Centre for Computational Finance and Economic Agents

WP004-06

Working Paper Series

Kyriakos Chourdakis

Lévy processes driven by stochastic volatility

March 2006

CCFEA

www.essex.ac.uk/ccfea

Lévy processes driven by stochastic volatility

Kyriakos Chourdakis

CCFEA, University of Essex, Colchester CO4 3SQ, United Kingdom
e-mail: kchour@essex.ac.uk
URL address: http://www.theponytail.net/

In this paper we extend option pricing under Lévy dynamics, by assuming that the volatility of the Lévy process is stochastic. We therefore develop the analog of the standard stochastic volatility models, when the underlying process is not a standard (unit variance) Brownian motion, but rather a standardized Lévy process. We present a methodology that allows one to compute option prices, under virtually any set of diffusive dynamics for the parameters of the volatility process. First, we use 'local consistency' arguments to approximate the volatility process with a finite, but sufficiently dense Markov chain; we then use this regime switching approximation to efficiently compute option prices using Fourier inversion. A detailed example, based on a generalization of the popular stochastic volatility model of Heston (1993, RFS), is used to illustrate the implementation of the algorithms.

1 Introduction

Over the last thirty years, a vast number of pricing models have been proposed as an alternative to the classic Black and Scholes (1973) and Merton (1973) approach (henceforth BMS), driven by the well documented biases of the BMS formula. Moving into the implied volatility space, observed option prices exhibit well defined patterns across moneyness and maturity. In some markets, such as the currency markets, a convex implied volatility smile is observed across moneyness, while in others, such as the stock index options markets, the pattern resembles more a downward sloping skew. Across maturities, the term structure of implied volatilities can be downward or upward sloping, similar to the term structure of interest rates. In order to accommodate for these biases, some of the pivotal BMS assumptions have been relaxed over the years.

There are two key assumptions that need to be made, in order to price derivatives in the BMS world: that returns are subject to a single source of uncertainty, and that asset prices follow continuous sample paths. Under these two assumptions, a continuously rebalanced portfolio can be used to perfectly hedge an options position, thus determining a unique price for the option. Relaxing the assumption of a unique source of uncertainty leads to the *stochastic volatility* family of models, where the volatility parameter follows a separate diffusion, for example the processes in Hull and White (1987), and Heston (1993). Relaxing the assumption of continuous sample paths, leads to the *jump diffusion* models, where infrequent jumps are added to the standard BMS diffusion, in the spirit of Merton (1976). In both cases the market is incomplete, and option prices are not determined in a unique way. A multitude of risk-adjusted probability measures exists, and the information conveyed in the price process is not sufficient to distinguish between them. State-of-the-art pricing models combine the two approaches, producing models that incorporate both stochastic volatility and jumps (the most common being the ones proposed in Bates (2000) and Duffie, Pan, and Singleton (2000)). A growing literature is dedicated in reconciling the 'objective' and 'risk-adjusted' parameters, and in identifying the various prices of risk (see

for example Bates (1996, 1998), Andersen, Benzoni, and Lund (1998), Pan (2002), and Eraker, Johannes, and Polson (2003), *inter alia*).

It has been recently recognized that, although the presence of jumps is pivotal, allowing for continuous sample paths is not required to correctly price option contracts (see for example Geman, Madan, and Yor (2001) and Ané and Geman (2000)). The family of pure jump models, where the underlying asset is subject to (perhaps infinitely dense) jumps, suffices to capture the stylized facts of asset returns, such as the negative skewness and excess kurtosis. The theory of Lévy processes links such specifications to the characteristic functions of their respective densities, and provides the probabilistic framework that offers option prices. Broadly speaking, a Lévy log-price will exhibit a sequence of independent and identically distributed increments, and offers a more general framework than the lognormal assumption of BMS. The geometric Brownian motion, which is the cornerstone of the BMS analysis, is therefore relegated to a special case.

More importantly, assuming a jump process breaks the link between the statistical properties of the underlying asset and the risk neutral properties, since a replicating hedge cannot be constructed. Therefore, the 'risk-adjusted' parameters under Lévy dynamics can be separated from the 'objective' set of parameters of the underlying process. Thus, the analysis of the underlying price process in the 'risk-adjusted' world is completely disentangled from the corresponding analysis in the 'real' world. Therefore, derivative contracts can be priced in isolation from their underlying assets.

Although Lévy processes are very versatile, from a distribution point of view, they fail to capture the time series properties of asset returns. In fact, since the densities assumed are i.i.d., a Lévy process cannot generate the time-varying higher moments, or the volatility clusters that are typical of financial time series. This problem has been recognized in the literature, and has been addressed in two ways: Konikov and Madan (2001) consider a two-state Markov chain that drives the Lévy process, while more recently Carr, Geman, Madan, and Yor (2003) (henceforth CGMY) introduce a stochastic clock that induces time-varying higher moments.[1] Carr and Wu

[1] Lévy processes are the result of subordinating a Brownian motion to a stochastic clock,

(2004b) generalize the CGMY approach, by introducing correlations between the asset prices and the activity rates.

The starting point of this paper is a standardized (unit variance over a unit time interval) Lévy process L^1, and attempts to 'create' the necessary volatility variation using the process $L_t = L_0 + \sum_{0<s<t} \sqrt{v_s} \cdot \Delta L_s^1$, where the variance process v is diffusive. This (more traditional) approach, follows the existing literature on stochastic volatility[2] more closely than the subordinating approach of CGMY. Given a parameter set that describes the log price dynamics, we turn into computing the prices of European options.

In order to retrieve the option prices, we discretize the volatility state-space, following the approach of Chourdakis (2004). As an example, we use the popular square-root diffusion of Heston (1993), to show in detail how this methodology can be implemented.

For a set of parameters, the implementation consists of three steps: First, we construct a grid that reflects the possible volatility movements over the life of the option. Using 'local consistency' arguments, we approximate the volatility diffusion, using a Markov chain that lives on this particular grid. Finally, we show how the characteristic function of the approximating model can be computed in closed form.

The corresponding prices can now be computed using standard Fourier inversion, using an FFT library.

Based on a set of observed contracts, we calibrate a model that exhibits stochastic Variance-Gamma dynamics. We find that both diffusive and pure-jump components are necessary, in order to capture the dynamics of the smile. In particular, we find that Brownian motion are responsible for the leverage effect, while the discontinuous component is responsible for the tail behavior.

---

which is considered to increase in an i.i.d. fashion. CGMY subordinate some of the most common Lévy processes once more, this time to the cumulative square-root diffusion of Cox, Ingersoll, and Ross (1985).

[2] The 'traditional' stochastic volatility approach would model the noise element as $\int_0^t \sqrt{v_s} \cdot dB_s$, with B a standard Brownian motion.

