OPERATIONS RESEARCH
Vol. 58, No. 6, November–December 2010, pp. 1611–1623
ISSN 0030-364X | EISSN 1526-5463 | 10 | 5806 | 1611

DOI 10.1287/opre.1100.0837

Sensitivity Estimates from Characteristic Functions

Paul Glasserman
Columbia Business School, Columbia University, New York, New York 10027, pg20@columbia.edu

Zongjian Liu
IEOR Department, Columbia University, New York, New York 10027, zl2115@columbia.edu

The likelihood ratio method (LRM) is a technique for estimating derivatives of expectations through simulation. LRM estimators are constructed from the derivatives of probability densities of inputs to a simulation. We investigate the application of the likelihood ratio method for sensitivity estimation when the relevant densities for the underlying model are known only through their characteristic functions or Laplace transforms. This problem arises in financial applications, where sensitivities are used for managing risk and where a substantial class of models have transition densities known only through their transforms. We quantify various sources of errors arising when numerical transform inversion is used to sample through the characteristic function and to evaluate the density and its derivative, as required in LRM. This analysis provides guidance for setting parameters in the method to accelerate convergence.

Subject classifications: simulation.
Area of review: Simulation.
History: Received January 2008; revisions received July 2008, July 2009, November 2009; accepted December 2009.

1. Introduction

Simulation-based derivative estimates are useful in sensitivity analysis and as inputs to optimization. In financial applications, sensitivity estimates are important in measuring and managing risk. Financial intermediaries hedge the risks they incur in buying and selling options by trading in the underlying assets; the hedging strategy is determined by the sensitivity of the option price to the underlying assets and other parameters.

The simulation literature provides several approaches to parameter sensitivity estimation; see Asmussen and Glynn (2007, Chapter 7) for an overview and references. The likelihood ratio method (LRM) constructs derivative estimators from the derivatives of probability densities associated with the simulation model. In a discrete-event simulation, the relevant densities would usually be those of interarrival or processing times; in financial applications, the relevant densities are typically those associated with the transition laws of the underlying assets. LRM is particularly attractive when there are discontinuities in the payoff or performance associated with simulated paths or discontinuities in the dynamics of the underlying model.

This paper develops and analyzes the use of LRM when the relevant densities are known only through their characteristic functions or Laplace transforms. Our primary motivation for investigating this problem lies in financial applications, where many important classes of models are characterized through Laplace or Fourier transforms. This is often the case for models driven by Lévy processes (see, e.g., Cont and Tankov 2004, Schoutens 2003) and also the affine class of jump-diffusion models studied in Duffie et al. (2000), which includes many of the most widely used models of asset prices. In other application domains, including telecommunications (e.g., Mikosch et al. 2002), stable laws arise in models with self-similarity; with few exceptions, the densities of stable laws are known only through their characteristic functions. Lévy-driven models also arise in queueing (Kella and Whitt 1992), insurance risk (Klüppelberg et al. 2004), and other areas of applied probability.

We analyze LRM estimators calculated through numerical transform inversion using, in particular, the method of Abate and Whitt (1992). The random weight used in the LRM estimator—the score function—involves both a probability density and its derivative; we numerically invert the transform of each. We analyze the various sources of error that arise in estimating expectations and their sensitivities using this approach. In particular, we focus on the simulation impact of three sources of numerical error: the grid spacing used in building an approximation to a distribution, the discretization error in the transform inversion integral, and the truncation error in the inversion integral. Building on the error analysis of Abate and Whitt (1992), we arrive at the following conclusions for the bias resulting from the combination of transform inversion and simulation in the problems we consider:

- the bias from approximating the distribution is quadratic in the grid spacing;
- the bias from the discretization error is exponentially small in the integration step size; and
- the magnitude of the bias from truncation depends on the decay of the modulus of the complex Laplace transform of the density or its derivative.

This analysis provides guidance on the optimal allocation of computational effort across the various parts of the algorithm in order to balance the three sources of error. We will see that the decay rate of the transform can vary widely across models. Moreover, the relevant rate can be different for an expectation and its sensitivity. Thus, our main conclusion is that to understand the computational cost of this method, the key step is to examine the decay of the transform of the density and its derivative. We provide both theoretical and experimental support for our conclusions.

The implementation of simulation-based LRM estimators involves the underlying probability densities—and thus, in our setting, the use of transform inversion—in two ways: we need to know how to sample from the relevant densities in order to simulate the model, and we need to evaluate the density and its derivative in order to calculate the score function that defines the LRM estimators. There are general methods for sampling from transforms (see Devroye 1981) and specific methods for specific distributions that do not require numerical inversion—see, for example, the methods in §1.7 of Samorodnitsky and Taqqu (1994) for stable distributions. However, these methods do not address the problem of computing the score function. In order to keep our investigation generic, we analyze a method in which numerical transform inversion is used both for sampling and for evaluation of the score. We precompute the cumulative distribution function on a grid and use this approximate distribution function in generating samples. The spacing in this grid is the first of the three sources of error listed above.

The problem of numerical inversion of Fourier and Laplace transforms has been studied extensively. Relative to this literature, our contribution lies in analyzing how the errors in transform inversion affect simulation results when transforms are used for sampling and as an ingredient in sensitivity estimation. There are many inversion techniques to choose from and many choices to make in combining these techniques with simulation—for example, whether to invert the transform for a density or the transform for its distribution function. We will see that the application to simulation motivates certain choices. Given our focus on applying LRM, we have also imposed an important consistency condition between the sampling method and the method used to approximate the score function—namely, that the expectation of the approximate score function be zero with respect to the approximating distribution, just as the exact score function has zero mean with respect to the exact distribution. This constraint drives some of the details of our approach. Indeed, this constraint would not ordinarily hold if one combined an arbitrary sampling method with a score function approximated through transform inversion.

The rest of this paper is organized as follows. Section 2 provides more detailed background and motivation for our investigation, and it introduces examples to which we return in our numerical tests. Section 3 lays out the steps in our method, including numerical transform inversion, approximate sampling, and sensitivity estimation. Section 5 presents our error analysis for both expectations and their sensitivities. Section 6 provides a summary algorithm and illustrates our results through numerical examples. Longer proofs are deferred to an electronic companion, which is available as part of the online version at http://or.journal.informs.org/. Some of the results in this paper appear in preliminary form in Glasserman and Liu (2010).

2. Background and Motivation

To fix ideas, let the random variable X have a density g_θ that depends on a parameter θ, and consider the expectation of $V(X)$, for some nonnegative function V. For example, X may represent the price of an underlying asset (or its logarithm) at some future date, and $V(X)$ the discounted payoff of an option on the asset. To stress the dependence of the expected value of $V(X)$ on the parameter θ, we subscript the expectation by θ and write

$$E_\theta[V(X)] = \int V(x) g_\theta(x)\, dx.$$

In using LRM to estimate the sensitivity with respect to θ, the key LRM identity is

$$\frac{d}{d\theta} E_\theta[V(X)] = E_\theta\left[V(X)\frac{\dot{g}_\theta(X)}{g_\theta(X)}\right], \tag{1}$$

with $\dot{g}_\theta$ denoting the derivative of g_θ with respect to the parameter θ. This follows from bringing the derivative inside the integral and then multiplying and dividing by g_θ. When this identity holds (and it does under mild regularity conditions—see Asmussen and Glynn 2007, Proposition 7.3.5), the expression inside the expectation on the right in (1) provides an unbiased estimator of the sensitivity on the left. We will write this estimator as $V(X)S_\theta(X)$ with

$$S_\theta(x) = \dot{g}_\theta(x)/g_\theta(x)$$

the score function. For a function $V(X_1,\ldots,X_m)$ of independent random variables $X_1,\ldots,X_m$, each with density g_θ, the LRM estimator of the derivative of $E_\theta[V(X_1,\ldots,X_m)]$ is

$$V(X_1,\ldots,X_m)\mathbf{S}_\theta(X_1,\ldots,X_m),$$
$$\mathbf{S}_\theta(x_1,\ldots,x_m) = \sum_{i=1}^{m} \frac{\dot{g}_\theta(x_i)}{g_\theta(x_i)}. \tag{2}$$

The LRM estimator is easy to use when the density g_θ and its derivative are readily available. Moreover, it has the important practical feature of allowing the estimation of sensitivities for many different functions V using a single score function S_θ, and it does not require differentiability of V. The drawback of LRM is that it often suffers from large variance. As explained in the introduction, here we

investigate the application of LRM when the density is not explicitly available but is known through its characteristic function or through its Laplace transform.

Financial applications provide an important source of motivation for this investigation. In the most extensively studied models of asset prices with jumps, the price of an asset at time T is given by $A_T = A_0 \exp(aT + X_T)$, where X_T is the time-T value of a Lévy process with $X_0 = 0$, and A_0 and a are constants. A Lévy process has stationary independent increments, so its increments have infinitely divisible distributions; such distributions are often specified through their characteristic functions, via the Lévy-Khinchine formula (as in, e.g., Sato 1999, p. 37). For example, in the variance gamma (VG) model of Madan et al. (1998) with parameters ρ, ν, and θ, the Laplace transform of X_T is given by

$$L_{\mathrm{VG}}(t) = E[e^{-tX_T}] = \left(\frac{1}{1+\theta\nu t - \rho^2\nu t^2/2}\right)^{T/\nu} \tag{3}$$

for all complex t in a neighborhood of the origin.

