ARTICLE IN PRESS

Available online at www.sciencedirect.com

ScienceDirect

Stochastic Processes and their Applications ▮ (▮▮▮▮) ▮▮▮–▮▮▮

stochastic processes and their applications

www.elsevier.com/locate/spa

Importance sampling and statistical Romberg method for Lévy processes

Mohamed Ben Alaya, Kaouther Hajji, Ahmed Kebaier*

Université Paris 13, Sorbonne Paris Cité, LAGA, CNRS (UMR 7539), France

Received 4 August 2014; received in revised form 10 November 2015; accepted 23 December 2015

Abstract

An important family of stochastic processes arising in many areas of applied probability is the class of Lévy processes. Generally, such processes are not simulatable especially for those with infinite activity. In practice, it is common to approximate them by truncating the jumps at some cut-off size ε ($\varepsilon \searrow 0$). This procedure leads us to consider a simulatable compound Poisson process. This paper first introduces, for this setting, the statistical Romberg method to improve the complexity of the classical Monte Carlo method. Roughly speaking, we use many sample paths with a coarse cut-off ε^{β}, $\beta \in (0, 1)$, and few additional sample paths with a fine cut-off ε. Central limit theorems of Lindeberg–Feller type for both Monte Carlo and statistical Romberg method for the inferred errors depending on the parameter ε are proved with explicit formulas for the limit variances. This leads to an accurate description of the optimal choice of parameters. Afterwards, the authors propose a stochastic approximation method in order to find the optimal measure change by Esscher transform for Lévy processes with Monte Carlo and statistical Romberg importance sampling variance reduction. Furthermore, we develop new adaptive Monte Carlo and statistical Romberg algorithms and prove the associated central limit theorems. Finally, numerical simulations are processed to illustrate the efficiency of the adaptive statistical Romberg method that reduces at the same time the variance

* Corresponding author.

E-mail addresses: mba@math.univ-paris13.fr (M.B. Alaya), hajji@math.univ-paris13.fr (K. Hajji), kebaier@math.univ-paris13.fr (A. Kebaier).

http://dx.doi.org/10.1016/j.spa.2015.12.008

Please cite this article in press as: M.B. Alaya, et al., Importance sampling and statistical Romberg method for Lévy processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2015.12.008

and the computational effort associated to the effective computation of option prices when the underlying asset process follows an exponential pure jump CGMY model.

MSC: 60E07; 60G51; 60F05; 62L20; 65C05; 60H35

Keywords: Lévy processes; Esscher transform; Monte Carlo; Statistical Romberg; Variance reduction; Central limit theorems; CGMY model

1. Introduction

Lévy processes arise in many areas of applied probability and specially in mathematical finance, where they become very fashionable since they can describe the observed reality of financial markets in a more accurate way than models based on Brownian motion (see e.g. Cont and Tankov [10] and Schoutens [32]). In particular in the pricing of financial securities we are interested in the computation of the real quantity $\mathbb{E}F(L_T)$, $T > 0$, where $(L_t)_{0\le t\le T}$ is a $\mathbb{R}^d$-valued pure jump Lévy process, $d \ge 1$ and $F : \mathbb{R}^d \to \mathbb{R}$ is a given function. In the literature, the computation of this quantity may be carried out using three different methods: Fourier transform methods, numerical methods for partial integro-differential equations and Monte Carlo methods. It is well known that the two first methods cannot cope with high dimensional problems. This gives a competitive edge for Monte Carlo methods in this setting. Therefore, the focus of this work is to study improved Monte Carlo methods using the statistical Romberg algorithm and the importance sampling technique. The statistical Romberg method is known for reducing the time complexity and the importance sampling technique is aimed at reducing the variance.

The Monte Carlo method consists of two steps. In the first step, we approximate the Lévy process $(L_t)_{0\le t\le T}$ by a simulatable Lévy process $(L^\varepsilon_t)_{0\le t\le T}$ with $\varepsilon > 0$. If ν denotes the Lévy measure of the Lévy process under consideration, then it is common to take $(L^\varepsilon_t)_{0\le t\le T}$ with Lévy measure $\nu_{|\{|x|\ge\varepsilon\}}$ and $\varepsilon \searrow 0$. This approximation is nothing but a compound Poisson process. In the second step, we approximate $\mathbb{E}\, F\left(L^\varepsilon_T\right)$ by $\frac{1}{N}\sum_{i=1}^N F(L^\varepsilon_{T,i})$, where $(L^\varepsilon_{T,i})_{1\le i\le N}$ is a sample of N independent copies of L^ε_T. Therefore, this Monte Carlo method (MC) is affected respectively by an approximation error and a statistical error:

$$\mathcal{E}_1(\varepsilon) := \mathbb{E}\left(F(L^\varepsilon_T) - F(L_T)\right) \quad \text{and} \quad \mathcal{E}_2(N) := \frac{1}{N}\sum_{i=1}^N F(L^\varepsilon_{T,i}) - \mathbb{E}F(L^\varepsilon_T).$$

On one hand, for a Lipschitz function F we have $\mathcal{E}_1(\varepsilon) = O(\sigma(\varepsilon))$, where $\sigma^2(\varepsilon) = \mathbb{E}|L_1 - L^\varepsilon_1|^2$ (see relation (6) for more details). On the other hand, the statistical error is controlled by the central limit theorem with order $1/\sqrt{N}$. Hence, optimizing the choice of the sample size N in the Monte Carlo method leads to $N = O(\sigma^{-2}(\varepsilon))$. Moreover, if we choose $N = \sigma^{-2}(\varepsilon)$ we prove a central limit theorem of Lindeberg–Feller type (see Theorem 3.1). Therefore, if we denote by $\mathcal{K}(\varepsilon)$ the cost of a single simulation of L^ε_T, then the mean total cost necessary to achieve the precision $\sigma(\varepsilon)$ is given by $C_{MC} = O(\mathcal{K}(\varepsilon)\sigma^{-2}(\varepsilon))$ (see Section 3.3).

In order to improve the performance of this method we use the idea of the statistical Romberg method introduced by Kebaier [22] in the setting of Euler Monte Carlo methods for stochastic differential equations driven by a standard Brownian Motion which is also related to the well known Romberg's method introduced by Talay and Tubaro in [33]. Inspired by this technique, we introduce a novel method for the computation of our initial target. The main idea of this new method is to consider two cut-off sizes ε and ε^{β}, $\beta \in (0, 1)$ and then approximate $\mathbb{E}F(L_T)$ by

$$\frac{1}{N_1}\sum_{i=1}^{N_1} F(\hat{L}^{\varepsilon^{\beta}}_{T,i}) + \frac{1}{N_2}\sum_{i=1}^{N_2} F(L^{\varepsilon}_{T,i}) - F(L^{\varepsilon^{\beta}}_{T,i}).$$

The samples $(L^{\varepsilon}_{T,i})_{1\le i\le N_2}$ and $(L^{\varepsilon^{\beta}}_{T,i})_{1\le i\le N_2}$ have to be independent of $(\hat{L}^{\varepsilon}_{T,i})_{1\le i\le N_1}$. Moreover, for $1 \le i \le N_2$, the process $(L^{\varepsilon}_{t,i})_{0\le t\le T}$ is nothing else the sum of $(L^{\varepsilon^{\beta}}_{t,i})_{0\le t\le T}$ and an independent Lévy process $(L^{\varepsilon,\varepsilon^{\beta}}_{t,i})_{0\le t\le T}$ with Lévy measure $\nu_{|\{\varepsilon\le|x|\le\varepsilon^{\beta}\}}$ which is also simulatable as a compound Poisson process. This new method will be referred as the statistical Romberg method (SR). Additionally, like for the MC method, we prove a central limit theorem of Lindeberg–Feller type for the SR algorithm with $N_1 = \sigma^{-2}(\varepsilon)$ and $N_2 = \sigma^{-2}(\varepsilon)\sigma^{2}(\varepsilon^{\beta})$ (see Theorem 3.2). Then, according to Section 3.3, the total time complexity necessary to achieve the precision $\sigma(\varepsilon)$ is given by $C_{SR} = \left(\mathcal{K}(\varepsilon^{\beta}) + \mathcal{K}(\varepsilon)\sigma^{2}(\varepsilon^{\beta})\right)\sigma^{-2}(\varepsilon)$. It turns out that the complexity ratio C_{SR}/C_{MC} vanishes as ε goes to zero.

Since the efficiency of the Monte Carlo simulation considerably depends on the smallness of the variance in the estimation, many variance reduction techniques were developed in the recent years. Among these methods appears the technique of importance sampling very popular for its efficiency. For the Gaussian setting, the importance sampling technique was studied by Arouna [2], Glasserman, Heidelberger and Shahabuddin [19] for MC method and by Ben Alaya, Hajji and Kebaier [4] for SR method. Concerning Lévy processes without a Brownian component, Kawai [21] has already applied this technique for MC algorithm using the Esscher transform which is nothing but the well known exponential tilting of laws. From a practical point of view, his approach is exploitable only when the Lévy process $(L_t)_{0\le t\le T}$ is simulatable without any approximation. Note also that in his study there is no result on the rate of convergence of the obtained algorithm. Furthermore, we can cite the recent works of Dereich and Li [11], Giles and Xia [34,35] and Abdulle, Blumenthal and Buckwar [1] on the Multilevel Monte Carlo method and references there. All these papers apply this method to Lévy processes. In [11], the authors consider a one-dimensional stochastic differential equation driven by a Lévy process $dX_t = a(X_t)dL_t$, for $t \ge 0$ and obtain a central limit theorem on the associated Multilevel Monte Carlo method combined with their proposed approximation scheme, see Theorem 1.6 in [11]. The limit variance in this result is zero when the coefficient a is a constant function. So, their result is not optimal when $(X_t)_{t\ge 0}$ is itself a Lévy process and another study is needed to obtain the appropriate rate of converge. This is the aim of our Theorem 3.2. The three other papers are rather interested in the study of the mean square error as in the original work of Giles [18] and on the numerical performance of the Multilevel Monte Carlo method in the setting of Lévy processes. Elsewhere, there is no importance sampling approach in the algorithms proposed by [1,11,34,35].

The main aim of the present work is to apply the idea of [21] to the approximation Lévy process $(L^{\varepsilon}_t)_{0\le t\le T}$ for both MC and SR algorithms and to study the inferred error in terms of the cut-off ε; a question which has not been addressed in previous research. Roughly speaking, thanks to the Esscher transform we produce a parametric transformation such that for all $\theta \in K$ we have $\mathbb{E}F(L^{\varepsilon}_T) = \mathbb{E}G(\theta, L^{\varepsilon}_T)$, where K is a suitable subset of $\mathbb{R}^d$ and $(\theta, x) \mapsto G(\theta, x)$ is a real function taking values in $\mathbb{R}^d \times \mathbb{R}^d$. Concerning the MC method it looks natural to implement the method with $\theta^*_{1,\varepsilon} = \arg\min_{\theta\in K}\mathbb{E}G^2(\theta, L^{\varepsilon}_T)$. However, for the SR method the inferred error is controlled by $\mathrm{Var}(G(\theta, L^{\varepsilon}_T)) + T\mathbb{E}(|\nabla_x G(\theta, L^{\varepsilon}_T)|^2)$. Then, in this case, it is natural to implement the first (resp. the second) empirical mean appearing in the SR estimator with $\theta^*_{1,\varepsilon}$ (resp. $\theta^*_{2,\varepsilon} = \arg\min_{\theta\in K}\mathbb{E}(|\nabla_x G(\theta, L^\varepsilon_T)|^2)$). But what about the effective computation of $(\theta^*_{i,\varepsilon})_{i\in\{1,2\}}$? To answer this question, we use a constrained version of the well-known stochastic approximation Robbins–Monro. All these ideas led us to introduce two new methods based on adaptive approximations. The first method concerns a combination of an adaptive importance sampling technique and the MC method that will be called Importance Sampling Monte Carlo method (ISMC) (see relation (24)). The second one concerns an original combination of an adaptive importance sampling technique with the SR algorithm that will be referred as Importance Sampling Statistical Romberg method (ISSR) (see relation (28)). The main point in favor of the ISSR method is that it inherits the variance reduction from the Importance sampling procedure and the complexity reduction from the SR method. A complexity analysis is also provided.

From a technical point of view, in this paper we deal with a pure jump Lévy process, so many specific problems related to the nature of jumps appear and we need new techniques completely different from the continuous Brownian diffusion case as in [4,5,22]. Unlike the Brownian motion, Lévy processes may not have finite exponential moments. This issue makes the study of the error terms quite complicated especially when we combine the Monte Carlo or the statistical Romberg method with the importance sampling procedure.

The rest of the paper is organized as follows. Section 2 introduces the general framework and recalls some useful results. In Section 3, the central limit theorems of Lindeberg–Feller type are proved for both MC and SR methods (see Theorems 3.1 and 3.2). Similar results are derived for the setting of an exponential Lévy model (see Corollaries 3.1 and 3.2). A complexity analysis is included. In Section 4, we recall the Esscher transform and the principle of importance sampling technique for the SR method. For $i \in \{1, 2\}$ and $\varepsilon \searrow 0$, we prove the convergence of the optimal choice $\theta^*_{i,\varepsilon}$ to the optimal choice associated to the limit model (see Theorem 4.1). In Section 5, we first study, for $i \in \{1, 2\}$, the almost sure convergence of the stochastic recursive constrained Robbins–Monro algorithm given by the double indexed sequence $\theta_{i,\varepsilon,n}$ as $\varepsilon \searrow 0$ and $n \nearrow \infty$ (see Theorems 5.1 and 5.2 and Corollary 5.1). The rest of this section is devoted to prove the central limit theorems of Lindeberg–Feller type for both adaptive ISMC and ISSR methods (see Theorems 5.3 and 5.4). Section 6 illustrates the superiority of the ISSR method over all the other ones via numerical examples for both one and two-dimensional Carr, Geman, Madan and Yor (CGMY) process [8]. Finally, the last section is devoted to discuss some future openings.

2. General framework

We denote by $(\Omega, \mathcal{F}, \mathbb{P})$ our underlying probability space. A stochastic process $(L_t)_{t\geq 0}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ with values in $\mathbb{R}^d$ such that $L_0 = 0$ is a Lévy process if it has independent and stationary increments. We endow the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with the canonical filtration $(\mathcal{F}_t)_{0\leq t\leq T}$ where $\mathcal{F}_t = \sigma(L_s, s \leq t)$. The characteristic function of a Lévy process L with generating triplet (γ, A, ν) is given by the well known Lévy–Khintchine representation

$$\mathbb{E}e^{iu.L_t} = \exp\left\{t\left(i\gamma.u - \frac{1}{2}u.Au + \int_{\mathbb{R}^d}(e^{iu.x} - 1 - iu.x\mathbf{1}_{|x|\leq 1})\nu(dx)\right)\right\}, \quad u \in \mathbb{R}^d,$$

where $\gamma \in \mathbb{R}^d$, A is a symmetric non-negative-definite $d \times d$ matrix and ν is a Lévy measure on $\mathbb{R}^d \setminus \{0\}$ verifying $\int_{\mathbb{R}^d\setminus\{0\}}(|x|^2 \wedge 1)\nu(dx) < \infty$. (Given vectors x and $y \in \mathbb{R}^d$, $x.y$ denotes the inner product of x and y associated to the Euclidean norm $|\cdot|$). In this paper, we are interested in studying pure-jump Lévy processes, that is, we set $A \equiv 0$ throughout all the paper. Then, $(L_t)_{t\geq 0}$

is a Lévy process with generating triplet $(\gamma, 0, \nu)$. The simulation of a Lévy process with infinite Lévy measure is not straightforward. From the Lévy–Itô decomposition (see e.g. Theorem 19.2 in Sato [31]), we know that L can be represented as a sum of a compound Poisson process and an almost sure limit of compensated compound Poisson process $L_t = \lim_{\varepsilon\to 0} L_t^{\varepsilon}$ a.s. where for $0 < \varepsilon < 1$

$$L_t^{\varepsilon} = \gamma t + \sum_{0<s\le t} \Delta L_s \mathbf{1}_{|\Delta L_s|>1} + \left(\sum_{0<s\le t} \Delta L_s \mathbf{1}_{\varepsilon<|\Delta L_s|\le 1} - t\int_{\varepsilon<|x|\le 1} x\nu(dx) \right), \quad t \ge 0. \tag{1}$$

Note that without the compensation $t\int_{\varepsilon<|x|\le 1} x\nu(dx)$, the sum of jumps $\sum_{0<s\le t} \Delta L_s \mathbf{1}_{\varepsilon<|\Delta L_s|\le 1}$ may not converge as ε goes to zero. We denote the approximation error by

$$R^{\varepsilon} = L - L^{\varepsilon}. \tag{2}$$

The process R^{ε} is also a Lévy process independent of L^{ε} with characteristic function

$$\mathbb{E}e^{iu.R_t^{\varepsilon}} = \exp\left\{ t\int_{|x|\le\varepsilon} (e^{iu.x} - 1 - iu.x)\nu(dx) \right\}.$$

Consequently, $\mathbb{E}[R_t^{\varepsilon}] = 0$ and the variance–covariance matrix $\mathbb{E}[R_t^{\varepsilon}(R_t^{\varepsilon})'] = t\Sigma_{\varepsilon}$ where

$$\Sigma_{\varepsilon} = \int_{|x|\le\varepsilon} xx'\nu(dx).$$

(Here, we denote A' the transpose of a given matrix A). The asymptotic behavior of the distribution of R^{ε} is firstly studied by Asmussen and Rosiński [3] in the one-dimensional case and later extended to the multidimensional case by Cohen and Rosiński [9]. Throughout this paper $W = (W_t)_{t\ge 0}$ is a standard Brownian motion in $\mathbb{R}^d$ independent of $(L_t)_{t\ge 0}$.

Theorem 2.1. *Assume that the matrix Σ_{ε}, defined above, is invertible for every $\varepsilon \in (0, 1]$. Then, as $\varepsilon \to 0$,*

$$\Sigma_{\varepsilon}^{-1/2} R^{\varepsilon} \Rightarrow W,$$

if and only if for each $k > 0$

$$\lim_{\varepsilon\to 0} \int_{\langle \Sigma_{\varepsilon}^{-1}x, x\rangle > k} \langle \Sigma_{\varepsilon}^{-1}x, x\rangle \mathbf{1}_{|x|\le\varepsilon}\nu(dx) = 0. \tag{3}$$

Here "$\Rightarrow$" stands for the convergence in distribution.

