7

Stochastic Volatility

Modeling and Asymptotic Approaches to Option Pricing and Portfolio Selection

Matthew Lorig[1] and Ronnie Sircar[2]
[1] *University of Washington, USA*
[2] *Princeton University,USA*

7.1 Introduction

Understanding and measuring the inherent uncertainty in market volatility are crucial for portfolio optimization, risk management, and derivatives trading. The problem is made difficult since volatility is not directly observed. Rather, volatility is a statistic of the observable returns of, for example, a stock, and so estimates of it are at best noisy. Among the major empirical challenges have been separating contributions of diffusive and jump components of log returns, typical timescales of fluctuation, and memory effects. Until recently, data were limited to low frequencies, typically daily. The availability of high-frequency data over the past 20 years brings with it issues of deciphering market microstructure effects such as the bid–ask bounce, which contaminate the potential usefulness of such large datasets, and we refer to the recent book by Aït-Sahalia and Jacod (2014) for an overview of the difficulties.

The major problem that has been the driver of *stochastic volatility* models is the valuation and hedging of derivative securities. This market grew in large part from the landmark paper by Black and Scholes (1973), which showed how to value simple options contracts when volatility is constant. Even at the time of their paper, Black and Scholes realized that the constant volatility assumption was a strong idealization. In an empirical paper by Black and Scholes (1972), the authors tested their option price formulas and concluded: "we found that using past data to estimate the variance caused the model to overprice options on high-variance stocks and underprice options on low-variance stocks." Indeed the overwhelming evidence from time-series data reveals that volatility exhibits unpredictable variation. In addition, as we will describe

Financial Signal Processing and Machine Learning, First Edition.
Edited by Ali N. Akansu, Sanjeev R. Kulkarni and Dmitry Malioutov.

in this chapter, option prices exhibit a significant departure from the Black–Scholes (constant volatility) theory, the *implied volatility skew*, which can be explained by allowing volatility to vary randomly in time.

A second important problem is portfolio optimization: namely, how to optimally invest capital between a risky stock and a riskless bank account. In a continuous time stochastic model with constant volatility, the pioneering work was by Robert Merton (Merton, 1969, 1971; reprinted in Merton, 1992). Since then, understanding the effect of volatility uncertainty, stochastic growth rate, transaction costs, price impact, illiquidity, and other frictions on the portfolio choice problem has generated considerable research. Here, we will focus on the effect of stochastic volatility and present some new results in Section 7.3.

The increased realism obtained by allowing volatility to be stochastic comes with increased computational difficulties. Moreover, there is no broad consensus concerning how to best model volatility. Here, we will discuss some computationally efficient approaches, focusing particularly on asymptotic approximations.

7.1.1 Options and Implied Volatility

The most liquidly traded derivatives contracts are call and put options, which give the option holder the right to buy (in the case of a call) or sell (in the case of a put) one unit of the underlying security at a *fixed strike price* K on a *fixed expiration date* T. Here we are focusing specifically on *European*-style options (i.e., no early exercise), which are typically traded on indices such as the S&P 500. If S_t represents the price of a stock or index at time t, then a European-style derivative has a payoff at time T, which is a function h of S_T. In the case of calls and puts, the payoff functions are $h(S) = (S - K)^+$ and $h(S) = (K - S)^+$, respectively.

7.1.1.1 Black–Scholes model

In the Black–Scholes model, the stock price S is a geometric Brownian motion described by the following stochastic differential equation (SDE):

$$\frac{\mathrm{d}S_t}{S_t} = \mu \, \mathrm{d}t + \sigma \, \mathrm{d}W_t, \tag{7.1}$$

where W is a standard Brownian motion with respect to a historical (or real-world, or physical) probability measure $\mathbb{P}$. Here, the parameters are the expected growth rate μ and the volatility σ, both assumed constant. The remarkable finding of Black and Scholes (1973) is that the no-arbitrage price of an option does not depend on μ, and so, to price an option, the only parameter that needs to be estimated from data is the volatility σ. Unless otherwise stated, we shall assume throughout this article that interest rates are zero.

It will be convenient to introduce the following notation:

$$\tau := T - t, \qquad x := \log S_t, \qquad k := \log K,$$

where t is the current time; and K and T are the strike and expiration date, respectively, of a call or put option. Then, for fixed (t, T, x, k), the Black–Scholes pricing formula for a call

option with time to expiration $\tau > 0$ is given by

$$u^{\mathrm{BS}}(\sigma) := \mathrm{e}^{x}\mathcal{N}(d_{+}(\sigma)) - \mathrm{e}^{k}\mathcal{N}(d_{-}(\sigma)), \qquad d_{\pm}(\sigma) := \frac{1}{\sigma\sqrt{\tau}}\left(x - k \pm \frac{\sigma^2\tau}{2}\right), \tag{7.2}$$

where $\mathcal{N}$ is the CDF of a standard normal random variable, and we have stressed the volatility argument σ in the notation.

It turns out that the Black–Scholes price (7.2) can be expressed as the expected payoff of the option, but where the expectation is taken with respect to a different probability measure $\mathbb{Q}$ under which the stock price is a martingale (that is, it is a pure fluctuation process with no trend or growth rate). This means that there is a so-called risk-neutral world in which the stock price follows the dynamics

$$\frac{\mathrm{d}S_t}{S_t} = \sigma\ \mathrm{d}W_t^{\mathbb{Q}},$$

where $W^{\mathbb{Q}}$ is a Brownian motion under $\mathbb{Q}$, and the call option price (7.2) can be expressed as the conditional expectation

$$u^{\mathrm{BS}}(\sigma) = \mathbb{E}^{\mathbb{Q}}[(S_T - K)^{+} \mid \log\ S_t = x],$$

where $\mathbb{E}^{\mathbb{Q}}$ denotes that the expectation is taken under the probability measure $\mathbb{Q}$.

7.1.1.2 Implied Volatility

The *implied volatility* of a given call option with price u (which is either observed in the market or computed from a model) is the unique positive solution I of

$$u^{\mathrm{BS}}(I) = u. \tag{7.3}$$

It is the volatility parameter that has to be put into the Black–Scholes formula to match the observed price u.

Note that the implied volatility I depends implicitly on the maturity date T and the log strike k as the option price u will depend on these quantities. The map $(T, k) \mapsto I(T, k)$ is known as the *implied volatility surface*. If market option prices reflected Black–Scholes assumptions, I would be constant and equal to the stock's historical volatility σ. However, in equities data, the function $I(T, \cdot)$ exhibits downward-sloping behavior in k, whose slope varies with the option maturities T, as illustrated in Figure 7.1. This downward slope is known as the *implied volatility skew*.

These features of the implied volatility surface can be reproduced by enhancing the Black–Scholes model (7.1) with stochastic volatility and/or jumps. One focus of this chapter will be to survey some approaches taken to capturing the implied volatility skew.

7.1.2 Volatility Modeling

While the overwhelming evidence from time-series and option price data indicates that the volatility σ in (7.1) should be allowed to vary stochastically in time:

$$\frac{\mathrm{d}S_t}{S_t} = \mu\ \mathrm{d}t + \sigma_t\ \mathrm{d}W_t,$$

there is no consensus as to how exactly the (stochastic) volatility σ_t should be modeled.

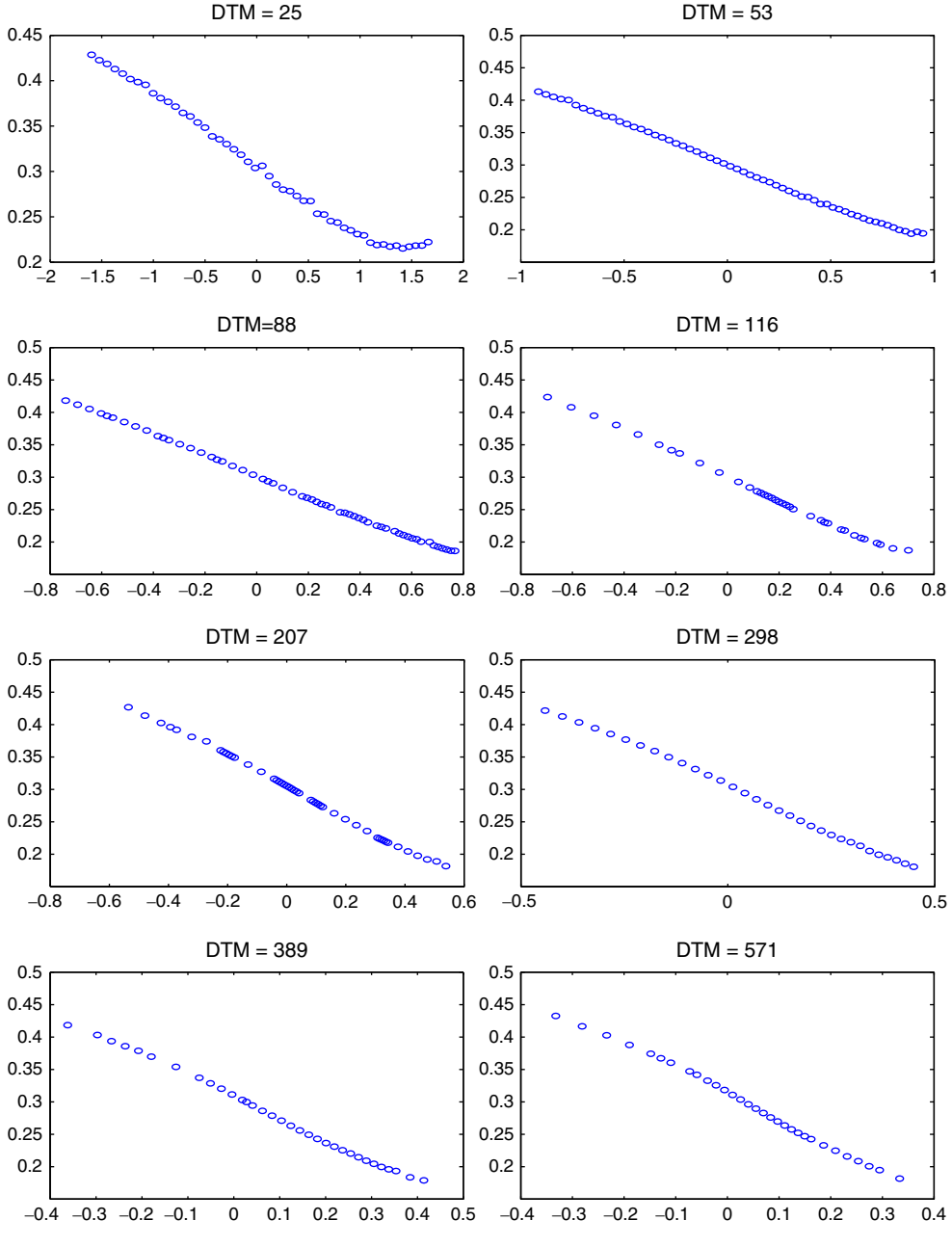

Figure 7.1 Implied volatility from S&P 500 index options on May 25, 2010, plotted as a function of log-moneyness to maturity ratio: $(k - x)/(T - t)$. DTM, days to maturity.

7.1.2.1 No-arbitrage Pricing and Risk-Neutral Measure

In standard option-pricing theory, it is assumed that markets do not admit arbitrage. No arbitrage pricing implies that all traded asset prices (after discounting) are martingales under some probability measure $\mathbb{Q}$, typically referred to as a *risk-neutral* measure. Consequently, the price u_t of an option at time t with payoff $\varphi(S_T)$ at time T is given by

$$u_t = \mathbb{E}^{\mathbb{Q}}[\varphi(S_T)|\mathcal{F}_t], \tag{7.4}$$

where $\mathcal{F}_t$ is the history of the market up to time t. Typically there are nontraded sources of randomness such as jumps or stochastic volatility. As a result, there exist infinitely many risk-neutral measures. The nonuniqueness of these measures is often referred to as *market incompleteness*, meaning not every derivative asset can be perfectly hedged. In practice, one assumes that the market has chosen a specific risk-neutral measure, which is consistent with observed option prices.