As a simple example of a discounted payoff function, we could take

$$\begin{aligned} &e^{-rT}\max\{0, A_T - K\} \\ &\quad = e^{-rT}\max\{0, A_0 e^{aT+X_T} - K\} \equiv V(X_T). \end{aligned} \tag{4}$$

Of particular interest are sensitivities with respect to A_0 (the option's delta) and ρ, which is a type of volatility parameter. As discussed in Madan et al. (1998), θ determines the mean and the sign of the skewness of the distribution, ν determines its kurtosis. Through a change of variables, we can move A_0 and a into the density and thus into the Laplace transform.

In the normal inverse Gaussian (NIG) model introduced by Barndorff-Nielsen (1998), we have with model parameters α, β, μ, and θ,

$$\begin{aligned} L_{\mathrm{NIG}}(t) &= E[e^{-tX_T}] \\ &= \exp\bigl(\theta T\bigl[\sqrt{\alpha^2-\beta^2} - \sqrt{\alpha^2-(\beta-t)^2}\bigr] - \mu T t\bigr) \end{aligned} \tag{5}$$

for all complex t in a neighborhood of the origin. We will examine sensitivities with respect to θ and A_0. We will also note an important contrast between the VG and NIG models that is relevant to our analysis: along a vertical line in the complex plane, the modulus of L_{NIG} decays exponentially, whereas that of L_{VG} has a power decay.

The marginal distributions of Lévy processes are sufficiently flexible to allow fits to many sources of data, including asset returns. However, because Lévy processes necessarily have independent increments, they cannot produce serial dependence or clustering as models of log asset prices. Barndorff-Nielsen and Shephard (2001) introduce serial dependence in a process Y driven by a Lévy process X through a stochastic differential equation of the form

$$dY_t = -\lambda Y_t\, dt + dX_t, \tag{6}$$

λ a constant. When X is Brownian motion, this defines an ordinary Ornstein-Uhlenbeck (OU) process; using other Lévy processes for X provides greater flexibility in shaping the marginals of Y. In this setting, even the characteristic function of Y_T may be unknown. In practice, one would ordinarily simulate a discrete-time approximation to (6). By viewing Y as a function of the increments of X, we may then estimate sensitivities of $E[V(Y_T)]$ using the score function for these increments. Thus, the methods we develop using the characteristic functions for VG and NIG models extend to OU-type processes driven by these Lévy processes.

In another extension of these models, Carr et al. (2003) apply random time changes to Lévy processes; here as well, tractability is limited to transforms of the transition laws. The same is true for the broad class of affine jump-diffusion models characterized in Duffie et al. (2000). Therefore, our method is relevant to these models as well.

It is sometimes possible to find specific methods or approximations to estimate sensitivities in specific models. The VG and NIG processes, for example, can both be represented as time-changed Brownian motions, so a score function can be constructed from the normal distribution of the Brownian motion and the mixing distribution of the random time change. Broadie and Kaya (2006) develop an exact method for certain affine processes. However, the objective of this paper is to develop and analyze a general-purpose method that requires little more than a characteristic function as input. A generic method is useful in the absence of special features of a model; it also serves as a benchmark against which to compare approximations and other methods. In subsequent work (Glasserman and Liu 2010), we have tested approximations and alternative methods using the method developed here as a benchmark.

3. Overview of the Method

In this section, we present the key steps in the method we analyze, explaining how we sample from the characteristic function and how we evaluate the score function. We first detail the one-dimensional case, then discuss the multidimensional generalization.

Let G_θ denote the distribution function associated with the density g_θ. If we knew G_θ, we could sample from this distribution by setting $X = G_\theta^{-1}(U)$, with U uniformly distributed on $[0, 1]$. Similarly, if we knew g_θ and $\dot{g}_\theta$, we could evaluate the score function. Thus, we need methods to accomplish the following tasks, starting from the characteristic function or Laplace transform of g_θ: We need to generate samples from G_θ, and we need to approximate g_θ and $\dot{g}_\theta$ in order to evaluate the score function.

With these objectives in mind, we can outline our procedure as follows:

- On a finite grid of x values, we precompute values of G_θ, g_θ, and $\dot{g}_\theta$.

—Using the Abate-Whitt (1992) algorithm, each transform inversion is approximated using a finite weighted sum of transform values, given in (13).

—Each inversion is subject to two types of error: a discretization error (defined in (14)) and a truncation error (defined in (15)).

- At each simulation step, we

—generate $\hat{X} = \hat{G}_\theta^{-1}(U)$, $U \sim$Unif$[0,1]$, from our approximation $\hat{G}_\theta$, as in (9); and

—evaluate the approximate score function $\hat{S}_\theta = \dot{\hat{g}}_\theta / \hat{g}_\theta$ at $\hat{X}$, using (8) and (10).

- At the end of each path, we evaluate the function V and the sum of the approximate score values, as in (2).

The details of this procedure are specified in this section and the next and then summarized in an algorithm in §6.1. The rest of the paper is devoted to analyzing the bias introduced in the simulation results through the numerical approximations used in the algorithm. The error analysis provides guidance on allocating computational effort to the various pieces of the algorithm—how fine a grid to use, how small a discretization step to use, and how large a truncation point to use. We will see that the truncation point is the parameter that most depends on the underlying transform and drives the optimal allocation of computational effort. These statements will be made precise in §5 and illustrated in §6.

In implementing the overall approach outlined above, we face many choices. For example, we can invert the transform for g_θ and then integrate to get G_θ, or we can invert the transform for G_θ or $1 - G_\theta$ directly. We face similar choices in evaluating g_θ and its derivative. Some of our choices are driven by computational considerations, some are driven by a desire for simplicity. Two general principles are worth highlighting: First, because we sample from an approximation to G_θ, we want the numerical approximation to conform to a genuine probability distribution (a function that increases from 0 to 1) as much as possible. Second, for LRM, we want to impose a consistency condition between the sampling mechanism and the approximate score. For the exact score function S_θ and a sample X from the exact distribution G_θ, we would have $E[S_\theta(X)] = 0$. Similarly, we want to ensure $E[\hat{S}_\theta(\hat{X})] = 0$, when $\hat{S}_\theta$ is our numerical approximation to the score function and $\hat{X}$ is sampled from our approximation to G_θ.

Finally, it should be noted that our need to evaluate the score function for LRM estimation largely eliminates the additional overhead associated with tabulating values of G_θ—we use these function values in approximating the score. Thus, even in cases where other mechanisms might be available for sampling from the transform of g_θ, using a numerical approximation to $\hat{G}_\theta^{-1}$ becomes attractive in our context.

We could restrict ourselves to working with the characteristic function, but there is some practical and expository advantage to working with the full Laplace transform in the complex plane. We use Re and Im to denote the real and imaginary parts of a complex number; thus, for complex $t = \sigma + i\omega$, we have $\mathrm{Re}(t) = \sigma$ and $\mathrm{Im}(t) = \omega$. The two-sided Laplace transform of a function f is given by

$$L_f(t) = \int_{-\infty}^{\infty} e^{-tx} f(x)\, dx,$$

where t is a complex variable. This transform is two-sided because the lower limit of integration is $-\infty$ rather than zero. For background, see Widder (1941, Ch. VI).

For the transform L_{g_θ} of g_θ, we suppose that the region of convergence in the complex plane includes a real interval (σ_l, σ_u), where $\sigma_l < 0$ and $\sigma_u > 0$. (Of the examples in §2, this excludes stable distributions, although it includes tempered stable distributions.) By Widder (1941 Theorem 5b, p. 242), we have $L_{G_\theta}(t) = L_{g_\theta}(t)/t$ for $\mathrm{Re}(t) \in (0, \sigma_u)$, and we have $L_{\bar{G}_\theta}(t) = -L_{g_\theta}(t)/t$ for $\mathrm{Re}(t) \in (\sigma_l, 0)$, where $\bar{G}_\theta = 1 - G_\theta$. The transform of $\dot{g}_\theta$ is the derivative of the transform of g_θ under, for example, the condition in the following lemma, which follows from Proposition 7.3.5 in Asmussen and Glynn (2007).

Lemma 1. *If there is an integrable function H for which $e^{-\sigma x}|\dot{g}_\theta(x)| \leqslant H(x)$ for all real x for all θ in a neighborhood of θ_0, then*

$$\begin{aligned} L_{\dot{g}_\theta}(t) &= \int_{-\infty}^{\infty} e^{-tx} \frac{\partial}{\partial\theta} g_\theta(x)\, dx \\ &= \frac{\partial}{\partial\theta} \int_{-\infty}^{\infty} e^{-tx} g_\theta(x)\, dx = \frac{\partial}{\partial\theta} L_{g_\theta}(t) \end{aligned}$$

at $\theta = \theta_0$ where $\mathrm{Re}(t) = \sigma \in (\sigma_l, \sigma_u)$.

Thus, from the transform of g_θ, we get the transforms of G_θ, $\bar{G}_\theta$, and $\dot{g}_\theta$. By restricting to a subinterval if necessary, we assume that the region of convergence of $L_{\dot{g}_\theta}$ includes (σ_l, σ_u).

Using numerical transform inversion (discussed in detail in §4), we can approximate the value of $G_\theta(x)$ at any x. Our overall approach will be to build a piecewise linear approximation $\hat{G}_\theta$ to the function G_θ, extract from that a piecewise-constant approximation $\hat{g}_\theta$ to the density g_θ, and then differentiate to get an approximation $\dot{\hat{g}}_\theta$ to the differentiated density $\dot{g}_\theta$. At each step, we seek to ensure that $\hat{G}_\theta$ and $\hat{g}_\theta$ have the properties of a probability distribution and density, respectively, and that the approximate score function $\dot{\hat{g}}_\theta / \hat{g}_\theta$ will have expected value zero when evaluated at samples drawn from $\hat{G}_\theta$.

