Simulation of the drawdown and its duration in Lévy models via stick-breaking Gaussian approximation

Jorge González Cázares[1] · Aleksandar Mijatović[1]

Abstract

We develop a computational method for expected functionals of the drawdown and its duration in exponential Lévy models. It is based on a novel simulation algorithm for the joint law of the state, supremum and time the supremum is attained of the Gaussian approximation for a general Lévy process. We bound the bias for various locally Lipschitz and discontinuous payoffs arising in applications and analyse the computational complexities of the corresponding Monte Carlo and multilevel Monte Carlo estimators. Monte Carlo methods for Lévy processes (using Gaussian approximation) have been analysed for Lipschitz payoffs, in which case the computational complexity of our algorithm is up to two orders of magnitude smaller when the jump activity is high. At the core of our approach are bounds on certain Wasserstein distances, obtained via the novel stick-breaking Gaussian (SBG) coupling between a Lévy process and its Gaussian approximation. Numerical performance, based on the implementation in Cázares and Mijatović (SBG approximation. GitHub repository. Available online at https://github.com/jorgeignaciogc/SBG.jl (2020)), exhibits a good agreement with our theoretical bounds. Numerical evidence suggests that our algorithm remains stable and accurate when estimating Greeks for barrier options and outperforms the "obvious" algorithm for finite-jump-activity Lévy processes.

✉ J. González Cázares
jorge.i.gonzalez-cazares@warwick.ac.uk

A. Mijatović
a.mijatovic@warwick.ac.uk

[1] Department of Statistics, University of Warwick and The Alan Turing Institute, London, UK

1 Introduction

1.1 Setting and motivation

Lévy processes are increasingly popular for the modelling of market prices of risky assets. They naturally address the shortcoming of diffusion models by allowing large (often heavy-tailed) sudden movements of the asset price as observed in the markets; see Schoutens [63, Chap. 4], Kou [43, Sect. 1], Cont and Tankov [20, Chap. 1]. While Lévy models with finite jump activity (i.e., jump-diffusions) can model large instantaneous movements (jumps) and small-time fluctuations (Gaussian component), infinite activity Lévy models often provide better parsimonious statistical and/or risk-neutral descriptions of asset returns (e.g. Carr et al. [13]) and exhibit a more flexible implied volatility behaviour over short time horizons (e.g. Mijatović and Tankov [54]). For risk management, it is thus crucial to quantify the probabilities of rare and/or extreme events under *all* Lévy models. Of particular interest in this context are the distributions of the drawdown (the current decline from a historical peak) and its duration (the elapsed time since the historical peak); see e.g. Sornette [65, Chap. 1], Večeř [67], Carr et al. [15], Baurdoux et al. [5], Landriault et al. [48]. Together with the hedges for barrier options (Avram et al. [3], Schoutens [64], Kudryavtsev and Levendorskiĭ [45], Giles and Xia [30]) and ruin probabilities in insurance (Mordecki [55], Klüppelberg et al. [41], Li et al. [49]), the expected drawdown and its duration constitute risk measures dependent on the following random vector, which is a statistic of the path of a Lévy process X: a historic maximum $\overline{X}_T$ at a time T, the time $\overline{\tau}_T(X) \leq T$ at which this maximum was attained, and the value X_T of the process at T. Since neither the distribution of the drawdown $1 - \exp(X_T - \overline{X}_T)$ nor of its duration $T - \overline{\tau}_T(X)$ is analytically tractable for a general X, simulation provides a natural alternative. The main objective of the present paper is to develop and analyse a novel practical simulation algorithm for the joint law of $(X_T, \overline{X}_T, \overline{\tau}_T(X))$ which is applicable to a general Lévy process X.

Exact simulation of the drawdown of a Lévy process is currently out of reach except for the stable (González Cázares et al. [33]) and jump-diffusion cases. However, even in the stable case, it is not known how to jointly simulate any two components of $(X_T, \overline{X}_T, \overline{\tau}_T(X))$. Among the approximate simulation algorithms, the recently developed stick-breaking approximation in González Cázares et al. [35] is the fastest in terms of its computational complexity, as it samples from the law of $(X_T, \overline{X}_T, \overline{\tau}_T(X))$ with a geometrically decaying bias. However, like most approximate simulation algorithms for a statistic of the entire trajectory, it is only valid for a Lévy process whose increments can be sampled. Such a requirement does not hold for large classes of widely used Lévy processes, including the general CGMY (aka KoBoL) model in Carr et al. [14]. Moreover, nonparametric estimation of Lévy processes typically yields Lévy measures whose transitions cannot be sampled (Neumann and Reiß [56], Chen et al. [17], Comte and Genon-Catalot [19], Cai et al. [11], Qin and Todorov [59]), which again makes a direct application of the algorithm in [35] infeasible.

If the increments of X cannot be sampled, a general approach is to use the Gaussian approximation introduced by Asmussen and Rosiński [2], which substitutes the

small-jump component of the Lévy process by a Brownian motion. Thus the Gaussian approximation process is a jump-diffusion, and the exact sample of the random vector (consisting of the state of the process, the supremum and the time the supremum is attained) can be obtained by applying Devroye's sampler, see Devroye [23, Alg. MAXLOCATION], between the consecutive jumps. However, little is known about how close these quantities are to the vector $(X_T, \overline{X}_T, \overline{\tau}_T(X))$ that is being approximated, in either Wasserstein or Kolmogorov distances. Indeed, bounds on the distances between the marginals of the Gaussian approximation and X_T have been considered by Dia [24] and recently improved by Mariucci and Reiß [50] and Carpentier et al. [12]. A Wasserstein bound on the supremum is given by Dia [24], but so far, no improvement analogous to the marginal case has been established. Moreover, to the best of our knowledge, there are no corresponding results for either the joint law of $(X_T, \overline{X}_T)$ or the time $\overline{\tau}_T(X)$. Furthermore, as explained in Remark 4.1 below, the exact simulation algorithm for the supremum and the time of the supremum of a Gaussian approximation based on Devroye [23, Alg. MAXLOCATION] is unsuitable for the multilevel Monte Carlo estimation.

The main motivation for the present work is to provide an operational framework for Lévy processes which allows us to settle the issues raised in the previous paragraph, develop a general simulation algorithm for $(X_T, \overline{X}_T, \overline{\tau}_T(X))$ and analyse the computational complexity of its Monte Carlo (MC) and multilevel Monte Carlo (MLMC) estimators.

1.2 Contributions in the present paper

Our main contributions are as follows.

(Ia) We establish bounds on the Wasserstein and Kolmogorov distances between the vector $\overline{\chi}_T = (X_T, \overline{X}_T, \overline{\tau}_T(X))$ and its Gaussian approximation, denoted by $\overline{\chi}_T^{(\kappa)} = (X_T^{(\kappa)}, \overline{X}_T^{(\kappa)}, \overline{\tau}_T(X^{(\kappa)}))$, where $X^{(\kappa)}$ is a jump-diffusion equal to the Lévy process X with the jumps smaller than $\kappa \in (0, 1]$ substituted by a Brownian motion (see (2.5) below) and $\overline{X}_T^{(\kappa)}$ (resp. $\overline{\tau}_T(X^{(\kappa)})$) is the supremum of $X^{(\kappa)}$ (resp. the time $X^{(\kappa)}$ attains the supremum) over the time interval $[0, T]$.

(Ib) These results enable us to control the bias $|\mathbb{E}[f(\overline{\chi}_T)] - \mathbb{E}[f(\overline{\chi}_T^{(\kappa)})]|$ for a discontinuous and/or locally Lipschitz payoff f, a fundamental advance in the area of Gaussian approximation of Lévy processes.

(II) We introduce a simple and fast algorithm, SBG-Alg, which samples exactly the vector of interest for the Gaussian approximation of any Lévy process X, develop an MLMC estimator based on SBG-Alg (see González Cázares and Mijatović [31] for an implementation in Julia), and analyse its complexity for discontinuous and locally Lipschitz payoffs arising in applications.

Before discussing contributions (Ia)+ (Ib) and (II) in more detail, note that the role in the present paper of the main algorithm from González Cázares et al. [35] is analogous to the role the simulation of Brownian increments plays in the sampling of Euler scheme chains approximating the law of the solutions of stochastic differential equations. Differently put, the present paper analyses the convergence of a *family of algorithms*, indexed by the cutoff parameter $\kappa > 0$ and essentially given by the algorithm in [35] for each approximation $\overline{\chi}_T^{(\kappa)}$, analogous to the analysis of

the convergence of Euler schemes indexed by the step size $h > 0$ that controls the volatility in the simulation algorithm for Brownian increments. The underlying ideas in the present paper and in [35] are based on the theory of convex minorants of general Lévy processes (see e.g. González Cázares and Mijatović [32]). While [35] has a theoretical obstruction as it is applicable only if the increments of X can be sampled, the results and algorithms in the present paper are applicable to essentially all Lévy processes. This requires fundamental advances in both bias and level variance control, which we now discuss briefly.

(Ia) In Theorem 3.4 (see also Corollary 3.5), we establish novel bounds on the Wasserstein distance between $\overline{\chi}_T$ and $\overline{\chi}_T^{(\kappa)}$ (as κ tends to 0) under weak assumptions, typically satisfied by the models used in applications. The proof of Theorem 3.4 has two main ingredients. First, in Sect. 7.2 below, we construct a novel *stick-breaking Gaussian (SBG) coupling* between $\overline{\chi}_T$ and $\overline{\chi}_T^{(\kappa)}$, based on the stick-breaking (SB) representation of $\overline{\chi}_T$ in (2.1) and the minimal transport coupling between the increments of X and its approximation $X^{(\kappa)}$. The second ingredient consists of new bounds on the Wasserstein and Kolmogorov distances between the laws of X_t and $X_t^{(\kappa)}$ for any $t > 0$, given in Theorems 3.1 and 3.3, respectively. The improvement of our bounds on the existing literature on the Gaussian approximation for the marginals of a Lévy process is reflected both in the bounds of Theorem 3.4 (for which there is no comparison in the literature) as well as in the performance of SBG-Alg. Moreover, even though the bounds in Theorem 3.4 are of Wasserstein type, the estimates in Theorem 3.3 of the Kolmogorov distance of the marginals of a Gaussian approximation are crucial in the proof of Theorem 3.4 because of the presence of the indicator functions in SB representations (2.1).

(Ib) In Sect. 3.2, we give novel bounds on the bias of locally Lipschitz and barrier payoffs of $\overline{\chi}_T$; see Propositions 3.7, 3.9 and 3.12. Their proofs are based on Theorem 3.4 and Lemma 7.5, which essentially converts the Wasserstein distance into the Kolmogorov distance for sufficiently regular distributions. In particular, note that Theorem 3.4 is used to control the distance between a non-Lipschitz functional $\overline{\tau}_T(X)$ of the path of X and $\overline{\tau}_T(X^{(\kappa)})$. Thus Proposition 3.12 bounds the bias of a discontinuous payoff of a non-Lipschitz functional of X. Applications related to the duration of the drawdown and the risk-management of barrier options require bounding the bias of certain discontinuous functions of $\overline{\chi}_T$. We thus develop explicit general sufficient conditions on the characteristic triplet of the Lévy process X (see Proposition 3.15 below) which guarantee the applicability of the results of Sect. 3.2 to models typically used in practice. Finally, Propositions 3.9 and 3.12 yield new bounds on the Kolmogorov distance between the components of $(\overline{X}_T, \overline{\tau}_T(X))$ and $(\overline{X}_T^{(\kappa)}, \overline{\tau}_T(X^{(\kappa)}))$ (see Corollary 3.14 below) which we hope are of independent interest, complementing the Wasserstein bounds of Corollary 3.5.

(II) Our main simulation algorithm, SBG-Alg, samples jointly coupled Gaussian approximations of $\overline{\chi}_T$ at distinct approximation levels (i.e., two different values of the cutoff κ). The coupling in SBG-Alg exploits the following simple observations:

– Any Gaussian approximation $\overline{\chi}_T^{(\kappa)}$ has an SB representation in (2.2), where the law of Y in (2.2) must equal that of $X^{(\kappa)}$.

– For any two Gaussian approximations, the stick-breaking process in (2.2) can be shared.
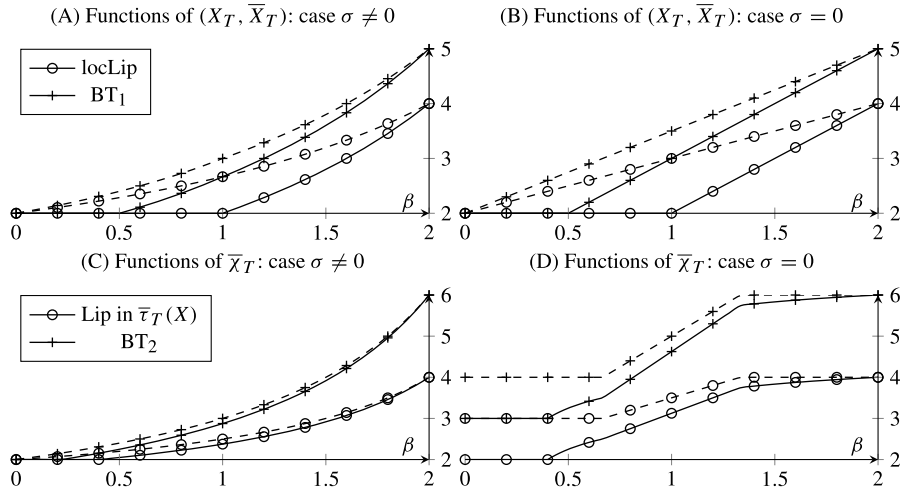
Fig. 1 Dashed (resp. solid) line plots the power of ϵ^{-1} in the computational complexity of an MC (resp. MLMC) estimator, as a function of the BG index β defined in (2.6), for discontinuous functions in BT$_1$ (3.12) and BT$_2$ (3.14), locally Lipschitz payoffs as well as Lipschitz functions of $\overline{\tau}_T(X)$. The cases are split according to whether X is with ($\sigma \neq 0$) or without ($\sigma = 0$) a Gaussian component. The pictures are based on Tables 1 and 2 under assumptions typically satisfied in applications; see Sect. 4.2 below for details

– The increments in (2.2) over the shared sticks can be coupled using the definition (2.5) of the Gaussian approximation $X^{(\kappa)}$.

We analyse the computational complexity of the MLMC estimator based on SBG-Alg for a variety of payoff functions arising in applications. Figure 1 shows the leading power of the resulting MC and MLMC complexities, summarised in Tables 1 and 2 below (see Theorem 7.17 for full details), for locally Lipschitz and discontinuous payoffs used in practice. To the best of our knowledge, neither locally Lipschitz nor discontinuous payoffs have been previously considered in the context of MLMC estimation under Gaussian approximation.

A key component of the analysis of the complexity of an MLMC estimator is the rate of decay of level variances (see Appendix A.2 for the definition). In the case of SBG-Alg, the rate of decay is given in Theorem 7.10 below for locally Lipschitz and discontinuous payoffs of interest. The analysis in the proof of Theorem 7.10 of the coupling simulated in SBG-Alg (between two Gaussian approximations at distinct cutoff levels) relies on the SBG coupling (between a Gaussian approximation and its limit) to control the rate of decay of level variances. Moreover, the proof of Theorem 7.10 shows that the decay of the level variances for Lipschitz payoffs under SBG-Alg is asymptotically equal to that of Algorithm 1, which samples jointly the increments at two distinct levels only. The principal reason for this equality between the "marginal" Algorithm 1 and the "path-dependent" SBG-Alg lies in the fact that it is the increments that dominate in the SB representation in (2.2). Furthermore, an improved coupling in Algorithm 1 for the increments of the Gaussian approximations (cf. the final of the three observations listed in the previous paragraph) would further reduce the computational complexity of the MLMC estimator for all payoffs consid-

ered in this paper (including the discontinuous ones). However, such a coupling is currently out of reach; cf. the discussion following Algorithm 1 in Sect. 4.1.1 below. To the best of our knowledge, SBG-Alg is the first exact simulation algorithm for coupled Gaussian approximations of $\overline{\chi}_T$ with *vanishing level variances* for a general Lévy process X; see Remark 4.1 below for further details.

In Sect. 6, using the code in the repository González Cázares and Mijatović [31], we test our theoretical findings against numerical results. In Sect. 6.1, we run SBG-Alg for models in the tempered stable and Watanabe classes. The former is a widely used class of processes whose increments cannot be sampled for all parameter values, and the latter is a well-known class of processes with infinite activity but singular continuous increments. In both cases, we find a good agreement between the theoretical prediction and the estimated decays of the bias and level variance; see Figs. 4 and 5 below.

In the context of MC estimation, a direct simulation algorithm based on Devroye's sampler [23, Alg. MAXLOCATION] (Algorithm 2 below) can be used instead of SBG-Alg. In Sect. 6.2, we compare numerically its cost with that of SBG-Alg. In the examples we considered, the speedup of SBG-Alg over Algorithm 2 is about 50, see Fig. 6, remaining significant even for processes with finite jump activity; see Fig. 7. In fact, these examples demonstrate that SBG-Alg (with $\kappa = 0$) is preferable for jump-diffusions as it significantly outperforms Algorithm 2.

In Sect. 6.3, we provide numerical evidence demonstrating that SBG-Alg (combined with central finite differences) remains very stable and fast for computing Delta and Gamma of barrier options under exponential Lévy models. Interestingly, the error of the Delta remains bounded all the way to the barrier (see Fig. 10 below), a property crucial in practice and very rare for Monte Carlo algorithms.

1.3 Organisation

The remainder of the paper is organised as follows. Section 2 recalls the SB representation (2.1), (2.2) and the Gaussian approximation (2.5) developed in González Cázares et al. [35] and Asmussen and Rosiński [2], respectively. Section 3 presents bounds on the Wasserstein and Kolmogorov distances between $\overline{\chi}_T$ and its Gaussian approximation $\overline{\chi}_T^{(\kappa)}$ and the biases of certain payoffs arising in applications. Section 3 also provides simple sufficient conditions, in terms of the Lévy triplet, under which these bounds hold. Section 4 constructs our main algorithm, SBG-Alg, and presents the computational complexity of the corresponding MC and MLMC estimators for all payoffs considered in this paper. Having stated our main results, we present a thorough comparison with the literature in Sect. 5. In Sect. 6, we illustrate numerically our results for a widely used class of Lévy models. The proofs and technical results are found in Sect. 7. Appendix A.1 gives a brief account of the complexity analysis of MC and MLMC (introduced in Heinrich [37] and Giles [28]) estimators.

2 The stick-breaking representation and the Gaussian approximation

Let $f : [0, \infty) \to \mathbb{R}$ be a right-continuous function with left limits. For $t \in (0, \infty)$, define $\underline{f}_t := \inf_{s \in [0,t]} f(s)$, $\overline{f}_t := \sup_{s \in [0,t]} f(s)$ and let $\underline{\tau}_t(f)$ (resp. $\overline{\tau}_t(f)$) be the

last time before t that the infimum $\underline{f}_t$ (resp. supremum $\overline{f}_t$) is attained. Throughout, $X = (X_t)_{t \geq 0}$ denotes a Lévy process, i.e., a stochastic process started at the origin with independent, stationary increments and right-continuous paths with left limits; see Bertoin [7, Chap. 1], Kyprianou [47, Chaps. 1 and 2] and Sato [62, Chaps. 1 and 2] for background on Lévy processes. In mathematical finance, the risky asset price $S = (S_t)_{t \geq 0}$ under an exponential Lévy model is given by $S_t := S_0 e^{X_t}$. The price S_t, its *drawdown* $1 - S_t/\overline{S}_t$ (resp. *drawup* $S_t/\underline{S}_t - 1$) and *duration* $t - \underline{\tau}_t(S)$ (resp. $t - \overline{\tau}_t(S)$) at time t can be recovered from the vector $\overline{\chi}_t := (X_t, \overline{X}_t, \overline{\tau}_t(X))$ (resp. $\underline{\chi}_t := (X_t, \underline{X}_t, \underline{\tau}_t(X))$). Because $Z := -X$ is a Lévy process and we have $\overline{\chi}_t = (-Z_t, -\underline{Z}_t, \underline{\tau}_t(Z))$, it is sufficient to analyse the vector $\underline{\chi}_t$.

2.1 The stick-breaking (SB) representation

Given a Lévy process X and a time horizon $t > 0$, there exist a coupling (X, Y), where $Y \overset{d}{=} X$ (throughout the paper, $\overset{d}{=}$ denotes equality in law), and a stick-breaking process $\ell = (\ell_n)_{n \in \mathbb{N}}$ on $[0, t]$ based on the uniform law $U(0, 1)$ (i.e., $L_0 := t$, $L_n := L_{n-1} U_n$, $\ell_n := L_{n-1} - L_n$ for $n \in \mathbb{N}$, where $(U_n)_{n \in \mathbb{N}}$ is an i.i.d. sequence following $U_n \sim U(0, 1)$), which is independent of Y, such that a.s.,

$$\underline{\chi}_t = \sum_{k=1}^{\infty} \left(Y_{L_{k-1}} - Y_{L_k}, \min\{Y_{L_{k-1}} - Y_{L_k}, 0\}, \ell_k \mathbb{1}_{\{Y_{L_{k-1}} - Y_{L_k} \leq 0\}} \right). \tag{2.1}$$

The coupling (X, Y) and ℓ can be constructed easily using the equality in law in González Cázares and Mijatović [32, Theorem 11]. For a construction of this coupling, see González Cázares et al. [35, Sect. 4.1]. Since given L_n, $(\ell_k)_{k>n}$ is a stick-breaking process on $[0, L_n]$, for any $n \in \mathbb{N}$, (2.1) implies that

$$\begin{aligned}
\underline{\chi}_t \overset{d}{=} & \left(Y_{L_n}, \underline{Y}_{L_n}, \underline{\tau}_{L_n}(Y) \right) \\
& + \sum_{k=1}^{n} \left(Y_{L_{k-1}} - Y_{L_k}, \min\{Y_{L_{k-1}} - Y_{L_k}, 0\}, \ell_k \mathbb{1}_{\{Y_{L_{k-1}} - Y_{L_k} \leq 0\}} \right).
\end{aligned} \tag{2.2}$$

Observe that the vector $(Y_{L_n}, \underline{Y}_{L_n}, \underline{\tau}_{L_n}(Y))$ and the sum on the right-hand side of the identity in (2.2) are conditionally independent given L_n: the former (resp. latter) is a function of $(Y_s)_{s \in [0, L_n]}$ (resp. $(Y_s - Y_{L_n})_{s \in [L_n, t]}$), cf. Fig. 2. The vector $\underline{\chi}_t$ of interest is thus represented by the corresponding vector $(Y_{L_n}, \underline{Y}_{L_n}, \underline{\tau}_{L_n}(Y))$ over an exponentially small interval (since $\mathbb{E}[L_n] = 2^{-n} t$) and n independent increments of the Lévy process over random intervals independent of Y. In (2.2) and throughout, $\mathbb{1}_A$ is the indicator function of the set A.

We stress that (2.1) and (2.2) reduce the analysis of the path-functional $\underline{\chi}_t$ to that of the increments of X, since the "error term" $(Y_{L_n}, \underline{Y}_{L_n}, \underline{\tau}_{L_n}(Y))$ in (2.2) is typically exponentially small in n. For an arbitrary Lévy process X', the vector $(X'_t, \underline{X}'_t, \underline{\tau}_t(X'))$ has a representation as in (2.1) for a Lévy process $Y' \overset{d}{=} X'$ independent of the stick-breaking process ℓ. Thus more generally, the laws of the vectors $\underline{\chi}_t$ and $(X'_t, \underline{X}'_t, \underline{\tau}_t(X'))$ will be close if the laws of the increments of Y and Y' over
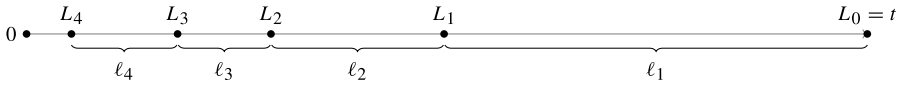
Fig. 2 The figure illustrates the first $n = 4$ sticks of a stick-breaking process. The increments of Y in (2.2) are taken over the intervals $[L_k, L_{k-1}]$ of length ℓ_k. Crucially, the time L_n featuring in the vector $(Y_{L_n}, \underline{Y}_{L_n}, \underline{\tau}_{L_n}(Y))$ in (2.2) is exponentially small in n

the intervals $[L_k, L_{k-1}]$ are close. Indeed, by the identity in law (2.1), in order to quantify the distance between the vectors $\underline{\chi}_t$ and $(X'_t, \underline{X}'_t, \underline{\tau}_t(X'))$, it suffices to couple the increments of Y and Y' over the intervals $[L_k, L_{k-1}]$, $k \in \mathbb{N}$, with a common stick-breaking process ℓ independent of (Y, Y') and compare the corresponding sums appearing on the right-hand side of (2.1). This observation constitutes a key step in the construction of the coupling used in the proof of Theorem 3.4 below, which in turn plays a crucial role in controlling the bias (see the subsequent results of Sect. 3) of our main simulation algorithm SBG-Alg described in Sect. 4 below. SBG-Alg is based on (2.2) with X' being the Gaussian approximation of a general Lévy process X introduced in Asmussen and Rosiński [2] and recalled briefly in the next subsection.

2.2 The Gaussian approximation

The law of a Lévy process $X = (X_t)_{t \geq 0}$ is uniquely determined by the law of its marginal X_t (for any $t > 0$), which is in turn given by the Lévy–Khintchine formula [62, Theorem 8.1]: for $u \in \mathbb{R}$,

$$\frac{1}{t} \log \mathbb{E}[e^{iuX_t}] = iub - \frac{1}{2}u^2\sigma^2 + \int_{\mathbb{R}\setminus\{0\}} \left(e^{iux} - 1 - iux\mathbb{1}_{(-1,1)}(x)\right)\nu(dx). \quad (2.3)$$

The *Lévy measure* ν is required to satisfy $\int_{\mathbb{R}\setminus\{0\}} \min\{x^2, 1\}\nu(dx) < \infty$, while $\sigma \geq 0$ specifies the volatility of the Brownian component of X. Note that the drift $b \in \mathbb{R}$ depends on the cutoff function $x \mapsto \mathbb{1}_{(-1,1)}(x)$. Thus the Lévy triplet (σ^2, ν, b), with respect to the cutoff function $x \mapsto \mathbb{1}_{(-1,1)}(x)$, determines the law of X. All Lévy triplets in the present paper use this cutoff function.

The *Lévy–Itô decomposition* at level $\kappa \in (0, 1]$ (see [62, Theorems 19.2 and 19.3]) is given by

$$X_t = b_\kappa t + \sigma B_t + J_t^{1,\kappa} + J_t^{2,\kappa}, \qquad t \geq 0, \quad (2.4)$$

where $b_\kappa := b - \int_{(-1,1)\setminus(-\kappa,\kappa)} x\nu(dx)$, $B = (B_t)_{t \geq 0}$ is a standard Brownian motion and the processes $J^{1,\kappa} = (J_t^{1,\kappa})_{t \geq 0}$ and $J^{2,\kappa} = (J_t^{2,\kappa})_{t \geq 0}$ are Lévy with triplets $(0, \nu|_{(-\kappa,\kappa)}, 0)$ and $(0, \nu|_{\mathbb{R}\setminus(-\kappa,\kappa)}, b - b_\kappa)$, respectively. The processes B, $J^{1,\kappa}$, $J^{2,\kappa}$ in (2.4) are independent, $J^{1,\kappa}$ is an L^2-bounded martingale with jumps of magnitude less than κ and $J^{2,\kappa}$ is a driftless (i.e., piecewise constant) compound Poisson process with intensity $\overline{\nu}(\kappa) := \nu(\mathbb{R} \setminus (-\kappa, \kappa))$ and jump distribution $\nu|_{\mathbb{R}\setminus(-\kappa,\kappa)}/\overline{\nu}(\kappa)$.

In applications, the main problem lies in the user's inability to simulate the increments of $J^{1,\kappa}$ in (2.4), i.e., the small jumps of the Lévy process X. Instead of ignoring

this component for a small value of κ, the Gaussian approximation in Asmussen and Rosiński [2],

$$X_t^{(\kappa)} := b_\kappa t + \sqrt{\overline{\sigma}_\kappa^2 + \sigma^2}\, W_t + J_t^{2,\kappa}, \qquad \text{where } \overline{\sigma}_\kappa^2 := \int_{(-\kappa,\kappa)\setminus\{0\}} x^2 \nu(dx), \quad (2.5)$$

substitutes the martingale $\sigma B + J^{1,\kappa}$ in (2.4) with a Brownian motion with variance $\overline{\sigma}_\kappa^2 + \sigma^2$. In (2.5), the standard Brownian motion $W = (W_t)_{t\geq 0}$ is independent of $J^{2,\kappa}$. Let $\overline{\sigma}_\kappa$ denote the nonnegative square root of $\overline{\sigma}_\kappa^2$. The *Gaussian approximation* of X at level κ, given by the Lévy process $X^{(\kappa)} = (X_t^{(\kappa)})_{t\geq 0}$, is natural in the following sense: the weak convergence $\overline{\sigma}_\kappa^{-1} J^{1,\kappa} \xrightarrow{d} W$ (in the Skorokhod space $\mathcal{D}[0,\infty)$) as $\kappa \to 0$ holds if and only if $\overline{\sigma}_{\min\{K\overline{\sigma}_\kappa,\kappa\}}/\overline{\sigma}_\kappa \to 1$ for every $K > 0$ (see [2]). This condition holds if $\overline{\sigma}_\kappa/\kappa \to \infty$, and the two conditions are equivalent if ν has no atoms in a neighbourhood of zero; see [2, Proposition 2.2].

Since $J^{2,\kappa}$ has an average of $\overline{\nu}(\kappa)t$ jumps on $[0,t]$, the expected cost of simulating the increment $X_t^{(\kappa)}$ is a constant multiple of $1 + \overline{\nu}(\kappa)t$ (see Algorithm 1 below). Moreover, the user need only be able to sample from the normalised tails of ν, which can typically be achieved in multiple ways (see e.g. Rosiński [61]). The behaviour of $\overline{\nu}(\kappa)$ and $\overline{\sigma}_\kappa$ as $\kappa \downarrow 0$, key in the analysis of the MC/MLMC complexity, can be described in terms of the *Blumenthal–Getoor (BG) index β* from Blumenthal and Getoor [9], defined by

$$\beta := \inf\{p > 0 : I_0^p < \infty\}, \qquad \text{where } I_0^p := \int_{(-1,1)\setminus\{0\}} |x|^p \nu(dx) \text{ for } p \geq 0. \quad (2.6)$$

Note that $\beta \in [0,2]$, since $I_0^2 < \infty$ for any Lévy measure ν. Furthermore, $I_0^1 < \infty$ if and only if the paths of $J^{1,\kappa}$ have finite variation. Moreover, $I_0^p < \infty$ for any $p > \beta$, but I_0^β can be either finite or infinite. If $q \in [0,2]$ satisfies $I_0^q < \infty$, we have for all $\kappa \in (0,1]$ the inequalities (see e.g. [35, Lemma 9])

$$\overline{\sigma}_\kappa^2 \leq I_0^q \kappa^{2-q} \qquad \text{and} \qquad \overline{\nu}(\kappa) \leq \overline{\nu}(1) + I_0^q \kappa^{-q}. \quad (2.7)$$

Finally, we stress that the dependence between W in (2.5) and $\sigma B + J^{1,\kappa}$ in (2.4) has not been specified. This coupling can vary greatly, depending on the circumstance (e.g. the analysis of the Wasserstein distance between functionals of X and $X^{(\kappa)}$ (Sect. 3) or the minimisation of level variances in MLMC (Sect. 4)). Thus unless otherwise stated, no explicit form for the dependence between $\sigma B + J^{1,\kappa}$ and W is assumed.

3 Distance between the laws of $\underline{\chi}_t$ and its Gaussian approximation $\underline{\chi}_t^{(\kappa)}$

In this section, we present bounds on the distance between the laws of $\underline{\chi}_t$, defined in Sect. 2 above, and its Gaussian approximation $\underline{\chi}_t^{(\kappa)} := (X_t^{(\kappa)}, \underline{X}_t^{(\kappa)}, \underline{\tau}_t(X^{(\kappa)}))$, based

on the Lévy process $X^{(\kappa)}$ in (2.5). The results in this section are crucial for the analysis of the computational complexity of the MC and MLMC estimators based on SBG-Alg discussed in Sect. 4 below.

Our bounds on the Wasserstein distance (see Theorem 3.4 and Corollary 3.5) are based on the SBG coupling constructed in Sect. 7.2 below, which in turn draws on the coupling in (2.1). Theorems 3.1 and 3.3 provide fundamental improvements (discussed in Sect. 3.1.1 below) for the bounds on the Wasserstein and Kolmogorov distances of the marginals X_t and $X_t^{(\kappa)}$, which play a key role in the proof of Theorem 3.4. In Sect. 3.2 below, Theorem 3.4 is applied to control the bias of some discontinuous and non-Lipschitz payoffs of $\underline{\chi}_t$ arising in applications as well as the Kolmogorov distances between the components of $(\underline{X}_t, \underline{\tau}_t(X))$ and $(\underline{X}_t^{(\kappa)}, \underline{\tau}_t(X^{(\kappa)}))$.

3.1 The Wasserstein distance between the vectors $\underline{\chi}_t$ and $\underline{\chi}_t^{(\kappa)}$

In order to study the Wasserstein distance between $\underline{\chi}_t$ and $\underline{\chi}_t^{(\kappa)}$ via (2.1), (2.2), we have to quantify the Wasserstein and Kolmogorov distances between the increments X_s and $X_s^{(\kappa)}$ for any time $s > 0$. With this in mind, we start with Theorems 3.1 and 3.3, which play a key role in the proofs of the main results of the subsection, Theorem 3.4 and Corollary 3.5 below, and are of independent interest.