If ν is given in polar coordinates by $\nu(dr, du) = \mu(dr|u)\lambda(du)$, $r > 0$, $u \in S^{d-1}$, where $\{\mu(\cdot|u) : u \in S^{d-1}\}$ is a measurable family of Lévy measures on $(0, \infty)$ and λ is a finite measure on the unit sphere S^{d-1}, then

$$\Sigma_{\varepsilon} = \int_{S^{d-1}} \int_0^{\varepsilon} r^2 uu'\mu(dr|u)\lambda(du).$$

If we define $\sigma^2(\varepsilon, u) := \int_0^{\varepsilon} r^2\mu(dr|u)$ and $\sigma^2(\varepsilon) := \int_{S^{d-1}} \sigma^2(\varepsilon, u)\lambda(du)$, then

$$\mathbb{E}|L_t - L_t^{\varepsilon}|^2 = t\mathrm{Tr}(\Sigma_{\varepsilon}) = t\sigma^2(\varepsilon). \tag{4}$$

Please cite this article in press as: M.B. Alaya, et al., Importance sampling and statistical Romberg method for Lévy processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2015.12.008

Remark. In the one-dimensional case Asmussen and Rosiński [3] have obtained the convergence of $\sigma^{-1}(\varepsilon)R^{\varepsilon}$ to a standard Brownian motion if and only if for each $k > 0$, $\sigma(k\sigma(\varepsilon) \wedge \varepsilon) \sim \sigma(\varepsilon)$ which is satisfied as soon as $\lim_{\varepsilon\to 0} \frac{\sigma(\varepsilon)}{\varepsilon} = \infty$ (see Theorem 2.1 and Proposition 2.1 in [3]). An extension of this condition to the multidimensional case is given by Theorem 2.5 in Cohen and Rosiński [9]. Suppose that the support of the measure λ is not contained in any proper linear subspace of $\mathbb{R}^d$, they proved that if

$$\lim_{\varepsilon\to 0} \frac{\sigma(\varepsilon, u)}{\varepsilon} = \infty, \lambda - a.e. \tag{5}$$

then Σ_ε is invertible and condition (3) of Theorem 2.1 holds.

The CGMY process. One example of Lévy process that satisfies our forthcoming theoretical assumptions and on which we carry out our numerical tests in Section 6 is the CGMY process. This real-valued pure jump process has for generating triplet $(0, 0, \nu)$ with

$$\nu(dx) = \frac{Ce^{-Mx}}{x^{1+Y}}\mathbf{1}_{x>0}dx + \frac{Ce^{-G|x|}}{|x|^{1+Y}}\mathbf{1}_{x<0}dx,$$
$$\text{where } C > 0, G > 0, M > 0, \text{ and } 0 < Y < 2.$$

In this case $\sigma^2(\varepsilon) = O(\varepsilon^{2-Y})$. So $\lim\limits_{\varepsilon\to 0} \frac{\sigma(\varepsilon)}{\varepsilon} = \infty$ and according to the above remark, Theorem 2.1 applies. In the multidimensional setting, we consider the d-dimensional independent CGMY processes, so that

$$\nu(dx_1, \dots, dx_d) = \sum_{i=1}^{d} \delta_0(dx_1)\dots\delta_0(dx_{i-1})\nu_i(dx_i)\delta_0(dx_{i+1})\dots\delta_0(dx_d)$$

where for $i \in \{1, \dots, d\}$ ν_i is the Lévy measure of a one-dimensional CGMY process with parameters $C_i > 0$, $G_i > 0$, $M_i > 0$, and $0 < Y_i < 2$. For this multidimensional setting, we have $\Sigma_\varepsilon = \mathrm{Diag}(\sigma_1^2(\varepsilon), \dots, \sigma_d^2(\varepsilon))$ where for $i \in \{1, \dots, d\}$, $\sigma_i^2(\varepsilon) = \int_{|x_i|\le\varepsilon} x_i^2 \nu_i(dx_i)$ which is of order ε^{2-Y_i}. Meanwhile, condition (3) reduces to $\lim_{\varepsilon\to 0}\sum_{i=1}^{d}\int_{\{\sqrt{k}\sigma_i(\varepsilon)\wedge\varepsilon\le|x_i|\le\varepsilon\}} \sigma_i^{-2}(\varepsilon)$ $x_i^2\nu_i(dx_i) = 0$, which is satisfied as we have $\lim\limits_{\varepsilon\to 0}\frac{\sigma_i(\varepsilon)}{\varepsilon} = \infty$ (see the above remark) and then the result of Theorem 2.1 holds. Another way to obtain this result is to apply Theorem 2.1 and Proposition 2.1 in [3] for each component of the d-dimensional CGMY process using the independence structure.

On the other hand, according to Proposition 2.1 of Dia [12], we have a L^q-upper bound of the error approximation in the one-dimensional case for any real $q > 0$. This result on the strong error approximation remains valid for the multidimensional case. More precisely, if we consider the d-dimensional error Lévy process R^ε given by relation (2), then we can easily deduce that

$$\mathbb{E}|R_t^\varepsilon|^q \le K_{q,T}\sigma_0(\varepsilon)^q, \quad \text{where } K_{q,T} > 0 \text{ and } \sigma_0(\varepsilon) = \sigma(\varepsilon) \vee \varepsilon. \tag{SE}$$

Concerning the weak error, if F denotes a real valued Lipschitz continuous function with Lipschitz constant $C > 0$, then it is easy to see that

$$|\mathbb{E}F(L_T) - \mathbb{E}F(L_T^\varepsilon)| \le C\sqrt{T}\sigma(\varepsilon). \tag{6}$$

Moreover, under some regularity conditions on the function F we obtain an expansion of the weak error. This result extends Proposition 2.2 of [12] to the multidimensional setting.

Proposition 2.1. 1. *Let* $F : \mathbb{R}^d \to \mathbb{R}$ *be a* $\mathscr{C}^1$ *function satisfying* $\mathbb{E}|\nabla F(L_T^\varepsilon)| < \infty$. *If there exists* $p > 1$ *such that* $\sup_{\varepsilon \in (0,1]} \mathbb{E}^{\frac{1}{p}} \left|\nabla F(L_T^\varepsilon + \theta R_T^\varepsilon) - \nabla F(L_T^\varepsilon)\right|^p$ *is finite and integrable with respect to* $\theta \in [0, 1]$, *then*

$$\mathbb{E}F(L_T) - \mathbb{E}F(L_T^\varepsilon) = o(\sigma_0(\varepsilon)).$$

2. *Let* $F : \mathbb{R}^d \to \mathbb{R}$ *be a* $\mathscr{C}^2$ *function satisfying* $\mathbb{E}|\nabla F(L_T^\varepsilon)| < \infty$. *Assume that there exists a definite positive matrix* Σ *such that* $\lim_{\varepsilon\to 0} \sigma^{-2}(\varepsilon)\Sigma_\varepsilon = \Sigma$. *If there exists* $p > 1$ *such that* $\sup_{\varepsilon\in(0,1]} \mathbb{E}|Hess(F(L_T^\varepsilon))|^p$ *and* $\sup_{\varepsilon\in(0,1]} \mathbb{E}^{\frac{1}{p}} \left|Hess(F(L_T^\varepsilon + \theta R_T^\varepsilon)) - Hess(F(L_T^\varepsilon))\right|^p$ *are finite and integrable with respect to* $\theta \in [0, 1]$, *then*

$$\mathbb{E}\left(F(L_T) - F(L_T^\varepsilon)\right) = \frac{T\sigma^2(\varepsilon)}{2}\mathrm{Tr}\left(\Sigma\mathbb{E}\left[Hess(F(L_T))\right]\right) + o(\sigma_0^2(\varepsilon)).$$

Proof. For the first assertion, we use Taylor's expansion for order one to get

$$F(L_T) - F(L_T^\varepsilon) = \nabla F(L_T^\varepsilon).R_T^\varepsilon + \int_0^1 \Big(\nabla F(L_T^\varepsilon + \theta R_T^\varepsilon) - \nabla F(L_T^\varepsilon)\Big).R_T^\varepsilon\, d\theta.$$

Thanks to the independence between L_T^ε and R_T^ε the expectation of the first term in the right hand side of the above equality is equal zero. For $1 < p' < p$, by Hölder's inequality and assumption (SE) there exists $C > 0$ such that

$$\mathbb{E}\left|\Big(\nabla F(L_T^\varepsilon + \theta R_T^\varepsilon) - \nabla F(L_T^\varepsilon)\Big).R_T^\varepsilon\right| \le C\sigma_0(\varepsilon)\mathbb{E}^{\frac{1}{p'}}\left|\nabla F(L_T^\varepsilon + \theta R_T^\varepsilon) - \nabla F(L_T^\varepsilon)\right|^{p'}.$$

Since, R_T^ε converges almost surely to zero as ε tends to zero, we obtain the desired result using our uniform integrability assumption. Concerning the second assertion, we use Taylor's expansion for order two to get

$$\begin{aligned}\mathbb{E}\left(F(L_T) - F(L_T^\varepsilon)\right) &= \mathbb{E}\left(\nabla F(L_T^\varepsilon).R_T^\varepsilon\right) + \frac{1}{2}\mathbb{E}\left(R_T^\varepsilon.\mathrm{Hess}(F(L_T^\varepsilon))R_T^\varepsilon\right)\\ &\quad + \mathbb{E}\left(\int_0^1 (1-\theta)R_T^\varepsilon.\left[\mathrm{Hess}(F(L_T^\varepsilon + \theta R_T^\varepsilon)) - \mathrm{Hess}(F(L_T^\varepsilon))\right]R_T^\varepsilon d\theta\right). \qquad (7)\end{aligned}$$

Using once again the independence between L_T^ε and R_T^ε, the first term in the right hand side of the above relation is equal to zero and

$$\begin{aligned}\mathbb{E}\left(R_T^\varepsilon.\mathrm{Hess}(F(L_T^\varepsilon))R_T^\varepsilon\right) &= \mathrm{Tr}\left(\mathbb{E}\left[R_T^\varepsilon(R_T^\varepsilon)'\right]\mathbb{E}\left[\mathrm{Hess}(F(L_T^\varepsilon))\right]\right)\\ &= T\,\mathrm{Tr}\left(\Sigma_\varepsilon\mathbb{E}\left[\mathrm{Hess}(F(L_T^\varepsilon))\right]\right).\end{aligned}$$

Now, as $\lim_{\varepsilon\to 0}\sigma^{-2}(\varepsilon)\Sigma_\varepsilon = \Sigma$ and $\sup_{\varepsilon\in(0,1]}\mathbb{E}|\mathrm{Hess}(F(L_T^\varepsilon))|^p < \infty$ we deduce that

$$\mathrm{Tr}\left(\Sigma_\varepsilon\mathbb{E}\left[\mathrm{Hess}(F(L_T^\varepsilon))\right]\right) = \sigma^2(\varepsilon)\mathrm{Tr}\left(\Sigma\mathbb{E}\left[\mathrm{Hess}(F(L_T))\right]\right) + o(\sigma_0^2(\varepsilon)).$$

Similarly to the proof of the first assertion, we use our uniform integrability assumption to obtain that the remaining term in (7) is of order $o(\sigma_0^2(\varepsilon))$. This completes the proof. □

The above result illustrates some relation existing between the regularity of F and the rate of the weak convergence. In fact, if F has bounded derivatives or F is an exponential element-wise function with a Lévy process having finite exponential moments, then the conditions on the

Please cite this article in press as: M.B. Alaya, et al., Importance sampling and statistical Romberg method for Lévy processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2015.12.008

above proposition involving the function F are satisfied. So, it is worth introducing the following assumption: there exist $C_F \in \mathbb{R}$ and $\upsilon_\varepsilon \searrow 0$ as $\varepsilon \searrow 0$ such that

$$\upsilon_\varepsilon^{-1} \left(\mathbb{E}F(L_T^\varepsilon) - \mathbb{E}F(L_T)\right) \to C_F \quad \text{as } \varepsilon \searrow 0. \tag{WE$_{\upsilon_\varepsilon}$}$$

Remark. According to the above proposition, in the first case one can chose $\upsilon_\varepsilon = \sigma_0(\varepsilon)$ and $C_F = 0$ and in the second case $\upsilon_\varepsilon = \sigma^2(\varepsilon)$ and $C_F = \frac{T}{2}\mathrm{Tr}\left(\Sigma\mathbb{E}\left[\mathrm{Hess}(F(L_T))\right]\right)$. Nevertheless, if F is only a Lipschitz-continuous function, then thanks to relation (6), we can take $\upsilon_\varepsilon = \sigma^{1-\eta}(\varepsilon)$, $\eta \in (0,1)$, and $C_F = 0$. Note also that condition $\lim_{\varepsilon\to 0} \sigma^{-2}(\varepsilon)\Sigma_\varepsilon = \Sigma$ is satisfied for the d-dimensional CGMY process introduced in Section 2 provided that for all $i \in \{1, \ldots, d\}$, $Y_i = Y \in (0, 2)$.

3. Statistical Romberg method and Lévy process

In this section, we establish two central limit theorems of Lindeberg–Feller type, for the inferred errors associated to MC and SR algorithms, in terms of the cut-off ε. Similar results are derived for the setting of an exponential Lévy model. We also provide a complexity analysis for both algorithms. In the sequel, for $x \in \mathbb{R}$ we denote by $\lfloor x \rfloor$ the biggest integer less or equal than x.

3.1. Central limit theorem for the MC method

Theorem 3.1. *Let $F : \mathbb{R}^d \to \mathbb{R}$ be a continuous function satisfying assumption (WE$_{\upsilon_\varepsilon}$). If $\sup_{0<\varepsilon\leq 1} \mathbb{E}\left[F^{2a}(L_T^\varepsilon)\right] < +\infty$ for $a > 1$, then for $N = \lfloor \upsilon_\varepsilon^{-2} \rfloor$ we have*

$$\upsilon_\varepsilon^{-1} \left(\frac{1}{N}\sum_{i=1}^N F(L_{T,i}^\varepsilon) - \mathbb{E}F(L_T)\right) \xrightarrow{\mathcal{L}} \mathcal{N}(C_F, \mathrm{Var}(F(L_T))) \quad \textit{as } \varepsilon \searrow 0. \tag{8}$$

Proof. At first, we write the total error as follows

$$\frac{1}{N}\sum_{i=1}^N F(L_{T,i}^\varepsilon) - \mathbb{E}F(L_T) = \frac{1}{N}\sum_{i=1}^N F(L_{T,i}^\varepsilon) - \mathbb{E}F(L_T^\varepsilon) + \left(\mathbb{E}F(L_T^\varepsilon) - \mathbb{E}F(L_T)\right).$$

Assumption (WE$_{\upsilon_\varepsilon}$) ensures that $\lim_{\varepsilon\to 0} \upsilon_\varepsilon^{-1}\mathbb{E}\left(F(L_T^\varepsilon) - F(L_T)\right) = C_F$. Concerning the first term on the right hand side of the above relation, as N depends on ε we plan to apply the Lindeberg–Feller central limit theorem for independent random variables (see e.g. [6]). In order to do that, we set $X_{i,\varepsilon} := \frac{\upsilon_\varepsilon^{-1}}{N}\left(F(L_{T,i}^\varepsilon) - \mathbb{E}F(L_T^\varepsilon)\right)$, and we split the proof into two steps.

Step 1. First, we easily check that $\sum_{i=1}^N \mathbb{E}(X_{i,\varepsilon}^2) \underset{\varepsilon\to 0}{\simeq} \mathrm{Var}(F(L_T^\varepsilon))$. Then, by the almost sure convergence of L_T^ε towards L_T, the continuity of function F and the uniform integrability condition given by $\sup_{0<\varepsilon\leq 1} \mathbb{E}\left[F^{2a}(L_T^\varepsilon)\right] < +\infty$, we obtain

$$\lim_{\varepsilon\to 0} \sum_{i=1}^N \mathbb{E}(X_{i,\varepsilon}^2) = \lim_{\varepsilon\to 0} \mathrm{Var}(F(L_T^\varepsilon)) = \mathrm{Var}(F(L_T)). \tag{9}$$

Step 2. Concerning the Lyapunov condition, for $1 < \tilde{a} < a$, we have

$$\sum_{i=1}^N \mathbb{E}\left[|X_{i,\varepsilon}|^{2\tilde{a}}\right] \underset{\varepsilon\to 0}{\simeq} \upsilon_\varepsilon^{2(\tilde{a}-1)}\mathbb{E}\left|F(L_T^\varepsilon) - \mathbb{E}F(L_T^\varepsilon)\right|^{2\tilde{a}}.$$

Please cite this article in press as: M.B. Alaya, et al., Importance sampling and statistical Romberg method for Lévy processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2015.12.008

Once again by the same arguments used in the previous step we prove the convergence of $\mathbb{E}\left|F(L_T^\varepsilon)-\mathbb{E}F(L_T^\varepsilon)\right|^{2\tilde{a}}$ towards $\mathbb{E}\,|F(L_T)-\mathbb{E}F(L_T)|^{2\tilde{a}}$ as ε tends to zero. Since $\upsilon_\varepsilon^{2(\tilde{a}-1)} \underset{\varepsilon\to 0}{\longrightarrow} 0$, we obtain

$$\lim_{\varepsilon\to 0}\sum_{i=1}^{N}\mathbb{E}\left[|X_{i,\varepsilon}|^{2\tilde{a}}\right]=0. \tag{10}$$

By (9) and (10), we obtain thanks to the Lindeberg–Feller central limit theorem for independent random variables the desired convergence in law. □

In the corollary below, we will treat the special case where $F(x)=f(e^{x_1},\ldots,e^{x_d})$ for all $x=(x_1,\ldots,x_d)\in\mathbb{R}^d$ and $f:\mathbb{R}_+^d\to\mathbb{R}$ is a Lipschitz continuous function. In finance this model is well known as an exponential Lévy model.

Corollary 3.1. *Assume that $\int_{|z|>1}e^{2a|z|}\nu(dz)$ is finite for $a>1$. Then, in the setting of an exponential Lévy model there is $C>0$ such that $\left|\mathbb{E}F(L_T)-\mathbb{E}F(L_T^\varepsilon)\right|\le C\sigma(\varepsilon)$. Moreover, if we choose $N=\lfloor\sigma^{-2+\eta}(\varepsilon)\rfloor$, with $0<\eta<2$, then*

$$\sigma^{-1+\eta/2}(\varepsilon)\left(\frac{1}{N}\sum_{i=1}^{N}F(L_{T,i}^\varepsilon)-\mathbb{E}F(L_T)\right)\xrightarrow{\mathcal{L}}\mathcal{N}(0,\mathrm{Var}(F(L_T)))\quad \text{as } \varepsilon\searrow 0. \tag{11}$$

It is worth noticing that this corollary ensures that Theorem 3.1 assumptions are satisfied for a large class of functions with any Lévy process satisfying $\int_{|z|>1}e^{2a|z|}\nu(dz)<\infty$, $a>1$. For the d-dimensional CGMY process introduced in Section 2, this condition is equivalent to choose for all $i\in\{1,\ldots,d\}M_i>2$ and $G_i>2$. Furthermore, this condition is obviously satisfied for any Lévy measure ν with a compact support or even with sufficiently decreasing behavior like it is the case for the Gaussian measure.

