $g(\chi) = \max\{S_T - K, 0\}\mathbb{1}\{\overline{S}_T \leq M\}$ using $N = 10^6$ samples with $T \in \left\{\frac{30}{365}, \frac{90}{365}\right\}$, $K = 95$, and $M = 102$. *Dotted* lines (the symmetric bands around each solid line) indicate the CI of the estimates with confidence level 95%. These CIs are very tight, making them hard to discern in the plot.
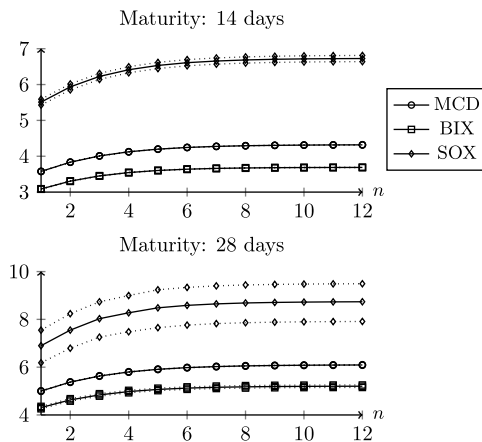

FIGURE 3. *Solid* lines indicate the estimated value of the ulcer index (UI) $100\sqrt{\mathbb{E}_\lambda[g(\chi_n)]}$ using $N = 10^7$ samples, where $g(\chi) = (S_T/\overline{S}_T - 1)^2$. *Dotted* lines (the bands around each solid line) indicate the 95% CI of the estimates.

(see e.g. [1, 7]). The values of the other parameters, given in Table 3 below, are either calibrated or estimated: in (I) we are pricing an option and thus use calibrated parameters with respect to the risk-neutral measure obtained in [1, Table 3] for the USD/JPY foreign exchange (FX) rate, and in (II) we are assessing risks under the real-world probability measure and hence use the maximum likelihood estimates in [7, Table 2] for various time series of historic asset returns. In both (I) and (II), it is assumed that $\alpha_\pm = \alpha$; (II) additionally assumes that $c_+ = c_-$ (i.e., X is a CGMY process). The stocks MCD, BIX, and SOX in (II) were chosen with $\alpha > 1$ to stress-test the performance of TSB-Alg when $\mu_{2\lambda} - 2\mu_\lambda$ is big, forcing the variance of the estimator to be large; see the discussion at the end of Subsection 4.1 above and Figure 1.

TABLE 3. Calibrated/estimated parameters of the tempered stable model. The first two sets of parameters were calibrated in [1, Table 3] based on FX option data. The bottom three sets of parameters were estimated using the maximum likelihood estimators in [7, Table 2], based on a time series of equity returns.

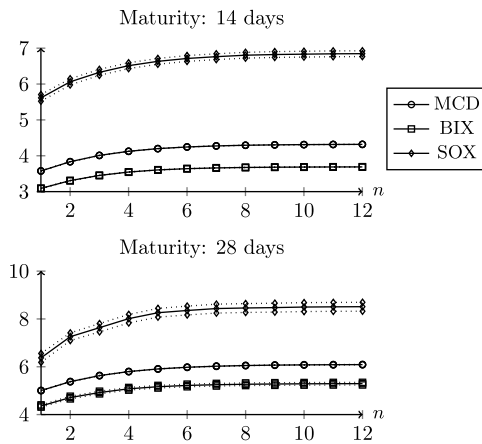| Stock	σ	α	c_+	c_-	λ_+	λ_-	$\mu_{2\lambda} - 2\mu_\lambda$
USD/JPY (v1)	0.0007	0.66	0.1305	0.0615	6.5022	3.0888	0.9658
USD/JPY (v2)	0.0001	1.5	0.0069	0.0063	1.932	0.4087	0.0395
MCD	0	1.50683	0.08	0.08	25.4	25.4	41.47
BIX	0	1.2341	0.32	0.32	37.42	47.76	96.6
SOX	0	1.3814	0.44	0.44	34.73	34.76	196.81

FIGURE 4. *Solid* lines indicate the estimated value of the ulcer index (UI) $100\sqrt{\mathbb{E}_\lambda g(\chi_n)}$ using $N = 10^7$ samples and the control variates method of Subsection 3.5, where $g(\chi) = (S_T/\overline{S}_T - 1)^2$. *Dotted* lines (the bands around each solid line) indicate the CI of the estimates with confidence level 95%.

The visible increasein the variance of the estimates for the SOX (see Figure 3) as the maturity increases from 14 to 28 days is a consequence of the unusually large value of $\mu_{2\lambda} - 2\mu_\lambda$; see Table 3 above. If $\mu_{2\lambda} - 2\mu_\lambda$ is not large, we may take long time horizons and relatively few samples to obtain an accurate MC estimate. In fact this appears to be the case for calibrated risk-neutral parameters (we were unable to find calibrated parameter values in the literature with a large value of $\mu_{2\lambda} - 2\mu_\lambda$). However, if $\mu_{2\lambda} - 2\mu_\lambda$ is large, then for any moderately large time horizon, an accurate MC estimate would require a large number of samples. In such cases, the control variates method from Subsection 3.5 becomes very useful.

To illustrate the added value of the control variates method from Subsection 3.5, we apply it in the setting of Figure 3, where we observed the widest CIs. (Recall that the CIs derived in this paper do not account for model uncertainty or the uncertainty in the estimation or calibration of the parameters.) Figure 4 displays the resulting estimators and CIs, showing that this method is beneficial in the case where the variance of the Radon–Nikodym derivative Υ_λ is large, i.e., when $(\mu_{2\lambda} - 2\mu_\lambda)T$ is large. The CIs for the SOX asset at $n = 12$ shrank by a factor of 4.23; in other words, the variance became 5.58% of its original value.

4.2.2. *Multilevel Monte Carlo estimators.* We will consider the MLMC estimator in (8) with parameters $n, N_1, \ldots, N_n \in \mathbb{N}$ given by [16, Equations (A.1)–(A.2)]. In this example we
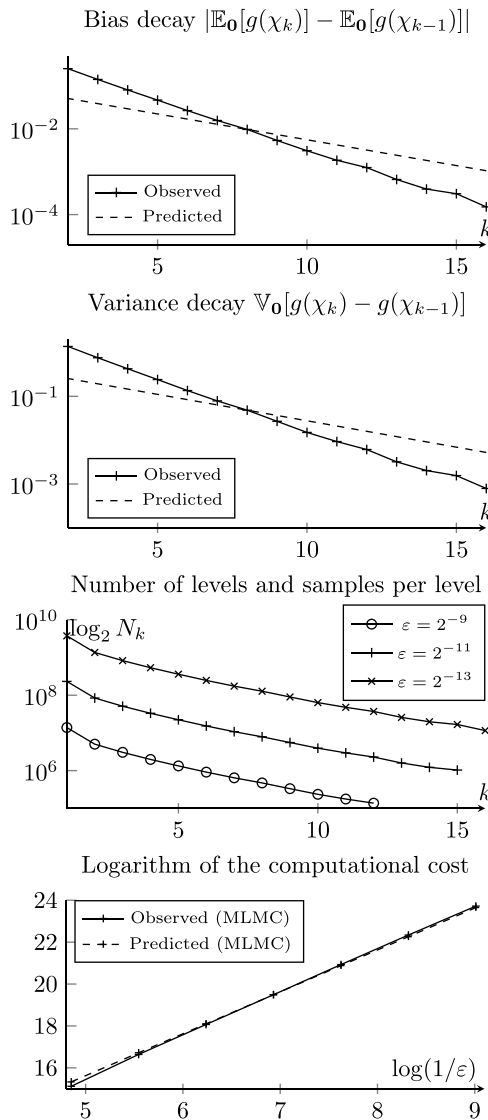
Bias decay $|\mathbb{E}_\mathbf{0}[g(\chi_k)] - \mathbb{E}_\mathbf{0}[g(\chi_{k-1})]|$

