Variance-Gamma and Monte Carlo

Michael C. Fu

Robert H. Smith School of Business
Department of Decision & Information Technologies
University of Maryland
College Park, MD 20742, USA
mfu@rhsmith.umd.edu

Summary. The Variance-Gamma (VG) process was introduced by Dilip B. Madan and Eugene Seneta as a model for asset returns in a paper that appeared in 1990, and subsequently used for option pricing in a 1991 paper by Dilip and Frank Milne. This paper serves as a tutorial overview of VG and Monte Carlo, including three methods for sequential simulation of the process, two bridge sampling methods, variance reduction via importance sampling, and estimation of the Greeks.

Key words: Variance-Gamma process; Lévy processes; Monte Carlo simulation; bridge sampling; variance reduction; importance sampling; Greeks; perturbation analysis; gradient estimation.

1 Introduction

1.1 Reflections on Dilip

Dilip and I have been colleagues since 1989, when I joined the faculty of the Business School at the University of Maryland, and just one year after Dilip himself had returned to his (Ph.D.) alma mater after twelve years on the faculty at the University of Sydney, Australia, following the earning of his two Ph.D.s in pure math (1975) and economics (1971). However, because we were in different departments—he in the Finance Department and I in what was then called the Management Science & Statistics Department—we did not really start collaborating until the mid-1990s when I first became interested in financial applications. Since then, Dilip and I have co-chaired five Ph.D. student dissertations (Tong Wang, Rongwen Wu, Xing Jin, Yi Su, Sunhee Kim), and each of us serves regularly on dissertation committees of the other's Ph.D. students. We team-taught a course on computational finance twice (1997, 1999). We have led a research group on computational/mathematical finance since the late 1990s, which became known as the Mathematical Finance Research Interactions Team (RIT) under the interdisciplinary Applied Mathematics and

Scientific Computing program at the University of Maryland. From this group have graduated scores of Ph.D. students (mostly supervised by Dilip; see the photo in the preface) who have almost all gone on to high-octane "quant" positions at financial institutions on Wall Street and in the Washington, D.C. area. Dilip has always been an inspiration to us all and an ideal academic colleague, always energetic and enthusiastic and full of ideas, whether it be at the research meetings with students or during our regular lunch strolls to downtown College Park.

As Dilip has told me numerous times in our discussions, on Wall Street there are really two main numerical models/techniques employed for pricing derivative securities: Monte Carlo simulation or numerical solution of partial differential equations. The former depends on the first fundamental theorem of asset pricing: the existence of a martingale measure so that the price can be expressed as the expectation of an appropriately discounted payoff function. For exotic, hybrid, and complicated path-dependent derivatives, however, simulation is often the only method available. Of course, Dilip's main research interests do not lie in simulation, but he is a regular consumer of the method, and as a result, we collaborated on two papers in the area, on control variates for pricing continuous arithmetic Asian options [12] and on pricing American-style options [11]. Furthermore, many of his recent papers address simulation of processes, for example, [21]. This brief tutorial is intended to be my small tribute to Dilip and also to provide a bridge or segue between the historical perspective on the VG process provided in the opening article by Eugene Seneta [27] and the remarkable properties of the gamma process described in the subsequent piece by Marc Yor [28], who provides a far more advanced and sophisticated view of gamma processes than the rudimentary presentation here.

1.2 The Variance-Gamma Process

The Variance-Gamma (VG) process was introduced to the finance community as a model for asset returns (log-price increments) and options pricing in Madan and Seneta [20], with further significant developments in Madan and Milne [19] and Madan et al. [18]. More recently, it has been applied to American-style options pricing in [15] and [17], the latter using the fast Fourier transform introduced in [6]. For more history on the earlier pioneering years, see the paper in this volume by Eugene Seneta [27], which also discusses distributions similar to VG proposed earlier as a possible alternative to the normal distribution (see also [22], [23, p. 166]).

The VG process is a Lévy process, that is, it is a process of independent and stationary increments (see the appendix for a review of basic definitions). A Lévy process can be represented as the sum of three independent components (cf. [26]): a deterministic drift, a continuous Wiener process, and a pure-jump process. Brownian motion is a special case where the latter is zero, and the Poisson process is a special case on the other end where the first

two components are zero. Like the Poisson process, the VG process is pure jump; that is, there is no continuous (Brownian motion and deterministic) component, and thus it can be expressed in terms of its Lévy density, the simplest version with no parameters being

$$k(x) = \frac{1}{|x|} e^{-\sqrt{2}|x|}.$$

For the VG process with the usual (θ, σ, ν) parameterization, the Lévy density is given by

$$k(x) = \frac{1}{\nu|x|} \exp\left(\frac{\theta}{\sigma^2} x - \frac{1}{\sigma} \sqrt{\frac{2}{\nu} + \frac{\theta^2}{\sigma^2}} |x| \right), \tag{1}$$

where $\nu, \sigma > 0$.

