Exponential Lévy models with stochastic volatility and stochastic jump-intensity

Matthew Lorig *

This version: July 13, 2012

Abstract

We consider the problem of valuing a European option written on an asset whose dynamics are described by an exponential Lévy-type model. Both the volatility and jump-intensity of the Lévy process are allowed to vary stochastically in time through common driving factors. Using results from the spectral theory of normal operators and singular perturbation theory, we derive an explicit formula for the approximate price of any European-style derivative. Additionally, we establish the accuracy of our pricing approximation. Lastly, as an example of our framework, we extend the jump-diffusion model of Merton (1976) to include stochastic volatility and stochastic jump-intensity.

Key words: spectral theory, normal operator, Lévy process, stochastic volatility, stochastic jump-intensity.

1 Introduction

An exponential Lévy model is an equity model in which an underlying $S = e^X$ is described by the exponential of a Lévy process $X$. Such models extend the geometric Brownian motion description of Black and Scholes (1973) by allowing the underlying $S$ to experience jumps, the need for which is well-documented in literature (see, Eraker (2004) and references therein). In addition to allowing the underlying $S$ to jump, exponential Lévy models are important because they capture many of the stylized features of asset prices, such as heavy tails, high-kurtosis and asymmetry of log returns.

Several well-known models fit within the class of exponential Lévy models: the jump-diffusion model of Merton (1976), the pure jump models of Mandelbrot (1963), the variance gamma model of Madan, Carr, and Chang (1998) and the double exponential model of Kou (2002). Lewis (2001); Lipton (2002) show that

* ORFE Department, Princeton University, Princeton, USA. Work partially supported by NSF grant DMS-0739195
all of the above-mentioned models allow for fast and easy computation of European option prices via one-dimensional Fourier transforms. A comprehensive reference on the subject of option-pricing in an exponential Lévy setting can be found in Cont and Tankov (2004), Chapter 11.

Despite their success, exponential Lévy models have some shortcomings. For example, because the log returns of any exponential Lévy process are independent and identically distributed, these models cannot exhibit volatility clustering or the leverage effect. To address these shortcomings, Carr and Wu (2004) add stochastic volatility (with correlation to the underlying) by stochastically time-changing the Lévy process. Notably, the models described in Carr and Wu (2004) maintain the analytic tractability that makes the class of exponential Lévy processes attractive.

In this paper, we address the need for volatility clustering and the leverage effect by allowing the Lévy triplet that describes the returns process \( X \) to be stochastic: \((\gamma, \sigma, \nu) \rightarrow (\gamma_t, \sigma_t, \nu_t)\). We then use results from the spectral theory of normal operators \(^1\) and singular perturbation theory to derive an explicit formula for the approximate price of any European-style derivative. The use of spectral theory for self-adjoint operators in finance is extensive (see, e.g. Linetsky (2007) and references therein). To our knowledge, this is the first time the spectral theory of normal operators has been applied in a financial setting.

Much like geometric Brownian motion arises as special case of an exponential Lévy process, the class of fast mean-reverting and multiscale stochastic volatility models considered in Fouque, Papanicolaou, and Sircar (2000) and Fouque, Papanicolaou, Sircar, and Solna (2011) arise as a special subset of the class of models we consider. In fact, by removing jumps from our framework, one recovers the spectral representation of the European option pricing formulas derived in Fouque, Papanicolaou, and Sircar (2000) and Fouque, Papanicolaou, Sircar, and Solna (2011).

The rest of this paper proceeds as follows. In section 2 we introduce a class of exponential Lévy models in which the volatility and jump-intensity are stochastically driven by a common factor. In section 3 we derive an expression for the approximate price of a European option (Theorem 3.1) when the underlying is described by the class of models introduced in section 2. We also quantify the accuracy of our pricing approximation (Theorem 3.2). In section 4, as an example of our framework, we extend the jump-diffusion model of Merton (1976) to include stochastic volatility and stochastic jump-intensity. We also compute (numerically) the implied volatility surface generated by this example. In section 5 we briefly describe how the class of models described in section 2 can be extended to allow for multiple driving factors of volatility and jump-intensity. Proofs, as well as some background material on the spectral theory normal operators

\(^1\)A normal operator is an operator that commutes with its Hermitian adjoint. See Appendix A for a review of the spectral theory of normal operators in a Hilbert space.
and the Fredholm alternative, are provided in an appendix.

2 Lévy processes with stochastic volatility and jump intensity

Let $(Ω, ℱ, ˜P)$ be a probability space endowed with a filtration $ℱ = {ℱ_t, t ≥ 0}$, which satisfies the usual conditions. Here, ˜P is the risk-neutral pricing measure, which we assume is chosen by the market. The filtration ℱ represents the history of the market. For simplicity, we assume that the risk-free rate of interest is zero so that all non-dividend paying assets are $(˜P, ℱ)$-martingales. All of our results can easily be extended to include constant or deterministic interest rates.

We consider a non-dividend paying asset $S$ whose dynamics under ˜P are described by the following Itô-Lévy stochastic differential equation (SDE)

\[
\begin{align*}
    dS_t &= \sigma(Y_t)S_t d\tilde{W}_t + S_t \int_{\mathbb{R}} (e^z - 1) d\tilde{N}_t(Y_t, dz), \quad S_0 = \log x, \\
    dY_t &= \left( \frac{1}{\varepsilon^2} \alpha(Y_t) - \frac{1}{\varepsilon} \Lambda(Y_t) \beta(Y_t) \right) dt + \frac{1}{\varepsilon} \beta(Y_t) d\tilde{B}_t, \quad Y_0 = y, \\
    d(\tilde{W}, \tilde{B}) &= \rho dt, \quad |\rho| ≤ 1.
\end{align*}
\]  

(2.1)

Here $\tilde{W}$ and $\tilde{B}$ are correlated Brownian motions and $\tilde{N}(Y, dz)$ is a compensated Poisson random measure

\[
d\tilde{N}_t(Y, dz) = dN_t(Y_t, dz) - \zeta(Y_t)\nu(dz)dt, \quad \mathbb{E}[dN_t(Y_t, dz)|Y_t] = \zeta(Y_t)\nu(dz)dt.
\]

We require that the measure $\nu$ satisfy

\[
\int_{\mathbb{R}} \min(1, z^2)\nu(dz) < \infty, \quad \int_{|z| ≥ 1} e^z\nu(dz) < \infty, \quad \text{and} \quad \int_{|z| ≥ 1} |z|\nu(dz) < \infty.
\]

The first integrability condition must be satisfied by all Lévy measures. The second integrability condition is needed to ensure $\mathbb{E}[S_t] < \infty$ for all $t ∈ \mathbb{R}^+$. The last integrability condition can be removed. However, keeping this condition vastly simplifies the formulas that follow, as it will allow us to replace the indicator function that usually appears in the Lévy-Kintchine formula with the constant one (i.e., $I_{|z|<1} → 1$). Although we do not require it, a correlation of $\rho < 0$ between $\tilde{W}$ and $\tilde{B}$ would be consistent with the leverage effect (i.e. a drop in the value of $S$ will usually be accompanied by an increase in volatility).