Variance decay $\mathbb{V}_\mathbf{0}[g(\chi_k) - g(\chi_{k-1})]$

Number of levels and samples per level

Logarithm of the computational cost

FIGURE 5. The top two panels show the bias and level variance decay as a function of k for the payoff $g(\chi) = \max\{S_T - K, 0\}\mathbb{1}\{\overline{S}_T \leq M\}$ with $S_0 = 100$, $T = 90/365$, $K = 95$, $M = 102$ and the parameter values for the USD/JPY FX rate (v2) in Table 3. The theoretical predictions (*dashed*) are based on Proposition 3.1 for barrier-type-1 payoffs. The bottom two panels show the corresponding values of the complexities and parameters $n, N_1, \ldots, N_n$ associated to the precision levels $\varepsilon \in \{2^{-9}, 2^{-10}, \ldots, 2^{-19}\}$.

chose (I) the payoff $g(\chi) = \max\{S_T - 95, 0\}\mathbb{1}\{\overline{S}_T \leq 102\}$ from Subsection 4.2.1 above and (II) the payoff $g(\chi) = (S_T/\overline{S}_T - 1)^2\mathbb{1}_{\{\tau_T < T/2\}}$, associated to the modified ulcer index (MUI) $100\sqrt{\mathbb{E}_\lambda[g(\chi)]}$, a risk measure which weighs trends more heavily than short-time fluctuations (see [16] and the references therein).

The payoff in (I) is that of a barrier up-and-out option, so it is natural to use the risk-neutral parameters for the USD/JPY FX rate (see (v2) in Table 3) over the time horizon $T = 90/365$. For (II) we take the parameter values of the MCD stock in Table 3 with $T = 28/365$. In both
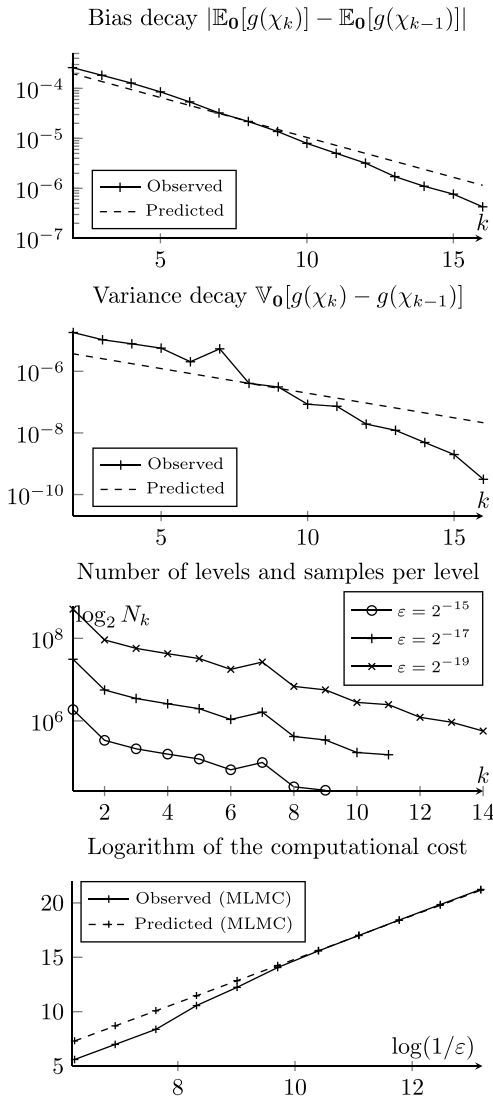
Bias decay $|\mathbb{E}_\mathbf{0}[g(\chi_k)] - \mathbb{E}_\mathbf{0}[g(\chi_{k-1})]|$

Variance decay $\mathbb{V}_\mathbf{0}[g(\chi_k) - g(\chi_{k-1})]$

Number of levels and samples per level

Logarithm of the computational cost

FIGURE 6. The top two panels show the bias and level variance decay as a function of k for the payoff $g(\chi) = (S_T/\overline{S}_T - 1)^2 \mathbb{1}_{\{\tau_T < T/2\}}$ with $S_0 = 100$, $T = 28/365$ and the parameter values for MCD in Table 3. The theoretical predictions (*dashed*) are based on Proposition 3.1 for barrier-type-2 payoffs. The bottom two panels show the corresponding values of the complexities and parameters $n, N_1, \ldots, N_n$ associated to the precision levels $\varepsilon \in \left\{2^{-9}, 2^{-10}, \ldots, 2^{-13}\right\}$.

cases we set $S_0 = 100$. Figure 5 (resp. Figure 6) shows the decay of the bias and level variance, the corresponding value of the constants $n, N_1, \ldots, N_n$, and the growth of the complexity for the first (resp. second) payoff.

In Figure 7 we plot the estimator $\hat{\theta}_{\mathrm{MC}}^{g,n}$ in Theorem 3.1 for (I), parametrised by $\varepsilon \to 0$. To further illustrate the CLT in Theorem 3.1, the figure also shows the CIs of confidence level 95% constructed using the CLT.
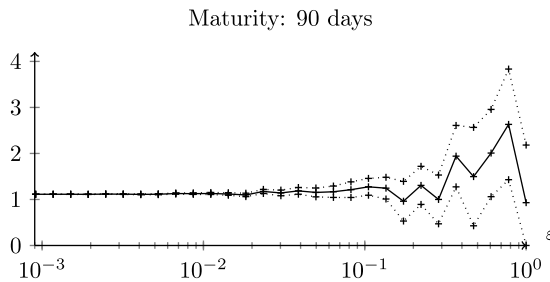
Maturity: 90 days