The VG is a special case of the CGMY model of [4, 5], whose Lévy density is given by

$$k(x) = \begin{cases} C \exp(-G|x|)/|x|^{1+Y} & x < 0, \\ C \exp(-M|x|)/|x|^{1+Y} & x > 0, \end{cases}$$

where VG is obtained by taking $Y = 0, C = 1/\nu, G = 2\mu_+/\sigma^2$, and $M = 2\mu_-/\sigma^2$, where $\mu_\pm$ are defined below in (4). Unlike the Poisson process, the VG process may have an infinite number of (infinitesimally small) jumps in any interval, making it a process of infinite activity. Unlike Brownian motion, the VG process has finite variation, so it is in some sense less erratic in its behavior.

The representation of the VG process presented above hides its roots, which come from the following well-known alternative representations:

1. Time-changed (subordinated) Brownian motion, where the subordinator is a gamma process.
2. Difference of two gamma processes.

As described in detail in [27], these are the original ways in which the process was introduced and proposed as a model for asset returns. The CGMY process can also be expressed as time-changed Brownian motion (see Madan and Yor [21]), but it seems that there is no simple representation as the difference of two increasing processes, although in principle such a representation does exist, because all processes with finite variation can be so expressed.

To express the process in terms of the two representations above, let $\{W_t\}$ denote standard Brownian motion (Wiener process), $B_t^{(\mu,\sigma)} \equiv \mu t + \sigma W_t$ denote Brownian motion with constant drift rate μ and volatility σ, $\gamma_t^{(\mu,\nu)}$ the gamma process with drift parameter μ and variance parameter ν, and $\gamma_t^{(\nu)}$ the gamma process with unit drift ($\mu = 1$) and variance parameter ν. Letting $\phi_X(u) = E[e^{iuX}]$ denote the characteristic function (c.f.) of random variable X, the c.f. for the VG process is given simply by

$$\phi_{X_t}(u) = (1 - iu\theta\nu + \sigma^2 u^2 \nu/2)^{-t/\nu},$$

which can be expressed in two forms

$$\phi_{X_t}(u) = [(1 - iu\nu\mu_+)(1 + iu\nu\mu_-)]^{-t/\nu},$$
$$\phi_{X_t}(u) = [1 - i\nu(u\theta + i\sigma^2 u^2/2)]^{-t/\nu},$$

reflecting the two representations above. In particular, using the notation introduced above, the representation of the VG process as time-changed Brownian motion is given by

$$X_t = B^{(\theta,\sigma)}_{\gamma_t^{(\nu)}} = \theta\gamma_t^{(\nu)} + \sigma W_{\gamma_t^{(\nu)}}, \tag{2}$$

whereas the difference-of-gammas representation is given by

$$X_t = \gamma_t^{(\mu_+,\nu_+)} - \gamma_t^{(\mu_-,\nu_-)}, \tag{3}$$

where the two gamma processes are independent (but defined on a common probability space) with parameters

$$\mu_\pm = (\sqrt{\theta^2 + 2\sigma^2/\nu} \pm \theta)/2, \tag{4}$$
$$\nu_\pm = \mu_\pm^2 \nu.$$

In general, there is no unique martingale measure for a Lévy process, due to the jumps, and thus this is the case for the (pure jump) VG process. Assume the asset price dynamics for a Lévy process $\{X_t\}$ (with no dividends and constant risk-free interest rate) are given by

$$S_t = S_0 \exp((r + \omega)t + X_t), \tag{5}$$

where the constant ω is such that the discounted asset price is a martingale; that is, it must satisfy

$$E[e^{-rt} S_t] = S_0,$$

which leads to the condition

$$e^{-\omega} = \phi(-i),$$

where ϕ denotes the characteristic function of the Lévy process. In the case of VG, we have

$$\omega = \ln(1 - \theta\nu - \sigma^2\nu/2)/\nu. \tag{6}$$

2 Monte Carlo Simulation

General books on Monte Carlo (MC) simulation for financial applications include [16], [14] and [23]; see also [2]. Monte Carlo simulation is most fruitful for "high-dimensional" problems, which in finance are prevalent in path-dependent options, such as Asian, lookback, and barrier options, all of which

are considered in [1], from which most of our discussion on bridge sampling for simulating VG is taken. The importance sampling discussion following that is based on [29]. Efficiently estimating the Greeks from simulation is included in [10], [16], [14], and [23]. Although extensions such as the CGMY process [4] are not treated explicitly here, comments are occasionally made on how the methods extend to that more general setting.

We begin by presenting sequential sampling and bridge sampling techniques for constructing sample paths of a VG process. Sequential sampling is called incremental path construction in [16], which also includes another technique for the Wiener process based on a spectral decomposition using an orthogonal Hilbert basis.