To accomplish these objectives, we construct (in §4) an equally spaced grid $\{x_j, j \in J\}$ on the x-axis, where J is a finite index set and $x_j - x_{j-1} = \delta$ for $j \in J$. At each grid point x_j, we use numerical transform inversion to calculate $G_j = \hat{G}_\theta(x_j)$. Between grid points, we use linear interpolation: for any $x \in (x_{j-1}, x_j)$, we set

$$\hat{G}_\theta(x) = \frac{x - x_{j-1}}{\delta} G_j + \frac{x_j - x}{\delta} G_{j-1}. \tag{7}$$

To ensure that $\hat{G}_\theta$ is increasing, we require $G_j \geqslant G_{j-1}$, for all $j \in J$. This is not automatically guaranteed because

of numerical error in transform inversion; we enforce this property when we discuss inversion in §4. Also, we choose grid points to make $G_{\min} \approx 0$ and $G_{\max} \approx 1$, where $G_{\min} \equiv G_{j_{\min}}$, $G_{\max} \equiv G_{j_{\max}}$ with $j_{\min} = \min\{j \in J\}$, $x_{\min} = x_{j_{\min}}$, and $j_{\max}$ and $x_{\max}$ defined accordingly. For $x < x_{\min}$, we set $\hat{G}_\theta(x) = 0$; for $x > x_{\max}$, we set $\hat{G}_\theta(x) = 1$.

We use a uniform grid for simplicity of implementation and error analysis. It may be possible to achieve better results with a grid tailored to a specific G_θ (or V), but this would require information about the distribution (or payoff) that might not be readily available. Our results are likely to carry over to settings in which the largest and smallest grid spacing remain within a constant multiple of each other as the grid is refined. One could also consider approximating the tail of the distribution by, e.g., an exponential beyond $x_{\min}$ and $x_{\max}$.

We denote by $\hat{X}$ a random variable with distribution $\hat{G}_\theta(x)$. The density of $\hat{X}$ is denoted by $\hat{g}_\theta(x)$ and is given by a piecewise constant function,

$$\hat{g}_\theta(x) = \begin{cases} (G_j - G_{j-1})/\delta, & \text{if } x \in [x_{j-1}, x_j),\ j \in J; \\ 0, & \text{if } x < x_{\min} \text{ or } x > x_{\max}. \end{cases} \tag{8}$$

To sample from $\hat{G}_\theta$, we generate U uniformly over $(G_{\min}, G_{\max})$ and set

$$\hat{X} = \frac{U\delta + x_{j-1}G_j - x_j G_{j-1}}{G_j - G_{j-1}}, \tag{9}$$

with j the index for which $G_{j-1} \leqslant U < G_j$.

By sampling from $\hat{G}_\theta(x)$, we can estimate $E_\theta[V(\hat{X})]$, with E_θ indicating that $\hat{X} \sim \hat{G}_\theta$. In order to estimate the sensitivity $E_\theta[V(\hat{X})\hat{S}_\theta(\hat{X})]$, where $\hat{S}_\theta(x) = \dot{\hat{g}}_\theta(x)/\hat{g}_\theta(x)$ and $\dot{\hat{g}}_\theta(x) = \partial \hat{g}_\theta(x)/\partial\theta$, we compute $\dot{\hat{g}}_\theta(x)$ as follows:

$$\dot{\hat{g}}_\theta(x) = \begin{cases} (\dot{G}_j - \dot{G}_{j-1})/\delta, & \text{if } x \in [x_{j-1}, x_j),\ j \in J; \\ 0, & \text{if } x < x_{\min} \text{ or } x > x_{\max}, \end{cases} \tag{10}$$

where $\dot{G}_j \approx \dot{G}_\theta(x_j)$ is calculated through numerical inversion of the transform of $\dot{G}_\theta = dG_\theta/d\theta$. So, as we compute each G_j to construct the approximation $\hat{G}_\theta$, we also compute $\dot{G}_j$ to be able to evaluate $\dot{\hat{g}}_\theta(x)$. We fix the values $\dot{G}_{j_{\min}} = 0$ and $\dot{G}_{j_{\max}} = 0$; this ensures that $\dot{\hat{g}}_\theta$ integrates to zero over $[x_{\min}, x_{\max}]$, and thus that the expected score is zero in the sense that

$$\int_{-\infty}^{\infty} \hat{S}_\theta(x)\hat{g}_\theta(x)\,dx = \int_{x_{\min}}^{x_{\max}} \dot{\hat{g}}_\theta(x)\,dx = 0.$$

Once these values are computed and stored, sampling is easy and fast (it is $O(\log|J|)$ using binary search), so the key question is the quality of the approximation; i.e., the difference between $E_\theta[V(\hat{X})]$ and $E_\theta[V(X)]$, and the difference between $E_\theta[V(\hat{X})\hat{S}_\theta(\hat{X})]$ and $E_\theta[V(X)S_\theta(X)]$. These differences have several sources, including numerical errors in transform inversion and discretization errors in the approximation $\hat{G}_\theta$.

In the multidimensional case, X is a vector $(X_1, \ldots, X_m)$ of independent random variables. For example, in a Lévy-driven model, these could be the increments of an underlying Lévy process. We sample each X_i, $i = 1, \ldots, m$, from its distribution using (9). Because the X_i are independent, their joint density is the product of their marginal densities, and the score function $\mathbf{S}_\theta$ for the vector is the sum of the score functions of the individual components.

4. The Fourier-Series Method for Laplace Inversion

Abate and Whitt (1992) derived and analyzed a Fourier-series inversion formula for the one-sided Laplace transform, and we follow their approach and build on their analysis. Cai et al. (2007) extend the Abate-Whitt algorithm to two-sided Laplace transforms, which yields the following expression for a continuous function f in terms of its transform L_f, for any σ in the interval of convergence (σ_l, σ_u):

$$f(x) = \frac{e^{\sigma x}}{\pi}\int_0^\infty \big(\mathrm{Re}[L_f(\sigma + i\omega)]\cos(\omega x) - \mathrm{Im}[L_f(\sigma + i\omega)]\sin(\omega x)\big)\,d\omega. \tag{11}$$

We abbreviate this formula as $f(x) = I_{\sigma,x}(L_f)$. We can (and will) allow the choice of σ to depend on x.

Employing the trapezoidal rule to numerically evaluate the infinite integral in (11) with a step size h gives the approximation

$$I^h_{\sigma,x}(L_f) = \frac{he^{\sigma x}}{2\pi}L_f(\sigma) + \frac{he^{\sigma x}}{\pi}\sum_{k=1}^{\infty}\big(\mathrm{Re}[L_f(\sigma + ikh)]\cos(khx) - \mathrm{Im}[L_f(\sigma + ikh)]\sin(khx)\big). \tag{12}$$

See Abate and Whitt (1992, pp. 18–19) for a discussion of the advantages of the trapezoidal rule over higher-order quadrature rules in this setting. As in Abate and Whitt (1992), we truncate the infinite sum in (12); keeping only the first N terms of the series in (12) yields

$$I^{N,h}_{\sigma,x}(L_f) = \frac{he^{\sigma x}}{2\pi}L_f(\sigma) + \frac{he^{\sigma x}}{\pi}\sum_{k=1}^{N}\big(\mathrm{Re}[L_f(\sigma + ikh)]\cdot\cos(khx) - \mathrm{Im}[L_f(\sigma + ikh)]\sin(khx)\big). \tag{13}$$

We call $T_p = Nh$ the *truncation point*.

By applying this method to L_{g_θ}, we obtain the approximations $I^h_{\sigma,x}(L_{g_\theta})$ and $I^{N,h}_{\sigma,x}(L_{g_\theta})$. The discretization error at x resulting from step size h is

$$e^d_\sigma(x) = I^h_{\sigma,x}(L_{g_\theta}) - g_\theta(x); \tag{14}$$

we show in Appendix B that $e^d_\sigma(x) \geqslant 0$. The truncation error is

$$e^t_\sigma(x) = I^{N,h}_{\sigma,x}(L_{g_\theta}) - I^h_{\sigma,x}(L_{g_\theta}). \tag{15}$$

Thus, $I^{N,h}_{\sigma,x}(L_{g_\theta}) = g_\theta(x) + e^d_\sigma(x) + e^t_\sigma(x)$. We similarly define the discretization error $\dot{e}^d_\sigma(x) = I^h_{\sigma,x}(L_{\dot{g}_\theta}) - \dot{g}_\theta(x)$ and the truncation error $\dot{e}^t_\sigma(x) = I^{N,h}_{\sigma,x}(L_{\dot{g}_\theta}) - I^h_{\sigma,x}(L_{\dot{g}_\theta})$ in approximating $\dot{g}_\theta(x)$. Note that the truncation error could be negative, so $I^{N,h}_{\sigma,x}(L_{g_\theta})$, as well as $I^{N,h}_{\sigma,x}(L_{G_\theta})$ and $I^{N,h}_{\sigma,x}(L_{\bar{G}_\theta})$, could be negative too.