FIGURE 7. The *solid* line indicates the estimated value of the expectation $\mathbb{E}_\lambda g(\chi_n)$ for the payoff $g(\chi) = \max\{S_T - K, 0\}\mathbb{1}\{\overline{S}_T \leq M\}$ with $S_0 = 100$, $T = 90/365$, $K = 95$, $M = 102$ and the parameter values for the USD/JPY FX rate (v2) in Table 3. We use the MLMC estimator $\hat{\theta}_{\mathrm{ML}}^{g,n}$ in (8), and the CIs (*dotted* lines) are constructed using Theorem 3.1 with confidence level 95%.
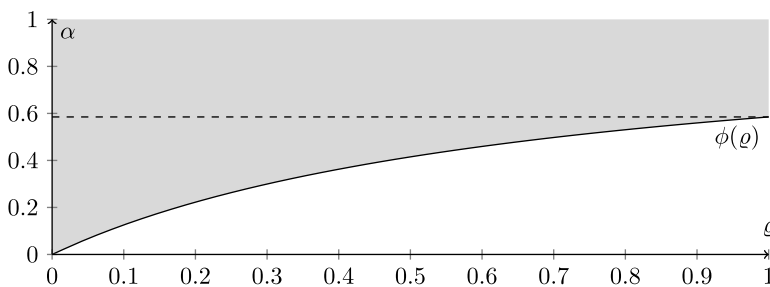

FIGURE 8. The picture shows the map $\varrho \mapsto \phi(\varrho)$, $\varrho \in [0, 1]$. Assuming $\alpha_\pm = \alpha$ and defining $\varrho := \min\{c_+\lambda_+^\alpha, c_-\lambda_-^\alpha\}/\max\{c_+\lambda_+^\alpha, c_-\lambda_-^\alpha\}$, TSB-Alg is preferable to SB-Alg when (ϱ, α) lies in the shaded region.

4.3. Comparing TSB-Alg with existing algorithms for tempered stable processes

In this subsection we take the analysis from Subsection 3.4 and apply it to the tempered stable case.

4.3.1. *Comparison with SB-Alg* Recall from Subsection 3.4.2 that SB-Alg is only applicable when $\alpha_\pm < 1$, and under Regime (II), SB-Alg is preferable over TSB-Alg for all sufficiently large T if and only if $\max\{\gamma_\lambda^{(+)}, \gamma_\lambda^{(-)}\} \leq \mu_{2\lambda} - 2\mu_\lambda$, where $\gamma_\lambda^{(\pm)} = -c_\pm\lambda_\pm^{\alpha_\pm}\Gamma(-\alpha_\pm) \geq 0$ is defined in (13). By the formulae in Subsection 4.1, it is easily seen that $\max\{\gamma_\lambda^{(+)}, \gamma_\lambda^{(-)}\} \leq \mu_{2\lambda} - 2\mu_\lambda$ is equivalent to

$$\min\{c_+\lambda_+^{\alpha_+}\Gamma(-\alpha_+), c_-\lambda_-^{\alpha_-}\Gamma(-\alpha_-)\}$$
$$\geq c_+\lambda_+^{\alpha_+}(2 - 2^{\alpha_+})\Gamma(-\alpha_+) + c_-\lambda_-^{\alpha_-}(2 - 2^{\alpha_-})\Gamma(-\alpha_-).$$

Assuming that $\alpha_\pm = \alpha$, the inequality simplifies to $\alpha \leq \phi(\varrho)$, where we define $\phi(x) := \log_2\left(1 + \frac{x}{1+x}\right)$ and $\varrho := \min\{c_+\lambda_+^\alpha, c_-\lambda_-^\alpha\}/\max\{c_+\lambda_+^\alpha, c_-\lambda_-^\alpha\}$. In particular, a symmetric Lévy measure yields $\varrho = 1$ and $\phi(1) = \log_2(3/2) = 0.58496\ldots$, and a one-sided Lévy measure gives $\varrho = 0$ and $\phi(0) = 0$.

4.3.2. *Comparison with SBG-Alg* Recall from Subsection 3.4.3 that TSB-Alg is preferable to SBG-Alg when $\max\{\alpha_+, \alpha_-\} \geq 1$ (as it is equivalent to $\beta_* \geq 1$). On the other hand, if $\alpha_\pm < 1$
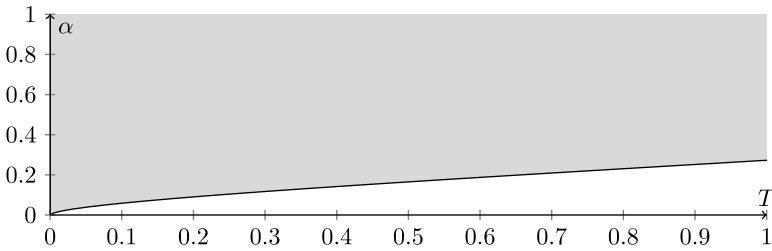
FIGURE 9. The shaded region is the set of points (T, α) where TSB-Alg is preferable to SBG-Alg assuming $\alpha_{\pm} = \alpha \in (0, 1)$ and all other parameters are as in the USD/JPY (v1) currency pair in Table 3.
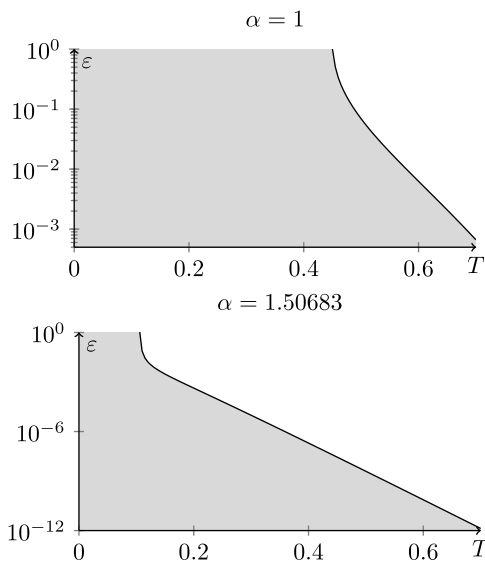

FIGURE 10. The shaded region is the set of points (T, ε) where TSB-Alg is preferable to SBG-Alg when $\alpha_{\pm} = \alpha \in [1, 2)$ and all other parameters are as in the MCD stock in Table 3.

(or equivalently, $\beta_* < 1$), TSB-Alg outperforms SB-Alg if and only if $(1 + T)e^{(\mu_{2\lambda} - 2\mu_\lambda)T} \leq C_1 + C_2 T$. For large enough T, SB-Alg will outperform TSB-Alg; however, it is generally hard to determine when this happens. In Figure 9 we illustrate the region of parameters (T, α) where TSB-Alg is preferable, assuming $\alpha_{\pm} = \alpha \in (0, 1)$ and all other parameters are as in the USD/JPY (v1) currency pair in Table 3.