Theorem 3.1 *There exist universal constants $K_1 := 1/2$ and $K_p > 0$, $p \in (1, 2]$, independent of (σ^2, ν, b), such that for any $t > 0$ and $\kappa \in (0, 1]$, there exists a coupling $(X_t, X_t^{(\kappa)})$ satisfying for all $p \in [1, 2]$ that*

$$\mathbb{E}[|X_t - X_t^{(\kappa)}|^p]^{1/p} \leq \min\{\sqrt{2t}\overline{\sigma}_\kappa, K_p \kappa \varphi_\kappa^{2/p}\}, \qquad where \ \varphi_\kappa := \frac{\overline{\sigma}_\kappa}{\sqrt{\overline{\sigma}_\kappa^2 + \sigma^2}}. \quad (3.1)$$

Bounds on the Kolmogorov distance may require the following generalisation of Orey's condition, which makes the distribution of X_t sufficiently regular (see Sato [62, Proposition 28.3]).

Assumption 3.2 *We have $\inf_{u \in (0,1]} u^{\delta-2}(\overline{\sigma}_u^2 + \sigma^2) > 0$ for some $\delta \in (0, 2]$.*

Note that if $\sigma \neq 0$, Assumption 3.2 holds with $\delta = 2$. If $\sigma = 0$ and δ satisfies the inequality in Assumption 3.2, we must have $\beta \geq \delta$, where β is the Blumenthal–Getoor index defined in (2.6) above. In fact, models typically used in applications either have $\sigma \neq 0$ or Assumption 3.2 holds with $\delta = \beta$. However, for Orey's process (defined in [62, Example. 41.23]), Assumption 3.2 holds for some $\delta < \beta$, but not for $\delta = \beta$ (see details in Bang et al. [4, Example. 2.7]).

Theorem 3.3 (a) *There exists a constant $C_{BE} \in (0, \frac{1}{2})$ such that for any $t > 0$, $\kappa \in (0, 1]$, we have*

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P}[X_t \leq x] - \mathbb{P}[X_t^{(\kappa)} \leq x] \right| \leq \frac{C_{BE} \kappa \varphi_\kappa^3}{\overline{\sigma}_\kappa \sqrt{t}}. \quad (3.2)$$

(b) *Let Assumption* 3.2 *hold. Then for every* $T > 0$, *there exists a constant* $C > 0$, *depending only on* (T, δ, σ, ν), *such that for any* $\kappa \in (0, 1]$ *and* $t \in (0, T]$, *we have*

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P}[X_t \leq x] - \mathbb{P}[X_t^{(\kappa)} \leq x] \right| \leq (Ct^{-1/\delta} \min\{\sqrt{t}\overline{\sigma}_\kappa, \kappa\varphi_\kappa\})^{2/3}. \quad (3.3)$$

Denote $x^+ := \max\{x, 0\}$ for $x \in \mathbb{R}$. The next result quantifies the Wasserstein distance between the laws of the vectors $\underline{\chi}_t$ and $\underline{\chi}_t^{(\kappa)}$.

Theorem 3.4 *For any* $\kappa \in (0, 1]$ *and* $t > 0$, *there exists a coupling between* X *and* $X^{(\kappa)}$ *on the interval* $[0, t]$ *such that we have for* $p \in \{1, 2\}$ *the inequalities*

$$\mathbb{E}[\max\{|X_t - X_t^{(\kappa)}|, |\underline{X}_t - \underline{X}_t^{(\kappa)}|\}^p]^{1/p} \leq \mu_p(\kappa, t), \quad (3.4)$$

where

$$\mu_1(\kappa, t) := \min\{2\sqrt{2t}\overline{\sigma}_\kappa, \kappa\varphi_\kappa^2\}\left(1 + \log^+\left(2\sqrt{2t}(\overline{\sigma}_\kappa/\kappa)\varphi_\kappa^{-2}\right)\right),$$

$$\mu_2(\kappa, t) := \sqrt{2}\mu_1(\kappa, t)$$
$$+ \min\{\sqrt{2t}\overline{\sigma}_\kappa, K_2\kappa\varphi_\kappa\}\sqrt{1 + 2\log^+\left(K_2^{-1}\sqrt{2t}(\overline{\sigma}_\kappa/\kappa)\varphi_\kappa^{-1}\right)}, \quad (3.5)$$

with $\varphi_\kappa = \overline{\sigma}_\kappa/\sqrt{\overline{\sigma}_\kappa^2 + \sigma^2}$ *and the universal constant* K_2 *from Theorem* 3.1, *and*

$$\mathbb{E}[|\underline{\tau}_t(X) - \underline{\tau}_t(X^{(\kappa)})|] \leq \mu_0^\tau(\kappa, t) := \sqrt{t}(\kappa/\overline{\sigma}_\kappa)\varphi_\kappa^3. \quad (3.6)$$

Moreover, under Assumption 3.2 *for some* $\delta \in (0, 2]$, *there exists for every* $T > 0$ *a constant* $C > 0$, *dependent only on* (T, δ, σ, ν), *such that for all* $t \in [0, T]$ *and* $\kappa \in (0, 1]$, *we have*

$$\mathbb{E}[|\underline{\tau}_t(X) - \underline{\tau}_t(X^{(\kappa)})|] \leq \mu_\delta^\tau(\kappa, t), \quad (3.7)$$

where $\psi_\kappa := C\kappa\varphi_\kappa$ *and*

$$\mu_\delta^\tau(\kappa, t) := \begin{cases} \min\{t, \psi_\kappa^\delta\} + t^{1-\frac{2}{3\delta}}\psi_\kappa^{\frac{2}{3}}(1 - \min\{1, t^{-\frac{1}{\delta}}\psi_\kappa\}^{\delta-\frac{2}{3}}), & \delta \neq \frac{2}{3}, \\ \min\{t, \psi_\kappa^{\frac{2}{3}}\}(1 + \log^+(t\psi_\kappa^{-\frac{2}{3}})), & \delta = \frac{2}{3}. \end{cases} \quad (3.8)$$

The proof of Theorem 3.4, given in Sect. 7.2 below, constructs the *SBG coupling* $(X, X^{(\kappa)})$, satisfying the above inequalities, in terms of the distribution functions of the marginals X_s and $X_s^{(\kappa)}$ (for $s > 0$) and the coupling used in (2.1); see González Cázares et al. [35] for the latter. The key idea is to couple $\underline{\chi}_t$ and $\underline{\chi}_t^{(\kappa)}$ so that they share the stick-breaking process in their respective SB representations (2.1), while the increments of the associated Lévy processes over each interval $[L_n, L_{n-1}]$ are coupled so that they minimise appropriate Wasserstein distances. This coupling produces a bound on the distance between $\underline{\chi}_t$ and $\underline{\chi}_t^{(\kappa)}$ that depends only

on the distances between the marginals of X_s and $X_s^{(\kappa)}$, $s > 0$, so that Theorems 3.1 and 3.3 can be applied. We stress that the bound in (3.4) cannot be obtained from Doob's L^2-maximal inequality and Theorem 3.1: if the processes X and $X^{(\kappa)}$ are coupled in such a way that $X_t - X_t^{(\kappa)}$ satisfies the inequality in (3.1), the difference process $(X_s - X_s^{(\kappa)})_{s \in [0,t]}$ need not be a martingale.

Inequality (3.4) holds without assumptions on X and is at most a logarithmic factor worse than the marginal inequality (3.1) for $p \in \{1, 2\}$, with the upper bound satisfying $\mu_p(\kappa, t) \le 2\kappa \log(1/\kappa)$ for all sufficiently small κ. Moreover, by Jensen's inequality, the SBG coupling satisfies for all $1 < p < 2$ the inequality

$$\mathbb{E}[\max\{|X_t - X_t^{(\kappa)}|, |\underline{X}_t - \underline{X}_t^{(\kappa)}|\}^p]^{1/p} \le \mu_2(\kappa, t).$$

In the absence of a Brownian component (i.e., $\sigma = 0$), we have $\varphi_\kappa = 1$, making the upper bound $\mu_2(\kappa, t)$ proportional to $\mu_1(\kappa, t)$ as $\kappa \to 0$. If $\sigma > 0$, then $\mu_1(\kappa, t) \le 2\kappa \overline{\sigma}_\kappa^2 \log(1/(\kappa \overline{\sigma}_\kappa))/\sigma^2$ for all small κ and typically, $\mu_2(\kappa, t)$ is proportional to $\kappa \overline{\sigma}_\kappa \sqrt{\log(1/(\kappa \overline{\sigma}_\kappa))}$ as $\kappa \to 0$, which dominates $\mu_1(\kappa, t)$.

The bound in (3.6) holds without assumptions on the Lévy process X, while (3.7) requires Assumption 3.2 and is the sharper the larger the value of $\delta \in (0, 2]$ satisfying Assumption 3.2. If $\sigma > 0$, the inequality in (3.6) is sharper than (3.7), i.e. $\mu_0^\tau(t, \kappa) \le \mu_2^\tau(t, \kappa)$ for all small $\kappa > 0$. However, if $\sigma = 0$ and $\delta \in (0, 2)$ satisfies Assumption 3.2, then typically $\mu_0^\tau(\kappa, t)$ is proportional to $\kappa^{\delta/2}$, while $\mu_\delta^\tau(\kappa, t)$ behaves as $\kappa^{\min\{2/3,\delta\}}(1 + \log(1/\kappa)\mathbb{1}_{\{2/3\}}(\delta))$ as $\kappa \to 0$, implying that (3.7) is sharper than (3.6) for $\delta < 4/3$. The smallest of the upper bounds in (3.6) and (3.7) is

$$\mu_*^\tau(\kappa, t) := \min\left\{\mu_0^\tau(\kappa, t), \inf\{\mu_\delta^\tau(\kappa, t) : \delta \in (0, 2] \text{ satisfies Assumption 3.2}\}\right\}.$$

Under Assumption 3.2, for some constant $c_t > 0$ and all $\kappa \in (0, 1]$, we have

$$\mu_*^\tau(\kappa, t) \le c_t \kappa^{\max\{\delta/2, \min\{2/3,\delta\}\}}\left(1 + \log(1/\kappa)\mathbb{1}_{\{2/3\}}(\delta)\right). \tag{3.9}$$

For any $a \in \mathbb{R}^d$, let $|a| := \sum_{i=1}^d |a_i|$ denote its ℓ^1-norm. Recall that for $p \ge 1$, the L^p-Wasserstein distance (Villani [68, Definition 6.1]) between the laws of random vectors ξ and ζ in $\mathbb{R}^d$ can be defined as

$$\mathcal{W}_p(\xi, \zeta) := \inf\{\mathbb{E}[|\xi' - \zeta'|^p]^{1/p} : \xi' \overset{d}{=} \xi, \zeta' \overset{d}{=} \zeta\}. \tag{3.10}$$

Theorem 3.4 implies a bound on the L^p-Wasserstein distance between the vectors $\underline{\chi}_t$ and $\underline{\chi}_t^{(\kappa)}$, extending the bound on the distance between the laws of the marginals $\overline{X}_t$ and $X_t^{(\kappa)}$ in Mariucci and Reiß [50, Theorem 9].

Corollary 3.5 *Fix $\kappa \in (0, 1]$ and $t > 0$. Then*

$$\mathcal{W}_p((X_t, \underline{X}_t), (X_t^{(\kappa)}, \underline{X}_t^{(\kappa)})) \le 2\left(\mathbb{1}_{\{p=1\}}\mu_1(\kappa, t) + \mathbb{1}_{\{1<p\le2\}}\mu_2(\kappa, t)\right), \quad p \in [1, 2],$$

$$\mathcal{W}_p(\underline{\tau}_t(X), \underline{\tau}_t(X^{(\kappa)})) \le t^{1-1/p}\mu_*^\tau(\kappa, t)^{1/p}, \qquad\qquad p \ge 1.$$

Moreover, $\mathcal{W}_p(\underline{\chi}_t, \underline{\chi}_t^{(\kappa)})$ is bounded by twice the sum of both bounds for $p \in [1, 2]$.

Given the bounds in Corollary 3.5 and Theorem 3.3, it is natural to inquire about the convergence in the Kolmogorov distance of the components of the vector $(\underline{X}_t^{(\kappa)}, \underline{\tau}_t(X^{(\kappa)}))$ to those of $(\underline{X}_t, \underline{\tau}_t(X))$ as $\kappa \to 0$. This question is addressed by Corollary 3.14 below.

Dereich [21, Theorem 6.1] used the famous Komlós–Major–Tusnády (KMT) coupling to bound the L^2-Wasserstein distance between the paths of X and $X^{(\kappa)}$ on $[0, t]$ in the supremum norm, implying a bound on $\mathcal{W}_2((X_t, \underline{X}_t), (X_t^{(\kappa)}, \underline{X}_t^{(\kappa)}))$ proportional to $\kappa \log(1/\kappa)$ as $\kappa \to 0$; cf. [21, Corollary 6.2]. If $\sigma > 0$, $\mu_2(\kappa, t)$ in (3.4) is bounded by a multiple of $\kappa \overline{\sigma}_\kappa \log(1/(\kappa \overline{\sigma}_\kappa))$ for small κ and is thus smaller than a multiple of $\kappa^{2-q/2}$ for any $q \in (\beta, 2)$ (where β is the BG index defined in (2.6)). As mentioned above, $\mu_2(\kappa, t)$ is bounded by a multiple of $\kappa \log(1/\kappa)$ for small κ. Unlike the SBG coupling which underpins Theorem 3.4, the KMT coupling does not imply a bound on the distance between the times of the infima $\underline{\tau}_t(X)$ and $\underline{\tau}_t(X^{(\kappa)})$ as these are not Lipschitz functionals of the trajectories with respect to the supremum norm.

Remark 3.6 The bounds on $\mathbb{E}[|\underline{\tau}_t(X) - \underline{\tau}_t(X^{(\kappa)})|]$ in Theorem 3.4 and Corollary 3.5, based on the SB representation in (2.1), require a control on the expected difference between the signs of the components of $(X_s, X_s^{(\kappa)})$ as either s or κ tend to zero. This is achieved via the minimal transport coupling (see (7.1) and Lemma 7.2 below) and a general bound in Theorem 3.3 on the Kolmogorov distance. However, further improvements seem possible in the finite variation case if the *natural drift* (i.e., the drift of X when small jumps are not compensated) is nonzero. Intuitively, the sign of the natural drift determines the sign of both components of $(X_s, X_s^{(\kappa)})$ with overwhelming likelihood as $s \to 0$. This suggestion is left for future research.

3.1.1 Why are Theorems 3.1 and 3.3 an improvement on the existing bounds on the distance between the laws of X_t and its Gaussian approximation $X_t^{(\kappa)}$?

Theorem 3.1 bounds the Wasserstein distance between X_t and $X_t^{(\kappa)}$. The inequality in (3.1) sharpens the bound $\mathbb{E}[|X_t - X_t^{(\kappa)}|^p]^{1/p} \leq \min\{\sqrt{2t}\overline{\sigma}_\kappa, K_p\kappa\}$, established by Mariucci and Reiß [50, Theorem 9], as follows. The factor $\varphi_\kappa^{2/p} \in [0, 1]$ tends to zero (as $\kappa \to 0$) as a constant multiple of $\overline{\sigma}_\kappa^{2/p}$ if the Brownian component is present (i.e., $\sigma > 0$) and is equal to 1 when $\sigma = 0$. The bound in (3.1) cannot be improved in general in the sense that there exists a Lévy process for which, up to constants, the reverse inequality holds (see Mariucci and Reiß [50, Remark 3] and Fournier [27, Sect. 4]).

The proof of Theorem 3.1, given in Sect. 7.1 below, decomposes the increment $M_t^{(\kappa)}$ of the Lévy martingale $M^{(\kappa)} := \sigma B + J^{1,\kappa}$ into a sum of m i.i.d. copies of $M_{t/m}^{(\kappa)}$ and applies a Berry–Esseen-type bound established by Rio [60] for the Wasserstein distance in the context of a central limit theorem (CLT) as $m \to \infty$. The small-time moment asymptotics of $M_{t/m}^{(\kappa)}$ in Figueroa-López [26] imply that $M_t^{(\kappa)}$ is much closer to the Gaussian limit in the CLT if the Brownian component is present than if $\sigma = 0$. This explains a vastly superior rate in (3.1) in the case $\sigma^2 > 0$.

The proof of Theorem 3.3 is in Sect. 7.1 below. Part (a) follows a similar strategy as in the proof of Theorem 3.1, applying the Berry–Esseen theorem (instead of [60,

Theorem 4.1]) to bound the Kolmogorov distance. By the same reasoning applied in the previous paragraph to the bound in (3.1), the rate in (3.2) is far better if $\sigma^2 > 0$. The proof of Theorem 3.3 (b) bounds the density of X_t using results due to Picard [58] and applies (3.1).

Note that no assumption is made on the Lévy process X in Theorem 3.3 (a). In particular, Assumption 3.2 is not required in part (a); however, if Assumption 3.2 is not satisfied, implying in particular that $\sigma = 0$, it is possible for the bound in (3.2) not to vanish as $\kappa \to 0$ even if the Lévy process has infinite activity, i.e., $\nu(\mathbb{R} \setminus \{0\}) = \infty$. In fact, if $\sigma = 0$, the bound in (3.2) vanishes (as $\kappa \to 0$) if and only if $\overline{\sigma}_\kappa / \kappa \to \infty$, which is also a necessary and sufficient condition for the weak limit $\overline{\sigma}_\kappa^{-1} J^{1,\kappa} \xrightarrow{d} W$ to hold whenever ν has no atoms in a neighbourhood of 0 (see Asmussen and Rosiński [2, Proposition 2.2]).

If X has a Brownian component (i.e., $\sigma \neq 0$), the bound on the total variation distance between the laws of X_t and $X_t^{(\kappa)}$ established in Mariucci and Reiß [50, Proposition 8] implies the bound

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P}[X_t \leq x] - \mathbb{P}[X_t^{(\kappa)} \leq x] \right| \leq \frac{\min\{\sqrt{8t}\,\overline{\sigma}_\kappa, \kappa\}}{\sqrt{2\pi\sigma^2 t}}$$

on the Kolmogorov distance. This inequality is both generalised and sharpened (as $\kappa \to 0$) by the bound in (3.2). Further improvements to the bound on the total variation were made in Carpentier et al. [12], but the implied rates for the Kolmogorov distance are worse than the ones in Theorem 3.3 and require model restrictions when $\sigma = 0$ (beyond those of Theorem 3.3 (b)) that can be hard to verify (see [12, Sect. 2.1.1]).

We stress that the dependence on t in the bounds of Theorem 3.3 is explicit. This is crucial in the proof of Theorem 3.4 as we need to apply (3.2), (3.3) over intervals of small random lengths. A related result proved by Dia [24, Proposition 10] contains similar bounds which are non-explicit in t and suboptimal in κ.

If Assumption 3.2 is satisfied, the parameter δ in part (b) of Theorem 3.3 should be taken as large as possible to get the sharpest inequality in (3.3). If $\sigma \neq 0$ (equivalently $\delta = 2$), the bound in part (a) has a faster decay in κ than the bound in part (b). If $\sigma = 0$ (equivalently $0 < \delta < 2$), it is possible for the bound in part (a) to be sharper than that in part (b) or vice versa. Indeed, it is easy to construct a Lévy measure ν such that $\delta \in (0, 2)$ in Theorem 3.3 (b) satisfies $\lim_{u\downarrow 0} u^{\delta-2}\overline{\sigma}_u^2 = \inf_{u\in(0,1]} u^{\delta-2}\overline{\sigma}_u^2 = 1$. Then the bound in (3.2) is a multiple of $t^{-1/2}\kappa^{\delta/2}$ as $t, \kappa \to 0$, while that in (3.3) behaves as $t^{-2/(3\delta)}\kappa^{2/3}\min\{1, t^{1/3}\kappa^{-\delta/3}\}$. Hence one bound may be sharper than the other depending on the value of δ, as t and/or κ tend to zero. In fact, we use the bound in part (b) only when the maximal δ satisfying the assumption of Theorem 3.3 (b) is smaller than 4/3. In that case, the activity of the Lévy measure around 0 is bounded away from its maximal possible activity, which would correspond to δ being close to 2.

3.2 The bias of locally Lipschitz and discontinuous functions of $\underline{\chi}_t$ and the Kolmogorov distance between the vectors $\underline{\chi}_t$ and $\underline{\chi}_t^{(\kappa)}$

The main tool for studying the bias of locally Lipschitz and discontinuous payoff functions of $\underline{\chi}_t$ is the SBG coupling underpinning Theorem 3.4. The Lipschitz case

follows trivially from Theorem 3.4. Indeed, for any $d \in \mathbb{N}$, let $\mathrm{Lip}_K(\mathbb{R}^d)$ denote the space of real-valued Lipschitz functions on $\mathbb{R}^d$ (under the ℓ^1-norm given above (3.10)) with Lipschitz constant $K > 0$ and note that the triangle inequality and Theorem 3.4 imply for any time horizon $T > 0$ the bounds on the bias

$$
\begin{aligned}
|\mathbb{E}[f(X_T, \underline{X}_T)] - \mathbb{E}[f(X_T^{(\kappa)}, \underline{X}_T^{(\kappa)})]| &\leq 2K\mu_1(\kappa, T), \\
\left|\mathbb{E}[g(\underline{\tau}_T)] - \mathbb{E}\left[g\left(\underline{\tau}_T(X^{(\kappa)})\right)\right]\right| &\leq K'\mu_*^{\tau}(\kappa, T),
\end{aligned}
\tag{3.11}
$$

where $f \in \mathrm{Lip}_K(\mathbb{R}^2)$ satisfies $\mathbb{E}[|f(X_T, \underline{X}_T)|] < \infty$ and $g \in \mathrm{Lip}_{K'}(\mathbb{R})$.

In applications, the process X is often used to model log-returns of a risky asset $(S_0 e^{X_t})_{t \geq 0}$. It is thus important to understand the bias of a Monte Carlo estimator for the class $\mathrm{locLip}_K(\mathbb{R}^2)$ of locally Lipschitz functions, with the defining property $f \in \mathrm{locLip}_K(\mathbb{R}^2)$ if and only if

$$
|f(x, y) - f(x', y')| \leq K(|e^x - e^{x'}| + |e^y - e^{y'}|) \qquad \text{for any } x, x', y, y' \in \mathbb{R}
$$

(equivalently $(x, y) \mapsto f(\log x, \log y)$ is in $\mathrm{Lip}_K((0, \infty) \times (0, \infty))$). Such payoffs arise in risk management (e.g. absolute drawdown) and in the pricing of hindsight call, perpetual American call and lookback put options.

Proposition 3.7 *Let $f \in \mathrm{locLip}_K(\mathbb{R}^2)$ and assume $\int_{[1,\infty)} e^{2x} \nu(dx) < \infty$, where ν is the Lévy measure of X. For any $T > 0$ and $\kappa \in (0, 1]$ and $\mu_2(\kappa, T)$ defined in (3.5), the SBG coupling satisfies*

$$
\mathbb{E}[|f(X_T, \underline{X}_T) - f(X_T^{(\kappa)}, \underline{X}_T^{(\kappa)})|] \leq 4K\mathbb{E}[e^{2X_T}]^{1/2}(1 + e^{\bar{\sigma}_\kappa^2 T})\mu_2(\kappa, T).
$$

The assumption $\int_{[1,\infty)} e^{2x} \nu(dx) < \infty$ is equivalent to $\mathbb{E}[e^{2X_T}] < \infty$ (see Sato [62, Theorem 25.3]), which is a natural requirement as the asset price model $(S_0 e^{X_t})_{t \geq 0}$ ought to have finite variance. Moreover, via the Lévy–Khintchine formula, an explicit bound on the expectation $\mathbb{E}[e^{2X_T}]$ (and hence the constant in the inequality of Proposition 3.7) in terms of the Lévy triplet of X can be obtained.

The bound in Proposition 3.7 does not via duality imply an analogous bound involving a function of the supremum $\overline{X}_T$, since the assumption is not symmetric in the Lévy measure. However, for $f(X_T, \overline{X}_T)$ the proof of Proposition 3.7 in Sect. 7.3 yields

$$
\mathbb{E}[|f(X_T, \overline{X}_T) - f(X_T^{(\kappa)}, \overline{X}_T^{(\kappa)})|] \leq 4K\left(\mathbb{E}[\exp(2\overline{X}_T)] + \mathbb{E}[\exp(2\overline{X}_T^{(\kappa)})]\right)^{1/2}\mu_2(\kappa, T),
$$

where both expectations $\mathbb{E}[e^{2\overline{X}_T}]$ and $\mathbb{E}[e^{2\overline{X}_T^{(\kappa)}}]$ are finite under our natural assumption $\int_{[1,\infty)} e^{2x} \nu(dx) < \infty$ and can be bounded explicitly in terms of the Lévy triplet of X; see e.g. the proof of [35, Proposition 2]. Thus the bias for $f \in \mathrm{locLip}_K(\mathbb{R}^2)$ (applied to either $(X_T, \overline{X}_T)$ or $(X_T, \underline{X}_T)$) is at most a multiple of $\kappa \log(1/\kappa)$, as is by (3.11) the case for $f \in \mathrm{Lip}_K(\mathbb{R}^2)$; see the discussion after Theorem 3.4.

In financial markets, the class of barrier-type functions arises naturally. For constants $K, M \geq 0$, $y < 0$, define

$$
\mathrm{BT}_1(y, K, M) := \{f : f(x, z) = h(x)\mathbb{1}_{[y,\infty)}(z), h \in \mathrm{Lip}_K(\mathbb{R}), 0 \leq h \leq M\}. \tag{3.12}
$$

Note that $\mathbb{1}_{[y,\infty)}$ lies in $\mathrm{BT}_1(y, 0, 1)$ and satisfies $\mathbb{E}[\mathbb{1}_{[y,\infty)}(\underline{X}_T)] = \mathbb{P}[\underline{X}_T \geq y]$. Moreover, a down-and-out put option payoff $x \mapsto \max\{e^k - e^x, 0\}\mathbb{1}_{[y,\infty)}(x)$, for some constants $y < 0 < k$, is in $\mathrm{BT}_1(y, e^k, e^k - e^y)$. Bounding the bias of the estimators for functions in $\mathrm{BT}_1(y, K, M)$ requires the following regularity of the distribution of $\underline{X}_T$ at y.

Assumption 3.8 *Given $C, \gamma > 0$ and $y < 0$, we have the inequality*

$$|\mathbb{P}[\underline{X}_T \leq x + y] - \mathbb{P}[\underline{X}_T \leq y]| \leq C|x|^\gamma \qquad \text{for all } x \in \mathbb{R}.$$

Proposition 3.9 *Let $f \in \mathrm{BT}_1(y, K, M)$ for some $K, M \geq 0$ and $y < 0$. If y and some $C, \gamma > 0$ satisfy Assumption 3.8, then for any $T > 0$ and $\kappa \in (0, 1]$, the SBG coupling satisfies*

$$\mathbb{E}[|f(X_T, \underline{X}_T) - f(X_T^{(\kappa)}, \underline{X}_T^{(\kappa)})|]$$

$$\leq K\mu_1(\kappa, T) + M' \min\left\{\mu_1(\kappa, T)^{\frac{\gamma}{1+\gamma}}, \mu_2(\kappa, T)^{\frac{2\gamma}{2+\gamma}}\right\}, \qquad (3.13)$$

where $M' = M \max\{(1 + 1/\gamma)(2C\gamma)^{1/(1+\gamma)}, (1 + 2/\gamma)(C\gamma)^{2/(2+\gamma)}\}$.

Remark 3.10 Because we have $\mu_1(\kappa, T) \to 0$ and $\mu_2(\kappa, T) \to 0$ as $\kappa \to 0$ and $\gamma/(1 + \gamma) < 2\gamma/(2 + \gamma)$ for all $\gamma > 0$, the bound in (3.13) is typically dominated by a multiple of $\mu_1(\kappa, T)^{\gamma/(1+\gamma)}$, if $\sigma \neq 0$ and $\beta < 2 - \gamma$ (recall the definition of the BG index β in (2.6)), or by $\mu_2(\kappa, T)^{2\gamma/(1+\gamma)}$, otherwise. By Hölder's inequality, f in (3.13) need not be bounded if appropriate moments of X exist.

The proof of Proposition 3.9 is in Sect. 7.3 below. Assumption 3.8 with $\gamma = 1$ requires the distribution function of $\underline{X}_T$ to be locally Lipschitz at y. By the Lebesgue differentiation theorem (see Cohn [18, Theorem 6.3.3]), any distribution function is differentiable Lebesgue-a.e., implying that Assumption 3.8 holds for $\gamma = 1$ and a.e. $y < 0$. However, there indeed exist Lévy processes that do not satisfy Assumption 3.8 with $\gamma = 1$ for countably many levels y; see the example in González Cázares et al. [35, App. B]. (In fact, that example shows that Assumption 3.8 may fail at countably many levels y for any $\gamma \in (0, 1]$.) Proposition 3.15 below provides simple sufficient conditions, in terms of the Lévy triplet of X, for Assumption 3.8 to hold with $\gamma = 1$ for all $y < 0$. In particular, this is the case if $\sigma \neq 0$.

The next class of payoffs arises in the analysis of the duration of a drawdown. For $K, M \geq 0$, $s \in (0, T)$, let

$$\mathrm{BT}_2(s, K, M) := \{f : f(x, t) = h(x)\mathbb{1}_{(s,T]}(t), h \in \mathrm{Lip}_K(\mathbb{R}^2), 0 \leq h \leq M\}. \quad (3.14)$$

The biases of these functions clearly include $|\mathbb{P}[\underline{\tau}_T(X) > s] - \mathbb{P}[\underline{\tau}_T(X^{(\kappa)}) > s]|$. Analogously to Proposition 3.9, we require the following regularity from the distribution function of $\underline{\tau}_T(X)$.

Assumption 3.11 *Given $C, \gamma > 0$ and $s \in (0, T)$, we have the inequality*

$$|\mathbb{P}[\underline{\tau}_T(X) \leq s] - \mathbb{P}[\underline{\tau}_T(X) \leq s + t]| \leq C|t|^\gamma \qquad \text{for all } t \in \mathbb{R}.$$

Proposition 3.12 *Let Assumption* 3.11 *hold for some* $s \in (0, T)$ *and* $C, \gamma > 0$. *Let* $f \in \mathrm{BT}_2(s, K, M)$ *for some* $K, M \geq 0$. *Then for all* $\kappa \in (0, 1]$, *the SBG coupling satisfies*

$$\mathbb{E}[|f(\underline{\chi}_T) - f(\underline{\chi}_T^{(\kappa)})|] \leq 2K\mu_1(\kappa, T)$$
$$+ M(2C\gamma)^{1/(1+\gamma)}(1 + 1/\gamma)\mu_*^\tau(\kappa, T)^{\gamma/(1+\gamma)}. \quad (3.15)$$

Remark 3.13 As in Remark 3.10, the bound in (3.15) is asymptotically proportional to $\mu_*^\tau(\kappa, T)^{\gamma/(1+\gamma)}$ as $\kappa \to 0$. The inequality (3.15) can be generalised to unbounded functions f if appropriate moments of X exist.

If X is not a compound Poisson process, Assumption 3.11 holds with $\gamma = 1$ for all $s \in (0, T)$ since by Lemma 7.7 below, $\underline{\tau}_T(X)$ has a locally bounded density, making the distribution function of $\underline{\tau}_T(X)$ locally Lipschitz on $(0, T)$. Assumption 3.11 is satisfied if either $\nu(\mathbb{R} \setminus \{0\}) = \infty$ or $\sigma \neq 0$. In particular, Assumption 3.2 implies Assumption 3.11. The proof of Proposition 3.12 is in Sect. 7.3 below.

The classes $\mathrm{BT}_1(y, K, M)$ and $\mathrm{BT}_2(s, K, M)$ of payoffs in (3.12) and (3.14), respectively, consist of bounded functions. We stress that boundedness of the payoffs is not essential in Propositions 3.9 and 3.12. It can be substituted by a combination of a local Lipschitz assumption and a moment bound on the tails of the Lévy measure (cf. Proposition 3.7), typically satisfied in applications. The crucial Lemma 7.5, applied in the proofs of Propositions 3.9 and 3.12, can be generalised to unbounded payoffs by substituting the almost sure bound on the payoff in its current proof (see (7.11) below) with a bound on its expected value via Hölder's inequality. The details are omitted for ease of exposition.

As a consequence of Proposition 3.9 (resp. 3.12), if Assumption 3.8 (resp. 3.11) holds uniformly over y for fixed $C, \gamma > 0$, then $\underline{X}_T^{(\kappa)}$ (resp. $\underline{\tau}_T(X^{(\kappa)})$) converges to $\underline{X}_T$ (resp. $\underline{\tau}_T(X)$) in the Kolmogorov distance as $\kappa \to 0$.

Corollary 3.14 (a) *Suppose* $C, \gamma > 0$ *satisfy Assumption* 3.8 *for all* $y < 0$. *Then for* $\kappa \in (0, 1]$,

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P}[\underline{X}_T \leq x] - \mathbb{P}[\underline{X}_T^{(\kappa)} \leq x] \right| \leq M' \min \left\{ \mu_1(\kappa, T)^{\frac{\gamma}{1+\gamma}}, \mu_2(\kappa, T)^{\frac{2\gamma}{2+\gamma}} \right\},$$

where $M' = \max\{(1 + 1/\gamma)(2C\gamma)^{1/(1+\gamma)}, (1 + 2/\gamma)(C\gamma)^{2/(2+\gamma)}\}$.

(b) *Suppose* $C, \gamma > 0$ *satisfy Assumption* 3.11 *for all* $s \in [0, T]$. *Then for* $\kappa \in (0, 1]$, *we have*

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P}[\underline{\tau}_T(X) \leq x] - \mathbb{P}[\underline{\tau}_T(X^{(\kappa)}) \leq x] \right| \leq (2C\gamma)^{\frac{1}{1+\gamma}}(1 + 1/\gamma)\mu_*^\tau(\kappa, T)^{\frac{\gamma}{1+\gamma}}.$$

Proposition 3.15 gives sufficient conditions (in terms of the triplet (σ^2, ν, b)) for Assumptions 3.8 and 3.11 to hold for all $y < 0$ and $s \in [0, T]$, respectively. Recall that a function $f(x)$ is said to be regularly varying with index r as $x \to 0$ if $\lim_{x \to 0} f(\lambda x)/f(x) = \lambda^r$ for every $\lambda > 0$.