2.1 Sequential Sampling

There are three main methods to simulate VG (cf. [1, 29]). The first two are based on the two representations presented in the previous section and are "exact" in the sense of having the correct distribution. The third method for simulating VG is to approximate it by a compound Poisson process. The main advantage of the third method is its generality, in that it can be used for any Lévy process, in particular in those settings where a representation as subordinated Brownian motion or as the difference of two other easily simulated processes is not readily available. Representation of the CGMY process as time-changed Brownian motion is treated in [21], which also includes procedures for simulating using that representation; see also [29].

Figure 1 presents the three different algorithms for sequentially generating VG sample paths on $[0, T]$ at time points $0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = T$, where the time spacings $\Delta t_i, i = 1, \ldots, N$ are given as inputs, along with VG parameters. In the compound Poisson process approximation, the positive and negative jumps are separated, and there is a cutoff of ε for the magnitude of the jumps. The Poisson rates are calculated by integrating the Lévy density over the appropriate range, and the jump sizes are then sampled from the corresponding renormalized Lévy density. Smaller jumps (which occur infinitely often) are incorporated into a diffusion process with zero drift and volatility estimated by the second moment of the Lévy density integrated over the range $[-\varepsilon, +\varepsilon]$.

2.2 Bridge Sampling

In [25], bridge sampling for the time-changed Brownian motion representation is introduced, along with stratified sampling and quasi-Monte Carlo to further reduce variance. In [1], bridge sampling for the difference-of-gammas representation is introduced and combined with randomized quasi-Monte Carlo, and bounds (upper and lower) on the discretization error for pricing certain forms of path-dependent options are derived. Here, we present these two bridge sampling approaches, but without the additional variance reduction techniques, to make the exposition easy to follow.

Simulating VG as Gamma Time-Changed Brownian Motion

Input: VG parameters θ, σ, ν; time spacing $\Delta t_1, ..., \Delta t_N$ s.t. $\sum_{i=1}^N \Delta t_i = T$.

Initialization: Set $X_0 = 0$.

Loop from $i = 1$ to N:

1. Generate $\Delta G_i \sim \Gamma(\Delta t_i/\nu, \nu), Z_i \sim \mathcal{N}(0,1)$ independently and independent of past r.v.s.
2. Return $X_{t_i} = X_{t_{i-1}} + \theta \Delta G_i + \sigma\sqrt{\Delta G_i} Z_i$.

Simulating VG as Difference of Gammas

Input: VG parameters θ, σ, ν; time spacing $\Delta t_1, ..., \Delta t_N$ s.t. $\sum_{i=1}^N \Delta t_i = T$.

Initialization: Set $X_0 = 0$.

Loop from $i = 1$ to N:

1. Generate $\Delta\gamma_i^- \sim \Gamma(\Delta t_i/\nu, \nu\mu_-), \Delta\gamma_i^+ \sim \Gamma(\Delta t_i/\nu, \nu\mu_+)$ independently and independent of past r.v.s.
2. Return $X_{t_i} = X_{t_{i-1}} + \Delta\gamma_i^+ - \Delta\gamma_i^-$.

Simulating VG as (Approximate) Compound Poisson Process

Input: VG parameters θ, σ, ν; time spacing $\Delta t_1, ..., \Delta t_N$ s.t. $\sum_{i=1}^N \Delta t_i = T$.

Initialization: Set $X_0 = 0$;

$$\sigma_\varepsilon^2 = \int_{-\varepsilon}^{+\varepsilon} x^2 k(x)dx, \quad \lambda_\varepsilon^+ = \int_{+\varepsilon}^{\infty} k(x)dx, \quad \lambda_\varepsilon^- = \int_{-\infty}^{-\varepsilon} k(x)dx,$$
$$k_\varepsilon^+(x) = k(x)\mathbf{1}_{\{x \geq \varepsilon\}}/\lambda_\varepsilon^+, \quad k_\varepsilon^-(x) = k(x)\mathbf{1}_{\{x \leq -\varepsilon\}}/\lambda_\varepsilon^-.$$

Loop from $i = 1$ to N:

1. Generate number of positive and negative jumps in Δt_i (N_i^+ and N_i^-, respectively) and corresponding size of jumps using Lévy density (everything independent of each other and of past generated samples):

$$N_i^+ \sim \text{Poisson}(\lambda_\varepsilon^+ \Delta t_i), \quad N_i^- \sim \text{Poisson}(\lambda_\varepsilon^- \Delta t_i),$$
$$X_{i,j}^+ \sim \{k_\varepsilon^+\}, j = 1, ..., N_i^+, \quad X_{i,j}^- \sim \{k_\varepsilon^-\}, j = 1, ..., N_i^-,$$
$$Z_i \sim \mathcal{N}(0,1).$$

2. Return $X_{t_i} = X_{t_{i-1}} + Z_i\sigma_\varepsilon\sqrt{\Delta t_i} + \sum_{j=1}^{N_i^+} X_{i,j}^+ + \sum_{j=1}^{N_i^-} X_{i,j}^-$.

Fig. 1. Algorithms for sequentially simulating VG process on $[0, T]$.