Note that both the volatility of $S$, given by $\sigma(Y) > 0$, and the Poisson random measure $N(Y, dz)$, which controls the jumps of $S$, are driven by a common stochastic process $Y$. We could have chosen for the volatility and jump measure to be controlled by separate stochastic processes. However, as an increase in volatility would likely be accompanied by an increase in jump-intensity, we think it makes sense for these two processes to be driven by the same factor.
The process \( Y \) is fast-varying in the following sense. Under the physical measure \( \mathbb{P} \), the dynamics of \( Y \) are described by
\[
dY_t = \frac{1}{\varepsilon^2} \alpha(Y_t) dt + \frac{1}{\varepsilon} \beta(Y_t) dB_t \quad \text{(under } \mathbb{P})
\]
where \( B_t = \tilde{B}_t - \int_0^t \Lambda(Y_s) ds \) is a \( \mathbb{P} \)-Brownian motion. The generator of \( Y \) under \( \mathbb{P} \) is scaled by a factor of \( 1/\varepsilon^2 \)
\[
A_{\varepsilon}^Y = \frac{1}{\varepsilon^2} \left( \frac{1}{2} \beta^2(y) \partial_{yy} + \alpha(y) \partial_y \right).
\]
Thus, \( Y \) operates with an intrinsic time-scale \( \varepsilon^2 \). We assume \( \varepsilon^2 \ll 1 \) so that the intrinsic time-scale of \( Y \) is small. Thus, \( Y \) is fast-varying. We shall further assume that (i) under \( \mathbb{P} \), the process \( Y \) is ergodic and has a unique invariant distribution \( F_Y \) (ii) the functions \( \sigma(y) \) and \( \zeta(y) \) satisfy \( \sigma^2, \zeta \in L^2(\mathbb{R}, F_Y) \) (iii) \( A_{\varepsilon}^Y \) is a Fredholm operator \(^2\) on the Hilbert space \( L^2(\mathbb{R}, F_Y) \) (iv) the smallest non-zero eigenvalue of \( -A_{\varepsilon}^Y \) is strictly positive. Typical processes which satisfy the above assumptions are
\[
\text{OU process : } \alpha(y) = m - y, \quad \beta(y) = \nu, \quad F_Y \sim \text{Normal},
\]
\[
\text{CIR process : } \alpha(y) = m - y, \quad \beta(y) = \nu \sqrt{y}, \quad F_Y \sim \text{Gamma}.
\]
Finally, we will assume that the market price of volatility risk \( \Lambda \) is bounded above and below. As mentioned in the introduction, the class of models described by (2.1) are a natural extension of those considered in Fouque, Papanicolaou, and Sircar (2000). The key differences between the class of models we consider and those considered in Fouque, Papanicolaou, and Sircar (2000) are that (i) we allow for the underlying \( S \) to jump and (ii) we allow for the jump intensity to be stochastic.

3 Option pricing

We wish to price a European-style option, which pays \( H(S_t) \) at the maturity date \( t > 0 \). It will be convenient to introduce the returns process \( X = \log S \). Using Itô’s formula for Itô-Lévy processes (see Øksendal and Sulem (2005), Theorem 1.14) one derives
\[
dx_t = \gamma(Y_t) dt + \sigma(Y_t) d\tilde{W}_t + \int_{\mathbb{R}} z d\tilde{N}_t(Y_t, dz), \quad X_0 = x,
\]
where the drift \( \gamma(Y_t) \) is given by
\[
\gamma(Y_t) = -\frac{1}{2} \sigma^2(Y_t) - \zeta(Y_t) \int_{\mathbb{R}} (e^z - 1 - z) \nu(dz).
\]
\(^2\)A Fredholm operator is an operator that has the Fredholm alternative property. See Appendix B for a discussion of the Fredholm alternative.
Using risk-neutral pricing, the value \( u^\varepsilon(t,x,y) \) of the European option under consideration is

\[
u^\varepsilon(t,x,y) = \tilde{E}_{x,y}[h(X_t)], \quad h(x) := H(e^x).
\]

From the Kolmogorov backward equation we find that \( u^\varepsilon(t,x,y) \) satisfies the following partial integro-differential equation (PIDE) and boundary condition (BC)

\[
(-\partial_t + \mathcal{A}^\varepsilon) u^\varepsilon(t,x,y) = 0, \quad u^\varepsilon(0,x,y) = h(x).
\]

Here, the partial integro-differential operator \( \mathcal{A}^\varepsilon \) is the generator of \( (X,Y) \):

\[
\mathcal{A}^\varepsilon = \frac{1}{\varepsilon^2} \mathcal{A}_0 + \frac{1}{\varepsilon} \mathcal{A}_1 + \mathcal{A}_2,
\]

\[
\mathcal{A}_0 = \mathcal{A}^1_Y = \frac{1}{2} \beta^2(y) \partial^2_{yy} + \alpha(y) \partial_y,
\]

\[
\mathcal{A}_1 = \rho \beta(y) \sigma(y) \partial^2_{xy} - \Lambda(y) \beta(y) \partial_y,
\]

\[
\mathcal{A}_2 = \gamma(y) \partial_x + \frac{1}{2} \sigma^2(y) \partial^2_{xx} + \zeta(y) \int_\mathbb{R} (e^{\varepsilon \partial_x} - 1 - z \partial_x) \nu(dz),
\]

where the domain of \( \mathcal{A}^\varepsilon \) is defined as

\[
\text{dom}(\mathcal{A}^\varepsilon) = \left\{ f : \mathbb{R}^2 \to \mathbb{R} : \lim_{t \searrow 0} \frac{1}{t} \left( \tilde{E}_{x,y}[f(X_t,Y_t)] - f(x,y) \right) \text{ exists} \right\}.
\]

In (3.4), we have used the fact that for any smooth function \( f \in C^\infty \), we have

\[
e^{\varepsilon \partial_x} f(x) = \sum_{n=0}^{\infty} \frac{1}{n!} (\varepsilon \partial_x)^n f(x) = f(x + z),
\]

by Taylor’s theorem. That is, the operator \( \partial_x \) is the generator of the translation group \( \{ e^{z \partial_x}, z \in \mathbb{R} \} \). Note that, although the payoff function \( h(e^x) \) is often not smooth, as long as \( H(s) \) satisfies

\[
|H(s_1) - H(s_2)| \leq c |s_1 - s_2|, \quad \forall s_1, s_2 \in \mathbb{R}^+,
\]

for some \( c > 0 \), then the option price \( u^\varepsilon \) is the classical solution to (3.1) with \( \mathcal{A}^\varepsilon \) given by (3.2) (see Cont and Tankov (2004), Proposition 12.1). Condition (3.5) is obviously satisfied by calls and puts.