2 The Lévy component

Consider a Lévy process L^1 with unit variance, having a characteristic triplet $(\alpha, \beta^2, \Gamma)$. This implies that the characteristic exponent (see Bertoin (1996) for details) is given by the Lévy-Khintchine formula

$$\Psi_{L^1}(u) = \mathtt{i}\alpha u - \frac{1}{2}\beta^2 u^2 + \int (1 - \exp \mathtt{i}us + \mathtt{i}us 1_{|s|<1})\Gamma(s)ds$$

The three members of the triplet (α, β, Γ) represent the drift, the volatility of the diffusive part, and the jump structure, respectively. In particular, the function $\Gamma(s)$ gives the arrival intensity of jumps of size equal to s. Regularity demands that Γ integrates $1 \wedge |s|^2$. The characteristic function of L_t^1 is given by

$$\Phi_t(u) = E_0 \exp \mathtt{i}L_t^1 u = \exp t\Psi_{L^1}(u)$$

We can decompose the Lévy process in a purely continuous, and a purely discontinuous part,[3] as follows

$$L_t^1 = \beta \cdot W_t + \sqrt{1 - \beta^2} \cdot J_t^1 \tag{2.1}$$

We assume that J^1 is normalized to exhibit unit variance, and that W is a standard Brownian motion. Then, the above decomposition ensures that L^1 is a standardized Lévy process.

Explicit knowledge of the characteristic function of a Lévy process, allows us to compute the corresponding characteristics of the standardized process. We will now review some widely used Lévy processes, and show how they can be standardized, in order to play the role of J^1 in equation (2.1).

2.1 The Normal Inverse Gaussian process

The Normal Inverse Gaussian (NIG) process of Barndorff-Nielsen (1998) is based on the inverse Gaussian process, which is compiled of the times, T_t, when a Brownian motion with drift, $B_t^* = a^* \cdot t + W_t^*$, first crosses the level t.

[3] See for example Carr and Wu (2004a) for a similar decomposition.

This process will define the relationship between 'trading' time and calendar time. A rapidly increasing B^* will imply a faster trading clock, while a rapidly decreasing B^* will imply that trading intensity slows down. The NIG process is constructed by subordinating a separate Brownian motion with drift, $B_t = b^* \cdot t + d \cdot W_t$, to the times T_t, that is to say $J_t = B_{T_t}$. The characteristic exponent of the NIG process is given by

$$\Psi_{\text{NIG}}(u) = -d\left(\sqrt{a^2 - (b + \mathtt{i}u)^2} - \sqrt{a^2 - b^2}\right)$$

Simple differentiation of the characteristic function will produce the moments of the process J. In particular, the variance is equal to $\frac{a^2 d}{(a^2-b^2)^{3/2}}$; it is therefore straightforward to verify that imposing the constraint

$$d \to \frac{(a^2 - b^2)^{3/2}}{a^2}$$

will ensure that this pure jump process exhibits unit variance, and can play the role of J^1 in (2.1). It is also of interest to examine how the higher moments of this standardized Lévy process behave, after the restriction has been imposed. As pointed out in Konikov and Madan (2001), for all Lévy processes, the skewness has an inverse relationship with the square root of the maturity, while the excess kurtosis is inversely related to the maturity. Table 1 gives the relevant formulas. We can observe that the skewness of the standardized distribution is determined by the parameter b, while a controls the kurtosis.

2.2 The Variance-Gamma process

The Variance-Gamma (VG) process is defined as a Brownian motion with drift θ and volatility σ, $B_t = b^* \cdot t + d \cdot W_t$, evaluated at a random Gamma time, G_t. The Gamma process G has (normalized) mean rate equal to one, and variance rate ν, and is assumed independent of the Brownian motion W. The VG process is thus defined as $J_t = B_{G_t}$. Madan, Carr, and Chang

(1998) show that the characteristic exponent of J is given by

$$\Psi_{\text{VG}}(u) = -\frac{1}{\nu}\log\left(1 - \mathtt{i}u\theta\nu + \frac{1}{2}u^2\sigma^2\nu\right)$$

Broadly speaking, out of the three VG parameters, σ controls for the volatility, θ is responsible for skewness and ν generates kurtosis. The standardization restriction for the VG process, and the corresponding skewness and kurtosis are given in table 1.

2.3 The Carr-Geman-Madan-Yor process

The CGMY process generalizes the VG process, introducing a fourth parameter that controls the finiteness of the activity rate and the variation of the sample paths. In that sense, the CGMY process nests the Poisson, the NIG, and the VG processes.

The characteristic function and the corresponding risk adjustments are given in table 1.

2.4 Processes with a continuous component

In this subsection we discuss the characteristic exponent of processes of the form (2.1), where both a continuous and a purely discontinuous component is present. Assuming that the characteristic exponent of the (standardized) pure jump process is $\Psi_{J^1}(u)$, the characteristic function can be written as

$$E\exp \mathtt{i}uL^1 = \exp\left(-\frac{1}{2}\beta^2u^2 + \Psi_{J^1}\left(\sqrt{1-\beta^2}u\right)\right)$$

Thus, the characteristic exponent of the standardized Lévy process will be

$$\Psi_{L^1}(u) = -\frac{1}{2}\beta^2u^2 + \Psi_{J^1}\left(\sqrt{1-\beta^2}u\right)$$

It is also straightforward to verify that the characteristic exponent of $L =$

$\sqrt{V}L^1$, which will have variance equal to V, will be equal to

$$\Psi_L(u|V) = -\frac{1}{2}\beta^2 V^2 u^2 + \Psi_{J^1}\left(\sqrt{V}\sqrt{1-\beta^2}u\right) \tag{2.2}$$

3 The volatility process

Having established the standardized Lévy process, L, we now turn to the volatility process. We will assume a general diffusive form, namely that the variance of the log-asset price is given by the following expression

$$v_t = v_0 + \int_0^t \mu(v_s)\cdot ds + \int_0^t \sigma(v_s)\cdot dW_s^v \tag{3.1}$$

At this stage, the only assumption we need to impose on the functional forms of μ and σ, is that they are bounded and continuous functions. The boundedness assumption can be further relaxed, if the process v spends 'negligible' time outside a compact set $\mathsf{S} \subset \mathbb{R}$. This, for instance, can be the case if the volatility process is stationary, which is a standard assumption in financial literature. Then, it suffices to assume that μ and σ are bounded over S.

Standard processes for the variance include the square-root diffusion of Feller (1971) and Cox et al. (1985), and the Ornstein-Uhlenbeck process for the logarithm of the variance, of Melino and Turnbull (1990). The numerical examples of this paper will concentrate on the square-root diffusion.

We also allow the two diffusions W and W^v to be correlated, with correlation coefficient ρ. This correlation will represent the well documented *leverage effect*. Unfortunately, as pointed out in Carr and Wu (2004b), diffusive components can only be correlated with other diffusive components, which forces us to assume that W^v is independent of the jump component. Therefore, although a stochastic volatility pure jump process can be accommodated in our framework, we can only consider leverage-neutral processes of this kind.

