Proceedings of the 2007 Winter Simulation Conference
S. G. Henderson, B. Biller, M.-H. Hsieh, J. Shortle, J. D. Tew, and R. R. Barton, eds.

SENSITIVITY ESTIMATES FROM CHARACTERISTIC FUNCTIONS

Paul Glasserman

Columbia Business School
Uris Hall, 3022 Broadway, Room 101
New York, N.Y. 10027, U.S.A.

Zongjian Liu

Department of IE & OR, Columbia University
S.W. Mudd Building, 500 West 120th Street, Room 313A
New York, N.Y. 10027, U.S.A.

ABSTRACT

We investigate the application of the likelihood ratio method (LRM) for sensitivity estimation when the relevant density for the underlying model is known only through its characteristic function or Laplace transform. 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.

1 INTRODUCTION

Stochastic simulation is used widely in the financial industry for the pricing and hedging of options and other derivative securities. Under standard conditions, the price of a derivative security can be represented as the expectation of its discounted payoff. A typical pricing simulation involves simulating paths of the underlying asset or assets, evaluating the discounted payoff on each path, and averaging over the paths.

Such simulations are often used as much for hedging as for pricing, and hedging requires calculation of sensitivities of prices with respect to model parameters, including the initial values of the underlying assets. For sensitivity calculations, the likelihood ratio method (LRM) (or score function methd) is attractive when the payoff is discontinuous in the parameters. To fix ideas, let $V(X)$ denote a (discounted) payoff, which is a function of the random variable X, and suppose X has a density g_θ depending on a parameter θ. The key LRM identity is

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

with E_θ denoting expectation with respect to g_θ, and $\dot{g}_\theta$ denoting the derivative of g_θ with respect to the parameter θ. When this identity holds (as it does under mild regularity conditions), the expression inside the expectation on the right 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.

The application of (1) requires evaluation of the density g_θ and its derivative, and this can limit the scope of the method. 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. This problem arises for broad classes of models used in financial applications, including models driven by Lévy processes (see, e.g., Cont and Tankov 2004) and the affine class of jump-diffusion models studied in Duffie et al. (2000).

An example of a Lévy-driven model is one that models the price of the underlying asset through a process $S_T = S_0\exp(aT + X_T)$, in which X_T is the time-T value of a Lévy process with $X_0 = 0$, and S_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). An extensively studied case of a Lévy-driven model is the Variance Gamma model; in the notation of Madan, Carr and Chang (1998), with parameters ρ, ν, and θ, the Laplace transform of X_T is given by

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

for t in a neighborhood of the origin. There is no closed-form expression for the density of X_T.

We analyze a method in which numerical transform inversion (the Fourier series method of Abate and Whitt

1992) is used both to sample through a Laplace transform (or characteristic function) and to compute an LRM estimator. We quantify various sources of errors in order to provide guidance for setting parameters to accelerate convergence. There are general methods for sampling from transforms (see Devroye 1981) and specific methods for specific distributions that do not require numerical inversion, but these do not address the problem of evaluating the score function. In separate work, we investigate alternatives to numerical transform inversion based on approximations to the score function; a side benefit of the method we discuss here is that it can serve as a benchmark for approximations.

The rest of this paper is organized as follows. In Section 2, we specify a sampling method in which we precompute a table of values of the cumulative distribution function (CDF); this involves discretization and truncation of the domain of the CDF. In Section 3, we review the method of Abate and Whitt (1992) and discuss its application to our problem. Section 4 summarizes the error in calculating prices, and Section 5 does the same for price sensitivities. We illustrate the results numerically in Section 6. We outline a proof of our error analysis in an appendix; complete proofs of all our results will be provided in a full-length article.

2 OUTLINE OF THE METHOD

For simplicity, we limit our discussion to scalar X in (1). Let G_θ denote the CDF associated with g_θ. Our first task is to sample X from G_θ when the distribution is known only through a transform. We will accomplish this by tabulating values of $G_\theta(x)$ calculated through numerical transform inversion, and then using the table to generate samples. We could restrict ourselves to working with the characteristic function; there is little practical difference between shifting the integration contour to invert a characteristic function and working directly with the Laplace transform in the complex plane, so we present the inversion steps using the latter. The two-sided Laplace transform of a function f is given by

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

where $t = \sigma + i\omega$ is a complex variable. This transform is two-sided because the lower limit of integration is $-\infty$ rather than zero. For background on two-sided Laplace transforms, see Widder (1941), Chapter VI.