Proof. We denote by e^x the exponential function element-wise of the vector $x=(x_1,\ldots,x_d)\in\mathbb{R}^d$, $e^x=(e^{x_1},\ldots,e^{x_d})$. Let C_f denote the Lipschitz constant of function f, since L_T^1 and $(L_T-L_T^1, L_T^\varepsilon-L_T^1)$ are independent we obtain by standard calculations

$$\begin{aligned}\left|\mathbb{E}F(L_T)-\mathbb{E}F(L_T^\varepsilon)\right| &\le C_f\mathbb{E}e^{|L_T^1|}\mathbb{E}\left|L_T-L_T^\varepsilon\right|(e^{|L_T-L_T^1|}+e^{|L_T^\varepsilon-L_T^1|})\\ &\le C_f\sigma(\varepsilon)\mathbb{E}e^{|L_T^1|}\left(\|e^{|L_T-L_T^1|}\|_2+\|e^{|L_T^\varepsilon-L_T^1|}\|_2\right).\end{aligned}$$

Now, on the one hand thanks to Theorem 25.3 in [31], the assumption $\int_{|z|>1}e^{2a|z|}\nu(dz)<+\infty$ ensures the finiteness of $\mathbb{E}e^{|L_T^1|}$. On the other hand by virtue of Lemmas 25.6 and 25.7 in Sato [31] we have the boundedness of $\|e^{|L_T-L_T^1|}\|_2$. Concerning the term $\|e^{|L_T^\varepsilon-L_T^1|}\|_2$, we have $e^{|x|}\le\prod_{j=1}^{d}(e^{x_j}+e^{-x_j})$, this last upper bound can be written as a sum of finite number of exponential functions evaluated at points which are a linear combination of the components of the vector x. Therefore, there exists a family of $\mathbb{R}^d$-valued vectors, $(b_j)_{1\le j\le 2^d}$ such that

$$\|e^{|L_T^\varepsilon-L_T^1|}\|_2^2\le\sum_{j=1}^{2^d}\exp\left\{T\int_{\varepsilon\le|x|\le 1}(e^{b_j.x}-1-b_j.x)\nu(dx)\right\}.$$

Note that the finiteness of the above upper bound is once again ensured by Lemmas 25.6 and 25.7 in Sato [31]. So, it follows that $\sup_{0<\varepsilon\le 1}\|e^{|L_T^\varepsilon-L_T^1|}\|_2$ is finite. Now, thanks to the linear

Please cite this article in press as: M.B. Alaya, et al., Importance sampling and statistical Romberg method for Lévy processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2015.12.008

growth of f and using the same arguments as above we check in the same manner the property $\sup_{0<\varepsilon\leq 1}\mathbb{E}\left[F^{2a}(L_T^{\varepsilon})\right]<+\infty$. Hence, if we choose $\upsilon_\varepsilon=\sigma^{1-\eta/2}(\varepsilon)$ then Theorem 3.1 applies and this completes the proof. □

3.2. Central limit theorem for the SR method

We use the SR method to approximate $\mathbb{E}[F(L_T)]$ by

$$Q_\varepsilon=\frac{1}{N_1}\sum_{i=1}^{N_1}F(\hat{L}_{T,i}^{\varepsilon^\beta})+\frac{1}{N_2}\sum_{i=1}^{N_2}\Big(F(L_{T,i}^{\varepsilon})-F(L_{T,i}^{\varepsilon^\beta})\Big).$$

Theorem 3.2. *Let $F:\mathbb{R}^d\to\mathbb{R}$ be a $\mathscr{C}^1$ function satisfying assumption* ($\mathrm{WE}_{\upsilon_\varepsilon}$) *and such that $\sup_{0<\varepsilon\leq 1}\mathbb{E}F^{2a}(L_T^{\varepsilon})$ and $\sup_{0<\varepsilon\leq 1}\mathbb{E}\left|\sigma^{-1}(\varepsilon)(F(L_T^{\varepsilon})-F(L_T))\right|^{2a}$ are finite, for $a>1$. Moreover, assume that*

$H1$. Condition (3) *in* Theorem 2.1 *holds and there exists a definite positive matrix Σ such that $\lim_{\varepsilon\to 0}\sigma^{-2}(\varepsilon)\Sigma_\varepsilon=\Sigma$.*
$H2$. For $0<\beta<1$, we have $\lim_{\varepsilon\to 0}\sigma(\varepsilon)\sigma^{-1}(\varepsilon^\beta)=0$ and $\lim_{\varepsilon\to 0}\upsilon_\varepsilon\sigma^{-1}(\varepsilon^\beta)=0$.

If we choose $N_1=\lfloor\upsilon_\varepsilon^{-2}\rfloor$ and $N_2=\lfloor\upsilon_\varepsilon^{-2}\sigma^2(\varepsilon^\beta)\rfloor$, then

$$\upsilon_\varepsilon^{-1}\left(Q_\varepsilon-\mathbb{E}F(L_T)\right)\xrightarrow{\mathcal{L}}\mathcal{N}\Big(C_F,\mathrm{Var}(F(L_T))+T\mathbb{E}(\nabla F(L_T).\Sigma\nabla F(L_T))\Big),\quad \text{as } \varepsilon\searrow 0.$$

Proof. At first we write the total error as $Q_\varepsilon-\mathbb{E}F(L_T)=Q_\varepsilon^1+Q_\varepsilon^2+\mathbb{E}F(L_T^{\varepsilon})-\mathbb{E}F(L_T)$, with

$$Q_\varepsilon^1=\frac{1}{N_1}\sum_{i=1}^{N_1}F(\hat{L}_{T,i}^{\varepsilon^\beta})-\mathbb{E}F(L_T^{\varepsilon^\beta})\quad\text{and}$$

$$Q_\varepsilon^2=\frac{1}{N_2}\sum_{i=1}^{N_2}F(L_{T,i}^{\varepsilon})-F(L_{T,i}^{\varepsilon^\beta})-\mathbb{E}\Big[F(L_T^{\varepsilon})-F(L_T^{\varepsilon^\beta})\Big].$$

So, assumption ($\mathrm{WE}_{\upsilon_\varepsilon}$) yields the convergence of $\upsilon_\varepsilon^{-1}\left(\mathbb{E}F(L_T^{\varepsilon})-\mathbb{E}F(L_T)\right)$ towards C_F as ε goes to zero and following step by step the proof of Theorem 3.1 the convergence in law of $\upsilon_\varepsilon^{-1}Q_\varepsilon^1$ to the normal distribution $\mathcal{N}(0,\mathrm{Var}(F(L_T)))$ is easily obtained. Concerning the term Q_ε^2, we plan to use the Lindeberg–Feller central limit theorem for independent random variables (see e.g. [6]). We set $X_{i,\varepsilon}:=\frac{\upsilon_\varepsilon^{-1}}{N_2}\Big(F(L_{T,i}^{\varepsilon})-F(L_{T,i}^{\varepsilon^\beta})-\Big(\mathbb{E}F(L_T^{\varepsilon})-\mathbb{E}F(L_T^{\varepsilon^\beta})\Big)\Big)$ and we split the proof into two steps.

Step 1. It is straightforward that $\sum_{i=1}^{N_2}\mathbb{E}(X_{i,\varepsilon}^2)\underset{\varepsilon\to 0}{\simeq}\sigma^{-2}(\varepsilon^\beta)\mathrm{Var}(F(L_T^{\varepsilon})-F(L_T^{\varepsilon^\beta}))$. Now applying Taylor–Young's expansion to the real valued $\mathscr{C}^1$ function F we get

$$F(L_T^{\varepsilon})-F(L_T^{\varepsilon^\beta})=\nabla F(L_T^{\varepsilon^\beta}).(L_T^{\varepsilon}-L_T^{\varepsilon^\beta})+(L_T^{\varepsilon}-L_T^{\varepsilon^\beta}).\epsilon(L_T^{\varepsilon}-L_T^{\varepsilon^\beta}),$$

where $\epsilon(L_T^{\varepsilon}-L_T^{\varepsilon^\beta})\xrightarrow{a.s.}0$ as $\varepsilon\to 0$. Now, by applying twice Theorem 2.1 to $L_T^{\varepsilon}-L_T$ and $L_T-L_T^{\varepsilon^\beta}$ and thanks to assumption $H2$ we obtain $\sigma^{-1}(\varepsilon^\beta)\big(L_T^{\varepsilon}-L_T^{\varepsilon^\beta}\big)\underset{\varepsilon\to 0}{\xrightarrow{\mathcal{L}}}\Sigma^{1/2}W_T$. Since $L_T^{\varepsilon^\beta}$

is independent from $L_T^\varepsilon - L_T^{\varepsilon^\beta}$ and $\nabla F(L_T^{\varepsilon^\beta}) \xrightarrow[\varepsilon\to 0]{a.s.} \nabla F(L_T)$, we obtain

$$\sigma^{-1}(\varepsilon^\beta)\left(F(L_T^\varepsilon) - F(L_T^{\varepsilon^\beta})\right) \xrightarrow[\varepsilon\to 0]{\mathcal{L}} \nabla F(L_T).\Sigma^{1/2}W_T. \tag{12}$$

For the second term, using the tightness of $\sigma^{-1}(\varepsilon^\beta)\big(L_T^\varepsilon - L_T^{\varepsilon^\beta}\big)$ we deduce that $\sigma^{-1}(\varepsilon^\beta)\big(L_T^\varepsilon - L_T^{\varepsilon^\beta}\big)\epsilon(L_T^\varepsilon - L_T^{\varepsilon^\beta}) \xrightarrow[\varepsilon\to 0]{a.s.} 0$. Thanks to the inequality $|x+y|^{2a} \le 2^{2a-1}(|x|^{2a} + |y|^{2a})$, for any $x, y \in \mathbb{R}$, we get $\sup_{0<\varepsilon\le 1}\mathbb{E}\left|\sigma^{-1}(\varepsilon)(F(L_T^\varepsilon) - F(L_T))\right|^{2a} < +\infty$ and $\lim_{\varepsilon\to 0}\sigma(\varepsilon)\sigma^{-1}(\varepsilon^\beta) = 0$ we deduce the uniform integrability of $\sigma^{-2}(\varepsilon^\beta)|F(L_T^\varepsilon) - F(L_T^{\varepsilon^\beta})|^2$. Therefore, we obtain the first condition needed to apply the Lindeberg–Feller central limit theorem

$$\lim_{\varepsilon\to 0}\sum_{i=1}^{N_2}\mathbb{E}(X_{i,\varepsilon})^2 = \mathrm{Var}(\nabla F(L_T).\Sigma^{1/2}W_T) = T\mathbb{E}(\nabla F(L_T).\Sigma\nabla F(L_T)).$$

Step 2. For the Lyapunov condition, let $1 < a' < a$, we get by standard evaluations

$$\sum_{i=1}^{N_2}\mathbb{E}|X_{i,\varepsilon}|^{2a'} \le C\upsilon_\varepsilon^{2(a'-1)}\sigma^{-2(a'-1)}(\varepsilon^\beta)\mathbb{E}\left|\sigma^{-1}(\varepsilon^\beta)(F(L_T^\varepsilon) - F(L_T^{\varepsilon^\beta}))\right|^{2a'},$$

with $C > 0$.

Once again we use the convergence in distribution given by relation (12) and the uniform integrability property $\sup_{0<\varepsilon\le 1}\mathbb{E}\left|\sigma^{-1}(\varepsilon^\beta)(F(L_T^\varepsilon) - F(L_T^{\varepsilon^\beta}))\right|^{2a} < +\infty$ to deduce the convergence of $\mathbb{E}\left|\sigma^{-1}(\varepsilon^\beta)(F(L_T^\varepsilon) - F(L_T^{\varepsilon^\beta}))\right|^{2a'}$ towards $\mathbb{E}\left|\nabla F(L_T).\Sigma^{1/2}W_T\right|^{2a'}$. Finally, since $\lim_{\varepsilon\to 0}\upsilon_\varepsilon\sigma^{-1}(\varepsilon^\beta) = 0$, we conclude that $\lim_{\varepsilon\to 0}\sum_{i=1}^{N_2}\mathbb{E}|X_{i,\varepsilon}|^{2a'} = 0$ with $a' > 1$. This gives the asymptotic normality of Q_ε^2 and completes the proof. □

Now, we get back to the exponential Lévy model setting introduced before Corollary 3.1 where $F(x) = f(e^{x_1}, \ldots, e^{x_d})$ for a given $\mathscr{C}^1$ Lipschitz continuous function f. Our aim is to deduce in this setting a central limit theorem for SR method.

Corollary 3.2. *Assume that $\int_{|z|>1} e^{2a|z|}\nu(dz)$ is finite for $a > 1$. In the setting of an exponential Lévy model there is $C > 0$ such that $\left|\mathbb{E}F(L_T) - \mathbb{E}F(L_T^\varepsilon)\right| \le C\sigma(\varepsilon)$. Moreover, assume that for $0 < \beta < 1$ there exists $0 < \eta < 2$ such that $\lim_{\varepsilon\to 0}\sigma^{1-\eta/2}(\varepsilon)\sigma^{-1}(\varepsilon^\beta) = 0$, $\sigma(\varepsilon) > \varepsilon$ for all $0 < \varepsilon < 1$ and condition $H1$ of Theorem 3.2 is satisfied. Then, if we choose $N_1 = \lfloor\sigma^{-2+\eta}(\varepsilon)\rfloor$ and $N_2 = \lfloor\sigma^{-2+\eta}(\varepsilon)\sigma^2(\varepsilon^\beta)\rfloor$ we obtain*

$$\sigma^{-1+\eta/2}(Q_\varepsilon - \mathbb{E}F(L_T)) \xrightarrow{\mathcal{L}} \mathcal{N}\Big(0, \mathrm{Var}(F(L_T)) + T\mathbb{E}(\nabla F(L_T).\Sigma\nabla F(L_T))\Big),$$

as $\varepsilon \searrow 0$.

Note that, this corollary gives a large class of functions for which assumptions involving F in Theorem 3.2 are satisfied. In addition, for these functions condition $H2$ is equivalent to check $\lim_{\varepsilon\to 0}\sigma^{1-\eta/2}(\varepsilon)\sigma^{-1}(\varepsilon^\beta) = 0$ since $\upsilon_\varepsilon = \sigma^{1-\eta/2}(\varepsilon)$. For the d-dimensional CGMY process condition $H1$ is already checked when $Y_i = Y \in (0, 2)$, $i \in \{1, \ldots, d\}$, see Section 2. Further, as mentioned after Corollary 3.1, we need to choose $M_i > 2$ and $G_i > 2$, $i \in \{1, \ldots, d\}$. Finally, condition $\lim_{\varepsilon \to 0} \sigma^{1-\eta/2}(\varepsilon)\sigma^{-1}(\varepsilon^\beta) = 0$ is satisfied as soon as $\eta < 2(1 - \beta)$.

Proof. According to Theorem 3.2 and Corollary 3.1 we only need to check that assumption $\sup_{0<\varepsilon\le 1} \mathbb{E}\left|\sigma^{-1}(\varepsilon)(F(L_T^\varepsilon) - F(L_T))\right|^{2a} < \infty$ is satisfied. Since f is Lipschitz it is sufficient to find an upper bound for $\mathbb{E}\left|e^{L_T^\varepsilon} - e^{L_T}\right|^{2a}$. To do so, we use the independence of L_T^1 and the couple $(L_T - L_T^1, L_T^\varepsilon - L_T^1)$ and Cauchy–Schwarz's inequality to get

$$\mathbb{E}\left|e^{L_T^\varepsilon} - e^{L_T}\right|^{2a} \le \mathbb{E}e^{2a|L_T^1|}\left\||L_T - L_T^\varepsilon|^{2a}\right\|_2 \left(\left\|e^{2a|L_T - L_T^1|}\right\|_2 + \left\|e^{2a|L_T^\varepsilon - L_T^1|}\right\|_2\right).$$

By the same arguments given in the proof of Corollary 3.1 we have the finiteness of $\mathbb{E}e^{2a|L_T^1|}$, $\left\|e^{2a|L_T - L_T^1|}\right\|_2$ and $\sup_{0<\varepsilon\le 1}\left\|e^{2a|L_T^\varepsilon - L_T^1|}\right\|_2$. Combining all these results together with assumption (SE) we deduce the existence of a constant $C > 0$ not depending on ε such that

$$\mathbb{E}\left|\sigma^{-1}(\varepsilon)(F(L_T^\varepsilon) - F(L_T))\right|^{2a} \le C\sigma^{-2a}(\varepsilon)\sigma_0^{2a}(\varepsilon).$$

This completes the proof since $\sigma_0(\varepsilon) = \sigma(\varepsilon)$, for $0 < \varepsilon < 1$. □

3.3. Complexity analysis

Thanks to the above limit results we are able now to provide a complexity analysis for both MC and SR algorithm. To keep things simple, we consider the particular case $d = 1$, $v_\varepsilon = \sigma(\varepsilon)$ and we assume that the measure ν has a density of the form $L(x)/|x|^{Y+1}$ for a small x, where $L(x)$ is a slowly varying as $x \to 0$ and $Y \in (0, 2)$. Observe that the positive (resp. negative) part of the approximation $(L_t^\varepsilon)_{0\le t\le T}$ is essentially a compound Poisson process with intensity $\nu([\varepsilon, +\infty))$ (resp. $\nu((-\infty, -\varepsilon])$). Then, the cost necessary of a single simulation is random, with expectation of order $\mathcal{K}(\varepsilon) = \nu(|x| \ge \varepsilon)$. Hence, according to Theorem 3.1 the time complexity of the MC method necessary to achieve a total error of order $\sigma(\varepsilon)$ is random with expectation of order

$$C_{MC} = \mathcal{K}(\varepsilon)N \underset{\varepsilon\to 0}{\simeq} \mathcal{K}(\varepsilon)\sigma^{-2}(\varepsilon).$$

In the same way, thanks to Theorem 3.2 the time complexity of the SR method necessary to achieve a total error of order $\sigma(\varepsilon)$ is random with expectation of order

$$C_{SR} = \mathcal{K}(\varepsilon^\beta)N_1 + \mathcal{K}(\varepsilon)N_2 \underset{\varepsilon\to 0}{\simeq} \left(\mathcal{K}(\varepsilon^\beta) + \mathcal{K}(\varepsilon)\sigma^2(\varepsilon^\beta)\right)\sigma^{-2}(\varepsilon).$$

By Karamata's theorem (see e.g. Bingham, Goldie and Teugels [7] or Feller [17])

$$\sigma^2(\varepsilon) = \int_{-\varepsilon}^{\varepsilon} |x|^{1-Y} L(x)dx \underset{\varepsilon\to 0}{\simeq} \frac{L(\varepsilon) + L(-\varepsilon)}{2 - Y}\varepsilon^{2-Y}.$$

Similarly we have

$$\mathcal{K}(\varepsilon) \underset{\varepsilon\to 0}{\simeq} \frac{L(\varepsilon) + L(-\varepsilon)}{Y}\varepsilon^{-Y}.$$

Please cite this article in press as: M.B. Alaya, et al., Importance sampling and statistical Romberg method for Lévy processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2015.12.008

Consequently, we compute the time complexity ratio given by

$$\frac{C_{SR}}{C_{MC}} \underset{\varepsilon\to 0}{\simeq} \frac{L(\varepsilon^\beta)+L(-\varepsilon^\beta)}{L(\varepsilon)+L(-\varepsilon^\beta)}\varepsilon^{Y(1-\beta)} + \frac{L(\varepsilon^\beta)+L(-\varepsilon^\beta)}{2-Y}\varepsilon^{\beta(2-Y)}.$$

If $L(\varepsilon)$ is constant in the neighborhood of zero, like for the CGMY model (see relation (2)), then we easily get

$$\frac{C_{SR}}{C_{MC}} = O\left(\varepsilon^{Y(1-\beta)} + \varepsilon^{\beta(2-Y)}\right).$$

Optimizing the order of this last quantity yields $\beta = Y/2$ which leads us to a gain of a complexity of order $\varepsilon^{Y(Y/2-1)}$ that asymptotically increases as soon as ε becomes small.