4.4. Concluding remarks

TSB-Alg is an easily implementable algorithm for which optimal MLMC (and even unbiased) estimators exist. TSB-Alg combines the best of both worlds: it is applicable to *all* tempered stable processes (as is SBG-Alg in [16]), while preserving the *geometric* convergence of SB-Alg in [18]. The only downside of TSB-Alg is the enlarged variance by the factor $\exp((\mu_{2\lambda} - 2\mu_\lambda)T)$. This factor is typically small (see the discussion in Subsection 4.2.1 above), and when it is large, easily implementable variance reduction techniques exist (see details in Subsection 3.5 above). These facts favour the use of the MLMC estimator based

on TSB-Alg over its competitors when $(\mu_{2\lambda} - 2\mu_\lambda)T$ is not large (for more concrete rules of thumb, see the concluding paragraphs of Subsections 3.4.1, 3.4.2, and 3.4.3 above).

In practice, in the implementation of the MC and MLMC estimators of TSB-Alg, the leading constants of the bias and variance decay are hard to compute and often overestimate the error. The general practice in this situation (see e.g. [13, Section 2]) is to numerically estimate these constants along with the integers $n(\varepsilon)$, N, and $(N_k)_{k\in\mathbb{N}}$ in Proposition 3.2. Such estimation requires few simulations for the first few levels and some extrapolation but typically performs well in practice. This is particularly true in our setting as the MLMC estimator is optimal; see [13, Section 3] for a detailed discussion and a generic MATLAB implementation.

5. Proofs

Let us introduce the geometric rate ϑ_p used in Proposition 3.1 above. Let β be the *Blumenthal–Getoor index* [6], defined as

$$\beta := \inf\{p > 0 : I_0^p < \infty\}, \quad \text{where} \quad I_0^p := \int_{(-1,1)} |x|^p \nu(\mathrm{d}x), \quad \text{for any } p \geq 0, \qquad (20)$$

and note that $\beta \in [0, 2]$ since $I_0^2 < \infty$. Moreover, $I_0^1 < \infty$ if and only if the jumps of X have finite variation. In the case of (tempered) stable processes, β is the greatest of the two activity indices of the Lévy measure. Note that $I_0^p < \infty$ for any $p > \beta$ but I_0^β can be either finite or infinite. If $I_0^\beta = \infty$ we must have $\beta < 2$ and can thus pick $\delta \in (0, 2 - \beta)$, satisfying $\beta + \delta < 1$ whenever $\beta < 1$, and define

$$\beta_* := \beta + \delta \cdot \mathbb{1}_{\left\{I_0^\beta = \infty\right\}} \in [\beta, 2]. \qquad (21)$$

The index β_* is either equal to β or arbitrarily close to it. In either case we have $I_0^{\beta_*} < \infty$. Define $\alpha \in [\beta, 2]$ and $\alpha_* \in [\beta_*, 2]$ by

$$\alpha := 2 \cdot \mathbb{1}_{\sigma \neq 0} + \mathbb{1}_{\sigma = 0} \begin{cases} 1, & I_0^1 < \infty \text{ and } b_0 \neq 0, \\ \beta, & \text{otherwise,} \end{cases} \quad \text{and} \quad \alpha_* := \alpha + (\beta_* - \beta) \cdot \mathbb{1}_{\alpha = \beta}. \qquad (22)$$

Finally, we may define

$$\vartheta_p := \log\left(1 + \mathbb{1}_{p > \alpha} + \frac{p}{\alpha_*} \cdot \mathbb{1}_{p \leq \alpha}\right) \in (0, \log 2], \quad \text{for any} \quad p > 0, \qquad (23)$$

and note that $\vartheta_p \geq \log(3/2)$ for $p \geq 1$.

In order to prove Proposition 3.1 for barrier-type-1 payoffs, we need to ensure that $\overline{X}_T$ has a sufficiently regular distribution function under $\mathbb{P}_{p\lambda}$. The following assumption will help us establish that in certain cases of interest.

Assumption 5.1. *Under $\mathbb{P}_0$, the Lévy process $X = (X_t)_{t\in[0,T]}$ is in the domain of attraction of an α-stable process as $t \to 0$ with $\alpha \in (1, 2]$. Put differently, there exists a positive function g such that the law of $X_t/g(t)$, under $\mathbb{P}_0$, converges in distribution to an α-stable law for some $\alpha \in (1, 2]$ as $t \to 0$.*

When X is tempered stable, Assumption 5.1 holds trivially if $\max\{\alpha_+, \alpha_-\} > 1$ or $\sigma \neq 0$. The index α in Assumption 5.1 necessarily agrees with the one in (22); see [5, Subsection 2.1]. For further sufficient and necessary conditions for Assumption 5.1, we refer the reader to

[5, 21]. In particular, Assumption 5 remains satisfied if the Lévy measure ν is modified away from 0 or the law of X is changed under an equivalent change of measure; see [5, Subsection 2.3.4].

Lemma 5.1. *For any Borel set $A \subset \mathbb{R} \times \mathbb{R}_+ \times [0, T]$ and $p > 1$ we have*

$$\mathbb{P}_{\boldsymbol{\lambda}}(\chi \in A) \leq e^{(\mu_{p\boldsymbol{\lambda}} - p\mu_{\boldsymbol{\lambda}})T/p} \mathbb{P}_{\boldsymbol{0}}(\chi \in A)^{1-1/p}, \tag{24}$$

where the constants $\mu_{\boldsymbol{\lambda}}$ and $\mu_{p\boldsymbol{\lambda}}$ are defined in (4). *Moreover, if $I_0^1 < \infty$, then we also have*

$$\mathbb{P}_{\boldsymbol{\lambda}}(\chi \in A) \leq e^{\left(\gamma_{\boldsymbol{\lambda}}^{(+)} + \gamma_{\boldsymbol{\lambda}}^{(-)}\right)T} \mathbb{P}_{\boldsymbol{0}}(\chi \in A), \tag{25}$$

where the constants $\gamma_{\boldsymbol{\lambda}}^{(\pm)}$ are defined in (13).

The proofs of Lemma 5.1 and Proposition 3.1 rely on the identity $\Upsilon_{\boldsymbol{\lambda}}^p = \Upsilon_{p\boldsymbol{\lambda}} e^{(\mu_{p\boldsymbol{\lambda}} - p\mu_{\boldsymbol{\lambda}})T}$, which is valid for any $\boldsymbol{\lambda} \in \mathbb{R}_+^2$ and $p \geq 1$.