For the transform L_{g_θ} of g_θ, we suppose that the region of convergence includes an interval (σ_l, σ_u), where $\sigma_l < 0$ and $\sigma_u > 0$. By Widder (1941), p.242, Theorem 5b, we have $L_{G_\theta}(t) = L_{g_\theta}(t)/t$ for $\mathrm{Re}(t) \in (0, \sigma_u)$, and we have $L_{\bar{G}_\theta}(t) = -L_{g_\theta}/t$ for $\mathrm{Re}(t) \in (\sigma_l, 0)$ and $\bar{G}_\theta = 1 - G_\theta$. Under mild condition on g_θ,

$$\begin{aligned} L_{\dot{g}_\theta} &= \int_{-\infty}^{\infty} e^{-tx} \frac{\partial}{\partial\theta} g_\theta(x)dx \\ &= \frac{\partial}{\partial\theta} \int_{-\infty}^{\infty} e^{-tx} g_\theta(x)dx = \frac{\partial}{\partial\theta} L_{g_\theta}. \end{aligned} \quad (3)$$

We assume that the region of convergence of $L_{\dot{g}_\theta}$ also includes (σ_l, σ_u).

Using numerical transform inversion, we can approximate the value of $G_\theta(x)$ at any x. We will build an approximation $\hat{G}_\theta$ to the function G_θ by inverting the transform at a fixed set of x values and interpolating between these values. In more detail, we calculate $\hat{G}_\theta$ as follows:

1. Pick a grid on the x-axis: $\{x_j, j \in J\}$ where J is a finite index set, $x_j - x_{j-1} = \delta$ for $j \in J$. Let $j_{\min} = \min\{j \in J\}$ and $x_{\min} = x_{j_{\min}}$. Define $j_{\max}$ and $x_{\max}$ accordingly.
2. Let G_j denote the approximation to $G_\theta(x_j)$ calculated through numerical transform inversion. Set $G_{\min} \equiv G_{j_{\min}} \approx 0$ and $G_{\max} \equiv G_{j_{\max}} \approx 1$.
3. For any $x \in [x_{j-1}, x_j]$, use piecewise linear interpolation to get $\hat{G}_\theta(x)$:
$$\hat{G}_\theta(x) = \frac{x - x_{j-1}}{\delta} G_j + \frac{x_j - x}{\delta} G_{j-1}. \quad (4)$$
4. For $x < x_{\min}$, let $\hat{G}_\theta(x) = 0$; for $x > x_{\max}$, let $\hat{G}_\theta(x) = 1$.

We defer the selection of δ, $x_{\min}$ and $x_{\max}$ for later discussion. To ensure that $\hat{G}_\theta$ is monotone increasing, we require $G_j \geq G_{j-1}$, for all $j \in J$. While this is not automatically guaranteed because of numerical error in transform inversion, we will enforce this property in the method of the next section.

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 equals $d\hat{G}_\theta(x)/dx$, which is a piecewise constant function:

$$\hat{g}_\theta(x) = \begin{cases} (G_j - G_{j-1})/\delta, & \text{if } x \in [x_{j-1}, x_j), j \in J; \\ 0, & \text{if } x < x_{\min} \text{ or } x > x_{\max}. \end{cases} \quad (5)$$

We sample from $\hat{G}_\theta(x)$ as follows:

1. Generate $U \sim U[0,\ 1]$.
2. Find the index j such that $G_{j-1} \leq U < G_j$.
3. Set
$$\hat{X} = \frac{U\delta + x_{j-1}G_j - x_j G_{j-1}}{G_j - G_{j-1}}. \quad (6)$$

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:

$$\dot{\hat{g}}_\theta(x) = \begin{cases} (\dot{G}_j - \dot{G}_{j-1})/\delta, & \text{if } x \in [x_{j-1}, x_j), j \in J; \\ 0, & \text{if } x < x_{\min} \text{ or } x > x_{\max}, \end{cases} \quad (7)$$

where $\dot{G}_j \approx \dot{G}_\theta(x_j)$ is calculated through numerical inversion of the transform of $\dot{G}_\theta$. So, as we compute each G_j to construct the approximation $\hat{G}_\theta$, we also compute $\dot{G}_j$ in order to be able to evaluate $\dot{\hat{g}}_\theta(x)$.

Once these values are computed and stored, sampling is easy and fast, 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 next section, we discuss transform inversion and the associated error analysis.