4. Importance sampling and statistical Romberg method

Let $\{L_t; t \geq 0\}$ be a Lévy process in $\mathbb{R}^d$ under the probability $\mathbb{P}$ with generating triplet $(\gamma, 0, \nu)$. We define the set

$$\Theta_1 := \left\{\theta \in \mathbb{R}^d : \mathbb{E}[e^{\theta.L_t}] < +\infty\right\} = \left\{\theta \in \mathbb{R}^d : \int_{|x|>1} e^{\theta.x}\nu(dx) < \infty\right\}, \tag{13}$$

where the second equality holds by Theorem 25.3 in [31]. Thanks to the convexity of the exponential function it is straightforward that the set Θ_1 is convex. In view to use importance sampling routine, based on exponential tilting, we define the family of $\{\mathbb{P}_\theta, \theta \in \Theta_1\}$, as all the equivalent probability measures with respect to $\mathbb{P}$ such that

$$\frac{d\mathbb{P}_\theta}{d\mathbb{P}}\Big|_{\mathcal{F}_t} = \frac{e^{\theta.L_t}}{\mathbb{E}[e^{\theta.L_t}]} = e^{\theta.L_t - t\kappa(\theta)}$$

where κ denotes the cumulant generating function given by $\kappa(\theta) = \ln \mathbb{E}\left[e^{\theta.L_1}\right]$. Under $\mathbb{P}_\theta$, the stochastic process $\{L_t; t \geq 0\}$ is still a Lévy process with the exponential tilted triplet $(\gamma_\theta, 0, \nu_\theta)$ where $\gamma_\theta = \gamma + \int_{|x|\leq 1} x(\nu_\theta - \nu)(dx)$ and $\nu_\theta(dx) = e^{\theta.x}\nu(dx)$ (see e.g. Cont and Tankov [10]). For $\varepsilon, \beta \in (0, 1)$, it is easy to check that under $\mathbb{P}_\theta$ the components of the family $\{L_t, L_t^\varepsilon, L_t^{\varepsilon^\beta}; t \geq 0\}$ are still three Lévy processes with characteristic triplets given respectively by $(\gamma_\theta, 0, \nu_\theta)$, $(\gamma_{\varepsilon,\theta}, 0, \nu_{\varepsilon,\theta})$ and $(\gamma_{\varepsilon^\beta,\theta}, 0, \nu_{\varepsilon^\beta,\theta})$ where for $\alpha \in (0, 1)$, $\gamma_{\alpha,\theta} = \gamma + \int_{\alpha<|x|\leq 1} x(\nu_\theta - \nu)(dx)$ and $\nu_{\alpha,\theta}(dx) = e^{\theta.x}\mathbf{1}_{\{|x|>\alpha\}}\nu(dx)$. It is worth noticing that under $\mathbb{P}_\theta$, $L - L^\varepsilon$ is still independent of L^ε and similarly $L^\varepsilon - L^{\varepsilon^\beta}$ is still independent of L^{ε^β}.

Then, we introduce the family of Lévy processes $\{L_t^\theta, L_t^{\varepsilon,\theta}, L_t^{\varepsilon^\beta,\theta}; t \geq 0\}$ such that under $\mathbb{P}$ it has the same law as $\{L_t, L_t^\varepsilon, L_t^{\varepsilon^\beta}; t \geq 0\}$ under $\mathbb{P}_\theta$ and we get

$$\mathbb{E}\left[F(L_T)\right] = \mathbb{E}_\theta\left[F(L_T)e^{-\theta.L_T + T\kappa(\theta)}\right] = \mathbb{E}\left[F(L_T^\theta)e^{-\theta.L_T^\theta + T\kappa(\theta)}\right].$$

Taking $F = 1$ in the above identity we obtain that $e^{T\kappa(\theta)} = \left(\mathbb{E}(e^{-\theta.L_T^\theta})\right)^{-1}$. Moreover, using the independence between $L_T^\theta - L_T^{\varepsilon,\theta}$ and $L_T^{\varepsilon,\theta}$ we obtain that $\mathbb{E}\left[e^{-\theta.\left(L_T^\theta - L_T^{\varepsilon,\theta}\right) + T(\kappa(\theta) - \kappa_\varepsilon(\theta))}\right] = 1$ with $e^{T\kappa_\varepsilon(\theta)} = \left(\mathbb{E}(e^{-\theta.L_T^{\varepsilon,\theta}})\right)^{-1}$. Note that $\kappa_\varepsilon(\theta)$ is also equal to $\ln \mathbb{E}\left[e^{\theta.L_1^\varepsilon}\right]$. Then, we deduce that

$$\mathbb{E}\left[F(L_T^\varepsilon)\right] = \mathbb{E}\left[F(L_T^{\varepsilon,\theta})e^{-\theta.L_T^\theta + T\kappa(\theta)}\right] = \mathbb{E}\left[F(L_T^{\varepsilon,\theta})e^{-\theta.L_T^{\varepsilon,\theta} + T\kappa_\varepsilon(\theta)}\right].$$

Similarly, we have $\mathbb{E}\left[e^{-\theta.\left(L_T^{\varepsilon,\theta}-L_T^{\varepsilon^\beta,\theta}\right)+T\left(\kappa_\varepsilon(\theta)-\kappa_{\varepsilon^\beta}(\theta)\right)}\right]=1$ since $L_T^{\varepsilon,\theta}-L_T^{\varepsilon^\beta,\theta}$ and $L_T^{\varepsilon^\beta,\theta}$ are independent. So, by the same arguments we deduce that

$$\mathbb{E}\left[F(L_T^{\varepsilon^\beta})\right]=\mathbb{E}\left[F(L_T^{\varepsilon^\beta,\theta})e^{-\theta.L_T^{\varepsilon^\beta,\theta}+T\kappa_{\varepsilon^\beta}(\theta)}\right]=\mathbb{E}\left[F(L_T^{\varepsilon^\beta,\theta})e^{-\theta.L_T^{\varepsilon,\theta}+T\kappa_\varepsilon(\theta)}\right].$$

Hence, we write now for $\theta_1,\theta_2\in\Theta_1$

$$\mathbb{E}\left[F(L_T^\varepsilon)\right]=\mathbb{E}\left[F(L_T^{\varepsilon^\beta,\theta_1})e^{-\theta_1.L_T^{\varepsilon^\beta,\theta_1}+T\kappa_{\varepsilon^\beta}(\theta_1)}\right]+\mathbb{E}\left[\left(F(L_T^{\varepsilon,\theta_2})-F(L_T^{\varepsilon^\beta,\theta_2})\right)e^{-\theta_2.L_T^{\varepsilon,\theta_2}+T\kappa_\varepsilon(\theta_2)}\right]. \tag{14}$$

The idea now is to approximate independently each expectation by the associated empirical mean which leads to a new version of the statistical Romberg method

$$\frac{1}{N_1}\sum_{k=1}^{N_1}F(L_{T,k}^{\varepsilon^\beta,\theta_1})e^{-\theta_1.L_{T,k}^{\varepsilon^\beta,\theta_1}+T\kappa_{\varepsilon^\beta}(\theta_1)}+\frac{1}{N_2}\sum_{k=1}^{N_2}(F(L_{T,k}^{\varepsilon,\theta_2})-F(L_{T,k}^{\varepsilon^\beta,\theta_2}))e^{-\theta_2.L_{T,k}^{\varepsilon,\theta_2}+T\kappa_\varepsilon(\theta_2)}.$$

Mimicking the proof of Theorem 3.2 we establish a central limit theorem with limit variance $\mathrm{Var}(F(L_T^{\theta_1})e^{-\theta_1L_T^{\theta_1}+T\kappa(\theta_1)})+T\mathbb{E}((\nabla F(L_T^{\theta_2}).\Sigma\nabla F(L_T^{\theta_2}))e^{-2\theta_2L_T^{\theta_2}+2T\kappa(\theta_2)})$. Since $L_T^{\theta_1}$ (resp. $L_T^{\theta_2}$) under $\mathbb{P}$ has the same law as L_T under $\mathbb{P}_{\theta_1}$ (resp. $\mathbb{P}_{\theta_2}$) we rewrite this variance using once again the Esscher transform as

$$\mathbb{E}\left[F^2(L_T)e^{-\theta_1.L_T+T\kappa(\theta_1)}\right]-[\mathbb{E}F(L_T)]^2+T\mathbb{E}\left[(\nabla F(L_T).\Sigma\nabla F(L_T))e^{-\theta_2L_T+T\kappa(\theta_2)}\right].$$

Hence, let us introduce for $i\in\{1,2\}$,

$$v_i(\theta):=\mathbb{E}\left[F_i(L_T)e^{-\theta L_T+T\kappa(\theta)}\right],\ \text{with } F_1\equiv F^2 \text{ and } F_2\equiv\nabla F.\Sigma\nabla F. \tag{15}$$

Our aim now is to minimize separately these two quantities. To do so, for $i\in\{1,2\}$, we introduce a first set

$$\Theta_{i,2}:=\Theta_1\cap\left\{\theta\in\mathbb{R}^d:\mathbb{E}\left[F_i(L_T)e^{-\theta.L_T}\right]<+\infty\right\}$$

to ensure the existence of $v_i(\theta)$ and a second set

$$\Theta_{i,3}:=\Theta_{i,2}\cap\left\{\theta\in\mathbb{R}^d:\mathbb{E}\left[|L_T|^2F_i(L_T)e^{-\theta.L_T}\right]<+\infty\right\}$$

to ensure the existence of the first and second derivatives of $v_i(\theta)$. In the sequel, we will assume that $Leb(\Theta_{i,3})>0$, for $i\in\{1,2\}$, so that Θ_1, $\Theta_{i,2}$ and $\Theta_{i,3}$ are nonempty sets. In addition, the convexity of sets $\Theta_{i,2}$ and $\Theta_{i,3}$ can be proved in a similar manner to the proof of Lemma 2.2 in [21]. Moreover, we prove the convexity of v_i, $i\in\{1,2\}$.

Proposition 4.1. *Let $i\in\{1,2\}$. Assume $\mathbb{P}(F_i(L_T)\neq 0)>0$. Then, $\theta\mapsto v_i(\theta)$ is a $\mathscr{C}^2$ strictly convex function on $\Theta_{i,3}$ and $\nabla v_i(\theta)=\mathbb{E}[H_i(\theta,L_T)]$ where*

$$H_i(\theta,L_T)=(T\nabla\kappa(\theta)-L_T)F_i(L_T)\exp(-\theta.L_T+T\kappa(\theta)). \tag{16}$$

Proof. For a fixed $i \in \{1, 2\}$, the function $\theta \mapsto F_i(L_T)e^{-\theta L_T + T\kappa(\theta)}$ is almost surely differentiable on Θ_1 with a first derivative equal to $H_i(\theta, L_T)$. Further, according to the properties of the moment generating function, the function $\theta \mapsto v_i(\theta)$ is finite for $\theta \in \Theta_{i,2}$ and is differentiable with $\nabla v_i(\theta) = \mathbb{E}[H_i(\theta, L_T)]$ provided that $\mathbb{E}[|H_i(\theta, L_T)|]$ is finite. Using Hölder's inequality, this last condition is satisfied as soon as $\theta \in \Theta_{i,3}$. In the same way, we prove that v_i is of class $\mathscr{C}^2$ on $\Theta_{i,3}$ and we get for all $u \in \mathbb{R}^d \setminus \{0\}$,

$$u.\text{Hess}(v_i(\theta))u = \mathbb{E}\left[\left(u.\text{Hess}(\kappa(\theta))u + (u.(T\nabla\kappa(\theta) - L_T))^2\right) F_i(L_T)e^{-\theta.L_T + T\kappa(\theta)}\right].$$

Note that $\text{Hess}(\kappa(\theta))$ is nothing but the variance–covariance matrix of the random vector L_T under the probability measure $\mathbb{P}_\theta$ and it is clearly definite positive. Finally, since $\mathbb{P}(F_i(L_T) \neq 0) > 0$, we conclude that v_i is strictly convex on $\Theta_{i,3}$. □

For $\varepsilon > 0$, the same result holds for the approximated Lévy process $(L_t^\varepsilon)_{t\geq 0}$ by considering the associated sets Θ_1^ε, $\Theta_{i,2}^\varepsilon$ and $\Theta_{i,3}^\varepsilon$ and functions κ_ε and $v_{i,\varepsilon}$, $i \in \{1, 2\}$, with the canonical filtration $(\mathcal{F}_t^\varepsilon)_{0\leq t\leq T}$ defined by $\mathcal{F}_t^\varepsilon = \sigma(L_s^\varepsilon, s \leq t)$. In the sequel, we assume that $Leb(\Theta_{i,3}^\varepsilon) > 0$.

Proposition 4.2. *Let $i \in \{1, 2\}$. Assume $\mathbb{P}(F_i(L_T^\varepsilon) \neq 0) > 0$ then the function $v_{i,\varepsilon}(\theta) = \mathbb{E}\left[F_i(L_T^\varepsilon)e^{-\theta L_T^\varepsilon + T\kappa_\varepsilon(\theta)}\right]$ is of class $\mathscr{C}^2$ and strictly convex on $\Theta_{i,3}^\varepsilon$ with $\nabla v_{i,\varepsilon}(\theta) = \mathbb{E}\left[H_i(\theta, L_T^\varepsilon)\right]$.*

Now, let us introduce for $i \in \{1, 2\}$

$$\theta_{i,\varepsilon}^* := \arg\min_{\theta\in\Theta_{i,3}^\varepsilon} v_{i,\varepsilon}(\theta) \quad \text{and} \quad \theta_i^* := \arg\min_{\theta\in\Theta_{i,3}} v_i(\theta). \tag{17}$$

Our aim now is to study for $i \in \{1, 2\}$ the convergence of $\theta_{i,\varepsilon}^*$ towards θ_i^* as ε tends to zero. For $q > 1$, we define the set

$$\Theta_q := \left\{\theta \in \mathbb{R}^d : \int_{|x|>1} |x|^{2q} e^{-q\theta.x} \nu(dx) < +\infty\right\}. \tag{18}$$

Remark. 1. It is worth noticing that for $0 \leq q' \leq 2q$ and $\theta \in \Theta_q$ we have $\int_{|x|>1} |x|^{q'} e^{-q\theta.x} \nu(dx) < +\infty$. We also have $\Theta_{q_2} \subset \Theta_{q_1}$ for all $q_1 \leq q_2$.
2. Further, for $i \in \{1, 2\}$, if $\mathbb{E}\left[F_i^a(L_T)\right]$, for any $a > 1$, is finite then by Hölder's inequality we easily get $\Theta_q \subset \Theta_{i,3}$ for all $q \geq a/a - 1$. The same result holds for the approximated Lévy process. Indeed, for $\varepsilon > 0$, we have $\Theta_q \subset \Theta_{i,3}^\varepsilon$ provided that $\mathbb{E}\left[F_i^a(L_T^\varepsilon)\right] < \infty$.

According the above remark, choosing $\theta \in \Theta_q$ with $q \geq a/a - 1$ ensures that θ will belong to the domain of convexity of both v_i and $v_{i,\varepsilon}$. On the other hand it also guarantees the finiteness of the quantity $\int_{|x|>1} |x|^q e^{-q\theta.x} \nu(dx)$ which will be needed in each proof assuming condition $\theta \in \Theta_q$.

In what follows, let $\mathring{E}$ denote the set of all interior points of a given set E. We have the following result.

Theorem 4.1. *Let $i \in \{1, 2\}$. Suppose that $x \mapsto F_i(x)$ is continuous, that is for the case $i = 1$ the function F is continuous and for $i = 2$ the function F is of class $\mathscr{C}^1$. Moreover, assume $\mathbb{P}(F_i(L_T) \neq 0) > 0$, $\mathbb{P}(F_i(L_T^\varepsilon) \neq 0) > 0$ for all $\varepsilon > 0$ and there exists $a > 1$ such that*

$\mathbb{E}\left[F_i^a(L_T)\right]$ and $\sup_{\varepsilon>0}\mathbb{E}\left[F_i^a(L_T^\varepsilon)\right]$ are finite. Let K be a nonempty compact set such that $K\subset\mathring{\Theta}_q$ with $q>\frac{a}{a-1}$ and assume that the sequence $(\theta^*_{i,\varepsilon})_{\varepsilon>0}\in K$. Then,

$$\theta^*_{i,\varepsilon}\longrightarrow\theta^*_i\in K,\qquad \text{as } \varepsilon\to 0.$$

We prove Theorem 4.1 after the following technical lemma.