Proposition 3.15 *Let $\overline{v}_+(x) := v([x,\infty))$ and $\overline{v}_-(x) := v((-\infty,-x])$ for $x > 0$ and let β be the BG index of X defined in* (2.6). *Suppose that either* (I) $\sigma > 0$ *or* (II) *the Lévy measure v satisfies the following conditions*: $\overline{v}_+(x)$ *is regularly varying with index $-\beta$ as $x \to 0$ and either*

- $\beta = 2$ *and* $\liminf_{x\to 0} \overline{v}_+(x)/\overline{v}_-(x) > 0$, *or*
- $\beta \in (1,2)$ *and* $\lim_{x\to 0} \overline{v}_+(x)/\overline{v}_-(x) \in (0,\infty]$.

Then there exist constants $\gamma > 0$ and C such that Assumption 3.11 *holds with γ, C for all $s \in [0,T]$. Moreover, for any compact $I \subseteq (-\infty, 0)$, Assumption* 3.8 *holds with $\gamma = 1$ and some constant C_I for all $y \in I$.*

Note that Proposition 3.15 holds if the roles of $\overline{v}_+$ and $\overline{v}_-$ are interchanged, i.e., $\overline{v}_-(x)$ is regularly varying and the limit conditions are satisfied by the quotients $\overline{v}_-(x)/\overline{v}_+(x)$. The assumptions of Proposition 3.15 are satisfied by most models used in practice that have infinite variation, including tempered stable and subordinated Brownian motion processes.

Proposition 3.15 is a consequence of a more general result, Proposition 7.9 below, stating that Assumptions 3.11 and 3.8 hold uniformly and locally uniformly, respectively, if over short time horizons, X is "attracted to" an α-stable process with non-monotone paths; see Sect. 7.3 below for details. In this case, $\rho := \lim_{t\to 0} \mathbb{P}[X_t > 0]$ exists in $(0,1)$, and γ in the conclusion of Proposition 3.15, satisfying Assumption 3.11 on $[0,T]$, can be arbitrarily chosen in $(0, \min\{\rho, 1 - \rho\})$. In contrast to Assumption 3.11, a simple sufficient condition for the uniform version of Assumption 3.8, required in Corollary 3.14(a), remains elusive beyond special cases such as stable or tempered stable processes with γ in the interval $(0, \alpha(1 - \rho))$, where α is the stability parameter and ρ is as above.

The reasoning in the previous paragraph does not apply if X is attracted to a linear drift, since the paths of the limit are monotone, implying $\rho \in \{0, 1\}$. This occurs if X is of finite variation with $b \neq \int_{(-1,1)} x v(dx)$ or if X is of infinite variation with $\beta = 1$ (see Ivanovs [38, Theorem 2] for details). In these cases, the uniform convergence in Corollary 3.14 may fail due to an atom in the limit.

4 Simulation and the computational complexity of MC and MLMC

In this section, we describe a method using MC or MLMC for simulating the vector $\underline{\chi}_T^{(\kappa)} = (X_T^{(\kappa)}, \underline{X}_T^{(\kappa)}, \underline{\tau}_T(X^{(\kappa)}))$ (SBG-Alg in Sect. 4.1) and analyse the computational complexities for various locally Lipschitz and discontinuous functions of $\underline{\chi}_T^{(\kappa)}$ (Sect. 4.2). The numerical performance of SBG-Alg, which is based on the SB representation in (2.1), (2.2) of $\underline{\chi}_T^{(\kappa)}$, is far superior to that of the "obvious" algorithm for jump-diffusions (see Algorithm 2 below), particularly when the jump intensity is large (cf. Sects. 4.1.2 and 4.1.3). Moreover, SBG-Alg is designed with MLMC in mind, which turns out not to be feasible in general for the "obvious" algorithm (see Sect. 4.1.2).

4.1 Simulation of $\underline{\chi}_T^{(\kappa)}$

The main aim of the subsection is to develop a simulation algorithm for the pair of vectors $(\underline{\chi}_T^{(\kappa)}, \underline{\chi}_T^{(\kappa')})$ at levels $\kappa, \kappa' \in (0,1]$ over a time horizon $[0,T]$ such that the

---

Algorithm 1 Simulation of the law $\Pi_t^{\kappa_1,\kappa_2}$

Require: Cutoff levels $1 \geq \kappa_1 > \kappa_2 > 0$ and time horizon $t > 0$.
1: Compute b_{κ_i} and $\overline{\sigma}_{\kappa_i}^2$ for $i \in \{1, 2\}$ and $\overline{v}(\kappa_2)$
2: Sample $W_t \sim \mathcal{N}(0, t)$, $N_t \sim \mathrm{Poi}(\overline{v}(\kappa_2)t)$ and $\lambda_k \sim v(\cdot \setminus (-\kappa_2, \kappa_2))/\overline{v}(\kappa_2)$ for $k \in \{1, \ldots, N_t\}$
3: Put $J_t^{2,\kappa_i} := \sum_{k=1}^{N_t} \lambda_k \mathbb{1}_{\{|\lambda_k| \geq \kappa_i\}}$ for $i \in \{1, 2\}$
4: **return** $(Z_t^{(\kappa_1)}, Z_t^{(\kappa_2)})$, where $Z_t^{(\kappa_i)} := b_{\kappa_i}t + \sqrt{\sigma^2 + \overline{\sigma}_{\kappa_i}^2}\, W_t + J_t^{2,\kappa_i}$ for $i \in \{1, 2\}$

L^2-distance between $\underline{\chi}_{-T}^{(\kappa)}$ and $\underline{\chi}_{-T}^{(\kappa')}$ tends to zero as $\kappa, \kappa' \to 0$. SBG-Alg below, based on the SB representation in (2.2), achieves this aim; it applies Algorithm 1 for the increments over the stick-breaking lengths that arise in (2.2) and Algorithm 2 for the "error term" over the time horizon $[0, L_n]$. By Theorem 7.10 below, the L^2-distance for the coupling given in SBG-Alg decays to zero, ensuring the feasibility of MLMC (see Theorem 7.17 for the computational complexity of MLMC).

4.1.1 Simulation of $(X_t^{(\kappa_1)}, X_t^{(\kappa_2)})$

A simulation algorithm for a coupling $(X_t^{(\kappa_1)}, X_t^{(\kappa_2)})$ of Gaussian approximations (at levels $1 \geq \kappa_1 > \kappa_2 > 0$) of X_t at an arbitrary time $t > 0$ is based on the following observation. The compound Poisson processes J^{2,κ_1} and J^{2,κ_2} in the Lévy–Itô decomposition in (2.4) can be simulated jointly, as the jumps of J^{2,κ_1} are precisely those of J^{2,κ_2} with modulus of at least κ_1. By choosing the same Brownian motion W in the representation (2.5) of $X_t^{(\kappa_1)}$ and $X_t^{(\kappa_2)}$, we obtain the coupling $(X_t^{(\kappa_1)}, X_t^{(\kappa_2)})$ with law $\Pi_t^{\kappa_1,\kappa_2}$ given in Algorithm 1.

Since $Z_t^{(\kappa_i)} \stackrel{d}{=} X_t^{(\kappa_i)}$, $i \in \{1, 2\}$, the definition in (3.10) and Proposition 7.11(a) below imply that the coupling $\Pi_t^{\kappa_1,\kappa_2}$ provides the bound

$$\mathcal{W}_2(X_t^{(\kappa_1)}, X_t^{(\kappa_2)}) \leq \mathbb{E}[(Z_t^{(\kappa_1)} - Z_t^{(\kappa_2)})^2]^{1/2} \leq \left(2t(\overline{\sigma}_{\kappa_1}^2 - \overline{\sigma}_{\kappa_2}^2)\right)^{1/2}. \qquad (4.1)$$

This bound is not optimal since the sum of the jumps $J_t^{2,\kappa_2} - J_t^{2,\kappa_1}$ of magnitude in the interval $(\kappa_2, \kappa_1]$ and the normal random variable W_t constructed in Algorithm 1, which appear in the difference $Z_t^{(\kappa_1)} - Z_t^{(\kappa_2)}$, are independent. The minimal transport coupling, with the L^2-distance equal to $\mathcal{W}_2(X_t^{(\kappa_1)}, X_t^{(\kappa_2)})$, is not accessible via simulation.

An important open problem in this context is to find an algorithm that samples $Z_t^{(\kappa_2)}$ as in Algorithm 1 and constructs

$$Z_t^{(\kappa_1)} = b_{\kappa_1}t + \sqrt{\sigma^2 + \overline{\sigma}_{\kappa_2}^2}\, W_t + \sqrt{\overline{\sigma}_{\kappa_1}^2 - \overline{\sigma}_{\kappa_2}^2}\, W_t' + J_t^{2,\kappa_2},$$

where W_t' is a normal variable with mean zero and variance t, independent of W_t and J_t^{2,κ_2}, but coupled with the difference $J_t^{2,\kappa_2} - J_t^{2,\kappa_1}$ in a way that reduces the second moment $\mathbb{E}[(Z_t^{(\kappa_1)} - Z_t^{(\kappa_2)})^2]$ asymptotically as $\kappa_1 \to 0$. Such an improvement in the bound in (4.1) on the L^2-Wasserstein distance would make the level variances

---

Algorithm 2 Simulation of the law $\underline{\Pi}_t^{\kappa_1,\kappa_2}$

Require: Cutoff levels $1 \geq \kappa_1 > \kappa_2 > 0$ and time horizon $t > 0$.

1: Compute b_{κ_i}, $\overline{\sigma}_{\kappa_i}^2$ and $\upsilon_{\kappa_i} := \sqrt{\sigma^2 + \overline{\sigma}_{\kappa_i}^2}$ for $i \in \{1, 2\}$ and $\overline{\nu}(\kappa_2)$, see (2.4), (2.5)

2: Sample $N_t \sim \text{Poi}(\overline{\nu}(\kappa_2)t)$ and $U_k \sim \text{U}(0, t)$ for $k \in \{1, \ldots, N_t + 1\}$

3: Set $s := \sum_{i=1}^{N_t+1} \log U_k$ and let $t_k := s^{-1} \sum_{i=1}^{k} \log U_i$ for $k \in \{0, \ldots, N_t + 1\}$

4: Set $(Z_0^{(\kappa_1)}, \underline{Z}_0^{(\kappa_1)}, \underline{\tau}_0^{(\kappa_1)}, Z_0^{(\kappa_2)}, \underline{Z}_0^{(\kappa_2)}, \underline{\tau}_0^{(\kappa_2)}) := (0, 0, 0, 0, 0, 0)$

5: **for** $k \in \{1, \ldots, N_t + 1\}$ **do**

6:    Sample $\lambda_k \sim \nu(\cdot \setminus (-\kappa_2, \kappa_2))/\overline{\nu}(\kappa_2)$ if $k \leq N_t$ and otherwise put $\lambda_k := 0$

7:    Independently sample $(\Delta_{k,i}^1, \Delta_{k,i}^2, \Delta_{k,i}^3) \sim \Phi_{\delta_k}(\upsilon_{\kappa_i}, b_{\kappa_i})$, $i \in \{1, 2\}$, where $\delta_k := t_k - t_{k-1}$

8:    **for** $i \in \{1, 2\}$ **do**

9:       **if** $\underline{Z}_{t_{k-1}}^{(\kappa_i)} > Z_{t_{k-1}}^{(\kappa_i)} + \Delta_{k,i}^2$ **then**

10:          Set $(Z_{t_k}^{(\kappa_i)}, \underline{Z}_{t_k}^{(\kappa_i)}, \underline{\tau}_{t_k}^{(\kappa_i)}) := (Z_{t_{k-1}}^{(\kappa_i)} + \Delta_{k,i}^1 + \lambda_k \mathbb{1}_{\{|\lambda_k| \geq \kappa_i\}}, Z_{t_{k-1}}^{(\kappa_i)} + \Delta_{k,i}^2, t_{k-1} + \Delta_{k,i}^3)$

11:       **else**

12:          Set $(Z_{t_k}^{(\kappa_i)}, \underline{Z}_{t_k}^{(\kappa_i)}, \underline{\tau}_{t_k}^{(\kappa_i)}) := (Z_{t_{k-1}}^{(\kappa_i)} + \Delta_{k,i}^1 + \lambda_k \mathbb{1}_{\{|\lambda_k| \geq \kappa_i\}}, \underline{Z}_{t_{k-1}}^{(\kappa_i)}, \underline{\tau}_{t_{k-1}}^{(\kappa_i)})$

13:       **end if**

14:    **end for**

15: **end for**

16: **return** $(\underline{\zeta}^{(\kappa_1)}, \underline{\zeta}^{(\kappa_2)})$, where $\underline{\zeta}^{(\kappa_i)} := (Z_t^{(\kappa_i)}, \underline{Z}_t^{(\kappa_i)}, \underline{\tau}_t^{(\kappa_i)})$ for $i \in \{1, 2\}$

---

in the MLMC estimator in (4.3), based on SBG-Alg, decay at a faster rate (because the increments in the SB representation in (2.2), and thus Algorithm 1, account for most of the output). Note, however, that sampling W_t' independently of $J_t^{2,\kappa_2} - J_t^{2,\kappa_1}$ would increase the second moment $\mathbb{E}[(Z_t^{(\kappa_1)} - Z_t^{(\kappa_2)})^2]$ compared to the output of Algorithm 1, which uses a single normal random variable in both coordinates.

Since the law $\text{Poi}(\overline{\nu}(\kappa_2)t)$ of the variable N_t in line 2 of Algorithm 1 is Poisson with mean $\overline{\nu}(\kappa_2)t$, the expected number of steps of Algorithm 1 is bounded by a constant multiple of $1 + \overline{\nu}(\kappa_2)t$, which is in turn bounded by a negative power of κ_2 by (2.7). Since the computational complexity of sampling the law of $X_t^{(\kappa_2)}$ is of the same order as that of the law $\Pi_t^{\kappa_1,\kappa_2}$, in the complexity analysis of SBG-Alg below, we may apply Algorithm 1 with $\Pi_t^{1,\kappa}$ to sample $X_t^{(\kappa)}$ for any $\kappa \in (0, 1]$.

4.1.2 Direct simulation of $(\underline{\chi}_t^{(\kappa_1)}, \underline{\chi}_t^{(\kappa_2)})$

Algorithm 2 samples from the law $\underline{\Pi}_t^{\kappa_1,\kappa_2}$ of a coupling $(\underline{\chi}_t^{(\kappa_1)}, \underline{\chi}_t^{(\kappa_2)})$ for levels $0 < \kappa_2 < \kappa_1 \leq 1$ and any $t > 0$. In particular, it requires the sampler from Devroye [23, Alg. MAXLOCATION] for the law $\Phi_t(v, \mu)$ of $(\hat{B}_t, \underline{\hat{B}}_t, \tau_t(\hat{B}))$, where the process $(\hat{B}_s)_{s \geq 0} = (vB_s + \mu s)_{s \geq 0}$ is a Brownian motion with drift $\mu \in \mathbb{R}$ and volatility $v > 0$.

Algorithm 2 samples the jump times and sizes of the compound Poisson process J^{2,κ_2} on the interval $(0, t)$ and prunes the jumps to get J^{2,κ_1}. Then it samples the increment, infimum and the time the infimum is attained for the Brownian motion with drift on each interval between the jumps of J^{2,κ_2}. The pair $(\underline{\zeta}^{(\kappa_1)}, \underline{\zeta}^{(\kappa_2)})$ clearly

satisfies $\zeta^{(\kappa_i)} \overset{d}{=} \underline{\chi}^{(\kappa_i)}$, $i \in \{1, 2\}$. Since [23, Alg. MAXLOCATION] samples the law $\Phi_t(v, \mu)$ with expected runtime uniformly bounded over the choice of parameters μ, v and t, the computational cost of sampling the pair of vectors $(\underline{\chi}_t^{(\kappa_1)}, \underline{\chi}_t^{(\kappa_2)})$ using Algorithm 2 is proportional to the cost of sampling $X_t^{(\kappa)}$ via Algorithm 1.

In principle, Algorithm 2 is an exact algorithm for the simulation of a coupling $(\underline{\chi}_t^{(\kappa_1)}, \underline{\chi}_t^{(\kappa_2)})$. However, as explained in Remark 4.1 below, it cannot be applied within an MLMC simulation scheme for a function of $\underline{\chi}_T^{(\kappa)}$ at a fixed time horizon T. SBG-Alg below circumvents this issue via the SB representation in (2.2), which also makes SBG-Alg *parallelisable* and thus much faster in practice even in the context of MC simulation (see the discussion after Corollary 7.14 below).

Remark 4.1 To the best of our knowledge, there is no simulation algorithm for the increment, the infima and the times the infima are attained of a Brownian motion under different drifts, i.e., of the vector

$$\left(B_t, \underline{B}_t^{(c_1)}, \underline{\tau}_t(B^{(c_1)}), \underline{B}_t^{(c_2)}, \underline{\tau}_t(B^{(c_2)})\right), \quad \text{where } c_1 \neq c_2 \text{ and } B_s^{(c)} = B_s + cs, \ s \geq 0.$$

Thus in line 7 of Algorithm 2, we are forced to take independent samples from $\Phi_{\delta_k}(v_{\kappa_1}, b_{\kappa_1})$ and $\Phi_{\delta_k}(v_{\kappa_2}, b_{\kappa_2})$ at each step k. In particular, the coupling of the marginals $X_t^{(\kappa_1)}$ and $X_t^{(\kappa_2)}$ of $\underline{\Pi}_t^{\kappa_1, \kappa_2}$ given in line 16 of Algorithm 2 amounts to taking two independent Brownian motions in the respective representations in (2.5) of $X_t^{(\kappa_1)}$ and $X_t^{(\kappa_2)}$. Thus unlike the coupling defined in Algorithm 1, here, by Proposition 7.11 (b), the squared L^2-norm satisfies $\mathbb{E}[(X_t^{(\kappa_1)} - X_t^{(\kappa_2)})^2] \geq 2t\sigma^2$ for all levels $1 \geq \kappa_1 > \kappa_2 > 0$, where σ^2 is the Gaussian component of X. Hence for a fixed time horizon, the coupling $\underline{\Pi}_t^{\kappa_1, \kappa_2}$ of $\underline{\chi}_t^{(\kappa_1)}$ and $\underline{\chi}_t^{(\kappa_2)}$ is not sufficiently strong for an MLMC scheme to be feasible if X has a Gaussian component, because the level variances do not decay to zero. However, by Proposition 7.11 (b), the L^2-distance between $\zeta^{(\kappa_1)}$ and $\zeta^{(\kappa_2)}$ constructed in Algorithm 2 does tend to zero with $t \to 0$. Thus SBG-Alg below, which applies Algorithm 2 over the time interval $[0, L_n]$ (recall $\mathbb{E}[L_n] = T/2^n$ from the SB representation (2.2)), circumvents this issue.

4.1.3 The SBG sampler

For a time horizon T, we can now define the coupling $\underline{\Pi}_{n,T}^{\kappa_1, \kappa_2}$ of the vectors $\underline{\chi}_T^{(\kappa_1)}$ and $\underline{\chi}_T^{(\kappa_2)}$ via the following algorithm.

By the SB representation (2.2), the law $\underline{\Pi}_{n,T}^{\kappa_1, \kappa_2}$ is indeed a coupling of the vectors $\underline{\chi}_T^{(\kappa_1)}$ and $\underline{\chi}_T^{(\kappa_2)}$ for any $n \in \mathbb{N} \cup \{0\}$. Note that if n equals zero, the set $\{1, \ldots, n\}$ in lines 1 and 2 of the algorithm is empty and the laws $\underline{\Pi}_{0,T}^{\kappa_1, \kappa_2}$ and $\underline{\Pi}_T^{\kappa_1, \kappa_2}$ coincide, implying that SBG-Alg may be viewed as a generalisation of Algorithm 2. The main advantage of SBG-Alg over Algorithm 2 is that it samples n increments of the Gaussian approximation over the interval $[L_n, T]$ using the fast Algorithm 1, with the "error term" contribution $\underline{\xi}_i$ being geometrically small in the number of sticks n.

The computational complexity of SBG-Alg and Algorithms 1 and 2 is simple to analyse. Assume throughout that all mathematical operations (addition, multiplication, exponentiation, etc.), as well as the evaluation of $\overline{v}(\kappa)$ and $\overline{\sigma}_\kappa^2$ for all $\kappa \in (0, 1]$

Algorithm 3 (SBG-Alg) Simulation of the coupling $(\underline{\chi}_T^{(\kappa_1)}, \underline{\chi}_T^{(\kappa_2)})$ with law $\underline{\Pi}_{n,T}^{\kappa_1,\kappa_2}$

Require: Cutoff levels $1 \geq \kappa_1 > \kappa_2 > 0$, number of sticks $n \in \mathbb{N} \cup \{0\}$ and $T > 0$.
1: Set $L_0 := T$, sample $U_k \sim U(0,1)$ and put $(\ell_k, L_k) := L_{k-1}(U_k, 1 - U_k)$ for $k \in \{1, \ldots, n\}$
2: Sample $(\xi_{k,1}, \xi_{k,2}) \sim \Pi_{\ell_k}^{\kappa_1,\kappa_2}$, $k \in \{1, \ldots, n\}$, and $(\underline{\xi}_1, \underline{\xi}_2) \sim \underline{\Pi}_{L_n}^{\kappa_1,\kappa_2}$
   $\rhd$ Algorithms 1 & 2
3: Put $\underline{\chi}_{n,T}^{(\kappa_i)} := \underline{\xi}_i + \sum_{k=1}^n (\xi_{k,i}, \min\{\xi_{k,i}, 0\}, \ell_k \mathbb{1}_{\{\xi_{k,i} \leq 0\}})$ for $i \in \{1, 2\}$
4: **return** $(\underline{\chi}_{n,T}^{(\kappa_1)}, \underline{\chi}_{n,T}^{(\kappa_2)})$

---

have constant computational cost. Moreover, assume that the simulation of any of the following random variables has constant expected cost: standard normal $\mathcal{N}(0,1)$, uniform $U(0,1)$, Poisson random variable (independently of its mean) and any jump with distribution $\nu|_{\mathbb{R}\setminus(-\kappa,\kappa)}/\overline{\nu}(\kappa)$ (independently of the cutoff level $\kappa \in (0,1]$). Recall that [23, Alg. MAXLOCATION] samples the law $\Phi_t(v, \mu)$ with uniformly bounded expected cost for all values of the parameters $\mu \in \mathbb{R}$, $v > 0$ and $t > 0$. The next statement follows directly from the algorithms.

Corollary 4.2 *Under the assumptions above, there exists a positive constant C_1 (resp. C_2; C_3), independent of $\kappa_1, \kappa_2 \in (0,1]$, $n \in \mathbb{N}$ and time horizon $t > 0$, such that the expected computational complexity of Algorithm 1 (resp. Algorithm 2; SBG-Alg) is bounded by $C_1(1 + \overline{\nu}(\kappa_2)t)$ (resp. $C_2(1 + \overline{\nu}(\kappa_2)t)$; $C_3(n + \overline{\nu}(\kappa_2)t)$).*

Up to a multiplicative constant, Algorithms 1 and 2 have the same expected computational complexity. However, Algorithm 2 requires not only the additional simulation of the jump times of $X^{(\kappa_2)}$ and a sample from $\Phi_t(v, \mu)$ using the sampler [23, Alg. MAXLOCATION] between any two consecutive jumps, but also a sequential computation of the output (the "for-loop" in lines 5–15) due to the condition in line 9 of Algorithm 2. This makes it hard to parallelise Algorithm 2. SBG-Alg avoids this issue by using the fast Algorithm 1 over the stick lengths in the SB representation (2.2) and calling Algorithm 2 only over the short time interval $[0, L_n]$, during which very few (if any) jumps of $X^{(\kappa_2)}$ occur. Moreover, SBG-Alg consists of several conditionally independent evaluations of Algorithm 1, which is parallelisable, leading to additional numerical benefits (see Sect. 6.2 below).

Remark 4.3 Line 2 of SBG-Alg contains the only call of Algorithm 2, which samples the coupling $(\underline{\xi}_1, \underline{\xi}_2)$ of the "error terms" in the SB representation (2.2) over the geometrically small time interval $[0, L_n]$. There are two natural modifications of SBG-Alg that avoid Algorithm 2 altogether, but retain the asymptotic properties (up to logarithmic factors) of the bias and level variances: (I) set $(\underline{\xi}_1, \underline{\xi}_2) = 0$ or (II) apply SBA introduced in González Cázares et al. [35] to approximate $\underline{\chi}_{L_m}^{(\kappa_1)}$ and $\underline{\chi}_{L_n}^{(\kappa_2)}$ as a function of the output of Algorithm 1 with $t = L_n$. Both of these choices would increase bias and level variance because unlike SBG-Alg, they are sampling from approximations to the laws of $\underline{\chi}_{L_n}^{(\kappa_1)}$ and $\underline{\chi}_{L_n}^{(\kappa_2)}$. This makes them slightly simpler to

implement, but theoretically less attractive. Moreover, in order to match the asymptotic properties of the bias under SBG-Alg, the number of sticks n in algorithms (I) and (II) would have to grow as a function of the decaying cutoff level κ.

4.2 MC and MLMC estimator based on SBG-Alg

This subsection gives an overview of the bounds on the computational complexity of the MC and MLMC estimators defined respectively in (4.2) and (4.3) below. Corollary 7.14 (for MC) and Theorem 7.17 (for MLMC) in Sect. 7.5 give the full analysis.

We suppose **throughout the subsection** that Assumption 3.2 holds with some $\delta \in (0, 2]$. As discussed in Sect. 3.1 above, we take δ as large as possible. In particular, if $\sigma \neq 0$ then $\delta = 2$. Let $q \in (0, 2]$ be as in (2.7) and thus $q \geq \delta$ if $\sigma = 0$. We take q as small as possible. For processes used in practice with $\sigma = 0$, we may typically take $\delta = q = \beta$, where β is the BG index defined in (2.6). Assumption 3.11, required for the analysis of the class BT_2 in (3.14) of discontinuous functions of $\underline{\tau}_T(X)$, holds with $\gamma = 1$ as Assumption 3.2 is satisfied (see the discussion following Proposition 3.12). When analysing the class of discontinuous functions BT_1 in (3.12), we suppose that Assumption 3.8 holds throughout with some $\gamma > 0$.

4.2.1 Monte Carlo

Pick $\kappa \in (0, 1]$ and let the sequence $\underline{\chi}_T^{\kappa,i}$, $i \in \mathbb{N}$, be i.i.d. (with the same distribution as $\underline{\chi}_T^{(\kappa)}$) simulated by SBG-Alg with $n \in \mathbb{N} \cup \{0\}$ sticks. The MC estimator based on $N \in \mathbb{N}$ independent samples is given by

$$\Upsilon_{\mathrm{MC}} := \frac{1}{N} \sum_{i=1}^{N} f(\underline{\chi}_T^{\kappa,i}). \tag{4.2}$$

The MC estimator is L^2-accurate at level $\epsilon > 0$ if its bias is smaller than $\epsilon/\sqrt{2}$ and the number N of independent samples is proportional to ϵ^{-2}; see Appendix A.1. Table 1 contains a summary of the values κ, as a function of ϵ, such that the bias of the estimator in (4.2) is at most $\epsilon/\sqrt{2}$, and of the associated Monte Carlo cost $\mathcal{C}_{\mathrm{MC}}(\epsilon)$ (up to a constant) for various classes of functions of $\underline{\chi}_T$ analysed in Sect. 3.2. Corollary 7.14 below contains the full details of the analysis.

The number $n \in \mathbb{N} \cup \{0\}$ of sticks in SBG-Alg affects neither the law of $\underline{\chi}_T^{(\kappa)}$ nor the asymptotic behaviour as $\epsilon \searrow 0$ of the computational complexity $\mathcal{C}_{\mathrm{MC}}(\epsilon)$. It only impacts the MC estimator in (4.2) through numerical stability and the reduction of the simulation cost by a constant factor. It is hard to determine the optimal choice for n. Clearly, the choice $n = 0$ (i.e., Algorithm 2) is not a good one as discussed in Sect. 4.1.3. A balance needs to be struck between (i) having a vanishingly small number of jumps in the time interval $[0, L_n]$, so that Algorithm 2 behaves in a numerically stable way, and (ii) not having too many sticks so that line 2 of SBG-Alg does not execute redundant computation of many geometrically small increments of $X^{(\kappa)}$, which are not detected in the final output. A good rule of thumb is $n = n_0 + \lceil \log_2(1 + \overline{\nu}(\kappa)T) \rceil$, where $\lceil x \rceil := \inf\{j \in \mathbb{Z} : j \geq x\}$, $x \in \mathbb{R}$, and the

Table 1 Asymptotic behaviour of the level κ and the complexity $\mathcal{C}_{\mathrm{MC}}(\epsilon)$ as $\epsilon \to 0$ for the MC estimator in (4.2)

Family of functions f	Case	κ	$\epsilon^2 \cdot \mathcal{C}_{\mathrm{MC}}(\epsilon)$
Lip in $(X_T, \underline{X}_T)$	$\sigma \neq 0$	$\epsilon^{1/(3-q)}\lvert\log\epsilon\rvert^{-1}$	$\epsilon^{-q/(3-q)}\lvert\log\epsilon\rvert^q$
locLip in $(X_T, \underline{X}_T)$	$\sigma \neq 0$	$\epsilon^{2/(4-q)}\lvert\log\epsilon\rvert^{-1/2}$	$\epsilon^{-2q/(4-q)}\lvert\log\epsilon\rvert^{q/2}$
Lip $\cup$ locLip in $(X_T, \underline{X}_T)$	$\sigma = 0$	$\epsilon\lvert\log\epsilon\rvert^{-1}$	$\epsilon^{-q}\lvert\log\epsilon\rvert^q$
BT$_1$ defined in (3.12)	$\sigma \neq 0$	$\max\left\{\frac{\epsilon^{3/(4-q)}}{\lvert\log\epsilon\rvert}, \frac{\epsilon^{2/(3-q)}}{\lvert\log\epsilon\rvert^{1/2}}\right\}$	$\min\left\{\frac{\lvert\log\epsilon\rvert^q}{\epsilon^{3q/(4-q)}}, \frac{\lvert\log\epsilon\rvert^{q/2}}{\epsilon^{2q/(3-q)}}\right\}$
	$\sigma = 0$	$\epsilon^{1/2+1/\gamma}\lvert\log\epsilon\rvert^{-1}$	$\epsilon^{-q(1/2+1/\gamma)}\lvert\log\epsilon\rvert^q$
Lip in $\underline{\tau}_T(X)$	$\sigma \neq 0$	$\epsilon^{1/(3-q)}$	$\epsilon^{-q/(3-q)}$
	$\delta \in (0,2)\setminus\{\frac{2}{3}\}$	$\epsilon^{\min\{2/\delta, \max\{3/2, 1/\delta\}\}}$	$\epsilon^{-q\min\{2/\delta, \max\{3/2, 1/\delta\}\}}$
	$\delta = \frac{2}{3}$	$\epsilon^{3/2}\lvert\log\epsilon\rvert^{-1}$	$\epsilon^{-3q/2}\lvert\log\epsilon\rvert^q$
BT$_2$ defined in (3.14)	$\sigma \neq 0$	$\epsilon^{2/(3-q)}$	$\epsilon^{-2q/(3-q)}$
	$\delta \in (0,2)\setminus\{\frac{2}{3}\}$	$\epsilon^{\min\{4/\delta, \max\{3, 2/\delta\}\}}$	$\epsilon^{-q\min\{4/\delta, \max\{3, 2/\delta\}\}}$
	$\delta = \frac{2}{3}$	$\epsilon^3\lvert\log\epsilon\rvert^{-1/2}$	$\epsilon^{-3q}\lvert\log\epsilon\rvert^{q/2}$

initial value n_0 is chosen so that some sticks are present if for large κ, the total expected number $\overline{\nu}(\kappa)T$ of jumps is small (e.g. $n_0 = 5$ works well in Sect. 6.2 for jump-diffusions with low activity; see Figs. 7 and 6), ensuring that the expected number of jumps in $[0, L_n]$ vanishes as $\epsilon \to 0$ (and hence $\kappa \to 0$). This choice keeps the complexity of the simulation nearly constant for varying values of κ when the increments of $X^{(\kappa)}$ can be sampled efficiently. Moreover, it typically satisfies the asymptotic assumptions of Theorem 7.17.