3 THE FOURIER-SERIES METHOD FOR LAPLACE INVERSION

Abate and Whitt (1992) defined and analyzed a Fourier-series inversion formula for the one-sided Laplace transform, and we follow their approach. Extending it to the two-sided case (see Cai, Kou and Liu 2007) yields, for a function f and its two-sided Laplace transform L_f,

$$f(x) = \frac{e^{\sigma x}}{\pi}\int_0^\infty \big[\mathrm{Re}[L_f(\sigma+i\omega)]\cos(\omega x) - \mathrm{Im}[L_f(\sigma+i\omega)]\sin(\omega x)\big]\,d\omega. \quad (8)$$

We abbreviate this formula as $f(x) = I_x(L_f)$.

Employing the trapezoidal rule to numerically evaluate the infinite integral in (8) with a step size h gives

$$I^h_{\sigma,x}(L_f) = \frac{he^{\sigma x}}{2\pi}L_f(\sigma) + \frac{he^{\sigma x}}{\pi}\sum_{k=1}^{\infty}\big[\mathrm{Re}[L_f(\sigma+ikh)]\cos(khx) - \mathrm{Im}[L_f(\sigma+ikh)]\sin(khx)\big], \quad (9)$$

where σ can be any point in (σ_l, σ_u) and can be chosen to depend on x.

As in Abate and Whitt (1992), we truncate the infinite sum in (9); let $I^{N,h}_{\sigma,x}(L_f)$ denote the truncation of the series in (9) to the first N terms. We call $T_p = Nh$ the *truncation point*.

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

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

we can show that $e^d_\sigma(x) \geq 0$ (see Appendix A). The truncation error is

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

Thus, $I^{N,h}_{\sigma,x}(L_{g_\theta}) = g_\theta(x) + e^d_\sigma(x) + e^t_\sigma(x)$. Likewise, we define $\dot{e}^d_\sigma(x) = I^h_{\sigma,x}(L_{\dot{g}_\theta}) - \dot{g}_\theta(x)$ and $\dot{e}^t_\sigma(x) = I^{N,h}_{\sigma,x}(L_{\dot{g}_\theta}) - I^h_{\sigma,x}(L_{\dot{g}_\theta})$.

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 $\sigma \in (0, \sigma_u)$,*

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

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

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

Proof: By looking at the formula of $I^{N,h}_{\sigma,x}(L_{G_\theta})$ and $I^{N,h}_{\sigma,x}(L_{\bar{G}_\theta})$, we have $I^{N,h}_{\sigma,x}(L_{G_\theta}) = O(e^{\sigma x})$ and $I^{N,h}_{\sigma,x}(L_{\bar{G}_\theta}) = O(e^{\sigma x})$, 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

$$G_j = \begin{cases} I^{N,h}_{\sigma_+,x_j}(L_{G_\theta}), & \text{if } x_j \leq 0; \\ 1 - I^{N,h}_{\sigma_-,x_j}(L_{\bar{G}_\theta}), & \text{if } x_j > 0. \end{cases} \quad (10)$$

For the monotonicity of the G_j, we will use the following property of the Fourier-series method, which can be verified by direct differentiation:

Proposition 2 *Let f be a density with CDF F. 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 $\sigma \in (0, \sigma_2)$,*

$$\frac{d}{dx}I^{N,h}_{\sigma,x}(L_F) = I^{N,h}_{\sigma,x}(L_f). \quad (11)$$

Similarly, if $\bar{F}(x)$ is the complementary CDF, then for any $\sigma \in (\sigma_1, 0)$,

$$\frac{d}{dx}I^{N,h}_{\sigma,x}(L_{\bar{F}}) = -I^{N,h}_{\sigma,x}(L_f). \quad (12)$$

Because $I^{N,h}_{\sigma,x}(L_{g_\theta}) = g_\theta(x) + e^d_\sigma(x) + e^t_\sigma(x)$ and $e^d_\sigma(x) \geq 0$, 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. The steps we use to construct the sequence G_j are as follows:

1. Let $x_0 = E_\theta[X] = -L'_{g_\theta}(0)$. We start from x_0 in constructing our grid. Compute G_0 by (10). For this value we use a very large truncation point to get an accurate value for G_0.
2. Let $x_j = x_0 + j\delta$ and and $x_{-j} = x_0 - j\delta$. Compute $G_{\pm j}$ by (10). After getting G_j and G_{-j}, we adjust their values by 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 $j = 1, 2, \dots$ until we find $j_{\max} > 0$ and $j_{\min} < 0$ such that $G_{j_{\max}} \approx 1$ or $x_{\max} \equiv x_{j_{\max}}$ is large enough, and $G_{j_{\min}} \approx 0$ or $x_{\min} \equiv x_{j_{\min}}$ is large enough in the negative direction. We will explain how to determine the magnitude of $x_{\max}$ and $x_{\min}$ in the next section.