NIG process
$\Psi(u) = -d\left(\sqrt{a^2-(b+\mathtt{i}u)^2}-\sqrt{a^2-b^2}\right)$
$d \to \dfrac{(a^2-b^2)^{3/2}}{a^2}$
$Skew = \dfrac{3b}{a^2-b^2}$
$XKurt = 3\dfrac{a^2+4b^2}{(a^2-b^2)^2}$
VG process
$\Psi(u) = -\dfrac{1}{\nu}\log\left(1-\mathtt{i}u\theta\nu+\dfrac{1}{2}u^2\sigma^2\nu\right)$
$\nu \to \dfrac{1-\sigma^2}{\theta^2}$
$Skew = \dfrac{(1-\sigma^2)(2+\sigma^2)}{\theta}$
$XKurt = 3\dfrac{(1-\sigma^2)(2-\sigma^4)}{\theta^2}$
CGMY process
$\Psi(u) = C\Gamma(-Y)\left((M-\mathtt{i}u)^Y+(G-\mathtt{i}u)^Y-M^Y-G^Y\right)$
$C \to \dfrac{G^2M^2}{\Gamma(-Y)(Y-1)Y}\cdot\dfrac{1}{G^YM^2+G^2M^Y}$
$Skew = \dfrac{Y-2}{GM}\cdot\dfrac{G^YM^3+G^3M^Y}{G^YM^2+G^2M^Y}$
$XKurt = \dfrac{Y-2}{GM}\cdot\dfrac{Y-3}{GM}\cdot\dfrac{G^YM^4+G^4M^Y}{G^YM^2+G^2M^Y}$

Tab. 1: Standardization of three Lévy process. The table presents the characteristic exponent of the NIG, VG and CGMY processes, and the restriction that has to be imposed on the parameters to ensure that the process exhibits unit variance. The table also gives a measure or the corresponding skewness and excess kurtosis of the Lévy process under this restriction. The actual skewness over a time interval τ is equal to $Skew/\sqrt{\tau}$, and the actual kurtosis is equal to $3+XKurt/\tau$.

4 The asset price process

This section combines the definitions of the previous parts, and constructs a stochastic volatility model for a Lévy process. We assume that the log-price of the underlying asset, say y_t follows a process of the form

$$y_t = y_0 + \int_0^t \mu_y(v_s) \cdot dt + \sum_{s<t} \sqrt{v_s} \cdot \Delta L_s^1 \tag{4.1}$$

Under the risk-adjusted measure, the functional form of μ_y has to be chosen, as to ensure that the discounted price process forms a martingale. In particular, we demand

$$\exp y_0 = \exp -rt \cdot E_0 \exp y_t \tag{4.2}$$

Conditioning on the path of the variance process $(v_s)_0^t$, we can rewrite the expectation as

$$E_0 \left[\exp \left(y_0 + \int_0^t \mu_y(v_s) \cdot dt \right) \cdot E_0 \left[\exp \left(\sum_{s<t} \sqrt{v_s} \cdot \Delta L_s^1 \right) \middle| (v_s)_0^t \right] \right]$$

Using now that fact that the increments ΔL_s^1 are independent, and using the definition of the characteristic exponent, we can express the innermost expectation as

$$\prod_{s<t} E_0 \left[\exp \left(\sqrt{v_s} \cdot \Delta L_s^1 \right) \middle| v_s \right] = \prod_{s<t} \exp \Psi_{L^1}(-\mathtt{i}\sqrt{v_s}) = \exp \int_0^t \Psi_{L^1}(-\mathtt{i}\sqrt{v_s}) ds$$

It is straightforward now to recognize that setting

$$\mu_y(v_s) = r - \Psi_{L^1}(-\mathtt{i}\sqrt{v_s}) = r - \Psi_L(-\mathtt{i}|v_s) \tag{4.3}$$

will satisfy (4.2), and render the discounted process a martingale.

To illustrate the above relationship, we can use the trivial example of the Brownian motion, used in the BMS paradigm. In this case, the process L^1 is just a standard Brownian motion, with characteristic exponent $\Psi_{L^1}(u) =$

$-\frac{1}{2}u^2$. In addition, the variance process is constant, $v_s = \sigma^2$. Then, from equation (4.3), the required drift of the risk-adjusted process is $\mu_y = r - \frac{1}{2}\sigma^2$, which is the familiar log-normality adjustment of the BMS formula.

As another example, which is more related to the current paper, we can consider the VG process under stochastic volatility. Then, following equation (4.3), we can identify that the risk-adjusted drift should be (where of course $\nu = \frac{1-\sigma^2}{\theta^2}$)

$$\mu_y(v_s) = r + \frac{1}{\nu}\log\left(1 - \sqrt{v_s}\theta\nu - \frac{1}{2}v_s^2\sigma^2\nu\right)$$

5 Approximating the volatility diffusion

As explained in detail in Kushner (1990) and Dupuis and Kushner (2001), a diffusion of the form (3.1) can be approximated arbitrarily well using a carefully selected Markov chain, that lives on a finite grid. The Markov chain has to satisfy the *local consistency* requirements, namely that it exhibits the same instantaneous drift and volatility as the diffusion in question.[4] In essence, the log-price process is approximated by a sufficiently dense regime switching model; Chourdakis (2004) shows how closed form characteristic functions for such regime switching models can be used to approximate option prices, and how correlations between the state process and the prices process can be introduced.

5.1 The local consistency concept

In order to implement the approximation method, we will need a variance grid $\Upsilon^h = \{V_0^h, \cdots, V_{N_h}^h\}$, indexed by $h > 0$. Assume that as $h \to 0$: (*i*) $|V_{j+1}^h - V_j^h| \to 0$ for all $j = 1, \cdots, N_h$, or in other words that the grid becomes progressively denser; (*ii*) $N_h \to \infty$, or that the grid becomes larger, and (*iii*) that the grid tends to cover the support of the diffusion (3.1). We

[4] Kushner and DiMasi (1978) generalize this approach for jump diffusions. Therefore, the same technique could be in principle employed in the case where the volatility was driven by a Lévy process as well. Of course, in this case the structure of the approximating chain will be more complicated, and, in particular, moving to non-neighboring variance states has to be allowed.

will denote the approximating chain with $v_t^h \in \Upsilon$, and the corresponding rate matrix with $\mathbf{Q}^h = [q_{i,j}^h]$.

Assume for a moment that at time t, the values of the variance process and the approximating chain coincide, $v_t = v_t^h$. It can be shown, for example in Aït-Sahalia (2002), that allowing switches to the neighboring states V_{j-1}^h and V_{j+1}^h is sufficient to ensure local consistency.

The local consistency conditions stipulate that, for $\delta > 0$

$$\begin{aligned} E_t^h(v_{t+\delta} - v_t) &= \mu(v_t)\delta + o(\delta) \\ E_t^h(v_{t+\delta} - v_t)^2 &= \sigma^2(v_t)\delta + o(\delta) \end{aligned} \tag{5.1}$$

The expectations operator E^h is applied with respect to the rate matrix $\mathbf{Q}^h$. Kushner shows that given local consistency, the Markov chains indexed by h will converge weakly to the diffusion 3.1, as $h \to 0$.