Instead of sequential sampling, which progresses chronologically forward in time, an alternative method for simulating asset price paths is to use bridge sampling, which samples the end of the path first, and then "fills in" the rest of the path as needed. This can lead to a more efficient simulation. In bridge sampling, fixed times are chosen, and the value of the process at such an arbitrary fixed time is a random variable with known (conditional) distribution. It is particularly effective in combination with quasi-Monte Carlo methods, because the sampling sequence usually means that the first samples are more critical than the latter ones, leading to a lower "effective dimension" than in the usual sequential sampling. It is in this setting that quasi-Monte Carlo methods show the greatest improvement over traditional MC methods, and gamma bridge sampling coupled with quasi-Monte Carlo is treated in [25] and [1].

The main idea of bridge sampling is that the conditional distribution of a stochastic process X_t at time $t \in (T_1, T_2)$, given X_{T_1} and X_{T_2} can be easily obtained; that is, for $T_1 \le t \le T_2$, one can apply Bayes' rule to get the necessary conditional distribution:

$$P(X_t|X_{T_1}, X_{T_2}) = \frac{P(X_{T_2}|X_{T_1}, X_t)P(X_t|X_{T_1})}{P(X_{T_2}|X_{T_1})}.$$

The following are the most well-known examples.

Poisson process $\{N_t\}$: Conditional on N_{T_1} and N_{T_2},

$$N_t \sim N_{T_1} + \mathbf{bin}(N_{T_2} - N_{T_1}, (t - T_1)/(T_2 - T_1)),$$

where $\mathbf{bin}(n, p)$ denotes the binomial distribution with parameters n and p (mean np and variance $np(1-p)$). Note that the conditional distribution has no dependence on the arrival rate.

Brownian motion $\{B_t^{(\mu,\sigma)}\}$: Conditional on B_{T_1} and B_{T_2},

$$B_t \sim \mathcal{N}(\alpha B_{T_1} + (1-\alpha)B_{T_2}, \alpha(t - T_1)\sigma^2),$$

where $\alpha = (T_2 - t)/(T_2 - T_1)$, and there is dependence on σ, but not on the drift μ.

Gamma process $\{\gamma_t^{(\mu,\nu)}\}$: Conditional on γ_{T_1} and γ_{T_2},

$$\gamma_t \sim \gamma_{T_1} + (\gamma_{T_2} - \gamma_{T_1})Y,$$

where $Y \sim \beta((t - T_1)/\nu, (T_2 - t)/\nu)$, which only depends on the parameter ν, where $\beta(\alpha_1, \alpha_2)$ denotes the beta distribution with mean $\alpha_1/(\alpha_1 + \alpha_2)$, variance $\alpha_1\alpha_2/[(\alpha_1 + \alpha_2)^2(\alpha_1 + \alpha_2 + 1)]$, and density $x^{\alpha_1 - 1}(1-x)^{\alpha_2 - 1}/B(\alpha_1, \alpha_2)$ on $[0, 1]$, where

$$B(x, y) = \int_0^1 t^{x-1}(1-t)^{y-1}dt = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}, \qquad \Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt.$$

Simulating VG via Brownian (Gamma Time-Changed) Bridge

Input: VG parameters θ, σ, ν; number of bridges $N = 2^M (T = t_N)$.

Initialization: Set $X_0 = 0, \gamma_0 = 0$.
Generate $\gamma_{t_N} \sim \Gamma(t_N/\nu, \nu), X_{t_N} \sim \mathcal{N}(\theta\gamma_{t_N}, \sigma^2\gamma_{t_N})$ independently.

Loop from $k = 1$ to M: $n \leftarrow 2^{M-k}$;
Loop from $j = 1$ to 2^{k-1}:
1. $i \leftarrow (2j-1)n$;
2. Generate $Y_i \sim \beta((t_i - t_{i-n})/\nu, (t_{i+n} - t_i)/\nu)$ independent of past r.v.s;
3. $\gamma_{t_i} = \gamma_{t_{i-n}} + [\gamma_{t_{i+n}} - \gamma_{t_{i-n}}]Y_i$;
4. Generate $Z_i \sim \mathcal{N}(0, [\gamma_{t_{i+n}} - \gamma_{t_i}]\sigma^2 Y_i)$ independent of past r.v.s;
5. Return $X_{t_i} = Y_i X_{t_{i+n}} + (1 - Y_i)X_{t_{i-n}} + Z_i$.

Simulating VG via Difference-of-Gammas Bridge

Input: VG parameters θ, σ, ν; number of bridges $N = 2^M (T = t_N)$.

Initialization: Set $\gamma_0^+ = \gamma_0^- = 0$.
Generate $\gamma_{t_N}^+ \sim \Gamma(t_N/\nu, \nu\mu_+), \gamma_{t_N}^- \sim \Gamma(t_N/\nu, \nu\mu_-)$ independently.