We then set $J = \{j_{\min}, j_{\min}+1, \dots, j_{\max}-1, j_{\max}\}$ and use $\{x_j, j \in J\}$ as our grid.

In the next two sections, we discuss the errors in estimating prices and sensitivities using the Fourier-series method. It will be important to keep in mind that we use $\sigma_- \in (\sigma_l, 0)$ in computing values at $x > 0$, and we use $\sigma_+ \in (0, \sigma_u)$ for all $x < 0$.

4 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

$$I^{N,h}_x(L_{g_\theta}) = \begin{cases} I^{N,h}_{\sigma_+,x}(L_{g_\theta}), \text{if } x \leq 0; \\ I^{N,h}_{\sigma_-,x}(L_{g_\theta}), \text{if } x > 0, \end{cases} \tag{13}$$

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

We can decompose the error using

$$\begin{aligned} &\left| E_\theta[V(\hat{X})] - E_\theta[V(X)] \right| \\ &= \left| \int_{x_{\min}}^{x_{\max}} V(x)\hat{g}_\theta(x)dx - \int_{-\infty}^{\infty} V(x) g_\theta(x) dx \right| \\ &\leq \left| \int_{x_{\min}}^{x_{\max}} V(x)\hat{g}_\theta(x)dx - \int_{-\infty}^{\infty} V(x) I^{N,h}_x(L_{g_\theta}) dx \right| \quad (14) \\ &\quad + \left| \int_{-\infty}^{\infty} V(x) \left(I^{N,h}_x(L_{g_\theta}) - g_\theta(x) \right) dx \right|. \quad (15) \end{aligned}$$

We will analyze (15) first, and then turn to (14). Note that

$$\begin{aligned} &\left| \int_{-\infty}^{\infty} V(x) \left(I^{N,h}_x(L_{g_\theta}) - g_\theta(x) \right) dx \right| \\ &= \left| \int_{-\infty}^{\infty} V(x)(e_d(x) + e_t(x)) dx \right| \\ &\leq \int_{-\infty}^{\infty} V(x) e_d(x) dx + \left| \int_{-\infty}^{\infty} V(x) e_t(x) dx \right| \quad (16) \end{aligned}$$

In order to bound the error, we need to impose some conditions. Our condition on L_{g_θ} is the following:

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 $\lambda \geq 1$, for some $\alpha_R > 1$ and $\alpha_I > 1$.

This assumption is not very restrictive. 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. See, e.g., Bingham, Goldie and Teugels (1987) for background on regular variation.

We impose the following condition on the payoff function V:

Assumption 2 *For $x > 0$, $0 \leq V(x) \leq C_v e^{v_+ x}$, and for $x < 0$, $0 \leq V(x) \leq C_v e^{v_- x}$, for some constants $C_v > 0$, $v_+ \in (0, -\sigma_l)$, and $v_- \in (-\sigma_u, 0)$.*

This assumption is more than sufficient to ensure that $E_\theta[V(X)]$ exists, and it is satisfied by many standard option payoffs.

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

$$M_\pm(T_p) = |L_{g_\theta}(\sigma_\pm + iT_p)|.$$

We now have the following:

Theorem 1 *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

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

Proof: See Appendix A.

Through (16), this result determines the order of (14). We turn next to (15) and decompose this error term as

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

For the last term, we have the following result:

Lemma 1 *If V is continuous 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:

Theorem 2 *If V is differentiable almost everywhere, then*

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

and there are positive constant $C_{\min}$ and $C_{\max}$ for which

$$\left|\int_{-\infty}^{x_{\min}} V(x)I_x^{N,h}(L_{g_\theta})dx\right| = O(e^{-C_{\min}|x_{\min}|}),$$

and

$$\left|\int_{x_{\max}}^{\infty} V(x)I_x^{N,h}(L_{g_\theta})dx\right| = O(e^{-C_{\max}|x_{\max}|}).$$

Proof: Given Lemma 1, we only need to establish the two tail errors. Since $I_x^{N,h}(L_{g_\theta}) = O(e^{\sigma_- x})$ when $x \to \infty$ and $I_x^{N,h}(L_{g_\theta}) = O(e^{\sigma_+ x})$ when $x \to -\infty$, the result follows. □

This result indicates that we can set $x_{\min}$ and $x_{\max}$ large enough in absolute value to make