Given a grid, equations (5.1) naturally lead to a system that can be solved to retrieve the approximating rate matrix. For convenience we drop the index h from the rest of the exposition, keeping in mind though that the results will only hold asymptotically. Assume that at time t, the variance is equal to V_j. Over a time interval δ, there are three possibilities: we can remain at V_j, move up by d_U to $V_{j+1} = V_j + d_U$, or move down by d_D to $V_{j-1} = V_j - d_D$. The local consistency conditions can therefore be restated as

$$\begin{aligned} -q_{j-1,j}d_D\delta + q_{j+1,j}d_U\delta &= \mu(V_j)\delta + o(\delta) \\ q_{j-1,j}d_D^2\delta + q_{j+1,j}d_U^2\delta &= \sigma^2(V_j)\delta + o(\delta) \end{aligned}$$

The above relationships can lead to the approximations schemes. Passing to the limit $\delta \to 0$, we can solve the above system to retrieve the rate matrix elements, for all $j = 2, \cdots, N-1$, as

$$\begin{aligned} q_{j-1,j} &= \tfrac{1}{d_D(d_U+d_D)}\left(\sigma^2(V_j) - d_U\mu(V_j)\right) \\ q_{j+1,j} &= \tfrac{1}{d_U(d_U+d_D)}\left(\sigma^2(V_j) + d_D\mu(V_j)\right) \\ q_{j,j} &= -q_{j-1,j} - q_{j+1,j} \end{aligned} \tag{5.2}$$

For $\mathbf{Q}$ to be a valid rate matrix, we have to ensure that the off-diagonal

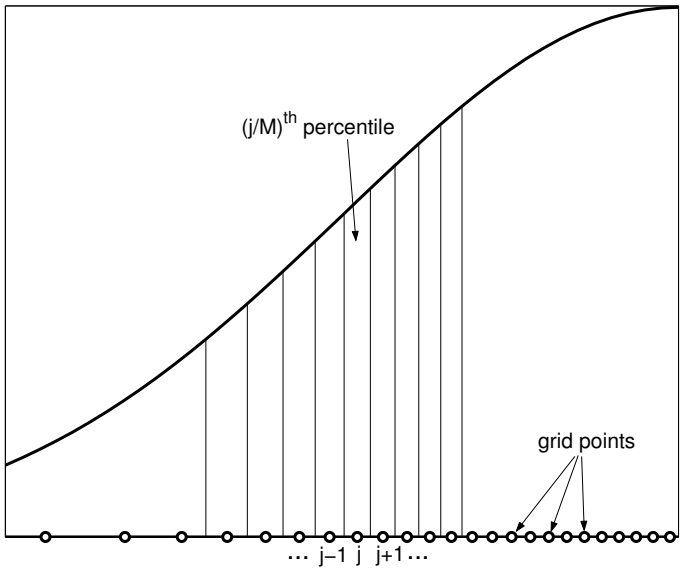

Fig. 1: Grid construction. Setting the grid points at the $\left(\frac{1}{2M} + (j-1)\frac{1}{M}\right)$-percentile of the unconditional distribution of the diffusion.

elements are non-negative. For this reason, when the equations (5.2) above produce invalid elements, one can replace them with a scheme that always remains valid

$$\begin{aligned} q_{j-1,j} &= \tfrac{1}{d_D(d_U+d_D)}\left(\sigma^2(V_j) + (d_D+d_U)\mu^-(V_j)\right) \\ q_{j+1,j} &= \tfrac{1}{d_U(d_U+d_D)}\left(\sigma^2(V_j) + (d_D+d_U)\mu^+(V_j)\right) \\ q_{j,j} &= -q_{j-1,j} - q_{j+1,j} \end{aligned} \tag{5.3}$$

The superscripts $\mu^{\pm}$ denote the positive and negative parts of μ. Both schemes match the instantaneous drift, but only the first one matches the instantaneous volatility too.[5]

We now turn in issues related to the grid construction.

5.2 The grid

The grid of the Markov chain has to be carefully chosen, keeping in mind that: (*i*) it has to be sufficiently dense, in order for the local consistency conditions to apply, and (*ii*) it has to be sufficiently wide, in order for the effect of truncating the support of the diffusion to be negligible. In this paper, we will use the unconditional density to create the grid: We can construct an M-point grid by setting V_j equal to the $\left(\frac{1}{2M} + (j-1)\frac{1}{M}\right)$-percentile of the unconditional distribution. Figure 1 illustrates this construction. This approach offers two major benefits: (*i*) the grid extends and covers the diffusion support as M increases, and (*ii*) the grid is tighter where it matters most, that is to say, where the diffusion is more likely to take values in the future. Of course other constructions are also possible; for example, if more extreme values are required, the symmetric regularized Beta function $B_{SR}(\cdot, \eta)$ can be used to 'spread out' the percentiles generated above.[6] Figure 2 gives examples of $B_{SR}(\cdot, \eta)$ for different values of the parameter η.

It is assumed that the grid has been selected in such a way, that the behavior of the chain on the boundaries is not important. Generally speaking, this will be the case if the probability of reaching the boundaries (in the relevant time frame) is negligible. These probabilities can be actually examined, since the time-t transition density of the Markov chain is given by the matrix exponential $\mathbf{P}(t) = \exp t\mathbf{Q}$. If this is the case, the boundaries can be made absorbent or reflective, without any significant impact on the approximation procedure.

[5] It is straightforward to verify that the instantaneous volatility of approximations (5.3) is $\sigma^2(V_j) + (d_D + d_U)|\mu(V_j)|$. This will asymptotically converge to the diffusion volatility σ as the grid becomes finer.

[6] The symmetric regularized Beta function maps the interval $[0,1] \to [0,1]$ in a sigmoid way, maintaining $\frac{1}{2} \to \frac{1}{2}$. It is given by $B_{SR}(\cdot, \eta) = \frac{B(\cdot,\eta,\eta)}{B(\eta,\eta)}$, where the numerator and denominator are the incomplete and complete Beta functions, respectively. It can also be computed as the cumulative density function of the Beta distribution, with both shape parameters equal to η. Setting $\eta = 1$ results in a linear (in probability) grid, while values of $\eta > 1$ spread out the grid points towards more extreme values.