We will apply the Fourier-series method in a way that ensures monotonicity of G_j, $j \in J$, and ensures that $G_{j_{\min}}$ approaches 0 and $G_{j_{\max}}$ approaches 1 as $x_{j_{\min}}$ and $x_{j_{\max}}$ approach $-\infty$ and $+\infty$, respectively. First, we make the following observation about the behavior of the inversion method at extreme values of x:

Proposition 1. *For any $\sigma \in (0, \sigma_u)$,*

$$I^{N,h}_{\sigma,x}(L_{G_\theta}) \to 0 \quad \text{as } x \to -\infty;$$

for any $\sigma \in (\sigma_l, 0)$,

$$I^{N,h}_{\sigma,x}(L_{\bar{G}_\theta}) \to 0 \quad \text{as } x \to \infty.$$

Proof. From the formulas for $I^{N,h}_{\sigma,x}(L_{G_\theta})$ and $I^{N,h}_{\sigma,x}(L_{\bar{G}_\theta})$, we see that $I^{N,h}_{\sigma,x}(L_{G_\theta}) = O(e^{\sigma x})$ as $x \to -\infty$ and $I^{N,h}_{\sigma,x}(L_{\bar{G}_\theta}) = O(e^{\sigma x})$ as $x \to \infty$, which yields the conclusion. □

From this result, we see that in order for the G_j to approach 0 and 1 at extreme values of $x_{j_{\min}}$ and $x_{j_{\max}}$, we can pick $\sigma_+ \in (0, \sigma_u)$ and $\sigma_- \in (\sigma_l, 0)$, and let

$$G_j = \begin{cases} I^{N,h}_{\sigma_+,x_j}(L_{G_\theta}), & \text{if } x_j \leqslant 0; \\ 1 - I^{N,h}_{\sigma_-,x_j}(L_{\bar{G}_\theta}), & \text{if } x_j > 0. \end{cases} \tag{16}$$

The transforms and their approximations may diverge if σ_+ is chosen equal to σ_u or σ_- is chosen equal to σ_l (recall that the interval (σ_l, σ_u) determines the region of convergence of L_{g_θ}). To ensure stability, the values of σ_+ and σ_- should therefore not be chosen right next to the boundaries; with this qualification, we have found that the results are not very sensitive to the choices of σ_+ and σ_-. The question of choosing the vertical line in the complex plane for the inversion integral is investigated in Cai et al. (2007).

For the monotonicity of the G_j, we will use the following property of the Fourier-series method.

Proposition 2. *Let f be a density and F its distribution function. Suppose the interval (σ_1, σ_2) is within the region of convergence of L_F and L_f, where $\sigma_1 < 0$ and $\sigma_2 > 0$. Then for any $\sigma \in (0, \sigma_2)$,*

$$\frac{d}{dx} I^{N,h}_{\sigma,x}(L_F) = I^{N,h}_{\sigma,x}(L_f). \tag{17}$$

Similarly, if $\bar{F} = 1 - F$, then for any $\sigma \in (\sigma_1, 0)$,

$$\frac{d}{dx} I^{N,h}_{\sigma,x}(L_{\bar{F}}) = -I^{N,h}_{\sigma,x}(L_f). \tag{18}$$

Because $I^{N,h}_{\sigma,x}(L_{g_\theta}) = g_\theta(x) + e^d_\sigma(x) + e^t_\sigma(x)$ and $e^d_\sigma(x) \geqslant 0$, we may conclude that $I^{N,h}_{\sigma,x}(L_{g_\theta})$ is nonnegative for all sufficiently large N at any point at which $g_\theta(x)$ is strictly positive. From Proposition 2, we see that nonnegativity of $I^{N,h}_{\sigma,x}(L_{g_\theta})$ implies monotonicity of $I^{N,h}_{\sigma,x}(L_{G_\theta})$ and $I^{N,h}_{\sigma,x}(L_{\bar{G}_\theta})$. In practice, we do not know how large N needs to be, so we apply the following rule: if it happens that $G_{j_0} < G_{j_0-1}$ for some j_0, we simply let $G_{j_0} = G_{j_0-1}$ to make G_j, $j \in J$ a monotonically increasing sequence. (Alternatively, one might increase N adaptively, but we have not explored this possibility.) The steps we use to construct the sequence G_j are as follows:

1. Let $x_0 = E_\theta[X] = -L'_{g_\theta}(0)$ and compute G_0 using (16).

2. Let $x_j = x_0 + j\delta$ and $x_{-j} = x_0 - j\delta$. Compute $G_{\pm j}$ by (16). At any x_j close to zero (say $|x_j| < \epsilon$ for some tolerance ϵ much smaller than δ), we use a very large truncation point to ensure accuracy as we switch from one case in (16) to the other. After getting G_j and G_{-j}, we enforce monotonicity by adjusting their values according to the following rule: If $G_j < G_{j-1}$, then set $G_j = G_{j-1}$; if $G_{-j} > G_{-(j-1)}$, then set $G_{-j} = G_{-(j-1)}$.

3. We continue for $j = 1, 2, \ldots$ until we find $j_{\max} > 0$ and $j_{\min} < 0$ such that $G_{j_{\max}} \approx 1$ and $G_{j_{\min}} \approx 0$; we set $J = \{j_{\min}, j_{\min}+1, \ldots, j_{\max}-1, j_{\max}\}$ and use $\{x_j, j \in J\}$ as our grid.

In the next section, we discuss the errors in estimating expectations $E_\theta[V(X)]$ (which we refer to generically as prices) and sensitivities $E_\theta[V(X)S_\theta(X)]$. It will be important to keep in mind that our calculation (16) uses $\sigma_- \in (\sigma_l, 0)$ in computing values at $x > 0$, and uses $\sigma_+ \in (0, \sigma_u)$ for all $x < 0$.

5. Error Analysis

5.1. Error Analysis for Prices

In this section, we analyze the error in estimating a price, i.e., the difference between $E_\theta[V(\hat{X})]$ and $E_\theta[V(X)]$. For simplicity, we let

$$I^{N,h}_x(L_{g_\theta}) = \begin{cases} I^{N,h}_{\sigma_+,x}(L_{g_\theta}), & \text{if } x \leqslant 0, \\ I^{N,h}_{\sigma_-,x}(L_{g_\theta}), & \text{if } x > 0, \end{cases} \tag{19}$$

and let $e_d(x) = \mathbf{1}\{x > 0\}e^d_{\sigma_-}(x) + \mathbf{1}\{x \leqslant 0\}e^d_{\sigma_+}(x)$ and $e_t(x) = \mathbf{1}\{x > 0\}e^t_{\sigma_-}(x) + \mathbf{1}\{x \leqslant 0\}e^t_{\sigma_+}(x)$, where $\mathbf{1}\{\cdot\}$ is the indicator function.

We can decompose the error using

$$\begin{aligned} &|E_\theta[V(\hat{X})] - E_\theta[V(X)]| \\ &\quad = \left| \int_{x_{\min}}^{x_{\max}} V(x)\hat{g}_\theta(x)\,dx - \int_{-\infty}^{\infty} V(x)g_\theta(x)\,dx \right| \\ &\quad \leqslant \left| \int_{x_{\min}}^{x_{\max}} V(x)\hat{g}_\theta(x)\,dx - \int_{-\infty}^{\infty} V(x)I^{N,h}_x(L_{g_\theta})\,dx \right| \qquad (20) \\ &\qquad + \left| \int_{-\infty}^{\infty} V(x)(I^{N,h}_x(L_{g_\theta}) - g_\theta(x))\,dx \right|. \qquad (21) \end{aligned}$$

We will analyze (21) first and then turn to (20). If $V \geqslant 0$, then $e_d \geqslant 0$ implies that

$$\begin{aligned}&\left|\int_{-\infty}^{\infty} V(x)(I_x^{N,h}(L_{g_\theta}) - g_\theta(x))\,dx\right| \\ &\quad = \left|\int_{-\infty}^{\infty} V(x)(e_d(x) + e_t(x))\,dx\right| \\ &\quad \leqslant \int_{-\infty}^{\infty} V(x)e_d(x)\,dx + \left|\int_{-\infty}^{\infty} V(x)e_t(x)\,dx\right|. \qquad (22)\end{aligned}$$

To bound the error, we impose the following condition on L_{g_θ}:

ASSUMPTION 1. *For any σ in (σ_l, σ_u), as $\omega \to \infty$,*

$$|\mathrm{Re}[L_{g_\theta}(\sigma + i\lambda\omega)]| = O(\lambda^{-\alpha_R}\mathrm{Re}[L_{g_\theta}(\sigma + i\omega)])$$

and

$$|\mathrm{Im}[L_{g_\theta}(\sigma + i\lambda\omega)]| = O(\lambda^{-\alpha_I}\mathrm{Im}[L_{g_\theta}(\sigma + i\omega)])$$

uniformly in $\lambda \geqslant 1$, for some $\alpha_R > 0$ and $\alpha_I > 0$.

This assumption is broadly applicable. For example, it holds if $\mathrm{Re}[L_{g_\theta}(\sigma + i\omega)]$ and $\mathrm{Im}[L_{g_\theta}(\sigma + i\omega)]$ are regularly varying functions (of ω) with negative indices, or if $-\log(\mathrm{Re}[L_{g_\theta}(\sigma + i\omega)]$ and $-\log(\mathrm{Im}[L_{g_\theta}(\sigma + i\omega)]$ are regularly varying functions with positive indices. Regular variation makes precise the notion of a power decay (see, e.g., Bingham et al. 1987 for background), so Assumption 1 is consistent with both a power decay and an exponential decay of the real and imaginary parts of the Laplace transform in the indicated directions. In particular, this assumption holds for the variance gamma and normal inverse Gaussian models we use in our numerical tests.

We impose the following condition on the payoff function V:

ASSUMPTION 2. *For $x > 0$, $0 \leqslant V(x) \leqslant C_v e^{v_+ x}$, and for $x < 0$, $0 \leqslant V(x) \leqslant C_v e^{v_- x}$, for some constants $C_v > 0$, $v_+ \in (0, -\sigma_l)$, and $v_- \in (-\sigma_u, 0)$. Also, we assume V to be differentiable at all but finitely many points x.*

This assumption is more than sufficient to ensure that $E_\theta[V(X)]$ exists, and it is satisfied by standard option payoffs, such as the call option payoff in (4). This assumption together with Lemma 1 also ensures the validity of the LRM identity in (1) through Proposition 7.3.5 of Asmussen and Glynn (2007).