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

and

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

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 Theorems 1 and 2 to quantify the pricing error:

Corollary 3 *Under the foregoing conditions,*

$$\begin{aligned}
&\left|E_\theta[V(\hat{X})] - E_\theta[V(X)]\right| = \\
&O(\delta^2) + O(e^{-C/h}) + O(\max\{M_-(T_p), M_+(T_p)\}).
\end{aligned}$$

5 ERROR ANALYSIS FOR SENSITIVITIES

In this section, we analyze the error in estimating the sensitivity, i.e., $|E_\theta[V(\hat{X})\hat{S}_\theta(\hat{X})] - E_\theta[V(X)S_\theta(X)]|$. Much as in the previous section, we define

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

and we let $\dot{e}_d(x) = \mathbf{1}\{x > 0\}\dot{e}_{\sigma_-}^d(x) + \mathbf{1}\{x \le 0\}\dot{e}_{\sigma_+}^d(x)$ and $\dot{e}_t(x) = \mathbf{1}\{x > 0\}\dot{e}_{\sigma_-}^t(x) + \mathbf{1}\{x \le 0\}\dot{e}_{\sigma_+}^t(x)$.

We bound the error in the sensitivity estimate as

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

The form of this bound is very similar to that used for the error in the price estimate, but now with derivatives of g_θ. We require that

$$\int_{-\infty}^{\infty} |\dot{g}_\theta(x)|dx < \infty,$$

and much as in Assumption 1, we impose

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

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

and

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

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

For fixed σ_- and σ_+, let

$$\dot{M}_\pm(T_p) = |L_{\dot{g}_\theta}(\sigma_\pm + iT_p)|.$$

With these assumptions and definitions, the analysis in the previous section goes through with appropriate modification, leading to the following result:

Theorem 4 *Under Assumptions 2 and 3, using the same* σ_- *and* σ_+ *as in Theorem 1,*

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

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

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

6 A NUMERICAL EXAMPLE

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

$$\mathrm{MSE}_{price} = \left(O(\max\{M_-(T_p),M_+(T_p)\}) + O(e^{-C/h}) + O(\delta^2)\right)^2 + \frac{\mathrm{Var}_{price}}{N_s},$$

and for the sensitivity, the MSE is

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

where Var_{price} and Var_{sen} denote the variance per replication of the price estimate and sensitivity estimate, respectively.

Several factors affect the two MSEs, including the truncation parameter T_p, the step size h, the grid parameter δ, and the number of paths N_s. To make each MSE converge to 0, we need to change all of these factors simultaneously, and, for efficiency, we should do so at rates consistent with their impact on the MSE. In this section, we use the Variance Gamma (VG) model (as in, e.g., Madan, Carr and Chang 1998) to illustrate how to change the values of the factors appropriately based on the error analysis.

The function we use is the discounted payoff for a European call option,

$$V(X) = e^{-rT}\max(S_T - K, 0),$$

where T is the maturity of the option and S_T follows formula (22) in Madan, Carr and Chang (1998), in which $S_T = S_0\exp(aT + X_T)$, X is a VG process, and

$$a = r + \frac{1}{\nu}\log(1-\theta\nu-\rho^2\nu/2), \tag{17}$$

with r a constant interest rate and ρ, ν, and θ parameters of the model. The Laplace transform of X_T appears in (2).

The region of convergence of the Laplace transform 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, $|L_{vg}(\sigma+i\omega)|$ has a power decay (as $\omega\to\infty$) with rate $2T/\nu$. Therefore, the MSE for the price in VG model is

$$\mathrm{MSE}_{price,vg} = \left(O(T_p^{-2T/\nu}) + O(e^{-C/h}) + O(\delta^2)\right)^2 + \frac{\mathrm{Var}_{price}}{N_s}. \tag{18}$$

To reduce the MSE, we need to increase T_p, decrease h, decrease δ, and increase N_s. The purpose of our error analysis is to guide the allocation of computational effort. We increase or decrease these parameters to equate the magnitude of the error reduction in each source of error. From (18), we see that if T_p increases by a factor of 10, then h should decrease by a factor of $C\nu/(2T\log 10)$, δ should decrease by a factor of $10^{T/\nu}$, and the number of replications should increase by a factor of $10^{4T/\nu}$. (Our choice of C is specified in the proof of Theorem 1 in the Appendix.) With these changes, the RMSE (the square root of the MSE) for the price estimate should decrease by a factor of $10^{2T/\nu}$.