In what follows, we will model asset dynamics under a unique risk-neutral pricing measure $\mathbb{Q}$, which we assume has been chosen by the market. Under $\mathbb{Q}$, we have

$$S_t = \mathrm{e}^{X_t}, \;\; \mathrm{d}X_t = -\frac{1}{2}\sigma_t^2 \mathrm{d}t + \sigma_t \mathrm{d}W_t^{\mathbb{Q}}, \tag{7.5}$$

which describes the dynamics of $X_t = \log\, S_t$. In the rest of this section, we review some of the most common models of volatility and discuss some of their advantages and disadvantages.

7.1.2.2 Local Volatility Models

In *local volatility* (LV) models, the volatility σ_t of the underlying is modeled as a deterministic function $\sigma(\cdot, \cdot)$ of time t, and the time-t value of the underlying X_t. That is,

$$\mathrm{d}X_t = -\tfrac{1}{2}\sigma^2(t, X_t)\, \mathrm{d}t + \sigma(t, X_t)\mathrm{d}W_t^{\mathbb{Q}}, \quad \text{(local volatility)}$$

Typically, one assumes that the function $\sigma(t, \cdot)$ increases as x decreases in order to capture the *leverage effect*, which refers to the tendency for the value of an asset to decrease as its volatility increases.

One advantage of local volatility models is that markets remain *complete*, meaning that derivatives written on S can be hedged perfectly – just as in the Black–Scholes model. While market completeness is convenient from a theoretical point of view, it is not necessarily a realistic property of financial markets. Indeed, if markets are complete, then one can ask: why do we need derivatives?

Another advantage of local volatility models is that they can provide a very tight fit to option prices quoted on the market. In fact, Dupire (1994) shows that there exists a local volatility model that can *exactly* match option prices quoted on the market and that there is an explicit formula for how to construct this model from observed call and put prices under the assumption that they can be interpolated across continuous strikes and maturities. However, a tight fit must be balanced with stability: local volatility models are notoriously bad at providing stability and typically need to be recalibrated hourly.

7.1.2.3 Stochastic Volatility Models

In a *stochastic volatility* (SV) model, promoted in the late 1980s by Hull and White (1987), Scott (1987), and Wiggins (1987), the volatility σ_t of the underlying is modeled as a deterministic function $\sigma(\cdot)$ of some auxiliary process Y, which is usually modeled as a diffusion:

$$\begin{aligned} \mathrm{d}X_t &= -\frac{1}{2}\sigma^2(Y_t)\,\mathrm{d}t + \sigma(Y_t)\,\mathrm{d}W_t^{\mathbb{Q}}, \\ \mathrm{d}Y_t &= \alpha(Y_t)\,\mathrm{d}t + \beta(Y_t)\,\mathrm{d}B_t^{\mathbb{Q}}, \qquad \text{(stochastic volatility)} \\ \mathrm{d}\langle W^{\mathbb{Q}}, B^{\mathbb{Q}}\rangle_t &= \rho\,\mathrm{d}t, \end{aligned} \tag{7.6}$$

with $|\rho| < 1$. Here, $B^{\mathbb{Q}}$ is a Brownian motion that is correlated with $W^{\mathbb{Q}}$. One typically takes the correlation ρ to be negative in order to capture the empirical observation that when volatility goes up, stock prices tend to go down, which is called the *leverage effect*. In a single-factor stochastic volatility setting such as that described by (7.6), derivatives written on S cannot be perfectly hedged by continuously trading a bond and the underlying S alone. However, a derivative written on S can be perfectly replicated by continuously trading a bond, the underlying S, and a single option on S. Thus, assuming options can be traded continuously can *complete the market*. However, as transaction costs on options are much higher than on stocks, and as their liquidity is typically lower, this assumption typically is not made. Unlike the local volatility case, there is no explicit formula for constructing Y dynamics and a volatility function $\sigma(\cdot)$ so that model-induced option prices fit observed market prices exactly.

It is common to assume that the volatility driving process Y is mean-reverting, or *ergodic*, meaning there exists a distribution Π such that the ergodic theorem holds:

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t g(Y_s)\mathrm{d}s = \int g(y)\Pi(\mathrm{d}y),$$

for all bounded functions g.

Equation (7.6) actually refers specifically to *one-factor* stochastic volatility models. One can always introduce another auxiliary process Z, and model the volatility σ_t as a function $\sigma(\cdot,\cdot)$ of both Y and Z. If S, Y, and Z are driven by three distinct Brownian motions, then continuously trading a bond, the underlying S, and two options on S would be required to perfectly hedge further options on S. *Multifactor* stochastic volatility models have the ability of fit option prices better than their one-factor counterparts. But, each additional factor of volatility brings with it additional computational challenges. Multifactor and multiscale stochastic volatility models are discussed at length in Fouque *et al*. (2011).

7.1.2.4 Local-Stochastic Volatility Models

As the name suggests, *local-stochastic volatility* (LSV) models combine features of both local volatility and stochastic volatility models by modeling the volatility σ_t as a function $\sigma(\cdot,\cdot,\cdot)$ of time t, the underlying X, and an auxiliary process Y (possibly multidimensional). For example,

$$\begin{aligned} \mathrm{d}X_t &= -\frac{1}{2}\sigma^2(t, X_t, Y_t)\,\mathrm{d}t + \sigma(t, X_t, Y_t)\,\mathrm{d}W_t^{\mathbb{Q}}, \\ \mathrm{d}Y_t &= f(t, X_t, Y_t)\,\mathrm{d}t + \beta(t, X_t, Y_t)\,\mathrm{d}B_t^{\mathbb{Q}}, \qquad \text{(local-stochastic volatility)} \\ \mathrm{d}\langle W^{\mathbb{Q}}, B^{\mathbb{Q}}\rangle_t &= \rho\,\mathrm{d}t, \end{aligned} \tag{7.7}$$

where $|\rho| < 1$. Note that the class of models described by (7.7) nests all LV models and all (one-factor) SV models. However, while LSV models offer more modeling flexibility than LV or SV models separately, these models also present new computational challenges. Indeed, while there exist LV and SV models for which option prices can be computed in closed form (or semiclosed form, e.g., up to Fourier inversion), explicit formulas for option prices are available in an LSV setting only when $\rho = 0$. We remark that having a closed-form formula or a fast approximation is crucial for the inverse problem of calibrating an LSV model from observed option prices.

7.1.2.5 Models with Jumps

Some authors argue that diffusion models of the form (7.5) are not adequate to capture the complex dynamics of stock price processes because diffusion models do not allow stock prices to jump. Discontinuities in the stock price process can be modeled by adding a jump term $\mathrm{d}J_t$ to the process (7.5) as follows:

$$\mathrm{d}X_t = \left(-\frac{1}{2}\sigma_t^2 - \lambda \int (\mathrm{e}^z - 1)F(\mathrm{d}z)\right)\mathrm{d}t + \sigma \mathrm{d}W_t^{\mathbb{Q}} + \mathrm{d}J_t,$$

and this type of model dates back to Merton (1976). Here, jumps arrive as a Poisson process with intensity λ and have distribution F, and the drift of X is compensated to ensure that $S = \mathrm{e}^X$ is a martingale under $\mathbb{Q}$. As with SV models, jumps render a market incomplete. Adding jumps also to volatility can help fit the strong implied volatility smile that is commonly observed for short maturity options, and we refer to Bakshi *et al*. (1997) for an analysis.

7.1.2.6 ARCH and GARCH Models

Although our focus will be on the continuous-time models, it is worth mentioning that discrete-time models for stock returns are widely studied in econometrics literature. A large class of discrete-time models are the autoregressive conditional heteroscedasticity (ARCH) processes introduced by Engle (1982), later generalized under the name GARCH. The discrete-time models that are closest to the type of continuous-time stochastic volatility models (7.6) driven by diffusions are the EGARCH models developed in Nelson (1991, 1990). Those papers also discuss convergence of the discrete-time EGARCH process to an exponential Ornstein–Uhlenbeck continuous-time stochastic volatility model.

7.2 Asymptotic Regimes and Approximations

Given a model and its parameters, computing expectations of the form (7.4) one time is straightforward using Monte Carlo methods or a numerical solution of the associated pricing partial integro-differential equation (PIDE). However, when the computation is part of an iterative procedure to calibrate the model to the observed implied volatility surface, it becomes important to have a fast method of computing option prices or model-induced implied volatilities. As a result, a number of different efficient approximation methods have developed.

Broadly speaking, there are two methods of setting up asymptotic expansions for option pricing and implied volatility. In *contract asymptotics*, one considers extreme regimes specific

to the option contract, in other words, large or small time-to-maturity ($T - t$), or large or small strikes K. In the *model asymptotics* approach, one views the complicated incomplete market model as a perturbation around a more tractable model, often the Black–Scholes model (7.1).

7.2.1 Contract Asymptotics

The foundational papers in this approach appeared in 2004. The main result from Lee (2004), commonly referred to as the *moment formula*, relates the implied volatility slope at extreme strikes to the largest finite moment of the stock price. We refer to the recent book by Gulisashvili (2012) for an overview of this and related asymptotics.

The approach pioneered by Berestycki *et al*. (2004) uses large-deviation calculations for the short-time regime. The regime where the time-to-maturity is large is studied by Tehranchi (2009). However, for the rest of this chapter, we concentrate on model asymptotics as they are adaptable to other option contracts and, moreover, are amenable to nonlinear portfolio optimization problems, as we discuss in Section 7.3.

7.2.2 Model Asymptotics

We will present the analysis in terms of the log-stock price $X_t = \log S_t$, and consider a European option with payoff $\varphi(X_T)$ at time T, where $\varphi(x) = h(e^x)$. There may be other factors such as stochastic volatility driving the stock price dynamics, and we denote these by the (possibly multi-dimensional) process Y. If (X, Y) is (jointly) a Markov process, then the time t price $u(t, x, y)$ of the European option is an expectation of the form

$$u(t, x, y) = \mathbb{E}^{\mathbb{Q}}[\varphi(X_T)|X_t = x, Y_t = y].$$

Here, we are using the Markov property of (X, Y) to replace the filtration $\mathcal{F}_t$ in (7.4) with the time-t values of (X, Y).

Under mild conditions on the processes (X, Y) and the payoff function φ, the function $u(t, x, y)$ is sufficiently smooth to be the solution of the *Kolmogorov backward equation* (KBE)

$$(\partial_t + \mathcal{A}(t))u = 0, \qquad u(T, x) = \varphi(x), \tag{7.8}$$

where the operator $\mathcal{A}(t)$ is the *generator* of (X, Y) (which may have t-dependence from the coefficients of (X, Y)). The operator $\mathcal{A}(t)$ is, in general, a second-order partial integro-differential operator. Unfortunately, equation (7.8) rarely has a closed-form solution – especially when we include realistic features such as jumps and stochastic volatility. As such, one typically seeks an approximate solution to (7.8). We will discuss an approach to this using *perturbation theory* (also referred to as asymptotic analysis).