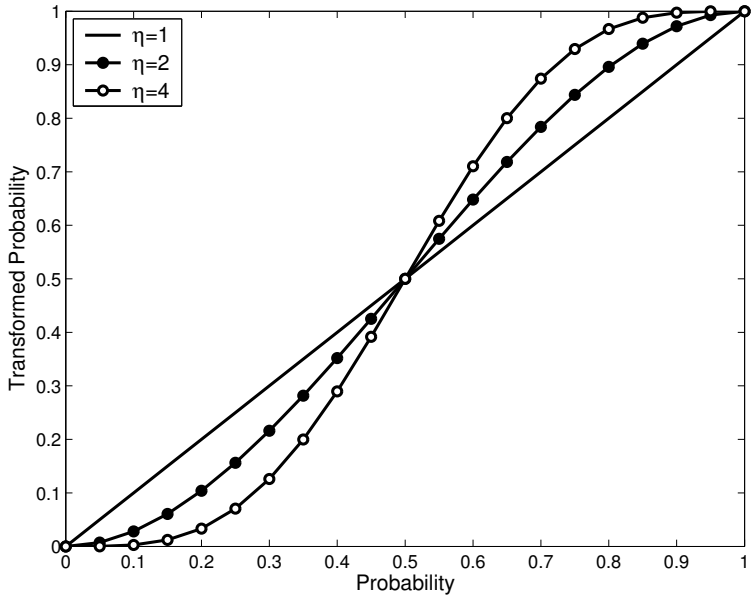

Fig. 2: The symmetric regularized Beta function, $B_{SR}(\cdot, \eta)$, for different values of η. This function can be used to 'spread out' the percentiles that lead to the grid construction.

6 The characteristic function

Chourdakis (2000, 2004) shows that, if some model parameters are considered stochastic, the (time-t) characteristic function of the log-asset price y_t can be approximated by the characteristic function of the discrete state process. This function, can in turn be computed in terms of a matrix exponential, which is based on the rate matrix of the approximating chain, and the conditional characteristic exponents. Correlations are accommodated, by allowing y_t to jump, whenever the volatility switches. In particular, we can rewrite the joint process of (y, v)

$$\begin{aligned} y_t &= y_0 + \int_0^t \mu_y(v_s) \cdot dt + \sum_{s<t} \sqrt{v_s} \cdot \Delta L_s^1 \\ v_t &= v_0 + \int_0^t \mu(v_s) \cdot dt + \int_0^t \sigma(v_s) \cdot dW_s^v \end{aligned}$$

Using the decompositions: $L^1 = \beta \cdot W + \sqrt{1-\beta^2} \cdot J^1$, of equation (2.1), and

$W = \rho \cdot W^v + \sqrt{1-\rho^2} \cdot Z$, for Z a standard Brownian motion, independent of all other processes, we can write

$$\begin{aligned} y_t &= y_0 + \int_0^t \mu_y(v_s) \cdot dt + \int_0^t \beta\rho\sqrt{v_s} \cdot dW_s^v \\ &\quad + \int_0^t \beta\sqrt{1-\rho^2}\sqrt{v_s} \cdot dZ_s + \sum_{s<t} \sqrt{1-\beta^2}\sqrt{v_s} \cdot \Delta J_s^1 \\ v_t &= v_0 + \int_0^t \mu(v_s) \cdot dt + \int_0^t \sigma(v_s) \cdot dW_s^v \end{aligned}$$

The second expression implies that $dW_s^v = \frac{1}{\sigma(v_s)} \cdot dv_s - \frac{\mu(v_s)}{\sigma(v_s)} \cdot dt$, which can be substituted into the first expression, to give

$$\begin{aligned} y_t = y_0 + \int_0^t \left(\mu_y(v_s) - \beta\rho\sqrt{v_s}\frac{\mu(v_s)}{\sigma(v_s)}\right) \cdot dt + \int_0^t \beta\rho\sqrt{v_s}\frac{1}{\sigma(v_s)} \cdot dv_s \\ + \int_0^t \beta\sqrt{1-\rho^2}\sqrt{v_s} \cdot dZ_s + \sum_{s<t} \sqrt{1-\beta^2}\sqrt{v_s} \cdot \Delta J_s^1 \end{aligned}$$

The process v_t can be now approximated by the Markov chain v_t^h, with rate matrix $\mathbf{Q}^h$, as described in the previous section. The process y_t^h, which is driven by this Markov chain, will also converge weakly to y_t

$$\begin{aligned} y_t^h = y_0 + \int_0^t \left(\mu_y(v_s^h) - \beta\rho\sqrt{v_s^h}\frac{\mu(v_s^h)}{\sigma(v_s^h)}\right) \cdot dt + \sum_{s<t} \beta\rho\sqrt{v_s^h}\frac{1}{\sigma(v_s^h)} \cdot \Delta v_s^h \\ + \int_0^t \beta\sqrt{1-\rho^2}\sqrt{v_s^h} \cdot dZ_s + \sum_{s<t} \sqrt{1-\beta^2}\sqrt{v_s^h} \cdot \Delta J_s^1 \end{aligned}$$

We can observe how the leverage effect is incorporated in this approximation: for $\rho < 0$, a volatility increase, $\Delta v^h > 0$, will cause the asset price to jump downwards. It is also important to note that if the correlation between the log-price and the variance process is equal to $\beta\rho$.

As Chourdakis (2004) shows, the approximating characteristic function is given by

$$\Phi_t^h(u) = \iota' \cdot \exp\{t\mathbf{\Psi}^h(u)\} \cdot \mathbf{v}_0 \tag{6.1}$$

where the matrix $\mathbf{\Psi}^h$ is tri-diagonal, with elements given by

$$\psi^h_{i,j}(u) = \begin{cases} q_{i,i} + F(u|V_i) & \text{, if} j = i \\ q_{i,i\pm1} G(u|V_i, V_{i\pm1}) & \text{, if} j = i \pm 1 \\ 0 & \text{, otherwise} \end{cases}$$

Here, ι is a vector of ones, and $\mathbf{v}_0$ is the initial (time-0) distribution of the parameter v^h_0. For example, if $v_0 = V_k$, then $\mathbf{v}_0$ would be a vector of zeros, with the k-th element equal to one.

The function $F(u|V_i)$ is the conditional characteristic exponent, associated with variance level V_i. Putting everything together, and applying (2.2), we have

$$F(u|V_i) = \mathtt{i}u\left(r + \frac{1}{2}\beta^2 V_i^2 - \Psi_{J^1}(-\mathtt{i}\sqrt{V_i}\sqrt{1-\beta^2}) - \beta\rho\sqrt{V_i}\frac{\mu(V_i)}{\sigma(V_i)}\right) - \frac{1}{2}u^2\beta^2(1-\rho^2)V_i + \Psi_{J^1}\left(\sqrt{V_i}\sqrt{1-\beta^2}u\right)$$

The functions $G(u|V_i, V_{i\pm1})$ are the characteristic functions of the jumps that induce correlations, namely

$$G(u|V_i, V_{i-1}) = \exp\left(-\mathtt{i}u\beta\rho\sqrt{v^h_s}\frac{1}{\sigma(v^h_s)} \cdot d_D\right)$$
$$G(u|V_i, V_{i+1}) = \exp\left(+\mathtt{i}u\beta\rho\sqrt{v^h_s}\frac{1}{\sigma(v^h_s)} \cdot d_U\right)$$

7 The pricing of European call options

Summarizing, in order to retrieve the characteristic function of the log-price, given a set of parameter values, we follow the three steps below:

1. Construct the volatility grid, $\Upsilon = \{V_1, \ldots, V_N\}$. The choice of the grid points can be based, for example, on the asymptotic behavior of the variance process.