The rate of decrease of the RMSE is constrained by the slowest rate in (18); if we were to change the parameters T_p, h, δ, and N_s without equating the overall rates of decrease in the corresponding error terms, we would be allocating too much computational effort to some parts of the algorithm, insufficient effort to others. All of these statements should be understood in the big-O sense provided by our results.

In our examples, we use the following values for the VG process and the call option payoff:

$$\begin{array}{ccc} S_0 = 100 & K = 100 & r = 0.05 \\ T = 1 & \rho = 0.2 & \theta = -0.15 \end{array}$$

We compare results at $\nu = 1$ and $\nu = 0.5$. Using the formula in Madan, Carr and Chang (1998) for the prices of European call options, we get the values in Table 1, against which we compare the simulation estimates.

Table 1: European call prices for VG model

ν	Call Price
1	11.2669
0.5	10.9292

To test our sensitivity estimates, we calculate sensitivities with respect to the model parameter ρ and the initial price S_0 of the underlying asset. By applying finite difference approximations to the formula for option prices, we get the derivative values in Table 2.

Table 2: Derivatives for VG model

Parameter	Derivatives	
	$\nu = 1$	$\nu = 0.5$
S_0	0.7282	0.6927
ρ	23.0434	28.5971

To apply LRM, we need to move the dependence on S_0 and ρ into the density; recall from (17) that a is a function of ρ. We therefore work with the random variable $\log S_0 + aT + X_T$, whose Laplace transform is $S_0^{-t}\exp(-aTt)L_{vg}(t)$. For the parameter S_0, the Laplace transform of the partial derivative is $-tL_{vg}(t)/S_0$; for the parameter ρ, the Laplace transform of the derivative is $\partial(S_0^{-t}\exp(-aTt)L_{vg}(t))/\partial\rho$. In both cases, the sensitivity MSE is

$$\mathrm{MSE}_{sen,vg} = \left(O(T_p^{-(2T/\nu)+1}) + O(e^{-C/h}) + O(\delta^2)\right)^2 + \frac{\mathrm{Var}_{sen}}{N_s}. \quad (19)$$

The impact of the truncation point T_p in the sensitivity MSE (19) differs from that in the price MSE (18) 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. A similar comparison applies with $\nu = 0.5$. In each case, we also change h, δ and N_s consistent with (19) and (18).

Table 3 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 (somewhat arbitrarily) by equating $T_p^{-2} = \delta^2 = e^{-C/h}$. In the "Error" column, we report the difference between the simulation mean and the formula price. In general agreement with our analysis, the 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.

Table 3: Results for prices, with $\nu = 1$.

T_p	δ	N_s	Mean	Error	Std
20	0.05	5E4	11.0835	-0.1834	0.0593
63.25	0.0158	5E6	11.2476	-0.0193	0.0058
200	0.005	5E8	11.2622	-0.0047	0.0006

Tables 4 and 5 show numerical results for the sensitivities with $\nu = 1$. The error decreases by approximately $\sqrt{10}$ from one row to the next, in line with our analysis.

Table 4: Results for sensitivities to S_0, with $\nu = 1$.

T_p	δ	N_s	Mean	Error	Std
20	0.05	5E4	0.8842	0.1560	0.0105
63.25	0.0158	5E6	0.7812	0.0530	0.0009
200	0.005	5E8	0.7508	0.0226	0.00008

Table 5: Results for sensitivities to ρ, with $\nu = 1$.

T_p	δ	N_s	Mean	Error	Std
20	0.05	5E4	19.9978	-3.0456	1.6666
63.25	0.0158	5E6	22.5957	-0.4478	0.1242
200	0.005	5E8	22.9118	-0.1316	0.0114

Tables 6, 7 and 8 show numerical results for $\nu = 0.5$. In this case, the modulus of the Laplace transform decays more quickly, so we start with a smaller value of T_p and increase it 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. The results in the tables are roughly in line with these predictions, though the convergence in Table 6 is a bit slower than expected.

7 SUMMARY

We have proposed and tested a method for estimating price sensitivities by simulation using the likelihood ratio method when the underlying density is known only through its characteristic function or Laplace transform. The method

Table 6: Results for prices, with $\nu = 0.5$.

T_p	δ	N_s	Mean	Error	Std
12	0.0069	5E4	10.7984	-0.1308	0.0621
22.33	0.0020	5E6	10.8670	-0.0622	0.0062
41.57	0.0006	5E8	10.9172	-0.0120	0.0006

Table 7: Results for sensitivities to S_0, with $\nu = 0.5$.