Lemma 4.1. *For compact set $K\subset\Theta_q$ with $q>1$, we have $\sup_{\theta\in K}\mathbb{E}\big[|L_T^\varepsilon|^q e^{-q\theta.L_T^\varepsilon}\big]$ is uniformly bounded in ε.*

Proof. Let us consider the two independent Lévy processes L^1 and $\tilde{L}^\varepsilon:=L^\varepsilon-L^1$ and the submultiplicative function $g_\theta(x):=(|x|\vee 1)^q e^{-q\theta.x}$. There exists $c_q>0$ depending only on q such that $g_\theta(x+y)\le c_q g_\theta(x)g_\theta(y)$ for any $\theta\in\mathbb{R}^d$ and

$$\mathbb{E}\big[|L_T^\varepsilon|^q e^{-q\theta.L_T^\varepsilon}\big]\le c_q\mathbb{E}\big[g_\theta(\tilde{L}_T^\varepsilon)\big]\mathbb{E}\big[g_\theta(L_T^1)\big].$$

Since the function $\theta\mapsto\mathbb{E}\big[g_\theta(L_T^1)\big]$ is continuous on Θ_q the second expectation on the right hand side is uniformly bounded on $\theta\in K$. Concerning the first expectation, we start by establishing the uniform convergence of $\tilde{\kappa}_\varepsilon$ towards $\tilde{\kappa}$, where $\tilde{\kappa}_\varepsilon$ and $\tilde{\kappa}$ denote the cumulant generating functions of respectively $\tilde{L}^\varepsilon=L^\varepsilon-L^1$ and $\tilde{L}=L-L^1$. According to the Lévy–Khintchine decomposition, we have $\tilde{\kappa}(\theta)-\tilde{\kappa}_\varepsilon(\theta)=\int_{|x|<\varepsilon}(e^{\theta.x}-1-\theta.x)\nu(dx)$ and thanks to Taylor's expansion we get

$$|\tilde{\kappa}(\theta)-\tilde{\kappa}_\varepsilon(\theta)|\le\frac{|\theta|^2}{2}e^{|\theta|}\sigma^2(\varepsilon).\tag{19}$$

This ensures the uniform convergence of the family functions $(\tilde{\kappa}_\varepsilon)_{0<\varepsilon<1}$ on any compact set of $\mathbb{R}^d$. Note that for all $x=(x_1,\dots,x_d)\in\mathbb{R}^d$ we have $(|x|\vee 1)^q\le ce^{|x|}\le c\prod_{j=1}^d(e^{x_j}+e^{-x_j})$ with some $c>0$ depending only on q. This last upper bound can be written as a sum of finite number of exponential functions evaluated at points which are a linear combination of the components of the vector x. Therefore there exists a family of deterministic $\mathbb{R}^d$-valued vectors, $(b_j)_{1\le j\le 2^d}$ such that

$$\mathbb{E}\big[g_\theta(\tilde{L}_T^\varepsilon)\big]\le c\sum_{j=1}^{2^d}\mathbb{E}\big[e^{(b_j-q\theta).L_T^\varepsilon}\big].$$

Each term in the above sum is nothing else $\exp(\tilde{\kappa}_\varepsilon(b_j-q\theta))$ which in turn converges to $\exp(\tilde{\kappa}(b_j-q\theta))$ as ε tends to zero. This gives us the desired claim. □

Proof of Theorem 4.1. Let $i\in\{1,2\}$ and $(\varepsilon_n)_{n\in\mathbb{N}}$ be a sequence decreasing to zero. Note that $(\theta^*_{i,\varepsilon_n})_{n\in\mathbb{N}}$ is a $\mathbb{R}^d$-bounded sequence. So, we only need to prove that for any subsequence $(\theta^*_{i,\varepsilon_{n_k}})_{k\in\mathbb{N}}$, if $\theta^*_{i,\varepsilon_{n_k}}\to\theta^*_{i,\infty}\in\mathbb{R}^d$ then $\theta^*_{i,\infty}=\theta^*_i$. According to Proposition 4.2 we have

$$\nabla v_{i,\varepsilon_{n_k}}(\theta^*_{i,\varepsilon_{n_k}})=\mathbb{E}\left[(\theta^*_{i,\varepsilon_{n_k}}T-L_T^{\varepsilon_{n_k}})F_i(L_T^{\varepsilon_{n_k}})e^{-\theta^*_{i,\varepsilon_{n_k}}.L_T^{\varepsilon_{n_k}}+T\kappa_{\varepsilon_{n_k}}(\theta^*_{i,\varepsilon_{n_k}})}\right]=0.$$

Now, let $\tilde{a}=\frac{aq}{a+q}$, it is easy to check that $1<\tilde{a}<a$, so by applying Hölder's inequality we get

$$\begin{aligned}&\mathbb{E}\left[\left|(\theta^*_{i,\varepsilon_{n_k}}T-L_T^{\varepsilon_{n_k}})F_i(L_T^{\varepsilon_{n_k}})e^{-\theta^*_{i,\varepsilon_{n_k}}.L_T^{\varepsilon_{n_k}}+T\kappa_{\varepsilon_{n_k}}(\theta^*_{i,\varepsilon_{n_k}})}\right|^{\tilde{a}}\right]\\&\quad\le\mathbb{E}^{(a-\tilde{a})/a}\left[\left|(\theta^*_{i,\varepsilon_{n_k}}T-L_T^{\varepsilon_{n_k}})e^{-\theta^*_{i,\varepsilon_{n_k}}.L_T^{\varepsilon_{n_k}}+T\kappa_{\varepsilon_{n_k}}(\theta^*_{i,\varepsilon_{n_k}})}\right|^{\tilde{a}a/(a-\tilde{a})}\right]\mathbb{E}^{\tilde{a}/a}\left[F_i^a(L_T^{\varepsilon_{n_k}})\right].\end{aligned}$$

Please cite this article in press as: M.B. Alaya, et al., Importance sampling and statistical Romberg method for Lévy processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2015.12.008

Note that $\sup_{\varepsilon>0} \mathbb{E}\left[F_i^a(L_T^\varepsilon)\right] < \infty$. Hence, to get the uniform integrability it is sufficient to prove that the first expectation on the right hand side of the above inequality is uniformly bounded on ε_{n_k} and $\theta^*_{i,\varepsilon_{n_k}}$. Indeed, using the almost sure convergence of L_T^ε towards L_T and the continuity of function F_i, we easily get

$$\nabla v_i(\theta^*_{i,\infty}) = \mathbb{E}\left[(\theta^*_{i,\infty}T - L_T)F_i(L_T)e^{-\theta^*_{i,\infty}.L_T+T\kappa(\theta^*_{i,\infty})}\right] = 0$$

and then we complete the proof using the uniqueness of the minimum ensured by Proposition 4.1. Consequently, noticing that $q = \tilde{a}a/(a-\tilde{a})$, it remains now to prove the uniform boundedness of the quantity $\mathbb{E}\left[\left|(\theta^*_{i,\varepsilon_{n_k}}T - L_T^{\varepsilon_{n_k}})e^{-\theta^*_{i,\varepsilon_{n_k}}.L_T^{\varepsilon_{n_k}}+T\kappa_{\varepsilon_{n_k}}(\theta^*_{i,\varepsilon_{n_k}})}\right|^q\right]$. To do so, we establish first the uniform convergence of κ_ε towards κ. According to the decomposition given by relation (2), we have that $\kappa(\theta) - \kappa_\varepsilon(\theta) = \int_{|x|<\varepsilon}(e^{\theta.x} - 1 - \theta.x)\nu(dx)$. By Taylor's expansion we deduce

$$|\kappa(\theta) - \kappa_\varepsilon(\theta)| \le \frac{|\theta|^2}{2}e^{|\theta|}\sigma^2(\varepsilon). \tag{20}$$

Hence, the family functions $(\kappa_\varepsilon)_{0<\varepsilon<1}$ is equicontinuous on any compact subset of Θ_1 and we deduce the convergence of $\kappa_{\varepsilon_{n_k}}(\theta^*_{i,\varepsilon_{n_k}})$ towards $\kappa(\theta^*_{i,\infty})$ when k tends to infinity. Noticing that $-qK \subset \Theta_1$, we use once again the equicontinuity of $(\kappa_\varepsilon)_{0<\varepsilon<1}$ on the compact set $-qK$ to get $\lim_{k\to\infty}\kappa_{\varepsilon_{n_k}}(-q\theta^*_{i,\varepsilon_{n_k}}) = \kappa(-q\theta^*_{i,\infty})$ and then the problem is reduced to prove the uniform boundedness of $\mathbb{E}\left[|L_T^{\varepsilon_{n_k}}|^q e^{-q\theta^*_{i,\varepsilon_{n_k}}.L_T^{\varepsilon_{n_k}}}\right]$ which is ensured by Lemma 4.1. □

5. The adaptive procedure

5.1. Stochastic algorithms

The aim now is to construct family sequences converging almost surely to the optimal limits $\theta^*_{1,\varepsilon}$ and $\theta^*_{2,\varepsilon}$ of the previous section. For this, let $(L_{T,n})_{n\ge1}$ (resp. $(L^\varepsilon_{T,n})_{n\ge1}$, $\varepsilon > 0$), be i.i.d copies of the $\mathbb{R}^d$-valued random variable L_T (resp. L_T^ε). Let K be a compact convex subset of $\Theta_1 \subset \mathbb{R}^d$ with $\{0\} \in K$. For fixed $i \in \{1,2\}$ and $\theta_{i,0} \in K$, we construct recursively the sequences of $\mathbb{R}^d$-valued random variables $(\theta_{i,n})_{n\in\mathbb{N}}$ and $(\theta_{i,\varepsilon,n})_{n\in\mathbb{N}}$ defined by the system

$$\begin{cases} \theta_{i,n+1} = \Pi_K\left[\theta_{i,n} - \gamma_{n+1}H_i(\theta_{i,n}, L_{T,n+1})\right] \\ \theta_{i,\varepsilon,n+1} = \Pi_K\left[\theta_{i,\varepsilon,n} - \gamma_{n+1}H_i(\theta_{i,\varepsilon,n}, L^\varepsilon_{T,n+1})\right] \end{cases} \tag{21}$$

where Π_K is the Euclidean projection onto the constraint set K, H_1 and H_2 are given by relation (16) and the gain sequence $(\gamma_n)_{n\ge1}$ is a decreasing sequence of positive real numbers satisfying

$$\sum_{n=1}^\infty \gamma_n = \infty \quad \text{and} \quad \sum_{n=1}^\infty \gamma_n^2 < \infty. \tag{22}$$

Theorem 5.1. *Let $i \in \{1,2\}$. Assume $\mathbb{P}(F_i(L_T) \neq 0) > 0$, $\mathbb{P}(F_i(L_T^\varepsilon) \neq 0) > 0$ for all $\varepsilon > 0$ and there exists $a > 1$ such that $\mathbb{E}\left[F_i^{2a}(L_T)\right]$ and $\sup_{\varepsilon>0}\mathbb{E}\left[F_i^{2a}(L_T^\varepsilon)\right]$ are finite. Let K be a nonempty compact set such that $K \subset \mathring{\Theta}_{2a/(a-1)}$ then the following assertions hold.*

- *If the unique* $\theta_i^* = \arg\min_{\theta\in\Theta_{i,3}} v_i(\theta)$ *satisfies* $\theta_i^* \in K$ *then the sequence* $\theta_{i,n} \underset{n\to+\infty}{\longrightarrow} \theta_i^*$ *a.s.*
- *If the unique* $\theta_{i,\varepsilon}^* = \arg\min_{\theta\in\Theta_{i,3}^\varepsilon} v_{i,\varepsilon}(\theta)$ *satisfies* $\theta_{i,\varepsilon}^* \in K$ *then the sequence* $\theta_{i,\varepsilon,n} \underset{n\to+\infty}{\longrightarrow} \theta_{i,\varepsilon}^*$ *a.s.*

Proof. Both items can be proved in the same way, so we choose to give the proof only for the first one. According to Theorem A.1 in Laruelle, Lehalle and Pagès [25] on truncated Robbins Monro algorithm (see also Kushner and Yin [24] for more details): in order to prove that $\theta_{i,n}^\varepsilon \underset{n\to+\infty}{\longrightarrow} \theta_{i,\varepsilon}^*$ *a.s.*, we need to check firstly the mean-reverting property, namely

$$\forall\theta \neq \theta_i^* \in K, \quad \langle \nabla v_i(\theta), \theta - \theta_i^*\rangle > 0.$$

This is satisfied using $\nabla v_i(\theta_i^*) = 0$ and the convexity of v_i ensured by Proposition 4.1. Secondly, we have to check the non explosion assumption given by

$$\exists C > 0 \quad \text{such that } \forall\theta \in K, \quad \mathbb{E}\left[|H_i(\theta, L_T)|^2\right] < C(1 + |\theta|^2).$$

In fact, using Hölder's inequality with the couple a and $a/(a-1)$, we obtain

$$\mathbb{E}|H_i(\theta, L_T)|^2 \leq \mathbb{E}^{\frac{1}{a}}\left[F_i^{2a}(L_T)\right]\mathbb{E}^{\frac{a-1}{a}}\left[|T\nabla\kappa(\theta) - L_T|^{2a/(a-1)}e^{-2a/(a-1)\theta.L_T}\right]e^{2T\kappa(\theta)}.$$

Since $\mathbb{E}\left[F_i^{2a}(L_T)\right]$ is finite and $\theta \in K \subset \Theta_{2a/(a-1)}$, we deduce that $\sup_{\theta\in K}\mathbb{E}|H_i(\theta, L_T)|^2 < \infty$ which completes the proof. $\square$

Theorem 5.2. *Considering the sequences given by relation (21), for* $i \in \{1, 2\}$, *we have for all* $n \in \mathbb{N}$

$$\theta_{i,\varepsilon,n} \underset{\varepsilon\to 0}{\longrightarrow} \theta_{i,n} \quad a.s.$$

Proof. We proceed by induction. The base case is trivial and for the inductive step we suppose that for $i \in \{1, 2\}$, $n \in \mathbb{N}$, $\theta_{i,\varepsilon,n}$ converges to $\theta_{i,n}$ *a.s.* as ε goes to 0 and we prove the statement for $n+1$. We have $\theta_{i,\varepsilon,n+1} = \Pi_K\left[\theta_{i,\varepsilon,n} - \gamma_{i+1}H_i(\theta_{i,\varepsilon,n}, L_{T,n+1}^\varepsilon)\right]$. By the continuity of the function H_i given by (16), the almost sure convergence of $L_{T,n+1}^\varepsilon$ to $L_{T,n+1}$ and the continuity of the projection function Π_K, we deduce that $\theta_{i,\varepsilon,n+1}$ converges to $\theta_{i,n+1}$ *a.s.* as ε goes to 0. $\square$

The following corollary follows immediately thanks to Theorems 4.1, 5.1 and 5.2.

Corollary 5.1. *Under assumptions of Theorem 5.1, the constrained algorithm given by routine (21) satisfies for* $i \in \{1, 2\}$

$$\lim_{\substack{\varepsilon\to 0\\ n\to\infty}} \theta_{i,\varepsilon,n} = \lim_{\varepsilon\to 0}(\lim_{n\to\infty}\theta_{i,\varepsilon,n}) = \lim_{n\to\infty}(\lim_{\varepsilon\to 0}\theta_{i,\varepsilon,n}) = \theta_i^*, \quad \mathbb{P}\text{-}a.s. \tag{23}$$

Remark. Suppose for a while that we omit assumptions $\theta_i^* \in K$ and $\theta_{i,\varepsilon}^* \in K$ in Theorem 5.1. According to Theorem 3.2 of Kawai [21] based on Theorem 2.1 of Kushner and Yin [24] there exist $\bar\theta_i$ and $\bar\theta_{i,\varepsilon}$ in K such that $\theta_{i,n} \underset{n\to+\infty}{\longrightarrow} \bar\theta_i$ *a.s.* and $\theta_{i,\varepsilon,n} \underset{n\to+\infty}{\longrightarrow} \bar\theta_{i,\varepsilon}$ *a.s.* Moreover, $v_i(\bar\theta_i) \leq v_i(\theta)$ and $v_{i,\varepsilon}(\bar\theta_{i,\varepsilon}) \leq v_{i,\varepsilon}(\theta)$ for all $\theta \in K$. In this case we can prove that the constrained algorithm given by routine (21) satisfies relation (23) with $\bar\theta_{i,\varepsilon}$ instead of θ_i^*.

5.2. Central limit theorems

In what follows, we consider the filtration $\mathcal{F}_{T,k} = \sigma(L^{\theta}_{t,\ell}, L^{\varepsilon,\theta}_{t,\ell}, 0 < \varepsilon < 1, t \leq T, \ell \leq k, \theta \in \mathbb{R}^d)$, where $(L^{\theta}_{\ell}, L^{\varepsilon,\theta}_{\ell})_{\ell \geq 1}$ are independent copies of $(L^{\theta}, L^{\varepsilon,\theta})$ supposed to be Borel-measurable on $\theta \in \mathbb{R}^d$. Let us assume that there exists a family of sequences $(\theta^{\varepsilon}_k)_{k\geq 0, 0<\varepsilon\leq 1}$ and $(\theta_k)_{k\geq 0}$ satisfying

$$(\mathcal{H}_\theta) \quad \begin{cases} \text{For each } \varepsilon > 0,\ (\theta^{\varepsilon}_k)_{k\geq 0} \text{ and } (\theta_k)_{k\geq 0} \text{ are } (\mathcal{F}_{T,k})_{k\geq 0}\text{-adapted} \\ \lim_{k\to\infty}(\lim_{\varepsilon\to 0}\theta^{\varepsilon}_k) = \lim_{k\to\infty}\theta_k = \lim_{\varepsilon\to 0}(\lim_{k\to\infty}\theta^{\varepsilon}_k) = \lim_{\varepsilon\to 0}\theta^*_{\varepsilon} = \theta^*, \quad \mathbb{P}\text{-a.s.,} \end{cases}$$

with deterministic limits θ^* and θ^*_{ε}.

At first, we start with studying the MC setting. We use the adaptive importance sampling algorithm for the MC method to approximate our initial quantity of interest $\mathbb{E}F(L_T)$ by

$$Q^{\text{ISMC}}_{\varepsilon} = \frac{1}{N}\sum_{k=1}^{N} F(L^{\varepsilon,\theta^{\varepsilon}_{k-1}}_{T,k}) e^{-\theta^{\varepsilon}_{k-1}.L^{\varepsilon,\theta^{\varepsilon}_{k-1}}_{T,k} + T\kappa_{\varepsilon}(\theta^{\varepsilon}_{k-1})}. \tag{24}$$

Our task now is to establish a central limit theorem for the adaptive importance sampling Monte Carlo method (ISMC).

Theorem 5.3. *Let $F : \mathbb{R}^d \to \mathbb{R}$ be a continuous function satisfying assumption $(\mathrm{WE}_{\upsilon_\varepsilon})$ and such that $\sup_{0<\varepsilon\leq 1}\mathbb{E}\left[F^{2a}(L^{\varepsilon}_T)\right] < +\infty$ for $a > 1$. Moreover, assume there exists a double indexed family $(\theta^{\varepsilon}_k)_{k\in\mathbb{N},\varepsilon>0}$ satisfying $(\mathcal{H}_\theta)$ and belonging to some nonempty compact set $K \subset \mathring{\Theta}_q$ with $q > a/(a-1)$. Then, if we choose $N = \lfloor \upsilon_{\varepsilon}^{-2} \rfloor$, the following convergence holds*

$$\upsilon_{\varepsilon}^{-1}\left(Q^{\text{ISMC}}_{\varepsilon} - \mathbb{E}F(L_T)\right) \xrightarrow{\mathcal{L}} \mathcal{N}(C_F, \sigma^2), \quad \text{as } \varepsilon \searrow 0, \tag{25}$$

where $\sigma^2 := \mathbb{E}\left[F^2(L_T)e^{-\theta^*.L_T + T\kappa(\theta^*)}\right] - (\mathbb{E}[F(L_T)])^2$.

It is worth noticing that assumptions involving F are the same as in Theorem 3.1. Moreover, Corollary 3.1 gives a large class of functions for which these assumptions are satisfied. For the d-dimensional CGMY process introduced in Section 2 and for $F(x) = f(e^{x_1}, \dots, e^{x_d})$ with $x = (x_1, \dots, x_d) \in \mathbb{R}^d$, by standard calculations for $q > \frac{a}{a-1}$ the set Θ_q is equal to $(-M\frac{a-1}{a}, G\frac{a-1}{a})$, for any $1 < a \leq \frac{M}{2} \wedge \frac{G}{2}$. Furthermore, if the Lévy measure ν is with a compact support or even with sufficiently decreasing behavior like for the Gaussian measure, then $\Theta_q = \mathbb{R}^d$, for any $q > 1$.