Proof of Lemma 5.1. Fix the Borel set A. By Theorem 2.1 and Hölder's inequality with $p > 1$, we get

$$\mathbb{P}_{\boldsymbol{\lambda}}(\chi \in A) = \mathbb{E}_{\boldsymbol{0}}\left[\Upsilon_{\boldsymbol{\lambda}} \mathbb{1}_{\{\chi \in A\}}\right] \leq \mathbb{E}_{\boldsymbol{0}}\left[\Upsilon_{\boldsymbol{\lambda}}^p\right]^{1/p} \mathbb{E}_{\boldsymbol{0}}\left[\mathbb{1}_{\{\chi \in A\}}^q\right]^{1/q} = e^{(\mu_{p\boldsymbol{\lambda}} - p\mu_{\boldsymbol{\lambda}})T/p} \mathbb{P}_{\boldsymbol{0}}(\chi \in A)^{1/q},$$

where $1/q = 1 - 1/p$, implying (24). If $I_0^1 < \infty$, then $\mu_{p\boldsymbol{\lambda}} - p\mu_{\boldsymbol{\lambda}} = p\left(\gamma_{\boldsymbol{\lambda}}^{(+)} + \gamma_{\boldsymbol{\lambda}}^{(-)}\right) - \left(\gamma_{p\boldsymbol{\lambda}}^{(+)} + \gamma_{p\boldsymbol{\lambda}}^{(-)}\right) \leq p\left(\gamma_{\boldsymbol{\lambda}}^{(+)} + \gamma_{\boldsymbol{\lambda}}^{(-)}\right)$. Thus, taking $p \to \infty$ (and hence $q \to 1$) in (24) yields (25). $\square$

Proof of Proposition 3.1. Theorem 2.1 implies that all the expectations in the statement of Proposition 3.1 can be replaced with the expectation $\mathbb{E}_{p\boldsymbol{\lambda}}[|g(\chi) - g(\chi_n)|]$. Since $\boldsymbol{\lambda} \neq \boldsymbol{0}$, implying that $\min\{\mathbb{E}_{p\boldsymbol{\lambda}}[\max\{X_t, 0\}], \mathbb{E}_{p\boldsymbol{\lambda}}[\max\{-X_t, 0\}]\} < \infty$, [18, Proposition 1] yields the result for Lipschitz payoffs. By the assumption in Proposition 3.1 for the locally Lipschitz case, the Lévy measure $\nu_{p\boldsymbol{\lambda}}$ in (1) satisfies the assumption in [18, Proposition 2], implying the result for locally Lipschitz payoffs. The result for barrier-type-2 payoffs follows from a direct application of [16, Lemmas 14 and 15] and [18, Theorem 2].

The result for barrier-type-1 payoffs follows from [18, Proposition 3 and Remark 6] if we show the existence of a constant K' satisfying $\mathbb{P}_{p\boldsymbol{\lambda}}(M < \overline{X}_T \leq M + x) \leq K'x^{\gamma}$ for all $x > 0$. If $\gamma \in (0, 1)$, such a K' exists by (24) in Lemma 5 above with $p = (1 - \gamma)^{-1} > 1$ and $A = \mathbb{R} \times (M, M + x] \times [0, T]$. If $\gamma = 1$ and $I_0^1 < \infty$, the existence of K' follows from the assumption in Proposition 3.1 and (25) in Lemma 5.1. If $\gamma = 1$ and Assumption 5.1 holds, then [5, Theorem 5.1] implies the existence of K'. $\square$

Lemma 5.2. *Let the payoff g be as in Proposition 3.1 with $p = 2$ and $\mathbb{E}_{\boldsymbol{0}}\left[g(\chi)^2 \Upsilon_{\boldsymbol{\lambda}}^2\right] < \infty$. Let $n = n(\varepsilon)$, N, and $N_1, \ldots, N_n$ be as in Proposition 3.2; then the following limits hold as $\varepsilon \to 0$:*

$$\varepsilon^2 N_k \to 2\sqrt{\mathbb{V}_{\boldsymbol{0}}[D_{k,1}^g]/k} \sum_{j=1}^{\infty} \sqrt{j \mathbb{V}_{\boldsymbol{0}}[D_{j,1}^g]} \in (0, \infty), \qquad k \in \mathbb{N}, \tag{26}$$

$$\frac{\mathbb{V}_{\boldsymbol{0}}[\hat{\theta}_{\mathrm{MC}}^{g,n(\varepsilon)}]}{\varepsilon^2/2} = \frac{\mathbb{V}_{\boldsymbol{0}}[\theta_1^{g,n(\varepsilon)}]}{\varepsilon^2 N/2} \to 1, \qquad \text{and} \qquad \frac{\mathbb{V}_{\boldsymbol{0}}[\hat{\theta}_{\mathrm{ML}}^{g,n(\varepsilon)}]}{\varepsilon^2/2} = \sum_{k=1}^{n} \frac{\mathbb{V}_{\boldsymbol{0}}[D_{k,1}^g]}{\varepsilon^2 N_k/2} \to 1. \tag{27}$$

Proof. Since $x + 1 \geq \lceil x \rceil \geq x$, we have $B_k(\varepsilon) \geq \varepsilon^2 N_k \geq B_{n,k}(\varepsilon)$, where

$$B_k(\varepsilon) := \varepsilon^2 + 2\sqrt{\mathbb{V}_{\mathbf{0}}[D_{k,1}^g]/k} \sum_{j=1}^{\infty} \sqrt{j\mathbb{V}_{\mathbf{0}}[D_{j,1}^g]}$$

$$\text{and} \quad B_{n,k}(\varepsilon) := 2\sqrt{\mathbb{V}_{\mathbf{0}}[D_{k,1}^g]/k} \sum_{j=1}^{n} \sqrt{j\mathbb{V}_{\mathbf{0}}[D_{j,1}^g]},$$

implying the limit in (26) (note that since $\mathbb{V}_{\mathbf{0}}[D_{k,1}^g] \leq 2\mathbb{E}_{\mathbf{0}}[(\Delta_k^g)^2 + (\Delta_{k-1}^g)^2] = \mathcal{O}(e^{-\vartheta_g k})$ for some $\vartheta_g > 0$, the limiting value in (26) is finite). The first limit in (27) follows similarly: $\varepsilon^2/2 + \mathbb{V}_{\mathbf{0}}[g(\chi_n)\Upsilon_{\lambda}] \geq \varepsilon^2 N/2 \geq \mathbb{V}_{\mathbf{0}}[g(\chi_n)\Upsilon_{\lambda}]$, where $\mathbb{V}_{\mathbf{0}}[\theta_1^{g,n}] = \mathbb{V}_{\mathbf{0}}[g(\chi_n)\Upsilon_{\lambda}] \to \mathbb{V}_{\mathbf{0}}[g(\chi)\Upsilon_{\lambda}] > 0$ as $\varepsilon \to 0$ by the convergence in L^2 of Proposition 3.1. By the same inequalities, we obtain

$$\sum_{k=1}^{n} \frac{\mathbb{V}_{\mathbf{0}}[D_{k,1}^g]}{B_k(\varepsilon)/2} \leq \sum_{k=1}^{n} \frac{\mathbb{V}_{\mathbf{0}}[D_{k,1}^g]}{\varepsilon^2 N_k/2} \leq \sum_{k=1}^{n} \frac{\mathbb{V}_{\mathbf{0}}[D_{k,1}^g]}{B_{n,k}(\varepsilon)/2} = 1.$$

The left-hand side converges to 1 by the monotone convergence theorem with respect to the counting measure, implying the second limit in (27) and completing the proof. □