Loop from $k = 1$ to M: $n \leftarrow 2^{M-k}$;
Loop from $j = 1$ to 2^{k-1}:
1. $i \leftarrow (2j-1)n$;
2. Generate $Y_i^+, Y_i^- \sim \beta((t_i - t_{i-n})/\nu, (t_{i+n} - t_i)/\nu)$ independently;
3. $\gamma_{t_i}^+ = \gamma_{t_{i-n}}^+ + [\gamma_{t_{i+n}}^+ - \gamma_{t_{i-n}}^+]Y_i^+, \gamma_{t_i}^- = \gamma_{t_{i-n}}^- + [\gamma_{t_{i+n}}^- - \gamma_{t_{i-n}}^+]Y_i^-$;
4. Return $X_{t_i} = \gamma_{t_i}^+ - \gamma_{t_i}^-$.

Fig. 2. Algorithms for simulating VG process on $[0, T]$ via bridge sampling.

As with the gamma distribution, there is no analytical closed form for the c.d.f.; also $\beta(1,1) = U(0,1)$. If either of the parameters is equal to 1, then the inverse transform method can be easily applied. Otherwise, the main methods to generate variates are acceptance–rejection, numerical inversion, or a special algorithm using the following known relationship with the gamma distribution: if $Y_i \sim \Gamma(\alpha_i, \psi)$ independently generated, then $Y_1/(Y_1+Y_2) \sim \beta(\alpha_1, \alpha_2)$. For more information on efficiently generating samples from the gamma and beta distributions, see [7]. MATLAB and other commercial software generally have built-in functions/subroutines to handle these distributions.

The two algorithms given in Figure 2, adapted from [1] and corresponding to the time-changed Brownian motion representation of (2) and the difference-of-gammas representation of (3), respectively, generate VG sample paths of progressively finer resolution by adding bisecting bridge points, so the number

of simulated points on the path should be a power of 2, called the dyadic partition in [1], where it is noted this allows the simulation efficiency to be further improved by using "a fast beta random-variate generator that exploits the symmetry." This partition also makes the methods easier to present in algorithmic form; however, the method is clearly just as easily applicable to general time steps as in the previous sequential versions.

For pricing path-dependent options that depend on the entire continuous sample path—such as continuous Asian, barrier, and lookback options—the advantage of the bridge sampling is that the first samples often capture most of the contribution to the expected payoff that is being estimated. In either case, Richardson extrapolation is one method that can be used to go from the discrete to the continuous case. This is discussed for the difference-of-gammas bridge sampling in [1]; it was also used in [12] in sequential Monte Carlo simulation.

2.3 Variance Reduction

Variance reduction techniques can lead to orders of magnitude of improvement in simulation efficiency, and thus are of practical importance. A simple example where it would be critical is a deep out-of-the-money barrier option, in which the payoff on most sample paths generated by simulation would be zero, so being in the money is essentially a "rare event" in simulation lingo. In this section, we briefly discuss the variance reduction technique of importance sampling, whereby simulation is carried out under a different measure than the one of interest, and then an adjustment is made to the payoff function by way of the Radon–Nikodym derivative (change of measure). In the barrier option example, the resulting measure change would lead to a great increase in the number of generated paths that are in the money. Other useful variance reduction techniques not discussed here that are effective in financial simulations include common random numbers (called "variate recycling" in [16]), conditional Monte Carlo, stratified sampling, and control variates; see [14] for more details. As mentioned earlier, quasi-Monte Carlo combined with gamma bridge sampling for the VG process is described in [25] and [1].

The general form for the Radon–Nikodym derivative of a Lévy process can be found in [26]; see also [28] for gamma processes. For pure-jump Lévy processes, it turns out that sufficient conditions for ensuring equivalence in the measure change are that the corresponding Lévy measures be equivalent plus a constraint relating the drifts and corresponding Lévy measures. For the VG process, under the difference-of-gammas representation, it turns out that the measure change in which the ν parameter is kept the same can be computed based on just the terminal values; that is, the intermediate values on the path are not needed (cf. [29, Propositions 2 and 3 in Chapter 2]). The Radon–Nikodym derivative needed to adjust for going from VG with parameters (θ, σ, ν) to the measure change of VG with parameters (θ', σ', ν) is given by the following exponential twisting:

$$\exp(-t\int_{-\infty}^{+\infty}(k(x)-k'(x))dx)\phi^{+}(\tilde{\gamma}_t^{+})\phi^{-}(-\tilde{\gamma}_t^{-}),$$

$$\phi^{\pm}(x)=e^{2(\mu_{\mp}^{'}/(\sigma^{'})^2-\mu_{\mp}/\sigma^2|x|)},$$

where $\tilde{\gamma}_t^{\pm}$ are the independent processes generated in the difference-of-gammas representation of VG with (VG) parameters $(\theta^{'},\sigma^{'},\nu)$, and $\mu_{\pm}$ is given by (4).