For fixed $\sigma_- \in (\sigma_l, 0)$ and $\sigma_+ \in (0, \sigma_u)$, let

$$M_\pm(T_p) = |L_{g_\theta}(\sigma_\pm + iT_p)| \quad \text{and}$$
$$M(T_p) = \max\{M_-(T_p), M_+(T_p)\}.$$

We can now state the following result, which is proved in Appendix B.

PROPOSITION 3. *Under Assumptions 1 and 2, we can find $\sigma_- \in (\sigma_l, 0)$ and $\sigma_+ \in (0, \sigma_u)$ such that*

$$\int_{-\infty}^{\infty} V(x)e_d(x)\,dx = O(e^{-C/h}),$$

for some constant $C > 0$, and

$$\left|\int_{-\infty}^{\infty} V(x)e_t(x)\,dx\right| = O(M(T_p)).$$

Through (22), this determines the order of (21). We next decompose (20) as

$$\begin{aligned}&\left|\int_{x_{\min}}^{x_{\max}} V(x)\hat{g}_\theta(x)\,dx - \int_{-\infty}^{\infty} V(x)I_x^{N,h}(L_{g_\theta})\,dx\right| \\ &\quad \leqslant \left|\int_{-\infty}^{x_{\min}} V(x)I_x^{N,h}(L_{g_\theta})\,dx\right| + \left|\int_{x_{\max}}^{\infty} V(x)I_x^{N,h}(L_{g_\theta})\,dx\right| \\ &\qquad + \left|\int_{x_{\min}}^{x_{\max}} V(x)(\hat{g}_\theta(x) - I_x^{N,h}(L_{g_\theta}))\,dx\right|.\end{aligned}$$

For the last term, we have the following result, proved in Appendix C:

LEMMA 2. *If V is bounded on the interval $[x_{j-1}, x_j]$, then*

$$\left|\int_{x_{j-1}}^{x_j} V(x)(\hat{g}_\theta(x) - I_x^{N,h}(L_{g_\theta}))\,dx\right| = O(\delta^2).$$

If, furthermore, V is differentiable, then

$$\left|\int_{x_{j-1}}^{x_j} V(x)(\hat{g}_\theta(x) - I_x^{N,h}(L_{g_\theta}))\,dx\right| = O(\delta^3).$$

Through this lemma, we arrive at the following result:

PROPOSITION 4. *If V is bounded on $[x_{\min}, x_{\max}]$ and differentiable at all but finitely many points,*

$$\left|\int_{x_{\min}}^{x_{\max}} V(x)(\hat{g}_\theta(x) - I_x^{N,h}(L_{g_\theta}))\,dx\right| = O(\delta^2).$$

For the two tail errors over $(-\infty, x_{\min})$ and $(x_{\max}, \infty)$, we have the following proposition, proved in Appendix D:

PROPOSITION 5. *If $G_{j_{\min}} = \epsilon$, then*

$$\int_{-\infty}^{x_{\min}} V(x)I_x^{N,h}(L_{g_\theta})\,dx = O(\epsilon^{(\sigma_+ + v_-)/\sigma_+});$$

if $G_{j_{\max}} = 1 - \epsilon$, then

$$\int_{x_{\max}}^{\infty} V(x)I_x^{N,h}(L_{g_\theta})\,dx = O(\epsilon^{(\sigma_- + v_+)/\sigma_-}).$$

Because both $(\sigma_+ + v_-)/\sigma_+$ and $(\sigma_- + v_+)/\sigma_-$ are strictly between 0 and 1, this result indicates that

$$\left|\int_{-\infty}^{x_{\min}} V(x)I_x^{N,h}(L_{g_\theta})\,dx\right| \quad \text{and} \quad \left|\int_{x_{\max}}^{\infty} V(x)I_x^{N,h}(L_{g_\theta})\,dx\right|$$

are negligible compared to

$$\left|\int_{x_{\min}}^{x_{\max}} V(x)(\hat{g}(x) - I_x^{N,h}(L_{g_\theta}))\,dx\right|.$$

With this specification, we can combine Propositions 3 and 4 to quantify the pricing error:

THEOREM 1. *Under Assumptions 1 and 2,*

$$|E_\theta[V(\hat{X})] - E_\theta[V(X)]| = O(\delta^2 + e^{-C/h} + M(T_p)).$$

5.2. Error Analysis for Sensitivities

We now extend our analysis to the error in estimating the sensitivities, bounding $|E_\theta[V(\hat{X})\hat{S}_\theta(\hat{X})]-E_\theta[V(X)S_\theta(X)]|$. Much as in the previous section, we define

$$I_x^{N,h}(L_{\dot{g}_\theta}) = \begin{cases} I^{N,h}_{\sigma_+,x}(L_{\dot{g}_\theta}), & \text{if } x \leqslant 0 \\ I^{N,h}_{\sigma_-,x}(L_{\dot{g}_\theta}), & \text{if } x > 0, \end{cases}$$

and we let $\dot{e}_d(x) = \mathbf{1}\{x>0\}\dot{e}^d_{\sigma_-}(x) + \mathbf{1}\{x\leqslant 0\}\dot{e}^d_{\sigma_+}(x)$ and $\dot{e}_t(x) = \mathbf{1}\{x>0\}\dot{e}^t_{\sigma_-}(x) + \mathbf{1}\{x\leqslant 0\}\dot{e}^t_{\sigma_+}(x)$.

We bound the error in the sensitivity estimate as

$$\begin{aligned} &|E_\theta[V(\hat{X})\hat{S}_\theta(\hat{X})] - E_\theta[V(X)S_\theta(X)]| \\ &\quad = \left|\int_{x_{\min}}^{x_{\max}} V(x)\dot{\hat{g}}_\theta(x)\,dx - \int_{-\infty}^{\infty} V(x)\dot{g}_\theta(x)\,dx\right| \\ &\quad \leqslant \left|\int_{x_{\min}}^{x_{\max}} V(x)\dot{\hat{g}}_\theta(x)\,dx - \int_{-\infty}^{\infty} V(x)I_x^{N,h}(L_{\dot{g}_\theta})\,dx\right| \\ &\qquad + \left|\int_{-\infty}^{\infty} V(x)(I_x^{N,h}(L_{\dot{g}_\theta}) - \dot{g}_\theta(x))\,dx\right|. \end{aligned}$$

This bound is very similar to the one we used for the error in the price estimate, but now with derivatives of g_θ. Much as in Assumption 1, we impose

ASSUMPTION 3. *For any σ in (σ_l, σ_u), as $\omega \to \infty$,*

$$\left|\mathrm{Re}[L_{\dot{g}_\theta}(\sigma + i\lambda\omega)]\right| = O\left(\lambda^{-\dot{\alpha}_R}\mathrm{Re}[L_{\dot{g}_\theta}(\sigma+i\omega)]\right)$$

and

$$\left|\mathrm{Im}[L_{\dot{g}_\theta}(\sigma + i\lambda\omega)]\right| = O\left(\lambda^{-\dot{\alpha}_I}\mathrm{Im}[L_{\dot{g}_\theta}(\sigma+i\omega)]\right)$$

uniformly in $\lambda \geqslant 1$, for some $\dot{\alpha}_R > 0$ and $\dot{\alpha}_I > 0$.

For fixed σ_- and σ_+, let

$$\dot{M}_\pm(T_p) = \left|L_{\dot{g}_\theta}(\sigma_\pm + iT_p)\right| \quad \text{and}$$
$$\dot{M}(T_p) = \max\{\dot{M}_-(T_p), \dot{M}_+(T_p)\}.$$

With these assumptions and definitions, the previous analysis carries over with appropriate modification (the function g_θ now replaced by $\dot{g}_\theta$), leading to the following result:

THEOREM 2. *Under Assumptions 2 and 3, using the same σ_- and σ_+ as in Proposition 3,*

$$\int_{-\infty}^{\infty} V(x)|\dot{e}_d(x)|\,dx = O(e^{-\dot{C}/h}),$$

for some positive constant $\dot{C}$,

$$\left|\int_{-\infty}^{\infty} V(x)\dot{e}_t(x)\,dx\right| = O(\dot{M}(T_p))$$

and

$$\begin{aligned} &\left|E_\theta[V(\hat{X})\hat{S}_\theta(\hat{X})] - E_\theta[V(X)S_\theta(X)]\right| \\ &\quad = O(\delta^2 + e^{-\dot{C}/h} + \dot{M}(T_p)). \end{aligned}$$

5.3. The Multidimensional Case

Similar conclusions can be reached for the multidimensional case on the error analysis. First, for the payoff function $V(x_1,\ldots,x_m)$, we assume

ASSUMPTION 4. *$V(x_1,\ldots,x_m) \leqslant V_1(x_1)\cdots V_m(x_m)$ for some functions $V_1,\ldots,V_m$ satisfying Assumption 2.*

Each V_i in this condition is bounded by the exponential of a linear function. When a Lévy process is used to model a log asset price, as in (4), Assumption 4 allows us to bound a path-dependent payoff using linear functions of the underlying asset, and this should cover most cases of practical interest. Under this assumption, the error analysis in the multidimensional case is very similar to that in the one-dimensional case. The details are in the electronic companion to this paper.

6. Algorithm and Numerical Examples

In this section, we encapsulate our analysis into an overall algorithm and then illustrate the theoretical results with numerical examples. We investigate the convergence rate of our method as we vary the grid spacing δ, the step size h, and the truncation parameter T_p, using our error estimates to guide the relative magnitudes of the changes in the parameters.