T_p	δ	N_s	Mean	Error	Std
12	0.0069	5E4	0.6731	-0.0196	0.0069
22.33	0.0020	5E6	0.6894	-0.0032	0.0008
41.57	0.0006	5E8	0.6922	-0.0004	0.00007

Table 8: Results for sensitivities to ρ, with $\nu = 0.5$.

T_p	δ	N_s	Mean	Error	Std
12	0.0069	5E4	26.0214	-2.5757	1.4963
22.33	0.0020	5E6	27.9911	-0.6060	0.1211
41.57	0.0006	5E8	28.4800	-0.1171	0.0116

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 in the Variance Gamma model. In this example, the main determinant of the overall convergence rate is the truncation point used in the transform inversion, and this parameter results in slower convergence of sensitivity estimates than of price estimates.

A THEOREM 1: SKETCH OF PROOF

In proving the first statement in the theorem, we use the Poisson summation formula in Abate and Whitt (1992), Section 5, and get

$$e_{\sigma}^{d}(x) = \sum_{k=-\infty, k\neq 0}^{\infty} \exp\left(\frac{-2\pi\sigma k}{h}\right) g_{\theta}\left(x + \frac{2\pi k}{h}\right). \quad (20)$$

Because g_θ is nonnegative, $e_\sigma^d(x) \geq 0$, and $e_\sigma^d(x) = 0$ if and only if $g_\theta(x + 2\pi k h^{-1}) = 0$ for all nonzero k.

To simplify notation, we now write g and L instead of g_θ and L_{g_θ}. Because L is finite on (σ_l, σ_u), for any σ in this interval we have $g(x) < e^{\sigma x}$ for all sufficiently large $|x|$. In particular, we can choose $\varepsilon > 0$ sufficiently small to have $\sigma_{l,\varepsilon} \equiv \sigma_l + \varepsilon < -\nu_+$ and $\sigma_{u,\varepsilon} \equiv \sigma_u - \varepsilon > -\nu_-$, and then have $g(x) < e^{\sigma_{l,\varepsilon} x}$, for all sufficiently large x, and $g(x) < e^{\sigma_{u,\varepsilon} x}$ for all sufficiently large $-x$.

Now suppose $x > 0$ and take $\sigma = \sigma_-$, a negative number. For sufficiently small $h > 0$, (20) gives

$$e_{\sigma_-}^{d}(x) \leq \sum_{k=1}^{\infty} \exp\left(\frac{-2\pi\sigma_- k}{h} + \sigma_{l,\varepsilon}\left(x + \frac{2\pi k}{h}\right)\right) + \sum_{k=-\infty}^{-1} \exp\left(\frac{-2\pi\sigma_- k}{h}\right) g\left(x + \frac{2\pi k}{h}\right). \quad (21)$$

The first term is less than or equal to

$$2\exp\left(\sigma_{l,\varepsilon} x - \frac{2\pi(\sigma_- - \sigma_{l,\varepsilon})}{h}\right),$$

if h is sufficiently small. It follows that

$$\begin{aligned}
&\int_0^\infty V(x) e_{\sigma_-}^d(x)dx \\
&\leq C_\nu \int_0^\infty \left[2e^{(\sigma_{l,\varepsilon}+\nu_+)x} e^{-\frac{2\pi(\sigma_- - \sigma_{l,\varepsilon})}{h}} \right. \\
&\quad \left. + \sum_{k=-\infty}^{-1} e^{\frac{-2\pi\sigma_- k}{h}} e^{\frac{-2\pi\nu_+ k}{h}} e^{\nu_+(x+2\pi k/h)} g(x + \frac{2\pi k}{h})\right] dx
\end{aligned}$$

The first term on the right can be integrated to get

$$\frac{-2C_\nu}{\sigma_{l,\varepsilon} + \nu_+} \exp\left(-\frac{2\pi(\sigma_- - \sigma_{l,\varepsilon})}{h}\right).$$

We bound the second term by

$$\begin{aligned}
&C_\nu \sum_{k=-\infty}^{-1} e^{\frac{-2\pi(\sigma_- + \nu_+)k}{h}} \int_0^\infty e^{\nu_+(x+2\pi k/h)} g(x + \frac{2\pi k}{h}) dx \\
&\leq C_\nu L(-\nu_+) \sum_{k=-\infty}^{-1} e^{\frac{-2\pi(\sigma_- + \nu_+)k}{h}} \\
&\leq 2C_\nu L(-\nu_+) e^{\frac{2\pi(\sigma_- + \nu_+)}{h}}.
\end{aligned}$$

So, if we let $\sigma_- = (\sigma_l - \nu_+)/2$, then

$$\int_0^\infty V(x) e_{\sigma_-}^d(x) dx = O(e^{-C_1/h}),$$

with $C_1 = -\pi(\sigma_l + \nu_+)$.