2.4 The Greeks

In many cases, it is possible to estimate sensitivities of derivatives prices with respect to various parameters directly using the same sample path that was used to estimate the price itself, that is, without resorting to resimulation. This was first demonstrated in [9] and [3]. In this section, we discuss only infinitesimal perturbation analysis (IPA), or what is called pathwise differentiation in [16] and [14]. We do not treat the likelihood ratio method (also known as the score function method or measure-valued differentiation); see the previous two referenced books [16, 14], or [8] for further discussion of this approach, which is also related to the Malliavin calculus approach proposed by some researchers (see [14]). An extension of IPA based on conditional Monte Carlo, which can handle discontinuous payoff functions, is treated in [10]. None of the previous references treat the estimation of Greeks explicitly in the VG (or general Lévy process) context.

The basis of IPA is quite simple: one differentiates the sample quantities of interest with respect to the parameter of interest. Specifically, one is usually interested in some payoff function h that may depend on the entire path of $\{S_t\}$. For simplicity of illustration here, we assume the payoff depends only on a single point in time, such as a call option with $h(x)=(x-K)^{+}$, where K is the strike price. Then, the price to be estimated by simulation is given by

$$E[e^{-rT}h(S_T)],$$

for maturity T (European option), whereas the sensitivity to be estimated is given by

$$\frac{dE[e^{-rT}h(S_T)]}{d\chi},$$

where χ is the parameter of interest. The IPA estimator is given by

$$\frac{d(e^{-rT}h(S_T))}{d\chi}=h(S_T)\frac{d(e^{-rT})}{d\chi}+e^{-rT}\left(\frac{\partial h}{\partial\chi}+h'(S_T)\frac{dS_T}{d\chi}\right),$$

so the applicability of the IPA estimator comes down to a question of whether exchanging the operations of expectation (integration) and differentiation (limit) is justified (according to the dominated convergence theorem); that is, whether

$$E\left[\frac{d(e^{-rT}h(S_T))}{d\chi}\right] = \frac{dE[e^{-rT}h(S_T)]}{d\chi},$$

which clearly depends both on the payoff function h and the representation of the VG process $\{X_t\}$, which enters $\{S_t\}$ through (5).

As an example, it can be easily seen from (5) that the current asset price is a scale parameter for a future stock price, so that we have

$$\frac{dS_t}{dS_0} = \frac{S_t}{S_0}.$$

Thus, for example, a call option with the usual payoff $(S_T - K)^+$, which would be estimated in simulation by

$$e^{-rT}(S_T - K)^+,$$

would have its delta estimated by

$$e^{-rT}\frac{S_T}{S_0}\mathbf{1}_{\{S_T>K\}},$$

because $h'(x) = \mathbf{1}_{\{x>K\}}$. However, a digital option with payoff $\mathbf{1}_{\{S_T>K\}}$ would lead to a biased estimator (identically zero). Thus, the IPA estimator for the gamma would be biased. Roughly speaking, if h is almost surely continuous with respect to the parameter of interest, then the IPA estimator will be unbiased. The call payoff function is continuous, with a "kink" at K, which leads to a discontinuity in its first derivative at K, just as for the digital option.

The example above was in some sense the simplest, because S_0 doesn't appear anywhere else in the expression for S_t given by (5) except as a scale factor. Other parameters of interest include time t, the interest rate r, and the VG parameters (θ, σ, ν), and these all make more complicated appearances in S_t, both directly through (5) and indirectly through X_t, where in the latter case the chain rule would be applied. The resulting quantity $dX_t/d\chi$ may depend on the representation of $\{X_t\}$, that is, in how the stochastic process $\{X_t\}$ is constructed, and we saw there are at least three different ways a VG process can be generated. Thus, for the simple call option, we have the following IPA estimators (for any exercise point t).

$$\begin{aligned}
\frac{dS_t}{dr} &= tS_t,\\
\frac{dS_t}{dt} &= S_t\left(r + \omega + \frac{dX_t}{dt}\right),\\
\frac{dS_t}{d\theta} &= S_t\left(t\frac{d\omega}{d\theta} + \frac{dX_t}{d\theta}\right),\\
\frac{dS_t}{d\sigma} &= S_t\left(t\frac{d\omega}{d\sigma} + \frac{dX_t}{d\sigma}\right),\\
\frac{dS_t}{d\nu} &= S_t\left(t\frac{d\omega}{d\nu} + \frac{dX_t}{d\nu}\right),
\end{aligned}$$

with (from differentiating (6))

$$\frac{d\omega}{d\theta} = -1/(1 - \theta\nu - \sigma^2\nu/2),$$
$$\frac{d\omega}{d\sigma} = -\sigma/(1 - \theta\nu - \sigma^2\nu/2),$$
$$\frac{d\omega}{d\nu} = -[(\theta + \sigma^2/2)/(1 - \theta\nu - \sigma^2\nu/2) + \omega]/\nu,$$

but the $dX_t/d\chi$ term in the IPA estimators for $dS_t/d\chi$ above undetermined without specifying a particular representation of $\{X_t\}$. For example, if we chose the parameter of interest χ to be the θ parameter defining the VG process, then using the time-changed Brownian motion representation given by (2) leads to a very simple

$$\frac{dX_t}{d\theta} = \gamma_t^{(\nu)},$$

whereas for the difference-of-gammas representation given by (3),

$$\frac{dX_t}{d\theta} = \frac{d\gamma_t^{(\mu_+,\nu_+)}}{d\theta} - \frac{d\gamma_t^{(\mu_-,\nu_-)}}{d\theta},$$

which is more complicated to compute. In this case, it is likely that both lead to unbiased IPA estimators, but the resulting estimators would have very different variance properties.