### 3.1 Formal asymptotic analysis

For general \( (\sigma, \zeta, \alpha, \beta, \Lambda) \) there is no analytic solution to (3.1). We notice, however, that terms containing \( \varepsilon \) in (3.1) are diverging in the small-\( \varepsilon \) limit, giving rise to a singular perturbation about the \( \mathcal{O}(1) \) operator \((-\partial_t + \mathcal{A}_2)\). This special form suggests that we seek an asymptotic solution to PIDE (3.1). Thus, following Fouque, Papanicolaou, and Sircar (2000), we expand \( u^\varepsilon \)

\[
u^\varepsilon = \sum_{n=0}^{\infty} \varepsilon^n u_n.
\]
Our goal will be to find an approximation \( u^\varepsilon = u_0 + \varepsilon u_1 + O(\varepsilon^2) \). The choice of expanding in integer powers of \( \varepsilon \) is natural given the form of \( A^\varepsilon \). We will justify this expansion when we prove the accuracy of our pricing approximation in Theorem 3.2.

In the formal asymptotic analysis that follows, we insert expansion (3.6) into PIDE (3.1) and collect terms of like powers of \( \varepsilon \), starting at the lowest order. The \( O(1/\varepsilon^2) \) and \( O(1/\varepsilon) \) terms are

\[
\begin{align*}
\mathcal{O}(1/\varepsilon^2) : & \quad 0 = A_0 u_0, \\
\mathcal{O}(1/\varepsilon) : & \quad 0 = A_1 u_0 + A_0 u_1.
\end{align*}
\]

Noting that all terms in \( A_0 \) and \( A_1 \) take derivatives with respect to \( y \), we choose \( u_0 = u_0(t,x) \) and \( u_1 = u_1(t,x) \). Continuing the asymptotic analysis, the \( O(1) \) and \( O(\varepsilon) \) terms are

\[
\begin{align*}
(3.7) \quad & \mathcal{O}(1) : \quad 0 = (-\partial_t + A_2) u_0 + A_0 u_2, \\
(3.8) \quad & \mathcal{O}(\varepsilon) : \quad 0 = (-\partial_t + A_2) u_1 + A_1 u_2 + A_0 u_3,
\end{align*}
\]

where we have used the fact that \( A_1 u_1 = 0 \) in the \( O(1) \) equation. Equations (3.7) and (3.8) are equations of the form

\[
A_0 u = \chi.
\]

Since \( A_0 \) is a Fredholm operator, by the Fredholm alternative, a solution \( u \in L^2(\mathbb{R}, E_Y) \) to (3.9) exists if and only if \( \chi \) satisfies the centering condition

\[
\langle \chi \rangle := \int \chi dF_Y = 0.
\]

Applying the centering condition to (3.7) and (3.8) yields

\[
\begin{align*}
(3.11) \quad & \mathcal{O}(1) : \quad 0 = (-\partial_t + \langle A_2 \rangle) u_0, \\
(3.12) \quad & \mathcal{O}(\varepsilon) : \quad 0 = (-\partial_t + \langle A_2 \rangle) u_1 + \langle A_1 u_2 \rangle.
\end{align*}
\]

Note, from and (3.7) and (3.11) we have

\[
A_0 u_2 = (-\partial_t + A_2) u_0 + (-\partial_t + \langle A_2 \rangle) u_0 = - (A_2 - \langle A_2 \rangle) u_0
\]

\[
= -\frac{1}{2} \left( \sigma^2 - \langle \sigma^2 \rangle \right) (\partial_{xx}^2 - \partial_x^2) u_0 - (\zeta - \langle \zeta \rangle) \left( -\int_{\mathbb{R}} (e^z - 1 - z) \nu(dz) \partial_x + \int_{\mathbb{R}} (e^z \partial_x - 1 - z \partial_x) \nu(dz) \right) u_0
\]

\[
= -A_0 \left( \frac{1}{2} \eta (\partial_{xx}^2 - \partial_x^2) - \xi \int_{\mathbb{R}} (e^z - 1 - z) \nu(dz) \partial_x + \xi \int_{\mathbb{R}} (e^z \partial_x - 1 - z \partial_x) \nu(dz) \right) u_0,
\]

(3.13)
where we have introduced \( \eta(y) \) and \( \xi(y) \) as solutions to

\[
A_0 \eta = \sigma^2 - \langle \sigma^2 \rangle, \quad A_0 \xi = \zeta - \langle \zeta \rangle.
\]

Thus, from (3.12) and (3.13) we find

\[
O(\varepsilon) : (-\partial_t + \langle A_2 \rangle)u_1 = -B u_0,
\]

where the operator \( B \) is given by

\[
B = \left\langle -A_1 \left( \frac{1}{2} \eta(y) (\partial_x^2 - \partial_z) - \xi \int_{\mathbb{R}} (e^z - 1 - z) \nu(dz) \partial_x + \xi \int_{\mathbb{R}} (e^{z\partial_x} - 1 - z \partial_x) \nu(dz) \right) \right\rangle
\]

\[
= V_3 (\partial_{xx}^3 - \partial_{zz}^2) + U_3 \left( - \int_{\mathbb{R}} (e^z - 1 - z) \nu(dz) \partial_{xx}^2 + \int_{\mathbb{R}} (e^{z\partial_x} - 1 - z \partial_x) \partial_x \nu(dz) \right)
\]

\[
+ V_2 (\partial_{xx}^3 - \partial_x) + U_2 \left( - \int_{\mathbb{R}} (e^z - 1 - z) \nu(dz) \partial_x + \int_{\mathbb{R}} (e^{z\partial_x} - 1 - z \partial_x) \nu(dz) \right),
\]

and the constants \((V_3, U_3, V_2, U_2)\) are defined as

\[
V_3 = -\frac{\rho}{2} \langle \beta \sigma \partial_y \eta \rangle, \quad U_3 = -\rho \langle \beta \sigma \partial_y \xi \rangle, \quad V_2 = \frac{1}{2} \langle \beta \Lambda \partial_y \eta \rangle, \quad U_2 = \langle \beta \Lambda \partial_y \xi \rangle.
\]

This is as far as we will take the asymptotic analysis. To review, we have found that \( u_0(t,x) \) and \( u_1(t,x) \) satisfy PIDEs (3.11) and (3.14) respectively. We also impose the following BCs

\[
O(1) : \quad u_0(0,x) = h(x),
\]

\[
O(\varepsilon) : \quad u_1(0,x) = 0.
\]

### 3.2 Explicit solution for \( u_0(t,x) \) and \( u_1(t,x) \)

In order to find explicit formulas for \( u_0(t,x) \) and \( u_1(t,x) \), we note that the operator \( \langle A_2 \rangle \) is a normal operator in the Hilbert space \( L^2(\mathbb{R}, dx) \)

\[
\langle A_2 \rangle = \langle \gamma \rangle \partial_x + \frac{1}{2} \langle \sigma^2 \rangle \partial_x^2 + \langle \zeta \rangle \int_{\mathbb{R}} (e^{z\partial_x} - 1 - z \partial_x) \nu(dz).
\]

Thus, we may apply results from the spectral theory of normal operators to solve PIDEs (3.11) and (3.14). The discrete spectrum of \( \langle A_2 \rangle \) is empty: \( \sigma_d(\langle A_2 \rangle) = \emptyset \). The (improper) eigenfunctions \( \psi_\lambda(x) \) and eigenvalues \( \phi_\lambda \) of the operator \( \langle A_2 \rangle \) satisfy

\[
\langle A_2 \rangle \psi_\lambda = \phi_\lambda \psi_\lambda, \quad \lim_{x \to \pm \infty} |\psi_\lambda(x)| < \infty, \quad \lambda \in \sigma_c(\langle A_2 \rangle),
\]