2. Compute the rate matrix of the approximating Markov chain, $\mathbf{Q}$. The rate matrix will ensure that the chain is locally consistent with the

variance process.

3. Compute the matrix valued function $\Psi(u)$, that enables us to compute the approximate characteristic function, following equation (6.1).

Given the characteristic function, Carr and Madan (1999) show how European option prices can be computed. A set of option prices can be simultaneously computed, by application of the Fast Fourier Transform (FFT). Chourdakis (2005) shows how the computations can be substantially sped up. A complete volatility surface can be computed in a matter of seconds.

In particular, given a control parameter ζ, call prices can be computed, for log-strikes m, as the following Fourier transforms

$$C_t(m;\zeta) = \frac{\exp(-\zeta m)}{\pi}\int_0^\infty \exp(-\mathtt{i}mu)\hat{\Phi}_t(u;\zeta)\cdot du, \text{ where}$$
$$\hat{\Phi}_t(u;\zeta) = \frac{\exp(-rt)\Phi_t(u-(\zeta+1)\mathtt{i})}{\zeta^2+\zeta-u^2+(2\zeta+1)u\mathtt{i}}$$

See Carr and Madan (1999), Lee (2004) and Chourdakis (2005) for technical and implementation details.

8 Example: The Variance-Gamma square-root stochastic volatility process

We now focus on the following specification

$$y_t = y_0 + \int_0^t \mu_y(v_s)\cdot dt + \sum_{s<t}\sqrt{v_s}\cdot \Delta L_s^1 \tag{8.1}$$
$$v_t = v_0 + \int_0^t \theta_v(\bar{v}-v_s)\cdot dt + \int_0^t \phi_v\sqrt{v_s}\cdot dW_s^v \tag{8.2}$$

This specification is analogous to the popular Heston (1993) stochastic volatility model. Our model reduces to the Heston model, if the Lévy process L is a Brownian motion. In general, we will assume that L has a diffusive and a jump component, the latter following the Variance-Gamma process of Madan et al. (1998).

Following the standard decomposition (2.1), we can rewrite the log-price process as

$$y_t = y_0 + \int_0^t \mu_y(v_s) \cdot dt + \int_0^t \beta\sqrt{v_s} \cdot dW_s + \sum_{s<t} \sqrt{1-\beta^2}\sqrt{v_s} \cdot \Delta J_s^{1,\text{VG}}$$

We assume that the correlation between the Brownian motion is equal to ρ, and that $J^{1,\text{VG}}$ follows the standardized Variance-Gamma process of section 2.2. Based on this process, we can establish that the (conditional on the variance) characteristic exponent of the Lévy parts, L is

$$\Psi_L(u|V) = -\frac{1}{2}\beta^2 V u^2 - \frac{1}{\nu}\log\left(1 - \mathtt{i}\sqrt{1-\beta^2}\sqrt{V}u\theta\nu + \frac{1}{2}(1-\beta^2)V\sigma^2u^2\nu\right)$$

In the above relationship, the restriction $\nu = \frac{1-\sigma^2}{\theta^2}$ is also applied. Thus, we can determine the risk-adjusted drift by applying (4.3)

$$\mu_y(v_s) = r - \frac{1}{2}\beta^2 v_s + \frac{1}{\nu}\log\left(1 - \sqrt{1-\beta^2}\sqrt{v_s}\theta\nu - \frac{1}{2}(1-\beta^2)v_s\sigma^2\nu\right)$$

The volatility grid is based on information extracted from the unconditional distribution of the volatility process. We assume in this example a grid size of $M = 21$ points; experimentation showed little differences for various sizes above 15. Since the variance follows the square-root diffusion, it is unconditionally distributed as a Gamma random variable with (scale and shape) parameters equal to $\frac{2\theta_v\bar{v}}{\phi_v^2}$ and $\frac{\phi_v^2}{2\theta_v}$. Our base 21-point grid is based on the points with cumulative probability $\frac{1}{2\times 21} + \frac{1}{21} \times j$ for $j = 0, \ldots, 20$. The minimum and maximum volatility values were approximately 0.4% and 61.2%, annualized; thus, the approximation method truncates the support of the volatility process, with $\sqrt{v_s} \in [0.4\%, 62.1\%]$. To ensure that the more extreme values are attainable, we transform these base probabilities, using the regularized Beta function, with transform parameter $\eta = 3$. The resulting bounds are now $7 \times 10^{-4}\%$ and 106.5%, respectively. Figure 3 gives the volatility grid, together with the long run volatilities. Overall, a large number of experiments conducted, suggest that the results are very robust with

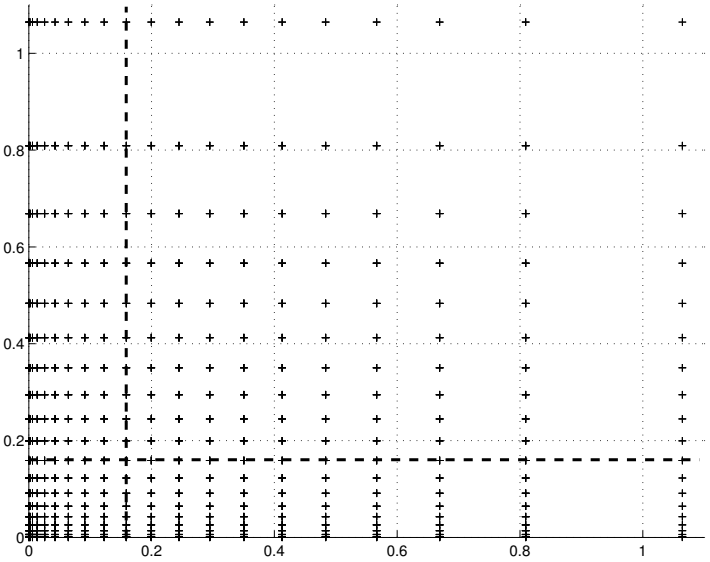

Fig. 3: The grid constructed for the 21-point variance process. The axes measure the corresponding annualized volatility values. The dashed lines indicate the long run volatilities.

respect to the grid, given that the support covers the bulk of the distribution.