3 Conclusions

In the context of Monte Carlo simulation, there is a large body of work on variance reduction techniques and gradient estimation techniques for the usual Gaussian/diffusion setting in finance, but with the exception of a few recent results such as [25] and [1], the more general Lévy process setting is relatively untouched. After reviewing the three main methods for simulating a VG process and presenting two recently developed bridge sampling approaches, we just scratched the surface here in presenting the importance sampling change of measure and introducing IPA sensitivity estimators for the Greeks in the VG setting. In the latter case, the choice of representation plays a key role in determining the applicability of IPA; see [13] and [10] for more on this theme.

Acknowledgments

This work was supported in part by the National Science Foundation under Grant DMI-0323220, and by the Air Force Office of Scientific Research under Grant FA95500410210. I thank Eugene Seneta, Philip Protter, Andy Hall, Scott Nestler, Qing Xia, and Huiju Zhang for their comments and corrections.

References

1. A.N. Avramidis and P. L'Ecuyer. Efficient Monte Carlo and quasi-Monte Carlo option pricing under the Variance-Gamma model. *Management Science*, 52:1930–1944, 2006.
2. P. Boyle, M. Broadie, and P. Glasserman. Monte Carlo methods for security pricing. *Journal of Economic Dynamics and Control*, 21:1267–1321, 1997.
3. M. Broadie and P. Glasserman. Estimating security price derivatives using simulation. *Management Science*, 42:269–285, 1996.
4. P. Carr, H. Geman, D.B. Madan, and M. Yor. The fine structure of asset returns: An empirical investigation. *Journal of Business*, 75:305–332, 2002.
5. P. Carr, H. Geman, D.B. Madan, and M. Yor. Stochastic volatility for Lévy processes. *Mathematical Finance*, 2:87–106, 2003.
6. P. Carr and D. Madan. Option valuation using the fast Fourier transform. *Journal of Computational Finance*, 2:61–73, 1999.
7. G.S. Fishman. *Monte Carlo Methods: Concepts, Algorithms, and Applications*. Springer, 1996.
8. M.C. Fu. Stochastic gradient estimation. Chapter 19 in *Handbooks in Operations Research and Management Science: Simulation*, eds. S.G. Henderson and B.L. Nelson, Elsevier, 2006.
9. M.C. Fu and J.Q. Hu, Sensitivity analysis for Monte Carlo simulation of option pricing. *Probability in the Engineering and Informational Sciences*, 9:417–446, 1995.
10. M.C. Fu and J.Q. Hu. *Conditional Monte Carlo: Gradient Estimation and Optimization Applications*. Kluwer Academic, 1997.
11. M.C. Fu, S.B. Laprise, D.B. Madan, Y. Su, and R. Wu. Pricing American options: A comparison of Monte Carlo simulation approaches. *Journal of Computational Finance*, 4:39–88, 2001.
12. M.C. Fu, D.B. Madan and T. Wang. Pricing continuous Asian options: A comparison of Monte Carlo and Laplace transform inversion methods. *Journal of Computational Finance*, 2:49–74, 1999.
13. P. Glasserman. *Gradient Estimation Via Perturbation Analysis*. Kluwer Academic, 1991.
14. P. Glasserman. *Monte Carlo Methods in Financial Engineering*. Springer, 2003.
15. A. Hirsa and D.B. Madan. Pricing American options under variance gamma. *Journal of Computational Finance*, 7:63–80, 2004.
16. P. Jäckel. *Monte Carlo Methods in Finance*. Wiley, 2002.
17. S. Laprise. *Stochastic Dynamic Programming: Monte Carlo Simulation and Applications to Finance*. Ph.D. dissertation, Department of Mathematics, University of Maryland, 2002.
18. D.B. Madan, P. Carr, and E.C. Chang. The variance gamma process and option pricing. *European Finance Review*, 2:79–105, 1998.
19. D.B. Madan and F. Milne. Option pricing with V.G. martingale components. *Mathematical Finance*, 1:39–55, 1991.
20. D.B. Madan and E. Seneta. The Variance-Gamma (V.G.) model for share market returns. *Journal of Business*, 63:511–524, 1990.
21. D.B. Madan and M. Yor. Representing the CGMY and Meixner processes as time changed Brownian motions. manuscript, 2006.
22. D.L. McLeish. A robust alternative to the normal distribution. *Canadian Journal of Statistics*, 10:89–102, 1982.