4.2.2 Multilevel Monte Carlo

The key ingredient of any MLMC estimator is the coupling between two consecutive levels of approximation that can be sampled efficiently. SBG-Alg is constructed with this in mind, returning a sample from a joint law for two different cutoff levels. (Note that the coupling constructed in SBG-Alg is different from the SBG coupling between $\underline{\chi}_T$ and its Gaussian approximation $\underline{\chi}_T^{(\kappa)}$, used in Sect. 3 to control the distances the two laws.) More precisely, the MLMC estimator in (4.3), based on the coupling in SBG-Alg, is given as follows. Let $(\kappa_j)_{j\in\mathbb{N}}$ (resp. $(n_j)_{j\in\mathbb{N}\cup\{0\}}$) be a decreasing (resp. increasing) sequence in $(0, 1]$ (resp. $\mathbb{N}$) satisfying $\lim_{j\to\infty}\kappa_j = 0$. Let $\underline{\chi}^{0,i} \overset{d}{=} \underline{\chi}_T^{(\kappa_1)}$ and $(\underline{\chi}_1^{j,i}, \underline{\chi}_2^{j,i}) \sim \underline{\Pi}_{n_j,T}^{\kappa_j, \kappa_{j+1}}$, $i, j \in \mathbb{N}$, be independent draws constructed by SBG-Alg. Recall that the sequence (n_j) appears as a parameter in the coupling $\underline{\Pi}_{n_j,T}^{\kappa_j, \kappa_{j+1}}$ (which is the law that the pair of vectors $(\underline{\chi}_1^{j,i}, \underline{\chi}_2^{j,i})$ follow). The number n_j specifies the number n of sticks used in SBG-Alg for the level j. Then for the parameters $m, N_0, \ldots, N_m \in \mathbb{N}$, the MLMC estimator takes the form

$$\Upsilon_{\mathrm{ML}} := \sum_{j=0}^{m}\frac{1}{N_j}\sum_{i=1}^{N_j}D_j^i, \quad \text{where} \quad D_j^i := \begin{cases} f(\underline{\chi}_2^{j,i}) - f(\underline{\chi}_1^{j,i}), & j \geq 1, \\ f(\underline{\chi}^{0,i}), & j = 0. \end{cases} \quad (4.3)$$

Table 2 The table presents the power of ϵ^{-1} in $\epsilon^2 \mathcal{C}_{\mathrm{ML}}(\epsilon)$ as $\epsilon \to 0$, neglecting only the logarithmic factors (see Theorem 7.17 below for the complete result). The parameter a in the table determines the decreasing sequence of cutoff levels $(\kappa_j)_{j \in \mathbb{N}}$ as follows: $\kappa_j = (1 + |a|/q)^{-2(j-1)/|a|}$ if $a \neq 0$, and $\kappa_j = \exp(-(2/q)(j-1))$ otherwise. The corresponding increasing number n_j of sticks in the definition of the law $\underline{\Pi}_{n_j,T}^{\kappa_j,\kappa_{j+1}}$ can be taken to grow asymptotically as $\log^2(1 + \overline{v}(\kappa_j)T)$ for large j; see Theorem 7.17 below

Family of functions f	Case	a	The power of ϵ^{-1} in $\epsilon^2 \mathcal{C}_{\mathrm{ML}}(\epsilon)$
Lip in $(X_T, \underline{X}_T)$	$\sigma \neq 0$	$2(q-1)$	$2(q-1)^+/(3-q)$
locLip in $(X_T, \underline{X}_T)$	$\sigma \neq 0$	$2(q-1)$	$4(q-1)^+/(4-q)$
Lip $\cup$ locLip in $(X_T, \underline{X}_T)$	$\sigma = 0$	$2(q-1)$	$2(q-1)^+$
BT$_1$ defined in (3.12)	$\sigma \neq 0$	$2(2q-1)/3$	$(2q-1)^+ \min\{2/(4-q), 4/(9-3q)\}$
	$\sigma = 0$	$2(q(1+\gamma) - \gamma)/(2+\gamma)$	$(q(1+1/\gamma)-1)^+$
Lip in $\underline{\tau}_T(X)$	$\sigma \neq 0$	$\frac{5}{4}q - \frac{1}{2}$	$(\frac{5}{4}q - \frac{1}{2})^+$
	$\sigma = 0$	$q - (1 - \frac{q}{2}) \min\{\frac{1}{2}, \frac{2\delta}{2-\delta}\}$	$\dfrac{(2q - (2-q) \min\{1, 4\delta/(2-\delta)\})^+}{\max\{\delta, \min\{4/3, 2\delta\}\}}$
BT$_2$ defined in (3.14)	$\sigma \neq 0$	$\frac{9}{8}q - \frac{1}{4}$	$(\frac{9}{4}q - \frac{1}{2})^+$
	$\sigma = 0$	$q - (1 - \frac{q}{2}) \min\{\frac{1}{4}, \frac{\delta}{2-\delta}\}$	$\dfrac{(2q - (2-q) \min\{1/2, 2\delta/(2-\delta)\})^+}{\max\{\delta/2, \min\{2/3, \delta\}\}}$

Given a coupling between consecutive levels, the integer parameters $m, N_0, \ldots, N_m$ in the estimator Υ_{ML} are chosen in a well-known optimal way; see Appendix A.2 for details. Table 2 summarises the resulting MLMC complexity up to logarithmic factors, with complete results available in Theorem 7.17 below.

There are two key ingredients in the proof of Theorem 7.17: (I) the bounds in Theorem 7.10 on the L^2-distance (i.e., the level variance, see Appendix A.2) between the functions of the marginals of the coupling $\underline{\Pi}_{n_j,T}^{\kappa_j,\kappa_{j+1}}$ constructed by SBG-Alg; (II) the bounds on the bias of various functions in Sect. 3. The number m of levels in the MLMC estimator in (4.3) is chosen to ensure that its bias, equal to the bias of $\underline{\chi}_T^{(\kappa_m)}$ at the top cutoff level κ_m, is bounded by $\epsilon/\sqrt{2}$. Thus the value of m can be expressed in terms of ϵ using Table 1 and the explicit formula for the cutoff κ_j, given in the caption of Table 2. The formula for κ_j at level j in the MLMC estimator in (4.3) is established in the proof of Theorem 7.17 by minimising the multiplicative constant in the computational complexity $\mathcal{C}_{\mathrm{ML}}(\epsilon)$ over all possible rates of the geometric decay of the sequence $(\kappa_j)_{j \in \mathbb{N}}$.

We stress that the analysis of the level variances for the various payoff functions of the coupling $\underline{\Pi}_{n_j,T}^{\kappa_j,\kappa_{j+1}}$ in Theorem 7.10 is carried out directly for locally Lipschitz payoffs; see Propositions 7.11. However, in the case of the discontinuous payoffs in BT$_1$ (see (3.12)) and BT$_2$ (see (3.14)), the analysis requires a certain regularity (uniformly in the cutoff levels) of the coupling $(\underline{\chi}_T^{(\kappa_j)}, \underline{\chi}_T^{(\kappa_{j+1})})$. This leads to a construction of a further coupling $(\underline{\chi}_T^{(\kappa_j)}, \underline{\chi}_T^{(\kappa_{j+1})}, \underline{\chi}_T)$ where the components of $(\underline{\chi}_T^{(\kappa_j)}, \underline{\chi}_T^{(\kappa_{j+1})})$ can be compared to the limiting object $\underline{\chi}_T$, which can be shown to possess the necessary regularity (see Proposition 7.13 below for details).

(A) X with Gaussian component ($\sigma \neq 0$)   (B) X without Gaussian component ($\sigma = 0$)
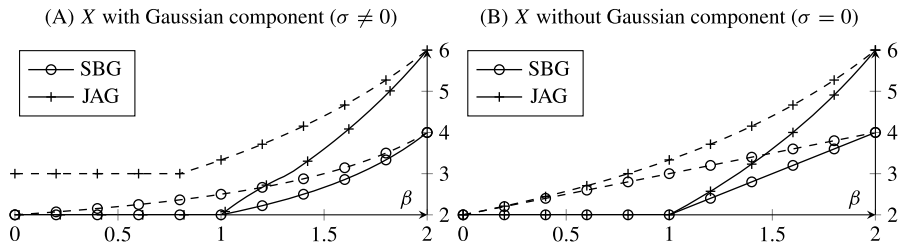
Fig. 3 Dashed (resp. solid) lines represent the power of ϵ^{-1} in the computational complexity of the MC (resp. MLMC) estimator for the expectation of a Lipschitz functional $f(X_T, \overline{X}_T)$, plotted as a function of the BG index β defined in (2.6). The SBG plots are based on Tables 1 and 2. The JAG plots are based on Dereich [21, Corollary 3.2] for the MC cost, and on [21, Corollary 1.2] if $\beta \geq 1$ (resp. Dereich and Heidenreich [22, Corollary 1] if $\beta < 1$) for the MLMC cost

5 Comparison with the literature

Approximations of the pair $(X_T, \overline{X}_T)$ abound. They include the random walk approximation, a Wiener–Hopf based approximation (Kuznetsov et al. [46], Ferreiro-Castilla et al. [25]), the jump-adapted Gaussian (JAG) approximation (Dereich and Heidenreich [22, 21]) and more recently, the SB approximation (González Cázares et al. [35]). The SB approximation converges the fastest as its bias decays geometrically in its computational cost. However, the JAG approximation is the only method known to us that does not require the ability to simulate the increments of the Lévy process X. Indeed, the JAG approximation simulates all jumps above a cutoff level, together with their jump times, and then samples the transitions of the Brownian motion from the Gaussian approximation on a random grid containing all the jump times. In contrast, in the present paper, we approximate the vector $\overline{\chi}_T = (X_T, \overline{X}_T, \overline{\tau}_T(X))$ with an *exact* sample from the law of the Gaussian approximation $\overline{\chi}_T^{(\kappa)} = (X_T^{(\kappa)}, \overline{X}_T^{(\kappa)}, \overline{\tau}_T(X^{(\kappa)}))$.

The JAG approximation has been analysed for Lipschitz payoffs applied to the pair $(X_T, \overline{X}_T)$ in Dereich and Heidenreich [22], Dereich [21]. The discontinuous and locally Lipschitz payoffs arising in applications and considered in this paper (see Fig. 1) have, to the best of our knowledge, not been analysed for the JAG approximation, nor have the payoffs involving the time $\overline{\tau}_T(X)$ at which the supremum is attained. Within the class of Lipschitz payoffs of $(X_T, \overline{X}_T)$, the computational complexities of the MC and MLMC estimators based on SBG-Alg are asymptotically smaller than those based on the JAG approximation; see Fig. 3. In fact, SBG-Alg applied to discontinuous payoffs outperforms the JAG approximation applied to Lipschitz payoffs by up to an order of magnitude in computational complexity; cf. Figs. 1(A), (B) and 3.

In order to understand where the differences in Fig. 3 come from, we summarise in Table 3 the bias and level variance for SBG-Alg and the JAG approximation as a function of the cutoff level $\kappa \in (0, 1]$ in the Gaussian approximation; cf. (2.5).

Table 3 shows that both bias and level variance decay at least as fast (and typically faster) for SBG-Alg than for the JAG approximation. The large improvement

Table 3 The rates (as $\kappa \to 0$) of decay of bias and level variance for Lipschitz payoffs of $(X_T, \overline{X}_T)$ under the JAG approximation are based on Dereich [21, Corollary 3.2] and Dereich and Heidenreich [22, Theorem 2], respectively. The rates on the bias and level variance for the SBG-Alg are given in Theorems 3.4 and 7.10

Gaussian component	Approximation	Bias	Level variance
With ($\sigma \neq 0$)	JAG	$\max\{\kappa^{1-\beta/4}, \kappa^{\beta/2}\}\lvert\log\kappa\rvert^{1/2}$	$\max\{\kappa^{2-\beta}, \kappa^{\beta}\lvert\log\kappa\rvert\}$
	SBG	$\kappa^{3-\beta}\lvert\log\kappa\rvert$	$\kappa^{2-\beta}$
Without ($\sigma = 0$)	JAG	$\max\{\kappa^{1-\beta/4}\lvert\log\kappa\rvert^{1/2}, \kappa^{\beta}\}$	$\max\{\kappa^{2-\beta}, \kappa^{2\beta}\}$
	SBG	$\kappa\lvert\log\kappa\rvert$	$\kappa^{2-\beta}$

in the computational complexity of the MC estimator in Fig. 3 is due to the faster decay of the bias under SBG-Alg. Put differently, the SBG coupling constructed in the present paper controls the Wasserstein distance much better than the KMT-based coupling in Dereich [21]. For a BG index $\beta > 1$, the improvement in the computational complexity of the MLMC estimator is mostly due to a faster bias decay. For $\beta < 1$, Fig. 3 (A) suggests that the computational complexity of the MLMC estimator under both algorithms is optimal. However, in this case, Table 3 and the equality in (A.3) imply that the MLMC estimator based on the JAG approximation has a computational complexity proportional to $\epsilon^{-2}\log^3(1/\epsilon)$, while that of SBG-Alg is proportional to ϵ^{-2}. This improvement is due solely to the faster decay of level variance under SBG-Alg. The numerical experiments in Sect. 6.1 suggest that our bounds for Lipschitz and locally Lipschitz functions are sharp; see the graphs (A) and (C) in Figs. 4 and 5.

To the best of our knowledge, there are no directly comparable results in the literature to either Theorem 3.4 or Proposition 3.12. Partial results in the direction of Theorem 3.4 are given in Dia [24], Mariucci and Reiß [50], Carpentier et al. [12]. Improvements in our Theorems 3.1 and 3.3 on the existing bounds on the distance between the marginals X_t and $X_t^{(\kappa)}$ have been discussed in detail in Sect. 3.1.1. The rate of the bound in [24, Theorem 2] on the Wasserstein distance between the suprema $\overline{X}_t$ and $\overline{X}_t^{(\kappa)}$ is worse than that implied by the bound in Corollary 3.5 on the Wasserstein distance between the joint laws of $(X_t, \overline{X}_t)$ and $(X_t^{(\kappa)}, \overline{X}_t^{(\kappa)})$. Proposition 3.7 bounds the bias of locally Lipschitz functions, generalising [24, Proposition 9] and providing a faster decay rate. Proposition 3.9 and Corollary 3.14 (a) cover a class of discontinuous payoffs, including the up-and-out digital option considered in [24, Proposition 10 (part 3)], and provide a faster rate of decay as $\kappa \to 0$ if either X has a Gaussian component or a BG index $\beta > 2/3$.

6 Numerical examples

In this section, we study numerically the performance of SBG-Alg. All the results are based on the code available in the repository [31]. In Sect. 6.1, we apply SBG-Alg to two families of Lévy models (tempered stable and Watanabe processes) and

verify numerically the decay of the bias (established in Sect. 3.2) and level variance (see Theorem 7.10 below) of the Gaussian approximations. In Sect. 6.2, we study numerically the cost reduction of SBG-Alg, when compared to Algorithm 2, for the simulation of the vector $\underline{\chi}_T^{(\kappa)}$. In Sect. 6.3, we numerically demonstrate the stability and effectiveness of SBG-Alg for the Monte Carlo estimation of Delta and Gamma of barrier options under Lévy models.

6.1 Numerical performance of SBG-Alg for tempered stable and Watanabe models

To illustrate numerically our results, we consider two classes of exponential Lévy models $S = S_0 e^X$. The first is the tempered stable class, containing the CGMY (or KoBoL) model, a process widely used for modelling risky assets in financial mathematics (see e.g. Cont and Tankov [20, Remarks 4.3 and 4.4] and the references therein), which satisfies the regularity assumptions from Sect. 3.2. The second is the Watanabe class, which has diffuse but singular transition laws (Sato [62, Theorem 27.19]), making it a good candidate to stress test our results.

We numerically study the decay of the bias and level variance of the MLMC estimator in (4.3) for the prices of a lookback put $\mathbb{E}[\overline{S}_T - S_T]$ and an up-and-out call $\mathbb{E}[(S_T - K)^+ \mathbb{1}_{\{\overline{S}_T \leq M\}}]$ as well as the values of the ulcer index (UI) given by $100 \mathbb{E}[(S_T/\overline{S}_T - 1)^2]^{1/2}$ (see Martin and McCann [51, Sect. 6.2]) and a modified ulcer index (MUI) defined by $100 \mathbb{E}[(S_T/\overline{S}_T - 1)^2 \mathbb{1}_{\{\overline{\tau}_T(S) < T/2\}}]^{1/2}$. The first three quantities are commonplace in applications; see Cont and Tankov [20, Sect. 11.3] and Martin and McCann [51, Sect. 6.2]. The MUI refines the UI by incorporating the information on the drawdown duration, and weights trends more heavily than short-time fluctuations.

In Sects. 6.1.1 and 6.1.2, we use $N = 10^5$ independent samples to estimate the means and variances of the variables D_j^1 in (4.3) (with $\underline{\chi}_T^{(\kappa_j)}$ substituted by $\overline{\chi}_T^{(\kappa_j)}$), where the choice of the parameters $n_j = \lceil \max\{j, \log^2(1 + \overline{\nu}(\kappa_{j+1}))\} \rceil$ and $\kappa_j = e^{-r(j-1)}$, $j \in \mathbb{N}$, is discussed in Sect. 7.5 below.

6.1.1 Tempered stable model

The characteristic triplet (σ^2, ν, b) of the tempered stable Lévy process X is given by $\sigma = 0$, drift $b \in \mathbb{R}$ and Lévy measure $\nu(dx) = |x|^{-1-\alpha_{\mathrm{sgn}(x)}} c_{\mathrm{sgn}(x)} e^{-\lambda_{\mathrm{sgn}(x)}|x|} dx$, where $\alpha_\pm \in [0, 2)$, $c_\pm \geq 0$ and $\lambda_\pm > 0$; cf. (2.3). Exact simulation of increments is currently out of reach if either $\alpha_+ > 1$ or $\alpha_- > 1$ (see e.g. Grabchak [36]) and requires the Gaussian approximation.

Figure 4 suggests that our bounds are close to the exhibited numerical behaviour for continuous payoff functions. In the discontinuous case, $\overline{\chi}_T^{(\kappa_j)}$ appears to be much closer to $\overline{\chi}_T$ (resp. $\overline{\chi}_T^{(\kappa_{j+1})}$), than predicted by Propositions 3.9 and 3.12 (resp. Theorem 7.10 (b) and (d)).
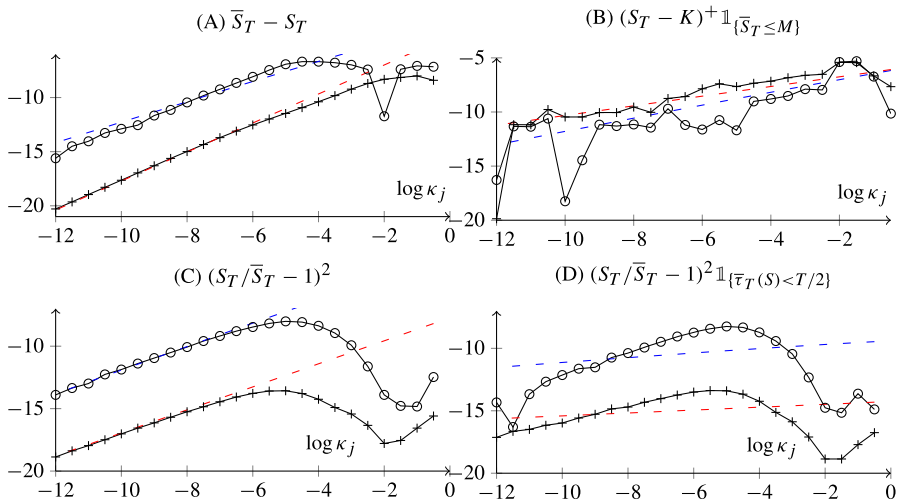
Fig. 4 Gaussian approximation of a tempered stable process: log–log plot of the bias and level variance for various payoffs as a function of $\log \kappa_j$. Circle (∘) and plus (+) correspond to $\log |\mathbb{E}[D_j^1]|$ and $\log \mathbb{V}[D_j^1]$, respectively, where D_j^1 is given in (4.3) with $\kappa_j = \exp(-r(j-1))$ for $r = 1/2$. The dashed lines in all the graphs plot the rates of the theoretical bounds in Sect. 3.2 (blue for the bias) and Theorem 7.10 (red for level variances). In plots (A)–(D), the initial value of the risky asset is normalised to $S_0 = 1$ and the time horizon is set to $T = 1/6$. In plot (B), we set $K = 1$ and $M = 1.2$. The model parameters are given in Table 4 below

Table 4 The parameters used for Fig. 4. The first set of parameters corresponds to the risk-neutral calibration to vanilla options on the USD/JPY exchange rate; see Andersen and Lipton [1, Table 3]. The second set is the maximum likelihood estimate based on the real-world S&P stock prices; see Kim et al. [40, Table 1]

Parameter set	b	α_+	α_-	c_+	c_-	λ_+	λ_-	Graphs in Fig. 4
1	0	0.66	0.66	0.1305	0.0615	6.5022	3.0888	(A) and (B)
2	0.1274	1.0781	1.0781	0.41077	0.41077	49.663	59.078	(C) and (D)

6.1.2 Watanabe model

The characteristic triplet (σ, ν, b) of the Watanabe process is given by $\sigma = 0$, the Lévy measure ν equals $\sum_{n \in \mathbb{N}} (c_+ \delta_{a^{-n}} + c_- \delta_{-a^{-n}})$, where $a \in \mathbb{N} \setminus \{1\}$ and δ_x is the Dirac measure at x, and the drift $b \in \mathbb{R}$ is arbitrary. The increments of the Watanabe process are diffuse, but have no density (see Sato [62, Theorem 27.19]). Since the process has very little jump activity, the bound in Proposition 3.12 (see also (3.6)) is non-vanishing and the bounds in Theorem 7.10 (c) and (d) are not applicable, meaning that we have no theoretical control on the approximation of $\overline{\tau}_T(S)$. This is not surprising as such acute lack of jump activity makes the Gaussian approximation unsuitable (cf. Asmussen and Rosiński [2, Proposition 2.2]).

The pictures in Fig. 5 (A) and (C) suggest that our bounds on the bias and level variance in Sect. 3.2 and Theorem 7.10 are robust for continuous payoff functions
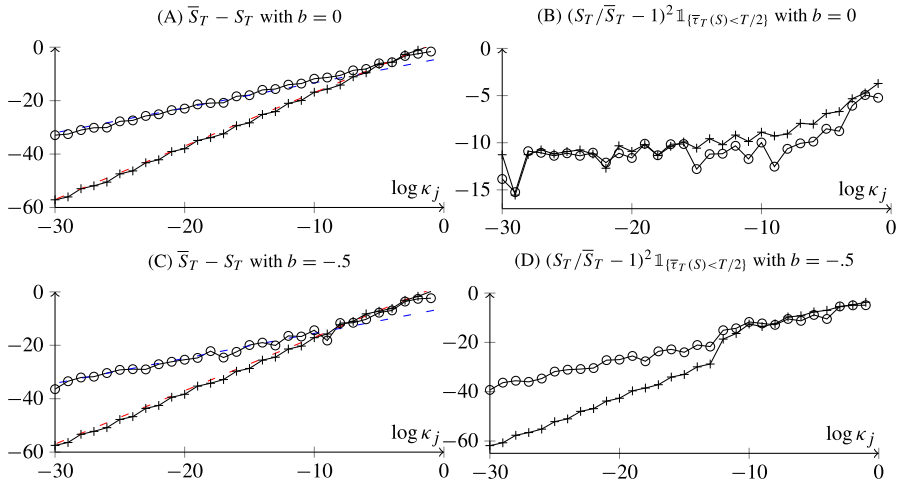
Fig. 5 Gaussian approximation of a Watanabe process: log–log plot of the bias and level variance for various payoffs as a function of $\log \kappa_j$. Circle (○) and plus (+) correspond to $\log |\mathbb{E}[D_j^1]|$ and $\log \mathbb{V}[D_j^1]$, respectively, where D_j^1 is given in (4.3) with $\kappa_j = \exp(-r(j-1))$ for $r = 1$. The dashed lines in graphs (A) and (C) plot the rates of the theoretical bounds in Sect. 3.2 (blue for the bias) and Theorem 7.10 (red for level variances). In plots (A)–(D), the initial value of the risky asset is normalised to $S_0 = 1$ and the time horizon is set to $T = 1$. The model parameters are given by $a = 2$, $c_+ = c_- = 1$

even if the underlying Lévy process has no transition densities. There are no dashed lines in Fig. 5 (B) and (D) as there are no results for discontinuous functions of $\overline{\tau}_T(S)$ in this case. In fact, Fig. 5 (B) suggests that the decay rate of the bias and level variance for functions of $\overline{\tau}_T(S)$ can be arbitrarily slow if the process does not have sufficient activity. Figure 5 (D), however, suggests that this decay is still fast if the underlying finite variation process X has a nonzero natural drift (see also Remark 3.6).

6.2 The cost reduction of SBG-Alg over Algorithm 2

Recall that Algorithm 2 and SBG-Alg both draw exact samples of a Gaussian approximation $\chi_{\overline{T}}^{(\kappa)}$. However, in practice, SBG-Alg may be many times faster than Algorithm 2; Fig. 6 plots the speedup factor in the case of a tempered stable process, defined in Sect. 6.1.1, as a function of κ. In conclusion, one should use SBG-Alg instead of Algorithm 2 for the MC estimator in (4.2). Even though the efficiency of SBG-Alg over Algorithm 2 for $\kappa > 0.1$ is not as drastic, one should still use SBG-Alg because Algorithm 2 is neither parallelisable nor suitable for MLMC, as discussed in Sect. 4.1.2.

If the Lévy process X is a jump-diffusion, i.e., $\nu(\mathbb{R} \setminus \{0\}) < \infty$, we may apply Algorithms 1 and 2 and SBG-Alg with $\kappa_1 = \kappa_2 = 0$. In that case, SBG-Alg still outperforms Algorithm 2 by a constant factor, with computational benefits being more pronounced when the total expected number of jumps $\lambda := \nu(\mathbb{R} \setminus \{0\})T$ is large; see Fig. 7. The cost reduction is most drastic when λ is large, but the improvement is already significant for $\lambda = 2$; see Fig. 7.
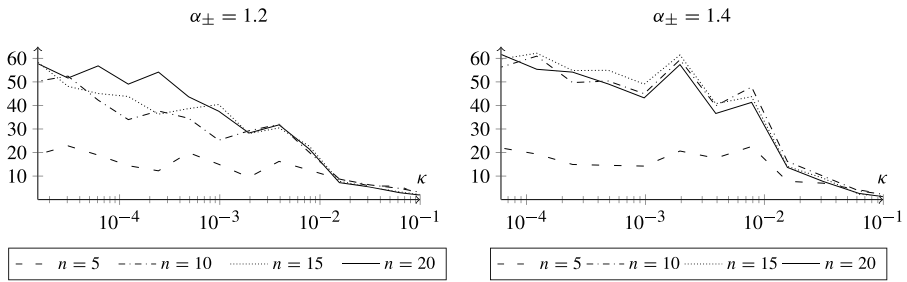
Fig. 6 The pictures show the ratio of the cost of Algorithm 2 over the cost of SBG-Alg (both in seconds) for the Gaussian approximations of a tempered stable process as a function of the cutoff level κ. The parameters used are $\lambda_\pm = 5$, $c_\pm = 2$. The number of sticks n in SBG-Alg varies between 5 and 20. The ratio for $n = 20$ is 57.8 (resp. 61.7) in the case $\alpha_\pm = 1.2$ (resp. $\alpha_\pm = 1.4$) for $\kappa = 2^{-16}$ (resp. $\kappa = 2^{-14}$)
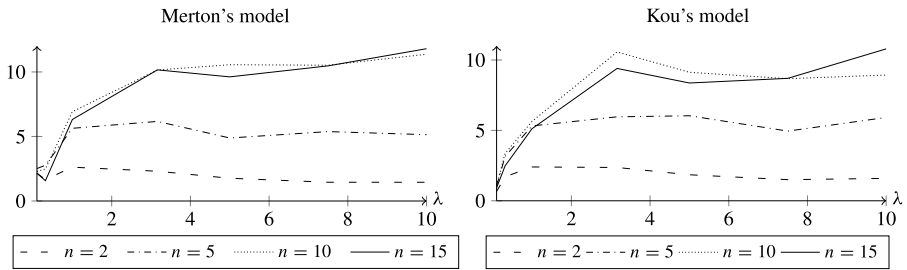

Fig. 7 The pictures show, for multiple number of sticks n, the ratio of the cost of Algorithm 2 over the cost of SBG-Alg (both in seconds) for jump-diffusions as a function of the mean number of jumps $\lambda = \nu(\mathbb{R} \setminus \{0\})T$. The ratio for $n = 15$ is 11.8 (resp. 10.8) in Merton's (resp. Kou's) model when $\lambda = 10$, see Merton [52] and Kou [44]

6.3 Estimating Greeks: Delta and Gamma for barrier options in Lévy models

A fundamental problem in mathematical finance is to compute the sensitivity of the price of a derivative security to the various underlying parameters in order to construct appropriate hedging strategies. These sensitivities are known as the *Greeks* and are in practice given by the partial derivatives of the option price $e^{-rT}\mathbb{E}[P]$ (where r is the discount rate over the time horizon T and P is a random payoff). The most common of the Greeks are Delta and Gamma, given by the first and second derivatives of the price $e^{-rT}\mathbb{E}[P]$ with respect to the spot S_0.

If the risk-neutral dynamics of the risky asset is described by an exponential Lévy model $S = S_0 e^X$, SBG-Alg provides a simple procedure for the Monte Carlo estimation of Delta and Gamma for any payoff $P = g(\overline{\chi}_T)$ (where we recall that the derivatives $f'(x)$ and $f''(x)$ of a function $f(x)$ are approximated by the quotients $(f(x + h) - f(x - h))/(2h)$ and $(f(x + h) - 2f(x) + f(x - h))/h^2$, respectively, for a small $h > 0$). This approach, widely used in practice, requires a Monte Carlo evaluation using SBG-Alg of $e^{-rT}\mathbb{E}[g(\overline{\chi}_T)]$ on a grid of current spot prices S_0, where the simulated stick-breaking sequence and the corresponding increments of $X^{(\kappa)}$ can be re-used over the grid points of S_0.

(A) Price (estimation)

(B) Delta (estimation)   (C) Gamma (estimation)
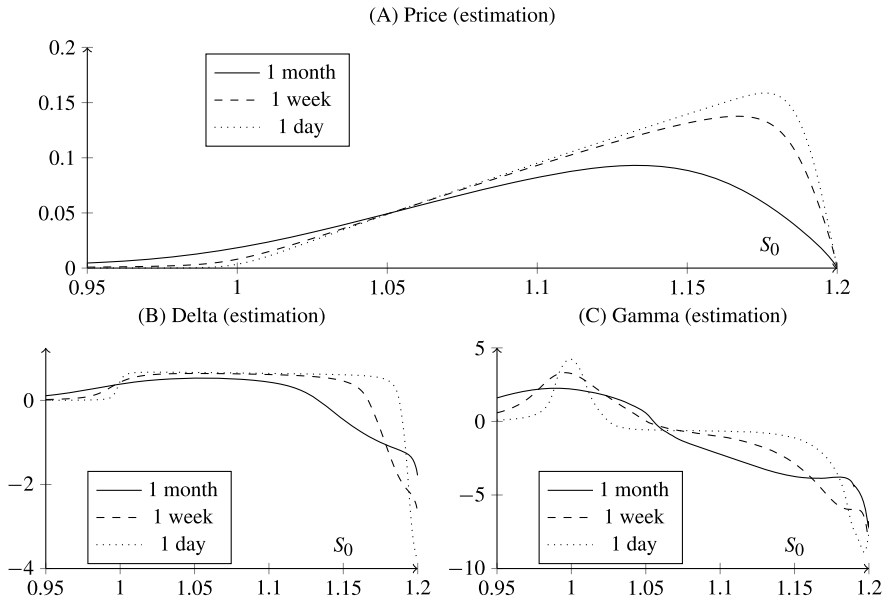
Fig. 8 Monte Carlo estimation of the price, Delta and Gamma for the up-and-out call option with payoff $g(\overline{\chi}_T) = (S_T - K)^+ \mathbb{1}_{\{\overline{S}_T \leq M\}}$ (see the first paragraph of Sect. 6.3.1 for parameter values) based on SBG-Alg. Solid lines (resp. dashed lines; dotted lines) correspond to the option maturity of one month (resp. one week; one day). Since the orders of magnitude (close to the barrier) of the Deltas and Gammas vary significantly with the maturity and all the Greeks in this example take both positive and negative values, we plot in graphs (B) and (C) F(Delta) and F(Gamma), where $F(x) := \text{sgn}(x)\log(1 + |x|)$, $x \in \mathbb{R}$, is a bijection on $\mathbb{R}$ preserving the sign of x but reducing its magnitude to logarithmic scale

6.3.1 Delta and Gamma for up-and-out call in a tempered stable (CGMY) model

We demonstrate the numerical stability of SBG-Alg in this context by computing Delta and Gamma for the up-and-out call payoff $g(\overline{\chi}_T) = (S_T - K)^+ \mathbb{1}_{\{\overline{S}_T \leq M\}}$ considered in Sect. 6.1.1; the model parameters, calibrated from USD/JPY foreign exchange options data, are given as parameter set 1 in Table 4, with strike $K = 1$, barrier level $M = 1.2$ and interest rate $r = 0.05$. To stress test the model, we increased the activity index by taking $\alpha_\pm = 1.16$ instead of 0.66. Figure 8 reports the results of the Monte Carlo estimation of the up-and-out call option price, together with its Delta and Gamma, with the spot ranging over the interval $S_0 \in [0.95, 1.2)$ for the maturities of 1 month, 1 week and 1 day. The grid spacing in the interval $[0.95, 1.2)$ was set at $h = 0.001$ with jump cutoff $\kappa = 0.001$. We used $n = 10$ sticks and $N = 10^7$ samples, which resulted in a total simulation time[1] of 12 minutes for all spot values S_0 on the grid and all maturities.