6.1. Algorithm

In the previous sections, we have focused on the bias in estimating prices and sensitivities. As a measure of overall simulation error, we use mean square error (MSE), which is the sum of the squared bias and the estimator variance. If we use N_s simulation trials, then the MSE for the price estimate is

$$\mathrm{MSE}_{\mathrm{price}} = (O(\delta^2 + e^{-C/h} + M(T_p)))^2 + \frac{\mathrm{Var}_{\mathrm{price}}}{N_s}, \tag{23}$$

and for the sensitivity, the MSE is

$$\mathrm{MSE}_{\mathrm{sen}} = (O(\delta^2 + e^{-\dot{C}/h} + \dot{M}(T_p)))^2 + \frac{\mathrm{Var}_{\mathrm{sen}}}{N_s}, \tag{24}$$

where $\mathrm{Var}_{\mathrm{price}}$ and $\mathrm{Var}_{\mathrm{sen}}$ denote the variance per replication of the price estimate and sensitivity estimate, respectively. To drive each MSE to 0, we need to decrease δ and h and increase T_p and N_s. For efficiency, we should do so at rates consistent with their impact on the MSE.

Algorithm

INITIALIZE δ, h, T_p, and N_s
REPEAT until stopping condition is satisfied:

1. Build grid $\{x_j\}$ and compute $\widehat{G}_\theta$, $\hat{g}_\theta$, and $\dot{\hat{g}}_\theta$ at these grid points;
2. FOR replications $n = 1,\ldots,N_s$
   (a) Generate $\hat{X}$;
   (b) Evaluate $V(\hat{X})$ and $\hat{S}_\theta(\hat{X})$;

3. RETURN sample means of $V(\hat{X})$ and $V(\hat{X})\hat{S}_\theta(\hat{X})$ over N_s replications;
4. UPDATE parameters: $\delta \to \delta'$, $h \to h'$, $T_p \to T_p'$, $N_s \to N_s'$

The grid construction is detailed in Steps 1–3 in §4; the function evaluations and sampling mechanism are specified in (7)–(10). The theoretical analysis in §5 allows us to specify the parameter update rule in the last step to maximize the rate of decrease of the mean square error. For example, from (23) and (24), we see that if we increase the number of replications by a factor of c, (so $N_s \to cN_s$), then we should refine the grid size by a factor of $c^{1/4}$ (so $\delta \to \delta/c^{1/4}$). However, the relative allocation of effort between these parameters and the truncation point T_p depends on the underlying model through $M(T_p)$ and $\dot{M}(T_p)$; we illustrate this point in our examples, where we make the update rule explicit.

The iterative refinement of the computational parameters requires a stopping condition. Ideally, one would like to stop when a target level of accuracy and precision is reached. The variance per replication is easily estimated using the sample variance. The bias is more difficult to estimate because the exact value of the price or sensitivity is unknown; a simple and standard approach is to use the difference in estimates from one iteration to the next as an estimate of the remaining bias. The bias and variance estimates can then be combined to estimate the MSE, and the algorithm stopped once the MSE falls below a desired tolerance.

Our asymptotic results do not provide much guidance on initializing the computational parameters in the first step of the algorithm. Similar issues arise across many applications of simulation. In discretizing a diffusion, for example, one seldom if ever has solid grounds for choosing the initial size of the time step. Even in the absence of bias, choosing an initial number of replications requires some problem-specific information about the magnitude of the variance. Although theoretical support for initial parameter values can be difficult to find, practical considerations often dictate at least rough values for the number of replications, the number of grid points, and integration parameters h and T_p, as in our examples below.

We use the VG and NIG models as test cases. These two models exhibit qualitatively different dependence on T_p; this will be reflected in the rates we use to vary the error parameters and in the overall convergence rate of the MSE. The VG and NIG models have been studied extensively in previous work and offer some tractability that makes them useful as test cases.

6.2. VG Model

We begin by considering the VG model with the call option payoff in (4). From a VG process X, we define an asset price by setting $A_T = A_0 \exp(aT + X_T)$, for some initial price A_0 and parameter a. Formula (22) in Madan et al. (1998) gives

$$a = r + \frac{1}{\nu}\log(1 - \theta\nu - \rho^2\nu/2), \tag{25}$$

with ρ, ν, and θ parameters of the model. Madan et al. (1998) provide an option pricing formula for the VG model, which is useful in evaluating the error in our simulation estimates.

The Laplace transform of X_T appears in (3). Its region of convergence is the vertical strip in the complex plane that intersects the real axis on the interval

$$\left(\frac{\theta\nu - \sqrt{\theta^2\nu^2 + 2\rho^2\nu}}{\rho^2\nu},\ \frac{\theta\nu + \sqrt{\theta^2\nu^2 + 2\rho^2\nu}}{\rho^2\nu}\right).$$

For any σ in this interval, $|L_{\mathrm{VG}}(\sigma + i\omega)|$ has a power decay (as $\omega \to \infty$) with rate $2T/\nu$. Thus,

$$M(T_p) = O(T_p^{-2T/\nu}), \tag{26}$$

and the MSE for the price estimate is therefore

$$\mathrm{MSE}_{\mathrm{price,\,VG}} = (O(\delta^2 + e^{-C/h} + T_p^{-2T/\nu}))^2 + \frac{\mathrm{Var}_{\mathrm{price}}}{N_s}. \tag{27}$$

To reduce the MSE, we would like, ideally, to increase T_p and N_s and decrease δ and h in a way that equates the magnitude of the error reduction in each source of error. Otherwise, we would be devoting too much computational effort to some parts of the algorithm and insufficient effort to others. This yields the following update rule for the algorithm in §6.1:

$$\text{UPDATE:}\quad T_p \to \sqrt{10}T_p,\quad \delta \to 10^{-T/2\nu}\delta,$$
$$N_s \to 10^{2T/\nu}N_s,\quad h^{-1} \to h^{-1} + T\log 10/C\nu.$$

(Our choice of C is specified in the proof of Proposition 3.) With these changes, the RMSE (the square root of the MSE) for the price estimate should decrease by a factor of $10^{T/\nu}$. The number 10 could be replaced with any other constant in the update rule and in this conclusion. All of these statements should be understood in the big-O sense provided by our results.

Although our focus is on the impact of T_p and the other parameters on the convergence of the RMSE, the tail behavior identified in (26) also provides information about the density of X_T: the decay rate of the transform characterizes the smoothness of the density. In particular, from (26) and Feller (1971, Lemma 4, p. 517), we find that if $2T/\nu \leqslant 1$, then the density of X_T does not have an integrable derivative and may fail to be differentiable. In this case, LRM would not be applicable, even if we had a convenient expression for the density. We therefore require $2T/\nu > 1$.

In our examples, we use the following values for the VG process and the call option payoff: $A_0 = 100$, $r = 5\%$,

Table 1. European call prices and derivatives for VG model.

ν	Call price	A_0-derivative	ρ-derivative
1	11.2669	0.7282	23.0434
0.5	10.9292	0.6927	28.5971

$T = 1$, $\rho = 0.2$, $\theta = -0.15$, and we set K at 100 or 125. We compare results at $\nu = 1$ and $\nu = 0.5$. Using the formula in Madan et al. (1998) for the prices of European call options, we get the values in Table 1, against which we compare the simulation estimates.

To test our sensitivity estimates, we calculate sensitivities with respect to the model parameter ρ and the initial price A_0 of the underlying asset. By applying finite difference approximations to the formula for option prices, we get the derivative values in Table 1.

To apply LRM, we need to move the dependence on A_0 and ρ into the density; recall from (25) that a is a function of ρ. We therefore work with the random variable $\log A_0 + aT + X_T$, whose Laplace transform is $A_0^{-t}\exp(-aTt)L_{\mathrm{VG}}(t)$. For the parameter A_0, the Laplace transform of the partial derivative is $-tL_{\mathrm{VG}}(t)/A_0$; for the parameter ρ, the Laplace transform of the derivative is $\partial(A_0^{-t}\exp(-aTt)L_{\mathrm{VG}}(t))/\partial\rho$. In both cases, we get $\dot{M}(T_p) = O(T_p^{-(2T/\nu)+1})$, so the sensitivity MSE is

$$\mathrm{MSE}_{\mathrm{sen,VG}} = (O(\delta^2 + e^{-C/h} + T_p^{-(2T/\nu)+1}))^2 + \frac{\mathrm{Var}_{\mathrm{sen}}}{N_s}. \quad (28)$$

Recall that we have required $2T/\nu > 1$ for the differentiability of the density, and this ensures that the exponent of T_p is negative.

The impact of the truncation point T_p in the sensitivity MSE (28) differs from that in the price MSE (27) and results in a slower overall rate of convergence. For example, with $\nu = 1$, we get $2T/\nu = 2$, so the optimal RMSE for the price is $O(T_p^{-2})$, whereas for the sensitivity it is $O(T_p^{-1})$. Thus, to decrease the price RMSE by a factor of 10, we increase the truncation point by a factor of $\sqrt{10}$, but to decrease the sensitivity RMSE by a factor of 10 we increase the truncation point by a factor of 10:

$$\text{UPDATE:}\quad T_p \to 10T_p, \quad \delta \to 10^{-1/2}\delta,$$
$$N_s \to 10^2 N_s, \quad h^{-1} \to h^{-1} + \log 10/\dot{C}.$$

A corresponding rule applies for $\nu = 0.5$.