23. D.L. McLeish, *Monte Carlo Simulation and Finance*. Wiley, 2005.
24. P. Protter. *Stochastic Integration and Differential Equations*, 2nd edition. Springer-Verlag, 2005.
25. C. Ribeiro, and N. Webber. Valuing path-dependent options in the Variance-Gamma model by Monte Carlo with a gamma bridge. *Journal of Computational Finance*, 7:81–100, 2004.
26. K. Sato. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, 1999.
27. E. Seneta. The early years of the Variance-Gamma process. In this volume, 2007.
28. M. Yor. Some remarkable properties of gamma processes. In this volume, 2007.
29. B. Zhang. *A New Lévy Based Short-Rate Model for the Fixed Income Market and Its Estimation with Particle Filter*. Ph.D. dissertation (directed by Dilip B. Madan), Department of Mathematics, University of Maryland, 2006.

Appendix: Review of Basic Definitions

Here we provide the elementary characterizations of the Wiener, Poisson, and gamma processes. (For further properties on the gamma process, see Marc Yor's contribution [28] in this volume.) Each of these is a Lévy process, that is, a process of independent and stationary increments. We consider only homogeneous Lévy processes, in which case the increments are i.i.d. Thus, one way to differentiate among them is to specify the distribution of the increments. For standard Brownian motion $\{W_t\}$, also known as the Wiener process, the increments are normally distributed with zero mean and variance equal to the size of the increment; that is, for any t,

$$W_{t+\delta} - W_t \sim \mathcal{N}(0, \delta),$$

where $\mathcal{N}(\mu, \sigma^2)$ denotes the normal (Gaussian) distribution with mean μ and variance σ^2. Similarly, the gamma process $\{\gamma_t\}$ has gamma distributed increments; that is, for any t, $\gamma_{t+\delta} - \gamma_t \sim \Gamma(\delta, 1)$, where $\Gamma(\alpha, \beta)$ denotes the gamma distribution with mean $\alpha\beta$ and variance $\alpha\beta^2$. Unlike the Wiener process, the gamma process is discontinuous, and like the Poisson process, it is nondecreasing (since the gamma distribution has support on the positive real line). For the two-parameter gamma process $\{\gamma_t^{(\mu,\nu)}\}$, increments are gamma distributed with the mean μ and variance ν both multiplied by the size of the increment; that is, for any t,

$$\gamma_{t+\delta}^{(\mu,\nu)} - \gamma_t^{(\mu,\nu)} \sim \Gamma(\delta\mu^2/\nu, \nu/\mu).$$

The nondecreasing property makes the gamma distribution suitable as a subordinator (time change). Although the difference-of-gammas representation of VG uses the two-parameter gamma distribution, the time-changed representation of VG uses the one-parameter version (taking $\mu = 1$ in the two-parameter version), denoted here $\{\gamma_t^{(\nu)}\}$, where for any t,

$$\gamma_{t+\delta}^{(\nu)} - \gamma_t^{(\nu)} \sim \Gamma(\delta/\nu, \nu),$$

with a mean equal to the size of the increment, making it most suitable for a time change. This is also the version of the gamma process that is used in the following paper by Marc Yor [28] in this volume.

By the Lévy–Khintchine theorem (cf. [24, 26]), a Lévy process $\{X_t\}$, with finite-variation jump component can be specified by its unique characteristic function

$$\phi_{X_t}(u) = E[e^{iuX_t}] = \exp\left(iuat - u^2b^2t/2 + t\int_{-\infty}^{+\infty}(e^{iux} - 1)k(x)dx\right),$$

where $k(x)$ is the Lévy density, a is the drift rate, and b is the diffusion coefficient. The three components correspond to a deterministic drift, a continuous Wiener process, and a pure-jump process, where intuitively the Lévy density is a measure on the arrival rate of different jump sizes. Note that because a Lévy process is infinitely divisible, a single time point (e.g., $t = 1$) suffices to characterize the process, in terms of its marginal distribution at any time. In the VG process, the drift rate (a) and diffusion coefficient (b) are both zero, and the process is characterized by three parameters (θ, σ, ν) that appear in the Lévy density according to (1). The symmetric version of the VG process given by ($\theta = 0$),

$$X_t = \sigma W_{\gamma_t^{(\nu)}},$$

has mean 0, variance $\sigma^2 t$, skewness 0, and kurtosis $3(1+\nu)$, so the parameter ν gives the excess kurtosis over Brownian motion (which has a kurtosis of 3). In the symmetric version, the difference-of-gammas representation involves two (independent) gamma processes with the same distribution. In the asymmetric VG process, the parameter θ controls skewness, with a negative value giving a fatter (heavier) left tail. Brownian motion can be obtained as a limiting case of the VG process, because $\lim_{\nu\to 0}\gamma_t^{(\nu)} = t$.