We note that the option price, Delta and Gamma in Fig. 8 look remarkably smooth as functions of the spot S_0, given that all the values in the three graphs are obtained using Monte Carlo simulation. We further note that the spot grid spacing $h = 10^{-3}$

---

[1] We used an HP Pavilion laptop 15-cw0xxx containing an AMD Ryzen 5 2500U with Radeon Vega Mobile and 12 GB of RAM.
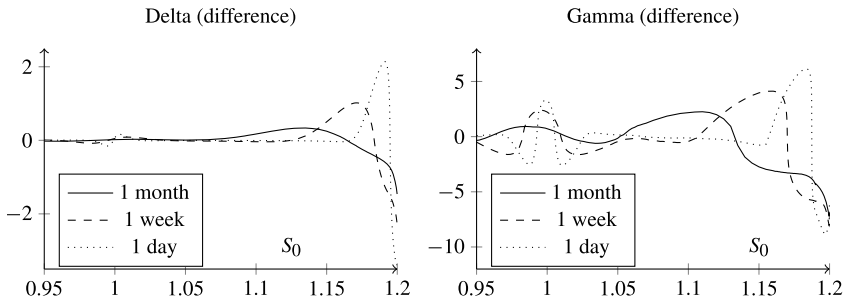
Fig. 9 Modulus of the difference between the estimates for Delta and Gamma for an up-and-out call option using a Monte Carlo algorithm based on SBG-Alg, where the difference reflects the improved accuracy of using the cutoff $\kappa = 0.001$ over the cutoff $\kappa = 0.1$. Away from the barrier, the numerical values for both cutoff levels are close to each other. As expected, the gap widens close to the barrier. Since the order of magnitude (close to the barrier) of the differences between Deltas and Gammas for the two cutoffs varies significantly, we plot in both graphs F(difference), where $F(x) := \mathrm{sgn}(x)\log(1 + |x|)$, $x \in \mathbb{R}$, is a bijection on $\mathbb{R}$ preserving the sign of x but reducing its magnitude to logarithmic scale. A positive value of F(difference) indicates that the estimate of the Greek based on the cutoff $\kappa = 0.001$ is larger than the one for $\kappa = 0.1$

should be small, but $h^2 N = 10$ should not be, which is consistent with the similar problem of density estimation via Monte Carlo (see e.g. Ben Abdellah et al. [6]). Having $h^2 N$ of moderate size ensures that the discontinuities of the function $S_0 \mapsto e^{-rT}(S_0 \exp(X_T) - K)^+ \mathbb{1}_{\{\overline{X}_T \le \log(M/S_0)\}}$ do not dominate in the estimation of the Greeks. For each realisation of X, this function has a single discontinuity at the random level $I = \exp(-\overline{X}_T)M$ with size of constant order. Every discontinuity of the Monte Carlo estimator will thus be of order $1/N$. The estimators of Delta and Gamma take a linear combination of such averages at different spot levels S_0 and divide by h and h^2, respectively, resulting in discontinuities (as a function of S_0) of orders $1/(hN)$ and $1/(h^2 N)$ for the Delta and Gamma, respectively. Hence h should be chosen to make $h^2 N$ of moderate size to keep these discontinuities from dominating in the estimation. The apparent smoothness of both Greeks further confirms our choice of h to be reasonable.

There is an inflection point (a switch from concavity to convexity) for our approximation of the Delta near the barrier, visible in Fig. 8 for the 1 day maturity; it corresponds to a local minimum of Gamma at $S_0 = 1.1971$. For the larger maturities, this inflection point is much closer to the barrier and is not visible in the graph, but can be observed in the figures coming out of the simulation. This inflection point close to the barrier was a persistent feature of our numerical scheme, observed after repeating the simulation with different random seeds a number of times. In contrast to the diffusion setting in Mijatović [53], we are not aware of any theoretical or empirical studies documenting such phenomena for barrier options under models with jumps. Put differently, while our Monte Carlo method suggests that the approximation of the Delta exhibits inflection near the barrier, it is neither clear that inflection persists in the limit as the spacing tends to zero (i.e., $h \to 0$), nor that it is a genuine feature of the actual Delta in the model.

In order to understand the dependence on the jump cutoff κ, we compare the numerical results in Fig. 8 (where $\kappa = 10^{-3}$) with an identical Monte Carlo algorithm

based on SBG-Alg, but with $\kappa = 0.1$. In Fig. 9, we plot the difference between Delta and Gamma produced by two different values of κ. There is little difference between the cutoff levels away from the barrier, meaning that the Monte Carlo estimator based on SBG-Alg for Delta and Gamma is robust in the value of the cutoff κ. This is surprising because we are using a naive (i.e., not based on Malliavin calculus) Monte Carlo estimator to compute derivatives of an expectation with respect to the starting point. The difference of the estimates for Delta and Gamma, corresponding to the different cutoffs, increases close to the barrier. In the regime when S_0 tends to the barrier level M, it is hard to give an intuitive explanation for the difference between the approximations of the first and second derivatives of the price.

6.3.2 Comparison with an analytically tractable pricing formula

We are not aware of an analytically tractable approximation for the price of an up-and-out call option studied in Sect. 6.3.1 under a Lévy model with jumps of infinite activity (i.e., with Lévy measure of infinite total mass), much less of its sensitivities. In order to test numerically the performance of SBG-Alg against a tractable formula, we compute Delta and Gamma of an up-and-out digital option with barrier M and payoff $g(\overline{\chi}_T) = \mathbb{1}_{\{\overline{S}_T \le M\}} = \mathbb{1}_{\{\overline{X}_T \le \log(M/S_0)\}}$ under an exponential Lévy model with analytically tractable Delta and Gamma. Let X be an α-stable process of infinite variation without positive jumps. The strong Markov property of X at the first crossing time $\tau_x := \inf\{t \ge 0 : X_t > x\}$ yields

$$\mathbb{P}[X_T > x] = \int_{[0,T)} \mathbb{P}[X_{T-t} > 0] \, \mathbb{P}[\tau_x \in dt] = \mathbb{P}[X_T > 0] \, \mathbb{P}[\overline{X}_T > x]$$

for any $x > 0$, since the scaling property of X implies $\mathbb{P}[X_t > 0] = \mathbb{P}[X_T > 0]$ for all $t > 0$, and the absence of positive jumps yields $X_{\tau_x} = x$ a.s. for any $x > 0$ and $\{\tau_x < T\} = \{\overline{X}_T > x\}$ a.s. In turn, the formula above implies that $\overline{X}_T$ has the same law as X_T conditioned to be positive. The option price is thus given as function of the spot by

$$S_0 \mapsto e^{-rT} \mathbb{E}[g(\overline{\chi}_T)] = e^{-rT} \mathbb{P}[X_T \le \log(M/S_0)] / \mathbb{P}[X_T > 0]. \tag{6.1}$$

The corresponding Delta and Gamma, equal to the first and second derivatives of this function, can be expressed as a geometrically converging power series (see Uchaikin and Zolotarev [66, Chap. 4]) and computed numerically without the use of Monte Carlo.

The parameters for this test were chosen as follows: X has unit scale (in Zolotarev's (C) parametrisation; see Uchaikin and Zolotarev [66, Sect. 3.6]) and the stability parameter $\alpha = 1.7$, while the market data is $r = 0.05$, $T = 1/12$, $M = 1$ and $S_0 \in [0.85, 1)$. To highlight the importance of the cutoff level κ, we compare $\kappa = 0.1$ with $\kappa = 0.001$. We used $n = 25$ sticks, $N = 10^7$ samples and grid size $h = 0.005$. This resulted in a total simulation time of 12 minutes. The estimation of Delta and Gamma is accurate and numerically stable.

Surprisingly, the error in Delta – see graph (B) in Fig. 10 – remains bounded all the way to the barrier M for both values of κ. The error for $\kappa = 0.001$ is smaller than the one for $\kappa = 0.1$, but comes at a significant increase in cost due to the Blumenthal–Getoor index of X being equal to 1.7 (and in particular much greater than one).
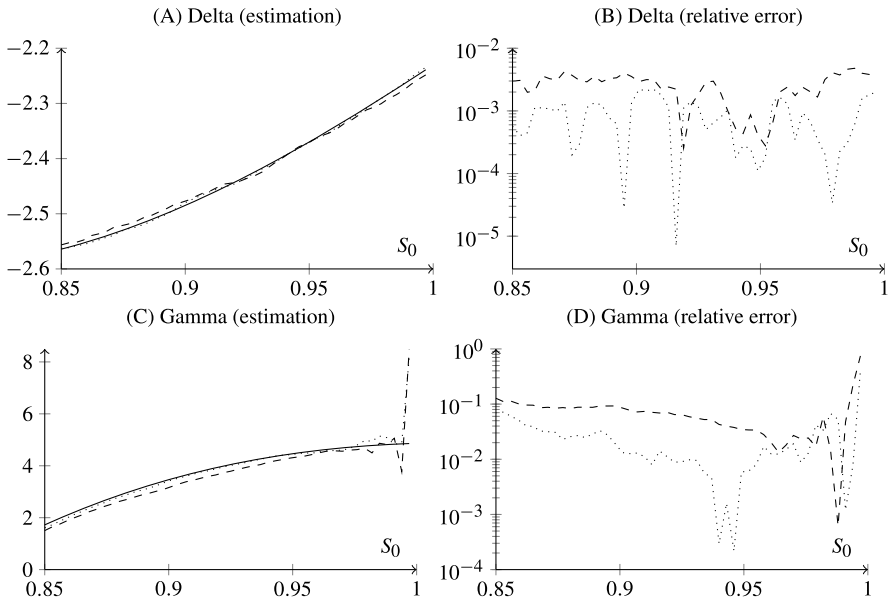
Fig. 10 Monte Carlo estimation of Delta and Gamma for the up-and-out digital option with payoff $g(\overline{X}_T) = \mathbb{1}_{\{\overline{S}_T \leq M\}}$ under an α-stable model with no positive jumps; see the second paragraph of Sect. 6.3.2 for the values of the model and market parameters. Dashed lines (resp. dotted lines) correspond to the output based on Monte Carlo estimation using SBG-Alg for $\kappa = 0.1$ (resp. $\kappa = 0.001$). Figures (A) and (C) contain a solid line corresponding to the true values of Delta and Gamma, computed using a geometrically convergent power series in Uchaikin and Zolotarev [66, Chap. 4] for the value function in (6.1)

7 Proofs

In the remainder of the paper, we use the notation $\underline{\tau}_t := \underline{\tau}_t(X)$, $\underline{\tau}_t^{(\kappa)} := \underline{\tau}_t(X^{(\kappa)})$ for all $t > 0$.

7.1 Proof of Theorems 3.1 and 3.3

In this subsection, we establish bounds on the Wasserstein and Kolmogorov distances between the increment X_t and its Gaussian approximation $X_t^{(\kappa)}$ in (2.5).

Proof of Theorem 3.1 Recall the Lévy–Itô decomposition of X at level κ in (2.4) and the martingale $M^{(\kappa)} = \sigma B + J^{1,\kappa}$. Set $Z := X - M^{(\kappa)}$ and note that we have $X^{(\kappa)} = Z + \sqrt{\overline{\sigma}_\kappa^2 + \sigma^2} \, W$, where W is a standard Brownian motion in (2.5), independent of Z. Hence any coupling $(W_t, M_t^{(\kappa)})$ yields a coupling of $(X_t, X_t^{(\kappa)})$ satisfying

$$\mathbb{E}[|X_t - X_t^{(\kappa)}|^p] = \mathbb{E}\left[\left|M_t^{(\kappa)} - \sqrt{\overline{\sigma}_\kappa^2 + \sigma^2} \, W_t\right|^p\right].$$

Setting $W := B$, which amounts to the independence coupling $(W, J^{1,\kappa})$, and applying Jensen's inequality for $p \in [1, 2]$ yields

$$\mathbb{E}[|X_t - X_t^{(\kappa)}|^p]^{2/p} \leq \mathbb{E}\left[\left|J_t^{1,\kappa} - \left(\sqrt{\overline{\sigma}_\kappa^2 + \sigma^2} - \sigma\right)W_t\right|^2\right] \leq 2t\overline{\sigma}_\kappa^2.$$

For any $m \in \mathbb{N}$, we have $M_t^{(\kappa)} \stackrel{d}{=} \sum_{i=1}^m \xi_i$, where $\xi_1, \ldots, \xi_m$ are i.i.d. with $\xi_1 \stackrel{d}{=} M_{t/m}^{(\kappa)}$. Hence Petrov [57, Theorem 16] and Rio [60, Theorem 4.1] imply the existence of universal constants K_p, $p \in [1, 2]$, with $K_1 = 1/2$, satisfying

$$\mathcal{W}_p^p\left(M_t^{(\kappa)}, \sqrt{\overline{\sigma}_\kappa^2 + \sigma^2}\,W_t\right) \leq K_p^p \frac{(t(\overline{\sigma}_\kappa^2 + \sigma^2))^{p/2}\mathbb{E}[|\xi_1|^{p+2}]}{m^{p/2}\mathbb{E}[\xi_1^2]^{(p+2)/2}}$$

$$= K_p^p \frac{(m/t)\mathbb{E}[|M_{t/m}^{(\kappa)}|^{p+2}]}{\overline{\sigma}_\kappa^2 + \sigma^2}$$

for all $m \in \mathbb{N}$. According to Figueroa-López [26, Theorem 1.1], the limit as $m \to \infty$ of the right-hand side of the display above equals

$$K_p^p \int_{(-\kappa,\kappa)} |x|^{p+2}\nu(dx)/(\overline{\sigma}_\kappa^2 + \sigma^2) \leq K_p^p \kappa^p \varphi_\kappa^2,$$

implying the claim in the theorem. $\square$

Proof of Theorem 3.3 (a) Set $d_\kappa := \sup_{x \in \mathbb{R}} |\mathbb{P}[M_t^{(\kappa)} \leq x] - \mathbb{P}[\sqrt{\overline{\sigma}_\kappa^2 + \sigma^2}\,W_t \leq x]|$ and note that

$$\left|\mathbb{P}[X_t \leq x] - \mathbb{P}[X_t^{(\kappa)} \leq x]\right|$$

$$= \left|\mathbb{E}\left[\mathbb{P}[M_t^{(\kappa)} \leq x - Z_t|Z_t] - \mathbb{P}[\sqrt{\overline{\sigma}_\kappa^2 + \sigma^2}\,W_t \leq x - Z_t|Z_t]\right]\right| \leq d_\kappa,$$

where the processes Z and $M^{(\kappa)}$ are as in the proof of Theorem 3.1. Since $M^{(\kappa)}$ is a Lévy process, for any $m \in \mathbb{N}$, we have $M_t^{(\kappa)} \stackrel{d}{=} \sum_{i=1}^m \xi_i$, where $\xi_1, \ldots, \xi_m$ are i.i.d. with $\xi_1 \stackrel{d}{=} M_{t/m}^{(\kappa)}$. By the Berry–Esseen inequality, see Korolev and Shevtsova [42, Theorem 1], there exists a constant $C_{\mathrm{BE}} \in (0, \frac{1}{2})$ such that

$$d_\kappa \leq \frac{C_{\mathrm{BE}}\mathbb{E}[|\xi_1|^3]}{\sqrt{m}\mathbb{E}[\xi_1^2]^{3/2}} = \frac{C_{\mathrm{BE}}t}{m\sqrt{m}} \frac{(m/t)\mathbb{E}[|M_{t/m}^{(\kappa)}|^3]}{(t/m)^{3/2}(\overline{\sigma}_\kappa^2 + \sigma^2)^{3/2}} = C_{\mathrm{BE}} \frac{(m/t)\mathbb{E}[|M_{t/m}^{(\kappa)}|^3]}{\sqrt{t}(\overline{\sigma}_\kappa^2 + \sigma^2)^{3/2}}$$

for all $m \in \mathbb{N}$. Taking $m \to \infty$ yields (a) since according to Figueroa-López [26, Theorem 1.1], the limit of the right-hand side of the display above equals

$$C_{\mathrm{BE}} \int_{(-\kappa,\kappa)} |x|^3\nu(dx)/\left(\sqrt{t}(\overline{\sigma}_\kappa^2 + \sigma^2)^{3/2}\right) \leq C_{\mathrm{BE}}(\kappa/\overline{\sigma}_\kappa)\varphi_\kappa^3/\sqrt{t}.$$

(b) By Picard [58, Theorem 3.1(a)], X_t has a smooth density f_t, and given $T > 0$, the constant $C' = \sup_{(t,x) \in (0,T] \times \mathbb{R}} t^{1/\delta} f_t(x)$ is finite. Applying (3.1) and (7.10) in Lemma 7.5 with $p = 2$ gives (3.3). $\square$

7.2 Proof of Theorem 3.4

We recall an elementary result for stick-breaking processes.

Lemma 7.1 *Let $(\varpi_n)_{n \in \mathbb{N}}$ be a stick-breaking process on $[0, 1]$ based on the uniform law. For any measurable function $\phi \geq 0$, we have*

$$\sum_{n \in \mathbb{N}} \mathbb{E}[\phi(\varpi_n)] = \int_0^1 \frac{\phi(x)}{x} dx.$$

In particular, for any $a_1, a_2 > 0$ and $b_1 < b_2$ with $b_2 > 0$, we have

$$\sum_{n \in \mathbb{N}} \mathbb{E}[\min\{a_1 \varpi_n^{b_1}, a_2 \varpi_n^{b_2}\}]$$

$$= \begin{cases} \frac{a_2}{b_2} \min\{1, \frac{a_1}{a_2}\}^{b_2/(b_2 - b_1)} + \frac{a_1}{b_1}(1 - \min\{1, \frac{a_1}{a_2}\}^{b_1/(b_2 - b_1)}), & b_1 \neq 0, \\ b_2^{-1} \min\{a_2, a_1\}(1 + \log^+ \frac{a_2}{a_1}), & b_1 = 0. \end{cases}$$

Proof The law of $-\log \varpi_n$ is gamma with shape n and scale 1. Applying Fubini's theorem implies

$$\sum_{n \in \mathbb{N}} \mathbb{E}[\phi(\varpi_n)] = \sum_{n \in \mathbb{N}} \int_0^\infty \frac{x^{n-1}}{(n-1)!} e^{-x} \phi(e^{-x}) dx = \int_0^\infty \phi(e^{-x}) dx = \int_0^1 \frac{\phi(x)}{x} dx.$$

The formula for $\phi(x) := \min\{a_1 x^{b_1}, a_2 x^{b_2}\}$ follows by a direct calculation. □

The L^p-Wasserstein distance, defined in (3.10), satisfies

$$\mathcal{W}_p^p(\xi, \xi_*) = \int_0^1 |F^{-1}(u) - F_*^{-1}(u)|^p du,$$

where F^{-1} (resp. F_*^{-1}) is the right inverse of the distribution function F (resp. F_*) of the real-valued random variable ξ (resp. ξ_*) (see Bobkov and Ledoux [10, Theorem 2.10]). Thus the *comonotonic* (or *minimal transport*) coupling defined by

$$(\xi, \xi_*) := (F^{-1}(U), F_*^{-1}(U)) \qquad \text{for some } U \sim \text{U}(0, 1) \tag{7.1}$$

attains the infimum in definition (3.10).

Lemma 7.2 *If ξ and ξ_* are real-valued and comonotonically coupled, then*

$$\mathbb{E}[|\mathbb{1}_{\{\xi \leq x\}} - \mathbb{1}_{\{\xi_* \leq x\}}|] = |\mathbb{E}[\mathbb{1}_{\{\xi \leq x\}} - \mathbb{1}_{\{\xi_* \leq x\}}]| \qquad \text{for any } x \in \mathbb{R}.$$

Proof Suppose $(\xi, \xi_*) = (F^{-1}(U), F_*^{-1}(U))$ for some $U \sim \text{U}(0, 1)$, where F and F_* are the distribution functions of ξ and ξ_*. Suppose $y := F(x) \leq F_*(x) =: y_*$. Since F^{-1} and F_*^{-1} are monotonic functions, it follows that $\mathbb{1}_{\{\xi \leq x\}} - \mathbb{1}_{\{\xi_* \leq x\}} \leq 0$ a.s. since this difference equals 0 or -1 according to whether $U \in (0, 1) \setminus (y, y_*]$ or $U \in (y, y_*]$, respectively. If $y \geq y_*$, we have $\mathbb{1}_{\{\xi \leq x\}} - \mathbb{1}_{\{\xi_* \leq x\}} \geq 0$ a.s. In either case, the result follows. □

For any $t > 0$, let G_t^κ denote the joint law of the comonotonic coupling of X_t and $X_t^{(\kappa)}$ defined in (7.1). Note that a coupling $(X_t, X_t^{(\kappa)})$ with law G_t^κ satisfies the inequality in Theorem 3.1. The following lemma is crucial in the proof of Theorem 3.4.

Lemma 7.3 *Let $\ell = (\ell_n)_{n \in \mathbb{N}}$ be a stick-breaking process on $[0, t]$ and $(\xi_n, \xi_n^{(\kappa)})$, $n \in \mathbb{N}$, a sequence of random vectors that conditionally on ℓ are independent and satisfy $(\xi_n, \xi_n^{(\kappa)}) \sim G_{\ell_n}^\kappa$ for all $n \in \mathbb{N}$. Then for any $p \in [1, 2]$ and $x \in \mathbb{R}$, we have*

$$\mathbb{E}\left[\left(\sum_{n=1}^{\infty} |\xi_n - \xi_n^{(\kappa)}|\right)^p\right]^{\frac{1}{p}} \leq \mu_p(\kappa, t), \tag{7.2}$$

$$\mathbb{E}\left[\sum_{n=1}^{\infty} \ell_n |\mathbb{1}_{\{\xi_n \leq x\}} - \mathbb{1}_{\{\xi_n^{(\kappa)} \leq x\}}|\right] \leq \mu_0^\tau(\kappa, t), \tag{7.3}$$

where μ_p and μ_0^τ are defined in (3.5) and (3.6), respectively. Moreover, if Assumption 3.2 holds, then for every $T > 0$, there exists a constant $C > 0$, dependent only on (T, δ, σ, ν), such that for all $t \in [0, T]$, $\kappa \in (0, 1]$ and $x \in \mathbb{R}$, we have

$$\mathbb{E}\left[\sum_{n=1}^{\infty} \ell_n |\mathbb{1}_{\{\xi_n \leq x\}} - \mathbb{1}_{\{\xi_n^{(\kappa)} \leq x\}}|\right] \leq \mu_\delta^\tau(\kappa, t), \tag{7.4}$$

where μ_δ^τ is defined in (3.8).

Proof Note that $\mu_p(\kappa, t) = \mu_2(\kappa, t)$ for all $p \in (1, 2]$. Hence by Jensen's inequality, in (7.3), we need only consider $p \in \{1, 2\}$. Pick $n \in \mathbb{N}$ and set $\kappa_p := K_p^p \kappa^p \varphi_\kappa^2$, $p \in \{1, 2\}$, where K_p and φ_κ are as in the statement of Theorem 3.1. Condition on ℓ_n and apply the bound in (3.1) to obtain

$$\mathbb{E}\left[|\xi_n^{(\kappa)} - \xi_n|^p \big| \ell_n\right] \leq \min\{2^{p/2} \overline{\sigma}_\kappa^p \ell_n^{p/2}, \kappa_p\}, \qquad p \in \{1, 2\}. \tag{7.5}$$

Applying (7.5) and Lemma 7.1 yields (7.2) for $p = 1$; indeed,

$$\sum_{n=1}^{\infty} \mathbb{E}[|\xi_n - \xi_n^{(\kappa)}|] \leq \sum_{n=1}^{\infty} \mathbb{E}[\min\{\sqrt{2\ell_n}\overline{\sigma}_\kappa, \kappa_1\}]$$

$$= 2\min\{\sqrt{2t}\overline{\sigma}_\kappa, \kappa_1\}\left(1 + \log^+(\sqrt{2t}\overline{\sigma}_\kappa/\kappa_1)\right).$$

Now consider the case $p = 2$. A simple expansion yields

$$\mathbb{E}\left[\left(\sum_{n=1}^{\infty} |\xi_n - \xi_n^{(\kappa)}|\right)^2\right] = \sum_{n=1}^{\infty} \mathbb{E}[(\xi_n - \xi_n^{(\kappa)})^2] + 2\sum_{n=1}^{\infty} \sum_{m=n+1}^{\infty} \mathbb{E}[|\xi_n - \xi_n^{(\kappa)}||\xi_m - \xi_m^{(\kappa)}|].$$

We proceed to bound both sums. Inequality (7.5) for $p = 2$ and Lemma 7.1 imply

$$\sum_{n=1}^{\infty} \mathbb{E}[(\xi_n - \xi_n^{(\kappa)})^2] \leq \sum_{n=1}^{\infty} \mathbb{E}[\min\{2\overline{\sigma}_\kappa^2 \ell_n, \kappa_2\}]$$

$$= \min\{2t\overline{\sigma}_\kappa^2, \kappa_2\}\big(1 + 2\log^+(\sqrt{2t}\overline{\sigma}_\kappa/\sqrt{\kappa_2})\big).$$

Define the σ-algebra $\mathcal{F}_n := \sigma(\ell_1, \ldots, \ell_n)$ and use the conditional independence to obtain

$$\mathbb{E}\big[|\xi_n - \xi_n^{(\kappa)}||\xi_m - \xi_m^{(\kappa)}|\big|\mathcal{F}_m\big] \leq \min\{\sqrt{2\ell_n}\overline{\sigma}_\kappa, \kappa_1\}\min\{\sqrt{2\ell_m}\overline{\sigma}_\kappa, \kappa_1\}, \qquad n < m.$$

Note that $(\ell_m/L_n)_{m=n+1}^{\infty}$ is a stick-breaking process on $[0, 1]$ independent of $\mathcal{F}_n$. Use the tower property and apply (3.1) and Lemma 7.1 to get

$$\sum_{m=n+1}^{\infty} \mathbb{E}\big[|\xi_n - \xi_n^{(\kappa)}||\xi_m - \xi_m^{(\kappa)}|\big|\mathcal{F}_n\big]$$

$$\leq \min\{\sqrt{2\ell_n}\overline{\sigma}_\kappa, \kappa_1\} \sum_{m=n+1}^{\infty} \mathbb{E}[\min\{\sqrt{2\ell_m}\overline{\sigma}_\kappa, \kappa_1\}|\mathcal{F}_n]$$

$$= 2\min\{\sqrt{2\ell_n}\overline{\sigma}_\kappa, \kappa_1\}\min\{\sqrt{2L_n}\overline{\sigma}_\kappa, \kappa_1\}\left(1 + \log^+ \frac{\sqrt{2L_n}\overline{\sigma}_\kappa}{\kappa_1}\right)$$

$$\leq 2\min\{2L_{n-1}\overline{\sigma}_\kappa^2, \kappa_1^2\}\big(1 + \log^+(\sqrt{2t}\,\overline{\sigma}_\kappa/\kappa_1)\big),$$

where $\max\{L_n, \ell_n\} \leq L_{n-1} \leq t$ is used in the last step. Since $\ell_n \overset{d}{=} L_n$, $n \in \mathbb{N}$, Lemma 7.1 yields

$$2\sum_{n=1}^{\infty}\sum_{m=n+1}^{\infty} \mathbb{E}[|\xi_n - \xi_n^{(\kappa)}||\xi_m - \xi_m^{(\kappa)}|]$$

$$\leq 4\sum_{n=1}^{\infty} \mathbb{E}[\min\{2L_{n-1}\overline{\sigma}_\kappa^2, \kappa_1^2\}]\big(1 + \log^+(\sqrt{2t}\overline{\sigma}_\kappa/\kappa_1)\big) = 2\mu_1(\kappa, t)^2.$$

Putting everything together yields (7.2) for $p = 2$.

Next we prove (7.3). By Lemma 7.2, we have

$$\mathbb{E}\big[|\mathbb{1}_{\{\xi_n \leq x\}} - \mathbb{1}_{\{\xi_n^{(\kappa)} \leq x\}}|\big|\ell_n\big] = \big|\mathbb{P}[X_{\ell_n} \leq x|\ell_n] - \mathbb{P}[X_{\ell_n}^{(\kappa)} \leq x|\ell_n]\big|. \qquad (7.6)$$

By (3.2) in Theorem 3.3 (a), $\ell_n|\mathbb{P}[X_{\ell_n} \leq x|\ell_n] - \mathbb{P}[X_{\ell_n}^{(\kappa)} \leq x|\ell_n]| \leq \frac{1}{2}(\kappa/\overline{\sigma}_\kappa)\varphi_\kappa^3 \ell_n^{1/2}$. By Fubini's theorem, conditioning in each summand on ℓ_n, using (7.6) and Lemma 7.1, we have

$$\mathbb{E}\left[\sum_{n \in \mathbb{N}} \ell_n|\mathbb{1}_{\{\xi_n \leq x\}} - \mathbb{1}_{\{\xi_n^{(\kappa)} \leq x\}}|\right] \leq \frac{1}{2}\sqrt{t}(\kappa/\overline{\sigma}_\kappa)\varphi_\kappa^3 \sum_{n \in \mathbb{N}} \mathbb{E}[(\ell_n/t)^{1/2}] = \mu_0^\tau(\kappa, t).$$

Let $\delta \in (0, 2]$ satisfy $\inf_{u \in (0,1]} u^{\delta-2}(\overline{\sigma}_u^2 + \sigma^2) > 0$. By (3.3) in Theorem 3.3 (b), we get $\ell_n |\mathbb{P}[X_{\ell_n} \le x | \ell_n] - \mathbb{P}[X_{\ell_n}^{(\kappa)} \le x | \ell_n]| \le \psi_\kappa^{2/3} \ell_n^{1-2/(3\delta)}$, where $\psi_\kappa = C\kappa\varphi_\kappa$ as defined in (3.7). Moreover, $\ell_n |\mathbb{P}[X_{\ell_n} \le x | \ell_n] - \mathbb{P}[X_{\ell_n}^{(\kappa)} \le x | \ell_n]| \le \ell_n$. Hence by (7.6) and Lemma 7.1, we obtain

$$\sum_{n=1}^{\infty} \mathbb{E}[\ell_n | \mathbb{1}_{\{\xi_n \le x\}} - \mathbb{1}_{\{\xi_n^{(\kappa)} \le x\}}|]$$

$$\le \sum_{n=1}^{\infty} \mathbb{E}[\min\{\ell_n, \psi_\kappa^{2/3} \ell_n^{1-2/(3\delta)}\}]$$

$$= \begin{cases} \min\{t, \psi_\kappa^\delta\} + \frac{3\delta}{3\delta-2}\psi_\kappa^{2/3} t^{1-\frac{2}{3\delta}}(1 - \min\{1, t^{-1/\delta}\psi_\kappa\}^{\delta-2/3}), & \delta \in (0, 2] \setminus \{\frac{2}{3}\}, \\ \min\{t, \psi_\kappa^{2/3}\}(1 + \log^+(t\psi_\kappa^{-2/3})), & \delta = \frac{2}{3}, \end{cases}$$

completing the proof. □

Proof of Theorem 3.4 Let $\ell = (\ell_n)_{n \in \mathbb{N}}$ and $(\xi_n, \xi_n^{(\kappa)})$, $n \in \mathbb{N}$, be as in Lemma 7.3. Define the vector

$$(\zeta_1, \zeta_2, \zeta_3, \zeta_1^{(\kappa)}, \zeta_2^{(\kappa)}, \zeta_3^{(\kappa)})$$

$$:= \sum_{n=1}^{\infty} (\xi_n, \min\{\xi_n, 0\}, \ell_n \mathbb{1}_{\{\xi_n \le 0\}}, \xi_n^{(\kappa)}, \min\{\xi_n^{(\kappa)}, 0\}, \ell_n \mathbb{1}_{\{\xi_n^{(\kappa)} \le 0\}}).$$

By (2.1) and (7.1), we have $(\zeta_1, \zeta_2, \zeta_3) \overset{d}{=} \underline{\chi}_t$ and $(\zeta_1^{(\kappa)}, \zeta_2^{(\kappa)}, \zeta_3^{(\kappa)}) \overset{d}{=} \underline{\chi}_t^{(\kappa)}$. Hence it suffices to show that these vectors satisfy (3.4), (3.6) and (3.7). Since the function $x \mapsto \min\{x, 0\}$ is in $\mathrm{Lip}_1(\mathbb{R})$, the inequalities

$$\max\{|\zeta_1 - \zeta_1^{(\kappa)}|, |\zeta_2 - \zeta_2^{(\kappa)}|\} \le \sum_{n=1}^{\infty} |\xi_n - \xi_n^{(\kappa)}|,$$

$$|\zeta_3 - \zeta_3^{(\kappa)}| \le \sum_{n=1}^{\infty} \ell_n |\mathbb{1}_{\{\xi_n \le 0\}} - \mathbb{1}_{\{\xi_n^{(\kappa)} \le 0\}}|,$$

follow from the triangle inequality. The theorem follows from Lemma 7.3. □

Remark 7.4 Let C_t and $C_t^{(\kappa)}$ denote the convex minorants of X and $X^{(\kappa)}$ on $[0, t]$, respectively. Couple X and $X^{(\kappa)}$ in such a way that the stick-breaking processes describing the lengths of the faces of their convex minorants (see González Cázares and Mijatović [32, Theorem 11] and González Cázares et al. [35, Sect. 4.1]) coincide. (The Skorokhod space $\mathcal{D}[0, t]$ and the space of sequences on $\mathbb{R}$ are both Borel spaces by Kallenberg [39, Theorems A1.1, A1.2 and A2.2]; so the existence of such a coupling is guaranteed by [39, Theorem 6.10].) By conditioning further, we may assume the increments $(\xi_n, \xi_n^{(\kappa)})$, $n \in \mathbb{N}$, are coupled as in Lemma 7.3, yielding an *SBG coupling* of the pair $(X, X^{(\kappa)})$ on the interval $[0, t]$.

Geometric arguments similar to González Cázares et al. [34, Sect. 5.1] show that the sequences of heights of the faces of the convex minorants $(\xi_n)_{n\in\mathbb{N}}$ and $(\xi_n^{(\kappa)})_{n\in\mathbb{N}}$ satisfy

$$\sup_{s\in[0,t]} |C_t(s) - C_t^{(\kappa)}(s)| \le \sum_{n=1}^{\infty} |\xi_n - \xi_n^{(\kappa)}|,$$

$$|\underline{\tau}_t - \underline{\tau}_t^{(\kappa)}| \le \sum_{n=1}^{\infty} \ell_n |\mathbb{1}_{\{\xi_n \le 0\}} - \mathbb{1}_{\{\xi_n^{(\kappa)} \le 0\}}|.$$

Hence the inequalities in (7.3) and (7.4) yield the same bounds as in Theorem 3.4, but in a stronger metric (namely, the distance between the convex minorants in the supremum norm), while retaining the control on the time of the infimum.