Perturbation theory is a classical tool developed to solve problems arising in physics and engineering. We describe its use here to find an approximate solution to (typically) a PIDE starting from the exact solution of a related PIDE. More specifically, suppose the integro-differential operator $\mathcal{A}(t)$ in (7.8) can be written in the form

$$\mathcal{A}(t) = \sum_{n=0}^{\infty} \varepsilon^n \mathcal{A}_n(t), \tag{7.9}$$

where each $\mathcal{A}_n(t)$ in the sequence $(\mathcal{A}_n(t))$ is an integro-differential operators, and $\varepsilon > 0$ is a (typically small) parameter. Formally, one seeks an approximate solution to (7.8) by expanding the function u as a power series in the parameter ε

$$u = \sum_{n=0}^{\infty} \varepsilon^n u_n. \tag{7.10}$$

Inserting the expansion (7.9) for $\mathcal{A}(t)$ and the expansion (7.10) for u into the PIDE (7.8), and collecting like powers of ε, one finds

$$\begin{aligned}
\mathcal{O}(1) &: \quad (\partial_t + \mathcal{A}_0(t))u_0 = 0, && u_0(T,x,y) = \varphi(x),\\
\mathcal{O}(\varepsilon) &: \quad (\partial_t + \mathcal{A}_0(t))u_1 = -\mathcal{A}_1(t)u_0, && u_n(T,x,y) = 0,\\
&\vdots \qquad\qquad \vdots && \vdots\\
\mathcal{O}(\varepsilon^n) &: \quad (\partial_t + \mathcal{A}_0(t))u_n = -\sum_{k=1}^{\infty} \mathcal{A}_k(t)u_{n-k}, && u_n(T,x,y) = 0.
\end{aligned}$$

The approximating sequence of functions (u_n) is then found by solving the above nested sequence of PIDEs.

This method is most useful when the *fundamental solution* Γ_0 (also referred to as *Green's function*), corresponding to the operator $\mathcal{A}_0(t)$, is available in closed form. It is the solution of

$$(\partial_t + \mathcal{A}_0(t))\Gamma_0(t,x,y;T,\xi,\omega) = 0, \qquad \Gamma_0(T,x,y;T,\xi,\omega) = \delta(x-\xi)\delta(y-\omega),$$

where $\delta(\cdot)$ is the Dirac delta function (or point mass at zero).

Upon finding Γ_0, the approximating sequence of functions (u_n) can be written down directly:

$$u_0(t,x,y) = \mathcal{P}_0(t,T)\varphi(x) := \int \mathrm{d}\xi \mathrm{d}\omega\, \Gamma_0(t,x,y;T,\xi,\omega)\varphi(\xi), \tag{7.11}$$

$$u_n(t,x,y) = \int_t^T \mathrm{d}t_1 \mathcal{P}_0(t,t_1) \sum_{k=1}^{n} \mathcal{A}_k u_{n-k}(t_1,x,y). \tag{7.12}$$

where the operator $\mathcal{P}_0(t,T)$ is referred to as the *semigroup* generated by $\mathcal{A}_0(t)$.

Finding an appropriate decomposition of the generator $\mathcal{A}(t) = \sum_{n=0}^{\infty} \mathcal{A}_n(t)$ is a bit of an art. In general, the most appropriate decomposition will depend strongly on the underlying process X (from which $\mathcal{A}(t)$ is derived). As a starting point, it will help to identify operators $\mathcal{A}_0$ for which the fundamental solution Γ_0 can be written in closed form or semiclosed form, and we shall discuss some examples in Section 7.2.4.

7.2.3 Implied Volatility Asymptotics

Models are typically calibrated to implied volatilities rather than to prices directly. As such, it is useful to have closed-form approximations for model-induced implied volatilities. In this section, we will show how to translate an expansion for option prices into an expansion for implied volatilities. Throughout this section, we fix a model for $X = \log S$, a time t, a maturity date $T > t$, the initial values $X_t = x$, and a call option payoff $\varphi(X_T) = (\mathrm{e}^{X_T} - \mathrm{e}^k)^+$. Our goal

is to find the implied volatility for *this particular call option*. To ease notation, we will suppress much of the dependence on (t, T, x, k). However, the reader should keep in mind that the implied volatility of the option under consideration *does* depend on (t, T, x, k), even if this is not explicitly indicated.

Assume that the option price u has an expansion of the form

$$u = u_0 + \sum_{n=1} \varepsilon^n u_n, \qquad \text{where} \qquad u_0 = u^{\mathrm{BS}}(\sigma_0), \text{ for some } \sigma_0 > 0. \tag{7.13}$$

We wish to find the implied volatility I corresponding to u, which is the unique positive solution of (7.3). To find the unknown implied volatility I, we expand it in powers of ε as follows:

$$I = I_0 + E^\varepsilon, \text{where} \quad E^\varepsilon = \sum_{n=1}^{\infty} \varepsilon^n I_n.$$

Expanding $u^{\mathrm{BS}}(I)$ about the point I_0, we find

$$\begin{aligned} u^{\mathrm{BS}}(I) &= u^{\mathrm{BS}}(I_0 + E^\varepsilon) \\ &= u^{\mathrm{BS}}(I_0) + \sum_{n=1}^{\infty} \frac{(E^\varepsilon)^n}{n!} \partial_\sigma u^{\mathrm{BS}}(I_0) \\ &= u^{\mathrm{BS}}(I_0) + \varepsilon I_1 \partial_\sigma u^{\mathrm{BS}}(I_0) + \varepsilon^2 \left(I_2 \partial_\sigma + \frac{1}{2!} I_1^2 \partial_\sigma^2 \right) u^{\mathrm{BS}}(I_0) \\ &\quad + \varepsilon^3 \left(I_3 \partial_\sigma + \frac{1}{2!} 2 I_1 I_2 \partial_\sigma^2 + \frac{1}{3!} I_1^3 \partial_\sigma^3 \right) u^{\mathrm{BS}}(I_0) + \cdots . \end{aligned} \tag{7.14}$$

Inserting the expansion (7.13) for u and the expansion (7.14) for $u^{\mathrm{BS}}(I)$ into equation (7.3), and collecting like powers of ε, we obtain

$$\begin{aligned} \mathcal{O}(1) : &\qquad u_0 = u^{\mathrm{BS}}(I_0), \\ \mathcal{O}(\varepsilon) : &\qquad u_1 = I_1 \partial_\sigma u^{\mathrm{BS}}(I_0), \\ \mathcal{O}(\varepsilon^2) : &\qquad u_2 = \left(I_2 \partial_\sigma + \frac{1}{2!} I_1^2 \partial_\sigma^2 \right) u^{\mathrm{BS}}(I_0) \\ \mathcal{O}(\varepsilon^3) : &\qquad u_3 = \left(I_3 \partial_\sigma + \frac{1}{2!} 2 I_1 I_2 \partial_\sigma^2 + \frac{1}{3!} I_1^3 \partial_\sigma^3 \right) u^{\mathrm{BS}}(I_0). \end{aligned}$$

Using $u_0 = u^{\mathrm{BS}}(\sigma_0)$, we can solve for the sequence (I_n) recursively. We have

$$\mathcal{O}(1) : \qquad I_0 = \sigma_0, \tag{7.15}$$

$$\begin{aligned} \mathcal{O}(\varepsilon) : &\qquad I_1 = \frac{1}{\partial_\sigma u^{\mathrm{BS}}(I_0)} u_1, \\ \mathcal{O}(\varepsilon^2) : &\qquad I_2 = \frac{1}{\partial_\sigma u^{\mathrm{BS}}(I_0)} \left(u_2 - \left(\frac{1}{2!} I_1^2 \partial_\sigma^2 \right) u^{\mathrm{BS}}(I_0) \right) \\ \mathcal{O}(\varepsilon^3) : &\qquad I_3 = \frac{1}{\partial_\sigma u^{\mathrm{BS}}(I_0)} \left(u_3 - \left(\frac{1}{2!} 2 I_1 I_2 \partial_\sigma^2 + \frac{1}{3!} I_1^3 \partial_\sigma^3 \right) u^{\mathrm{BS}}(I_0) \right). \end{aligned} \tag{7.16}$$

Note that the implied volatility expansion involves both *model-independent* and *model-dependent* terms:

$$\text{model independent:} \quad \frac{\partial_\sigma^n u^{\mathrm{BS}}(I_0)}{\partial_\sigma u^{\mathrm{BS}}(I_0)} \qquad \text{model dependent:} \quad \frac{u_n}{\partial_\sigma u^{\mathrm{BS}}(I_0)}. \tag{7.17}$$

The model-independent terms are always explicit. For example, using (7.2), a direct computation reveals

$$\frac{\partial_\sigma^2 u^{\mathrm{BS}}(\sigma)}{\partial_\sigma u^{\mathrm{BS}}(\sigma)} = \frac{(k-x)^2}{t\sigma^3} - \frac{t\sigma}{4}, \frac{\partial_\sigma^3 u^{\mathrm{BS}}(\sigma)}{\partial_\sigma u^{\mathrm{BS}}(\sigma)} = -\frac{t}{4} + \frac{(k-x)^4}{t^2\sigma^6} - \frac{3(k-x)^2}{t\sigma^4} - \frac{(k-x)^2}{2\sigma^2} + \frac{t^2\sigma^2}{16}.$$

Higher order terms are also explicit. Whether or not the model-dependent terms in (7.17) can be computed explicitly (i.e., meaning without numerical integration or special functions) depends on the specific form the sequence (u_n) takes.

7.2.4 Tractable Models

We say that a model X is *tractable* if its generator $\mathcal{A}_0(t)$ admits a closed-form (or semiclosed form) fundamental solution Γ_0. A large class of tractable models are the exponential Lévy models. In this class, a traded asset $S = \mathrm{e}^X$ is described by the following Lévy–Itô SDE:

$$\mathrm{d}X_t = \mu\ \mathrm{d}t + \sigma \mathrm{d}W_t^{\mathbb{Q}} + \int_{\mathbb{R}} z\ \mathrm{d}\widetilde{N}_t(\mathrm{d}z). \tag{7.18}$$

Here, $W^{\mathbb{Q}}$ is a standard Brownian motion and $\widetilde{N}$ is an independent compensated Poisson random measure:

$$\mathrm{d}\widetilde{N}_t(\mathrm{d}z) = \mathrm{d}N_t(\mathrm{d}z) - \nu(\mathrm{d}z)\mathrm{d}t.$$

The last term in (7.18) can be understood as follows: for any Borel set A, the process $N(A)$ is a Poisson process with intensity $\nu(A)$. Thus, the probability that X experiences a jump of size $z \in A$ in the time interval $[t, t + \mathrm{d}t)$ is $\nu(A)\mathrm{d}t$. In order for S to be a martingale, the drift μ must be given by

$$\mu = -\frac{1}{2}\sigma^2 - \int_{\mathbb{R}} \nu(\mathrm{d}z)(\mathrm{e}^z - 1 - z).$$

The generator $\mathcal{A}_0$ of X is given by

$$\mathcal{A}_0 = \mu\partial_x + \frac{1}{2}\sigma^2\partial_{xx}^2 + \int \nu(\mathrm{d}z)(\theta_z - 1 - z\partial_x), \tag{7.19}$$

where the operator θ_z is a *shift operator*: $\theta_z f(x) = f(x + z)$.