Proof of Theorem 3.1. We first establish the CLT for the MLMC estimator $\hat{\theta}_{\mathrm{ML}}^{g,n(\varepsilon)}$, where $n = n(\varepsilon)$ is as stated in the theorem and the numbers of samples $N_1, \ldots, N_n$ are given in (9). By (6) and (8) we have

$$\sqrt{2}\varepsilon^{-1}\left(\hat{\theta}_{\mathrm{ML}}^{g,n(\varepsilon)} - \mathbb{E}_{\lambda}[g(\chi)]\right) = \sqrt{2}\varepsilon^{-1}\mathbb{E}_{\mathbf{0}}[\Delta_{n(\varepsilon)}^g] + \sum_{k=1}^{n(\varepsilon)}\sum_{i=1}^{N_k} \zeta_{k,i},$$

where

$$\zeta_{k,i} := \frac{\sqrt{2}}{\varepsilon N_k}\left(D_{k,i}^g - \mathbb{E}_{\mathbf{0}}[D_{k,i}^g]\right).$$

By assumption we have $c_0 > 1/\vartheta_g$. Thus, the limit $\sqrt{2}\varepsilon^{-1}\mathbb{E}_{\mathbf{0}}[\Delta_{n(\varepsilon)}^g] = \mathcal{O}(\varepsilon^{-1+c_0\vartheta_g}) \to 0$ as $\varepsilon \to 0$ follows from Proposition 3.1. Hence the CLT in (11) for the estimator $\hat{\theta}_{\mathrm{ML}}^{g,n(\varepsilon)}$ follows if we prove

$$\sum_{k=1}^{n(\varepsilon)}\sum_{i=1}^{N_k} \zeta_{k,i} \xrightarrow{d} Z \quad \text{as } \varepsilon \to 0,$$

where Z is a normal random variable with mean zero and unit variance. Thus, by [22, Theorem 5.12], it suffices to note that the $\zeta_{k,i}$ have zero mean $\mathbb{E}_{\mathbf{0}}[\zeta_{k,i}] = 0$,

$$\sum_{k=1}^{n(\varepsilon)}\sum_{i=1}^{N_k} \mathbb{E}_{\mathbf{0}}[\zeta_{k,i}^2] = \sum_{k=1}^{n(\varepsilon)} \frac{2}{\varepsilon^2 N_k} \mathbb{V}_{\mathbf{0}}[D_{k,1}^g] \to 1 \quad \text{as } \varepsilon \to 0,$$

which holds by (27), and to establish the Lindeberg condition, which states that for any $r > 0$ the following limit holds: $\sum_{k=1}^{n(\varepsilon)}\sum_{i=1}^{N_k} \mathbb{E}_{\mathbf{0}}[\zeta_{k,i}^2 \mathbb{1}_{\{|\zeta_{k,i}|>r\}}] \to 0$ as $\varepsilon \to 0$.

To prove the Lindeberg condition, first note that

$$\sum_{i=1}^{N_k} \mathbb{E}_{\mathbf{0}}[\zeta_{k,i}^2 \mathbb{1}_{\{|\zeta_{k,i}|>r\}}] = N_k \mathbb{E}_{\mathbf{0}}[\zeta_{k,1}^2 \mathbb{1}_{\{|\zeta_{k,i}|>r\}}] \leq N_k \mathbb{E}_{\mathbf{0}}[\zeta_{k,1}^2] \quad \text{for any } k \in \mathbb{N}. \quad (28)$$

By (26), the bound $N_k\mathbb{E}_{\mathbf{0}}[\zeta_{k,1}^2] = 2\mathbb{V}_{\mathbf{0}}[D_{k,1}^g]/(\varepsilon^2 N_k)$ converges for all $k \in \mathbb{N}$ as $\varepsilon \to 0$ to some $c_k \geq 0$ and $\sum_{k=1}^n N_k\mathbb{E}_{\mathbf{0}}[\zeta_{k,1}^2] \to 1 = \sum_{k=1}^\infty c_k$. Lemma 5.2 also implies that $\varepsilon N_k \to \infty$ and $\varepsilon^2 N_k$ converges to a positive finite constant as $\varepsilon \to 0$. Since $\mathbb{V}_{\mathbf{0}}[D_{k,1}^g] < \infty$ for all $k \in \mathbb{N}$, the dominated convergence theorem implies

$$N_k\mathbb{E}_{\mathbf{0}}[\zeta_{k,1}^2 \mathbb{1}_{\{|\zeta_{k,i}|>r\}}] = \frac{2\mathbb{E}_{\mathbf{0}}\left[(D_{k,1}^g - \mathbb{E}_{\mathbf{0}}[D_{k,1}^g])^2 \mathbb{1}_{\{|D_{k,1}^g - \mathbb{E}_{\mathbf{0}}[D_{k,1}^g]|>r\varepsilon N_k/2\}}\right]}{\varepsilon^2 N_k} \to 0, \qquad \text{as } \varepsilon \to 0.$$

Thus, the inequality in (28) and the dominated convergence theorem [22, Theorem 1.21] with respect to the counting measure yield the Lindeberg condition, establishing the CLT for $\hat{\theta}_{\mathrm{ML}}^{g,n(\varepsilon)}$.

Let us now establish the CLT for the MC estimator $\hat{\theta}_{\mathrm{MC}}^{g,n(\varepsilon)}$, with the number of samples N given in Proposition 3.2. As before, by Proposition 3.1 and the definition of $n(\varepsilon)$ in the theorem, the bias satisfies $\sqrt{2}\varepsilon^{-1}\mathbb{E}_{\mathbf{0}}[\Delta_{n(\varepsilon)}^g] = \mathcal{O}(\varepsilon^{-1+c_0\vartheta_g}) \to 0$ as $\varepsilon \to 0$. Thus, by [22, Theorem 5.12], it suffices to show that $2\mathbb{V}_{\mathbf{0}}[g(\chi_n)\Upsilon_\lambda]/(\varepsilon^2 N) \to 1$ as $\varepsilon \to 0$ and the Lindeberg condition holds: for any $r > 0$,

$$C(\varepsilon) := \sum_{i=1}^N \mathbb{E}_{\mathbf{0}}\left[|\zeta_{i,n(\varepsilon)}'|^2 \mathbb{1}_{\{|\zeta_{1,n(\varepsilon)}'|>r\}}\right] \to 0, \qquad \text{as } \varepsilon \to 0,$$

where

$$\zeta_{i,n}' := \frac{\sqrt{2}}{\varepsilon N}(\theta_i^{g,n} - \mathbb{E}_{\mathbf{0}}[\theta_i^{g,n}]).$$