Proof. By assumption $(\mathrm{WE}_{\upsilon_\varepsilon})$ we only need to study the asymptotic behavior of the martingale arrays $(M^{\varepsilon}_k)_{k\geq 1}$ given by $M^{\varepsilon}_k := \frac{\upsilon_{\varepsilon}^{-1}}{N}\sum_{i=1}^{k}\left(F(L^{\varepsilon,\theta^{\varepsilon}_{i-1}}_{T,i})e^{-\theta^{\varepsilon}_{i-1}.L^{\varepsilon,\theta^{\varepsilon}_{i-1}}_{T,i} + T\kappa_{\varepsilon}(\theta^{\varepsilon}_{i-1})} - \mathbb{E}F(L^{\varepsilon}_T)\right)$. To do so, we plan to apply the Lindeberg–Feller central limit theorem for martingales arrays (see Theorem A.1 in the Appendix section). The proof is divided into two steps.

Step 1. The quadratic variation of the martingale arrays $(M^{\varepsilon}_k)_{k\geq 1}$ is given by

$$\begin{aligned}\langle M^{\varepsilon}\rangle_N = {} & \frac{\upsilon_{\varepsilon}^{-2}}{N^2}\sum_{k=1}^{N}\mathbb{E}\left[F^2(L^{\varepsilon,\theta^{\varepsilon}_{k-1}}_{T,k})e^{-2\theta^{\varepsilon}_{k-1}.L^{\varepsilon,\theta^{\varepsilon}_{k-1}}_{T,k} + 2T\kappa_{\varepsilon}(\theta^{\varepsilon}_{k-1})}|\mathcal{F}_{T,k-1}\right] \\ & - \frac{\upsilon_{\varepsilon}^{-2}}{N}\left(\mathbb{E}F(L^{\varepsilon}_T)\right)^2. \end{aligned} \tag{26}$$

Since θ^ε_{k-1} is $\mathcal{F}_{T,k-1}$-measurable and $(L^{\varepsilon,\theta}_{T,k})_{\theta\in\Theta_q} \perp\!\!\!\perp \mathcal{F}_{T,k-1}$, by Esscher transform we obtain

$$\langle M^\varepsilon\rangle_N = \frac{\upsilon_\varepsilon^{-2}}{N^2}\sum_{k=1}^N \gamma_\varepsilon(\theta^\varepsilon_{k-1})e^{T\kappa_\varepsilon(\theta^\varepsilon_{k-1})} - \frac{\upsilon_\varepsilon^{-2}}{N}\left(\mathbb{E}F(L^\varepsilon_T)\right)^2,$$

where for all $\theta\in\Theta_q$, $\gamma_\varepsilon(\theta) = \mathbb{E}\left[F^2(L^\varepsilon_T)e^{-\theta.L^\varepsilon_T}\right]$. On the one hand, using assumption (WE$_{\upsilon_\varepsilon}$), we have $\lim_{\varepsilon\to0}\mathbb{E}F(L^\varepsilon_T) = \mathbb{E}F(L_T)$. On the other hand, thanks to relation (20) we have the uniform equicontinuity of the family $(\kappa_\varepsilon)_{\varepsilon>0}$ on the compact set K. So, we only need to check this last property for the family $(\gamma_\varepsilon)_{\varepsilon>0}$ in view to use after that Lemma A.1 and then deduce the convergence of $\langle M^\varepsilon\rangle_N$ towards $\gamma(\theta^*) - (\mathbb{E}F(L_T))^2$ as $\varepsilon\searrow 0$, where $\gamma(\theta) := \mathbb{E}\left[F^2(L_T)e^{-\theta.L_T}\right]$.

Thus, it remains to prove the uniform equicontinuity of the family functions $(\gamma_\varepsilon)_{\varepsilon>0}$ defined on the compact set K. Using Hölder's inequality and the assumption $\sup_{\varepsilon>0}\mathbb{E}\left[F^{2a}(L^\varepsilon_T)\right] < +\infty$, there exists $c_1>0$ not depending on ε such that

$$\begin{aligned}\left|\gamma_\varepsilon(\theta)-\gamma_\varepsilon(\theta')\right| &\le \mathbb{E}\left[F^2(L^\varepsilon_T)\left|e^{-\theta.L^\varepsilon_T} - e^{-\theta'.L^\varepsilon_T}\right|\right]\\ &\le c_1\mathbb{E}^{1/q}\left[\left|e^{-\theta.L^\varepsilon_T} - e^{-\theta'.L^\varepsilon_T}\right|^q\right].\end{aligned}$$

By Taylor's expansion and standard calculations we easily get

$$|e^{-\theta.L^\varepsilon_T} - e^{-\theta'.L^\varepsilon_T}|^q \le |\theta-\theta'|^q\int_0^1 |L^\varepsilon_T|^q e^{-q(u\theta+(1-u)\theta').L^\varepsilon_T}du.$$

Therefore, we have

$$\left|\gamma_\varepsilon(\theta)-\gamma_\varepsilon(\theta')\right| \le c_1|\theta-\theta'|\sup_{\theta\in K}\mathbb{E}^{1/q}\left[|L^\varepsilon_T|^q e^{-q\theta.L^\varepsilon_T}\right].$$

Hence, according to Lemma 4.1 there exists a constant $c_2>0$ also not depending on and ε such that

$$\left|\gamma_\varepsilon(\theta)-\gamma_\varepsilon(\theta')\right| \le c_2|\theta-\theta'|. \tag{27}$$

This completes the proof of the first step.

Step 2. We check now the Lyapunov condition given by assumption A3 in Theorem A.1. So, let $\tilde a = \frac{aq+a}{2a+q}$, it is easy to check that $1<\tilde a<a$. Once again using the measurability properties of the family $(L^{\varepsilon,\theta}_{T,k})_{\theta\in\Theta_q}$ and the sequence $(\theta^\varepsilon_k)_{k\ge0}$, we get using the Esscher transform

$$\begin{aligned}&\sum_{k=1}^N\mathbb{E}\left[\left|M^\varepsilon_k - M^\varepsilon_{k-1}\right|^{2\tilde a}|\mathcal{F}_{T,k-1}\right]\\ &\quad= \frac{\upsilon_\varepsilon^{-2\tilde a}}{N^{2\tilde a}}\sum_{k=1}^N\mathbb{E}\left[\left|F(L^{\varepsilon,\theta^\varepsilon_{k-1}}_{T,k})e^{-\theta^\varepsilon_{k-1}.L^{\varepsilon,\theta^\varepsilon_{k-1}}_{T,k}+T\kappa_\varepsilon(\theta^\varepsilon_{k-1})} - \mathbb{E}F(L^\varepsilon_T)\right|^{2\tilde a}|\mathcal{F}_{T,k-1}\right]\\ &\quad\le \frac{2^{2\tilde a-1}\upsilon_\varepsilon^{-2\tilde a}}{N^{2\tilde a}}\sum_{k=1}^N\gamma_{\tilde a,\varepsilon}(\theta^\varepsilon_{k-1})e^{(2\tilde a-1)T\kappa_\varepsilon(\theta^\varepsilon_{k-1})} + \frac{2^{2\tilde a-1}\upsilon_\varepsilon^{-2\tilde a}}{N^{2\tilde a-1}}\left|\mathbb{E}F(L^\varepsilon_T)\right|^{2\tilde a},\end{aligned}$$

where for all $\theta\in\Theta_q$, $\gamma_{\tilde a,\varepsilon}(\theta) = \mathbb{E}\left[F^{2\tilde a}(L^\varepsilon_T)e^{-(2\tilde a-1)\theta.L^\varepsilon_T}\right]$. Then, by Hölder's inequality we get

$$\gamma_{\tilde a,\varepsilon}(\theta) \le \mathbb{E}^{\tilde a/a}\left[F^{2a}(L^\varepsilon_T)\right]\mathbb{E}^{(a-\tilde a)/a}\left[e^{-(2\tilde a-1)a/(a-\tilde a)\theta.L^\varepsilon_T}\right].$$

Please cite this article in press as: M.B. Alaya, et al., Importance sampling and statistical Romberg method for Lévy processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2015.12.008

Noticing that $q = (2\tilde{a}-1)a/(a-\tilde{a})$, it results from assumption $\sup_{0<\varepsilon\le 1}\mathbb{E}\left[F^{2a}(L_T^\varepsilon)\right] < +\infty$ that $\gamma_{\tilde{a},\varepsilon}$ is uniformly bounded on the compact set $K \subset \Theta_q$. Moreover, using once again relation (20) we deduce the uniform boundedness of the family $(\kappa_\varepsilon)_{\varepsilon>0}$ on the compact set K. Hence, combining all these results together with assumption (WE$_{\upsilon_\varepsilon}$), we deduce the existence of $c_3 > 0$ not depending on ε such that $\sum_{k=1}^{N}\mathbb{E}\left[\left|M_k^\varepsilon - M_{k-1}^\varepsilon\right|^{2\tilde{a}}|\mathcal{F}_{T,k-1}\right] \le \frac{c_3}{N^{\tilde{a}-1}}$. This completes the proof. □

Remark. If one have in mind to reduce the variance by using an adaptive crude Monte Carlo method, it appears clear that the natural choice is

$$\theta_1^* = \arg\min_{\theta\in\Theta_{1,3}} v_1(\theta) \quad \text{and} \quad \theta_{1,\varepsilon}^* = \arg\min_{\theta\in\Theta_{1,3}^\varepsilon} v_{1,\varepsilon}(\theta) \quad \text{for } \varepsilon > 0,$$

where v_1 and $v_{1,\varepsilon}$ are presented in Section 4. The construction of stochastic sequences converging almost surely to these desired targets and satisfying $(\mathcal{H}_\theta)$ is ensured by Corollary 5.1.

Now based on the telescoping identity (14), we introduce the adaptive importance sampling statistical Romberg method (ISSR) to approximate our initial quantity of interest $\mathbb{E}F(L_T)$ by

$$\begin{aligned} Q_\varepsilon^{\text{ISSR}} &:= \frac{1}{N_1}\sum_{k=1}^{N_1} F\left(L_{T,k}^{\varepsilon^\beta,\theta_{1,k-1}^{\varepsilon^\beta}}\right) e^{-\theta_{1,k-1}^{\varepsilon^\beta}.L_{T,k}^{\varepsilon^\beta,\theta_{1,k-1}^{\varepsilon^\beta}}+T\kappa_{\varepsilon^\beta}(\theta_{1,k-1}^{\varepsilon^\beta})} \\ &\quad + \frac{1}{N_2}\sum_{k=1}^{N_2}\left(F\left(L_{T,k}^{\varepsilon,\theta_{2,k-1}^\varepsilon}\right) - F\left(L_{T,k}^{\varepsilon^\beta,\theta_{2,k-1}^\varepsilon}\right)\right) e^{-\theta_{2,k-1}^\varepsilon.L_{T,k}^{\varepsilon,\theta_{2,k-1}^\varepsilon}+T\kappa_\varepsilon(\theta_{2,k-1}^\varepsilon)}. \end{aligned} \tag{28}$$

Our second result is a central limit theorem for the adaptive ISSR method.

Theorem 5.4. *Let $F : \mathbb{R}^d \to \mathbb{R}$ be a $\mathscr{C}^1$ function satisfying assumption (WE$_{\upsilon_\varepsilon}$) and such that $\sup_{0<\varepsilon\le 1}\mathbb{E}F^{2a}(L_T^\varepsilon)$ and $\sup_{0<\varepsilon\le 1}\mathbb{E}\left|\sigma^{-1}(\varepsilon)(F(L_T^\varepsilon) - F(L_T))\right|^{2a}$ are finite, for $a > 1$. Suppose also that the following assumptions are satisfied.*

H1. *Condition (3) in Theorem 2.1 holds and there exists a definite positive matrix Σ such that $\lim_{\varepsilon\to 0}\sigma^{-2}(\varepsilon)\Sigma_\varepsilon = \Sigma$.*
H2. *For $0 < \beta < 1$, we have $\lim_{\varepsilon\to 0}\sigma(\varepsilon)\sigma^{-1}(\varepsilon^\beta) = 0$ and $\lim_{\varepsilon\to 0}\upsilon_\varepsilon\sigma^{-1}(\varepsilon^\beta) = 0$.*

Moreover, assume there exists a double indexed family $(\theta_{i,k}^\varepsilon)_{k\in\mathbb{N},\varepsilon>0}$, for $i \in \{1,2\}$, satisfying $(\mathcal{H}_\theta)$ and belonging to some nonempty compact set $K_i \subset \mathring{\Theta}_q$ with $q > a/(a-1)$. If we choose $N_1 = \lfloor \upsilon_\varepsilon^{-2}\rfloor$ and $N_2 = \lfloor \upsilon_\varepsilon^{-2}\sigma^2(\varepsilon^\beta)\rfloor$, then

$$\upsilon_\varepsilon^{-1}\left(Q_\varepsilon^{\text{ISSR}} - \mathbb{E}F(L_T)\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(C_F, \sigma^2 + \tilde{\sigma}^2\right), \quad \textit{as } \varepsilon \to 0,$$

where

$$\begin{aligned} \sigma^2 &= \mathbb{E}\left[F^2(L_T)e^{-\theta^*.L_T+T\kappa(\theta^*)}\right] - [\mathbb{E}F(L_T)]^2 \quad \textit{and} \\ \tilde{\sigma}^2 &= T\mathbb{E}\left[(\nabla F(L_T).\Sigma\nabla F(L_T))e^{-\theta^*.L_T+T\kappa(\theta^*)}\right]. \end{aligned}$$

Note that in the above theorem, assumptions involving F, conditions H1. and H2. are the same as in Theorem 3.2. Moreover, Corollary 3.2 exhibits examples of functions for which these assumptions are satisfied. Finally, according to the comment below Theorem 5.3, the

Please cite this article in press as: M.B. Alaya, et al., Importance sampling and statistical Romberg method for Lévy processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2015.12.008

d-dimensional CGMY process and the examples given there still satisfy the assumed hypotheses in the above theorem.

Proof. By assumption (WE$_{\upsilon_\varepsilon}$) we only need to study the asymptotic behavior of $\upsilon_\varepsilon^{-1} Q_{1,\varepsilon}^{\text{ISSR}} + \upsilon_\varepsilon^{-1} Q_{2,\varepsilon}^{\text{ISSR}}$ with

$$Q_{1,\varepsilon}^{\text{ISSR}} = \frac{1}{N_1}\sum_{k=1}^{N_1}\Big(F\Big(L_{T,k}^{\varepsilon^\beta,\theta_{1,k-1}^{\varepsilon^\beta}}\Big)e^{-\theta_{1,k-1}^{\varepsilon^\beta}.L_{T,k}^{\varepsilon^\beta,\theta_{1,k-1}^{\varepsilon^\beta}}+T\kappa_{\varepsilon^\beta}(\theta_{1,k-1}^{\varepsilon^\beta})} - \mathbb{E}F(L_T^{\varepsilon^\beta})\Big)$$

and

$$\begin{aligned}Q_{2,\varepsilon}^{\text{ISSR}} = \frac{1}{N_2}\sum_{k=1}^{N_2}\Big(&\big[F(L_{T,k}^{\varepsilon,\theta_{2,k-1}^{\varepsilon}}) - F(L_{T,k}^{\varepsilon^\beta,\theta_{2,k-1}^{\varepsilon}})\big]e^{-\theta_{2,k-1}^{\varepsilon}.L_{T,k}^{\varepsilon,\theta_{2,k-1}^{\varepsilon}}+T\kappa_\varepsilon(\theta_{2,k-1}^{\varepsilon})}\\ &-\mathbb{E}\big[F(L_T^\varepsilon) - F(L_T^{\varepsilon^\beta})\big]\Big).\end{aligned}$$

An application of Theorem 5.3 yields $\upsilon_\varepsilon^{-1} Q_{1,\varepsilon}^{\text{ISSR}} \xrightarrow{\mathcal{L}} \mathcal{N}\big(0,\sigma^2\big)$, as $\varepsilon \to 0$. For the second term, we aim to apply Theorem A.1. So, we introduce the martingale arrays $(M_k^\varepsilon)_{k\geq 1}$

$$\begin{aligned}M_k^\varepsilon := \frac{\upsilon_\varepsilon^{-1}}{N_2}\sum_{\ell=1}^{k}\Big(&\big(F(L_{T,\ell}^{\varepsilon,\theta_{2,\ell-1}^{\varepsilon}}) - F(L_{T,\ell}^{\varepsilon^\beta,\theta_{2,\ell-1}^{\varepsilon}})\big)e^{-\theta_{2,\ell-1}^{\varepsilon}.L_{T,\ell}^{\varepsilon,\theta_{2,\ell-1}^{\varepsilon}}+T\kappa_\varepsilon(\theta_{2,\ell-1}^{\varepsilon})}\\ &-\mathbb{E}\big[F(L_T^\varepsilon) - F(L_T^{\varepsilon^\beta})\big]\Big).\end{aligned}$$

Step 1. Thanks to assumption $(\mathcal{H}_\theta)$ and the Esscher transform, the quadratic variation of M evaluated at N_2 is equal to

$$\langle M^\varepsilon\rangle_{N_2} = \frac{\upsilon_\varepsilon^{-2}}{N_2^2}\sum_{k=1}^{N_2}\xi_\varepsilon(\theta_{2,k-1}^{\varepsilon})e^{T\kappa_\varepsilon(\theta_{2,k-1}^{\varepsilon})} - \frac{\upsilon_\varepsilon^{-2}}{N_2}\Big(\mathbb{E}\big[\sigma^{-1}(\varepsilon^\beta)(F(L_T^\varepsilon) - F(L_T^{\varepsilon^\beta}))\big]\Big)^2,$$

where for all $\theta \in \Theta_q$, $\xi_\varepsilon(\theta) = \sigma^{-2}(\varepsilon^\beta)\mathbb{E}\left(\left|F(L_T^\varepsilon) - F(L_T^{\varepsilon^\beta})\right|^2 e^{-\theta.L_T^\varepsilon}\right)$. Using the convergence in law given by relation (12), the assumption $\sup_{0<\varepsilon\leq 1}\mathbb{E}\left|\sigma^{-1}(\varepsilon)(F(L_T^\varepsilon) - F(L_T))\right|^{2a} < +\infty$ and the independence of L_T and W_T, we deduce that the second term on the right hand side of the above equation vanishes when ε tends to zero. Concerning the first one, we aim to use Lemma A.1. So, we only need to prove the equicontinuity of the family $(\xi_\varepsilon)_{\varepsilon>0}$ on any compact subset of Θ_q. First, we prove the simple convergence of ξ_ε to ξ with $\xi(\theta) = \mathbb{E}\Big(\big|\nabla F(L_T).\Sigma^{\frac{1}{2}}W_T\big|^2 e^{-\theta.L_T}\Big)$. For this, we can proceed analogously to the proof of relation (12). More precisely, we use Taylor–Young's expansion with function F, the convergence in law given by (12), the independence of $L_T^\varepsilon - L_T^{\varepsilon^\beta}$ and $L_T^{\varepsilon^\beta}$ and Slutsky's theorem to get

$$\sigma^{-2}(\varepsilon^\beta)\big|F(L_T^\varepsilon) - F(L_T^{\varepsilon^\beta})\big|^2 e^{-\theta.L_T^\varepsilon}\xrightarrow[\varepsilon\to 0]{\mathcal{L}}\big|\nabla F(L_T).\Sigma^{\frac{1}{2}}W_T\big|^2 e^{-\theta.L_T}.$$

Now, applying Hölder's inequality with $\tilde{a} = \frac{aq}{a+q}$ yields

$$\begin{aligned}&\mathbb{E}\big|\sigma^{-2}(\varepsilon^\beta)\big(F(L_T^\varepsilon) - F(L_T^{\varepsilon^\beta})\big)^2 e^{-\theta.L_T^\varepsilon}\big|^{\tilde{a}}\\ &\quad\leq \mathbb{E}^{\tilde{a}/a}\big|\sigma^{-1}(\varepsilon^\beta)\big(F(L_T^\varepsilon) - F(L_T^{\varepsilon^\beta})\big)\big|^{2a}\mathbb{E}^{(a-\tilde{a})/a}e^{-\frac{\tilde{a}a}{a-\tilde{a}}\theta.L_T^\varepsilon}.\end{aligned}$$

Please cite this article in press as: M.B. Alaya, et al., Importance sampling and statistical Romberg method for Lévy processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2015.12.008

Using assumptions H2 and $\sup_{0<\varepsilon\leq 1}\mathbb{E}\left|\sigma^{-1}(\varepsilon)(F(L_T^\varepsilon)-F(L_T))\right|^{2a}<+\infty$, it is easy to check the uniform boundedness with respect to ε of the first term on the right hand side of the above inequality. Concerning the second one, since $q=\frac{\tilde{a}a}{a-\tilde{a}}$ we use relation (20) to deduce the same result. Hence, we have the simple convergence of ξ_ε towards ξ when ε tends to zero. Therefore, it remains to prove the equicontinuity of the family functions $(\xi_\varepsilon)_{\varepsilon>0}$ on any compact set $K\subset\Theta_q$. Replacing $F(L_T^\varepsilon)$ by $\sigma^{-1}(\varepsilon^\beta)\big(F(L_T^\varepsilon)-F(L_T^{\varepsilon^\beta})\big)$ in the steps of the proof of relation (27) and using assumptions $H2$ and $\sup_{0<\varepsilon\leq 1}\mathbb{E}\left|\sigma^{-1}(\varepsilon)(F(L_T^\varepsilon)-F(L_T))\right|^{2a}<+\infty$ we prove the existence of a constant $c>0$ not depending on ε such that

$$\left|\xi_\varepsilon(\theta)-\xi_\varepsilon(\theta')\right|\leq c|\theta-\theta'|. \tag{29}$$

Thus, under assumption $(\mathcal{H}_\theta)$, we get the almost sure convergence of $\xi_\varepsilon(\theta^\varepsilon_{2,k})$ towards $\xi(\theta^*)$ as k goes to infinity and ε vanishes. We complete the proof of the first step using the almost sure convergence of $\kappa_\varepsilon(\theta^\varepsilon_{2,k})$ towards $\kappa(\theta^*)$ as k goes to infinity and ε vanishes. This last convergence is obtained thanks to relation (20).