Table 2 shows numerical results for price estimates with $\nu = 1$. From each row to the next, we multiply T_p by $\sqrt{10}$ and change the other parameters at the corresponding rates. The initial values are set to make the modulus of $L_{g_\theta}(\sigma_\pm + iT_p)$ reasonably small (some number between 0.1 and 0.5) and J have several points. In the "Error" column, we report the absolute difference between the simulation mean and the formula price. The column labeled "SE" shows the standard error of each estimate. To estimate the standard errors precisely, we simulate all rows for 5E7 replications. We then multiply the standard error obtained by 100 for the first row, 10 for the second row and 1 for the third row. In general agreement with our analysis, the order of magnitude of the price error decreases by roughly a factor of 10 from each row to the next. In order to get reliable estimates for our comparison, we use a larger number of replications than would be optimal under our analysis. In practice, we would try to set the value of N_s to make the standard error approximately equal to the bias.

Our results are calculated using MATLAB running on a laptop computer with a 1.6 GHz Pentium M processor and 1 G of RAM. Even at our largest values of $|J|$, building the grid takes less than a tenth of a second. Once the grid is built, we can run 100,000 replications in under a second. Because we tabulate the transform inversion results, the only computational parameter that affects the computing time per replication is the number of grid points $|J|$. As a point of comparison, with $|J| = 87$ (as in the third row of Table 2), sampling from the grid takes approximately 3.7 times as long as sampling from the normal distribution using MATLAB's built-in implementation of the inverse normal distribution. The sampling time could potentially be accelerated using a cut-point method or alias table.

Table 2 also shows numerical results for the sensitivities with $\nu = 1$. The sensitivity errors decrease by approximately $\sqrt{10}$ from one row to the next, in line with our analysis, although the convergence of the derivative with respect to ρ is a bit faster than expected.

Table 3 shows results for $\nu = 0.5$. In this case, the modulus of the Laplace transform decays more quickly, so we increase T_p by a factor of $\sqrt[4]{10}$ from one row to the next. This should decrease the price error by a factor of 10 and the sensitivity error by a factor of $10^{3/4} \approx 5.6$ in each case.

Table 2. Results for prices and derivatives on A_0 and ρ in VG model, with $\nu = 1$.

				Price			A_0-derivative			ρ-derivative		
T_p	δ	N_s	$\lvert J\rvert$	Mean	Error	SE	Mean	Error	SE	Mean	Error	SE
10	0.5	5E3	12	13.073	1.806	0.2999	0.559	0.169	0.0325	31.557	8.514	2.6898
31.6	0.16	5E5	37	11.588	0.321	0.0201	0.772	0.043	0.0025	21.928	1.115	0.3377
100	0.05	5E7	87	11.300	0.032	0.0018	0.743	0.014	0.0002	22.797	0.246	0.0341
10	0.5	5E3	12	4.120	2.505	0.2051	0.379	0.189	0.0235	29.202	6.949	1.9321
31.6	0.16	5E5	37	2.188	0.573	0.0098	0.240	0.050	0.0012	20.759	−1.493	0.2371
100	0.05	5E7	87	1.663	0.049	0.0009	0.197	0.007	0.0001	22.186	−0.067	0.0232

Note. The top panel is for $K = 100$, and the bottom panel is for $K = 125$.

Table 3. Results for prices and derivatives on A_0 and ρ in VG model, with $\nu = 0.5$.

				Price			A_0-derivative			ρ-derivative		
T_p	δ	N_s	$\lvert J\rvert$	Mean	Error	SE	Mean	Error	SE	Mean	Error	SE
10	0.5	5E3	8	13.179	2.250	0.2862	0.529	0.163	0.0296	37.472	8.875	3.6388
17.8	0.16	5E5	23	11.307	0.378	0.0209	0.752	0.060	0.0024	26.670	1.927	0.3584
31.6	0.05	5E7	67	10.964	0.035	0.0020	0.704	0.011	0.0002	28.461	0.136	0.0360
10	0.5	5E3	8	4.066	2.129	0.1945	0.344	0.133	0.0216	35.835	10.288	2.7194
17.8	0.16	5E5	23	2.399	0.462	0.0105	0.254	0.043	0.0012	24.129	−1.418	0.2510
31.6	0.05	5E7	67	1.982	0.045	0.0009	0.217	0.007	0.0001	25.453	−0.094	0.0243

Note. The top panel is for $K = 100$, and the bottom panel is for $K = 125$.

The results in the tables are roughly in line with these predictions.

6.3. NIG Model

For an NIG process X at time T with parameters $(\alpha, \beta, \theta, \mu)$, the Laplace transform is given in (5). The region of convergence is the vertical strip in the complex plane that intersects the real line on the interval $[\beta - \alpha, \beta + \alpha]$. Whereas the Laplace transform of the VG process exhibits a power decay, that of the NIG process exhibits an exponential decay; more precisely, $|L_{\text{NIG}}(\sigma + i\omega)| \sim \exp(-\theta T\omega)$ as $\omega \to \infty$, for any fixed σ in the convergence interval.

To define an asset price A_T from X_T as in (4), the appropriate value of a is

$$a = r - \mu - \theta\left(\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta+1)^2}\right). \tag{29}$$

For the parameters, we use the following values from Këllezi and Webber (2004):

$$A_0 = 100 \quad K = 100 \quad r = 0.1 \quad T = 1 \quad \alpha = 28.42141$$
$$\beta = -15.08623 \quad \theta = 0.31694 \quad \mu = 0.05851.$$

We also consider the case $K = 125$. As in Këllezi and Webber (2004), numerical integration produces a call price of 11.3599 for these parameters. Numerical differentiation yields 0.8124 as the derivative with respect to A_0 and 5.8165 as the derivative with respect to θ.

By following the same steps as in the VG example, we find that the MSEs for the price and the two sensitivities for the NIG process all have the same order of convergence, given by

$$\text{MSE}_{\{\text{price, sen}\},\,\text{NIG}} = \left(O(\delta^2 + e^{-C/h} + e^{-\theta T T_p})\right)^2 + \frac{\text{Var}_{\{\text{price, sen}\},\text{NIG}}}{N_s}. \tag{30}$$

Notice that the dependence on T_p now shows an exponential decay rather than a power decay. So, to target a factor of 10 reduction in the RMSE, we use the following rule:

$$\text{UPDATE:} \quad T_p \to T_p + \log 10/\theta T, \quad \delta \to 10^{-1/2}\delta,$$
$$N_s \to 10^2 N_s, \quad h^{-1} \to h^{-1} + \log 10/C.$$

The resulting increment in T_p is around 7.265 in our example.

Table 4 shows results for prices and derivatives using this update rule. In this example, we start δ at 0.25 rather than 0.5 so that $|J|$ is not too small. The values of T_p, h are varied in accordance with (30). The values in the "Error" column decrease by roughly a factor of 10 from each row to the next, in agreement with our analysis. Computing times in this example are consistent with those reported for the VG model in the previous section.

6.4. A Path-Dependent Option

We now turn to a multidimensional example and consider the pricing of a discretely monitored Asian option in the

Table 4. Results for prices and derivatives on A_0 and θ in NIG model.

				Price			A_0-derivative			θ-derivative		
T_p	δ	N_s	$\lvert J\rvert$	Mean	Error	SE	Mean	Error	SE	Mean	Error	SE
20	0.25	1E4	9	13.082	1.722	0.1315	0.7641	0.0483	0.0179	2.075	3.741	0.7400
27.3	0.079	1E6	22	11.501	0.141	0.0107	0.8048	0.0076	0.0019	5.463	0.353	0.0665
34.5	0.025	1E8	66	11.373	0.014	0.0010	0.8116	0.0008	0.0002	5.784	0.032	0.0068
20	0.25	1E4	9	2.234	1.208	0.0542	0.2148	0.0298	0.0104	4.356	−1.231	0.4404
27.3	0.079	1E6	22	1.148	0.123	0.0036	0.1931	0.0081	0.0008	5.345	−0.242	0.0318
34.5	0.025	1E8	66	1.038	0.013	0.0003	0.1858	0.0008	0.0001	5.557	−0.030	0.0032

Note. The top panel is for $K = 100$, and the bottom panel is for $K = 125$.

Table 5. Asian call prices and derivatives in NIG model.

				Price			A_0-derivative			θ-derivative		
T_p	δ	N_s	$\lvert J\rvert$	Mean	Error	SE	Mean	Error	SE	Mean	Error	SE
50	0.1	1E4	10	10.727	4.392	0.5470	0.876	0.123	0.0316	−3.51	7.22	0.7048
137	0.032	1E6	31	6.491	0.156	0.0061	0.735	0.017	0.0034	3.42	0.29	0.0728
224	0.01	1E8	95	6.348	0.013	0.0006	0.751	0.002	0.0004	3.69	0.02	0.0078

NIG model. The payoff of the Asian option is a function of the average level of the underlying asset over m equally spaced averaging dates and is given by

$$\max\left(\frac{1}{m}\sum_{k=1}^{m} A_{k\Delta t} - K, 0\right),$$

where

$$A_{k\Delta t} = A_0 \exp\left(ak\Delta t + \sum_{l=1}^{k} X_{\Delta t,l}\right) = A_{(k-1)\Delta t}\exp(a\Delta t + X_{\Delta t,k})$$

and the $X_{\Delta t,k}$, $k = 1,\ldots,m$, are independent increments of the NIG process. To formulate the problem in the setting of §5.3, we set

$$V(x_1,x_2,\ldots,x_m) = \max\left(\frac{A_0}{m}\sum_{k=1}^{m}\exp\left(ak\Delta t + \sum_{l=1}^{k} x_l\right) - K, 0\right).$$

It is easy to see that this function satisfies Assumption 4. As in Equation (30), the MSE is

$$\mathrm{MSE}^{\mathrm{Mult}}_{\{\mathrm{price,\,sen}\},\mathrm{NIG}} = \left(O(\delta^2 + e^{-C/h} + e^{-\theta\Delta t T_p})\right)^2 + \frac{\mathrm{Var}^{\mathrm{Mult}}_{\{\mathrm{price,\,sen}\},\mathrm{NIG}}}{N_s}. \quad (31)$$

In order to measure the error in our method, we need benchmark values of the prices and sensitivities against which to compare. We estimate accurate values using 10^8 simulation replications and exact sampling from the NIG distribution using its representation as a continuous mixture of normals. We use finite difference approximations with common random numbers to estimate sensitivities. This would be prohibitively time consuming in practice but is necessary to get accurate benchmark values for comparison. We use $m = 12$ steps in the average, which corresponds to monthly price fixings with a maturity of $T = 1$ year. We set $K = 100$. We estimate a price of 6.335, a sensitivity to A_0 of 0.7525, and a sensitivity to θ of 3.71. We believe these estimates are accurate to the number of decimal places reported.