By a similar argument, when when $x < 0$, we can let $\sigma_+ = (\sigma_u - \nu_-)/2$ and $C_2 = \pi(\sigma_u + \nu_-)$ to get

$$\int_{-\infty}^0 V(x) e_{\sigma_+}^d(x) dx = O(e^{-C_2/h}).$$

Setting $C = \min\{C_1, C_2\}$ concludes the proof of the first statement in the theorem.

To illustrate the argument for the second part of the theorem, we simplify to $V \equiv 1$. Note that, using Assumption 1,

$$
\begin{aligned}
&\Big|\int_0^\infty e^t_{\sigma_-}(x)dx\Big| \\
&\le \frac{h}{\pi}\sum_{k=1}^{\infty}\Big(\big|\mathrm{Re}[L(\sigma_- + iT_p + ikh)]\int_0^\infty e^{\sigma_- x}\cos((kh+T_p)x)dx\big| \\
&\quad + \big|\mathrm{Im}[L(\sigma_- + iT_p + ikh)]\int_0^\infty e^{\sigma_- x}\sin((kh+T_p)x)dx\big|\Big) \\
&\le \frac{h}{\pi}\sum_{k=1}^{\infty}\Big(\big|\mathrm{Re}[L(\sigma_- + iT_p(1+\frac{k}{N}))]\big|\frac{-\sigma_-}{\sigma_-^2+(kh+T_p)^2} \\
&\quad + \big|\mathrm{Im}[L(\sigma_- + iT_p(1+\frac{k}{N}))]\big|\frac{kh+T_p}{\sigma_-^2+(kh+T_p)^2}\Big) \\
&\le \frac{h}{\pi}\sum_{k=1}^{\infty}\Big(\big|\mathrm{Re}[L(\sigma_- + iT_p)]\big|(1+\frac{k}{N})^{-\alpha_R}\frac{-\sigma_-}{\sigma_-^2+(kh+T_p)^2} \\
&\quad + \big|\mathrm{Im}[L(\sigma_- + iT_p)]\big|(1+\frac{k}{N})^{-\alpha_I}\frac{kh+T_p}{\sigma_-^2+(kh+T_p)^2}\Big) \\
&\le \frac{h}{\pi}|L(\sigma_- + iT_p)|\sum_{k=1}^{\infty}\Big((1+\frac{k}{N})^{-\alpha_R}\frac{-\sigma_-}{\sigma_-^2+(kh+T_p)^2} \\
&\quad + (1+\frac{k}{N})^{-\alpha_I}\frac{kh+T_p}{\sigma_-^2+(kh+T_p)^2}\Big) \\
&= O(|L(\sigma_- + iT_p)|)
\end{aligned}
$$

Similarly, the integral from $-\infty$ to 0 is $O(|L(\sigma_+ + iT_p)|)$. This conclusion continues to hold for V satisfying Assumption 2. □

ACKNOWLEDGMENTS

This research is supported in part by NSF grants DMS007463 and DMI0300044.

REFERENCES

Abate, J. and W. Whitt. 1992. The Fourier-series method for inverting transforms of probability distributions. *Queueing Systems: Theory and Applications* 10:5–88.

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

Cai, N., S.G. Kou and Z. Liu. 2007. Manuscript in preparation.

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

Devroye, L.. 1981. On the computer generation of random variables with a given characteristic function. *Computers and Mathematics with Applications* 7:547–552.

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

Madan, D. and P. Carr and E. Chang. 1998. The variance gamma process and option pricing. *European Finance Review* 2:79–105.

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

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

AUTHOR BIOGRAPHIES

PAUL GLASSERMAN is the Jack R. Anderson Professor in the Decision, Risk, and Operations Division of Columbia Business School. His research interests include modeling and computational problems in risk management and the pricing of derivative securities. He is author of the book *Monte Carlo Methods in Financial Engineering* (Springer, 2004). His e-mail address is ⟨pg20@columbia.edu⟩ and his web page is ⟨www.paulglasserman.net⟩.

ZONGJIAN LIU is a PhD candidate in the Department of Industrial Engineering and Operations Research at Columbia University. His research interests include security pricing and portfolio management. His e-mail address is ⟨zl2115@columbia.edu⟩ and his web page is ⟨http://www.columbia.edu/~zl2115/⟩.