\(^3\text{To see that } \langle A_2 \rangle \text{ is normal note that } (\partial_x^n)\ast = (-1)^n \partial_x^n \text{ and } \partial_x^n \partial_x^m = \partial_x^m \partial_x^n.\)
and are given explicitly by

\[
\psi_\lambda(x) = \frac{1}{\sqrt{2\pi}} e^{i\lambda x}, \quad \phi_\lambda = i\langle \gamma \rangle \lambda - \frac{1}{2} (\sigma^2) \lambda^2 + \langle \zeta \rangle \int_\mathbb{R} (e^{i\lambda z} - 1 - i\lambda z) \nu(dz), \quad \lambda \in \mathbb{R}.
\]

Note that the eigenfunctions are normalized according to

\[
(\psi_\lambda, \psi_\lambda') := \int_{\mathbb{R}} \overline{\psi_\lambda(x)} \psi_\lambda'(x) dx = \delta(\lambda - \lambda').
\]

**Theorem 3.1.** The solution \(u_0(t, x)\) to PIDE (3.11) with BC (3.15) is

\[
u_0(t, x) = \int_{\mathbb{R}} e^{t\phi_\lambda} (\psi_\lambda, h) \psi_\lambda(x) d\lambda,
\]

and the solution \(u_1(t, x)\) to PIDE (3.14) with BC (3.16) is

\[
u_1(t, x) = \int_{\mathbb{R}} t e^{t\phi_\lambda} (\psi_\lambda, h) B(\lambda) \psi_\lambda(x) d\lambda,
\]

where \(\psi_\lambda(x)\) and \(\phi_\lambda\) are given in (3.17) and

\[
B(\lambda) = V_3 (-i\lambda^3 + \lambda^2) + U_3 \left( -\lambda^2 \int_{\mathbb{R}} (e^z - 1 - z) \nu(dz) + i\lambda \int_{\mathbb{R}} (e^{i\lambda z} - 1 - i\lambda z) \nu(dz) \right) + V_2 (-\lambda^2 - i\lambda) + U_2 \left( -i\lambda \int_{\mathbb{R}} (e^z - 1 - z) \nu(dz) + \int_{\mathbb{R}} (e^{i\lambda z} - 1 - i\lambda z) \nu(dz) \right).
\]

**Proof.** Formally, the solutions to PIDEs (3.11) and (3.14) with BCs (3.15) and (3.16) respectively are

\[
u_0(t, x) = e^{t\langle A_2 \rangle} h(x), \quad \nu_1(t, x) = \int_0^t e^{(t-s)\langle A_2 \rangle} (B_0(s, x)) ds.
\]

Now, using equation (A.3) one derives

\[
u_0(t, x) = \int_{\mathbb{R}} e^{t\phi_\lambda} (\psi_\lambda, h) \psi_\lambda(x) d\lambda,
\]

\[
u_1(t, x) = \int_{\mathbb{R}} t e^{t\phi_\lambda} (\psi_\lambda, h) B(\lambda) \psi_\lambda(x) d\lambda.
\]

The last line follows by exchanging the order of integration and noting that \((\psi_\lambda, B u_0(s, \cdot)) = e^{s\phi_\lambda} (\psi_\lambda, h) B(\lambda)\).

One can easily verify directly that \(u_0(t, x)\) (3.18) and \(u_1(t, x)\) (3.19) satisfy PIDEs (3.11) and (3.14) respectively, as well as BCs (3.15) and (3.16).

### 3.3 Accuracy of the approximation

We have now derived an approximation \(u^\varepsilon \approx u_0 + \varepsilon u_1\) for the price of any European option. However, this derivation relied on formal singular perturbation arguments. In what follows, we establish the accuracy of our pricing approximation. For our accuracy result, we shall need the following assumption:
- The payoff function $h(x)$ and all of its derivatives are smooth and bounded.

Obviously, many common derivatives – e.g., call and put options – do not fit this assumption. To prove the accuracy of our pricing approximation for calls and puts would require regularizing the option payoff as is done in Fouque, Papanicolaou, Sircar, and Sølna (2003). The regularization procedure is beyond the scope of this paper. As such, we limit our analysis to options with smooth and bounded payoffs. Our accuracy result is as follows:

**Theorem 3.2.** For fixed $(t, x, y)$, there exists a constant $C$ such that for any $\varepsilon \leq 1$ we have

$$|u^\varepsilon - (u_0 + \varepsilon u_1)| \leq C \varepsilon^2.$$  

**Proof.** See appendix C.

Theorem 3.2 gives us information about how our pricing approximation behaves as $\varepsilon \to 0$. In practice $\varepsilon$ is small, but fixed (it does not go to zero). Without knowing what the constant $C$ is in theorem 3.2, it is difficult to gauge exactly how good the pricing approximation is. As such, in the example provided in section 4, we will compare the approximate price $u_0 + \varepsilon u_1$ of a derivative-asset, calculated using the formulas in Theorem 3.1, to the full price $u^\varepsilon$, calculated via Monte Carlo simulation.

---

**4 Example: Merton jump-diffusion with stochastic volatility and jump-intensity**

In this section we provide one specific example within the class of models described in section 2. Specifically, we extend the jump-diffusion model of Merton (1976) to include stochastic volatility and jump-intensity. In the Merton jump-diffusion model, jumps are log-normally distributed. Thus, we let the measure $\nu$ be given by

$$\nu(dz) = \frac{1}{\sqrt{2\pi s^2}} \exp\left(\frac{-(z-m)^2}{2s^2}\right) dz.$$  

Under this specification, we have

$$\langle \gamma \rangle = -\frac{1}{2} \langle \sigma^2 \rangle - \langle \zeta \rangle \left(e^{m+s/2} - m - 1\right),$$

$$\phi_\lambda = i \langle \gamma \rangle \lambda - \frac{1}{2} \langle \sigma^2 \rangle \lambda^2 + \langle \zeta \rangle \left(e^{i\lambda m - \frac{1}{2} s^2 \lambda^2} - i\lambda m - 1\right),$$

$$B(\lambda) = V_3 (-i\lambda^3 + \lambda^2) + U_3 \left(\lambda^2 \left(e^{m+s/2} - 1 - m\right) + i\lambda \left(e^{i\lambda m - s^2 \lambda^2/2} - 1 - i\lambda m\right)\right)$$

$$+ V_2 (-\lambda^2 - i\lambda) + U_2 \left(-i\lambda \left(e^{m+s^2/2} - 1 - m\right) + \left(e^{i\lambda m - s^2 \lambda^2/2} - 1 - i\lambda m\right)\right).$$
For a European call option with payoff \( h(X_t) = (e^{X_t} - e^k)^+ \), the inner product

\[
\langle \psi, h \rangle = \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-i\lambda x} (e^x - e^k)^+ \, dx = -\frac{e^{k-ik\lambda}}{\sqrt{2\pi} (i\lambda + \lambda^2)},
\]

converges only if we fix \( \text{Im}(\lambda) < -1 \). Thus, when evaluating integrals (3.18) and (3.19), we must set \( \lambda = \lambda_0 + i\lambda_i \) with \( \lambda_i < -1 \) and integrate along a contour parallel to the real axis in the complex plane (i.e., \( d\lambda = d\lambda_r \)).