7.3 The proofs of Propositions 3.7, 3.9, 3.12 and 3.15

The Lévy–Khintchine formula for X_t in (2.3), the definition of $X_t^{(\kappa)}$ in (2.5) and the inequality $e^z \ge 1 + z$ (for all $z \in \mathbb{R}$) imply for any $u \in \mathbb{R}$, $t > 0$ and $\kappa \in (0, 1]$ that

$$t^{-1} \log \mathbb{E}[e^{uX_t^{(\kappa)}}] = bu + \frac{(\sigma^2 + \overline{\sigma}_\kappa^2)u^2}{2} + \int_{\mathbb{R}\setminus(-\kappa,\kappa)} \left(e^{ux} - 1 - ux\mathbb{1}_{(-1,1)}(x)\right)\nu(dx)$$

$$\le \overline{\sigma}_\kappa^2 u^2/2 + t^{-1} \log \mathbb{E}[e^{uX_t}]. \tag{7.7}$$

Thus $\mathbb{E}[\exp(uX_t^{(\kappa)})] \le \mathbb{E}[\exp(uX_t)]\exp(\overline{\sigma}_\kappa^2 u^2 t/2)$, and in particular, the Gaussian approximation $X^{(\kappa)}$ has as many exponential moments as the Lévy process X.

Proof of Proposition 3.7 By Villani [68, Theorem 6.16], there exists a coupling between $(\xi, \zeta) \stackrel{d}{=} (X_T, \underline{X}_T)$ and $(\xi', \zeta') \stackrel{d}{=} (X_T^{(\kappa)}, \underline{X}_T^{(\kappa)})$ such that

$$\mathbb{E}[(|\xi - \xi'| + |\zeta - \zeta'|)^2]^{1/2} = \mathcal{W}_2\big((X_T, \underline{X}_T), (X_T^{(\kappa)}, \underline{X}_T^{(\kappa)})\big).$$

The identity $e^b - e^a = \int_a^b e^z dz$ implies that for $x \ge y$ and $x' \ge y'$, we have

$$|f(x, y) - f(x', y')| \le K(|e^x - e^{x'}| + |e^y - e^{y'}|)$$

$$\le K(|x - x'| + |y - y'|)e^{\max\{x,x'\}}. \tag{7.8}$$

Apply this inequality, the Cauchy–Schwarz inequality, the elementary inequalities $(a + b)^2 \le 2(a^2 + b^2)$ and $(a + b)^{1/2} \le a^{1/2} + b^{1/2}$, which hold for all $a, b \ge 0$, and the bound in (7.7) to obtain

$$\mathbb{E}[|f(\xi, \zeta) - f(\xi', \zeta')|] \le K\mathbb{E}[(|\xi - \xi'| + |\zeta - \zeta'|)^2]^{1/2}\mathbb{E}[(e^\xi + e^{\xi'})^2]^{1/2}$$

$$\le 2^{1/2}K\mathcal{W}_2\big((X_T, \underline{X}_T), (X_T^{(\kappa)}, \underline{X}_T^{(\kappa)})\big)\mathbb{E}[e^{2\xi} + e^{2\xi'}]^{1/2}$$

$$\le 2K\mathcal{W}_2\big((X_T, \underline{X}_T), (X_T^{(\kappa)}, \underline{X}_T^{(\kappa)})\big)\mathbb{E}[e^{2X_T}]^{1/2}(1 + e^{\overline{\sigma}_\kappa^2 T}).$$

Applying Corollary 3.5 gives the desired inequality, concluding the proof of the proposition. □

We now introduce a tool that uses the L^p-distance $\mathbb{E}[|\zeta - \zeta'|^p]^{1/p}$ between random variables ζ and ζ' to bound the L^1-distance $\mathbb{E}[|\mathbb{1}_{[y,\infty)}(\zeta) - \mathbb{1}_{[y,\infty)}(\zeta')|]$ between the indicators.

Lemma 7.5 *Let* (ξ, ζ) *and* (ξ', ζ') *be random vectors in* $\mathbb{R}^n \times \mathbb{R}$. *Fix* $y \in \mathbb{R}$ *and let* $h \in \mathrm{Lip}_K(\mathbb{R}^n)$ *satisfy* $0 \leq h \leq M$ *for some constants* $K, M \geq 0$. *Denote* $f(x,z) = h(x)\mathbb{1}_{[y,\infty)}(z)$. *Then for any* $p, r > 0$, *we have*

$$\mathbb{E}[|f(\xi, \zeta) - f(\xi', \zeta')|] \leq K\mathbb{E}[|\xi - \xi'|] + M\mathbb{P}[|\zeta - y| \leq r]$$
$$+ Mr^{-p}\mathbb{E}[|\zeta - \zeta'|^p]. \qquad (7.9)$$

In particular, if $|\mathbb{P}[\zeta \leq y] - \mathbb{P}[\zeta \leq y + r]| \leq C|r|^\gamma$ *for some* $C, \gamma > 0$ *and all* $r \in \mathbb{R}$, *then*

$$\mathbb{E}[|f(\xi, \zeta) - f(\xi', \zeta')|] \leq K\mathbb{E}[|\xi - \xi'|]$$
$$+ M\left(2C\frac{\gamma}{p}\right)^{\frac{p}{p+\gamma}}\left(1 + \frac{p}{\gamma}\right)\mathbb{E}[|\zeta - \zeta'|^p]^{\frac{\gamma}{p+\gamma}}. \qquad (7.10)$$

Remark 7.6 An analogous bound to the one in (7.9) holds for the indicator $\mathbb{1}_{(-\infty, y]}$. Moreover, it follows from the proof below that the boundedness of the function h in Lemma 7.5 may be replaced with a moment assumption $\xi, \xi' \in L^q$ for some $q > 1$. In that case, Hölder's inequality could be invoked to obtain an analogue to (7.11) below. Similar arguments may be used to simultaneously handle multiple indicators.

Proof of Lemma 7.5 Applying the local γ-Hölder-continuity of the distribution function of ζ to (7.9) and optimising over $r > 0$ yields (7.10). Thus it remains to establish (7.9).

Elementary set manipulation yields

$$|\mathbb{1}_{\{y \leq \zeta\}} - \mathbb{1}_{\{y \leq \zeta'\}}| = |\mathbb{1}_{\{\zeta' < y \leq \zeta\}} - \mathbb{1}_{\{\zeta < y \leq \zeta'\}}|$$
$$\leq \mathbb{1}_{\{|\zeta - \zeta'| > r, \zeta' < y \leq \zeta\}} + \mathbb{1}_{\{|\zeta - \zeta'| \leq r, \zeta' < y \leq \zeta\}}$$
$$+ \mathbb{1}_{\{|\zeta - \zeta'| > r, \zeta < y \leq \zeta'\}} + \mathbb{1}_{\{|\zeta - \zeta'| \leq r, \zeta < y \leq \zeta'\}}$$
$$\leq \mathbb{1}_{\{|\zeta - \zeta'| > r\}} + \mathbb{1}_{\{|\zeta - y| \leq r\}}.$$

Hence the triangle inequality and the Lipschitz property give

$$|f(\xi, \zeta) - f(\xi', \zeta')| \leq |h(\xi)||\mathbb{1}_{\{\zeta \geq y\}} - \mathbb{1}_{\{\zeta' \geq y\}}| + |h(\xi) - h(\xi')|\mathbb{1}_{\{\zeta' \geq y\}}$$
$$\leq M(\mathbb{1}_{\{|\zeta - y| \leq r\}} + \mathbb{1}_{\{|\zeta - \zeta'| > r\}}) + K|\xi - \xi'|. \qquad (7.11)$$

Taking expectations and using Markov's inequality $\mathbb{P}[|\zeta - \zeta'| > r] \leq r^{-p}\mathbb{E}[|\zeta - \zeta'|^p]$ yields (7.9). □

Proof of Proposition 3.9 Theorem 3.4 and (7.10) in Lemma 7.5 (with C and γ given in Assumption 3.8 and $p = 2$) applied to $(X_T, \underline{X}_T)$ and $(X_T^{(\kappa)}, \underline{X}_T^{(\kappa)})$ under the SBG coupling give the claim. □

Proof of Proposition 3.12 Analogously to the proof of Proposition 3.9, applying Theorem 3.4 and (7.10) in Lemma 7.5 (with C and γ given in Assumption 3.11 and $p = 1$) gives the result. □

Lemma 7.7 *Suppose X is not a compound Poisson process. Then the law of $\underline{\tau}_T$ is absolutely continuous on $(0, T)$ and its density is bounded on compact subsets of $(0, T)$.*

Proof If X or $-X$ is a subordinator, then $\underline{\tau}_T$ is a.s. 0 or T, respectively. In either case, the result follows immediately. Suppose now that neither X nor $-X$ is a subordinator. Denote by $\overline{n}(\zeta > \cdot)$ (resp. $\underline{n}(\zeta > \cdot)$) the intensity measures of the lengths ζ of the excursions away from 0 of the Markov process $\overline{X} - X$ (resp. $X - \underline{X}$). Then by Chaumont [16, Theorem 6], the law of $\underline{\tau}_T$ can only have atoms at 0 or T, is absolutely continuous on $(0, T)$ and its density is given by $s \mapsto \underline{n}(\zeta > s)\overline{n}(\zeta > T - s)$, $s \in (0, T)$. The maps $s \mapsto \underline{n}(\zeta > s)$ and $s \mapsto \overline{n}(\zeta > s)$ are nonincreasing, and so the density is bounded on any compact subset of $(0, T)$, completing the proof. □

In preparation for the next result, we introduce the following assumption.

Assumption 7.8 *There exists some function $a : (0, \infty) \to (0, \infty)$ such that $X_t/a(t)$ converges in distribution to an α-stable law as $t \to 0$.*

Proposition 7.9 *Let Assumption 7.8 hold for some $\alpha \in (0, 2]$.*

(a) If $\alpha > 1$, then Assumption 3.8 holds uniformly on compact subsets of $(-\infty, 0)$ with $\gamma = 1$.

(b) Suppose $\rho := \lim_{t\to 0} \mathbb{P}[X_t > 0] \in (0, 1)$. Then for any $\gamma \in (0, \min\{\rho, 1 - \rho\})$, there exists a constant $C > 0$ such that Assumption 3.11 holds for all $s \in [0, T]$.

Note that ρ is well defined under Assumption 7.8 and that $X_t/a(t)$ can only have a nonzero weak limit as $t \to 0$ if the limit is α-stable. Moreover, in that case, a is necessarily regularly varying at 0 with index $1/\alpha$, and α is given in terms of the Lévy triplet (σ^2, ν, b) of X by

$$\alpha := \begin{cases} 2, & \sigma \neq 0, \\ 1, & \beta \in (0, 1) \text{ and } b \neq \int_{(-1,1)} x\nu(dx), \\ \beta, & \text{otherwise}, \end{cases}$$

where β is the BG index introduced in (2.6). In fact, the assumptions of Proposition 3.15 imply Assumption 7.8 for $\alpha > 1$ by Bisewski and Ivanovs [8, Proposition 2.3], so that Proposition 7.9 generalises Proposition 3.15. We refer the reader to Ivanovs [38, Sects. 3 and 4] for conditions that are equivalent to Assumption 7.8.

Assumption 7.8 allows the cases $\rho = 0$ or $\rho = 1$ when $\alpha \leq 1$, which correspond to the stable limit being a.s. negative or a.s. positive, respectively. In these cases, the

distribution of $\overline{\tau}_T(X)$ may have an atom at 0 or T, while the law of $\overline{\tau}_T(X^{(\kappa)})$ is absolutely continuous, making the convergence in the Kolmogorov distance impossible. This is the reason for excluding $\rho \in \{0, 1\}$ in Proposition 7.9.

Proof of Proposition 7.9 By [8, Lemma 5.7], under the assumptions in part (a) of the proposition, $\underline{X}_T$ has a continuous density on $(-\infty, 0)$, implying the conclusion in (a).

Since $\rho = \lim_{t \to 0} \mathbb{P}[X_t > 0] \in (0, 1)$, 0 is regular for both half-lines by Rogozin's criterion; see Sato [62, Theorem 47.2]. Chaumont [16, Theorem 6] then asserts that the law of $\underline{\tau}_T$ is absolutely continuous with density given by $s \mapsto \underline{n}(\zeta > s)\overline{n}(\zeta > T-s)$, $s \in (0, T)$. The maps $s \mapsto \underline{n}(\zeta > s)$ and $s \mapsto \overline{n}(\zeta > s)$ are nonincreasing and by [8, Proposition 3.5] regularly varying with indices $\rho - 1$ and $-\rho$, respectively. Thus for any $\gamma \in (0, \min\{\rho, 1-\rho\})$, there exists some constant $C > 0$ such that the inequality $\underline{n}(\zeta > s)\overline{n}(\zeta > T - s) \leq Cs^{\gamma-1}(T-s)^{\gamma-1}$ holds for all $s \in (0, T)$. Thus for any $s, t \in [0, T/2]$ with $t \geq s$, we have

$$\mathbb{P}[\underline{\tau}_T \leq t] - \mathbb{P}[\underline{\tau}_T \leq s] \leq \int_s^t Cu^{\gamma-1}(T-u)^{\gamma-1}du \leq C\int_s^t u^{\gamma-1}(T/2)^{\gamma-1}du$$
$$\leq C\gamma^{-1}(T/2)^{\gamma-1}(t^\gamma - s^\gamma) \leq C\gamma^{-1}(T/2)^{\gamma-1}(t-s)^\gamma,$$

since the map $x \mapsto x^\gamma$ is concave. A similar bound holds for $s, t \in [T/2, T]$. Moreover, when $s \in [0, T/2]$ and $t \in [T/2, T]$, we have

$$\mathbb{P}[\underline{\tau}_T \leq t] - \mathbb{P}[\underline{\tau}_T \leq s] \leq \mathbb{P}[\underline{\tau}_T \leq t] - \mathbb{P}[\underline{\tau}_T \leq T/2] + \mathbb{P}[\underline{\tau}_T \leq T/2] - \mathbb{P}[\underline{\tau}_T \leq s]$$
$$\leq C\gamma^{-1}(T/2)^{\gamma-1}\big((T/2-s)^\gamma + (t-T/2)^\gamma\big)$$
$$\leq C\gamma^{-1}(T/2)^{2\gamma-2}(t-s)^\gamma.$$

This gives part (b) of the proposition. $\qquad\square$

7.4 Level variances under SBG-Alg

In the present subsection, we establish bounds on the level variances under the coupling $\underline{\Pi}_{n,T}^{\kappa_1,\kappa_2}$ (constructed in SBG-Alg) for Lipschitz, locally Lipschitz and discontinuous payoff functions (see BT_1 in (3.12) and BT_2 in (3.14)) of $\underline{\chi}_T$.

Theorem 7.10 *Fix $T > 0$, $n \in \mathbb{N}$ and $1 \geq \kappa_1 > \kappa_2 > 0$. Let $(Z_{n,T}^{(\kappa_i)}, \underline{Z}_{n,T}^{(\kappa_i)}, \underline{\tau}_{n,T}^{(\kappa_i)})$ denote the vector $\underline{\chi}_{n,T}^{(\kappa_i)}$, $i \in \{1, 2\}$, where the vector $(\underline{\chi}_{n,T}^{(\kappa_1)}, \underline{\chi}_{n,T}^{(\kappa_2)})$ constructed in SBG-Alg follows the law $\underline{\Pi}_{n,T}^{\kappa_1,\kappa_2}$.*

(a) *For any Lipschitz function $f \in \mathrm{Lip}_K(\mathbb{R}^2)$, $K > 0$, we have*

$$\mathbb{E}\big[\big(f(Z_{n,T}^{(\kappa_2)}, \underline{Z}_{n,T}^{(\kappa_2)}) - f(Z_{n,T}^{(\kappa_1)}, \underline{Z}_{n,T}^{(\kappa_1)})\big)^2\big] \leq K^2 T(27\sigma^2 2^{-n} + 40\overline{\sigma}_{\kappa_1}^2). \qquad (7.12)$$

For $f \in \mathrm{locLip}_K(\mathbb{R}^2)$ as defined in Sect. 3.2, if $\int_{[1,\infty)} e^{4x}\nu(dx) < \infty$, then there exists a constant $C > 0$ independent of (n, κ_1, κ_2) such that

$$\mathbb{E}\big[\big(f(Z_{n,T}^{(\kappa_2)}, \underline{Z}_{n,T}^{(\kappa_2)}) - f(Z_{n,T}^{(\kappa_1)}, \underline{Z}_{n,T}^{(\kappa_1)})\big)^2\big]$$
$$\leq C\left(\left(\frac{2}{3}\right)^{\frac{n}{2}}\mathbb{1}_{\mathbb{R}\setminus\{0\}}(\sigma) + \overline{\sigma}_{\kappa_1}^2 + \overline{\sigma}_{\kappa_1}\kappa_1\right). \qquad (7.13)$$

(b) *Suppose Assumption 3.8 is satisfied by some $y < 0$ and $C, \gamma > 0$. Then for any $f \in \mathrm{BT}_1(y, K, M)$, $K, M \geq 0$, there exists some $K' > 0$ independent of (n, κ_1, κ_2) such that*

$$\mathbb{E}\big[\big(f(Z_{n,T}^{(\kappa_2)}, \underline{Z}_{n,T}^{(\kappa_2)}) - f(Z_{n,T}^{(\kappa_1)}, \underline{Z}_{n,T}^{(\kappa_1)})\big)^2\big] \leq K'(\sigma^2 2^{-n} + \overline{\sigma}_{\kappa_1}^2)^{\frac{\gamma}{2+\gamma}}.$$

(c) *If $\delta \in (0, 2]$ satisfies Assumption 3.2, then there exists some $C > 0$ such that for any $K > 0$, $f \in \mathrm{Lip}_K(\mathbb{R})$, $n \in \mathbb{N}$, $\kappa_1 > \kappa_2$ and $p \in \{1, 2\}$, we have*

$$\mathbb{E}[|f(\underline{\tau}_{n,T}^{(\kappa_1)}) - f(\underline{\tau}_{n,T}^{(\kappa_2)})|^p]$$
$$\leq 2K^p T^p \Big(2^{-n} + C\overline{\sigma}_{\kappa_1}^{\min\{\frac{2\delta}{2-\delta}, \frac{1}{2}\}}\big(1 + |\log\kappa_1| \mathbb{1}_{\{2/5\}}(\delta)\big)\Big). \qquad (7.14)$$

(d) *Fix $s \in (0, T)$ and let Assumption 3.2 hold for some $\delta \in (0, 2]$. Then for any $f \in \mathrm{BT}_2(s, K, M)$, $K, M \geq 0$, there exists a constant $C > 0$ such that for any $n \in \mathbb{N}$, $p \in \{1, 2\}$ and $\kappa_1 > \kappa_2$, we have*

$$\mathbb{E}[|f(\underline{\chi}_{n,T}^{(\kappa_1)}) - f(\underline{\chi}_{n,T}^{(\kappa_2)})|^p] \leq C\Big(2^{-n/2} + \overline{\sigma}_{\kappa_1}^{\min\{\frac{\delta}{2-\delta}, \frac{1}{4}\}}\big(1 + \sqrt{|\log\kappa_1|} \mathbb{1}_{\{2/5\}}(\delta)\big)\Big).$$

The synchronous coupling of the large jumps of the Gaussian approximations implicit in SBG-Alg ensures that no moment assumption on the large jumps of X is necessary for (7.12) to hold. For locally Lipschitz payoffs, however, the function may magnify the distance when a large jump occurs. This leads to the moment assumption $\int_{[1,\infty)} e^{4x} \nu(dx) < \infty$ for $f \in \mathrm{locLip}_K(\mathbb{R}^2)$.

The proof of Theorem 7.10 requires bounds on certain moments of the differences of the components of the output of Algorithms 1 and 2 and SBG-Alg. They are given in Proposition 7.11.

Proposition 7.11 *For any $1 \geq \kappa_1 > \kappa_2 > 0$, $t > 0$ and $n \in \mathbb{N}$, the following statements hold*:

(a) *The pair $(Z_t^{(\kappa_1)}, Z_t^{(\kappa_2)}) \sim \Pi_t^{\kappa_1, \kappa_2}$ constructed in Algorithm 1 satisfies the inequalities*

$$\mathbb{E}[(Z_t^{(\kappa_1)} - Z_t^{(\kappa_2)})^2] \leq 2(\overline{\sigma}_{\kappa_1}^2 - \overline{\sigma}_{\kappa_2}^2)t,$$
$$\mathbb{E}[(Z_t^{(\kappa_1)} - Z_t^{(\kappa_2)})^4] \leq 12(\overline{\sigma}_{\kappa_1}^2 - \overline{\sigma}_{\kappa_2}^2)^2 t^2 + (\overline{\sigma}_{\kappa_1}^2 - \overline{\sigma}_{\kappa_2}^2)\kappa_1^2 t.$$

Moreover, we have $\mathbb{E}[(\underline{Z}_t^{(\kappa_1)} - \underline{Z}_t^{(\kappa_2)})^{2p}] \leq 4\mathbb{E}[(Z_t^{(\kappa_1)} - Z_t^{(\kappa_2)})^{2p}]$ for $p \in \{1, 2\}$.

(b) *The vector $(Z_t^{(\kappa_1)}, \underline{Z}_t^{(\kappa_1)}, \underline{\tau}_t^{(\kappa_1)}, Z_t^{(\kappa_2)}, \underline{Z}_t^{(\kappa_2)}, \underline{\tau}_t^{(\kappa_2)}) \sim \underline{\Pi}_t^{\kappa_1, \kappa_2}$ constructed in Algorithm 2 satisfies the inequalities*

$$\mathbb{E}[(Z_t^{(\kappa_1)} - Z_t^{(\kappa_2)})^2] = 2(\sigma^2 + \overline{\sigma}_{\kappa_1}^2)t,$$
$$\mathbb{E}[(Z_t^{(\kappa_1)} - Z_t^{(\kappa_2)})^4] \leq 12(\sigma^2 + \overline{\sigma}_{\kappa_1}^2)^2 t^2 + (\overline{\sigma}_{\kappa_1}^2 - \overline{\sigma}_{\kappa_2}^2)\kappa_1^2 t.$$

Moreover, we have $\mathbb{E}[(\underline{Z}_t^{(\kappa_1)} - \underline{Z}_t^{(\kappa_2)})^{2p}] \leq 4\mathbb{E}[(Z_t^{(\kappa_1)} - Z_t^{(\kappa_2)})^{2p}]$ for $p \in \{1, 2\}$.

(c) *The coupling $(\underline{\chi}_{n,t}^{(\kappa_1)}, \underline{\chi}_{n,t}^{(\kappa_2)}) \sim \underline{\Pi}_{n,t}^{\kappa_1,\kappa_2}$ constructed in* SBG-Alg *with components* $\underline{\chi}_{n,t}^{(\kappa_i)} = (Z_{n,t}^{(\kappa_i)}, \underline{Z}_{n,t}^{(\kappa_i)}, \underline{\tau}_{n,t}^{(\kappa_i)})$, $i \in \{1, 2\}$, *satisfies the inequalities*

$$\mathbb{E}[(Z_{n,t}^{(\kappa_1)} - Z_{n,t}^{(\kappa_2)})^2] \leq 2(\sigma^2 2^{-n} + \overline{\sigma}_{\kappa_1}^2)t, \tag{7.15}$$

$$\mathbb{E}[(Z_{n,t}^{(\kappa_1)} - Z_{n,t}^{(\kappa_2)})^4] \leq (25\overline{\sigma}_{\kappa_1}^4 + 24\sigma^4 3^{-n})t^2 + \overline{\sigma}_{\kappa_1}^2 \kappa_1^2 t, \tag{7.16}$$

$$\mathbb{E}[(\underline{Z}_{n,t}^{(\kappa_1)} - \underline{Z}_{n,t}^{(\kappa_2)})^2] \leq (2 + 3\pi)(\sigma^2 + \overline{\sigma}_{\kappa_1}^2)2^{-n}t + (2 + 5\pi)\overline{\sigma}_{\kappa_1}^2 t, \tag{7.17}$$

$$\mathbb{E}[(\underline{Z}_{n,t}^{(\kappa_1)} - \underline{Z}_{n,t}^{(\kappa_2)})^4] \leq 2 \times 10^3 \big((\sigma^2 + \overline{\sigma}_{\kappa_1})^2 3^{-n} + \overline{\sigma}_{\kappa_1}^4\big)t^2$$
$$+ 2\pi \overline{\sigma}_{\kappa_1}^{5/2} \kappa_1^{3/2} t^{5/4} + 4\overline{\sigma}_{\kappa_1}^2 \kappa_1^2 t. \tag{7.18}$$

Remark 7.12 (i) Applying Proposition 7.11, we see that the L^2-norms of the differences $Z_{n,t}^{(\kappa_1)} - Z_{n,t}^{(\kappa_2)}$ and $\underline{Z}_{n,t}^{(\kappa_1)} - \underline{Z}_{n,t}^{(\kappa_2)}$ of the components of $(\underline{\chi}_{n,t}^{(\kappa_1)}, \underline{\chi}_{n,t}^{(\kappa_2)})$ constructed in SBG-Alg decay at the same rate as the L^2-norm of $Z_t^{(\kappa_1)} - Z_t^{(\kappa_2)}$ constructed in Algorithm 1. Indeed, assume that $\kappa_1 = c\kappa_2$ for some $c > 1$, $\kappa_2 \to 0$ and for some $c', r > 0$ and all $x > 0$, we have $\overline{\nu}(x) = \nu(\mathbb{R} \setminus (-x, x)) \geq c'x^{-r}$. Then for $n = \lceil \log^2(1 + \overline{\nu}(\kappa_2)) \rceil$, we have $2^{-n} \leq \overline{\sigma}_{\kappa_1}^2$ for all sufficiently small κ_1, implying the claim by Proposition 7.11 (a) and (c). Moreover, by Corollary 4.2, the corresponding expected computational complexities of Algorithm 1 and SBG-Alg are proportional as $\kappa_2 \to 0$. Furthermore, since the decay of the bias of SBG-Alg is by Theorem 3.4 at most a logarithmic factor away from that of Algorithm 1, the MLMC estimator based on Algorithm 1 for $\mathbb{E}[f(X_t)]$ has the same computational complexity (up to logarithmic factors) as the MLMC estimator for $\mathbb{E}[f(X_t, \underline{X}_t)]$ based on SBG-Alg (see Table 2 for the complexity of the latter).

(ii) The proof of Proposition 7.11 implies that an improvement in Algorithm 1 (i.e., a simulation procedure for a coupling with a smaller L^2-norm of $Z_t^{(\kappa_1)} - Z_t^{(\kappa_2)}$) would result in an improvement in SBG-Alg for the simulation of a coupling $(\underline{\chi}_t^{(\kappa_1)}, \underline{\chi}_t^{(\kappa_2)})$. Interestingly, this holds in spite of the fact that SBG-Alg calls Algorithm 2 whose coupling $\underline{\Pi}_t^{\kappa_1,\kappa_2}$ is inefficient in terms of the L^2-distance, but is applied over the short interval $[0, L_n]$.

(iii) A nontrivial bound on the moments of the difference $\underline{\tau}_t^{(\kappa_1)} - \underline{\tau}_t^{(\kappa_2)}$ under the coupling of Algorithm 2, which would complete the statement in Proposition 7.11 (b), appears to be out of reach. By the SB representation in (2.2), such a bound is not necessary for our purposes. The corresponding bound on the moments of the difference $\underline{\tau}_{n,t}^{(\kappa_1)} - \underline{\tau}_{n,t}^{(\kappa_2)}$ constructed in SBG-Alg follows from Proposition 7.13 below; see (7.24).

(iv) The bounds on the fourth moments in (7.16) and (7.18) are required to control the level variances of the MLMC estimator in the case of locally Lipschitz payoff functions and are applied in the proof of Theorem 7.10 (a).

Proof of Proposition 7.11 (a) The difference $Z_t^{(\kappa_1)} - Z_t^{(\kappa_2)}$ constructed by Algorithm 1 equals by (2.5) a sum of the two independent martingales

$$\big((\overline{\sigma}_{\kappa_1}^2 + \sigma^2)^{1/2} - (\overline{\sigma}_{\kappa_2}^2 + \sigma^2)^{1/2}\big)W_t, \qquad J_t^{2,\kappa_1} - J_t^{2,\kappa_2} + (b_{\kappa_1} - b_{\kappa_2})t.$$

Thus we obtain the identity

$$\mathbb{E}[(Z_t^{(\kappa_1)} - Z_t^{(\kappa_2)})^2] = \left(\left(\sqrt{\sigma^2 + \overline{\sigma}_{\kappa_1}^2} - \sqrt{\sigma^2 + \overline{\sigma}_{\kappa_2}^2}\right)^2 + \overline{\sigma}_{\kappa_1}^2 - \overline{\sigma}_{\kappa_2}^2\right)t.$$

The first inequality follows since $0 < (\sigma^2 + \overline{\sigma}_{\kappa_1}^2)^{1/2} - (\sigma^2 + \overline{\sigma}_{\kappa_2}^2)^{1/2} \leq (\overline{\sigma}_{\kappa_1}^2 - \overline{\sigma}_{\kappa_2}^2)^{1/2}$. Since $Z_t^{(\kappa_1)} - Z_t^{(\kappa_2)}$ is a Lévy process, differentiating its Lévy–Khintchine formula in (2.3) yields the identity

$$\mathbb{E}[(Z_t^{(\kappa_1)} - Z_t^{(\kappa_2)})^4] = 3\left(\left(\sqrt{\sigma^2 + \overline{\sigma}_{\kappa_1}^2} - \sqrt{\sigma^2 + \overline{\sigma}_{\kappa_2}^2}\right)^2 + \overline{\sigma}_{\kappa_1}^2 - \overline{\sigma}_{\kappa_2}^2\right)^2 t^2$$

$$+ t \int_{(-\kappa_1, \kappa_1) \setminus (-\kappa_2, \kappa_2)} x^4 \nu(dx),$$

implying the second inequality. Since $|\underline{Z}_t^{(\kappa_1)} - \underline{Z}_t^{(\kappa_2)}| \leq \sup_{s \in [0,t]} |Z_s^{(\kappa_1)} - Z_s^{(\kappa_2)}|$, Doob's maximal martingale inequality applied to the martingale $(Z_s^{(\kappa_1)} - Z_s^{(\kappa_2)})_{s \in [0,t]}$ yields

$$\mathbb{E}[|\underline{Z}_t^{(\kappa_1)} - \underline{Z}_t^{(\kappa_2)}|^p] \leq (1 - 1/p)^{-p} \mathbb{E}[|Z_t^{(\kappa_1)} - Z_t^{(\kappa_2)}|^p], \qquad p > 1.$$

The corresponding inequalities follow because $(p/(p-1))^p \leq 4$ for $p \in \{2, 4\}$.

(b) Analogously to part (a), the difference $Z_t^{(\kappa_1)} - Z_t^{(\kappa_2)}$ constructed in Algorithm 2 is a sum of two independent martingales $(\overline{\sigma}_{\kappa_1}^2 + \sigma^2)^{1/2} B_t - (\overline{\sigma}_{\kappa_2}^2 + \sigma^2)^{1/2} W_t$ and $J_t^{2,\kappa_1} - J_t^{2,\kappa_2} + (b_{\kappa_1} - b_{\kappa_2})t$, where B and W are independent standard Brownian motions. Thus the statements follow as in part (a).

(c) Let $(\xi_{1,k}, \xi_{2,k}) \sim \Pi_{\ell_k}^{\kappa_1, \kappa_2}$, $k \in \{1, \ldots, n\}$, and $(\underline{\zeta}_1, \underline{\zeta}_2) \sim \underline{\Pi}_{L_n}^{\kappa_1, \kappa_2}$ be independent draws as in line 2 of SBG-Alg. Denote by $(\xi_{i,n+1}, \underline{\xi}_{i,n+1})$ the first two coordinates of $\underline{\zeta}_i$, $i \in \{1, 2\}$. Since the variables $\xi_{1,k} - \xi_{2,k}$, $k = 1, \ldots, n+1$, have zero mean and are uncorrelated, by conditioning on ℓ_k, $k = 1, \ldots, n$, and L_n and applying parts (a) and (b), we obtain

$$\mathbb{E}[(Z_{n,t}^{(\kappa_1)} - Z_{n,t}^{(\kappa_2)})^2] = \mathbb{V}[Z_{n,t}^{(\kappa_1)} - Z_{n,t}^{(\kappa_2)}]$$

$$= \mathbb{V}[\xi_{1,n+1} - \xi_{2,n+1}] + \sum_{k=1}^{n} \mathbb{V}[\xi_{1,k} - \xi_{2,k}]$$

$$\leq 2(\sigma^2 + \overline{\sigma}_{\kappa_1}^2) \mathbb{E}[L_n] + 2\overline{\sigma}_{\kappa_1}^2 \sum_{k=1}^{n} \mathbb{E}[\ell_k]$$

$$= 2(\sigma^2 + \overline{\sigma}_{\kappa_1}^2) 2^{-n} t + 2\overline{\sigma}_{\kappa_1}^2 (1 - 2^{-n}) t,$$

implying (7.15). Similarly, by conditioning on ℓ_k, $k = 1, \ldots, n$, and L_n, we deduce that the expectations of

$$(\xi_{1,k_1} - \xi_{2,k_1})^3 (\xi_{1,k_2} - \xi_{2,k_2}), \quad (\xi_{1,k_1} - \xi_{2,k_1})^2 \prod_{i=2}^{3} (\xi_{1,k_i} - \xi_{2,k_i}), \quad \prod_{i=1}^{4} (\xi_{1,k_i} - \xi_{2,k_i})$$

vanish for any distinct $k_1, k_2, k_3, k_4 \in \{1, \ldots, n+1\}$. Thus expanding yields

$$\mathbb{E}[(Z_{n,t}^{(\kappa_1)} - Z_{n,t}^{(\kappa_2)})^4] = \sum_{k=1}^{n+1} \mathbb{E}[(\xi_{1,k} - \xi_{2,k})^4]$$

$$+ 6 \sum_{m=1}^{n} \sum_{k=m+1}^{n+1} \mathbb{E}[(\xi_{1,m} - \xi_{2,m})^2 (\xi_{1,k} - \xi_{2,k})^2].$$

The summands in the first sum are easily bounded by parts (a) and (b). To bound the summands of the second sum, condition on ℓ_k, $k = 1, \ldots, n$, and L_n and apply parts (a) and (b) to get

$$\mathbb{E}[(\xi_{1,k} - \xi_{2,k})^2 (\xi_{1,m} - \xi_{2,m})^2] \leq \begin{cases} 4\overline{\sigma}_{\kappa_1}^4 \mathbb{E}[\ell_m \ell_k], & m < k \leq n, \\ 4(\sigma^2 + \overline{\sigma}_{\kappa_1}^2)\overline{\sigma}_{\kappa_1}^2 \mathbb{E}[\ell_m L_n], & m < k = n+1. \end{cases}$$

Inequality (7.16) follows since $\mathbb{E}[\ell_m \ell_k] = 3^{-m} 2^{m-k-1} t^2$, $\mathbb{E}[\ell_k L_n] = 3^{-k} 2^{k-n-1} t^2$ for $m < k \leq n$ and $\sigma^2 2^{-n} \overline{\sigma}_\kappa^2 \leq \sigma^2 3^{-n/2} \overline{\sigma}_\kappa^2 \leq (\sigma^4 3^{-n} + \overline{\sigma}_\kappa^4)/2$.