We set the initial values $\delta = 0.1$ and $T_p = 50$ to make the modulus of the Laplace transform reasonably small. The initial value of h is obtained by setting $e^{-C/h} = \delta^2$. From each row to the next, we use the following rule:

$$\text{UPDATE:}\quad T_p \to T_p + \log 10/\theta\Delta t, \quad \delta \to 10^{-1/2}\delta,$$
$$N_s \to 10^2 N_s, \quad h^{-1} \to h^{-1} + \log 10/C.$$

The results are shown in Table 5 and are in general agreement with our theoretical analysis. The derivative with respect to θ in the first row has a very large error; we attribute this to the small value of $|J|$ in the first row.

6.5. Integrated OU Process

The Asian option gives rise to a multidimensional problem because the payoff of the option is path dependent. We now consider an example that is multidimensional because of the dynamics of the underlying asset, rather than the form of the payoff. As in Barndorff-Nielsen and Shephard (2001), we introduce a process Y with the dynamics in (6), in which X is a Lévy process. In our example, we take X to be an NIG process and interpret Y as the logarithm of the price of an underlying asset on which we price a standard European call option. This reduces to a problem of estimating the expectation of a function of Y_T and its sensitivities.

We use a Euler scheme for the evolution of Y, and thus replace (6) with the recursion

$$Y_{t+\Delta t} = (1 - \lambda\Delta t)Y_t + X_{t+\Delta t} - X_t.$$

If we set $\Delta t = T/m$, then

$$Y_T = \sum_{k=1}^{m}(1-\lambda\Delta t)^{k-1} X_{\Delta t,k}, \quad (32)$$

where $X_{\Delta t,k}$, $k = 1,\ldots,m$, are i.i.d. increments of X over consecutive intervals of length Δt. Because (32) represents Y_T as a linear combination of independent NIG increments, it would allow us to write the Laplace transform of the Euler approximation to Y_T in terms of the NIG Laplace transform; however, for the purpose of testing our results, we will treat this as a multidimensional problem with a payoff that depends on $X_{\Delta t,k}$, $i = 1,\ldots,m$.

We use the NIG parameters from §6.3 with $K = 100$, and we take $\lambda = 2$ and $m = 12$. By forcing $E[e^{aT+Y_T}] = 1$, we get

$$a = r - \frac{\theta\Delta t}{T}\sum_{k=1}^{m}\left(\sqrt{\alpha^2-\beta^2} - \sqrt{\alpha^2 - (\beta + (1-\lambda\Delta t)^{k-1})^2}\right).$$

The payoff function is $V(x_1,x_2,\ldots,x_m) = \max\{A_0\exp(aT + y_T) - K, 0\}$ where $y_T = \sum_{k=1}^{m}(1-\lambda\Delta t)^{k-1}x_k$. This V satisfies Assumption 4, which leads us to

$$\mathrm{MSE}^{\mathrm{OU}}_{\{\mathrm{price,\,sen}\},\mathrm{NIG}} = (O(\delta^2 + e^{-C/h} + e^{-\theta\Delta t T_p}))^2 + \frac{\mathrm{Var}^{\mathrm{OU}}_{\{\mathrm{price,\,sen}\},\mathrm{et\ al.NIG}}}{N_s}. \quad (33)$$

As in §6.4, we estimate benchmark values through extensive simulation with 10^8 replications. This gives a price of 9.896, a sensitivity to A_0 of 0.9261, and a sensitivity to θ of 1.61. Table 6 illustrates the convergence of the price and derivative estimates computed through transform inversion, and these are in general agreement with our theoretical analysis.

Table 6. European call prices and derivatives in OU model.

				Price			A_0-derivative			θ-derivative		
T_p	δ	N_s	$\lvert J\rvert$	Mean	Error	SE	Mean	Error	SE	Mean	Error	SE
50	0.1	1E4	10	13.155	3.259	0.0517	1.052	0.126	0.0341	−3.31	4.91	0.9648
137	0.032	1E6	31	9.958	0.062	0.0063	0.905	0.021	0.0044	1.37	0.24	0.1392
224	0.01	1E8	95	9.902	0.006	0.0006	0.924	0.002	0.0005	1.59	0.02	0.0151

7. Summary

We have analyzed and tested a method for estimating sensitivities by simulation using the likelihood ratio method when the underlying density is known only through its characteristic function or Laplace transform. The method uses numerical transform inversion and incurs several types of error; we have presented results on the convergence rates of these errors and illustrated these results through numerical examples. The main determinant of the overall convergence rate is the truncation point used in the transform inversion. The convergence rate determined by the truncation point can vary qualitatively between models and, within a given model, can be different for expectations and their sensitivities. In our examples, the variance gamma model exhibits polynomial convergence for both expectations and sensitivities, with a slower rate for the sensitivities; the normal inverse Gaussian model exhibits exponential convergence for both expectations and sensitivities.

8. Electronic Companion

An electronic companion to this paper is available as part of the online version that can be found at http://or.journal.informs.org/.

Acknowledgments

The authors thank the associate editor and referees for many comments and suggestions that helped improve the clarity of this paper. This research is supported in part by NSF grant DMS-0410234.

References

Abate, J., W. Whitt. 1992. The Fourier-series method for inverting transforms of probability distributions. *Queueing Systems: Theory Appl.* **10**(1–2) 5–88.

Asmussen, S., P. W. Glynn. 2007. *Stochastic Simulation*. Springer, New York.

Barndorff-Nielsen, O. E. 1998. Processes of normal inverse Gaussian type. *Finance Stochastics* **2**(1) 41–68.

Barndorff-Nielsen, O. E., N. Shephard. 2001. Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics (with discussion). *J. Roy. Statist. Soc. Ser. B* **63**(2) 167–241.

Bingham, N. H., C. M. Goldie, J. L. Teugels. 1987. *Regular Variation*. Cambridge University Press, Cambridge, UK.

Broadie, M., O. Kaya. 2006. Exact simulation of stochastic volatility and other affine jump diffusion processes. *Oper. Res.* **54**(2) 217–231.

Cai, N., S. G. Kou, Z. Liu. 2007. Two-sided Euler inversion algorithm with a shift parameter. Working paper, Columbia University, New York.

Carr, P., H. Geman, D. Madan, M. Yor. 2003. Stochastic volatility Lévy processes. *Math. Finance* **13**(3) 345–382.

Cont, R., P. Tankov. 2004. *Financial Modelling with Jump Processes*. Chapman & Hall/CRC, Boca Raton, FL.

Devroye, L. 1981. On the computer generation of random variables with a given characteristic function. *Comput. Math. Appl.* **7**(6) 547–552.

Duffie, D., J. Pan, K. Singleton. 2000. Transform analysis and option pricing for affine jump-diffusions. *Econometrica* **68**(6) 1343–1376.

Feller, W. 1971. *An Introduction to Probability Theory and Its Applications*, 2nd ed., Vol. II. Wiley, New York.

Glasserman, P., Z. Liu. 2007. Sensitivity estimates from characteristic functions. *Proc. Winter Simulation Conf.*, IEEE Press, Washington, DC, 932–940.

Glasserman, P., Z. Liu. 2010. Estimating Greeks in simulating Lévy-driven models. *J. Comput. Finance*. Forthcoming.

Kella, O., W. Whitt. 1992. Useful martingales for stochastic storage processes with Lévy input. *J. Appl. Probab.* **29**(2) 396–403

Këllezi, E., N. Webber, 2004. Valuing Bermudan options when asset returns are Lévy processes. *Quant. Finance* **4**(1) 87–100.

Klüppelberg, C., A. Kyprianou, R. Maller. 2004. Ruin probabilities and overshoots for general Lévy insurance risk processes. *Ann. Appl. Probab.* **14**(4) 1766–1801.

Madan, D. B., P. Carr, E. C. Chang. 1998. The variance gamma process and option pricing. *Eur. Finance Rev.* **2**(1) 79–105.

Mikosch, T., S. Resnick, H. Roótzen, A. Stegeman. 2002. Is network traffic approximated by stable Lévy motion or fractional Brownian motion? *Ann. Appl. Probab.* **12**(1) 23–68.

Samorodnitsky, G., M. S. Taqqu. 1994. *Stable Non-Gaussian Random Processes.* Chapman & Hall, London.

Sato, K.-I. 1999. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, Cambridge, UK.

Schoutens, W. 2003. *Lévy Processes in Finance: Pricing Financial Derivatives.* Wiley, New York.

Widder, D. V. 1941. *The Laplace Transform*. Princeton University Press, Princeton, NJ.