The values of \( (\langle \sigma^2 \rangle, \langle \zeta \rangle, V_3, U_3, V_2, U_2) \) depend on the particular choice of \( \sigma(y) \) and \( \zeta(y) \) as well as a specific choice for the \( Y \) process. In the numerical examples below we let \( \alpha(y) = -y, \beta(y) = \beta, \) and \( \Lambda(y) = \Lambda \) so that

\[
dY_t = \left( -\frac{1}{\varepsilon^2} Y_t - \frac{1}{\varepsilon} \Lambda \beta \right) \, dt + \frac{1}{\varepsilon} \beta \, dB_t,
\]

and we choose \( \sigma(y) = ae^y \) and \( \zeta(y) = be^y \). With these choices the invariant distribution of \( Y \) under the physical measure \( \mathbb{P} \) is normal \( F_Y \sim \mathcal{N}(0, \frac{\beta^2}{2}) \) and we can compute explicitly

\[
\langle \sigma^2 \rangle = a^2 e^{\beta^2}, \quad \langle \zeta \rangle = be^\beta,
\]

\[
V_3 = \frac{\beta}{\beta} a^2 e^{\frac{\beta^2}{2}} \left( e^{\beta^2} - 1 \right), \quad U_3 = \frac{\beta}{\beta} 2ab \left( e^{\beta^2} - e^{\frac{\beta^2}{2}} \right),
\]

\[
V_2 = -\beta \Lambda a^2 e^{\beta^2}, \quad U_2 = -\beta \Lambda be^\beta.
\]

In figure 1 we fix the time to maturity at \( t = 1/10 \) and we plot the Black-Scholes implied volatility smile induced by the approximate price of European calls \( u_0 + \varepsilon u_1 \) for different values of \( \varepsilon = \{0.1, 0.033, 0.01\} \). For comparison, we also plot the implied volatility smile induced by the full price \( u^c \) (computed using Monte Carlo simulation). As expected, as \( \varepsilon \) goes to zero, the implied volatility induced by the approximate price \( u_0 + \varepsilon u_1 \) converges to the implied volatility induced by the full price \( u^c \).

## 5 Extension to multiscale stochastic volatility and jump intensity

It is worth mentioning that the results of this paper can be extended in a straightforward manner to include multiscale stochastic volatility and jump intensity. We briefly describe how this may be done. Our intent in this section is not to be rigorous, but rather to give a flavor of the computations involved in this extension.

To begin, we modify the dynamics of of \( S \) slightly. Letting \( S = e^X \) we have

\[
\begin{align*}
\frac{dX_t}{\mathbb{P}} &= \gamma(Y_t, Z_t) \, dt + \sigma(Y_t, Z_t) \, d\tilde{W}_t^x + \int_{\mathbb{R}} s \, d\tilde{N}_t(Y_t, Z_t, ds), \quad X_0 = x, \\
\frac{dY_t}{\mathbb{P}} &= \left( \frac{1}{\varepsilon^2} \alpha(Y_t) - \frac{1}{\varepsilon} \Lambda(Y_t, Z_t) \beta(Y_t) \right) \, dt + \frac{1}{\varepsilon} \beta(Y_t) \, d\tilde{W}_t^y, \quad Y_0 = y, \\
\frac{dZ_t}{\mathbb{P}} &= \left( \delta^2 c(Z_t) - \delta \Gamma(Y_t, Z_t) g(Z_t) \right) \, dt + \delta g(Z_t) \, d\tilde{W}_t^z, \quad Z_0 = z.
\end{align*}
\]

\text{(under } \mathbb{P} \text{)}
Here, $Z$ is a slow-varying factor, in the sense that its infinitesimal generator under $\mathbb{P}$ is scaled by $\delta^2$, which is assumed to be a small parameter: $\delta^2 << 1$. The Brownian motions $\tilde{W}^x, \tilde{W}^y, \tilde{W}^z$ have correlations $\rho_{xy}, \rho_{xz}$ and $\rho_{yz}$ (which must be such that the covariance matrix is positive semi-definite), the compensated Poisson random measure $\tilde{N}(Y, Z, ds)$ satisfies

$$d\tilde{N}_t(Y_t, Z_t, ds) = dN_t(Y_t, Z_t, ds) - \zeta(Y_t, Z_t)\nu(ds)dt,$$

and the drift $\gamma(Y_t, Z_t)$ is given by

$$\gamma(Y_t, Z_t) = -\frac{1}{2}\sigma^2(Y_t, Z_t) - \zeta(Y_t, Z_t) \int_\mathbb{R} (e^s - 1 - s)\nu(ds).$$

Using risk-neutral pricing, the value $u^{\varepsilon, \delta}(t, x, y, z)$ of a European option in this setting is

$$u^{\varepsilon, \delta}(t, x, y, z) = \tilde{E}_{x,y,z}[h(X_t)],$$

where the partial integro-differential operator $A^{\varepsilon, \delta}$ is the generator of $(X, Y, Z)$. The operator $A^{\varepsilon, \delta}$ has the following form

$$A^{\varepsilon, \delta} = \frac{1}{\varepsilon^2}A_0 + \frac{1}{\varepsilon}A_1 + A_2 + \frac{\delta}{\varepsilon}M_1 + \delta M_2 + \delta^2 M_2.$$

Terms containing $\delta$ in (5.1) are small in the small-$\delta$ limit, giving rise to a regular perturbation. Thus, (5.1) has the form of a combined singular-regular perturbation about the $O(1)$ operator $(-\partial_t + A_2)$. Following Fouque, Papanicolaou, Sircar, and Solna (2011) we seek a solution of the form

$$u^{\varepsilon, \delta} = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \varepsilon^n \delta^m u_{n,m}.$$