When X has no jump component (i.e., when $\nu \equiv 0$), the generator $\mathcal{A}_0$ has a fundamental solution Γ_0 that can be written in closed form as a Gaussian kernel:

$$\Gamma_0(t, x; T, y) = \frac{1}{\sqrt{2\pi\sigma^2(T-t)}} \exp\left(-\frac{(x - y + \mu(T-t))^2}{2\sigma^2(T-t)}\right).$$

More generally, when jumps are present, the generator $\mathcal{A}_0$ has a fundamental solution Γ_0, which is available in semiclosed form as a Fourier transform:

$$\Gamma_0(t,x;T,y) = \frac{1}{2\pi}\int \mathrm{d}\xi\ \mathrm{e}^{\mathrm{i}\xi(x-y)+(T-t)\Phi_0(\xi)}, \tag{7.20}$$

where

$$\Phi_0(\xi) = \mathrm{i}\mu\xi - \frac{1}{2}\sigma^2\xi^2 + \int \nu(\mathrm{d}z)(\mathrm{e}^{\mathrm{i}\xi z} - 1 - \mathrm{i}\xi z).$$

The function Φ_0 is referred to as the *characteristic exponent* of X, since it satisfies

$$\mathbb{E}^{\mathbb{Q}}[\mathrm{e}^{\mathrm{i}\xi X_T}|X_t = x] = \mathrm{e}^{\mathrm{i}\xi x+(T-t)\Phi_0(\xi)}.$$

Using (7.11) and (7.20), we can express the action of the semigroup operator $\mathcal{P}_0(t,T)$ on a general function ψ as follows:

$$\begin{aligned}\mathcal{P}_0(t,T)\psi(x) &= \int \mathrm{d}y\ \Gamma_0(t,x;T,y)\psi(y) = \frac{1}{2\pi}\int \mathrm{d}y\ \int \mathrm{d}\xi\ \mathrm{e}^{i\xi(x-y)+(T-t)\Phi_0(\xi)}\psi(y)\\ &= \frac{1}{2\pi}\int \mathrm{d}\xi\ \mathrm{e}^{i\xi x+(T-t)\Phi_0(\xi)}\hat{\psi}(\xi),\end{aligned}$$

where $\hat{\psi}$ is the Fourier transform of ψ:

$$\hat{\psi}(\xi) := \int \mathrm{d}y\ \mathrm{e}^{-\mathrm{i}\xi y}\psi(y).$$

Because the semigroup operator $\mathcal{P}_0(t,T)$ corresponding to $\mathcal{A}_0$ is well understood in the exponential Lévy setting, if one can write the generator $\mathcal{A}$ of a process X as $\mathcal{A} = \sum_{n=0}^{\infty}\varepsilon^n\mathcal{A}_n$ with $\mathcal{A}_0$ given by (7.19), then one can use (7.11)–(7.12) to find approximate solutions to the full pricing PIDE $(\partial_t + \mathcal{A})u = 0$.

7.2.5 Model Coefficient Polynomial Expansions

Model coefficient expansions are developed in a series of papers by Lorig *et al*. (2014, 2015a, b, c, d). The authors' method (which we shall henceforth refer to as LPP) can be used to find closed-form asymptotic approximations for option prices and implied volatilities in a general d-dimensional Markov setting. Here, for simplicity, we focus on two simple cases: (i) general two-dimensional *local-stochastic volatility* (LSV) models, and (ii) general scalar Lévy-type models.

7.2.5.1 LSV models

We consider a general class of models, in which an asset $S = \mathrm{e}^X$ is modeled as the exponential of a Markov diffusion process X that satisfies the SDEs

$$\begin{aligned}\mathrm{d}X_t &= -\frac{1}{2}\sigma^2(X_t,Y_t)\,\mathrm{d}t + \sigma(X_t,Y_t)\,\mathrm{d}W_t^{\mathbb{Q}},\\ \mathrm{d}Y_t &= f(X_t,Y_t)\,\mathrm{d}t + \beta(X_t,Y_t)\,\mathrm{d}B_t^{\mathbb{Q}},\\ \mathrm{d}\langle W^{\mathbb{Q}},B^{\mathbb{Q}}\rangle_t &= \rho\,\mathrm{d}t,\end{aligned}$$

where $W^{\mathbb{Q}}$ and $B^{\mathbb{Q}}$ are correlated Brownian motions. This is as in the model discussed in Section 7.1.2.4 except we remove the explicit t-dependence in the coefficients for simplicity. Note that the drift of X is $-\frac{1}{2}\sigma^2$, which ensures that $S = \mathrm{e}^X$ is a $\mathbb{Q}$-martingale and so the stock price is arbitrage free.

The generator $\mathcal{A}$ of X is given by

$$\mathcal{A} = a(x,y)(\partial_{xx}^2 - \partial_x) + f(x,y)\partial_y + b(x,y)\partial_{yy}^2 + c(x,y)\partial_{xy}^2, \tag{7.21}$$

where we have defined

$$a(x,y) := \frac{1}{2}\sigma^2(x,y), \qquad b(x,y) := \frac{1}{2}\beta^2(x,y), \qquad c(x,y) := \rho\sigma(x,y)\beta(x,y).$$

For general coefficients (σ, β, f), there is no closed-form (or even semiclosed-form) expression of Γ, the fundamental solution corresponding to $\mathcal{A}$. Thus, we seek a decomposition of $\mathcal{A} = \sum_{n=0}^{\infty} \mathcal{A}_n$ for which the order-zero operator $\mathcal{A}_0$ admits a closed-form fundamental solution Γ_0.

The LPP approach is to expand the coefficients of $\mathcal{A}$ in polynomial basis functions where the zeroth-order terms in the expansion are constant. Specifically,

$$\chi(x,y) = \sum_{n=0}^{\infty} \chi_n(x,y), \qquad \chi \in \{a,b,c,f\},$$

where χ_0 is a constant and $\chi_n(x,y)$ depends polynomially on x and y for every $n \geq 1$. For example, Taylor series:

$$\chi_n(x,y) = \sum_{k=0}^{n} \mathcal{X}_{k,n-k}(x-\bar{x})^k(y-\bar{y})^{n-k}, \qquad \mathcal{X}_{k,n-k} = \frac{1}{k!(n-k)!}\partial_x^k\partial_y^{n-k}\chi(\bar{x},\bar{y}), \tag{7.22}$$

or Hermite polynomials:

$$\chi_n(x,y) = \sum_{k=0}^{n} \mathcal{X}_{k,n-k}H_{k,n-k}(x,y), \qquad \mathcal{X}_{k,n-k} = \langle H_{k,n-k}, \chi\rangle\langle H_{k,n-k}, \chi\rangle.$$

In the Taylor series example, $(\bar{x},\bar{y})$ is a fixed point in $\mathbb{R}^2$. In the Hermite polynomial example, the brackets $\langle\cdot,\cdot\rangle$ indicate an L^2 inner product with a Gaussian weighting. The Hermite polynomials $(H_{n,m})$ form a complete basis in this space and (properly weighted) are orthonormal $\langle H_{n,m}, H_{i,j}\rangle = \delta_{n,i}\delta_{m,j}$.

Upon expanding the coefficients of $\mathcal{A}$ in polynomial basis functions, one can formally write the operator $\mathcal{A}$ as

$$\mathcal{A} = \sum_{n=0}^{\infty} \mathcal{A}_n, \quad \text{with} \quad \mathcal{A}_n = \sum_{k=0}^{n} \mathcal{A}_{k,n-k}, \tag{7.23}$$

where

$$\mathcal{A}_{k,n-k} = a_{k,n-k}(x,y)(\partial_{xx}^2 - \partial_x) + f_{k,n-k}(x,y)\partial_y + b_{k,n-k}(x,y)\partial_{yy}^2 + c_{k,n-k}(x,y)\partial_{xy}^2.$$

Note that the operator $\mathcal{A}$ in (7.23) is of the form (7.9) if one sets $\varepsilon = 1$.

Moreover, the order-zero operator

$$\mathcal{A}_0 = \mathcal{A}_{0,0} = a_{0,0}(\partial_{xx}^2 - \partial_x) + f_{0,0}\partial_y + b_{0,0}\partial_{yy}^2 + c_{0,0}\partial_{xy}^2.$$

has a fundamental solution Γ_0, which is a Gaussian density:

$$\Gamma_0(t,x,y;T,\xi,\omega) = \frac{1}{2\pi\sqrt{|\mathbf{C}|}} \exp\left(-\frac{1}{2}(\boldsymbol{\eta}-\mathbf{m})^{\mathrm{T}}\mathbf{C}^{-1}(\boldsymbol{\eta}-\mathbf{m})\right), \qquad \boldsymbol{\eta} = \begin{pmatrix}\xi\\ \omega\end{pmatrix},$$

with covariance matrix $\mathbf{C}$ and mean vector $\mathbf{m}$ given by

$$\mathbf{C} = (T-t)\begin{pmatrix}2a_{0,0} & c_{0,0}\\ c_{0,0} & 2b_{0,0}\end{pmatrix}, \qquad \mathbf{m} = \begin{pmatrix}x-(T-t)a_{0,0}\\ y+(T-t)f_{0,0}\end{pmatrix}.$$

Since Γ_0 is available in closed form, one can use (7.11) and (7.12) to find u_0 and the sequence of functions $(u_n)_{n\geq 1}$, respectively. After a bit of algebra, one can show that

$$u_n(t,x,y) = \mathcal{L}_n u_0(t,x,y), \tag{7.24}$$

where the operator $\mathcal{L}_n$ is of the form

$$\mathcal{L}_n = \sum_{k,m} \eta^{(n)}_{k,m}(t,x,y)\partial_y^k \partial_x^m (\partial_x^2 - \partial_x).$$

The precise form of the coefficients $(\eta^{(n)}_{k,m})$ will depend on the choice of polynomial basis function. If the coefficients of $\mathcal{A}$ are smooth and bounded, then the small-time-to-maturity accuracy of the price approximation is

$$\sup_{x,y} |u(t,x,y) - \bar{u}_n(t,x,y)| = \mathcal{O}((T-t)^{(n+3)/2}), \qquad \bar{u}_n(t,x,y) := \sum_{k=0}^{n} u_k(t,x,y).$$

The proof can be found in Lorig *et al*. (2015a).

The LPP price expansion also leads to closed-form expressions for implied volatility. Indeed, for European call options, one can easily show that $u_0 = u^{\mathrm{BS}}(\sqrt{2a_{0,0}})$. Therefore, the series expansion for u is of the form (7.13), and hence (I_0, I_1, I_2, I_3) can be computed using (7.15)–(7.16). Moreover, the model-dependent terms $(u_n/\partial_\sigma u^{\mathrm{BS}})$ appearing in (7.15)–(7.16) can be computed explicitly with no numerical integration. This is due to the fact that u_n can be written in the form (7.24). For details, we refer the reader to Lorig *et al*. (2015b).

Example 7.1 (Heston Model) *Consider the Heston (1993) model, under which the risk-neutral dynamics of X are given by*

$$\begin{aligned} dX_t &= -\frac{1}{2}e^{Y_t}dt + e^{Y_t/2}dW_t^{\mathbb{Q}},\\ dY_t &= \left((\kappa\theta - \frac{1}{2}\delta^2)e^{-Y_t} - \kappa\right)dt + \delta\, e^{-Y_t/2}dB_t^{\mathbb{Q}},\\ d\left\langle W^{\mathbb{Q}}, B^{\mathbb{Q}}\right\rangle_t &= \rho\, dt. \end{aligned}$$

The generator of (X,Y) *is given by*

$$\mathcal{A} = \frac{1}{2}e^y(\partial_x^2 - \partial_x) + \left((\kappa\theta - \frac{1}{2}\delta^2)e^{-y} - \kappa\right)\partial_y + \frac{1}{2}\delta^2 e^{-y}\partial_y^2 + \rho\,\delta\partial_x\partial_y.$$

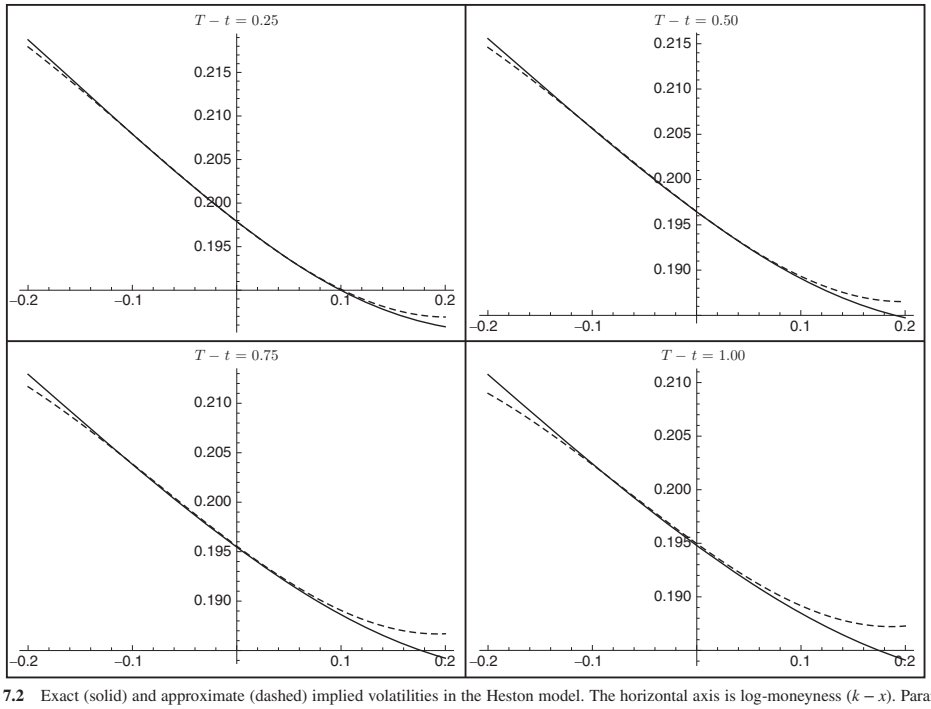

Figure 7.2 Exact (solid) and approximate (dashed) implied volatilities in the Heston model. The horizontal axis is log-moneyness ($k - x$). Parameters: $\kappa = 1.15$, $\theta = 0.04$, $\delta = 0.2$, $\rho = -0.40$ $x = 0.0$, $y = \log\ \theta$.