The limit $2\mathbb{V}_{\mathbf{0}}[g(\chi_n)\Upsilon_\lambda]/(\varepsilon^2 N) \to 1$ as $\varepsilon \to 0$ follows from Lemma 5.2. To establish the Lindeberg condition, let $\widetilde{\theta}_\varepsilon := \theta_1^{g,n(\varepsilon)} - \mathbb{E}_{\mathbf{0}}[\theta_1^{g,n(\varepsilon)}]$ and note that

$$C(\varepsilon) = N\mathbb{E}\left[|\zeta_{1,n(\varepsilon)}'|^2 \mathbb{1}_{\{|\zeta_{1,n(\varepsilon)}'|>r\}}\right] = 2\mathbb{E}[|\widetilde{\theta}_\varepsilon|^2 \mathbb{1}_{\{|\widetilde{\theta}_\varepsilon|>r\varepsilon N/2\}}]/(\varepsilon^2 N).$$

Since $\theta_1^{g,n} \overset{L^2}{\to} g(\chi)\Upsilon_\lambda$ as $n \to \infty$ by Proposition 3.1, we get $\widetilde{\theta}_\varepsilon \overset{L^2}{\to} g(\chi)\Upsilon_\lambda - \mathbb{E}_{\mathbf{0}}[g(\chi)\Upsilon_\lambda]$ as $\varepsilon \to 0$. By Lemma 5.2, $\varepsilon^2 N \to 2\mathbb{V}_{\mathbf{0}}[g(\chi)\Upsilon_\lambda] > 0$, and the indicator function in the integrand vanishes in probability since $\varepsilon N \to \infty$. Thus, the dominated convergence theorem [22, Theorem 1.21] yields $C(\varepsilon) \to 0$, completing the proof. $\qquad \square$

Proof of Lemma 3.1. (a) Note that the density of ℓ_n is given by $x \mapsto \log(1/x)^{n-1}/(n-1)!$, $x \in (0,1)$. Note that $x^{-1} = e^{\log(1/x)} = \sum_{k=0}^\infty \log(1/x)^k/k!$; hence $\int_0^1 x^{-1}\phi(x)\mathrm{d}x = \sum_{k=1}^\infty \mathbb{E}[\phi(\ell_k)]$. This yields

$$n + \int_0^1 \frac{1}{x}(e^{cx}-1)\mathrm{d}x - \sum_{k=1}^n \mathbb{E}[e^{c\ell_k}] = \sum_{k=n+1}^\infty \mathbb{E}[e^{c\ell_k}-1] = \sum_{k=n+1}^\infty \sum_{j=1}^\infty \frac{c^j}{j!}\mathbb{E}[\ell_k^j]$$

$$= \sum_{j=1}^\infty \frac{c^j}{j!}\sum_{k=n+1}^\infty (1+j)^{-k} = \sum_{j=1}^\infty \frac{c^j}{j!j}(1+j)^{-n}.$$

Since $\int_0^1 x^{-1}(e^{cx}-1)\mathrm{d}x = \sum_{j=1}^\infty c^j/(j!j)$ and $(1+j)^{-n} \leq 2^{-n}$ for all $j \geq 1$, the inequality in (14) holds, implying (a).

(b) Note that $\int_0^1 x^{-1}(e^{cx}-1)\mathrm{d}x = \int_0^c x^{-1}(e^x-1)\mathrm{d}x$ and apply l'Hôpital's rule. $\qquad \square$

Appendix A. Simulation of stable laws

In this section we adapt the Chambers–Mallows–Stuck simulation of the increments of a Lévy process Z with generating triplet $(0, v, b)$, where

$$\frac{v(\mathrm{d}x)}{\mathrm{d}x} = \frac{c_+}{x^{\alpha+1}} \cdot \mathbb{1}_{(0,\infty)}(x) + \frac{c_-}{|x|^{\alpha+1}} \cdot \mathbb{1}_{(-\infty,0)}(x),$$

for arbitrary $(c_+, c_-) \in \mathbb{R}_+^2 \setminus \{\mathbf{0}\}$ and $\alpha \in (0, 2)$. First, we introduce the constant v, given as follows (see (14.20)–(14.21) in [35, Lemma 14.11]):

$$v = \int_1^\infty x^{-2} \sin(x)\mathrm{d}x + \int_0^1 x^{-2}(\sin(x) - x)\mathrm{d}x = 1 - \gamma.$$

Then the characteristic function of Z_t is given by the following (see [35, Theorem 14.15]):

$$\mathbb{E}\big[e^{iuZ_t}\big] = \exp\big(t\Psi(u)\big), \qquad u \in \mathbb{R},$$

$$\Psi(u) = i\mu u - \begin{cases} \varsigma|u|^\alpha\big(1 - i\theta \tan\big(\frac{\pi\alpha}{2}\big)\mathrm{sgn}(u)\big), & \alpha \in (0, 2) \setminus \{1\}, \\ \varsigma|u|\big(1 + i\theta\frac{2}{\pi}\mathrm{sgn}(u) \log|u|\big), & \alpha = 1, \end{cases}$$

where $\theta = (c_+ - c_-)/(c_+ + c_-)$ and the constants μ and ς are given by

$$(\mu, \varsigma) = \begin{cases} \big(b + \frac{c_- - c_+}{1-\alpha}, -(c_+ + c_-)\Gamma(-\alpha)\cos\big(\frac{\pi\alpha}{2}\big)\big), & \alpha \in (0, 2) \setminus \{1\}, \\ \big(b + v(c_+ - c_-), \frac{\pi}{2}(c_+ + c_-)\big), & \alpha = 1. \end{cases} \tag{29}$$

Finally, we define Zolotarev's function,

$$A_{a,r}(u) = \big(1 + \theta^2 \tan^2\big(\tfrac{\pi\alpha}{2}\big)\big)^{\frac{1}{2\alpha}} \frac{\sin(a(r+u)) \cos(ar + (a-1)u)^{1/a-1}}{\cos(u)^{1/a}}, \qquad u \in \big(-\tfrac{\pi}{2}, \tfrac{\pi}{2}\big).$$

Algorithm 2. (Chambers–Mallows–Stuck)

Simulation of Z_t with triplet $(0, v, b)$

Require: Parameters $(c_\pm, \alpha, b)$ and time horizon $t > 0$

1: Compute $\theta = (c_+ - c_-)/(c_+ + c_-)$ and (μ, ς) in (29)

2: Sample $U \sim \mathrm{U}\big(-\frac{\pi}{2}, \frac{\pi}{2}\big)$ and $E \sim \mathrm{Exp}(1)$

3: **if** $\alpha \neq 1$ **then**

4:     Compute $\delta = \arctan\big(\theta \tan\big(\frac{\pi\alpha}{2}\big)\big)/\alpha$ and **return** $(\varsigma t)^{1/\alpha} A_{\alpha,\delta}(U) E^{1-1/\alpha} + \mu t$

5: **else**

6:     **return** $\frac{2}{\pi}\varsigma t\big(\big(\frac{\pi}{2} + \theta U\big) \tan(U) - \theta \log\big(\frac{\pi E \cos(U)}{\varsigma t(\pi + 2\theta U)}\big)\big) + \mu t$

7: **end if**

Acknowledgement

Funding information

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] ANDERSEN, L. AND LIPTON, A. (2013). Asymptotics for exponential Lévy processes and their volatility smile: survey and new results. *Internat. J. Theoret. Appl. Finance* 16, article no. 1350001, 98 pp.