Our goal is to find an approximation $u^{\varepsilon, \delta} = u_{0,0,0} + \varepsilon u_{1,0} + \delta u_{0,1} + O(\varepsilon^2 + \delta^2)$. A formal asymptotic analysis yields the following PIDEs for $u_{0,0}, u_{1,0}$ and $u_{0,1}$

\begin{align*}
\mathcal{O}(1) & : \quad (-\partial_t + (A_2))u_{0,0} = 0, & u_{0,0}(0, x, z) = h(x), \\
\mathcal{O}(\varepsilon) & : \quad (-\partial_t + (A_2))u_{1,0} = -2u_{0,0}, & u_{1,0}(0, x, z) = 0, \\
\mathcal{O}(\delta) & : \quad (-\partial_t + (A_2))u_{0,1} = -(M_1)u_{0,0}, & u_{0,1}(0, x, z) = 0.
\end{align*}
where, as in section 3.1, the $y$-dependence has disappeared from $u_{0.0}$, $u_{1.0}$ and $u_{0.1}$. The operators $\langle A_2 \rangle$, $\mathcal{B}$ and $\langle M_1 \rangle$ are given by

$$\langle A_2 \rangle = \langle \gamma(\cdot, z) \rangle \partial_z + \frac{1}{2} \langle \sigma^2(\cdot, z) \rangle \partial_{xx}^2 + \langle \zeta(\cdot, z) \rangle \int_{\mathbb{R}} \left( e^{s \partial_z} - 1 - s \partial_z \right) \nu(ds),$$

$$\mathcal{B} = V_3(z) \left( \partial_{xxx}^2 - \partial_{xx}^2 \right) + U_3(z) \left( - \int_{\mathbb{R}} \left( e^s - 1 - s \right) \nu(ds) \partial_{xx}^2 + \int_{\mathbb{R}} \left( e^{s \partial_z} - 1 - s \partial_z \right) \partial_z \nu(ds) \right)$$

$$+ V_2(z) \left( \partial_{xx}^2 - \partial_z \right) + U_2(z) \left( - \int_{\mathbb{R}} \left( e^s - 1 - s \right) \nu(ds) \partial_z + \int_{\mathbb{R}} \left( e^{s \partial_z} - 1 - s \partial_z \right) \nu(ds) \right),$$

$$\langle M_1 \rangle = -g(z) \langle \Gamma(\cdot, z) \rangle \partial_z + g(z) \rho_{xz} \langle \sigma(\cdot, z) \rangle \partial^2_{xz},$$

where the $z$-dependent parameters $(V_3(z), U_3(z), V_2(z), U_2(z))$ are

$$V_3(z) = -\frac{\rho_{xy}}{2} \langle \beta(\cdot) \sigma(\cdot, z) \partial_y \eta(\cdot, z) \rangle, \quad U_3(z) = -\rho_{xy} \langle \beta(\cdot) \sigma(\cdot, z) \partial_y \xi(\cdot, z) \rangle,$$

$$V_2(z) = \frac{1}{2} \langle \beta(\cdot) \Lambda(\cdot, z) \partial_y \eta(\cdot, z) \rangle, \quad U_2(z) = \langle \beta(\cdot) \Lambda(\cdot, z) \partial_y \xi(\cdot, z) \rangle.$$

The expressions for $u_{0.0}$ and $u_{1.0}$ are analogous to those given for $u_0$ and $u_1$ in Theorem 3.1. An expression for $u_{0.1}$ is obtained using the spectral theory of normal operators in a Hilbert space

$$u_{0.1}(t, x, z) = \int_0^t \int_{\mathbb{R}} e^{(t-s)\langle A_2 \rangle} \left( \langle M_1 \rangle u_{0.0}(s, x, z), \psi_\lambda(x, z) \right) d\lambda \, ds,$$

where $\psi_\lambda(x, z)$ are the eigenfunctions of $\langle A_2 \rangle$ and $\phi_\lambda(z)$ are the corresponding eigenvalues. Note, care must be taken when computing $\langle M_1 \rangle u_{0.0}$ as both terms in $\langle M_1 \rangle$ contain the operator $\partial_z$ and $u_{0.0}$ depends on $z$ through both $\langle \sigma^2(\cdot, z) \rangle$ and $\langle \zeta(\cdot, z) \rangle$. A careful computation shows that $u_{0.1}$ is linear in the following four parameters

$$V_1(z) = g(z) \rho_{xz} \langle \sigma(\cdot, z) \rangle \partial_z \langle \sigma^2(\cdot, z) \rangle, \quad V_0(z) = -g(z) \langle \Gamma(\cdot, z) \rangle \partial_z \langle \sigma^2(\cdot, z) \rangle,$$

$$U_1(z) = g(z) \rho_{xz} \langle \sigma(\cdot, z) \rangle \partial_z \langle \zeta(\cdot, z) \rangle, \quad U_0(z) = -g(z) \langle \Gamma(\cdot, z) \rangle \partial_z \langle \zeta(\cdot, z) \rangle.$$

Finally, the accuracy of the multiscale pricing approximation $u_{0.0} + \varepsilon u_{1.0} + \delta u_{0.1}$ is as follows: for fixed $(t, x, y, z)$ there exists a constant $C$ such that for any $\varepsilon \leq 1, \delta \leq 1$ we have

$$|u^{\varepsilon, \delta} - (u_{0.0} + \varepsilon u_{1.0} + \delta u_{0.1})| \leq C(\varepsilon^2 + \delta^2).$$

The proof of this error bound is analogous to the proof found in chapter 4 of Fouque, Papanicolaou, Sircar, and Solna (2011).
6 Conclusion

In this paper, we have introduced a class of exponential Lévy models in which the volatility and jump-intensity vary stochastically in time. Using techniques from the spectral theory of normal operators and singular perturbation theory we have derived a general formula for the approximate price of any European-style derivative. Furthermore, we have quantified the accuracy of our pricing approximation both theoretically (see Theorem 3.2) and numerically (see figure 1). We hope this work motivates further research into applications of normal operators in finance. It would be interesting, for example, to find a class of models with local volatility and local jump-intensity, whose generators are normal on some Hilbert space. A possible extension of this paper in particular would be to allow the jump sizes (rather than just the jump intensity) to vary stochastically in time.

A Spectral theory of normal operators in a Hilbert space

In this appendix we summarize the theory of normal operators acting on a Hilbert space. A detailed exposition on this topic (including proofs) can be found in Reed and Simon (1980) and Rudin (1973).

Let $\mathcal{H}$ be a Hilbert space with inner product $(\cdot, \cdot)$. A linear operator is a pair $(\text{dom}(A), A)$ where $\text{dom}(A)$ is a linear subset of $\mathcal{H}$ and $A$ is a linear map $A : \text{dom}(A) \rightarrow \mathcal{H}$. The adjoint of an operator $A$ is an operator $A^*$ such that $(Af, g) = (f, A^*g), \forall f \in \text{dom}(A), g \in \text{dom}(A^*)$, where

$$\text{dom}(A^*) := \{g \in \mathcal{H} : \exists h \in \mathcal{H} \text{ such that } (Af, g) = (f, h) \forall f \in \text{dom}(A)\}.$$  

An operator $(\text{dom}(A), A)$ is said to be self-adjoint in $\mathcal{H}$ if

$$\text{dom}(A) = \text{dom}(A^*), \quad (Af, g) = (f, Ag) \quad \forall f, g \in \text{dom}(A).$$

Throughout this appendix, for any self-adjoint operator $A$, we will assume that $\text{dom}(A)$ is a dense subset of $\mathcal{H}$. A densely defined self-adjoint operator is closed (see Rudin (1973), Theorem 13.9). An operator $(\text{dom}(A), A)$ is said to be normal in $\mathcal{H}$ if it is closed, densely defined and commutes with its adjoint: $A^*A = AA^*$. Clearly, every self-adjoint operator is a normal operator.