Thus, using (7.21), we identify

$$a(x,y)=\frac{1}{2}e^{y},\quad b(x,y)=\frac{1}{2}\delta^{2}e^{-y},\quad c(x,y)=\rho\delta,\quad f(x,y)=\left((\kappa\theta-\frac{1}{2}\delta^{2})e^{-y}-\kappa\right).$$

We fix a time to maturity τ *and* log*-strike* k*. Assuming a Taylor series expansion (7.22) of the coefficients of* $\mathcal{A}$ *with* $(\bar{x},\bar{y})=(x,y)$*, the time* t *levels of* (X,Y)*, one computes*

$$\begin{aligned}
I_0&=e^{y/2},\\
I_1&=\frac{1}{8}e^{-y/2}\tau(-\delta^2+2(-e^{y}+\theta)\kappa+e^{y}\delta\rho)+\frac{1}{4}e^{-y/2}\delta\rho(k-x),\\
I_2&=\frac{-e^{-3y/2}}{128}\tau^2(\delta^2-2\theta\kappa)^2+\frac{e^{y/2}}{96}\tau^2(5\kappa^2-5\delta\kappa\rho+\delta^2(-1+2\rho^2))\\
&\quad+\frac{e^{-y/2}}{192}\tau(-4\tau\theta\kappa^2-\tau\delta^3\rho+2\tau\delta\theta\kappa\rho+2\delta^2(8+\tau\kappa+\rho^2))\\
&\quad+\frac{1}{96}e^{-3y/2}\tau\delta\rho(5\delta^2+2(e^{y}-5\theta)\kappa-e^{y}\delta\rho)(k-x)+\frac{1}{48}e^{-3y/2}\delta^2(2-5\rho^2)(k-x)^2.
\end{aligned}$$

The expression for I_3 *is also explicit, but omitted for brevity. In Figure 7.2, we plot the approximate implied volatility* $(I_0+I_1+I_2+I_3)$ *as well as the exact implied volatility* I*, which can be computed using the pricing formula given in Heston (1993) and then inverting the Black–Scholes formula numerically.*

7.2.5.2 Lévy-type models

In this section, we explore how the LPP method can be applied to compute approximate option prices in a one-dimensional Lévy-type setting. Specifically, we consider an asset $S=\mathrm{e}^{X}$, where X is a scalar Lévy-type Markov process. Under some integrability conditions on the size and intensity of jumps, every scalar Markov process on $\mathbb{R}$ can be expressed as the solution of a Lévy–Itô SDE of the form:

$$\mathrm{d}X_t=\mu(X_t)\mathrm{d}t+\sqrt{2a(X_t)}\mathrm{d}W_t^{\mathbb{Q}}+\int z\mathrm{d}\widetilde{N}_t(X_{t-},\mathrm{d}z),$$

where $W^{\mathbb{Q}}$ is a standard Brownian motion; and $\widetilde{N}$ is a state-dependent compensated Poisson random measure

$$\mathrm{d}\widetilde{N}_t(x,\mathrm{d}z)=\mathrm{d}N_t(x,\mathrm{d}z)-\nu(x,\mathrm{d}z)\mathrm{d}t.$$

Note that jumps are now described by a *Lévy kernel* $\nu(x,\mathrm{d}z)$, which is a Lévy measure for every $x\in\mathbb{R}$. In order for S to be a martingale, the drift μ must be given by

$$\mu(x)=-a(x)-\int_{\mathbb{R}}\nu(x,\mathrm{d}z)(\mathrm{e}^{z}-1-z).$$

The generator $\mathcal{A}$ of X is

$$\mathcal{A}=\mu(x)\partial_x+a(x)\partial_{xx}^2+\int\nu(x,\mathrm{d}z)(\theta_z-1-z\partial_x).$$

For general drift $\mu(x)$ variance $a(x)$ and Lévy kernel $\nu(x, \mathrm{d}z)$, there is no closed-form (or even semiclosed-form) expression for the fundamental solution Γ corresponding to $\mathcal{A}$. Thus, we seek a decomposition of generator $\mathcal{A} = \sum_n \mathcal{A}_n$ for which $\mathcal{A}_0$ has a fundamental solution Γ_0 available in semiclosed form.

Once again, the LPP approach of expanding the coefficients and Levy kernel of $\mathcal{A}$ in polynomial basis functions will lead to the desired form for the expansion of $\mathcal{A}$. Specifically, let

$$\mu(x) = \sum_{n=0}^{\infty} \mu_n(x), \qquad a(x) = \sum_{n=0}^{\infty} a_n(x), \qquad \nu(x, \mathrm{d}z) = \sum_{n=0}^{\infty} \nu_n(x, \mathrm{d}z),$$

where $\mu_0(x) = \mu_0$; $a_0(x) = a_0$; and $\nu_0(x, \mathrm{d}z) = \nu_0(\mathrm{d}z)$; and higher order terms $\mu_n(x)$, $a_n(x)$, and $\nu_n(x, \mathrm{d}z)$ depend polynomially on x. For example,

$$\begin{array}{rl} \text{Taylor series:} & a_n(x) = \frac{1}{n!}\partial_x^n a(\bar{x}) \cdot (x - \bar{x})^n, \\ \text{Hermite polynomial:} & a_n(x) = \langle H_n, a \rangle \cdot H_n(x), \end{array}$$

and similarly for μ and ν. Upon expanding the coefficients of $\mathcal{A}$ in polynomial basis functions, one can formally write the operator $\mathcal{A}$ as $\mathcal{A} = \sum_{n=0}^{\infty} \mathcal{A}_n$, where $\mathcal{A}_0$ is given by

$$\mathcal{A}_0 = \mu_0 \partial_x + a_0 \partial_{xx}^2 + \int \nu_0(\mathrm{d}z)(\theta_z - 1 - z\partial_x), \tag{7.25}$$

and each $\mathcal{A}_n$ for $n \geq 0$ is of the form

$$\mathcal{A}_n = \mu_n(x)\partial_x + a_n(x)\partial_{xx}^2 + \int \nu_n(x, \mathrm{d}z)(\theta_z - 1 - z\partial_x).$$

Comparing (7.25) with (7.19), we see that $\mathcal{A}_0$ is the generator of a Lévy process. Thus, using (7.20), the fundamental solution Γ_0 corresponding to $\mathcal{A}_0$ can be written as a Fourier integral

$$\Gamma_0(t, x; T, y) = \frac{1}{2\pi} \int \mathrm{d}\xi \; \mathrm{e}^{\mathrm{i}\xi(x-y) + (T-t)\Phi_0(\xi)},$$

where

$$\Phi_0(\xi) = \mathrm{i}\mu_0\xi - a_0\xi^2 + \int \nu_0(\mathrm{d}z)(\mathrm{e}^{\mathrm{i}\xi z} - 1 - \mathrm{i}\xi z).$$

Since Γ_0 is available in closed form, one can use (7.11) and (7.12) to find u_0 and u_n, respectively. Explicit computations are carried out in Lorig *et al.* (2015c).

In this case, it is convenient to express $u_n(t, x)$ as an (inverse) Fourier transform of $\hat{u}_n(t, x)$. Defining

$$\begin{array}{rl} \text{Fourier transform:} & \hat{u}_n(t, \xi) = \mathcal{F}[u_n(t, \cdot)](\xi) := \int \mathrm{d}x \; \mathrm{e}^{-\mathrm{i}\xi x} u_n(t, x), \\ \text{Inverse transform:} & u_n(t, x) = \mathcal{F}^{-1}[\hat{u}_n(t, \cdot)](x) := \frac{1}{2\pi} \int \mathrm{d}x \; \mathrm{e}^{\mathrm{i}\xi x} \hat{u}_n(t, \xi), \end{array}$$

we have

$$\begin{aligned} \hat{u}_n(t, \xi) = &\sum_{k=1}^{n} \int_t^T \mathrm{d}t_1 \; \mathrm{e}^{(t_1 - t)\Phi_0(\xi)} \int_{\mathbb{R}_0^d} \nu_k(\mathrm{i}\partial_\xi, \mathrm{d}z)(\mathrm{e}^{\mathrm{i}\xi z} - 1 - \mathrm{i}\xi z)\hat{u}_{n-k}(t_1, \xi) \\ &+ \sum_{k=1}^{n} \int_t^T \mathrm{d}t_1 \; \mathrm{e}^{(t_1 - t)\Phi_0(\xi)} (\mu_k(\mathrm{i}\partial_\xi)(\mathrm{i}\xi)\hat{u}_{n-k}(t_1, \xi) + a_k(\mathrm{i}\partial_\xi)(\mathrm{i}\xi)^2 \hat{u}_{n-k}(t_1, \xi)), \end{aligned}$$

where $\hat{u}_0(t,\xi)$ is given by

$$\hat{u}_0(t,\xi) = \mathrm{e}^{\mathrm{i}\xi x + (T-t)\Phi_0(\xi)}\hat{\phi}(\xi).$$

7.2.6 Small "Vol of Vol" Expansion

Lewis (2000) considers a stochastic volatility model of the form

$$\begin{aligned} \mathrm{d}X_t &= -\frac{1}{2}Y_t\,\mathrm{d}t + \sqrt{Y_t}\,\mathrm{d}W_t^{\mathbb{Q}}, \\ \mathrm{d}Y_t &= \alpha(Y_t)\,\mathrm{d}t + \varepsilon\beta(Y_t)\,\mathrm{d}B_t^{\mathbb{Q}}, \\ \mathrm{d}\langle W^{\mathbb{Q}}, B^{\mathbb{Q}}\rangle_t &= \rho\,\mathrm{d}t. \end{aligned}$$

The parameter ε is referred to as the *volatility of volatility*, or *vol of vol* for short. The generator $\mathcal{A}$ of (X,Y) is given by

$$\mathcal{A} = \mathcal{A}_0 + \varepsilon\mathcal{A}_1 + \varepsilon^2\mathcal{A}_2,$$

where

$$\mathcal{A}_0 = \frac{y}{2}(\partial_x^2 - \partial_x) + \alpha(y)\partial_y, \qquad \mathcal{A}_1 = \rho\sqrt{y}\beta(y)\partial_x\partial_y, \qquad \mathcal{A}_2 = \frac{1}{2}\beta^2(y)\partial_y^2.$$

Thus, $\mathcal{A}$ is of the form (7.9). Moreover, the solution of

$$(\partial_t + \mathcal{A}_0)\Gamma_0 = 0, \qquad \Gamma_0(T,x,y;T,\xi) = \delta(x-\xi),$$

is given by a Gaussian

$$\Gamma_0(t,x,y;T,\xi,\zeta) = \frac{1}{\sqrt{2\pi\sigma^2(t,T)}}\exp\left(-\frac{(\xi - x - \sigma^2(t,T)/2)^2}{2\sigma^2(t,T)}\right)\delta(\zeta - \eta(t)),$$

where $\sigma^2(t,T) = \int_t^T \eta(s)\,\mathrm{d}s$, and $\eta(s)$ is the solution of the following ODE:

$$\tfrac{\mathrm{d}\eta(s)}{\mathrm{d}s} = \alpha(\eta(s)), \qquad \eta(t) = y.$$

Since Γ_0 is available in closed form up to finding η, one can use (7.11) and (7.12) to find u_0 and the sequence of functions (u_n), respectively. In particular, for European calls, one finds

$$u_0 = u^{\mathrm{BS}}(\bar{\sigma}(t,T)), \qquad \bar{\sigma}(t,T) = \sqrt{\tfrac{\sigma^2(t,T)}{T-t}}$$

As the order-zero price is equal to the Black–Scholes price, computed with volatility $\bar{\sigma}(t,T)$, one can apply the implied volatility asymptotics of Section 7.2.3 to find approximate implied volatilities. Explicit expressions for approximate implied vols are given in Lewis (2000) for the case $\alpha(y) = \kappa(\theta - y)$ and $\beta(y) = y^{\gamma}$.