The representation in line 3 of SBG-Alg and an appeal to the elementary inequality $|a - b| \geq |\min\{a, 0\} - \min\{b, 0\}|$ (for all $a, b \in \mathbb{R}$) imply

$$\mathbb{E}[(\underline{Z}_{n,t}^{(\kappa_1)} - \underline{Z}_{n,t}^{(\kappa_2)})^2] \leq \mathbb{E}\left[(\underline{\xi}_{1,n+1} - \underline{\xi}_{2,n+1})^2 + \sum_{k=1}^{n} (\xi_{1,k} - \xi_{2,k})^2 \right] \qquad (7.19)$$

$$+ 2\mathbb{E}\left[\sum_{k=1}^{n} |\underline{\xi}_{1,n+1} - \underline{\xi}_{2,n+1}| |\xi_{1,k} - \xi_{2,k}| \right.$$

$$\left. + \sum_{m=1}^{n-1} \sum_{k=m+1}^{n} |\xi_{1,m} - \xi_{2,m}| |\xi_{1,k} - \xi_{2,k}| \right].$$

The first term on the right-hand side of this inequality is easily bounded via the inequalities in parts (a) and (b). To bound the second term, condition on ℓ_k, $k = 1, \ldots, n$, and L_n, apply the Cauchy–Schwarz inequality, denote $\upsilon := \sqrt{\sigma^2 + \overline{\sigma}_{\kappa_1}^2}$ and observe that for $m < k \leq n$, we get

$$\mathbb{E}[|\underline{\xi}_{1,n+1} - \underline{\xi}_{2,n+1}| |\xi_{1,k} - \xi_{2,k}|] \leq \mathbb{E}\left[\sqrt{16(\sigma^2 + \overline{\sigma}_{\kappa_1}^2)\overline{\sigma}_{\kappa_1}^2 \ell_k L_n} \right]$$

$$= \pi \upsilon \overline{\sigma}_{\kappa_1} \left(\frac{2}{3}\right)^n \left(\frac{3}{4}\right)^k t,$$

$$\mathbb{E}[|\xi_{1,m} - \xi_{2,m}| |\xi_{1,k} - \xi_{2,k}|] \leq \mathbb{E}\left[\sqrt{4\overline{\sigma}_{\kappa_1}^4 \ell_m \ell_k} \right] = \pi \overline{\sigma}_{\kappa_1}^2 (1/2)^{m+1} (2/3)^{k-m} t,$$

where the equalities follow from the definition of the stick-breaking process (see Sect. 2.1). By (7.19), we have

$$\mathbb{E}[(\underline{Z}_{n,t}^{(\kappa_1)} - \underline{Z}_{n,t}^{(\kappa_2)})^2] \leq \upsilon^2 2^{1-n} t + 2\overline{\sigma}_{\kappa_1}^2 t \sum_{k=1}^{\infty} 2^{-k} + 2\pi \upsilon \overline{\sigma}_{\kappa_1} \left(\frac{2}{3}\right)^n t \sum_{k=1}^{\infty} \left(\frac{3}{4}\right)^k$$

$$+ \pi \overline{\sigma}_{\kappa_1}^2 t \sum_{m=1}^{\infty} \sum_{k=1}^{\infty} 2^{-m} \left(\frac{2}{3}\right)^k.$$

So (7.17) follows from the inequalities $\upsilon(2/3)^n \overline{\sigma}_\kappa \leq \upsilon 2^{-n/2} \overline{\sigma}_\kappa \leq (\upsilon^2 2^{-n} + \overline{\sigma}_\kappa^2)/2$.

As before, $|\min\{a, 0\} - \min\{b, 0\}| \leq |a - b|$ for $a, b \in \mathbb{R}$ yields the inequality

$$\mathbb{E}[(\underline{Z}_{n,t}^{(\kappa_1)} - \underline{Z}_{n,t}^{(\kappa_2)})^4] \leq \mathbb{E}\left[\left(|\underline{\xi}_{1,n+1} - \underline{\xi}_{2,n+1}| + \sum_{k=1}^{n} |\xi_{1,k} - \xi_{2,k}|\right)^4\right]. \qquad (7.20)$$

By Jensen's inequality, $\mathbb{E}[|\vartheta|^3] \leq \mathbb{E}[\vartheta^4]^{3/4}$ and $\mathbb{E}[\vartheta] \leq \sqrt{\mathbb{E}[\vartheta^2]}$ for any random variable ϑ. Hence we may bound the first and third conditional moments of the differences $|\xi_{1,k} - \xi_{2,k}|$ and $|\underline{\xi}_{1,n+1} - \underline{\xi}_{2,n+1}|$ given ℓ_k, $k = 1, \ldots, n$, and L_n. Thus by expanding (7.20), conditioning on ℓ_k, $k = 1, \ldots, n$, and L_n and using elementary estimates as in all the previously developed bounds, we obtain (7.18). □

In order to control the level variances of the MLMC estimator in (4.3) for discontinuous payoffs of $\underline{\chi}_t$ and functions of $\underline{\tau}_t$, we should need to apply Lemma 7.5 to the components of $(\underline{\chi}_{n,t}^{(\kappa_1)}, \underline{\chi}_{n,t}^{(\kappa_2)})$ constructed in SBG-Alg. In particular, the assumption in Lemma 7.5 requires a control on the constants in the local Lipschitz property of the distribution functions of the various components of $(\underline{\chi}_{n,t}^{(\kappa_1)}, \underline{\chi}_{n,t}^{(\kappa_2)})$ in terms of the cutoff levels κ_1 and κ_2. As such a uniform bound in the cutoff level appears to be out of reach, we establish Proposition 7.13 which allows us to compare the sampled quantities $\underline{\chi}_{n,t}^{(\kappa_1)}$ and $\underline{\chi}_{n,t}^{(\kappa_2)}$ with their limit $\underline{\chi}_t$ (as $\kappa_1, \kappa_2 \to 0$). Since under mild assumptions, the distribution functions of the components of the limit $\underline{\chi}_t$ possess the necessary regularity and do not depend on the cutoff level, the application of Lemma 7.5 in the proof of Theorem 7.10 becomes feasible by using Proposition 7.13.

Proposition 7.13 *There exists a coupling between the vector $\underline{\chi}_t = (X_t, \underline{X}_t, \underline{\tau}_t)$ and the pair of vectors $(\underline{\chi}_{n,t}^{(\kappa_1)}, \underline{\chi}_{n,t}^{(\kappa_2)}) \sim \underline{\Pi}_{n,t}^{\kappa_1;\kappa_2}$ such that for $i \in \{1, 2\}$ and any $p \geq 1$, the vector $(Z_{n,t}^{(\kappa_i)}, \underline{Z}_{n,t}^{(\kappa_i)}, \underline{\tau}_{n,t}^{(\kappa_i)}) = \underline{\chi}_{n,t}^{(\kappa_i)}$ satisfies*

$$\mathbb{E}[(X_t - Z_{n,t}^{(\kappa_i)})^2] \leq \left(4\sigma^2 2^{-n} \mathbb{1}_{\{1\}}(i) + 2\overline{\sigma}_{\kappa_i}^2\right)t, \qquad (7.21)$$

$$\mathbb{E}[(\underline{X}_t - \underline{Z}_{n,t}^{(\kappa_i)})^2] \leq \left(48\sigma^2 2^{-n} \mathbb{1}_{\{1\}}(i) + 42\overline{\sigma}_{\kappa_i}^2\right)t. \qquad (7.22)$$

Moreover, if $\delta \in (0, 2]$ satisfies Assumption 3.2, we have

$$\mathbb{E}[|\underline{\tau}_t - \underline{\tau}_{n,t}^{(\kappa_i)}|^p] \leq 2^{-n} t^p + t^{p-1} \theta(t, \kappa_i), \qquad (7.23)$$

where given $T \geq t$, there exists a constant $C > 0$ dependent only on (T, σ^2, ν, b) such that for all $\kappa \in (0, 1]$, the function $\theta(t, \kappa)$ is defined as

$$\theta(t, \kappa) := \begin{cases} \min\{1, \sqrt{C\overline{\sigma}_\kappa}\}t, & \delta = 2, \\[2mm] \min\{t, (C\overline{\sigma}_\kappa)^{\frac{2\delta}{2-\delta}}\} \\ \quad + \frac{4\delta}{5\delta-2}\sqrt{C\overline{\sigma}_\kappa}(t^{\frac{5\delta-2}{4\delta}} - \min\{t, (C\overline{\sigma}_\kappa)^{\frac{4\delta}{4-2\delta}}\}^{\frac{5\delta-2}{4\delta}}), & \delta \in (0, 2) \setminus \{\frac{2}{5}\}, \\[2mm] \min\{t, \sqrt{C\overline{\sigma}_\kappa}\} + \sqrt{C\overline{\sigma}_\kappa}\log^+(t/\sqrt{C\overline{\sigma}_\kappa}), & \delta = \frac{2}{5}. \end{cases}$$

A simple consequence of (7.23) (with $p = 1$) in Proposition 7.13 and the elementary inequality $|\underline{\tau}_{n,t}^{(\kappa_1)} - \underline{\tau}_{n,t}^{(\kappa_2)}| \leq t$ is that the coupling in SBG-Alg satisfies

$$\mathbb{E}[|\underline{\tau}_{n,t}^{(\kappa_1)} - \underline{\tau}_{n,t}^{(\kappa_2)}|^p] \leq 2^{1-n}t^p + 2t^{p-1}\theta(t, \kappa_1) \qquad \text{for any } p \geq 1. \tag{7.24}$$

The bounds in (7.21) and (7.22) of Proposition 7.13 imply the inequalities in (7.15) and (7.17) of Proposition 7.11 (c) with slightly worse constants.

Proof The proof and construction of the random variables is analogous to that of Proposition 7.11 (c) where for $i \in \{1, 2\}$, we compare the increment $Z_s^{(\kappa_i)}$ defined in Algorithm 1 with the Lévy–Itô decomposition $X_s = bs + \sigma W_s + J_s^{1,\kappa_i} + J_s^{2,\kappa_i}$ (W is as in Algorithm 1, independent of J^{1,κ_i} and J^{2,κ_i}) over the time horizons $s \in \{\ell_1, \ldots, \ell_{n-1}\}$. Similarly, we compare the pair of vectors $(\underline{\chi}_s^{(\kappa_1)}, \underline{\chi}_s^{(\kappa_2)})$ produced by Algorithm 2 with $\underline{\chi}_s$ for $s = L_n$, where we assume that the (standardised) Brownian component of X equals that of $\underline{\chi}_s^{(\kappa_2)}$ (and is thus independent of the one in $\underline{\chi}_s^{(\kappa_1)}$) and all jumps in $J^{2,\kappa}$ are synchronously coupled.

Denote the first and fourth components of the vector $(\underline{\chi}_s^{(\kappa_1)}, \underline{\chi}_s^{(\kappa_2)})$ by $Z_s^{(\kappa_1)}$ and $Z_s^{(\kappa_2)}$, respectively. Hence it is enough to obtain the analogous bounds and identities to those presented in parts (a) and (b) for the expectations $\mathbb{E}[(X_t - Z_t^{(\kappa_i)})^2]$, $i \in \{1, 2\}$, under both couplings $\Pi_t^{\kappa_1,\kappa_2}$ and $\underline{\Pi}_t^{\kappa_1,\kappa_2}$. Such bounds may be obtained by using the proofs of parts (a) and (b), resulting in the following estimates: for $i \in \{1, 2\}$, we have

$$\mathbb{E}[(X_t - Z_t^{(\kappa_i)})^2] = \left(\left(\sqrt{\sigma^2 + \overline{\sigma}_{\kappa_i}^2} - \sigma\right)^2 + \overline{\sigma}_{\kappa_i}^2\right)t \leq 2\overline{\sigma}_{\kappa_i}^2 t \qquad \text{under } \Pi_t^{\kappa_1,\kappa_2}, \tag{7.25}$$

$$\mathbb{E}[(X_t - Z_t^{(\kappa_i)})^2] = 2(\sigma^2 \mathbb{1}_{\{1\}}(i) + \overline{\sigma}_{\kappa_1}^2)t \qquad \text{under } \underline{\Pi}_t^{\kappa_1,\kappa_2}. \tag{7.26}$$

Thus Doob's martingale inequality and elementary inequalities give (7.21), (7.22).

By the construction of the law $\underline{\Pi}_{n,t}^{\kappa_1,\kappa_2}$ in SBG-Alg, there exist random variables ξ_k', $k = 1, \ldots, n$, such that for $k \in \{1, \ldots, n\}$, conditionally on $\ell_k = s$ and independently of ℓ_j, $j \neq k$, the distributional equality $(\xi_k', \xi_{1,k}, \xi_{2,k}) \stackrel{d}{=} (X_s, Z_s^{(\kappa_1)}, Z_s^{(\kappa_2)})$ holds, where $(Z_t^{(\kappa_1)}, Z_t^{(\kappa_2)}) \sim \Pi_t^{\kappa_1,\kappa_2}$ and W in Algorithm 1 equals the Brownian component of X in (2.4). Note that by (2.2), we have

$$|\underline{\tau}_t - \underline{\tau}_{n,t}^{(\kappa_i)}| \leq L_n + \sum_{k=1}^{n} \ell_k |\mathbb{1}_{\{\xi_k' < 0\}} - \mathbb{1}_{\{\xi_{i,k} < 0\}}| \qquad \text{for } i \in \{1, 2\}. \tag{7.26}$$

Let $\delta \in (0, 2]$ be as in the statement of the proposition. By Picard [58, Theorem 3.1 (a)], as in the proof of Theorem 3.4, we know that the density f_t of X_t exists, is smooth and given $T > 0$, the constant $C' := 2^{3/2} \sup_{(s,x) \in (0,T] \times \mathbb{R}} s^{1/\delta} f_s(x)$ is finite. Thus (7.10) in Lemma 7.5 (with constants $\gamma = 1$, $C = 2^{-3/2} \ell_k^{-1/\delta} C'$ and $M = 1$, $K = 0$, $p = 1$) gives

$$\mathbb{E}\big[|\mathbb{1}_{\{\xi'_k < 0\}} - \mathbb{1}_{\{\xi_{i,k} < 0\}}| \big| \ell_k\big] \leq \min\big\{1, 2^{-1/4} \sqrt{C'} \ell_k^{-\frac{1}{2\delta}} \mathbb{E}\big[|\xi'_k - \xi_{i,k}| \big| \ell_k\big]^{1/2}\big\}$$

$$\leq \min\{1, 2^{-1/4} \sqrt{C'} \ell_k^{-\frac{1}{2\delta}} (2\overline{\sigma}_{\kappa_i}^2 \ell_k)^{1/4}\},$$

for $i \in \{1, 2\}$ and any $k \in \{1, \ldots, n\}$, where the second inequality follows from Jensen's inequality and (7.25). Hence elementary inequalities, (7.26) and Lemma 7.1 imply that for $i \in \{1, 2\}$,

$$\mathbb{E}[|\underline{\tau}_t - \underline{\tau}_{n,t}^{(\kappa_i)}|] \leq \mathbb{E}[L_n] + \sum_{k=1}^n \mathbb{E}[\ell_k |\mathbb{1}_{\{\xi'_k < 0\}} - \mathbb{1}_{\{\xi_{i,k} < 0\}}|]$$

$$\leq 2^{-n} t + \sum_{k=1}^{\infty} \mathbb{E}[\min\{\sqrt{C' \overline{\sigma}_{\kappa_i}} \ell_k^{\frac{5}{4} - \frac{1}{2\delta}}, \ell_k\}]$$

$$\leq 2^{-n} t + \theta(t, \kappa_i).$$

For $p > 1$, the result follows from the case $p = 1$ and by using the inequality $|\underline{\tau}_t - \underline{\tau}_{n,t}^{(\kappa_i)}|^p \leq t^{p-1} |\underline{\tau}_t - \underline{\tau}_{n,t}^{(\kappa_i)}|$. $\qquad \square$

Proof of Theorem 7.10 (a) Proposition 7.11 (c) and elementary inequalities yield (7.12). So it remains to consider the case $f \in \mathrm{locLip}_K(\mathbb{R}^2)$. As in the proof of Proposition 3.7, by the inequality in (7.8) and the Cauchy–Schwarz inequality, we have

$$\mathbb{E}\big[\big(f(Z_{n,T}^{(\kappa_1)}, \underline{Z}_{n,T}^{(\kappa_1)}) - f(Z_{n,T}^{(\kappa_2)}, \underline{Z}_{n,T}^{(\kappa_2)})\big)^2\big]^2$$

$$\leq K^4 K' \mathbb{E}[(|Z_{n,T}^{(\kappa_1)} - Z_{n,T}^{(\kappa_2)}| + |\underline{Z}_{n,T}^{(\kappa_1)} - \underline{Z}_{n,T}^{(\kappa_2)}|)^4],$$

where $K' := \mathbb{E}[(\exp(Z_{n,T}^{(\kappa_1)}) + \exp(Z_{n,T}^{(\kappa_2)}))^4] \leq 8\mathbb{E}[\exp(4X_T^{(\kappa_1)}) + \exp(4X_T^{(\kappa_2)})]$. Applying (7.7), we get $\mathbb{E}[\exp(4X_T^{(\kappa_i)})] \leq \mathbb{E}[\exp(4X_T)] \exp(4T\overline{\sigma}_{\kappa_i}^2)$ and $\overline{\sigma}_{\kappa_i}^2 \leq \overline{\sigma}_1^2$, $i \in \{1, 2\}$, where $\mathbb{E}[\exp(4X_T)]$ is finite since $\int_{[1,\infty)} e^{4x} \nu(dx) < \infty$. The concavity of $x \mapsto \sqrt{x}$ and the inequalities (7.16) and (7.18) in Proposition 7.11 (c) imply the existence of a constant $C > 0$ satisfying

$$\sqrt{\mathbb{E}[(|Z_{n,T}^{(\kappa_1)} - Z_{n,T}^{(\kappa_2)}| + |\underline{Z}_{n,T}^{(\kappa_1)} - \underline{Z}_{n,T}^{(\kappa_2)}|)^4]}$$

$$\leq C(2/3)^{n/2} + 11T\overline{\sigma}_{\kappa_1}^2 + \sqrt{2\pi} T^{5/8} \overline{\sigma}_{\kappa_1}^{5/4} \kappa_1^{3/4} + \sqrt{5T} \overline{\sigma}_{\kappa_1} \kappa_1.$$

The inequality (7.13) then follows since $\overline{\sigma}_{\kappa_1}^{1/4} \kappa_1^{3/4} \leq \max\{\overline{\sigma}_{\kappa_1}, \kappa_1\} \leq \overline{\sigma}_{\kappa_1} + \kappa_1$.

(b) Let $(\underline{\chi}_T, \underline{\chi}_{n,T}^{(\kappa_1)}, \underline{\chi}_{n,T}^{(\kappa_2)})$ be as in Proposition 7.13, where $\underline{\chi}_T = (X_T, \underline{X}_T, \underline{\tau}_T)$ and $\underline{\chi}_{n,T}^{(\kappa_i)} = (Z_{n,T}^{(\kappa_i)}, \underline{Z}_{n,T}^{(\kappa_i)}, \underline{\tau}_{n,T}^{(\kappa_i)})$, $i \in \{1, 2\}$. The triangle inequality and the inequalities $0 \le f \le M$ give

$$\mathbb{E}\big[\big(f(Z_{n,T}^{(\kappa_1)}, \underline{Z}_{n,T}^{(\kappa_1)}) - f(Z_{n,T}^{(\kappa_2)}, \underline{Z}_{n,T}^{(\kappa_2)})\big)^2\big]$$

$$\le M\mathbb{E}[|f(Z_{n,T}^{(\kappa_1)}, \underline{Z}_{n,T}^{(\kappa_1)}) - f(Z_{n,T}^{(\kappa_2)}, \underline{Z}_{n,T}^{(\kappa_2)})|]$$

$$\le M \sum_{i=1}^{2} \mathbb{E}[|f(Z_{n,T}^{(\kappa_i)}, \underline{Z}_{n,T}^{(\kappa_i)}) - f(X_T, \underline{X}_T)|].$$

Apply (7.10) in Lemma 7.5 with C and γ from Assumption 3.8 to $(X_T, \underline{X}_T)$ and $(Z_{n,T}^{(\kappa_i)}, \underline{Z}_{n,T}^{(\kappa_i)})$ to get

$$\mathbb{E}[|f(X_T, \underline{X}_T) - f(Z_{n,T}^{(\kappa_i)}, \underline{Z}_{n,T}^{(\kappa_i)})|]$$

$$\le K\mathbb{E}[|Z_{n,T}^{(\kappa_i)} - X_T|] + M(1 + 2/\gamma)(C^2\gamma^2\mathbb{E}[|\underline{Z}_{n,T}^{(\kappa_1)} - \underline{X}_T|^2]^\gamma)^{\frac{1}{2+\gamma}}$$

$$\le K\sqrt{T\big(4\sigma^2 2^{-n}\mathbb{1}_{\{1\}}(i) + 2\overline{\sigma}_{\kappa_i}^2\big)} + K''\big(\sigma^2 2^{-n}\mathbb{1}_{\{1\}}(i) + \overline{\sigma}_{\kappa_i}^2\big)^{\frac{\gamma}{2+\gamma}}$$

for $i \in \{1, 2\}$, where $K'' := M(1 + 2/\gamma)(48C^2\gamma^2 T^\gamma)^{1/(2+\gamma)}$. In the second inequality, we used the bounds (7.21) and (7.22). Since $\overline{\sigma}_{\kappa_1} \ge \overline{\sigma}_{\kappa_2}$, the result follows.

(c) Recall that (7.24) follows from (7.23). The inequality (7.14) in the proposition is a direct consequence of the Lipschitz property and (7.24).

(d) The proof follows along the same lines as in part (b): apply (7.10) in Lemma 7.5 with C and γ from Assumption 3.11 and the bounds (7.21)–(7.23) from Proposition 7.13. $\qquad\square$

7.5 Computational complexity of the MC and MLMC estimators

In this subsection, we address the application of our previous results to estimate the expectation $\mathbb{E}[f(\underline{\chi}_T)]$ for various real-valued functions f satisfying $\mathbb{E}[f(\underline{\chi}_T)^2] < \infty$. By definition, an estimator Υ of $\mathbb{E}[f(\underline{\chi}_T)]$ has L^2-accuracy of level $\epsilon > 0$ if it satisfies $\mathbb{E}[(\Upsilon - \mathbb{E}[f(\underline{\chi}_T)])^2] < \epsilon^2$. We assume in this subsection that X has jumps of infinite activity, i.e., $\nu(\mathbb{R} \setminus \{0\}) = \infty$. If the jumps of X are of finite activity, both Algorithm 2 and SBG-Alg are exact with the latter outperforming the former in practice by a constant factor, which is a function of the total number of jumps, $T\nu(\mathbb{R} \setminus \{0\}) < \infty$; see Sect. 6.2 for a numerical example.

7.5.1 MC complexity

As above, for any $\kappa \in (0, 1]$, let the sequence $\underline{\chi}_T^{\kappa,i}$, $i \in \mathbb{N}$, be i.i.d. with the same distribution as $\underline{\chi}_T^{(\kappa)}$ simulated by SBG-Alg with $n \in \mathbb{N} \cup \{0\}$ sticks. Recall that the

MC estimator in (4.2), based on $N \in \mathbb{N}$ independent samples, is given by

$$\Upsilon_{\mathrm{MC}} = \frac{1}{N} \sum_{i=1}^{N} f(\underline{\chi}_T^{\kappa,i}).$$

The requirements on the bias and variance of the estimator Υ_{MC} (see Appendix A.1), together with Theorem 3.4 and the bounds in (3.11) as well as Propositions 3.7, 3.9 and 3.12, imply Corollary 7.14. By expressing κ in terms of ϵ via Corollary 7.14 and (3.5), (3.8), (3.9), the formulae for the expected computational complexity $\mathcal{C}_{\mathrm{MC}}(\epsilon)$ in Table 1 (see Sect. 4.2) follow.

Corollary 7.14 *For any $\epsilon \in (0, 1)$, define first κ for given f as in one of* (a)–(d) *below and then $N := \lceil 2\epsilon^{-2}\mathbb{V}[f(\underline{\chi}_T^{(\kappa)})] \rceil$ as in Appendix A.1. Then the MC estimator Υ_{MC} of $\mathbb{E}[f(\underline{\chi}_T)]$ has L^2-accuracy of level ϵ and expected cost $\mathcal{C}_{\mathrm{MC}}(\epsilon)$ bounded by a constant multiple of $(1 + \overline{\nu}(\kappa)T)N$.*

(a) *For any $K > 0$, $g \in \mathrm{Lip}_K(\mathbb{R}^2)$ (resp. $g \in \mathrm{locLip}_K(\mathbb{R}^2)$) and the function $f : (x, z, t) \mapsto g(x, z)$, set*

$$\kappa := \sup\{\kappa' \in (0, 1] : 2\mu_1(\kappa', T) < \epsilon/\sqrt{2}\}$$

$$\left(\text{resp. } \kappa := \sup\left\{\kappa' \in (0, 1] : 8K^2\mu_2(\kappa', T)\left(1 + \exp(2T\overline{\sigma}_{\kappa'}^2)\right)\right.\right.$$
$$\left.\left. \times \mathbb{E}[\exp(2X_T)] < \epsilon^2/2\right\}\right).$$

(b) *Pick $y < 0$ and let Assumption 3.8 hold for some $C, \gamma > 0$. Suppose that $f : \mathbb{R}^3 \to \mathbb{R}$ is given by $f(x, z, t) = h(x)\mathbb{1}_{[y,\infty)}(z)$, where $h \in \mathrm{Lip}_K(\mathbb{R})$ and $0 \le h \le M$ for some $K, M > 0$. Then*

$$\kappa := \sup\left\{\kappa' \in (0, 1] : M(C\gamma)^{2/(2+\gamma)}\left(1 + \frac{2}{\gamma}\right)\mu_2(\kappa', T)^{2\gamma/(2+\gamma)}\right.$$
$$\left. + K\mu_1(\kappa', T) < \epsilon/\sqrt{2}\right\}.$$

(c) *Let $\delta \in (0, 2]$ satisfy Assumption 3.2. Let $f(x, z, t) = g(t)$, $g \in \mathrm{Lip}_K(\mathbb{R})$, $K > 0$. Then*

$$\kappa := \sup\{\kappa' \in (0, 1] : K\mu_*^\tau(\kappa', T) < \epsilon/\sqrt{2}\}.$$

(d) *Fix $s \in (0, T)$ and let $\delta \in (0, 2]$ satisfy Assumption 3.2. Then there exists a constant $C > 0$ such that for $f \in \mathrm{BT}_2(s, K, M)$, $K, M > 0$, we have*

$$\kappa := \sup\{\kappa' \in (0, 1] : C\sqrt{K\mu_*^\tau(\kappa', T)} < \epsilon/\sqrt{2}\}.$$

The constant C in Corollary 7.14 (d) depends on a value of the density of $\underline{\tau}_t$ at s which is not known a priori. The purpose of Corollary 7.14 (d) is to specify the dependence of κ on ϵ up to the constant C for all considered payoffs. In practice, κ is determined adaptively from some initial preliminary samples, and Corollary 7.14 (d) only serves as a theoretical guarantee (see e.g. Giles [29, Sect. 2.1]).

7.5.2 MLMC complexity

Let $(\kappa_j)_{j\in\mathbb{N}}$ (resp. $(n_j)_{j\in\mathbb{N}\cup\{0\}}$) be a decreasing (resp. increasing) sequence in $(0, 1]$ (resp. $\mathbb{N}$) such that $\lim_{j\to\infty}\kappa_j = 0$. Let $\underline{\chi}^{0,i} \stackrel{d}{=} \underline{\chi}_T^{(\kappa_1)}$ and $(\underline{\chi}_1^{j,i}, \underline{\chi}_2^{j,i}) \sim \underline{\Pi}_{n_j,T}^{\kappa_j,\kappa_{j+1}}$, $i, j \in \mathbb{N}$, be independent draws constructed by SBG-Alg. As in Sect. 4.2.2, recall that the sequence (n_j) appears as a parameter in the coupling $\underline{\Pi}_{n_j,T}^{\kappa_j,\kappa_{j+1}}$ (which is the law that the pair of vectors $(\underline{\chi}_1^{j,i}, \underline{\chi}_2^{j,i})$ follow). The number n_j specifies the number n of sticks used in SBG-Alg for the level j. Recall that for the parameters $m, N_0, \ldots, N_m \in \mathbb{N}$, the MLMC estimator in (4.3) takes the form

$$\Upsilon_{\mathrm{ML}} = \sum_{j=0}^{m} \frac{1}{N_j} \sum_{i=1}^{N_j} D_j^i, \qquad \text{where } D_j^i := \begin{cases} f(\underline{\chi}_2^{j,i}) - f(\underline{\chi}_1^{j,i}), & j \geq 1, \\ f(\underline{\chi}^{0,i}), & j = 0. \end{cases}$$

The bias of the MLMC estimator is equal to that of the MC estimator in (4.2) with $\kappa = \kappa_m$. Given the sequences $(n_j)_{j\in\mathbb{N}\cup\{0\}}$ and $(\kappa_j)_{j\in\mathbb{N}}$ which determine the simulation algorithms used in the estimator (4.3), Appendix A.2 derives the asymptotically optimal (as $\epsilon \searrow 0$) values for the integers m and $(N_j)_{j=0}^{m}$ minimising the expected computational complexity of (4.3) under the constraint that the L^2-accuracy of Υ_{ML} is of level ϵ. The key quantities are the bounds $B(j)$, $V(j)$ and $C(j)$ on the bias, level variance and the computational complexity of SBG-Alg at level j (i.e., run with parameters κ_j and n_j). The number m of levels in (4.3) is determined by the bound on the bias $B(j)$, while the number N_j of samples used at level j is given by the bounds on the complexity and level variances; see the formulas in (A.1), (A.2) below. Proposition 7.15, which is a consequence of Theorem 3.4 and Propositions 3.7, 3.9 and 3.12 (for the bias), Theorem 7.10 (for the level variance) and Corollary 4.2 (for the complexity), summarises the relevant bounds $B(j)$, $V(j)$ and $C(j)$ established in this paper (suppressing the unknown constants as we are only interested in the asymptotic behaviour as $\epsilon \searrow 0$).

Proposition 7.15 *Given sequences $(\kappa_j)_{j\in\mathbb{N}}$ and $(n_j)_{j\in\mathbb{N}\cup\{0\}}$ as above, we define $C(j) := n_j + \overline{\nu}(\kappa_{j+1})T$. Then the following choices of functions B and V ensure that for any $\epsilon > 0$, the MLMC estimator Υ_{ML}, with integers m and N_j, $j = 0, \ldots, m$, given by (A.1)–(A.2), has L^2-accuracy of level ϵ with complexity asymptotically proportional to $\mathcal{C}_{\mathrm{ML}}(\epsilon) = 2\epsilon^{-2}(\sum_{j=0}^{m} \sqrt{C(j)V(j)})^2$.*

(a) If $K > 0$, $g \in \mathrm{Lip}_K(\mathbb{R}^2)$ (resp. $g \in \mathrm{locLip}_K(\mathbb{R}^2)$) and $f(x, z, t) = g(x, z)$, then for any $j \in \mathbb{N}$,

$$B(j) := \mu_1(\kappa_j, T) \quad \text{and} \quad V(j) := \sigma^2 2^{-n_j} + \overline{\sigma}_{\kappa_j}^2$$

$$\left(\text{resp. } B(j) := \mu_2(\kappa_j, T) \quad \text{and} \quad V(j) := (2/3)^{n_j/2}\mathbb{1}_{\mathbb{R}\setminus\{0\}}(\sigma) + \overline{\sigma}_{\kappa_j}^2 + \overline{\sigma}_{\kappa_j}\kappa_j\right).$$

(b) Pick $y < 0$ and suppose that Assumption 3.8 holds for some $C, \gamma > 0$. If $f \in \mathrm{BT}_1(y, K, M)$, $K, M > 0$, then for any $j \in \mathbb{N}$,

$$B(j) := \min\left\{\mu_1(\kappa_j, T)^{\frac{\gamma}{1+\gamma}}, \mu_2(\kappa_j, T)^{\frac{2\gamma}{2+\gamma}}\right\} \quad \text{and} \quad V(j) := \sigma^{\frac{2\gamma}{2+\gamma}} 2^{-\frac{n_j\gamma}{2+\gamma}} + \overline{\sigma}_{\kappa_j}^{\frac{2\gamma}{2+\gamma}}.$$

(c) *Let Assumption 3.2 hold for some $\delta \in (0, 2]$ and $f(x, z, t) = g(t)$ for some $g \in \mathrm{Lip}_K(\mathbb{R})$, $K > 0$. Then for any $j \in \mathbb{N}$,*

$$B(j) := \mu_*^\tau(\kappa_j, T) \quad and \quad V(j) := 2^{-n_j} + \overline{\sigma}_{\kappa_j}^{\min\{1/2, 2\delta/(2-\delta)\}} \left(1 + |\log \kappa_j| \mathbb{1}_{\{2/5\}}(\delta)\right).$$

(d) *Let $f \in \mathrm{BT}_2(s, K, M)$ for some $s \in (0, T)$ and $K, M \geq 0$. If $\delta \in (0, 2]$ satisfies Assumption 3.2, then for any $j \in \mathbb{N}$,*

$$B(j) := \sqrt{\mu_*^\tau(\kappa_j, T)} \quad and \quad V(j) := 2^{-n_j/2} + \overline{\sigma}_{\kappa_j}^{\min\{\frac{1}{4}, \frac{\delta}{2-\delta}\}} \left(1 + \sqrt{|\log \kappa_j|} \mathbb{1}_{\{2/5\}}(\delta)\right).$$

Remark 7.16 By (3.5) and (A.2), we note that κ_m in Proposition 7.15 (a) is bounded by (and typically proportional to) $C_0 \epsilon / |\log \epsilon|$. Moreover, if $\kappa_m = e^{-r(m-1)}$ for some $r > 0$, then the constant C_0 does not depend on the rate r. A similar statement holds for (b), (c) and (d); see Table 1 in Sect. 4.2.