Given a linear operator $A$, the resolvent set $\rho(A)$ is defined as the set of $\lambda \in \mathbb{C}$ such that the mapping $(A - \text{Id} \lambda)$ is one-to-one and $R_\lambda := (A - \text{Id} \lambda)^{-1}$ is continuous with $\text{dom}(R_\lambda) = \mathcal{H}$. The operator $R_\lambda : \mathcal{H} \rightarrow \mathcal{H}$ is called the resolvent. The spectrum $\sigma(A)$ of an operator $A$ is defined as $\sigma(A) := \mathbb{C} \setminus \rho(A)$. We say that $\lambda \in \sigma(A)$ is an eigenvalue of $A$ if there exists $\psi \in \text{dom}(A)$ such that the eigenvalue equation is satisfied

$$(A.1) \quad A \psi = \lambda \psi.$$
A function $\psi$ that solves (A.1) is called an eigenfunction of $A$ corresponding to $\lambda$. The multiplicity of an eigenvalue $\lambda$ is the number of linearly independent eigenfunctions for which equation (A.1) is satisfied. The spectrum of an operator $A$ can be decomposed into two disjoint sets called the discrete and essential spectra: $\sigma(A) = \sigma_d(A) \cup \sigma_e(A)$. For a normal operator $A$, a number $\lambda \in \mathbb{C}$ belongs to $\sigma_d(A)$ if and only if $\lambda$ is an isolated point of $\sigma(A)$ and $\lambda$ is an eigenvalue of finite multiplicity (see Rudin (1973), Theorem 12.29).

A projection-valued measure on the measure space $(\mathbb{C}, \mathcal{B}(\mathbb{C}))$ is a family of bounded linear operators $\{E(B), B \in \mathcal{B}(\mathbb{C})\}$ in $\mathcal{H}$ that satisfies:

1. $E(\emptyset) = 0$ and $E(\mathbb{C}) = \text{Id}$.
2. $E(B)$ is an orthogonal projection. That is, $E^2(B) = E(B)$ and $E(B)$ is self-adjoint: $E^*(B) = E(B)$.
3. $E(A \cap B) = E(A)E(B)$.
4. If $B = \bigcup_{i=1}^{\infty} B_i$ and $B_i \cap B_j = \emptyset$ for $i \neq j$ then $E(B) = \lim_{n \to \infty} \sum_{i=1}^{n} E(B_j)$, where the limit is in the strong operator topology.
5. For every $f, g \in \mathcal{H}$ the set function $\mu_{f,g}(B) := (f, E(B)g)$ is a complex measure on $\mathcal{B}(\mathbb{C})$.

**Theorem A.1** (Spectral Representation Theorem). There is a one-to-one correspondence between normal operators $A$ and projection-valued measures $E$ on $\mathcal{H}$, the correspondence being given by

$$A = \int_{\sigma(A)} \lambda E(d\lambda).$$

If $g(\cdot)$ is a Borel function on $\mathbb{C}$ then

$$(A.2) \quad g(A) = \int_{\sigma(A)} g(\lambda) E(d\lambda), \quad \text{dom}(g(A)) = \{f \in \mathcal{H} : \int_{\sigma(A)} |g(\lambda)|^2 \mu_{f,f}(d\lambda) < \infty\}.$$ 

**Proof.** See Rudin (1973) Theorems 12.21 and 13.33.

As a practical matter, if $A$ is an integro-differential operator acting on a Hilbert space $L^2(\mathbb{R}, m(x)dx)$, then the operators defined by (A.2) can be constructed by solving the proper and improper eigenvalue problems

<table>
<thead>
<tr>
<th>Proper</th>
<th>$A \psi_n = \phi_n \psi_n$, $\phi_n \in \sigma_d(A)$, $\psi_n \in \mathcal{H}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Improper</td>
<td>$A \psi_\lambda = \phi_\lambda \psi_\lambda$, $\phi_\lambda \in \sigma_e(A)$, $\psi_\lambda \notin \mathcal{H}$</td>
</tr>
</tbody>
</table>

4The essential spectrum may be further decomposed into the continuous spectrum and the residual spectrum. It can be shown that the residual spectrum of an ordinary differential operator is empty (see Roach (1982), page 184).
5The term “improper” is used because the improper eigenvalues $\lambda \notin \sigma_d(A)$ and the improper eigenfunctions $\psi_\lambda \notin \mathcal{H}$ since $(\psi_\lambda, \psi_\lambda) = \infty$.  

14
For the improper eigenvalue problem one extends the domain of $A$ to include functions all functions $f$ for which $Af$ makes sense and for which the following boundedness conditions are satisfied

$$
\lim_{x \to \pm \infty} |f(x)|^2 m(x) < \infty.
$$

After normalizing, the proper and improper eigenfunctions $A$ satisfy the following orthogonality relations

$$
(\psi_n, \psi_m) = \delta_{n,m}, \quad (\psi_\lambda, \psi_\lambda') = \delta(\lambda - \lambda'), \quad (\psi_n, \psi_\lambda) = 0.
$$

The operator $g(A)$ in (A.2) is constructed as follows (see Hanson and Yakovlev (2002), section 5.3.2)

(A.3) 

$$
g(A)f = \sum_{\lambda \in \sigma_d(A)} g(\phi_\lambda) (\psi_\lambda, f) \psi_\lambda + \int_{\sigma_e(A)} g(\phi_\lambda) (\psi_\lambda, f) \psi_\lambda d\lambda.
$$

It is not always easy to evaluate divergent integrals of the form $(\psi_\lambda, \psi_\lambda')$ and verify that they are in fact delta functions $\delta(\lambda - \lambda')$. A method for directly obtaining properly normalised improper eigenfunctions can be found on page 238 of Friedman (1956).

**B The Fredholm alternative and the centering condition**

Throughout this appendix, we let $A$ be a compact self-adjoint operator in a Hilbert space $\mathcal{H}$. As $A$ is compact and self-adjoint, its eigenfunctions $\{\psi_n\}$ form a countable orthonormal basis of $\mathcal{H}$. We consider the following problem: find, $u \in \mathcal{H}$ such that

(B.1) 

$$(A - \lambda)u = \chi,$$

where the function $\chi \in \mathcal{H}$ and the constant $\lambda \in \mathbb{R}$ are given.

**Theorem B.1** (Fredholm Alternative). If $A$ is a compact self-adjoint operator in a Hilbert space $\mathcal{H}$, then one of the following is true:

1. Either $\lambda$ is not an eigenvalue of $A$, in which case equation (B.1) has a unique solution

$$
\begin{align*}
&u = R_\lambda \chi = \sum_n \frac{(\psi_n, \chi)}{\lambda_n - \lambda} \psi_n.
\end{align*}
$$

2. Or, $\lambda$ is an eigenvalue of $A$. Suppose this is the case. Let $\lambda = \lambda_1 = \lambda_2 = \cdots = \lambda_m$ (i.e., the eigenvalue $\lambda$ has multiplicity $m$). Then (B.1) has a solution if and only if $(\psi_n, \chi) = 0$ for all $n \leq m$. Assuming $(\psi_n, \chi) = 0$ for all $n \leq m$, a solution to (B.1) has the form

$$
\begin{align*}
u &= \sum_{n > m} \frac{(\psi_n, \chi)}{\lambda_n - \lambda_k} \psi_n + \sum_{n \leq m} c_n \psi_n, \\
c_n &\in \mathbb{R}.
\end{align*}
$$
Proof. See Reed and Simon (1980), Theorem VI.14 and the ensuing corollary.