The second step is to identify the rate matrix of the approximating Markov chain. This can be readily computed using equation (5.2) or (5.3),

$$\begin{aligned} q_{j-1,j} &= \frac{\phi_v^2 V_j - d_U \theta_v(\bar{v} - V_j)}{d_D(d_U + d_D)}, \text{ or } \frac{\phi_v^2 V_j + (d_D + d_U)\theta_v(\bar{v} - V_j)^-}{d_D(d_U + d_D)} \\ q_{j+1,j} &= \frac{\phi_v^2 V_j + d_D \theta_v(\bar{v} - V_j)}{d_U(d_U + d_D)}, \text{ or } \frac{\phi_v^2 V_j + (d_D + d_U)\theta_v(\bar{v} - V_j)^+}{d_D(d_U + d_D)} \\ q_{j,j} &= -q_{j+1,j} - q_{j+1,j} \end{aligned}$$

The quantities on the right are used if the quantities on the left are negative.

The process y^h, which approximates the log-price y, is given by

$$\begin{aligned} y_t^h = y_0 + \int_0^t \left(\mu_y(v_s^h) - \frac{\beta\rho\theta_v}{\phi_v}(\bar{v} - v_s^h) \right) \cdot dt + \sum_{s<t} \frac{\beta\rho}{\phi_v} \cdot \Delta v_s^h \\ + \int_0^t \beta\sqrt{1-\rho^2}\sqrt{v_s^h} \cdot dZ_s + \sum_{s<t} \sqrt{1-\beta^2}\sqrt{v_s^h} \cdot \Delta J_s^{1,\mathrm{VG}} \end{aligned}$$

Therefore, the matrix function $\mathbf{\Psi}^h$, which enables us to compute the approximate characteristic function, will have elements

$$\psi^h_{i,j}(u) = \begin{cases} q_{i,i} + \mathtt{i}u\left(\mu_y(V_i) - \frac{\beta\rho\theta_v}{\phi_v}(\bar{v} - V_i)\right) \\ \quad -\frac{1}{2}u^2\beta^2(1-\rho^2)V_i + \Psi_{J^{1,\text{VG}}}\left(\sqrt{V_i}\sqrt{1-\beta^2}u\right) & , \text{if} j = i \\ q_{i,i-1}\exp\left(-\mathtt{i}u\frac{\beta\rho}{\phi_v}d_D\right) & , \text{if} j = i-1 \\ q_{i,i+1}\exp\left(+\mathtt{i}u\frac{\beta\rho}{\phi_v}d_U\right) & , \text{if} j = i+1 \\ 0 & , \text{otherwise} \end{cases}$$

where $\Psi_{J^{1,\text{VG}}}$ is the characteristic exponent of the standardized Variance-Gamma process, given in section 2.2.

Finally, the characteristic function is inverted using the FFT, outlined in Carr and Madan (1999). In particular, the fractional version of Chourdakis (2005) is used. The parameter ζ, which dampens the function to be integrated was chosen in the range of $\frac{3}{4}$ to 3, for options with different maturities. For options with shorter maturities, higher values of ζ were used. Each maturity was integrated using 100 to 250 sampling points, and the integrand was assumed equal to zero after the seventh decimal place. This determined the upper integration bound.

The model was calibrated using the closing prices of the SP500 index options, on the 4th January 2005. Options with deltas outside the 10% to 90% range were discarded. Implied volatilities of out-of-the-money contracts were used, and option prices were computed based on these values. Zero dividends and a flat interest rate of 3% were assumed. In total, our dataset consists of 123 prices, which span a variety of maturities, ranging from two weeks to two years.

The general specification (8.1-8.2) admits some interesting special cases. These were used to assess the performance of the model, in terms of the computational speed and the quality of the numerical approximations involved. We found that the procedures outlined here provide a very robust pricing mechanism. The special cases are the following:

- For $\beta = 1$, the model converges to the familiar stochastic volatility

model of Heston (1993), since there are no discontinuities present.

- If we further let the volatility to be constant, e.g. by setting $\phi = 0$ and $\bar{v} = v_0$, the process becomes the familiar BMS geometric Brownian motion.

- Setting $\rho = \pm 1$ gives us the special case where the two Brownian motions are identical, that is to say, the continuous component of the volatility and price paths is common.

- Setting $\beta = 0$ results into a price process that is purely discontinuous. Such a process is related, but not identical, to the processes discussed in Carr et al. (2003) (there, volatility variations were achieved with further subordinations).

- By setting a constant variance, we arrive to the classic VG model of Madan et al. (1998).

The models were calibrated using a weighted sum of errors as an objective function. The pricing errors were weighted with $\frac{1}{\sqrt{t}}$, to reflect the differences of error variability with maturity (see Carr and Wu (2004a) for details). Figure 4 gives the observed and fitted prices, together with the corresponding pricing errors, for the full specification. Table 2 gives the calibrated parameters, together with the RMSE (root mean squared error).

One can make the following general remarks: As expected, and in terms of the RMSE, the fully fledged specification clearly outperforms the rest. The correlation coefficient between the two Brownian motions is -0.90, indicating that perhaps one diffusion might suffice to capture both the price and the volatility dynamics. Indeed, when such a specification is calibrated, we can observe that the RMSE does not deteriorate substantially.

The correlation coefficient appears to be an important component of the price dynamics, generating the necessary skewness in the log-prices. It is important to note, that when the price process is augmented with jumps, the jump distribution appears to be positively skewed. It also appears that, when a pure-jump process is considered, the lack of the leverage effect is