Step 2. The second step of this proof consists on checking the Lyapunov condition A3 of Theorem A.1. We proceed in the same way as in the second step of the proof of Theorem 5.3. We take $\tilde{a}=\frac{aq+a}{2a+q}$ and we get using the same arguments that $\sum_{k=1}^{N_2}\mathbb{E}\big[\left|M_k^\varepsilon-M_{k-1}^\varepsilon\right|^{2\tilde{a}}|\mathcal{F}_{T,k-1}\big]$ is bounded by

$$\frac{2^{2\tilde{a}-1}\upsilon_\varepsilon^{-2\tilde{a}}}{N_2^{2\tilde{a}}}\sum_{k=1}^{N_2}\xi_{a,\varepsilon}(\theta^\varepsilon_{2,k-1})e^{(2\tilde{a}-1)T\kappa_\varepsilon(\theta^\varepsilon_{2,k-1})}$$
$$+\frac{2^{2\tilde{a}-1}\upsilon_\varepsilon^{-2\tilde{a}}}{N_2^{2\tilde{a}-1}}\left|\mathbb{E}\big[\sigma^{-1}(\varepsilon^\beta)(F(L_T^\varepsilon)-F(L_T^{\varepsilon^\beta}))\big]\right|^{2\tilde{a}}$$

where for all $\theta\in\Theta_q$, $\xi_{\tilde{a},\varepsilon}(\theta)=\mathbb{E}\Big[\left|\sigma^{-1}(\varepsilon^\beta)(F(L_T^\varepsilon)-F(L_T^{\varepsilon^\beta}))\right|^{2\tilde{a}}e^{-(2\tilde{a}-1)\theta.L_T^\varepsilon}\Big]$. Then replacing $F(L_T^\varepsilon)$ by $\sigma^{-1}(\varepsilon^\beta)\big(F(L_T^\varepsilon)-F(L_T^{\varepsilon^\beta})\big)$ in the second step of the proof of Theorem 5.3, the same arguments remain valid thanks to assumptions H2 and $\sup_{0<\varepsilon\leq 1}\mathbb{E}[|\sigma^{-1}(\varepsilon)(F(L_T^\varepsilon)-F(L_T))|^{2a}<+\infty$. So, we deduce the existence of $c>0$ not depending on ε such that

$$\sum_{k=1}^{N_2}\mathbb{E}\left[\left|M_k^\varepsilon-M_{k-1}^\varepsilon\right|^{2\tilde{a}}|\mathcal{F}_{T,k-1}\right]\leq\frac{c}{N_2^{\tilde{a}-1}}.$$

This completes the proof. $\square$

Remark. 1. Similarly as in the MC case, we still have in mind to reduce the variance associated now to the SR method. This goes back to optimize separately v_1 and v_2. Hence, the optimal choice corresponds to

$$\theta_i^*=\arg\min_{\theta\in\Theta_{1,3}}v_i(\theta)\quad\text{and}\quad\theta^*_{i,\varepsilon}=\arg\min_{\theta\in\Theta^\varepsilon_{i,3}}v_{i,\varepsilon}(\theta)\quad\text{for }\varepsilon>0\text{ and }i\in\{1,2\},$$

where v_i and $v_{i,\varepsilon}$ are presented in Section 4. In the same way, the construction of stochastic sequences converging almost surely to these desired targets and satisfying $(\mathcal{H}_\theta)$ is ensured by Corollary 5.1.

2. For the setting of Section 3.3, we have that the complexity of ISMC (resp. ISSR) algorithm is equal to the complexity of MC (resp. SR) method up to a multiplicative factor 2. Indeed, for a fixed k in relations (24) (resp. (28)) we need to simulate first θ^ε_{k-1} with a cost equal

to $\mathcal{K}(\varepsilon) = \nu(|x| \geq \varepsilon)$, then we simulate independently $L^{\varepsilon,\theta}_{T,k}$ for a fixed $\theta = \theta^{\varepsilon}_{k-1}$ with an additional cost equal to $\mathcal{K}_\theta(\varepsilon) = \nu_\theta(|x| \geq \varepsilon)$ which is of the same order as $\mathcal{K}(\varepsilon)$ when $\varepsilon \to 0$. Hence, the optimal gain is also given by

$$\frac{C_{ISSR}}{C_{ISMC}} = O\left(\varepsilon^{Y(Y/2-1)}\right).$$

6. Numerical results

Now, we present numerical simulations that illustrate the efficiency of the ISSR method throughout the pricing of vanilla options with an underlying asset following an exponential pure jump CGMY model. The CGMY process has been introduced by Carr, Geman, Madan and Yor [8] with the aim to develop a model for the dynamic of equity log-returns which is rich enough to accommodate jumps of finite or infinite activity, and finite or infinite variation. Monte Carlo simulation of the CGMY process has been tackled in the literature specifically by Madan and Yor [26], Poirot and Tankov [27] and Rosinski [30]. A CGMY process is a pure jump process with generating triplet $(0, 0, \nu)$ where for $C > 0$, $G > 0$, $M > 0$ and $Y \in (0, 2)$

$$\nu(dx) = \frac{Ce^{-Mx}}{x^{1+Y}}\mathbf{1}_{x>0}dx + \frac{Ce^{-G|x|}}{|x|^{1+Y}}\mathbf{1}_{x<0}dx. \tag{30}$$

Following the notations of [27], we consider the Lévy–Khintchine representation with a truncation function h and a characteristic exponent given by

$$\psi(u) = i\gamma_h u + \int_{\mathbb{R}}(e^{iux} - 1 - iuh(x))\nu(dx) \quad \text{with}$$
$$\gamma_h = \int_{\mathbb{R}}(h(x) - x\mathbf{1}_{\{|x|\leq 1\}})\nu(dx), \qquad u \in \mathbb{R}.$$

- For $1 < Y < 2$ and $h(x) = x$, we have $\gamma_h = \int_{|x|\geq 1} x\nu(dx)$ and
$$\psi(u) = iu\gamma_h + C\Gamma(-Y)$$
$$\times\left[M^Y\left(\left(1 - \frac{iu}{M}\right)^Y - 1 + \frac{iuY}{M}\right) + G^Y\left(\left(1 + \frac{iu}{G}\right)^Y - 1 - \frac{iuY}{G}\right)\right].$$
- For $0 < Y < 1$ and $h(x) = 0$, we have $\gamma_h = \int_{|x|\leq 1} x\nu(dx)$ and
$$\psi(u) = iu\gamma_h + C\Gamma(-Y)\left[M^Y\left(\left(1 - \frac{iu}{M}\right)^Y - 1\right) + G^Y\left(\left(1 + \frac{iu}{G}\right)^Y - 1\right)\right].$$

In what follows, we consider the risk neutral model with jumps generalizing the Black Scholes model by replacing the Brownian motion by $(L_t)_{0\leq t\leq T}$ the CGMY process with generating triplet $(\gamma, 0, \nu)$, $\gamma \in \mathbb{R}$ and define the asset price

$$S_t = S_0 \exp(rt + L_t), \qquad \text{where } r > 0 \text{ is the interest rate and } S_0 > 0.$$

To guarantee that $e^{-rt}S_t$ is a martingale we have to impose the condition $\int_{|x|\geq 1} e^x\nu(dx) < \infty$ (which is satisfied as soon as $M > 1$) and the condition

$$\gamma + \int_{\mathbb{R}}(e^y - 1 - y\mathbf{1}_{\{|y|\leq 1\}})\nu(dy) = 0, \tag{31}$$

or in other words $\gamma = -\psi(-i)$.

Please cite this article in press as: M.B. Alaya, et al., Importance sampling and statistical Romberg method for Lévy processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2015.12.008

Now, let us recall that for $0 < \varepsilon < 1$, the approximation $(L_t^\varepsilon)_{t\geq 0}$ of $(L_t)_{t\geq 0}$ is a Lévy process with generating triplet $(\gamma, 0, \nu_\varepsilon)$ where $\nu_\varepsilon(dx) := \mathbf{1}_{\{|x|\geq\varepsilon\}}\nu(dx)$. It is worth to note that $(L_t^\varepsilon)_{t\geq 0}$ can be seen as a compound Poisson process with drift $\gamma_\varepsilon := \gamma - \int_{\varepsilon\leq|x|\leq 1} x\nu(dx)$, see (1). This compound Poisson process can be represented as the difference of two independent processes namely the positive part and the negative one. More precisely, the positive part (resp. the negative part) is a compound Poisson process with jump size $\nu_\varepsilon^+ = \mathbf{1}_{\{x\geq\varepsilon\}}\frac{\nu(dx)}{\nu([\varepsilon,+\infty[)}$ (resp. $\nu_\varepsilon^- = \mathbf{1}_{\{x\leq-\varepsilon\}}\frac{\nu(dx)}{\nu(]-\infty,-\varepsilon])}$) and intensity $\nu([\varepsilon, +\infty[)$ (resp. $\nu(]-\infty, -\varepsilon])$). To simulate these compound Poisson processes, we can use either the classical rejection method as described in Cont and Tankov [10] or an improved method used by Madan and Yor [26]. Indeed, when we simulate the positive part we choose $\nu_{0,\varepsilon}^+$ so that $\frac{d\nu_\varepsilon^+}{d\nu_{0,\varepsilon}^+}(x) = e^{-Mx}\mathbf{1}_{\{x>\varepsilon\}} \leq 1$. Then, according to Rosinski [29] we may simulate the paths of ν_ε^+ from those of $\nu_{0,\varepsilon}^+$ by only accepting all jumps x in the paths of $\nu_{0,\varepsilon}^+$ for which $\frac{d\nu_\varepsilon^+}{d\nu_{0,\varepsilon}^+}(x) > u$ where u is an independent draw from uniform distribution. Hence, we use following algorithm

Algorithm 1 Simulating the positive jump size Z of the CGMY process using Rosinski's rejection

Require: U_1 and U_2 are uniform random variables and $Z = \varepsilon U_1^{-\frac{1}{Y}}$
  if $U_2 > \exp -M.Z$ **then**
    $Z = 0$
  end if
  return Z

In the same way, we simulate the negative jump part by replacing in the above algorithm the parameter M by G.

Our aim is to test our approximation methods for computing the price of a vanilla option with payoff F. To do so, we use the importance sampling technique, introduced in Section 4, to approximate the price $e^{-rT}\mathbb{E}F(S_T)$ by

$$e^{-rT}\mathbb{E}\left[F(S_T^{\varepsilon,\theta})e^{-\theta.L_t^{\varepsilon,\theta}+T\kappa_\varepsilon(\theta)}\right],$$

$$\text{with } S_T^{\varepsilon,\theta} = S_0\exp(rt + L_t^{\varepsilon,\theta}),\ \text{ for } \theta \in \Theta_1 = [-G, M], \tag{32}$$

where $L_T^{\varepsilon,\theta}$ is also a Lévy process with generating triplet $(\gamma_{\varepsilon,\theta}, 0, \nu_{\varepsilon,\theta})$, where $\nu_{\varepsilon,\theta} = e^{\theta.x}\nu_\varepsilon(dx)$ and $\gamma_{\theta,\varepsilon} = \gamma_\varepsilon + \int_{-1}^1 x(e^{\theta.x} - 1)\nu_\varepsilon(dx)$. The choice of θ depends on using the classical MC method or the SR one. According to relation (17), $\theta_{1,\varepsilon}^*$ is the optimal choice for the MC method. However, for the SR method, we optimize separately each quantity appearing in the associated variance and the optimal choice is given by the couple $(\theta_{1,\varepsilon}^*, \theta_{2,\varepsilon}^*)$ (see relation (17)). To compute these optimal terms, we use the constrained algorithms introduced in the system (21). It is worth to note that in practice it is easier to use $\kappa(\theta)$ instead of $\kappa_\varepsilon(\theta)$.

Concerning the numerical illustrations, we decide to test our algorithms for a range of parameters M and G commonly used by practitioners without taking into account some theoretical restrictions. In fact, we simply choose the default parameters given by the Premia platform (https://www.rocq.inria.fr/mathfi/Premia/index.html).

6.1. One-dimensional CGMY process

In this setting we consider the European call option with payoff $F(x) = (x - \text{Strike})_+$. The parameters of the CGMY model are chosen as follows: $S0 = 100$, Strike $= 100$, $C = 0.0244$, $G = 0.0765$, $M = 7.5515$, $Y = 1.2945$, the free interest rate $r = \log(1.1)$ and maturity time $T = 1$. In this case, we have $\sigma^2(\varepsilon) \underset{\varepsilon\to 0}{\simeq} 2C\varepsilon^{2-Y}/(2-Y)$ and as it is mentioned in the proofs of Corollaries 3.1 and 3.2 we have $\upsilon_\varepsilon = \sigma^{1-\eta/2}(\varepsilon)$, for $\eta \in (0,2)$ and we finally choose $\eta = 0.2$. We run 50000 iteration for the constrained algorithm with the compact set $[-G, M]$. The obtained optimal values are given by $(\theta^*_{1,\varepsilon}, \theta^*_{2,\varepsilon}) = (5.3, 2.5)$ (see Fig. 1).

In order to compare the ISMC algorithm (24) (with $N = \lfloor \varepsilon^{0.9(Y-2)} \rfloor$) and the ISSR one (28) (with $N_1 = \lfloor \varepsilon^{0.9(Y-2)} \rfloor$ and $N_2 = \lfloor \varepsilon^{(Y-2)(0.9-Y/2)} \rfloor$), we use the couple $(\theta^*_{1,\varepsilon}, \theta^*_{2,\varepsilon})$ computed above. Note that we take the optimal sample sizes N, N_1 and N_2 given in the remark after Theorem 5.4. For this, we compute for each method the CPU time (per second) (the computations are done on a PC with a 2.5 GHz Intel core i5 processor) and an error measure given by the mean squared error (MSE) which is defined by

$$\text{MSE} = \frac{1}{30} \sum_{i=1}^{30} (\text{Real value} - \text{Simulated value})^2. \tag{33}$$

The real value is obtained using the Fourier-cosine method introduced by Fang and Oosterlee [16] for a one-dimensional CGMY with an accuracy of order 10^{-10}. This method is available in the free online version of Premia platform. For this setting, our ISSR algorithm (28) is now available in the latest premium version of Premia.

For different values of ε, we give in Fig. 2 the log–log plot of the obtained MSE versus the CPU time for the classical Monte Carlo (MC), the statistical Romberg (SR), the importance sampling Monte Carlo (ISMC) and the importance sampling statistical Romberg (ISSR) methods.

According to Table 1 and for a fixed MSE of order $6 \cdot 10^{-3}$, the ISSR method reduces the CPU time by a factor of 8, 73 compared to the ISMC one. Clearly the ISSR method is the most efficient compared to the other ones.

6.2. Two-dimensional CGMY process

We focus now on the computation of a price of the form $e^{-rT}\mathbb{E}F(S^1_T, S^2_T)$, where $F(x, y) = (x + y - \text{Strike})_+$ and the couple $(S^1_t, S^2_t)_{0\le t\le T}$ denotes the underlying asset process. In this setting we choose $(S^1_t, S^2_t) = (S_0 e^{rt+L^1_t}, S_0 e^{rt+L^2_t})$ where $(L^1_t)_{0\le t\le T}$ and $(L^2_t)_{0\le t\le T}$ are two independent CGMY processes with generating triplets $(\gamma_1, 0, \nu_1)$ and $(\gamma_2, 0, \nu_2)$ such that the processes $(e^{-rt}S^1_t)_{0\le t\le T}$ and $(e^{-rt}S^2_t)_{0\le t\le T}$ are two martingales. So, it amounts to select γ_1 and γ_2 as in relation (31).