[2] AVRAM, F., CHAN, T. AND USABEL, M. (2002). On the valuation of constant barrier options under spectrally one-sided exponential Lévy models and Carr's approximation for American puts. *Stoch. Process. Appl.* **100**, 75–107.

[3] BAURDOUX, E., PALMOWSKI, Z. AND PISTORIUS, M. (2017). On future drawdowns of Lévy processes. *Stoch. Process. Appl.* **127**, 2679–2698.

[4] BEN ALAYA, M. AND KEBAIER, A. (2015). Central limit theorem for the multilevel Monte Carlo Euler method. *Ann. Appl. Prob.* **25**, 211–234.

[5] BISEWSKI, K. AND IVANOVS, J. (2020). Zooming-in on a Lévy process: failure to observe threshold exceedance over a dense grid. *Electron. J. Prob.* **25**, 33 pp.

[6] BLUMENTHAL, R. M. AND GETOOR, R. K. (1961). Sample functions of stochastic processes with stationary independent increments. *J. Math. Mech.* **10**, 493–516.

[7] CARR, P., GEMAN, H., MADAN, D. AND YOR, M. (2002). The fine structure of asset returns: an empirical investigation. *J. Business* **75**, 305–332.

[8] CARR, P., ZHANG, H. AND HADJILIADIS, O. (2011). Maximum drawdown insurance. *Internat. J. Theoret. Appl. Finance* **14**, 1195–1230.

[9] CHAMBERS, J. M., MALLOWS, C. L. AND STUCK, B. W. (1976). A method for simulating stable random variables. *J. Amer. Statist. Assoc.* **71**, 340–344.

[10] CONT, R. AND TANKOV, P. (2015). *Financial Modelling with Jump Processes*, 2nd edn. CRC Press, Boca Raton.

[11] DOWNEY, L. (2009). Ulcer Index (UI): what it is, how it works. Available at https://www.investopedia.com/terms/u/ulcerindex.asp.

[12] FIGUEROA-LÓPEZ, J. E. AND TANKOV, P. (2014). Small-time asymptotics of stopped Lévy bridges and simulation schemes with controlled bias. *Bernoulli* **20**, 1126–1164.

[13] GILES, M. B. (2015). Multilevel Monte Carlo methods. *Acta Numer.* **24**, 259–328.

[14] GILES, M. B. AND XIA, Y. (2017). Multilevel Monte Carlo for exponential Lévy models. *Finance Stoch.* **21**, 995–1026.

[15] GLASSERMAN, P. (2004). *Monte Carlo Methods in Financial Engineering*. Springer, New York.

[16] GONZÁLEZ CÁZARES, J. AND MIJATOVIĆ, A. (2022). Simulation of the drawdown and its duration in Lévy models via stick-breaking Gaussian approximation. *Finance Stoch.* **26**, 671–732.

[17] GONZÁLEZ CÁZARES, J. I. AND MIJATOVIĆ, A. (2021). TSB-Algorithm (video). Available at https://youtu.be/FJG6A3zk2lI. Published on Prob-AM YouTube channel.

[18] GONZÁLEZ CÁZARES, J. I., MIJATOVIĆ, A. AND URIBE BRAVO, G. (2022). Geometrically convergent simulation of the extrema of Lévy processes. *Math. Operat. Res.* **47**, 1141–1168.

[19] GRABCHAK, M. (2019). Rejection sampling for tempered Lévy processes. *Statist. Comput.* **29**, 549–558.

[20] HOEL, H. AND KRUMSCHEID, S. (2019). Central limit theorems for multilevel Monte Carlo methods. *J. Complexity* **54**, article no. 101407, 16 pp.

[21] IVANOVS, J. (2018). Zooming in on a Lévy process at its supremum. *Ann. Appl. Prob.* **28**, 912–940.

[22] KALLENBERG, O. (2002). *Foundations of Modern Probability*, 2nd edn. Springer, New York.

[23] KAWAI, R. AND MASUDA, H. (2011). On simulation of tempered stable random variates. *J. Comput. Appl. Math.* **235**, 2873–2887.

[24] KIM, K.-K. AND KIM, S. (2016). Simulation of tempered stable Lévy bridges and its applications. *Operat. Res.* **64**, 495–509.

[25] KLÜPPELBERG, C., KYPRIANOU, A. E. AND MALLER, R. A. (2004). Ruin probabilities and overshoots for general Lévy insurance risk processes. *Ann. Appl. Prob.* **14**, 1766–1801.

[26] KOU, S. (2014). Lévy processes in asset pricing. In *Encyclopedia of Quantitative Risk Analysis and Assessment*, John Wiley, New York. Available at https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118445112.stat03738.

[27] KUDRYAVTSEV, O. AND LEVENDORSKI, S. (2009). Fast and accurate pricing of barrier options under Lévy processes. *Finance Stoch.* **13**, 531–562.

[28] LANDRIAULT, D., LI, B. AND ZHANG, H. (2017). On magnitude, asymptotics and duration of drawdowns for Lévy models. *Bernoulli* **23**, 432–458.

[29] LI, P., ZHAO, W. AND ZHOU, W. (2015). Ruin probabilities and optimal investment when the stock price follows an exponential Lévy process. *Appl. Math. Comput.* **259**, 1030–1045.

[30] MCLEISH, D. (2011). A general method for debiasing a Monte Carlo estimator. *Monte Carlo Meth. Appl.* **17**, 301–315.

[31] MORDECKI, E. (2003). Ruin probabilities for Lévy processes with mixed-exponential negative jumps. *Theory Prob. Appl.* **48**, 188–194.

[32] PITMAN, J. AND URIBE BRAVO, G. (2012). The convex minorant of a Lévy process. *Ann. Prob.* **40**, 1636–1674.

[33] POIROT, J. AND TANKOV, P. (2006). Monte Carlo option pricing for tempered stable (CGMY) processes. *Asia-Pacific Financial Markets* **13**, 327–344.

[34] RHEE, C.-H. AND GLYNN, P. W. (2015). Unbiased estimation with square root convergence for SDE models. *Operat. Res.* **63**, 1026–1043.

[35] SATO, K.-I. (2013). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press.

[36] SCHOUTENS, W. (2003). *Levy Processes in Finance: Pricing Financial Derivatives*. John Wiley, New York.

[37] SCHOUTENS, W. (2006). Exotic options under Lévy models: an overview. *J. Comput. Appl. Math.* **189**, 526–538.

[38] SORNETTE, D. (2003). *Why Stock Markets Crash: Critical Events in Complex Financial Systems*. Princeton University Press.

[39] VECER, J. (2006). Maximum drawdown and directional trading. *Risk* **19**, 99–92.

[40] VIHOLA, M. (2018). Unbiased estimators and multilevel Monte Carlo. *Operat. Res.* **66**, 448–462.