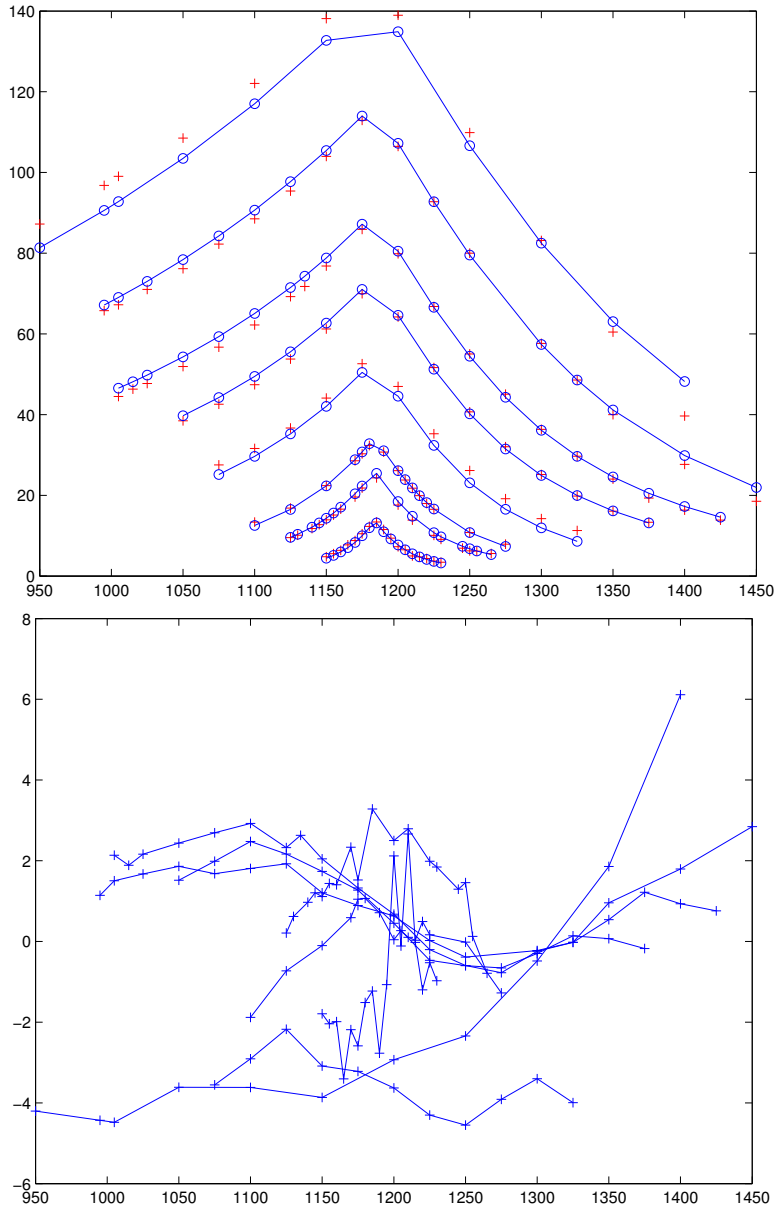

Fig. 4: The pricing performance of the stochastic volatility model with Variance-Gamma discontinuities. The upper graph gives the observed and fitted intrinsic values, for all maturities. The lower graph gives the weighted errors.

	Full	Heston	Single W	Pure Jump
$\sqrt{v_0}$	0.1631	0.1549	0.1608	0.1833
θ_v	0.2607	0.4331	0.3642	0.2489
$\sqrt{\bar{v}}$	0.2976	0.2676	0.2591	0.2561
ϕ_v	0.3937	0.4167	0.3441	0.4435
β	0.6931	1	0.6755	0
ρ	-0.9012	-0.4834	-1	
σ	0.6670		0.7532	0.8476
θ	1.2989		1.1387	-1.1938
RMSE	2.0812	2.4289	2.1602	6.3146
$Skew$	1.045	0	0.976	-0.641
$XKurt$	1.779	0	1.681	0.880

Tab. 2: Calibrated parameters for a set of models.

the reason of its very bad performance. The RMSE of a pure jump process is nearly three times the RMSE of the models that also include a diffusive component.

9 Conclusions

This paper introduced a family of models that generalize the standard Lévy processes, in the sense that they allow for stochastic volatilities. A generic decomposition of the Lévy process, together with a discretization of the volatility process, allow us to compute European option prices, with an arbitrary degree of accuracy. An example, using the familiar square-root diffusion, illustrates step-by-step the implementation of the various algorithms.

We find that the methodology presented in this paper offers a robust alternative to standard stochastic volatility models. The parameters are easily calibrated, using a standard notebook. Of course, a full analysis of the usefulness of this approach, would examine the pricing accuracy of the calibrated models using prices of more exotic contracts. We leave such an approach for further research.

References

Aït-Sahalia, Y. (2002). Telling from discrete data whether the underlying continuous-time model is a diffusion. *The Journal of Finance 57*, 2075–2112.

Andersen, T. G., L. Benzoni, and J. Lund (1998). Estimating jump diffusions for equity returns. Technical report, Kellogg Graduate School of Management.

Ané, T. and H. Geman (2000). Order flow, transaction clock, and the normality of asset returns. *Journal of Finance 11*, 79–96.

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

Bates, D. S. (1996). Jumps and stochastic volatility: Exchange rate processes implicit in deutschemark options. *Review of Financial Studies 9*, 69–108.

Bates, D. S. (1998). Pricing options under jump diffusion processes. Technical Report 37/88, The Wharton School, University of Pennsylvania.

Bates, D. S. (2000). Post-'87 crash fears in S&P500 futures options. *Journal of Econometrics 94*, 181–238.

Bertoin, J. (1996). *Lévy Processes.* Cambridge Tracks in Mathematics. Cambridge, U.K.: Cambridge University Press.

Black, F. and M. Scholes (1973, May-June). The pricing of options and corporate liabilities. *Journal of Political Economy 81*, 637–659.

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

Carr, P. and D. Madan (1999). Option valuation using the Fast Fourier Transform. *Journal of Computational Finance 3*, 463–520.

Carr, P. and L. Wu (2004a). Stochastic skew in currency options. Unpublished manuscript.

Carr, P. and L. Wu (2004b). Stochastic skew models for FX options. Technical report.

Chourdakis, K. (2000). Stochastic volatility and jumps driven by continuous time Markov chains. Technical Report 430, Queen Mary, University of London.

Chourdakis, K. (2004). Non-affine option pricing. *Journal of Derivatives 11*(3), 10–25.

Chourdakis, K. (2005). Option pricing using the Fractional FFT. *Journal of Computational Finance*.

Cox, J. C., J. E. Ingersoll, and S. A. Ross (1985). A theory of the term structure of interest rates. *Econometrica 53*, 385–407.

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

Dupuis, P. G. and H. J. Kushner (2001). *Numerical Methods for Stochastic Control Problems in Continuous Time.*, Volume 24 of *Applications of Mathematics*. New York, N.Y.: Springer Verlag.

Eraker, B., M. Johannes, and N. Polson (2003). The impact of jumps in returns and volatility. *Journal of Finance 53*, 1269–1300.

Feller, W. E. (1971). *An Introduction to Probability Theory and its Applications*. (2nd ed.). New York, NY: John Wiley and Sons.

Geman, H., D. Madan, and M. Yor (2001). Time changes for lévy processes. *Mathematical Finance 11*, 79–96.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies 6*, 327–344.

Hull, J. C. and A. White (1987). The pricing of options with stochastic volatilities. *The Journal of Finance 42*, 281–300.

Konikov, M. and D. Madan (2001). Stochastic volatility via Markov chains. Technical report, Robert J. Smith School of Business, University of Maryland.

Kushner, H. J. (1990). Numerical methods for stochastic control problems in continuous time. *SIAM Journal of Control and Optimization 28*(5), 999–1048.

Kushner, H. J. and G. DiMasi (1978). Approximations of functionals and optimal control problems on jump diffusion processes. *Journal of Mathematical Analysis and Applications 63*, 772–800.

Lee, R. (2004). Option pricing by transform methods: extensions, unification and error control. *Journal of Computational Finance 7*(3), 51–86.

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

Melino, A. and M. Turnbull (1990). Pricing foreign currency options with stochastic volatility. *Journal of Econometrics 45*, 239–265.

Merton, R. (1976). Option pricing when the underlying stock returns are discontinuous. *Journal of Financial Economics 4*, 125–144.

Merton, R. C. (1973). Theory of rational option pricing. *Bell Journal of Economics and Management Sciences 4*, 141–183.

Pan, J. (2002). The jump risk-premia implicit in options: Evidence from an integrated time series study. *Journal of Financial Economics 63*, 3–50.