Since the Fourier-cosine method with high accuracy cannot be used for the two-dimensional setting, the “Benchmark” price is obtained by running the classical MC algorithm with a very small value of ε. Indeed, for $\varepsilon = 10^{-6}$ the “Benchmark” price is 21.0782 with a CPU time of 24718 s. The parameters of the considered two CGMY processes defined by (C, G_1, M_1, Y) and (C, G_2, M_2, Y) are chosen as follows: $C = 0.0244$, $G_1 = 0.0765$, $M_1 = 7.55015$, $G_2 = 2$, $M_2 = 5$, $Y = 0.9$, $S0 = 100$, Strike $= 200$, $r = \log(1.1)$ and the maturity time $T = 1$. In this case, we have $\Sigma_\varepsilon = \sigma^2(\varepsilon) I_{2\times 2}$ with $\sigma^2(\varepsilon) \underset{\varepsilon\to 0}{\simeq} 2C\varepsilon^{2-Y}/(2-Y)$ and $\upsilon_\varepsilon = \sigma^{0.9}(\varepsilon)$. Using

Please cite this article in press as: M.B. Alaya, et al., Importance sampling and statistical Romberg method for Lévy processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2015.12.008

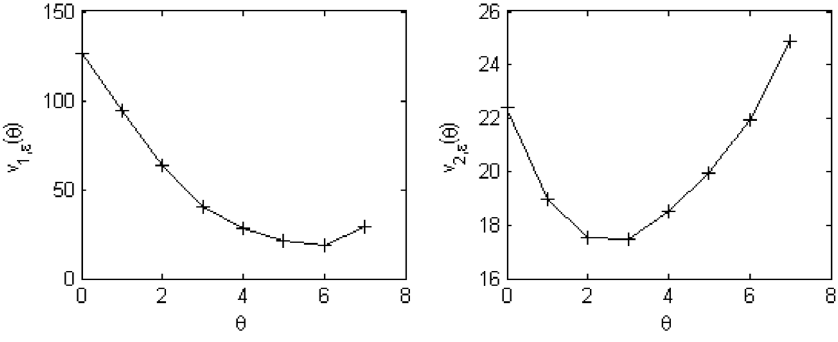

Fig. 1. Variances $v_{1,\varepsilon}$ and $v_{2,\varepsilon}$ versus θ in the one-dimensional setting.

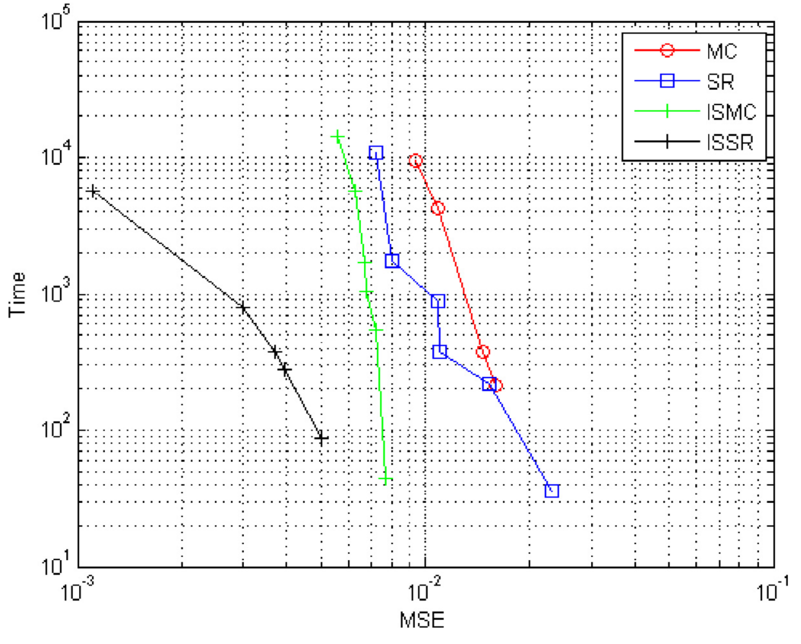

Fig. 2. CPU time versus MSE in the one-dimensional setting.

Table 1
Time complexity reduction (ISSR versus ISMC).

Time complexity reduction		
MSE	ISMC CPU time	ISSR CPU time
$7 \cdot 10^{-3}$	$7 \cdot 10^{2}$	$5 \cdot 10^{2}$
$6,5 \cdot 10^{-3}$	$2 \cdot 10^{3}$	$6 \cdot 10^{2}$
$6 \cdot 10^{-3}$	$5,5 \cdot 10^{3}$	$6,3 \cdot 10^{2}$
$5,5 \cdot 10^{-3}$	$15 \cdot 10^{3}$	$7 \cdot 10^{2}$

the constrained algorithms (21), we obtain the values of the optimal two-dimensional vectors given by relation (17) and we get $\theta^*_{1,\varepsilon} = (4, 3.5)$ and $\theta^*_{2,\varepsilon} = (3.5, 1.1)$. In Fig. 3, we plot the

Please cite this article in press as: M.B. Alaya, et al., Importance sampling and statistical Romberg method for Lévy processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2015.12.008

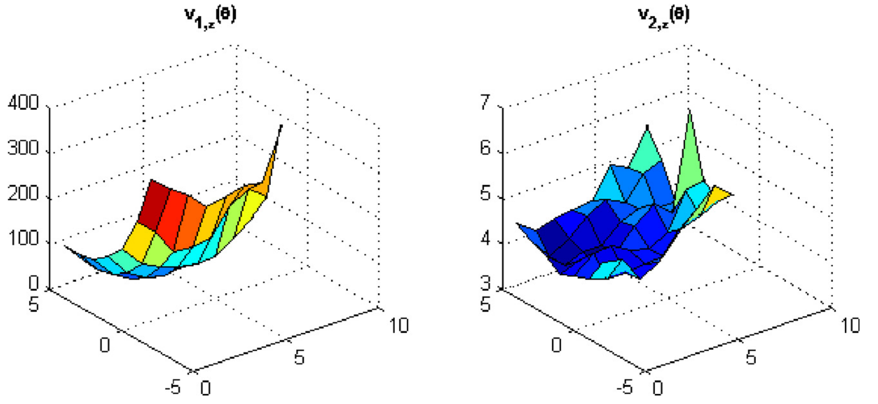

Fig. 3. Variances $v_{1,\varepsilon}$ and $v_{2,\varepsilon}$ versus θ in the two-dimensional setting.

Table 2
Time complexity reduction ISSR versus ISMC.

Time complexity reduction		
MSE	ISMC CPU time	ISSR CPU time
10^{-3}	40	20
$6 \cdot 10^{-4}$	100	30
$4 \cdot 10^{-4}$	250	60
$3 \cdot 10^{-4}$	450	80

evolution of both variances $v_{1,\varepsilon}$ and $v_{2,\varepsilon}$ in terms of $\theta = (\theta_1, \theta_2) \in [-G_1, M_1] \times [-G_2, M_2]$. Now we proceed as in the one-dimensional case with the same sample sizes to compare the different methods. Fig. 4 confirms the superiority of the ISSR method over the other ones and this holds even when we compare it to the ISMC method. Indeed, for a given MSE, the ISSR spends less time than the other methods to compute the desired option price. The difference in terms of computational time becomes more significant as soon as the MSE becomes very small, which corresponds to low values of ε (see Fig. 4).

According to Table 2 and for a fixed MSE of order 10^{-3}, the ISSR reduces the CPU time of the considered option price by a factor 2 in comparison to the ISMC method. Moreover, this factor becomes more important when we consider a smaller MSE. In fact, for a fixed MSE of order $3 \cdot 10^{-4}$, the ISSR reduces the CPU time by a factor > 5 in comparison to the ISMC one.

7. Conclusion

In this paper, we highlight the superiority of the ISSR method over the classical Monte Carlo approach for the setting of Lévy processes. It may be of interest to extend this study to the setting of Euler discretization schemes for Lévy driven diffusions developed by Protter and Talay [28], by Jacod, Kurtz, Méléard and Protter [20] and more recently by Kohatsu-Higa and Tankov [23]. Also, a next natural question consists on developing analogous results for path dependent options in exponential Lévy models in the spirit of the works of Dia and Lamberton [13,14]. These two points will be the object of a forthcoming works.

Please cite this article in press as: M.B. Alaya, et al., Importance sampling and statistical Romberg method for Lévy processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2015.12.008

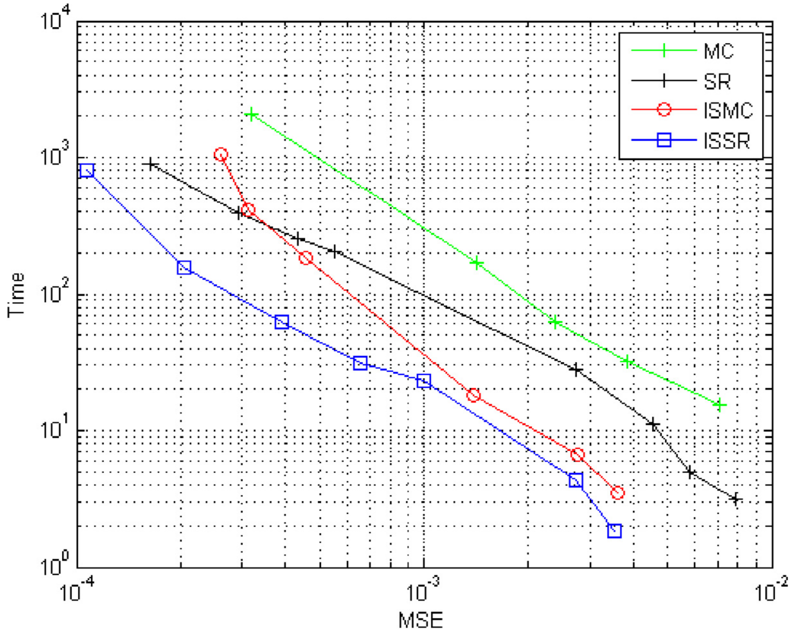

Fig. 4. CPU time versus MSE in the two-dimensional setting.

Acknowledgments

The first and third author's research was supported by Laboratory of Excellence MME-DII http://labex-mme-dii.u-cergy.fr/. This research benefited from the support of the chair " Risques Financiers", Fondation du Risque.

Appendix

We recall first the Lindeberg–Feller Central Limit Theorem for martingales arrays.

Theorem A.1 (*Central Limit Theorem for martingales arrays [15]*). *Suppose that $(\Omega, \mathbb{F}, \mathbb{P})$ is a probability space and that for each n, we have a filtration $\mathbb{F}_n = (\mathcal{F}_k^n)_{k\geq 0}$, a sequence $k_n \longrightarrow \infty$ as $n \longrightarrow \infty$ and a real square integrable vector martingale $M^n = (M_k^n)_{k\geq 0}$ which is adapted to $\mathbb{F}_n$ and has quadratic variation denoted by $(\langle M\rangle_k^n)_{k\geq 0}$. We make the following two assumptions.*

A1. *There exists a deterministic symmetric positive semi-definite matrix Γ, such that*

$$\langle M\rangle_{k_n}^n = \sum_{k=1}^{k_n} \mathbb{E}\left[|M_k^n - M_{k-1}^n|^2 | \mathcal{F}_{k-1}^n\right] \xrightarrow[n\to\infty]{\mathbb{P}} \Gamma.$$

A2. *Lindeberg's condition holds: that is, for all $\varepsilon > 0$,*

$$\sum_{k=1}^{k_n} \mathbb{E}\left[|M_k^n - M_{k-1}^n|^2 1_{\{|M_k^n - M_{k-1}^n| > \varepsilon\}} | \mathcal{F}_{k-1}^n\right] \xrightarrow[n\to\infty]{\mathbb{P}} 0.$$

Then

$$M_{k_n}^n \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Gamma) \quad \text{as } n \to \infty.$$

Please cite this article in press as: M.B. Alaya, et al., Importance sampling and statistical Romberg method for Lévy processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2015.12.008

Remark. The following assumption known as the Lyapunov condition, implies the Lindberg's condition A2,

A3. There exists a real number $a > 1$, such that

$$\sum_{k=1}^{k_n} \mathbb{E}\left[|M_k^n - M_{k-1}^n|^{2a} | \mathcal{F}_{k-1}^n\right] \xrightarrow[n\to\infty]{\mathbb{P}} 0.$$

Moreover, we give a double indexed version of the Toeplitz lemma. For a proof of this result see Lemma 4.1 in [4].

Lemma A.1. *Let $(a_i)_{1\le i\le k_n}$ be a sequence of real positive numbers, where $k_\varepsilon \uparrow \infty$ as ε tends to 0, and $(x_i^\varepsilon)_{i\ge 1, 0<\varepsilon\le 1}$ a double indexed sequence such that*

(i) $\lim_{\varepsilon\to 0} \sum_{1\le i\le k_\varepsilon} a_i = \infty$
(ii) $\lim_{\substack{i\to+\infty\\ \varepsilon\to 0}} x_i^\varepsilon = \lim_{i\to+\infty}(\lim_{\varepsilon\to 0} x_i^\varepsilon) = \lim_{\varepsilon\to 0}(\lim_{i\to+\infty} x_i^\varepsilon) = x.$

Then

$$\lim_{\varepsilon\to 0} \frac{\sum_{i=1}^{k_\varepsilon} a_i x_i^\varepsilon}{\sum_{i=1}^{k_\varepsilon} a_i} = x.$$

References

[1] A. Abdulle, A. Blumenthal, E. Buckwar, The multilevel monte carlo method for stochastic differential equations driven by jump-diffusion processes, Technical report, 2011.
[2] B. Arouna, Adaptative Monte Carlo method, a variance reduction technique, Monte Carlo Methods Appl. 10 (1) (2004) 1–24.
[3] S. Asmussen, J. Rosiński, Approximations of small jumps of lévy processes with a view towards simulation, J. Appl. Probab. 38 (2) (2001) 482–493.
[4] M. Ben Alaya, K. Hajji, A. Kebaier, Importance sampling and statistical romberg method, Bernoulli 21 (4) (2015) 1947–1983. MR3378456.
[5] M. Ben Alaya, A. Kebaier, Central limit theorem for the multilevel Monte Carlo Euler method, Ann. Appl. Probab. 25 (1) (2015) 211–234.
[6] P. Billingsley, Convergence of Probability Measures, John Wiley & Sons Inc., New York, 1968.
[7] N.H. Bingham, C.M. Goldie, J.L. Teugels, Regular Variation, in: Encyclopedia of Mathematics and its Applications, vol. 27, Cambridge University Press, Cambridge, 1987.
[8] P. Carr, H. Geman, D.B. Madan, M. Yor, The fine structure of asset returns: An empirical investigation, 2000.
[9] S. Cohen, J. Rosiński, Gaussian approximation of multivariate Lévy processes with applications to simulation of tempered stable processes, Bernoulli 13 (1) (2007) 195–210.
[10] R. Cont, P. Tankov, Financial Modelling With Jump Processes, in: Chapman & Hall/CRC Financial Mathematics Series, Chapman & Hall/CRC, Boca Raton, FL, 2004.
[11] S. Dereich, S. Li, Multilevel monte carlo for lévy driven sde: central limit theorems for adaptive euler schemes, Ann. Appl. Probab. 26 (1) (2016) 136–185. MR3449315.
[12] E.H.A. Dia, Error bounds for small jumps of Lévy processes, Adv. Appl. Probab. 45 (1) (2013) 86–105.
[13] E.H.A. Dia, D. Lamberton, Connecting discrete and continuous lookback or hindsight options in exponential lévy models, Adv. Appl. Probab. 43 (4) (2011) 1136–1165.
[14] E.H.A. Dia, D. Lamberton, Continuity correction for barrier options in jump-diffusion models, SIAM J. Financial Math. 2 (1) (2011) 866–900.
[15] M. Duflo, Random Iterative Models, in: Applications of Mathematics (New York), vol. 34, Springer-Verlag, Berlin, 1997, Translated from the 1990 French original by Stephen S. Wilson and revised by the author.

Please cite this article in press as: M.B. Alaya, et al., Importance sampling and statistical Romberg method for Lévy processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2015.12.008

[16] F. Fang, C.W. Oosterlee, A novel pricing method for European options based on Fourier-cosine series expansions, SIAM J. Sci. Comput. 31 (2) (2008/09) 826–848.
[17] W. Feller, An Introduction to Probability Theory and Its Applications. Vol. II, Second edn., John Wiley & Sons, Inc., New York-London-Sydney, 1971.
[18] M.B. Giles, Multilevel monte Carlo path simulation, Oper. Res. 56 (3) (2008) 607–617.
[19] P. Glasserman, P. Heidelberger, P. Shahabuddin, Asymptotically optimal importance sampling and stratification for pricing path-dependent options, Math. Finance 9 (2) (1999) 117–152.
[20] J. Jacod, T.G. Kurtz, S. Méléard, P. Protter, The approximate Euler method for Lévy driven stochastic differential equations, Ann. Inst. H. Poincaré Probab. Statist. 41 (3) (2005) 523–558.
[21] R. Kawai, Optimal importance sampling parameter search for Lévy processes via stochastic approximation, SIAM J. Numer. Anal. 47 (1) (2008/09) 293–307.
[22] A. Kebaier, Statistical Romberg extrapolation: a new variance reduction method and applications to option pricing, Ann. Appl. Probab. 15 (4) (2005) 2681–2705.
[23] A. Kohatsu-Higa, P. Tankov, Jump-adapted discretization schemes for Lévy-driven SDEs, Stochastic Process. Appl. 120 (11) (2010) 2258–2285.
[24] H.J. Kushner, G.G. Yin, Stochastic Approximation Algorithms and Applications, in: Applications of Mathematics (New York), vol. 35, Springer-Verlag, New York, 1997.
[25] S. Laruelle, C. Lehalle, G. Pagès, Optimal posting price of limit orders: learning by trading, Math. Financ. Econ. 7 (3) (2013) 359–403.
[26] D.B. Madan, M. Yor, Representing the CGMY and Meixner Lévy processes as time changed Brownian motions, J. Comput. Finance 12 (1) (2008) 27–47.
[27] J. Poirot, P. Tankov, Monte carlo option pricing for tempered stable (cgmy) processes, J. Asia-Pacific Financial Markets 13 (4) (2006) 327–344.
[28] P. Protter, D. Talay, The Euler scheme for Lévy driven stochastic differential equations, Ann. Probab. 25 (1) (1997) 393–423.
[29] J. Rosiński, Series representations of Lévy processes from the perspective of point processes, in: Lévy Processes, Birkhäuser Boston, Boston, MA, 2001, pp. 401–415.
[30] J. Rosiński, Tempering stable processes, Stochastic Process. Appl. 117 (6) (2007) 677–707.
[31] K. Sato, Lévy Processes and Infinitely Divisible Distributions, in: Cambridge Studies in Advanced Mathematics, vol. 68, Cambridge University Press, Cambridge, 1999, Translated from the 1990 Japanese original, Revised by the author.
[32] W. Schoutens, Levy Processes in Finance: Pricing Financial Derivatives, in: Wiley Series in Probability and Statistics, Wiley, 2003.
[33] D. Talay, L. Tubaro, Expansion of the global error for numerical schemes solving stochastic differential equations, Stoch. Anal. Appl. 8 (4) (1991) 483–509. 1990.
[34] Y. Xia, M.B. Giles, Multilevel path simulation for jump-diffusion SDEs, in: Monte Carlo and quasi-Monte Carlo methods 2010, in: Springer Proc. Math. Stat., vol. 23, Springer, Heidelberg, 2012, pp. 695–708.
[35] Y. Xia, M.B. Giles, Multilevel monte carlo for exponential lévy models, preprint, 2014.

Please cite this article in press as: M.B. Alaya, et al., Importance sampling and statistical Romberg method for Lévy processes, Stochastic Processes and their Applications (2016), http://dx.doi.org/10.1016/j.spa.2015.12.008