It remains to choose the parameters $(n_j)_{j \in \mathbb{N} \cup \{0\}}$ and $(\kappa_j)_{j \in \mathbb{N}}$ for the estimator in (4.3). Since we require the bias to vanish geometrically fast, we set $\kappa_j = e^{-r(j-1)}$ for $j \in \mathbb{N}$ and some $r > 0$. The value of the rate r in Theorem 7.17 below is obtained by minimising the multiplicative constant in the complexity $\mathcal{C}_{\mathrm{ML}}(\epsilon)$. Note that n_j does not affect the bias (nor the bound $B(j)$) of Υ_{ML}. By Proposition 7.15, n_j may be as small as a multiple of $\log(1/\overline{\sigma}_{\kappa_j}^2)$ without affecting the asymptotic behaviour of the level variances $V(j)$, and as large as $\overline{\nu}(\kappa_{j+1})$ without increasing the asymptotic behaviour of the cost $C(j)$ of each level. Moreover, to ensure that the term $\sigma^2 2^{-n_j}$ in the level variances (see Theorem 7.10) decays geometrically, it suffices to let n_j grow at least linearly in j. In short, there is a large interval within which we may choose n_j without it having any effect on the asymptotic performance of the MLMC estimation (see Theorem 7.17 below). The choice $n_j = n_0 + \lceil \max\{j, \log^2(1 + \overline{\nu}(\kappa_{j+1})T)\} \rceil$ for $j \in \mathbb{N}$ in the numerical examples of Sect. 6 falls within this interval.

Theorem 7.17 *Suppose $q \in (0, 2]$ and $c > 0$ satisfy $\overline{\nu}(\kappa) \leq c\kappa^{-q}$ and $\overline{\sigma}_\kappa^2 \leq c\kappa^{2-q}$ for all $\kappa \in (0, 1]$. Pick $r > 0$, set $\kappa_j := e^{-r(j-1)}$ and assume that for some $C > 0$ and all sufficiently large $j \in \mathbb{N}$, we have $\max\{j, \log_{2/3}(\overline{\sigma}_{\kappa_j}^4)\} \leq n_j \leq C\overline{\nu}(\kappa_{j+1})$. Then in the cases (a)–(d) below, there exists a constant C_r such that for $\epsilon \in (0, 1)$, the MLMC estimator Υ_{ML} defined in (4.3), with parameters given by (A.1), (A.2) below, is L^2-accurate at level ϵ with the stated expected computational complexity $\mathcal{C}_{\mathrm{ML}}(\epsilon)$. Moreover, C_r is minimal for $r := (2/|a|) \log(1 + |a|/q) \mathbb{1}_{\mathbb{R} \setminus \{0\}}(a) + (2/q) \mathbb{1}_{\{0\}}(a)$, with $a \in \mathbb{R}$ given explicitly in each case (a)–(d).*

(a) *Let $g \in \mathrm{Lip}_K(\mathbb{R}^2) \cup \mathrm{locLip}_K(\mathbb{R}^2)$ for $K > 0$ and $f(x, z, t) = g(x, z)$. Define*

$$b := \mathbb{1}_{\{\sigma=0\}} + \mathbb{1}_{\{\sigma \neq 0\}} \left(\mathbb{1}_{\{g \in \mathrm{Lip}_K(\mathbb{R}^2)\}} \frac{1}{3-q} + \mathbb{1}_{\{g \notin \mathrm{Lip}_K(\mathbb{R}^2)\}} \frac{2}{4-q}\right)$$

and $a := 2(q - 1)$. Then

$$\mathcal{C}_{\mathrm{ML}}(\epsilon) \leq \frac{C_r}{\epsilon^{2+a+b}} \left(1 + (\log^2 \epsilon) \mathbb{1}_{\{1\}}(q) + |\log \epsilon|^{(a/2)(1+\mathbb{1}_{\{g \in \mathrm{Lip}_K(\mathbb{R}^2)\}})} \mathbb{1}_{(1,2]}(q)\right).$$

(b) *Let* $f(x, z, t) = g(x, z)$, *where* $g \in \mathrm{BT}_1(y, K, M)$ *for some* $y < 0 \le K, M$, *such that Assumption* 3.8 *is satisfied by* y *and some* $C, \gamma > 0$. *Define*

$$b := \left(\frac{1}{2} + \frac{1}{\gamma}\right)\left(\mathbb{1}_{\{\sigma=0\}} + \mathbb{1}_{\{\sigma\neq0,q<1\}}\frac{4}{9-3q} + \mathbb{1}_{\{\sigma\neq0,q\geq1\}}\frac{2}{4-q}\right)$$

and $a := 2\frac{q(1+\gamma)-\gamma}{2+\gamma} \in (-\frac{2\gamma}{2+\gamma}, 2]$. *Then*

$$\mathcal{C}_{\mathrm{ML}}(\epsilon) \le \frac{C_r}{\epsilon^{2+a+b}}\big(1 + (\log^2 \epsilon)\mathbb{1}_{\{q=\gamma/(1+\gamma)\}}$$
$$+ |\log\epsilon|^a \mathbb{1}_{\{q\in(\gamma/(1+\gamma),1)\}} + |\log\epsilon|^{a/2}\mathbb{1}_{\{q\in[1,2]\}}\big).$$

(c) *Let* $f(x, z, t) = g(t)$, $g \in \mathrm{Lip}_K(\mathbb{R})$, $K > 0$, *and let Assumption* 3.2 *hold for some* $\delta \in (0, 2]$. *Set* $a := q - (1 - \frac{q}{2})\min\{\frac{1}{2}, \frac{2\delta}{2-\delta}\}$ *and* $b := \min\{2/\delta, \max\{3/2, 1/\delta\}\}$. *Then*

$$\mathcal{C}_{\mathrm{ML}}(\epsilon) \le \frac{C_r}{\epsilon^{2+a+b}}\begin{cases} 1 + (\log^2 \epsilon)(\mathbb{1}_{\{\delta<2/5,q=\delta\}} + \mathbb{1}_{\{\delta=2,q=2/5\}}), & \delta \notin \{\frac{2}{5}, \frac{2}{3}\}, \\ |\log\epsilon|\mathbb{1}_{\{q\in(2/5,2]\}} + |\log\epsilon|^3\mathbb{1}_{\{q=2/5\}}, & \delta = 2/5, \\ |\log\epsilon|^a, & \delta = 2/3. \end{cases}$$

(d) *Fix* $s \in (0, T)$ *and let* $\delta \in (0, 2]$ *satisfy Assumption* 3.2. *Define the constants* $a := q - (1 - \frac{q}{2})\min\{\frac{1}{4}, \frac{\delta}{2-\delta}\}$ *and* $b := \min\{4/\delta, \max\{3, 2/\delta\}\}$. *Then for* $K, M \ge 0$ *and* $f \in \mathrm{BT}_2(s, K, M)$,

$$\mathcal{C}_{\mathrm{ML}}(\epsilon) \le \frac{C_r}{\epsilon^{2+a+b}}\begin{cases} 1 + (\log^2 \epsilon)\mathbb{1}_{\{q=2/9\}}, & \delta = 2, \\ 1 + \sqrt{|\log\epsilon|}\mathbb{1}_{\{\delta=2/5\}} + |\log\epsilon|^{\frac{a}{2}}\mathbb{1}_{\{\delta=2/3\}}, & \delta \in (0, 2). \end{cases}$$

Remark 7.18 For most models, either $\beta = \delta$ or $\sigma > 0$, implying $a+b \in [0, 2]$ in parts (a) and (c), $a+b \in [0, 2(1/2 + 1/\gamma)]$ in part (b) (with γ typically equal to 1) and $a+b \in [0, 4]$ in part (d).

Proof of Theorem 7.17 Note that $\kappa_1 = 1$ by definition independently of $r > 0$, thus making both the variance $\mathbb{V}[D_0^i]$ and the cost of sampling of D_0^i independent of r. We may thus ignore the 0th term in the bound $\epsilon^{-2}(\sum_{j=0}^m \sqrt{V(j)C(j)})^2$ on the complexity $\mathcal{C}_{\mathrm{ML}}(\epsilon)$ derived in Appendix A.2. Since m is given by (A.1) below, by Table 1 and Remark 7.16, the function $\overline{m} : (0, 1) \to (0, \infty)$ given by

$$\overline{m}(\epsilon) := \begin{cases} (b|\log\epsilon| + c\log|\log\epsilon|)/r & \text{in (a) and (b) and,} \\ & \text{if } \delta = \frac{2}{3}, \text{ in (c) and (d),} \\ b|\log\epsilon|/r & \text{in (c) and (d) if } \delta \neq \frac{2}{3}, \end{cases} \qquad (7.27)$$

where

$$c = \begin{cases} 1 & \text{in (a) and (c),} \\ 1/2 & \text{in (b) and (d),} \end{cases}$$

satisfies $m \leq \overline{m}(\epsilon) + C'/r$ for all $\epsilon \in (0, 1)$ and $r > 0$, where the constant $C' > 0$ is independent of $r > 0$. Thus we need only study the growth rate of

$$\phi(\epsilon) := \sum_{j=1}^{\lceil \overline{m}(\epsilon) \rceil} \sqrt{C(j)V(j)} = \sum_{j=1}^{\lceil \overline{m}(\epsilon) \rceil} \sqrt{\left(n_j + \overline{v}(\kappa_{j+1})T\right)V(j)} \qquad \text{as } \epsilon \to 0,$$

because $\mathcal{C}_{\mathrm{ML}}(\epsilon)$ is bounded by a constant multiple of $\epsilon^{-2}\phi(\epsilon)^2$. In the cases where $V(j)$ contains a term of the form e^{-sn_j} for some $s > 0$ (only possible if $\sigma \neq 0$), the product $n_j e^{-sn_j} \leq e^{-sn_j/2}$ vanishes geometrically fast since $n_j \geq j$ for all large j. Thus the corresponding component in $\phi(\epsilon)$ is bounded as $\epsilon \to 0$ and may thus be ignored. By Proposition 7.15, we may assume in all cases that $V(j)$ is bounded by a multiple of a power of $\overline{\sigma}_{\kappa_j}^2$ and $C(j)$ is dominated by a multiple of $\overline{v}(\kappa_{j+1})$.

Since $\overline{v}(\kappa) \leq c\kappa^{-q}$ and $\overline{\sigma}_\kappa^2 \leq c\kappa^{2-q}$ for $\kappa \in (0, 1]$, Proposition 7.15 yields

$$\phi(\epsilon) \leq K_* \begin{cases} \sum_{j=1}^{\lceil \overline{m}(\epsilon) \rceil} (\kappa_{j+1}^{-q}\kappa_j^{2-q})^{\frac{1}{2}} & \text{in (a),} \\[2mm] \sum_{j=1}^{\lceil \overline{m}(\epsilon) \rceil} (\kappa_{j+1}^{-q}\kappa_j^{(2-q)\gamma/(2+\gamma)})^{\frac{1}{2}} & \text{in (b),} \\[2mm] \sum_{j=1}^{\lceil \overline{m}(\epsilon) \rceil} (\kappa_{j+1}^{-q}\kappa_j^{(2-q)\min\{\frac{1}{2}, \frac{2\delta}{2-\delta}\}})(1 + |\log \kappa_j|\mathbb{1}_{\{\delta=2/5\}}))^{\frac{1}{2}} & \text{in (c),} \\[2mm] \sum_{j=1}^{\lceil \overline{m}(\epsilon) \rceil} (\kappa_{j+1}^{-q}\kappa_j^{(2-q)\min\{\frac{1}{4}, \frac{\delta}{2-\delta}\}})(1 + \sqrt{|\log \kappa_j|}\mathbb{1}_{\{\delta=2/5\}}))^{\frac{1}{2}} & \text{in (d),} \end{cases}$$

for some constant $K_* > 0$ independent of r and all $\epsilon \in (0, 1)$, where we used in part (a) the fact that $\overline{\sigma}_\kappa\kappa \leq \sqrt{c}\kappa^{2-q/2}$ for all $\kappa \in (0, 1]$.

(a) Recall that $\kappa_j = e^{-r(j-1)}$ and $\kappa_{j+1} = e^{-r(j-1)-r}$, implying

$$\kappa_{j+1}^{-q}\kappa_j^{2-q} = e^{rq}e^{ar(j-1)} \qquad \text{for all } j \in \mathbb{N}, \text{ where } a = 2(q-1). \qquad (7.28)$$

Suppose first $a < 0$, implying $q \in (0, 1)$. By (7.28), the sequence $(\kappa_{j+1}^{-q}\kappa_j^{2-q})_{j \in \mathbb{N}}$ decays geometrically fast. This implies that $\lim_{\epsilon \downarrow 0} \phi(\epsilon) < \infty$ and gives the desired result. Moreover, the leading constant C_r, as a function of r, is proportional to $e^{rq}/(1 - e^{ar/2})^2$ as $\epsilon \downarrow 0$. Since $a \neq 0$ for $q \in (0, 1)$, the minimal value of C_r is attained when $r = (2/|a|)\log(1 + |a|/q)$.

If $a = 0$, hence $q = 1$, we have $\phi(\epsilon) \leq K_* e^{r/2}(b|\log \epsilon| + \log|\log \epsilon|)/r$ by (7.28) and (7.27), giving the desired result. As before, the leading constant C_r, as a function of r, is proportional to e^r/r^2 as $\epsilon \to 0$, attaining its minimum at $r = 2$.

Finally, suppose $a > 0$, implying $q \in (1, 2]$. By (7.28) and (7.27), it follows that

$$\phi(\epsilon)^2 \leq \frac{K_*^2 e^{rq}}{(e^{ar/2} - 1)^2}e^{a(b|\log \epsilon|+\log|\log \epsilon|)} = \frac{K_*^2 e^{rq}}{(e^{ar/2} - 1)^2}\epsilon^{-ab}|\log \epsilon|^a.$$

The corresponding result follows easily, where the leading constant C_r, as a function of r, is proportional to $e^{rq}/(e^{ar/2} - 1)^2$ as $\epsilon \downarrow 0$ and attains its minimum at $r = (2/a)\log(1 + a/q)$, concluding the proof of (a).

(b) As before, we have

$$\kappa_{j+1}^{-q}\kappa_j^{(2-q)\gamma/(2+\gamma)} = e^{rq}e^{ar(j-1)} \qquad \text{for } j \in \mathbb{N}, \text{ with } a = 2\frac{q(1+\gamma) - \gamma}{2+\gamma}. \qquad (7.29)$$

Suppose $a < 0$, implying $q < \gamma/(1+\gamma)$. Then $\lim_{\epsilon\downarrow 0}\phi(\epsilon) < \infty$ by (7.29), implying the claim. Moreover, C_r is minimal for $r = (2/|a|)\log(1 + |a|/q)$ as in part (a).

Now suppose $a = 0$, implying $q = \gamma/(1+\gamma)$. Then

$$\phi(\epsilon)^2 \leq K_*^2 r^{-2} e^{rq} (b|\log\epsilon| + \log|\log\epsilon|/2)^2,$$

and the leading constant is minimised when $r = 2/q = 2 + 2/\gamma$.

Finally, suppose $a > 0$, implying $q > \gamma/(1+\gamma)$. By (7.29), we have

$$\phi(\epsilon)^2 \leq \frac{K_*^2 e^{rq}}{(e^{ar/2}-1)^2} e^{a(b|\log\epsilon|+\log|\log\epsilon|/2)} = \frac{K_*^2 e^{rq}}{(e^{ar/2}-1)^2}\epsilon^{-ab}|\log\epsilon|^{a/2},$$

and the leading constant is minimal for $r = (2/a)\log(1 + a/q)$.

In parts (c) and (d), note that $a < 0$ if and only if $\delta = 2$ (i.e., $\sigma \neq 0$). Analogous arguments as in (a) and (b) complete the proof of the theorem. $\qquad\square$

Appendix: MC and MLMC estimators

A.1 Monte Carlo estimator

Consider square-integrable random variables $P, P_1, P_2, \ldots$ Let $(P_k^i)_{k,i\in\mathbb{N}}$ be independent with $P_k^i \overset{d}{=} P_k$ for $k, i \in \mathbb{N}$. Suppose $|\mathbb{E}[P] - \mathbb{E}[P_k]| \leq B(k)$ for all $k \in \mathbb{N}$ and assume that $C(n)$ bounds the expected computational cost of simulating a single value of P_n. Pick arbitrary $\epsilon > 0$ and define $m := \inf\{k \in \mathbb{N} : B(k) < \epsilon/\sqrt{2}\}$, $N := \lceil 2\mathbb{V}[P_m]/\epsilon^2\rceil$. Then the Monte Carlo estimator

$$\hat{P} := \frac{1}{N}\sum_{i=1}^{N} P_m^i \quad \text{of } \mathbb{E}[P] \text{ is } L^2\text{-accurate at level } \epsilon, \text{ i.e., } \mathbb{E}\big[(\hat{P} - \mathbb{E}[P])^2\big]^{1/2} < \epsilon,$$

since $\mathbb{E}[(\hat{P} - \mathbb{E}[P])^2] = \mathbb{V}[\hat{P}] + (\mathbb{E}[P_m] - \mathbb{E}[P])^2$ and $\mathbb{V}[\hat{P}] < \epsilon^2/2$ by the definition of N, while $(\mathbb{E}[P_m] - \mathbb{E}[P])^2 < \epsilon^2/2$ by the definition of m. Thus if the bound $B(m)$ on the bias is asymptotically sharp, the formulae for the integers $m, N \in \mathbb{N}$ above result in the computational complexity given by the formula

$$\mathcal{C}_{\mathrm{MC}}(\epsilon) = NC(m) = \lceil 2\mathbb{V}[P_m]/\epsilon^2\rceil C(m).$$

Although in practice one does not have access to the variance $\mathbb{V}[P_m]$, it is typically close to $\mathbb{V}[P]$ (which often has an a priori bound) or can be estimated via simulation.

A.2 Multilevel Monte Carlo estimator

This section is based on the works of Heinrich and Giles [37] and Giles [28]. Let $P, P_1, P_2, \ldots$ be square-integrable random variables and set $P_0 := 0$. Suppose that the array of independent random variables $(D_k^i)_{k\in\mathbb{N}\cup\{0\}, i\in\mathbb{N}}$ satisfies $D_k^i \overset{d}{=} D_k^1$ and $\mathbb{E}[D_k^i] = \mathbb{E}[P_{k+1} - P_k]$ for $k \in \mathbb{N}\cup\{0\}$ and $i \in \mathbb{N}$. For $k \in \mathbb{N}\cup\{0\}$, assume

that the bias and level variance satisfy the inequalities $B(k) \geq |\mathbb{E}[P] - \mathbb{E}[P_k]|$ and $V(k) \geq \mathbb{V}[D_k^1]$ for some functions $k \mapsto B(k)$ and $k \mapsto V(k)$, respectively, and let $C(k)$ bound the expected computational complexity of simulating a single value of D_k^1. For $m \in \mathbb{N} \cup \{0\}$ and $N_0, \ldots, N_m \in \mathbb{N}$, the MLMC estimator

$$\hat{P} := \sum_{k=0}^{m} \frac{1}{N_k} \sum_{i=1}^{N_k} D_k^i$$

satisfies $\mathbb{E}[(\hat{P} - \mathbb{E}[P])^2] = \mathbb{V}[\hat{P}] + (\mathbb{E}[P_m] - \mathbb{E}[P])^2$, since $\mathbb{E}[\hat{P}] = \mathbb{E}[P_m]$. Thus for any $\epsilon > 0$, the inequality $\mathbb{E}[(\hat{P} - \mathbb{E}[P])^2] < \epsilon^2$ holds if the number of levels in $\hat{P}$ equals

$$m := \inf\{k \in \mathbb{N} \cup \{0\} : B(k) < \epsilon/\sqrt{2}\}, \tag{A.1}$$

and the variance is bounded by $\mathbb{V}[\hat{P}] = \sum_{k=0}^{m} \mathbb{V}[D_k^1]/N_k \leq \sum_{k=0}^{m} V(k)/N_k \leq \epsilon^2/2$. Since the computational complexity of $\hat{P}$, $\mathcal{C}_{\mathrm{ML}}(\epsilon) = \sum_{k=0}^{m} C(k)N_k$, is linear in the number of samples N_k on each level k, we only require that the variance $\mathbb{V}[\hat{P}]$ be of the same order as $\epsilon^2/2 = \sum_{k=0}^{m} V(k)/N_k$. Then by the Cauchy–Schwarz inequality, we have

$$\mathcal{C}_{\mathrm{ML}}(\epsilon)\epsilon^2/2 = \left(\sum_{k=1}^{m} C(k)N_k\right)\left(\sum_{k=0}^{m} \frac{V(k)}{N_k}\right) \geq \left(\sum_{k=0}^{m} \sqrt{C(k)V(k)}\right)^2,$$

where the lower bound does not depend on $N_0, \ldots, N_m$ and is attained if and only if

$$N_k := \left\lceil 2\epsilon^{-2} \sqrt{\frac{V(k)}{C(k)}} \sum_{j=0}^{m} \sqrt{C(j)V(j)} \right\rceil \qquad \text{for } k \in \{0, \ldots, n\}, \tag{A.2}$$

ensuring that the expected cost is a multiple of

$$\mathcal{C}_{\mathrm{ML}}(\epsilon) = 2\epsilon^{-2}\left(\sum_{k=0}^{m} \sqrt{C(k)V(k)}\right)^2. \tag{A.3}$$

Moreover, if B, V and C are asymptotically sharp, the formulae in (A.2), up to constants, minimise the expected computational complexity. Consequently, the computational complexity analysis of the MLMC estimator is reduced to the analysis of the behaviour of $\sum_{j=0}^{m} \sqrt{C(j)V(j)}$ as $\epsilon \downarrow 0$.

References

1. Andersen, L., Lipton, A.: Asymptotics for exponential Lévy processes and their volatility smile: survey and new results. Int. J. Theor. Appl. Finance **16**, 1350001 (2013)
2. Asmussen, S., Rosiński, J.: Approximations of small jumps of Lévy processes with a view towards simulation. J. Appl. Probab. **38**, 482–493 (2001)
3. Avram, F., Chan, T., Usabel, M.: 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 (2002)
4. Bang, D., González Cázares, J.I., Mijatović, A.: When is the convex hull of a Lévy path smooth? Preprint (2022). Available online at arXiv:2205.14416
5. Baurdoux, E., Palmowski, Z., Pistorius, M.: On future drawdowns of Lévy processes. Stoch. Process. Appl. **127**, 2679–2698 (2017)
6. Ben Abdellah, A., L'Ecuyer, P., Owen, A.B., Puchhammer, F.: Density estimation by randomized quasi-Monte Carlo. SIAM/ASA J. Uncertain. Quantificat. **9**, 280–301 (2021)
7. Bertoin, J.: Lévy Processes. Cambridge University Press, Cambridge (1996)
8. Bisewski, K., Ivanovs, J.: Zooming-in on a Lévy process: failure to observe threshold exceedance over a dense grid. Electron. J. Probab. **25**, 1–33 (2020)
9. Blumenthal, R.M., Getoor, R.K.: Sample functions of stochastic processes with stationary independent increments. J. Math. Mech. **10**, 493–516 (1961)
10. Bobkov, S., Ledoux, M.: One-dimensional empirical measures, order statistics and Kantorovich transport distances. Mem. Am. Math. Soc. **261**, 1–126 (2019)
11. Cai, C., Guo, J., You, H.: Non-parametric estimation for a pure-jump Lévy process. Ann. Actuar. Sci. **12**, 338–349 (2018)
12. Carpentier, A., Duval, C., Mariucci, E.: Total variation distance for discretely observed Lévy processes: a Gaussian approximation of the small jumps. Ann. Inst. Henri Poincaré Probab. Stat. **57**, 901–939 (2021)
13. Carr, P., Geman, H., Madan, D., Yor, M.: The fine structure of asset returns: an empirical investigation. J. Bus. **75**, 305–332 (2002)
14. Carr, P., Geman, H., Madan, D.B., Yor, M.: Stochastic volatility for Lévy processes. Math. Finance **13**, 345–382 (2003)
15. Carr, P., Zhang, H., Hadjiliadis, O.: Maximum drawdown insurance. Int. J. Theor. Appl. Finance **14**, 1195–1230 (2011)
16. Chaumont, L.: On the law of the supremum of Lévy processes. Ann. Probab. **41**, 1191–1217 (2013)
17. Chen, S.X., Delaigle, A., Hall, P.: Hadjiliadis nonparametric estimation for a class of Lévy processes. J. Econom. **157**, 257–271 (2010)
18. Cohn, D.L.: Measure Theory, 2nd edn. Springer, New York (2013)
19. Comte, F., Genon-Catalot, V.: Estimation for Lévy processes from high frequency data within a long time interval. Ann. Stat. **39**, 803–837 (2011)
20. Cont, R., Tankov, P.: Financial Modelling with Jump Processes, 2nd edn. Taylor & Francis, London (2015)
21. Dereich, S.: Multilevel Monte Carlo algorithms for Lévy-driven SDEs with Gaussian correction. Ann. Appl. Probab. **21**, 283–311 (2011)
22. Dereich, S., Heidenreich, F.: A multilevel Monte Carlo algorithm for Lévy-driven stochastic differential equations. Stoch. Process. Appl. **121**, 1565–1587 (2011)

23. Devroye, L.: On exact simulation algorithms for some distributions related to Brownian motion and Brownian meanders. In: Devroye, L., et al. (eds.) Recent Developments in Applied Probability and Statistics, pp. 1–35. Physica, Heidelberg (2010)
24. Dia, E.H.A.: Error bounds for small jumps of Lévy processes. Adv. Appl. Probab. **45**, 86–105 (2013)
25. Ferreiro-Castilla, A., Kyprianou, A.E., Scheichl, R., Suryanarayana, G.: Multilevel Monte Carlo simulation for Lévy processes based on the Wiener–Hopf factorisation. Stoch. Process. Appl. **124**, 985–1010 (2014)
26. Figueroa-López, J.E.: Small-time moment asymptotics for Lévy processes. Stat. Probab. Lett. **78**, 3355–3365 (2008)
27. Fournier, N.: Simulation and approximation of Lévy-driven stochastic differential equations. ESAIM Probab. Stat. **15**, 233–248 (2011)
28. Giles, M.B.: Multilevel Monte Carlo path simulation. Oper. Res. **56**, 607–617 (2008)
29. Giles, M.B.: Multilevel Monte Carlo methods. Acta Numer. **24**, 259–328 (2015)
30. Giles, M.B., Xia, Y.: Multilevel Monte Carlo for exponential Lévy models. Finance Stoch. **21**, 995–1026 (2017)
31. Cázares, J.I.G., Mijatović, A.: SBG approximation. GitHub repository (2020). Available online at https://github.com/jorgeignaciogc/SBG.jl
32. Cázares, J.I.G., Mijatović, A.: Convex minorants and the fluctuation theory of Lévy processes. ALEA Lat. Am. J. Probab. Math. Stat. **19**, 983–999 (2022)
33. González Cázares, J.I., Mijatović, A., Uribe Bravo, G.: Exact simulation of the extrema of stable processes. Adv. Appl. Probab. **51**, 967–993 (2019)
34. González Cázares, J.I., Mijatović, A., Uribe Bravo, G.: ε-strong simulation of the convex minorants of stable processes and meanders. Electron. J. Probab. **25**, 1–33 (2020)
35. González Cázares, J.I., Mijatović, A., Uribe Bravo, G.: Geometrically convergent simulation of the extrema of Lévy processes. Math. Oper. Res. **47**, 1141–1168 (2022)
36. Grabchak, M.: Rejection sampling for tempered Lévy processes. Stat. Comput. **29**, 549–558 (2019)
37. Heinrich, S.: Multilevel Monte Carlo methods. In: Margenov, S., et al. (eds.) Large-Scale Scientific Computing. Lecture Notes in Computer Science, vol. 2179, pp. 58–67. Springer, Berlin (2001)
38. Ivanovs, J.: Zooming in on a Lévy process at its supremum. Ann. Appl. Probab. **28**, 912–940 (2018)
39. Kallenberg, O.: Foundations of Modern Probability, 2nd edn. Springer, New York (2002)
40. Kim, Y.S., Rachev, S.T., Bianchi, M.L., Fabozzi, F.J.: A new tempered stable distribution and its application to finance. In: Bol, G., et al. (eds.) Risk Assessment, pp. 77–109. Physica-Verlag, Heidelberg (2009)
41. Klüppelberg, C., Kyprianou, A.E., Maller, R.A.: Ruin probabilities and overshoots for general Lévy insurance risk processes. Ann. Appl. Probab. **14**, 1766–1801 (2004)
42. Korolev, V., Shevtsova, I.: An improvement of the Berry–Esseen inequality with applications to Poisson and mixed Poisson random sums. Scand. Actuar. J. **2**, 81–105 (2012)
43. Kou, S.: Lévy Processes in Asset Pricing. Wiley, New York (2008)
44. Kou, S.: A jump-diffusion model for option pricing. Manag. Sci. **48**, 1086–1101 (2002)
45. Kudryavtsev, O., Levendorskiǐ, S.: Fast and accurate pricing of barrier options under Lévy processes. Finance Stoch. **13**, 531–562 (2009)
46. Kuznetsov, A., Kyprianou, A.E., Pardo, J.C., van Schaik, K.: A Wiener–Hopf Monte Carlo simulation technique for Lévy processes. Ann. Appl. Probab. **21**, 2171–2190 (2011)
47. Kyprianou, A.E.: Introductory Lectures on Fluctuations of Lévy Processes with Applications. Springer, Berlin (2006)
48. Landriault, D., Li, B., Zhang, H.: On magnitude, asymptotics and duration of drawdowns for Lévy models. Bernoulli **23**, 432–458 (2017)
49. Li, P., Zhao, W., Zhou, W.: Ruin probabilities and optimal investment when the stock price follows an exponential Lévy process. Appl. Math. Comput. **259**, 1030–1045 (2015)
50. Mariucci, E., Reiß, M.: Wasserstein and total variation distance between marginals of Lévy processes. Electron. J. Stat. **12**, 2482–2514 (2018)
51. Martin, P., McCann, B.: The Investors Guide to Fidelity Funds, 2nd edn. Venture Catalyst, Redmond, Washington (1989)
52. Merton, R.C.: Option pricing when underlying stock returns are discontinuous. J. Financ. Econ. **3**, 125–144 (1976)
53. Mijatović, A.: Local time and the pricing of time-dependent barrier options. Finance Stoch. **14**, 13–48 (2010)
54. Mijatović, A., Tankov, P.: A new look at short-term implied volatility in asset price models with jumps. Math. Finance **26**, 149–183 (2016)

55. Mordecki, E.: Ruin probabilities for Lévy processes with mixed-exponential negative jumps. Theory Probab. Appl. **48**, 170–176 (2004)
56. Neumann, M.H., Reiß, M.: Nonparametric estimation for Lévy processes from low-frequency observations. Bernoulli **15**, 223–248 (2009)
57. Petrov, V.V.: Sums of Independent Random Variables. Springer, New York (1975)
58. Picard, J.: Density in small time at accessible points for jump processes. Stoch. Process. Appl. **67**, 251–279 (1997)
59. Qin, L., Todorov, V.: Nonparametric implied Lévy densities. Ann. Stat. **47**, 1025–1060 (2019)
60. Rio, E.: Upper bounds for minimal distances in the central limit theorem. Ann. Inst. Henri Poincaré Probab. Stat. **45**, 802–817 (2009)
61. Rosiński, J.: Series representations of Lévy processes from the perspective of point processes. In: Barndorff-Nielsen, O.E., et al. (eds.) Lévy Processes, pp. 401–415. Birkhäuser, Boston (2001)
62. Sato, K-i.: Lévy Processes and Infinitely Divisible Distributions. Cambridge University Press, Cambridge (2013)
63. Schoutens, W.: Levy Processes in Finance: Pricing Financial Derivatives. Wiley Series in Probability and Statistics. Wiley, New York (2003)
64. Schoutens, W.: Exotic options under Lévy models: an overview. J. Comput. Appl. Math. **189**, 526–538 (2006)
65. Sornette, D.: Why Stock Markets Crash: Critical Events in Complex Financial Systems. Princeton University Press, Princeton (2003)
66. Uchaikin, V.V., Zolotarev, V.M.: Chance and Stability. Modern Probability and Statistics. VSP, Utrecht (1999)
67. Večeř, J.: Maximum drawdown and directional trading. Risk **19**(12), 88–92 (2006)
68. Villani, C.: Optimal Transport. Springer, Berlin (2009)