7.2.7 Separation of Timescales Approach

In Fouque *et al*. (2000, 2011), the authors develop analytic price approximations for interest rate, credit, and equity derivatives. The authors' approach (henceforth referred to as FPSS) exploits the separation of timescales that is observed in volatility time-series data. More

specifically, the authors consider a class of *multiscale* diffusion models in which the volatility of an underlying is driven by two factors, Y and Z, operating on fast and slow timescales, respectively. Lorig and Lozano-Carbassé (2015) extend the FPSS method to models with jumps. Here, for simplicity, we consider an exponential Lévy-type model with a single slowly varying factor, which drives both the volatility and jump-intensity.

Consider a model for a stock $S = \mathrm{e}^X$, where X is modeled under the risk-neutral pricing measure as the solution of the following Lévy–Itô SDE

$$\begin{aligned}
\mathrm{d}X_t &= \mu(Z_t)\,\mathrm{d}t + \sigma(Z_t)\,\mathrm{d}W_t^{\mathbb{Q}} + \int_{\mathbb{R}} s\,\mathrm{d}\tilde{N}_t(Z_{t-}, \mathrm{d}s), \\
\mathrm{d}Z_t &= (\varepsilon^2 c(Z_t) - \varepsilon\,\Gamma(Z_t)\,g(Z_t))\,\mathrm{d}t + \varepsilon\,g(Z_t)\,\mathrm{d}B_t^{\mathbb{Q}}, \qquad \text{(risk-neutral measure } \mathbb{Q}) \\
\mathrm{d}\langle W^{\mathbb{Q}}, B^{\mathbb{Q}}\rangle_t &= \rho\,\mathrm{d}t.
\end{aligned}$$

where $W^{\mathbb{Q}}$ and $B^{\mathbb{Q}}$ are correlated Brownian motions; and $\widetilde{N}$ is a compensated state-dependent Poisson random measure with which we associate a Lévy kernel $\zeta(z)\nu(\mathrm{d}s)$. The drift $\mu(z)$ is fixed by the Lévy kernel and volatility so that $S = \mathrm{e}^X$ is a martingale:

$$\mu(z) = -\frac{1}{2}\sigma^2(z) - \zeta(z)\int_{\mathbb{R}} \nu(\mathrm{d}s)(\mathrm{e}^s - 1 - s).$$

Note that the volatility σ and Lévy kernel $\zeta\nu$ are driven by Z, which is *slowly varying* in the following sense. Under the physical measure $\mathbb{P}$, the dynamics of Z are given by

$$\mathrm{d}Z_t = \varepsilon^2 c(Z_t)\,\mathrm{d}t + \varepsilon\,g(Z_t)\,\mathrm{d}\tilde{B}_t, \qquad \text{(physical measure } \mathbb{P})$$

where $\tilde{B}_t = B_t^{\mathbb{Q}} - \int_0^t \Gamma(Z_s)\,\mathrm{d}s$ is a standard $\mathbb{P}$-Brownian motion (Γ is known as the *market price of risk* associated with $\tilde{B}$). The infinitesimal generator of Z under the physical measure

$$\mathcal{A}_Z = \varepsilon^2\left(\frac{1}{2}g^2(z)\partial_z^2 + c(z)\partial_z\right)$$

is scaled by ε^2, which is assumed to be a small parameter: $\varepsilon^2 << 1$. Thus, Z fluctuates over an intrinsic timescale $1/\varepsilon^2$, which is large.

Due to the separation of timescales, the (X, Z) process has a generator that naturally factors into three terms of different powers of ε, just as in (7.9):

$$\mathcal{A} = \mathcal{A}_0 + \varepsilon\mathcal{A}_1 + \varepsilon^2\mathcal{A}_2,$$

where $\mathcal{A}_0$, $\mathcal{A}_1$, and $\mathcal{A}_2$ are given by

$$\begin{aligned}
\mathcal{A}_0 &= \mu(z)\partial_x + \frac{1}{2}\sigma^2(z)\partial_x^2 + \zeta(z)\int_{\mathbb{R}} \nu(\mathrm{d}s)(\theta_s - 1 - s\partial_x), \qquad (7.26) \\
\mathcal{A}_1 &= -\Gamma(z)g(z)\partial_z + \rho g(z)\sigma(z)\partial_x\partial_z, \\
\mathcal{A}_2 &= \frac{1}{2}g^2(z)\partial_z^2 + c(z)\partial_z.
\end{aligned}$$

Here, the shift operator θ_s acts on the x variable. Comparing equations (7.19) and (7.26), we observe that, for fixed z, the operator $\mathcal{A}_0$ above is the generator of a Lévy process. As the

operator $\mathcal{A}_0$ in (7.26) acts only on the variable x, the variable z serves only as a parameter. Thus, we have

$$\Gamma_0(t,x,z;T,y) = \frac{1}{2\pi}\int \mathrm{d}\xi\ \mathrm{e}^{\mathrm{i}\xi(x-y)+(T-t)\Phi_0(\xi,z)},$$

where

$$\Phi_0(\xi,z) := \mathrm{i}\mu(z)\xi - \frac{1}{2}\sigma^2(z)\xi^2 + \zeta(z)\int \nu(\mathrm{d}s)(\mathrm{e}^{\mathrm{i}\xi s} - 1 - \mathrm{i}\xi s).$$

Because $\mathcal{A}$ is of the form (7.9) (set $\mathcal{A}_n = 0$ for $n \geq 3$), and since the fundamental solution Γ_0 corresponding to $\mathcal{A}_0$ is known, we can use (7.11) and (7.12) to write u_0 and u_1 explicitly. For an option with payoff $\varphi(X_T)$, we have

$$u_0(t,x,z) = \int \mathrm{d}x\ \varphi(y)\frac{1}{2\pi}\int \mathrm{d}\xi\ \mathrm{e}^{\mathrm{i}\xi(x-y)+(T-t)\Phi_0(\xi,z)} = \frac{1}{2\pi}\int \mathrm{d}\xi\ \mathrm{e}^{\mathrm{i}\xi(x-y)+(T-t)\Phi_0(\xi,z)}\hat{\varphi}(\xi),$$

where $\hat{\varphi}$ is the Fourier transform of ϕ.

A similar computation yields the following expression for u_1:

$$u_1(t,x,z) = \frac{(T-t)/2}{2\pi}\int_{\mathbb{R}} d\xi\ \mathrm{e}^{\mathrm{i}\xi x + t\Phi_0(\xi,z)}\hat{\varphi}(\xi)M_\xi(z),$$

where

$$\begin{aligned} M_\xi(z) = {} & V_1(z)(-\mathrm{i}\xi^3 + \xi^2) + U_1(z)\left(\xi^2\int_{\mathbb{R}}\nu(\mathrm{d}s)(\mathrm{e}^s - 1 - s) + \mathrm{i}\xi\int_{\mathbb{R}}\nu(ds)(\mathrm{e}^{i\xi s} - 1 - \mathrm{i}\xi s)\right) \\ & + V_0(z)(-\xi^2 - i\xi) + U_0(z)\left(-i\xi\int_{\mathbb{R}}\nu(\mathrm{d}s)(\mathrm{e}^s - 1 - s) + \int_{\mathbb{R}}\nu(ds)(\mathrm{e}^{\mathrm{i}\xi s} - 1 - \mathrm{i}\xi s)\right), \end{aligned}$$

and

$$\begin{aligned} V_1(z) &= \frac{1}{2}g(z)\rho\sigma(z)\partial_z\sigma^2(z), & V_0(z) &= -\frac{1}{2}g(z)\Gamma(z)\partial_z\sigma^2(z), \\ U_1(z) &= g(z)\rho\sigma(z)\partial_z\zeta(z), & U_0(z) &= -g(z)\Gamma(z)\partial_z\zeta(z). \end{aligned}$$

If the coefficients of $\mathcal{A}$ and the payoff function are smooth and bounded, then one can establish the following accuracy for the first-order price approximation

$$|u(t,x,z) - (u_0(t,x,z) + \varepsilon u_1(t,x,z))| = \mathcal{O}(\varepsilon^2),$$

which holds pointwise. Note that, to compute the approximate price of an option $u_0 + \varepsilon u_1$, one does not need full knowledge of $(g,\sigma,\Gamma,\zeta,\rho,z,\varepsilon)$. Rather, at order ε, the information contained in these five functions and two variables is entirely captured by four *group parameters* $(\varepsilon U_1(z), \varepsilon U_0(z), \varepsilon V_1(z), \varepsilon V_0(z))$. The values of these four parameters can be obtained by calibrating to observed call and put prices, as described in detail in Lorig and Lozano-Carbassé (2015).

7.2.8 Comparison of the Expansion Schemes

The approximation methods presented in Sections 7.2.5, 7.2.6, and 7.2.7 exploit different small parameters in $\mathcal{A}$ and therefore work best (i.e., provide the most accurate approximations) in different regimes. Rigorous accuracy results must be obtained on a case-by-case

basis. However, it is generally true that the accuracy of $\bar{u}_n := \sum_{k=0}^{n} u_k$ depends on how well $\bar{\mathcal{A}}_n := \sum_{k=0}^{n} A_k$ approximates $\mathcal{A}$.

For example, the Taylor series expansion described in Section 7.2.5 works best if the coefficients of $\mathcal{A}$ are slowly varying. In this case, the coefficients can be well approximated by a Taylor series of low order. Thus, a highly accurate approximation of $\mathcal{A}$ can be obtained with only a few terms.

The approximation considered in Section 7.2.6 works best when the diffusion coefficient of the volatility-driving process Y is small in comparison to the drift coefficient of Y and the drift and diffusion coefficients of X.

Finally, the approximation considered in Section 7.2.7 is when the drift coefficient of the volatility-driving process Z is small in comparison to the diffusion coefficient of Z, which in turn is small in comparison to the drift and diffusion coefficients of X. As mentioned in this chapter, this means that the intrinsic timescale of Z must be slow in comparison to the intrinsic timescale of X.