If a self-adjoint operator $A$ is not compact, Fredholm alternative may still apply. We call any operator for which Theorem B.1 applies a Fredholm operator.

To see how the Fredholm alternative relates to the centering condition (3.10), we note that $\lambda = 0$ is an eigenvalue of $A_0$ (3.3), which is a self-adjoint operator in $L^2(\mathbb{R}, F_Y)$. The (properly normalized) eigenfunction of $A_0$ corresponding to the eigenvalue $\lambda = 0$ is the constant $\psi_1 = 1$. Since we have required that $A_0$ be a Fredholm operator, in order for equations of the form $A_0 u = \chi$ to have at least one solution $u \in L^2(\mathbb{R}, F_Y)$ the function $\chi$ must satisfy $(1, \chi) = \int \chi dF_Y =: \langle \chi \rangle = 0$, which is precisely the centering condition given in (3.10).

C Proof of accuracy

Before establishing our main accuracy result – Theorem 3.2 – we shall need the following lemma.

Lemma C.1. Suppose $J(y)$ is at most polynomially growing. Then, for every $y$ and $s < t$, there exists a positive constant $C < \infty$ such that for any $\varepsilon \leq 1$, we have the following inequality

$$\mathbb{E}_y [|J(Y_s)|] \leq C.$$  

Proof of Lemma C.1. It is enough to prove the result for $J(y) = y^k$ for any $k \in \mathbb{N}$. To begin, we define a time-rescaled process via $Y^1_t := Y_{\varepsilon t}$. The process $Y^1$ has infinitesimal generator $A^1_0 = A_0$ under $\mathbb{P}$. By the assumptions of section 2, the process $Y^1$ is ergodic and has a unique invariant distribution $F_Y$. The operator $A_0$ has a strictly positive spectral gap and is self-adjoint on the Hilbert space $L^2(\mathbb{R}, F_Y)$. These properties ensure (see Fouque, Papanicolaou, Sircar, and Solna (2011), p. 139) that for all $k \in \mathbb{N}$ there exists a constant $C(k) < \infty$ such that

(C.1) $$\sup_t \mathbb{E} [|Y^1_t|^k] \leq C(k).$$

Hence, under the physical measure $\mathbb{P}$ we have

$$\mathbb{E} [|Y_s|^k] = \mathbb{E} \left[ |Y^{(1)}_{s/\varepsilon}|^k \right] \leq \sup_{\varepsilon \leq 1} \mathbb{E} \left[ |Y^{(1)}_{s/\varepsilon}|^k \right] \leq C(k),$$

by (C.1). Now we define an exponential martingale $M^{(A)}$, which relates the dynamics of $Y$ under the risk-neutral measure $\mathbb{P}$ to its dynamics under the physical measure $\mathbb{P}$. We have

$$M^{(A)}_t := \exp \left( -\int_0^t \lambda(Y_s) \, dB_s - \frac{1}{2} \int_0^t \lambda^2(Y_s) \, ds \right) = \frac{d\mathbb{P}}{d\mathbb{P}|_{Y_t}}.$$
The $\mathbb{P}$-expectation of $|Y_s|^k$ can be found as follows:

\[
\mathbb{E}\left[|Y_s|^k\right] = E\left[|Y_s|^k M_s^{(2\Lambda)}\right]
\leq \left(E\left[|Y_s|^{2k} \exp\left(\int_0^s \Lambda^2(Y_u)\,du\right)\right]\right)^{1/2} \left(E\left[M_s^{(2\Lambda)}\right]\right)^{1/2} \quad \text{(by Cauchy-Schwarz)}
\leq \left(E\left[|Y_s|^{2k} \exp\left(\int_0^s \Lambda^2(Y_u)\,du\right)\right]\right)^{1/2} \left(E\left[M_s^{(2\Lambda)}\right]\right)^{1/2} \quad \text{($M^{(2\Lambda)}$ is a $\mathbb{P}$-martingale)}
\leq \left(E\left[|Y_s|^{1/2} \exp\left(s ||\Lambda||_\infty^2\right)\right]\right)^{1/2} \leq C,
\]

where we have used the fact that $\Lambda$ is bounded in the last line.

We are now in a position to prove Theorem 3.2. We begin by defining a remainder term $R^\varepsilon$ by

\[ u^\varepsilon = u_0 + \varepsilon u_1 + \varepsilon^2 u_2 + \varepsilon^3 u_3 + R^\varepsilon. \]

From the equations satisfied by $u_0$, $u_1$, $u_2$ and $u_3$ in section 3.1, one derives

\[ (-\partial_t + A^\varepsilon) R^\varepsilon = -\varepsilon^2 F^\varepsilon, \quad R^\varepsilon(0, x, y) = -\varepsilon^2 G^\varepsilon(x, y). \]

where we have defined

\[ F^\varepsilon(t, x, y) := (A_1 u_3 + A_2 u_2) + \varepsilon A_2 u_3, \quad G^\varepsilon(x, y) := u_2(0, x, y) + \varepsilon u_3(0, x, y). \]

Thus, $R^\varepsilon$ has the following probabilistic representation

\[ R^\varepsilon(t, x, y) = \varepsilon^2 \mathbb{E}_{x,y} \left[ -G^\varepsilon(X_1, Y_1) + \int_0^t F^\varepsilon(s, X_s, Y_s)\,ds \right]. \]

From the assumptions of section 2 and the boundedness of the option payoff, one can deduce that the functions $F^\varepsilon$ and $G^\varepsilon$ are bounded in $x$ and at most polynomially growing in $y$ (see Fouque, Papanicolaou, Sircar, and Solna (2011)). Hence, by Lemma C.1 there exists a constant $C_1 < \infty$ such that

\[ |R^\varepsilon| \leq \varepsilon^2 C_1. \]

Finally

\[ |u^\varepsilon,\delta - (u_0 + \varepsilon u_1)| \leq |R^\varepsilon| + |\varepsilon^2 u_{2,0} + \varepsilon^3 u_{3,0}| \leq \varepsilon^2 C_1 + \varepsilon^2 |u_{2,0} + \varepsilon u_{3,0}| \leq \varepsilon^2 C, \]

for some constant $C < \infty$. This concludes the proof of accuracy.
References


Figure 1: Using the model described in section 4, we plot the implied volatility induced by the price of European call option as a function of the strike price $K$. In each plot, the dashed blue line corresponds to the implied volatility induced by the full price $u^\varepsilon$ (computed via Monte Carlo simulation) and the solid green line corresponds to the implied volatility induced by our approximation $u_0 + \varepsilon u_1$. For all plots we use the following parameter values: $t = 1/10$, $e^\varepsilon = 50$, $m = -0.2$, $s = 0.2$, $\rho = -0.7$, $a = 0.2$, $b = 1.5$, $\beta = 1.0$ and $\Lambda = 0.25$. 