7.3 Merton Problem with Stochastic Volatility: Model Coefficient Polynomial Expansions

A landmark pair of papers on optimal investment strategy by Robert Merton analyzed the problem of how an investor should optimally allocate his wealth between a riskless bond and some risky assets, in order to maximize his expected utility of wealth. This problem and its variation are now referred to as the *Merton problem*.

In the original papers, each of the risky assets follows geometric Brownian motion with constant volatility. This modeling assumption is convenient from the standpoint of analytic tractability but is not realistic in practice, as it does not allow for stochastic volatility. Because of this, there has been much interest in analyzing how an investor's optimal investment strategy changes in the presence of stochastic volatility. Here, we focus on asymptotic methods for analyzing the stochastic control problem associated with portfolio optimization.

Analysis with multiscale stochastic volatility models described in Section 7.2.7 was presented in the case of simple power utilities in Fouque *et al*. (2000 Section 10.1), and expansions for a hedging problem in Jonsson and Sircar (2002a, b) in the dual optimization problem, both for *fast mean-reverting* stochastic volatility. In Fouque *et al*. (2013), expansions are constructed directly in the primal problem under both fast and slow volatility fluctuations. Indifference pricing approximations with *exponential utility* and fast volatility were studied in Sircar and Zariphopoulou (2005). In this section, we present some new approximations for the Merton problem with stochastic volatility.

7.3.1 Models and Dynamic Programming Equation

The polynomial expansion techniques outlined in Section 7.2.5 can be extended to find explicit asymptotic solutions to the Merton problem in a general local-stochastic volatility (LSV) setting, that is, $\sigma_t = \sigma(t, S_t, Y_t)$. To fix ideas, however, we consider the simpler stochastic volatility (SV) setting, in which a risky asset S, under the physical probability measure $\mathbb{P}$, is the solution

of the following SDE:

$$\mathrm{d}S_t = \mu(Y_t)S_t\,\mathrm{d}t + \sigma(Y_t)S_t\,\mathrm{d}B_t^S, \qquad \mathrm{d}Y_t = c(Y_t)\,\mathrm{d}t + \beta(Y_t)\,\mathrm{d}B_t^Y, \qquad \mathrm{d}\langle B^S, B^Y\rangle_t = \rho\mathrm{d}t,$$

where B^S and B^Y are standard Brownian motions with correlation ρ. Let W denote the wealth process of an investor who holds π_t units worth of currency in S at time t, and has $(W_t - \pi_t)$ units of currency in a bond. For simplicity, we assume the risk-free rate of interest is zero. As such, the wealth process W satisfies

$$\mathrm{d}W_t = \frac{\pi_t}{S_t}\,\mathrm{d}S_t = \pi_t\mu(Y_t)\,\mathrm{d}t + \pi_t\sigma(Y_t)\,\mathrm{d}B_t^S.$$

Observe that S does not appear in the dynamics of the wealth process W.

An investor chooses π_t to maximize his expected utility of wealth at a time T in the future, where utility is measured by a smooth, increasing, and strictly concave function $U : \mathbb{R}_+ \to \mathbb{R}$, and the objective to maximize is $\mathbb{E}\,U(W_T)$. Increasing describes a preference for more wealth than less, whereas concavity captures risk aversion, with more concave being more risk averse. The analysis is illustrated with power utility functions in Section 7.3.3.

We define the investor's *value function* u by

$$u(t, y, w) := \sup_{\pi\in\Pi} \mathbb{E}[U(W_T)|Y_t = y, W_t = w],$$

where Π is the set of *admissible strategies*:

$$\Pi := \left\{ \pi \text{ adapted: } \mathbb{E}\int_0^T \pi_t^2\sigma^2(Y_t)\,\mathrm{d}t < \infty \text{ and } W_t \geq 0 \text{ a.s.} \right\},$$

where adapted means adapted to the filtration generated by (B^S, B^Y).

Assuming that $u \in C^{1,2}([0,T], \mathbb{R}, \mathbb{R}_+)$, the value function solves the Hamilton–Jacobi–Bellman partial differential equation (HJB-PDE) problem:

$$(\partial_t + \mathcal{A}^Y)u + \max_{\pi\in\mathbb{R}} \mathcal{A}^\pi u = 0, \qquad u(T, y, w) = U(w), \tag{7.27}$$

where $(\mathcal{A}^Y + \mathcal{A}^\pi)$ is the generator of (Y, W), assuming a Markov investment strategy $\pi_t = \pi(t, Y_t, W_t)$. Specifically, the operators $\mathcal{A}^Y$ and $\mathcal{A}^\pi$ are given by

$$\mathcal{A}^Y = c(y)\partial_y + \frac{1}{2}\beta^2(y)\partial_y^2,$$

$$\mathcal{A}^\pi = \pi(t, y, w)\mu(y)\partial_w + \frac{1}{2}\pi^2(t, y, w)\sigma^2(y)\partial_w^2 + \pi(t, y, w)\rho\sigma(y)\beta(y)\partial_y\partial_w.$$

The optimal strategy π^* is given by

$$\pi^* = \arg\max_{\pi\in\mathbb{R}} \mathcal{A}^\pi u = -\frac{\mu(\partial_w u) + \rho\beta\sigma(\partial_y\partial_w u)}{\sigma^2(\partial_w^2 u)}, \tag{7.28}$$

where, for simplicity (and from now on), we have omitted the arguments (t, y, w).

Inserting the optimal strategy π^* into the HJB-PDE (7.27) yields

$$(\partial_t + \mathcal{A}^Y)u + \mathcal{N}(u) = 0, \tag{7.29}$$

where $\mathcal{N}(u)$ is a nonlinear term:

$$\mathcal{N}(u) = -\frac{1}{2}\lambda^2\frac{(\partial_w u)^2}{\partial_w^2 u} - \rho\beta\lambda\frac{(\partial_w u)(\partial_y\partial_w u)}{\partial_w^2 u} - \frac{1}{2}\rho^2\beta^2\frac{(\partial_y\partial_w u)^2}{\partial_w^2 u}.$$

Here, we have introduced the *Sharpe ratio* $\lambda(y) := \mu(y)/\sigma(y)$.

7.3.2 Asymptotic Approximation

For general $\{\beta, c, \lambda\}$, there is no closed-form solution of (7.29). Hence, we seek an asymptotic approximation for u. To this end, using equation (7.22) from Section 7.2.5.1 as a guide, we expand the coefficients in (7.29) in a Taylor series about an arbitrary point $\bar{y}$. Specifically, for any function $\chi : \mathbb{R} \to \mathbb{R}$, we may formally write

$$\chi(y) = \sum_{n=0}^{\infty} \varepsilon^n \chi_n(y), \qquad \chi_n(y) := \frac{1}{n!}\partial_y^n \chi(\bar{y})\cdot(y-\bar{y})^n, \qquad \varepsilon = 1, \tag{7.30}$$

where we have once again introduced ε for purposes of accounting. We also expand the function u as a power series in ε

$$u = \sum_{n=0}^{\infty} \varepsilon^n u_n, \qquad \varepsilon = 1. \tag{7.31}$$

Now, for each group of coefficients appearing in (7.29), we insert an expansion of the form (7.30), and we define

$$\mathcal{A}_n := c_n\partial_y + (\tfrac{1}{2}\beta^2)_n\partial_y^2, \qquad n \in \{0\} \cup \mathbb{N}. \tag{7.32}$$

We also insert into (7.29) our expansion (7.31) for u.

Next, collecting terms of like powers of ε, we obtain at lowest order

$$(\partial_t + \mathcal{A}_0)u_0 - (\frac{1}{2}\lambda^2)_0\frac{(\partial_w u_0)^2}{\partial_w^2 u_0} - (\rho\beta\lambda)_0\frac{(\partial_w u_0)(\partial_y\partial_w u_0)}{\partial_w^2 u_0} - (\frac{1}{2}\rho^2\beta^2)_0\frac{(\partial_y\partial_w u_0)^2}{\partial_w^2 u_0} = 0,$$

with $u_0(T, y, w) = U(w)$. We can look for a solution $u_0 = u(t, w)$ that is independent of y, and then we have

$$\partial_t u_0 - (\frac{1}{2}\lambda^2)_0\frac{(\partial_w u_0)^2}{\partial_w^2 u_0} = 0, \qquad u_0(T, w) = U(w). \tag{7.33}$$

We observe that (7.33) is the same *nonlinear* PDE problem that arises when one considers an underlying that has a constant drift $\mu_0 = \mu(\bar{y})$, diffusion coefficient $\sigma_0 = \sigma(\bar{y})$, and Sharpe ratio $\lambda_0 = \lambda(\bar{y}) = \mu(\bar{y})/\sigma(\bar{y})$.

It is convenient to define the *risk-tolerance* function

$$R_0 := \frac{-\partial_w u_0}{\partial_w^2 u_0},$$

and the operators

$$\mathcal{D}_k = R_0^k\partial_w^k, \quad k = 1, 2, \cdots.$$

We now proceed to the order $\mathcal{O}(\varepsilon)$ terms. Using $u_0 = u_0(t, w)$ and (7.33), we obtain

$$(\partial_t + \mathcal{A}_0)u_1 + (\frac{1}{2}\lambda_0^2)\mathcal{D}_2 u_1 + \lambda_0^2 \mathcal{D}_1 u_1 + (\rho\beta\lambda)_0 \mathcal{D}_1 \partial_y u_1 = -(\frac{1}{2}\lambda^2)_1 \mathcal{D}_1 u_0, \tag{7.34}$$

$$u_1(T, y, w) = 0, \tag{7.35}$$

which is a *linear* PDE problem for u_1.

We can rewrite equations (7.34)–(7.35) more compactly as

$$(\partial_t + \mathcal{A}_0 + \mathcal{B}_0(t))u_1 + H_1 = 0, \quad u_1(T, y, w) = 0, \tag{7.36}$$

where the linear operator $\mathcal{B}(t)$ and the source term H_1 are given by

$$\mathcal{B}_0(t) = \tfrac{1}{2}\lambda_0^2 \mathcal{D}_2 + \lambda_0^2 \mathcal{D}_1 + (\rho\beta\lambda)_0 \mathcal{D}_1 \partial_y, \quad H_1 = (\tfrac{1}{2}\lambda^2)_1 R_0 \partial_w u_0.$$

Observe that (7.36) is a linear PDE for u_1.

The following change of variables (see Fouque *et al.*, 2013) will be useful for solving the PDE problem (7.36). Define

$$u_1(t, y, w) = q_1(t, y, z(t, w)), \quad z(t, w) = -\log\, \partial_w u_0(t, w) + \tfrac{1}{2}\lambda_0^2 (T - t). \tag{7.37}$$

Inserting (7.37) into (7.36), we find that q_1 satisfies

$$0 = (\partial_t + \mathcal{A}_0 + \mathcal{C}_0)q_1 + Q_1, \quad q_1(T, y, z) = 0, \tag{7.38}$$

where the operator $\mathcal{C}_0$ is given by

$$\mathcal{C}_0 = \frac{1}{2}\lambda_0^2 \partial_z^2 + (\rho\beta\lambda)_0 \partial_y \partial_z, \tag{7.39}$$

and the function Q_1 satisfies $H_1(t, y, w) = Q_1(t, y, z(t, w))$.

Now, from (7.32) and (7.39), we observe that the operator $(\mathcal{A}_0 + \mathcal{C}_0)$ is the infinitesimal generator of a diffusion in $\mathbb{R}^2$ whose drift vector and covariance matrix are constant. The *semigroup* $\mathcal{P}_0(t, t')$ generated by $(\mathcal{A}_0 + \mathcal{C}_0)$ is given by

$$\mathcal{P}_0(t, T)G(y, z) := \int_{\mathbb{R}^2} \mathrm{d}\eta\, \mathrm{d}\zeta\, \Gamma_0(t, y, z; T, \eta, \zeta)G(\eta, \zeta),$$

where Γ_0, the *fundamental solution* corresponding to $(\partial_t + \mathcal{A}_0 + \mathcal{C}_0)$, is a Gaussian kernel:

$$\Gamma_0(t, y, z; T, \eta, \zeta) = \frac{1}{\sqrt{(2\pi)^3 |\mathbf{C}|}} \exp\left(-\frac{1}{2}\mathbf{m}^{\mathrm{T}} \mathbf{C}^{-1} \mathbf{m}\right),$$

with covariance matrix $\mathbf{C}$ and vector $\mathbf{m}$ given by

$$\mathbf{C} = (T - t)\begin{pmatrix} (\beta^2)_0 & (\rho\beta\lambda)_0 \\ (\rho\beta\lambda)_0 & (\lambda^2)_0 \end{pmatrix}, \qquad \mathbf{m} = \begin{pmatrix} \eta - y - (T - t)c_0 \\ \zeta - z \end{pmatrix}.$$

By Duhamel's principle, the unique classical solution to (7.38) is given by

$$q_1(t) = \int_t^T \mathrm{d}s\, \mathcal{P}_0(t, s)Q_1(s),$$

In the case of a general utility function, (7.33) is easily solved numerically, for instance by solving the fast diffusion (or Black's) equation for the risk tolerance function R_0 (see Fouque *et al*., 2013). Then, u_1 can also be computed numerically using the formulas above. In the case of power utility, there are explicit formulas, as given in Section 7.3.3.

Having obtained an approximation for the value function $u \approx u_0 + u_1$, we now seek an expansion for the optimal control $\pi^* \approx \pi_0^* + \pi_1^*$. Inserting the expansion (7.30) of the coefficients and the expansion (7.31) for u into (7.28), and collecting terms of like powers of ε, we obtain

$$\mathcal{O}(1): \qquad \pi_0 = -\frac{\mu_0(\partial_w u_0)}{(\sigma^2)_0(\partial_w^2 u_0)}, \tag{7.40}$$

$$\mathcal{O}(\varepsilon): \qquad \pi_1 = -\pi_0 \frac{(\sigma^2)_1}{(\sigma^2)_0} - \pi_0 \frac{(\partial_w^2 u_1)}{(\partial_w^2 u_0)} - \mu_1 \frac{(\partial_w u_0)}{(\sigma^2)_0(\partial_w^2 u_0)} - \mu_0 \frac{(\partial_w u_1)}{(\sigma^2)_0(\partial_w^2 u_0)} - (\rho\beta\sigma)_0 \frac{(\partial_y \partial_w u_1)}{(\sigma^2)_0(\partial_w^2 u_0)}. \tag{7.41}$$

Higher order terms for both the value function u and the optimal control π^* can be obtained in the same manner as u_1 and π_1. Analysis of the asymptotic formulas for different utility functions and stochastic volatility models is presented in more detail in Lorig and Sircar (2015).

7.3.3 Power Utility

Finally, we consider a utility function U from the constant relative risk aversion (CRRA), or power family:

$$\text{CRRA utility:} \qquad U(w) := \frac{w^{1-\gamma}}{1-\gamma}, w > 0, \qquad \gamma > 0, \gamma \neq 1,$$

where γ is called the risk aversion coefficient. Here, all the quantities above can be computed explicitly.

The explicit solution u_0 to (7.33) is

$$u_0(t,w) = U(w) \exp\left(\frac{1}{2}\lambda_0^2 \left(\frac{1-\gamma}{\gamma}\right)(T-t)\right).$$

The risk-tolerance function is $R_0 = \frac{w}{\gamma}$, and the transformation in (7.37) is then

$$z(t,w) = \gamma w + (T-t)\left(\frac{2\gamma - 1}{\gamma}\right)\frac{1}{2}\lambda_0^2.$$

An explicit computation reveals that u_1 is given by

$$u_1(t,y,w) = q_1(t,y,z(t,w)) = \frac{1-\gamma}{\gamma} u_0(t,w) \left(\frac{1}{2}\lambda^2(\bar{y})\right)' \left((T-t)(y-\bar{y}) + \frac{1}{2}(T-t)^2\left(c_0 + \frac{1-\gamma}{\gamma}\rho\beta_0\lambda_0\right)\right).$$

For the specific case $\bar{y} = y$, the above expression simplifies to

$$u_1(t,y,w) = \frac{1-\gamma}{\gamma} u_0(t,w) \left(\frac{1}{2}\lambda^2(y)\right)' \left(\frac{1}{2}(T-t)^2 \left(c_0 + \frac{1-\gamma}{\gamma}\rho\beta_0\lambda_0\right)\right).$$

Using these explicit representations of u_0 and u_1, the expressions (7.40) and (7.41) for the optimal stockholding approximations become

$$\pi_0^* = \frac{\mu_0}{\gamma\sigma_0^2},$$

$$\pi_1^*(t,y) = (y-\bar{y})\left(\frac{\mu'(\bar{y})}{\gamma\sigma_0^2} - \frac{\mu_0}{\gamma\sigma_0^4}\left(\sigma^2(\bar{y})\right)'\right) + \frac{(1-\gamma)(T-t)}{\gamma\sigma_0}\left(\rho\beta_0\frac{1}{\gamma}\left(\frac{1}{2}\lambda^2(\bar{y})\right)'\right).$$

For the specific case $\bar{y} = y$, the formula for π_1^* simplifies to

$$\pi_1^*(t,y) = \frac{(1-\gamma)(T-t)}{\gamma\sigma_0}\rho\beta_0\frac{1}{\gamma}\left(\frac{1}{2}\lambda^2(y)\right)'.$$

7.4 Conclusions

Asymptotic methods can be used to analyze and simplify pricing and portfolio optimization problems, and we have presented some examples and methodologies. A key insight is to perturb problems with stochastic volatility around their constant volatility counterparts to obtain the principle effect of volatility uncertainty.

These approaches reduce the dimension of the effective problems that have to be solved, and often lead to explicit formulas that can be analyzed for intuition. In the context of portfolio problems accounting for stochastic volatility, recent progress has been made in cases where there are transaction costs (Bichuch and Sircar, 2014, Kallsen and Muhle-Karbe, 2013) or stochastic risk aversion that varies with market conditions (Dong and Sircar, 2014), and under more complex local-stochastic volatility models (Lorig and Sircar, 2015).

Acknowledgements

RL's work is partially supported by NSF grant DMS-0739195. RS's work is partially supported by NSF grant DMS-1211906.

References

Aït-Sahalia, Y. and Jacod, J. (2014). *High-frequency financial econometrics*. Princeton: Princeton University Press.

Bakshi, G., Cao, C. and Chen, Z. (1997). Empirical performance of alternative option pricing models. *Journal of Finance*, 52 (5), 2003–2049.

Berestycki, H., Busca, J. and Florent, I. (2004). Computing the implied volatility in stochastic volatility models. *Communications on Pure and Applied Mathematics*, 57 (10), 1352–1373.

Bichuch, M. and Sircar, R. (2014). Optimal investment with transaction costs and stochastic volatility. Submitted.

Black, F. and Scholes, M. (1972). The valuation of option contracts and a test of market efficiency. *Journal of Finance*, 27, 399–417.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *The Journal of Political Economy*, 81(3), 637–654.

Dong, Y. and Sircar, R. (2014). Time-inconsistent portfolio investment problems. In *Stochastic analysis and applications 2014 – in honour of Terry Lyons* (ed. Crisan, D. Hambly, B. and Zariphopoulou T.). Berlin: Springer.

Dupire, B. (1994). Pricing with a smile. *Risk*, 7 (1), 18–20.

Engle, R. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50 (4), 987–1008.

Fouque, J.-P., Papanicolaou, G. and Sircar, R. (2000). *Derivatives in financial markets with stochastic volatility*. Cambridge: Cambridge University Press.

Fouque, J.-P., Papanicolaou, G., Sircar, R. and Solna, K. (2011). *Multiscale stochastic volatility for equity, interest rate, and credit derivatives*. Cambridge: Cambridge University Press.

Fouque, J.-P., Sircar, R. and Zariphopoulou, T. (2013). Portfolio optimization & stochastic volatility asymptotics. *Mathematical Finance*, forthcoming.

Gulisashvili, A. (2012). *Analytically tractable stochastic stock price models*. New York: Springer.

Heston, S. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6 (2), 327–343.

Hull, J. and White, A. (1987). The pricing of options on assets with stochastic volatilities. *Journal of Finance*, 42 (2), 281–300.

Jonsson, M. and Sircar, R. (2002a). Optimal investment problems and volatility homogenization approximations. In *Modern methods in scientific computing and applications* (ed. Bourlioux, A., Gander, M. and Sabidussi, G.), NATO Science Series II vol. 75. New York: Kluwer, pp. 255–281.

Jonsson, M. and Sircar, R. (2002b). Partial hedging in a stochastic volatility environment. *Mathematical Finance*, 12 (4), 375–409.

Kallsen, J. and Muhle-Karbe, J. (2013). The general structure of optimal investment and consumption with small transaction costs. Submitted.

Lee, R.W. (2004). The moment formula for implied volatility at extreme strikes. *Mathematical Finance*, 14 (3), 469–480.

Lewis, A. (2000). *Option valuation under stochastic volatility*. Newport Beach, CA: Finance Press.

Lorig, M. and Lozano-Carbassé, O. (2015). Multiscale exponential Lévy models. *Quantitative Finance*, 15 (1), 91–100.

Lorig, M., Pagliarani, S. and Pascucci, A. (2014). A Taylor series approach to pricing and implied vol for LSV models. *Journal of Risk*, 17, 1–17.

Lorig, M., Pagliarani, S. and Pascucci, A. (2015a). Analytical expansions for parabolic equations. *SIAM Journal on Applied Mathematics*, 75, 468–491.

Lorig, M., Pagliarani, S. and Pascucci, A. (2015b). Explicit implied volatilities for multifactor local-stochastic volatility models. *Mathematical Finance*, forthcoming.

Lorig, M., Pagliarani, S. and Pascucci, A. (2015c). A family of density expansions for Lévy-type processes with default. *Annals of Applied Probability*, 25 (1), 235–267.

Lorig, M., Pagliarani, S. and Pascucci, A. (2015d). Pricing approximations and error estimates for local Lévy-type models with default. *Computers and Mathematics with Applications*, forthcoming.

Lorig, M. and Sircar, R. (2015). Portfolio optimization under local-stochastic volatility: coefficient Taylor series approximations & implied Sharpe ratio. Submitted.

Merton, R. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3 (1), 125–144.

Merton, R. (1992). *Continuous-time finance*. Oxford: Blackwell.

Merton, R.C. (1969). Lifetime portfolio selection under uncertainty: the continuous-time case. *Review of Economics and Statistics*, 51 (3), 247–257.

Merton, R.C. (1971). Optimum consumption and portfolio rules in a continuous-time model. *Journal of Economic Theory*, 3 (1–2), 373–413.

Nelson, D. (1990). ARCH models as diffusion approximations. *Journal of Econometrics*, 45 (1–2), 7–38.

Nelson, D. (1991). Conditional heteroskedasticity in asset returns: a new approach. *Econometrica*, 59 (2), 347–370.

Scott, L. (1987). Option pricing when the variance changes randomly: theory, estimation, and an application. *Journal of Financial and Quantitative Analysis*, 22 (4), 419–438.

Sircar, R. and T. Zariphopoulou (2005). Bounds and asymptotic approximations for utility prices when volatility is random. *SIAM Journal of Control and Optimization*, 43 (4), 1328–1353.

Tehranchi, M.R. (2009). Asymptotics of implied volatility far from maturity. *Journal of Applied Probability*, 629–650.

Wiggins, J. (1987). Option values under stochastic volatility. *Journal of Financial Economics*, 19 (2), 351–372.